

Algèbre

Chapitre ALG.1 Nombres complexes & Trigonométrie 1

1	Définition de \mathbb{C}	1
2	Forme exponentielle	4
2.1	Module	4
2.2	Exponentielle imaginaire	6
2.3	Exponentielle générale e^z et forme exponentielle	7
2.4	Complément – racines n -ièmes d'un complexe [H.P]	10
3	Trigonométrie & Applications des nombres complexes	14
3.1	En trigonométrie	14
3.2	Techniques calculatoires	17
4	Exercices	20
4.1	Généralités	20
4.2	Résolution d'équations & Racines	21
4.3	Trigonométrie	26

Chapitre ALG.2 Polynômes 1

1	Définition de $\mathbb{K}[X]$	1
1.1	Généralités	2
1.2	Propriétés du degré	3
2	Polynôme dérivé & primitif	4
2.1	Généralités	5
3	Racines	7
3.1	Généralités	7
3.2	Existence de racines & Comptage	9
3.3	Cas du second degré	12

4	Exercices	15
4.1	Généralités, Racines, Factorisation	15
4.2	Équations fonctionnelles polynomiales	17
4.3	Familles classiques	18

Chapitre ALG.3 Espaces Vectoriels 1

1	Structure d'espace vectoriel	2
1.1	Généralités	2
1.2	Combinaisons linéaires & Sous-espaces vectoriels	6
2	Familles de vecteurs	15
2.1	Famille libre	16
2.2	Familles génératrices	20
2.3	Base	21
2.4	Extraction & Complétion	23
3	Dimension	26
3.1	Généralités	26
3.2	Sous-espaces et dimension	29
3.3	Rang	29
4	Applications linéaires	30
4.1	Généralités	30
4.2	Structure de \mathbb{K} -espace vectoriel sur $\mathcal{L}_{\mathbb{K}}(E, F)$ et puissances	34
4.3	Image & Noyau	36
4.4	Isomorphismes	41
4.5	Cas particulier de la dimension finie	42
5	Exercices	46
5.1	Structure d'espace vectoriel	46
5.2	Familles de vecteurs et dimension	47
5.3	Applications linéaires	53

Chapitre ALG.4 Matrices 1

1	Matrices & Opérations	2
1.1	Définition	2
1.2	Opérations sur les matrices	3
1.3	Échelonnement et algorithme de GAUß-JORDAN des matrices	11

1.4	Matrices remarquables	20
1.5	Et en Python?	21
2	Représentation matricielle d'objets linéaires	23
2.1	Matrice d'un vecteur, d'une famille	23
2.2	Matrice d'une application linéaire	28
2.3	Opérations endomorphiques & opérations sur les matrices	34
2.4	Changements de base	37
3	Exercices	45
3.1	Calcul matriciel	45
3.2	Représentation matricielle	50
3.3	Changements de base	55

Chapitre ALG.5 Diagonalisation 1

1	Éléments propres	2
1.1	Pour un endomorphisme	2
1.2	Pour une matrice	3
1.3	Calcul effectif en dimension finie	6
1.4	Utilisation d'un polynôme annulateur	9
1.5	Et en Python?	10
2	Critère de diagonalisation	11
2.1	Familles de vecteurs propres	11
2.2	Diagonalisation	12
3	Premières applications	18
3.1	Calculs des puissances d'une matrice	18
3.2	Calculs de commutants	19
4	Exercices	21
4.1	Calculs d'éléments propres	21
4.2	Pour des matrices	21
4.3	Pour des endomorphismes	26
4.4	Polynôme annulateur	27

Chapitre ALG.6 Produit scalaire euclidien 1

1	Généralités	2
1.1	Espaces euclidiens	2

1.2	Norme & Distance	4
2	Orthogonalité	9
2.1	Orthogonal d'une partie	9
2.2	Familles orthogonales & orthonormales	11
2.3	Coordonnées d'un vecteur dans une base orthonormale	12
2.4	Diagonalisation des matrices symétriques réelles	14
3	Projection orthogonale	17
3.1	Projection orthogonale	17
4	Exercices	24
4.1	Inégalités classiques	24
4.2	Diagonalisation de matrices symétriques	26
4.3	Projection orthogonale	31

Analyse

Chapitre ANA.7 Fonctions de la variable réelle 1

1	Généralités	2
2	Limites et continuité	6
2.1	Limite d'une fonction en un point	6
2.2	Continuité	11
3	Dérivation	17
3.1	Généralités et premières propriétés	17
3.2	Dérivées d'ordre supérieur	20
3.3	<i>Extrema</i> et théorème de ROLLE	21
4	Développements limités	25
4.1	Généralités	25
4.2	Développement géométrique	28
4.3	Développements obtenus par la formule de TAYLOR-YOUNG	28
4.4	Développements obtenus par primitivation	30

4.5	Développements obtenus par produits	32
4.6	Composition de développements limités	33
5	Exercices	35
5.1	Continuité	35
5.2	Dérivabilité	37
5.3	Fonctions usuelles	41
5.4	Développements limités	43
5.5	Méthode numérique	44
5.6	Du côté de l'Algèbre	46

Chapitre ANA.8 Fonctions de plusieurs variables 1

1	Généralités	2
1.1	Premiers exemples	2
1.2	Vocabulaire	3
2	Limite et continuité	4
2.1	Limite	5
2.2	Continuité	6
3	Dérivabilité	8
3.1	Dérivées directionnelles	8
3.2	Gradient	12
3.3	Fonctions \mathcal{C}^1	13
3.4	Fonctions \mathcal{C}^2	17
3.5	<i>Extrema</i>	19
4	Exercices	23
4.1	Généralités	23
4.2	Dérivation & Primitivation	24
4.3	Règle de la chaîne et équation aux dérivées partielles	26
4.4	En Physique	31
4.5	Optimisation	32

Chapitre ANA.9 Dynamique : Équations différentielles & Suites numériques 1

1	Dynamique discrète : suites numériques	2
1.1	Généralités	2

1.2	Limite d'une suite	6
1.3	Théorèmes de convergence par majoration, minoration et encadrement	8
1.4	Suites extraites des termes pairs et impairs	9
1.5	Théorèmes de convergence par monotonie	10
1.6	Suites remarquables	13
2	Dynamique continue : équations différentielles	23
2.1	Généralités	23
2.2	Équations différentielles linéaires scalaires	26
2.3	Résolution approchée par la méthode d'EULER	37
2.4	Systèmes différentiels linéaires à coefficients constants	40
3	Modélisation de phénomènes dynamiques	42
3.1	Taux d'évolution et dynamiques linéaires	42
3.2	Modélisation de systèmes présentant une dynamique	43
3.3	Dynamique des populations	45
4	Exercices	50
4.1	Suites numériques	50
4.2	Équations différentielles	60
4.3	Modélisation discrète et continue	64

Chapitre ANA.10 Séries Numériques 1

1	Généralités	2
1.1	Définitions	2
1.2	Propriétés	5
2	Séries usuelles	6
2.1	Géométrique	6
2.2	Exponentielle	8
2.3	de RIEMANN	10
3	Séries de signe constant & Convergence absolue	11
3.1	Séries de signe constant	11
3.2	Séries à termes quelconques & Convergence absolue	17
3.3	Résumé : plan d'étude d'une série numérique	19
4	Séries doubles	19

5	Exercices	22
5.1	Séries simples	22
5.2	Séries doubles	28
Chapitre ANA.11 Intégration		1
1	Primitives & Intégration sur un segment	2
1.1	Primitives	2
1.2	Intégrale sur un segment	2
1.3	Propriétés	5
1.4	Lien entre primitive et intégrale	7
1.5	Calculs d'intégrales	9
1.6	Sommes de RIEMANN & Intégration Numérique	14
2	Intégration sur un intervalle quelconque	17
2.1	Généralités	18
2.2	Propriétés des intégrales convergentes	25
2.3	Calculs d'intégrales	26
2.4	Intégrales de fonctions de signe constant	29
2.5	Fonctions de signe quelconque & Convergence absolue	32
2.6	Plan d'étude d'une intégrale	34
3	Exercices	35
3.1	Intégrales sur un segment	35
3.2	Intégrales impropres	44

Aléatoire

Chapitre ALEA.12 Dénombrement, Espaces probabilisés, Variables aléatoires		1
1	Dénombrement	2
1.1	Cardinal d'un ensemble fini	4
1.2	Listes, Permutations, Combinaisons	7

2	Axiomatique des probabilités	11
2.1	Espace probabilisé	12
2.2	Univers & Espace probabilisable	13
2.3	Espace probabilisé	15
2.4	Résultat d'existence de probabilités	18
2.5	Conditionnement & Indépendance d'évènements	20
3	Variables aléatoires	26
3.1	Généralités	27
3.2	Fonction de répartition & Loi	28
4	Exercices	33
4.1	Dénombrement	33
4.2	Espaces probabilisés	36
Chapitre ALEA.13 Variables aléatoires discrètes		1
1	Généralités	2
1.1	Définitions	2
1.2	Loi & fonction de répartition	5
1.3	Propriétés des variables aléatoires réelles discrètes	9
1.4	Indépendance	12
2	Espérance, Variance, Moments	14
2.1	Espérance	14
2.2	Moments d'ordre supérieur	19
3	Lois discrètes usuelles	22
3.1	Loi uniforme discrète	24
3.2	Loi de BERNOULLI & binomiale	26
3.3	Loi Hypergéométrique	30
3.4	Loi Géométrique & Absence de mémoire	32
3.5	Loi de POISSON	36
3.6	Bilan des lois discrètes	38
4	Exercices	39
4.1	Généralités	39
4.2	Expérience aléatoire	45
4.3	Matrices aléatoires	54
4.4	Récurrence & Chaîne de MARKOV	55

Chapitre ALEA.14	Variables aléatoires à densité	1
1	Généralités	2
1.1	Définitions	2
1.2	Loi & fonction de répartition	4
1.3	Fonctions de variables aléatoires à densité	11
1.4	Indépendance	14
2	Espérance, Variance, Moments	17
2.1	Espérance d'une variable à densité	17
2.2	Moments d'ordre supérieur	22
3	Lois usuelles	24
3.1	Loi uniforme sur $[a, b]$	24
3.2	Loi exponentielle	27
3.3	La loi normale	30
3.4	Bilan des lois continues usuelles	35
4	Exercices	38
4.1	Généralités	38
4.2	Études de lois	40
4.3	Fonctions de variables aléatoires	41
4.4	Discret vS. Continu	41
4.5	Autour des min et max	45
Chapitre ALEA.15	Couples aléatoires	1
1	Couples aléatoires	2
1.1	Généralités	2
1.2	Fonction de répartition & Loi	3
2	Couples aléatoires discrets	4
2.1	Système complet associé	4
2.2	Loi marginale, conjointe, conditionnelle	5
2.3	Espérance, Covariance	14
3	Sommes de variables aléatoires indépendantes & convolution	21
3.1	Cas discret	21
3.2	Cas à densité	24

4	Exercices	28
4.1	Calculs déterministes	28
4.2	Généralités	28
4.3	Étude de lois	30
4.4	Somme de variables aléatoires & Stabilité de lois	33

Chapitre ALEA.16	Théorèmes limites probabilistes	1
1	Inégalités de concentration	2
2	Théorèmes limites	5
2.1	Moyenne et variance empirique	5
2.2	Loi faible des grands nombres	6
2.3	Loi limite de la centrée/réduite : le théorème central limite	12
3	Approximations de lois	17
3.1	Conséquences du théorème central limite	17
3.2	Conséquences de calculs directs	18
4	Exercices	24
4.1	Inégalités de concentration	24
4.2	Théorèmes limites & Approximations	24


Chapitre ALEA.17	Statistiques Descriptives & Inférentielles	1
1	La Statistique : position du problème	2
1.1	Où apparaît l'aléatoire?	2
1.2	Statistiques descriptive	2
1.3	Statistique inférentielle	3
2	Statistiques descriptives	3
2.1	Univariées	4
2.2	Bivariées	15
3	Statistiques inférentielles	24
3.1	Estimation ponctuelle	24
3.2	Estimation par Intervalle de confiance	30
3.3	Test de conformité à la moyenne	36
4	Exercices	42
4.1	Descriptives	42
4.2	Estimateurs	43

4.3	Intervalles de confiance	46
4.4	Tests	50

Annexes

Chapitre ANN.18	L'alphabet grec	1
Chapitre ANN.19	Résumé des méthodes	3
1	En Algèbre	3
2	En Analyse	10
3	En Aléatoire	15
Chapitre ANN.20	Annexe – Questions de cours posées au concours	
	Agro–Véto	21
1	Algèbre	21
2	Géométrie	24
3	Analyse	24
4	Dénombrement, Probabilités & Statistiques	27
Chapitre ANN.3	Synthèses des lois probabilistes	1

QUELQUES REMARQUES À PROPOS DE L'UTILISATION DE CE POLYCOPIÉ

1. Les énoncés et faits hors-programme, mais très classiques parfois, seront indiqués par le logo [H.P]. Si vous souhaitez les utiliser à un concours, il faut donc en connaître la preuve ou la méthode mise en jeu.
2. Les preuves déjà tapées sont généralement des démonstrations non exigibles en BCPST, mais qui peuvent être lues en seconde (voir troisième) lecture, ou en première pour les plus à l'aise/ et 5/2.
3. Quelques éléments¹ relatifs à l'application informatique des Mathématiques (simulations, calculs matriciels, approximations d'intégrales ou de solutions d'équations différentielles par le logo  .
4. En revanche, l'Informatique (Algorithmique notamment) fera l'objet d'un polycopié séparé.
5. liste de questions de cours émises par le SCAV est disponible en fin de polycopié. Nous nous en servons surtout en fin d'année, mais elle peut dorénavant déjà vous aider à cibler les points les plus importants du cours. Elle sera actualisée en Novembre afin de tenir compte du rapport mis à disposition.

LES COMMANDEMENTS DU COURS DE MATHÉMATIQUES (ET DES DEVOIRS ASSOCIÉS)

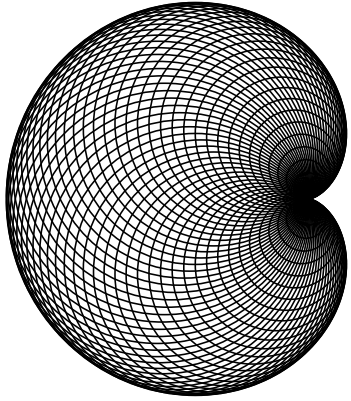
1. Tu n'oublieras jamais les liens logiques entre les lignes.
2. Tu ne diviseras jamais sans t'assurer que ce n'est pas par zéro.
3. Tu placeras bien les parenthèses et les barres de fractions dans tes calculs.
4. Tu ne multiplieras pas une inégalité par un nombre dont tu ne connais pas le signe.
5. Tu ne compareras jamais deux complexes.
6. Tu n'écriras jamais de logarithme d'un nombre complexe.
7. Tu ne confondras jamais f et $f(x)$, (u_n) et u_n . En revanche, en BPCST, tu pourras confondre P et $x \mapsto P(x)$ si P est un polynôme.
8. Tu n'écriras jamais de probabilités strictement négatives ou strictement supérieures à 1.
9. Tu n'appliqueras jamais de formule des probabilités totales sans préciser de système complet d'évènements.
10. Tu n'écriras jamais de nouvelle variable sans l'avoir introduite au préalable (même si ce sont celles de ton cours!).
11. Tu ne parleras pas des variables muettes comme des autres.
12. Tu écriras des phrases complètes et sensées.
13. Tu utiliseras une écriture lisible.
14. Tu répondras à la question et tu encadreras le résultat final.

Copyright ©2022

Ce document est mis à disposition selon les termes de la licence Creative Commons "Attribution – Pas d'utilisation commerciale – Partage dans les mêmes conditions 3.0 non transposé".



¹Tous les détails sont à retrouver dans le polycopié d'Informatique



Version du 8 mars 2022



Première partie

Algèbre

Chapitre ALG.1.

Nombres complexes & Trigonométrie

Résumé & Plan

L'objectif de ce chapitre est de revoir certaines propriétés de première année sur les nombres complexes. Quelques compléments seront présentés en fin de chapitre, notamment sur les solutions de l'équation $z^n = 1$ avec $n \in \mathbf{N}$ appelées *Racines n -ièmes de l'unité*.

1	Définition de \mathbf{C}	1
2	Forme exponentielle	4
2.1	Module	4
2.2	Exponentielle imaginaire	6
2.3	Exponentielle générale e^z et forme exponentielle	7
2.4	Complément – racines n -ièmes d'un complexe [H.P]	10
3	Trigonométrie & Applications des nombres complexes	14
3.1	En trigonométrie	14
3.2	Techniques calculatoires	17

4	Exercices	20
4.1	Généralités	20
4.2	Résolution d'équations & Racines	21
4.3	Trigonométrie	26

L'histoire des nombres complexes commence vers le milieu du XV ième siècle avec une première apparition en 1545, dans l'œuvre de CARDAN, d'une expression contenant la racine carrée d'un nombre négatif, nombre qu'il appelle «sophistiqué». C'est Raphaël BOMBELLI qui met en place les règles de calcul sur ces quantités que l'on appelle alors impossibles avant de leur donner le nom d'imaginaires.

— **Le saviez-vous ?**

1. DÉFINITION DE \mathbf{C}

On souhaite construire un ensemble de nombres, appelé ensemble des *nombres complexes* dans la suite, dans lequel certaines équations admettent une solution (alors

que ce n'est pas le cas dans \mathbf{R}) comme par exemple

$$x^2 + 1 = 0. \tag{1.1}$$

Même si nous n'insisterons pas trop là-dessus : il ne suffit pas de prétendre son existence pour qu'il existe, *i.e.* la phrase «soit \mathbf{C} un ensemble contenant \mathbf{R} et possédant un élément i tel que $i^2 = -1$ » n'a aucune légitimité mathématique. Un *principe de construction* en Mathématiques — par exemple de construction d'ensemble ici — consiste en la démarche suivante : on part d'un ensemble déjà connu, \mathbf{R} en l'occurrence (mais attention, cet ensemble aussi on ne vous l'a jamais construit!) et on en définit un autre possédant les propriétés souhaitées. C'est ce qui est fait dans le paragraphe ci-après. Dans le suivant, nous oublierons déjà le fait que les éléments de \mathbf{C} peuvent être vus comme des couples $(x, y) \in \mathbf{R}^2$, et nous reprenons la notation $x + iy$ habituelle.

PRINCIPE DE CONSTRUCTION DE \mathbf{C} COMME \mathbf{R}^2 MUNI DE DEUX LOIS. [H.P] Rappelons que $\mathbf{R}^2 = \{(x, y), x \in \mathbf{R}, y \in \mathbf{R}\}$, c'est un ensemble bien défini que nous pouvons utiliser pour construire \mathbf{C} . Alors on note \mathbf{C} l'ensemble \mathbf{R}^2 muni des lois $+$ et \times suivantes :

1. **(Somme de nombres complexes)** $\forall (x, y, x', y') \in \mathbf{R}^4, (x, y) + (x', y') = (x + x', y + y')$,
2. **(Produit de nombres complexes)** $\forall (x, y, x', y') \in \mathbf{R}^4, (x, y) \cdot (x', y') = (xx' - yy', xy' + x'y)$.¹

Les éléments de \mathbf{C} sont plutôt représentés de la manière suivante : l'élément (x, y) est noté $x + iy$, et les assertions précédentes deviennent :

1. **(Somme de nombres complexes)** $\forall (x, y, x', y') \in \mathbf{R}^4, (x + iy) + (x' + iy') = (x + x') + i(y + y')$,
2. **(Produit de nombres complexes)** $\forall (x, y, x', y') \in \mathbf{R}^4, (x + iy) \cdot (x' + iy') = xx' - yy' + i(xy' + x'y)$,

¹Les coordonnées du couple correspondent aux parties réelles et imaginaires de $(x + iy)(x' + iy')$ avec la règle de calcul $i^2 = -1$

impliquant en particulier que $i^2 = (0, 1) \cdot (0, 1) = (-1, 0)$ en faisant $y = 1, y' = 1, x = 0$ et $x' = 0$ dans les définitions de loi produit précédentes, *i.e.* en notation complexe

$$i^2 = -1.$$

On a donc construit un élément noté i et un ensemble \mathbf{C} , où cet élément i est une solution dans \mathbf{C} de $x^2 + 1 = 0$. C'est ce qu'on voulait. Classiquement on exige certaines propriétés supplémentaire sur les lois $+, \times$ (associativité, inverse, *etc.*) qui sont vérifiées ici.

Les lois précédentes permettent d'additionner et de multiplier deux complexes. On peut aussi multiplier tout complexe $(x, y) = x + iy$ par $\lambda \in \mathbf{R}$ (*resp.* \mathbf{C}) :

$$\lambda \cdot (x + iy) = (\lambda + i0)(x + iy) = (\lambda x) + i(\lambda y), \quad \lambda \in \mathbf{R} \quad (\text{resp. } \mathbf{C}).$$

En particulier, on peut montrer que cette opération externe $\mathbf{R} \times \mathbf{C} \rightarrow \mathbf{C}$ (*resp.* $\mathbf{C} \times \mathbf{C} \rightarrow \mathbf{C}$) définit une structure de \mathbf{R} -espace vectoriel sur \mathbf{C} (*resp.* \mathbf{C} -espace vectoriel), voir le **Chapter ALG.3**.

Définition ALG.1.1 | Définition d'un nombre complexe

- ▶ Les éléments de \mathbf{C} sont appelés *nombres complexes*. Si $z \in \mathbf{C}$, alors $z = x + iy$ avec $x \in \mathbf{R}$ et $y \in \mathbf{R}$. On appellera *partie réelle* de z le réel x (noté $\text{Re}(z)$), et y la partie imaginaire (notée $\text{Im}(z)$).
- ▶ L'écriture $z = x + iy$ est appelée *forme algébrique* du nombre complexe z et elle est unique.
Si $y = 0$, on dit que x est réel. Si $x = 0$, on dit que z est *imaginaire pur* (ensemble noté $i\mathbf{R}$).

Méthode Unicité de l'écriture algébrique et identification

Une reformulation de l'unicité de l'écriture $z = x + iy$ est la suivante :

$$x + iy = x' + iy' \iff x = x', \quad y = y', \quad (x, y, x', y') \in \mathbf{R}^4.$$

On peut donc *identifier* partie réelle et partie imaginaire.

Attention La partie imaginaire d'un nombre complexe est, par définition, une quantité réelle.

Remarque 1.1 — On a l'inclusion immédiate suivante :

$$\mathbf{R} \times \{0\} \subset \mathbf{C}.$$

Mais² comme $\mathbf{R} \times \{0\}$ peut être *identifié* à \mathbf{R} (grâce à la bijection $x \in \mathbf{R} \mapsto (x, 0) \in \mathbf{R} \times \{0\}$), on note en général

$$\mathbf{R} \subset \mathbf{C}.$$

Définition ALG.1.2 | Complexe conjugué

Si $z \in \mathbf{C}$, on appelle *conjugué de* $z = x + iy$ le complexe $\bar{z} = x - iy$.

La construction précédente de \mathbf{C} nous permet d'assimiler tout complexe à un unique point de $\mathbf{R}^2 = \{(x, y), x \in \mathbf{R}, y \in \mathbf{R}\}$. Ceci n'est pas surprenant : les points géométriques de \mathbf{R}^2 possèdent deux coordonnées, et les complexes sont caractérisés par deux scalaires : partie réelle et partie imaginaire. Ce constat nous mène tout droit à la définition suivante.

Définition ALG.1.3 | Affixe

Soit $M = (x, y) \in \mathbf{R}^2$, l'élément $z = x + iy \in \mathbf{C}$ est appelé *affixe* de M .
Soit $u = (x, y)$ un vecteur de \mathbf{R}^2 , l'élément $z = x + iy \in \mathbf{C}$ est appelé *affixe* de u .

La notion d'affixe permet donc de relier la géométrie du plan dans \mathbf{R}^2 aux complexes ; on pourra donc se servir largement des complexes pour traiter des problèmes de géométrie.

Proposition ALG.1.1 | Quelques propriétés des complexes

Soient $z, z' \in \mathbf{C}$.

²i.e. l'ensemble des complexes de partie imaginaire nulle

- (Existence d'un élément inverse)** Si $z \neq 0$, alors z est inversible dans \mathbf{C} i.e. il existe $z' \in \mathbf{C}$ tel que $zz' = z'z = 1$.
- (\mathbf{R} -linéarité de la partie réelle/imaginaire)** pour tous $\lambda \in \mathbf{R}$ et $z \in \mathbf{C}$, alors

$$\operatorname{Re}(\lambda z) = \lambda \operatorname{Re}(z), \quad \operatorname{Im}(\lambda z) = \lambda \operatorname{Im}(z).$$

- (Conjugué d'une somme/d'un produit)**

$$\overline{z + z'} = \bar{z} + \bar{z}', \quad \overline{z \times z'} = \bar{z} \times \bar{z}'.$$

- (Partie réelle/imaginaire en fonction du conjugué)**

$$\operatorname{Re}(z) = \frac{1}{2}(z + \bar{z}), \quad \operatorname{Im}(z) = \frac{1}{2i}(z - \bar{z}).$$

- $z \in \mathbf{R} \iff \bar{z} = z$ et $z \in i\mathbf{R} \iff \bar{z} = -z$.

Preuve

- On anticipe légèrement sur la définition qui suit, si $z = x + iy \neq 0$ alors $x \neq 0$ ou $y \neq 0$. On vérifie aisément que z' défini par $z' = \frac{x}{x^2+y^2} + i \left(\frac{-y}{x^2+y^2} \right)$ est un inverse pour le complexe z .
- (\mathbf{R} -linéarité de la partie réelle/imaginaire)** soient $\lambda \in \mathbf{R}$ et $z \in \mathbf{C}$, alors $\operatorname{Re}(\lambda z) = \operatorname{Re}(\lambda(x + iy)) = \operatorname{Re}(\lambda x + i\lambda y) = \lambda x = \lambda \operatorname{Re}(z)$ et de la même manière $\operatorname{Im}(\lambda z) = \lambda \operatorname{Im}(z)$.
- (Conjugué d'une somme)** le conjugué de $z + z' = (x + x') + i(y + y')$ avec $z = x + iy$ et $z' = x' + iy'$, $(x, x', y, y') \in \mathbf{R}^4$ est $(x + x') - i(y + y') = (x - iy) + (x' - iy')$ c'est donc aussi la somme des conjugués.
- (Partie réelle/imaginaire en fonction du conjugué)** calcul explicite direct : en sommant un complexe et son conjugué on fait disparaître la partie réelle, de même pour la partie imaginaire.
- Par exemple pour caractériser les réels, avec les mêmes notations que précédemment : $z = \bar{z}$ si et seulement si $x = x$ et $y = -y$ donc si et seulement si $y = 0$.

Définition ALG.1.4 | Complexe inverse

Si $z = x + iy \in \mathbf{C}^*$, alors on définit³ le complexe $\frac{1}{z} = \frac{1}{x+iy}$ comme étant :

$$\frac{1}{z} = \frac{x}{x^2 + y^2} + i \left(\frac{-y}{x^2 + y^2} \right).$$



Méthode Expression conjuguée

Dans la preuve précédente on a utilisé une technique classique pour obtenir l'inverse d'un nombre complexe écrit sous forme algébrique. On peut la résumer comme suit :

$$\frac{1}{x + iy} = \frac{x - iy}{(x + iy)(x - iy)} = \frac{x}{x^2 + y^2} + i \left(\frac{-y}{x^2 + y^2} \right).$$

2. FORME EXPONENTIELLE

Nous avons vu que \mathbf{C} et \mathbf{R}^2 sont deux ensembles très proches, et même en bijection. De la même manière qu'un point de \mathbf{R}^2 peut être repéré par ses coordonnées cartésiennes et polaires, un complexe peut être écrit en forme algébrique ou comme nous allons le voir de suite sous *forme trigonométrique*.

2.1. Module

Soient $z \in \mathbf{C}$, $x = \text{Re}(z)$ et $y = \text{Im}(z)$. Remarquons que :

$$z\bar{z} = (x + iy)(x - iy) = x^2 + y^2 \geq 0.$$

Ainsi, la quantité $z\bar{z}$ est réelle positive, on peut donc considérer sa racine carrée.

³Il s'agit bien d'une définition, bien entendu inspirée de la technique de l'« expression conjuguée »

Définition ALG.1.5 | Module

On appelle *module* de z le réel positif $\sqrt{z\bar{z}} = \sqrt{x^2 + y^2}$, que l'on note $|z|$.

Pour $z \in \mathbf{R}$, on retrouve la valeur absolue : $z = x \in \mathbf{R}$, $|z| = \sqrt{x^2} = |x|$. Notons aussi au passage la factorisation suivante :

$$x^2 + y^2 = (x + iy)(x - iy).$$

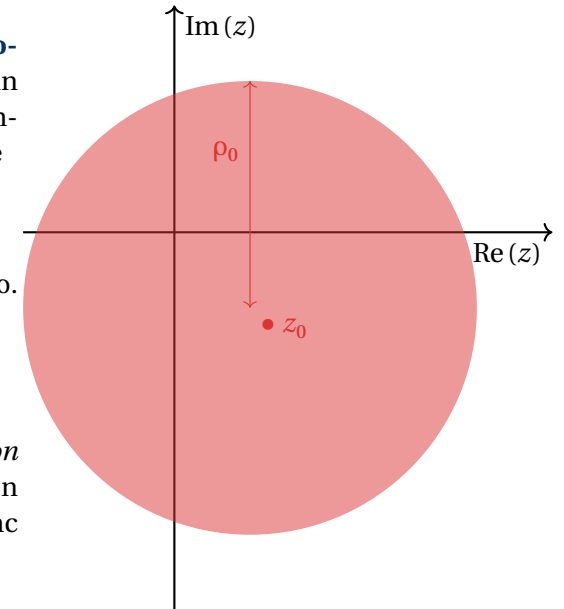
Remarque 2.1 – Interprétation géométrique du module Soient M_0 un point d'affixe z_0 et $\rho_0 \in \mathbf{R}_+^*$. Alors l'ensemble des points $M(z) \in \mathbf{R}^2$ tels que

$$|z - z_0| = \rho_0$$

est le *cercle de centre M_0 et de rayon ρ* . De même,

$$\{z \in \mathbf{C}, |z - z_0| \leq \rho_0\}$$

est le *disque de centre M_0 et de rayon ρ_0* . En particulier, prenant $z_0 = 0$ on obtient $|z| = d(O, M(z))$, c'est donc la distance entre $M(z)$ et l'origine.



Proposition ALG.1.2 | Propriétés du module

Soient $z, z' \in \mathbf{C}$.

1. (Multiplicativité du module) $|zz'| = |z| \cdot |z'|$. De plus si $z \neq 0$, alors

$$\frac{1}{z} = \frac{\bar{z}}{|z|^2}.$$

2. (Développement du module au carré)

$$|z + z'|^2 = |z|^2 + 2 \text{Re}(z\bar{z}') + |z'|^2.$$

3. (Expression du produit scalaire euclidien dans \mathbf{R}^2) Notons $u = (x, y)$, $u' = (x', y')$, avec $(x, x', y, y') \in \mathbf{R}^4$ tel que $z = x + iy$, $z' = x' + iy'$,

$$\langle u | u' \rangle = \operatorname{Re}(z\bar{z}')$$

Preuve Calculs directs.

Nous rappelons également sans démonstration l'inégalité triangulaire vue en première année.

Théorème ALG.1.1 | Inégalité triangulaire

1. Soient z et z' deux nombres complexes. Alors :

$$|z + z'| \leq |z| + |z'|.$$

De plus, l'égalité est réalisée si et seulement s'il existe $\lambda \in \mathbf{R}_+$ tel que $z' = \lambda z$ ou $z = \lambda z'$. On dit aussi que z et z' sont *positivement liés*.

2. (**Version généralisée**) Soit $(z_i)_{1 \leq i \leq n}$ une famille de complexes. Alors :

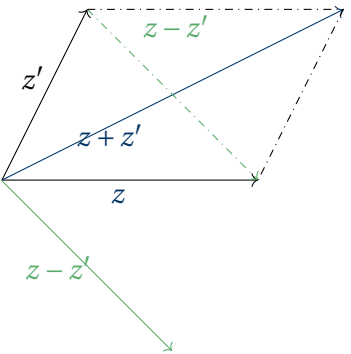
$$\left| \sum_{i=1}^n z_i \right| \leq \sum_{i=1}^n |z_i|.$$

Proposition ALG.1.3 | Identité du parallélogramme

Soit $(z, z') \in \mathbf{C}^2$, alors :

$$|z + z'|^2 + |z - z'|^2 = 2(|z|^2 + |z'|^2).$$

L'interprétation géométrique est la suivante : la somme des carrés des longueurs des diagonales d'un parallélogramme est la somme des carrés des longueurs des côtés. Ce résultat peut se retrouver avec le théorème de PYTHAGORE dans le cas d'un rectangle.



Preuve



Dans les résultats qui précèdent, nous avons souvent eu besoin de développer des modules de somme au carré. Cette formule de développement doit être bien comprise et mémorisée, elle résumée dans la méthode ci-après.



Méthode Développement d'une norme de somme au carré

Soit $|z + z'|^2$ avec $z, z' \in \mathbf{C}$.

1. Écrire la quantité en fonction du conjugué : $|z + z'|^2 = (z + z')(\overline{z + z'})$.
2. Développer.

En revanche, on oublie de suite le développement suivant.

Attention

On oublie la formule *archi*-fausse suivante :

$$|z + z'|^2 \neq |z|^2 + |z'|^2 + 2|z||z'|.$$

L'objectif est à présent d'arriver à la définition de la forme exponentielle d'un nombre complexe, on commence pour cela par définir :

- ▶ l'exponentielle d'un imaginaire pur,
- ▶ puis l'exponentielle générale d'un nombre complexe.

2.2. Exponentielle imaginaire

Notation Modulo

Soient $x, y, z \in \mathbf{C}$, alors « $x \equiv y [z]$ » signifie $x = y + kz$ pour un certain $k \in \mathbf{Z}$.

Définition ALG.1.6 | Nombre complexe $e^{i\theta}$

Pour tout $\theta \in \mathbf{R}$, on note $e^{i\theta}$ le nombre complexe de forme algébrique

$$\cos\theta + i \sin\theta.$$

Ce nombre est appelé *exponentielle imaginaire* de $\theta \in \mathbf{R}$. On note de plus $j = e^{\frac{2i\pi}{3}}$.

Attention

Le complexe j utilisé par les physiciens (pour éviter les confusions avec l'intensité électrique) est le complexe i de ce chapitre.

Pour l'instant, $e^{i\theta}$ n'est donc qu'une notation! Les propriétés de cette notation, qui permettront d'effectuer des calculs, sont données dans la proposition suivante. Par exemple, la quantité e^{1+3i} n'est pas encore définie. Mais si l'on note tout ceci avec une exponentielle c'est qu'elle va sûrement hériter des mêmes propriétés que l'exponentielle réelle connue depuis le lycée. Les voici.

Proposition ALG.1.4 | Propriétés de l'exponentielle imaginaire

Soient $\theta, \theta' \in \mathbf{R}$ deux réels. Alors :

1. $|e^{i\theta}| = 1, \quad \overline{e^{i\theta}} = e^{-i\theta} = \frac{1}{e^{i\theta}}.$
2. $e^{i\theta} = 1 \iff \theta \equiv 0 [2\pi] \iff \exists k \in \mathbf{Z}, \theta = 2k\pi.$
3. $e^{i\pi/2} = i, \quad e^{i\pi} = -1.$
4. $e^{i\theta} = e^{i\theta'} \iff \theta \equiv \theta' [2\pi] \iff \exists k \in \mathbf{Z}, \theta = \theta' + 2k\pi.$
5. Pour tout $(\theta, \theta') \in \mathbf{R}^2$, on a : $e^{i(\theta+\theta')} = e^{i\theta} \cdot e^{i\theta'}, \quad e^{i(\theta-\theta')} = \frac{e^{i\theta}}{e^{i\theta'}},$
6. **(Formule de MOIVRE)** $\forall \theta \in \mathbf{R}, \forall n \in \mathbf{Z}, \quad e^{ni\theta} = (e^{i\theta})^n.$

Preuve Voir le cours de première année.

Les formules ci-après paraissent anecdotiques au premier abord, mais elles seront d'intérêt capital pour toutes les applications des nombres complexes en trigonométrie.

Proposition ALG.1.5 | Formules d'EULER

$$\forall \theta \in \mathbf{R}, \quad \cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2} = \operatorname{Re}(e^{i\theta}) \quad \text{et} \quad \sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i} = \operatorname{Im}(e^{i\theta}).$$

Preuve Immédiate en utilisant les propriétés de parité de cos et sin.

Proposition ALG.1.6 | Relèvement

Tout complexe $z \in \mathbf{U} = \{z \in \mathbf{C}, |z| = 1\}$ peut s'écrire sous la forme $z = e^{i\theta}$ pour un certain $\theta \in \mathbf{R}$. Autrement dit, l'application

$$\begin{array}{l} \mathbf{R} \longrightarrow \mathbf{U} \\ \theta \longrightarrow e^{i\theta} \end{array} \quad \text{est surjective.}$$

Preuve Admis.

2.3. Exponentielle générale e^z et forme exponentielle

Nous avons vu dans la partie précédente que $e^{i\theta}$ est le complexe $\cos \theta + i \sin \theta$. Il est de module un. Maintenant nous allons chercher à exprimer les autres nombres complexes en fonction de cette exponentielle.

Définition ALG.1.7 | Argument d'un nombre complexe

Soit z un nombre complexe **non nul**.

Alors $\frac{z}{|z|} \in \mathbf{U}$, donc il existe $\theta \in \mathbf{R}$ tel que $\frac{z}{|z|} = e^{i\theta}$. Un tel réel θ est appelé *un argument* de z et il est noté $\text{Arg}(z)$. L'ensemble des arguments de z est alors $\theta + 2\pi\mathbf{Z} = \{\theta + 2k\pi, k \in \mathbf{Z}\}$, on note cela

$$\text{Arg}(z) \equiv \theta \pmod{2\pi}.$$

On utilise en général pas le signe « \equiv » à cause de la non-unicité de l'argument.

Définition ALG.1.8 | Forme exponentielle

Soit z un nombre complexe non nul. L'écriture

$$z = |z| e^{i\text{Arg}(z)}$$

est appelée *forme exponentielle* (ou *forme trigonométrique*) de z .

Avant de poursuivre, on introduit un complexe relativement usuel donné sous forme exponentielle, noté j^4 , qui interviendra dans une prochaine section.

Notation Complexe j

On note généralement $j = e^{\frac{2i\pi}{3}} = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$.

Proposition ALG.1.7 | Propriétés de j

- ▶ $j^3 = 1, \quad \bar{j} = j^2,$
- ▶ $1 + j + j^2 = 0.$

⁴attention, il ne s'agit pas « $du j$ » utilisé par les physiciens, qui quant à lui désigne notre i — eh oui, la vie est mal faite!

Preuve



Méthode Mettre sous forme exponentielle un nombre complexe

Soit $z \neq 0$.

1. Calculer $|z|$, puis $\frac{z}{|z|}$.
2. Chercher $\theta \in [0, 2\pi[$ tel que : $\frac{z}{|z|} = e^{i\theta}$, *i.e.* tel que


$$\cos(\theta) = \frac{\text{Re}(z)}{|z|}, \quad \sin(\theta) = \frac{\text{Im}(z)}{|z|}.$$

La forme exponentielle est alors : $z = |z| e^{i\theta}$. Il arrive parfois que l'angle θ


Méthode Technique de l'angle moitié (forme trigonométrique d'une somme d'exponentielles imaginaires)

Soient deux nombres complexes z, z' de module un donnés sous forme trigonométrique : $z = e^{i\theta}, z' = e^{i\theta'}$ avec $(\theta, \theta') \in [0, 2\pi[$. Alors la forme trigonométrique de $z + z'$ s'obtient par le calcul suivant :

$$\begin{aligned} z + z' &= e^{i\theta} + e^{i\theta'} = e^{i\frac{\theta+\theta'}{2}} \left(e^{i\frac{\theta-\theta'}{2}} + e^{-i\frac{\theta-\theta'}{2}} \right) \\ &= 2e^{i\frac{\theta+\theta'}{2}} \cos\left(\frac{\theta-\theta'}{2}\right). \end{aligned}$$

 La méthode s'adapte à $z - z'$ en faisant apparaître un sinus. On obtient alors facilement module et argument :

$$|z + z'| = 2 \left| \cos \left(\frac{\theta - \theta'}{2} \right) \right|, \quad \text{Arg}(z + z') \equiv \frac{\theta + \theta'}{2} \quad [2\pi].$$

Exemple 1 — Déterminer module et argument de $1 + e^{i\theta}$ avec $\theta \in [0, \pi[$. Que dire si $\theta \in [\pi, 2\pi[$?  On met en facteur l'angle moitié, i.e. $e^{i\frac{\theta}{2}}$. Nous avons alors :

$$1 + e^{i\theta} = e^{i\frac{\theta}{2}} \left(e^{-i\frac{\theta}{2}} + e^{i\frac{\theta}{2}} \right) = 2 \cos \left(\frac{\theta}{2} \right) e^{i\frac{\theta}{2}}.$$

Alors $|1 + e^{i\theta}| = \left| 2 \cos \left(\frac{\theta}{2} \right) \right| = 2 \cos \left(\frac{\theta}{2} \right)$ puisque $\theta \in [0, \pi[$. Donc la forme trigonométrique de $1 + e^{i\theta}$ est

$$1 + e^{i\theta} = \boxed{2 \cos \left(\frac{\theta}{2} \right) e^{i\frac{\theta}{2}}}.$$

Si $\theta \in [\pi, 2\pi[$, alors $|1 + e^{i\theta}| = \left| 2 \cos \left(\frac{\theta}{2} \right) \right| = -2 \cos \left(\frac{\theta}{2} \right)$. Donc la forme exponentielle est :

$$\begin{aligned} 1 + e^{i\theta} &= \left(-2 \cos \left(\frac{\theta}{2} \right) \right) \left(-e^{i\frac{\theta}{2}} \right) \\ &= \left(-2 \cos \left(\frac{\theta}{2} \right) \right) \left(e^{i\pi} e^{i\frac{\theta}{2}} \right) \\ &= \boxed{\left(-2 \cos \left(\frac{\theta}{2} \right) \right) \left(e^{i\frac{\theta+2\pi}{2}} \right)}. \end{aligned}$$

Passons à quelques propriétés de l'argument d'un nombre complexe, qui découlent des propriétés déjà établies sur l'exponentielle imaginaire.

Proposition ALG.1.8 | Propriétés de l'argument

Soient z et z' deux complexes non nuls. Alors :

$$1. \text{Arg}(zz') \equiv \text{Arg}(z) + \text{Arg}(z') \quad [2\pi],$$

$$\begin{aligned} 2. \text{Arg} \left(\frac{z}{z'} \right) &\equiv \text{Arg}(z) - \text{Arg}(z') \quad [2\pi], \\ 3. \text{Arg}(\bar{z}) &\equiv -\text{Arg}(z) \quad [2\pi]. \end{aligned}$$

Preuve Découlent essentiellement des propriétés établies sur $e^{i\theta}$ avec $\theta \in \mathbf{R}$.

Proposition ALG.1.9 | Caractérisation de l'égalité de nombres complexes

Soient z, z' deux nombres complexes. Alors :

$$\begin{aligned} z = z' &\iff \begin{cases} \text{Re}(z) = \text{Re}(z') \\ \text{Im}(z) = \text{Im}(z') \end{cases} \\ &\iff \begin{cases} |z| = |z'| \\ \text{Arg}(z) \equiv \text{Arg}(z') \quad [2\pi]. \end{cases} \end{aligned}$$

Remarque 2.2 — Comment choisir la forme à utiliser? Lorsque l'on cherche à démontrer un résultat sur des nombres complexes, il ne faut **pas** systématiquement l'écrire sous forme algébrique :

$$z = \text{Re}(z) + i \text{Im}(z).$$

Cette forme est adaptée aux problèmes « additifs », où ce qui intervient est plutôt des sommes (ou plus généralement des combinaisons linéaires) de complexes. Les problèmes « multiplicatifs » se résolvent mieux en utilisant la forme exponentielle lorsque $z \neq 0$:

$$z = |z| e^{i \text{Arg}(z)}.$$

Définition ALG.1.9 | Exponentielle d'un complexe quelconque

Soit $z \in \mathbf{C}$. On définit l'*exponentielle* de z comme étant le complexe noté e^z suivant :

$$e^z = e^{\text{Re}(z)} e^{i \text{Im}(z)}.$$

Remarque 2.3 — Attention à bien comprendre la nature des exponentielles ci-dessus. Comme $\operatorname{Re}(z) \in \mathbf{R}$, la première est l'exponentielle réelle définie au lycée comme unique solution de l'équation différentielle $y' = y, y(0) = 1$. La seconde, en revanche, a été définie dans la sous-section 2.2.

En tenant compte des propriétés de l'exponentielle réelle d'une part, et de l'exponentielle imaginaire d'autre part, on obtient la proposition suivante.

Proposition ALG.1.10 | Propriétés de l'exponentielle complexe

1. Pour tout $(z, z') \in \mathbf{C}^2$, on a : $e^{z+z'} = e^z \cdot e^{z'}$, $e^{z-z'} = \frac{e^z}{e^{z'}}$.
2. Pour tout $z \in \mathbf{C}$, on a : $e^z = 1 \iff z \in 2i\pi\mathbf{Z}$.
3. Enfin, tout complexe non nul $Z \in \mathbf{C}^*$ peut s'écrire sous la forme $Z = e^z$ pour un certain $z \in \mathbf{C}$. Autrement dit, l'application $\begin{array}{l} \mathbf{C} \longrightarrow \mathbf{C}^* \\ z \longmapsto e^z \end{array}$ est surjective.

Attention Non-existence d'un logarithme complexe

La dernière assertion ne dit pas que $z \mapsto e^z$ est une bijection, nous n'avons donc pas construit de logarithme complexe.⁵

Preuve



Exemple 2 — Soit $z \in \mathbf{C}$.

1. Calculer module, argument, partie réelle et partie imaginaire de e^z .
2. Déterminer ensuite l'ensemble $\{z \in \mathbf{C}, e^z = 1\}$. Existe-t-il un logarithme complexe, i.e. $\exp : z \in \mathbf{C} \mapsto e^z$ est-elle une bijection ?



⁵Il en existe en fait une infinité, mais cela dépasse largement le cadre de notre programme.

2.4. Complément – racines n -ièmes d'un complexe [H.P]

Cette sous-section porte le label [H.P], cependant elle est très classique, et tombe régulièrement dans les sujets d'écrits.

De manière générale, on appelle « racine n -ième » d'un objet mathématique une quantité qui élevée à la puissance n donne cet objet (l'objet en question peut être un réel, un complexe ou même une matrice).

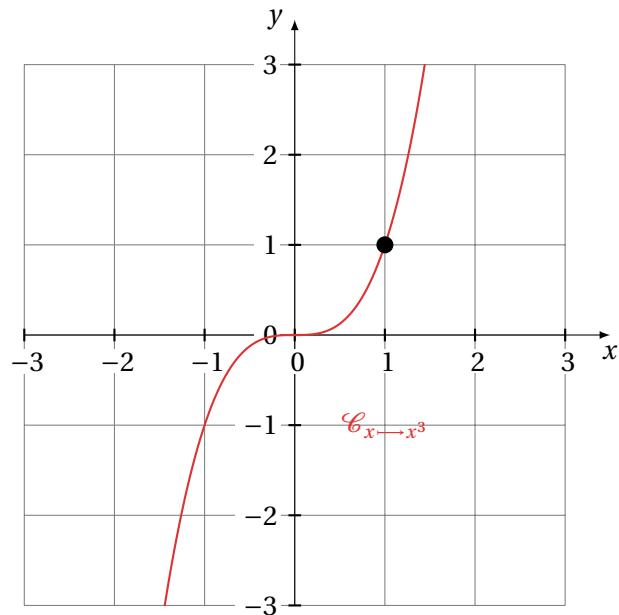
On résume les différentes propriétés, avantages et inconvénients de chacune des deux formes. La forme exponentielle est donc

	Forme cartésienne	Forme exponentielle
Lien	$x = r + iy$	$z = re^{i\theta}$
Égalité	$z = z' \iff x = x', y = y'$	$z = z' \iff r = r', \theta \equiv \theta' \pmod{2\pi}$
Somme	$z + z' = (x + x') + i(y + y')$	$\bar{z} = re^{-i\theta}$
Produit	$z \times z' = (xx' - yy') + i(xy' + x'y)$	$\tan \theta = \frac{y}{x} \quad \cos \theta = \frac{x}{r} \quad \sin \theta = \frac{y}{r}$
Puissance	...	$z \times z' = rr'e^{i(\theta+\theta')}$
Inverse	$\frac{1}{z} = \frac{\bar{z}}{z\bar{z}} = \frac{x-iy}{x^2+y^2}$	$z^n = r^n e^{in\theta}$
Quotient	...	$\frac{1}{z} = \frac{1}{r} e^{-i\theta}$

Regardons pour commencer un exemple. Notons $\mathbf{K} = \mathbf{R}$ ou \mathbf{C} et considérons l'équation $z^3 = 1$ avec $z \in \mathbf{K}$.

1. Si $\mathbf{K} = \mathbf{R}$, l'équation n'admet qu'une solution : 1.
2. Si $\mathbf{K} = \mathbf{C}$, on voit que $j = e^{2i\pi/3}$ convient, mais aussi $j^2 = e^{4i\pi/3}$ – et en fait nous allons montrer que ce sont les seules.

On constate que : l'ensemble des racines cubiques complexes de 1 contient l'ensemble des racines cubiques réelles de 1, et il y en a systématiquement au moins autant dans \mathbf{C} que dans \mathbf{R} .



Passons à présent au cas général.

Définition ALG.1.10 | Racines n -ième

Soit $n \in \mathbf{N}^*$ et $\alpha \in \mathbf{C}$. On appelle *racine n -ième* de α tout complexe $z \in \mathbf{C}$ tel que $z^n = \alpha$, i.e. une racine du polynôme $X^n - \alpha$.

- ▶ On note $\mathbf{U}_n(\alpha)$ l'ensemble des racines de $X^n - \alpha$. Si $\alpha = 1$, on parle de *racine n -ième de l'unité*, i.e. les racines de $X^n - 1$. On notera \mathbf{U}_n l'ensemble de ces complexes.
- ▶ Si $n = 2$, on parle de *racine carrée* de α , pour $n = 3$ de *racine cubique*.

Il s'agit d'un problème multiplicatif (avec des puissances), donc la bonne forme à adopter est la forme exponentielle, nous allons très largement nous en servir dans la suite.

Notation

- ▶ Les notations $\sqrt{\alpha}$ et $\sqrt[n]{\alpha}$ sont réservées à $\alpha \in \mathbf{R}_+$ (ou bien $\alpha \in \mathbf{R}$ si n est impair).



Les notations $\sqrt{\alpha}$ et $\sqrt[n]{\alpha}$ où $\alpha \in \mathbf{C} \setminus \mathbf{R}$ sont **interdites** (elles n'ont aucun sens car il n'y a pas unicité).

Que nous donne la théorie des polynômes? Le polynôme $X^n - \alpha$ étant de degré n , nous savons qu'il possède au plus n racines (voir le **Chapter ALG.2** pour plus de détails). On peut même encore préciser : comme $(X^n - \alpha)' = nX^{n-1}$ et que 0 n'est pas racine, on sait que toutes les racines de $X^n - \alpha$ sont simples. En conclusion : nous avons alors accès au nombre de racines n -ièmes de l'unité

$$\# \mathbf{U}_n = n.$$

On va donc déterminer à présent explicitement les racines d'un complexe α .

Théorème ALG.1.2 | Racines n -ième d'un complexe α [H.P]

Soient $n \in \mathbf{N} \setminus \{0\}$ et $\alpha \in \mathbf{C}^*$ un complexe non nul, $\rho = |\alpha|$ et θ un argument de α . L'ensemble des racines n -ièmes complexes de α est

$$\mathbf{U}_n(\alpha) = \left\{ \boxed{\rho^{\frac{1}{n}} \cdot e^{i \frac{\theta}{n}}} \times e^{\frac{2i\pi}{n}k}, k \in \llbracket 0, n-1 \rrbracket \right\}.$$

une racine n -ième de α

En faisant $\alpha = 1$ (i.e. $\rho = 1, \theta = 0$) dans l'énoncé précédent, nous déduisons le corollaire ci-après.

Corollaire ALG.1.1 | Racines n -ième de l'unité [H.P]

Soit $n \in \mathbf{N} \setminus \{0\}$. L'ensemble des racines n -ièmes de l'unité (i.e. de 1) est :

$$\mathbf{U}_n = \left\{ e^{\frac{2i\pi}{n}k}, k \in \llbracket 0, n-1 \rrbracket \right\}.$$

En particulier :

$$\mathbf{U}_2 = \{1, -1\} \quad \text{et} \quad \mathbf{U}_3 = \{1, j, j^2\} = \{1, j, \bar{j}\}.$$

On retrouve bien entendu le fait qu'il y a n racines n -ièmes, que ce soit pour un complexe non nul ou bien 1.

Attention

Les calculs doivent être refaits à chaque fois pour le n considéré, et vous ne pouvez pas vous servir de ce théorème tel quel.

Méthode Calculs de racines n -ième de complexes


On cherche donc les solutions de $z^n = \alpha$ avec $\alpha \neq 0$ (si $\alpha = 0$ il n'y a que zéro comme solution).

1. Calculer la forme trigonométrique de $\alpha = \rho e^{i\theta}$.
2. Chercher z sous la forme $z = \rho' e^{i\theta'}$.
3. En remplaçant, on obtient comme conditions $(\rho')^n = \rho$ et $n\theta' = \theta + 2k\pi$ avec $k \in \mathbf{Z}$. Résoudre ces deux équations puis conclure.

Dans la pratique, la méthode sous-entend que l'on est capable de trouver une racine n -ième de α .

Preuve (Point clef — Chercher les solutions sous forme exponentielle)


**Exemple 3** — Structure géométrique

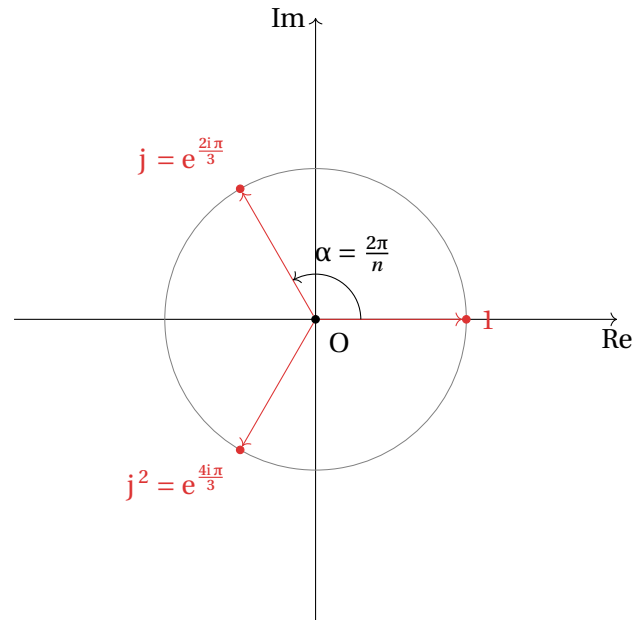
1. Déterminer \mathbf{U}_4 . On fera le calcul, puis on contrôlera le résultat avec l'énoncé précédent. 

2. Dessiner les points géométriques d'affixes les éléments de \mathbf{U}_2 , \mathbf{U}_3 et \mathbf{U}_4 . Que remarque-t-on? 

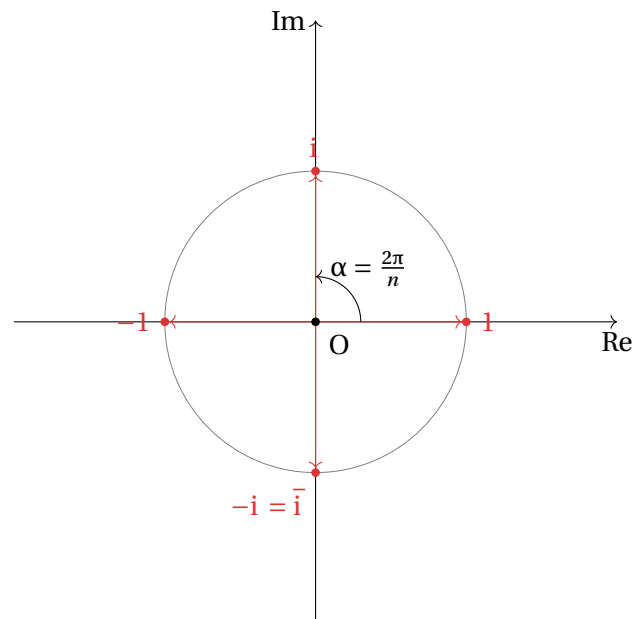
VISUALISATION GRAPHIQUE DES RACINES DE L'UNITÉ On propose deux situations, celui des racines troisièmes et quatrièmes.

3. Déterminer les racines quatrièmes de -16 . 

4. Déterminer les racines cubiques de $1 + i$. 



(a) Racines cubiques



(b) Racines quatrièmes

3. TRIGONOMÉTRIE & APPLICATIONS DES NOMBRES COMPLEXES

Les applications sont très nombreuses, nous ne précisons ici qu'une infime partie d'entre elles. Notamment celles en trigonométrie, ceci est dû à l'exponentielle complexe qui fournit un pont entre les deux domaines. Ces applications trigonométriques ont été largement utilisées dans de nombreux domaines, notamment en physique pour l'étude de circuits électriques en régime sinusoïdal forcé.

3.1. En trigonométrie

3.1.1. Formulaire

On rappelle dans cette sous-section les formules principales de trigonométrie à très bien connaître, ainsi qu'une roue de secours pour certaines d'entre elles faisant intervenir les nombres complexes. Les définitions de cos, sin, tan étant supposées connues.

Proposition ALG.1.11

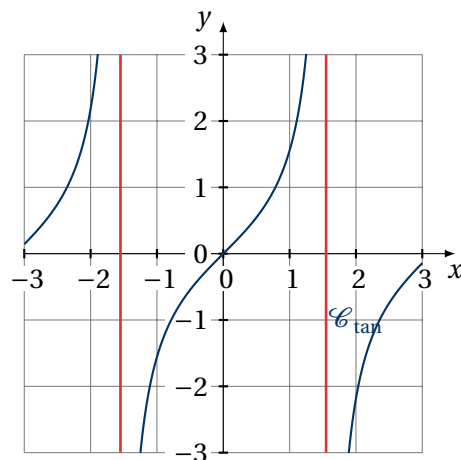
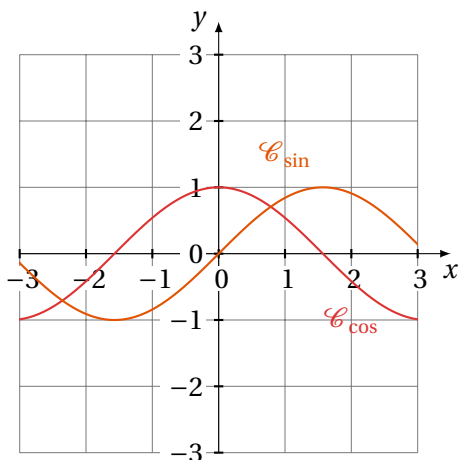
Les fonctions sinus et cosinus sont définies et dérivables sur \mathbf{R} et 2π -périodiques. La fonction sinus est impaire, et la fonction cosinus est paire. De plus :

$$\forall x \in \mathbf{R}, \quad \cos'(x) = -\sin(x), \quad \sin'(x) = \cos(x).$$

Proposition ALG.1.12

La fonction tangente est définie et dérivable sur $\mathcal{D}_{\tan} = \mathbf{R} \setminus \left\{ \frac{\pi}{2} + k\pi, k \in \mathbf{Z} \right\}$ et est π -périodique. La fonction tangente est impaire. De plus,

$$\forall x \in \mathcal{D}_{\tan}, \quad \tan'(x) = 1 + \tan^2(x) = \frac{1}{\cos^2(x)}.$$



Par simple lecture graphique sur le cercle trigonométrique (cf. le cours de 1ère année), nous déduisons les résultats ci-après.

Proposition ALG.1.13 | Résolution d'équations

1. Soit $(x, y) \in \mathbb{R}^2$.

$$\cos(x) = \cos(y) \iff \exists k \in \mathbb{Z}, (x = y + 2k\pi) \text{ ou } (x = -y + 2k\pi).$$

2. Soit $(x, y) \in \mathbb{R}^2$.

$$\sin(x) = \sin(y) \iff \exists k \in \mathbb{Z}, (x = y + 2k\pi) \text{ ou } (x = \pi - y + 2k\pi).$$

3. Soit $(x, y) \in \mathcal{D}_{\tan}^2$,

$$\tan(x) = \tan(y) \iff \exists k \in \mathbb{Z}, (x = y + k\pi).$$

Exemple 4 — Résoudre $\sin x = \cos x$ en $x \in \mathbb{R}$.

Proposition ALG.1.14 | Valeurs remarquables

x	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
$\cos(x)$	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0
$\sin(x)$	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1
$\tan(x)$	0	$\frac{\sqrt{3}}{3}$	1	$\sqrt{3}$	

Proposition ALG.1.15 | Formules d'addition

Soient $x, y \in \mathbb{R}$.

- ▶ $\cos(x \pm y) = \cos(x) \cos(y) \mp \sin(x) \sin(y)$,
- ▶ $\sin(x \pm y) = \sin(x) \cos(y) \pm \sin(y) \cos(x)$.

En cas de trou de mémoire sur l'une des formules ci-après, toujours refaire un dessin de cercle trigonométrique.

Corollaire ALG.1.2 | Formules de transformation d'angles

Soit $x \in \mathbb{R}$.

- | | |
|-------------------------------|--|
| ▶ $\cos(-x) = \cos(x),$ | ▶ $\cos\left(\frac{\pi}{2} - x\right) = \sin(x),$ |
| ▶ $\sin(-x) = -\sin(x),$ | ▶ $\sin\left(\frac{\pi}{2} - x\right) = \cos(x),$ |
| ▶ $\cos(\pi - x) = -\cos(x),$ | ▶ $\cos\left(\frac{\pi}{2} + x\right) = -\sin(x),$ |
| ▶ $\sin(\pi - x) = \sin(x),$ | ▶ $\sin\left(\frac{\pi}{2} + x\right) = \cos(x).$ |
| ▶ $\cos(\pi + x) = -\cos(x),$ | |
| ▶ $\sin(\pi + x) = -\sin(x),$ | |

Et enfin, les cas particuliers $x = y$ dans les formules d'addition fournissent les formules dites de duplication.

Corollaire ALG.1.3 | Formules de duplication / anti-linéarisation

Soit $x \in \mathbf{R}$.

- ▶ $\cos(2x) = \cos^2(x) - \sin^2(x) = 1 - 2\sin^2(x) = 2\cos^2(x) - 1.$
- ▶ $\sin(2x) = 2\cos(x)\sin(x).$

Ou encore, de manière équivalente, nous avons le corollaire suivant.

Corollaire ALG.1.4 | Formules de linéarisation

Soit $x \in \mathbf{R}$. Alors :

- ▶ $\cos^2(x) = \frac{1 + \cos(2x)}{2}$
- ▶ $\sin^2(x) = \frac{1 - \cos(2x)}{2}.$

Une application importante de ces formules est le calcul d'intégrales de la forme $\int_a^b \cos^2(t) dt, \int_a^b \sin^2(t) dt$ avec $a < b$ deux réels. Des exemples seront faits dans le [Chapter ANA.11](#). On obtient immédiatement le corollaire ci-après.

Une dernière conséquence des formules d'addition est la transformation d'expressions trigonométriques en cos ou sin.

Méthode Écriture d'une combinaison linéaire de fonctions trigonométriques sous « forme déphasée »



Soient $a, b, x \in \mathbf{R}$. On souhaite transformer l'expression $E(x) = a \cos x + b \sin x$ en $\rho \cos(x + \varphi)$, avec $\rho \in \mathbf{R}^+, \varphi \in \mathbf{R}$.

1. Mettre $\rho = \sqrt{a^2 + b^2}$ en facteur, de sorte que

$$E(x) = \rho \left(\frac{a}{\rho} \cos x + \frac{b}{\rho} \sin x \right).$$

2. Comme $\left(\frac{a}{\rho}, -\frac{b}{\rho}\right)$ est sur le cercle unité, puisque $\left(\frac{a}{\rho}\right)^2 + \left(-\frac{b}{\rho}\right)^2 = 1$, il existe $\varphi \in [0, 2\pi[$ tel que :

$$\frac{a}{\sqrt{a^2 + b^2}} = \cos \varphi, \quad -\frac{b}{\sqrt{a^2 + b^2}} = \sin \varphi.$$

3. Alors $E(x) = \cos x \cos \varphi - \sin x \sin \varphi = \cos(x + \varphi)$.
Une méthode analogue existe si l'on souhaite une forme déphasée de la forme $\rho \sin(x + \varphi)$, il suffit de choisir l'angle différemment.

Remarque 3.1 – Interprétation La somme deux signaux sinusoïdaux est encore un signal sinusoïdal.

Exemple 5 – Écrire sous forme d'un sinus puis d'un cosinus l'expression $\sqrt{2} \cos \theta + \sqrt{6} \sin \theta$ avec $\theta \in \mathbf{R}$.

3.2. Techniques calculatoires

Les nombres complexes, grâce à l'exponentielle complexe, fournissent une méthode beaucoup plus efficace que les formules précédentes pour linéariser et anti-linéariser des expressions trigonométriques compliquées. Rappelons la méthode.



Méthode Linéarisation & Antilinéarisation avec des complexes

1. (Pour linéariser $\cos^k \theta, \sin^k \theta$) écrire


$$\cos^k \theta = \left(\frac{e^{i\theta} + e^{-i\theta}}{2} \right)^k, \quad \sin^k \theta = \left(\frac{e^{i\theta} - e^{-i\theta}}{2i} \right)^k,$$

puis développer avec le binôme, regrouper les termes avec leur conjugué, utiliser les formules d'EULER.

2. (Pour antilinéariser $\cos(k\theta), \sin(k\theta)$) écrire

$$\begin{aligned} \cos(k\theta) &= \operatorname{Re} \left(e^{ik\theta} \right) \underset{\text{MOIVRE}}{=} \operatorname{Re} \left(\left(e^{i\theta} \right)^k \right) = \operatorname{Re} \left((\cos \theta + i \sin \theta)^k \right), \\ \sin(k\theta) &= \operatorname{Im} \left(e^{ik\theta} \right) \underset{\text{MOIVRE}}{=} \operatorname{Im} \left(\left(e^{i\theta} \right)^k \right) = \operatorname{Im} \left((\cos \theta + i \sin \theta)^k \right), \end{aligned}$$

puis développer avec le binôme et calculer les parties réelles et imaginaires.

Exemple 6 – Linéarisation Soit $x \in \mathbf{R}$. Linéariser $\cos^2 x$ et $\sin^3 x$ en utilisant les nombres complexes 

Exemple 7 – Anti-Linéarisation Soit $x \in \mathbf{R}$. Exprimer $\cos(4x)$ en fonction de $\cos x, \sin x$, en utilisant les nombres complexes. 

Les complexes peuvent rendre de multiples services en trigonométrie, y compris les calculs de sommes de fonctions trigonométriques, comme le montrent les exemples ci-après.



Méthode Calculs de sommes trigonométriques

1. Écrire \cos , \sin comme des parties réelles/imaginaires d'exponentielles complexes.
2. Utiliser la linéarité de $\operatorname{Re}(\dots)$, $\operatorname{Im}(\dots)$, *i.e.* : $\operatorname{Re}(\sum \dots) = \sum \operatorname{Re}(\dots)$, $\operatorname{Im}(\sum \dots) = \sum \operatorname{Im}(\dots)$.
3. Utiliser la formule donnant la somme de termes géométriques. Conclure.

Exemple 8 – Calculer pour tout $n \in \mathbf{N}$ les sommes $\sum_{k=0}^n \cos(kx)$ et $\sum_{k=0}^n \sin(kx)$ pour

$x \in \mathbf{R}$. 

Exemple 9 – Calculer pour tout $n \in \mathbf{N}$ la somme $\sum_{k=0}^n \frac{\cos kx}{\cos^k x}$ pour $x \in \mathbf{R}$ tel que $\cos x \neq 0$

0.

$$\begin{aligned}
 \sum_{k=0}^n \frac{\cos kx}{\cos^k x} &= \sum_{k=0}^n \frac{\operatorname{Re}(e^{ikx})}{(\cos x)^k} \\
 &= \sum_{k=0}^n \operatorname{Re}\left(\frac{e^{ikx}}{(\cos x)^k}\right) \\
 &= \operatorname{Re}\left(\sum_{k=0}^n \left(\frac{e^{ix}}{\cos x}\right)^k\right) \\
 &= \operatorname{Re}\left(\frac{1 - \left(\frac{e^{ix}}{\cos x}\right)^{n+1}}{1 - \frac{e^{ix}}{\cos x}}\right) \\
 &= \operatorname{Re}\left(\frac{1 - \frac{e^{i(n+1)x}}{\cos^{n+1} x}}{1 - \frac{e^{ix}}{\cos x}}\right) \\
 &= \operatorname{Re}\left(\frac{\frac{\cos^{n+1} x - e^{i(n+1)x}}{\cos^{n+1} x}}{\frac{\cos x - e^{ix}}{\cos x}}\right) \\
 &= \frac{1}{\cos^n x} \operatorname{Re}\left(\frac{\cos^{n+1} x - e^{i(n+1)x}}{\cos x - e^{ix}}\right) \\
 &= \frac{1}{\cos^n x} \operatorname{Re}\left(\frac{\cos^{n+1} x - (\cos(n+1)x + i \sin(n+1)x)}{\cos x - (\cos x + i \sin x)}\right) \\
 &= \frac{1}{\cos^n x} \operatorname{Re}\left(\frac{(\cos^{n+1} x - \cos(n+1)x) - i \sin(n+1)x}{-i \sin x}\right) \\
 &= \boxed{\frac{1}{\cos^n x} \cdot \frac{\sin(n+1)x}{\sin x}}.
 \end{aligned}$$

$(\cos x)^k$ est réel
 la partie réelle est linéaire
 somme de termes d'une suite géométrique, car $\cos x \neq e^{ix}$ pour les x qui nous intéressent
 réduction au même dénominateur
 $\frac{1}{\cos^n(x)} \in \mathbf{R}$
 forme algébrique des nombres complexes

4. EXERCICES

Exercice ALG.1.1 | Vrai ou Faux?

1. Soit $z \in \mathbf{C}$, la partie imaginaire de iz est égale à celle de z .
2. Soient $z_1, z_2 \in \mathbf{C}$. Alors $|z_1 - z_2| \leq |z_1| - |z_2|$.
3. Soit $z \in \mathbf{C}$. Alors le conjugué de e^z est $e^{\bar{z}}$.

4.1. Généralités

Exercice ALG.1.2 | Soit $\theta \in \mathbf{R} \setminus \left(\frac{\pi}{2} + k\pi, k \in \mathbf{Z}\right)$. Déterminer le module et un argument de $\frac{1}{1+i \tan \theta}$.

Solution (exercice ALG.1.2)

Il s'agit de trouver une forme exponentielle pour $1+i \tan \theta$. Nous avons $1+i \tan \theta = \sqrt{1+\tan^2 \theta} \left(\frac{1}{\sqrt{1+\tan^2 \theta}} + \frac{i \tan \theta}{\sqrt{1+\tan^2 \theta}} \right)$. Or, $1+\tan^2 \theta = \frac{1}{\cos^2 \theta}$, donc

$$1+i \tan \theta = \frac{1}{|\cos \theta|} (|\cos \theta| + i |\cos \theta| \tan \theta).$$

- si $\theta \in \left] -\frac{\pi}{2} + k\pi, \frac{\pi}{2} + k\pi \right[$ pour un certain $k \in \mathbf{Z}$. Alors

$$\begin{aligned} 1+i \tan \theta &= \frac{1}{\cos \theta} (\cos \theta + i \cos \theta \tan \theta) \\ &= \frac{1}{\cos \theta} (\cos \theta + i \sin \theta) = \boxed{\frac{1}{\cos \theta} e^{i\theta}}. \end{aligned}$$

En passant à l'inverse, on déduit alors :

$$\boxed{\left| \frac{1}{1+i \tan \theta} \right| = \cos \theta, \quad \text{Arg} \left(\frac{1}{1+i \tan \theta} \right) = -\theta \quad [2\pi].}$$

- si $\theta \in \left] \frac{\pi}{2} + k\pi, \frac{3\pi}{2} + k\pi \right[$ pour un certain $k \in \mathbf{Z}$. Alors

$$\begin{aligned} 1+i \tan \theta &= -\frac{1}{\cos \theta} (-\cos \theta - i \cos \theta \tan \theta) \\ &= -\frac{1}{\cos \theta} (\cos(\theta + \pi) + i \sin(\theta + \pi)) = \boxed{-\frac{1}{\cos \theta} e^{i(\theta + \pi)}}. \end{aligned}$$

En passant à l'inverse, on déduit alors :

$$\boxed{\left| \frac{1}{1+i \tan \theta} \right| = -\cos \theta, \quad \text{Arg} \left(\frac{1}{1+i \tan \theta} \right) = -\theta - \pi \quad [2\pi].}$$

Attention à ne pas simplifier le signe moins dans ce second cas, $1/\cos \theta$ étant négatif ce n'est pas le module.

Exercice ALG.1.3 | Soient $n \in \mathbf{N}$ et $x \in]-\pi, \pi[$. Déterminer le module et un argument de : $\alpha = (1 + \cos x + i \sin x)^n$ et $\beta = \left(\frac{1+i\sqrt{3}}{1-i} \right)^n$.

Solution (exercice ALG.1.3)

Soient $n \in \mathbf{N}$ et $x \in]-\pi, \pi[$. Il s'agit de déterminer les formes exponentielles des deux complexes.

$$\begin{aligned} 1 + \cos x + i \sin x &= 1 + e^{ix} \\ &= e^{i\frac{x}{2}} \left(e^{-i\frac{x}{2}} + e^{i\frac{x}{2}} \right) \\ &= e^{i\frac{x}{2}} 2 \cos \left(\frac{x}{2} \right) \quad \text{formules d'Euler.} \end{aligned}$$

technique de l'angle moitié

Attention, ici nous n'obtenons pas forcément le module et l'argument, cela dépend du signe de $2 \cos \left(\frac{x}{2} \right)$. Mais comme $x \in]-\pi, \pi[$, alors $2 \cos \left(\frac{x}{2} \right) \geq 0$, on déduit alors la forme trigonométrique de α :

$$\alpha = 2^n \cos^n \left(\frac{x}{2} \right) e^{i \frac{nx}{2}}.$$

L'argument est donc $2^n \cos^n\left(\frac{x}{2}\right)$, et un module est $\frac{nx}{2}$. Pour le second, on met sous forme trigonométrique :

$$\begin{aligned} \beta^n &= \left(\frac{1+i\sqrt{3}}{1-i}\right)^n \\ &= \left(2e^{i\frac{\pi}{3}}\sqrt{2}e^{-i\frac{\pi}{4}}\right)^n \\ &= \boxed{2^{\frac{n}{2}}e^{i\frac{n\pi}{12}}}. \end{aligned}$$

Le module est $2^{\frac{n}{2}}$, un argument est $\frac{n\pi}{12}$.

Exercice ALG.1.4 | Soient A et B deux points distincts du plan, d'affixes respectives a et b . Montrer qu'un point M d'affixe z appartient au cercle Γ de diamètre [AB] si et seulement si : $2z\bar{z} - (\bar{a} + \bar{b})z - (a + b)\bar{z} + a\bar{b} + \bar{a}b = 0$.

Solution (exercice ALG.1.4)

L'appartenance au cercle en question s'exprime à l'aide de l'affixe du centre et d'un module. L'affixe du centre du cercle est $\frac{a+b}{2}$, et la distance d'un point M(z) au centre est

$$\left|z - \frac{a+b}{2}\right|.$$

Ainsi, la condition d'appartenance au cercle est la suivante :

$$\begin{aligned} M(z) \in \Gamma &\iff \left|z - \frac{a+b}{2}\right| = \frac{|b-a|}{2} \\ &\iff \left|z - \frac{a+b}{2}\right|^2 = \frac{|b-a|^2}{4} \quad \text{élévation au carré} \\ &\iff \left(z - \frac{a+b}{2}\right)\overline{\left(z - \frac{a+b}{2}\right)} = \frac{1}{4}(b-a)\overline{(b-a)} \\ &\iff (2z - a + b)(2\bar{z} - \bar{a} + \bar{b}) = (b-a)\overline{(b-a)} \\ &\iff (2z - a - b)(2\bar{z} - \bar{a} - \bar{b}) = (b-a)\overline{(b-a)} \\ &\iff 4|z|^2 - 2z\bar{a}b - 2z\bar{b} - 2a\bar{z} + |a|^2 + a\bar{b} - 2b\bar{z} + b\bar{a} + |b|^2 \\ &\quad = |b|^2 + |a|^2 - b\bar{a} - \bar{b}a. \end{aligned}$$

En simplifiant et en divisant par deux, on trouve la condition de l'énoncé :

$$\boxed{2z\bar{z} - (\bar{a} + \bar{b})z - (a + b)\bar{z} + a\bar{b} + \bar{a}b = 0}.$$

4.2. Résolution d'équations & Racines

Exercice ALG.1.5 | Résoudre dans \mathbf{C} l'équation : $z^6 + 1 = i\sqrt{3}$.

Solution (exercice ALG.1.5)

On cherche z sous la forme exponentielle $z = \rho e^{i\theta}$ avec $\rho \geq 0$ et $\theta \in [0, 2\pi[$. On a

$$\begin{aligned} z^6 + 1 = i\sqrt{3} &\iff \rho^6 e^{6i\theta} = -1 + i\sqrt{3} \\ &\iff \rho^6 e^{6i\theta} = 2\left(-\frac{1}{2} + i\frac{\sqrt{3}}{2}\right) \\ &\iff \rho^6 e^{6i\theta} = 2e^{i\frac{2\pi}{3}} \\ &\iff \rho^6 = 2, \quad 6\theta = \frac{2\pi}{3} + 2k\pi, \quad k \in \mathbf{Z}, \\ &\iff \rho = \sqrt[6]{2}, \quad \theta = \frac{\pi}{9} + \frac{k\pi}{3}, \quad k \in \mathbf{Z}. \end{aligned}$$

L'ensemble des solutions est donc $\boxed{\left\{\sqrt[6]{2}e^{i\frac{\pi}{9} + \frac{k\pi}{3}}, k \in \mathbf{Z}\right\}}$. On peut aussi ne garder que les six premières valeurs de k .

Exercice ALG.1.6 | **Autour des racines 7-ièmes** Soient $u = e^{i\frac{2\pi}{7}}$, $S = u + u^2 + u^4$ et $T = u^3 + u^5 + u^6$.

1. Montrer que S et T sont conjugués, et que $\text{Im}(S) \geq 0$.
2. Calculer S + T et ST.

3. En déduire que :

$$\cos \frac{2\pi}{7} + \cos \frac{4\pi}{7} + \cos \frac{8\pi}{7} = -\frac{1}{2} \quad \text{et} \quad \sin \frac{2\pi}{7} + \sin \frac{4\pi}{7} + \sin \frac{8\pi}{7} = \frac{\sqrt{3}}{2}.$$

Solution (exercice ALG.1.6)

1. Constatons que $\bar{u} = e^{i\frac{2\pi}{7}} = e^{-i\frac{2\pi}{7}} = e^{-i\frac{2\pi}{7}+2i\pi} = e^{i\frac{12\pi}{7}} = u^6$. De-même :

$$\overline{u^2} = \bar{u}^2 = (u^6)^2 = u^{12} = u^7 u^5 = 1 \cdot u^5 = u^5,$$

et

$$\overline{u^4} = \bar{u}^4 = (u^6)^4 = u^{24} = u^{3 \times 7} u^3 = (u^7)^3 u^3 = 1 \cdot u^3.$$

Les termes de S, T sont donc conjuguées dans le même ordre, et par propriété de la conjugaison, on obtient $\overline{S} = T$. La deuxième partie est plus technique, on utilise la formule de Moivre, puis on calcule les puissances :

$$\begin{aligned} \text{Im}(S) &= \text{Im} \left(\cos \left(\frac{2\pi}{7} \right) + i \sin \left(\frac{2\pi}{7} \right) \right) \\ &+ \left(\cos \left(\frac{2\pi}{7} \right) + i \sin \left(\frac{2\pi}{7} \right) \right)^2 + \left(\cos \left(\frac{2\pi}{7} \right) + i \sin \left(\frac{2\pi}{7} \right) \right)^4 \\ &= \sin \left(\frac{2\pi}{7} \right) + 2 \sin \left(\frac{2\pi}{7} \right) \cos \left(\frac{2\pi}{7} \right) + 4 \cos^3 \left(\frac{2\pi}{7} \right) \sin \left(\frac{2\pi}{7} \right) - 4 \cos \left(\frac{2\pi}{7} \right) \sin^3 \left(\frac{2\pi}{7} \right) \\ &= \sin \left(\frac{2\pi}{7} \right) \left(1 + 2 \cos \left(\frac{2\pi}{7} \right) + 4 \cos^3 \left(\frac{2\pi}{7} \right) - 4 \cos \left(\frac{2\pi}{7} \right) \sin^2 \left(\frac{2\pi}{7} \right) \right) \end{aligned}$$

La partie imaginaire est donc du signe de la parenthèse car $\sin \left(\frac{2\pi}{7} \right) \geq 0$, et elle vaut

$$\begin{aligned} &1 + 2 \cos \left(\frac{2\pi}{7} \right) + 4 \cos^3 \left(\frac{2\pi}{7} \right) - 4 \cos \left(\frac{2\pi}{7} \right) \sin^2 \left(\frac{2\pi}{7} \right) \\ &= 1 + 2 \cos \left(\frac{2\pi}{7} \right) \left(1 + \cos^2 \left(\frac{2\pi}{7} \right) - \sin^2 \left(\frac{2\pi}{7} \right) \right) \\ &= 1 + 2 \cos \left(\frac{2\pi}{7} \right) 2 \cos^2 \left(\frac{2\pi}{7} \right) \geq 0 \end{aligned}$$

puisque $\cos \left(\frac{2\pi}{7} \right) \geq 0$. Donc finalement $\text{Im}(S) \geq 0$.

2.

$$S + T + 1 = \sum_{k=0}^6 u^k = \frac{1 - u^7}{1 - u} = \frac{1 - 1}{1 - u} = 0 \implies \boxed{S + T = -1}$$

Puis

$$\begin{aligned} ST &= (u + u^2 + u^4)(u^3 + u^5 + u^6) \\ &= u^4 + u^6 + u^7 + u^5 + u^7 + u^8 + u^7 + u^9 + u^{10} \\ &= u^4 + u^6 + 1 + u^5 + 1 + u + 1 + u^2 + u^3, \quad \text{en utilisant } u^7 = 1, \\ &= 2 + (S + T) = 2 + (-1) = \boxed{1}. \end{aligned}$$

3. Constatons que

$$\cos \frac{2\pi}{7} + \cos \frac{4\pi}{7} + \cos \frac{8\pi}{7} = \text{Re}(S) \quad \text{et} \quad \sin \frac{2\pi}{7} + \sin \frac{4\pi}{7} + \sin \frac{8\pi}{7} = \text{Im}(S).$$

Or, $T = \bar{S}$ et $S + T = -1$ donc $S + \bar{S} = -1 = 2 \text{Re}(S) = -1$, donc $\text{Re}(S) = -\frac{1}{2}$.

De plus, $ST = S\bar{S} = |S|^2 = 1$, donc S est de module un. Or, $\text{Re}(S)^2 + \text{Im}(S)^2 = 1$, donc $\text{Im}(S)^2 = 1 - \text{Re}(S)^2 = 1 - \frac{1}{4} = \frac{3}{4}$, on déduit que :

$$\text{Im}(S) = \frac{\sqrt{3}}{2}.$$

Exercice ALG.1.7 | Résoudre les équations suivantes dans \mathbf{C} :

1. $e^z = -7$.
2. $e^{2z} = -8 + 8i$.
3. $i + e^{z+2} = \sqrt{3}$.

Solution (exercice ALG.1.7)

Notons $z = x + iy$ avec $x, y \in \mathbf{R}$ dans tout l'exercice. Rappelons que $e^z = e^x e^{iy}$. Alors on résout

1.

$$\begin{aligned} e^z = -7 &\iff e^x e^{iy} = 7e^{i\pi} \\ &\iff e^x = 7, \quad y = \pi + 2k\pi, \quad k \in \mathbf{Z} \\ &\iff x = \ln 7, \quad y = \pi + 2k\pi, \quad k \in \mathbf{Z}. \end{aligned}$$

Les solutions sont donc tous les complexes de $\{\ln 7 + i(2k+1)\pi, k \in \mathbf{Z}\}$.

2.

$$\begin{aligned} e^z = -8 + 8i &\iff e^x e^{iy} = 8\sqrt{2}e^{i\frac{3\pi}{4}} \\ &\iff e^x = 8\sqrt{2}, \quad y = \frac{3\pi}{4} + 2k\pi, \quad k \in \mathbf{Z} \\ &\iff x = \ln(8\sqrt{2}) = \frac{7}{2}\ln 2, \quad y = \frac{3\pi}{4} + 2k\pi, \quad k \in \mathbf{Z}. \end{aligned}$$

Les solutions sont donc tous les complexes de $\{\frac{7}{2}\ln 2 + i(\frac{3\pi}{4} + 2k\pi), k \in \mathbf{Z}\}$.

3.

$$\begin{aligned} i + e^{z+2} = \sqrt{3} &\iff e^{(x+2)+iy} = \sqrt{3} - i = 2\left(\frac{\sqrt{3}}{2} - i\frac{1}{2}\right) = 2e^{i\frac{7\pi}{6}} \\ &\iff e^{(x+2)+iy} = 2e^{i\frac{7\pi}{6}} \\ &\iff e^{x+2} = 2, \quad y = \frac{7\pi}{6} + 2k\pi, \quad k \in \mathbf{Z} \\ &\iff x = \ln 2 - 2, \quad y = \frac{7\pi}{6} + 2k\pi, \quad k \in \mathbf{Z}. \end{aligned}$$

Les solutions sont donc tous les complexes de $\{\ln 2 - 2 + i(\frac{7\pi}{6} + 2k\pi), k \in \mathbf{Z}\}$.

Exercice ALG.1.8 | Soit $n \in \mathbf{N}$ tel que $n > 2$. On pose $z = e^{i\frac{2\pi}{n}}$.

1. On suppose que $k \in \llbracket 1, n-1 \rrbracket$. Déterminer le module et un argument du complexe $z^k - 1$.

2. On pose $S = \sum_{k=0}^{n-1} |z^k - 1|$. Montrer que $S = \frac{2 \cos \frac{\pi}{2n}}{\sin \frac{\pi}{2n}}$.

Solution (exercice ALG.1.8)

1. Soit $k \in \llbracket 1, n-1 \rrbracket$. Alors $z^k - 1 = e^{i\frac{2k\pi}{n}} - 1$. Mettons l'angle moitié en facteur. Il vient alors :

$$\begin{aligned} z^k - 1 &= e^{i\frac{k\pi}{n}} \left(2i \sin\left(\frac{k\pi}{n}\right) \right), \\ &= 2 \sin\left(\frac{k\pi}{n}\right) e^{i\pi\left(\frac{k}{n} + \frac{1}{2}\right)}. \end{aligned} \quad \left. \begin{array}{l} \\ \end{array} \right\} i = e^{i\frac{\pi}{2}}$$

Mais comme $\frac{k\pi}{n} \in [0, \pi]$, $\sin\left(\frac{k\pi}{n}\right) \geq 0$. Donc :

$$|z^k - 1| = 2 \sin\left(\frac{k\pi}{n}\right), \quad \text{Arg}(z^k - 1) = \pi\left(\frac{k}{n} + \frac{1}{2}\right).$$

$$\begin{aligned}
 2. S &= \sum_{k=0}^{n-1} |z^k - 1| \\
 &= 2 \sum_{k=0}^{n-1} \sin\left(\frac{k\pi}{n}\right) \\
 &= 2 \operatorname{Im}\left(\sum_{k=0}^{n-1} e^{i\frac{k\pi}{n}}\right) && \text{linéarité de la partie imaginaire} \\
 &= 2 \operatorname{Im}\left(\frac{1 - e^{i\frac{n\pi}{n}}}{1 - e^{i\frac{\pi}{n}}}\right) && \text{somme géométrique, } e^{i\pi} \neq 1 \\
 &= 2 \operatorname{Im}\left(\frac{2}{1 - e^{i\frac{\pi}{n}}}\right) && e^{i\pi} = -1 \\
 &= 4 \operatorname{Im}\left(\frac{1}{-2i \sin\left(\frac{\pi}{2n}\right)} e^{-i\frac{\pi}{2n}}\right) && \text{angle moitié} \\
 &= 2 \operatorname{Im}\left(\frac{i}{\sin\left(\frac{\pi}{2n}\right)} e^{-i\frac{\pi}{2n}}\right) \\
 &= \boxed{\frac{2}{\tan\frac{\pi}{2n}}}.
 \end{aligned}$$

Exercice ALG.1.9 | Soit $\varphi \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ fixé. On veut résoudre l'équation :

$$(E) \quad (1 + iz)^3(1 - i \tan \varphi) = (1 - iz)^3(1 + i \tan \varphi).$$

1. Montrer que si z est solution de (E) alors $|1 - iz| = |1 + iz|$. En déduire que z est réel.
2. Posons $z = \tan \theta$. Justifier ce changement d'inconnue, puis résoudre (E).

Solution (exercice ALG.1.9)

1. Soit z une solution, alors passons au module : $|1 + iz|^3 \frac{1}{|\cos \varphi|} = |1 - iz|^3 \frac{1}{|\cos \varphi|}$ puisque $1 + \tan^2 = \frac{1}{\cos^2}$. Donc en multipliant par $|\cos \varphi|$ et en utilisant la positivité

des modules, on obtient : $|1 + iz| = |1 - iz|$. Élevons ceci au carré, on a alors :

$$\begin{aligned}
 |1 + iz| &= |1 - iz| \\
 \iff |1 + iz|^2 &= |1 - iz|^2 \\
 \iff (1 + iz)(1 - i\bar{z}) &= (1 - iz)(1 + i\bar{z}) \\
 \iff 1 - i\bar{z} + iz + |z|^2 &= 1 - iz + i\bar{z} + |z|^2. \\
 \iff z &= \bar{z}.
 \end{aligned}$$

Donc $z \in \mathbf{R}$.

2. La fonction \tan réalise une bijection de $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ dans \mathbf{R} et nous avons montré que z est réel. Posons dès lors, puisque $z \in \mathbf{R}$, $z = \tan \theta$ et résolvons l'équation ci-dessous en $\theta \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$:

$$(1 + i \tan \theta)^3(1 - i \tan \varphi) = (1 - i \tan \theta)^3(1 + i \tan \varphi).$$

Elle est équivalente à

$$\left(\frac{1 + i \tan \theta}{1 - i \tan \theta}\right)^3 = \frac{1 + i \tan \varphi}{1 - i \tan \varphi},$$

ou encore, en multipliant par $\cos \theta, \cos \varphi$, au numérateur et dénominateur,

$$\begin{aligned}
 \left(\frac{\cos \theta + i \sin \theta}{\cos \theta - i \sin \theta}\right)^3 &= \left(\frac{e^{i\theta}}{e^{-i\theta}}\right)^3 = e^{6i\theta} \\
 &= \frac{\cos \varphi + i \sin \varphi}{\cos \varphi - i \sin \varphi} = \frac{e^{i\varphi}}{e^{-i\varphi}} = e^{2i\varphi}
 \end{aligned}$$

Donc on est amené à résoudre en $\theta \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$:

$$e^{6i\theta} = e^{2i\varphi}.$$

D'où $6\theta = 2\varphi + 2k\pi$ avec $k \in \mathbf{Z}$, i.e. $\theta = \frac{\varphi + k\pi}{3}$. On ne garde ensuite que les solutions dans $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$, l'ensemble des solutions est

$$\left\{ \tan\left(\frac{\varphi + k\pi}{3}\right), \frac{\varphi + k\pi}{3} \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[\right\}.$$

Exercice ALG.1.10 | Autour des racines 7-ièmes Soient $u = e^{i\frac{2\pi}{7}}$, $S = u + u^2 + u^4$ et $T = u^3 + u^5 + u^6$.

1. Montrer que S et T sont conjugués, et que $\text{Im}(S) \geq 0$.
2. Calculer S + T et ST.
3. En déduire que :

$$\cos \frac{2\pi}{7} + \cos \frac{4\pi}{7} + \cos \frac{8\pi}{7} = -\frac{1}{2} \quad \text{et} \quad \sin \frac{2\pi}{7} + \sin \frac{4\pi}{7} + \sin \frac{8\pi}{7} = \frac{\sqrt{3}}{2}.$$

Solution (exercice ALG.1.10)

1. Constatons que $\bar{u} = e^{-i\frac{2\pi}{7}} = e^{-i\frac{2\pi}{7} + 2i\pi} = e^{i\frac{12\pi}{7}} = u^6$. De-même :

$$\bar{u}^2 = \bar{u}^2 = (u^6)^2 = u^{12} = u^7 u^5 = 1 \cdot u^5 = u^5,$$

et

$$\bar{u}^4 = \bar{u}^4 = (u^6)^4 = u^{24} = u^{3 \times 7} u^3 = (u^7)^3 u^3 = 1 \cdot u^3.$$

Les termes de S, T sont donc conjugués dans le même ordre, et par propriété de la conjugaison, on obtient $\bar{S} = T$. La deuxième partie est plus technique, on utilise la formule de Moivre, puis on calcule les puissances :

$$\begin{aligned} \text{Im}(S) &= \text{Im} \left(\cos \left(\frac{2\pi}{7} \right) + i \sin \left(\frac{2\pi}{7} \right) \right) \\ &+ \left(\cos \left(\frac{2\pi}{7} \right) + i \sin \left(\frac{2\pi}{7} \right) \right)^2 + \left(\cos \left(\frac{2\pi}{7} \right) + i \sin \left(\frac{2\pi}{7} \right) \right)^4 \\ &= \sin \left(\frac{2\pi}{7} \right) + 2 \sin \left(\frac{2\pi}{7} \right) \cos \left(\frac{2\pi}{7} \right) + 4 \cos^3 \left(\frac{2\pi}{7} \right) \sin \left(\frac{2\pi}{7} \right) - 4 \cos \left(\frac{2\pi}{7} \right) \sin^3 \left(\frac{2\pi}{7} \right) \\ &= \sin \left(\frac{2\pi}{7} \right) \left(1 + 2 \cos \left(\frac{2\pi}{7} \right) + 4 \cos^3 \left(\frac{2\pi}{7} \right) - 4 \cos \left(\frac{2\pi}{7} \right) \sin^2 \left(\frac{2\pi}{7} \right) \right) \end{aligned}$$

La partie imaginaire est donc du signe de la parenthèse car $\sin \left(\frac{2\pi}{7} \right) \geq 0$, et elle vaut

$$\begin{aligned} &1 + 2 \cos \left(\frac{2\pi}{7} \right) + 4 \cos^3 \left(\frac{2\pi}{7} \right) - 4 \cos \left(\frac{2\pi}{7} \right) \sin^2 \left(\frac{2\pi}{7} \right) \\ &= 1 + 2 \cos \left(\frac{2\pi}{7} \right) \left(1 + \cos^2 \left(\frac{2\pi}{7} \right) - \sin^2 \left(\frac{2\pi}{7} \right) \right) \\ &= 1 + 2 \cos \left(\frac{2\pi}{7} \right) 2 \cos^2 \left(\frac{2\pi}{7} \right) \geq 0 \end{aligned}$$

puisque $\cos \left(\frac{2\pi}{7} \right) \geq 0$. Donc finalement $\boxed{\text{Im}(S) \geq 0}$.

2.

$$S + T + 1 = \sum_{k=0}^6 u^k = \frac{1 - u^7}{1 - u} = \frac{1 - 1}{1 - u} = \Rightarrow \boxed{S + T = -1}$$

Puis

$$\begin{aligned} ST &= (u + u^2 + u^4)(u^3 + u^5 + u^6) \\ &= u^4 + u^6 + u^7 + u^5 + u^7 + u^8 + u^7 + u^9 + u^{10} \\ &= u^4 + u^6 + 1 + u^5 + 1 + u + 1 + u^2 + u^3, \quad \text{en utilisant } u^7 = 1, \\ &= 2 + (S + T) = 2 + (-1) = \boxed{1}. \end{aligned}$$

3. Constatons que

$$\cos \frac{2\pi}{7} + \cos \frac{4\pi}{7} + \cos \frac{8\pi}{7} = \text{Re}(S) \quad \text{et} \quad \sin \frac{2\pi}{7} + \sin \frac{4\pi}{7} + \sin \frac{8\pi}{7} = \text{Im}(S).$$

Or, $T = \bar{S}$ et $S + T = -1$ donc $S + \bar{S} = -1 = 2 \text{Re}(S) = -1$, donc $\boxed{\text{Re}(S) = -\frac{1}{2}}$.

De plus, $ST = S\bar{S} = |S|^2 = 1$, donc S est de module un. Or, $\text{Re}(S)^2 + \text{Im}(S)^2 = 1$, donc $\text{Im}(S)^2 = 1 - \text{Re}(S)^2 = 1 - \frac{1}{4} = \frac{3}{4}$, on déduit que :

$$\text{Im}(S) = \frac{\sqrt{3}}{2}.$$

4.3. Trigonométrie

Exercice ALG.1.11 | À l'aide des nombres complexes, établir que pour tous $(x, y) \in \mathbf{R}^2$:

- $\cos(x) + \cos(y) = 2 \cos\left(\frac{x+y}{2}\right) \cos\left(\frac{x-y}{2}\right)$.
- $\sin(x) - \sin(y) = 2 \cos\left(\frac{x+y}{2}\right) \sin\left(\frac{x-y}{2}\right)$.

Solution (exercice ALG.1.11)

$$\begin{aligned} \cos x + \cos y &= \operatorname{Re}\left(e^{ix} + e^{iy}\right) \\ &= \operatorname{Re}\left(e^{i\frac{x+y}{2}} \left(e^{i\frac{x-y}{2}} + e^{-i\frac{x-y}{2}}\right)\right) \\ &= \operatorname{Re}\left(e^{i\frac{x+y}{2}} 2 \cos\left(\frac{x-y}{2}\right)\right) \\ &= \boxed{2 \cos\left(\frac{x-y}{2}\right) \cos\left(\frac{x+y}{2}\right)}. \end{aligned}$$

De-même :

$$\begin{aligned} \sin x - \sin y &= \operatorname{Im}\left(e^{ix} - e^{iy}\right) \\ &= \operatorname{Im}\left(e^{i\frac{x+y}{2}} \left(e^{i\frac{x-y}{2}} - e^{-i\frac{x-y}{2}}\right)\right) \\ &= \operatorname{Im}\left(e^{i\frac{x+y}{2}} 2i \sin\left(\frac{x-y}{2}\right)\right) \\ &= \boxed{2 \sin\left(\frac{x-y}{2}\right) \sin\left(\frac{x+y}{2}\right)}. \end{aligned}$$

Exercice ALG.1.12 | Soient $n \in \mathbf{N} \setminus \{0\}$ et $x \in \mathbf{R}$. Calculer : $\sum_{p=0}^n \cos^2(px)$.

Solution (exercice ALG.1.12)

En utilisant les formules d'Euler, nous avons

$$\begin{aligned} \sum_{p=0}^n \cos^2(px) &= \frac{1}{4} \sum_{p=0}^n \left(e^{ipx} + e^{-ipx}\right)^2 \\ &= \frac{1}{4} \sum_{p=0}^n \left(e^{2ipx} + e^{-2ipx} + 2\right) \\ &= \frac{1}{4} \left(\sum_{p=0}^n e^{2ipx} + \sum_{p=0}^n e^{-2ipx} + (n+1)\right) \\ &= \begin{cases} \frac{1}{4} \left(\frac{1-e^{2i(n+1)x}}{1-e^{2ix}} + \frac{1-e^{-2i(n+1)x}}{1-e^{-2ix}} + 2(n+1)\right) & \text{si } x \notin \mathbf{Z}, \\ \frac{3(n+1)}{4} & \text{sinon.} \end{cases} \end{aligned}$$

Or, par propriété de la conjugaison, on constate que :

$$\frac{1 - e^{2i(n+1)x}}{1 - e^{2ix}} = \overline{\left(\frac{1 - e^{-2i(n+1)x}}{1 - e^{-2ix}}\right)}.$$

Donc, si $x \notin \mathbf{N}$, on peut finir le calcul en utilisant la technique de l'angle moitié

$$\begin{aligned} \sum_{p=0}^n \cos^2(px) &= \frac{1}{4} \left(2 \operatorname{Re}\left(\frac{1 - e^{2i(n+1)x}}{1 - e^{2ix}}\right) + 2(n+1)\right) \\ &= \frac{1}{2} \left(\operatorname{Re}\left(\frac{e^{i(n+1)x} - 2 \cos((n+1)x)}{e^{ix} - 2 \cos(x)}\right) + (n+1)\right) \\ &= \boxed{\frac{1}{2} \left(\cos\left(\frac{nx}{2}\right) \frac{\cos((n+1)x)}{\cos(x)} + (n+1)\right)}. \end{aligned}$$

Exercice ALG.1.13 | Soient $n \in \mathbf{N}$ et $(\alpha, \beta) \in \mathbf{R}^2$. Calculer :

- $C = \sum_{k=0}^n \binom{n}{k} \cos(\alpha + k\beta)$,
- $S = \sum_{k=0}^n \binom{n}{k} \sin(\alpha + k\beta)$.

Solution (exercice ALG.1.13)

Commençons par calculer

$$\sum_{k=0}^n \binom{n}{k} e^{i(\alpha+k\beta)},$$

il suffira ensuite de calculer la partie réelle et imaginaire.

$$\begin{aligned} \sum_{k=0}^n \binom{n}{k} e^{i(\alpha+k\beta)} &= e^{i\alpha} \sum_{k=0}^n \binom{n}{k} (e^{i\beta})^k \\ &= e^{i\alpha} (1 + e^{i\beta})^n && \left. \begin{array}{l} \text{binôme} \\ \text{angle moitié} \end{array} \right\} \\ &= e^{i\alpha} \left(e^{i\frac{\beta}{2}} \left(2 \cos \left(\frac{\beta}{2} \right) \right) \right)^n, \\ &= e^{i\alpha + ni\frac{\beta}{2}} 2^n \cos^n \left(\frac{\beta}{2} \right). \end{aligned}$$

On déduit alors les parties réelles et imaginaires,

$$\begin{aligned} \sum_{k=0}^n \binom{n}{k} \cos(\alpha + k\beta) &= \operatorname{Re} \left(\sum_{k=0}^n \binom{n}{k} e^{i(\alpha+k\beta)} \right) \\ &= \boxed{2^n \cos^n \left(\frac{\beta}{2} \right) \cos \left(\alpha + n \frac{\beta}{2} \right)}, \end{aligned}$$

$$\begin{aligned} \sum_{k=0}^n \binom{n}{k} \sin(\alpha + k\beta) &= \operatorname{Im} \left(\sum_{k=0}^n \binom{n}{k} e^{i(\alpha+k\beta)} \right) \\ &= \boxed{2^n \cos^n \left(\frac{\beta}{2} \right) \sin \left(\alpha + n \frac{\beta}{2} \right)}. \end{aligned}$$

Chapitre ALG.2.

Polynômes

Résumé & Plan

L'objectif de ce chapitre est de revoir certaines propriétés de première année sur les polynômes. Quelques compléments seront présentés en fin de chapitre, notamment la formule de TAYLOR qui stipule que pour les fonctions polynomiales la formule de TAYLOR-YOUNG est en fait exacte — *i.e.* sans terme en $o()$.

1	Définition de $\mathbb{K}[X]$	1
1.1	Généralités	2
1.2	Propriétés du degré	3
2	Polynôme dérivé & primitivé	4
2.1	Généralités	5
3	Racines	7
3.1	Généralités	7
3.2	Existence de racines & Comptage	9
3.3	Cas du second degré	12

4	Exercices	15
4.1	Généralités, Racines, Factorisation	15
4.2	Équations fonctionnelles polynomiales	17
4.3	Familles classiques	18

114 est le plus petit nombre entier naturel dont on ne sait toujours pas s'il peut s'écrire comme une somme de trois cubes. Avant Septembre 2019, ce plus petit nombre était 42, mais finalement :

$$42 = (80435758145817515)^3 + (-80538738812075974)^3 + (12602123297335631)^3$$

— Le saviez-vous ?



Cadre

Dans tout le chapitre, l'ensemble \mathbb{K} désignera \mathbb{R} ou \mathbb{C} .

1. DÉFINITION DE $\mathbb{K}[X]$

Les polynômes sont définis généralement comme des suites d'éléments de \mathbb{K} qui comportent un nombre fini de termes non nuls, la suite des coefficients a_0, \dots, a_n

(avec $a_{n+1} = a_{n+2} = \dots = 0$) pour un certain $n \in \mathbf{N}$, que l'on appelle degré. Et on appelle fonction polynomiale associée la fonction $x \mapsto \sum_{k=0}^n a_k x^k$. Conformément au programme de BCPST, nous ne ferons pas la différence entre polynôme et fonction polynomiale associée¹. Ainsi, pour nous, les polynômes seront déjà des fonctions.

1.1. Généralités

Définition ALG.2.1 | Polynôme sur \mathbf{K}

1. Soit $n \in \mathbf{N}$. Un *monôme de degré n* est une fonction $P : \mathbf{K} \rightarrow \mathbf{K}$ de la forme $x \mapsto ax^n$ avec $n \in \mathbf{N}$ et $a \in \mathbf{K}$. L'entier n est appelé *degré de P* et généralement noté $\deg P$.
2. Soit $n \in \mathbf{N}$. Un *polynôme* est une fonction du type $P : \mathbf{K} \rightarrow \mathbf{K}$ de la forme $x \mapsto \sum_{k=0}^n a_k x^k$ avec $n \in \mathbf{N}$ et $a_0, \dots, a_n \in \mathbf{K}$ tel que $a_n \neq 0$. L'entier n est appelé *degré de P* et généralement noté $\deg P$, c'est le plus grand entier k tel que $a_k \neq 0$. La fonction $P : x \mapsto \sum_{k=0}^n a_k x^k$ est généralement notée

$$P = \sum_{k=0}^n a_k X^k.$$

On convient que $\deg 0 = -\infty$. Le coefficient $a_{\deg P}$ est appelé *coefficient dominant de P* . Si $a_{\deg P} = 1$ on dit que P est *unitaire*. Si $\mathbf{K} = \mathbf{R}$, on parle de polynôme à coefficients réels. Si $\mathbf{K} = \mathbf{C}$, on parle de polynôme à coefficients complexes.

Notation

- On note :
- ▶ $\mathbf{K}[X]$ l'ensemble des polynômes sur \mathbf{K} ,
 - ▶ $\mathbf{K}_n[X]$ l'ensemble des polynômes sur \mathbf{K} de degré inférieur à n ,
 - ▶ $\mathbf{K}_{=n}[X]$ l'ensemble des polynômes sur \mathbf{K} de degré égal à n .
 - ▶ Le polynôme $0_{\mathbf{K}_n[X]}$ est le polynôme nul, *i.e.* la fonction nulle pour nous.

¹de ce point de vue : il nous sera impossible de substituer $x \in \mathbf{K}$ par une matrice ou tout autre objet plus général.

Attention

à ne pas confondre $\sum_{k=0}^n a_k x^k$ pour un certain $x \in \mathbf{K}$ et $\sum_{k=0}^n a_k X^k$ qui dans le premier cas est un élément de \mathbf{K} (*i.e.* un réel ou un complexe), dans le second une **fonction**. Mais nous avons :

$$\forall x \in \mathbf{K}, \left(\sum_{k=0}^n a_k X^k \right) (x) = \sum_{k=0}^n a_k x^k.$$

Attention Dans un polynôme, la somme est finie

une fonction du type $x \mapsto \sum_{k=0}^{\infty} a_k x^k = \lim_{n \rightarrow \infty} \sum_{k=0}^n a_k x^k$, en cas d'existence, n'est **pas** un polynôme.² Par exemple, pour tout $n \in \mathbf{N}$, si $x \neq 1$,

$$\sum_{k=0}^n x^k = \frac{1 - x^{n+1}}{1 - x},$$

cette quantité tend vers $\frac{1}{1-x}$ si $|x| < 1$. Et la fonction $x \mapsto \frac{1}{1-x}$ n'est **pas** un polynôme en x .

Notation

Parfois le polynôme P est aussi noté $P(X)$, la notation $P(X)$ désigne donc encore une fonction.

Remarque 1.1 — Convention de degré. L'application \deg est donc à valeurs dans $\mathbf{N} \cup \{-\infty\}$:

$$\deg : \begin{cases} \mathbf{K}[X] & \rightarrow \mathbf{N} \cup \{-\infty\}, \\ P & \mapsto \deg P. \end{cases}$$

La convention $\deg 0 = -\infty$ est purement technique. Elle trouve son intérêt dans la formule de degré d'un produit que nous reverrons plus tard : soit $P \in \mathbf{K}[X]$, alors $\deg(0.P) = \deg(0) + \deg P$ d'une part, et d'autre part comme $0.P = 0$ on devrait avoir

²Dans ce cas on parle de série entière lorsque $\mathbf{K} = \mathbf{R}$, de fonction analytique lorsque $\mathbf{K} = \mathbf{C}$, leur étude n'est pas au programme de BCPST

$\deg 0 = -\infty + \deg P$. Cette formule n'est jamais vérifiée sauf si P est constant, d'où la convention précédente de sorte que

$$-\infty = -\infty + 0.$$

Les polynômes étant définis ici comme un sous-ensemble de l'espace des fonctions de \mathbf{K} dans \mathbf{K} , on peut réaliser plusieurs opérations sur eux comme pour les fonctions habituelles.

Définition ALG.2.2 | Opérations +, ×, ◦, .

Soient $P, Q \in \mathbf{K}[X]$, $\lambda \in \mathbf{K}$. On notera $P+Q$, PQ , $P \circ Q$ et λP les fonctions ci-dessous :

1. $P+Q \left| \begin{array}{l} \mathbf{K} \rightarrow \mathbf{K}, \\ x \mapsto P(x)+Q(x), \end{array} \right.$
2. $P \times Q \left| \begin{array}{l} \mathbf{K} \rightarrow \mathbf{K}, \\ x \mapsto P(x)Q(x), \end{array} \right.$
3. $P \circ Q \left| \begin{array}{l} \mathbf{K} \rightarrow \mathbf{K}, \\ x \mapsto P(x) \circ Q(x), \end{array} \right.$
4. $\lambda P \left| \begin{array}{l} \mathbf{K} \rightarrow \mathbf{K}, \\ x \mapsto \lambda P(x). \end{array} \right.$

Attention Piège de notation

Si $P \in \mathbf{K}[X]$ et $p \geq 0$, attention à la différence entre le polynôme $X^p P : x \in \mathbf{K} \mapsto x^p P(x)$ — un produit — et le polynôme $P(X^p) : x \in \mathbf{K} \mapsto P(x^p)$ — une composée.

Plus explicitement à l'aide de coefficients, soient $P = \sum_{k=0}^p a_k X^k$ et $Q = \sum_{k=0}^q b_k X^k$, avec la convention

- ▶ $a_k = 0$ si $k \notin \llbracket 0, p \rrbracket$,
- ▶ $b_k = 0$ si $k \notin \llbracket 0, q \rrbracket$.

Alors on a :

1. $P+Q = \sum_{n=0}^{\max(p,q)} (a_n + b_n) X^n,$
2. $PQ(X) = \sum_{n=0}^{p+q} \left(\sum_{k=0}^n a_k b_{n-k} \right) X^n,$
3. $P \circ Q(X) = \sum_{n=0}^p a_n \left(\sum_{k=0}^q b_k X^k \right)^n,$
4. $\lambda P = \sum_{n=0}^p \lambda a_n X^n.$

Proposition ALG.2.1 | Stabilité des opérations

Soient $n \geq 1$, $P, Q \in \mathbf{K}[X]$ (resp. $\mathbf{K}_n[X]$), $\lambda \in \mathbf{K}$, alors $P+Q$, PQ , $P \circ Q$ et λP sont des éléments de $\mathbf{K}[X]$ (resp. $\mathbf{K}_n[X]$). L'ensemble $\mathbf{K}[X]$ est donc stable par addition, produit, composition et multiplication par un scalaire.³

Preuve On vérifie sans peine que les précédentes fonctions sont encore des polynômes.

Attention

La proposition précédente est fautive pour $\mathbf{K}_{=n}[X]$.⁴ Par exemple, puisque $X^2 + (-X^2) = 0$, la somme de deux polynômes de degré deux n'est pas forcément un polynôme de degré deux.

1.2. Propriétés du degré

Proposition ALG.2.2

Soient $P, Q \in \mathbf{K}[X]$. On a :

1. (Degré d'une somme)

$$\deg(P+Q) \leq \max\{\deg P, \deg Q\}.$$

De plus, si $\deg P \neq \deg Q$, alors il y a égalité :

$$\deg(P+Q) = \max\{\deg P, \deg Q\}.$$

2. (Degré d'une multiplication scalaire) Si $\lambda \neq 0$,

$$\deg(\lambda P) = \deg P.$$

³Il s'agit aussi d'un espace vectoriel, cf. **Chapter ALG.3** pour le détail.

⁴Cet ensemble ne sera **pas** un espace vectoriel, cf. **Chapter ALG.3** pour le détail.

3. (Degré d'un produit)

$$\deg(PQ) = \deg P + \deg Q.$$

4. (Degré d'une composée)

$$\deg(P \circ Q) = \deg P \times \deg Q.$$

Preuve Voir le cours de 1ère année.

Remarque 1.2 — La distinction de cas de la première assertion est présente pour traiter le cas où les coefficients dominant se compensent : si $\deg P \neq \deg Q$, cela n'arrivera jamais, sinon on peut observer une chute de degré. Par exemple, considérons $P = -X^2 + X + 1$ et $Q = X^2 + 1$. Alors $P + Q = X + 2$ est de degré 1, et ce n'est pas le max des degrés qui est 2.

Proposition ALG.2.3 | Intégrité de l'ensemble des polynômes

Soient $P, Q \in \mathbf{K}[X]$. Alors : $PQ = 0_{\mathbf{K}[X]} \implies (P = 0_{\mathbf{K}[X]} \text{ ou } Q = 0_{\mathbf{K}[X]}).$

Preuve Supposons que $PQ = 0_{\mathbf{K}[X]}$, alors en prenant le degré nous avons : $\deg P + \deg Q = -\infty$, donc nécessairement un des deux degrés vaut $-\infty$ i.e. $P = 0_{\mathbf{K}[X]}$ ou $Q = 0_{\mathbf{K}[X]}$ — encore une bonne illustration de l'intérêt de la convention « $\deg(0_{\mathbf{K}[X]}) = -\infty$ »!^a

Attention
Ce résultat est faux pour deux fonctions quelconques non polynomiales.

Codage informatique d'un polynôme

Comment coder un polynôme en Python? Puisqu'il est entièrement déterminé par la donnée de ses coefficients, il suffit par exemple de les ranger par ordre croissant de degré et considérer en convention que le polynôme nul correspond à la liste vide. Par exemple, le polynôme $P = X^2 + 2$ peut être codé par la liste $L = [2, 0, 1]$. On peut en déduire alors facilement une fonction qui calcule le

^aOn a l'impression de n'avoir rien fait dans cette preuve : en fait le travail principal réside dans la propriété de degré d'un produit établie plus haut.

degré.

```
import numpy as np
# np.inf correspond à l'infini des maths
def degre(P):
    """
    retourne le degré de P
    """
    if len(P) == 0:
        return -np.inf
    else:
        return len(P) - 1
```

Comment coder à présent les opérations élémentaires sur les polynômes. Par exemple, voici comment s'y prendre pour la multiplication par X d'un polynôme (on constate que cela décale vers la droite tous les coefficients). Il faut donc simplement ajouter 0 en début de liste.

```
def mult(P):
    """
    retourne la liste correspond à X.P
    """
    return [0] + P
```

2. POLYNÔME DÉRIVÉ & PRIMITIVÉ

Les polynômes ont été définis dans ce chapitre comme un sous-ensemble de l'ensemble des fonctions, donc ils héritent en particulier de la notion de dérivation connue depuis longtemps si $\mathbf{K} = \mathbf{R}$. En revanche, pour $\mathbf{K} = \mathbf{C}$, comme nous ne savons pas dériver les fonctions de la variable complexe, la définition *infra* est finalement plus générale.

2.1. Généralités

Définition ALG.2.3 | Dérivation

Soit $P = \sum_{\ell=0}^n a_{\ell} X^{\ell} \in \mathbf{K}[X]$, avec $n \geq 1$.

1. On appelle *polynôme dérivé de P* le polynôme noté

$$P' = \sum_{\ell=1}^n \ell a_{\ell} X^{\ell-1},$$

i.e. la dérivée de la fonction P.

2. Soit $k \in \mathbf{N}$. On appelle *polynôme dérivé k fois de P* le polynôme

$$P^{(k)} = \sum_{\ell=k}^n \ell(\ell-1)\dots(\ell-k+1)a_{\ell} X^{\ell-k},$$

i.e. la dérivée k-fois de la fonction P.

Codage informatique du polynôme dérivé



On déduit directement de la définition, la propriété de degré d'une dérivée.

Proposition ALG.2.4 | Degré d'un polynôme dérivé

Si $P \in \mathbf{K}[X]$ et $k \in \mathbf{N}$, alors

$$\deg P^{(k)} = \begin{cases} \deg P - k & \text{si } k \leq \deg P, \\ -\infty & \text{si } k > \deg P. \end{cases}$$


Remarque 2.1 — Pour la dérivée première, peut-on écrire $k = 0$ ou $k = 1$ en premier indice? On prendra garde d'éviter l'expression $P' = \sum_{k=0}^n k a_k X^{k-1}$ même si le terme d'ordre $k = 0$ est nul. En effet, nous n'avons pas donné un sens à $0 \times \frac{1}{X}$, ce n'est pas un élément de $\mathbf{K}[X]$ ⁵.

Définition ALG.2.4 | Primitivation (en zéro)

Soit $P = \sum_{\ell=0}^n a_{\ell} X^{\ell} \in \mathbf{K}[X]$, avec $n \geq 1$. On appelle *polynôme primitivé de P* le polynôme noté

$$\int P = \sum_{\ell=0}^n a_{\ell} \frac{X^{\ell+1}}{\ell+1},$$

i.e. la primitive **qui s'annule en zéro** de la fonction P.

Exemple 1 — Calculer les dérivées successives de $P = X^4 - 3X^3 + iX^2 - 1$ et le polynôme primitif de P.  Puisque P est de degré 4, alors $P^{(k)} = 0$ dès que $k \geq 5$. Par ailleurs, $P' = 4X^3 - 9X^2 + 2iX$, $P'' = 12X^2 - 18X + 2i$, $P''' = 24X - 18$ et $P^{(4)} = 24$.

Proposition ALG.2.5 | Dérivées d'un monôme

Soient $a \in \mathbf{K}$ et $(n, k) \in \mathbf{N}^2$. Alors : $[(X - a)^n]^{(k)} = \begin{cases} \frac{n!}{(n-k)!} (X - a)^{n-k} & \text{si } k \leq n, \\ 0 & \text{si } k > n. \end{cases}$ ⁶

Preuve



Le polynôme $(X - a)^n$ est de degré n donc si $k > n$, alors $[(X - a)^n]^{(k)} =$

⁵mais de $\mathbf{K}(X)$, l'ensemble des fractions rationnelles

⁶Cette formule très classique est à savoir retrouver très rapidement.

0. Supposons que $k \leq n$, alors

$$[(X - a)^n]^{(k)} = n[(X - a)^{n-1}]^{(k-1)} = n(n-1)[(X - a)^{n-2}]^{(k-1)},$$

puis de manière générale on obtient :

$$[(X - a)^n]^{(k)} = n(n-1) \dots (n-k+1)(X - a)^{k-n} = \frac{n!}{(n-k)!} (X - a)^{n-k}.$$

Proposition ALG.2.6 | Dérivations et opérations

Soient $P, Q \in \mathbf{K}[X]$ et $\lambda \in \mathbf{K}$. Alors :

1. $(\lambda P)' = \lambda P'$,
2. $(P + Q)' = P' + Q'$,
3. $(PQ)' = P'Q + PQ'$,
4. $(P \circ Q)' = P' \circ Q + P \circ Q'$.

Preuve Si $\mathbf{K} = \mathbf{R}$: comme pour nous les polynômes sont des fonctions, les formules ci-dessus découlent donc des formules déjà connues pour les fonctions réelles de la variable réelle. Si $\mathbf{K} = \mathbf{C}$, il faut les vérifier à l'aide de coefficients [...].

UNICITÉ DES COEFFICIENTS. Posons-nous à présent la question suivante : existe-il plusieurs familles de coefficients possibles $(a_0, \dots, a_n) \in \mathbf{K}^n$ telles que

$$P = \sum_{k=0}^n a_k X^k \quad ?$$

La réponse est non, un polynôme est entièrement déterminé par la suite de ses coefficients (réels ou complexes).

Proposition ALG.2.7 | Un polynôme est déterminé par ses coefficients

Soit $P = \sum_{k=0}^n a_k X^k \in \mathbf{K}[X]$ avec $a_0, \dots, a_n \in \mathbf{K}$.

1. Pour tout entier $k \in \llbracket 0, n \rrbracket$, nous avons :

$$a_k = \frac{P^{(k)}(0)}{k!}.$$

2. Par conséquent :

- ▶ $P = 0 \iff a_0 = \dots = a_n = 0.$
- ▶ la suite de coefficients a_0, \dots, a_n est unique.

Preuve Exprimons déjà les coefficients de P en fonction des dérivées successives. En effet, pour tout entier $k \in \llbracket 0, n \rrbracket$,

$$P^{(k)}(X) = \left(\sum_{\ell=0}^n a_\ell X^\ell \right)^{(k)} = \sum_{\ell=k}^n a_\ell (\ell-1) \dots (\ell-k+1) X^{\ell-k}.$$

Nous obtenons :

$$P^{(k)}(0) = \sum_{\ell=k}^n a_\ell (\ell-1) \dots (\ell-k+1) x^{\ell-k} \Big|_{x=0} = k! a_k.$$



Les coefficients sont donc donnés par la formule suivante $a_k = \frac{P^{(k)}(0)}{k!}.$


- ▶ \Leftarrow Évident.
- ▶ \Rightarrow Supposons que $P = 0$, alors pour tout $k \in \llbracket 0, n \rrbracket$, $\frac{P^{(k)}(0)}{k!} = \frac{0}{k!} = 0 = a_k$. L'équivalence est donc démontrée.
- ▶ Supposons qu'il existe par ailleurs $b_0, \dots, b_n \in \mathbf{R}$ tels que : $P = \sum_{k=0}^n b_k X^k = \sum_{k=0}^n a_k X^k$. Alors $0 = \sum_{k=0}^n 0 \cdot X^k = \sum_{k=0}^n (b_k - a_k) X^k$, donc d'après ce qui précède $0 = b_k - a_k$ pour tout $k \in \llbracket 0, n \rrbracket$. C'est terminé.

Exemple 2 – Fonction génératrice d'une variable aléatoire réelle discrète. Soit X une variable aléatoire réelle discrète à support fini $\llbracket 0, n \rrbracket$, on note

$$g_X(t) = \mathbf{E} \left(t^X \right) = \sum_{k=0}^n \mathbf{P}(X = k) t^k$$

pour $t \in \mathbf{R}$ — l'égalité provient du théorème de transfert.

1. Que dire de la fonction g_X ?  La fonction g_X est un polynôme de degré n par définition d'un polynôme.
2. Comment obtenir la loi de X à partir de g_X ?  D'après la proposition précédente, pour tout $k \in \mathbf{N}$: $\mathbf{P}(X = k) = \frac{g_X^{(k)}(0)}{k!}.$

3. Déterminer g_X dans le cas $X \hookrightarrow \mathcal{U}[[0, n]]$.  Nous avons alors : $\forall k \in [[0, n], P(X = k) = \frac{1}{n+1}$. Donc pour tout $t \in \mathbf{R}$,

$$g_X(t) = \sum_{k=0}^n \frac{t^k}{n} = \frac{1}{n} \begin{cases} \frac{1-t^{n+1}}{1-t} & \text{si } t \neq 1, \\ \frac{n+1}{n+1} = 1 & \text{si } t = 1. \end{cases}$$

3. RACINES

On se préoccupe à présent des points d'annulation d'un polynôme.

1. Premier point : on voit très clairement quel est leur nombre pour des petits degrés ; une fonction affine (un polynôme de degré un) s'annule en un point, un trinôme s'annule toujours au plus deux fois dans \mathbf{C} . En fait, tout polynôme de degré n s'annulera au plus n fois. Des propriétés sur le nombre de racines semblent donc exister.
2. Second point : lorsque λ annule un polynôme, alors ledit polynôme sera de la forme $(X - \lambda)Q$ où $Q \in \mathbf{K}[X]$. Cette «factorisation» est uniquement vraie pour les polynômes — pas question d'écrire cela pour d'autres types de fonctions!

Exemple 3 — Contre- La fonction $x \in \mathbf{R}^+ \mapsto \sqrt{x}$ s'annule en zéro et pourtant n'est pas de la forme $x \mapsto xQ(x)$ avec $Q \in \mathbf{R}[X]$.

Commençons par une notation que l'on utilisera dans toute la section.

Définition ALG.2.5 | Relation de divisibilité

Soient $P, Q \in \mathbf{K}[X]$. Alors on dit que P *divise* Q s'il existe un polynôme $R \in \mathbf{K}[X]$ tel que : $PR = Q$. On notera $P \mid Q$.

3.1. Généralités

Définition ALG.2.6 | Racine

Soient $P \in \mathbf{K}[X]$ et $\lambda \in \mathbf{K}$. On dit que λ est une *racine* de P si :

$$P(\lambda) = 0.$$

Lorsque $\mathbf{K} = \mathbf{C}$: on parle de *racine réelle* si on a de plus $\lambda \in \mathbf{R}$.

Proposition ALG.2.8 | Existence d'une racine réelle pour un polynôme de degré impair

Soit $P \in \mathbf{K}[X]$ un polynôme de degré impair. Alors P possède une racine réelle.

Attention

Cette propriété est très classique : il faut bien en connaître la démonstration. Il n'y a en revanche pas nécessairement unicité.

Preuve Appliquons le théorème des valeurs intermédiaires à la fonction $x \mapsto P(x)$, notons $P(X) = a_{2p+1}X^{2p+1} + \dots + a_0$ avec $a_0, \dots, a_{2p+1} \in \mathbf{K}$, et supposons que $a_{2p+1} > 0$ — sinon on applique le même raisonnement à $-P$. Alors P est une fonction continue, et $\lim_{x \rightarrow \pm\infty} P(x) = \pm\infty$. Donc d'après le théorème des valeurs intermédiaires, il existe $\lambda \in \mathbf{R}$ tel que $P(\lambda) = 0$. Autrement dit, P possède une racine réelle.


Passons maintenant à la propriété de factorisation constatée plus haut sur des exemples.

Proposition ALG.2.9 | Caractérisation par factorisation

Soient $P \in \mathbf{K}[X]$ et $\lambda \in \mathbf{K}$. Alors :

$$\lambda \text{ est racine de } P \iff \exists Q \in \mathbf{K}[X], \quad (X - \lambda) \times Q = P.$$

Preuve

 Immédiat, évaluer l'identité en λ .

⇒ Soit λ une racine de P . Alors écrivons que $P(X) = P(X) - P(\lambda)$. Alors

$$\begin{aligned} P(X) &= P(X) - P(\alpha) = \sum_{k=0}^n a_k (X^k - \alpha^k) \\ &= \sum_{k=0}^n (X - \alpha) \sum_{i=0}^{k-1} X^i \alpha^{k-i-1} \\ &= (X - \alpha) \sum_{k=0}^n \sum_{i=0}^{k-1} X^i \alpha^{k-i-1}. \end{aligned}$$

) téléscopage

En notant $Q(X) = \sum_{k=0}^n \sum_{i=0}^{k-1} X^i \alpha^{k-i-1}$, qui est bien un polynôme, nous obtenons le résultat.

Dans la Proposition ALG.2.9, rien ne nous dit que Q n'est pas encore lui-même de la forme $(X - \lambda)\tilde{Q}$ avec $\tilde{Q} \in \mathbf{K}[X]$, auquel cas $P = (X - \lambda)^2\tilde{Q}$. Pour quantifier la puissance maximale apparaissant dans l'exposant de $X - \lambda$, on introduit la notion de *multiplicité*.

Définition/Proposition ALG.2.1 | Multiplicité

Soient $P \in \mathbf{K}[X]$, $P \neq 0_{\mathbf{K}[X]}$,⁷ et $\lambda \in \mathbf{K}$. On dit que λ est une *racine d'ordre k* (ou de *multiplicité k*) de P si l'une des propositions ci-dessous est vérifiée :

$$\begin{aligned} &(X - \lambda)^k \mid P \quad \text{et} \quad (X - \lambda)^{k+1} \nmid P, \\ \iff &\exists Q \in \mathbf{K}[X], (X - \lambda)^k Q = P, \quad Q(\lambda) \neq 0. \end{aligned}$$

Autrement dit, l'ordre d'une racine est la plus grande puissance k telle que P soit factorisable par $(X - \lambda)^k$ dans $\mathbf{K}[X]$. On note $k = \text{Mult}_\lambda(P)$.

Preuve Montrons l'équivalence des deux propositions.

⇐ Supposons que : $\exists Q \in \mathbf{K}[X], (X - \lambda)^k Q = P$ avec $Q(\lambda) \neq 0$. Alors $(X - \lambda)^k \mid P$. Supposons par l'absurde que $(X - \lambda)^{k+1} \mid P$, alors il existe $\tilde{Q} \in \mathbf{K}[X]$ tel que $(X - \lambda)^{k+1}\tilde{Q} = P$. Alors :

$$(X - \lambda)^k \left((X - \lambda)\tilde{Q} \right) = P = (X - \lambda)^k Q \implies (X - \lambda)^k \left((X - \lambda)\tilde{Q} - Q \right) = 0.$$

⁷Pourquoi supposer P non nul? Étant donné que $(X - \lambda)^k$ divise P pour tout k , toutes les racines de 0 serait de multiplicité « infinie ».

Donc : $(X - \lambda)\tilde{Q} = Q$ et $Q(\lambda) = 0$ — contradiction.

⇒ Notons $Q \in \mathbf{K}[X], (X - \lambda)^k Q = P$, montrons que $Q(\lambda) \neq 0$. Si ce n'était pas le cas, i.e. si $Q(\lambda) = 0$, alors il existerait $\tilde{Q} \in \mathbf{K}[X]$ tel que $Q = (X - \lambda)\tilde{Q}$, donc $(X - \lambda)^k Q = (X - \lambda)^{k+1}\tilde{Q} = P$ — contradiction.

Définition/Proposition ALG.2.2 | Multiplicité « au moins »

Soient $P \in \mathbf{K}[X]$, et $\lambda \in \mathbf{K}$. On dit que λ est une racine de *multiplicité au moins k* de P si elle est de multiplicité $\ell \geq k$, i.e. si l'une des propositions ci-dessous est vérifiée :

$$(X - \lambda)^k \mid P \iff \exists Q \in \mathbf{K}[X], (X - \lambda)^k Q = P.$$

Si l'on considère $P = X^3$, on voit que 0 est une racine de multiplicité 3. Alors :

$$P(0) = P'(0) = P''(0), \quad \text{alors que} \quad P'''(0) = 6 \neq 0.$$

Ce fait s'étend à n'importe quel polynôme et caractérise même les racines multiples.


Proposition ALG.2.10 | Caractérisation à l'aide du polynôme dérivé

Soient $P \in \mathbf{K}[X]$, $P \neq 0_{\mathbf{K}[X]}$, et $\lambda \in \mathbf{K}$. Alors :


$$\begin{aligned} &\lambda \text{ est une racine d'ordre } k \in \mathbf{N} \text{ de } P \\ \iff &\text{pour tout } i \in \llbracket 0, k - 1 \rrbracket, \quad P^{(i)}(\lambda) = 0 \quad \text{et} \quad P^{(k)}(\lambda) \neq 0. \end{aligned}$$


Preuve Nous admettons cette proposition.


Exemple 4 — Notons $P = X^7 - 3X^5 + 2X^4 - X^3$, alors 0 est racine de multiplicité trois.

 En effet, $P(0) = 0$ et $P' = 7X^6 - 15X^4 + 8X^3 - 3X^2, P'' = 42X^5 - 45X^3 + 24X^2 - 6X, P''' = 210X^4 - 135X^2 + 48X - 6$, donc $P'(0) = P''(0) = 0$ mais $P'''(0) = -6 \neq 0$.

Exemple 5 — Soit $n \in \mathbf{N}$. On note $P = \sum_{k=0}^n \frac{X^k}{k!}$.

1. $P' = P - \frac{X^n}{n!}$,  Par linéarité de la dérivation, nous avons $P' = \sum_{k=1}^n \frac{kX^{k-1}}{k!} = \sum_{k=1}^n \frac{X^{k-1}}{(k-1)!} = P - \frac{X^n}{n!}$.

2. on déduit que les racines de P sont simples.  Soit $\alpha \in \mathbf{R}$ une racine de P , alors si α était une racine multiple, nous aurions $P'(\alpha) = 0 = 0 - \frac{\alpha^n}{n!}$ d'après la première question, donc $\alpha = 0$, or 0 n'est pas racine de P . D'où une contradiction et la non-existence d'une racine multiple pour P .

Exemple 6 — Soit $n \in \mathbf{N}$. On note $P = (X^2 - 1)^n$. Calculs $P^{(\ell)}(\pm 1)$ pour tout $\ell \geq 0$. 

► Un polynôme est dit *irréductible*, s'il ne peut pas être factorisé en produit de polynômes de degrés strictement inférieurs au sien.

3.2. Existence de racines & Comptage

On termine à présent le chapitre par probablement l'argument qui revient le plus souvent dans les exercices : le comptage des racines et la comparaison au degré.

Théorème ALG.2.1 | Comptage de racines

Soit $P \in \mathbf{K}[X]$. Alors :

1. si $P \neq 0_{\mathbf{K}[X]} \implies P$ possède au plus $\deg P$ racines comptées avec multiplicité.
2. P possède au moins $\deg P + 1$ racines $\implies P = 0_{\mathbf{K}[X]}$.

Preuve Commençons par 2.. Si P possède au moins $\deg P + 1$, alors il existe un polynôme $Q \in \mathbf{K}[X]$ tel que $(X - \alpha_1) \dots (X - \alpha_p)Q = P$, avec $\alpha_1, \dots, \alpha_p \in \mathbf{K}$ distincts deux à deux et $p = \deg P + 1$. Donc en passant au degré, nous avons $p + \deg Q = \deg P$ donc on aurait $\deg P + 1 \geq \deg P$ — contradiction puisque $\deg P \neq -\infty$. Pour 1., contraposer 2..

Corollaire ALG.2.1 | Polynôme s'annulant sur un ensemble infini

Soit $P \in \mathbf{K}[X]$ et \mathcal{D} un sous-ensemble infini de \mathbf{K} . Alors :

$$P \text{ est nul sur } \mathcal{D} \implies P = 0_{\mathbf{K}[X]}.$$

Preuve Puisque P est nul sur \mathcal{D} , il possède une infinité de racines donc *a fortiori* en possède plus que son degré. Il est donc nul.

Nous verrons dans la pratique comment factoriser un polynôme, mais avant définissons deux notions.

Définition ALG.2.7 | Polynôme scindé/irréductible

► Un polynôme $P \in \mathbf{K}[X]$ est dit *scindé sur \mathbf{K}* s'il s'écrit sous la forme

$$P(X) = \alpha \prod_{i=1}^r (X - \lambda_i)^{m_i}$$

où pour tout $i \in \llbracket 1, r \rrbracket$, $m_i = \text{Mult}_{\lambda_i}(P)$, $m_1 + \dots + m_r = \deg P$ et $\alpha \in \mathbf{K}$.
Si $\mathbf{K} = \mathbf{C}$, on dit que P est *scindé sur \mathbf{R}* si $\lambda_i \in \mathbf{R}$ pour tout i .⁸

⊗ Attention

Ces résultats sont caractéristiques des polynômes, pas question de les utiliser pour d'autres fonctions. Par exemple, \cos et \sin s'annulent une infinité de fois et ne sont pourtant pas identiquement nulles.

⁸Si $\mathbf{K} = \mathbf{C}$, on a seulement $\lambda_i \in \mathbf{C}$ si on écrit «scindé» tout court.

Méthode Montrer qu'un polynôme est nul

Pour montrer qu'un polynôme est nul, on peut au choix :

1. montrer que tous ses coefficients sont nuls,
2. montrer qu'il admet plus de racines que son degré (en particulier s'il en admet une infinité).



Le plus souvent, on utilise **2** pour en déduire la nullité de tous les coefficients.


Méthode Montrer que deux polynômes sont égaux


Pour montrer que deux polynômes sont égaux, on peut au choix :

1. montrer que leurs coefficients sont identiques,
2. montrer que la différence admet plus de racines que son degré (en particulier si elle en admet une infinité).

Exemple 7 — Soient $P, Q \in \mathbf{K}[X]$. A-t-on $P = Q$ dans les cas suivants ?

1. $\forall x \in \mathbf{R}, P(x) = Q(x)$?  *Le polynôme $P - Q$ possède alors tous les réels comme racines, donc en possède une infinité et a fortiori plus que son degré. Donc $P - Q = 0$ et $P = Q$.*
2. $\forall x \in]a, b[, a < b, P(x) = Q(x)$? La différence $b - a$ est-elle importante?  *Le polynôme $P - Q$ possède alors tous éléments de $]a, b[$ comme racines, donc en possède une infinité et a fortiori plus que son degré. Donc $P - Q = 0$ et $P = Q$. La différence $b - a$ (i.e. la longueur de l'intervalle) n'est pas importante étant donné que $]a, b[$ est toujours un ensemble infini.*

Exemple 8 — Il n'existe pas de polynôme $P \in \mathbf{R}[X]$ tel que pour tout $n \in \mathbf{N}$: $P(n) = \sqrt[3]{n^2 + 1}$.  *Supposons qu'un tel polynôme existe. Alors $P^3(n) - (n^2 + 1) = 0$ pour tout $n \in \mathbf{N}$. Donc le polynôme $P^3 - X^2 - 1$ possède tous les entiers comme racines, donc en possède une infinité et a fortiori plus que son degré. Ainsi, $P^3 - X^2 - 1 = 0$ et $P^3 = X^2 + 1$. En passant au degré on trouve $3 \deg P = 2$ si $P \neq 0$ donc c'est une contradiction. Mais $P = 0$ ne convient pas non plus car $0 \neq \sqrt[3]{n^2 + 1}$ pour tout $n \in \mathbf{N}$. Il n'existe donc pas de polynôme comme annoncé.*

Exemple 9 — **La conjugaison n'est pas polynomiale** Il n'existe pas de polynôme $P \in \mathbf{C}[X]$ tel que pour tout $z \in \mathbf{C}$: $P(z) = \bar{z}$.  *Supposons qu'un tel polynôme existe.*

Alors $P(x) - x = 0$ pour tout $x \in \mathbf{R}$. Donc le polynôme $P - X$ possède tous les réels comme racines, donc en possède une infinité et a fortiori plus que son degré. Ainsi, $P - X = 0$ i.e. $P = X$. On aurait alors :

$$\forall z \in \mathbf{R}, P(z) = z = \bar{z}.$$

Ce qui est clairement une contradiction, car il existe des complexes non réels.

Théorème ALG.2.2 | D'ALEMBERT-GAUB

Tout polynôme non constant de $\mathbf{K}[X]$ possède au moins une racine dans \mathbf{C} . En particulier, tout polynôme non constant de $\mathbf{K}[X]$ est scindé sur \mathbf{C} .

Attention

C'est faux sur $\mathbf{R}[X]$, par exemple $X^2 + 1$ n'est pas scindé sur \mathbf{R} , mais il est scindé sur \mathbf{C} .

Preuve Ce résultat est admis. La plupart des preuves sont difficiles et dépassent largement les programmes de CPGE.

Le **Théorème ALG.2.2** a donc pour conséquence que tout polynôme est scindé sur \mathbf{C} . Comment obtenir une telle décomposition ? Et sur \mathbf{R} ? Commençons par un lemme.

Lemme ALG.2.1 | Structure des racines d'un polynôme à coefficients réels

Soit $P \in \mathbf{R}[X]$. Les racines de P sont conjuguées.⁹

Preuve Montrons ceci par exemple avec les polynômes dérivés. Notons $P = \sum_{k=0}^n a_k X^k$ avec $n = \deg P$, $a_k \in \mathbf{R}$ pour tout $k \in \llbracket 0, n \rrbracket$. Alors soit $\lambda \in \mathbf{K}$ une racine de

⁹On peut montrer de plus que les multiplicités sont les mêmes

P i.e. telle que $P(\lambda) = 0$.

$$\begin{aligned}
 P(\lambda) = 0 &\iff \overline{P(\lambda)} = 0 \\
 &\iff \sum_{k=0}^n a_k \lambda^k = 0 \\
 &\iff \sum_{k=0}^n \overline{a_k \lambda^k} = 0 \\
 &\iff \sum_{k=0}^n a_k \bar{\lambda}^k = 0 \\
 &\iff P(\bar{\lambda}) = 0.
 \end{aligned}$$

propriété de la conjugaison
 $a_k \in \mathbf{R}$ pour tout k

Théorème ALG.2.3 | Forme des polynômes irréductibles

1. **(Dans $\mathbf{C}[X]$)** Les polynômes irréductibles sur \mathbf{C} sont les polynômes de degré inférieur ou égal à un.
2. **(Dans $\mathbf{R}[X]$)** Les polynômes irréductibles sur \mathbf{R} sont les polynômes de degré inférieur ou égal à un, ou de degré deux à discriminant strictement négatif.

Preuve

1. Tout polynôme possède une racine complexe. Donc tout polynôme de degré supérieur à deux est factorisable par un polynôme de degré strictement inférieur au sien (un polynôme de degré 1).
2. Soit $P \in \mathbf{R}[X]$ un polynôme irréductible. Alors s'il est de degré supérieur à trois, d'après D'ALEMBERT-GAUß, il existe une racine $\lambda \in \mathbf{C}$. Mais alors d'après le lemme qui précède, $\bar{\lambda}$ est aussi racine, donc $(X - \lambda)(X - \bar{\lambda}) \mid P$. Or,

$$(X - \lambda)(X - \bar{\lambda}) = X^2 - 2 \operatorname{Re}(\lambda)X + |\lambda|^2 \in \mathbf{R}[X].$$

Donc P n'est pas irréductible sur \mathbf{R} puisqu'on l'a factorisé par un polynôme de degré strictement inférieur au sien. Donc nécessairement :

- ▶ soit P est de degré 1, dans ce cas il est irréductible.
- ▶ Soit P est de degré 2, et irréductible si et seulement si il n'a pas de racine réelle, donc si et seulement si son discriminant est strictement négatif.

FACTORISATION DE POLYNÔMES. Tout polynôme s'écrit comme produit de polynômes irréductibles sur \mathbf{C} , mais sur \mathbf{R} également. L'énoncé de ce résultat n'est pas au pro-

gramme, mais vous devez savoir comment obtenir une telle décomposition. Revoyns les méthodes avant de faire des exemples.

 **Méthode Factorisation d'un polynôme**


Soit $P \in \mathbf{K}[X]$. Pour transformer P en un produit de polynômes de degré 1 ou 2, on :

1. cherche une racine $\lambda \in \mathbf{K}$.
2. On écrit P sous la forme $(X - \lambda) \times Q = P$ avec $Q \in \mathbf{K}[X]$.
3. On recommence le processus avec Q .


En résumé : cela revient à chercher les racines de P .

 **Méthode Lien entre la factorisation sur \mathbf{C} et \mathbf{R}**

Pour décomposer un polynôme en produit d'irréductibles dans $\mathbf{R}[X]$, on peut déjà le décomposer dans $\mathbf{C}[X]$ (en cherchant ses racines), puis on regroupe les racines complexes conjuguées entre elles.

Exemple 10 — Factoriser $P = X^4 - 2X^3 - 16X^2 + 2X + 15$. On pourra commencer par chercher une racine évidente  On constate que $P(-1) = 0 = P(1)$, donc il existe $a, b \in \mathbf{R}$ tels que $P = (X - 1)(X + 1)(X^2 + aX + b) = (X^2 - 1)(X^2 + aX + b)$. On peut ensuite identifier les coefficients, par exemple le terme constant et le terme d'ordre trois, on obtient alors comme conditions $\begin{cases} 15 = -b, \\ -2 = a, \end{cases}$ ce qui fournit ensuite $\begin{cases} b = -15, \\ a = -2. \end{cases}$ Donc $P = (X - 1)(X + 1)(X^2 - 2X - 15) = (X - 1)(X + 1)(X - 5)(X + 3)$. À la dernière étape on peut par exemple chercher les racines à l'aide du discriminant.

Exemple 11 — Factoriser

1. $X^2 - 2X \cos \theta + 1$ avec $\theta \in]0, \pi[$ sur $\mathbf{C}[X]$. Que dire sur \mathbf{R} ?  Calculons pour commencer le discriminant Δ : nous avons $\Delta = 4 \cos^2 \theta - 4 = -4 \sin^2 \theta < 0$. Nous avons donc deux racines complexes conjuguées $\frac{2 \cos \theta \pm 2i |\sin \theta|}{2} = e^{\pm i \theta}$. Donc la décomposition dans $\mathbf{C}[X]$ est :

$$X^2 - 2X \cos \theta + 1 = (X - e^{i\theta})(X + e^{-i\theta}).$$

Le polynôme est irréductible sur \mathbf{R} donc est déjà factorisé au maximum.

2. $X^4 + 1$ dans $\mathbf{C}[X]$ puis dans $\mathbf{R}[X]$.  *Commençons par chercher les racines de $X^4 + 1$ en utilisant les techniques du **Chapter ALG.1** : on utilise la forme trigonométrique de $X = \rho e^{i\theta}$ avec $\rho \geq 0, \theta \in [0, 2\pi[$. Nous avons*

$$\begin{aligned} X^4 + 1 = 0 &\iff X^4 = -1 = e^{i\pi} = \left(e^{\frac{i\pi}{4}}\right)^4, \\ &\iff \rho^4 e^{4i\theta} = e^{i\pi}, \\ &\iff \rho^4 = 1, \quad 4\theta = \pi \pmod{2\pi}, \\ &\iff \rho = 1, \quad \theta = \frac{\pi}{4} \pmod{\frac{\pi}{2}}. \end{aligned}$$

Donc :

$$X^4 + 1 = 0 \iff X \in \left\{ e^{\frac{i\pi}{4}}, e^{\frac{3i\pi}{4}}, -e^{\frac{i\pi}{4}}, -e^{\frac{3i\pi}{4}} \right\}.$$

Ainsi, la décomposition du polynôme dans $\mathbf{C}[X]$ est :

$$X^4 + 1 = \left(X - e^{\frac{i\pi}{4}}\right) \left(X - e^{\frac{3i\pi}{4}}\right) \left(X - \overline{e^{\frac{i\pi}{4}}}\right) \left(X - \overline{e^{\frac{3i\pi}{4}}}\right).$$

On regroupe ensuite les racines avec leur conjugué pour obtenir celle dans $\mathbf{R}[X]$:

$$\begin{aligned} X^4 + 1 &= \left(X - e^{\frac{3i\pi}{4}}\right) \left(X - \overline{e^{\frac{3i\pi}{4}}}\right) \left(X - e^{\frac{i\pi}{4}}\right) \left(X - \overline{e^{\frac{i\pi}{4}}}\right) \\ &= (X^2 - 2\cos(\pi/4)X + 1)(X^2 - 2\cos(3\pi/4)X + 1) \\ &= \boxed{(X^2 - \sqrt{2}X + 1)(X^2 + \sqrt{2}X + 1)}. \end{aligned}$$

3.3. Cas du second degré

Considérons un trinôme $aX^2 + bX + c$ avec $a, b, c \in \mathbf{R}$. Vous avez appris en première comment on trouvait ses racines, en effet, mettons le trinôme sous forme canonique. Notons $\delta \in \mathbf{C}$ une racine complexe de $\Delta = b^2 - 4ac$:

- ▶ si $\Delta \geq 0$, $\delta = \sqrt{\Delta}$ convient car $(\sqrt{\Delta})^2 = \Delta$.
- ▶ Si $\Delta < 0$, $\delta = i\sqrt{-\Delta}$ convient car $(i\sqrt{-\Delta})^2 = -(-\Delta) = \Delta$.

$$\begin{aligned} aX^2 + bX + c &= a \left[\left(X^2 + \frac{b}{a}X\right) + \frac{c}{a} \right] \\ &= a \left[\left(X + \frac{b}{2a}\right)^2 - \frac{b^2}{4a^2} + \frac{c}{a} \right] \quad \left. \vphantom{aX^2 + bX + c} \right\} \text{forme canonique} \\ &= a \left[\left(X + \frac{b}{2a}\right)^2 - \frac{1}{4a^2}(b^2 - 4ac) \right] \\ &= a \left[\left(X + \frac{b}{2a}\right)^2 - \left(\frac{\delta}{2a}\right)^2 \right]. \end{aligned}$$

On reconnaît alors une identité remarquable du type « $a^2 - b^2$ ». On écrit alors :

$$aX^2 + bX + c = a \left(X - \frac{-b + \delta}{2a}\right) \left(X - \frac{-b - \delta}{2a}\right).$$

On a trouvé les racines du trinôme, ce qui nous mène tout droit au théorème suivant.

Théorème ALG.2.4 | Solutions d'une équation du second degré

Soit $(a, b, c) \in \mathbf{R}^3$ tel que $a \neq 0$. On considère l'équation $az^2 + bz + c = 0$ d'inconnue $z \in \mathbf{C}$.


1. On appelle *discriminant* du trinôme $az^2 + bz + c$ le réel $\Delta = b^2 - 4ac$.
2. Soit δ **une** racine carrée de Δ (i.e. $\delta = \sqrt{\Delta}$ si $\Delta \geq 0$ et $\delta = i\sqrt{-\Delta}$ si $\Delta < 0$). Alors les solutions de l'équation sont :

$$\frac{-b + \delta}{2a} \quad \text{et} \quad \frac{-b - \delta}{2a}.$$

De plus, l'équation admet deux racines distinctes si $\Delta \neq 0$ et une seule racine si $\Delta = 0$.

Exemple 12 — Résoudre les équations :

1. $z^2 + 2z + 4 = 0$,
2. $z^2 - 2\cos(\theta)z + 1 = 0$, avec $\theta \in [0, 2\pi[$.

 **1.** On a $\Delta = 4 - 16 = -12$. On cherche ensuite une racine carrée complexe δ i.e. un élément $\delta \in \mathbf{C}$ vérifiant $\delta^2 = \Delta$. Puisque $-12 = (4i)^2$ on choisit $\delta = 4i$. Ainsi les solutions sont

$$\frac{-2 \pm (4i)}{2} = \boxed{-1 \pm 2i}.$$

Pour **2.**, on a $\Delta = -4 \sin^2 \theta$. Nous avons plusieurs cas :

- ▶ si $\theta = \pi$: alors $\Delta = 0$ et on a une seule racine double $\boxed{\cos \theta}$.
- ▶ si $\theta \in [0, 2\pi[\setminus \{\pi\}$, comme nous avons $\Delta = (2i \sin \theta)^2$, nous obtenons une racine carrée $\delta = 2i \sin \theta$. D'où les racines $\frac{2 \cos \theta \pm 2i \sin \theta}{2} = \boxed{e^{\pm i \theta}}$.

Notez que le second cas est inclus dans le premier.

RELATIONS COEFFICIENTS/RACINES Les relations coefficients/racines relient, comme leur nom l'indique, les coefficients d'un polynôme aux racines. Inutile en revanche d'espérer trouver les racines par cette méthode puisque elles sont solution d'un système linéaire, comme le montre la proposition qui suit.

Proposition ALG.2.11

Soit $P = aX^2 + bX + c \in \mathbf{K}[X]$, avec $a \in \mathbf{K}^*$, $b, c \in \mathbf{K}$. Notons x_1, x_2 les deux racines de P complexes conjuguées.¹⁰ Alors :

$$x_1 + x_2 = -\frac{b}{a}, \quad x_1 x_2 = \frac{c}{a}.$$

Preuve

- ▶ **(Première méthode : sans utiliser l'expression des racines.)** Par définition d'une racine, nous avons $P = a(X - x_1)(X - x_2)$. En développant, on obtient $P = a(X^2 - (x_1 + x_2)X + x_1 x_2)$. Par identification on obtient immédiatement :

$$a(x_1 + x_2) = b, \quad a x_1 x_2 = c,$$

ce qui en divisant par a donne les relations de l'énoncé.

- ▶ **(Seconde méthode : en utilisant l'expression des racines, si $\mathbf{K} = \mathbf{R}$ uniquement.)** Notons δ une racine carrée complexe de $\Delta = b^2 - 4ac$. Alors

$x_1 = \frac{-b+\delta}{2a}$, $x_2 = \frac{-b-\delta}{2a}$. On obtient alors :

$$x_1 + x_2 = \frac{(-b + \delta) + (-b - \delta)}{2a} = \frac{-b}{a},$$

$$x_1 x_2 = \frac{-b + \delta}{2a} \times \frac{-b - \delta}{2a} = \frac{b^2 - \delta^2}{4a^2} = \frac{b^2 - (b^2 - 4ac)}{4a^2} = \frac{c}{a}.$$

Remarque 3.1 — Et pour un degré quelconque

- ▶ Dans le TD, nous établirons les mêmes relations mais pour le degré trois, mais en utilisant la première méthode car nous ne connaissons pas d'expression explicite des racines.¹¹
- ▶ Il est complètement illusoire d'espérer calculer, de manière générale, les racines d'un polynôme à l'aide des relations coefficients/racines : pour le degré $n \in \mathbf{N}^*$, nous avons un système de n équations, mais absolument pas linéaire.

Une conséquence de la preuve du théorème précédent est le corollaire qui suit, qui est en quelque sorte une réciproque des relations coefficients/racines : si l'on se fixe $s, p \in \mathbf{K}$ alors il est possible de trouver un polynôme dont les racines ont pour somme s et produit p .

Corollaire ALG.2.2 | Quantités de somme et produit fixés

Soient $s, p \in \mathbf{K}$. Alors :

$$\begin{cases} x_1 + x_2 = s, \\ x_1 x_2 = p, \end{cases} \iff x_1, x_2 \text{ sont les racines de } X^2 - sX + p.$$

Preuve

\Leftarrow Déjà montré : conséquence des relations coefficients/racines (avec $a = 1$).

\Rightarrow Supposons que $\begin{cases} x_1 + x_2 = s, \\ x_1 x_2 = p. \end{cases}$ Alors notons $P = X^2 - sX + p$. On a par hypothèse $P = X^2 - (x_1 + x_2)X + x_1 x_2 = (X - x_1)(X - x_2)$, autrement dit x_1, x_2 sont les deux racines de P .

¹⁰Cela inclut tous les cas, y compris celui de racines doubles réelles


¹¹Mais de telles formules existent, dues à CARDAN.


**Méthode** Système à somme et produit fixés

Ainsi, pour résoudre le système

$$\begin{cases} x_1 + x_2 = s, \\ x_1 x_2 = p \end{cases}$$

en $(x_1, x_2) \in \mathbf{K}^2$, il suffit de chercher les racines de $X^2 - sX + p$ à l'aide du discriminant.

Exemple 13 — Résoudre le système $\begin{cases} x + y = 1 \\ xy = -1 \end{cases}$ sans substitution.  (x, y) sont solutions du système si et seulement si x, y sont les racines de $X^2 - 1X + (-1) = X^2 - X - 1$. Le discriminant est $\Delta = 1 + 4 = 5$. Donc $x = \frac{1+\sqrt{5}}{2}, y = \frac{1-\sqrt{5}}{2}$ ou $x = \frac{1-\sqrt{5}}{2}, y = \frac{1+\sqrt{5}}{2}$.

Exemple 14 — Soit $f \left| \begin{array}{l} \mathbf{R}^2 \longrightarrow \mathbf{R}^2, \\ (x, y) \longmapsto \left(\frac{x+y}{2}, \frac{2xy}{x+y} \right) \end{array} \right.$. Déterminer $f(\mathbf{R}^2)$.  Par définition, $(x', y') \in f(\mathbf{R}^2)$ si et seulement si, il existe $(x, y) \in \mathbf{R}^2$, tel que :

$$\begin{cases} \frac{x+y}{2} = x', \\ \frac{2xy}{x+y} = y' \end{cases} .$$

Il s'agit donc de résoudre le système — non linéaire! — précédent en (x, y) , et éventuellement de trouver une condition sur (x', y') pour avoir l'existence d'une solution. L'ensemble image sera alors l'ensemble des couples (x', y') pour lesquels le système admet **au moins** une solution.

$$\begin{aligned} \begin{cases} \frac{x+y}{2} = x', \\ \frac{2xy}{x+y} = y' \end{cases} &\iff \begin{cases} x + y = 2x', \\ 2xy = 2x'y' \end{cases} &\iff \begin{cases} x + y = 2x', \\ xy = x'y' \end{cases} \\ &\iff x, y \text{ sont racines de } X^2 - (2x')X + x'y' \end{aligned}$$

Or, le discriminant de ce trinôme est $4x'^2 - 4x'y' = 4x'(x' - y')$. Donc $(x', y') \in f(\mathbf{R}^2)$ si et seulement si $\Delta = 4x'(x' - y') \geq 0$, donc en conclusion

$$f(\mathbf{R}^2) = \left\{ (x', y') \in \mathbf{R}^2, x'(x' - y') \geq 0 \right\} .$$

4. EXERCICES

Exercice ALG.2.1 | Vrai ou Faux ?

1. Pour tout $P \in \mathbf{R}[X]$, $\deg(-P) = \deg P$.
2. Pour tous $P, Q \in \mathbf{R}[X]$, $\deg(P - Q) \leq \deg P - \deg Q$.
3. Un polynôme constant est de degré nul.
4. Le polynôme $X - 2$ divise $X^5 - 3X^4 - 2X^3 + 3X^2 + 7X + 6$.
5. Le polynôme $X^{17} + X + 1$ est irréductible dans $\mathbf{R}[X]$.
6. Si $z \in \mathbf{C}$ est une racine de multiplicité n d'un polynôme P , alors $P^{(n)}(z) = 0$.

4.1. Généralités, Racines, Factorisation

Exercice ALG.2.2 | Polynôme des différences finies Soit $P \in \mathbf{K}[X]$. Exprimer le degré de $P(X + 1) - P(X)$ en fonction de celui de P .

Solution (exercice ALG.2.2)

Soit $n \geq 0$, $P = \sum_{k=0}^n a_k X^k$, avec $a_0, \dots, a_n \in \mathbf{K}$. Alors

$$P(X + 1) - P(X) = \sum_{k=0}^n a_k \left[(X + 1)^k - X^k \right].$$

Dès lors, cherchons le degré de $(X + 1)^n - X^n$, qui est égale à degré de $P(X + 1) - P(X)$. D'après le formule du binôme

$$(X + 1)^n - X^n = \sum_{\ell=0}^n \binom{n}{\ell} X^\ell - X^n,$$

les termes d'ordre n étant identiques, nous avons une simplification si $n \geq 1$:

$$(X + 1)^n - X^n = \sum_{\ell=0}^{n-1} \binom{n}{\ell} X^\ell,$$

et $\binom{n-1}{n} \neq 0$. Ainsi, le polynôme $P(X + 1) - P(X)$ est de degré

- ▶ $k - 1$ si $k \geq 1$.

- ▶ Si $k = 0$ ou $-\infty$, alors $P(X + 1) - P(X) = 0$ donc il est de degré $-\infty$.
- ▶

Exercice ALG.2.3 | Quel est l'ordre de multiplicité de 1 dans $P_n = nX^{n+2} - (n+2)X^{n+1} + (n+2)X - n$ pour tout $n \in \mathbf{N}$?

Solution (exercice ALG.2.3)

Calculons les dérivées successives de P .

$$\begin{aligned} P(1) &= 0, \\ P'(X) &= n(n+2)X^{n+1} - (n+2)(n+1)X^n + (n+2), \\ P'(1) &= 0. \end{aligned}$$

Pour pouvoir redériver, nous avons besoin de distinguer des cas sur n .

- ▶ Si $n = 0$, alors $P_0 = 0$, donc 1 est racine de multiplicité zéro (convention).
- ▶ Si $n \geq 1$, $P_n''(X) = n(n+1)(n+2)X^n - n(n+1)(n+2)X^{n-1} = n(n+1)(n+2)X^{n-1}(X-1)$. Donc 1 est racine de multiplicité 1 de P'' et donc $P''(1) = 0$ mais $P''''(1) \neq 0$. On déduit alors que 1 est racine de multiplicité 3 de P_n .

Exercice ALG.2.4 | Factorisation Déterminer pour chaque exemple si Q est factorisable par P , et déterminer le cas échéant un polynôme R tel que $Q = PR$.

- ▶ $P = X - 1$ et $Q = X^3 - 2X^2 + 3X - 2$,
- ▶ $P = X - 2$ et $Q = X^4 - 3X^3 + X + 1$,
- ▶ $P = X^2$ et $Q = (X + 1)^n - nX - 1$ avec $n \in \mathbf{N}$.

Solution (exercice ALG.2.4)

On commence par vérifier si les racines du diviseur testé sont bien des racines du polynôme P .

- ▶ $Q(1) = 1 - 2 + 3 - 2 = 0$, donc 1 est racine de P, et **P divise Q**. On cherche R sous la forme $R = X^2 + aX + b$ avec $a, b \in \mathbf{R}$, donc tel que :

$$X^3 - 2X^2 + 3X - 2 = (X - 1)(X^2 + aX + b),$$

donc en développant :

$$X^3 - 2X^2 + 3X - 2 = X^3 + X^2(a - 1) + X(b - a) - b,$$

on peut ensuite identifier coefficient par coefficient.

$$-2 = a - 1, \quad 3 = b - a, \quad -b = -2 \iff b = 2, \quad a = -1.$$

Donc **$R = X^2 - X + 2$** .

- ▶ $Q(2) = 2^4 - 3 \cdot 8 + 2 + 1 = -8 + 3 + 5 \neq 0$, donc **P ne divise pas Q**.
- ▶ Pour savoir si P divise Q, il s'agit de regarder si 0 est une racine de multiplicité 2 de Q.

$$Q(1) = 0 \quad \forall n \in \mathbf{N}.$$

Et $Q'(X) = n(X + 1)^{n-1} - n$, $Q'(0) = n - n = 0$ dès que $n \geq 1$, puis $Q''(X) = n(n - 1)X^{n-2}$, $Q''(0) = 0$ dès que $n \geq 2$.

- si $n = 0$: $Q = 1 + 1 = 2$ donc **P ne divise pas Q**.
- Si $n = 1$: $Q = (X + 1) - X - 1 = 0$ donc **P ne divise pas Q**.
- Si $n \geq 2$, alors d'après les calculs précédents **P divise Q**.

Exercice ALG.2.5 | Soit $P = (X + 1)^7 - X^7 - 1$. On rappelle que $j = e^{i \frac{2\pi}{3}}$.

1. Calculer $1 + j + j^2$, que dire de j^3 ? Montrer que $j^2 = \bar{j}$.
2. Montrer que : $P = \sum_{k=1}^6 \binom{7}{k} X^k$.
3. Montrer que P est divisible par $(X - j)^2$, puis que P est divisible par $(X - \bar{j})^2$.
4. Factoriser P en produit d'irréductibles dans $\mathbf{C}[X]$, puis dans $\mathbf{R}[X]$.

Solution (exercice ALG.2.5)

1. $1 + j + j^2 = 0$ (somme géométrique), et $j^3 = 1$. Par ailleurs $j^2 = e^{i \frac{4\pi}{3}} = e^{i \frac{4\pi}{3} - 2i\pi} = e^{-i \frac{2\pi}{3}} = \bar{j}$.

2. D'après la formule du binôme, on a : $P = \sum_{k=0}^7 \binom{7}{k} X^k - X^7 - 1 = \sum_{k=1}^6 \binom{7}{k} X^k$. Le

coefficient dominant est donc $\binom{7}{6} = 7$. Le degré est donc 6.

3. Il s'agit de montrer que $P(j) = P'(j) = 0$ — cela signifie que j est une racine de multiplicité au moins égale à deux. On a :

$$P(j) = (1 + j)^7 - j^7 - 1 = (-j^2)^7 - j^7 - 1 = -j^14 - j^7 - 1 = -j^2 - j - 1 = 0,$$

$$P'(j) = 7(1 + j)^6 - 7j^6 = 7(-j^2)^6 - 7 = 7 - 7 = 0.$$

Dès lors **$(X - j)^2 \mid P$** . Puisque $P \in \mathbf{R}[X]$, on sait d'après le cours que \bar{j} est également une racine de P et de même multiplicité. Donc **\bar{j} est une racine de multiplicité au moins deux**.

4. Nous avons déjà deux racines, chacune de multiplicité au moins deux. Or P est de degré six, il en manque donc deux. Constatons que 0, -1 sont deux racines évidentes. Or, le coefficient dominant de P est 7, donc

$$P = 7(X - j)^2(X - \bar{j})^2X(X + 1).$$

C'est la décomposition en irréductibles dans $\mathbf{C}[X]$. Pour obtenir des facteurs réels, on regroupe les parties complexes avec leur version conjuguée.

$$P = 7(X - j)^2(X - \bar{j})^2X(X + 1)$$

$$= 7(X^2 - 2 \operatorname{Re}(j)X + |j|^2)^2X(X + 1)$$

$$= 7(X^2 + X + 1)^2X(X + 1).$$

Exercice ALG.2.6 | Pour tout entier $p \in \llbracket 0, n \rrbracket$, on considère les polynômes :

$$T_p = X^p (1 - X)^p, \quad L_p = \frac{1}{p!} (T_p)^{(p)}$$

où $(T_p)^{(p)}$ désigne la dérivée d'ordre p de T_p . Fixons un entier $p \in \mathbf{N}$.

- Déterminer le degré et le coefficient dominant de L_p .
- Notons L_p sous la forme $L_p = \sum_{k=0}^p a_{k,p} X^k$. Établir que :

$$\forall k \in \llbracket 0, p \rrbracket \quad a_{k,p} = (-1)^k \binom{p}{k} \binom{p+k}{k}$$

- Déterminer une relation entre $a_{k+1,p}$ et $a_{k,p}$ pour $k \in \llbracket 0, p-1 \rrbracket$. Préciser la valeur de $a_{0,p}$.
- En vous appuyant sur la question 3, écrire une fonction informatique qui prend en argument un entier naturel p et renvoie la liste des coefficients $a_{0,p}, a_{1,p}, \dots, a_{p,p}$ de L_p .
Tester cette fonction dans le cas où $p \in \{0, 1, 2\}$.

Exercice ALG.2.7 | Relations coefficients/racines pour l'ordre trois.

- Soit le polynôme $P = aX^3 + bX^2 + cX + d \in \mathbf{C}[X]$ supposé scindé, avec $a \neq 0$ et de racines x, y, z . Exprimer $\sigma_1 = x + y + z$, $\sigma_2 = xy + xz + yz$ et $\sigma_3 = xyz$ en fonction de a, b, c, d . Il existe des relations similaires pour les autres degrés, appelées relations coefficients/racines.
- (Application à la résolution d'un système non linéaire)** Résoudre dans \mathbf{C}^3 les

$$\text{systemes : } \begin{cases} x + y + z = 2, \\ x^2 + y^2 + z^2 = 2, \\ xyz = 0. \end{cases}$$

Solution (exercice ALG.2.7)

- Par définition d'une racine, $P = a(X - x)(X - y)(X - z)$. Développons ce produit :

$$\begin{aligned} P &= a(X - x)(X - y)(X - z) \\ &= a(X^2 - (x + y)X + xy)(X - z) \\ &= a(X^3 - (x + y)X^2 + xyX - zX^2 + z(x + y)X - xyz) \\ &= aX^3 - a(x + y + z)X^2 + a(xy + xz + zy)X - axyz. \end{aligned}$$

En identifiant les coefficients, on trouve : $-a\sigma_1 = b, \quad a\sigma_2 = c \quad \text{et} \quad d = -a\sigma_3$.
De manière équivalente :

$$\sigma_1 = -\frac{b}{a}, \quad \sigma_2 = \frac{c}{a}, \quad \sigma_3 = -\frac{d}{a}.$$

- Notons $\sigma_1, \sigma_2, \sigma_3$ les trois fonctions de x, y, z définies dans l'énoncé. Constatons que $x^2 + y^2 + z^2 = (x + y + z)(x + y + z) - 2(xy + xz + yz) = \sigma_1^2 - 2\sigma_2$. Alors :

$$\begin{aligned} \begin{cases} x + y + z = 2, \\ x^2 + y^2 + z^2 = 2, \\ xyz = 0. \end{cases} &\iff \begin{cases} \sigma_1 = 2, \\ \sigma_1^2 - 2\sigma_2 = 2, \\ \sigma_3 = 0. \end{cases} \\ &\iff \begin{cases} \sigma_1 = 2, \\ \sigma_2 = 1, \\ \sigma_3 = 0. \end{cases} \\ &\iff x, y, z \text{ sont les racines de} \\ &\quad X^3 - \sigma_1 X^2 + \sigma_2 X - \sigma_3 = X^3 - 2X^2 + 2X = X(X^2 - 2X + 2). \end{aligned}$$

Les racines de ce polynôme sont composées de 0 et des racines de $X^2 - 2X + 2$ (ce dernier polynôme est de discriminant $4 - 8 = -4 = (2i)^2$). Donc :

$$x, y, z \in \{0, 1 - i, 1 + i\}.$$

4.2. Équations fonctionnelles polynomiales

- Exercice ALG.2.8 |** Résoudre dans l'ensemble des fonctions polynomiales l'équation $P \circ P = P$.

Solution (exercice ALG.2.8)

On commence par regarder le degré d'une éventuelle fonction solution. On constate que nécessairement $(\deg P)^2 = \deg P$. Ainsi $\deg P \in \{0, 1\}$. Les polynômes constants

sont tous solutions. Si P est de degré 1, il s'écrit a priori sous la forme $P = \alpha X + \beta$ et α, β vérifient pour tout $x \in \mathbf{R}$

$$\alpha^2 x + \alpha\beta + \beta = \alpha x + \beta.$$

Donc

$$\alpha^2 X + \alpha\beta + \beta = \alpha X + \beta.$$

Et $\alpha \in \{0, 1\}$. Le cas $\alpha = 0$ a déjà été traité, sinon $\alpha = 1$ et $\beta = 0$. Donc X est aussi solution. Mais

$$j\sqrt[3]{-q} \times j^2\sqrt[3]{-q} = \sqrt[3]{-q},$$

donc le polynôme $X^3 + p$ satisfait la condition.

Exercice ALG.2.9 |

1. Soit $P \in \mathbf{R}[X]$ vérifiant : $(P')^2 = 4P$. Montrer que si P n'est pas constant alors $\deg P = 2$.
2. En déduire l'ensemble des polynômes de $\mathbf{R}[X]$ vérifiant $(P')^2 = 4P$.

Solution (exercice ALG.2.9)

1. Supposons P non constant, i.e. $n = \deg P \geq 1$. Alors en passant au degré dans l'hypothèse vérifiée par P, nous obtenons

$$2(n - 1) = n, \quad \boxed{n = 2}.$$

2. Le seul polynôme constant solution est le polynôme nul. Supposons que $P = aX^2 + bX + c$ avec $a, b, c \in \mathbf{R}$. Alors

$$\begin{aligned} (P')^2 = 4P &\iff (2aX + b)^2 = 4(aX^2 + bX + c), \\ &\iff 4a^2X^2 + 4aX + b^2 = 4(aX^2 + bX + c), \\ &\iff a^2 = a, a = b, b^2 = c. \end{aligned}$$

Donc :

- ▶ soit $a = 0$, auquel cas $b = 0 = c$, donc $P = 0$ - ceci est exclu.
- ▶ Soit $a \neq 0$, et donc $a = 1$, ce qui libre $b = 1$ puis $c = 1$. Donc $P = X^2 + X + 1$.

L'ensemble des polynômes solution est donc $\boxed{\{0, X^2 + X + 1\}}$.

4.3. Familles classiques

Exercice ALG.2.10 | Autour des polynômes d'HERMITE On considère la suite (H_n) de polynômes, telle que $H_0(X) = 1$, et pour tout $n \in \mathbf{N}$,

$$H_{n+1} = H'_n - 2XH_n.$$

1. Justifier que pour tout $n \in \mathbf{N}$, $H_n \in \mathbf{R}[X]$, et donner H_1, H_2 .
2. Déterminer le degré de H_n pour tout $n \in \mathbf{N}$.
3. Déterminer le coefficient dominant de H_n pour tout $n \in \mathbf{N}$.

Solution (exercice ALG.2.10)

1. Le fait que $H_n \in \mathbf{R}[X]$ pour tout $n \in \mathbf{N}$ est une récurrence évidente puisque la dérivée d'un polynôme à coefficients réels est encore à coefficients réels. De plus, $H_1 = 1' - 2X1 = -2X$, puis $H_2 = (-2)' - 2X(-2X) = 4X^2 - 2 = \boxed{2(2X^2 - 1)}$.
2. Déterminons le degré de H_n pour tout $n \in \mathbf{N}$. Soit $n \in \mathbf{N}$, alors notant $d_n = \deg P_n$, on a en passant au degré puisque le degré de XH_n est nécessairement strictement supérieur à celui de H'_n :

$$d_{n+1} = d_n + 1, \quad d_0 = 0.$$

Donc $\boxed{d_n = n}$ pour tout $n \in \mathbf{N}$.

3. Notons a_n le coefficient dominant de P_n pour tout entier $n \in \mathbf{N}$. Pour la même raison qu'à la question précédente, le coefficient dominant de $H'_n - 2XH_n$ est égal à celui de $-2XH_n$. On obtient alors :

$$a_{n+1} = (-2)a_n.$$

Donc $\boxed{a_n = (-2)^n a_0 = (-2)^n}$ pour tout $n \in \mathbf{N}$.



Chapitre ALG.3.

Espaces Vectoriels

Résumé & Plan

Beaucoup de notions ci-dessous ont déjà été exposées en première année pour l'espace vectoriel \mathbf{K}^n avec $\mathbf{K} = \mathbf{R}$ ou \mathbf{C} . Vous aviez constaté que l'on pouvait additionner des vecteurs, les multiplier par élément de \mathbf{K} et donc plus généralement réaliser des combinaisons linéaires entre eux. Nous allons plus généralement réaliser ce type d'opérations sur des polynômes, des suites, des fonctions ... tous ces ensembles munis des lois d'addition et de multiplication de \mathbf{K} seront aussi appelés des *espaces-vectoriels*.

Nous allons considérer pour terminer des applications entre les espaces vectoriels, qui seront des applications entre les deux ensembles sous-jacents et vérifiant une propriété *ad hoc* dite de *linéarité*. Ces applications seront aussi des applications en tant qu'ensembles : il est donc important de revoir le chapitre de première année traitant du sujet (notion d'injectivité, surjectivité, bijectivité, d'application réciproque *etc.*).

1	Structure d'espace vectoriel	2
1.1	Généralités	2
1.2	Combinaisons linéaires & Sous-espaces vectoriels	6
2	Familles de vecteurs	15
2.1	Famille libre	16

2.2	Familles génératrices	20
2.3	Base	21
2.4	Extraction & Complétion	23
3	Dimension	26
3.1	Généralités	26
3.2	Sous-espaces et dimension	29
3.3	Rang	29
4	Applications linéaires	30
4.1	Généralités	30
4.2	Structure de \mathbf{K} -espace vectoriel sur $\mathcal{L}_{\mathbf{K}}(E, F)$ et puissances	34
4.3	Image & Noyau	36
4.4	Isomorphismes	41
4.5	Cas particulier de la dimension finie	42

5	Exercices	46
5.1	Structure d'espace vectoriel	46
5.2	Familles de vecteurs et dimension	47
5.3	Applications linéaires	53

*Le travail est l'activité vitale propre au travailleur,
l'expression personnelle de sa vie.*

— Emmanuel Kant

Commençons par introduire une notation, qui nous servira dans la plupart des chapitres d'Algèbre qui suivront.

Σ Notation Ensemble des applications

Soient E, F deux ensembles. On notera E^F (ou $\mathcal{F}(F, E)$) l'ensemble des applications de F dans E .

Exemple 1 — Par exemple, $\mathbf{C}^{\mathbf{N}}$ (*resp.* $\mathbf{R}^{\mathbf{N}}$) désigne l'ensemble des suites à valeurs complexes (*resp.* à valeurs réelles). Et $\mathbf{R}^{\mathbf{R}} = \mathcal{F}(\mathbf{R}, \mathbf{R})$ l'ensemble des fonctions de \mathbf{R} dans \mathbf{R} , $\mathbf{R}^{\mathbf{R}^2}$ l'ensemble des fonctions de deux variables à valeurs réelles.

⚙️ Cadre

Dans tout le chapitre, l'ensemble K désignera \mathbf{R} ou \mathbf{C} . Ainsi, tous les énoncés faisant intervenir K sont vrais que K soit \mathbf{R} ou \mathbf{C} .

INTÉRÊT DES RAISONNEMENTS ALGÈBRIQUES Ce chapitre — ainsi que quelques autres pendant l'année — s'inscrit dans le domaine de l'Algèbre en Mathématiques, dont la vocation principale est l'étude d'ensembles munis de lois, et la recherche d'un cadre commun à plusieurs objets mathématiques. Une fois ce cadre dégagé (voir la **Définition ALG.3.1** ci-dessous) nous l'étudierons en détail. Tous les résultats établis dans cette étude se transmettront donc automatiquement aux objets qui vérifient la **Définition ALG.3.1**.

COMMENT MÉMORISER FACILEMENT LES DÉFINITIONS? 🎉 Gardez à l'esprit que toutes les notions qui vont être présentées, quoique abstraites à première vue, sont des généralisations de quelque chose de concret que vous connaissez déjà (vecteurs, repères, coordonnées, ...).

1. STRUCTURE D'ESPACE VECTORIEL

En première année, vous aviez vu des opérations sur les éléments de \mathbf{K}^n avec $n \geq 1$ un entier, plus précisément :

- ▶ vous savez additionner deux vecteurs : si $X = (x_1, \dots, x_n), Y = (y_1, \dots, y_n) \in \mathbf{K}^n$, alors on pose :

$$X + Y = (x_1 + y_1, \dots, x_n + y_n) \in \mathbf{K}^n.$$

C'est l'addition « coordonnée par coordonnée », on l'appelle la *loi interne*.

- ▶ Vous savez multiplier un vecteur par un scalaire : si $X = (x_1, \dots, x_n) \in \mathbf{K}^n, \lambda \in \mathbf{K}$, alors on pose :

$$\lambda \cdot X = (\lambda x_1, \dots, \lambda x_n) \in \mathbf{K}^n.$$

C'est la multiplication « coordonnée par coordonnée », on l'appelle la *loi externe*.

De manière plus général on sait réaliser ce type d'opérations sur des fonctions, des polynômes, des complexes, des suites, *etc.* tous ces ensembles munis de ces lois seront qualifiés d'espaces-vectoriels, dont voici la définition.

1.1. Généralités

Définition ALG.3.1 | Espace vectoriel

On appelle *K-espace vectoriel* (ou *espace vectoriel sur K*) tout triplet $(E, +, \cdot)$ où :

1. $(E, +)$ est un *groupe commutatif i.e.* :

(i) $+$ est une *loi interne* :

$$+ \left\{ \begin{array}{l} E \times E \longrightarrow E, \\ (x, y) \longmapsto x + y. \end{array} \right.$$

(ii) $+$ est *associative* : $\forall x, y, z \in E, (x + y) + z = x + (y + z),$

(iii) **(Existence d'un élément neutre 0_E pour $+$)** $\forall x \in E, x + 0_E = 0_E + x,$

(iv) tout élément de E est inversible pour $+$: $\forall x \in E, \exists y \in E, x + y = y + x = 0_E.$ L'élément y inverse de x sera noté $-x.$

(v) La loi $+$ est *commutative, i.e.* $\forall x, y \in E, x + y = y + x.$

2. La loi \cdot est une *loi de externe*

$$\cdot \left\{ \begin{array}{l} K \times E \longrightarrow E \\ (\lambda, x) \longmapsto \lambda \cdot x \end{array} \right.$$

telle que pour tout $(\lambda, \mu) \in K^2$ et tout $(x, y) \in E^2$ on ait les *règles de calcul suivantes* entre $+$ et \cdot :

(i) $(\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x,$

(ii) $\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y,$

(iii) $(\lambda \times \mu) \cdot x = \lambda \cdot (\mu \cdot x),$

(iv) $1 \cdot x = x.$

Les éléments de K sont appelés les *scalaires*, les éléments de E sont les *vecteurs*, K est le *corps de base*, la loi $+$ est appelée *addition* et la loi \cdot est appelée *multiplication par un scalaire*. L'élément neutre de E pour la loi $+$ est appelé *vecteur nul*, et on le notera 0_E comme précédemment, ou simplement 0 s'il n'y a pas d'ambiguïté.

Toutes les propriétés précédentes sont des règles de calcul dans E , et sont comparables aux règles habituelles sur les nombres réels que vous utilisez depuis l'école primaire. Rassurez-vous, nous n'aurons pas la plupart du temps à vérifier l'ensemble des axiomes pour montrer qu'un ensemble est un espace vectoriel, nous aurons la notion de sous-espace vectoriel qui nous permettra de gagner en efficacité.

Attention

Dans un espace vectoriel, vous pouvez multiplier les vecteurs $x \in E$ par des **scalaires** $\lambda \in K$, **mais pas** multiplier deux vecteurs $x, y \in E$ entre eux, en règle générale.¹

PREMIERS EXEMPLES. Donnons sans plus tarder quelques exemples d'espaces vectoriels. La vérification complète des axiomes ne présente aucune difficulté et est laissée au lecteur, seuls quelques-uns seront précisés.

Exemple 2 — Cas de $E = K^n$ – Uplets Soient $X = (x_1, \dots, x_n), Y = (y_1, \dots, y_n) \in K^n$, et $\lambda \in K$, alors on définit :

$$X + Y = (x_1 + y_1, \dots, x_n + y_n), \quad \lambda X = (\lambda x_1, \dots, \lambda x_n).$$

Le triplet $(E, +, \cdot)$ est un espace vectoriel sur K . L'élément 0_E est ici le n -uplet nul $0_E = (0, \dots, 0)$, l'élément opposé pour $+$ de $X = (x_1, \dots, x_n)$ est $-X = (-x_1, \dots, -x_n)$ car $X + (-X) = 0_E.$

Exemple 3 — Cas de $E = \mathcal{M}_{n,p}(K)$ avec $n, p \in N^*$ – Matrices Soient $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}, B = (b_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} \in K^n$, et $\lambda \in K$, alors on définit :

$$A + B = (a_{i,j} + b_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}, \quad \lambda A = (\lambda a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}.$$

Le triplet $(E, +, \cdot)$ est un espace vectoriel sur K . L'élément 0_E est ici la matrice nulle de format $n \times p$ $0_E = (0)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$, l'élément opposé pour $+$ de $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$ est $-A = (-a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$ car $A + (-A) = 0_E.$

Exemple 4 — Cas de $E = \mathcal{F}(N, K) = K^N$ – Suites Soient $(u_n), (v_n) \in E$, et $\lambda \in K$, alors on définit :

$$(u_n) + (v_n) = (u_n + v_n), \quad \lambda (u_n) = (\lambda u_n).$$

¹ Même si dans certains espaces vectoriels c'est possible, par exemple il est possible de multiplier deux polynômes ou deux matrices compatibles entre elles.

Le triplet $(E, +, \cdot)$ est un espace vectoriel sur \mathbf{K} . L'élément 0_E est ici la suite nulle $0_E = (0)$, l'élément opposé pour $+$ de (u_n) est $(-u_n)$ car pour tout $n \in \mathbf{N}$, $u_n + (-u_n) = 0$.

Exemple 5 — Cas des polynômes Nous avons défini dans le **Chapter ALG.2** deux lois $+$, \cdot sur les polynômes. Le triplet $(E, +, \cdot)$ est un espace vectoriel sur \mathbf{K} . L'élément 0_E est ici le polynôme nul $0_E = 0$, l'élément opposé pour $+$ de P est $-P$ car $P + (-P) = 0$.

Exemple 6 — Cas de $E = \mathcal{F}(\mathbf{K}, \mathbf{K})$ – Fonctions Soient $f, g \in E$, et $\lambda \in \mathbf{K}$, alors on définit :

$$f + g \left| \begin{array}{l} \mathbf{K} \longrightarrow \mathbf{K}, \\ x \longrightarrow f(x) + g(x), \end{array} \right. \quad \lambda f \left| \begin{array}{l} \mathbf{K} \longrightarrow \mathbf{K}, \\ x \longrightarrow \lambda \cdot f(x). \end{array} \right.$$

Le triplet $(E, +, \cdot)$ est un espace vectoriel sur \mathbf{K} . L'élément 0_E est ici la fonction nulle $0_E : x \in \mathbf{K} \longrightarrow 0$, l'élément opposé pour $+$ de f est $-f : x \in \mathbf{K} \longrightarrow -f(x)$ car pour tout $x \in \mathbf{K}$, $f(x) + (-f(x)) = 0$.

L'**Exemple 6** peut se généraliser en la proposition suivante : l'espace de départ peut être un ensemble général F et l'ensemble d'arrivée un espace vectoriel E .

Proposition ALG.3.1 | Espaces de fonctions

Soient F un ensemble non vide et E un \mathbf{K} -espace vectoriel. L'ensemble E^F des applications de F dans E , muni des lois :

1. $\forall (f, g) \in (\mathcal{F}(F, E))^2, \quad (f + g) \left| \begin{array}{l} F \longrightarrow E, \\ x \longrightarrow f(x) + g(x), \end{array} \right.$
2. $\forall \lambda \in \mathbf{K}, \forall f \in \mathcal{F}(F, E), \quad (\lambda \cdot f) \left| \begin{array}{l} F \longrightarrow E \\ x \longrightarrow \lambda \cdot f(x) \end{array} \right.$

est un \mathbf{K} -espace vectoriel. Son élément neutre est donc $0_{\mathcal{F}(F, E)}$, i.e. la fonction

$$\text{nulle} \left| \begin{array}{l} F \longrightarrow E \\ x \longrightarrow 0_E \end{array} \right.$$

Preuve Vérifier les différents axiomes.

Remarque 1.1 — Il faut bien comprendre que dans la **Proposition ALG.3.1**, il est nécessaire que **l'espace d'arrivée** soit un espace vectoriel pour donner un sens à $f(x) + g(x)$, $\lambda f(x)$ pour tout $x \in F$, $(f, g) \in \mathcal{F}(F, E)$ et $\lambda \in \mathbf{R}$. En revanche, on voit bien que ce n'est pas indispensable pour l'espace de départ.

À PROPOS DU CORPS DE BASE. Lorsque $\mathbf{K} = \mathbf{C}$, on constate que l'on peut voir un \mathbf{C} -espace vectoriel comme un \mathbf{R} -espace vectoriel.

Proposition ALG.3.2 | Restriction du corps de base

Tout \mathbf{C} -espace vectoriel est *a fortiori* un \mathbf{R} -espace vectoriel.²

Preuve



Soit E un \mathbf{C} -espace vectoriel. Il suffit de restreindre la loi externe

$$\cdot : \mathbf{C} \times E \longrightarrow E, \quad \text{à} \quad \cdot|_{\mathbf{R} \times \mathbf{C}} : \mathbf{R} \times E \longrightarrow E,$$

et on vérifie ensuite facilement que cette restriction définit une \mathbf{R} -loi externe sur E . La loi $+$ est inchangée.

PRODUIT CARTÉSIEN D'ESPACES VECTORIELS. Soient E_1, \dots, E_n des ensembles. On rappelle que l'ensemble $E_1 \times \dots \times E_n$ est appelé le *produit cartésien* des $E_i, i \in \llbracket 1, n \rrbracket$, et ses éléments sont notés (x_1, \dots, x_n) avec $x_1 \in E_1, \dots, x_n \in E_n$. En d'autres termes :

$$E_1 \times \dots \times E_n = \{(x_1, \dots, x_n), x_1 \in E_1, \dots, x_n \in E_n\}.$$

²De manière plus informelle : si on sait multiplier un vecteur par un complexe, on sait en particulier le multiplier par un réel.

Proposition ALG.3.3 | Produit cartésien

Soient E_1, \dots, E_n des \mathbf{K} -espaces vectoriels. On munit $E = E_1 \times \dots \times E_n$ de deux lois $+$ et \cdot de la manière suivante :

1. (Addition coordonnée par coordonnée)

$$+ \begin{cases} E \times E & \longrightarrow E, \\ ((x_1, \dots, x_n), (y_1, \dots, y_n)) & \longmapsto (x_1 + y_1, \dots, x_n + y_n). \end{cases}$$

2. (Multiplication coordonnée par coordonnée par un scalaire)

$$\cdot \begin{cases} \mathbf{K} \times E & \longrightarrow E, \\ (\lambda, (y_1, \dots, y_n)) & \longmapsto (\lambda \cdot y_1, \dots, \lambda \cdot y_n). \end{cases}$$

Alors :

$(E_1 \times \dots \times E_n, +, \cdot)$ est un \mathbf{K} -espace vectoriel,³ avec pour élément neutre

$$0_{E_1 \times \dots \times E_n} = (0_{E_1}, \dots, 0_{E_n}).$$

Autrement dit, le neutre du produit cartésien est le n -uplet des neutres.

Preuve Simple vérification de tous les axiomes.

La proposition précédente généralise alors l'**Exemple 2**, lorsque

$$E_1 = \dots = E_n = \mathbf{K}.$$

RÈGLES DE CALCULS SECONDAIRES. De la définition d'un espace vectoriel découlent directement d'autres propriétés.

Proposition ALG.3.4 | Autres règles de calcul

Soit E un \mathbf{K} -espace vectoriel. Pour tout $(\lambda, \mu) \in \mathbf{K}^2$ et tout $(x, y) \in E^2$, on a :

1. (Développement d'une expression) $(\lambda - \mu) \cdot x = \lambda \cdot x - \mu \cdot x, \quad \lambda \cdot (x - y) = \lambda \cdot x - \lambda \cdot y,$

2. (Multiplication par zéro) $0_{\mathbf{K}} \cdot x = 0_E, \quad \lambda \cdot 0_E = 0_E,$


3. (Multiplication par l'opposé) $(-\lambda) \cdot x = -(\lambda \cdot x) = \lambda \cdot (-x),$

4. (Produit nul) $\lambda \cdot x = 0_E \iff (\lambda = 0_{\mathbf{K}} \text{ ou } x = 0_E).$

Remarque 1.2 — La dernière assertion n'a *a priori* rien d'évident, d'ailleurs elle est même fautive pour d'autres ensembles que $\mathbf{K} = \mathbf{R}$ ou \mathbf{C} , mais dans notre cadre (*cf.* début de chapitre) ce sera toujours le cas.

³On retrouve que \mathbf{K}^n est un \mathbf{K} -espace vectoriel.

Preuve Par exemple, pour 4) :

▶ si $\lambda \neq 0_{\mathbf{K}}$,  alors on peut multiplier l'hypothèse par $\frac{1}{\lambda}$ à gauche :

$$\frac{1}{\lambda} (\lambda x) = \frac{1}{\lambda} 0_E = 0_E$$

d'où $x = 0_E,$

▶ sinon $\lambda = 0_{\mathbf{K}}$.

Nous admettons le reste dont la preuve ne présente pas de difficulté.

Corollaire ALG.3.1

Soit E un \mathbf{K} -espace vectoriel. Alors si E n'est pas $\{0_E\}$, il n'est jamais de cardinal fini.

Preuve Par définition d'un espace vectoriel, E est stable par multiplication par un scalaire. Donc si $x \neq 0_E$ est dans E ,

$$D_x \stackrel{\text{(déf.)}}{=} \{\lambda x, \lambda \in \mathbf{K}\} \subset E,$$

mais tous les éléments de D_x sont distincts (car sinon il existe $\lambda, \mu \in \mathbf{K}$ tels que $\lambda x = \mu x$, ce qui implique d'après la propriété précédente que $\lambda = \mu$ puisque x est non nul). Donc D_x n'est pas un ensemble fini et E non plus.

Attention

On n'écrira donc jamais $\# E$ pour E un espace vectoriel non réduit à son élément neutre.

Maintenant que le cadre est posé, regardons ce que l'on peut faire avec les vecteurs, *i.e.* les éléments d'un espace vectoriel. En combinant les deux lois définies plus haut (additive et scalaire-multiplicative), on arrive directement à la notion de combinaison linéaire présentée ci-après.

1.2. Combinaisons linéaires & Sous-espaces vectoriels

Définition ALG.3.2 | Combinaison linéaire de vecteurs

Soit E un \mathbf{K} -espace vectoriel.

1. (**Famille finie**) Soient x_1, \dots, x_n des vecteurs de E . On appelle *combinaison linéaire* des vecteurs x_1, \dots, x_n tout vecteur $x \in E$ s'écrivant sous la forme

$$x = \sum_{k=1}^n \lambda_k x_k \quad \text{où pour tout } k \in \llbracket 1, n \rrbracket, \text{ on a } \lambda_k \in \mathbf{K}.$$

Si $\mathbf{K} = \mathbf{C}$, alors on appelle *\mathbf{R} -combinaison linéaire* tout vecteur x s'écrivant sous la forme

$$x = \sum_{k=1}^n \lambda_k x_k \quad \text{où pour tout } k \in \llbracket 1, n \rrbracket, \text{ on a } \lambda_k \in \mathbf{R}.$$

2. (**Famille quelconque**) Soit $(x_i)_{i \in I}$ une famille de vecteurs de E , où I est un ensemble quelconque. On appelle *combinaison linéaire* des vecteurs $(x_i)_{i \in I}$ toute combinaison linéaire **finie** de ces vecteurs.

Si $\mathbf{K} = \mathbf{C}$, alors on appelle *\mathbf{R} -combinaison linéaire* des vecteurs $(x_i)_{i \in I}$ toute **\mathbf{R} -combinaison linéaire finie** de ces vecteurs.

Σ Notation

On note généralement :

- ▶ $\text{Vect}_{\mathbf{K}}(x_1, \dots, x_n)$ l'ensemble des combinaisons linéaires d'une famille finie de vecteurs, ou plus simplement $\text{Vect}(x_1, \dots, x_n)$ s'il n'y a pas d'ambiguïté,
- ▶ $\text{Vect}_{\mathbf{K}}(x_i)_{i \in I}$ l'ensemble des combinaisons linéaires d'une famille quelconque de vecteurs, ou plus simplement $\text{Vect}(x_i)_{i \in I}$ s'il n'y a pas d'ambiguïté.

Les ensembles de **\mathbf{R} -combinaisons linéaires** seront notés $\text{Vect}_{\mathbf{R}}(\dots)$.

Proposition ALG.3.5

Soit E un \mathbf{C} -espace vectoriel, et $x_1, \dots, x_n \in E$ Alors :

$$\text{Vect}_{\mathbf{R}}(x_1, \dots, x_n) \subset \text{Vect}_{\mathbf{C}}(x_1, \dots, x_n).$$

Les **\mathbf{R} -combinaisons linéaires** sont des **\mathbf{C} -combinaisons linéaires**.⁴

Preuve


**Méthode Montrer l'appartenance « à un Vect »**


Pour montrer que $x \in \text{Vect}(x_i)_{i \in I}$, on cherche $\lambda_1, \dots, \lambda_n$ des scalaires, tels que : $x = \sum_{i=1}^n \lambda_i x_i$. En particulier, si $E = \mathbf{K}^n$, on tombe sur la résolution d'un système linéaire.


Exemple 7 — Est-on combinaison linéaire de ?

1. (**Dans \mathbf{R}^2**) Est-ce que le vecteur $u = (3, 3)$ est combinaison linéaire des vecteurs $a = (1, 1)$, $b = (1, 0)$ et $c = (0, 1)$? Y a-t-il unicité de l'écriture de u comme combinaison linéaire de a , b et c ?

⁴Rien d'étonnant à cela, un réel est un complexe!


2. **(Dans \mathbf{R}^3)** Le vecteur $u = (-1, 2, 2)$ de \mathbf{R}^3 est-il combinaison linéaire des vecteurs $a = (1, 1, 0)$, $b = (-2, 1, 3)$ et $c = (1, 0, -1)$? 

3. **(Dans $\mathfrak{M}_2(\mathbf{R})$)** La matrice $\begin{pmatrix} -1 & 2 \\ 2 & 0 \end{pmatrix}$ n'est pas combinaison linéaire des matrices $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} -2 & 1 \\ 3 & 2 \end{pmatrix}$ et $\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$. 

4. (Dans $\mathbf{R}^{\mathbf{R}}$) La fonction $x \mapsto \cos^2 x$ est-elle combinaison linéaire de $x \mapsto 1$ et $x \mapsto \cos(2x)$? 

5. (Dans $\mathbf{R}^{\mathbf{N}^*}$) La suite $(1)_{n \in \mathbf{N}^*}$ est-elle combinaison linéaire de $(n)_{n \in \mathbf{N}^*}$ et $(\frac{1}{n})_{n \in \mathbf{N}^*}$?



6. (Dans $\mathbf{R}[X]$) Soit $P \in \mathbf{R}[X]$. Peut-on avoir $P \in \text{Vect}(1, P')$? 

Attention Identification

On retiendra du premier exemple qu'*a priori* on ne peut pas identifier de manière systématique⁵ :

$$\left(\sum_{k=1}^n \lambda_k x_k = \sum_{k=1}^n \mu_k x_k \right) \Leftrightarrow (\forall k \in \llbracket 1, n \rrbracket, \lambda_k = \mu_k).$$

Exemple 8 — Toute fonction polynomiale P de degré inférieur ou égal à n est combinaison linéaire de 1, X, X², ..., Xⁿ, puisqu'il peut être écrit sous la forme $P = \sum_{k=0}^n a_k X^k$, où les a_k sont dans K. On peut donc résumer cela en :

$$\text{Vect}(1, X, \dots, X^n) = \mathbf{K}_n[X].$$

SOUS-ESPACES VECTORIELS. La structure de sous-espace vectoriel aura un intérêt pour justifier qu'un ensemble est un espace vectoriel. En effet, nous allons voir que si un ensemble est un sous-espace vectoriel d'un espace vectoriel connu, alors ce sous-ensemble sera aussi un espace vectoriel (et il est plus facile de vérifier la propriété de sous-espace vectoriel que d'espace vectoriel).

⁵Nous appellerons plus tard *famille libre* toute famille où c'est le cas.

Définition ALG.3.3 | Sous-espace vectoriel

On appelle *sous-espace vectoriel* d'un K-espace vectoriel E tout ensemble F tel que :

- ▶ F ⊂ E,
- ▶ 0_E ∈ F,
- ▶ F est *stable par combinaison linéaire* : $\forall (\lambda, \mu) \in \mathbf{K}^2, \forall (x, y) \in F, \lambda x + \mu y \in F.$

Voici la proposition principale, qui nous permettra de vérifier facilement que des ensembles sont des espaces vectoriels.

Proposition ALG.3.6

Soit (E, +, ·) un K-espace vectoriel, et F un sous-espace vectoriel de (E, +, ·). Alors :

$$(F, +, \cdot)^6 \text{ est un K-espace vectoriel.}$$

Preuve Simple vérification des axiomes de la **Définition ALG.3.1**.

Méthode Montrer qu'un ensemble n'est pas un espace vectoriel

Pour montrer qu'un ensemble n'est pas un espace vectoriel, on peut montrer qu'il n'est pas un sous-espace vectoriel d'un espace vectoriel de référence. En

- ▶ vérifiant qu'il ne contient pas le neutre dudit ensemble de référence.
- ▶ Ou, on montre qu'il n'est pas stable par combinaison linéaire.

Remarque 1.3 — Économie de rédaction L'énorme intérêt de la notion de sous-espace vectoriel réside dans la proposition précédente : une économie dans la rédaction. En effet, il nous suffira de vérifier qu'une structure est un sous-espace vectoriel d'un espace vectoriel du cours (fait que vous avez le droit d'utiliser sans argument supplémentaire), pour justifier la structure espace vectoriel. Ce qui semble beaucoup plus rapide que la vérification complète de la **Définition ALG.3.1**.

Avant de regarder quelques exemples, commençons par montrer une propriété très utile dans la pratique : «les Vect» sont des sous-espaces vectoriels!

⁶De manière plus rigoureuse, il faudrait noter $(F, +|_{F \times F}, \cdot|_{\mathbf{K} \times F})$ au lieu de (F, +, ·), puisque les ensembles de départ ne sont pas les mêmes pour les deux lois.

Définition/Proposition ALG.3.1 | Un «Vect» est un espace vectoriel.

Soient E un \mathbf{K} -espace vectoriel et $(x_i)_{i \in I}$ une famille de vecteurs de E , où I est un ensemble. Alors :

$\text{Vect}(x_i)_{i \in I}$ est le plus petit sous-espace vectoriel de E contenant $(x_i)_{i \in I}$.⁷

Lorsque la famille est réduite à un unique vecteur $u \in E$, on appelle $\text{Vect}(u)$ la *droite vectorielle* engendrée par u .

**Méthode Montrer qu'un ensemble est un espace vectoriel**

Deux options sont possibles.

1. Justifier qu'il est un sous-espace vectoriel d'un espace vectoriel de référence (\mathbf{K}^n , de polynômes, de fonctions, de suites ...) **ou**
2. Montrer que l'ensemble s'écrit comme Vect d'une famille.

Preuve (dans le cas d'une famille finie de deux vecteurs) Faisons la preuve dans le cas d'une famille de deux vecteurs (les autres cas ne présentent pas de difficulté supplémentaire). Soient $x_1, x_2 \in E$, montrons que $F = \text{Vect}(x_1, x_2)$ est le plus petit sous-espace vectoriel de E contenant x_1, x_2 .

- ▶ Montrons que $\text{Vect}(x_1, x_2)$ est un sous-espace vectoriel de E contenant x_1, x_2 .



- ▶ Montrons que $\text{Vect}(x_1, x_2)$ est le plus petit sous-espace vectoriel vérifiant cette propriété, *i.e.* soit F un sous-espace vectoriel de E tel que $x_1, x_2 \in F$ et montrons que $\text{Vect}(x_1, x_2) \subset F$.

⁷Et c'est donc en particulier un espace vectoriel.

Exemple 9 — dans \mathbf{K}^n

1. Un \mathbf{K} -espace vectoriel E admet toujours comme sous-espaces vectoriels $\{0_E\}$ et E lui-même.
2. L'ensemble $F = \{\lambda(2, 3) = (2\lambda, 3\lambda) \in \mathbf{R}^2, \lambda \in \mathbf{R}\} = \text{Vect}(2, 3)$ est un sous-espace vectoriel de \mathbf{R}^2 appelé *droite vectorielle de vecteur directeur* $(2, 3)$.⁸




3. L'ensemble $F = \{\lambda(1, 2, 0) = (\lambda, 2\lambda, 0) \in \mathbf{R}^3, \lambda \in \mathbf{R}\} = \text{Vect}(1, 2, 0)$ est un sous-espace vectoriel de \mathbf{R}^3 appelé *droite vectorielle de vecteur directeur* $(1, 2, 0)$.
4. L'ensemble $F = \{\lambda(1, 2, 0) + \mu(1, 0, 1) = (\lambda + \mu, 2\lambda, \mu) \in \mathbf{R}^3, (\lambda, \mu) \in \mathbf{R}^2\}$ est un sous-espace vectoriel de \mathbf{R}^3 appelé *plan vectoriel de vecteurs directeurs* $(1, 2, 0)$ et $(1, 0, 1)$.
 - ▶ **(1ère méthode : avec la définition)**




⁸Mais la Définition/Proposition ALG.3.1 généralise cela.

de $\mathcal{C}^1(\mathbf{R}, \mathbf{R})$.


- ▶ (1ère méthode : avec la définition) 

▶ (2ème méthode : c'est un Vect) $F = \text{Vect}((1, 2, 0), (1, 0, 1))$.


5. L'ensemble $F = \{(x, y, z, t) \in \mathbf{K}^4, x + y + z + t = 0, x - y = 0\}$ est un sous-espace vectoriel de \mathbf{K}^4 . Montrons qu'il s'agit d'un Vect. 

- ▶ (2ème méthode : c'est un Vect) 

Exemple 10 — dans des espaces de fonctions


1. Soient I un intervalle non trivial de \mathbf{R} et $n \in \mathbf{N} \cup \{\infty\}$. L'ensemble $\mathcal{C}^n(I, \mathbf{R})$ des fonctions \mathcal{C}^n est un sous-espace vectoriel de \mathbf{R}^I . 
- ▶ l'inclusion $\mathcal{C}^n(I, \mathbf{R}) \subset \mathbf{R}^I$ est évidente car une fonction \mathcal{C}^n est en particulier une fonction,
 - ▶ la fonction nulle est de classe \mathcal{C}^n ,
 - ▶ une combinaison linéaire de fonctions \mathcal{C}^n est \mathcal{C}^n d'après le cours de première année.
2. L'ensemble $F = \{f \in \mathcal{C}^1(\mathbf{R}, \mathbf{R}), \forall x \in \mathbf{R}, f'(x) = 3f(x)\}$ ⁹ est un sous-espace vectoriel

⁹on reconnaît ici l'ensemble des fonctions solution d'une équation différentielle linéaire d'ordre 1.

3. L'ensemble $F = \{f \in \mathcal{C}^1(\mathbf{R}_+, \mathbf{R}), \forall x \in \mathbf{R}_+, f'(x) = 3f(x)^2\}$ ¹⁰ est-il un sous-espace vectoriel de $\mathcal{C}^1(\mathbf{R}, \mathbf{R})$? *Indication* : Commencer par trouver une fonction non nulle dans cet ensemble. 



⁹Nous systématiserons la propriété d'espace vectoriel dans le [Chapter ANA.9](#)

¹⁰La présence d'un carré vous guide sur la réponse : non.

4. L'ensemble des solutions d'une équation différentielle **linéaire homogène** sur un intervalle I (du premier ou du second ordre) est un sous-espace vectoriel de $\mathcal{F}(I, \mathbf{K})$ ¹¹. Nous le montrerons dans le **Chapter ANA.9** de révisions.
5. Soit $E = \mathcal{F}(\mathbf{R}, \mathbf{R})$. Montrer que l'ensemble $F = \mathcal{P}$ des fonctions paires de \mathbf{R} dans \mathbf{R} et l'ensemble \mathcal{I} des fonctions impaires de \mathbf{R} dans \mathbf{R} sont des sous-espaces vectoriels de E . Nous le montrerons dans le **Chapter ANA.7** de révisions.
6. Soit $E = \mathcal{C}^0([0, 1], \mathbf{R})$. Montrer que l'ensemble $F = \left\{ f \in E, \int_0^1 f(t) dt = 0 \right\}$ est un sous-espace vectoriel de E . 


¹¹Ceci généralise donc l'exemple précédent

Exemple 11 — dans des espaces de suites

1. Notons F l'ensemble des suites réelles convergentes, alors F est un sous-espace vectoriel de $\mathbf{R}^{\mathbf{N}}$. 
2. Soit $(a, b) \in \mathbf{R}^2$. L'ensemble $F = \left\{ (u_n) \in \mathbf{R}^{\mathbf{N}}, \forall n \geq 0, u_{n+2} = au_{n+1} + bu_n \right\}$ est un sous-espace vectoriel de $\mathbf{R}^{\mathbf{N}}$.
 ▶ (1ère méthode : avec la définition) 

► (2ème méthode : c'est un Vect) 

3. L'ensemble $F = \{(u_n) \in \mathbf{R}^{\mathbf{N}}, u_n \underset{n \rightarrow \infty}{=} o(1)\}$ est un sous-espace vectoriel de $\mathbf{R}^{\mathbf{N}}$.


Peut-on remplacer $(1)_n$ (dans le $o(\cdot)$) par une autre suite? 


4. L'ensemble des suites vérifiant une **relation de récurrence linéaire** est un sous-espace vectoriel de $\mathbf{K}^{\mathbf{N}}$. Nous le montrerons dans le [Chapter ANA.9](#) de révisions.

Exemple 12 — *dans des espaces de matrices colonnes* L'ensemble des solutions d'un système **linéaire homogène** à coefficients dans \mathbf{K} et à p inconnues est un sous-espace vectoriel de \mathbf{K}^p . Notons $A \in \mathfrak{M}_{n,p}(\mathbf{K})$ la matrice du système homogène $AX = 0$.



Exemple 13 — *dans des espaces de polynômes*

1. Soit $n \geq 1$. Alors $\mathbf{K}_n[X]$ — l'ensemble des polynômes de degré inférieur à n — est un sous-espace vectoriel de $\mathbf{K}[X]$.  En effet, par définition d'un polynôme, $\mathbf{K}_n[X] = \text{Vect}(1, X, \dots, X^{n-1})$.

2. Soit $n \geq 1$. En revanche $\mathbf{K}_{=n}[X]$ — l'ensemble des polynômes de degré n — n'est pas un sous-espace vectoriel de $\mathbf{K}[X]$. Pourquoi?  $\mathbf{K}_{=n}[X]$ ne contient pas le polynôme nul, qui est de degré $-\infty$.


INTERSECTION & RÉUNION D'ESPACES VECTORIELS.

Proposition ALG.3.7

Soit E un \mathbf{K} -espace vectoriel.

1. Toute intersection de sous-espaces vectoriels de E est encore un sous-espace vectoriel de E .
2. Une réunion de sous-espaces vectoriels de E **n'est en général pas** un sous-espace vectoriel de E .

Preuve

1. Considérons, pour simplifier la rédaction, une famille finie F_1, \dots, F_n de sous-espaces vectoriels de E . Montrons que $F = F_1 \cap \dots \cap F_n$ est un sous-espace vectoriel de E . 

2. 

CAS PARTICULIER : DESCRIPTION DES SOUS-ESPACES VECTORIELS DE $\mathbf{K}^n = \mathbf{R}^n$ OU \mathbf{C}^n . De façon générale, il existe principalement trois types de description des sous-espaces de \mathbf{K}^n , que nous avons déjà rencontrés dans des précédents exemples. Vous devez savoir passer de l'une à l'autre. Par exemple, la droite D de \mathbf{R}^2 dirigée par $(1, 1)$ peut s'écrire :

- ▶ comme un Vect, c'est $F = \text{Vect}(1, 1) = \{(\lambda, \lambda), \lambda \in \mathbf{R}\}$. On parle aussi de *forme paramétrique*.
- ▶ Ou à l'aide d'une ou plusieurs équations cartésiennes, ici $F = \{(x, y) \in \mathbf{R}^2, y = x\}$. On parle de *forme cartésienne*.

Nous allons avoir comment faire pour, de manière générale, passer d'une forme à une autre. On pourrait montrer de manière générale que tout sous-espace vectoriel de \mathbf{K}^n peut être exprimé avec l'une ou l'autre des descriptions, mais nous nous contenterons d'exemples.


Méthode Lien entre paramétrisation & équations cartésiennes


- ▶ **Paramétrisation** → **Équations implicites / cartésiennes** : Résolution d'un système en les paramètres de la forme paramétrique (généralement notés λ, μ, ν, \dots). Le système présentera alors une condition de compatibilité qui sera l'équation cartésienne cherchée.
- ▶ **Équations implicites / cartésiennes** → **Paramétrisation** : Résolution du système linéaire en les inconnues (généralement notées x, y, z, \dots), l'ensemble des solutions correspond alors à la forme paramétrique cherchée.

Attention

Cette méthode doit être parfaitement maîtrisée.

Exemple 14 — Reprenons les deux espaces vectoriels définis plus haut, qui sont des sous-espaces vectoriels de $E = \mathbf{R}^4$.

1. $F = \text{Vect}(X, Y) \subset \mathbf{K}^4$ où $X = (1, 2, 1, 1)$ et $Y = (0, 1, 1, 1)$. Déterminer un système d'équations caractérisant F . 

2. $G = \{(x, y, z, t) \in \mathbf{K}^4, x + y + z + t = 0 \text{ et } x = y\} \subset \mathbf{K}^4$. Déterminer une forme paramétrique de G . 

D'autres exemples seront traités en TD.

2. FAMILLES DE VECTEURS

Notation Famille/Ensemble ?

Soit E un ensemble et $x_1, \dots, x_n \in E$ avec $n \in \mathbf{N}$. Rappelons les deux notations suivantes :

1. (x_1, \dots, x_n) désigne le n -uplet x_1, \dots, x_n donc *a priori*

$$(x_1, x_2, x_3, \dots, x_n) \neq (x_2, x_1, x_3, \dots, x_n).$$

On parle de *famille*, l'ordre des éléments a une importance.

2. $\{x_1, \dots, x_n\}$ désigne l'ensemble formé des éléments x_1, \dots, x_n donc *a priori*

$$\{x_1, x_2, x_3, \dots, x_n\} = \{x_2, x_1, x_3, \dots, x_n\}.$$

On parle d'*ensemble*, et dans ce cas l'ordre des éléments n'a aucune importance.

L'objectif de cette section est cette fois-ci d'abstraire la notion de repère du plan *i.e.* la faculté de *caractériser*, et de manière unique, les vecteurs par des coordonnées, sauf que maintenant nos vecteurs peuvent être des polynômes, des fonctions, des suites *etc....*

Le vocabulaire général pour les espaces vectoriels est plutôt le suivant :

- ▶ les repères seront appelés des *bases*,
- ▶ et le nombre d'éléments d'un repère sera appelé la *dimension*. Nous allons donc également devoir justifier que toutes les bases ont même cardinal afin que la définition soit bien posée.

2.1. Famille libre

Définition ALG.3.4 | Famille libre/liée

- (Famille finie)** Soit $(x_k)_{1 \leq k \leq n}$ une famille finie de vecteurs d'un \mathbf{K} -espace vectoriel E . On dit que $(x_k)_{1 \leq k \leq n}$ est une *famille libre* de vecteurs de E , ou que les vecteurs x_1, \dots, x_n sont *linéairement indépendants*, si :

$$\forall (\lambda_k)_{1 \leq k \leq n} \in \mathbf{K}^n, \left[\sum_{k=1}^n \lambda_k x_k = 0_E \implies \forall k \in \llbracket 1, n \rrbracket, \lambda_k = 0 \right].$$

- (Famille quelconque)** Soit $(x_i)_{i \in I}$ une famille de vecteurs de E indexée par un ensemble I . On dit que $(x_i)_{i \in I}$ est libre dans E si toute sous-famille **finie** est libre. Si ce n'est pas le cas, on dit que la famille est liée.
 - ▶ Dans \mathbf{R}^2 ou \mathbf{R}^3 , un couple (u, v) de vecteurs liés est un couple de vecteurs *colinéaires*.
 - ▶ Dans un \mathbf{R}^3 , un triplet (u, v, w) de vecteurs liés est un triplet de vecteurs *coplanaires*.

Remarque 2.1 — Lien avec l'identification Quitte à remplacer λ_i par $\lambda_i - \mu_i$ pour tout i dans la définition, nous pouvons aussi dire que : $(x_k)_{1 \leq k \leq n}$ une famille finie

de vecteurs d'un \mathbf{K} -espace vectoriel E si l'on peut identifier deux combinaisons linéaires

$$\forall (\lambda_k)_{1 \leq k \leq n}, (\mu_k)_{1 \leq k \leq n} \in \mathbf{K}^n, \left[\sum_{k=1}^n \lambda_k x_k = \sum_{k=1}^n \mu_k x_k \implies \forall k \in \llbracket 1, n \rrbracket, \lambda_k = \mu_k \right].$$

Remarque 2.2 — Négation de la liberté

$$\exists (\lambda_k)_{1 \leq k \leq n} \in \mathbf{K}^n, \sum_{k=1}^n \lambda_k x_k = 0_E, \text{ mais } \exists k \in \llbracket 1, n \rrbracket, \lambda_k \neq 0.$$

Cela signifie donc qu'il existe une combinaison linéaire nulle (des vecteurs x_1, \dots, x_n) dont tous les coefficients ne sont pas nuls.




Méthode Montrer la liberté/liaison d'une famille

- Pour montrer qu'une famille $(x_k)_{1 \leq k \leq n}$ est libre, on écrit :
 « Soit $(\lambda_k)_{1 \leq k \leq n} \in \mathbf{K}^n$ tel que $\sum_{k=1}^n \lambda_k x_k = 0_E$. Alors [...] donc les λ_k sont tous nuls. » En général, l'étape [...] consiste en les arguments suivants :
 - ▶ Dans $\mathbf{R}^n, \mathbf{C}^n$ ou $\mathbf{K}_n[X]$: on arrive à la résolution d'un système linéaire.
 - ▶ Dans $\mathbf{R}^1, \mathbf{R}^n$ on fait de l'analyse (limites, dérivation, *etc.*).
- Pour montrer qu'une famille $(x_k)_{1 \leq k \leq n}$ est liée, on écrit :
 « [...] Posons alors $\lambda_1 = \dots, \dots, \lambda_n = \dots$: on a alors $\sum_{k=1}^n \lambda_k x_k = 0_E$, mais les λ_k ne sont pas tous nuls. »


Exemple 15 —

- On considère les vecteurs de \mathbf{R}^3 suivants $x_1 = (1, 0, 2)$, $x_2 = (0, 1, 1)$ et $x_3 = (1, 0, 1)$. La famille (x_1, x_2, x_3) est-elle libre?

2. Montrons que (\cos, \sin) est libre dans $\mathcal{F}(\mathbf{R}, \mathbf{R})$.

▶ (1ère méthode : en évaluant en plusieurs x) 

3. Montrons que $(1, (2^n)_n, (3^n)_n)$ est libre dans $\mathbf{R}^{\mathbf{N}}$. Généraliser.

▶ (1ère méthode : en évaluant en plusieurs n) 

▶ (2ème méthode : avec des limites) 

▶ (2ème méthode : avec des développements limités) 

Attention Égalité à zéro dans un espace vectoriel de fonctions



L'égalité $\lambda \cos + \mu \sin = 0_{\mathcal{F}(\mathbf{R}, \mathbf{R})}$, signifie que :


$$\forall x \in \mathbf{R}, \quad \lambda \cos(x) + \mu \sin(x) = 0,$$

puisque la fonction $0_{\mathcal{F}(\mathbf{R}, \mathbf{R})}$ est la fonction nulle. Le quantificateur « \forall » est ici fondamental pour démontrer la liberté, attention aux oublis ou aux mélanges dans

 votre rédaction. La même remarque s'applique aussi aux suites.

Plus généralement, nous avons :

Égalité	Signification
Pour des uplets : $(x_1, \dots, x_n) = (y_1, \dots, y_n)$	$\forall i \in \{1, \dots, n\}, x_i = y_i,$
Pour des suites : $(u_n) = (v_n)$	$\forall n, u_n = v_n.$
Pour des polynômes : $\sum_{k=1}^n a_k X^k = \sum_{k=1}^n b_k X^k$	$\forall k \in \{1, \dots, n\}, a_k = b_k,$
Pour des fonctions : $f = g$	$\forall x, f(x) = g(x).$

 **Cadre**
 Dans la suite sur les familles libres nous travaillerons uniquement avec des familles finies, même si l'ensemble des résultats s'adaptent aux familles quelconques.

Proposition ALG.3.8 | Propriétés des familles liées

Soit $(x_k)_{1 \leq k \leq n}$ une famille finie de vecteurs d'un \mathbf{K} -espace vectoriel E.

1. Si l'un des vecteurs x_{k_0} est nul, alors la famille $(x_k)_{1 \leq k \leq n}$ est liée.
2. Si l'un des vecteurs x_{k_0} apparaît plus d'une fois dans la famille $(x_k)_{1 \leq k \leq n}$, alors celle-ci est liée.

Preuve

1. Supposons pour simplifier que $k_0 = 1$, donc que $x_0 = 0$. Alors

$$1 \cdot x_0 + 0 \cdot x_1 + \dots + 0 \cdot x_n = 0,$$

il s'agit donc d'une combinaison linéaire nulle, mais dont tous les coefficients ne sont pas nuls (le premier). La famille est donc liée.

2. Supposons pour simplifier que $k_0 = 1$, et que $x_0 = x_1$. Alors

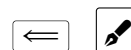
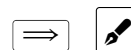
$$1 \cdot x_0 + (-1) \cdot x_1 + 0 \cdot x_2 + \dots + 0 \cdot x_n = 0,$$

il s'agit donc d'une combinaison linéaire nulle, mais dont tous les coefficients ne sont pas nuls (les deux premiers). La famille est donc liée.

Proposition ALG.3.9 | Caractérisation des familles liées

Soit $(x_k)_{1 \leq k \leq n}$ une famille finie de vecteurs d'un \mathbf{K} -espace vectoriel E. Alors :
 $(x_k)_{1 \leq k \leq n}$ est liée,
 \iff un de ses vecteurs s'exprime comme combinaison linéaire des autres vecteurs de la famille, i.e. $\exists k_0 \in \llbracket 1, n \rrbracket, \exists (\lambda_k)_{1 \leq k \leq n} \in \mathbf{K}^n, x_{k_0} = \sum_{\substack{1 \leq k \leq n \\ k \neq k_0}} \lambda_k x_k.$

Preuve





FAMILLE ÉCHELONNÉE DE POLYNÔMES. Passons à un exemple fondamental de famille libre de polynômes, celles où les degrés ont une forme particulière.

Définition ALG.3.5 | Famille échelonnée de polynômes

Une famille de polynômes est une *famille échelonnée* si les degrés des polynômes sont deux à deux distincts.

Exemple 16 — Les familles ci-dessous sont-elles échelonnées?

1. $(X^2, X - 1, 10)$, 

2. $(X(X - 1)(X - 2), X, (X + 1)^4 - X^4)$. 

On en vient maintenant à la propriété principale de ces familles.

Théorème ALG.3.1 | Familles échelonnées de polynômes

Toute famille finie de polynômes de degrés échelonnés non nuls de $\mathbf{K}[X]$ est libre.

Attention

La réciproque est fautive : il existe des familles libres non échelonnées de polynômes.¹²

¹² Consulter l'Exemple 17 pour un contre-exemple.


Preuve (*Point clef — Récurrence sur le nombre de polynômes*)

■ **Initialisation.** Soit (P_0) une famille d'un polynôme, avec $P_0 \neq 0$ (la famille est bien entendu échelonnée), et libre car $P_0 \neq 0$.

■ **Hérédité.** Supposons le théorème établi pour une famille de $n \in \mathbf{N}^*$ polynômes échelonnée. Et soit (P_0, \dots, P_n) une famille échelonnée de $n + 1$ polynômes non nuls. Quitte à renuméroter les polynômes, on peut supposer que

$$\deg P_0 < \dots < \deg P_n.$$



Exemple 17 — Montrer que la famille $(X(X - 1), X(X - 2), (X - 1)(X - 2))$ est une famille libre de $\mathbf{R}_2[X]$. Est-elle échelonnée? 

2.2. Familles génératrices

Définition ALG.3.6 | Famille génératrice

Soit F un sous-espace vectoriel d'un \mathbf{K} -espace vectoriel E . On appelle *famille génératrice* de F toute famille $(x_i)_{i \in I}$, où I est un ensemble, de vecteurs de F telle que

$$F = \text{Vect}(x_i)_{i \in I}.$$

Autrement dit, tout vecteur de F peut s'écrire comme combinaison linéaire **finie** d'éléments de $(x_i)_{i \in I}$. On dit aussi que F est engendré par les x_i pour $i \in I$.

Définition ALG.3.7 | Dimension finie

Soit F un sous-espace vectoriel d'un \mathbf{K} -espace vectoriel E . On dit que F est de *dimension finie* s'il existe une famille finie engendrant F .

Ainsi, pour montrer qu'un espace vectoriel est de dimension finie on exhibe une famille génératrice finie.

Attention Nous n'avons pas encore défini ce qu'est la dimension d'un espace de dimension finie, seulement la propriété de « dimension finie ».

Méthode Montrer qu'une famille est génératrice

Pour montrer qu'une famille $(x_k)_{1 \leq k \leq n}$ ¹³ de vecteurs de F est une famille génératrice de F , on écrit :



« Soit $x \in F$. Alors cherchons $\lambda_1, \dots, \lambda_n$ tels que $x = \sum_{k=1}^n \lambda_k x_k$. [...] On a donc déterminé des λ_i qui conviennent, la famille est génératrice. »

En général, l'étape [...] consiste en les arguments suivants :


- ▶ Dans \mathbf{R}^n ou \mathbf{C}^n : on résout un système linéaire.
- ▶ Dans \mathbf{R}^1 , \mathbf{R}^N on fait de l'analyse (limites, dérivation, etc.).

Exemple 18 —

1. Soit $E = \mathbf{R}^2$. On considère les vecteurs de \mathbf{R}^2 suivants $x_1 = (1, 1)$, $x_2 = (1, 2)$ et $x_3 = (1, 3)$. La famille (x_1, x_2, x_3) est-elle une famille génératrice de \mathbf{R}^2 ?

2. Soit $E = \mathcal{C}^1(\mathbf{R}, \mathbf{R})$. Donnons une famille génératrice de $F = \{y \in E, y' = 2y\}$.

¹³adapter la méthode pour une famille quelconque $\text{Vect}(x_i)_{i \in I}$

3. Soit $E = \mathbf{R}^{\mathbf{N}}$. Donnons une famille génératrice de $F = \{(u_n) \in E, \forall n \in \mathbf{N}, u_{n+1} = 5u_n\}$. 

Définition/Proposition ALG.3.2 | Base et décomposition en coordonnées. Coordonnées d'un vecteur.

- ▶ Une famille finie $(x_k)_{1 \leq k \leq n}$ de vecteurs de E est une base de E **si et seulement si** tout vecteur de E s'écrit de façon unique comme combinaison linéaire des x_k .
- ▶ Si $x \in E$, notons $(\lambda_k)_{1 \leq k \leq n}$ telle que

$$x = \sum_{k=1}^n \lambda_k x_k.$$

Alors, pour tout $k \in \llbracket 1, n \rrbracket$, la constante λ_k s'appelle **la** k -ième coordonnée de x dans la base (x_1, \dots, x_n) .

Preuve



2.3. Base

Nous arrivons à une notion qui généralise les repères que vous connaissez depuis que vous avez fait de la géométrie analytique (*i.e.* avec des coordonnées) en fin de collège.

Définition ALG.3.8 | Base

Soit E un \mathbf{K} -espace vectoriel. On appelle *base* de E toute famille de vecteurs de E qui est libre et génératrice de E .

Rappelons que :

- ▶ le caractère générateur garantit l'existence d'une combinaison linéaire,
- ▶ le caractère libre garantit l'unicité des coefficients.

Une base garantira donc l'existence et l'unicité des coefficients, c'est ce que nous précisons maintenant.

En Mathématiques, l'adjectif *canonique* après «base» signifie parfois «la plus simple».

Définition/Proposition ALG.3.3 | Bases canoniques de \mathbf{K}^n , $\mathbf{K}_n[x]$

Soit $n \geq 1$ un entier.

► **(Dans \mathbf{K}^n)** Notons


$$e_1 = (1, 0, 0, \dots, 0), \quad e_2 = (0, 1, 0, \dots, 0),$$

$$e_3 = (0, 0, 1, \dots, 0), \quad \dots, \quad e_n = (0, 0, 0, \dots, 1).$$


La famille $(e_k)_{1 \leq k \leq n}$ est une base de \mathbf{K}^n , appelée *base canonique de \mathbf{K}^n* .

► **(Dans $\mathbf{K}_n[X]$)** La famille $(X^k)_{0 \leq k \leq n} = (1, X, X^2, \dots, X^n)$ est une base de $\mathbf{K}_n[X]$, appelée *base canonique de $\mathbf{K}_n[X]$* .


Preuve Faisons la preuve pour $\mathbf{K}_n[X]$, elle est similaire pour \mathbf{K}^n . 

2. Montrer que la famille $(X^2 + X, X^2 + 1, X + 1)^{14}$ est une base de $\mathbf{R}_2[X]$. 

Exemple 19 —

1. Montrer que la famille $((1, 1), (1, -2))$ est une base de \mathbf{R}^2 . 

¹⁴Attention, elle n'est pas échelonnée

3. On note $E = \{y \in \mathcal{C}^2, y'' - 3y' + 2y = 0\}$. Déterminer une base de E . 

4. On note $E = \{(u_n) \in \mathbf{R}^{\mathbf{N}}, \forall n \in \mathbf{N}, u_{n+2} = -u_{n+1} - u_n\}$. Déterminer une base de E .



2.4. Extraction & Complétion

Il est possible de construire des bases à l'aide de familles libres (en complétant), et de familles génératrices (en extrayant). Nous allons établir deux faits principaux :

1. toute famille libre peut être complétée en une base (de-même, toute famille génératrice peut être diminuée en une base),
2. toutes les bases ont même cardinal; cet entier, nous allons l'appeler la *dimension*.

Proposition ALG.3.10 | Augmentation d'une famille libre finie

Soient $\mathcal{L} = (\ell_1, \dots, \ell_n)$ une famille libre d'un \mathbf{K} -espace vectoriel E , et $\ell_{n+1} \in E$.
Si $\ell_{n+1} \notin \text{Vect } \mathcal{L}$, alors :
 $\mathcal{L}' = (\ell_1, \dots, \ell_n, \ell_{n+1})$ est encore une famille libre.

Preuve



Exemple 20 — Dans \mathbf{R}^3 , on considère $\mathcal{L} = (\ell_1, \ell_2) = ((2, 1, 0), (1, -1, 0))$. Que dire de $\ell_3 = (2, 3, 1)$?

Proposition ALG.3.11 | Diminution d'une famille génératrice finie

Soit $\mathcal{G} = (g_1, \dots, g_{n+1})$ une famille génératrice d'un \mathbf{K} -espace vectoriel E .¹⁵
Si $g_{n+1} \in \text{Vect}(g_1, \dots, g_n)$, alors :
 $\mathcal{G}' = (g_1, \dots, g_n)$ est encore une famille génératrice de E .

Preuve Par hypothèse, il existe des constantes μ_1, \dots, μ_n telles que :

$$g_{n+1} = \sum_{i=1}^n \mu_i g_i. \quad (\star)$$

Soit donc $x \in E$. Puisque \mathcal{G} est génératrice de E , choisissons $\lambda_1, \dots, \lambda_n$ telles que :

$$x = \sum_{i=1}^n \lambda_i g_i + \lambda_{n+1} g_{n+1},$$

ainsi en remplaçant g_{n+1} par (\star) , on trouve :


$$x = \sum_{i=1}^n \lambda_i g_i + \sum_{i=1}^n \lambda_{n+1} \mu_i g_i = \sum_{i=1}^n (\lambda_i + \lambda_{n+1} \mu_i) g_i.$$

Ceci prouve que x est combinaison linéaire de g_1, \dots, g_n donc que \mathcal{G}' est génératrice.

Les propriétés précédentes nous apprennent donc que :

- ▶ l'on peut augmenter une famille libre finie (en une nouvelle famille libre) en ajoutant un vecteur qui n'est pas combinaison linéaire de ceux de la famille.
- ▶ D'autre part, on peut diminuer une famille génératrice finie (en une nouvelle famille génératrice) en retirant un vecteur qui est combinaison linéaire des autres.

¹⁵Dit autrement, l'espace vectoriel E est de dimension finie.

Exemple 21 — On considère le sous-espace de \mathbf{R}^3 suivant $E = \text{Vect}(g_1, g_2, g_3)$, où $g_1 = (2, 1, 3)$, $g_2 = (1, 0, 1)$. Que dire de $g_3 = (1, 1, 2)$? *Indication*: On pourra remarquer que $g_3 = g_1 - g_2$. 

ALGORITHME DE LA BASE INCOMPLÈTE.

Théorème ALG.3.2 | Algorithme de la base incomplète

Soient E un \mathbf{K} -espace vectoriel de dimension finie et non réduit à $\{0_E\}$, $\mathcal{L} = (\ell_1, \dots, \ell_p)$ une famille libre finie de E et $\mathcal{G} = (g_1, \dots, g_q)$ une famille génératrice finie de E . Alors il existe une base de la forme

$$\mathcal{B} = (\underbrace{\ell_1, \dots, \ell_p}_{\text{proviennent de } \mathcal{L}}, \underbrace{\ell_{p+1}, \dots, \ell_n}_{\text{proviennent de } \mathcal{G}}).$$

Le théorème nous apprend donc que l'on peut compléter une famille libre finie en une base en ajoutant des vecteurs puisés dans une famille génératrice finie.

La preuve ci-dessous fournit en outre un algorithme pour construire une base comme *supra*¹⁶. Si $E = \{0_E\}$ alors E ne possède aucune famille libre, le résultat précédent ne s'applique donc pas.

Preuve

- Construction algorithmique de la base \mathcal{B}** — ce théorème repose sur un algorithme simple et fondamental dit de la base incomplète. Nous allons complé-


¹⁶algorithme assez peu utilisé dans la pratique

ter peu à peu la famille libre $\mathcal{L} = (\ell_1, \dots, \ell_p)$ à l'aide de certains vecteurs parmi $\mathcal{G} = (g_1, \dots, g_q)$ en prenant soin de conserver la liberté à chaque ajout.

- ▶ Commençons par considérer $\mathcal{B} = \mathcal{L}$.
- ▶ Ensuite, pour $k \in \llbracket 1, q \rrbracket$, si (\mathcal{B}, g_k) est encore libre on modifie \mathcal{B} : $\mathcal{B} \leftarrow (\mathcal{B}, g_k)$ où (\mathcal{B}, g_k) désigne la famille \mathcal{B} augmentée de g_k . Sinon nous ne faisons rien.

Puisque k par k la famille est libre, la base \mathcal{B} obtenue à l'arrivée est encore libre.

- La famille \mathcal{B} ainsi construite convient**— en effet, nous savons déjà que \mathcal{B} est une famille libre, il reste donc à montrer qu'elle est génératrice.

 Or, par hypothèse, la famille \mathcal{G} génère E donc il suffit de montrer que chaque g_i pour $i \in \llbracket 1, q \rrbracket$ est combinaison linéaire des éléments de \mathcal{B} notés ℓ_i avec $i \in \llbracket 1, n \rrbracket$. Soit donc $i \in \llbracket 1, q \rrbracket$. Si $g_i \in \mathcal{B}$ alors évidemment il est combinaison linéaire de lui-même.

Si non, $g_i \notin \mathcal{B}$, cela signifie qu'à l'étape $k = i$ de l'algorithme nous n'avons pas ajouté g_i dans la famille \mathcal{B} parce qu'il était combinaison linéaire d'éléments de \mathcal{B} . C'est donc terminé.


Théorème ALG.3.3 | Base incomplète/extraite

Soit E un \mathbf{K} -espace vectoriel de dimension finie et non réduit à $\{0_E\}$.

- (Théorème de la base incomplète)** Toute famille libre finie de E peut être complétée en une base de E .
- (Théorème de la base extraite)** De toute famille génératrice finie de E on peut extraire une base de E . En particulier, E admet une base.

Preuve

- Si E est de dimension finie, alors E possède une famille génératrice finie \mathcal{G} . Ainsi, si \mathcal{L} est une famille libre finie, nous pouvons lui appliquer l'algorithme de la base incomplète donné par le **Théorème ALG.3.2** précédent.
- Soit \mathcal{G} une famille génératrice finie de E . Puisque $E \neq \{0_E\}$, nous pouvons choisir un élément $x \neq 0_E$ dans E . Il suffit alors d'appliquer le **Théorème ALG.3.2** précédent à la famille $(x) \cup \mathcal{G}$, qui donne le résultat.

Exemple 22 — La famille $((1, -5, 7), (2, 6, 8))$ est une base de $F = \text{Vect}((1, -5, 7), (2, 6, 8), (3, 1, 15), (1, 11, 1))$. 

Exemple 23 — On considère la famille de \mathbf{R}^4 suivante :

$$\mathcal{G} = ((1, 0, 2, 1), (2, 1, 3, 2), (1, 1, 1, 1), (0, 1, -1, -1), (2, 0, 4, 3)).$$

Déterminons une base de $\text{Vect } \mathcal{G}$, extraite de \mathcal{G} . 

3. DIMENSION

31. Généralités

Nous admettons dans cette dernière section ce résultat, appelé aussi *lemme de STEINITZ* dont la preuve est technique, et qui permet de comparer le nombre d'éléments d'une famille libre par rapport au nombre d'éléments d'une famille génératrice.

Lemme ALG.3.1 | de STEINITZ

Soient E un \mathbf{K} -espace vectoriel de dimension finie non réduit à $\{0_E\}$, \mathcal{L} une famille libre et \mathcal{G} une famille génératrice finie de E . Alors :

$$\# \mathcal{L} \leq \# \mathcal{G}.$$

Preuve Admis.

Définition/Proposition ALG.3.4 | Dimension

Soit E un espace vectoriel de dimension finie. Alors :

► si $E \neq \{0_E\}$: toutes les bases de E ont le même nombre d'éléments, on l'ap-

pelle la *dimension* de E , et est notée $\dim E$,

- ▶ si $E = \{0_E\}$, on pose comme convention : $\dim \{0_E\} = \dim E = 0$.

Si $\mathbf{K} = \mathbf{C}$, on notera $\dim_{\mathbf{R}} E$ la dimension de E vu comme un \mathbf{R} -espace vectoriel, et $\dim_{\mathbf{C}} E = \dim E$ vu comme un \mathbf{C} -espace vectoriel.

Preuve Montrons que dans le cas $E \neq \{0_E\}$, toutes les bases ont même nombre d'éléments. Soient \mathcal{B} et \mathcal{B}' deux bases de E ayant respectivement n et n' éléments.



2. $\dim_{\mathbf{C}} \mathbf{C}^2 = 2$, mais $\dim_{\mathbf{R}} \mathbf{C}^2 = 4$. Pourquoi?

3. $\dim_{\mathbf{K}} \mathbf{K}_n[X] = n + 1$.

4. Dans $E = \mathbf{K}^n$, tout sous-espace défini par une équation linéaire *non triviale*.

$$F = \left\{ (x_k)_{1 \leq k \leq n} \in E, a_1 x_1 + \dots + a_n x_n = 0 \right\} \quad (\text{où les } a_k \in \mathbf{K} \text{ ne sont pas tous nuls}),$$

est un hyperplan. Pour le prouver, il suffit d'en obtenir une famille génératrice.

Définition ALG.3.9 | Droite, Plan & Hyperplan

Soit E un espace vectoriel de dimension finie. Un sous-espace vectoriel F de E est appelé :

- ▶ *droite vectorielle* si $\dim F = 1$,
- ▶ *plan vectoriel* si $\dim F = 2$,
- ▶ *hyperplan* si $\dim E \geq 1$ et $\dim F = \dim E - 1$.



Attention Confusion cardinal/dimension


Ne pas confondre les notions de cardinal et de dimension ! Une famille a un certain cardinal mais pas de dimension, et par contre un espace vectoriel (de dimension finie) a une dimension mais est toujours de cardinal infini¹⁷ lorsque $\mathbf{K} = \mathbf{R}$ ou \mathbf{C} (sauf $\{0\}$). En revanche, on a toujours que si \mathcal{B} est une base de E


$$\dim E = \# \mathcal{B}.$$

Exemple 24 —

1. $\dim_{\mathbf{K}} \mathbf{K}^n = n$ (notamment $\dim_{\mathbf{K}} \mathbf{K} = 1$).

¹⁷En effet, si $x \in E$ est non nul, alors $\text{Vect}(x)$ est un ensemble de cardinal infini. Car si ce n'est pas le cas il existerait $\lambda \neq \lambda' \in \mathbf{K}$ tels que $\lambda x = \lambda' x \implies \lambda = \lambda'$ puisque $x \neq 0_E$ — contradiction.

5. L'ensemble des solutions réelles d'une équation différentielle linéaire homogène normalisée d'ordre un est un \mathbf{R} -espace vectoriel de dimension un. Nous le reverrons dans le [Chapter ANA.9](#). Par exemple, déterminer la dimension de $\{y \in \mathcal{D}^1(\mathbf{R}, \mathbf{R}), y' = 2y\}$. 

6. L'ensemble des solutions réelles d'une équation différentielle linéaire homogène d'ordre deux à coefficients constants est un \mathbf{R} -espace vectoriel de dimension deux. Nous le reverrons dans le [Chapter ANA.10](#). Par exemple, déterminer la dimension de $\{(u_n) \in \mathbf{R}^{\mathbf{N}}, \forall n \in \mathbf{N}, u_{n+2} = 3u_n\}$. 

Enfin, un résultat fondamental qui nous permet de gagner du temps en pratique : lorsque le nombre d'éléments d'une famille est égal à la dimension de l'espace en question (encore faut-il la connaître, on ne l'utilisera donc que dans ce cas), il suffit de prouver le caractère générateur **OU** libre.

Théorème ALG.3.4

Soient E un \mathbf{K} -espace vectoriel de **dimension finie** $n \neq 0$ et \mathcal{F} une famille finie telle que $\#\mathcal{F} = \dim E$. Alors :


$$\begin{aligned} \mathcal{F} \text{ est une base de } E &\iff \mathcal{F} \text{ est une famille libre de } E \\ &\iff \mathcal{F} \text{ est une famille génératrice de } E. \end{aligned}$$


Remarque 3.1 — Et en général, on préfère souvent montrer qu'une famille est libre... c'est souvent plus rapide.

Preuve Dans un \mathbf{K} -espace vectoriel de dimension finie $n \neq 0$, si une famille \mathcal{F} de n vecteurs est libre (*resp.* génératrice), on peut la compléter en une base d'après le théorème de la base incomplète (*resp.* en extraire une base d'après le théorème de la base extraite).

Le résultat est une famille de n vecteurs par définition de la dimension, ce qui veut dire qu'on a en fait ajouté (*resp.* ôté) aucun vecteur à \mathcal{B} . Conclusion : \mathcal{B} était une base dès le départ.

Exemple 25 —

1. La famille $((0, 1, 2), (1, 2, 0), (2, 0, 1))$ est une base de \mathbf{R}^3 . 

2. Pour $a \in \mathbf{K}$. La famille de polynômes $\mathcal{F} = (1, X - a, (X - a)^2, (X - a)^3)$ est une base de $\mathbf{K}_3[X]$. Déterminer les coordonnées d'un polynôme dans cette base. 

Notation

On notera plus simplement $\text{Rg}(x_1, \dots, x_n) = \dim \text{Vect}(x_1, \dots, x_n)$ lorsque le contexte est clair.

Exemple 26 –

1. Dans $\mathbf{K}[X]$, on a : $\text{Rg}(1, X, X^2, X^3) = 4$.
2. Dans $\mathbf{K}[X]$, on a : $\text{Rg}(0_{\mathbf{K}[X]}, X, 2X, 3X) = 1$.
3. Dans \mathbf{R}^3 , on a : $\text{Rg}((1, 1, 0), (0, 0, 1)) = 2$.
4. Dans \mathbf{R}^3 , on a : $\text{Rg}((1, 1, 0), (0, 0, 1), (1, 1, 1)) = 2$.
5. Dans \mathbf{R}^N , calculer : $\text{Rg}((2^n), (n2^n))$.



3.2. Sous-espaces et dimension

Proposition ALG.3.12

Soient E un \mathbf{K} -espace vectoriel de **dimension finie** et F un sous-espace vectoriel de E . Alors :

- ▶ F est de dimension finie, et : $\dim F \leq \dim E$.
- ▶ $\dim F = \dim E \iff F = E$.

Preuve Admise.

3.3. Rang

Définition ALG.3.10 | Rang d'une famille de vecteurs

Soit (x_1, \dots, x_n) une famille finie d'un \mathbf{K} -espace vectoriel E . On appelle *rang sur \mathbf{K} de la famille (x_1, \dots, x_n)* , et on note $\text{Rg}_{\mathbf{K}}(x_1, \dots, x_n)$, la dimension sur \mathbf{K} de $\text{Vect}_{\mathbf{K}}(x_1, \dots, x_n)$:

$$\text{Rg}_{\mathbf{K}}(x_1, \dots, x_n) = \dim_{\mathbf{K}} \text{Vect}_{\mathbf{K}}(x_1, \dots, x_n).$$

4. APPLICATIONS LINÉAIRES

Les espaces vectoriels sont en particulier des ensembles, on peut donc tout à fait considérer des applications entre eux. Mais pour pouvoir faire des calculs, il faut que notre définition tienne compte des opérations +, .. On va donc supposer qu'une *application linéaire* transforme les sommes en sommes et les produits avec un scalaire en produits.

Avant de passer à la définition, souvenez-vous que vous connaissez déjà des exemples d'applications linéaires. En l'occurrence celles vues en troisième : en effet, si $E = F = \mathbf{R}$ dans les définitions *infra*, alors $x \in E \rightarrow Kx \in F$ sera une application linéaire au sens de la définition qui suit, pour tout $K \in \mathbf{R}$.

4.1. Généralités

Définition ALG.3.11 | Application linéaire

Soient $(E, +_E, \cdot_E)$ et $(F, +_F, \cdot_F)$ deux \mathbf{K} -espaces vectoriels. On appelle *application \mathbf{K} -linéaire* (parfois aussi *morphisme linéaire*) de E dans F toute application $u : E \rightarrow F$ telle que :

$$\forall (x, y) \in E^2, \quad \forall (\lambda, \mu) \in \mathbf{K}^2, \quad u(\lambda \cdot_E x +_E \mu \cdot_E y) = \lambda \cdot_F u(x) +_F \mu \cdot_F u(y).$$

- ▶ Lorsque $u(E) \subset E$, *i.e.* pour tout $x \in E$, $u(x) \in E$, on dit que u est un *endomorphisme*.¹⁸
- ▶ Si $F = \mathbf{K}$ alors on dit que u est une *forme linéaire* sur E .

¹⁸Provient du grec *endon* qui signifie «à l'intérieur», que vous connaissez déjà des S.V.T. *via* le qualificatif «endogène».

Notation

On note :

- ▶ $\mathcal{L}_{\mathbf{K}}(E, F)$ l'ensemble des applications \mathbf{K} -linéaires de E dans F , ou plus simplement $\mathcal{L}(E, F)$ s'il n'y a pas d'ambiguïté.
- ▶ On note parfois $\mathcal{L}_{\mathbf{K}}(E, \mathbf{K}) = E^*$ l'ensemble des formes linéaires, cet ensemble est appelé *espace dual* de E .

Remarque 4.1 — De façon équivalente,¹⁹ l'application $u : E \rightarrow F$ est linéaire si et seulement si elle est compatible avec la somme et la multiplication par un scalaire :

$$\begin{aligned} \forall (x, y) \in E^2, \quad u(x +_E y) &= u(x) +_F u(y), \\ \forall \lambda \in \mathbf{K}, \quad \forall x \in E, \quad u(\lambda \cdot_E x) &= \lambda \cdot_F u(x). \end{aligned}$$

Remarque 4.2 — Abus de notation Puisqu'il s'agit de notre définition initiale, nous notons très correctement en indices les espaces vectoriels auxquels sont attachées les opérations $+_E, \cdot_E, +_F, \cdot_F$... À partir de maintenant, nous noterons simplement $+, \cdot$. Ainsi, avec cet abus, la définition *supra* se réécrit plus simplement :

$$\forall (x, y) \in E^2, \quad \forall (\lambda, \mu) \in \mathbf{K}^2, \quad u(\lambda \cdot x + \mu y) = \lambda \cdot u(x) + \mu \cdot u(y).$$

Proposition ALG.3.13 | Le neutre est envoyé sur le neutre

Soient $(E, +_E, \cdot_E)$ et $(F, +_F, \cdot_F)$ deux \mathbf{K} -espaces vectoriels et $u : E \rightarrow F$ une application *linéaire*. Alors :

$$u(0_E) = 0_F.$$

Preuve Faire simplement $\lambda = \mu = 0$ et $x = y = 0_E$ dans la définition d'une application linéaire.

¹⁸catif «endogène».

¹⁹Mais il est plus rapide de vérifier la définition

**Méthode** Montrer qu'une application n'est pas linéaire

Pour montrer qu'une application n'est pas linéaire, on peut :

- ▶ vérifier si l'égalité $u(0_E) = 0_F$ est vérifiée.
- ▶ Si c'est le cas, on cherche s'il existe $\lambda, \mu \in \mathbf{K}$ et $x, y \in E$ tels que : $u(\lambda.x + \mu.y) \neq \lambda.u(x) + \mu.u(y)$.

Exemple 27 – Contre... Les applications ci-après ne sont pas linéaires.

$$1. f : \begin{cases} \mathbf{R}^3 & \longrightarrow \mathbf{R}^2, \\ (x, \dots, z) & \longmapsto (x + y + z + 1, z), \end{cases} \quad \boxed{\text{✎}}$$

$$2. g : \begin{cases} \mathbf{R}^3 & \longrightarrow \mathbf{R}, \\ (x, y, z) & \longmapsto x^2 + y + z. \end{cases} \quad \boxed{\text{✎}}$$

Commençons par quelques premiers exemples.

Exemple 28 – avec des uplets Les applications suivantes sont linéaires.

$$1. f : \begin{cases} \mathbf{R}^3 & \longrightarrow \mathbf{R}, \\ (x, y, z) & \longmapsto (x - 2y + z, 2x + 3y - 5z, x + y + z) \end{cases} \cdot \boxed{\text{✎}}$$

$$2. \text{ Soit } n \geq 1 \text{ un entier. } \pi_i : \begin{cases} \mathbf{K}^n & \longrightarrow \mathbf{K} \\ (x_1, \dots, x_n) & \longmapsto x_i \end{cases} \text{ est, pour tout } i \in \llbracket 1, n \rrbracket, \text{ une forme} \\ \text{linéaire sur } \mathbf{K}^n \text{ appelée } i\text{-ème projection canonique.} \quad \boxed{\text{✎}}$$

$$1. D : \begin{cases} \mathbf{K}[X] & \longrightarrow & \mathbf{K}[X], \\ P & \longrightarrow & P(X^2 + 1), \end{cases} \quad \boxed{\text{✍}}$$

Plus généralement, l'application qui à un vecteur associe sa i -ème coordonnée dans une base fixée est toujours linéaire, nous le montrerons dans la proposition qui suit.

$$3. \text{ Soit } n \geq 1 \text{ un entier. } \mu : \begin{cases} \mathbf{K}^n & \longrightarrow & \mathbf{K} \\ (x_1, \dots, x_n) & \longrightarrow & \frac{1}{n} \sum_{i=1}^n x_i \end{cases} \text{ est une forme linéaire sur } \mathbf{K}^n \text{ appelée } \textit{moyenne} \text{ en statistiques. } \boxed{\text{✍}}$$

$$2. \varphi : \begin{cases} \mathcal{C}^0([0, 1], \mathbf{R}) & \longrightarrow & \mathbf{R} \\ f & \longrightarrow & \int_0^1 f(t) dt, \end{cases} \quad \boxed{\text{✍}}$$

Exemple 29 — *Avec des polynômes, suites, ...* Les applications suivantes sont linéaires :

$$3. \psi : \begin{cases} \mathcal{F}(\mathbf{R}, \mathbf{R}) & \longrightarrow & \mathbf{R}^3, \\ f & \longmapsto & (f(0), f(1), f(2)). \end{cases} \quad \text{✎}$$

$$4. \delta : \begin{cases} \mathbf{R}^{\mathbb{N}} & \longrightarrow & \mathbf{R}^{\mathbb{N}}, \\ (u_n)_{n \in \mathbb{N}} & \longmapsto & (u_{n+1} - u_n)_{n \in \mathbb{N}}. \end{cases} \quad \text{✎}$$

Définition/Proposition ALG.3.5 | Identité et homothétie

- ▶ $\text{Id}_E : \begin{cases} E & \longrightarrow & E \\ x & \longmapsto & x \end{cases}$ est un endomorphisme de E , appelé *endomorphisme identité* (ou *endomorphisme identité*).
- ▶ Pour tout $\lambda \in \mathbf{K}$, l'application λId_E est un endomorphisme de E appelé *homothétie de rapport* λ .

Preuve



Proposition ALG.3.14 | Dépendance linéaire des coordonnées d'un vecteur

Soit E un \mathbf{K} -espace vectoriel de dimension finie non nulle et $\mathcal{B} = (e_1, \dots, e_n)$ une base de E avec $n \geq 1$. Si $x \in E$, notons $\lambda_i(x)$ la i -ème coordonnée de x dans la base \mathcal{B} avec $i \in \llbracket 1, n \rrbracket$. Alors l'application

$$\pi_i^{\mathcal{B}} : \begin{cases} E & \longrightarrow & \mathbf{K} \\ x & \longmapsto & \lambda_i(x) \end{cases} \quad \text{est une forme linéaire.}$$

Preuve Supposons pour simplifier que $n = 2$. La preuve pour n quelconque est identique.

Proposition ALG.3.15 | Opérations sur les applications linéaires

Soient E, F et G des \mathbf{K} -espaces vectoriels.

- (Combinaison linéaire)** Soient $(\lambda, \mu) \in \mathbf{K}^2$ et $(u, v) \in \mathcal{L}(E, F)^2$. Alors $\lambda u + \mu v \in \mathcal{L}(E, F)$. Ces opérations définissent une structure de \mathbf{K} -espace vectoriel $(\mathcal{L}_{\mathbf{K}}(E, F), +, \cdot)$.
- (Composition)** Soient $u \in \mathcal{L}(E, F)$ et $v \in \mathcal{L}(F, G)$. Alors :

$$v \circ u \in \mathcal{L}(E, G).$$

- (Restriction)** Soient $u \in \mathcal{L}(E, F)$ et V un sous-espace vectoriel de E . Alors l'application

$$u|_V \begin{cases} V & \longrightarrow & F \\ x & \longmapsto & u(x) \end{cases}$$

est appelée *application restreinte*²⁰ de u à V . C'est une application linéaire de V dans F .

Preuve

1. 

4.2. Structure de \mathbf{K} -espace vectoriel sur $\mathcal{L}_{\mathbf{K}}(E, F)$ et puissances

2. 

²⁰Puisqu'une application est déterminée par : un ensemble de départ et d'arrivée, et une opération d'association, il est important d'introduire une autre notation pour la restriction. Nous avons changé l'ensemble de départ ici.


3. Évident.

Attention

Même si $u|_V$ et u coïncident sur V , en tant qu'application elles sont différentes (puisque leur espace de départ n'est pas le même).

Exemple 30 — On considère les applications :

$$u : \begin{cases} \mathbf{R}^2 & \longrightarrow & \mathbf{R}^2 \\ (x, y) & \longmapsto & (0, x) \end{cases} \quad \text{et} \quad v : \begin{cases} \mathbf{R}^2 & \longrightarrow & \mathbf{R}^2 \\ (x, y) & \longmapsto & (y, 0) \end{cases}$$

Calculons $u \circ v$ et $v \circ u$. Que constate-t-on? 

Attention

En général, on note plutôt uv pour des applications linéaires au lieu de $u \circ v$. Mais soyez méfiant vis-à-vis de cette notation, cela ne désigne en aucun cas l'application produit $x \mapsto u(x)v(x)$ qui n'a même aucun sens ici.²¹

PUISSANCES D'UN ENDOMORPHISME On se place désormais dans le cas d'endomorphismes, *i.e.* lorsque $E = F$.

Définition ALG.3.12 | Puissances

Soit E un \mathbf{K} -espace vectoriel, et $f \in \mathcal{L}(E)$. Alors on définit par récurrence l'endomorphisme f^k pour tout $k \in \mathbf{N}$:


$$\begin{cases} f^0 = \text{Id}_E, \\ \forall k \in \mathbf{N}, \quad f^{k+1} = f \circ f^k = f^k \circ f. \end{cases}$$

Remarque 4.3 — On suppose que f est un **endomorphisme**, tout simplement pour que les composées aient un sens. On pourrait faire sans cette hypothèse et supposer plus généralement que $f(E) \subset F$ si $f : E \rightarrow F$.

Exemple 31 —

1. On considère à nouveau


$$u : \begin{cases} \mathbf{R}^2 & \longrightarrow & \mathbf{R}^2 \\ (x, y) & \longmapsto & (0, x). \end{cases}$$

Calculer les puissances de u . 

²¹Pourquoi? Nous l'avons déjà vu : nous ne pouvons pas multiplier deux vecteurs entre eux.


2. On considère

$$T : \begin{cases} \mathbf{R}[X] & \longrightarrow & \mathbf{R}[X] \\ \mathbf{P} & \longrightarrow & \mathbf{P}(X+1). \end{cases}$$

Calculer les puissances de T . 

d'un vecteur de E par u . En d'autres termes, pour $y \in F$:

$$y \in \text{Im } u \iff \exists x \in E, \quad y = u(x).$$

Preuve Montrons que $u(V)$ est un sous-espace vectoriel de F . 

4.3. Image & Noyau

La notion d'image d'application a déjà été rencontrée en première année, elle est donc toujours en vigueur pour des applications en particulier linéaires, et fortement liée à la surjectivité. Nous allons voir que l'injectivité peut dans le cas d'applications linéaires être reformulée à l'aide de ce que nous appellerons le *noyau*; mais bien entendu il est toujours possible de recourir à la définition basique vue l'année dernière. C'est-à-dire *une application $u : E \rightarrow F$ entre deux ensembles E et F est injective si*

$$\forall (x, x') \in E^2, \quad u(x) = u(x') \implies x = x'.$$

Définition/Proposition ALG.3.6 | Image directe d'un espace vectoriel

Soient E et F des \mathbf{K} -espaces vectoriels et $u \in \mathcal{L}(E, F)$ et V un sous-espace vectoriel de E . Alors :

- ▶ l'ensemble $u(V) = \{u(x), x \in V\}$ est un sous-espace vectoriel de F , appelé *image directe de V par u* .
- ▶ On appelle *image de u* , et on note $\text{Im } u$, l'image directe de E — l'ensemble de départ — par l'application u :

$$\text{Im } u = u(E).$$

C'est le sous-espace vectoriel **de F** constitué des vecteurs qui sont l'image

Attention

Il est complètement faux de se passer de la linéarité dans l'énoncé précédent. Par exemple, $u : x \in \mathbf{R} \mapsto 1 + x$ n'est pas une application linéaire (zéro n'est pas envoyé sur zéro), et $u(\{0\}) = \{1\}$ n'est pas un sous-espace vectoriel de \mathbf{R} puisqu'il ne contient pas zéro.

Passons maintenant à la définition du noyau. Reprenons la définition de l'injectivité rappelée plus haut :

$$\forall (x, x') \in E^2, \quad u(x) = u(x') \implies x = x'.$$

Puisque u est linéaire, elle est équivalente à :

$$\forall (x, x') \in E^2, \quad u(x - x') = 0_F \implies x - x' = 0_E.$$

Par conséquent, une application injective linéaire vérifie (en renommant $y = x - x'$):

$$\forall y \in E, \quad u(y) = 0_F \implies y = 0_E.$$

Autrement dit, le seul vecteur y vérifiant $u(y) = 0_E$ est le vecteur nul. Cela nous incite alors à considérer l'ensemble des vecteurs qui s'envoient sur 0_F : nous appellerons cet ensemble le *noyau* de u .

Définition/Proposition ALG.3.7 | Noyau

On appelle *noyau* de u , et on le note $\ker u$, l'ensemble :

$$\ker u = \{x \in E, u(x) = 0_F\}.$$

C'est le sous-espace vectoriel de E constitué des vecteurs ayant 0_F pour image par u . Donc, pour $x \in E$:

$$x \in \ker u \iff u(x) = 0_F.$$

Preuve Montrons que $\ker(u)$ est un sous-espace vectoriel de E . 

Résumé Qui est dans quoi?

On retient donc :

- ▶ le noyau est un sous-espace vectoriel de l'espace de départ,
- ▶ l'image est un sous-espace vectoriel de l'espace d'arrivée.

Exemple 32 — Soient E et F des \mathbf{K} -espaces vectoriels. Alors, on a :

- ▶ $\text{Im}(0_{\mathcal{L}(E,F)}) = \{0_F\}, \ker(0_{\mathcal{L}(E,F)}) = E,$
- ▶ $\text{Im}(\text{Id}_E) = E, \ker(\text{Id}_E) = \{0_E\}.$

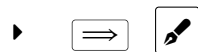


Proposition ALG.3.16 | Caractérisation de la surjectivité/injectivité pour les applications linéaires

Soient E et F des \mathbf{K} -espaces vectoriels et $u \in \mathcal{L}(E, F)$. Alors :

- ▶ u est injective $\iff \ker u = \{0_E\}$,
- ▶ u est surjective $\iff \text{Im } u = F$.

Preuve



- ▶ Pour la surjectivité, il n'y a rien à faire, c'est une conséquence du cours de 1ère année.

Attention

Notez bien que dans la preuve précédente, la linéarité fut **fondamentale** pour écrire pour tous $x, x' \in E$:

$$u(x) = u(x') \iff u(x - x') = 0_F \iff x - x' \in \ker u.$$

Pas question donc de calculer des noyaux pour montrer qu'une application **NON** linéaire est injective. Par exemple, il est clair que $x \in \mathbf{R} \mapsto x^2 \in \mathbf{R}$ n'est pas injective et pourtant son « noyau » est réduit à zéro.

Une dernière proposition avant de regarder des exemples.

Proposition ALG.3.17 | Image directe d'un «Vect»

Soient E et F des \mathbf{K} -espaces vectoriels, $u \in \mathcal{L}(E, F)$ et $(x_k)_{1 \leq k \leq n}$ une famille finie de vecteurs de E. Alors :

$$u(\text{Vect}(x_k)_{1 \leq k \leq n}) = \text{Vect}(u(x_k))_{1 \leq k \leq n}.$$

En particulier, si (x_1, \dots, x_n) est une famille génératrice finie de E, alors :

$(u(x_1), \dots, u(x_n))$ est génératrice de $\text{Im } u$, et est donc de dimension finie.

Méthode Image d'une application si une famille génératrice de l'espace de départ est connue



1. On commence par chercher une famille génératrice \mathcal{G} de l'ensemble de départ E.
2. On calcule les images de chacun des vecteurs de \mathcal{G} .
3. Si l'on souhaite une base, on cherche à extraire une sous-famille libre.

Preuve (Point clef — Exploiter la linéarité de u)

On a :

$$\begin{aligned}
 y \in u(\text{Vect}(x_k)_{1 \leq k \leq n}) &\iff \text{il existe } \lambda_1, \dots, \lambda_n, y = u\left(\sum_{i=1}^n \lambda_i x_i\right) \\
 &\stackrel{\text{linéarité}}{\iff} \text{il existe } \lambda_1, \dots, \lambda_n, y = \sum_{i=1}^n \lambda_i u(x_i) \\
 &\iff y \in \text{Vect}(u(x_k))_{1 \leq k \leq n}.
 \end{aligned}$$

Déterminons l'image de

Exemple 33 — Déterminer une base de l'image de

$$u : \begin{array}{l} \mathbf{R}^3 \longrightarrow \mathbf{R}^3 \\ (x, y, z) \longmapsto (x + 2y + z, 2x + y - z, x + 2y + z). \end{array}$$



Passons à présent à des exemples de calculs de noyaux et d'images.

Exemple 34 — Pour les applications ci-dessous, déterminer $\ker u$ et $\text{Im } u$ en exhibant une famille génératrice. *On admettra les linéarités dans cet exemple.*

$$1. \quad u : \begin{array}{l} \mathbf{R}^3 \longrightarrow \mathbf{R}^2 \\ (x, y, z) \longmapsto (x + y - z, x - y + 2z). \end{array} \quad \text{✎}$$

$$2. \quad v : \begin{array}{l} \mathbf{R}^2 \longrightarrow \mathbf{R}^3 \\ (x, y) \longmapsto (2x - y, y, -x + y). \end{array} \quad \text{✎}$$

$$3. D : \begin{cases} \mathcal{C}^1(I, \mathbf{R}) & \longrightarrow & \mathcal{C}^0(I, \mathbf{R}) \\ f & \longmapsto & f' \end{cases} \text{ où } I \text{ est un intervalle de } \mathbf{R}. \quad \boxed{\text{✎}}$$

$$5. T : \begin{cases} \mathbf{R}_2[X] & \longrightarrow & \mathbf{R}[X] \\ P & \longmapsto & P(X+1). \end{cases} \quad \boxed{\text{✎}}$$

$$4. \Psi : \begin{cases} \mathbf{R}^N & \longrightarrow & \mathbf{R}^N \\ (u_n) & \longmapsto & (au_n) \end{cases} \text{ où } a \in \mathbf{R}. \quad \boxed{\text{✎}}$$

4.4. Isomorphismes

Définition ALG.3.13 | Isomorphisme, Automorphisme

- ▶ On appelle *isomorphisme* entre deux espaces vectoriels E et F toute application linéaire bijective de E dans F, dans ce cas on dit que E et F sont isomorphes.
- ▶ On appelle *automorphisme* de E tout endomorphisme bijectif de E, c'est donc un isomorphisme de E dans E.

Définition ALG.3.14 | Groupe linéaire GL(E)

Soit E un **K**-espace vectoriel, l'ensemble des automorphismes linéaires de E est noté GL(E), et appelé *groupe linéaire sur E*.



Résumé

- ▶ **morphisme** = application linéaire,
- ▶ **endomorphisme** = application linéaire + entre mêmes espaces,
- ▶ **isomorphisme** = application linéaire + bijective,
- ▶ **automorphisme** = application linéaire + bijective + entre mêmes espaces.

Une application bijective $u \in \mathcal{F}(E, F)$ (*i.e.* injective et surjective de E dans F) possède (voir cours de 1^{ère} année) une application inverse $u^{-1} \in \mathcal{F}(F, E)$, *i.e.* satisfaisant :

$$u \circ u^{-1} = \text{Id}_F, \quad u^{-1} \circ u = \text{Id}_E.$$

Si celle de départ est linéaire, on montre que l'inverse l'est aussi comme le précise la **Proposition ALG.3.18**.

Proposition ALG.3.18 | Linéarité de l'inverse

Soient E et F des **K**-espaces vectoriels et $u \in \mathcal{L}(E, F)$. Si u est un isomorphisme, alors $u^{-1} \in \mathcal{L}(F, E)$ est un isomorphisme.

Preuve Il suffit de montrer que u^{-1} est linéaire elle aussi. En effet, soient $y, y' \in F$

et $\lambda, \mu \in \mathbf{K}$, alors :

$$\begin{aligned} & u^{-1}(\lambda y + \mu y') \\ &= u^{-1}(\lambda u(u^{-1}(y)) + \mu u(u^{-1}(y'))) \quad \left. \begin{array}{l} \downarrow u^{-1} \circ u = \text{Id} \\ \downarrow u \text{ linéaire} \end{array} \right\} \\ &= u^{-1} \circ u(\lambda u^{-1}(y) + \mu u^{-1}(y')) \\ &= \lambda u^{-1}(y) + \mu u^{-1}(y'). \end{aligned}$$

Cette dernière égalité prouve bien la linéarité de u^{-1} .

Proposition ALG.3.19 | Composée d'isomorphisme = isomorphisme

Soient $u : E \rightarrow F$ et $v : F \rightarrow G$ deux isomorphismes de **K**-espaces vectoriels. Alors :

- ▶ $v \circ u$ est un isomorphisme de E dans G,
- ▶ et : $(v \circ u)^{-1} = u^{-1} \circ v^{-1}$.²²

Preuve



²²Attention, l'ordre est inversé.

4.5. Cas particulier de la dimension finie

4.5.1. Rang d'une application linéaire

Définition ALG.3.15 | Rang d'une application linéaire

Soient E et F des \mathbf{K} -espaces vectoriels et $u \in \mathcal{L}(E, F)$. On dit que u est de rang fini si $\text{Im } u$ est de dimension finie. On appelle alors *rang de u sur \mathbf{K}* , et on note $\text{Rg}_{\mathbf{K}}(u)$, la dimension de $\text{Im } u$ sur \mathbf{K} :

$$\text{Rg}_{\mathbf{K}}(u) = \dim_{\mathbf{K}}(\text{Im } u)$$

Notation

On notera plus simplement $\text{Rg}(u) = \dim \text{Im } u$ lorsque le contexte est clair.

La notion de rang décrite ici est fortement reliée à la notion de rang d'une famille de vecteurs définie précédemment. Le lien entre les deux apparaît avec le théorème suivant, qui pourra aussi s'écrire ultérieurement avec des matrices.

Théorème ALG.3.5 | Lien entre les notions de rang

Soient E un \mathbf{K} -espace vectoriel de dimension finie $n \in \mathbf{N}^*$, F un \mathbf{K} -espace vectoriel de dimension quelconque, $\mathcal{B} = (x_1, \dots, x_n)$ une famille génératrice de E et $u \in \mathcal{L}(E, F)$. Alors, u est de rang fini et :

$$\text{Rg}(u) = \text{Rg}(u(\mathcal{B})) \quad (= \dim \text{Vect}(u(x_1), \dots, u(x_n))).$$

Notez bien que le membre de gauche du théorème précédent est la notion de rang d'application que l'on vient de définir, les deux membres de droite correspondent à la notion de rang d'une famille de vecteurs définie début de ce chapitre, comme dimension « du vect ».

Preuve Notons (x_1, \dots, x_n) une famille génératrice finie de E . Alors, nous avons déjà établi que :

$$u(E) = \text{Vect}(u(x_1), \dots, u(x_n)).$$

Il suffit alors de passer à la dimension :

$$(\text{Rg}(u) =) \dim(u(E)) = \dim \text{Vect}(u(x_1), \dots, u(x_n)).$$

Proposition ALG.3.20 | Invariance du rang par composition avec un isomorphisme

Soient E, E', F et F' des \mathbf{K} -espaces vectoriels, $v \in \mathcal{L}(E, F)$ une application linéaire de rang fini, $u \in \mathcal{L}(E', E)$ et $w \in \mathcal{L}(F, F')$ deux **isomorphismes**. Alors $v \circ u$ et $w \circ v$ sont de rangs finis et :

$$\text{Rg}(v) = \text{Rg}(v \circ u) = \text{Rg}(w \circ v).$$

En d'autres termes, le rang est invariant par composition avec un isomorphisme.

Preuve Admis.

Théorème ALG.3.6 | Théorème du rang

Soient E un \mathbf{K} -espace vectoriel de dimension finie, F un \mathbf{K} -espace vectoriel de dimension quelconque, et $u \in \mathcal{L}(E, F)$. Alors u est de rang fini et

$$\dim E = \text{Rg}(u) + \dim \ker u$$

Attention

Bien mettre la dimension de l'espace de **départ** dans le membre de droite.

Preuve Notons $n = \dim E$.

- (Complétion d'une base de $\ker u$)** D'après le théorème de la base incomplète, notant e_1, \dots, e_p une base de $\ker u$, il existe e_{p+1}, \dots, e_n des éléments de E tels que $(e_1, \dots, e_p, e_{p+1}, \dots, e_n)$ soit une base de E . Notons par ailleurs $S = \text{Vect}(e_{p+1}, \dots, e_n)$ l'espace vectoriel engendré par les vecteurs qui complètent.
- (Il suffit de montrer que $v = u|_S \in \mathcal{L}(S, \text{Im } u)$ est un isomorphisme.)** En effet, si c'est le cas, on a alors $\dim S = \dim F$, mais comme $\dim S + \dim \ker u = \dim E$, le résultat s'en suivra.

Il est immédiat que $v = u|_S \in \mathcal{L}(S, \text{Im } u)$, montrons donc le caractère bijectif. Soit $x \in \ker v$, i.e. $x \in S$ et $x \in \ker u$, or puisque $(e_1, \dots, e_p, e_{p+1}, \dots, e_n)$ est une base de E , il existe un unique couple $(x_K, x_S) \in \ker u \times S$ tel que $x = x_K + x_S$. Mais $x = x + 0$ et $x = 0 + x$ sont deux autres décompositions (puisque $x \in S$ et $x \in \ker u$) donc

par unicité : $x_K = 0 = x_S$ donc $x = 0_E$.

Reste à montrer la surjectivité. Soit $y \in \text{Im } u$, il existe $x \in E$ tel que $u(x) = y$, mais comme mentionné plus tôt il existe un unique couple $(x_K, x_S) \in \ker u \times S$ tel que $x = x_K + x_S$, donc $u(x) = 0 + u(x_S)$, et finalement $u(x_S) = y$ justifie l'existence d'un antécédent dans S pour y .

Remarque 4.4 – Majoration sur le rang On en déduit notamment que si $u \in \mathcal{L}(E, F)$, $\dim E = n \in \mathbf{N}$ et $\dim F = p \in \mathbf{N}$, alors

$$\text{Rg } u \leq \min(n, p).$$

En effet, 

Ce résultat paraît anecdotique, mais il est d'importance capitale dans la pratique : il permet de réduire bon nombre de problèmes à une preuve d'injectivité d'application linéaire (souvent relativement aisée). Voyons deux exemples.


Exemple 35 – Existence d'un polynôme interpolateur Soient un entier $n \in \mathbf{N}$ et une famille de $n + 1$ réels distincts $(x_k)_{0 \leq k \leq n} \in \mathbf{R}^{n+1}$. Montrons que pour toute famille de réels $(y_k)_{0 \leq k \leq n} \in \mathbf{R}^{n+1}$ il existe un unique $P \in \mathbf{R}_n[X]$ tel que :

$$\forall k \in \llbracket 0, n \rrbracket, \quad P(x_k) = y_k.$$

Le théorème précédent permet de montrer l'existence et l'unicité sans effort. En effet, considérons l'application

$$\Phi : \begin{cases} \mathbf{R}_n[X] & \longrightarrow & \mathbf{R}^{n+1} \\ P & \longmapsto & (P(x_0), \dots, P(x_n)). \end{cases}$$




 **Attention**
L'énoncé est faux en dimension infinie!

Théorème ALG.3.7

Soient E et F deux \mathbf{K} -espaces vectoriels de dimension finie telle que $\dim E = \dim F$ et $u \in \mathcal{L}(E, F)$. Alors :

$$(i) \ u \text{ est bijective} \iff (ii) \ u \text{ est injective} \iff (iii) \ u \text{ est surjective.}$$

Preuve D'après le théorème du rang, puisque E est de dimension finie, on a $\dim E = \text{Rg } u + \dim \ker u$. Montrons alors que **(ii)** \iff **(iii)** (l'équivalence

(i) \iff **(ii)** s'en suivra alors sans difficulté). 

4.5.2. Image de familles de vecteurs

Théorème ALG.3.8 | Construction d'une application linéaire à l'aide d'une base

Soient E un \mathbf{K} -espace vectoriel de dimension finie non nulle n et de **base** $\mathcal{B} = (e_1, \dots, e_n)$, F un \mathbf{K} -espace vectoriel de dimension quelconque, et $\mathcal{F} = (f_1, \dots, f_n)$ une famille **quelconque** de n vecteurs de F . Alors :

il existe une et une seule application linéaire $u \in \mathcal{L}(E, F)$ telle que


$$\forall i \in \llbracket 1, n \rrbracket, \quad u(e_i) = f_i,$$

De plus, avec u ainsi définie,

1. u est injective si et seulement si \mathcal{F} est une famille libre de F ,
2. u est surjective si et seulement si \mathcal{F} est une famille génératrice de F ,
3. u est un isomorphisme si et seulement si \mathcal{F} est une base de F .

Preuve

Commençons par construire u . 

La propriété **3)** est immédiate en combinant **1)** et **2)**. Montrons **1)**. 

Remarque 4.5 — Si l'on veut rendre le raisonnement plus explicite, on introduira les *polynômes d'interpolation de LAGRANGE*.

De-même pour 2).

$$\begin{aligned}
 u \text{ est surjective} &\iff \forall y \in F, \exists x = \sum_{i=1}^n \lambda_i(x) e_i, y = u(x) = \sum_{i=1}^n \lambda_i(x) f_i, \\
 &\iff \forall y \in F, \exists \mu_1, \dots, \mu_n \in \mathbf{K}, y = \sum_{i=1}^n \mu_i f_i, \\
 @ &\iff \mathcal{F} \text{ est génératrice.}
 \end{aligned}$$

Exemple 36 — Soit $n \in \mathbf{N}$, et notons $E = \mathbf{R}_n[X]$, $F = \mathbf{R}_{n+1}$.



Méthode Construction d'applications linéaires à l'aide d'une base

À la question « construisez une application linéaire entre E et F », si vous connaissez une base $\mathcal{B} = (e_1, \dots, e_n)$ de E, vous pouvez répondre :

je pose $u(e_1) = \text{Truc}_1 \in F$, ..., je pose $u(e_n) = \text{Truc}_n$,

en découlera alors automatiquement $u(x)$ pour tout $x \in E$ par linéarité.

Remarque 4.6 — On retiendra notamment deux idées importantes :

- ▶ une application linéaire est entièrement déterminée par l'image d'une base ; ce fait est important à analyser pour comprendre la notion de matrice d'application linéaire du prochain chapitre,
- ▶ u est un isomorphisme si et seulement si l'image **d'une** base de E est une base de F.

La première partie de la remarque nous conduit tout droit au chapitre suivant : comment représenter les applications linéaires notamment par des matrices ?

★ ★ ★ **Fin du chapitre** ★ ★ ★

5. EXERCICES

Exercice ALG.3.1 | Vrai ou Faux? En cas de réponse fausse, on donnera un contre-exemple.

1. La famille $x \mapsto x, x \mapsto -x, x \mapsto |x|$ est libre.
2. L'intersection de deux sous-espaces vectoriels est un sous-espace vectoriel.
3. Tout élément du sous-espace vectoriel engendré par des vecteurs x_1, \dots, x_n est une combinaison linéaire de deux de ces vecteurs.
4. Si u, v sont deux applications linéaires telles que $u \circ v = 0$, alors $u = 0$ ou $v = 0$.
5. Si E est de dimension finie, et $x, y \in E$, alors il existe un endomorphisme u de E tel que $u(x) = y$.

5.1. Structure d'espace vectoriel

Exercice ALG.3.2 | Être ou ne pas être un espace vectoriel dans \mathbf{K}^n Les ensembles ci-dessous, sont-ils, munis de l'addition et la multiplication par un scalaire des vecteurs, des espaces vectoriels? On les représentera graphiquement lorsque cela est possible.

1. $E_1 = \{(x, y) \in \mathbf{R}_+^2, x = y\}$, $E'_1 = \{(x, y) \in \mathbf{R}^2, x \neq y\}$, $E''_1 = \{(x, y) \in \mathbf{R}^2, x \geq y\}$,
2. $E_2 = (Ox) \cup (Oy)$, $E'_2 = \{(x, y) \in \mathbf{R}^2, y = 1\}$,
3. $E_3 = \{(x, y) \in \mathbf{R}^2, y - x = 1\}$,
4. $E_4 = \{(x, y) \in \mathbf{R}^2, x^2 - y^2 = 0\}$, $E'_4 = \{(x, y) \in \mathbf{R}^2, x^2 + y^2 = 0\}$,
5. $E_5 = \{(x, y, z) \in \mathbf{R}^3, x + y + z = 0, x - y + z = 0\}$, puis trouver une famille génératrice de E_5 (dit autrement, on demande d'écrire l'espace sous forme paramétrique),
6. l'ensemble E_6 des matrices triangulaires supérieures réelles d'ordre $n \geq 1$, $E'_6 = \{M \in \mathfrak{M}_{n,1}(\mathbf{R}), {}^T M + M = 0\}$.

Solution (exercice ALG.3.2)

1. $E_1 = \{(x, y) \in \mathbf{R}_+^2, x = y\}$ n'est pas un espace vectoriel car il n'est pas stable par opposé. En effet, $(1, 1) \in E_1$ mais $-(1, 1) = (-1, -1) \notin E_1$. L'ensemble E'_1 n'est pas un sous-espace vectoriel de \mathbf{R}^2 puisqu'il ne contient pas $(0, 0)$. L'ensemble E''_1 n'est pas non plus un sous-espace vectoriel de \mathbf{R}^2 , en effet $(1, 0) \in E''_1$, $(-1, -3)$ alors que $(1, 0) - (-1, -3) = (2, 3) \notin E''_1$ car $3 > 2$.
2. $E_2 = (Ox) \cup (Oy) = \text{Vect}(1, 0) \cup \text{Vect}(0, 1)$. Le vecteur $(1, 1) = (1, 0) + (0, 1)$ n'est pas dans E_2 , donc E_2 n'est pas stable par somme, ce n'est donc pas un espace vectoriel, de-même pour E'_2 car il ne contient pas $(0, 0)$.
3. E_3 n'est pas un espace vectoriel car il ne contient pas $(0, 0)$.
4. $E_4 = \{(x, y) \in \mathbf{R}^2, x^2 - y^2 = 0\}$ n'est pas un espace vectoriel. En effet, le vecteur $(1, 1)$ est un élément de E_4 , de-même pour $(2, -2)$, mais pas $(2, -2) + 2(1, 1) = (4, 0)$. L'ensemble E'_4 est égal à $\{(0, 0)\}$ puisque $x^2 + y^2 = 0 \iff x = y = 0$ pour $(x, y) \in \mathbf{R}^2$.
5. $E_5 = \{(x, y, z) \in \mathbf{R}^3, x + y + z = 0, x - y + z = 0\}$. Exprimons cet espace sous forme d'un Vect. Soit $(x, y, z) \in \mathbf{R}^3$, on a :

$$\begin{aligned} & \begin{cases} x + y + z = 0 \\ x - y + z = 0 \end{cases} \\ \iff & \begin{cases} x + y + z = 0 \\ -2y = 0 \end{cases} \quad L_2 \leftarrow L_2 - L_1 \\ \iff & \begin{cases} x = -z \\ y = 0 \end{cases} \\ \iff & (x, y, z) = (x, 0, -x) = x(1, 0, -1). \end{aligned}$$

- Donc $E_5 = \text{Vect}(1, 0, -1)$, c'est en particulier un sous-espace vectoriel de \mathbf{R}^3 .
6. E_6 est un sous-espace vectoriel de $\mathfrak{M}_{n,n}(\mathbf{R})$ puisque la matrice nulle est en particulier triangulaire supérieure, et que toute combinaison linéaire de matrices triangulaires supérieures est triangulaire supérieure. Le second est aussi un sous-espace vectoriel de $\mathfrak{M}_{n,n}(\mathbf{R})$.
 - ▶ ${}^T 0 + 0 = 0$, où 0 désigne ici la matrice nulle de format $n \times n$.
 - ▶ L'ensemble en question est bien inclus dans l'espace vectoriel des matrices $\mathfrak{M}_{n,n}(\mathbf{R})$.
 - ▶ Soient deux matrices $M, N \in \mathfrak{M}_{n,n}(\mathbf{R})$ vérifiant ${}^T M + M = 0$ et ${}^T N + N = 0$ ainsi

que $\lambda, \mu \in \mathbf{R}$. On a, par linéarité de la transposition :

$$\begin{aligned} {}^T(\lambda M + \mu N) + \lambda M + \mu N &= \lambda {}^T M + \mu {}^T N + \lambda M + \mu N \\ &= \lambda ({}^T M + M) + \mu ({}^T N + N) = \lambda 0 + \mu 0 = 0. \end{aligned}$$

Donc E_6 est un sous-espace vectoriel de $\mathfrak{M}_{n,n}(\mathbf{R})$.

Exercice ALG.3.3 | Être ou ne pas être un espace vectoriel de suites, fonctions, ...

Pour chacun des ensembles suivants, indiquer avec justification si c'est un espace vectoriel ou pas, en précisant le corps de base.

1. L'ensemble des fonctions de \mathbf{R} dans \mathbf{R} qui sont nulles en 1 ou nulles en 4.
2. L'ensemble des fonctions f croissantes sur \mathbf{R} .
3. L'ensemble des suites arithmétiques de $\mathbf{R}^{\mathbf{N}}$. Soit $r \in \mathbf{R}$, même question avec l'ensemble des suites arithmétiques de raison r .
4. L'ensemble des suites géométriques de $\mathbf{R}^{\mathbf{N}}$. Soit $\rho \in \mathbf{R}$, même question avec l'ensemble des suites géométriques de raison ρ .
5. L'ensembles des fonctions réelles définies sur $] - 1, 1[$, continues, positives ou nulles.

Solution (exercice ALG.3.3)

1. Faux, considérons $f : x \mapsto x - 4$, $g : x \mapsto x - 1$, alors les deux fonctions appartiennent à l'ensemble considéré, et pourtant $f + g$ ne s'annule ni en 1 ni en 4 ($(f + g)(1) = -3$, $(f + g)(0) = -5$). L'ensemble n'est donc pas stable par somme, *a fortiori* ce n'est pas un espace vectoriel.
2. Faux. Par exemple g définies précédemment est croissante, et pourtant $-g$ ne l'est pas.
3. \blacktriangleright Vrai dans le premier cas. En effet, notons $r, r' \in \mathbf{R}$ et (u_n) une suite arithmétique de raison r , (v_n) une suite arithmétique de raison r' . Soient $\lambda, \mu \in \mathbf{R}$, alors pour tout $n \in \mathbf{N}$,

$$\lambda u_{n+1} + \mu v_{n+1} = \lambda(u_n + r) + \mu(v_n + r') = \lambda u_n + \mu v_n + (\lambda r + \mu r').$$

Ceci prouve que $(\lambda u_n + \mu v_n)_n$ est encore arithmétique, de raison $\lambda r + \mu r'$.

- \blacktriangleright En revanche, si l'on impose la même raison r pour les deux suites, cela ne fonctionne plus. En effet, la suite $(u_n) = (1 + nr)_n$, $(v_n) = (nr)_n$ sont arithmétiques de raison r , et pourtant $(u_n) + (v_n) = (1 + n(2r))$ est arithmétique de raison $2r$. L'ensemble n'est pas donc pas stable par somme.
4. Pour les suites géométriques, c'est curieusement l'inverse. Si l'on impose la raison, l'ensemble forme un espace vectoriel, et par contre ce n'est pas le cas dans le cas contraire.
 - \blacktriangleright Faux, si l'on impose pas de raison. En effet, (2^n) est géométrique de raison 2, (1) est géométrique de raison 1. Et pourtant $(2^n + 1)$ n'est pas géométrique car $\left(\frac{2^{n+1}+1}{2^{n+1}}\right)$ n'est pas une suite constante (valeurs en 0 et 1 distinctes par exemple).
 - \blacktriangleright Notons $(u_n), (v_n)$ deux suites géométriques de raison ρ . Soient $\lambda, \mu \in \mathbf{R}$, alors pour tout $n \in \mathbf{N}$,

$$\lambda u_{n+1} + \mu v_{n+1} = \lambda \rho u_n + \mu \rho v_n = \rho(\lambda u_n + \mu v_n).$$

Ceci prouve que $(\lambda u_n + \mu v_n)_n$ est géométrique de raison ρ . L'ensemble des suites géométriques de raison ρ est donc un sous-espace vectoriel de l'ensemble des suites.

5. Faux. En effet, $x \mapsto x$ est positive continue ou nulle sur $] - 1, 1[$, mais $x \mapsto -x$ ne l'est pas. L'ensemble n'est donc pas stable par opposé et n'est *a fortiori* pas un espace vectoriel.

5.2. Familles de vecteurs et dimension

Exercice ALG.3.4 | Études de liberté/génération dans K^n

1. Les familles suivantes sont-elles libres ?
 - 1.1) $\mathcal{L}_1 = ((1, 1, 1), (2, 2, 2))$,
 - 1.2) $\mathcal{L}_2 = ((1, 0, 1), (0, 2, 2), (3, 3, 0))$,
 - 1.3) $\mathcal{L}_3 = ((-1, -2, 2), (4, -3, -2), (2, -1, -1))$.
2. Les familles suivantes sont-elles génératrices de \mathbf{R}^3 ?

2.1) $\mathcal{G}_1 = ((1, 0, 0), (0, 1, 0), (0, 0, 2)),$

2.2) $\mathcal{G}_2 = ((0, 1, 1), (1, 0, 1), (1, 1, 0)).$

Solution (exercice ALG.3.4)

1. 1.1) Clairement non, puisque $(2, 2, 2) = 2(1, 1, 1).$

1.2) Soient $\lambda, \mu, \nu \in \mathbf{R}$ tels que $\lambda(1, 0, 1) + \mu(0, 2, 2) + \nu(3, 3, 0) = (0, 0, 0).$ Alors

$$\begin{cases} \lambda + 3\nu = 0, \\ 2\mu + 3\nu = 0, \\ \lambda + 2\mu = 0 \end{cases} . \text{ En faisant } L_3 \rightarrow L_3 - L_1 \text{ et } L_3 \leftarrow L_3 + L_2, \text{ on a alors :}$$

$$\begin{cases} \lambda + 3\nu = 0, \\ 2\mu + 3\nu = 0, \\ 2\mu - 3\nu = 0 \end{cases} \iff \begin{cases} \lambda + 3\nu = 0, \\ 2\mu + 3\nu = 0, \\ 4\mu = 0 \end{cases}$$

On déduit alors $\lambda = \mu = \nu = 0,$ la famille est libre.

1.3) Soient $\lambda, \mu, \nu \in \mathbf{R}$ tels que $\lambda(-1, -2, 2) + \mu(4, -3, -2) + \nu(2, -1, -1) = (0, 0, 0).$

Alors $\begin{cases} -\lambda + 4\mu + 2\nu = 0, \\ -2\lambda - 3\mu - \nu = 0, \\ 2\lambda - 2\mu - \nu = 0 \end{cases} .$ Faisons les opérations $L_2 \leftarrow L_2 - 2L_1, L_3 \leftarrow L_3 + 2L_1,$ on obtient

$$\begin{cases} -\lambda + 4\mu + 2\nu = 0, \\ -11\mu - 5\nu = 0, \\ 6\mu + 3\nu = 0 \end{cases} .$$

Avec les deux dernières lignes, on obtient $\mu = \nu = 0,$ puis $\lambda = 0$ dans la première ligne.

2. Les familles suivantes sont-elles génératrices de \mathbf{R}^3 ?

2.1) Puisque $\text{Vect}(\mathcal{G}_1) = \text{Vect}((1, 0, 0), (0, 1, 0), (0, 0, 2)) = \text{Vect}((1, 0, 0), (0, 1, 0), (0, 0, 1))$ et que $((1, 0, 0), (0, 1, 0), (0, 0, 1))$ est génératrice (c'est la base canonique), la famille \mathcal{G}_1 est elle aussi génératrice.

2.2) Soit $(x, y, z) \in \mathbf{R}^3.$ Alors cherchons λ, μ, ν tels que

$$(x, y, z) = \lambda(0, 1, 1) + \mu(1, 0, 1) + \nu(1, 1, 0).$$

D'où le système $\begin{cases} x = \mu + \nu, \\ y = \lambda + \nu, \\ z = \lambda + \mu \end{cases} .$ Faisons $L_3 \leftarrow L_3 - L_2.$ Donc

$$\begin{cases} x = \mu + \nu, \\ y = \lambda + \nu, \\ z = \mu - \nu \end{cases} .$$

En considérant les deux premières lignes nous avons : $\mu = \frac{x+z}{2}, \nu = \frac{x-z}{2}.$ Donc $\lambda = y - \frac{x-z}{2}.$ On a donc, pour tout (x, y, z) une solution $(\lambda, \mu, \nu),$ donc la famille est génératrice de $\mathbf{R}^3.$

Exercice ALG.3.5 | Avec un paramètre Pour $m \in \mathbf{R},$ on note $u = (4 - m, 4, 4), v = (3, 3 - m, 6)$ et $w = (3, 6, 3 - m).$ Pour quelle(s) valeur(s) du paramètre m la famille (u, v, w) est-elle une base de \mathbf{R}^3 ?

Solution (exercice ALG.3.5)

Soient λ, μ, ν telles que

$$\lambda(4 - m, 4, 4) + \mu(3, 3 - m, 6) + \nu(3, 6, 3 - m) = 0,$$

on obtient alors le système suivant, que l'on souhaite résoudre en (λ, μ, ν) par la méthode du pivot. Commençons par échanger les lignes 1 et 3 afin d'avoir un pivot indépendant de λ en position $(1, 1).$

$$\begin{aligned} & \begin{cases} (4 - m)\lambda + 3\mu + 3\nu = 0, \\ 4\lambda + (3 - m)\mu + 6\nu = 0, \\ 4\lambda + 6\mu + (3 - m)\nu = 0, \end{cases} \\ & \stackrel{\sim}{L} \begin{cases} 4\lambda + 6\mu + (3 - m)\nu = 0, \\ 4\lambda + (3 - m)\mu + 6\nu = 0, \\ (4 - m)\lambda + 3\mu + 3\nu = 0, \end{cases} \quad \begin{matrix} \left. \begin{matrix} L_1 \leftrightarrow L_3 \\ L_2 \leftarrow L_2 - L_1, \\ L_3 \leftarrow 4L_3 - (4 - m)L_1 \end{matrix} \right\} \end{matrix} \\ & \stackrel{\sim}{L} \begin{cases} 4\lambda + 6\mu + (3 - m)\nu = 0, \\ -(m + 3)\mu + (m + 3)\nu = 0, \\ 6(m - 2)\mu + m(7 - m)\nu = 0, \end{cases} \end{aligned}$$

Deux cas se présentent alors :

► si $m = -3$, alors le système devient

$$\begin{cases} 2\lambda + 3\mu + 3\nu = 0, \\ 6(-5)\mu + 10(-3)\nu = 0, \end{cases} \iff \begin{cases} 2\lambda = 0, \\ \mu = -\nu. \end{cases}$$

Ainsi, le triplet $(0, 1, -1)$ est une solution non nulle du système donc la famille n'est pas libre.

► Si $m \neq -3$, alors le système est

$$\begin{cases} 4\lambda + (9 - m)\mu = 0, \\ \mu = \nu, \\ (6(m - 2) + m(7 - m))\mu = 0, \end{cases}$$

mais $6(m - 2) + m(7 - m) = -m^2 + 13m - 12 = -(m - 1)(m - 12)$. Donc si $m \in \{1, 12\}$, on obtient $\begin{cases} 4\lambda + (9 - m)\mu = 0, \\ \mu = \nu, \end{cases}$, il existe alors des solutions non nulles puisque $9 - m \neq 0$. La famille n'est pas libre. En revanche, si $m \neq 1, 12$, le système devient

$$\begin{cases} 4\lambda + (9 - m)\mu = 0, \\ \mu = \nu, \\ \mu = 0, \end{cases} \iff \lambda = \mu = \nu = 0.$$

La famille est donc libre dans ce cas.

A-t-on trois vecteurs dans cette famille? $u \neq v$ car $4 \neq 6$, et $v = w$ si et seulement si $m = -3$. Donc lorsque $m \neq -3$, la famille considérée est de cardinal 3. En conclusion

$$(u, v, w) \text{ est une base de } \mathbf{R}^3 \text{ si et seulement si } m \neq -3, 1, 12.$$

Exercice ALG.3.6 | Pour $a \neq b$ deux complexes et $n \geq 1$, on définit pour tout $k \in \llbracket 0, n \rrbracket$, $P_k = (X - a)^k (X - b)^{n-k}$. Montrer que (P_0, \dots, P_n) est libre dans $\mathbf{C}[X]$.

Solution (exercice ALG.3.6)

Faisons une récurrence sur n , le nombre de polynômes -1 de la famille.

■ **Initialisation.** Pour $n = 0$, puisque $P_0 = 1$ n'est pas le polynôme nul, la famille est clairement libre.

■ **Hérédité.** Supposons la famille libre pour $n + 1$ polynômes. Alors considérons $(P_0, \dots, P_n, P_{n+1})$ et montrons que cette famille est libre. Soient $\lambda_0, \dots, \lambda_{n+1} \in \mathbf{C}$ tels que :

$$\lambda_0(X - b)^{n+1} + \lambda_1(X - a)(X - b)^n + \dots + \lambda_n(X - a)^n(X - b) + \lambda_{n+1}(X - a)^{n+1} = 0.$$

Alors évaluons en b , on obtient puisque $a - b \neq 0$

$$\lambda_0 0 + \lambda_1 0 + \dots + \lambda_n 0 + \lambda_{n+1}(b - a)^{n+1} = 0 \implies \lambda_{n+1} = 0.$$

Donc :

$$(X - b)(\lambda_0(X - b)^n + \lambda_1(X - a)(X - b)^{n-1} + \dots + \lambda_n(X - a)^n) = 0.$$

Comme $X - b \neq 0_{\mathbf{C}[X]}$, on a :

$$\lambda_0(X - b)^n + \lambda_1(X - a)(X - b)^{n-1} + \dots + \lambda_n(X - a)^n = 0.$$

Il reste à appliquer l'hypothèse de récurrence à la famille (P_0, \dots, P_n) , et l'on obtient :

$$\lambda_0 = \dots = \lambda_n = 0.$$

Donc la famille est libre.

Exercice ALG.3.7 | Étude de liberté de familles de fonctions

1. Pour $k = 0, \dots, n$ on définit $e_k : x \mapsto e^{kx}$. Montrer que $(e_k)_{0 \leq k \leq n}$ est libre dans $\mathcal{F}(\mathbf{R}, \mathbf{R})$.
2. Pour $k = 0, \dots, n$ on définit, $f_k : x \mapsto |x - k|$. Montrer que $(f_k)_{0 \leq k \leq n}$ est une famille libre de $\mathcal{F}(\mathbf{R}, \mathbf{R})$.

Solution (exercice ALG.3.7)

1. Montrons le résultat par récurrence sur $n \in \mathbf{N}$. Notons \mathcal{P}_n la propriété «la famille $e_k : x \mapsto e^{kx}$ pour $k = 0, \dots, n$ est libre.

■ **Initialisation.** Pour $n = 0$ c'est immédiat, puisque l'exponentielle est non nulle.

■ **Hérédité.** Supposons \mathcal{P}_n vraie pour un certain entier $n \in \mathbf{N}$. Alors soit $(\lambda_0, \lambda_1, \dots, \lambda_{n+1}) \in \mathbf{R}^{n+2}$ telle que :

$$\forall x \in \mathbf{R}, \quad \lambda_0 + \lambda_1 e^x + \dots + \lambda_n e^{nx} + \lambda_{n+1} e^{(n+1)x} = 0.$$

Alors en factorisant par $e^{(n+1)x}$ dans l'identité précédente, nous obtenons :

$$\forall x \in \mathbf{R}, \quad e^{-(n+1)x} \left(\frac{\lambda_0 + \lambda_1 e^x + \dots + \lambda_n e^{nx}}{e^{(n+1)x}} + \lambda_{n+1} \right) = 0.$$

Si $\lambda_{n+1} \neq 0$, alors en faisant $x \rightarrow \infty$ nous obtenons une limite infinie dans le membre de gauche, ce qui est contradictoire. Donc $\lambda_{n+1} = 0$. Et par hypothèse de récurrence, les autres constantes sont nulles également.

En conclusion, $(e_k)_{0 \leq k \leq n}$ est libre.

2. Soit $(\lambda_0, \lambda_1, \dots, \lambda_n) \in \mathbf{R}^{n+1}$ telle que :

$$\forall x \in \mathbf{R}, \quad \lambda_0 |x| + \lambda_1 |x - 1| + \dots + \lambda_n |x - n| = 0.$$

Constatons que pour $k = 0, \dots, n$ la fonction $f_k : x \mapsto |x - k|$ est dérivable sur \mathbf{R} sauf en $x = k$. Ainsi, si l'un des λ_k est non nul, par exemple λ_0 , on a :

$$\forall x \in \mathbf{R}, \quad |x| = \frac{1}{\lambda_0} (-\lambda_1 |x - 1| + \dots + \lambda_n |x - n|) = 0.$$

Or la fonction $x \mapsto \frac{1}{\lambda_0} (-\lambda_1 |x - 1| + \dots + \lambda_n |x - n|)$ est dérivable en zéro, alors que $x \mapsto |x|$ ne l'est pas — contradiction. Ainsi, $\lambda_0 = 0$, il suffit ensuite de recommencer l'opération pour tous les autres λ_i . En conclusion, $(f_k)_{0 \leq k \leq n}$ est libre.

Exercice ALG.3.8 | Soient les deux ensembles ci-dessous.

$$F = \{(x, y, z) \in \mathbf{R}^3, y = z\} \quad \text{et} \quad G = \{(x, y, z) \in \mathbf{R}^3, x + y + 2z = 0\}.$$

1. Montrer que F et G sont des sous-espaces vectoriels de \mathbf{R}^3 , et donner une base de F et de G.

2. Déterminer une base de $F \cap G$.

Solution (exercice ALG.3.8)

1. Nous savons, d'après le cours de géométrie, que ce sont des équations cartésiennes d'hyperplans de \mathbf{R}^3 .



$$\begin{aligned} (x, y, z) \in F &\iff y = z \\ &\iff (x, y, z) = (x, y, y) = x(1, 0, 0) + y(0, 1, 1). \end{aligned}$$

Ainsi $F = \text{Vect}((1, 0, 0), (0, 1, 1))$, c'est donc un sous-espace vectoriel de \mathbf{R}^3 et $((1, 0, 0), (0, 1, 1))$ en est une famille génératrice. Elle est clairement libre donc c'est une base de F, $\dim F = 2$.



$$\begin{aligned} (x, y, z) \in G &\iff x + y + 2z = 0 \\ &\iff (x, y, z) = (-y - 2z, y, z) = y(-1, 1, 0) + z(-2, 0, 1). \end{aligned}$$

Ainsi $F = \text{Vect}((-1, 1, 0), (-2, 0, 1))$, c'est donc un sous-espace vectoriel de \mathbf{R}^3 et $((-1, 1, 0), (-2, 0, 1))$ en est une famille génératrice. Elle est clairement libre donc c'est une base de G, $\dim G = 2$.

$$\begin{aligned} 2. (x, y, z) \in F \cap G &\iff \begin{cases} x + y + 2z = 0, \\ y = z \end{cases} \\ &\iff x = -3z, y = z \iff (x, y, z) = (-3z, z, z) = z(-3, 1, 1). \end{aligned}$$

Donc : $F \cap G = \text{Vect}((-3, 1, 1))$.

Exercice ALG.3.9 | **Un hyperplan de \mathbf{R}^n** Soient a_1, \dots, a_{n-1} des réels et F le sous-ensemble de \mathbf{R}^n défini par :

$$F = \{(x_1, \dots, x_n) \in \mathbf{R}^n, a_1 x_1 + \dots + a_{n-1} x_{n-1} + x_n = 0\}.$$

1. Montrer, de deux manières, que F est un sous-espace vectoriel de \mathbf{R}^n .
2. Déterminer une base et la dimension de F.

Solution (exercice ALG.3.9)

1. ▶ Première méthode :

$$(x_1, \dots, x_n) \in F$$

$$\iff a_1 x_1 + \dots + a_{n-1} x_{n-1} + x_n = 0$$

$$\iff (x_1, \dots, x_n) = (x_1, \dots, x_{n-1}, -a_1 x_1 - \dots - a_{n-1} x_{n-1})$$

$$\iff (x_1, \dots, x_n) = x_1(1, 0, \dots, 0, -a_1) +$$

$$x_2(0, 1, 0, \dots, 0, -a_2) + \dots + x_{n-1}(0, \dots, 0, 1, -a_{n-1}).$$

Donc F est engendrés par les vecteurs ci-après :

$$(1, 0, \dots, 0, -a_1), (0, 1, 0, \dots, 0, -a_2), \dots, (0, \dots, 0, 1, -a_{n-1}).$$

C'est donc un sous-espace vectoriel de \mathbf{R}^n .

- ▶ Soient $X = (x_1, \dots, x_n), Y = (y_1, \dots, y_n) \in F$, et $\lambda, \mu \in \mathbf{R}$. Alors

$$Z = \lambda X + \mu Y = (\lambda x_1 + \mu y_1, \dots, \lambda x_n + \mu y_n) \in F$$

car

$$a_1(\lambda x_1 + \mu y_1) + \dots + a_n(\lambda x_n + \mu y_n)$$

$$= \lambda(a_1 x_1 + \dots + a_{n-1} x_{n-1} + x_n) + \mu(a_1 y_1 + \dots + a_{n-1} y_{n-1} + y_n)$$

$$= \lambda \cdot 0 + \mu \cdot 0 = 0.$$

Donc F est un sous-espace vectoriel de \mathbf{R}^3 .

2. De ce qui précède, une famille génératrice est donnée par

$$(1, 0, \dots, 0, -a_1), (0, 1, 0, \dots, 0, -a_2), \dots, (0, \dots, 0, 1, -a_{n-1}).$$

Les 0 étant dans des positions différentes, la famille est donc clairement libre, c'est une base et $\dim F = n - 1$.

Exercice ALG.3.10 | Soit E l'espace des applications de $] - 1; 1[$ dans \mathbf{R} . Soit $f_1 : x \mapsto \frac{1}{x-1}$ et $f_2 : x \mapsto \frac{1}{x+1}$.

1. Montrer que la famille (f_1, f_2) est libre.
2. Montrer que $g : x \mapsto \frac{1}{x^2-1}$ appartient à $\text{Vect}(f_1, f_2)$.

Solution (exercice ALG.3.10)

1. Soient λ_1, λ_2 tels que :

$$\forall x \in] - 1, 1[, \quad \frac{\lambda_1}{x-1} + \frac{\lambda_2}{x+1} = \frac{\lambda_2(x+1) + \lambda_1(x-1)}{(x-1)(x+1)} = 0,$$

donc en identifiant les termes en x et constants au numérateur, on déduit

$$\lambda_1 + \lambda_2 = 0, \quad \lambda_2 - \lambda_1 = 0,$$

d'où l'on tire facilement $\lambda_1 = \lambda_2 = 0$.

2. Montrons que $g : x \mapsto \frac{1}{x^2-1}$ appartient à $\text{Vect}(f_1, f_2)$. On cherche λ, μ telles que pour tout $x \in] - 1, 1[$, on a

$$\frac{1}{x^2-1} = \frac{\lambda}{x-1} + \frac{\mu}{x+1} = \frac{\lambda_2(x+1) + \lambda_1(x-1)}{x^2-1},$$

en identifiant les termes en x et constants au numérateur, on trouve que

$$\lambda = \frac{1}{2}, \mu = -\frac{1}{2} \text{ convient.}$$

Exercice ALG.3.11 | Soit $E = \mathcal{D}^2(\mathbf{R}, \mathbf{R})$ et

$$S = \{y \in \mathcal{D}^2(\mathbf{R}, \mathbf{R}), 2y'' + 2y' + y = 0\}.$$

1. Montrer que S est un sous-espace vectoriel de E.
2. Déterminer S et sa dimension.

3. Soit α fixé dans \mathbf{R} et $H_\alpha = \{f \in S, f(0) = f(\alpha) = 0\}$. Déterminer H_α ainsi que sa dimension.

Solution (exercice ALG.3.11)

1. S est inclus dans E par construction, de plus, $2.0'' + 2.0' + 0 = 0$ donc la fonction nulle est dans \mathcal{S} . Soient deux solutions $y_1, y_2 \in \mathcal{S}$, et $\lambda, \mu \in \mathbf{R}$,

$$\begin{aligned} & 2(\lambda y_1 + \mu y_2)'' + 2(\lambda y_1 + \mu y_2)' + (\lambda y_1 + \mu y_2) \\ &= \lambda(2y_1'' + 2y_1' + y_1) + \mu(2y_2'' + 2y_2' + y_2) \end{aligned} \quad \left. \begin{array}{l} \\ \end{array} \right\} \text{linéarité de la dérivation}$$

$$= \lambda 0 + \mu 0 = 0.$$

Donc $\lambda y_1 + \mu y_2$ est solution et \mathcal{S} est bien un sous-espace vectoriel de S .

2. Notons (EC) $2x^2 + 2x + 1 = 0$ l'équation caractéristique associée, qui a pour discriminant $4 - 8 = -4$ et donc pour racines

$$\frac{-2 \pm 2i}{4} = -\frac{1}{2} \pm i \frac{1}{2}.$$

Ainsi, les solutions sont les fonctions de la forme

$$x \mapsto e^{-x/2} (A \cos(x/2) + B \sin(x/2)), A, B \in \mathbf{R},$$

autrement dit

$$\mathcal{S} = \text{Vect}(x \mapsto e^{-x/2} \cos(x/2), x \mapsto e^{-x/2} \sin(x/2)).$$

Montrons que la famille est libre, soit donc $\lambda, \mu \in \mathbf{R}$ tels que

$$\forall x \in \mathbf{R}, \quad \lambda e^{-x/2} \cos(x/2) + \mu e^{-x/2} \sin(x/2) = 0,$$

dès lors, en simplifiant par l'exponentielle, on déduit

$$\forall x \in \mathbf{R}, \quad \lambda \cos(x/2) + \mu \sin(x/2) = 0.$$

Il reste par exemple à prendre deux valeurs particulières de x :

$$\lambda = 0 \quad (x = 0), \quad \mu = 0 \quad (x = \pi).$$

Donc la famille proposée est libre et de cardinal 2, donc $\dim \mathcal{S} = 2$.

3. Soit α fixé dans \mathbf{R} et $H_\alpha = \{f \in S, f(0) = f(\alpha) = 0\}$. Soit $y \in \mathcal{S}$, alors choisissons A, B réels de sorte que $y(x) = e^{-x/2} (A \cos(x/2) + B \sin(x/2))$ pour tout $x \in \mathbf{R}$.

$$\begin{aligned} y \in H_\alpha &\iff y(0) = y(\alpha) = 0, \\ &\iff \begin{cases} A = 0, \\ e^{-\alpha/2} B \sin(\alpha/2) = 0 \end{cases} \\ &\iff \begin{cases} A = 0, \\ B \sin(\alpha/2) = 0 \end{cases} \end{aligned}$$

Deux cas se présentent alors.

- ▶ si $\sin(\alpha/2) = 0$, alors B est quelconque et $H_\alpha = \text{Vect}(x \mapsto e^{-x/2} \sin(x/2))$ donc $\dim H_\alpha = 1$.
- ▶ Si $\sin(\alpha/2) \neq 0$, alors $B = 0$ et $H_\alpha = \{0\}$ donc $\dim H_\alpha = 0$.

Exercice ALG.3.12 | Soit $E = \mathbf{R}^N$ et q_1, q_2 deux réels non nuls tels que $|q_1| < |q_2|$. On pose

$$F = \text{Vect}((q_1^n), (q_2^n)).$$

1. Justifier que F est de dimension finie et déterminer sa dimension.
2. Notons $F_0 = \{(u_n) \in F, u_0 = u_1 = 0\}$. Déterminer F_0 ainsi que sa dimension.

Solution (exercice ALG.3.12)

1. La famille $((q_1^n), (q_2^n))$ est une famille génératrice de F donc F est de dimension finie. Montrons qu'elle est libre. Soient λ, μ deux réels, tels que

$$\forall n \in \mathbf{N}, \quad \lambda q_1^n + \mu q_2^n = 0.$$

Donc comme $|q_1| < |q_2|$, $|\frac{q_1}{q_2}| < 1$. Ainsi, $\lim_{n \rightarrow \infty} \frac{q_1}{q_2} = 0$. Donc

$$\lambda \left(\frac{q_1}{q_2}\right)^n + \mu = 0,$$

puis en passant à la limite on obtient $\mu = 0$. On peut ensuite faire $n = 1$ dans l'hypothèse initiale, ce qui livre $\lambda = 0$. Donc la famille considérée est une base, et $\dim F = 2$.

2. Si $(u_n) \in F_0$, alors il existe λ, μ tels que

$$\forall n \in \mathbf{N}, \quad u_n = \lambda q_1^n + \mu q_2^n.$$

La condition $u_0 = u_1 = 0$ donne

$$\lambda 0 + \mu = 0, 0 = \lambda q_1 + \mu q_2.$$

Ou de manière équivalente,

$$0 = \lambda(q_1 - q_2),$$

or $q_1 \neq q_2$ donc $\lambda = 0$ et $\mu = 0$. Donc $F_0 = \{(0)\}, \dim F_0 = 0$.

5.3. Applications linéaires

Exercice ALG.3.13 | Études de linéarité Dire si les applications suivantes sont des applications linéaires, le montrer le cas échéant.

1. $f: \mathbf{R} \rightarrow \mathbf{R}, x \mapsto 2x^2$,
2. $g: \mathbf{R} \rightarrow \mathbf{R}, x \mapsto 4x - 3$,
3. $h: \mathbf{R} \rightarrow \mathbf{R}^3, x \mapsto (2x, x/\pi, x\sqrt{2})$,
4. $S: \mathbf{R} \rightarrow \mathbf{R}^2, x \mapsto (u, v)$, où (u, v) est l'unique solution de $\begin{cases} 3u - v = x, \\ 6u + 2v = y, \end{cases}$
5. $D: \mathcal{C}^\infty(\mathbf{R}, \mathbf{R}) \rightarrow \mathcal{C}^\infty(\mathbf{R}, \mathbf{R}), f \mapsto ff''$,
6. $E: \mathcal{C}^\infty(\mathbf{R}, \mathbf{R}) \rightarrow \mathcal{C}^\infty(\mathbf{R}, \mathbf{R}), f \mapsto f(x^3)$,
7. $I: \mathcal{C}^\infty(\mathbf{R}, \mathbf{R}) \rightarrow \mathcal{C}^\infty(\mathbf{R}, \mathbf{R}), f \mapsto \left(x \mapsto \int_0^x (x-t)^2 f(t) dt\right)$.

Solution (exercice ALG.3.13)

1. L'application f n'est pas linéaire, à cause du carré. En effet, $f(2 \times 1) = f(2) = 2 \times 4 = 8 \neq 2f(1) (= 2.2)$.
2. Puisque $g(0) = -3 \neq 0$, l'application g n'est pas non plus linéaire.
3. Cette fois-ci c'est bon. Soient $x, x' \in \mathbf{R}$ et $\lambda, \mu \in \mathbf{R}$, alors

$$\begin{aligned} h(\lambda x + \mu x') &= (2(\lambda x + \mu x'), (\lambda x + \mu x')/\pi, (\lambda x + \mu x')\sqrt{2}) \\ &= (2\lambda x + 2\mu x', \lambda/\pi x + \mu/\pi x', \lambda\sqrt{2}x + \mu\sqrt{2}x') \\ &= \lambda(2x, x/\pi, x\sqrt{2}) + \mu(2x', x'/\pi, x'\sqrt{2}) \\ &= \lambda h(x) + \mu h(x'). \end{aligned}$$

4. Résolvons le système. On a, en notant $M = \begin{pmatrix} 2 & -1 \\ 6 & 2 \end{pmatrix}$, inversible d'inverse $M^{-1} = \frac{1}{12} \begin{pmatrix} 3 & 1 \\ -6 & 3 \end{pmatrix}$:

$$\begin{cases} 3u - v = x, \\ 6u + 2v = y, \end{cases} \iff M \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \iff \begin{pmatrix} u \\ v \end{pmatrix} = \frac{1}{12} \begin{pmatrix} 3 & 1 \\ -6 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Donc : $u = \frac{1}{12}(3x + y), v = \frac{1}{12}(-6x + 3y)$. On vérifie ensuite sans difficulté que les applications $(x, y) \mapsto \frac{1}{12}(3x + y)$ et $(x, y) \mapsto \frac{1}{12}(-6x + 3y)$ sont linéaires de \mathbf{R}^2 dans \mathbf{R} .

5. L'application n'est pas linéaire à cause du produit. Prenons $f = X^2$. Alors $D(2f)(x) = (2x^2) \times 4 = 8x^2 \neq 2D(f)(x) = 2x^2$ pour tout $x \in \mathbf{R}$ (au moins non nul).
6. L'application E est linéaire, par linéarité de l'évaluation en x^3 pour tout x . En effet, soient $f, g \in \mathcal{C}^\infty(\mathbf{R}, \mathbf{R})$ et $\lambda, \mu \in \mathbf{R}$. Alors :

$$E(\lambda f + \mu g) = (\lambda f + \mu g)(x^3) = \lambda f(x^3) + \mu g(x^3) = \lambda E(f) + \mu E(g).$$

Donc E est linéaire. Notons que $E(\mathcal{C}^\infty(\mathbf{R}, \mathbf{R})) \subset \mathcal{C}^\infty(\mathbf{R}, \mathbf{R})$, comme le prétend l'énoncé, puisqu'une composée d'applications \mathcal{C}^∞ et \mathcal{C}^∞ .

7. L'application I est linéaire, par linéarité de de l'intégration. En effet, soient $f, g \in$

$\mathcal{C}^\infty(\mathbf{R}, \mathbf{R})$ et $\lambda, \mu \in \mathbf{R}$. Alors, pour tout $x \in \mathbf{R}$,

$$\begin{aligned} I(\lambda f + \mu g)(x) &= \int_0^x (x-t)^2(\lambda f + \mu g)(t) dt \\ &= \int_0^x (\lambda(x-t)^2 f(t) + \mu(x-t)^2 g(t)) dt \\ &= \lambda I(f)(x) + \mu I(g)(x). \end{aligned}$$

$\left. \begin{array}{l} \text{linéarité de l'évaluation} \\ \text{linéarité de l'intégration} \end{array} \right\}$

Donc I est linéaire. Reste à justifier (même si l'énoncé le prétend!) que I(f) est de classe \mathcal{C}^∞ .

Il s'agit d'une intégrale à borne variable, cependant, l'intégrande (la fonction que l'on intègre) dépend elle aussi de x. Développons le carré : soit $x \in \mathbf{R}$,

$$I(f)(x) = x^2 \int_0^x f(t) dt - 2 \int_0^x t f(t) dt + \int_0^x t^2 f(t) dt.$$

Notons F (resp. G, resp. H) une primitive de f (resp. Id f, resp. Id² f), ces trois primitives existent car chacune des fonctions citée est continue. Or, chaque fonction F, G, H est \mathcal{C}^∞ puisque leur dérivée première est égale à une fonction \mathcal{C}^∞ . Ainsi, par somme/produit de fonctions \mathcal{C}^∞ , I(f) l'est aussi.

Exercice ALG.3.14 | On note $E = \mathbf{R}_2[X]$ et $F = \{P \in \mathbf{R}_2[X], P(1) = 0\}$.

1. Montrer que F est un sous-espace vectoriel de E.
2. Déterminer la dimension de F, en exhibant une base de F.
3. Déterminer une forme linéaire ϕ sur E, telle que : $F = \ker \phi$. Retrouver les résultats précédents en utilisant ϕ .

Solution (exercice ALG.3.14)

1. $F \subset E$, et $0(1) = 0$ donc F contient le polynôme nul. De plus, si $\lambda, \mu \in \mathbf{R}$ et $P, Q \in F$, alors

$$(\lambda P + \mu Q)(1) = \lambda P(1) + \mu Q(1) = \lambda \cdot 0 + \mu \cdot 0 = 0.$$

Ainsi, F est un sous-espace vectoriel de E.

2. Nous avons $P \in F$ si et seulement si il existe $\alpha, K \in \mathbf{R}$ tels que

$$P = K(X-1)(X-\alpha) = KX(X-1) - K\alpha(X-1), \quad (\star)$$

mais comme $X, X(X-1)$ sont deux polynômes de F, que la famille $(X, X(X-1))$ est échelonnée, il vient alors que $(X, X(X-1))$ est une base de F (elle est génératrice d'après l'égalité (\star)) et donc $\dim F = 2$.

3. On peut poser $\Phi : P \in \mathbf{R}_2[X] \mapsto P(1)$. Alors on vérifie sans difficulté que Φ est linéaire. D'après le théorème du rang (puis $\dim \mathbf{R}_2[X] < \infty$), on a :

$$\dim \ker \phi + \text{Rg} \phi = \dim \mathbf{R}_2[X] = 3.$$

Mais $\text{Rg}(\phi) \leq \dim \mathbf{R} = 1$ (car $\text{Im}(\phi)$ est un sous-espace vectoriel de \mathbf{R}). Or, $\text{Rg}(\phi) \neq 0$ car sinon ϕ est l'application nulle et ce n'est pas le cas (puisque par exemple $\phi(X) = 1$). Donc

$$\dim \ker \phi = 3 - 1 = 2.$$

Exercice ALG.3.15 | Soient $n \geq 3$ et $b \in \mathbf{R}$. On considère l'application

$$\varphi \begin{cases} \mathbf{R}_n[X] & \longrightarrow & \mathbf{R}[X], \\ P & \longmapsto & (X-b)(P' + P'(b)) - 2(P - P(b)). \end{cases}$$

1. Justifier que φ est un endomorphisme de $\mathbf{R}_n[X]$.
2. Soit $F = \{P \in \mathbf{R}_n[X], \exists Q \in \mathbf{R}_n[X], P = (X-b)^3 Q\}$.
 - 2.1) Justifier que $F = \{P \in \mathbf{R}_n[X], \exists Q \in \mathbf{R}_{n-3}[X], P = (X-b)^3 Q\}$.
 - 2.2) Déterminer une base puis la dimension de F. En utilisant $\varphi(P)''$, démontrer que $\text{Im} \varphi \subset F$.
 - 2.3) En utilisant à nouveau $\varphi(P)''$, démontrer que $\ker \varphi \subset \mathbf{R}_2[X]$.
 - 2.4) Déterminer noyau et image de φ .

Solution (exercice ALG.3.15)

1. Soient $P, Q \in \mathbf{R}_n[X]$, et $\lambda, \mu \in \mathbf{R}$. Alors

$$\begin{aligned} \varphi(\lambda P + \mu Q) &= (X - b)((\lambda P + \mu Q)' + (\lambda P + \mu Q)'(b)) - 2((\lambda P + \mu Q) - (\lambda P + \mu Q)(b)) \\ &= \lambda(X - b)(P' + P'(b)) - 2(P - P(b)) + \mu(X - b)(Q' + Q'(b)) - 2(Q - Q(b)) \end{aligned}$$

par linéarité de la dérivation, et de l'évaluation en b . De plus, $\deg((X - b)(P' + P'(b))) = n = \deg(P - P(b))$. Donc $\Phi(P) \in \mathbf{R}_n[X]$. En conclusion,

Φ est un endomorphisme de $\mathbf{R}_n[X]$.

2. 2.1) Condition de degré, avec les notations de l'énoncé, si $P \in F$ alors on a $\deg Q + 3 = \deg P$ donc $\deg Q \leq n - 3$.

2.2) Soit $P \in F$, alors il existe $Q \in \mathbf{R}_{n-3}[X]$ tel que $P = (X - b)^3 Q$. Il existe $\lambda_0, \dots, \lambda_{n-3} \in \mathbf{R}$ tels que $Q = \lambda_0 + \lambda_1 X + \dots + \lambda_{n-3} X^{n-3}$. Donc

$$P = \lambda_0(X - b)^3 + \lambda_1(X - b)^3 X + \dots + \lambda_{n-3}(X - b)^3 X^{n-3}.$$

Autrement dit, la famille $\mathcal{F} = ((X - b)^3, X(X - b)^3, \dots, X^{n-3}(X - b)^3)$ est une famille génératrice de F . Elle est de plus libre car échelonnée. Donc

\mathcal{F} est une base de F . Elle est de cardinal $n - 2$, donc $\dim F = n - 2$. Calculons $\varphi(P)''$: $\varphi(P)' = P' + P'(b) + (X - b)P'' - 2P'$ puis $\varphi(P)'' = P'' + P'' + (X - b)P'''' - 2P'' = (X - b)P''''$. Constatons que $\varphi(P)(b) = \varphi(P)'(b) = \varphi(P)''(b)$ donc b est au moins une racine de multiplicité au moins trois de $\varphi(P)$, c'est exactement dire que $\varphi(P) \in F$.

2.3) Recyclons une nouvelle fois le calcul précédent. Si $P \in \ker \varphi$, alors $\varphi(P) = 0$ donc $\varphi(P)'' = 0 = (X - b)P''''$. Par propriété du cours, comme $X - b \neq 0_{\mathbf{R}_n[X]}$, nous avons $P'''' = 0$, i.e. $P \in \mathbf{R}_2[X]$ en primitivant deux fois l'égalité $P'''' = 0$.

2.4) Appliquons le théorème du rang à φ puisque $\mathbf{R}_n[X]$ est de dimension finie $n + 1$. Nous avons

$$\dim \ker \varphi + \text{Rg} \varphi = n + 1.$$

Or, d'après les questions précédentes, $\dim \ker \varphi \leq 3 = \dim \mathbf{R}_2[X]$ et $\text{Rg} \varphi \leq \dim F = n - 3 + 1 = n - 2$. Donc comme $3 + n - 1 = n + 1$, on a nécessairement $\dim \ker \varphi = 3$ et $\text{Rg} \varphi = n - 2$. Par égalité des dimensions, les inclusions deviennent alors des égalités :

$$\ker \varphi = \mathbf{R}_2[X], \quad \text{Im}(\varphi) = F.$$

Exercice ALG.3.16 | Endomorphisme aux différences finies de $\mathbf{C}_n[X]$. Soient $n \in \mathbf{N}^*$ et Δ_n l'application :

$$\Delta_n \left| \begin{array}{l} \mathbf{C}_n[X] \longrightarrow \mathbf{C}_n[X], \\ P \longmapsto P(X + 1) - P(X). \end{array} \right.$$

1. Calculer $\Delta(1), \Delta(X), \Delta(X^2)$.
2. Montrer que Δ_n est un endomorphisme de $\mathbf{C}_n[X]$.
3. Soit $P \in \ker \Delta_n$ tel que $\deg(P) \geq 1$. Montrer que $P - P(0)$ a une infinité de racines. En déduire que $\ker \Delta_n = \mathbf{C}_0[X]$.
4. En déduire que $\text{Im}(\Delta_n) = \mathbf{C}_{n-1}[X]$.
5. On définit dans cette question une application Δ comme

$$\Delta \left| \begin{array}{l} \mathbf{C}[X] \longrightarrow \mathbf{C}[X] \\ P \longmapsto P(X + 1) - P(X) \end{array} \right.$$

À l'aide des questions précédentes, déterminer $\ker \Delta$, et montrer que Δ est surjective à l'aide de la définition de la surjectivité. Commenter.

Solution (exercice ALG.3.16)

1. $\Delta(1) = 1 - 1 = 0, \Delta(X) = X + 1 - X = 1, \Delta(X^2) = (X + 1)^2 - X^2 = 2X + 1$. Le degré semble chuter de un.
2. Soient $P, Q \in \mathbf{C}_n[X]$, et $\lambda, \mu \in \mathbf{C}$. Alors :

$$\begin{aligned} \Delta_n(\lambda P + \mu Q) &= (\lambda P + \mu Q)(X + 1) - (\lambda P + \mu Q)(X) \\ &= \lambda(P(X + 1) - P(X)) + \mu(Q(X + 1) - Q(X)) \\ &= \lambda \Delta_n(P) + \mu \Delta_n(Q). \end{aligned}$$

De plus, si $\deg P \leq n$ alors $\deg \Delta_n(P) \leq n$ puisque c'est une différence de polynômes de degré au plus n . Finalement, on a bien montré que

Δ_n est un endomorphisme de $\mathbf{C}_n[X]$.

3. Soit $P \in \ker \Delta_n$. Alors pour tout $x \in \mathbf{R}$, $P(x+1) = P(x)$. En particulier, $P(n) = P(0)$ pour tout $n \in \mathbf{N}$. Ainsi, $P - P(0)$ possède une infinité de racines, donc il est nul, et $P = P(0)$ est le polynôme constant. Donc $\ker \Delta_n \subset \mathbf{C}_0[X]$. Inversement, montrons que $\mathbf{C}_0[X] \subset \ker \Delta_n$. Soit $P = C \in \mathbf{C}$ un polynôme constant, alors $\Delta(P) = C - C = 0$.

En conclusion : $\ker \Delta_n = \mathbf{C}_0[X]$.

4. D'après le théorème du rang, $\dim \ker \Delta_n + \text{Rg} \Delta_n = \dim \mathbf{C}_n[X] = n+1$, donc comme $\dim \ker \Delta_n = 1$ d'après la question précédente, il vient $\text{Rg} \Delta_n = n+1-1 = n$. Il suffit alors de montrer que $\text{Im} \Delta_n \subset \mathbf{C}_{n-1}[X]$. Si tel que le cas, puisque $\dim \text{Im} \Delta_n = \dim \mathbf{C}_{n-1}[X] = n$, on aura l'égalité $\text{Im} \Delta_n = \mathbf{C}_{n-1}[X]$ par égalité des dimensions. Soit $P = \sum_{k=0}^n a_k X^k \in \mathbf{C}_n[X]$, alors par linéarité de Δ_n et d'après la formule du binôme,

$$\Delta_n(P) = \sum_{k=0}^n a_k \Delta_n(X^k) = \sum_{k=0}^n a_k \left(\sum_{\ell=0}^k \binom{k}{\ell} X^\ell - X^k \right).$$

On constate que les termes d'ordre $\ell = k$ dans la somme interne sont nuls pour tout k , donc

$$\Delta_n(P) = \sum_{k=0}^n a_k \sum_{\ell=0}^{k-1} \binom{k}{\ell} X^\ell.$$

Sous cette forme, on voit alors $\deg \Delta_n(P) = n-1$. On a bien montré que $\text{Im} \Delta_n \subset \mathbf{C}_{n-1}[X]$. Et donc par égalité des deux dimensions : $\text{Im}(\Delta_n) = \mathbf{C}_{n-1}[X]$.

5. Soit $Q \in \mathbf{C}[X]$, alors notons $q = \deg Q$. Nous avons donc $Q \in \mathbf{C}_q[X]$, alors puisque Δ_{q+1} est surjective d'après le début de l'exercice (et que Δ_{q+1} est à valeurs dans $\mathbf{C}_q[X]$), il existe $P \in \mathbf{C}_{q+1}[X]$ tel que $\Delta_{q+1}(P) = Q = \Delta(P)$. Donc : Δ est surjective. En revanche, Δ n'est pas injective puisque le noyau est encore une fois constitué des polynômes constants. Puisque $\mathbf{C}[X]$ n'est pas de dimension finie, cela n'est pas étonnant, l'injectivité n'est dans ce cadre pas nécessairement équivalente à la surjectivité.

Chapitre ALG.4.

Matrices

Résumé & Plan

Ce chapitre est d'abord constitué de révisions de première année sur les matrices, vues comme des tableaux de nombres sur lesquels des opérations sont autorisées. Nous constatons alors que l'ensemble des matrices, muni de deux opérations (addition, multiplication scalaire), est un espace vectoriel au sens du **Chapter ALG.3**. Enfin, nous reverrons enfin la méthode du Pivot de Gauß. Dans un second temps, nous allons voir comment les matrices peuvent nous aider à étudier les applications linéaires du **Chapter ALG.3** définies en dimension finie.

1	Matrices & Opérations	2
1.1	Définition	2
1.2	Opérations sur les matrices	3
1.3	Échelonnement et algorithme de GAUß-JORDAN des matrices	11
1.4	Matrices remarquables	20
1.5	Et en Python?	21
2	Représentation matricielle d'objets linéaires	23
2.1	Matrice d'un vecteur, d'une famille	23
2.2	Matrice d'une application linéaire	28

2.3	Opérations endomorphiques & opérations sur les matrices	34
2.4	Changements de base	37
3	Exercices	45
3.1	Calcul matriciel	45
3.2	Représentation matricielle	50
3.3	Changements de base	55

L'Algèbre est généreuse, elle donne souvent plus qu'on ne lui demande.

— J. D'Alembert



Cadre

Dans tout le chapitre, l'ensemble K désignera \mathbb{R} ou \mathbb{C} .

Commençons par introduire une notation importante que nous utiliserons dans le chapitre.



Notation Symbole de KRONECKER

Soient x, y deux éléments d'un ensemble E , alors le *symbole de KRONECKER* de

Σ x, y est défini par :

$$\delta_{x,y} = \begin{cases} 1 & \text{si } x = y, \\ 0 & \text{sinon.} \end{cases}$$

Exemple 1 — Soit $E = \llbracket 1, n \rrbracket$ avec $n \in \mathbf{N}^*$, et $i, j, k \in \llbracket 1, n \rrbracket$.

- ▶ $\delta_{i,j} = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{sinon.} \end{cases}$
- ▶ $1 - \delta_{i,j} = \begin{cases} 1 - 0 & \text{si } i \neq j, \\ 1 - 1 & \text{sinon.} \end{cases} = \begin{cases} 1 & \text{si } i \neq j, \\ 0 & \text{sinon.} \end{cases}$
- ▶ $\delta_{i,j} \times \delta_{j,k} = \begin{cases} 1 & \text{si } i = j = k, \\ 0 & \text{sinon.} \end{cases}$

1. MATRICES & OPÉRATIONS

1.1. Définition

Définition ALG.4.1 | Matrice

On appelle *matrice* $n \times p$ à coefficients dans \mathbf{K} toute famille d'éléments de \mathbf{K} indexée par $\llbracket 1, n \rrbracket \times \llbracket 1, p \rrbracket$. On note une telle matrice sous la forme $A = (a_{i,j})_{\substack{1 \leq i \leq n, \\ 1 \leq j \leq p}}$ ou plus simplement $(a_{i,j})$ et on explicite les coefficients dans un tableau à n lignes et p colonnes :

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,p} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,p} \end{pmatrix}.$$

- ▶ Si $p = 1$, on parle de *vecteur colonne*.
- ▶ Si $n = 1$, on parle de *vecteur ligne*.

Σ **Notation**

- ▶ On note $\mathfrak{M}_{n,p}(\mathbf{K})$ l'ensemble des matrices $n \times p$ à coefficients dans \mathbf{K} .
- ▶ Si $A \in \mathfrak{M}_{n,p}(\mathbf{K})$, on notera $[A]_{i,j}$ son terme général d'ordre (i, j) .
- ▶ Lorsque $n = p$, on note plus simplement $\mathfrak{M}_n(\mathbf{K})$ au lieu de $\mathfrak{M}_{n,n}(\mathbf{K})$.

Σ **Notation Convention**

Si $n = p$, on note en général plus simplement $A = (a_{i,j})_{1 \leq i, j \leq n}$. Dans les notations précédentes, le premier indice désignera toujours le numéro de ligne et le second le numéro de colonne. Le scalaire $a_{i,j}$, noté aussi $A_{i,j}$ ou parfois $[A]_{i,j}$, est appelé « coefficient de A de la ligne i et de la colonne j ».

Σ **Notation Lignes et colonnes d'une matrice**

Si $A \in \mathfrak{M}_{n,p}(\mathbf{K})$, alors on notera en ligne ou en colonnes de la manière suivante :

$$A = \left(\begin{array}{c|ccc|c} C_1(A) & & & & C_p(A) \end{array} \right) = \begin{pmatrix} L_1(A) \\ \vdots \\ L_n(A) \end{pmatrix}.$$

Exemple 2 — Pour $A = \begin{pmatrix} -2 & 0 & 1 \\ 3 & 2 & 1 \end{pmatrix}$, on a :

$$C_1(A) = \begin{pmatrix} -2 \\ 3 \end{pmatrix}, \quad C_2(A) = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \quad C_3(A) = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

$$L_1(A) = \begin{pmatrix} -2 & 0 & 1 \end{pmatrix}, \quad L_2(A) = \begin{pmatrix} 3 & 2 & 1 \end{pmatrix}.$$

Exemple 3 — $(2^i + 3j)_{\substack{1 \leq i \leq 3 \\ 1 \leq j \leq 2}} = \begin{pmatrix} 2+3 & 2+6 \\ 4+3 & 4+6 \\ 8+3 & 8+6 \end{pmatrix} = \begin{pmatrix} 5 & 8 \\ 7 & 10 \\ 11 & 14 \end{pmatrix} \in \mathfrak{M}_{3,2}(\mathbf{R})$.

Exemple 4 — On considère la matrice $A = \begin{pmatrix} 2 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \cdots & 1 & 2 \end{pmatrix}$.

Écrire le coefficient (i, j) de la matrice A en fonction de $\delta_{i,j}$ pour tout $(i, j) \in \llbracket 1, n \rrbracket^2$.



On définit le *produit des matrices* A et B comme étant la matrice $n \times p$ suivante :

$$A \times B = \begin{pmatrix} L_1(A) \times C_1(B) & L_1(A) \times C_2(B) & \cdots & L_1(A) \times C_p(B) \\ L_2(A) \times C_1(B) & L_2(A) \times C_2(B) & \cdots & L_2(A) \times C_p(B) \\ \vdots & \vdots & \ddots & \vdots \\ L_n(A) \times C_1(B) & L_n(A) \times C_2(B) & \cdots & L_n(A) \times C_p(B) \end{pmatrix}.$$

Autrement dit, le terme général de $A \times B$ est donné par :

$$\forall (i, j) \in \llbracket 1, n \rrbracket \times \llbracket 1, p \rrbracket, \quad [A \times B]_{i,j} = L_i \times C_j = \sum_{k=1}^q [A]_{i,k} [B]_{k,j}.$$

Si pour tout $(i, j) \in \llbracket 1, n \rrbracket \times \llbracket 1, p \rrbracket$, on note $c_{i,j}$ le coefficient (i, j) de AB, alors le produit matriciel peut être illustré de la façon suivante.

1.2. Opérations sur les matrices

On considère $n, p, q \in \mathbb{N}^*$ trois entiers naturels non nuls. Commençons par définir quelques opérations sur les matrices.

Définition ALG.4.2 | Somme matricielle

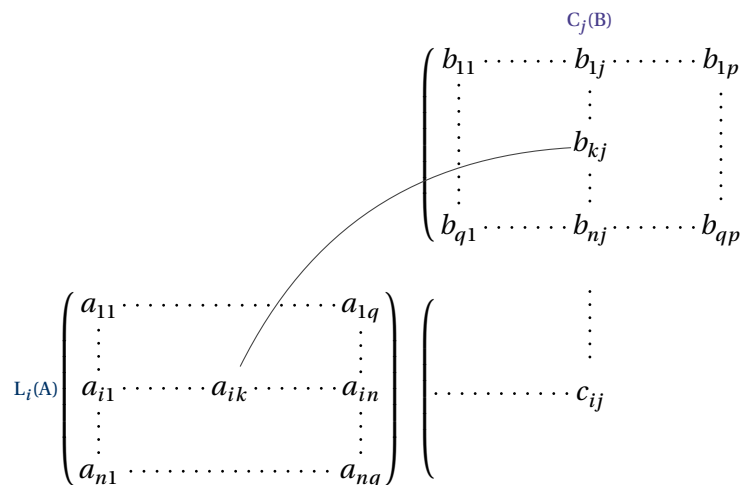
Soient $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \in \mathfrak{M}_{n,n}(\mathbf{K})$ et $B = (b_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \in \mathfrak{M}_{n,n}(\mathbf{K})$. On note $A + B \in \mathfrak{M}_{n,n}(\mathbf{K})$ la matrice définie par :

$$(A + B)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} = (a_{i,j} + b_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}.$$

Définition ALG.4.3 | Produit matriciel

Soient $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq q}} \in \mathfrak{M}_{n,q}(\mathbf{K})$ et $B = (b_{i,j})_{\substack{1 \leq i \leq q \\ 1 \leq j \leq p}} \in \mathfrak{M}_{q,p}(\mathbf{K})$. On nomme les lignes de A et les colonnes de B :

$$A = \left(\begin{array}{c|c|c} L_1(A) & & \\ \vdots & & \\ L_n(A) & & \end{array} \right) \quad \text{et} \quad B = \left(\begin{array}{c|c|c} C_1(B) & \dots & C_p(B) \end{array} \right).$$



Attention Non commutativité du produit matriciel

Le produit matriciel n'est pas commutatif (sauf pour $n = 1$). Calculer par exemple



$A \times B$ et $B \times A$ pour

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Remarque 1.1 — On remarque que le nombre de colonnes de A doit obligatoirement être égal au nombre de lignes de B . On pourra retenir le schéma suivant type « relation de CHASLES » pour connaître le format de la matrice produit :

$$\boxed{n \times q} \quad \times \quad \boxed{q \times p}$$

On peut également multiplier une matrice par un scalaire, et ainsi définir une opération externe sur $\mathfrak{M}_{n,p}(\mathbf{K})$.

Définition ALG.4.4 | Multiplication par un scalaire d'une matrice

Soit $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} \in \mathfrak{M}_{n,p}(\mathbf{K})$ et $\lambda \in \mathbf{K}$. Alors la matrice λA est définie par :

$$[\lambda.A]_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} = \lambda a_{i,j}.$$

STRUCTURE D'ESPACE VECTORIEL ET DIMENSION DE $\mathfrak{M}_{n,p}(\mathbf{K})$. Le produit matriciel permet donc de définir une opération dite interne¹ :

$$\left| \begin{array}{l} \mathfrak{M}_{n,q}(\mathbf{K}) \times \mathfrak{M}_{q,p}(\mathbf{K}) \longrightarrow \mathfrak{M}_{n,p}(\mathbf{K}) \\ (A, B) \longrightarrow A \times B \end{array} \right.$$

La **Définition ALG.4.4** définit en revanche une opération externe telle qu'introduite dans le **Chapter ALG.3**, et donc $\mathfrak{M}_{n,p}(\mathbf{K})$ deviendra *potentiellement* un espace vectoriel. C'est ce que confirme la proposition suivante.

¹elle prend en argument deux matrices, et retourne une autre matrice

Proposition ALG.4.1

Le triplet $(\mathfrak{M}_{n,p}(\mathbf{K}), +, \cdot)$ est un \mathbf{K} -espace vectoriel, où :

- ▶ $+$ désigne l'addition des matrices,
- ▶ \cdot la multiplication externe par un élément de \mathbf{K} .

L'élément neutre pour $+$ est la *matrice nulle*, i.e. la matrice dont tous les coefficients sont nuls.

Comme il nous est impossible de voir $(\mathfrak{M}_{n,p}(\mathbf{K}), +, \cdot)$ comme un sous-espace vectoriel d'un espace vectoriel de référence, il s'agit ici de vérifier la définition d'un espace vectoriel.

Preuve Vérification de toutes les propriétés définissant un espace vectoriel.

On constate notamment que l'élément opposé d'une matrice $M \in \mathfrak{M}_{n,p}(\mathbf{K}) =$

$$\left(m_{i,j} \right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} \text{ est la matrice } -M = \left(-m_{i,j} \right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}.$$

Comment définir une base canonique sur $(\mathfrak{M}_{n,p}(\mathbf{K}), +, \cdot)$? Pour \mathbf{K}^n , nous avons pris des vecteurs qui possèdent un seul 1 en une position donnée, et des 0 ailleurs. Il est naturel d'en faire de-même pour les matrices, c'est ce que l'on appelle les *matrices élémentaires*, et qui sont définies dans le paragraphe suivant.

MATRICES USUELLES. Pour terminer, définissons quelques matrices usuelles. La terminologie associée aux deux premières n'est pas anodine, il y a un lien avec les applications linéaires identiques et les homothéties définies dans le **Chapter ALG.3**, ce lien sera explicité plus tard dans ce chapitre.

Définition ALG.4.5 | Matrice identité

On appelle *matrice identité* de $\mathfrak{M}_{n,n}(\mathbf{K})$ la matrice $n \times n$ suivante :

$$I_n = \left(\delta_{i,j} \right)_{1 \leq i,j \leq n} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}.$$

Définition ALG.4.6 | Matrice homothétique

On appelle *matrice homothétique* de $\mathcal{M}_{n,n}(\mathbf{K})$ de rapport $\lambda \in \mathbf{K}$ la matrice $n \times n$ suivante :

$$\lambda I_n = (\lambda \delta_{i,j})_{1 \leq i,j \leq n} = \begin{pmatrix} \lambda & 0 & \dots & 0 \\ 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda \end{pmatrix}.$$

Définition ALG.4.7 | Matrices élémentaires (ou base canonique)

Pour tout $k \in \llbracket 1, n \rrbracket$, et $\ell \in \llbracket 1, p \rrbracket$, on appelle *matrice élémentaire d'indice* (k, ℓ) , notée $E_{k,\ell}$, la matrice de $\mathcal{M}_{n,p}(\mathbf{K})$ constituée de zéros partout sauf pour le coefficient en ligne k et colonne ℓ , qui vaut un.

Remarque 1.2 — Réécriture avec le symbole de KRONECKER Autrement dit,

$$E_{k,\ell} = (\delta_{i,k} \delta_{j,\ell})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$$

En effet, tous les coefficients sont nuls, sauf si $\delta_{i,k}$ et $\delta_{j,\ell}$ valent 1, c'est-à-dire si le coefficient considéré est sur la ligne k et la colonne ℓ .

Exemple 5 — pour $n = 2, p = 3$ Dans $\mathcal{M}_{2,3}(\mathbf{R})$, on a :

$$E_{1,1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad E_{1,2} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad E_{1,3} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$E_{2,1} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad E_{2,2} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad E_{2,3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Dans $\mathcal{M}_{2,3}(\mathbf{R})$, on constate qu'il y a six matrices élémentaires, et donc on **conjecture** que $\dim \mathcal{M}_{2,3}(\mathbf{R}) = 6$.

Définition/Proposition ALG.4.1 | Dimension de $\mathcal{M}_{n,p}(\mathbf{K})$

L'espace vectoriel $\mathcal{M}_{n,p}(\mathbf{K})$ est un \mathbf{K} -espace vectoriel de dimension

$$n \times p$$

dont la famille $(E_{k,\ell})_{(k,\ell) \in \llbracket 1, n \rrbracket \times \llbracket 1, p \rrbracket}$ est une base, appelée *base canonique* de $\mathcal{M}_{n,p}(\mathbf{K})$.

Preuve On vérifie sans difficulté qu'elle est libre et génératrice.

Il nous reste d'autres opérations sur les matrices à définir et à analyser : la transposition et l'inversion.

PUISSANCES D'UNE MATRICE. Le calcul des puissances itérées d'une matrice est généralement difficile, nous voyons ici quelques techniques pour y parvenir.

Définition ALG.4.8 | Puissance p -ième

Soit $A \in \mathcal{M}_{n,n}(\mathbf{K})$ avec $n \geq 1$ et $p \geq 0$. On définit par récurrence la matrice A^p pour tout $p \in \mathbf{N}$ comme étant :

$$\begin{cases} A^0 = I_n, \\ \forall p \in \mathbf{N}, \quad A^{p+1} = A \times A^p = A^p \times A. \end{cases}$$

2

Proposition ALG.4.2 | Binôme de NEWTON pour les matrices

Soit $(A, B) \in \mathcal{M}_{n,n}(\mathbf{K})^2$ tel que $AB = BA$. Alors on a la formule du *binôme de NEWTON* :

$$\forall p \in \mathbf{N}, \quad (A + B)^p = \sum_{k=0}^p \binom{p}{k} A^k B^{p-k}.$$

Preuve Voir cours de première année.

Exemple 6 — Soit $A = \begin{pmatrix} a & b & 0 \\ 0 & a & b \\ 0 & 0 & a \end{pmatrix} \in \mathcal{M}_{3,3}(\mathbf{R})$. Calculer A^n pour tout $n \in \mathbf{N}$.



²De manière plus explicite, il s'agit d'un produit de p matrices, toutes égales à A

Proposition ALG.4.3 | Formule de BERNOULLI [H.P]

Soit $(A, B) \in \mathfrak{M}_{n,n}(\mathbf{K})^2$ tel que $AB = BA$.

$$\forall p \in \mathbf{N}^*, \quad A^p - B^p = (A - B) \left(\sum_{k=0}^{p-1} A^k B^{p-1-k} \right).$$

En particulier,

$$\forall p \in \mathbf{N}^*, \quad I_n - A^p = (I_n - A) \times \sum_{k=0}^{p-1} A^k.$$

En conséquence

$$A \text{ nilpotente} \implies I_n - A \text{ inversible.}$$

Remarque 1.3 — Dans les deux dernières égalités, les factorisations par $A - B$ et $I_n - A$ respectivement peuvent se faire à droite.

Preuve (Point clef — *Télescopage*)

**Définition ALG.4.9 | Matrice nilpotente**

Soit $A \in \mathfrak{M}_{n,n}(\mathbf{K})$. Alors A est dite *nilpotente* s'il existe $k \in \mathbf{N}$ tel que $A^k = 0$.

Seule la preuve de la prochaine proposition est à connaître.

Exemple 7 — Considérons la matrice $B = \begin{pmatrix} 1 & -a & 0 & 0 \\ 0 & 1 & -a & 0 \\ 0 & 0 & 1 & -a \\ 0 & 0 & 0 & 1 \end{pmatrix} \in \mathfrak{M}_4(\mathbf{K})$. On peut dé-

composer $B = I_4 - aJ$ où $J = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. Alors la matrice B est inversible et on peut

donner son inverse. 

Résumons les techniques précédentes sous forme d'une méthode.

Méthode Binôme et calculs des puissances

Si on arrive à écrire une matrice comme somme d'une matrice D diagonale et d'une matrice nilpotente N (*i.e.* telle que $N^p = 0$ pour un certain $p \in \mathbf{N}$), qui **commutent**, on utilise la formule du binôme matricielle :

$$(D + N)^p = \sum_{k=0}^p \binom{p}{k} D^{p-k} N^k.$$

TRANSPOSITION MATRICIELLE. L'opération de transposition est une opération qui réalise une «symétrie d'axe $i = j$ » dans les coefficients de la matrice.

Définition ALG.4.10 | Transposée

Soit $A = (a_{i,j}) \in \mathfrak{M}_{n,p}(\mathbf{K})$ une matrice. On appelle *transposée* de A la matrice de $\mathfrak{M}_{p,n}(\mathbf{K})$, notée ${}^T A$, telle que pour tout $i \in \llbracket 1, p \rrbracket$ et tout $j \in \llbracket 1, n \rrbracket$

$$[{}^T A]_{i,j} = a_{j,i},$$

i.e. le coefficient $(i,j) \in \llbracket 1, p \rrbracket \times \llbracket 1, n \rrbracket$ de la matrice ${}^T A$ est le coefficient (j,i) de A .

Attention

Parfois certains sujets et ouvrages notent la transposition d'une matrice à droite (*i.e.* A^T au lieu de ${}^T A$ dans ce cours).

Exemple 8 — On a : ${}^T \begin{pmatrix} -2 & 0 & 1 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 3 \\ 0 & 2 \\ 1 & 1 \end{pmatrix}$.

Proposition ALG.4.4

- L'application $\mathcal{T}(\cdot) \left| \begin{array}{l} \mathfrak{M}_{n,p}(\mathbf{K}) \longrightarrow \mathfrak{M}_{p,n}(\mathbf{K}) \\ \mathbf{A} \longmapsto \mathbf{A}^T \end{array} \right.$ est un *isomorphisme de \mathbf{K} -espaces vectoriels*, i.e. une application linéaire bijective.
- **(Transposée d'un produit)** Si A et B sont deux matrices $n \times n$:

$$\mathcal{T}(A \times B) = \mathcal{T}B \mathcal{T}A.$$

Preuve

En nota


2. On se place dans le cas $n = 2$ et $\mathbf{K} = \mathbf{R}$. On considère $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ avec $\lambda_1, \lambda_2 \in \mathbf{R}$.


On a l'équivalence :

$$\lambda_1, \lambda_2 \in \mathbf{R}^+ \iff \forall X \in \mathfrak{M}_{2,1}(\mathbf{R}), \quad q(X) \stackrel{\text{(défi.)}}{=} \mathcal{T}XDX \geq 0.$$

**Exemple 9** — « *Forme quadratique* » associée à une matrice diagonale réelle

1. Soit $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathfrak{M}_{n,1}(\mathbf{K})$ avec $n \geq 1$ et $x_i \in \mathbf{K}$ pour tout $i \in \llbracket 1, n \rrbracket$. Quel est le format

de $\mathcal{T}X.X$? de $X.\mathcal{T}X$? Exprimer le coefficient général de chacune des matrices. 


3. Établir la même équivalence, mais en remplaçant D par $M = P \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} P^{-1}$. 

Dans ce cas, B est aussi inversible, on l'appelle la *matrice inverse de A*, et on la note $B = A^{-1}$.

Σ Notation

On note $GL_n(\mathbf{K})$ l'ensemble des matrices inversibles, appelé *groupe linéaire de* $\mathfrak{M}_n(\mathbf{K})$.

Remarque 1.4 — En fait on peut même montrer — mais ce n'est pas du tout trivial! — que l'existence d'un inverse à droite, *i.e.* $B \in \mathfrak{M}_n(\mathbf{K})$ telle que $A \times B = I_n$ suffit pour garantir l'inversibilité.

Exemple 10 — Les matrices homothétiques sont-elles inversibles? 

Les éléments précédents s'étendent naturellement à des matrices de taille $n \times n$.

INVERSION MATRICIELLE ET DÉTERMINANT EN DIMENSION DEUX.

Définition ALG.4.11 | Matrice inversible & Groupe linéaire

Une matrice $A \in \mathfrak{M}_n(\mathbf{K})$ est dite *inversible* s'il existe une matrice $B \in \mathfrak{M}_n(\mathbf{K})$ telle que :

$$A \times B = B \times A = I_n.$$

⊗ Attention

L'ensemble $GL_n(\mathbf{K})$ des matrices inversibles n'est pas un espace vectoriel. Si $n = 2$. On peut considérer par exemple $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ et $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$, ces deux matrices sont inversibles mais leur somme ne l'est pas, car égale à la matrice nulle qui n'est pas inversible.

Proposition ALG.4.5 | Inversion d'un produit et d'une transposée

▶ **(Inverse d'un produit)** Si A et B sont deux matrices $n \times n$ inversibles, alors $A \times B$ est inversible et :

$$(A \times B)^{-1} = B^{-1} \times A^{-1}.$$

▶ **(Transposition et inversion commutent)** Soit $A \in GL_n(\mathbf{K})$ une matrice inversible. Alors ${}^T A$ est inversible aussi et l'on a :


$$({}^T A)^{-1} = {}^T (A^{-1}).$$

Preuve (Point clef — Vérifier la définition d'une matrice inverse)



$$AB = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \times \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ad-bc & 0 \\ 0 & ad-bc \end{pmatrix} = \det(A)I_2$$

- ▶ Supposons $\det(A) \neq 0$, alors $A \times \left(\frac{1}{\det(A)}B\right) = I_2$. A est donc inversible et $A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$
- ▶ Supposons que $\det(A) = 0$. Si A était inversible, alors $AB = 0_{2,2}$ entraînerait $A^{-1}AB = 0_{2,2}$, c'est-à-dire $B = 0_{2,2}$ ce qui est clairement absurde. Ainsi A n'est pas inversible.

Exemple 11 — Soit $A = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$ et $\lambda \in \mathbf{K}$. Quand est-ce que la matrice $A - \lambda I_2$ est inversible? 

CAS PARTICULIER DE LA DIMENSION DEUX.

Définition/Proposition ALG.4.2 | Inversibilité d'une matrice 2×2 & Déterminant

Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathfrak{M}_{2,2}(\mathbf{K})$. Alors :


$$A \text{ est inversible} \iff ad - bc \neq 0.$$

En cas d'inversibilité, nous avons :

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

On appelle *déterminant* de A noté $\det A$ la quantité $\det A = ad - bc$.

Preuve Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ et $B = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. Calculons AB.


Exemple 12 — Soit $A \in \mathfrak{M}_{2,2}(\mathbf{C})$ une matrice inversible. Exprimer $\det(A^{-1})$ en fonction de $\det A$. 

CALCUL D'UN INVERSE À L'AIDE D'UN POLYNÔME ANNULATEUR. Nous allons revoir la classique méthode du «miroir» pour le calcul de l'inverse dans le prochain paragraphe, mais d'abord, voyons-en une basée sur l'existence d'une relation polynomiale en la matrice.

Exemple 13 — Soit $M = \begin{pmatrix} 2 & -2 & 1 \\ 2 & -3 & 2 \\ -1 & 2 & 0 \end{pmatrix}$.

- ▶ La matrice M vérifie la relation $M^2 + 2M - 3I_3 = 0_3$. 

- ▶ On déduit alors que M est inversible et on peut calculer son inverse.

Indication : On montrera que $M^{-1} = \begin{pmatrix} \frac{4}{3} & \frac{-2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{-1}{3} & \frac{2}{3} \\ \frac{-1}{3} & \frac{2}{3} & \frac{2}{3} \end{pmatrix}$. 



Méthode Inverse matriciel à l'aide d'un polynôme annulateur

Supposons qu'il existe $a_0, \dots, a_p \in \mathbf{K}$ tel que $a_0 \neq 0$, et soit $A \in \mathfrak{M}_{n,n}(\mathbf{K})$ une matrice carrée vérifiant

$$a_0 I_n + a_1 A + \dots + a_p A^p = 0_n. \quad (1)$$

On dit que $P = a_0 + a_1 X + \dots + a_p X^p$ est *annulateur de A*. Alors (1) est équivalente à

$$a_1 A + \dots + a_p A^p = -a_0 I_n,$$

puis étant donné que a_0 est non nul,

$$A \left(-\frac{a_1}{a_0} I_n + \dots - \frac{a_p}{a_0} A^{p-1} \right) = I_n.$$

La matrice A est alors inversible d'inverse $-\frac{a_1}{a_0} I_n + \dots - \frac{a_p}{a_0} A^{p-1}$.

1.3.

Échelonnement et algorithme de GAUß-JORDAN des matrices

1.3.1.

Principe

Nous reprenons rapidement l'algorithme de GAUß-JORDAN d'intérêt majeur dans :

1. la recherche du rang d'une matrice (et par conséquent l'analyse de l'inversibilité d'une matrice carrée, en comparant le rang à son format),

2. la résolution de systèmes linéaires.

OPÉRATIONS ÉLÉMENTAIRES AUTORISÉES SUR LES LIGNES. Cet algorithme fait intervenir trois types d'opération sur les lignes que nous définissons dès à présent.

Définition ALG.4.12 | Opérations élémentaires

Soit $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbf{K})$ une matrice écrite en lignes $A = \begin{pmatrix} L_1(A) \\ \vdots \\ L_n(A) \end{pmatrix}$. On appelle

opération élémentaire sur les lignes de A une des opérations suivantes :

- ▶ **(Permutation des lignes i, j)**

$$L_i(A) \leftrightarrow L_j(A),$$

- ▶ **(Dilatation de paramètre $\lambda \in \mathbf{K} \setminus \{0\}$)**

$$L_i(A) \leftarrow \lambda L_i(A),$$

- ▶ **(Transvection de paramètre $\mu \in \mathbf{K}$)**

$$L_i(A) \leftarrow L_i(A) + \mu L_j(A)$$

avec $i \neq j$.

- ▶ **(Combinaison dilatation/transvection)** si $\lambda \in \mathbf{K} \setminus \{0\}$ et $\mu \in \mathbf{K}$,

$$L_i(A) \leftarrow \lambda L_i(A) + \mu L_j(A)$$

avec $i \neq j$.

De-même, on définit les *opération élémentaire sur les colonnes de A* de manière analogue.

Attention

Retenez que si l'on autorisait $\lambda = 0$ dans l'un des cas encadrés :

- ▶ pour les dilatations, alors on supprimerait une ligne, aïe...
- ▶ Pour une combinaison dilatation/transvection, on «écraserait» une ligne



par une autre, re-aïe...

Peu de chances que ces deux opérations servent à quelque chose dans la suite.

Remarque 1.5 — Lien entre la matrice de départ et la matrice modifiée via une des opérations Chacune des ces opérations se traduit matriciellement par la multiplication **à gauche** par une matrice inversible. Plus précisément, elles correspondent à la multiplication à gauche par les matrices suivantes (portant le même nom) :

- ▶ on appelle *matrice de permutation* de $\mathcal{M}_{n,n}(\mathbf{K})$ toute matrice de la forme $P_{i,j} = I_n + E_{i,j} + E_{j,i} - E_{i,i} - E_{j,j} \in \mathcal{M}_{n,n}(\mathbf{K})$ où $1 \leq i < j \leq n$.
- ▶ On appelle *matrice de dilatation* de $\mathcal{M}_{n,n}(\mathbf{K})$ toute matrice de la forme $D_i(\lambda) = I_n + (\lambda - 1)E_{i,i} \in \mathcal{M}_{n,n}(\mathbf{K})$ où $\lambda \in \mathbf{K} \setminus \{0\}$, $i \in \llbracket 1, n \rrbracket$.
- ▶ On appelle *matrice de transvection* de $\mathcal{M}_{n,n}(\mathbf{K})$ toute matrice de la forme $T_{i,j}(\lambda) = I_n + \lambda E_{i,j}$ où $\lambda \in \mathbf{K}$, $(i, j) \in \llbracket 1, n \rrbracket^2$ et $i \neq j$.

Si l'on multiplie **à droite** par ces mêmes matrices, il est possible de constater que l'on réalise les mêmes opérations, mais sur les colonnes cette fois-ci.

Définition ALG.4.13 | Équivalence en ligne/colonne

Soient $A, B = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbf{K})$ deux matrices. On dit que A, B sont

1. *équivalentes en lignes* et l'on note $A \underset{L}{\sim} B$, si A peut être transformée en B via une suite d'opérations élémentaires sur les lignes,
2. *équivalentes en colonnes* et l'on note $A \underset{C}{\sim} B$, si A peut être transformée en B via une suite d'opérations élémentaires sur les colonnes.

Remarque 1.6 — D'après la remarque précédente, l'équivalence en ligne de deux matrices A, B signifie aussi qu'il existe une matrice C vérifiant : $CA = B$, où C est un produit de matrices de dilatation, transvections et permutations. La matrice C est par conséquent inversible (car toutes les matrices élémentaires le sont et qu'un produit de matrices inversibles est inversible).

Cette remarque est d'intérêt fondamentale, car, puisqu'on ne change pas le rang en multipliant à droite/gauche par une matrice inversible, alors : $\text{Rg}A = \text{Rg}B$ — le rang est ainsi préservé après une suite d'opérations élémentaires. Cela légitime l'algorithme de GAUß-JORDAN que nous allons détailler ci-après.

ALGORITHME DU PIVOT DE GAUß-JORDAN : TOUTE MATRICE EST ÉQUIVALENTE EN LIGNE À UNE MATRICE ÉCHELONNÉE EN LIGNES.

Définition ALG.4.14 | Échelonnement en ligne

Une matrice est dite *échelonnée par lignes* si :

- ▶ si une ligne est nulle, alors toutes les lignes suivantes sont nulles,
- ▶ à partir de la deuxième ligne, dans chaque ligne non nulle, le premier coefficient non nul à partir de la gauche est situé à droite strictement du premier coefficient non nul de la ligne précédente.

Dans ce cas, on appelle *pivot* le premier coefficient non nul de chaque ligne.

Voici la forme typique d'une matrice échelonnée, les pivots étant indiqués par le symbole \oplus .

$$\begin{pmatrix} \oplus & * & * & * & * & * & * & * & * \\ 0 & 0 & \oplus & * & * & * & * & * & * \\ 0 & 0 & 0 & \oplus & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \oplus & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \oplus \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Exemple 14 — Les matrices suivantes sont-elles échelonnées par ligne? Déterminer les pivots, le cas échéant.

$$A = \begin{pmatrix} 3 & 4 & 0 & 5 & 6 \\ 0 & 0 & -8 & 10 & 1 \\ 0 & 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 3 & 4 & 0 & 5 & 6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 1 & 1 \\ 0 & 0 & 5 & 6 & 7 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 5 & 6 & 7 \\ 0 & 8 & 0 & 9 \end{pmatrix}$$



Définition ALG.4.15 | Échelonnement réduit en ligne

Une matrice est dite *échelonnée réduite par lignes* si :

- ▶ elle est nulle,
- ▶ ou elle est échelonnée par lignes, a tous ses pivots égaux à 1, et les pivots sont les seuls coefficients non nuls de leur colonne.

Exemple 15 — Par exemple, la matrice A ci-dessus n'est pas échelonnée réduite par lignes, mais la matrice suivante l'est :

$$D = \begin{pmatrix} \boxed{1} & 2 & 0 & 0 & 3 \\ 0 & 0 & \boxed{1} & 0 & 4 \\ 0 & 0 & 0 & \boxed{1} & 5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Méthode Passer d'une forme échelonnée en ligne à une forme échelonnée réduite en ligne



On élimine de droite à gauche, colonne par colonne, tous les coefficients au-dessus de la diagonale à l'aide de transvections, donc d'opérations de la forme $L_i \leftarrow L_i + \mu L_j$ avec $j > i$ ³ et $\mu \in \mathbf{R}$.

Exemple 16 — Déterminer la réduite échelonnée de $E = \begin{pmatrix} 1 & -1 & 0 & -1 & -1 \\ 0 & 1 & -1 & 2 & 3 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}$. Encadrer les pivots au préalable.



³Afin de ne pas ajouter des termes non nuls dans les colonnes à gauche du pivot de la ligne i

le change en 1 par l'opération $L_1 \leftarrow \frac{1}{\alpha} L_1$.

2. (**Élimination**) On annule les autres coefficients a_{i,j_1} de la colonne j_1 par des transvections, i.e. des opérations de la forme $L_i \leftarrow L_i - a_{i,j_1} L_1$. On donc

$$A \underset{L}{\sim} \begin{pmatrix} 0 \dots 0 & \boxed{1} & * & \dots & * \\ 0 & \dots & 0 & & \\ \vdots & & \vdots & & \\ 0 & \dots & 0 & & \end{pmatrix} \quad \begin{matrix} \\ \\ \\ B \end{matrix}$$

Où $\boxed{1}$ est donc en colonne j_1 .

3. On recommence avec le bloc B s'il n'est pas déjà nul; dans la suite on effectue que des opérations sur L_2, \dots, L_n . Et ainsi la partie précédente n'est pas modifiée.

La preuve précédente est en fait un algorithme à connaître : il est à savoir reproduire parfaitement sur des exemples.

Remarque 1.7 — Dans l'algorithme, on avance. On retiendra bien que dès que l'on a obtenu un pivot égal à un dans une colonne et que l'on a annulé les autres coefficients de cette colonne, alors les opérations élémentaires suivantes ne pourront plus modifier ni cette colonne, ni les précédentes.

Définition/Proposition ALG.4.3 | Rang

On appelle *rang d'une matrice* A le nombre de pivots non nuls d'une des deux formes échelonnées.

Exemple 17 — Soit $A = \begin{pmatrix} 1 & \sqrt{2} & 0 & 2\sqrt{2} \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. Déterminer $\text{Rg}(A)$, encadrer les pivots.



Théorème ALG.4.1 | Algorithme de GAUß-JORDAN

Toute matrice est équivalente par lignes à :

1. une matrice échelonnée en ligne,
2. une matrice échelonnée réduite par lignes.

Les deux formes ont le même nombre de pivots non nuls.

Preuve (Point clef — Si un coefficient est non nul sur une colonne, on peut éliminer tous les autres sur la même colonne à l'aide d'opérations élémentaires)

1. (**Placement correct d'un pivot**) Si la matrice A est nulle, alors il n'y a rien à faire et l'algorithme est terminé. Sinon, soit j_1 l'indice de la première colonne non nulle. On choisit un pivot (c'est-à-dire un coefficient α non nul) dans cette colonne, on le place en première ligne par un échange de lignes $L_1 \leftrightarrow L_i$, et on

Proposition ALG.4.6 | Lien entre rang et inversibilité

Soit $A \in \mathcal{M}_{n,n}(\mathbf{K})$. Alors :

$$A \text{ est inversible} \iff \text{Rg}A = n.$$

Preuve Constatons qu'être de rang n pour une matrice carrée signifie précisément que le nombre de 1 dans la forme échelonnée-réduite est n , donc finalement que la forme échelonnée-réduite est I_n , et que réaliser une série d'opérations sur les lignes revient à multiplier à gauche la matrice de départ par une matrice inversible.

$$\begin{aligned} \text{Rg}A = n &\iff A \underset{L}{\sim} I_n, \\ &\iff A = E \times I_n = E, E \text{ inversible}, \\ &\iff A \text{ inversible}. \end{aligned}$$

Remarque 1.8 — On retiendra surtout la méthode pour obtenir la réduite d'une matrice A , connue sous le nom d'*algorithme de GAUß-JORDAN*.

Exemple 18 — Étudier l'inversibilité de $A = \begin{pmatrix} 2 & 7 & 3 \\ 3 & 9 & 4 \\ 1 & 5 & 3 \end{pmatrix}$. On montrera que $A \underset{L}{\sim}$

$$\begin{pmatrix} 2 & 7 & 3 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Exemple 19 — Calculer le rang de $A = \begin{pmatrix} 0 & 0 & 2 & 1 & 4 \\ 0 & 2 & -1 & 0 & 7 \\ 0 & 4 & 2 & 2 & 6 \\ 0 & 0 & 6 & 3 & 1 \end{pmatrix}$. On montrera que sa réduite

échelonnée est $A \underset{L}{\sim} \begin{pmatrix} 0 & \boxed{1} & 0 & \frac{1}{4} & 0 \\ 0 & 0 & \boxed{1} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$. En commençant par échanger L_1 et L_3 ,

on a :

$$\begin{array}{l}
 A \sim_L \begin{pmatrix} 0 & 4 & 2 & 2 & 6 \\ 0 & 2 & -1 & 0 & 7 \\ 0 & 0 & 2 & 1 & 4 \\ 0 & 0 & 6 & 3 & 1 \end{pmatrix} \\
 \sim_L \begin{pmatrix} 0 & 4 & 9 & 2 & 6 \\ 0 & 0 & -4 & -2 & 8 \\ 0 & 0 & 2 & 1 & 4 \\ 0 & 0 & 0 & 0 & -11 \end{pmatrix} \\
 \sim_L \begin{pmatrix} 0 & 4 & 2 & 2 & 6 \\ 0 & 0 & -4 & -2 & 8 \\ 0 & 0 & 0 & 0 & 16 \\ 0 & 0 & 0 & 0 & -11 \end{pmatrix} \\
 \sim_L \begin{pmatrix} 0 & 2 & 1 & 1 & 3 \\ 0 & 0 & 2 & 1 & -4 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}
 \end{array}$$

$L_2 \leftarrow 2L_2 - L_1$
 $L_4 \leftarrow L_4 - 3L_3$
 $L_3 \leftarrow 2L_3 + L_2$
 $L_1 \leftarrow \frac{1}{2}L_1$
 $L_2 \leftarrow -\frac{1}{2}L_2$
 $L_3 \leftarrow \frac{1}{16}L_3,$
 $L_4 \leftarrow 16L_4 + 11L_3$

En faisant la chaîne d'opérations $L_1 \leftarrow L_1 - 3L_3, L_2 \leftarrow L_2 + 4L_3, L_1 \leftarrow L_1 - \frac{1}{2}L_2$, on trouve la forme échelonnée réduite.

1.3.2. Deux applications

Aux deux applications présentées *infra*, on peut aussi ajouter la résolution de systèmes linéaires, cf. votre cours de 1^{ère} Année.

MÉTHODE DU « MIROIR » POUR CALCULER L'INVERSE D'UNE MATRICE. On en déduit l'inversion d'une matrice inversible par la *méthode du pivot de Gauss*. Soit $A \in GL_n(\mathbf{K})$ une matrice donnée, dont on cherche la matrice inverse. On peut multiplier à gauche les deux membres de l'égalité

$$A = I_n A$$

par des matrices élémentaires notées E_k où k est dans un certain ensemble d'entiers, afin d'effectuer des **opérations élémentaires simultanées** sur les lignes de A (dans le membre de gauche) et celles de I_n (membre de droite).

Méthode du miroir

Supposons que l'on arrive, après avoir effectué certaines opérations élémentaires sur les lignes de A (*i.e.* après avoir multiplié par $E = \dots E_3 E_2 E_1$, un produit de matrices d'opérations élémentaires, à gauche la matrice A), à transformer A en I_n . Alors on obtient l'égalité matricielle :

$$\begin{array}{c} EA \\ I_n \end{array} = \begin{array}{c} EI_n A \\ A^{-1} \end{array}, \quad \text{qui fournit ainsi l'inverse de } A,$$

la matrice inverse sera alors $EI_n = E$.

Remarque 1.9 — La méthode fonctionne toujours Si A est inversible, alors A possède trois pivots non nuls. Ainsi, la forme échelonnée réduite de A est nécessairement l'identité. Ce constat légitime donc la méthode : on arrivera nécessairement à l'identité par une suite d'opérations élémentaires sur les lignes.

Remarque 1.10 — Analogue sur les colonnes En partant de $A = AI_n$ puis en effectuant que des opérations sur les colonnes, on peut aussi déduire l'inverse de A .

Attention

Pas de mélanges entre les deux types d'opérations!

En effet, si vous faites une série d'opérations sur les lignes (qui correspondent à la multiplication à gauche par une matrice inversible E), et sur les colonnes (qui correspondent à la multiplication à droite par une matrice inversible F), alors nous avons la modification suivante :

$$A = I_n \times A \quad \rightarrow \quad (EAF) = (EI_n F)A.$$


Cette égalité n'a bien sûr aucune raison d'être vraie, puisque F, A ne commutent pas forcément. Vous pouvez aussi essayer en partant de $A = A \times I_n$, cela ne fonctionne pas non plus!

Exemple 20 — Calculons l'inverse de $A = \begin{pmatrix} 1 & 1 & 3 \\ 2 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$.

$$\begin{aligned} & \begin{pmatrix} \boxed{1} & 1 & 3 \\ 2 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot A \\ \Rightarrow & \begin{pmatrix} 1 & 1 & 3 \\ 0 & -1 & -6 \\ 0 & 0 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \cdot A & \begin{cases} L_2 \leftarrow L_2 - 2L_1 \\ L_3 \leftarrow L_3 - L_1 \end{cases} \\ \Rightarrow & \begin{pmatrix} 1 & 1 & 3 \\ 0 & \boxed{1} & 6 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ 1/2 & 0 & -1/2 \end{pmatrix} \cdot A & \begin{cases} L_2 \leftarrow -L_2 \\ L_3 \leftarrow -L_3/2 \end{cases} \\ \Rightarrow & \begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & 6 \\ 0 & 0 & \boxed{1} \end{pmatrix} = \begin{pmatrix} -1 & 1 & 0 \\ 2 & -1 & 0 \\ 1/2 & 0 & -1/2 \end{pmatrix} \cdot A & \begin{cases} L_1 \leftarrow L_1 - L_2 \end{cases} \\ \Rightarrow & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1/2 & 1 & -3/2 \\ -1 & -1 & 3 \\ 1/2 & 0 & -1/2 \end{pmatrix} \cdot A & \begin{cases} L_1 \leftarrow L_1 + 3L_3 \\ L_2 \leftarrow L_2 - 6L_3 \end{cases} \end{aligned}$$

On a donc $A^{-1} = \begin{pmatrix} 1/2 & 1 & -3/2 \\ -1 & -1 & 3 \\ 1/2 & 0 & -1/2 \end{pmatrix}$.

Exemple 21 — Calculons l'inverse de $A = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 3 & -2 \\ 3 & 5 & 8 \end{pmatrix}$ par la même méthode. On

montrera que $A^{-1} = \begin{pmatrix} -17 & 3 & 5 \\ 7 & -1 & -2 \\ 2 & -\frac{1}{2} & -\frac{1}{2} \end{pmatrix}$ 

$$\begin{aligned} & \begin{pmatrix} 1 & 2 & 2 \\ 1 & 3 & -2 \\ 3 & 5 & 8 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot A \\ \Rightarrow & \begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & -4 \\ 0 & -1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix} \cdot A & \begin{cases} L_2 \leftarrow L_2 - L_1 \\ L_3 \leftarrow L_3 - 3L_1 \end{cases} \\ \Rightarrow & \begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & -4 \\ 0 & 0 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -4 & 1 & 1 \end{pmatrix} \cdot A & \begin{cases} L_3 \leftarrow L_3 + L_2 \end{cases} \\ \Rightarrow & \begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & -\frac{1}{2} & -\frac{1}{2} \end{pmatrix} \cdot A & \begin{cases} L_3 \leftarrow L_3 / (-2) \end{cases} \\ \Rightarrow & \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -3 & 1 & 1 \\ 7 & -1 & -2 \\ 2 & -\frac{1}{2} & -\frac{1}{2} \end{pmatrix} \cdot A & \begin{cases} L_2 \leftarrow L_2 + 4L_3 \\ L_1 \leftarrow L_1 - 2L_3 \end{cases} \\ \Rightarrow & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -17 & 3 & 5 \\ 7 & -1 & -2 \\ 2 & -\frac{1}{2} & -\frac{1}{2} \end{pmatrix} \cdot A & \begin{cases} L_1 \leftarrow L_1 - 2L_2 \end{cases} \end{aligned}$$

RECHERCHE DES ÉLÉMENTS PROPRES D'UNE MATRICE. Pour le moment on ne se préoccupe pas de l'utilité de la définition suivante, elle fera l'objet d'un chapitre ultérieur (le **Chapter ALG.5**). On suppose dans cette partie que $n = p$, i.e. que la matrice A est carrée.

Définition ALG.4.16 | Éléments propres d'une matrice carrée

Soit $A = (a_{i,j}) \in \mathcal{M}_{n,n}(\mathbf{K})$ une matrice.

- On appelle *valeur propre* de A tout scalaire $\lambda \in \mathbf{K}$ tel que :
 $A - \lambda I_n$ ne soit pas inversible.

On note $\text{Spec}(A)$ l'ensemble des valeurs propres — éventuellement complexes.

- ▶ Soit $\lambda \in \text{Spec}A$. Alors on appelle *espace propre* associé à la valeur propre λ , l'ensemble des solutions (en $X \in \mathfrak{M}_{n,1}(\mathbf{K})$) du système $(A - \lambda I_n)X = 0$. On note $E_\lambda(A)$ l'ensemble des vecteurs propres associés à la valeur propre $\lambda \in \text{Spec}A$.

Déterminer les éléments propres d'une matrice consiste à déterminer les valeurs propres et les vecteurs propres de ladite matrice.

Proposition ALG.4.7 | Caractérisation

Soit $A = (a_{i,j}) \in \mathfrak{M}_{n,n}(\mathbf{K})$ une matrice. Alors :

$$\lambda \in \text{Spec}(A) \iff \text{Rg}(A - \lambda I_n) < n.$$

Preuve On a, pour toute matrice $B \in \mathfrak{M}_{n,n}(\mathbf{K})$,

$$B \text{ inversible} \iff \text{Rg}(B) = n.$$

Donc pour $\lambda \in \mathbf{K}$,

$$A - \lambda I_n \text{ non inversible} \iff \text{Rg}(A - \lambda I_n) < n.$$

Il s'agit donc d'appliquer l'algorithme du pivot de Gauß à la matrice $A - \lambda I_n$. Du fait de la présence d'un paramètre λ , il est nécessaire de prendre garde aux opérations réalisées, je vous conseille les suivantes.

Méthode Opérations pour la recherche d'éléments propres

- ▶ Si le coefficient (3, 1) n'est pas nul :

1. l'opération optimale à effectuer en premier pour des matrices de taille 3×3 est la permutation $L_1 \longleftrightarrow L_3$ qui permettra d'obtenir un coefficient indépendant de λ en position (1, 1). On élimine alors avec celui-ci les coefficients (2, 1), (3, 1).

2. Ensuite, en position (2, 2), nous avons un coefficient affine en λ et que l'on souhaite utiliser en nouveau pivot afin d'éliminer le coefficient (3, 2). Pour éliminer λ en (2, 2), on peut faire une opération simple en fonction


de L_3 .

3. Un pivot indépendant de λ est alors obtenu en (2, 2), on peut alors éliminer le coefficient (3, 2).

- ▶ Si le coefficient (3, 1) est nul :

1. on fait la permutation $L_1 \longleftrightarrow L_2$ qui permettra d'obtenir un coefficient indépendant de λ en position (1, 1). On élimine alors avec celui-ci le coefficient (3, 1).
2. En positions (2, 2), nous avons un coefficient indépendant de λ qui sert à éliminer le coefficient (3, 2).

Exemple 22 — Déterminer les éléments propres des matrices ci-dessous.

1. $A = \begin{pmatrix} -2 & -2 & 1 \\ -2 & 1 & -2 \\ 1 & -2 & -2 \end{pmatrix}$. On montrera que $\text{Spec}A = \{-3, 3\}$. 

2. $B = \begin{pmatrix} 11 & -5 & 5 \\ -5 & 3 & -3 \\ 5 & -3 & 3 \end{pmatrix}$. On montrera que $\text{Spec } B = \{0, 1, 16\}$. Soit $\lambda \in \mathbf{K}$, on cherche λ de sorte que $A - \lambda I_3$ soit non inversible.

$$\begin{array}{l}
 A - \lambda I_3 = \begin{pmatrix} 11 - \lambda & -5 & 5 \\ -5 & 3 - \lambda & -3 \\ 5 & -3 & 3 - \lambda \end{pmatrix} \\
 \sim_{\mathbf{L}} \begin{pmatrix} 5 & -3 & 3 - \lambda \\ -5 & 3 - \lambda & -3 \\ 11 - \lambda & -5 & 5 \end{pmatrix} \quad \left. \begin{array}{l} \\ \\ \end{array} \right\} \begin{array}{l} L_1 \leftrightarrow L_3 \\ \\ \end{array} \\
 \sim_{\mathbf{L}} \begin{pmatrix} 5 & -3 & 3 - \lambda \\ 0 & -\lambda & -\lambda \\ 0 & -3\lambda + 8 & -\lambda^2 + 14\lambda - 8 \end{pmatrix} \quad \left. \begin{array}{l} \\ \\ \end{array} \right\} \begin{array}{l} L_2 \leftarrow L_2 + L_1, \\ L_3 \leftarrow 5L_3 - (11 - \lambda)L_1 \\ \\ \end{array} \\
 \sim_{\mathbf{L}} \begin{pmatrix} 5 & -3 & 3 - \lambda \\ 0 & -8 & \lambda^2 - 17\lambda + 8 \\ 0 & -3\lambda + 8 & -\lambda^2 + 14\lambda - 8 \end{pmatrix} \quad \left. \begin{array}{l} \\ \\ \end{array} \right\} \begin{array}{l} L_2 \leftarrow 3L_2 - L_3 \\ \\ \end{array} \\
 \sim_{\mathbf{L}} \begin{pmatrix} 5 & -3 & 3 - \lambda \\ 0 & -8 & \lambda^2 - 17\lambda + 8 \\ 0 & 0 & P(\lambda) \end{pmatrix} \quad \left. \begin{array}{l} \\ \\ \end{array} \right\} \begin{array}{l} \\ \\ L_3 \leftarrow -8L_3 - (-3\lambda + 8)L_2 \end{array}
 \end{array}$$

avec $P(\lambda) = -8(-\lambda^2 + 14\lambda - 8) - (-3\lambda + 8)(\lambda^2 - 17\lambda + 8)$. C'est un polynôme de degré trois, de coefficient dominant 3, de racines évidentes 0, 1, et il existe λ_0 tel que

$$P = 3X(X - 1)(X - \lambda_0),$$

on obtient en évaluant par exemple en $X = -1$:

$$-2(\lambda_0 + 1) = -34,$$

soit $\lambda_0 = 16$. On conclut alors que $\text{Spec } B = \{0, 1, 16\}$. Et

► $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_0(B)$ si et seulement si :

$$\begin{cases} 5x - 3y + 3z = 0 \\ -8y + 8z = 0 \end{cases} \iff X = \begin{pmatrix} 0 \\ y \\ y \end{pmatrix} \in \text{Vect} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix},$$

donc $E_1(B) = \text{Vect} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$

► $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_1(B)$ si et seulement si :

$$\begin{cases} 5x - 3y + 2z = 0 \\ -8y - 8z = 0 \end{cases} \iff \begin{cases} x = -z \\ y = -z \end{cases} \iff X = \begin{pmatrix} -z \\ -z \\ 1 \end{pmatrix} \in \text{Vect} \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix},$$

donc $E_1(B) = \text{Vect} \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}.$

► $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_{16}(B)$ si et seulement si :

$$\begin{cases} 5x - 3y - 13z = 0 \\ -8y - 8z = 0 \end{cases} \iff X = \begin{pmatrix} 2z \\ -z \\ z \end{pmatrix} \in \text{Vect} \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix},$$

donc $E_{16}(B) = \text{Vect} \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}.$

Nous verrons dans le **Chapter ALG.5** que la connaissance des valeurs propres va nous permettre, dans certains cas, de déterminer A^n . Autre méthode, qui sera présentée dans ce chapitre, est l'utilisation de la formule du binôme de NEWTON matriciel.

1.4. Matrices remarquables

Définition ALG.4.17 | Homothétie


On appelle *matrice d'homothétie* de $\mathfrak{M}_{n,n}(\mathbb{K})$ toute matrice de la forme

$$\begin{pmatrix} \lambda & 0 & \dots & 0 \\ 0 & \lambda & & \\ \vdots & & \ddots & \\ 0 & \dots & 0 & \lambda \end{pmatrix} = \lambda I_n.$$

Proposition ALG.4.8 | Structure des homothéties

L'ensemble des matrices d'homothétie de $\mathfrak{M}_{n,n}(\mathbb{K})$ est un sous-espace vectoriel de $\mathfrak{M}_{n,n}(\mathbb{K})$ de dimension 1.

Preuve

► C'est un sous-espace vectoriel de $\mathfrak{M}_{n,n}(\mathbb{K})$. 

► Il est de dimension 1, puisque c'est $\text{Vect}(I_n)$ donc (I_n) en est une base.

Définition ALG.4.18 | Matrice diagonale


On appelle *matrice diagonale* de $\mathfrak{M}_{n,n}(\mathbb{K})$ toute matrice de la forme


$$\begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & & \\ \vdots & & \ddots & \\ 0 & \dots & 0 & a_n \end{pmatrix}.$$

Proposition ALG.4.9 | Structure des matrices diagonales

L'ensemble des matrices diagonales de $\mathfrak{M}_{n,n}(\mathbf{K})$ est un sous-espace vectoriel de $\mathfrak{M}_{n,n}(\mathbf{K})$ de dimension n .

Preuve

- ▶ C'est un sous-espace vectoriel de $\mathfrak{M}_{n,n}(\mathbf{K})$. 

- ▶ Il est de dimension n , puisque $(E_{1,1}, \dots, E_{n,n})$ en est une base. 

Définition ALG.4.19 | Matrice triangulaire

On appelle *matrice triangulaire inférieure* (resp. *supérieure*) de $\mathfrak{M}_{n,n}(\mathbf{K})$ toute matrice $A = (a_{i,j})$ telle que

$$\forall (i,j) \in \llbracket 1, n \rrbracket^2, \quad i < j \implies a_{i,j} = 0 \quad (\text{resp. } i > j \implies a_{i,j} = 0),$$

i.e. telle que :

$$A = \begin{pmatrix} \star & 0 & \dots & 0 \\ \star & \star & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \star & \dots & \star & \star \end{pmatrix}, \quad \text{resp. } A = \begin{pmatrix} \star & \star & \dots & \star \\ 0 & \star & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \star \end{pmatrix}.$$

Notation

On note $\text{TS}_n(\mathbf{K})$ (resp. $\text{TI}_n(\mathbf{K})$) l'ensemble des matrices triangulaires supérieures (resp. inférieures).

Remarque 1.11 —

- ▶ A est triangulaire supérieure si et seulement si ${}^T A$ est triangulaire inférieure. Et donc la transposition est un isomorphisme⁴ entre l'espace vectoriel des matrices triangulaires supérieures et l'espace vectoriel des triangulaires inférieures.
- ▶ L'ensemble des matrices triangulaires supérieures et inférieures sont des espaces vectoriels de dimension $\frac{n(n+1)}{2}$. Nous le montrerons en TD.

1.5. Et en Python ?

Un TP sera consacré partiellement aux manipulations de matrices en Python, et aux principales fonctions existantes. Nous faisons une brève synthèse ci-dessous, observez les résultats.

Rappels sur les manipulations de matrices en Python

```
>>> import numpy as np
>>>
>>> # Création
>>> A = np.array([[1,2,3], [4,5,6]])
>>> B = np.array([[2,3,4], [5,6,7]])
>>> # dans certains sujets le module matrix subsiste, il est
- désormais obsolète et ne devrait plus être utilisé dans les
- sujets à venir
>>> type(A)
<class 'numpy.ndarray'>
>>> #on accède aux éléments comme on le ferait avec des listes de
- listes classiques
>>> A[1][2]
```

⁴déjà vu

```

6
>>> #Mais on peut aussi utiliser une notation plus « mathématique
↳ » et se passer des doubles crochets
>>> A[1, 2]
6
>>> # Récupérer le format d'une matrice
>>> n,p = A.shape
>>> #Extraction : fonctionnement identique aux listes. Les :
↳ signifient que l'on impose aucune condition sur la position où
↳ il est placé (ici, cela signifie « peu importe la ligne »,
↳ donc on renvoie une colonne)
>>> A[:, 2]
array([3, 6])
>>> #Somme
>>> C = A + B
>>> C
array([[ 3,  5,  7],
       [ 9, 11, 13]])
>>> #Transposition
>>> C.transpose()
array([[ 3,  9],
       [ 5, 11],
       [ 7, 13]])
>>>
>>> #Produit
>>> np.dot(A, B.transpose())
array([[20, 38],
       [47, 92]])
>>>
>>> # Ou encore
>>> A@(B.transpose())
array([[20, 38],
       [47, 92]])

```

```

>>> #le symbole @ est d'utilisation plus aisée que la fonction
↳ np.dot
>>>
>>> #Matrices usuelles
>>> np.zeros((1, 4))
array([[0., 0., 0., 0.]])
>>> np.ones((2, 1)) #Attention : des tuples sont requis pour ces
↳ deux fonctions
array([[1.],
       [1.]])
>>> np.eye(3, 3)
array([[1., 0., 0.],
       [0., 1., 0.],
       [0., 0., 1.]])
>>> #Mais cela fonctionne aussi :
>>> np.eye(3)
array([[1., 0., 0.],
       [0., 1., 0.],
       [0., 0., 1.]])
>>> # Éléments propres
>>> T = np.array([[1, 1], [0, 2]])
>>> np.linalg.eig(T) #produit un tuple dont la première coordonnée
↳ est la liste de valeurs propres, la seconde la une liste de
↳ vecteurs propres
(array([1., 2.]), array([[1.          , 0.70710678],
                        [0.          , 0.70710678]]))
>>> np.linalg.eig(T)[0] # liste des valeurs propres
array([1., 2.])

```

Remarque 1.12 — La notion de « vecteur propre » sera vue dans le [Chapter ALG.5](#).

2. REPRÉSENTATION MATRICIELLE D'OBJETS LINÉAIRES

Les matrices ont pour le moment été étudiées comme des tableaux de nombres (entiers ou réels par exemple, voir complexes). Nous allons voir en quoi elles peuvent être utiles pour étudier les vecteurs, les applications linéaires *etc.*

Nous allons découvrir plusieurs choses importantes :

1. pour connaître une application linéaire $u : E \rightarrow F$ où E, F sont deux \mathbf{K} -espace vectoriels de dimension finie (*i.e.* connaître les $u(x)$ pour tout $x \in E$), il suffit de se donner une matrice à $\dim F$ lignes et $\dim E$ colonnes, donc à $\dim F \times \dim E$ coefficients.
2. Les opérations énoncées précédemment sur les matrices comme l'addition, la multiplication, l'inversion, correspondront à des matrices d'autres applications linéaires comme l'addition, la composition et l'inversion d'applications linéaires — que le monde est bien fait! 😊⁵

Regardons deux premiers exemples.

- Supposons que $f \in \mathcal{L}(\mathbf{K}^n, \mathbf{K})$, f est donc une forme linéaire sur \mathbf{K}^n et soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de \mathbf{K}^n . Alors nous avons déjà vu (dans le [Chapter ALG.3](#)) qu'il suffit de connaître les vecteurs (ici ce sont des scalaires car f est à valeurs dans \mathbf{K})

$$f(e_1), \dots, f(e_n)$$

pour entièrement déterminer f . Cette famille de réels ou complexes peut donc se mettre sous forme d'un tableau :

$$(f(e_1) \ f(e_2) \ \dots \ f(e_n)) \in \mathcal{M}_{1,n}(\mathbf{K}) \underset{\text{(nota.)}}{=} \underset{\mathcal{B}}{\mathcal{M}at}(f),$$

i.e. une matrice à une ligne et n colonnes.

⁵Il en manque une à l'appel : la transposition ! Elle aussi, elle possède son analogue « endomorphique ». On l'appelle l'endomorphisme *transposé*, mais là on dépasse largement le cadre du programme.

- Si plus généralement $f \in \mathcal{L}(\mathbf{K}^n, \mathbf{K}^p)$ avec $p \geq 1$ les $f(e_i)$ sont alors des vecteurs de \mathbf{K}^p . Si une base $\mathcal{F} = (f_1, \dots, f_p)$ de \mathbf{K}^p est fixée, chaque $f(e_i)$ possèdent p coordonnées (*cf.* la encore le [Chapter ALG.3](#)) que l'on peut écrire en colonne, comme ci-dessous :

$$\begin{pmatrix} e_i^1 \\ e_i^2 \\ \vdots \\ e_i^p \end{pmatrix}, \text{ pour rappel, cela signifie que : } f(e_i) = \sum_{j=1}^p e_j^i f_j.$$

Faisant cela pour tout $i \in \llbracket 1, n \rrbracket$, nous obtenons une matrice à n colonnes et p lignes, dont les coefficients e_i^j pour tous $(i, j) \in \llbracket 1, p \rrbracket \times \llbracket 1, n \rrbracket$ déterminent entièrement f .

Nous allons ainsi successivement caractériser tous les objets linéaires par des tableaux de scalaires (des matrices donc) : les vecteurs, les familles de vecteurs et les applications linéaires. On constate dans les deux exemples que les notions de matrices dépendent de choix de bases, ce fait apparaîtra dans les prochaines définitions.

2.1. Matrice d'un vecteur, d'une famille

Définition ALG.4.20 | Matrice d'un vecteur

Soit E un \mathbf{K} -espace vectoriel de dimension finie n , muni d'une base $\mathcal{B} = (e_1, \dots, e_n)$. Soit x un vecteur de E de coordonnées x_1, \dots, x_n dans la base \mathcal{B} , i.e. :

$$x = \sum_{k=1}^n x_k e_k.$$

On appelle *matrice du vecteur x relativement à la base \mathcal{B}* la matrice **colonne** suivante :


$$\mathcal{M}_{\mathcal{B}}(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathfrak{M}_{n,1}(\mathbf{K}).$$

Lorsque $\mathcal{B} = \mathcal{B}^{\text{can}}$ est la base canonique de $E = \mathbf{K}^n$ ou $\mathbf{K}_n[X]$, on dit que $\mathcal{M}_{\mathcal{B}}(x)$ est la *matrice canoniquement associée à x* .

De manière générale, toutes les notions de matrice qui suivront dépendront du choix d'une base. Nous aurons des moyens simples de passer d'une base à l'autre dans un second temps. Par unicité des coordonnées d'un vecteur dans une base, l'objet précédent est bien défini.

Exemple 23 —

1. Considérons $x = (1, 2, 3) \in \mathbf{R}^3$, et munissons \mathbf{R}^3 de deux bases : la base canonique \mathcal{B}^{can} et la base $\mathcal{C} = (a, b, c)$, où $a = (1, 1, 1)$, $b = (0, 1, 1)$ et $c = (0, 1, -1)$. Alors, on

a : $\mathcal{M}_{\mathcal{B}}(x) = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$. Que vaut $\mathcal{M}_{\mathcal{C}}(x)$? 

2. Considérons $P = 1 + X + X^2 \in \mathbf{R}_2[X]$, et munissons $\mathbf{R}_2[X]$ de deux bases : la base canonique \mathcal{B}^{can} et la base $\mathcal{C} = (1, X - 1, (X - 1)^2)$. Alors, on a : $\mathcal{M}_{\mathcal{B}}(P) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

Que vaut $\mathcal{M}_{\mathcal{C}}(P)$? 

Et dans le sens inverse? Peut-on voir toute colonne comme la matrice d'un certain vecteur dans une certaine base? La réponse est oui.

Proposition ALG.4.10 | La matrice dans une base caractérise le vecteur

Soit E un \mathbf{K} -espace vectoriel de dimension n , muni d'une base \mathcal{B} . L'application

$$\varphi : \begin{cases} E & \longrightarrow \mathfrak{M}_{n,1}(\mathbf{K}) \\ x & \longmapsto \underset{\mathcal{B}}{\mathcal{M}at}(x) \end{cases}$$

est un isomorphisme de \mathbf{K} -espaces vectoriels. En particulier,

- ▶ étant donnée une base \mathcal{B} fixée, il y a une correspondance bijective entre les colonnes de $\mathfrak{M}_{n,1}(\mathbf{K})$ et les vecteurs de E .
- ▶ Pour tout $(x, y) \in E^2, (\lambda, \mu) \in \mathbf{K}^2$, on a :

$$\underset{\mathcal{B}}{\mathcal{M}at}(\lambda x + \mu y) = \lambda \underset{\mathcal{B}}{\mathcal{M}at}(x) + \mu \underset{\mathcal{B}}{\mathcal{M}at}(y).$$


On peut alors considérer la définition ci-après.

Définition ALG.4.21 | Vecteur canoniquement associé à une colonne


Soit $E = \mathbf{K}^n$ ou $\mathbf{K}_{n-1}[X]$ munis l'une ou l'autre de leur base canonique \mathcal{B}^{can} , et $X = \mathfrak{M}_{n,1}(\mathbf{K})$. Alors on appelle *vecteur canoniquement associé* à X l'unique $x \in E$ tel que

$$\underset{\mathcal{B}^{can}}{\mathcal{M}at}(x) = X.$$

Preuve

- ▶ On admet que φ est linéaire, la preuve ne présente aucune difficulté.
- ▶ Passons au caractère isomorphe. Puisque $\dim E = \dim \mathfrak{M}_{n,1}(\mathbf{K}) = n$, il suffit d'établir l'injectivité en calculant le noyau. 

Remarque 2.1 — Surjectivité «à la main» Soit $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathfrak{M}_{n,1}(\mathbf{K})$. Posons 

Exemple 24 — Un cas spécial Soient $n \geq 1$ et $E = \mathbf{K}^n$. Que dire de l'application $\begin{cases} \mathbf{K}^n & \longrightarrow \mathfrak{M}_{n,1}(\mathbf{K}) \\ (x_1, \dots, x_n) & \longmapsto \underset{\mathcal{B}^{can}}{\mathcal{M}at}((x_1, \dots, x_n)) \end{cases} ?$ 

Définition ALG.4.22 | Matrice d'une famille


Soit E un \mathbf{K} -espace vectoriel de dimension finie n , muni d'une base $\mathcal{B} = (e_1, \dots, e_n)$. Soit une famille $\mathcal{F} = (x_1, \dots, x_p)$ de p vecteurs de E ⁶. On appelle *matrice de la famille \mathcal{F} relativement à la base \mathcal{B}* la matrice

$$\mathcal{M}_{\mathcal{B}}(\mathcal{F}) = \left(\begin{array}{c|c|c} \mathcal{M}_{\mathcal{B}}(x_1) & \dots & \mathcal{M}_{\mathcal{B}}(x_p) \end{array} \right) \in \mathfrak{M}_{n,p}(\mathbf{K}).$$

Autrement dit, c'est la matrice où la i -ème colonne est la matrice du i -ème vecteur de la famille.


Exemple 25 – avec des uplets Reprenons les données de l'exemple précédent. Alors :


$$\mathcal{M}_{\mathcal{B}}(\mathcal{B}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{et} \quad \mathcal{M}_{\mathcal{B}}(\mathcal{C}) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & -1 \end{pmatrix}.$$

Calculer $\mathcal{M}_{\mathcal{C}}(\mathcal{C})$ et $\mathcal{M}_{\mathcal{C}}(\mathcal{B})$. 

⁶Si $p = 1$, $\mathcal{F} = (x_1)$, on retombe sur la définition de matrice d'un vecteur vue précédemment.

Exemple 26 – avec des fonctions On note $F = \text{Vect}(\cos, \sin)$ dans $E = \mathbf{R}^{\mathbf{R}}$. On considère $f : x \mapsto \cos(x + 1)$ et $g : x \mapsto \sin(x - 1)$. Montrer que $f, g \in F$ et déterminer

$$\mathcal{M}_{(\cos, \sin)}((f, g)).$$
 

Exemple 27 – avec des polynômes Nous avons déjà montré que $\mathcal{F} = (X(X-1), X(X-2), (X-1)(X-2))$ est une base de $\mathbf{R}_2[X]$. On note $P = X^2 + X + 1$ et $Q = X^2 - 1$. Déterminer $\underset{\mathcal{F}}{\text{Mat}}(P, Q)$. 

Proposition ALG.4.11 | Lien rang d'une matrice de famille / rang de la famille

Soit E un \mathbf{K} -espace vectoriel de dimension finie n , muni d'une base $\mathcal{B} = (e_1, \dots, e_n)$. Soit une famille $\mathcal{F} = (x_1, \dots, x_p)$ de p vecteurs de E. Alors :

$$\underbrace{\text{Rg} \left(\underset{\mathcal{B}}{\text{Mat}}(\mathcal{F}) \right)}_{\text{« nombre de pivots de ... »}} = \underbrace{\text{Rg}(\mathcal{F})}_{\text{« dim du vect de ... »}} \quad ^7.$$

Par conséquent, si $n = p$, \mathcal{F} est une base de E si et seulement si $\underset{\mathcal{B}}{\text{Mat}}(\mathcal{F})$ est inversible.

Preuve Admis.

Méthode Montrer qu'une famille est une base à l'aide de matrice

Soit \mathcal{B} une base de E un espace vectoriel, et \mathcal{F} une famille telle que $\# \mathcal{B} = \# \mathcal{F}$.

- ▶ Calculer $M = \underset{\mathcal{B}}{\text{Mat}}(\mathcal{F})$.
- ▶ Montrer que M est inversible.

2.2. Matrice d'une application linéaire

Nous avons déjà vu en introduction comment voir une application linéaire comme une matrice. Formalisons ceci dans une définition.

Définition ALG.4.23 | Matrice d'une application linéaire

Soient E et F deux espaces vectoriels de dimensions finies, $p = \dim E$, $n = \dim F$, $\mathcal{B} = (e_1, \dots, e_p)$ une base de E et \mathcal{C} une base de F, et $u \in \mathcal{L}(E, F)$.

On appelle *matrice de l'application linéaire u relativement aux bases \mathcal{B} et \mathcal{C}* la

⁷ rang d'une matrice à gauche, rang d'une famille de vecteurs définis dans le [Chapter ALG.3](#) à droite

matrice :

$$\underset{\mathcal{B}, \mathcal{C}}{\text{Mat}}(u) = \underset{\mathcal{C}}{\text{Mat}}(u(\mathcal{B})) \in \mathfrak{M}_{n,p}(\mathbf{K}).$$

Lorsque \mathcal{B} et \mathcal{C} sont les bases canoniques de E et F respectivement (avec E, F qui sont égaux à \mathbf{K}^n ou $\mathbf{K}_n[X]$), on dit que $\underset{\mathcal{B}, \mathcal{C}}{\text{Mat}}(u)$ est la *matrice canoniquement associée à u*.

Si l'on note (e_1, \dots, e_n) une base de E, alors on range dans les colonnes les coordonnées des vecteurs $u(e_i)$, $i \in \llbracket 1, n \rrbracket$ dans une base (f_1, \dots, f_q) de F.

$$A = \begin{pmatrix} \overset{u(e_1)}{\star} & \dots & \overset{u(e_n)}{\star} \\ \vdots & \ddots & \vdots \\ \star & \dots & \star \end{pmatrix} \begin{matrix} f_1 \\ \vdots \\ f_q \end{matrix}$$

Notation Cas d'un endomorphisme

Si $E = F$ et $\mathcal{B} = \mathcal{C}$ alors la matrice de l'endomorphisme u relativement à \mathcal{B} est :

$$\underset{\mathcal{B}}{\text{Mat}}(u) \stackrel{\text{(nota.)}}{=} \underset{\mathcal{B}, \mathcal{B}}{\text{Mat}}(u) \in \mathfrak{M}_{n,n}(\mathbf{K}).$$

Dans ce cas, c'est une matrice carrée de format $n \times n$.

Méthode Calculer une matrice

Pour déterminer $\underset{\mathcal{B}, \mathcal{C}}{\text{Mat}}(u)$, il faut donc :

1. calculer les vecteurs de $u(\mathcal{B})$, i.e. les $u(e_j)$ pour tout $j \in \llbracket 1, p \rrbracket$.
2. Pour tout $i \in \llbracket 1, p \rrbracket$, calculer les coordonnées de $u(e_j)$ dans la base image \mathcal{C} , i.e. chercher les $\lambda_{i,j}$ tels que :

$$u(e_j) = \sum_{i=1}^n \lambda_{i,j} f_i,$$



où (f_1, \dots, f_n) est une base de \mathcal{F} . Ces coordonnées existent bien, et sont uniques, puisque \mathcal{F} est une base de F .


3. Conclure : la j -ème colonne de $\underset{\mathcal{B}, \mathcal{C}}{\text{Mat}}(u)$ sera donc : $\begin{pmatrix} \lambda_{1,j} \\ \vdots \\ \lambda_{n,j} \end{pmatrix}$.

Exemple 28 — Avec des uplets Soit :

$$u : \begin{cases} \mathbf{R}^3 & \longrightarrow & \mathbf{R}^2, \\ (x, y, z) & \longmapsto & (3x + y - z, -x + 2z). \end{cases}$$

Déterminer la matrice canoniquement associée à u .


Exemple 29 — Avec des polynômes Soit $D : \begin{cases} \mathbf{R}_3[X] & \longrightarrow & \mathbf{R}_2[X] \\ \mathcal{P} & \longrightarrow & \mathcal{P}'. \end{cases}$ Déterminer la matrice canoniquement associée à D .

2. Déterminons la matrice de D relativement à \mathcal{B} . 

Exemple 30 — Avec des fonctions Soient $E = \mathcal{C}^\infty(\mathbf{R}, \mathbf{R})$ et $F = \text{Vect}(g, h)$, où g et h sont les applications de \mathbf{R} dans \mathbf{R} définies par: $\forall x \in \mathbf{R}, g(x) = e^x \cos x$ et $h(x) = e^x \sin x$ respectivement.

1. La famille $\mathcal{B} = (g, h)$ est une base de F , et

$$D : \begin{cases} E & \longrightarrow & E \\ f & \longrightarrow & f' \end{cases}$$

définit un endomorphisme de F . 

Exemple 31 — Avec des suites Soient $E = \mathbf{R}^{\mathbf{N}}$, α et β deux réels distincts. On note $u = (\alpha^n) \in E$, $v = (\beta^n) \in E$ et $F = \text{Vect}(u, v)$.

1. La famille $\mathcal{B} = (u, v)$ est une base de F . Nous l'admettons.
2. Soit ensuite \mathcal{C} la base canonique de \mathbf{R}^3 . Déterminons la matrice relativement aux bases \mathcal{B} et \mathcal{C} de :

$$\varphi : \begin{cases} F & \longrightarrow & \mathbf{R}^3, \\ (x_n)_{n \in \mathbf{N}} & \longrightarrow & (x_0 + x_2, x_1, -x_0 + x_3). \end{cases}$$



Exemple 32 – Identité Soit E un \mathbf{K} -espace vectoriel de dimension $n \in \mathbf{N}^*$. Soit \mathcal{B} une base quelconque de E . Alors $\mathcal{M}_{\mathcal{B}}(\text{Id}_E) = I_n$, où :

$$I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} \in \mathfrak{M}_{n,n}(\mathbf{K})$$

(des 1 sur la diagonale et des 0 partout ailleurs). De même, pour tout $\lambda \in \mathbf{K}$, on établit que $\mathcal{M}_{\mathcal{B}}(\lambda \text{Id}_E) = \lambda I_n$.

⊗ Attention


Si l'espace vectoriel ambiant est un espace de


1. polynômes, on ne **met pas** « **de X** » dans les matrices!
2. suites, on ne **met pas** « **de n** » dans les matrices!
3. fonctions, on ne **met pas** « **de x, t, etc.** » dans les matrices!

Par définition d'une matrice, les éléments qui la constitue sont **des scalaires** *i.e.* des éléments de \mathbf{K} .

Traisons à présent le cas d'un formes linéaires, *i.e.* des applications linéaires dont l'espace d'arrivée est \mathbf{K} : leur matrice dans une base donnée est une matrice ligne.

Exemple 33 – Deux formes linéaires

1. Déterminons la matrice de $u : \begin{cases} \mathbf{R}^3 & \longrightarrow \mathbf{R} \\ (x, y, z) & \longmapsto 3x - y + 2z \end{cases}$ relativement à la base canonique \mathcal{B}^{can} de \mathbf{R}^3 après avoir justifié la linéarité. 

2. Déterminons la matrice de $\varphi : \begin{cases} \mathbf{R}_3[X] \longrightarrow \mathbf{R} \\ P \longmapsto P'(0) + P(1) + \int_0^1 P(t) dt \end{cases}$ relativement à la base canonique \mathcal{B}^{can} de $\mathbf{R}_3[X]$ après avoir justifié la linéarité. 

le sens retour est aussi possible. Nous pouvons associer à n'importe quel tableau de nombres et n'importe quel couple de bases fixé une unique application linéaire. On peut résumer cela en «une application linéaire est entièrement déterminée par sa matrice relativement à un couple de bases». Voyons comment.

Théorème ALG.4.2 | La matrice dans deux bases caractérise l'application

Soient E et F des \mathbf{K} -espaces vectoriels de dimensions respectives p et n , \mathcal{B} une base de E , \mathcal{C} une base de F . Alors l'application


$$\varphi : \begin{cases} \mathcal{L}(E, F) \longrightarrow \mathfrak{M}_{n,p}(\mathbf{K}) \\ u \longmapsto \underset{\mathcal{B}, \mathcal{C}}{\text{Mat}}(u) \end{cases}$$

est un isomorphisme de \mathbf{K} -espaces vectoriels. En particulier,

- ▶ $\dim \mathcal{L}(E, F) = n \times p$.
- ▶ Étant données deux bases \mathcal{B} de E et \mathcal{C} de F fixées, il y a une correspondance bijective entre les matrices de $\mathfrak{M}_{n,p}(\mathbf{K})$ et les applications linéaires de E dans F .
- ▶ Pour tout $(u, v) \in \mathcal{L}(E, F)^2$, $(\lambda, \mu) \in \mathbf{K}^2$, on a :

$$\underset{\mathcal{B}}{\text{Mat}}(\lambda u + \mu v) = \lambda \underset{\mathcal{B}}{\text{Mat}}(u) + \mu \underset{\mathcal{B}}{\text{Mat}}(v).$$

Preuve

- ▶ On admet que φ est linéaire, la preuve ne présente aucune difficulté.
- ▶ Passons au caractère isomorphe. Nous allons montrer l'injectivité et la surjectivité. 

Nous avons associé à toute application linéaire une matrice dans deux bases fixées,

► Nous avons $\underset{\mathcal{B}^{\text{can}}}{\text{Mat}}(u) = M$. 

Corollaire ALG.4.1 | Application canoniquement associée (dans $\mathfrak{M}_{n,1}(\mathbf{K})$) à une matrice

Soit $M \in \mathfrak{M}_{n,p}(\mathbf{K})$. Il existe une unique application linéaire $u \in \mathcal{L}(\mathfrak{M}_{p,1}(\mathbf{K}), \mathfrak{M}_{n,1}(\mathbf{K}))$ telle que $\underset{\mathcal{B}^{\text{can}}, \mathcal{B}^{\text{can}}}{\text{Mat}}(u) = M$. On l'appelle *l'application linéaire canoniquement associée* à M . Elle est définie par

$$u \left| \begin{array}{l} \mathfrak{M}_{p,1}(\mathbf{K}) \longrightarrow \mathfrak{M}_{n,1}(\mathbf{K}) \\ X \longrightarrow MX. \end{array} \right.$$


Preuve

► L'application u ainsi définie est linéaire. 

Définition ALG.4.24 | Noyau et image d'une matrice

On appelle *noyau de la matrice* M noté $\ker M$ (*resp. image de la matrice* M noté $\text{Im } M$) le noyau de u (*resp. l'image de* u) où $u \in \mathcal{L}(\mathfrak{M}_{p,1}(\mathbf{K}), \mathfrak{M}_{n,1}(\mathbf{K}))$ est l'application linéaire canoniquement associée à M . Plus explicitement, on définit :

- $\ker M = \{X \in \mathfrak{M}_{p,1}(\mathbf{K}), MX = 0\}$,
- $\text{Im } M = \{Y \in \mathfrak{M}_{n,1}(\mathbf{K}), \exists X \in \mathfrak{M}_{p,1}(\mathbf{K}), MX = Y\}$.

Exemple 34 — Déterminer l'application linéaire u canoniquement associée à $M = \begin{pmatrix} 3 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix}$. Préciser le noyau et l'image de M . 

Théorème ALG.4.3 | Écriture matricielle de $y = u(x)$

Soient E et F des \mathbf{K} -espaces vectoriels de dimensions finies, \mathcal{B} une base de E, \mathcal{C} une base de F, $u \in \mathcal{L}(E, F)$, $x \in E$. Alors :

$$\mathcal{M}_{\mathcal{C}}(u(x)) = \mathcal{M}_{\mathcal{B}, \mathcal{C}}(u) \times \mathcal{M}_{\mathcal{B}}(x).$$

Attention aux bases

La formule est intuitivement claire. Si vous choisissez la base \mathcal{B} comme base de départ de l'application u , alors u ne peut « manger » que des vecteurs x écrits dans \mathcal{B} également. Si \mathcal{F} est la base d'arrivée, alors $u(x)$ est lui aussi écrit dans la base d'arrivée.

Preuve Soient $p = \dim E$ et $n = \dim F$. On explicite les bases de E et F :

$$\mathcal{B} = (e_1, \dots, e_p) \quad \text{et} \quad \mathcal{C} = (f_1, \dots, f_n).$$

Posons :

$$Y = \mathcal{M}_{\mathcal{C}}(y) = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad U = \mathcal{M}_{\mathcal{B}, \mathcal{C}}(u) = \begin{pmatrix} u_{1,1} & \dots & u_{1,p} \\ \vdots & \ddots & \vdots \\ u_{n,1} & \dots & u_{n,p} \end{pmatrix}$$

$$\text{et} \quad X = \mathcal{M}_{\mathcal{B}}(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}.$$

Alors on calcule :

$$\begin{aligned} \sum_{i=1}^n y_i f_i &= y = u(x) = u\left(\sum_{j=1}^p x_j e_j\right) = \sum_{j=1}^p x_j u(e_j) \\ &= \sum_{j=1}^p x_j \sum_{i=1}^n u_{i,j} f_i \\ &= \sum_{i=1}^n \left(\sum_{j=1}^p u_{i,j} x_j\right) f_i. \end{aligned}$$

Par unicité des coordonnées de y dans la base \mathcal{C} , on en déduit, pour tout $i \in \llbracket 1, n \rrbracket$:

$$y_i = \sum_{j=1}^p u_{i,j} x_j. \text{ Ceci se traduit exactement par l'égalité matricielle } Y = U \times X.$$

Proposition ALG.4.12 | Consistance des définitions du rang

Soit $M \in \mathcal{M}_{n,p}(\mathbf{K})$ et u l'application linéaire canoniquement associée. Alors :

$$\text{Rg } u = \text{Rg } M.^8$$

Preuve Admis.

2.3. Opérations endomorphiques & opérations sur les matrices

Nous allons voir le lien entre les opérations matricielles d'une part (somme et produit/inversion), et les opérations fonctionnelles d'autre part (somme et composée/inversion d'applications linéaires).

⁸Pour rappel, $\text{Rg } M$ est le nombre de pivots non nuls dans la forme réduite de M , $\text{Rg } u$ est la dimension de $\text{Im } u$.

Théorème ALG.4.4 | Matrice d'une composée

Soient E, F et G trois \mathbf{K} -espaces vectoriels de dimensions finies, et soient \mathcal{B} une base de E, \mathcal{B}' une base de F, \mathcal{B}'' une base de G, soient $v \in \mathcal{L}(E, F)$ et $u \in \mathcal{L}(F, G)$. Alors :

$$\underset{\mathcal{B}, \mathcal{B}''}{\text{Mat}}(u \circ v) = \underset{\mathcal{B}', \mathcal{B}''}{\text{Mat}}(u) \times \underset{\mathcal{B}, \mathcal{B}'}{\text{Mat}}(v).$$

Lemme ALG.4.1

Soient A et B deux matrices $n \times p$. Alors :

$$A = B \iff (\forall X \in \mathfrak{M}_{p,1}(\mathbf{K}) \quad A \times X = B \times X).$$

Preuve (du lemme) L'implication directe est immédiate. Réciproquement, supposons que $A \times X = B \times X$ pour toute matrice colonne X de taille $p \times 1$. Notons C_j (resp. C'_j) les colonnes de A (resp. B), où $j \in \llbracket 1, p \rrbracket$. On constate alors qu'avec

$$X = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

on a $A \times X = C_1$ et $B \times X = C'_1$. On en déduit que $C_1 = C'_1$. On procède de même pour montrer que les autres colonnes de A et B sont deux à deux égales, d'où $A = B$.

Preuve Pour tout $x \in E$, on a $u \circ v(x) = u(v(x))$. En appliquant le théorème donnant la matrice d'une composée, on obtient l'égalité matricielle :

$$\begin{aligned} \underset{\mathcal{B}, \mathcal{B}''}{\text{Mat}}(u \circ v) \times \underset{\mathcal{B}}{\text{Mat}}(x) &= \underset{\mathcal{B}', \mathcal{B}''}{\text{Mat}}(u) \times \left(\underset{\mathcal{B}, \mathcal{B}'}{\text{Mat}}(v) \times \underset{\mathcal{B}}{\text{Mat}}(x) \right) \\ &= \left(\underset{\mathcal{B}', \mathcal{B}''}{\text{Mat}}(u) \times \underset{\mathcal{B}, \mathcal{B}'}{\text{Mat}}(v) \right) \times \underset{\mathcal{B}}{\text{Mat}}(x) \end{aligned}$$

Or, lorsque x parcourt E, la matrice $\underset{\mathcal{B}}{\text{Mat}}(x)$ parcourt $\mathfrak{M}_{p,1}(\mathbf{K})$. Le lemme permet alors de conclure :

$$\underset{\mathcal{B}, \mathcal{B}''}{\text{Mat}}(u \circ v) = \underset{\mathcal{B}', \mathcal{B}''}{\text{Mat}}(u) \times \underset{\mathcal{B}, \mathcal{B}'}{\text{Mat}}(v).$$

Exemple 35 — avec des uplets Soient

$$u : \begin{cases} \mathbf{R}^3, & \longrightarrow \mathbf{R}^2 \\ (x, y, z) & \longmapsto (x - y, x + y + z), \end{cases} \quad v : \begin{cases} \mathbf{R}^2 & \longrightarrow \mathbf{R}^3, \\ (x, y) & \longmapsto (x + y, x + 2y, x - y). \end{cases}$$


On note \mathcal{B} et \mathcal{C} les bases canoniques respectives de \mathbf{R}^3 et \mathbf{R}^2 . Calculer $\underset{\mathcal{B}, \mathcal{C}}{\text{Mat}}(u)$, $\underset{\mathcal{C}, \mathcal{B}}{\text{Mat}}(v)$, $\underset{\mathcal{B}, \mathcal{B}}{\text{Mat}}(v \circ u)$ et $\underset{\mathcal{C}, \mathcal{C}}{\text{Mat}}(u \circ v)$. Donner l'expression analytique de $u \circ v$ et de $v \circ u$.



Exemple 36 — avec des polynômes Soient

$$u : \begin{array}{l} \mathbf{R}_2[X], \longrightarrow \mathbf{R}_2[X] \\ \mathbf{P} \longrightarrow \mathbf{P}(X+1) \end{array} \quad v : \begin{array}{l} \mathbf{R}_2[X], \longrightarrow \mathbf{R}_2[X] \\ \mathbf{P} \longrightarrow \mathbf{P}(X-1). \end{array}$$

On note \mathcal{B} et \mathcal{C} les bases canoniques respectives de $\mathbf{R}^2[X]$ et \mathbf{R}^2 . Calculer $\underset{\mathcal{B}, \mathcal{C}}{\text{Mat}}(u)$,

$\underset{\mathcal{C}, \mathcal{B}}{\text{Mat}}(v)$, $\underset{\mathcal{B}, \mathcal{B}}{\text{Mat}}(v \circ u)$ et $\underset{\mathcal{C}, \mathcal{C}}{\text{Mat}}(u \circ v)$. Était-ce prévisible? 

On en déduit maintenant facilement que la matrice de u^{-1} , si u est un isomorphisme, est l'inverse de la matrice de u .

Proposition ALG.4.13 | Matrice d'une application inverse

Soient E un \mathbf{K} -espace vectoriel de dimension n , \mathcal{B} une base de E et $u \in \mathcal{L}(E)$ un endomorphisme de E . Alors

$$u \text{ est inversible} \iff \mathcal{M}_{\mathcal{B}}(u) \in \mathfrak{M}_{n,n}(\mathbf{K}) \text{ est inversible,}$$

et dans ce cas :

$$\mathcal{M}_{\mathcal{B}}(u^{-1}) = \left(\mathcal{M}_{\mathcal{B}}(u) \right)^{-1}.$$

Preuve




MATRICES DE PASSAGE.

Définition ALG.4.25

Soit E un \mathbf{K} -espace vectoriel de dimension n , muni de deux bases \mathcal{B} et \mathcal{B}' . On appelle *matrice de passage de \mathcal{B} à \mathcal{B}'* la matrice notée $P^{\mathcal{B} \rightarrow \mathcal{B}'}$ suivante :

$$P^{\mathcal{B} \rightarrow \mathcal{B}'} = \mathcal{M}_{\mathcal{B}'}(\mathcal{B}) = \mathcal{M}_{\mathcal{B}', \mathcal{B}}(\text{Id}_E) \in \mathfrak{M}_{n,n}(\mathbf{K}).$$

Exemple 37 — avec des uplets Soient $E = \mathbf{R}^2$, $\mathcal{B} = (e_1, e_2)$ la base canonique de E et $\mathcal{B}' = (e'_1, e'_2)$, où $e'_1 = (1, 2)$ et $e'_2 = (1, 3)$. Montrer que \mathcal{B}' est une base.

Déterminer les matrices de passage $P^{\mathcal{B} \rightarrow \mathcal{B}'}$ et $P^{\mathcal{B}' \rightarrow \mathcal{B}}$. 

2.4. Changements de base

Le but de cette dernière section est de faire lien entre les différentes matrices (d'applications, de vecteurs, de familles de vecteurs, ...) écrites dans une certaine base \mathcal{B} et celles écrites dans une autre base \mathcal{B}' . Nous allons voir qu'une quantité clef apparaîtra dans les formules : il s'agit de la matrice de passe de \mathcal{B} à \mathcal{B}' ou de \mathcal{B}' à \mathcal{B} .

Proposition ALG.4.14

Soient \mathcal{B} , \mathcal{B}' et \mathcal{B}'' trois bases d'un \mathbf{K} -espace vectoriel E de dimension finie. Alors :

1. **(Composition)** $P^{\mathcal{B} \rightarrow \mathcal{B}'} \cdot P^{\mathcal{B}' \rightarrow \mathcal{B}''} = P^{\mathcal{B} \rightarrow \mathcal{B}''}$,
2. **(Inversion)** $P^{\mathcal{B} \rightarrow \mathcal{B}'}$ est inversible et $(P^{\mathcal{B} \rightarrow \mathcal{B}'})^{-1} = P^{\mathcal{B}' \rightarrow \mathcal{B}}$.

Preuve



Exemple 38 — avec des polynômes Soient $E = \mathbf{R}_2[X]$. On rappelle que $\mathcal{B} = (1, X, X^2)$ est une base de E , et que $\mathcal{B}' = (1, X - 1, X^2 + X)$ en est également une.⁹ Déterminer les matrices de passage $P^{\mathcal{B} \rightarrow \mathcal{B}'}$ et $P^{\mathcal{B}' \rightarrow \mathcal{B}}$.

FORMULES DE CHANGEMENT DE BASE.**Théorème ALG.4.5 | Changement de base pour un vecteur**

Soient E un \mathbf{K} -espace vectoriel de dimension finie, \mathcal{B} et \mathcal{B}' deux bases de E et $x \in E$. Alors :


$$\underset{\mathcal{B}}{\text{Mat}}(x) = P^{\mathcal{B} \rightarrow \mathcal{B}'} \underset{\mathcal{B}'}{\text{Mat}}(x).$$

Preuve Elle tient en deux lignes (tout le travail a été fait dans la formule de composition) :

$$\begin{aligned} \underset{\mathcal{B}}{\text{Mat}}(x) &= \underset{\mathcal{B}}{\text{Mat}}(\text{Id}_E(x)) \\ &= \underset{\mathcal{B}', \mathcal{B}}{\text{Mat}}(\text{Id}_E) \underset{\mathcal{B}'}{\text{Mat}}(x) \\ &= P^{\mathcal{B} \rightarrow \mathcal{B}'} \underset{\mathcal{B}'}{\text{Mat}}(x). \end{aligned}$$

⁹ car échelonnée donc libre, et de cardinal $3 = \dim \mathbf{R}_2[X]$.

Exemple 39 — cas des uplets On reprend les notations de l'Exemple 37. On considère

$x = (4, 1) \in E$. Déterminer $\mathcal{M}_{\mathcal{B}}(x)$ et $\mathcal{M}_{\mathcal{B}'}(x)$. 

Exemple 40 — cas des polynômes On reprend les notations de l'Exemple 38. On considère

$P = 1 - X + X^2 \in E$. Déterminer $\mathcal{M}_{\mathcal{B}}(P)$ et $\mathcal{M}_{\mathcal{B}'}(P)$. 

Exemple 41 — cas d'une rotation de repère Prenons $E = \mathbf{R}^2$ et (e_1, e_2) sa base canonique.

1. On note (e_1^θ, e_2^θ) la base (e_1, e_2) « tournée d'un angle $\theta \in [0, 2\pi[$ » dans le sens trigonométrique. Donnée l'expression de (e_1^θ, e_2^θ) .
2. Si $u \in \mathbf{R}^2$ est tel que $\mathcal{M}_{(e_1, e_2)}(u) = \begin{pmatrix} x \\ y \end{pmatrix}$, déterminer $\mathcal{M}_{(e_1^\theta, e_2^\theta)}(u)$.



Preuve On a

$$\begin{aligned} \underset{\mathcal{B}', \mathcal{C}'}{\mathcal{M}at}(u) &= \underset{\mathcal{B}', \mathcal{C}'}{\mathcal{M}at}(\text{Id}_F \circ u \circ \text{Id}_E) \\ &= \underset{\mathcal{C}', \mathcal{C}'}{\mathcal{M}at}(\text{Id}_F) \underset{\mathcal{B}', \mathcal{C}}{\mathcal{M}at}(u) \underset{\mathcal{B}', \mathcal{B}}{\mathcal{M}at}(\text{Id}_E) \\ &= \underset{\mathcal{C}', \mathcal{C}}{P} \underset{\mathcal{B}', \mathcal{C}}{\mathcal{M}at}(u) \underset{\mathcal{B}', \mathcal{B}}{P}. \end{aligned}$$

Exemple 42 — On reprend les notations de l'Exemple 37, et on introduit $F = \mathbf{R}^3$, $\mathcal{C} = (f_1, f_2, f_3)$ la base canonique de F , $\mathcal{C}' = (f'_1, f'_2, f'_3)$, où $f'_1 = (0, 1, 1)$, $f'_2 = (1, 0, 1)$ et $f'_3 = (1, 1, 0)$ et on note

$$u : \begin{cases} E & \longrightarrow & F, \\ (x, y) & \longmapsto & (x + y, x + 2y, -2x + y). \end{cases}$$

On admet que \mathcal{C}' est une base de F , et que u est linéaire. Déterminer $\underset{\mathcal{B}', \mathcal{C}}{\mathcal{M}at}(u)$ et

$$\underset{\mathcal{B}', \mathcal{C}'}{\mathcal{M}at}(u). \quad \boxed{\text{✎}}$$

Proposition ALG.4.15 | Changement de base pour une application linéaire

Soient E et F deux \mathbf{K} -espaces vectoriels de dimensions finies, $u \in \mathcal{L}(E, F)$, $\mathcal{B}, \mathcal{B}'$ deux bases de E , et $\mathcal{C}, \mathcal{C}'$ deux bases de F . Alors :

$$\begin{aligned} \underset{\mathcal{B}', \mathcal{C}'}{\mathcal{M}at}(u) &= \underset{\mathcal{C}', \mathcal{C}}{P} \underset{\mathcal{B}', \mathcal{C}}{\mathcal{M}at}(u) \underset{\mathcal{B}', \mathcal{B}}{P} \\ &= \underset{\mathcal{C}', \mathcal{C}'}{\mathcal{M}at}(\text{Id}_F) \underset{\mathcal{B}', \mathcal{C}}{\mathcal{M}at}(u) \underset{\mathcal{B}', \mathcal{B}}{\mathcal{M}at}(\text{Id}_E). \end{aligned}$$


Proposition ALG.4.16 | Changement pour un endomorphisme

Soient E un \mathbf{K} -espace vectoriel de dimension finie, $u \in \mathcal{L}(E)$, et \mathcal{B} et \mathcal{B}' deux bases de E . Alors :

$$\begin{aligned} \mathcal{M}_{\mathcal{B}'}(u) &= P^{\mathcal{B}' \rightarrow \mathcal{B}} \cdot \mathcal{M}_{\mathcal{B}}(u) \cdot P^{\mathcal{B} \rightarrow \mathcal{B}'} \\ &= (P^{\mathcal{B} \rightarrow \mathcal{B}'})^{-1} \cdot \mathcal{M}_{\mathcal{B}}(u) \cdot P^{\mathcal{B} \rightarrow \mathcal{B}'} \end{aligned}$$

Exemple 43 — On reprend une dernière fois l'Exemple 37, et on introduit :

$$v : \begin{cases} E & \longrightarrow & E, \\ (x, y) & \longrightarrow & (2x + y, x - y). \end{cases}$$

Déterminer $\mathcal{M}_{\mathcal{B}}(v)$ et $\mathcal{M}_{\mathcal{B}'}(v)$. 

MATRICES SEMBLABLES ET APPLICATION AUX CALCULS DE PUISSANCES.

Définition ALG.4.26 | Matrices semblables

Soient A et B deux matrices de $\mathfrak{M}_{n,n}(\mathbf{K})$. Alors A et B sont dites *semblables* s'il existe une matrice inversible $P \in GL_n(\mathbf{K})$ telle que :

$$A = PBP^{-1}.$$

Si $\mathbf{K} = \mathbf{C}$, deux matrices sont dites *semblables sur \mathbf{R}* s'il existe une matrice inversible $P \in GL_n(\mathbf{R})$ telle que $A = PBP^{-1}$.

Notation

On notera $A \sim B$ lorsque deux matrices sont semblables.

Proposition ALG.4.17 | Calcul de puissances

Soient A et B deux matrices semblables de $\mathfrak{M}_{n,n}(\mathbf{K})$ et $P \in GL_n(\mathbf{K})$ telle que $A = PBP^{-1}$. Alors :

$$\forall k \in \mathbf{N}, \quad A^k = PB^kP^{-1}.$$

Preuve

(Point clef — *Réurrence sur k – preuve très importante*)



Proposition ALG.4.18 | Similitude et expression d'un même endomorphisme dans deux bases différentes

Soient $A, B \in \mathcal{M}_{n,n}(\mathbf{K})$. Alors :

$$A \sim B \iff \begin{cases} \exists E \text{ espace vectoriel de dimension } n, \\ \exists \mathcal{B}, \mathcal{B}' \text{ deux bases de } E, \\ \exists u \in \mathcal{L}(E), A = \underset{\mathcal{B}}{\mathcal{M}at}(u) \quad \text{et} \quad B = \underset{\mathcal{B}'}{\mathcal{M}at}(u). \end{cases}$$

Autrement dit, deux matrices sont semblables si et seulement si elles sont les matrices d'un même endomorphisme, mais exprimé dans deux bases différentes.

Preuve

\Rightarrow Notons P inversible telle que :

$$A = PBP^{-1}.$$

On utilise l'endomorphisme canoniquement associé, *i.e.* on pose

$$E = \mathcal{M}_{n,1}(\mathbf{K}), \quad u : X \in E \mapsto BX, \quad \mathcal{B} \text{ base canonique de } E$$

puis on considère une base \mathcal{B}' telle que

$$\underset{\mathcal{B}, \mathcal{B}'}{\mathcal{M}at}(\text{Id}_E) = P = \left(P^{\mathcal{B}' \rightarrow \mathcal{B}} \right)^{-1}.$$

Alors d'après la formule de changement de base pour u :

$$\underset{\mathcal{B}, \mathcal{B}'}{\mathcal{M}at}(u) = P \underset{\mathcal{B}, \mathcal{B}}{\mathcal{M}at}(u) P^{-1},$$

mais comme u est l'endomorphisme canoniquement associé à B , $\underset{\mathcal{B}, \mathcal{B}}{\mathcal{M}at}(u) = B$ et on a

$$\underset{\mathcal{B}, \mathcal{B}'}{\mathcal{M}at}(u) = PBP^{-1} = A.$$

C'est terminé.

\Leftarrow Conséquence directe de la formule de changement de base, une matrice P qui convient est alors $P = P^{\mathcal{B} \rightarrow \mathcal{B}'}$.

À l'aide de cette proposition, on en déduit la méthode ci-après pour montrer que deux matrices sont semblables.



Méthode Montrer que deux matrices sont semblables

Soient $A, B \in \mathcal{M}_{n,n}(\mathbf{K})$. Alors pour montrer $A \sim B$:

- ▶ **(Si $n = 2$)** on cherche P sous la forme $P = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_{2,2}(\mathbf{R})$ de sorte que $A \cdot P = P \cdot B$. On peut utiliser aussi la méthode ci-après.
- ▶ **(Si $n = 3$)** une seule méthode est à privilégier : celle du changement de base. Notons

$$B = \begin{pmatrix} b_{1,1} & \cdots & b_{1,n} \\ \vdots & & \vdots \\ b_{n,1} & \cdots & b_{n,n} \end{pmatrix}.$$

L'idée est de chercher $P = \left(e_1 \mid \cdots \mid e_n \right)$ inversible, de sorte que

$$A = P \cdot B \cdot P^{-1} = P \cdot \begin{matrix} \xrightarrow{A} e_1 & \cdots & \xrightarrow{A} e_n \\ \begin{pmatrix} b_{1,1} & \cdots & b_{1,n} \\ \vdots & & \vdots \\ b_{n,1} & \cdots & b_{n,n} \end{pmatrix} \end{matrix} \cdot P^{-1}.$$

D'après la formule de changement de base¹⁰, les $e_i, i \in \llbracket 1, n \rrbracket$ doivent vérifier :

$$Ae_i = b_{1,i}e_1 + \cdots + b_{n,i}e_n, \quad \forall i \in \llbracket 1, n \rrbracket.$$



Cette condition est simplement un ensemble de systèmes linéaires à résoudre. Dans la pratique, ces calculs seront rapides à mener puisque la matrice B aura très souvent que quelques coefficients non nuls.

Exemple 44 – Diagonalisation

1. Montrer que les matrices $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ et $D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ sont semblables. Préciser alors une matrice P telle que $A = P \cdot D \cdot P^{-1}$.

▶ (En cherchant avec des coefficients)

▶ (En utilisant un changement de base)

Dans ces deux exemples, nous avons en fait montré que A est semblable à une matrice diagonale, on dit que A est diagonalisable, et nous verrons dans le [Chapter ALG.5](#) comment diagonaliser une matrice, et surtout quand cela est possible.

Exemple 45 – Trigonalisation Montrer que $A = \begin{pmatrix} 0 & 4 & 4 \\ \frac{3}{4} & 0 & 0 \\ 0 & \frac{2}{3} & 0 \end{pmatrix}$ est semblable à la matrice

$T = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$ en utilisant un changement de base.

¹⁰appliquée à $f : X \in \mathfrak{M}_{n,1}(\mathbf{K}) \longrightarrow AX$ l'endomorphisme canoniquement associé à A, $\mathcal{B} = (e_1, \dots, e_n)$, P correspondant alors à $P^{\mathcal{B} \text{ can} \rightarrow \mathcal{B}}$

Dans cet exemple, nous avons en fait montré que A est semblable à une matrice triangulaire supérieure, on dit que A est trigonalisable mais cette notion ne sera pas étudiée dans la suite de l'année.

Remarque 2.2 — Nous verrons dans le [Chapter ALG.5](#) une méthode pour trouver en pratique une matrice D diagonale (si elle existe). Si c'est le cas, on dit que l'on a «diagonalisé» la matrice A *i.e.* trouvé une base dans laquelle la matrice de l'endomorphisme canoniquement associé à A est diagonale.

*** **Fin du chapitre** ***

3. EXERCICES

Exercice ALG.4.1 | Vrai ou Faux ?

1. Le produit de deux matrices triangulaires est triangulaire.
2. L'ensemble des matrices inversibles est un espace vectoriel.
3. Deux matrices semblables ont même rang.
4. Si E, F sont deux espaces vectoriels de dimension finie, alors $\dim \mathcal{L}(E, F) = (\dim F)^{\dim E}$.

3.1. Calcul matriciel

Exercice ALG.4.2 | Soit $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$, et $n \in \mathbf{N}$. Calculer A^n de deux manières différentes.

Solution (exercice ALG.4.2)

► **(1ère Méthode : récurrence)** On montre par récurrence que pour tout $n \in \mathbf{N}$,

$$A^n = \begin{pmatrix} 0 & n & \frac{n(n-1)}{2} \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix}.$$

- **Initialisation.** Elle est vraie pour $n = 1$.
- **Hérédité.** Supposons-là vraie au rang $n \in \mathbf{N}$. Alors

$$A^{n+1} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & n & \frac{n(n-1)}{2} \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & n+1 & \frac{n(n-1)}{2} + n \\ 0 & 1 & n+1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & n+1 & \frac{n(n+1)}{2} \\ 0 & 1 & n+1 \\ 0 & 0 & 1 \end{pmatrix}.$$

La formule est donc vraie pour tout $n \in \mathbf{N}$, d'après le principe de récurrence.

► **(2ème Méthode : binôme)** D'autre part, on peut séparer en une partie nilpotente et une partie diagonale :

$$A = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{=I_3} + \underbrace{\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}}_{:=N}.$$

Comme les matrices I_3 et N commutent, le binôme de NEWTON nous donne pour $n \geq 2$ (puisque $N^3 = 0$) :

$$A^n = I + nN + \frac{n(n-1)}{2}N^2.$$

Donc :

$$\forall n \in \mathbf{N}, \quad A^n = \begin{pmatrix} 0 & n & \frac{n(n-1)}{2} \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix}.$$

Exercice ALG.4.3 | Application matricielle Soit f l'application

$$f \left| \begin{array}{l} \mathbf{R} \longrightarrow \mathfrak{M}_3(\mathbf{R},) \\ x \longrightarrow \begin{pmatrix} 1 & 0 & x \\ -x & 1 & -\frac{x^2}{2} \\ 0 & 0 & 1 \end{pmatrix}. \end{array} \right.$$

On rappelle que $GL_3(\mathbf{R})$ désigne l'ensemble des matrices inversibles de format 3×3 .

1. L'application f est-elle injective ?
2. Soit $(x, y) \in \mathbf{R}^2$. Calculer $f(x)f(y)$ et montrer que $f(x)f(y) \in f(\mathbf{R})$.
3. En déduire :

3.1) $\begin{pmatrix} 1 & 0 & x \\ -x & 1 & -\frac{x^2}{2} \\ 0 & 0 & 1 \end{pmatrix}^n$ pour tout $x \in \mathbf{R}$.

3.2) Que $\begin{pmatrix} 1 & 0 & x \\ -x & 1 & -\frac{x^2}{2} \\ 0 & 0 & 1 \end{pmatrix}$ est inversible pour tout $x \in \mathbf{R}$ et donner son inverse. A-t-on $f(\mathbf{R}) = \text{GL}_3(\mathbf{R})$?

Solution (exercice ALG.4.3)

- Oui. Soient $x, x' \in \mathbf{R}$, tels que $f(x) = f(x')$. Alors $-x = -x'$, $x = x'$ et $-\frac{x^2}{2} = -\frac{x'^2}{2}$, ceci implique bien $x = x'$. Donc f est injective.
- Soient x, y deux réels. Alors un simple produit matriciel prouve que $f(x)f(y) = f(x+y) \in f(\mathbf{R})$ car $x+y \in \mathbf{R}$. Donc $f(x)f(y) \in f(\mathbf{R})$ pour tout $x, y \in \mathbf{R}$.

3. 3.1) Par récurrence évidente : $f(x)^n = f(nx) = \begin{pmatrix} 1 & 0 & nx \\ -nx & 1 & -\frac{n^2x^2}{2} \\ 0 & 0 & 1 \end{pmatrix}$.

3.2) Premier point : est-ce que toute matrice de la forme $f(x)$, $x \in \mathbf{R}$ est inversible? Oui, puisqu'en choisissant $y = -x$ dans la question précédente, on déduit

$$f(x)f(-x) = f(x-x) = f(0) = I_3.$$

Donc $f(x)$ est inversible et $f(x)^{-1} = f(-x) = \begin{pmatrix} 1 & 0 & -x \\ +x & 1 & -\frac{x^2}{2} \\ 0 & 0 & 1 \end{pmatrix}$ pour tout $x \in \mathbf{R}$.

Deuxième point : est-ce que toute matrice inversible est de la forme $f(x)$?

Non, considérer par exemple $T = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. Alors T est inversible car de rang 3 et n'est pas de la forme $f(x)$ puisque le coefficient (1,2) est non nul. Donc

$$f(\mathbf{R}^3) \neq \text{GL}_3(\mathbf{R}).$$

Exercice ALG.4.4 | Soit $A = \begin{pmatrix} 0 & 1/a & 1/a^2 & 1/a^3 \\ a & 0 & 1/a & 1/a^2 \\ a^2 & a & 0 & 1/a \\ a^3 & a^2 & a & 0 \end{pmatrix}$ avec $a \in \mathbf{R}^*$.

- Calculer A^2 et montrer que A^2 est combinaison linéaire de A et I_3 .
- En déduire que A est inversible et calculer A^{-1} .
- Montrer qu'il existe pour tout $n \in \mathbf{N}$, a_n, b_n vérifiant $A^n = a_n A + b_n I_3$. Que peut-on dire de la suite $(a_n - b_n)$?
- Soit $c_n = (-1)^n b_n$, trouver une relation de récurrence entre c_{n+1} et c_n . En déduire A^n .

Solution (exercice ALG.4.4)

1. Par simple calcul, on a

$$A^2 = \begin{pmatrix} 0 & 1/a & 1/a^2 & 1/a^3 \\ a & 0 & 1/a & 1/a^2 \\ a^2 & a & 0 & 1/a \\ a^3 & a^2 & a & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1/a & 1/a^2 & 1/a^3 \\ a & 0 & 1/a & 1/a^2 \\ a^2 & a & 0 & 1/a \\ a^3 & a^2 & a & 0 \end{pmatrix} = \begin{pmatrix} 3 & \frac{2}{a} & \frac{2}{a^2} & \frac{2}{a^3} \\ 2a & 3 & \frac{2}{a} & \frac{2}{a^2} \\ 2a^2 & 2a & 3 & \frac{2}{a} \\ 2a^3 & 2a^2 & 2a & 3 \end{pmatrix},$$

d'où l'on tire $A^2 = 2A + 3I_3$.

2. On obtient alors $A(A - 2I_3) = 3I_3$ soit

$$A \left(\frac{A}{3} - \frac{2I_3}{3} \right) = \left(\frac{A}{3} - \frac{2I_3}{3} \right) A = I_3.$$

Ainsi, A est inversible d'inverse $\frac{A}{3} - \frac{2I_3}{3}$ soit

$$A^{-1} = \begin{pmatrix} \frac{-2}{3} & \frac{1}{3a} & \frac{1}{3a^2} & \frac{1}{3a^3} \\ \frac{a}{3} & \frac{-2}{3} & \frac{1}{3a} & \frac{1}{3a^2} \\ \frac{a^2}{3} & \frac{a}{3} & \frac{-2}{3} & \frac{1}{3a} \\ \frac{a^3}{3} & \frac{a^2}{3} & \frac{a}{3} & \frac{-2}{3} \end{pmatrix}.$$

3. Montrons l'existence par récurrence sur $n \in \mathbf{N}$.

■ **Initialisation.** Pour $n = 0$, on pose $a_0 = 0, b_0 = 1$ car $A^0 = I_3$.

■ **Hérédité.** Supposons l'existence de a_n, b_n en un certain rang n , ainsi

$$A^n = a_n A + b_n I_3, \implies A^{n+1} = a_n A^2 + b_n A = a_n (2A + 3I_3) + b_n A,$$

soit en factorisant $A^{n+1} = (2a_n + b_n)A + 3a_n I_3$. On pose alors

$$a_{n+1} = 2a_n + b_n, \quad b_{n+1} = 3a_n.$$

D'où la propriété au rang $n + 1$, et pour tout $n \in \mathbf{N}$, il existe a_n, b_n vérifiant $A^n = a_n A + b_n I_3$. Que peut-on dire de la suite $(a_n - b_n)$?

Pour tout $n \in \mathbf{N}$,

$$a_{n+1} - b_{n+1} = 2a_n + b_n - 3a_n = -(a_n - b_n),$$

donc la suite $(b_n - a_n)$ est géométrique de raison 1, avec $a_0 - b_0 = -1$, d'où

$$\forall n \in \mathbf{N}, \quad b_n - a_n = (-1)^{n+1}.$$

4. Soit $c_n = (-1)^n b_n$, et $n \in \mathbf{N}$. Alors

$$\begin{aligned} c_{n+1} &= (-1)^{n+1} b_{n+1} \\ &= (-1)^{n+1} (3a_n) \\ &= 3(-1)^{n+1} (b_n - (-1)^{n+1}) \\ &= 3(-1)^{n+1} b_n - 3 \\ &= -3c_n - 3. \end{aligned}$$

Il s'agit alors d'une suite arithmético-géométrique. On cherche alors ρ de sorte que

$$\rho = -3\rho - 3 \iff \rho = -\frac{3}{4}.$$

Donc pour tout $n \in \mathbf{N}$, en faisant la différence des deux lignes du système

$$\begin{cases} c_{n+1} = -3c_n - 3, \\ \rho = -3\rho - 3 \end{cases},$$

on trouve que $(c_n - \rho)$ est géométrique de raison -3 . Donc

$$c_n - \rho = (-3)^n (c_0 - \rho),$$

soit

$$c_n = 3^n (-1)^n \frac{7}{4} - \frac{3}{4}.$$

D'où :

$$b_n = (-1)^n c_n = 3^n \frac{7}{4} - (-1)^n \frac{3}{4}, \quad a_n = \frac{1}{3} b_{n+1} = 3^n \frac{7}{4} - (-1)^{n+1} \frac{1}{4}.$$

En conclusion

$$A^n = \left(3^n \frac{7}{4} - (-1)^{n+1} \frac{1}{4} \right) A + \left(3^n \frac{7}{4} - (-1)^n \frac{3}{4} \right) I_3,$$

ce qui fournit après de longs calculs :

$$A^n = \begin{pmatrix} \frac{3 \cdot (-1)^n + 3^n}{(-1)^n \cdot a + 3^n \cdot a} & \frac{-(-1)^n + 3^n}{3 \cdot (-1)^n + 3^n} & \frac{-(-1)^n + 3^n}{(-1)^n + 3^n} & \frac{-(-1)^n + 3^n}{(-1)^n + 3^n} \\ \frac{-(-1)^n \cdot a^2 + 3^n \cdot a^2}{(-1)^n \cdot a + 3^n \cdot a} & \frac{-(-1)^n \cdot a + 3^n \cdot a}{3 \cdot (-1)^n + 3^n} & \frac{4 \cdot a^2}{(-1)^n + 3^n} & \frac{4 \cdot a^3}{(-1)^n + 3^n} \\ \frac{-(-1)^n \cdot a^3 + 3^n \cdot a^3}{(-1)^n \cdot a^2 + 3^n \cdot a^2} & \frac{-(-1)^n \cdot a^2 + 3^n \cdot a^2}{(-1)^n \cdot a + 3^n \cdot a} & \frac{4 \cdot a}{3 \cdot (-1)^n + 3^n} & \frac{4 \cdot a^2}{(-1)^n + 3^n} \\ \frac{-(-1)^n \cdot a^3 + 3^n \cdot a^3}{(-1)^n \cdot a^2 + 3^n \cdot a^2} & \frac{-(-1)^n \cdot a^2 + 3^n \cdot a^2}{(-1)^n \cdot a + 3^n \cdot a} & \frac{4 \cdot a}{3 \cdot (-1)^n + 3^n} & \frac{4 \cdot a}{(-1)^n + 3^n} \end{pmatrix}.$$

Exercice ALG.4.5 | Matrice d'ATTILA On note $J_n = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix}$ de format $n \times n$ avec $n \geq 1$.

1. Calculer J_n^k pour tout $k \in \mathbf{N}$.
2. La matrice J_n est-elle inversible?
3. Décrire l'image et le noyau de J_n .

Solution (exercice ALG.4.5)

1. On constate que $J_n^2 = nJ_n$ puis on conjecture que $J_n^k = n^{k-1}J_n$ pour $k \geq 1$.

■ **Initialisation.** La propriété est vraie au rang $k = 1$.

■ **Hérédité.** Supposons que $J_n^k = n^{k-1}J_n$ pour un certain $k \geq 1$. Alors

$$J_n^{k+1} = J_n^k J_n = n^{k-1} J_n^2 = n^k J_n,$$

en utilisant $J_n^2 = nJ_n$. Donc la propriété est vraie par principe de récurrence :

$$\forall k \geq 1, \quad J_n^k = n^{k-1} J_n.$$

2. La matrice J_n est-elle inversible? Non, car elle est clairement de rang $1 < n$, et donc est non inversible.

3. Nous avons

$$\begin{aligned} X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \ker(J_n) &\iff J_n X = 0 \iff \begin{cases} x_1 + \dots + x_n = 0, \\ \dots = 0, \\ x_1 + \dots + x_n = 0 \end{cases} \\ &\iff X = \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \\ -x_1 - \dots - x_{n-1} \end{pmatrix} \\ &\iff X = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix} + \dots + x_{n-1} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ -1 \end{pmatrix}. \end{aligned}$$

Donc

$$\ker J_n = \text{Vect} \left(\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ -1 \end{pmatrix} \right).$$

On sait déjà que l'image de J_n est de dimension 1, il suffit donc de trouver un vecteur dans l'image. Par exemple

$$J_n \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \text{Im}(J_n), \quad \text{donc} \quad \text{Im}(J_n) = \text{Vect} \left(\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \right).$$

Exercice ALG.4.6 | Équation dans $\mathfrak{M}_2(\mathbf{R})$

1. Soit $A \in \mathfrak{M}_2(\mathbf{R})$ qui commute avec une matrice diagonale $D \in \mathfrak{M}_2(\mathbf{R})$ à coefficients deux à deux distincts. Montrer que A est également diagonale.
2. On note $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. Trouver une matrice $P \in \mathfrak{M}_2(\mathbf{R})$ inversible, telle que $P^{-1}AP = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}$.
3. Résoudre l'équation $X^2 - 2X = A$ d'inconnue $X \in \mathfrak{M}_2(\mathbf{R})$. *Indication : On pourra introduire la matrice inconnue $Y \in \mathfrak{M}_2(\mathbf{R})$ telle que $X = PYP^{-1}$.*

Solution (exercice ALG.4.6)

1. On note $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ avec $\lambda_1 \neq \lambda_2$ et $A = (a_{i,j})_{1 \leq i,j \leq 2}$, les coefficients de A . L'égalité $AD = DA$ donne en particulier :

$$a_{1,2}\lambda_2 = \lambda_1 a_{1,2}, \quad a_{2,1}\lambda_1 = \lambda_2 a_{2,1},$$

soit comme $\lambda_1 \neq \lambda_2$:

$$a_{1,2} = a_{2,1} = 0.$$

Cela signifie que A est diagonale.

2. Prenant une matrice $P = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ avec $(a, b, c, d) \in \mathbf{R}^4$ et résolvant $AP = P \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}$,

on trouve $P = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$ qui convient.

3. Notons $D = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}$. Alors l'équation se réécrit $(P^{-1}XP)^2 - 2P^{-1}XP = D$, ce qui est

donc une équation en l'inconnue $Y = P^{-1}XP$. Prenons Y sous la forme $Y = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$,

on résout alors $Y^2 - 2Y = D = Y(Y - 2I_2)$. Si Y est une solution alors $YD = DY$ puisque D est un polynôme en Y . Donc Y est également une matrice diagonale de la forme

$\begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}$. En injectant dans l'identité on trouve $\alpha^2 - 2\alpha = -1$ et $\beta^2 - 2\beta = 3$. Donc

$$\alpha = 1 \text{ et } \beta \in \{3, -1\}.$$

Exercice ALG.4.7 | Une propriété des polynômes de degré trois Soit f l'application définie par :

$$f \left| \begin{array}{l} \mathbf{R}_3[X] \\ P \end{array} \right. \begin{array}{l} \longrightarrow \\ \longrightarrow \end{array} \begin{array}{l} \mathbf{R}^4 \\ (P(0), P'(1), P''(2), P^{(3)}(3)) \end{array}.$$

1. Montrer que f est linéaire et donner sa matrice M relativement aux bases canoniques.
2. M est-elle inversible? Si oui, donner M^{-1} .
3. Soit $P \in \mathbf{R}_3[X]$. Exprimer alors $P'(1)$, $P''(2)$ et $P^{(3)}(3)$ en fonction de $P'(0)$ et $P''(0)$ en utilisant la question précédente.

Solution (exercice ALG.4.7)

1. Soient $P, Q \in \mathbf{R}_3[X]$ et $\lambda, \mu \in \mathbf{R}$, alors

$$\begin{aligned} f(\lambda P + \mu Q) &= ((\lambda P + \mu Q)(0), (\lambda P + \mu Q)'(1), (\lambda P + \mu Q)''(2), (\lambda P + \mu Q)^{(3)}(3)) \\ &= (\lambda P(0) + \mu Q(0), \lambda P'(1) + \mu Q'(1), \lambda P''(2) + \mu Q''(2), \lambda P^{(3)}(3) + \mu Q^{(3)}(3)) \\ &= \lambda f(P) + \mu f(Q). \end{aligned}$$

Donc f est linéaire et $f(1) = (1, 0, 0, 0), f(X) = (0, 1, 0, 0), f(X^2) = (0, 2, 2, 0), f(X^3) = (0, 1, 12, 6)$, donc

$$\mathcal{M}_f^{\text{Mat}(\mathcal{B}^{\text{can}})} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 2 & 12 \\ 0 & 0 & 0 & 6 \end{pmatrix}.$$

2. La matrice M est inversible car échelonnée de rang 4. Après application de la méthode du miroir, nous déduisons

$$M^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & \frac{11}{6} \\ 0 & 0 & \frac{1}{2} & -1 \\ 0 & 0 & 0 & \frac{1}{6} \end{pmatrix}.$$

3. Soit $P = \sum_{k=0}^3 \frac{P^{(k)}(0)}{k!} X^k \in \mathbf{R}_3[X]$. Alors par définition de la matrice inverse :

$$M^{-1} \times \mathcal{M}_P^{\text{Mat}(\mathcal{B}^{\text{can}})} = M^{-1} \times \begin{pmatrix} P(0) \\ P'(0) \\ \frac{1}{2}P''(0) \\ \frac{1}{6}P^{(3)}(0) \end{pmatrix} = \begin{pmatrix} P(0) \\ P'(1) \\ P''(2) \\ P^{(3)}(3) \end{pmatrix}.$$

Ainsi, en effectuant le produit matriciel puis en analysant les trois dernières coordonnées, on trouve les relations suivantes :

$$P'(1) = P'(0) - \frac{P''(0)}{2} + \frac{11}{36}P^{(3)}(0),$$

$$P''(2) = \frac{P''(0)}{4} - \frac{1}{6}P^{(3)}(0),$$

$$P^{(3)}(3) = \frac{1}{36}P^{(3)}(0).$$

3.2. Représentation matricielle

Exercice ALG.4.8 | On admet la linéarité des applications ci-dessous. Déterminer leur rang, ainsi que les valeurs propres lorsque la matrice associée est carrée.

- $f_1 : (x, y, z) \in \mathbf{R}^3 \longrightarrow (x - y + z, 2z)$,
- $f_2 : (x, y) \in \mathbf{R}^2 \longrightarrow (4x - 6y, 2x - 3y)$.
- $f_3 : (x, y) \in \mathbf{R}^2 \longrightarrow (4x + y, x - y, 2x + 3y)$.

Solution (exercice ALG.4.8)

- Puisque $f_1(1, 0, 0) = (1, 0), f_1(0, 1, 0) = (-1, 0), f_1(0, 0, 1) = (1, 2)$, il vient

$\mathcal{M}_{\mathcal{B}^{\text{can}}}(f_1) = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 0 & 2 \end{pmatrix}$. La matrice est donc de rang 2 car possède deux pivots.

- Puisque $f_2(1, 0) = (4, 2), f_2(0, 1) = (-6, -3)$, il vient $\mathcal{M}_{\mathcal{B}^{\text{can}}}(f_2) = \begin{pmatrix} 4 & 2 \\ -6 & -3 \end{pmatrix}$. La matrice est donc de rang 1 car non nulle et non inversible. Soit $\lambda \in \mathbf{C}$, on résout alors $\det(A - \lambda I_2) = 0$ pour trouver les valeurs propres.

$$\det(A - \lambda I_2) = 0 \iff (4 - \lambda)(-3 - \lambda) + 12 = 0,$$

l'équation est alors $\lambda^2 - \lambda = \lambda(\lambda - 1) = 0$, donc $\text{Spec}A = \{0, 1\}$.

- Puisque $f_3(1, 0) = (4, 1, 2), f_3(0, 1) = (1, -1, 3)$, il vient $\mathcal{M}_{\mathcal{B}^{\text{can}}}(f_3) = \begin{pmatrix} 4 & 1 \\ 1 & -1 \\ 2 & 3 \end{pmatrix}$. On échelonne ensuite la matrice en faisant $L_2 \leftarrow 4L_2 - L_1, L_3 \leftarrow 2L_3 - L_1$

$$\begin{pmatrix} 4 & 1 \\ 1 & -1 \\ 2 & 3 \end{pmatrix} \underset{L}{\sim} \begin{pmatrix} 4 & 1 \\ 0 & -5 \\ 0 & 5 \end{pmatrix},$$

on obtient alors une matrice de rang deux puisque $L_2 = -L_3$. Donc $\text{Rg}(f_3) = 2$.

Exercice ALG.4.9 | Soit $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. On définit u par :

$$\forall M \in \mathfrak{M}_2(\mathbf{R}), \quad u(M) = AM.$$

- Montrer que u est un endomorphisme de $\mathfrak{M}_2(\mathbf{R})$.
- Déterminer la matrice de u dans la base canonique de $\mathfrak{M}_2(\mathbf{R})$.
- Déterminer le rang, le noyau et l'image de u .

Solution (exercice ALG.4.9)

- Soient $M, M' \in \mathfrak{M}_2(\mathbf{R})$, et λ, λ' deux réels. Alors

$$u(\lambda M + \mu M') = A(\lambda M + \mu M') = \lambda AM + \mu AM' = \lambda u(M) + \mu u(M'),$$

donc u est un endomorphisme de $\mathfrak{M}_2(\mathbf{R})$.

- Rappelons que la base canonique de $\mathfrak{M}_2(\mathbf{R})$ est $(E_{1,1}, E_{1,2}, E_{2,1}, E_{2,2})$, et que

$$u(E_{1,1}) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$$

$$u(E_{1,2}) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$$

$$u(E_{2,1}) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$$

$$u(E_{2,2}) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$$

donc

$$\mathcal{M}_{\mathcal{B}^{\text{can}}}(u) = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}.$$

3. Déterminer le rang, le noyau et l'image de u .

Exercice ALG.4.10 | Soient A, B deux matrices de $\mathfrak{M}_{n,n}(\mathbf{K})$ avec $n \geq 1$ un entier.

Montrer que

$$\text{Rg}(AB) \leq \min \{ \text{Rg}(A), \text{Rg}(B) \}.$$

Solution (exercice ALG.4.10)

Il s'agit de montrer que

$$\text{Rg}(AB) \leq \text{Rg} A \quad \text{et} \quad \text{Rg}(AB) \leq \text{Rg} B.$$

► Commençons par montrer que $\text{Rg}(AB) \leq \text{Rg} A$. Nous avons

$$\text{Im}(AB) = \{ ABX, X \in \mathfrak{M}_{n,1}(\mathbf{K}) \} \subset \{ AY, Y \in \mathfrak{M}_{n,1}(\mathbf{K}) \} = \text{Im}(A),$$

donc en passant à la dimension on déduit que

$$\text{Rg}(AB) \leq \text{Rg}(A).$$

► Poursuivons avec $\text{Rg}(AB) \leq \text{Rg} B$. Commençons par appliquer le théorème du rang :

$$\text{Rg}(AB) = n - \dim \ker(AB), \quad \text{Rg}(B) = n - \dim \ker(B),$$

donc l'inégalité est équivalente à

$$n - \dim \ker(AB) \leq n - \dim \ker(B),$$

soit

$$\dim \ker(AB) \geq \dim \ker(B).$$

Mais $\ker B \subset \ker(AB)$ puisque si $X \in \ker B$, alors $BX = 0$, donc $ABX = 0$ et donc $X \in \ker(AB)$. On a donc bien en passant à la dimension : $\dim \ker(AB) \geq \dim \ker(B)$. On déduit alors

$$\text{Rg}(AB) \leq \text{Rg} B.$$

En conclusion

$$\boxed{\text{Rg}(AB) \leq \min \{ \text{Rg}(A), \text{Rg}(B) \}}.$$

Exercice ALG.4.11 | Soit $E = \mathbf{R}[X]$, on définit

$$\varphi \begin{cases} \mathbf{R}[X] & \longrightarrow & \mathbf{R}[X], \\ P & \longrightarrow & (2X + 1)P(X) + (1 - X^2)P'(X). \end{cases}$$

1. Montrer que φ est linéaire et déterminer $\varphi(X^k)$ pour tout $k \in \mathbf{N}$.
2. Soit $P \in E, P \neq 0$, comparer les degrés de P et de $\varphi(P)$.
3. L'application φ est-elle injective?
4. Soit u la restriction de φ à $\mathbf{R}_2[X]$.
 - 4.1) Montrer que $u \in \mathcal{L}(\mathbf{R}_2[X])$.
 - 4.2) Écrire la matrice U de u dans la base canonique de $\mathbf{R}^2[X]$. L'application u est-elle injective? bijective?

Solution (exercice ALG.4.11)

1. Soient $P, Q \in \mathbf{R}[X]$ et $\lambda, \mu \in \mathbf{R}$. Alors :

$$\begin{aligned} \varphi(\lambda P + \mu Q) &= (2X + 1)(\lambda P + \mu Q)(X) + (1 - X^2)(\lambda P + \mu Q)'(X) \\ &= (2X + 1)(\lambda P + \mu Q)(X) + (1 - X^2)(\lambda P'(X) + \mu Q'(X)) \\ &= \lambda \varphi(P) + \mu \varphi(Q). \end{aligned}$$

Donc φ est linéaire. Par ailleurs,

$$\varphi(X^k) = \begin{cases} (2X + 1)X^k + (1 - X^2)kX^{k-1} = (2 - k)X^{k+1} + X^k + kX^{k-1} & \text{si } k \geq 1, \\ 2X + 1 & \text{si } k = 0. \end{cases}$$

2. Soit $P \in E, P \neq 0$, comparons les degrés de P et de $\varphi(P)$. On a $\deg((2X + 1)P(X)) = \deg P + 1$ et $\deg((1 - X^2)P'(X)) = \deg P - 1 + 2 = \deg P + 1$. Donc le degré de $\varphi(P)$ est inférieur ou égal à $\deg P + 1$. Puisque l'on veut une formule, il faut être plus précis

et considérer $P = \sum_{k=0}^n a_k X^k$ avec $n = \deg P, a_n \neq 0$ et les coefficients étant réels. Alors, par linéarité,

$$\varphi(P) = \sum_{k=0}^n a_k \varphi(X^k).$$

Mais, d'après la question 1, $\varphi(X^k)$ est toujours de degré $k + 1$, sauf pour $k = 2$ où il est de degré 2. Donc par propriété du degré,

▶ si $n = \deg P \neq 2$, $\deg \varphi(P) = \deg \varphi(X^n) = n + 1 = \boxed{\deg P + 1}$,

▶ si $\deg P = 2$, alors $\deg \varphi(P) = \deg \varphi(X^2) = 2 = \boxed{\deg P}$.

3. Soit $P \in \ker \varphi$, i.e. $\varphi(P) = 0$. Alors $\deg \varphi(P) = -\infty$, deux cas se présentent alors.

▶ Soit $\deg P = 2$, auquel cas $\deg \varphi(P) = \deg P = -\infty$ donc aucun polynôme de degré deux n'est dans le noyau.

▶ Si $\deg P \neq 2$, alors $\deg \varphi(P) = \deg P + 1 = -\infty$, donc P est nul.

Donc $\boxed{\varphi \text{ est injective.}}$

4. Soit u la restriction de φ à $\mathbf{R}_2[X]$.

4.1) L'application u est encore linéaire par restriction d'une telle application. De plus, d'après ce qui précède si $P \in \mathbf{R}_2[X]$ alors $\deg \varphi(P) \leq 2$ donc φ est un $\boxed{\text{endomorphisme de } \mathbf{R}_2[X]}$.

4.2) En recyclant les calculs de la première question pour $k = 0, 1, 2$, on obtient

$$\mathcal{M}_{\mathcal{B}^{\text{can}}}^{\text{can}}(u) = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix}.$$

Après échelonnement, on constate que A est une matrice de rang 3, elle est donc inversible et $\boxed{u \text{ est une bijection de } \mathbf{R}_2[X]}$.

Exercice ALG.4.12 | Soit E l'espace vectoriel des fonctions polynômes réelles définies sur $\mathbf{R}^{+\ast}$ de degré inférieur ou égal à 4. On considère

$$\varphi \begin{cases} E \longrightarrow \mathcal{F}(\mathbf{R}^{+\ast}, \mathbf{R}), \\ P \longrightarrow (x \longmapsto P(x) + 2x^4 P(1/x)). \end{cases}$$

1. Montrer que φ est un endomorphisme de E .

- 2. 2.1) Exprimer φ^2 en fonction de φ et de l'identité.
- 2.2) En déduire que φ est inversible et déterminer son inverse.
- 2.3) Déterminer la matrice de φ canoniquement associée, et retrouver l'inversibilité démontrée avant.

Solution (exercice ALG.4.12)

1. Soient $P, Q \in E$ et λ, μ deux réels, ainsi que $x \in \mathbf{R}^{+\ast}$. On a

$$\begin{aligned} \varphi(\lambda P + \mu Q)(x) &= (\lambda P + \mu Q)(x) + 2x^4(\lambda P + \mu Q)(1/x) \\ &= (\lambda P(x) + \mu Q(x)) + 2x^4(\lambda P(1/x) + \mu Q(1/x)) \\ &= \lambda \varphi(P)(x) + \mu \varphi(Q)(x). \end{aligned}$$

Il s'agit ensuite de constater que $\varphi(P)$ est encore un polynôme de degré inférieur à 4. Notons $P = \sum_{k=0}^4 a_k X^k \in E$. Alors

$$\forall x \in \mathbf{R}^{+\ast}, \quad \varphi(P)(x) = \sum_{k=0}^4 a_k x^k + 2x^4 \sum_{k=0}^4 a_k \frac{1}{x^k},$$

alors en rentrant la puissance de x , on obtient :

$$\forall x \in \mathbf{R}^{+\ast}, \quad \varphi(P)(x) = \sum_{k=0}^4 a_k x^k + 2 \sum_{k=0}^4 a_k x^{4-k},$$

donc $\varphi(P) \in E$. Finalement $\boxed{\varphi \text{ est un endomorphisme de } E}$.

2. 2.1) Soit $P \in E$ ainsi que $x \in \mathbf{R}^{+\ast}$. On calcule

$$\begin{aligned} \varphi \circ \varphi(P)(x) &= \varphi(P)(x) + 2x^4 \cdot \varphi(P)(1/x) \\ &= P(x) + 2x^4 P(1/x) + 2x^4 (P(1/x) + 2(1/x)^4 P(x)) \\ &= 5P(x) + 4x^4 P(1/x) \\ &= 3P(x) + 2\varphi(P)(x). \end{aligned}$$

Autrement dit, $\boxed{\varphi^2 = 2\varphi + 3 \text{Id}_E}$.

2.2) La question précédente fournit : $\varphi \circ \left(\frac{\varphi - 2\text{Id}_E}{3}\right) = \text{Id}_E = \left(\frac{\varphi - 2\text{Id}_E}{3}\right) \circ \varphi$. Donc φ est inversible et d'inverse $\boxed{\varphi^{-1} = \frac{\varphi - 2\text{Id}_E}{3}}$.

2.3) On calcule pour tout $x \in \mathbf{R}$,

$$\begin{aligned}\varphi(1)(x) &= 1 + 2x^4 = 2x^4 + 1, \\ \varphi(X)(x) &= x + 2x^3 = 2x^3 + x, \\ \varphi(X^2)(x) &= x^2 + 2x^2 = 3x^2, \\ \varphi(X^3)(x) &= x^3 + 2x = x^3 + 2x, \\ \varphi(X^4)(x) &= x^4 + 2.\end{aligned}$$

Autrement dit

$$\varphi(1) = 2X^4 + 1, \varphi(X) = 2X^3 + X, \varphi(X^2) = 3X^2, \varphi(X^3) = X^3 + 2X, \varphi(X^4) = X^4 + 2.$$

On déduit alors la matrice dans le base canonique

$$\mathcal{M}_{\varphi}(\mathcal{B}^{\text{can}}) = \begin{pmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Après échelonnement, on constate qu'elle est de rang 5 donc inversible.

Exercice ALG.4.13 | Dans le \mathbf{R} -espace vectoriel $E = \mathcal{F}(\mathbf{R}, \mathbf{R})$ des fonctions définies sur \mathbf{R} et à valeurs réelles, on considère les fonctions

$$f_1 : x \mapsto e^{-x}, \quad f_2 : x \mapsto xe^{-x}, \quad f_3 : x \mapsto x^2e^{-x},$$

et on note $F = \text{Vect}(f, g, h)$.

1. Montrer que la famille (f_1, f_2, f_3) est une base de F . On la notera \mathcal{B} .
2. On considère l'application $\Phi : f \in F \mapsto f'$. Montrer que Φ est un endomorphisme de F , puis écrire $A = \mathcal{M}_{\mathcal{B}}(\Phi)$.
3. Calculer A^n en écrivant A sous la forme $-I_n + N_n$ avec N_n une matrice de format $n \times n$.

Solution (exercice ALG.4.13)

1. la famille \mathcal{B} est par construction une famille génératrice de F . Soient λ, μ, ν trois réels tels que

$$\forall x \in \mathbf{R}, \quad \lambda e^{-x} + \mu x e^{-x} + \nu x^2 e^{-x} = 0,$$

en multipliant par e^x , on déduit :

$$\forall x \in \mathbf{R}, \quad \lambda + \mu x + \nu x^2 = 0.$$

Cela signifie que le polynôme $\lambda + \mu X + \nu X^2$ est nul, donc $\lambda = \mu = \nu = 0$, la famille (f_1, f_2, f_3) est donc une base de F .

2. On considère l'application $\Phi : f \in F \mapsto f'$. L'application Φ est bien linéaire, par linéarité de la dérivation. Soit maintenant $g = \lambda f_1 + \mu f_2 + \nu f_3$ un élément de F avec λ, μ, ν trois réels. Alors pour tout $x \in \mathbf{R}$,

$$\begin{aligned}g'(x) &= \lambda f_1'(x) + \mu f_2'(x) + \nu f_3'(x) \\ &= -\lambda e^{-x} + \mu(e^{-x} - x e^{-x}) + \nu(2x e^{-x} - x^2 e^{-x}) \\ &= (\mu - \lambda)f_1(x) + (2\nu - \mu)f_2(x) - \nu f_3(x).\end{aligned}$$

Donc $D(g) \in F$ et D est bien un endomorphisme de F . On a par ailleurs

$$\begin{aligned}D(f_1) &= -f_1 \\ D(f_2) &= f_1 - f_2 \\ D(f_3) &= 2f_2 - f_3.\end{aligned}$$

D'où :

$$\mathcal{M}_{\mathcal{B}}(\Phi) = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 2 \\ 0 & 0 & -1 \end{pmatrix}.$$

3. Posons $N_n = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$ de sorte que $A = -I_n + N_n$. Alors $-I_n$ et N_n commutent, et

$N_n^2 = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, $N_n^3 = 0$. On peut donc utiliser la formule du binôme

$$A^n = \sum_{k=0}^n \binom{n}{k} (-I_n)^{n-k} N_n^k = \sum_{k=0}^n \binom{n}{k} (-1)^{n-k} N_n^k.$$

Alors,

$$A^n = \begin{cases} (-1)^n I_n + n(-1)^{n-1} N_n + \frac{n(n-1)}{2} (-1)^{n-2} N_n^2 & \text{si } n \geq 2, \\ A & \text{si } n = 1, \\ I_n & \text{si } n = 0. \end{cases}$$

Les trois formules se réunissent en une seule :

$$\forall n \geq 0, \quad A^n = (-1)^n I_n + n(-1)^{n-1} N_n + \frac{n(n-1)}{2} (-1)^{n-2} N_n^2$$

d'où

$$A^n = \begin{pmatrix} (-1)^n & n(-1)^{n-1} & n(n-1)(-1)^n \\ 0 & (-1)^n & 2n(-1)^{n-1} \\ 0 & 0 & (-1)^n \end{pmatrix},$$

car $(-1)^{n-2} = (-1)^n$ car les entiers $n, n-2$ ont même parité.

Exercice ALG.4.14 | Soient $p \geq 3$ et

$$D_p = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & \vdots \\ 0 & & & & 1 \\ 1 & 0 & \dots & 0 & 0 \end{pmatrix} \in \mathfrak{M}_{p,p}(\mathbb{C}).$$

1. On souhaite étudier dans cette question l'inversibilité de D_p .

- 1.1) Inverser les matrices D_3 et D_4 .
- 1.2) Prouver que D_p est inversible et donner son inverse. On émettra une conjecture que l'on cherchera à démontrer.
- 2. On souhaite dans cette question retrouver le résultat précédent dans le cas $p = 4$. Soit $\mathcal{B} = (e_1, e_2, e_3, e_4)$ une base d'un espace vectoriel E . On définit $f \in \mathcal{L}(E)$ par $f(e_i) = e_{i+1}$ pour $1 \leq i \leq 3$ et $f(e_4) = e_1$.
 - 2.1) Justifier, sans calcul, que f est un automorphisme.
 - 2.2) Déterminer la matrice de f dans \mathcal{B} que l'on notera A .
 - 2.3) Déterminer l'application réciproque de f , et en déduire D_4^{-1} .
 - 2.4) Comment pourrait-on généraliser ce qui précède à D_p , en utilisant un endomorphisme?

Solution (exercice ALG.4.14)

1. On souhaite étudier dans cette question l'inversibilité de D_p . Avec la méthode du miroir on trouve sans difficulté que

$$D_3^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad D_4^{-1} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

On conjecture alors que D_p est inversible, et d'inverse :

$$D_p^{-1} = \begin{pmatrix} 0 & \dots & 0 & 1 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & & & 1 \\ 0 & \dots & 1 & 0 \end{pmatrix}.$$

Pour justifier cela, on effectue simplement le produit matriciel $D_p \times D_p^{-1} = D_p^{-1} \times D_p$, on constate qu'ils sont égaux à I_p .

- 2. Soit $\mathcal{B} = (e_1, e_2, e_3, e_4)$ une base d'un espace vectoriel E . On définit $f \in \mathcal{L}(E)$ par $f(e_i) = e_{i+1}$ pour $1 \leq i \leq 3$ et $f(e_4) = e_1$.
 - 2.1) Par définition, f envoie la base \mathcal{B} sur $f(e_2, e_3, e_4, e_1)$ qui est encore une base, donc f est un isomorphisme.

2.2) On obtient immédiatement

$$\mathcal{M}_{\mathcal{B}}(f) = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

2.3) On constate facilement que f^{-1} définie par

$$\forall i \in \llbracket 2, 4 \rrbracket, f^{-1}(e_i) = e_{i-1}, \quad f^{-1}(e_1) = e_4$$

définie une application linéaire qui est la réciproque de f . En effet, pour tout $i \in \llbracket 1, 4 \rrbracket$,

$$f^{-1} \circ f(e_i) = \begin{cases} f^{-1}(e_{i+1}) = e_i & \text{si } i \in \llbracket 1, 3 \rrbracket, \\ f^{-1}(e_1) = e_4 & \text{si } i = 4. \end{cases}$$

Donc $f^{-1} \circ f(e_i) = e_i = \text{Id}(e_i)$ pour tout $i \in \llbracket 1, 4 \rrbracket$. De-même on vérifie que $f \circ f^{-1}(e_i) = e_i = \text{Id}(e_i)$ si $i \in \llbracket 1, 4 \rrbracket$. On déduit alors que $f \circ f^{-1} = f^{-1} \circ f = \text{Id}$: soit donc $x \in E$, il existe $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ trois réels tels que $x = \sum_{i=1}^4 \lambda_i e_i$ et par linéarité

$$f^{-1} \circ f(x) = \sum_{i=1}^4 \lambda_i f^{-1} \circ f(e_i) = \sum_{i=1}^4 \lambda_i e_i = x = \text{Id}(x),$$

et de-même $f \circ f^{-1}(x) = x$ pour tout x . Donc f est inversible d'inverse f^{-1} . D'après le cours, A est alors inversible d'inverse

$$\mathcal{M}_{\mathcal{B}}(f^{-1}) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} = D_4.$$

On déduit alors en conséquence que D_4 est inversible d'inverse $(A^{-1})^{-1} = A$.

2.4) Pour généraliser, il suffirait alors de considérer $\mathcal{B} = (e_1, \dots, e_p)$ une base d'un espace vectoriel E . On définit $f \in \mathcal{L}(E)$ par $f(e_i) = e_{i+1}$ pour $1 \leq i \leq p-1$ et $f(e_p) = e_1$.

3.3. Changements de base

Exercice ALG.4.15 |

1. Soient $f_1 = (1, 2), f_2 = (-3, 1)$ et $u = (-1, 4)$.

1.1) Justifier que $\mathcal{B} = (f_1, f_2)$ est une base de \mathbf{R}^2 .

1.2) Donner les coordonnées de u dans la base \mathcal{B} .

2. Soit f l'application de \mathbf{R}^2 dans \mathbf{R}^3 ayant pour matrice $\begin{pmatrix} 1 & 1 \\ 2 & 1 \\ -1 & 3 \end{pmatrix}$ dans les bases ca-

noniques \mathcal{B}_c et \mathcal{B}'_c respectivement de \mathbf{R}^2 et \mathbf{R}^3 .

2.1) Déterminer $f(u)$, et $\ker f$.

2.2) Écrire la matrice de f dans les bases \mathcal{B} et \mathcal{B}'_c et l'utiliser pour retrouver $f(u)$.

Solution (exercice ALG.4.15)

1. 1.1) La famille \mathcal{B} est une base de \mathbf{R}^2 si et seulement si son rang vaut deux. Comme nous avons deux vecteurs en dimension deux, ceci est équivalent à $\begin{pmatrix} 1 & 2 \\ -3 & 1 \end{pmatrix}$ inversible. Puisque $\begin{vmatrix} 1 & 2 \\ -3 & 1 \end{vmatrix} = 7 \neq 0$, nous obtenons bien que

\mathcal{B} est une base de \mathbf{R}^2 .

1.2) De plus on cherche λ, μ dans \mathbf{R}^2 tels que : $u = \lambda e_1 + \mu e_2$ soit donc le système suivant à résoudre :

$$\begin{cases} -1 = \lambda - 3\mu \\ 4 = -3\lambda + \mu \end{cases} \iff \mu = \frac{6}{7}, \lambda = \frac{11}{7}.$$

Les coordonnées de u dans la base \mathcal{B} sont donc : $\mu = \frac{6}{7}, \lambda = \frac{11}{7}$.

2. 2.1) Par définition de la matrice, nous avons :

$$f(u) = f(-1(1, 0) + 4(0, 1)) = -1f(1, 0) + 4f(0, 1) = -1(1, 2, -1) + 4(1, 1, 3) = (3, 2, 13)$$

où nous avons utilisé les trois colonnes de la matrice pour calculer $f(1, 0)$ et $f(0, 1)$. Donc $f(u) = (3, 2, 13)$. On résout tout d'abord en (x, y) le système suivant :

$$\begin{pmatrix} 1 & 1 \\ 2 & 1 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff \begin{cases} x+y = 0 \\ 2x+y = 0 \\ -x+3y = 0 \end{cases} \iff x = y = 0. \text{ Donc}$$

$$\ker f = \{(0, 0)\}.$$

2.2) On a : $\text{Mat}_{\mathcal{B}, \mathcal{B}'_c}(f) = \text{Mat}_{\mathcal{B}'_c, \mathcal{B}'_c}(\text{Id}_{\mathbb{R}^2}) \text{Mat}_{\mathcal{B}_c, \mathcal{B}'_c}(f) \text{Mat}_{\mathcal{B}, \mathcal{B}_c}(\text{Id}_{\mathbb{R}^2}) = \text{Mat}_{\mathcal{B}_c, \mathcal{B}'_c}(f) \text{Mat}_{\mathcal{B}, \mathcal{B}_c}(\text{Id}_{\mathbb{R}^2}) =$

$$\begin{pmatrix} 1 & 1 \\ 2 & 1 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 1 & -3 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 3 & -2 \\ 4 & -5 \\ 5 & 6 \end{pmatrix}. \text{ On en déduit alors } f(u) \text{ de matrice } \begin{pmatrix} 3 \\ 2 \\ 13 \end{pmatrix} \text{ dans}$$

la base canonique \mathcal{B}'_c comme nous l'avons prouvé avant.

$$\text{On a : } \text{Mat}_{\mathcal{B}'_c}(f(u)) = \text{Mat}_{\mathcal{B}, \mathcal{B}'_c}(f) \text{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} 3 & -2 \\ 4 & -5 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 11/7 \\ 6/7 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 13 \end{pmatrix}. \text{ d'après la}$$

première question. Le dernier vecteur est écrit dans \mathcal{B}'_c , la base canonique de \mathbb{R}^3 .

$$\text{On retrouve bien } f(u) = (3, 2, 13).$$

Exercice ALG.4.16 | Changement de base polynomial On note $E = \mathbb{R}_2[X]$ et soit

$$\Phi \begin{cases} E \longrightarrow \mathbb{R}^3, \\ P \longrightarrow (P(0), P(1), P(2)). \end{cases}$$

1. Montrer que Φ est une application linéaire.
2. Montrer que Φ est un isomorphisme.
3. On note $P_0 = (X-1)(X-2), P_1 = X(X-2), P_2 = X(X-1)$. Montrer que $\mathcal{B} = (P_0, P_1, P_2)$ est une base de E et calculer la matrice de passage de la base canonique vers \mathcal{B} .
4. Déterminer la matrice de Φ dans les bases $\mathcal{B}, \mathcal{B}^{\text{can}}$, en déduire la matrice canoniquement associée à Φ .

Solution (exercice ALG.4.16)

1. Soient $P, Q \in E$ et $\lambda, \mu \in \mathbb{R}$, alors

$$\begin{aligned} \Phi(\lambda P + \mu Q) &= ((\lambda P + \mu Q)(0), (\lambda P + \mu Q)(1), (\lambda P + \mu Q)(2)) \\ &= ((\lambda P + \mu Q)(0), (\lambda P + \mu Q)(1), (\lambda P + \mu Q)(2)) \\ &= \lambda \Phi(P) + \mu \Phi(Q). \end{aligned}$$

Donc Φ est bien linéaire.

2. Soit $P \in \ker \Phi$, alors $(P(0), P(1), P(2)) = (0, 0, 0)$ donc P admet trois racines, mais comme P est de degré 2, on déduit que $P = 0$. Ainsi, Φ est un isomorphisme, et comme $\dim E = 3 = \dim \mathbb{R}^3$ l'application Φ est alors un isomorphisme.
3. On note $P_0 = (X-1)(X-2), P_1 = X(X-2), P_2 = X(X-1)$. Montrons que la famille est libre. Soient donc λ, μ, ν trois réels tels que

$$\lambda P_0 + \mu P_1 + \nu P_2 = 0,$$

alors en évaluant en $X = 1$, puis $X = 2$ et $X = 0$, on trouve successivement $\lambda = \mu = \nu = 0$, donc la famille est libre. Elle est en plus de cardinal 3, donc est une base de E .

Pour la matrice de passage, il s'agit d'exprimer les polynômes P_0, P_1, P_2 en fonction de $1, X, X^2$.

$$P_0 = (X-1)(X-2) = X^2 - 3X + 2,$$

$$P_1 = X(X-2) = X^2 - 2X + 0.1$$

$$P_2 = X(X-1) = X^2 - X + 0.1.$$

Donc si on note P la matrice de passage de la base canonique vers \mathcal{B} , on a

$$P = \begin{pmatrix} 1 & 1 & 1 \\ -3 & -2 & -1 \\ 2 & 0 & 0 \end{pmatrix}.$$

Après application de la méthode du miroir, on trouve que

$$P^{-1} = \begin{pmatrix} -1 & \frac{-1}{2} & \frac{1}{4} \\ 1 & 0 & \frac{-1}{2} \\ 1 & \frac{1}{2} & \frac{1}{4} \end{pmatrix}.$$

4. On a $\Phi(P_0) = (2, 0, 0)$, $\Phi(P_1) = (0, -1, 0)$, $\Phi(P_2) = (0, 0, 2)$, d'où :

$$\text{Mat}_{\mathcal{B}, \mathcal{B}^{\text{can}}}(\Phi) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

puis d'après la formule de changement de base :

$$\begin{aligned} \text{Mat}_{\mathcal{B}^{\text{can}}, \mathcal{B}^{\text{can}}}(\Phi) &= P^{-1} \text{Mat}_{\mathcal{B}, \mathcal{B}^{\text{can}}}(\Phi) P \\ &= \begin{pmatrix} -2 & \frac{1}{2} & \frac{1}{2} \\ 2 & 0 & -1 \\ 2 & -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ -3 & -2 & -1 \\ 2 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} -\frac{5}{2} & -3 & -\frac{5}{2} \\ 0 & 2 & 2 \\ \frac{9}{2} & 3 & \frac{5}{2} \end{pmatrix} \end{aligned}$$

Exercice ALG.4.17 | Changement de base hyperbolique On note

$$f : x \mapsto e^x, g : x \mapsto e^{-x}, \text{ch} : x \mapsto \frac{e^x + e^{-x}}{2}, \text{sh} : x \mapsto \frac{e^x - e^{-x}}{2}.$$

1. Montrer que $\mathcal{B} = (f, g)$ et $\mathcal{B}' = (\text{ch}, \text{sh})$ sont des familles libres, puis que

$$\text{Vect}(f, g) = \text{Vect}(\text{ch}, \text{sh}).$$

On notera $F = \text{Vect}(f, g)$ dans la suite.

- Montrer que la dérivation $D : f \in F \mapsto f'$ définit un endomorphisme de F . Déterminer la matrice de D dans les bases $\mathcal{B}, \mathcal{B}'$ d'une part, et $\mathcal{B}', \mathcal{B}$ d'autre part.
- Retrouver le résultat de la question précédente à l'aide d'un changement de base.

Solution (exercice ALG.4.17)

1. Montrons que $\mathcal{B} = (f, g)$ et $\mathcal{B}' = (\text{ch}, \text{sh})$ sont des familles libres. Soient λ, μ deux réels tels que

$$\forall x \in \mathbf{R}, \quad \lambda e^x + \mu e^{-x} = 0,$$

alors en divisant par e^x puis en faisant tendre x vers $+\infty$, on déduit que $\lambda = 0$, donc pour tout $x \in \mathbf{R}$, $\mu e^{-x} = 0$ ce qui donne $\mu = 0$ en multipliant par e^x . Donc (f, g) est libre. Soient λ, μ deux réels tels que

$$\forall x \in \mathbf{R}, \quad \lambda \text{ch} x + \mu \text{sh} x = 0,$$

cela donne encore

$$\forall x \in \mathbf{R}, \quad \frac{1}{2}(\lambda + \mu)e^x + \frac{1}{2}(\lambda - \mu)e^{-x} = 0,$$

or nous avons déjà montré que la famille (f, g) est libre, on obtient alors :

$$\frac{1}{2}(\lambda + \mu) = 0, \quad \frac{1}{2}(\lambda - \mu) = 0,$$

d'où l'on tire facilement $\lambda = \mu = 0$. Donc (ch, sh) est libre. On a clairement $\text{Vect}(\text{ch}, \text{sh}) \subset \text{Vect}(f, g)$ car toute combinaison linéaire de ch, sh est une combinaison linéaire de f, g . Montrons à présent que $\text{Vect}(f, g) \subset \text{Vect}(\text{ch}, \text{sh})$. Nous avons

$$\begin{cases} \text{ch} = \frac{f+g}{2}, \\ \text{sh} = \frac{f-g}{2} \end{cases} \iff \begin{cases} 2 \text{ch} = f + g, \\ 2 \text{sh} = f - g, \end{cases}$$

on obtient alors $f = \text{ch} + \text{sh}$, et $g = \text{ch} - \text{sh}$ donc toute combinaison linéaire de f, g est combinaison linéaire de ch, sh i.e. $\text{Vect}(f, g) \subset \text{Vect}(\text{ch}, \text{sh})$. En conclusion

$$F = \text{Vect}(f, g) = \text{Vect}(\text{ch}, \text{sh}).$$

On notera $F = \text{Vect}(f, g)$ dans la suite.

2. Notons $D : f \in F \mapsto f'$. Alors D est une application linéaire par linéarité de la dérivation. On a $D(f) = f$ et $D(g) = -g$, donc pour tout $\lambda, \mu \in \mathbf{R}$

$$D(\lambda f + \mu g) = \lambda f - \mu g \in F,$$

ceci prouve que D est un endomorphisme de F . D'après les calculs de dérivées précédents,

$$\mathcal{M}_{\mathcal{B}}(\text{D}) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

par ailleurs, par linéarité, $D(\text{ch}) = \frac{1}{2}(D(f) + D(g)) = \frac{1}{2}(f - g) = \text{sh}$ et $D(\text{sh}) = \frac{1}{2}(D(f) - D(g)) = \frac{1}{2}(f + g) = \text{ch}$, donc

$$\mathcal{M}_{\mathcal{B}'}(\text{D}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

3. Notons P la matrice de passage de \mathcal{B} vers \mathcal{B}' . Alors, comme $\text{ch} = \frac{1}{2}f + \frac{1}{2}g$ et $\text{sh} = \frac{1}{2}f - \frac{1}{2}g$, il vient

$$P = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

et $\mathcal{M}_{\mathcal{B}}(\text{D}) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Donc

$$\mathcal{M}_{\mathcal{B}'}(\text{D}) = P^{-1} \mathcal{M}_{\mathcal{B}}(\text{D}) P = 2 \begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

en effectuant le produit matriciel on retrouve

$$\mathcal{M}_{\mathcal{B}'}(\text{D}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

de la question précédente.

Chapitre ALG.5.

Diagonalisation

Résumé & Plan

Dans beaucoup de contextes, nous pourrions avoir besoin de calculer par exemple les puissances d'une matrice. Nous allons voir dans ce chapitre que si la matrice considérée est semblable à une matrice diagonale, alors un calcul explicite peut être mis en place. Avant de commencer la lecture de ce chapitre, revoyez la méthode vue dans le [Chapter ALG.4](#) pour les calculs de valeurs propres d'une matrice, ainsi que les formules de changement de base.

1	Éléments propres	2
1.1	Pour un endomorphisme	2
1.2	Pour une matrice	3
1.3	Calcul effectif en dimension finie	6
1.4	Utilisation d'un polynôme annulateur	9
1.5	Et en Python?	10
2	Critère de diagonalisation	11
2.1	Familles de vecteurs propres	11
2.2	Diagonalisation	12

3	Premières applications	18
3.1	Calculs des puissances d'une matrice	18
3.2	Calculs de commutants	19
4	Exercices	21
4.1	Calculs d'éléments propres	21
4.2	Pour des matrices	21
4.3	Pour des endomorphismes	26
4.4	Polynôme annulateur	27

Les Mathématiques sont un jeu que l'on exerce selon des règles simples en manipulant des symboles ou des concepts qui n'ont en soi, aucune importance particulière.

— David Hilbert



Cadre

Dans tout le chapitre,

- ▶ **K désignera \mathbb{R} ou \mathbb{C} ,**
- ▶ **E est un espace vectoriel sur K .**
- ▶ **sauf mention du contraire, f désignera un endomorphisme de E , et A est**



une matrice de $\mathcal{M}_n(\mathbf{K})$.

Supposons dans cette introduction que E est de dimension finie $n \in \mathbf{N}$. L'objectif sera de trouver une base dans laquelle la matrice de f est « simple », *i.e.* soit diagonale soit triangulaire. En BCPST, vous devrez savoir répondre à la première question.

DIAGONALISER : POURQUOI ? La première application est le calcul des puissances d'une matrice ou d'un endomorphisme (qui elle-même en induit beaucoup d'autres). Par exemple,

1. supposons que nous ayons trouvé une base \mathcal{B} telle que : $\mathcal{M}_{\mathcal{B}}(f) = D$ avec D diagonale. Alors $\mathcal{M}_{\mathcal{B}}(f^n) = D^n$, et D^n se calcule simplement puisqu'elle est diagonale !
2. Nous pouvons aussi adopter un point de vue purement matriciel. Soit $A \in \mathcal{M}_n(\mathbf{K})$, alors notons

$$f \in \mathcal{L}(\mathcal{M}_{n,1}(\mathbf{K}), \mathcal{M}_{n,1}(\mathbf{K}))$$

l'endomorphisme canoniquement associé à A, *i.e.* tel que $\mathcal{M}_{\mathcal{B}^{\text{can}}}(f) = A$ où \mathcal{B}^{can} désigne la base canonique de $\mathcal{M}_{n,1}(\mathbf{K})$. S'il existe une base \mathcal{B} telle que $\mathcal{M}_{\mathcal{B}}(f) = D$ avec D diagonale, alors il existe une matrice inversible $P \in \text{GL}_n(\mathbf{K})$ ¹ telle que :

$$A = PDP^{-1} \implies \forall n \in \mathbf{N}, A^n = PD^nP^{-1}.$$

Donc A est semblable à une matrice diagonale. Nous verrons que le calcul des puissances d'une matrice est riche d'applications.

Étant donnée une matrice, tout l'enjeu est donc de proposer dans ce chapitre une méthode pour trouver les matrices P et D indiquées *supra*. Notons que si $\mathcal{B} = (e_1, \dots, e_n)$ avec $e_i \in E$ pour tout i , alors nécessairement, puisque $\mathcal{M}_{\mathcal{B}}(f) = D$ est diagonale, les e_i vérifient :

$$f(e_i) = \lambda_i e_i, \quad \text{avec } \lambda_i \in \mathbf{K} \quad \text{pour tout } i.$$

Nous dirons que e_i un *vecteur propre* de f (ou A) associé à la *valeur propre* λ_i . Notons également que les e_i sont nécessairement non nuls, puisque \mathcal{B} est une base de E.

¹qui n'est autre que la matrice de passage de \mathcal{B}^{can} à \mathcal{B} .

1. ÉLÉMENTS PROPRES

Dans la définition qui suit, E ne sera pas forcément de dimension finie. On donne dans cette section la définition générale d'un élément propre (valeur ou vecteur), vous noterez que la définition pour les matrices avait déjà été donnée dans un chapitre précédent.

1.1. Pour un endomorphisme

Définition ALG.5.1 | Éléments propres d'un endomorphisme

Soit $f \in \mathcal{L}(E)$.

- ▶ On dit que $\lambda \in \mathbf{K}$ est une *valeur propre* de f si :

$$\exists x \in E \setminus \{0_E\}, f(x) = \lambda x.$$

Un tel vecteur x est appelé *vecteur propre* de f associé à la valeur propre λ .

- ▶ L'ensemble des valeurs propres de f est appelé *spectre* de f , noté $\text{Spec } f$, *i.e.*

$$\text{Spec } f = \{\lambda \in \mathbf{K}, \exists x \in E \setminus \{0_E\}, f(x) = \lambda x\}.$$

Si $\mathbf{K} = \mathbf{C}$, on appelle *spectre réel* de f , noté $\text{Spec}_{\mathbf{R}}(f)$, l'ensemble des valeurs propres réelles de f , *i.e.*

$$\text{Spec}_{\mathbf{R}} f = \{\lambda \in \mathbf{R}, \exists x \in E \setminus \{0_E\}, f(x) = \lambda x\}.$$

- ▶ On appelle *espace propre* associé à la valeur propre $\lambda \in \mathbf{K}$ l'ensemble

$$E_{\lambda}(f) = \ker(f - \lambda \text{Id}_E) = \{x \in E, f(x) = \lambda x\}.$$



Attention

Un vecteur propre est non nul, par définition.

Proposition ALG.5.1 | Ensemble des vecteurs propres

- ▶ L'ensemble des vecteurs propres de f associés à la valeur propre $\lambda \in \text{Spec } f$ est :

$$E_\lambda(f) \setminus \{0_E\}.$$

- ▶ Si λ n'est pas une valeur propre, alors $E_\lambda(f)$ est réduit au vecteur nul.

Preuve Immédiat à l'aide des définitions.

Remarque 1.1 – À propos de \mathbf{K}

1. *A priori* on peut avoir $\text{Spec}_{\mathbf{R}}(f) = \emptyset$, nous verrons en revanche plus tard que $\text{Spec}_{\mathbf{C}}(f)$ est systématiquement non vide.
2. Lorsque $\mathbf{K} = \mathbf{R}$, on a par définition $\text{Spec}_{\mathbf{R}}(f) = \text{Spec}(f)$.

Étant donné que $E_\lambda(f)$ est le noyau d'un certain endomorphisme, c'est en particulier un sous-espace vectoriel de E . Les résultats du **Chapter ALG.3** nous permettent alors d'aboutir à la caractérisation suivante.

Proposition ALG.5.2 | Caractérisation des éléments propres d'un endomorphisme

Supposons que E est de dimension finie n , et soit $\lambda \in \mathbf{K}$.

1. $E_\lambda(f)$ est un sous-espace vectoriel de E ,
2. **(Caractérisation des éléments propres)**

λ est valeur propre de f

$$\begin{aligned} \iff \ker(f - \lambda \text{Id}_E) \neq \{0_E\} & \iff \dim \ker(f - \lambda \text{Id}_E) \geq 1 \\ \iff f - \lambda \text{Id}_E \text{ non injective} & \iff f - \lambda \text{Id}_E \text{ non bijective} \\ \iff f - \lambda \text{Id}_E \text{ non surjective} & \\ \iff \text{Rg}(f - \lambda \text{Id}_E) < n. & \end{aligned}$$

Preuve (Point clef — *En dimension finie, pour des endomorphismes : « surjectif \iff injectif \iff bijectif »*)

²à ne pas confondre avec l'espace propre

1. $E_\lambda(f)$ est le noyau d'une application linéaire, c'est donc un sous-espace vectoriel de E d'après le **Chapter ALG.3**.



1.2. Pour une matrice

Définition ALG.5.2 | Éléments propres d'une matrice

Soit $A \in \mathfrak{M}_n(\mathbf{R})$ avec $n \geq 1$.

- ▶ On dit que $\lambda \in \mathbf{K}$ est une *valeur propre* de A si :

$$\exists X \in \mathfrak{M}_{n,1}(\mathbf{K}) \setminus \{0_E\}, AX = \lambda X.$$

Un tel vecteur X est appelé *vecteur propre* de A associé à la valeur propre λ .

- ▶ L'ensemble des valeurs propres de A est appelé *spectre* de A , noté $\text{Spec } A$, *i.e.*

$$\text{Spec } A = \left\{ \lambda \in \mathbf{K}, \exists X \in \mathfrak{M}_{n,1}(\mathbf{K}) \setminus \{0_{\mathfrak{M}_{n,1}(\mathbf{K})}\}, AX = \lambda X \right\}.$$

Si $\mathbf{K} = \mathbf{C}$, on appelle *spectre réel* de A , noté $\text{Spec}_{\mathbf{R}}(A)$, l'ensemble des valeurs propres réelles de A , *i.e.*

$$\text{Spec}_{\mathbf{R}} A = \left\{ \lambda \in \mathbf{R}, \exists X \in \mathfrak{M}_{n,1}(\mathbf{K}) \setminus \{0_{\mathfrak{M}_{n,1}(\mathbf{K})}\}, AX = \lambda X \right\}.$$

► On appelle *espace propre associé à la valeur propre* $\lambda \in \mathbf{K}$ l'ensemble

$$E_\lambda(A) = \ker(A - \lambda I_n) = \{X \in \mathfrak{M}_{n,1}(\mathbf{K}), AX = \lambda X\}.$$

Proposition ALG.5.3 | Caractérisation des éléments propres d'une matrice

Soit $\lambda \in \mathbf{K}$.

1. $E_\lambda(A)$ est un sous-espace vectoriel de E ,
2. **(Caractérisation des éléments propres)**

$$\begin{aligned} \lambda \text{ est valeur propre de } A &\iff \text{le système } (A - \lambda I_n)X = 0, \\ &\text{n'est pas de CRAMER, en } X \in \mathfrak{M}_{n,1}(\mathbf{K}), \\ &\iff A - \lambda I_n \text{ non inversible,} \\ &\iff \text{Rg}(A - \lambda I_n) < n. \end{aligned}$$

Rappelons que l'assertion

«le système $(A - \lambda I_n)X = 0$, en $X \in \mathfrak{M}_{n,1}(\mathbf{K})$, n'est pas de CRAMER

signifie que le système associé admet une solution non nulle.

Preuve Découle directement de la proposition analogue pour les endomorphismes, en considérant l'endomorphisme canoniquement associé à A .

Proposition ALG.5.4 | Lien entre les deux définitions

Supposons qu'il existe une base \mathcal{B} de E , telle que $A = \mathcal{M}_{\mathcal{B}}(f)$. Alors :

1. $\lambda \in \text{Spec}(f) \iff \lambda \in \text{Spec}(A)$.
2. Soit alors $x \in E$ et $X = \mathcal{M}_{\mathcal{B}}(x)$, alors $x \in E_\lambda(f) \iff X \in E_\lambda(A)$.

L'espace vectoriel E est donc supposé de dimension finie ici dans la proposition précédente.

Preuve

1. Si $\lambda \in \text{Spec}(f)$, alors il existe $x \in E \setminus \{0_E\}$. En posant $X = \mathcal{M}_{\mathcal{B}}(x)$ on a $AX = \lambda X$ avec $X \neq 0$, donc $\lambda \in \text{Spec}(A)$. Inversement, si X est un vecteur propre non nul de A , on a $f(x) = \lambda x$ avec x l'unique vecteur de E tel que $X = \mathcal{M}_{\mathcal{B}}(x)$.

2. En observant la première partie, on constate que 2) a également été prouvée.

Proposition ALG.5.5 | Bijectivité & Valeur propre nulle

- $0 \in \text{Spec}(f) \iff f$ non bijective.
- $0 \in \text{Spec}(A) \iff A$ non inversible.

Preuve




Attention

Cette proposition est **très** importante dans la pratique : dès que le calcul des éléments propres est achevé, le problème d'inversibilité d'une matrice ou d'un endomorphisme devient trivial.

Pour les endomorphismes en dimension finie, *i.e.* les matrices, nous avons déjà vu comment calculer les valeurs propres dans le **Chapter ALG.4**. Voyons deux premiers exemples avec des fonctions et suites.

Exemple 1 – Sur un espace de fonctions Soit $E = \mathcal{C}^\infty(\mathbf{R})$ et D l'endomorphisme de E qui à f associe f' . Déterminer les valeurs propres de D et les sous-espaces propres associés.

Exemple 2 – Sur un espace de suites Soit $E = \mathbf{R}^{\mathbf{N}}$ et D l'endomorphisme de E qui à $(u_n)_{n \in \mathbf{N}}$ associe $(u_{n+1})_{n \in \mathbf{N}}$. Déterminer les valeurs propres de D et les sous-espaces propres associés. 

2. 3. 

Proposition ALG.5.6 | Propriétés des éléments propres

Soient $A, B \in \mathfrak{M}_n(\mathbf{K})$. Alors :

1. **(Translation)** si $a \in \mathbf{K}$, $\lambda \in \text{Spec} A \iff \lambda + a \in \text{Spec}(A + aI_n)$.
2. **(Transposition)** $\text{Spec}(A) = \text{Spec}({}^t A)$.
3. **(Similitude)** Si A est semblable à B , i.e. s'il existe P inversible de format $n \times n$ telle que : $A = PBP^{-1}$, alors $\text{Spec}(A) = \text{Spec}(B)$.

Preuve

1. 

1.3. Calcul effectif en dimension finie

Nous supposons dans cette section que E est dimension finie, auquel cas nous pouvons utiliser des matrices pour trouver les éléments propres.

Pour vérifier les conditions de la **Proposition ALG.5.2**, on a recours à une matrice associée dans une certaine base.

CAS PARTICULIER DE LA DIMENSION TROIS : ÉCHELONNEMENT. Il s'agit donc d'appliquer l'algorithme du pivot de Gauß à la matrice $A - \lambda I_n$. Rappelons la méthode ci-dessous vue dans le **Chapter ALG.4**. On suppose ici que A est de format 3×3 .

Méthode Opérations pour la recherche d'éléments propres

- ▶ Si le coefficient $(3, 1)$ n'est pas nul :
 1. l'opération optimale à effectuer en premier pour des matrices de taille 3×3 est la permutation $L_1 \longleftrightarrow L_3$ qui permettra d'obtenir un coefficient indépendant de λ en position $(1, 1)$. On élimine alors avec celui-ci les coefficients $(2, 1), (3, 1)$.
 2. Ensuite, en position $(2, 2)$, nous avons un coefficient affine en λ et que l'on souhaite utiliser en nouveau pivot afin d'éliminer le coefficient $(3, 2)$. Pour éliminer λ en $(2, 2)$, on peut faire une opération simple en fonction de L_3 .
 3. Un pivot indépendant de λ est alors obtenu en $(2, 2)$, on peut alors éliminer le coefficient $(3, 2)$.
- ▶ Si le coefficient $(3, 1)$ est nul :
 1. on fait la permutation $L_1 \longleftrightarrow L_2$ qui permettra d'obtenir un coefficient indépendant de λ en position $(1, 1)$. On élimine alors avec celui-ci le coefficient $(3, 1)$.
 2. En positions $(2, 2)$, nous avons un coefficient indépendant de λ qui sert à éliminer le coefficient $(3, 2)$.

CAS PARTICULIER DE LA DIMENSION DEUX : ÉCHELONNEMENT OU DÉTERMINANT. En dimension deux, l'inversibilité peut être reformulée à l'aide du déterminant. On peut aussi utiliser la méthode du pivot de Gauß, adaptée à la dimension 2.

Méthode Éléments propres en dimension 2 l'aide du déterminant

Soit $A \in \mathcal{M}_{2,2}(\mathbf{K})$. Alors :

$$\lambda \in \text{Spec}(A) \iff \det(A - \lambda I_2) = 0.$$

Ainsi,

1. il suffit de résoudre en λ l'équation du second degré $\det(A - \lambda I_2) = 0$.
2. On calcule ensuite $E_\lambda(A) = \ker(A - \lambda I_2)$ pour chaque solution trouvée précédemment.

Cette méthode a l'avantage de donner rapidement les valeurs propres (elles sont racines d'un polynôme de degré deux). En revanche, vous devez calculer les espaces propres à partir de la matrice de départ (non échelonnée). Mieux vaut donc utiliser, même en dimension deux, la méthode d'échelonnement lorsque l'on vous demande également les vecteurs propres.

Méthode Éléments propres en dimension 2 à l'aide du pivot de Gauß

- ▶ Si le coefficient $(2, 1)$ est nul : la matrice $A - \lambda I_2$ est triangulaire, donc c'est terminé.
- ▶ Sinon, on fait la permutation $L_1 \longleftrightarrow L_2$ qui permettra d'obtenir un coefficient indépendant de λ en position $(1, 1)$. On élimine alors avec celui-ci le coefficient $(2, 1)$. La matrice $A - \lambda I_2$ est alors elle aussi échelonnée.

Exemple 3 — Trouver les éléments propres de $A = \begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix}$.

1. (1ère méthode : avec le déterminant.) 

2. (2ème méthode : par échelonnement.)



CALCUL DES ÉLÉMENTS PROPRES EN UTILISANT LA DÉFINITION. Lorsqu'une matrice possède beaucoup de zéros, et qu'elle est de taille importante, il peut être plus judicieux de passer par la définition d'un vecteur propre pour les calculer, plutôt que le pivot.



Méthode Calcul d'éléments propres d'une avec la définition

Soit $A \in \mathcal{M}_{n,n}(\mathbf{K})$. Alors pour trouver les éléments propres de A , on peut, si elle possède beaucoup de zéros :

- ▶ considérer le système $AX = \lambda X$,
- ▶ le résoudre en les $(\lambda, X) \in \mathbf{K} \times (\mathcal{M}_{n,1}(\mathbf{K}) \setminus \{0_{\mathcal{M}_{n,1}(\mathbf{K})}\})$.
- ▶ Conclure.

Voyons un exemple.³

Exemple 4 — Notons $M = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & 0 \\ 1 & 0 & \dots & \dots & 0 \end{pmatrix} \in \mathcal{M}_{n,n}(\mathbf{R})$ avec $n \geq 2$. Déterminons ses éléments propres. La matrice possède énormément de zéros, il est donc préférable de considérer le système $(A - \lambda I_3)X = 0$ plutôt qu'un algorithme d'échelonnement.

Soit $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ un vecteur et λ un complexe. On a :

$$D_p X = \lambda X \iff \begin{cases} x_2 = \lambda x_1 \\ x_3 = \lambda x_2 \\ \vdots \\ x_p = \lambda x_{p-1} \\ x_1 = \lambda x_p \end{cases} \iff \begin{cases} x_2 = \lambda x_1 \\ x_3 = \lambda^2 x_1 \\ \vdots \\ x_p = \lambda^{p-1} x_1 \\ x_1 = \lambda^p x_1 \end{cases}$$

Le système est alors équivalent à


$$X = 0 \text{ ou } \lambda^p = 1 \text{ et } X = \begin{pmatrix} 1 \\ \lambda \\ \lambda^2 \\ \vdots \\ \lambda^{p-1} x_1 \end{pmatrix}$$

Les valeurs propres sont donc les complexes λ tels que $\lambda^p = 1$, donc les p -ièmes de l'unité (cf. [Chapter ALG.1](#)) qui sont au nombre de p et sont distinctes. Et l'espace propre E_λ

associé à un tel λ est de dimension 1 et engendré par le vecteur $\begin{pmatrix} 1 \\ \lambda \\ \lambda^2 \\ \vdots \\ \lambda^{p-1} \end{pmatrix}$.

³Extrait de l'épreuve de Calculs & Raisonnements 2020

CAS D'UNE MATRICE TRIANGULAIRE. Puisqu'une matrice triangulaire est déjà échelonnée, il est facile de trouver ses éléments propres.

► (avec matrice) 

Proposition ALG.5.7 | Les valeurs propres d'une matrice triangulaire se lisent sur la diagonale

$$\text{Si } T = \begin{pmatrix} \star & \dots & \star \\ 0 & \dots & \star \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \star \end{pmatrix} \in \mathfrak{M}_{n,n}(\mathbf{K}), \quad \text{alors : } \text{Spec } T = \{\lambda_1, \dots, \lambda_n\}.$$

Preuve



► (sans matrice) 

POUR UN ENDOMORPHISME. Il y a dans ce cas deux méthodes :

1. calculer sa matrice, puis appliquer les méthodes précédentes.
2. Travailler directement avec l'expression analytique de l'endomorphisme.


Exemple 5 — Dans $E = \mathbf{R}_3[X]$, on considère l'application φ définie par : $\varphi(P) = P + P'$. On vérifie sans difficulté qu'il s'agit d'un endomorphisme de E . Déterminons les éléments propres de φ , avec une matrice puis sans matrice.

Le théorème suivant est admis dans le cas général, nous le prouvons pour $n = 2$.


Théorème ALG.5.1 | Existence de valeurs propres

- Soit $A \in \mathfrak{M}_n(\mathbf{C})$, alors : $\text{Spec}(A) \neq \emptyset$. En d'autres termes, toute matrice complexe possède au moins une valeur propre complexe.
- Si $A \in \mathfrak{M}_n(\mathbf{R})$, alors : A possède au moins une valeur propre complexe.

Preuve (Point clef — *Tout polynôme possède au moins une racine dans \mathbf{C}*)

Supposons que n est égale à 2. 

On peut cependant établir l'existence dans certains cas très particuliers de matrices, par exemple pour les matrices symétriques réelles. Nous verrons ce résultat dans un prochain chapitre.

Exemple 6 – Trouver les valeurs propres réelles et complexes de $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. 

On termine cette sous-section de calculs effectifs par une méthode efficace lorsque les sommes sur les lignes d'une matrice sont constantes.



Méthode Valeur propre et somme sur chaque ligne constante


Lorsque la somme des coefficients sur les lignes d'une matrice est constante égale à $\lambda \in \mathbf{K}$,


le vecteur $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ est un vecteur propre de valeur propre associée λ .

C'est une simple vérification. Notons $A = (a_{i,j})_{1 \leq i,j \leq n}$ une telle matrice, et notons λ la constante commune de la somme des coefficients de chaque ligne. On a alors

$$A \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1,j} \\ \vdots \\ \sum_{j=1}^n a_{n,j} \end{pmatrix} = \begin{pmatrix} \lambda \\ \vdots \\ \lambda \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Exemple 7 –

► 4 est une valeur propre de $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$, puisque 

► 4 est une valeur propre de $A = \begin{pmatrix} 2 & 4 & -2 \\ 4 & -1 & 1 \\ 4 & 1 & -1 \end{pmatrix}$, puisque 


1.4.


Utilisation d'un polynôme annulateur

Afin de simplifier la présentation, nous traitons uniquement le cas matriciel, même si la notion de polynôme d'endomorphisme existe aussi. Nous avons déjà vu à quoi pouvait servir un polynôme annulateur pour trouver l'inverse d'une matrice, voyons en quoi il peut nous aider à déterminer les éléments propres sur un exemple.

Exemple 8 – Fondamental On considère une matrice $A \in \mathfrak{M}_{n,n}(\mathbf{K})$ telle que :


$$A^2 = A.$$

1. Soit $X \in \mathfrak{M}_{n,1}(\mathbf{K})$ un vecteur propre associé à la valeur propre $\lambda \in \mathbf{K}$. Montrer que $\lambda^2 = \lambda$. 

2. En déduire les valeurs propres possibles pour la matrice A. 


Exemple 9 – Cas d'une matrice nilpotente On considère une matrice $A \in \mathfrak{M}_{n,n}(\mathbf{K})$ nilpotente, *i.e.* il existe $p \in \mathbf{N}$ tel que :

$$A^p = 0.$$

Déterminer les valeurs propres possibles pour la matrice A. 

Exemple 10 – Cas d'une matrice cyclique On considère une matrice $A \in \mathfrak{M}_{n,n}(\mathbf{K})$ cyclique, *i.e.* il existe $p \in \mathbf{N}$ tel que :

$$A^p = I_n.$$

Déterminer les valeurs propres possibles pour la matrice A. 

Remarque 1.2 – Cas général [H.P] Les conditions précédentes $A^2 = A$ et $A^p = 0$ sont des conditions de la forme « $P(A) = 0$ » où P est un polynôme.⁴ On a finalement montré dans ces deux exemples que les valeurs propres sont à chercher parmi les racines de P (celles de $X^2 - X = X(X - 1)$ pour le premier, et X^p pour le second).

1.5. Et en Python ?

Il faut garder à l'esprit que les commandes qui suivent fourniront des valeurs approchées des éléments propres d'une matrice : des algorithmes existent pour les obtenir et c'est précisément ceux-là que le module numpy utilise pour les renvoyer.

Calculs d'éléments propres avec le module numpy

```
>>> import numpy as np
>>> A = np.array([[1,0],[0,2]])
>>>
>>> # POUR UNE MATRICE QUELCONQUE
>>> # renvoie les éléments propres de C
>>> # d'abord les vp ensuite les vecp :
>>> vap, vep = np.linalg.eig(A)
>>> vap
array([1., 2.])
>>> vep
array([[1., 0.],
       [0., 1.]])
>>> # pour avoir le premier vecteur propre, on extrait la première
- colonne
```

⁴Dans le premier cas, $P = X^2 - X$ et dans le second $P = X^p$, mais encore faudrait-il définir la notation $P(A)$ de manière générale.

```

>>> vep[:,0]
array([1., 0.])
>>> for i in range(len(vap)):
...     X = vep[:,i]
...     print(A@X-vap[i]*X)
...
[0. 0.]
[0. 0.]
>>> # on obtient bien que des vecteurs proches du vecteur nul

```



Remarque 1.3 — Vous remarquerez que tout est compté avec multiplicité *i.e.* si un espace propre est de dimension 2, la valeur propre associée sera répétée deux fois dans le résultat : ce sont des listes ayant autant d'éléments que de colonnes (ou lignes) dans la matrice initiale.


2. CRITÈRE DE DIAGONALISATION

Nous avons vu dans la précédente sous-section comment déterminer les éléments propres d'un endomorphisme, à présent on énonce les différents résultats qui permettent de conclure quant à la diagonalisation, ou la non-diagonalisation. Rappelons que l'enjeu est de trouver une base de vecteurs propres de E (ou $\mathcal{M}_{n,1}(\mathbf{R})$ pour les matrices). Nous verrons tout d'abord que la liberté est automatique, il suffira ensuite de se demander si la famille obtenue est génératrice.

2.1. Familles de vecteurs propres

Proposition ALG.5.8 | Liberté d'une famille propre

Soient $e_1, \dots, e_p \in E$ une famille de p vecteurs propres de $f \in \mathcal{L}(E)$ (ou $A \in \mathcal{M}_n(\mathbf{K})$), associés à des valeurs propres **distinctes** $\lambda_1, \dots, \lambda_p$, alors :
la famille (e_1, \dots, e_p) est libre.

Preuve Nous faisons la preuve dans le cas d'un endomorphisme et pour $p = 2$, le résultat se déduit alors automatiquement pour les matrices en utilisant l'endomorphisme canoniquement associé et le cas p général se traite par récurrence. 

Corollaire ALG.5.1 | Nombre maximal de valeurs propres

- ▶ Tout endomorphisme $f \in \mathcal{L}(E)$ d'un espace vectoriel E de dimension n possède au plus n valeurs propres distinctes.
- ▶ Toute matrice de $A \in \mathcal{M}_n(\mathbf{K})$ admet au plus n valeurs propres distinctes.

Preuve



Proposition ALG.5.9 | Concaténation de familles propres

Soit $f \in \mathcal{L}(E)$ (ou $A \in \mathcal{M}_n(\mathbf{K})$) un endomorphisme d'un espace vectoriel E (ou une matrice carrée) admettant pour valeurs propres $\lambda_1, \dots, \lambda_p \in \mathbf{K}$. Soient $\mathcal{B}_1, \dots, \mathcal{B}_p$ des bases de respectivement $E_{\lambda_1}, \dots, E_{\lambda_p}$.

Alors :
la famille de vecteurs obtenue en mettant bout à bout les vecteurs des bases
 $\mathcal{B}_1, \dots, \mathcal{B}_p$ est une famille libre.

Preuve Notons $\mathcal{B}_i = (e_1^{(i)}, \dots, e_{n_i}^{(i)})$, avec $e_k^{(i)} \in E$ pour tout $i \in \llbracket 1, p \rrbracket$, $k \in \llbracket 1, n_i \rrbracket$ et $n_i = \dim E_i$. Soient $\lambda_1^{(i)}, \dots, \lambda_{n_i}^{(i)} \in \mathbf{K}$ pour tout $i \in \llbracket 1, p \rrbracket$ tels que :

$$(\lambda_1^{(1)} e_1^{(1)} + \dots + \lambda_{n_1}^{(1)} e_{n_1}^{(1)}) + \dots + (\lambda_1^{(p)} e_1^{(p)} + \dots + \lambda_{n_p}^{(p)} e_{n_p}^{(p)}) = 0_E.$$

Montrons que $\forall i \in \llbracket 1, p \rrbracket : \lambda_1^{(i)} = 0, \dots, \lambda_{n_i}^{(i)} = 0$.

Puisque les espaces propres sont des sous-espaces vectoriels, nous avons :

$$\forall i \in \llbracket 1, p \rrbracket, f_i = \lambda_1^{(i)} e_1^{(i)} + \dots + \lambda_{n_i}^{(i)} e_{n_i}^{(i)} \in E_{\lambda_i},$$

(défi.)

car ce sont des combinaisons linéaires d'éléments de E_{λ_i} , qui est un espace vectoriel, et

$$f_1 + \dots + f_p = 0_E,$$

donc (f_1, \dots, f_p) est liée. Or chacun des f_i est soit nul, soit un vecteur propre associé à la valeur propre λ_i . Or, nous avons montré que toute famille de vecteurs propres associés à des valeurs propres distinctes est libre, donc tous les f_i sont forcément nuls, *i.e.*

$$\forall i \in \llbracket 1, p \rrbracket, \lambda_1^{(i)} e_1^{(i)} + \dots + \lambda_{n_i}^{(i)} e_{n_i}^{(i)} = 0_E.$$

Mais $(e_1^{(i)}, \dots, e_{n_i}^{(i)})$ est une base donc une famille libre pour tout $i \in \llbracket 1, p \rrbracket$, donc $\lambda_1^{(i)} = 0, \dots, \lambda_{n_i}^{(i)} = 0$ et ce pour tout $i \in \llbracket 1, p \rrbracket$.

Corollaire ALG.5.2 | Majoration sur la somme des dimensions des espaces propres

Soit $f \in \mathcal{L}(E)$ (ou $A \in \mathfrak{M}_n(\mathbf{K})$) un endomorphisme d'un espace vectoriel E (ou une matrice carrée) admettant pour valeurs propres $\lambda_1, \dots, \lambda_p \in \mathbf{K}$. On suppose de plus que E est de dimension finie. Soient $\mathcal{B}_1, \dots, \mathcal{B}_p$ des bases de respectivement

$E_{\lambda_1}, \dots, E_{\lambda_p}$. Alors :

$$\sum_{i=1}^p \dim(E_{\lambda_i}) \leq \dim(E).$$

Preuve



2.2. Diagonalisation

Dans cette section, l'espace vectoriel E est de **de dimension finie** $n \in \mathbf{N}$ pour faire sens à la notion de diagonalisabilité.

Définition ALG.5.3 | Diagonalisabilité

- ▶ Soit $f \in \mathcal{L}(E)$. On dit que f est *diagonalisable* lorsqu'il existe une base \mathcal{B} de E telle que

$$\underset{\mathcal{B}}{\text{Mat}}(f) \text{ soit diagonale.}$$

Diagonaliser un endomorphisme c'est déterminer une telle base.

- ▶ Soit $A \in \mathfrak{M}_n(\mathbf{K})$. On dit que A est *diagonalisable* dans $\mathfrak{M}_n(\mathbf{K})$ si elle est semblable à une matrice diagonale D de $\mathfrak{M}_n(\mathbf{K})$, *i.e.* s'il existe P inversible telle que :

$$A = PDP^{-1}, \text{ avec } D \text{ diagonale.}$$

Diagonaliser une matrice c'est déterminer les matrices D, P précédentes.

Si $\mathbf{K} = \mathbf{C}$, f (ou A) est *diagonalisable dans* $\mathfrak{M}_n(\mathbf{R})$ (ou *diagonalisable sur* \mathbf{R}) s'il existe D, P comme précédemment mais à coefficients réels.

Exemple 11 – Interpréter un changement de base en terme de diagonalisation Dans chaque cas, donner les valeurs propres et une base de chaque espace propre.

1. Soit $A \in \mathcal{M}_{3,3}(\mathbf{R})$ telle que $A = P \cdot D \cdot P^{-1}$ avec :

$$P = \begin{pmatrix} 1 & -1 & 1 \\ -2 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} -3 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$



2. Soit $A \in \mathcal{M}_{3,3}(\mathbf{R})$ telle que $A = P \cdot D \cdot P^{-1}$ avec :

$$P = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \\ 1 & 0 & -1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$



Théorème ALG.5.2 | Lien entre les deux définitions

Soit \mathcal{B} une base de E et $f \in \mathcal{L}(E)$, $A = \underset{\mathcal{B}}{\text{Mat}}(f)$. Alors :

A est diagonalisable $\iff f$ est diagonalisable.

En particulier,

A est diagonalisable, $\iff f^5 : X \in \mathcal{M}_{n,1}(\mathbf{K}) \mapsto AX$ est diagonalisable,

où $n = \dim E$.

Preuve



Proposition ALG.5.10 | Diagonalisabilité et inversion/transposition

Soit $A \in \mathcal{M}_n(\mathbf{K})$.

1. Alors :

$A \in \mathcal{M}_n(\mathbf{K})$ est diagonalisable $\iff {}^t A$ est diagonalisable.

2. Supposons A inversible. Alors :

$A \in \mathcal{M}_n(\mathbf{K})$ est diagonalisable $\iff A^{-1}$ est diagonalisable.

Preuve

1.

⁵i.e. l'endomorphisme canoniquement associé à A

2. **Théorème ALG.5.3 | Diagonalisation & Existence d'une base propre**

- Soit $f \in \mathcal{L}(E)$. Alors :
 f est diagonalisable \iff il existe une base de E formée de vecteurs propres de f .
- Soit $A \in \mathfrak{M}_n(\mathbf{K})$. Alors :
 A est diagonalisable \iff il existe une base de $\mathfrak{M}_{n,1}(\mathbf{K})$ formée de vecteurs propres de A .

Preuve1. 2. Notons f l'endomorphisme canoniquement associé à A .

$$\begin{aligned}
 A \text{ est diagonalisable} &\iff f \text{ est diagonalisable} \\
 &\iff f \text{ est diagonalisable} \\
 &\iff \text{il existe une base de } E \\
 &\quad \text{formée de vecteurs propres de } f \\
 &\iff \text{il existe une base de } E \\
 &\quad \text{formée de vecteurs propres de } A.
 \end{aligned}$$

Justifions la dernière équivalence : si (e_1, \dots, e_p) est une base de vecteurs propres de f associés aux valeurs propres $\lambda_1, \dots, \lambda_p$, alors en prenant la matrice associée on obtient des vecteurs propres pour A :

$$f(e_i) = \lambda_i e_i \iff AX_i = \lambda_i X_i, \quad \text{où } X_i = \underset{e_i}{\text{Mat}}(\mathcal{B}^{\text{can}}),$$

Il est facile de vérifier que (e_1, \dots, e_p) est une base de E si et seulement si (X_1, \dots, X_p) est une base de $\mathfrak{M}_{p,1}(\mathbf{K})$.

Passons sans plus tarder au critère principal de diagonalisation.

Théorème ALG.5.4 | Condition nécessaire et suffisante de diagonalisation

- Soit $f \in \mathcal{L}(E)$ un endomorphisme dont les valeurs propres sont $\lambda_1, \dots, \lambda_p$. Alors :

$$\sum_{i=1}^p \dim(E_{\lambda_i}(f)) = \dim(E) \iff f \text{ est diagonalisable.}$$

- Soit $A \in \mathfrak{M}_n(\mathbf{K})$ une matrice dont les valeurs propres sont $\lambda_1, \dots, \lambda_p$. Alors :

$$\sum_{i=1}^p \dim(E_{\lambda_i}(A)) = n \iff A \text{ est diagonalisable.}$$

Preuve



Corollaire ALG.5.3 | Condition suffisante de diagonalisation

Soit $f \in \mathcal{L}(E)$ (ou $A \in \mathcal{M}_n(\mathbf{K})$) tel que f (resp. A) admette n valeurs propres distinctes. Alors :

- ▶ f (resp. A) est diagonalisable,
- ▶ tous les sous-espace propres sont de dimension un.

En particulier, toute matrice triangulaire supérieure ou inférieure dont les coefficients diagonaux sont deux à deux distincts est diagonalisable.

Preuve




← Notons \mathcal{B} une base de vecteurs propres de f , et α_j le nombre de vecteurs propres de \mathcal{B} associés à la valeur propre λ_j .

Pour la version matricielle, considérer l'endomorphisme f canoniquement associé à A .

Attention

La réciproque est fautive! Par exemple, l'application Id_E est diagonalisable mais a pour unique valeur propre 1.

Exemple 12 — Diagonaliser $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix}$. *Indication : On montrera que $A = P \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \cdot P^{-1}$, avec $P = \begin{pmatrix} 1 & 1 & \frac{3}{2} \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$, $P^{-1} = \begin{pmatrix} 1 & -1 & \frac{1}{2} \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$.*


Exemple 13 – Matrices de rotations Soit $\theta \in]0, \pi[$. Montrer que la matrice $R_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ est diagonalisable sur \mathbf{C} mais pas sur \mathbf{R} . 

Exemple 14 – Matrice d'ATILA⁶ Soient $n \geq 2$ et $J_n = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{pmatrix}$. Montrer que J_n est diagonalisable, en commençant par calculer J_n^2 , préciser les valeurs propres possibles.

1. (Valeurs propres possibles) 

2. (Les valeurs propres sont $0, n$) 


⁶Exemple très très classique, à bien maîtriser

3. (Conclusion) La matrice J_n est diagonalisable. 


CAS D'UNE UNIQUE VALEUR PROPRE. La définition de la diagonalisabilité nous permet de traiter très facilement les cas où l'endomorphisme (ou la matrice) ne possède qu'une seule valeur propre.


Méthode Diagonalisabilité des matrices possédant une unique valeur propre — Important

- ▶ Si une matrice⁷ A possède une unique valeur propre λ , alors $A = PDP^{-1}$, où $D = \lambda I$. On en déduit que $A = \lambda I$.
- ▶ **(Conséquence)** Si A est une matrice possédant une unique valeur propre $\lambda \in \mathbf{K}$ et n'est pas égale à λI_n , alors elle **n'est pas** diagonalisable.




Exemple 15 — La matrice $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ n'est pas diagonalisable.  En effet, $\text{Spec} A = \{1\}$. Si A était diagonalisable, alors il existerait P inversible telle que $A = P I_2 P^{-1} = I_2$, et ce n'est pas le cas.

⁷Le principe est naturellement le même pour un endomorphisme, en passant par une matrice associée qui aura elle aussi une unique valeur propre.

Exemple 16 — Soit $(a, b) \in \mathbf{R}^2$. La matrice $A = \begin{pmatrix} 1 & 0 & a^2 + b^2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ peut-elle être diagonalisable?  En effet, $\text{Spec} A = \{1\}$. Si A était diagonalisable, alors il existerait P inversible telle que $A = P(1 \times I_3)P^{-1} = I_3$, donc $a^2 + b^2 = 0$ soit $a = b = 0$. Donc A est diagonalisable si et seulement si $a = b = 0$.

Exemple 17 — L'endomorphisme φ de l'Exemple 5 n'est pas diagonalisable. Pourquoi?  Nous avons trouvé $\text{Spec} \varphi = \{1\}$ — c'était une matrice triangulaire supérieure dans la base canonique, avec une diagonale qui ne comporte que des 1. Or, si φ était diagonalisable, alors sa matrice dans la base canonique serait donc I_4 — contradiction.

Exemple 18 — Connaisant un polynôme annulateur Soit $A \in \mathfrak{M}_n(\mathbf{K})$.

1. Déterminer les matrices nilpotentes diagonalisables. 
2. Soit une matrice carrée A diagonalisable vérifiant la relation $A^2 = 4I_n$ et telle que $A \neq 2I_n$ et $A \neq -2I_n$. Quelles sont les valeurs propres de A ? 
3. On suppose que $A^2 - 10A + 25I_n = 0$. À quelle condition nécessaire et suffisante A est-elle diagonalisable? 

3.1. Calculs des puissances d'une matrice

On termine cette section avec un théorème admis sur la diagonalisabilité des matrices symétriques réelles : les outils pour le démontrer ne nous sont pas accessibles, nous admettons donc ce résultat et il sera revu dans le [Chapter ALG.6](#).

Théorème ALG.5.5 | Théorème spectral

Soit A une matrice *symétrique réelle*. Alors A est diagonalisable, et il existe une matrice D diagonale réelle, et une matrice $P \in GL_n(\mathbf{R})$, telles que :

$$A = PDP^{-1} \quad \text{et} \quad {}^tP = P^{-1}{}^8.$$

3. PREMIÈRES APPLICATIONS

On termine ce chapitre par quelques applications de la diagonalisation de matrice, elles sont citées de manière non exhaustive. Nous détaillons les applications algébriques dans ce chapitre (puissances de matrices, commutants). Les applications en analyse (suites récurrentes linéaires, systèmes différentiels linéaires) seront détaillées dans un prochain chapitre.

⁸Cette condition possède une interprétation géométrique : la base de vecteurs propres est orthonormale, nous le verrons dans le [Chapter ALG.6](#)

 **Méthode** Comment trouver les puissances d'une matrice diagonalisable ?

1. Diagonaliser la matrice A . On obtient $A = PDP^{-1}$ avec D diagonale et P inversible.
2. Chercher D^n pour tout $n \in \mathbf{N}$.
3. $A^n = PD^nP^{-1}$, que l'on montre généralement par récurrence, on en déduit A^n .

Exemple 19 — On reprend la matrice de l'Exemple 12. Calculer A^n en fonction de n .

Indication : On montrera que pour tout $n \in \mathbf{N}$, $A^n = \begin{pmatrix} 1 & -1 + 2^n & \frac{1}{2} - 2^{n+1} + \frac{3^{n+1}}{2} \\ 0 & 2^n & -2^{n+1} + 2 \times 3^n \\ 0 & 0 & 3^n \end{pmatrix}$.



3.2. Calculs de commutants

Définition ALG.5.4 | Commutant


Soit $n \geq 1$ et $A \in \mathfrak{M}_n(\mathbf{K})$. Alors on appelle *commutant de A*, l'ensemble noté $\mathcal{C}(A)$ et défini par :

$$\mathcal{C}(A) = \{B \in \mathfrak{M}_n(\mathbf{K}), BA = AB\}.$$


C'est donc l'ensemble des matrices qui commutent avec A.


En diagonalisant la matrice A, on peut alors faire le lien entre le commutant de A et celui de sa partie diagonale, qui est très simple.

Exemple 20 — On reprend la matrice de l'Exemple 12 : $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix}$. On rappelle que l'on avait déjà diagonalisé la matrice A, on note P inversible et D diagonale de sorte que $A = PDP^{-1}$.

1. Démontrer que $\mathcal{C}(A)$ est un sous-espace vectoriel de $\mathfrak{M}_n(\mathbf{K})$. 

2. Calculer $\mathcal{C}(D)$. 

3. Soit $N \in \mathfrak{M}_3(\mathbf{R})$. Montrer que : $PNP^{-1} \in \mathcal{C}(A) \iff N \in \mathcal{C}(D)$. En déduire un isomorphisme entre $\mathcal{C}(A)$ et $\mathcal{C}(D)$. 

4. En déduire $\mathcal{C}(A)$ et $\dim \mathcal{C}(A)$. 



*** **Fin du chapitre** ***

4. EXERCICES

4.1. Calculs d'éléments propres

Exercice ALG.5.1 | Soit $E = \mathbf{R}^{\mathbf{N}}$, et Φ l'endomorphisme de E qui à une suite (u_n) associe la suite (v_n) définie par $v_0 = u_0$ et pour tout $n \geq 1$: $v_n = \frac{u_n + u_{n-1}}{2}$. Déterminer les éléments propres de Φ .

Solution (exercice ALG.5.1)

Soit $\lambda \in \mathbf{R}$, cherchons les (u_n) telles que : $\Phi((u_n)) = \lambda(u_n)$. l'égalité est équivalente à

$$\begin{cases} \forall n \in \mathbf{N}^*, \frac{u_n + u_{n-1}}{2} = \lambda u_n, \\ u_0 = \lambda u_0. \end{cases}$$

Ceci est encore équivalent à

$$\begin{cases} \forall n \in \mathbf{N}^*, u_{n-1} = (2\lambda - 1)u_n, \\ (1 - \lambda)u_0 = 0. \end{cases}$$

Maintenant, il faut faire des disjonctions de cas.

- ▶ Si $\lambda = 1$, alors on obtient que pour tout $n \in \mathbf{N}^*$, $u_{n-1} = u_n$ donc que (u_n) est constante et en choisissant une constante non nulle, on montre que 1 est bien une valeur propre. Ainsi $1 \in \text{Spec}(\Phi)$ et $E_1(\Phi) = \text{Vect}((1))$.
- ▶ Si $\lambda = \frac{1}{2}$, alors pour tout $n \in \mathbf{N}^*$, $u_{n-1} = 0$ et $u_0 = 0$, donc la suite (u_n) est la suite nulle. Donc $1/2 \notin \text{Spec}(\Phi)$.
- ▶ Supposons que $\lambda \notin \{1/2, 1\}$. Alors les conditions deviennent : pour tout $n \in \mathbf{N}^*$, $u_n = \frac{1}{2\lambda - 1} u_{n-1}$, et $u_0 = 0$. Donc par récurrence immédiate, (u_n) est la suite nulle (c'est une suite géométrique de terme initial nul).

En conclusion, $\text{Spec}(\Phi) = \{1\}$ et $E_1(\Phi) = \text{Vect}((1))$.

Exercice ALG.5.2 | Soit $E = \mathcal{C}^0(\mathbf{R}, \mathbf{R})$ et I l'endomorphisme de E qui à une fonction $f \in E$ associe l'unique primitive de f qui s'annule en 0. Déterminer les valeurs propres de I .

Solution (exercice ALG.5.2)

Soit f une fonction non nulle, vecteur propre associé à λ . Cela signifie que $I(f) = F = \lambda f$ en notant F l'unique primitive de f s'annulant en zéro. Alors en dérivant, on obtient $f = \lambda f'$.

- ▶ si $\lambda \neq 0$, f est solution d'une équation différentielle linéaire d'ordre 1, à coefficients constants, dont les solutions sont :

$$\forall x \in \mathbf{R}, f(x) = Ke^{\frac{1}{\lambda}x}, K \in \mathbf{R}.$$

Il existe donc des solutions non nulles, donc $\lambda \neq 0$ est valeur propre pour tout $\lambda \in \mathbf{R}^*$.

- ▶ Si $\lambda = 0$. Alors $f = 0$ et donc 0 n'est pas valeur propre.

Donc $\text{Spec } I = \mathbf{R}^*$.

4.2. Pour des matrices

Exercice ALG.5.3

1. Étudier la diagonalisabilité de $A = \begin{pmatrix} 5 & -3 \\ 6 & -4 \end{pmatrix}$.
2. Soient (u_n) et (v_n) les deux suites définies par $u_{n+1} = 5u_n - 3v_n, v_{n+1} = 6u_n - 4v_n$ pour tout $n \in \mathbf{N}$ et $u_0 = -1, v_0 = 1$. Exprimer u_n, v_n en fonction de $n \in \mathbf{N}$.

Solution (exercice ALG.5.3)

1. On trouve après calculs : $\text{Spec } A = \{2, -1\}$, et $E_2(A) = \text{Vect} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, E_{-1}(A) = \text{Vect} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.
2. En notant $X_n = \begin{pmatrix} u_n \\ v_n \end{pmatrix}$ pour tout $n \in \mathbf{N}$, nous avons la relation $X_{n+1} = AX_n$. Par récurrence immédiate, $X_n = A^n X_0 = A^n \begin{pmatrix} -1 \\ 1 \end{pmatrix} = P \begin{pmatrix} 2^n & 0 \\ 0 & (-1)^n \end{pmatrix} P^{-1} X_0$ avec $P = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$. Il ne reste plus qu'à effectuer le produit matriciel [...].

Exercice ALG.5.4 | Soient $a, b \in [0, 1]$ tels que $a + b \in [0, 1]$ et

$$A = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}.$$

- Déterminer une condition nécessaire et suffisante pour que A soit inversible.
- Montrer que 1 est valeur propre. La matrice A est-elle diagonalisable?
- On suppose que $a + b \neq 0$, on pose :

$$P = \frac{1}{a+b} \begin{pmatrix} b & a \\ b & a \end{pmatrix} \quad \text{et} \quad Q = \frac{1}{a+b} \begin{pmatrix} a & -a \\ -b & b \end{pmatrix}.$$

Calculer P^2, Q^2, PQ, QP et montrer que $A = P + (1 - a - b)Q$.

- Exprimer A^n en fonction de a, b, n pour tout $n \in \mathbf{N}$. Que peut-on dire des coefficients de A^n quand n tend vers $+\infty$?
- On suppose maintenant que $0 \leq a + b < 1$. Montrer que, pour tout n entier naturel, il existe une matrice R telle que l'on ait $R^n = A$. On dit que R est une «racine n -ième» de A . Trouver une matrice R telle que

$$R^3 = \begin{pmatrix} 5/8 & 3/8 \\ 1/2 & 1/2 \end{pmatrix}.$$

Solution (exercice ALG.5.4)

- On a $\det A = (1-a)(1-b) - ab = 1 - (a+b)$. Donc :
 A est inversible si et seulement si $a + b \neq 1$.
- Nous avons $A \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Donc $1 \in \text{Spec} A$. Soit $\lambda \in \mathbf{C}$, puisqu'on ne demande pas de diagonaliser de manière effective, passons par le déterminant. On a :

$$\det(A - \lambda I_2) = \det \begin{pmatrix} 1-a-\lambda & a \\ b & 1-b-\lambda \end{pmatrix},$$

après développement, on trouve

$$\det(A - \lambda I_2) = \lambda^2 + \lambda(a+b-2) + (1 - (a+b)).$$

Nous savons déjà que ce polynôme en λ possède une racine réelle, et donc il en possède deux. La forme factorisée est alors

$$\det(A - \lambda I_2) = (\lambda - 1)(\lambda - (1 - (a+b))).$$

On déduit alors que $\text{Spec} A = \{1, 1 - (a+b)\}$. Deux cas se présentent alors :

- Si $a + b \neq 1$, A possède deux valeurs propres distinctes, et donc A est diagonalisable.
- Si $a + b = 1$, A possède une unique valeur propre qui est 1, et donc si elle était diagonalisable, on aurait $A = I_2$, i.e. $a = b = 0$. Donc, dans ce cas, A est diagonalisable si et seulement si $a = b = 0$.

- On suppose que $a + b \neq 0$, on pose :

$$P = \frac{1}{a+b} \begin{pmatrix} b & a \\ b & a \end{pmatrix} \quad \text{et} \quad Q = \frac{1}{a+b} \begin{pmatrix} a & -a \\ -b & b \end{pmatrix}.$$

On trouve après calculs $P^2 = P, Q^2 = Q, PQ = 0 = QP$. Et

$$\begin{aligned} P + (1 - a - b)Q &= \frac{1}{a+b} \begin{pmatrix} b & a \\ b & a \end{pmatrix} + (1 - a - b) \frac{1}{a+b} \begin{pmatrix} a & -a \\ -b & b \end{pmatrix}, \\ &= \frac{1}{a+b} \begin{pmatrix} b + (1 - a - b)a & a + (1 - a - b)(-a) \\ b + (1 - a - b)(-b) & a + (1 - a - b)b \end{pmatrix} \\ &= \frac{1}{a+b} \begin{pmatrix} (1-a)(a+b) & a(a+b) \\ b(a+b) & (1-b)(a+b) \end{pmatrix} \\ &= A. \end{aligned}$$

- On a constaté que les matrices P, Q commutent. Donc d'après la formule du binôme, on déduit

$$\forall n \in \mathbf{N}, \quad A^n = \sum_{k=0}^n \binom{n}{k} P^k (1 - a - b)^{n-k} Q^{n-k}.$$

▶ supposons que $n \geq 1$, alors $P^k = P, Q^{n-k} = Q$ si $k \in \llbracket 1, n-1 \rrbracket$ donc

$$A^n = \sum_{k=1}^{n-1} \binom{n}{k} (1-a-b)^{n-k} PQ + (1-a-b)^n Q^n + P^n.$$

Mais comme $PQ = 0$, on déduit que

$$\forall n \in \mathbf{N}^*, \quad A^n = (1-a-b)^n Q + P.$$

▶ Si $n = 0$, on a évidemment $A^n = I_2$.

La formule obtenue dans le cas $n \geq 1$ est encore vraie dans le cas $n = 0$, donc conclut que pour tout $n \in \mathbf{N}$,

$$A^n = (1-a-b)^n Q + P = \frac{1}{a+b} \begin{pmatrix} a(1-a-b)^n + b & -a(1-a-b)^n + a \\ (1-a-b)^n(-b) + b & (1-a-b)^n b + a \end{pmatrix}.$$

Lorsque n tend vers $+\infty$, nous notons $\lim_{n \rightarrow \infty} A^n$ la limite coefficient par coefficient de la matrice A^n . Comme $0 < a+b \leq 1$ car $a+b = 0$ est exclu, nous avons $\lim_{n \rightarrow \infty} A^n = \frac{1}{a+b} \begin{pmatrix} b & a \\ b & a \end{pmatrix} = P$, d'après le résultat de convergence des suites géométriques. En conclusion

$$\boxed{\lim_{n \rightarrow \infty} A^n = P.}$$

5. On suppose maintenant que $0 \leq a+b < 1$. Montrer que, pour tout n entier naturel, il existe une matrice R telle que l'on ait $R^n = A$. On dit que R est une «racine n -ième» de A . Trouver une matrice R telle que

$$R^3 = \begin{pmatrix} 5/8 & 3/8 \\ 1/2 & 1/2 \end{pmatrix}.$$

Exercice ALG.5.5 | Décomposition en somme de matrices inversibles Soit $M \in \mathfrak{M}_{2,2}(\mathbf{K})$ avec $\mathbf{K} = \mathbf{R}$ ou \mathbf{C} .

1. Montrer que λ est une valeur propre de M si et seulement si λ est racine d'un polynôme du second degré à coefficients dans \mathbf{K} .
2. En déduire que M est la somme d'une matrice inversible et d'une matrice d'homothétie également inversible.
3. ▶_🐍 En déduire une fonction Python d'en-tête `decompose(M)` qui étant donné un tableau numpy M retourne les deux matrices apparaissant dans la somme précédente.

4. **(Application)** Calculer une telle somme pour la matrice $M = \begin{pmatrix} 0 & 1 \\ 0 & 2 \end{pmatrix}$.

Exercice ALG.5.6 | Soient $(x_0, x_1, x_2) \in \mathbf{R}^3$ et $A = \begin{pmatrix} x_0 & x_1 \\ x_2 & x_0 \end{pmatrix}$.

1. Trouver un polynôme χ de degré deux dont les valeurs propres de A sont les racines.
2. Démontrer que si $x_1 x_2 > 0$ alors A est diagonalisable dans $\mathfrak{M}_2(\mathbf{R})$. Est-elle diagonalisable dans $\mathfrak{M}_2(\mathbf{C})$?
3. Démontrer que si $x_1 x_2 < 0$ alors A n'est pas diagonalisable dans $\mathfrak{M}_2(\mathbf{R})$ mais qu'elle l'est dans $\mathfrak{M}_2(\mathbf{C})$.
4. Etudier la diagonalisabilité de A dans $\mathfrak{M}_2(\mathbf{R})$ et dans $\mathfrak{M}_2(\mathbf{C})$ lorsque $x_1 x_2 = 0$.

Solution (exercice ALG.5.6)

1. Soit $\lambda \in \mathbf{C}$. Alors

$$\lambda \in \text{Spec} A \iff \det \begin{pmatrix} x_0 - \lambda & x_1 \\ x_2 & x_0 - \lambda \end{pmatrix} = (x_0 - \lambda)^2 - x_1 x_2 = 0.$$

Donc : les racines de $\chi = X^2 - 2x_0 X + (x_0^2 - x_1 x_2)$ sont les valeurs de propres de A .

2. Supposons que $x_1 x_2 > 0$. Comme le discriminant associé de ce polynôme est $\Delta = 4x_0^2 - 4(x_0^2 - x_1 x_2) = 4x_1 x_2$, nous avons $\Delta > 0$ dans le contexte de cette question. Ainsi, les racines de χ (i.e. les valeurs propres de A) sont réelles et distinctes, et donc A est diagonalisable sur \mathbf{R} , a fortiori sur \mathbf{C} , puisque $\mathbf{R} \subset \mathbf{C}$.
3. On a ici $\Delta < 0$, donc les racines de χ (i.e. les valeurs propres de A) sont complexes et distinctes, dès lors A est diagonalisable sur \mathbf{C} .

4. Si $x_1 = 0$ ou $x_2 = 0$, alors A est de la forme $\begin{pmatrix} x_0 & 0 \\ x_2 & x_0 \end{pmatrix}$ ou $\begin{pmatrix} x_0 & x_1 \\ 0 & x_0 \end{pmatrix}$. Dans les deux cas, elle est triangulaire donc les valeurs propres se lisent sur diagonale. Et $\text{Spec}(A) = \{x_0\}$. Or une matrice possédant une unique valeur propre est diagonalisable si et seulement si elle est de la forme $x_0 I_2$, donc si et seulement si $x_2 = 0$ ou $x_1 = 0$. En conclusion

A est diagonalisable si et seulement si $x_2 = 0$ ou $x_1 = 0$.

Exercice ALG.5.7 | Diagonaliser

$$A = \begin{pmatrix} 0 & 2 & -1 \\ 3 & -2 & 0 \\ -2 & 2 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 3 & 2 \\ -2 & 5 & 2 \\ 2 & -3 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & -1 & 2 \end{pmatrix}.$$

Solution (exercice ALG.5.7)

On trouve après calculs :

1. pour A : $\text{Spec} A = \{-4, 2, 1\}$, et $E_{-4}(A) = \text{Vect} \begin{pmatrix} 2 \\ -3 \\ 2 \end{pmatrix}$, $E_2(A) = \text{Vect} \begin{pmatrix} -4 \\ -3 \\ 2 \end{pmatrix}$, $E_1(A) = \text{Vect} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

2. pour B : $\text{Spec} A = \{2, 1\}$, et $E_2(A) = \text{Vect} \left(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix} \right)$, $E_1(A) = \text{Vect} \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$.

3. pour C : $\text{Spec} A = \{2, 1\}$, et $E_2(A) = \text{Vect} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, $E_1(A) = \text{Vect} \left(\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right)$.

Exercice ALG.5.8 | Co-diagonalisation & application Soient

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & -1 \\ 1 & -1 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

1. Chercher les éléments propres de A, B. Montrer qu'elles sont diagonalisables avec la même matrice de passage P que l'on précisera.
2. En déduire les valeurs propres de $M(a, b) = \begin{pmatrix} b & -b & a \\ -b & b & -a \\ a & -a & 2b-a \end{pmatrix}$ pour tout $a, b \in \mathbf{R}$.

Solution (exercice ALG.5.8)

1. Soit $\lambda \in \mathbf{R}$. Alors, avec la chaîne d'opérations $L_1 \leftrightarrow L_3, L_3 \leftarrow L_3 + \lambda L_1, L_3 \leftarrow L_3 - L_2$:

$$A - \lambda I_3 \sim_L \begin{pmatrix} 1 & -1 & -1-\lambda \\ 0 & -\lambda & -1 \\ 0 & 0 & -\lambda^2 - \lambda + 2 \end{pmatrix}$$

Ainsi, $\text{Spec}(A) = \{-2, 0, 1\}$. La matrice possède trois valeurs propres donc est diagonalisable. Après calculs d'espace propre, on trouve : $E_{-2}(A) = \text{Vect} \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}$, $E_0(A) = \text{Vect} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $E_1(A) = \text{Vect} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$. De-même, on montre avec la chaîne d'opérations $L_1 \leftrightarrow L_2, L_2 \leftarrow L_2 + (1-\lambda)L_1$,

$$B - \lambda I_3 \sim_L \begin{pmatrix} -1 & 1-\lambda & 0 \\ 0 & (1-\lambda)^2 - 1 = -\lambda(2-\lambda) & 0 \\ 0 & 0 & 2-\lambda \end{pmatrix}$$

Ainsi, $\text{Spec}(B) = \{0, 2\}$. La matrice possède trois valeurs propres donc est diagonalisable. Après calculs d'espace propre, on trouve :

$$E_0(B) = \text{Vect} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad E_2(B) = \text{Vect} \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Par opérations élémentaires, puisqu'on ne change pas un Vect en changeant les vecteurs par des combinaisons linéaires de ceux-ci, on constate que :

$$\begin{aligned} \text{Vect} \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} &= \text{Vect} \left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} = \text{Vect} \left\{ \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} \\ &= \text{Vect} \left\{ \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, 2 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\} = \text{Vect} \left\{ \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix} \right\}. \end{aligned}$$

. Donc $E_2(B) = \text{Vect} \left\{ \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix} \right\}$. Ceci montre que, en choisissant $P =$


$$\begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ 0 & 2 & 1 \end{pmatrix}, \text{ on a les relations :}$$

$$A = P \text{Diag}(0, -2, 1) P^{-1}, \quad B = P \text{Diag}(0, 2, 2) P^{-1}.$$

2. Soit a, b deux réels, alors $M(a, b) = aA + bB = P(\text{Diag}(0, -2a, a) + \text{Diag}(0, 2b, 2b))P^{-1} = \boxed{P \text{Diag}(0, 2(b-a), a+2b)P^{-1}}$. Cette identité nous dit alors que $M(a, b)$ est diagonalisable, que $\text{Spec}(M(a, b)) = \{0, 2(b-a), a+2b\}$ et une base des espaces propres est donnée par les colonnes de P .

Exercice ALG.5.9 | Trace d'une matrice carrée Soit $A = (a_{i,j})$ une matrice de $\mathcal{M}_n(\mathbb{K})$. On définit la trace de A par

$$\text{Tr}(A) = \sum_{i=1}^n a_{i,i}.$$

1. Montrer que Tr est une forme linéaire.
2.  Écrire un script Python qui prend en argument un tableau numpy carré, et renvoie la valeur de la trace.

3. Soient A et B deux matrices carrées d'ordre n .
 - 3.1) Montrer que $\text{Tr}(AB) = \text{Tr}(BA)$. En déduire que l'équation $AB - BA = I_n$ n'admet pas de solution.
 - 3.2) On suppose que A et B sont semblables. Montrer que $\text{Tr}(A^k) = \text{Tr}(B^k)$ pour tout $k \in \mathbb{N}^*$.
4. En déduire que l'on peut définir la trace d'un endomorphisme f d'un espace vectoriel de dimension finie.
5. On suppose que A est diagonalisable. Que vaut la somme des valeurs propres?
6. Soit $A = \begin{pmatrix} 2 & -17 & \pi \\ 2 & -17 & \pi \\ 2 & -17 & \pi \end{pmatrix}$. Trouver sans calcul ses valeurs propres.

Solution (exercice ALG.5.9)

1. Soient $A, B \in \mathcal{M}_{n,n}(\mathbb{K})$, et $\lambda, \mu \in \mathbb{K}$. Alors :

$$\begin{aligned} \text{Tr}(\lambda A + \mu B) &= \sum_{i=1}^n (\lambda a_{i,i} + \mu b_{i,i}), \\ &= \lambda \sum_{i=1}^n a_{i,i} + \mu \sum_{i=1}^n b_{i,i}, \\ &= \lambda \text{Tr}(A) + \mu \text{Tr}(B). \end{aligned}$$

De plus, $\text{Tr}(\cdot)$ est à valeurs dans \mathbb{R} , donc c'est une forme linéaire.

2. 

```
def trace(A):
    """
    retourne la somme des coefficients sur la diagonale
    """
    trace = 0
    for i in range(n):
        trace += A[i, i]
    return trace
```

3. Soient A et B deux matrices carrées d'ordre n .

3.1) Soient $A = (a_{i,j}), B = (b_{i,j}) \in \mathcal{M}_{n,n}(\mathbf{K})$. Montrons que $\text{Tr}(AB) = \text{Tr}(BA)$. Rappelons que le coefficient (i, j) de la matrice AB est $p_{i,j} = \sum_{k=1}^n a_{i,k} b_{k,j}$, de même le coefficient (i, j) de la matrice BA est $q_{i,j} = \sum_{k=1}^n b_{i,k} a_{k,j}$. Donc :

$$\begin{aligned} \text{Tr}(AB) &= \sum_{i=1}^n p_{i,i}, \\ &= \sum_{i=1}^n \sum_{k=1}^n a_{i,k} b_{k,i}, \\ &= \sum_{k=1}^n \sum_{i=1}^n a_{i,k} b_{k,i}, \quad \left. \begin{array}{l} \\ \end{array} \right\} \text{permutation de sommes finies,} \\ &= \sum_{k=1}^n \sum_{i=1}^n b_{k,i} a_{i,k} = \sum_{k=1}^n q_{k,k}. \end{aligned}$$

Donc : $\boxed{\text{Tr}(AB) = \text{Tr}(BA)}$. Supposons qu'il existe deux matrices A, B telles que : $AB - BA = I_n$. Alors en prenant la trace, on obtient en exploitant la linéarité : $\text{Tr}(AB) - \text{Tr}(BA) = \text{Tr}(I_n) = n$. Donc, comme $\text{Tr}(AB) = \text{Tr}(BA)$, on obtiendrait $0 = n$, c'est bien entendu une contradiction.

3.2) Il existe une matrice inversible P telle que : $A = PBP^{-1}$. Soit $k \in \mathbf{N}^*$, alors par récurrence immédiate $A^k = PB^kP^{-1}$. Donc :

$$\text{Tr}(A^k) = \text{Tr}(P(B^kP^{-1})) = \text{Tr}((B^kP^{-1})P) = \text{Tr}(B^k).$$

Donc $\boxed{\text{Tr}(A^k) = \text{Tr}(B^k)}$ pour tout $k \in \mathbf{N}^*$.

4. La notion de trace est définie avec une matrice. Si l'on veut définir la notion de trace pour un endomorphisme, il faut par exemple se donner une matrice associée, dans une base fixée de E , appelons-là \mathcal{B} et notons $A = \mathcal{M}_{\mathcal{B}}(f)$. La question est : si je choisis une autre base, aurais-je une autre trace ? La réponse est non. En effet, si \mathcal{B}' est une autre base de E , alors en notant $B = \mathcal{M}_{\mathcal{B}'}(f)$, on sait par formule de changement de base qu'il existe une matrice P inversible telle que $A = PBP^{-1}$ et donc $\text{Tr}(A) = \text{Tr}(B)$ d'après la question précédente pour $k = 1$. Ainsi, $\boxed{\text{on appelle trace d'un endomorphisme } f \text{ la trace d'une matrice associée dans une certaine base.}}$

5. Notons $D = \text{Diag}(\lambda_1, \dots, \lambda_n)$ avec $\lambda_1, \dots, \lambda_n \in \mathbf{K}$ les valeurs propres de A . Alors il existe une matrice P inversible telle que $A = PDP^{-1}$.

Nous avons déjà vu que $\text{Tr}(A) = \text{Tr}(D)$, donc $\text{Tr}(A) = \sum_{i=1}^n \lambda_i$.

$\boxed{\text{La somme des valeurs propres est donc égale à la trace.}}$

6. Sans calcul, on voit que la matrice est de rang 1, donc d'après le théorème du rang, 0 est une valeur propre et d'espace propre associé de dimension 2. Il ne reste donc plus éventuellement qu'une valeur propre à trouver, qui sera d'espace propre associé au maximum 1. Soit λ cette valeur propre, alors $\text{Tr}(A) = -15 + \pi = 0 + 0 + \lambda$ d'après l'exercice. Ainsi, $\lambda = \pi - 15$ et $\boxed{\text{Spec}(A) = \{0, \pi - 15\}}$.

4.3. Pour des endomorphismes

Exercice ALG.5.10 | Soit E l'espace vectoriel des polynômes à coefficients dans \mathbf{K} ($\mathbf{K} = \mathbf{R}$ ou $\mathbf{K} = \mathbf{C}$) de degré inférieur ou égal à n . Soit f l'endomorphisme de E défini par :

$$\forall P \in E, \quad f(P) = P - P'.$$

1. Démontrer que f est bijectif de deux manières :

- 1.1)** sans utiliser de matrice de f ,
- 1.2)** en utilisant une matrice de f .

2. L'endomorphisme f est-il diagonalisable ?

Solution (exercice ALG.5.10)

1. Démontrons que f est bijectif de deux manières :

- 1.1)** Tout d'abord, f est linéaire par linéarité de la dérivation, et puisque pour tout $P \in E$, $\deg P' < \deg P$, alors $\deg f(P) \leq n$. Donc f est un endomorphisme de E , et en dimension finie, donc il suffit de montrer l'injectivité. Soit donc $P \in E$ tel que $P = P'$. Alors, si $d = \deg P \neq -\infty$, on a $d = d - 1$ — c'est une contradiction. Donc $\deg P = -\infty$ et $P = 0$ donc $\boxed{f \text{ est bijectif.}}$

1.2) Soit $k \in \llbracket 0, n \rrbracket$, alors $f(X^k) = X^k - kX^{k-1}$, et $f(1) = 1$ donc :

$$\mathcal{M}_{\mathcal{B}^{\text{can}}}(f) = \begin{pmatrix} 1 & -1 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix}.$$

C'est donc une matrice triangulaire supérieure, de rang $n + 1$ donc

f est bijectif.

2. On a, d'après la première question, $\text{Spec}(f) = \{1\}$ donc si f était diagonalisable, alors $\mathcal{M}_{\mathcal{B}^{\text{can}}}(f) = P I_{n+1} P^{-1}$ avec P inversible donc $\mathcal{M}_{\mathcal{B}^{\text{can}}}(f)$ serait la matrice identité, ce qui n'est pas le cas. Donc f n'est pas diagonalisable.

Exercice ALG.5.11 | Soit $n \geq 2$. On note E l'espace vectoriel des fonctions continues de \mathbf{R} dans \mathbf{R} , et $F = \mathbf{R}_n[X]$. Pour tout $f \in E$, on note $T(f)$ l'application de \mathbf{R} dans \mathbf{R} définie par :

$$\forall x \in \mathbf{R}, \quad T(f)(x) = \int_{x-1}^{x+1} f(t) dt.$$

1. Soit $g : x \mapsto |x|$, déterminer $T(g)$.
2. Montrer que pour tout f de E , $T(f)$ est de classe \mathcal{C}^1 , et déterminer sa dérivée.
3. Montrer que T est une application linéaire de E dans E , T est-elle surjective? injective? bijective?
4. Montrer que T est aussi un endomorphisme de F , et que cette fois T est même un automorphisme. Est-il diagonalisable?

4.4. Polynôme annulateur

Exercice ALG.5.12 | Soit $\mathcal{A} = \{M \in \mathfrak{M}_{2,2}(\mathbf{R}), (M + I_2)(M + 2I_2) = O_2\}$.

1. Soit $M = \begin{pmatrix} -3 & 2 \\ -1 & 0 \end{pmatrix}$. Montrer que $M \in \mathcal{A}$. Déterminer les valeurs propres de M et les sous-espaces propres associés.

2. Soit $M \in \mathcal{A}$.

- 2.1) Montrer que $\text{Spec } M \subset \{-2, -1\}$. Que dire si $\text{Spec } M = \{-2, -1\}$? Est-il possible que ni -1 ni -2 ne soient valeur propre de M ?
- 2.2) On suppose que -1 est la seule valeur propre de M . Montrer que $M = -I$. Conclure quant à l'ensemble \mathcal{A} sous cette condition.

Exercice ALG.5.13 | Soient $a, b, c \in \mathbf{R}^*$, et $M = \begin{pmatrix} 1 & \frac{a}{b} & \frac{a}{c} \\ \frac{b}{a} & 1 & \frac{b}{c} \\ \frac{c}{a} & \frac{c}{b} & 1 \end{pmatrix}$.

1. Calculer M^2 , puis montrer que $\text{Spec}(M) \subset \{0, 3\}$.
2. Calculer $M \begin{pmatrix} a \\ b \\ c \end{pmatrix}$.
3. Déterminer $\text{Spec } M$ et les sous-espaces propres de M . La matrice M est-elle diagonalisable? Si oui, la diagonaliser.

Solution (exercice ALG.5.13)

1. Après calculs, on trouve que $M^2 - 3M = 0$. Ainsi, le polynôme $X^2 - 3X$ est annulateur de M et donc d'après le cours, $\text{Spec}(M) \subset \{0, 3\}$.
2. On trouve $M \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 3 \begin{pmatrix} a \\ b \\ c \end{pmatrix}$.
3. Soit $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathfrak{M}_{3,1}(\mathbf{R})$. Alors résolvons $MX = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$. Cela donne le système ci-après :

$$\begin{cases} x + y\frac{a}{b} + z\frac{a}{c} = 0 \\ x\frac{b}{a} + y + z\frac{b}{c} = 0 & L_1 \leftarrow L_1/a, L_2 \leftarrow L_2/b, L_3 \leftarrow L_3/c. \\ x\frac{c}{a} + y\frac{c}{b} + z = 0 \\ \frac{1}{a}x + \frac{1}{b}y + \frac{1}{c}z = 0. \end{cases}$$

Ainsi, il suffit de trouver deux vecteurs non colinéaires de $\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \frac{1}{a}x + \frac{1}{b}y + \frac{1}{c}z = 0 \right\}$, donc

$$E_0(M) = \text{Vect} \left(\begin{pmatrix} 0 \\ b \\ -c \end{pmatrix}, \begin{pmatrix} a \\ 0 \\ -c \end{pmatrix} \right).$$

À ce stade, on peut alors déjà déduire que **M est diagonalisable** puisque $\dim E_0(M) = 2$ et on a toujours $\dim E_0(M) + \dim E_3(M) \leq 3$ donc $\dim E_0(M) + \dim E_3(M) = 3$, et $\dim E_3(M) = 1$. Or, d'après la question précédente $\begin{pmatrix} a \\ b \\ c \end{pmatrix} \in E_3(M)$,

donc $E_3(M) = \text{Vect} \begin{pmatrix} a \\ b \\ c \end{pmatrix}$. Finalement, on a montré que :

$$M = P \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix} P^{-1}, \quad P = \begin{pmatrix} 0 & a & a \\ b & 0 & b \\ -c & -c & c \end{pmatrix}.$$

4. Soit $N \in \mathfrak{M}_{3,3}(\mathbb{R})$. Considérons la matrice N' telle que $N = PN'P^{-1}$ et notons $N' = \begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{pmatrix}$. Alors :

$$N \in \mathcal{C} \iff MN - NM = 3N \iff PDP^{-1}PN'P^{-1} - PN'P^{-1}PDP^{-1} = 3PN'P^{-1}.$$

Donc, en simplifiant les matrices P :

$$N \in \mathcal{C} \iff DN' - N'D = 3N'.$$



En écrivant cette condition avec les coefficients de N' , on trouve alors :

$$N \in \mathcal{C} \iff \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 3\alpha_3 & 3\beta_3 & 3\gamma_3 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 3\gamma_1 \\ 0 & 0 & 3\gamma_2 \\ 0 & 0 & 3\gamma_3 \end{pmatrix} = 3 \begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{pmatrix}.$$

En identifiant les coefficients, on a $\alpha_1 = \beta_1 = \alpha_2 = \beta_2 = 0$, puis $6\gamma_1 = 6\gamma_2 = 0$ et enfin $0 = 3\gamma_3$ donc tous les coefficients sont nuls sauf α_3, β_3 . Donc l'ensemble des solutions en N' est $\text{Vect}(A, B)$ où $A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$. Donc :

$$\mathcal{C} = \text{Vect}(PAP^{-1}, PBP^{-1}), \text{ c'est donc un espace vectoriel de dimension deux.}$$

Exercice ALG.5.14 | A-ENV 2018 Soit $n \in \mathbb{N}^*$. Soit Φ définie sur $\mathbb{R}_n[X]$ par $\Phi(P) = \sum_{k=0}^n \binom{n}{k} P^{(k)}$.

1. **1.1)** Déterminer le lien entre degré de P et degré de $P^{(k)}$.
1. **1.2)** Montrer que Φ est un endomorphisme de $\mathbb{R}_n[X]$. On note $A = (a_{i,j})_{1 \leq i, j \leq n+1}$ la matrice canoniquement associée à Φ .
2. Soit $\Delta \in \mathcal{L}(\mathbb{R}_n[X]) : P \mapsto P'$ avec P un polynôme quelconque de $\mathbb{R}_n[X]$.
 - 2.1) Écrire la matrice D de Δ dans la base canonique de $\mathbb{R}_n[X]$.
 - 2.2)  On rappelle que dans le module numpy on peut coder des matrices sous forme de tableau de listes en utilisant la commande array. Exemple : `np.array([[1, 2, 3], [7, 8, 9]])` code la matrice $\begin{pmatrix} 1 & 2 & 3 \\ 7 & 8 & 9 \end{pmatrix}$. La commande `np.zeros((2, 3))` code la matrice $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. La commande `np.eye(3)` code la matrice identité I_3 . La commande `np.dot(A, B)` effectue le produit de la matrice A par B sous réserve qu'elles soient compatibles et codées comme des array. Coder en langage Python la matrice D.
 - 2.3) Montrer que $A = (I_{n+1} + D)^n$.
 - 2.4)  Écrire une fonction `Phi(coordonneesP)` en langage Python qui retourne les coordonnées de $\Phi(P)$ dans la base canonique sous forme de liste en prenant en argument les coordonnées de P dans la base canonique sous forme de liste.
3. **3.1)** En utilisant le degré, montrer que Φ est injective.
3. **3.2)** Montrer que Φ est bijective.

- 4. Montrer que 1 est l'unique valeur propre de Φ et déterminer l'espace propre associé.
- 5. L'endomorphisme Φ est-il diagonalisable? (On justifiera avec 2 arguments différents)

Solution (exercice ALG.5.14)

- 1. **1.1)** C'est une question de cours. En effet, $\deg P^{(k)} = \deg P - k$ si $k \leq \deg P$, et $-\infty$ sinon.
- 1.2)** La dérivation d'un polynôme donne un polynôme donc $\Phi(P) \in \mathbf{R}[X]$. Par ailleurs, si $P \in \mathbf{R}_n[X]$, alors $\deg \Phi(P) = \deg P$ puisque d'après la question précédente, tous les autres termes pour $k = 1, \dots, n$ sont de degré strictement inférieur à celui de P . Ainsi, $\deg \Phi(P) \leq n$. Soient maintenant $\lambda, \mu \in \mathbf{R}$, et $P, Q \in \mathbf{R}_n[X]$. Alors

$$\begin{aligned} \Phi(\lambda P + \mu Q) &= \sum_{k=0}^n \binom{n}{k} (\lambda P + \mu Q)^{(k)} \\ &= \lambda \sum_{k=0}^n \binom{n}{k} P^{(k)} + \mu \sum_{k=0}^n \binom{n}{k} Q^{(k)} \\ &= \lambda \Phi(P) + \mu \Phi(Q). \end{aligned}$$

$\left. \begin{array}{l} \text{linéarité de la dérivation} \\ \text{linéarité de la somme} \end{array} \right\}$

Donc Φ est un endomorphisme de $\mathbf{R}_n[X]$.

- 2. **2.1)** La base canonique de $\mathbf{R}_n[X]$ est $(1, X, X^2, \dots, X^n)$ et pour tout $k \in \llbracket 1, n \rrbracket$, $(X^k)' = kX^{k-1}$. Par ailleurs $1' = 0$. On en déduit la forme

de D : $D = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ \vdots & 0 & 2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & n \\ 0 & 0 & \dots & \dots & \dots & 0 \end{pmatrix}$

```
2.2) import numpy as np
def D(n):
    """
```



```
en entree un entier n plus grand ou egal a 1
en sortie la matrice D
"""
M = np.zeros((n+1, n+1))
for i in range(n):
    # modif de la sur-diagonale
    M[i, i+1] = i+1
return M

>>> D(4)
array([[0., 1., 0., 0., 0.],
       [0., 0., 2., 0., 0.],
       [0., 0., 0., 3., 0.],
       [0., 0., 0., 0., 4.],
       [0., 0., 0., 0., 0.]])
```

- 2.3) Pour tout $P \in \mathbf{R}_n[X]$, et $k \in \llbracket 0, n \rrbracket$, on a

$$P^{(k)} = \Delta^k(P).$$

Donc nous avons l'égalité ci-après :

$$\Phi = \sum_{k=0}^n \binom{n}{k} \Delta^k.$$

Précisons ci-dessus que la puissance ci-dessus désigne la composition. Par ailleurs, d'après le cours, la matrice de Δ^k est D^k et le symbole matrice est

une application linéaire. Donc

$$\begin{aligned}
 A &= \mathcal{M}_{\mathcal{B}^{\text{can}}}(\Phi) \\
 &= \mathcal{M}_{\mathcal{B}^{\text{can}}}\left(\sum_{k=0}^n \binom{n}{k} \Delta^k\right) \\
 &= \sum_{k=0}^n \binom{n}{k} \mathcal{M}_{\mathcal{B}^{\text{can}}}(\Delta^k) \\
 &= \sum_{k=0}^n \binom{n}{k} I_n^{n-k} D^k \\
 &= (I_{n+1} + D)^n.
 \end{aligned}$$

linéarité du symbole matrice
 $I_n^{n-k} = I_n$ pour tout k
binôme matriciel

2.4) On commence par écrire un programme qui calcule A en utilisant la question précédente.

```

def A(n):
    I = np.eye(n+1)
    M = I+D(n)
    A = I #on initialise A puis on la multiplie autant de
    - fois que nécessaire
    for _ in range(n):
        A = A@M
    return A

>>> A(4)
array([[ 1.,  4., 12., 24., 24.],
       [ 0.,  1.,  8., 36., 96.],
       [ 0.,  0.,  1., 12., 72.],
       [ 0.,  0.,  0.,  1., 16.],
       [ 0.,  0.,  0.,  0.,  1.]])
    
```

Puis on définit une fonction prenant aussi comme paramètre n :

```

def Phi(n, coordonneesP):
    return A(n)@coordonneesP
    
```

```

>>> coordonneesP = np.array([1, 2, 3])
>>> Phi(2, coordonneesP)
array([11., 14., 3.])
    
```

On peut remarquer que la commande @ et dot dans numpy acceptent le produit d'une matrice sous forme de array avec une liste, considérée alors comme un vecteur colonne. Le résultat est donné sous forme de liste mais matriciellement correspond aussi à une colonne.

- 3. 3.1)** Soit $P \in \mathbf{R}_n[X]$ tel que $\Phi(P) = 0$. Alors en passant au degré on a $\deg P = -\infty$ donc P est le polynôme nul. Donc P est injectif.
- 3.2)** Φ est un endomorphisme injectif en dimension finie, donc Φ est bijectif.
- 4. 4.1)** Le plus simple est de revenir à A la matrice associée à Φ qui est une combinaison linéaire de matrices triangulaires supérieures, donc est triangulaire supérieure. Dans ce cas, on sait que les valeurs propres de A se lisent sur sa diagonale.

Or par construction de D, seul $D^0 = I_{n+1}$ a une diagonale non nulle, et $\binom{n}{0} = 1$ donc la diagonale de A n'est composée que de 1. Donc :

Φ , comme A, a 1 comme unique valeur propre

Cherchons alors $E_1(\Phi)$ en revenant à la définition : $P \in E_1(\Phi)$ si et seulement si

$$\Phi(P) = P \iff P + \sum_{k=1}^n \binom{n}{k} P^{(k)} = P.$$

Donc on résout

$$\sum_{k=1}^n \binom{n}{k} P^{(k)} = 0. \quad (\star)$$

Or la famille $(P', P'', \dots, P^{(n)})$ est échelonnée, donc libre, sauf si $P' = 0$. Mais (\star) nous dit que la famille $(P', P'', \dots, P^{(n)})$ est liée, donc $P' = 0$ et P est un polynôme constant. Inversement, on vérifie que $\Phi(C) = C$ pour tout $C \in \mathbf{R}_0[X]$. Donc $E_1(\Phi) = \mathbf{R}_0[X]$.

- 4.2)** Φ a une unique valeur propre : 1, et l'espace propre associé est de dimension $1 \neq n + 1$ qui est la dimension de $\mathbf{R}_n[X]$, donc Φ n'est pas diagonalisable.

- ▶ Autre argument : on s'intéresse cette fois à la matrice A . Comme A a une unique valeur propre 1 , elle est diagonalisable si et seulement si elle est semblable à la matrice $1 \times I_{n+1} = I_{n+1} : A = P^{-1}I_{n+1}P = I_{n+1}P^{-1}P = I_{n+1}$, ce qui n'est pas vrai.

Donc A n'est pas diagonalisable, donc Φ n'est pas diagonalisable

Chapitre ALG.6.

Produit scalaire euclidien

Résumé & Plan

L'objectif de ce chapitre est de définir une notion d'orthogonalité et de distance sur les espaces vectoriels $E = \mathbf{R}^n$ et $E = \mathcal{M}_{n,1}(\mathbf{R})$ correspondant aux vecteurs lignes et colonnes. Ainsi le but sera entre autres de généraliser le produit scalaire usuel de \mathbf{R}^2 définie par

$$\langle X|Y \rangle = x_1y_1 + x_2y_2, \quad X = (x_1, x_2) \in \mathbf{R}^2, \quad Y = (y_1, y_2) \in \mathbf{R}^2.$$

Rappelons que d'après le théorème de PYTHAGORE énoncé au collège dans \mathbf{R}^2 , la notion de distance associée entre $O(0, 0)$ et $M(X) = M(x_1, x_2)$ est :

$$d(O, M(X)) = \sqrt{\langle X|X \rangle} = \sqrt{x_1^2 + x_2^2}.$$

Que devient cette distance dans \mathbf{R}^n ? Nous définirons la notion générale de distance à une partie dans \mathbf{R}^n , comme la borne inférieure des distances à tout vecteur de la partie, et nous verrons que cette notion possède de nombreuses applications. Par exemple le problème d'existence de la droite des moindres carrés. Nous l'avons déjà obtenue pour rappel à l'aide d'outils analytiques sur les fonctions de deux variables.

1	Généralités	2
1.1	Espaces euclidiens	2
1.2	Norme & Distance	4

2	Orthogonalité	9
2.1	Orthogonal d'une partie	9
2.2	Familles orthogonales & orthonormales	11
2.3	Coordonnées d'un vecteur dans une base orthonormale	12
2.4	Diagonalisation des matrices symétriques réelles	14
3	Projection orthogonale	17
3.1	Projection orthogonale	17
4	Exercices	24
4.1	Inégalités classiques	24
4.2	Diagonalisation de matrices symétriques	26
4.3	Projection orthogonale	31

Le produit scalaire et les normes sont utilisés de façon très importante dans les jeux vidéo, par exemple pour déterminer des collisions entre objets ou pour calculer des éclairages.

— Le saviez-vous ?

Cadre
 Dans tout ce chapitre, sauf mention contraire, n désigne un entier strictement positif.

1. GÉNÉRALITÉS

On commence par définir la notion générale de produit scalaire, sur un espace vectoriel quelconque. On étudiera ensuite plus en détail le produit scalaire dit « euclidien » sur \mathbf{R}^n et $\mathfrak{M}_{n,1}(\mathbf{R})$.

1.1. Espaces euclidiens

Définition ALG.6.1 | Espace euclidien et produit scalaire

Soit E un \mathbf{R} -espace vectoriel. On appelle *produit scalaire* sur E toute « forme bilinéaire symétrique définie positive ». C'est-à-dire toute application $\langle \cdot | \cdot \rangle : E \times E \rightarrow \mathbf{R}$ telle que :

1. (Bilinéarité) $\forall (x, y, z) \in E^3, \quad \forall (\lambda, \mu) \in \mathbf{R}^2,$

$$\langle \lambda x + \mu y | z \rangle = \lambda \langle x | z \rangle + \mu \langle y | z \rangle, \quad \langle z | \lambda x + \mu y \rangle = \lambda \langle z | x \rangle + \mu \langle z | y \rangle.$$

2. (Symétrie) $\forall (x, y) \in E^2, \quad \langle x | y \rangle = \langle y | x \rangle.$

3. (Définie) $\forall x \in E, \quad \langle x | x \rangle = 0 \iff x = 0_E.$

4. (Positive) $\forall x \in E, \quad \langle x | x \rangle \geq 0.$

Un espace vectoriel muni d'un produit scalaire $(E, \langle \cdot | \cdot \rangle)$ est appelé *espace préhilbertien*. Si E est de plus de dimension finie, alors $(E, \langle \cdot | \cdot \rangle)$ est appelé *espace euclidien*.

Dans toute la suite, nous travaillerons quasi intégralement avec le produit scalaire défini ci-dessous, sur $E = \mathbf{R}^n$ ou plus rarement sur $E = \mathfrak{M}_{n,1}(\mathbf{R})$, que l'on appelle produit scalaire *canonique* ou *euclidien*, il généralise la version sur \mathbf{R}^2 rappelée en introduction.

Définition/Proposition ALG.6.1 | Produit scalaire « canonique », ou « euclidien »

On appelle :


► *produit scalaire canonique (ou produit scalaire euclidien) sur \mathbf{R}^n* l'application :

$$\langle \cdot | \cdot \rangle \left| \begin{array}{l} \mathbf{R}^n \times \mathbf{R}^n \quad \longrightarrow \quad \mathbf{R}, \\ ((x_1, \dots, x_n), (y_1, \dots, y_n)) \longrightarrow \sum_{i=1}^n x_i y_i, \end{array} \right.$$

► *produit scalaire canonique (ou produit scalaire euclidien) sur $\mathfrak{M}_{n,1}(\mathbf{R})$* l'application :

$$\langle \cdot | \cdot \rangle \left| \begin{array}{l} \mathfrak{M}_{n,1}(\mathbf{R}) \times \mathfrak{M}_{n,1}(\mathbf{R}) \quad \longrightarrow \quad \mathbf{R}, \\ \left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right) \longrightarrow \sum_{i=1}^n x_i y_i. \end{array} \right.$$

Pour $n = 1$, il correspond au produit de deux réels, pour $n = 2$ au produit scalaire du lycée.

Preuve (Il s'agit bien d'un produit scalaire) Faisons la preuve dans le cas $E = \mathbf{R}^n$, elle est strictement identique pour $E = \mathfrak{M}_{n,1}(\mathbf{R})$. 

1. Soit $(x, y, z) \in (\mathbf{R}^n)^3, (\lambda, \mu) \in \mathbf{R}^2$, en notant x_i, y_i, z_i les coordonnées des trois vecteurs, on obtient :

$$\begin{aligned} \langle \lambda x + \mu y | z \rangle &= \langle (\lambda x_1 + \mu y_1, \dots, \lambda x_n + \mu y_n) | (z_1, \dots, z_n) \rangle = \sum_{i=1}^n (\lambda x_i + \mu y_i) z_i \\ &= \lambda \sum_{i=1}^n x_i z_i + \mu \sum_{i=1}^n y_i z_i \\ &= \lambda \langle x | z \rangle + \mu \langle y | z \rangle. \end{aligned} \quad \left. \begin{array}{l} \text{linéarité} \\ \text{de la} \\ \text{somme} \end{array} \right\}$$

On a donc montré qu'elle était linéaire à gauche.


2. Avec les notations de la définition, puisque pour tout $i \in \llbracket 1, n \rrbracket, x_i y_i = y_i x_i$, il vient la propriété de symétrie. D'où finalement la linéarité à droite.

3. Soit $x \in \mathbf{R}^n$, alors comme $\langle x | x \rangle = \sum_{i=1}^n x_i^2$, il vient immédiatement $\langle x | x \rangle \geq 0$.

4. Comme une somme de termes positifs est nulle si et seulement si tous les termes sont nuls on obtient $\langle x | x \rangle = 0$ si et seulement si $x = 0_{\mathbf{R}^n}$.

En toute rigueur, il faudrait noter différemment les deux produits scalaires, plutôt qu'avec un symbole générique « $\langle \cdot | \cdot \rangle$ ». En tant qu'applications ce sont deux objets différents puisque les ensembles de départ ne sont pas les mêmes.

Dans la pratique, nous ne le ferons pas, car cette confusion de notation n'est pas particulièrement source d'erreurs. Les sujets de concours généralement font pire et introduisent le problème avec une phrase du type «on assimilera les vecteurs lignes aux vecteurs colonnes».

 **Cadre**
 Dans tout ce chapitre, E désignera dorénavant \mathbf{R}^n ou $\mathfrak{M}_{n,1}(\mathbf{R})$, et $\langle \cdot | \cdot \rangle$ le produit scalaire euclidien associé, comme défini précédemment.


Proposition ALG.6.1 | Expression matricielle dans \mathcal{B}^{can} du produit scalaire

1. (Si $E = \mathbf{R}^n$) Soient $x, y \in \mathbf{R}^n$, notons $X = \mathcal{M}_{\mathcal{B}^{\text{can}}}(x)$, $Y = \mathcal{M}_{\mathcal{B}^{\text{can}}}(y)$. Alors :


$$\langle x | y \rangle = {}^t X \times Y.$$

2. (Si $E = \mathfrak{M}_{n,1}(\mathbf{R})$) Soient $X, Y \in \mathfrak{M}_{n,1}(\mathbf{R})$. Alors :

$$\langle X | Y \rangle = {}^t X \times Y.$$

 **Attention**
 Afin de détecter toute erreur, bien vérifier les formats des matrices quand vous utilisez ces formules.

Preuve

1. (Si $E = \mathbf{R}^n$) 

2. (Si $E = \mathfrak{M}_{n,1}(\mathbf{R})$) Même preuve que précédemment.

Remarque 1.1 — Dans la suite, nous allons généraliser la Proposition ALG.6.1. Plus précisément, nous montrerons qu'elle est encore vraie en remplaçant \mathcal{B}^{can} par n'importe quelle autre base dite *orthonormée*.

Nous avons donc vu que $\langle \cdot | \cdot \rangle$ est bilinéaire, *i.e.*

- ▶ pour tout $y \in E$, $x \mapsto \langle x | y \rangle$ est linéaire (on parle de «linéarité à gauche»),
- ▶ et pour tout $x \in E$, $y \mapsto \langle x | y \rangle$ également (on parle de «linéarité à droite»).

Cela peut se formuler comme ci-dessous.¹

Corollaire ALG.6.1 | Écriture explicite de la bilinéarité

Soient p et q deux entiers strictement positifs, $(\lambda_1, \dots, \lambda_p) \in \mathbf{R}^p$ et $(\mu_1, \dots, \mu_q) \in \mathbf{R}^q$. Soient (x_1, \dots, x_p) une famille de p vecteurs de E et (y_1, \dots, y_q) une famille de q vecteurs de E . Alors :

$$\left\langle \sum_{i=1}^p \lambda_i x_i \mid \sum_{j=1}^q \mu_j y_j \right\rangle = \sum_{i=1}^p \sum_{j=1}^q \lambda_i \mu_j \langle x_i | y_j \rangle.$$

Preuve

$$\begin{aligned} \left\langle \sum_{i=1}^p \lambda_i x_i \mid \sum_{j=1}^q \mu_j y_j \right\rangle &= \sum_{i=1}^p \lambda_i \left\langle x_i \mid \sum_{j=1}^q \mu_j y_j \right\rangle && \text{par linéarité à gauche,} \\ &= \sum_{i=1}^p \sum_{j=1}^q \lambda_i \mu_j \langle x_i | y_j \rangle && \text{par linéarité à droite.} \end{aligned}$$

Remarque 1.2 — D'autres produits scalaires

¹Vous noterez l'analogie avec la formule de développement d'une covariance, normal puisque les deux quantités étudiées sont bilinéaires.

- ▶ Ce n'est pas le seul produit scalaire que l'on peut attacher à E, néanmoins le programme se contente de celui-ci. Par exemple, l'application

$$\begin{cases} E \times E \longrightarrow \mathbf{R}, \\ (x, y) \longrightarrow 2\langle x|y \rangle, \end{cases}$$

est encore un produit scalaire sur E.

- ▶ À titre culturel, voici deux applications définissant des produits scalaires sur d'autres espaces vectoriels que \mathbf{R}^n et $\mathcal{M}_{n,1}(\mathbf{R})$.
 1. Si $E = \mathcal{C}^0([a, b], \mathbf{R})$ avec $a < b$. Alors $(f, g) \longrightarrow \int_a^b f(t)g(t) dt$ est un produit scalaire sur E.
 2. Si E est $\mathbf{R}_n[X]$ avec $n \geq 1$, alors $(P, Q) \longrightarrow \sum_{k=0}^n P(k)Q(k)$ est un produit scalaire sur E.

Produit scalaire euclidien

```
def produit_scalaire_eucl(X, Y):
    """
    X,Y deux vecteurs->produit scalaire euclidien des deux
    """
    S = 0
    for i in range(len(X)):
        S += X[i]*Y[i]
    return S
```

Par exemple,

```
>>> produit_scalaire_eucl([1,0,1], [-2,3,5])
3
>>> produit_scalaire_eucl([0,0,1,2,2], [3,3,3,0,2])
7
```

1.2. Norme & Distance

La notion de produit scalaire sert à mesurer l'orthogonalité, mais aussi et surtout des distances. Voyons comment.

Définition ALG.6.2 | Norme d'un vecteur

Soit $x \in E$. Alors on appelle *norme euclidienne de x* la quantité notée $\|x\|$, définie par :

$$\|x\| = \sqrt{\langle x|x \rangle}.$$

Remarque 1.3 — Notez que la norme est bien définie puisque $\langle x|x \rangle \geq 0$ pour tout $x \in E$, comme nous l'avons déjà montré.


Remarque 1.4 — Interprétation géométrique si $n = 2$ Géométriquement, cela correspond à la distance entre le point M de coordonnées celles du vecteurs x , et l'origine $O(0, 0)$.

Proposition ALG.6.2 | Identités de polarisation.

Soit $(x, y) \in E^2$, alors nous avons :

1. $\langle x|y \rangle = \frac{1}{2} (\|x + y\|^2 - \|x\|^2 - \|y\|^2)$,
2. $\langle x|y \rangle = \frac{1}{2} (\|x\|^2 + \|y\|^2 - \|x - y\|^2)$,
3. $\langle x|y \rangle = \frac{1}{4} (\|x + y\|^2 - \|x - y\|^2)$.

Preuve (Point clef — Développer les normes en les exprimant à l'aide du produit scalaire)

Rédigeons uniquement la première (démonstration identique pour les autres). 
On développe le carré de la norme :

$$\|x + y\|^2 = \langle x + y|x + y \rangle = \langle x|x \rangle + \langle x|y \rangle + \langle y|x \rangle + \langle y|y \rangle = \|x\|^2 + 2\langle x|y \rangle + \|y\|^2.$$

En isolant $\langle x|y \rangle$, on obtient 1).

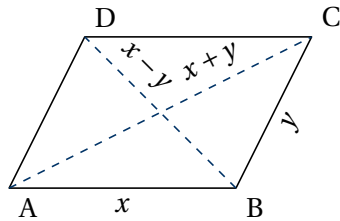
²Cette première formule est analogue à « $\text{Cov}(X, Y) = \frac{1}{2} (\text{Var}(X + Y) - \text{Var}(X) - \text{Var}(Y))$ »

Proposition ALG.6.3 | Identité du parallélogramme.

Soit $(x, y) \in E^2$, alors nous avons :

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

L'interprétation géométrique est la suivante : la somme des carrés des longueurs des diagonales d'un parallélogramme $A(0)B(y)C(x+y)D(x-y)$ est la somme des carrés des longueurs des côtés. Ce résultat peut se retrouver avec le théorème de PYTHAGORE dans le cas d'un rectangle.



Preuve



On développe le carré de la norme :

$$\begin{aligned} \|x + y\|^2 + \|x - y\|^2 &= \|x\|^2 + 2\langle x|y \rangle + \|y\|^2 + \|x\|^2 - 2\langle x|y \rangle + \|y\|^2 \\ &= 2(\|x\|^2 + \|y\|^2). \end{aligned}$$

Les deux techniques mises en jeu dans les preuves précédentes doivent être retenues comme une méthode, car elles sont fondamentales dans la pratique.

Méthode Développement d'une norme de somme au carré

Soit $\|x + y\|^2$ avec $x, y \in E$.

1. Écrire la quantité en fonction du produit scalaire : $\|x + y\|^2 = \langle x + y | x + y \rangle$.
2. Développer en utilisant la bilinéarité du produit scalaire.



Attention

On oublie de suite la formule *archi-fausse*³ suivante :

$$\|x + y\|^2 \neq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\|.$$

Proposition ALG.6.4 | Inégalité de CAUCHY-SCHWARZ

► **(Inégalité)** Pour tout $(x, y) \in E^2$:

$$|\langle x|y \rangle| \leq \|x\|\|y\|.$$

► **(Cas d'égalité)** De plus,

$$|\langle x|y \rangle| = \|x\|\|y\| \iff (x, y) \text{ est liée.}$$

Preuve (Point clef — Introduire la fonction $P : \lambda \in \mathbf{R} \mapsto \|x + \lambda y\|^2$, c'est un polynôme en λ)

Soit $\lambda \in \mathbf{R}$, alors $P(\lambda) = \langle x + \lambda y | x + \lambda y \rangle = \|x\|^2 + 2\lambda\langle x|y \rangle + \lambda^2\|y\|^2$. C'est un polynôme en λ de degré 1 ou 2.

- 1er cas : si $\|y\| = 0$, alors l'inégalité est triviale ($0 \leq 0$).
- 2ème cas : si $\|y\| > 0$, alors P est un trinôme, positif pour tout $\lambda \in \mathbf{R}$, donc de discriminant $\Delta = 4\langle x|y \rangle^2 - 4\|x\|^2\|y\|^2 \leq 0$, ce qui est équivalent à $\langle x|y \rangle^2 \leq \|x\|^2\|y\|^2$, il ne reste plus qu'à prendre la racine de chaque côté pour obtenir l'inégalité souhaitée.


Enfin, le cas d'égalité est obtenu lorsque :

$$\begin{aligned} \Delta = 0 &\iff P \text{ possède une racine double,} \\ &\iff P \text{ s'annule sur } \mathbf{R} \text{ (car positif),} \\ &\iff \exists \lambda \in \mathbf{R}, P(\lambda) = 0 = \|x + \lambda y\|^2 \iff x = \lambda y. \\ &\iff \exists \lambda \in \mathbf{R}, (x, y) \text{ liée.} \end{aligned}$$

Exemple 1 — Utilisations de l'inégalité de CAUCHY-SCHWARZ

1. Montrer que : pour tout $(x, y, z) \in \mathbf{R}^3$, $x + y + z \leq \sqrt{3}\sqrt{x^2 + y^2 + z^2}$.

³Elle est d'ailleurs vraie uniquement dans les cas où (x, y) réalise l'égalité dans CAUCHY-SCHWARZ

 Nous avons, en notant $X = (x, y, z)$ et $\mathbb{1} = (1, 1, 1)$, $x + y + z = \langle X | \mathbb{1} \rangle$. Donc par l'inégalité de CAUCHY-SCHWARZ, on obtient :

$$\langle X | \mathbb{1} \rangle \leq |\langle X | \mathbb{1} \rangle| \leq \|\mathbb{1}\| \|X\| = \sqrt{3} \sqrt{x^2 + y^2 + z^2}.$$

2. (Utilisation du cas d'égalité) Résoudre le système non linéaire ci-après :

$$\begin{cases} x_1 + x_2 + \dots + x_n = n, \\ x_1^2 + x_2^2 + \dots + x_n^2 = n. \end{cases}$$



■ Analyse – Soit $x = (x_1, \dots, x_n) \in \mathbf{R}^n$ une solution du système. Alors d'après l'égalité de CAUCHY-SCHWARZ :

$$n = x_1 + \dots + x_n = \langle x | (1, \dots, 1) \rangle \leq |\langle x | (1, \dots, 1) \rangle| \leq \|x\| \|(1, \dots, 1)\| = \sqrt{x_1^2 + \dots + x_n^2} \sqrt{n}.$$

En utilisant la seconde ligne du système, il vient :

$$n = x_1 + \dots + x_n = \langle x | (1, \dots, 1) \rangle \leq |\langle x | (1, \dots, 1) \rangle| = \sqrt{n} \sqrt{n} = n.$$

Il y a donc égalité dans l'inégalité de CAUCHY-SCHWARZ. Ainsi, les vecteurs x et $(1, \dots, 1)$ sont liés, i.e. : $\exists \lambda \in \mathbf{R}, x_1 = x_2 = \dots = \lambda$.

■ Synthèse. Inversement, soit $x = (\lambda, \dots, \lambda)$ avec $\lambda \in \mathbf{R}$. Alors x est solution du système si et seulement si

$$\begin{cases} n\lambda = n, \\ n\lambda^2 = n, \end{cases} \iff \begin{cases} \lambda = 1, \\ n = n. \end{cases}$$

En résumé, l'unique solution du système est $(1, \dots, 1)$.



Méthode Majorer des sommes avec CAUCHY-SCHWARZ

Interpréter une somme comme un produit scalaire. On retiendra l'exemple très classique suivant : si $(x_1, \dots, x_n) \in \mathbf{R}^n$,

$$\left| \sum_{i=1}^n x_i \right| = |\langle (1, \dots, 1) | (x_1, \dots, x_n) \rangle| \leq \sqrt{n} \sum_{i=1}^n x_i^2.$$

Proposition ALG.6.5 | Propriétés de norme


L'application $\|\cdot\| : \begin{cases} E \rightarrow \mathbf{R} \\ x \mapsto \|x\| \end{cases}$ est une *norme* sur E , i.e.

1. **(Homogénéité)** $\forall \lambda \in \mathbf{R}, x \in E, \|\lambda x\| = |\lambda| \|x\|$.
2. **(Positivité)** $\forall x \in E, \|x\| \geq 0$.
3. **(Inégalité triangulaire)** $\forall x, y \in E, \|x + y\| \leq \|x\| + \|y\|$. De plus, l'égalité est réalisée si et seulement si (x, y) est *positivement liée* : $\exists \lambda \in \mathbf{R}^+$, $x = \lambda y$.
4. **(Séparation)** $\forall x \in E, \|x\| = 0 \iff x = 0$.

Preuve Dans toute la preuve, nous nous plaçons dans le cas $E = \mathbf{R}^n$, l'autre se montrant de la même manière. Soient $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n)$ deux vecteurs de \mathbf{R}^n , et $\lambda \in \mathbf{R}$.

1. 

$$\|\lambda x\| = \sqrt{\sum_{i=1}^n (\lambda x_i)^2} = \sqrt{\lambda^2 \sum_{i=1}^n x_i^2} = |\lambda| \|x\|.$$

2. 

$$\|x\| = \sqrt{\sum_{i=1}^n x_i^2} \geq 0.$$

3. (Point clef – **Montrer l'inégalité élevée au carré**)



Nous avons, puisque $\sqrt{\|x + y\|^2} = \|\|x + y\|\| = \|x + y\|$ et que la fonction racine est croissante :

$$\|x + y\| \leq \|x\| + \|y\| \iff \|x + y\|^2 \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| \quad (\star).$$

Il suffit donc de montrer (\star) . Mais comme,

$$\begin{aligned} \|x + y\|^2 &= \|x\|^2 + \|y\|^2 + 2\langle x | y \rangle \\ &\leq \|x\|^2 + \|y\|^2 + 2|\langle x | y \rangle| \leq_{\text{CAUCHY-SCHWARZ}} \|x\|^2 + \|y\|^2 + 2\|x\|\|y\|, \end{aligned}$$

il vient (\star) . L'inégalité triangulaire est démontrée.

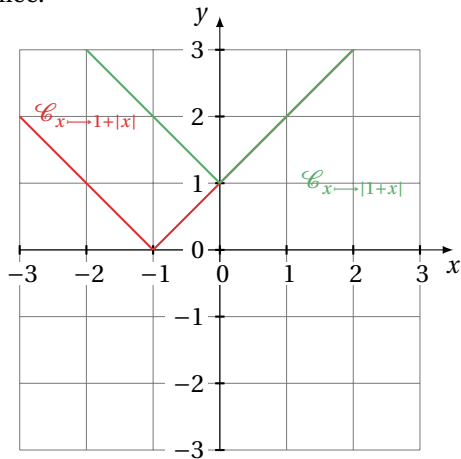
Reste à présent le cas d'égalité, il y a égalité si et seulement si $\langle x | y \rangle = \|x\|\|y\|$ — attention ceci n'est pas le cas d'égalité dans CAUCHY-SCHWARZ, du fait de l'absence de valeurs absolues autour du produit scalaire.

Supposons que $x = \lambda y$ avec $\lambda \in \mathbf{R}^+$, alors on vérifie sans difficulté que (x, y) réalise l'égalité dans l'inégalité triangulaire : $(1 + \lambda)\|y\| = (1 + \lambda)\|y\|$.

Inversement, supposons que $\langle x | y \rangle = \|x\|\|y\|$. Alors : $|\langle x | y \rangle| = \|x\|\|y\|$ en prenant la valeur absolue, donc (x, y) réalise l'égalité dans l'inégalité de CAUCHY-SCHWARZ. Donc il existe $\lambda \in \mathbf{R}, x = \lambda y$. Mais en injectant dans l'égalité triangulaire, on trouve :

$$\|1 + \lambda\| \|y\| = (1 + |\lambda|) \|y\|.$$

Si $\|y\| = 0$ (i.e. $y = 0$), alors il y a évidemment égalité.
 Si $y \neq 0$, alors $|1 + \lambda| = 1 + |\lambda|$ ce qui implique $\lambda \geq 0$. Donc (x, y) est positivement liée.



4.

$$\|x\| = 0 \iff \sqrt{\sum_{i=1}^n x_i^2} = 0 \iff \sum_{i=1}^n x_i^2 = 0 \iff \forall i \in \llbracket 1, n \rrbracket, x_i = 0.$$

Maintenant que l'on sait mesurer des vecteurs, on peut s'intéresser à comment on mesure l'écart entre deux vecteurs.

Définition ALG.6.3 | Distance entre deux vecteurs

Soient $x, y \in E$. Alors on appelle *distance euclidienne entre x et y* la quantité notée $d(x, y)$, définie par :

$$d(x, y) = \|x - y\|.$$

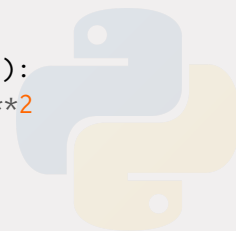
En particulier, $d(x, 0_E) = \|x\|$.

Distance euclidienne

```
import math as ma
def distance_eucl(X, Y):
```



```
"""
X,Y deux vecteurs->distance euclidienne entre les deux
"""
D = 0
for i in range(len(X)):
    D += (X[i]-Y[i])**2
return ma.sqrt(D)
```



Par exemple :

```
>>> distance_eucl([5, 5, -3],[5, 2, 7])
10.44030650891055
```

Enfin, nous arrivons à l'expression générale de la distance à une partie de E.

Définition/Proposition ALG.6.2 | Distance entre un vecteur et une partie de E

Soient $x \in E$ et $A \subset E$ **non vide**. Alors on appelle *distance euclidienne entre x et A* la quantité notée $d(x, A)$, définie par :

$$d(x, A) = \inf_{y \in A} d(x, y) = \inf_{y \in A} \|x - y\|.$$

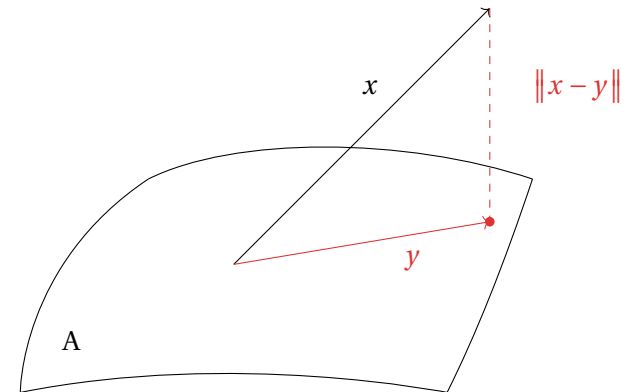



FIG. ALG.6.1. : La distance entre x et A est la plus petite des distances en pointillés.

Remarque 1.5 — On rappelle que la borne inférieure d'une partie est, par définition,

le plus grand minorant de la partie (s'il existe, c'est le cas lorsque la partie est non vide et minorée). De plus, $\delta = \inf_{y \in A} \|x - y\|$ si et seulement si :

- ▶ δ est un minorant de $\{\|x - y\|, y \in A\}$,
- ▶ il existe une suite (y_n) de A telle que : $\|x - y_n\| \xrightarrow{n \rightarrow \infty} \delta$.

Preuve Justifions l'existence de cette borne inférieure.  Toute partie non vide minorée admet une borne inférieure. L'ensemble

$$\{\|x - y\|, y \in A\}$$


est une partie minorée par zéro, elle est non vide car A est non vide, donc $d(x, A)$ existe.

Dans la prochaine section, nous verrons que, si A est un sous-espace vectoriel de E alors on a une expression explicite de $d(u, A)$ pour tout x : cette expression fera intervenir la *projection orthogonale de u sur A*. Voyons deux exemples où A n'en est pas un.

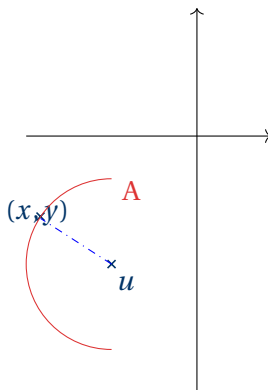
Exemple 2 — Distances à des cercles

1. Soient $u = (-2, -3)$ et

$$A = \left\{ (x, y) \in \mathbf{R}^2, (x + 2)^2 + (y + 3)^2 = 4, \quad x \leq -2 \right\},$$

on a : $d(u, A) = 2$.  Géométriquement, on constate que tous les points de A sont à distance 2 de u, en effet : si $(x, y) \in A$ vérifie $(x + 2)^2 + (y + 3)^2 = 4$, alors $\sqrt{(x + 2)^2 + (y + 3)^2} = \|(x, y) - u\| = 2$. Autrement dit, pour tout $(x, y) \in A$, $d((x, y), u) = 2$ donc $d(u, A) = 2$.

2. Soient $u = (-3, 0)$ et $B = \left\{ (x, y) \in \mathbf{R}^2, \sqrt{x^2 + y^2} < 2 \right\}$, on a : $d(u, B) = 1$.



L'ensemble B est le disque ouvert $\mathcal{C}(0, 2)$, l'ouverture signifiant que l'on enlève la frontière. Nous avons

$$d(u, B) = \inf_{(x, y) \in B} \|(-3, 0) - (x, y)\| = \inf_{(x, y) \in \mathbf{R}^2, \sqrt{x^2 + y^2} < 2} \sqrt{(-3 - x)^2 + y^2}.$$

Géométriquement, on conjecture que la distance (verte sur le dessin) vaut 1, et que la borne inférieure n'est pas atteinte (car la frontière est exclue). Pour montrer que $1 = d(u, B)$ il faut et il suffit d'établir que : 1 est un minorant, et qu'il existe une suite de points de B notée (b_n) telle que $\|u - b_n\| \xrightarrow{n \rightarrow \infty} 1$.

Le nombre 1 est un minorant des distances, en effet, d'après le théorème de PYTHAGORE :

$$\forall y \in B, \quad \|u - y\| \geq \|u - z\|,$$

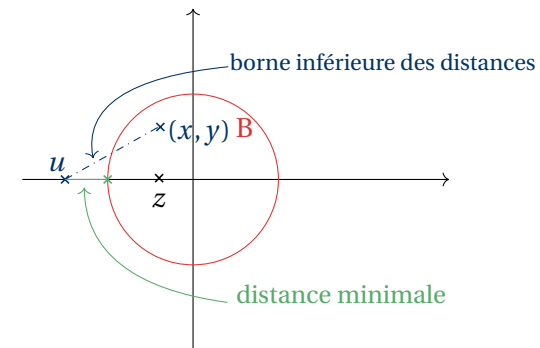
où z est défini sur le dessin, or $\|u - z\| \geq 1$. Donc :

$$\forall y \in B, \quad \|u - y\| \geq 1,$$

ceci prouve que 1 est un minorant de notre partie. Notons

$$b_n = \left(-\left(2 - \frac{1}{n}\right), 0 \right) = \left(\frac{1}{n} - 2, 0 \right),$$

alors $b_n \in B$ pour tout $n \in \mathbf{N}^*$, et $\|u - b_n\| = \sqrt{\left(-3 + 2 - \frac{1}{n}\right)^2 + 0^2} \xrightarrow{n \rightarrow \infty} 1$. C'est terminé.



2. ORTHOGONALITÉ

Définition ALG.6.4 | Vecteurs orthogonaux

Soient x et y deux vecteurs de E . Alors x et y sont dits *orthogonaux* si :

$$\langle x|y \rangle = 0.$$

Σ Notation

On note généralement $x \perp y$ pour désigner l'orthogonalité de deux vecteurs de E .

Définition ALG.6.5 | Parties orthogonales

Soit $A, B \subset E$ deux parties de E . Alors A et B sont dites *orthogonales* si :

$$\forall (x, y) \in A \times B, \quad x \perp y.$$

Σ Notation

On note généralement $A \perp B$ pour désigner l'orthogonalité de deux parties de E .

2.1. Orthogonal d'une partie


Définition/Proposition ALG.6.3 | Orthogonal d'une partie de E

Soit $A \subset E$ une partie de E . Alors

$$A^\perp = \{x \in E, \forall a \in A, \langle x|a \rangle = 0\}$$

est un sous-espace vectoriel de E . On l'appelle *l'orthogonal de A* dans E , et on a $A \perp A^\perp$.


Le point remarquable est que même si A n'est pas un sous-espace vectoriel de E , alors son orthogonal en sera un.


Preuve  Soient $x, y \in A^\perp$ et $\lambda, \mu \in \mathbf{R}$, montrons que $\lambda x + \mu y \in A^\perp$. Considérons $a \in A$, alors :

$$\langle \lambda x + \mu y | a \rangle = \lambda \langle x | a \rangle + \mu \langle y | a \rangle = \lambda 0 + \mu 0 = 0.$$

Donc A^\perp est stable par combinaison linéaire. De plus $0_E \in A^\perp$ car pour tout $a \in A$, $\langle 0_E | a \rangle = 0$, donc A^\perp est un sous-espace vectoriel de E .

Exemple 3 —

1. $E^\perp = \{0_E\}$. 
En effet, si $x \in E^\perp$, alors en particulier $\langle x|x \rangle = 0 = \|x\|^2$, ceci implique $x = 0$ par propriété de la norme euclidienne. Inversement, $0_E \in E^\perp$ puisque E^\perp est un espace vectoriel. Conclusion : $E^\perp = \{0_E\}$.

2. $\{0_E\}^\perp = E$. 
En effet,

$$x \in \{0_E\}^\perp \iff \forall y \in \{0_E\}, \langle x|y \rangle = 0 \iff x \in E,$$

puisque $\langle x|0 \rangle = 0$ pour tout $x \in E$. D'où l'égalité : $\{0_E\}^\perp = E$.

Proposition ALG.6.6 | Si on connaît une famille génératrice de A

Supposons que $A = \text{Vect}(e_1, \dots, e_m)$, avec $m \geq 1$, en particulier A est un sous-espace vectoriel de \mathbf{R}^n . Alors :

$$x \in A^\perp \iff x \perp e_i \quad \forall i \in \llbracket 1, m \rrbracket.$$

Cette propriété se comprend aisément sur le dessin ci-après. Pour être dans l'orthogonal d'une partie engendrée par (e_1, e_2) , il faut et il suffit d'être orthogonal aux deux vecteurs e_1, e_2 .

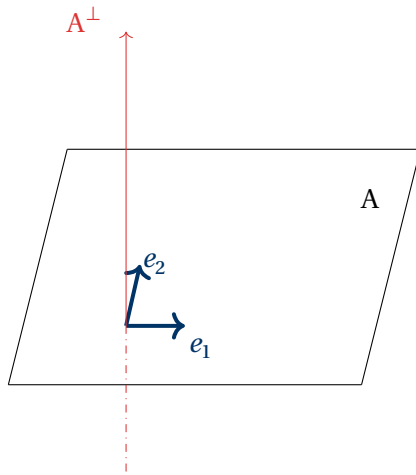



FIG. ALG.6.2. : Orthogonal d'un Vect



une famille génératrice de chaque espace.  Notons $e_1 = (1, 2, 0)$ et $e_2 = (0, 1, 3)$.



$$\begin{aligned}
 (x, y, z) \in E^\perp &\iff \begin{cases} (x, y, z) \perp e_1 \\ (x, y, z) \perp e_2 \end{cases} \\
 &\iff \begin{cases} x + 2y = 0, \\ y + 3z = 0 \end{cases} \\
 &\iff \begin{cases} x = -2y \\ y = -3z \end{cases} \\
 &\iff \begin{cases} x = 6z, \\ y = -3z \end{cases} \\
 &\iff (x, y, z) = (6z, -3z, z) = z(6, -3, 1) \\
 &\iff (x, y, z) \in \text{Vect}(6, -3, 1).
 \end{aligned}$$

Donc $F^\perp = \text{Vect}(6, -3, 1)$. Puis $(x, y, z) \in F^\perp$ si et seulement si $(x, y, z) \perp (6, -3, 1)$. Cela signifie que $6x - 3y + z = 0$. C'est l'équation d'un plan de \mathbf{R}^3 , une famille génératrice est donnée par

$$F^{\perp\perp} = \text{Vect}((0, 1, 3), (1, 0, -6)) = .$$

Preuve

  Soit $x \in A^\perp$, alors x est orthogonal à tous les éléments de A , donc en particulier à e_1, \dots, e_m .

  Soit $x \in E$ tel que $\langle x | e_i \rangle = 0, \quad \forall i \in \llbracket 1, m \rrbracket$ et montrons que $x \in A^\perp$. Alors soit $a = \sum_{i=1}^m \lambda_i e_i$ un élément de A , avec $\lambda_i \in \mathbf{R}$ pour tout $i \in \llbracket 1, m \rrbracket$. Alors :

$$\langle x | a \rangle = \left\langle x \left| \sum_{i=1}^m \lambda_i e_i \right. \right\rangle = \sum_{i=1}^m \lambda_i \langle x | e_i \rangle = 0. \quad \text{Donc } x \in A^\perp.$$

Exemple 4 — Calculer F^\perp puis $F^{\perp\perp} = (F^\perp)^\perp$ où $F = \text{Vect}((1, 2, 0), (0, 1, 3))$. On donnera

$$\dim F + \dim F^\perp = \dim E = n,$$

Remarque 2.1 — Dans $E = \mathbf{R}^n$ ou $E = \mathfrak{M}_{n,1}(\mathbf{R})$, nous avons toujours

i.e. $2 + 1 = 3$ sur la figure précédente. Cette propriété se comprend aisément géométriquement mais nous n'avons pas les outils mathématiques pour la démontrer pour le moment. En revanche, rien ne vous empêche de vous en servir pour contrôler *a posteriori* vos résultats.

2.2. Familles orthogonales & orthonormales

Définition ALG.6.6 | Familles orthogonales et orthonormales


Soit $(e_i)_{i \in I}$ avec I un ensemble une famille de vecteurs de E .

1. On dit que $(e_i)_{i \in I}$ est une famille orthogonale si :

$$\forall i \neq j \in I, \quad \langle e_i | e_j \rangle = 0.$$

2. On dit que $(e_i)_{i \in I}$ est une famille orthonormale si :

- ▶ elle est orthogonale,
- ▶ et pour tout $i \in I$, $\|e_i\| = 1$.

Exemple 6 — La base canonique de E est orthonormale La base canonique \mathcal{B}^{can} de E est-elle orthogonale? orthonormale?  Oui. Notons (e_1, \dots, e_n) la base canonique de E . Si $i \neq j$, alors $\langle e_i | e_j \rangle = 0$ étant donné que « les 1 » sont à des positions différentes. La famille \mathcal{B}^{can} est donc orthogonale. De plus, pour $i \in \llbracket 1, n \rrbracket$, $\|e_i\|^2 = 0^2 + \dots + 0^2 + 1^2 + 0^2 = 1^2$. C'est donc en résumé une base orthonormée.


Rappelons que toute famille de vecteurs propres associés à des valeurs propres distinctes d'un endomorphisme est libre. Nous voyons dans ce chapitre une autre condition suffisante garantissant la liberté : l'orthogonalité.

Proposition ALG.6.7 | Liberté d'une famille orthogonale

Soient $m \geq 1$, (e_1, \dots, e_m) une famille de vecteurs de E . On suppose que :

1. pour tout $i \in \llbracket 1, m \rrbracket$, $e_i \neq 0$,
 2. (e_1, \dots, e_m) est orthogonale,⁴
- alors la famille (e_1, \dots, e_m) est libre.

⁴Inutile qu'elle soit orthonormée, le fait que les vecteurs soient non nuls suffit

Preuve  Soient $\lambda_1, \dots, \lambda_m \in \mathbf{R}$ telles que $\sum_{i=1}^m \lambda_i e_i = 0$. Alors soit $j \in \llbracket 1, m \rrbracket$. Nous avons alors en prenant le produit scalaire avec e_j :

$$\left\langle \sum_{i=1}^m \lambda_i e_i \middle| e_j \right\rangle = \sum_{i=1}^m \lambda_i \langle e_i | e_j \rangle = \lambda_j \langle e_j | e_j \rangle = 0$$

en utilisant l'orthogonalité de la famille. Donc comme $\langle e_j | e_j \rangle = \|e_j\|^2 \neq 0$ (sinon $e_j = 0$ et c'est une contradiction par rapport aux hypothèses du théorème), on obtient

$$\lambda_j = 0.$$

Faisant cela pour tout j , on obtient la liberté de la famille.

Théorème ALG.6.1 | Pythagore

Soient (e_1, \dots, e_m) une famille orthogonale de m vecteurs de E . Alors :

$$\|e_1 + \dots + e_m\|^2 = \sum_{i=1}^m \|e_i\|^2.$$

De plus, pour deux vecteurs (e_1, e_2) , on a l'équivalence suivante :

$$\|e_1 + e_2\|^2 = \|e_1\|^2 + \|e_2\|^2 \iff e_1 \perp e_2.$$

Preuve 

$$\begin{aligned} \|e_1 + \dots + e_m\|^2 &= \left\langle \sum_{i=1}^n e_i \middle| \sum_{j=1}^m e_j \right\rangle \\ &= \sum_{i=1}^n \sum_{j=1}^m \langle e_i | e_j \rangle \\ &= \sum_{i=1}^n \langle e_i | e_i \rangle + 0 \\ &= \sum_{i=1}^n \|e_i\|^2. \end{aligned} \quad \left. \begin{array}{l} \text{bilinéarité du produit scalaire} \\ \langle e_i | e_j \rangle = 0 \text{ si } i \neq j \end{array} \right\}$$

Pour la dernière partie du théorème, on écrit

$$\begin{aligned} \|e_1 + e_2\|^2 &= \langle e_1 + e_2 | e_1 + e_2 \rangle = \|e_1\|^2 + \|e_2\|^2 \\ \iff \|e_1\|^2 + \|e_2\|^2 + 2\langle e_1 | e_2 \rangle &= \|e_1\|^2 + \|e_2\|^2, \\ \iff 2\langle e_1 | e_2 \rangle &= 0 \\ \iff e_1 \perp e_2. \end{aligned}$$

2.3. Coordonnées d'un vecteur dans une base orthonormale

Théorème ALG.6.2 | Expression du produit scalaire euclidien en base orthonormée

Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base **orthonormale** de E. Soient $x, y \in E$, on note :

$$X = \underset{\mathcal{B}}{\text{Mat}}(x), \quad Y = \underset{\mathcal{B}}{\text{Mat}}(y).$$

Alors : $\langle x | y \rangle = {}^T X \times Y$ et en particulier $\|x\| = \sqrt{{}^T X X}$.

Ce théorème généralise la Proposition ALG.6.1 à une base quelconque orthonormée.

Preuve Notons $X = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$ et $Y = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}$ les matrices données dans l'énoncé, où pour tout $i \in \{1, \dots, n\}$, $\lambda_i \in \mathbf{R}, \mu_i \in \mathbf{R}$. Par définition d'une matrice de vecteur, cela signifie que :

$$X = \sum_{i=1}^n \lambda_i e_i, \quad Y = \sum_{j=1}^n \mu_j e_j.$$

Alors calculons $\langle x | y \rangle$ en utilisant ces relations :

$$\begin{aligned} \langle x | y \rangle &= \left\langle \sum_{i=1}^n \lambda_i e_i \left| \sum_{j=1}^n \mu_j e_j \right. \right\rangle && \left. \begin{array}{l} \text{définition d'une matrice de vecteur} \\ \text{bilinéarité du produit scalaire} \\ \text{orthonormalité de la famille, produit scalaire nul dès} \\ \text{que } i \neq j \end{array} \right. \\ &= \sum_{i=1}^n \sum_{j=1}^n \lambda_i \mu_j \langle e_i | e_j \rangle \\ &= \sum_{i=1}^n \lambda_i \mu_i \\ &= {}^T X \cdot Y. \end{aligned}$$

La formule est donc établie.

Théorème ALG.6.3 | Expression/Norme d'un vecteur en base orthonormée

1. Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base **orthonormale** de E et $x \in E$. Alors les coordonnées de x dans \mathcal{B} s'expriment en terme de produit scalaire :

$$x = \sum_{i=1}^n \langle x | e_i \rangle e_i \quad \iff \quad (\|x\|^2 = \sum_{i=1}^n |\langle x | e_i \rangle|^2).$$

Ainsi, si \mathcal{B} est une base orthonormale,


$$\underset{\mathcal{B}}{\text{Mat}}(x) = \begin{pmatrix} \langle x | e_1 \rangle \\ \vdots \\ \langle x | e_n \rangle \end{pmatrix}.$$

2. Soit $\mathcal{B} = (e_1, \dots, e_p)$ une base orthonormale d'un sous-espace vectoriel F de E, et $x \in F$. Alors les coordonnées de x dans \mathcal{B} s'expriment en terme de produit scalaire :

$$x = \sum_{i=1}^p \langle x | e_i \rangle e_i \quad \iff \quad (\|x\|^2 = \sum_{i=1}^p |\langle x | e_i \rangle|^2).$$

Preuve

1. Premièrement, il existe des coordonnées : $\exists (\lambda_i)_{i=1}^n, \quad x = \sum_{i=1}^n \lambda_i e_i$. Deuxièmement, prenant le produit scalaire avec $e_j, j \in \llbracket 1, n \rrbracket$ de chaque côté de l'égalité,

nous obtenons : 

$$\langle e_j | x \rangle = \left\langle e_j \left| \sum_{i=1}^n \lambda_i e_i \right. \right\rangle = \sum_{i=1}^n \lambda_i \langle e_j | e_i \rangle = \lambda_j \times 1$$

puisque \mathcal{B} est orthonormale. D'où l'expression de x donnée dans l'énoncé. L'égalité avec des normes au carré découle directement du théorème de PYTHAGORE puisque la famille $(\langle x | e_1 \rangle e_1, \dots, \langle x | e_n \rangle e_n)$ est orthogonale et que $\|e_i\|^2 = 1$ pour tout i .

2. Preuve identique à la première.

Proposition ALG.6.8 | Orthonormalité d'une famille et matrice de passage

Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base orthonormale de E et $\mathcal{C} = (f_1, \dots, f_n)$ une base quelconque de E . Alors :

$$\mathcal{C} \text{ est orthonormale} \iff {}^T(P^{\mathcal{B} \rightarrow \mathcal{C}}) \cdot P^{\mathcal{B} \rightarrow \mathcal{C}} = I_n.$$

Plus généralement, une matrice $P \in \mathfrak{M}_{n,n}(\mathbf{R})$ satisfaisant ${}^T P P = I_n$ est appelée *matrice orthogonale*.

Preuve Notons dans la preuve $P = P^{\mathcal{B} \rightarrow \mathcal{C}}$. Par définition d'une matrice de passage et par bilinéarité de $\langle \cdot | \cdot \rangle$:

$$\langle f_i | f_j \rangle = \left\langle \sum_{k=1}^n P_{k,i} e_k \left| \sum_{\ell=1}^n P_{\ell,j} e_\ell \right. \right\rangle = \sum_{k=1}^n \sum_{\ell=1}^n P_{k,i} P_{\ell,j} \langle e_k | e_\ell \rangle.$$

Rappelons de plus que la famille (e_1, \dots, e_n) est orthonormée, ainsi

$$\langle f_i | f_j \rangle = \sum_{k=1}^n P_{k,i} P_{k,j} \mathbf{1} + 0.$$

D'autre part, le terme (i, j) avec $(i, j) \in \llbracket 1, n \rrbracket^2$ de ${}^T P \cdot P$ est :

$$({}^T P \cdot P)_{i,j} = \sum_{k=1}^n [{}^T P]_{i,k} P_{k,j} = \sum_{k=1}^n P_{k,i} P_{k,j} = \langle f_i | f_j \rangle.$$

Ainsi, notant $\delta_{i,j} = 1$ si $i = j$ et zéro sinon, nous obtenons le critère du théorème :


$$\begin{aligned} \mathcal{C} \text{ est orthonormale} &\iff \forall i, j \in \llbracket 1, n \rrbracket, \langle f_i | f_j \rangle = \delta_{i,j} \\ &\iff \forall i, j \in \llbracket 1, n \rrbracket, [{}^T P \cdot P]_{i,j} = \delta_{i,j} \\ &\iff {}^T P \cdot P = I_n. \end{aligned}$$

Corollaire ALG.6.2 | Propriétés d'une matrice de passage entre bases orthonormées

Soient $\mathcal{B} = (e_1, \dots, e_n)$ et $\mathcal{C} = (f_1, \dots, f_n)$ deux bases orthonormales de E , et $P = P^{\mathcal{B} \rightarrow \mathcal{C}}$. Alors :

$$P \text{ est inversible, et } P^{-1} = {}^T P.$$

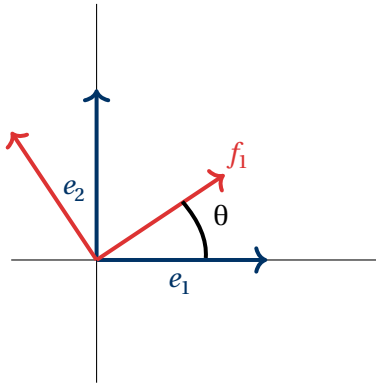
Preuve Conséquence directe de la proposition précédente.

Exemple 7 — On considère $e_1 = (1, 0)$ et $e_2 = (0, 1)$ la base canonique de \mathbf{R}^2 et $f_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, $f_2 = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$. Vérifier que (f_1, f_2) est une base orthonormée, calculer la matrice de passage $P^{(e_1, e_2) \rightarrow (f_1, f_2)}$ et vérifier qu'elle est orthogonale. 

2.4. Diagonalisation des matrices symétriques réelles

Exemple 8 – Les matrices de rotation sont orthogonales.

On considère $e_1 = (1, 0)$ et $e_2 = (0, 1)$ la base canonique de \mathbb{R}^2 et (f_1, f_2) les vecteurs e_1, e_2 tournés d'un angle $\theta \in [0, 2\pi[$ dans le sens trigonométrique.



1. Calculer les coordonnées de f_1 et f_2 dans la base canonique.
2. Expliciter la matrice de passage de (e_1, e_2) à (f_1, f_2) , on l'appelle *matrice de rotation d'angle θ* , notée R_θ , qu'en dire? Calculer son inverse. Est-ce cohérent géométriquement?

1. Nous avons, avec un peu de trigonométrie, $f_1 = (\cos\theta, \sin\theta)$ et $f_2 = (-\sin\theta, \cos\theta)$.
2. Par définition d'une matrice de changement de base, nous avons

$$P^{(e_1, e_2) \rightarrow (f_1, f_2)} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \stackrel{\text{(déf.)}}{=} R_\theta.$$

Ainsi, comme $\det(P^{(e_1, e_2) \rightarrow (f_1, f_2)}) = \cos^2\theta + \sin^2\theta = 1$, on obtient l'inverse de $P^{(e_1, e_2) \rightarrow (f_1, f_2)}$:

$$P^{(e_1, e_2) \rightarrow (f_1, f_2)}^{-1} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = {}^T P^{(e_1, e_2) \rightarrow (f_1, f_2)}.$$

La matrice de passage précédente est donc une matrice orthogonale.

On a

$$(P^{(e_1, e_2) \rightarrow (f_1, f_2)})^{-1} = P^{(f_1, f_2) \rightarrow (e_1, e_2)} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{pmatrix} = R_{-\theta}.$$

On a appliqué une rotation d'angle θ pour « passer de » (e_1, e_2) , le calcul qui précède nous apprend que pour « passer de » (f_1, f_2) à (e_1, e_2) , on applique une rotation d'angle $-\theta$: c'est donc cohérent géométriquement.

Nous reprenons le résultat du chapitre de réduction, en comprenant un peu plus la signification de « ${}^T P = P^{-1}$ » qui intervenait dans le résultat.

Proposition ALG.6.9 | Orthogonalité des espaces propres

1. Soit $f \in \mathcal{L}(\mathbb{R}^n)$ de matrice A symétrique réelle dans la base canonique \mathcal{B}^{can} de \mathbb{R}^n , possédant au moins deux valeurs propres distinctes, alors :

$$E_\lambda(f) \perp E_\mu(f).$$

2. Soit A une matrice symétrique possédant au moins deux valeurs propres distinctes. Si $\lambda \neq \mu$ sont deux valeurs propres de A , alors

$$E_\lambda(A) \perp E_\mu(A).$$

Preuve

1. Soient $\lambda \neq \mu \in \text{Spec } f$, et $x \in E_\lambda(f), y \in E_\mu(f)$. Alors $f(x) = \lambda x$ et $f(y) = \mu y$. Il faut montrer que $\langle x|y \rangle = 0$. Notons de plus $X = \mathcal{M}^{\text{can}}(x), Y = \mathcal{M}^{\text{can}}(y)$:

$$\begin{aligned} \lambda \langle x|y \rangle &= \langle \lambda x|y \rangle = \langle f(x)|y \rangle \\ &= {}^T (AX) Y \quad \left. \begin{array}{l} \text{)} \mathcal{B}^{\text{can}} \text{ orthonormale} \\ \text{)} A \text{ symétrique} \end{array} \right\} \\ &= {}^T X ({}^T AY) \\ &= {}^T X (AY) \\ &= \langle x|f(y) \rangle = \mu \langle x|y \rangle. \quad \left. \begin{array}{l} \text{)} A \text{ symétrique} \\ \text{)} \mathcal{B}^{\text{can}} \text{ orthonormale} \end{array} \right\} \end{aligned}$$

Donc : $(\lambda - \mu) \langle x|y \rangle = 0$, or $\lambda \neq \mu$ donc $\langle x|y \rangle = 0$. Finalement deux vecteurs propres associés à deux valeurs propres distinctes sont orthogonaux.

2. Même preuve que précédemment, ou considérer l'endomorphisme canoniquement associé à A .

Théorème ALG.6.4 | Théorème spectral

1. Soit A une matrice **symétrique réelle**. Alors A est diagonalisable, et il existe une matrice D diagonale réelle, et une matrice $P \in \text{GL}_n(\mathbb{R})$, telles que :

$$A = PDP^{-1} \quad \text{et} \quad {}^T P = P^{-1}.$$

2. Soit $f \in \mathcal{L}(\mathbf{R}^n)$ de matrice A symétrique réelle dans la base canonique \mathcal{B}^{can} de \mathbf{R}^n . Alors :

$$\exists \mathcal{B} \text{ base orthonormée de } \mathbf{R}^n, \mathcal{M}_{\mathcal{B}}(f) \text{ soit diagonale réelle.}^5$$

En particulier, les espaces propres de A (ou f) sont orthogonaux.

Attention Le résultat est faux pour les matrices symétriques complexes.⁶


Preuve

- Résultat admis.
- Notons $A = \mathcal{M}_{\mathcal{B}^{\text{can}}}(f)$. Alors d'après 1), puisque A est symétrique réelle, il existe $P \in GL_n(\mathbf{R})$, telles que :

$$A = PDP^{-1} \quad \text{et} \quad {}^tP = P^{-1}.$$

Ainsi, puisque \mathcal{B}^{can} est orthonormale pour le produit scalaire euclidien et que ${}^tP = P^{-1}$, la Proposition ALG.6.8 nous dit que \mathcal{B} est elle aussi orthonormale. De plus D est la matrice de f dans \mathcal{B} d'après les formules de changement de base. C'est terminé : nous avons bien montré que f est diagonalisable en base orthonormée.

Remarque 2.2 — Il est donc possible de retenir le théorème spectral de la manière suivante : « toute matrice symétrique réelle est diagonalisable en base orthonormée (ou plus précisément son endomorphisme canoniquement associé) et est à valeurs propres réelles ».

Exemple 9 — Contre-exemple La matrice $A = \begin{pmatrix} i & 1 \\ 1 & -i \end{pmatrix}$ n'est pas diagonalisable à valeurs propres réelles.  Soit $\lambda \in \mathbf{R}$. Alors $\det(A - \lambda I_2) = \begin{vmatrix} i - \lambda & 1 \\ 1 & -i - \lambda \end{vmatrix} = (i - \lambda)(-i - \lambda)$

⁵On dit que f est diagonalisable en base orthonormée

⁶Pour les matrices complexes S , la bonne notion à considérer n'est pas ${}^tS = S$ mais $\overline{{}^tS} = S$: on parle dans ce cas de matrices hermitiennes plutôt que symétriques, mais leur étude n'est plus au programme de CPGE.

$\lambda) - 1 = \lambda^2$, donc A possède une unique valeur propre qui est zéro. Donc si elle était diagonalisable, nous aurions $A = 0$ et ceci est faux. Le théorème spectral ne s'applique donc pas.

Exemple 10 — Diagonaliser en base orthonormée la matrice $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$. 

— **Calcul des éléments propres.** Avec la chaîne d'opérations élémentaires $L_1 \leftrightarrow L_3, L_2 \leftarrow L_2 - L_1, L_3 \leftarrow L_3 - (2 - \lambda)L_1$ réalisée sur $A - \lambda I_3$ avec $\lambda \in \mathbf{R}$, nous obtenons :

$$A - \lambda I_3 \underset{L}{\sim} \begin{pmatrix} 1 & 1 & 2 - \lambda \\ 0 & 1 - \lambda & \lambda - 1 \\ 0 & 0 & -\lambda^2 + 5\lambda - 4 \end{pmatrix}.$$

On constate alors que :

$$\text{Rg}(A - \lambda I_3) < 3 \iff \lambda = 1, 4.$$

En résolvant le système associé (pour rappel on garde le système précédent, qui est déjà échelonné), on obtient les espaces propres :

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_4(A) \iff \begin{cases} x + y - 2z = 0 \\ -3y + 3z = 0 \end{cases} \iff \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \text{Vect} \left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right).$$

Donc $E_4(A) = \text{Vect} \left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right)$. De même on trouve $E_1(A) = \text{Vect} \left(\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right)$.

— **Orthonormalisation de la base propre obtenue.** Notons $e_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, e_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, e_3 =$

$\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$ les trois vecteurs obtenus précédemment. On constate que la base (e_1, e_2, e_3) n'est pas orthonormée. En revanche, normalisons déjà les deux premiers vecteurs :

$$f_1 = \frac{e_1}{\|e_1\|} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad f_2 = \frac{e_2}{\|e_2\|} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix},$$

et on cherche $f_3 \in E_1(A)$ de sorte que $f_3 \perp f_1$ et $f_3 \perp f_2$. Le vecteur f_3 est de la forme $\begin{pmatrix} x \\ y \\ -x-y \end{pmatrix}$ (ce doit être un élément de $E_1(A)$), et on le souhaite orthogonal à f_1, f_2 , i.e. :

$$\frac{1}{\sqrt{3}}(x+y+(-x-y)) = 0, \quad \frac{1}{\sqrt{2}}(x-(x-y)) = 0.$$

On obtient une solution $f_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ en choisissant par exemple $x = 1, y = -2, z = -x - y = 1$ et en normalisant le vecteur.

— **Conclusion** : la famille (f_1, f_2, f_3) étant une base orthonormée de vecteurs propres, nous obtenons


$$A = PD^T P, \quad D = \text{Diag}(4, 1, 1), \quad P = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{pmatrix}.$$

Remarque 2.3 —

- ▶ Pour ces choix de vecteurs (f_1, f_2, f_3) , vous serez toujours guidés dans les exercices.
- ▶ Rappelons que nous avons montré que les vecteurs propres associés à des valeurs propres distinctes sont orthogonaux : donc c'est surtout « à l'intérieur de » chaque espace propre de dimension > 1 qu'il faut « orthonormaliser » les vecteurs propres obtenus.

Exemple 11 — **Sur la matrice** ${}^T A.A$ — **Extrait Agro—Véto 2019** Soit $A \in \mathfrak{M}_{n,p}(\mathbf{R})$ ⁷ avec n, p deux entiers. Alors :

⁷On rappelle qu'en raison du format $n \times p$ imposé, cette matrice est canoniquement associée à un endomorphisme de $\mathfrak{M}_{p,1}(\mathbf{R})$ (nombre de colonnes : p), vers $\mathfrak{M}_{n,1}(\mathbf{R})$ (nombre de lignes : n)


1. $\ker A = \ker({}^T AA)$.  On procède par double inclusion. Commençons par constater que ${}^T A.A \in \mathfrak{M}_{n,n}(\mathbf{R})$.

\subset Soit $X \in \mathfrak{M}_{p,1}(\mathbf{R})$ tel que $AX = 0$. Alors ${}^T AAX = {}^T A0 = 0$ donc $X \in \ker({}^T AA)$.

\supset Inversement, soit $X \in \ker({}^T AA)$, i.e. ${}^T A(AX) = 0$. Multiplions cette égalité à gauche par ${}^T X$. On obtient

$${}^T X {}^T A(AX) = {}^T (AX)(AX) = 0 = \|AX\|^2$$

donc, par propriété sur la norme, $AX = 0$, c'est-à-dire $X \in \ker A$. On a bien montré par double inclusion l'égalité entre les deux ensembles.

2. ${}^T AA$ est diagonalisable et toutes ses valeurs propres sont positives ou nulles. 


On constate que ${}^T({}^T AA) = {}^T A {}^T({}^T A) = {}^T AA$. Ainsi, la matrice ${}^T AA$ est symétrique, donc d'après le théorème spectral elle est diagonalisable à valeurs propres réelles. Il reste à montrer que les valeurs propres sont positives ou nulles, soit donc $\lambda \in \mathbf{R}$, et X un vecteur propre associé. Alors

$${}^T AAX = \lambda X,$$

alors en multipliant par ${}^T X$ à gauche on trouve

$$\|AX\|^2 = \lambda \|X\|^2.$$

Comme $X \neq 0$, c'est un vecteur propre, on a $\lambda = \frac{\|AX\|^2}{\|X\|^2}$, on obtient alors immédiatement $\lambda \geq 0$.

3. (**Inversibilité de** ${}^T AA$) ${}^T AA$ est inversible $\iff \text{Rg}(A) = p$. **Indication** : Appliquer le théorème du rang à ${}^T AA$ d'une part, et A d'autre part.  Il s'agit d'abord de bien analyser le format de la matrice : ${}^T AA$ est de format $p \times p$, appliquons-lui le théorème du rang, ainsi qu'à la matrice A :

$$p = \dim \ker({}^T AA) + \text{Rg}({}^T AA), \quad p = \dim(\ker(A)) + \text{Rg}(A).$$

Donc, puisque $\ker A = \ker({}^TAA)$,

$$\begin{aligned} \operatorname{Rg}(A) = p &\iff p - \dim(\ker(A)) = p - \dim \ker({}^TAA) = p \\ &\iff \dim \ker({}^TAA) = 0, \\ &\iff \operatorname{Rg}({}^TAA) = p \\ &\iff {}^TAA \text{ inversible,} \end{aligned}$$

la dernière équivalence provient du fait que TAA est carrée.

3. PROJECTION ORTHOGONALE

La notion de projection orthogonale va être un objet clef pour les calculs de distance entre un vecteur et un sous-espace vectoriel. Rappelons que la distance à une partie est définie pour des parties quelconques, alors que l'on va parler de *projection orthogonale* uniquement sur des parties qui sont des sous-espaces vectoriels de E.

3.1. Projection orthogonale

3.1.1. Généralités

Définition/Proposition ALG.6.4 | Existence de la projection orthogonale

Soit F un sous-espace vectoriel de E non réduit à zéro. Alors il existe un unique endomorphisme $p_F \in \mathcal{L}(E)$ tel que :

$$\forall x \in E, \quad p_F(x) \in F, \quad x - p_F(x) \in F^\perp.$$

On l'appelle *la projection orthogonale sur F*. Si $x \in E$, le vecteur $p_F(x)$ est appelé le *projeté orthogonal de x sur F*.

Proposition ALG.6.10 | Propriétés de la projection orthogonale

Soit F un sous-espace vectoriel de E et p_F la projection orthogonale sur F, alors :

1. $p_F \circ p_F = p_F$,
2. $F = \operatorname{Im} p_F = \ker(p_F - \operatorname{Id}_E)$.

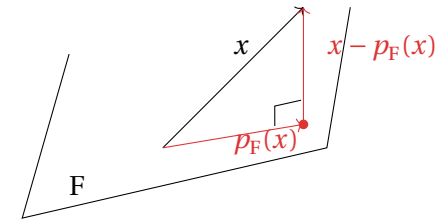


FIG. ALG.6.3. : Projection orthogonale

Preuve Nous démontrons les deux propriétés en même temps. La première contient plusieurs éléments : la linéarité, l'existence et l'unicité. Commençons par l'unicité.

■ **Unicité.** Supposons qu'il existe deux vecteurs $p_1(x)$ et $p_2(x)$ de F satisfaisant les conditions de l'énoncé. Nous souhaitons montrer que $p_1(x) = p_2(x)$, il suffit pour cela d'établir que $\|p_1(x) - p_2(x)\| = 0$. Puisque $p_1(x) \in F$ et $p_2(x) \in F$, on a donc aussi $p_1(x) - p_2(x) \in F$ car F est un sous-espace vectoriel de E. De plus, $x - p_1(x) \in F^\perp$ par hypothèse (de même que $x - p_2(x)$), donc d'après le théorème de PYTHAGORE,

$$\|x - p_1(x)\|^2 = \|x - p_2(x)\|^2 + \|p_2(x) - p_1(x)\|^2.$$


Avec le même argument (puisque $x - p_2(x) \in F^\perp$), on établit que

$$\|x - p_2(x)\|^2 = \|x - p_1(x)\|^2 + \|p_1(x) - p_2(x)\|^2.$$

En sommant et en simplifiant, $2\|p_1(x) - p_2(x)\|^2 = 0$, ce qui implique $p_1(x) = p_2(x)$. L'unicité est établie.

■ **Existence.** Soit (e_1, \dots, e_q) , $q = \dim F$, une base de F. On peut la compléter d'après le théorème de la base incomplète en une base $(e_1, \dots, e_q, e_{q+1}, \dots, e_n)$ de E. Par ailleurs, on peut la supposer orthonormée quitte à l'orthonormaliser (fait admis, consulter la **Remarque 2** pour plus de détails). Définissons l'application p ainsi :

$$p_F : \begin{cases} E & \longrightarrow F, \\ u = \sum_{i=1}^n \lambda_i e_i & \longrightarrow \sum_{i=1}^q \lambda_i e_i. \end{cases}$$


1. L'application est bien définie puisque les coordonnées λ_i sont uniques pour $u \in E$ fixé (la famille (e_1, \dots, e_n) est une base donc en particulier une famille libre de E).
2. L'application p_F est linéaire.  Soient $x = \sum_{i=1}^n \lambda_i e_i$ et $y = \sum_{i=1}^n \mu_i e_i$ avec $\lambda_i, \mu_i \in \mathbf{R}$ pour tout i . Alors :

$$\lambda x + \mu y = \sum_{i=1}^n (\lambda \lambda_i + \mu \mu_i) e_i.$$

Alors :

$$p_F(\lambda x + \mu y) = \sum_{i=1}^q (\lambda \lambda_i + \mu \mu_i) e_i = \lambda \sum_{i=1}^q \lambda_i e_i + \mu \sum_{i=1}^q \mu_i e_i = \lambda p_F(x) + \mu p_F(y).$$

Ainsi l'application p_F est bien linéaire.

3. Pour tout $x \in E$, $p_F(x) \in F$.  Immédiat, car $p_F(x)$ est une combinaison linéaire d'éléments de F qui est un espace vectoriel donc stable par combinaison linéaire.
4. Pour tout $x = \sum_{i=1}^n \lambda_i e_i \in E$, $\lambda_i \in \mathbf{R}$ pour tout i , montrons que $x - p_F(x) \in F^\perp$.

Soit $y \in F$, alors il existe $\mu_1, \dots, \mu_q \in \mathbf{R}$ tels que $y = \sum_{i=1}^q \mu_i e_i$. Montrons que


$$\langle x - p_F(x) | y \rangle = 0. \quad \text{En effet,}$$

$$\begin{aligned} \langle x - p_F(x) | y \rangle &= \left\langle \sum_{i=1}^n \lambda_i e_i - \sum_{i=1}^q \lambda_i e_i \mid \sum_{j=1}^q \mu_j e_j \right\rangle \\ &= \left\langle \sum_{i=q+1}^n \lambda_i e_i \mid \sum_{j=1}^q \mu_j e_j \right\rangle \\ &= \sum_{i=q+1}^n \sum_{j=1}^q \lambda_i \mu_j \underbrace{\langle e_i | e_j \rangle}_{=0} = 0. \end{aligned} \quad \left. \begin{array}{l} \\ \\ \end{array} \right) \text{ bilinéarité du produit scalaire}$$

L'existence et l'unicité de $p_F(x)$ sont donc établies. Montrons maintenant les deux propriétés. On se fixe $x = \sum_{i=1}^n \lambda_i e_i \in E$, $\lambda_i \in \mathbf{R}$ pour tout i .

1. 

$$p_F \circ p_F(x) = p_F \left(\sum_{i=1}^q \lambda_i e_i \right) = \sum_{i=1}^q \lambda_i e_i = p_F(x).$$

2.  On a évidemment $\text{Im}(p_F) \subset F$ d'après ce qui précède, il reste à établir que $F \subset \text{Im}(p_F)$. Supposons que $x \in F$, alors $x = \sum_{i=1}^q \lambda_i e_i = p_F \left(\sum_{i=1}^q \lambda_i e_i \right) = p_F(x) \in \text{Im}(p_F)$. Et enfin, il reste à établir que $F = \ker(p_F - \text{Id}_E)$ est l'ensemble des points fixes de p_F . Nous avons encore évidemment $\ker(p_F - \text{Id}_E) \subset F$: si $x \in \ker(p_F - \text{Id}_E)$ alors $x = p_F(x) \in \text{Im}(p_F)$. Inversement, montrons que $F \subset \ker(p_F - \text{Id}_E)$: si $x = \sum_{i=1}^q \lambda_i e_i \in F$, alors $p_F(x) = p_F \left(\sum_{i=1}^q \lambda_i e_i \right) = \sum_{i=1}^q \lambda_i e_i = x$ donc $x \in \ker(p_F - \text{Id}_E)$.


Proposition ALG.6.11 | Propriétés de la projection orthogonale

Soit F un sous-espace vectoriel de E non réduit à zéro. Alors $\text{Id}_E - p_F$ est la projection orthogonale sur F^\perp .

- Preuve** Rappelons une inclusion déjà établie : $F \subset F^{\perp\perp} = (F^\perp)^\perp$. Alors :
- ▶ $\text{Id}_E - p_F$ est une application linéaire en tant que différence de telles applications.
 - ▶ Pour tout $x \in E$, $(\text{Id}_E - p_F)(x) \in F^\perp$ par définition de la projection orthogonale sur F .
 - ▶ Et pour tout $x \in E$, $(\text{Id}_E - (\text{Id}_E - p_F))(x) = p_F(x) \in F \subset F^{\perp\perp}$.
- De tout cela, nous déduisons que $\text{Id}_E - p_F$ est la projection orthogonale sur F^\perp .

Exemple 12 — Soit F un sous-espace vectoriel de E non réduit à zéro, notons p_F la projection orthogonale sur F .

1. Montrer que $\text{Spec}(p_F) \subset \{0, 1\}$. 


2. Notons $M = \mathcal{M}_{\mathcal{B}}(p_F)$ où \mathcal{B} est une base de E . Que vaut M^2 ? 

Corollaire ALG.6.3 | Diagonalisation d'une projection orthogonale

Soit F un sous-espace vectoriel de E non réduit à zéro, notons p_F la projection orthogonale sur F . Notons (f_1, \dots, f_r) une base de $\text{Im}(p_F) = \ker(\text{Id}_E - p_F)$ avec $r = \text{Rg}f$ et (e_1, \dots, e_{n-r}) une base de $\ker(p_F)$. Alors :

1. $\mathcal{B} = (e_1, \dots, e_{n-r}, f_1, \dots, f_r)$ est une base de vecteurs propres de E ,
2. p_F est diagonalisable, et $\text{Spec } p_F = \{0, 1\}$.

Preuve

1. Puisque $\# \mathcal{B} = n$, il suffit de montrer la liberté de la famille. On aura ainsi montré que $\mathcal{B} = (e_1, \dots, e_{n-r}, f_1, \dots, f_r)$ est une base de vecteurs propres. 

2. On a montré précédemment que $\mathcal{B} = (e_1, \dots, e_{n-r}, f_1, \dots, f_r)$ est une base de vecteurs propres, donc p_F est diagonalisable et ses valeurs propres sont 0 et 1 (cf. exemple précédent).

Proposition ALG.6.12 | Expression du projeté orthogonal si l'on connaît une base orthonormale

Soit F un sous-espace vectoriel de E non réduit à zéro. Alors pour toute base orthonormale (e_1, \dots, e_q) de F , $q = \dim F$, on a :

$$\forall x \in E, \quad p_F(x) = \sum_{i=1}^q \langle x | e_i \rangle e_i.$$


Preuve D'après le théorème de la base incomplète, le famille (e_1, \dots, e_q) peut être complétée en une base $(e_1, \dots, e_q, e_{q+1}, \dots, e_n)$ de E . On peut de plus la supposer orthonormée quitte à l'orthonormaliser (fait admis en BCPST). Si $x = \sum_{i=1}^n \lambda_i e_i \in E$, avec $\lambda_i \in \mathbf{R}$, nous avons déjà vu que, puisque (e_1, \dots, e_n) est une base orthonormée, $\lambda_i = \langle x | e_i \rangle$ pour tout $i \in \llbracket 1, n \rrbracket$. Donc $x = \sum_{i=1}^n \langle x | e_i \rangle e_i \in E$ et par définition d'une projection orthogonale $p_F(x) = \sum_{i=1}^q \langle x | e_i \rangle e_i$.

Remarque 3.1 — Dès que l'on connaît un vecteur directeur d'une droite la projection orthogonale associée se calcule très facilement.

Corollaire ALG.6.4 | Cas de la projection orthogonale sur une droite vectorielle

Soit $v \in E$ non nul et $x \in E$, alors :

$$p_{\text{Vect}(v)}(x) = \frac{1}{\|v\|^2} \langle x | v \rangle v.$$

Preuve  la famille (v) est une base de $\text{Vect}(v)$ et $\left(\frac{v}{\|v\|}\right)$ en est une base orthonormée de $\text{Vect}(v)$, donc la proposition précédente s'applique :

$$p_{\text{Vect}(v)}(x) = \left\langle u \left| \frac{v}{\|v\|} \right. \right\rangle \frac{v}{\|v\|} = \frac{1}{\|v\|^2} \langle x | v \rangle v.$$

Remarque 3.2 — Orthonormalisation d'une famille de deux vecteurs. La Proposition ALG.6.12 s'applique dès qu'une famille orthonormée de l'espace sur lequel on projette est connue. Il est possible d'«orthonormaliser» toute base, i.e. de transformer toute base en une version orthonormale qui engendre le même espace vectoriel

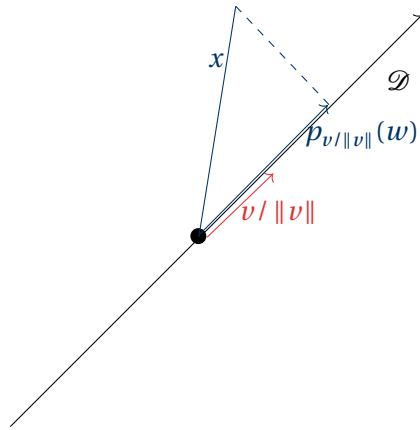
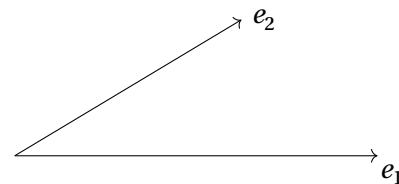


FIG. ALG.6.4. : Projection orthogonale sur une droite

in fine. Le procédé général n'est pas au programme, mais on peut retenir le cas particulier de deux vecteurs, très utile en pratique et qui se mémorise très facilement à l'aide d'un dessin.

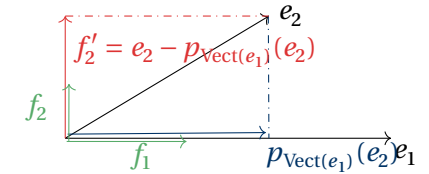


1) On part d'une base *a priori* non orthonormée (e_1, e_2) , on normalise le premier (*i.e.* on le divise par sa norme) en posant

$$f_1 = \frac{e_1}{\|e_1\|}.$$

2) On calcule le projeté orthogonal de e_2 sur la droite $\text{Vect}(e_1)$, puis $e_2 - p_{\text{Vect}(e_1)}(e_2)$ qui permet de « redresser » le second vecteur, et enfin on le normalise en posant

$$f_2 = \frac{e_2 - p_{\text{Vect}(e_1)}(e_2)}{\|e_2 - p_{\text{Vect}(e_1)}(e_2)\|}.$$



On peut retrouver ceci de manière plus analytique sans dessin. Notons $f_1 = \frac{e_1}{\|e_1\|}$. Soit $f_2' = e_2 + \lambda f_1$ avec $\lambda \in \mathbf{R}$. On cherche λ de sorte que $\langle f_1 | f_2' \rangle = 0$:


$$\begin{aligned} \langle f_1 | f_2' \rangle = 0 &\iff \langle f_1 | e_2 + \lambda f_1 \rangle = 0 \\ &\iff \langle f_1 | e_2 \rangle + \lambda \langle f_1 | f_1 \rangle = 0 = \langle f_1 | e_2 \rangle + \lambda \|f_1\|^2 = \langle f_1 | e_2 \rangle + \lambda. \end{aligned}$$

Le scalaire $\lambda = -\langle f_1 | e_2 \rangle$ convient. Ainsi, avec ce choix, la famille (f_1, f_2') est orthogonale. Il reste ensuite à former $f_2 = \frac{f_2'}{\|f_2'\|}$ pour obtenir une base orthonormée (f_1, f_2) : on justifie ensuite que cette famille est encore une base de \mathbf{R}^2 .

Méthode Orthonormalisation d'une famille de deux vecteurs

Lors (e_1, e_2) est une base quelconque (non forcément orthonormée) de F , une version orthonormée est obtenue en posant :

$$\begin{aligned} f_1 &= \frac{e_1}{\|e_1\|}, \\ f_2 &= \frac{e_2 - p_{\text{Vect}(e_1)}(e_2)}{\|e_2 - p_{\text{Vect}(e_1)}(e_2)\|} \\ &= \frac{e_2 - \frac{1}{\|e_1\|^2} \langle x | e_1 \rangle e_1}{\|e_2 - \frac{1}{\|e_1\|^2} \langle x | e_1 \rangle e_1\|}. \end{aligned}$$

 Le vecteur $e_2 - p_{\text{Vect}(e_1)}(e_2)$ est une version « redressée » du vecteur e_2 , orthogonale à e_1 (ou f_1).

3.1.2. Calculs pratiques de projections orthogonales

Avant de passer à des exemples de calculs, précisons les deux méthodes à disposition pour calculer une projection orthogonale.

Méthode Calcul d'une projection orthogonale

Deux méthodes pour calculer un projeté orthogonal sur F :

1. (En utilisant la définition) Si on ne connaît pas une base orthonormale de F, mais une famille génératrice (e_1, \dots, e_q) de F : soit $x \in E$, alors on cherche l'unique vecteur $p_F(x)$ vérifiant

$$p_F(x) \in F, \quad x - p_F(x) \in F^\perp.$$

On caractérise $x' = p_F(x)$ de la manière suivante :

$$x' = p_F(x) \iff x' \in F, \quad \text{et} \quad \langle x - x' | e_i \rangle = 0, \quad \forall i \in \llbracket 1, q \rrbracket.$$


2. (En utilisant la formule dans une base orthonormale) Si une base orthonormale (e_1, \dots, e_q) de F est connue, alors :

$$\forall u \in E, \quad p_F(u) = \sum_{i=1}^q \langle u | e_i \rangle e_i.$$

Lorsque F est de dimension 1 ou 2, on peut se ramener facilement à une base orthonormée et donc utiliser **2**, en dimension 3 ou plus : si aucune base orthonormée n'est donnée, on utilisera **1**.

Exemple 13 — Une base orthonormale est donnée On pose $F = \{(x, y, x), (x, y) \in \mathbf{R}^2\}$ et on note p la projection orthogonale sur F. Une base est donnée par $((1, 0, 1), (0, 1, 0))$, c'est une base orthogonale, et $(\frac{1}{\sqrt{2}}(1, 0, 1), (0, 1, 0))$ en est une

version orthonormée. On considère ci-après deux autres bases orthonormées de F.

1. Vérifier que $\mathcal{F} = \left(\left(\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2} \right), (0, 1, 0) \right)$ et $\mathcal{G} = \left(\left(\frac{\sqrt{6}}{4}, \frac{1}{2}, \frac{\sqrt{6}}{4} \right), \left(-\frac{\sqrt{2}}{4}, \frac{\sqrt{3}}{2}, -\frac{\sqrt{2}}{4} \right) \right)$ sont deux bases orthonormées de F.  *Calcul évident : on vérifie d'abord que les deux vecteurs constituant chaque base sont dans F, puis on calcule leur norme et leur produit scalaire.*

2. Qu'en déduit-on sur p ?  Soit $x = (x_1, x_2, x_3) \in \mathbf{R}^3$, alors :


$$\begin{aligned} p_F(x) &= \left\langle (x_1, x_2, x_3) \left| \left(\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2} \right) \right. \right\rangle \left(\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2} \right) + \left\langle (x_1, x_2, x_3) \left| (0, 1, 0) \right. \right\rangle (0, 1, 0) \\ &= \frac{x_1 + x_3}{\sqrt{2}} \left(\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2} \right) + x_2(0, 1, 0) \\ &= \left(\frac{x_1 + x_3}{2}, x_2, \frac{x_1 + x_3}{2} \right) \end{aligned}$$

De-même :


$$\begin{aligned} p_F(x) &= \left\langle (x_1, x_2, x_3) \left| \left(\frac{\sqrt{6}}{4}, \frac{1}{2}, \frac{\sqrt{6}}{4} \right) \right. \right\rangle \left(\frac{\sqrt{6}}{4}, \frac{1}{2}, \frac{\sqrt{6}}{4} \right) \\ &+ \left\langle (x_1, x_2, x_3) \left| \left(-\frac{\sqrt{2}}{4}, \frac{\sqrt{3}}{2}, -\frac{\sqrt{2}}{4} \right) \right. \right\rangle \left(-\frac{\sqrt{2}}{4}, \frac{\sqrt{3}}{2}, -\frac{\sqrt{2}}{4} \right) \\ &= \left(\frac{1}{16} \left(6(x_1 + x_3) + 2\sqrt{6}x_2 - 2\sqrt{2}\sqrt{3}x_2 + 2(x_1 + x_3) \right), [\dots], [\dots] \right) \\ &= \left(\frac{x_1 + x_3}{2}, x_2, \frac{x_1 + x_3}{2} \right). \end{aligned}$$

Les coordonnées deux et trois se calculent comme la première. Donc peu importe la base orthonormale choisie, on obtient la même expression analytique pour p_F : à mettre en lien avec la phrase « p_F est bien définie » dans la propriété d'existence de la projection orthogonale.

Exemple 14 — Espace donné « sous forme d'un Vect » Écrire la matrice dans la base canonique de \mathbf{R}^3 de la projection orthogonale sur $F = \text{Vect}(1, 1, 1)$.

1. (1ère Méthode) En utilisant la Proposition ALG.6.12, i.e. en cherchant une base orthonormale.  Une base de F est $(e_1) = ((1, 1, 1))$. Une base orthonormée de F est alors $(f_1) = \frac{1}{\sqrt{3}}(1, 1, 1)$. En utilisant la Proposition ALG.6.12 vient alors pour tous $(x, y, z) \in \mathbf{R}^3$,

$$p_F(x, y, z) = \langle (x, y, z) | f_1 \rangle f_1 = \frac{1}{(\sqrt{3})^2} (x + y + z)(1, 1, 1) = \frac{1}{3}(x + y + z)(1, 1, 1).$$

2. (2ème Méthode) En se ramenant à la définition.  Soit $(x, y, z) \in \mathbf{R}^3$, alors $(x', y', z') = p_F(x, y, z)$ si et seulement si

$$\begin{cases} (x', y', z') \in F \\ (x', y', z') - (x, y, z) \in F^\perp \end{cases} \iff \begin{cases} x' = y' = z' \\ (x' - x, y' - y, z' - z) \perp (1, 1, 1) \end{cases}$$

$$\iff \begin{cases} x' = y' = z' \\ (x' - x) + (y' - y) + (z' - z) = 0. \end{cases}$$

On souhaite résoudre le système précédent en (x', y, z') , on trouve alors

$$x' = \frac{1}{3}(x + y + z), \quad y' = z' = \frac{1}{3}(x + y + z).$$

C'est le même résultat que précédemment.

Dans les deux cas on obtient :

$$p_F \begin{cases} \mathbf{R}^3 & \longrightarrow & \mathbf{R}^3, \\ (x, y, z) & \longrightarrow & \frac{1}{3}(x + y + z, x + y + z, x + y + z). \end{cases}$$

On en déduit alors la matrice :


$$\mathcal{M}_{\mathcal{B}^{can}}(p_F) = \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}^B.$$

Exemple 15 — Espace donné sous forme d'équation implicite On considère le sous-espace vectoriel

$$F = \{(x, y, z) \in \mathbf{R}^3, x + y - z = 0\}$$

de \mathbf{R}^3 . On note p la projection orthogonale sur F. Déterminer la matrice A de p dans la base canonique de \mathbf{R}^3 .

On commence par déterminer une expression analytique de $p = p_F(x, y, z)$ pour tout $(x, y, z) \in \mathbf{R}^3$.

1. (1ère Méthode)  On a $F = \text{Vect}((1, 0, 1), (0, 1, 1))$. Cette base de F n'est pas orthonormée, rendons-là donc orthonormée. Posons $f_1 = \frac{1}{\sqrt{2}}(1, 0, 1)$. Ensuite, $p_{\text{Vect}(1,0,1)}(0, 1, 1) = \frac{1}{2} \langle (0, 1, 1) | (1, 0, 1) \rangle (1, 0, 1) = \frac{1}{2}(1, 0, 1)$. Ainsi,

$$(0, 1, 1) - p_{\text{Vect}(1,0,1)}(0, 1, 1) = (0, 1, 1) - \frac{1}{2}(1, 0, 1) = (-1/2, 1, 1/2),$$

et on pose


$$f_2 = \frac{(-1/2, 1, 1/2)}{\|(-1/2, 1, 1/2)\|} = \frac{1}{\sqrt{3/2}}(-1/2, 1, 1/2).$$

On vérifie aisément que la famille (f_1, f_2) ainsi construite est une base orthonormée de \mathbf{R}^3 . Nous obtenons alors l'expression de la projection :

$$\begin{aligned} p(x, y, z) &= \langle (x, y, z) | f_1 \rangle f_1 + \langle (x, y, z) | f_2 \rangle f_2 \\ &= \frac{1}{2}(x + z)(1, 0, 1) + \frac{2}{3}(-x/2 + y + z/2)(-1/2, 1, 1/2). \end{aligned}$$

On en déduit alors la matrice en remplaçant le vecteur (x, y, z) par ceux de la base canonique :

$$\mathcal{M}_{\mathcal{B}^{can}}(p) = \begin{pmatrix} 2/3 & -1/3 & 1/3 \\ -1/3 & 2/3 & 1/3 \\ 1/3 & 1/3 & 2/3 \end{pmatrix}.$$

2. (2ème Méthode) En se ramenant à la définition. 

$$\begin{aligned}
 (x', y', z') \\
 = p_F(x, y, z) &\iff \begin{cases} (x', y', z') \in \mathcal{P} \\ (x, y, z) - (x', y', z') = (x - x', y - y', z - z') \in \mathcal{P}^\perp \end{cases} \\
 &\iff \begin{cases} x' + y' - z' = 0 \\ (x - x', y - y', z - z') \perp (1, 0, 1) \\ (x - x', y - y', z - z') \perp (0, 1, 1) \end{cases} \iff \begin{cases} x' + z' = x + z \\ y' + z' = y + z \\ x' + y' - z' = 0. \end{cases}
 \end{aligned}$$

C'est un système à trois équations, trois inconnues, à résoudre en (x', y', z') : les solutions, qui dépendent de manière linéaire de x, y et z correspondent aux coordonnées de $p_F(x, y, z)$. On retrouve la solution précédente après résolution.

3.1.3. Projection orthogonale et distance à un sous-espace vectoriel

Revenons au but initial de cette section : trouver la valeur de la distance entre un vecteur et un sous-espace vectoriel de E , elle est donnée dans la proposition suivante.

Proposition ALG.6.13 | Distance d'un vecteur à un sous-espace vectoriel

Soit F un **sous-espace vectoriel** de E non réduit à zéro et $u \in E$. Alors :

$$d(u, F) = \inf_{v \in F} \|u - v\| = \|u - p_F(u)\|.$$

En particulier, la borne inférieure est atteinte, et est un minimum.

 **Attention**

L'hypothèse de structure d'espace vectoriel sur F est fondamentale : nous avons déjà rencontré des exemples où la borne inférieure n'était pas atteinte lorsque F n'est pas un espace vectoriel (par exemple un disque privé de son bord dans l'Exemple 2).

Preuve Rappelons que $d(u, F) = \inf_{v \in F} \|u - v\|$.

1. Le réel $\|u - p_F(u)\|$ minore $\{\|u - v\|, v \in F\}$. En effet, si $v \in F$, alors comme $u - p_F(u) \in F^\perp$ et $p_F(u) - p_F(v) \in F$, le théorème de PYTHAGORE nous livre :


$$\begin{aligned}
 \|u - v\|^2 &= \|(u - p_F(u)) + (p_F(u) - v)\|^2 = \|u - p_F(u)\|^2 + \|p_F(u) - p_F(v)\|^2 \\
 &\geq \|u - p_F(u)\|^2.
 \end{aligned}$$

Donc $\|u - p_F(u)\|$ est un minorant de $\{\|u - v\|, v \in F\}$.

2. De plus, ce minorant est atteint par $p_F(u) \in F$.

D'où le résultat.

Exemple 16 — Calculer $d((1, 1, 1), \mathcal{P})$ dans l'Exemple 15.

 D'après le cours :

$$\begin{aligned}
 d((1, 1, 1), \mathcal{P}) &= \|(1, 1, 1) - p_{\mathcal{P}}(1, 1, 1)\| \\
 &= \left\| (1, 1, 1) - (1, 0, 1) - \frac{2}{3}(-1/2, 1, 1/2) \right\| = \|(1/3, 1/3, -1/3)\| \\
 &= \sqrt{\frac{1}{3}}.
 \end{aligned}$$

*** **Fin du chapitre** ***

4. EXERCICES

Exercice ALG.6.1 | Vrai ou Faux? Dans \mathbf{R}^n muni de son produit scalaire canonique, x, y, z, t désigne des vecteurs de \mathbf{R}^n , et $(\alpha, \beta) \in \mathbf{R}^2$. Dire quelles affirmations sont vérifiées :

1. $\alpha \langle x|y \rangle + \beta \langle z|y \rangle = \langle \alpha x + \beta z|y \rangle$,
2. $\langle x|y \rangle + \langle z|t \rangle = \langle x + z|y + t \rangle$,
3. $\langle x|y \rangle = 0 \implies x = 0$ ou $y = 0$,
4. $\langle x|x \rangle = 0 \implies x = 0$,
5. $(\forall z \in \mathbf{R}^n, \langle x, z \rangle = 0) \implies x = 0$
6. $\forall (x, y) \in (\mathbf{R}^n)^2, \|x + y\| = \|x\| + \|y\|$,
7. toute famille de vecteurs non nuls orthogonaux est libre.
8. Si A est symétrique réelle et si \mathcal{C} est une base de vecteurs propres de A , alors \mathcal{C} est une base orthonormée de \mathbf{R}^n .

Solution (exercice ALG.6.1)

1. C'est vrai : propriété de linéarité à gauche du produit scalaire.
2. C'est faux, prendre par exemples $n = 1, x = y = z = t = 1$, on $1 + 1 \neq 2 \times 2$.
3. C'est faux pour $n > 1$: prendre par exemple $x = (1, 0), y = (0, 1)$. Alors $\langle x|y \rangle = 0$ et pourtant $x \neq 0$ et $y \neq 0$. En revanche pour $n = 1$ c'est vrai car un produit de deux nombres réels est nul si et seulement si l'un des deux réels est nul.
4. C'est vrai : propriété (« défini ») du produit scalaire, montrée au tout début du chapitre.
5. C'est vrai, faire simple $z = x$ dans l'hypothèse. On obtient alors $\|x\|^2 = 0 \implies x = 0$.
6. C'est faux : prendre par exemple $x = (1, 0), y = (0, 1)$.
7. C'est vrai : la propriété a été vue dans le cours.
8. C'est faux : cette affirmation est une « permutation » des éléments du théorème spectral. Nous avons vu dans le cours que la matrice $\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ fournit un contre

exemple : nous avons obtenu une base propre en diagonalisant, mais ce n'était pas une base orthonormée.

4.1. Inégalités classiques

Exercice ALG.6.2 | Soit a_1, a_2, \dots, a_n des nombres réels. Montrer que :

$$\left(\frac{a_1 + \dots + a_n}{n} \right)^2 \leq \frac{a_1^2 + \dots + a_n^2}{n}.$$

Indication : On pourra se servir de l'inégalité de CAUCHY-SCHWARZ.

Solution (exercice ALG.6.2)

Nous avons, d'après l'inégalité de CAUCHY-SCHWARZ élevée au carré :

$$\left(\frac{a_1 + \dots + a_n}{n} \right)^2 = \left\langle (a_1, \dots, a_n) \middle| \left(\frac{1}{n}, \dots, \frac{1}{n} \right) \right\rangle^2 \leq \left(\sum_{i=1}^n a_i^2 \right) \left(\sum_{i=1}^n \frac{1}{n^2} \right) = \frac{1}{n} \sum_{i=1}^n a_i^2.$$

C'est l'inégalité cherchée.

Exercice ALG.6.3 |

1. Montrer que pour tout $n \geq 2$: $\sum_{k=1}^n k\sqrt{k} \leq \frac{n(n+1)\sqrt{2n+1}}{2\sqrt{3}}$.
2. Montrer que pour tout $n \geq 2$: $\sum_{k=1}^{n-1} \frac{k}{(n-k)^2} \geq \frac{2}{n(n-1)} \left(\sum_{k=1}^{n-1} \frac{k}{n-k} \right)^2$.

Solution (exercice ALG.6.3)

1. C'est une conséquence de l'inégalité de CAUCHY-SCHWARZ :

$$\begin{aligned} \sum_{k=1}^n k\sqrt{k} &= \left\langle (1, 2, \dots, n) \middle| (\sqrt{1}, \sqrt{2}, \dots, \sqrt{n}) \right\rangle \leq \sqrt{\sum_{i=1}^n i^2} \sqrt{\sum_{i=1}^n i} \\ &= \sqrt{\frac{n(n+1)(2n+1)}{6}} \sqrt{\frac{n(n+1)}{2}}. \end{aligned}$$

On obtient le majorant cherché.

2. Constatons que $\frac{n(n-1)}{2} = \sum_{k=1}^{n-1} k$, pour rendre les choses plus claires. Il s'agit donc de montrer que :

$$\left(\sum_{k=1}^{n-1} \frac{k}{n-k}\right)^2 \leq \sum_{k=1}^{n-1} \frac{k}{(n-k)^2} \times \sum_{k=1}^{n-1} k.$$

Il suffit ensuite simplement d'appliquer l'inégalité de CAUCHY-SCHWARZ :

$$\begin{aligned} \left(\sum_{k=1}^{n-1} \frac{k}{n-k}\right)^2 &= \left(\sum_{k=1}^{n-1} \sqrt{k} \frac{\sqrt{k}}{n-k}\right)^2 \\ &= \left\langle (\sqrt{1}, \dots, \sqrt{n-1}) \middle| \left(\frac{\sqrt{1}}{n-1}, \dots, \frac{\sqrt{n-1}}{n-(n-1)}\right) \right\rangle^2 \\ &\leq \left(\sum_{k=1}^{n-1} k\right) \left(\sum_{k=1}^{n-1} \frac{k}{(n-k)^2}\right). \end{aligned}$$

C'est l'inégalité cherchée.

Exercice ALG.6.4 | CAUCHY-SCHWARZ et minimisation d'une fonction de plusieurs variables

- Soient x_1, \dots, x_n des nombres réels strictement positifs. Montrer que : $\left(\sum_{k=1}^n \frac{1}{x_k}\right) \left(\sum_{k=1}^n x_k\right) \geq n^2$. *Indication* : On pourra remarquer que $\sqrt{x_k^2} = x_k$ pour tout $k \in \llbracket 1, n \rrbracket$.
- Trouver le(s) minimum(s) global (ou globaux) de $f : (x_1, \dots, x_n) \in (\mathbf{R}^{+*})^n \mapsto \left(\sum_{k=1}^n \frac{1}{x_k}\right) \left(\sum_{k=1}^n x_k\right)$.

Solution (exercice ALG.6.4)

1. C'est une conséquence de l'inégalité de CAUCHY-SCHWARZ. En effet :

$$\begin{aligned} \left| \left\langle (\sqrt{x_1}, \dots, \sqrt{x_n}) \middle| \left(\frac{1}{\sqrt{x_1}}, \dots, \frac{1}{\sqrt{x_n}}\right) \right\rangle \right| &\leq \sqrt{\sum_{k=1}^n \left(\frac{1}{\sqrt{x_k}}\right)^2} \sqrt{\sum_{k=1}^n (\sqrt{x_k})^2} \\ &= \sqrt{\sum_{k=1}^n \frac{1}{x_k}} \sqrt{\sum_{k=1}^n x_k}. \end{aligned}$$

Il ne reste plus qu'à élever l'inégalité au carré, on obtient la majoration souhaitée. D'où :

$$\boxed{\left(\sum_{k=1}^n \frac{1}{x_k}\right) \left(\sum_{k=1}^n x_k\right) \geq n^2.}$$

2. On constate qu'en prenant $(x_1, \dots, x_n) = (1, \dots, 1)$, nous avons égalité dans la première question, il vient que n^2 est la valeur minimale de la fonction f puisque $(1, \dots, 1) \in (\mathbf{R}^{+*})^n$.

La valeur n^2 est atteinte en $(1, \dots, 1)$, plus généralement en tous les $(x_1, \dots, x_n) \in (\mathbf{R}^{+*})^n$ qui réalise le cas d'égalité dans la question précédente, *i.e.* tels que les vecteurs $(\sqrt{x_1}, \dots, \sqrt{x_n})$ et $(\frac{1}{\sqrt{x_1}}, \dots, \frac{1}{\sqrt{x_n}})$ soient liés. Ceci est équivalent à l'existence de $\lambda \in \mathbf{R}$ tel que :

$$\forall i \in \llbracket 1, n \rrbracket, \quad \frac{1}{\sqrt{x_i}} = \lambda \sqrt{x_i} \iff 1 = \lambda x_i,$$

car $x_i \neq 0$. Donc finalement $\lambda = \frac{1}{x_i}$ pour tout i et tous les x_i sont nécessairement égaux, donc $(x_1, \dots, x_n) = (\mu, \dots, \mu)$ avec $\mu > 0$. En remplaçant, on trouve la valeur de μ :

$$f(\mu, \dots, \mu) = n^2 \iff \frac{n}{\mu} n\mu = n^2 \iff \mu = 1.$$

En conclusion, $(1, \dots, 1)$ est l'unique minimum global de f .

4.2. Diagonalisation de matrices symétriques

Exercice ALG.6.5 | Soit $A = \begin{pmatrix} 1 & \sqrt{6} \\ \sqrt{6} & 2 \end{pmatrix}$.

Justifier que A est diagonalisable en base orthonormée, et trouver une base orthonormée de vecteurs propres.

Solution (exercice ALG.6.5)

La matrice A est diagonalisable en base orthonormée puisque elle est symétrique réelle. Autrement dit, il existe une matrice $D \in \mathfrak{M}_2(\mathbf{R})$ et $P \in GL_2\mathbf{R}$ telles que : $A = PD^T P$.

Regardons les éléments propres avec Python avant de faire les calculs.

```
>>> import numpy as np
>>> A = np.array([[1, np.sqrt(6)], [np.sqrt(6), 2]])
>>> np.linalg.eig(A)
(array([-1., 4.]), array([[ -0.77459667, -0.63245553],
[ 0.63245553, -0.77459667]]))
```

Soit $\lambda \in \mathbf{R}$. Pour trouver les valeurs propres de A , on résout en λ l'équation suivante :

$$\det \begin{pmatrix} 1-\lambda & \sqrt{6} \\ \sqrt{6} & 2-\lambda \end{pmatrix} = 0 \iff (1-\lambda)(2-\lambda) - 6 = \lambda^2 - 3\lambda - 4 = 0.$$

On vérifie sans peine que les solutions de l'équation sont 4 et -1 . Ainsi, puisque nous avons deux valeurs propres distinctes, la matrice A est diagonalisable (mais nous le savions déjà d'après le théorème spectral), et donc $\text{Spec}(A) = \{-1, 4\}$, et diagonalisable dans une base orthonormée car symétrique réelle. Calculons $E_{-1}(A)$ pour commencer :

$$(x, y) \in E_{-1}(A) \iff \begin{cases} 2x + \sqrt{6}y = 0 \\ \sqrt{6}x + 3y = 0 \end{cases}.$$

Or, $\frac{\sqrt{6}}{2}(2x + \sqrt{6}y) = \sqrt{6}x + 3y = 0$, donc les deux lignes du système sont liées. Nous obtenons alors : $E_{-1}(A) = \text{Vect} \begin{pmatrix} -1/2 \\ 1/\sqrt{6} \end{pmatrix}$. On procède de même avec $E_4(A)$:

$$E_{-1}(A) = \text{Vect} \begin{pmatrix} -1/2 \\ 1/\sqrt{6} \end{pmatrix}, \quad E_4(A) = \text{Vect} \begin{pmatrix} 1/\sqrt{6} \\ 1/2 \end{pmatrix}.$$

Les deux vecteurs sont orthogonaux mais pas normés. Il reste donc à les diviser par leur norme pour obtenir une matrice de passage orthogonale. Nous avons :

$$\|(-1/2, 1/\sqrt{6})\| = \sqrt{\frac{1}{4} + \frac{1}{6}} = \sqrt{\frac{5}{12}} = \frac{\sqrt{5}}{2\sqrt{3}},$$

on procède de même avec le second vecteur

$$\|(-1/\sqrt{6}, 1/2)\| = \sqrt{\frac{1}{6} + \frac{1}{4}} = \sqrt{\frac{5}{12}} = \frac{\sqrt{5}}{2\sqrt{3}}.$$

D'où :

$$E_{-1}(A) = \text{Vect} \begin{pmatrix} -\frac{1}{2} \frac{2\sqrt{3}}{\sqrt{5}} \\ \frac{1}{\sqrt{6}} \frac{2\sqrt{3}}{\sqrt{5}} \end{pmatrix} = \text{Vect} \begin{pmatrix} -\sqrt{3}/\sqrt{5} \\ \sqrt{2}/\sqrt{5} \end{pmatrix}, \quad E_4(A) = \text{Vect} \begin{pmatrix} \sqrt{2}/\sqrt{5} \\ \sqrt{3}/\sqrt{5} \end{pmatrix}.$$

Donc :

$$A = PD^T P, \quad P = \begin{pmatrix} -\sqrt{3}/\sqrt{5} & \sqrt{2}/\sqrt{5} \\ \sqrt{2}/\sqrt{5} & \sqrt{3}/\sqrt{5} \end{pmatrix}, \quad D = \begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix}.$$

On peut vérifier *a posteriori* les calculs de vecteurs propres.

```
>>> import numpy as np
>>> A = np.array([[1, np.sqrt(6)], [np.sqrt(6), 2]])
>>> X = np.array([[ -np.sqrt(3)/np.sqrt(5)],
- [np.sqrt(2)/np.sqrt(5)]])
```

```
>>> Y =
  ← np.array([[np.sqrt(2)/np.sqrt(5)],[np.sqrt(3)/np.sqrt(5)])]
>>> A@X+X
array([[0.00000000e+00],
       [3.33066907e-16]])
>>> A@Y-4*Y
array([[ -4.4408921e-16],
       [ 0.0000000e+00]])
```



On obtient bien deux vecteurs très proches du vecteur nul : ceci est dû à l'approximation de la racine carrée qu'utilise numpy.

Exercice ALG.6.6 | Soit $A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$.

Justifier que A est diagonalisable en base orthonormée, et trouver une base orthonormée de vecteurs propres.

Solution (exercice ALG.6.6)

Soit $A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$. On constate tout d'abord qu'elle est symétrique réelle donc d'après le théorème spectral, elle est diagonalisable en base orthonormée, cherchons

à présent ses valeurs propres et des vecteurs propres associés. Soit $\lambda \in \mathbf{R}$, alors :

$$A - \lambda I_3 = \begin{pmatrix} 1-\lambda & 0 & 2 \\ 0 & -1-\lambda & 0 \\ 2 & 0 & 1-\lambda \end{pmatrix} \begin{matrix} \\ \\ \left. \begin{matrix} L_1 \leftrightarrow L_3 \end{matrix} \right\} \\ \left. \begin{matrix} L_3 \leftrightarrow 2L_3 - (1-\lambda)L_1 \end{matrix} \right\} \end{matrix}$$

$$\sim \begin{pmatrix} 2 & 0 & 1-\lambda \\ 0 & -1-\lambda & 0 \\ 1-\lambda & 0 & 2 \end{pmatrix}$$

$$\sim \begin{pmatrix} 2 & 0 & 1-\lambda \\ 0 & -1-\lambda & 0 \\ 0 & 0 & 4-(1-\lambda)^2 \end{pmatrix}$$

$$= \begin{pmatrix} 2 & 0 & 1-\lambda \\ 0 & -1-\lambda & 0 \\ 0 & 0 & (1+\lambda)(3-\lambda) \end{pmatrix}.$$

On déduit alors que $\text{Spec}A = \{-1, 3\}$. Puis on calcule les espaces propres :

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_{-1}(A) \iff \begin{cases} 2x + 2z = 0 \\ 0 = 0 \\ 0 = 0 \end{cases} \iff \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \text{Vect} \left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right).$$

De-même

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_3(A) \iff \begin{cases} 2x - 2z = 0 \\ -4y = 0 \\ 0 = 0 \end{cases} \iff \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \text{Vect} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Par analyse des dimensions, nous avons que la matrice A est diagonalisable ($1+2=3$). Notons dans la suite $f_1 = (0, 1, 0)$, $f_2 = (1, 0, -1)$ et $f_3 = (1, 0, 1)$. Finalement, on constate que (f_1, f_2, f_3) est une famille orthogonale, il ne reste donc plus qu'à la normaliser en posant :

$$e_1 = f_1, \quad e_2 = \frac{1}{\sqrt{2}}(1, 0, -1), \quad e_3 = \frac{1}{\sqrt{2}}(1, 0, 1).$$

Donc :

$$A = PD^T P, \quad P = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Exercice ALG.6.7 | Dans cet exercice, \mathbf{R}^3 est muni de son produit scalaire euclidien et de sa base canonique notée \mathcal{C} . Soit A la matrice : $A = \begin{pmatrix} 11 & -5 & 5 \\ -5 & 3 & -3 \\ 5 & -3 & 3 \end{pmatrix}$ et f l'endomorphisme de \mathbf{R}^3 de matrice A dans la base canonique \mathcal{C} .

1. L'endomorphisme f est-il diagonalisable ?
2. Déterminer le noyau $K = \ker(f)$ puis une équation et une base \mathcal{B}' de $P = K^\perp$.
3. Vérifier que P est stable par f, c'est-à-dire que $f(P) \subset P$.
4. Déterminer la matrice B de l'endomorphisme g de P induit par f, dans la base \mathcal{B}' .
5. Montrer plus généralement que si F est un sous-espace vectoriel de \mathbf{R}^3 stable par f, alors F^\perp est également stable par f.
6. Déterminer les valeurs propres et les vecteurs propres de g.

Solution (exercice ALG.6.7)

1. La matrice A est symétrique réelle, donc diagonalisable et f est diagonalisable en base orthonormée, et est à valeurs propres.

2. On résout $A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff \begin{cases} 11x - 5y + 5z = 0 \\ -5x + 3y - 3z = 0 \\ 5x - 3y + 3z = 0 \end{cases} \iff \begin{cases} x = 0 \\ y = z \end{cases}$. Donc :

$$K = \ker f = \text{Vect}(0, 1, 1) = \text{Vect}(u).$$

Par conséquent $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in P = (\text{Vect } u)^\perp \iff \langle X | u \rangle = 0 \iff y + z = 0$.

Ainsi : $P = (\text{Vect}(u))^\perp$ est le plan d'équation $y + z = 0$ et on peut choisir pour base \mathcal{B}' de P la famille (v, w) avec $v = (1, 0, 0)$ et $w = (0, 1, -1)$.

3. Comme $P = \text{Vect}(v, w)$, soit $x \in P$, il s'agit de montrer que $Ax \in P$. Par hypothèse, il existe $\mu, \nu \in \mathbf{R}$ tels que $x = \lambda v + \mu w$. Donc $Ax = A(\lambda v + \mu w) = \lambda Av + \mu Aw$. On a

par calculs $Av = \begin{pmatrix} 11 \\ -5 \\ 5 \end{pmatrix} = 11v - 5w$. On lit sur les coordonnées $Av = 11v - 5w$. Et

$$Aw = \begin{pmatrix} 11 & -5 & 5 \\ -5 & 3 & -3 \\ 5 & -3 & 3 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -10 \\ 6 \\ -6 \end{pmatrix} = -10v + 6w.$$

Donc : $Ax = \lambda(11v - 5w) + \mu(-10v + 6w) = (11\lambda - 10\mu)v + (-5\lambda + 6\mu)w \in P$. Donc P est stable par A.

4. L'application $g \begin{matrix} \Big| \\ P \longrightarrow P \\ x \longmapsto f(x) \end{matrix}$ est donc bien définie d'après la question précédente, c'est un élément de $\mathcal{L}(P)$, on l'appelle l'«endomorphisme induit» de f sur P. De plus, les calculs précédents montrent aussi que :

$$B = \mathcal{M}_{\mathcal{B}'}(g) = \begin{pmatrix} 11 & -10 \\ -5 & 6 \end{pmatrix}.$$

5. Plus généralement, soit F un sous-espace stable par f : i.e. $\forall x \in F, f(x) \in F$. Soit maintenant $y \in F^\perp, \forall z \in F, \langle z | y \rangle = 0$. On veut montrer $f(y) \in F^\perp$. Soit donc $x \in F$, notons X, Y la matrice des coordonnées dans la base canonique de x, y : $\langle x | f(y) \rangle = {}^T X (AY) = ({}^T X) A Y = {}^T_{\text{Asym}} (AX) Y = \langle f(x) | y \rangle$, puisque $f(x) \in F$ est orthogonal à y par hypothèse, on obtient :

$$\langle x | f(y) \rangle = \langle f(x) | y \rangle = 0.$$



Ainsi $f(y)$ est orthogonal à F donc appartient à F^\perp pour tout y, donc F^\perp est stable par f.

6. $B \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix} \iff \begin{cases} (11 - \lambda)x - 10y = 0 \\ -5x + (6 - \lambda)y = 0 \end{cases}$

Le déterminant de ce système vaut $(11 - \lambda)(6 - \lambda) - 50 = \lambda^2 - 17\lambda + 16$. Les valeurs propres de B (donc de g) sont 1 et 16 et on obtient après calculs :

$$E_1(B) = \text{Vect} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad E_{16}(B) = \text{Vect} \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

Exercice ALG.6.8 | Endomorphismes orthogonaux Soit f un endomorphisme non nul de \mathbf{R}^n tel que : $\forall x, y \in \mathbf{R}^n, \langle f(x)|f(y) \rangle = \langle x|y \rangle$. On dit que f *préserve le produit scalaire*.

1. Montrer que f est orthogonal si et seulement si : $\forall x \in \mathbf{R}^n, \|f(x)\| = \|x\|$. Autrement dit, un endomorphisme est orthogonal si et seulement s'il préserve la norme.
Indication : On pourra, pour un sens, utiliser $\|f(x+y)\|^2$ avec $x, y \in \mathbf{R}^n$.
2. Déterminer $\ker(f)$, qu'en déduire sur f ?
3. Montrer que si λ est valeur propre de f , alors $|\lambda| = 1$. En déduire les valeurs propres (réelles) possibles pour f .
4. Dans cette question, on s'intéresse à la matrice M canoniquement associée à un endomorphisme f .
 - 4.1) Montrer que M vérifie ${}^TMM = M{}^TM = I_n$. On dit que M est une matrice *orthogonale*. Qu'en déduire sur M ? Était-ce prévisible?
 - 4.2)  En utilisant des fonctions du module numpy, créer une fonction d'entête est_orthogonale(M) qui prend en argument un tableau numpy et renvoie **True** si ce tableau est carré et correspond à une matrice orthogonale, **False** dans le cas contraire.
 - 4.3)  On se place dans la suite dans le cas $n = 2$. Déterminer, à l'aide de Python, les matrices orthogonales de $\mathfrak{M}_{2,2}(\mathbf{R})$ à coefficients entiers dans $[-5, 5]$.
 - 4.4) Déterminer, par le calcul, les matrices $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathfrak{M}_{2,2}(\mathbf{R})$ orthogonales de $\mathfrak{M}_{2,2}(\mathbf{R})$.

Solution (exercice ALG.6.8)

1. \Rightarrow Supposons f orthogonal. Alors en faisant $x = y \in \mathbf{R}^n$ dans la définition, on obtient $\|f(x)\|^2 = \|x\|^2$ puis le résultat s'en déduit en passant à la racine.

\Leftarrow Inversement, on a par hypothèse $\|f(x+y)\|^2 = \|x+y\|^2$ pour tous $x, y \in \mathbf{R}^n$. Alors ceci est, par linéarité de f , équivalent à

$$\begin{aligned} \langle f(x) + f(y) | f(x) + f(y) \rangle &= \langle x + y | x + y \rangle \\ \langle f(x) | f(x) \rangle + 2 \langle f(x) | f(y) \rangle + \langle f(y) | f(y) \rangle &= \langle x | x \rangle + 2 \langle x | y \rangle + \langle y | y \rangle \\ \langle x | x \rangle + 2 \langle f(x) | f(y) \rangle + \langle y | y \rangle &= \langle x | x \rangle + 2 \langle x | y \rangle + \langle y | y \rangle \\ \langle f(x) | f(y) \rangle &= \langle x | y \rangle. \end{aligned}$$

$\left. \begin{array}{l} f \\ \text{le} \\ \text{produit} \\ \text{scalaire} \end{array} \right\}$

Donc en conclusion, f est orthogonal si et seulement si : $\forall x \in \mathbf{R}^n, \|f(x)\| = \|x\|$.

2.

Exercice ALG.6.9 | Endomorphisme positif Soit $E = \mathbf{R}^n$, et $f \in \mathcal{L}(E)$. On note A la matrice de f dans la base canonique, et ${}^T f$ l'endomorphisme canoniquement associé à ${}^T A$. On souhaite déterminer les $f \in \mathcal{L}(E)$ tels que

$$\forall x \in E, \langle f(x) | x \rangle \geq 0 \tag{P}$$

1. Démontrer que : $\forall x \in \mathbf{R}^n, \langle x | f(x) \rangle = \langle x | {}^T f(x) \rangle$.
2. En déduire que f vérifie (P) si et seulement si $f + {}^T f$ vérifie (P).
3. Justifier qu'il existe une matrice $Q \in \mathfrak{M}_n(\mathbf{R})$ telle que ${}^T Q = Q^{-1}$ et une matrice diagonale $D \in \mathfrak{M}_n(\mathbf{R})$ telle que : $A + {}^T A = QDQ^{-1}$.
4. En déduire que f vérifie (P) si et seulement si $\text{Spec}(f + {}^T f) \in \mathbf{R}^+$.

Exercice ALG.6.10 | Décomposition de CHOLESKY Soit $n \geq 1$.

1. Soit $A \in \mathfrak{M}_n(\mathbf{R})$, on pose $S = A{}^T A$, montrer que S est diagonalisable et que ses valeurs propres sont positives.
2. Soit S une matrice de $\mathfrak{M}_n(\mathbf{R})$ qui est de plus symétrique et à valeurs propres positives. Montrer qu'il existe une matrice $A \in \mathfrak{M}_n(\mathbf{R})$ telle que $S = A{}^T A$.
3. **(Application)** On considère la matrice $S = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. Déterminer une matrice $A \in \mathfrak{M}_2(\mathbf{R})$ telle que $S = A{}^T A$.

Exercice ALG.6.11 | Matrice de dispersion d'un couple aléatoire Soient $Z = (X, Y)$ deux variables aléatoires admettant un moment d'ordre deux. On appelle *matrice de covariance* de $Z = (X, Y)$ la matrice

$$K_Z = \begin{pmatrix} \mathbf{Var}(X) & \mathbf{Cov}(X, Y) \\ \mathbf{Cov}(X, Y) & \mathbf{Var}(Y) \end{pmatrix}.$$

On définit l'espérance d'un vecteur ou plus généralement d'une matrice comme le vecteur (ou la matrice) des espérances.

1. Justifier que K_Z est une matrice diagonalisable.
2. Soit $(a, b) \in \mathbf{R}^2$. Montrer que

$$\mathbf{Var}(aX + bY) = \begin{pmatrix} a & b \end{pmatrix} \cdot K_Z \cdot \begin{pmatrix} a \\ b \end{pmatrix}.$$

En déduire que les valeurs propres de K_Z sont positives.

3. En déduire l'existence de deux matrices P, D telles que :
 - ▶ D soit diagonale,
 - ▶ P soit inversible d'inverse ${}^T P$,
 - ▶ $K_Z = PD^2P^{-1}$.

Solution (exercice ALG.6.11)

1. Puisque $\mathbf{Cov}(X, Y) = \mathbf{Cov}(Y, X)$, la matrice K_Z est une matrice symétrique réelle. Donc d'après le théorème spectral, elle est diagonalisable et est à valeurs propres réelles. Une matrice de passage associée est même orthogonale, ce fait sera exploité plus tard.

2. Soit $(a, b) \in \mathbf{R}^2$. Alors :

$$\begin{aligned} \mathbf{Var}(aX + bY) &= \mathbf{Cov}(aX + bY, aX + bY) \\ &= \mathbf{Cov}(aX, aX) + \mathbf{Cov}(aX, bY) + \mathbf{Cov}(bY, aX) + \mathbf{Cov}(bY, bY), \\ &= a^2\mathbf{Var}(X) + 2ab\mathbf{Cov}(X, Y) + b^2\mathbf{Var}(Y). \end{aligned}$$

Par un calcul de produit matriciel, on constate ensuite que

$$\begin{pmatrix} a & b \end{pmatrix} K_Z \begin{pmatrix} a \\ b \end{pmatrix} = a^2\mathbf{Var}(X) + 2ab\mathbf{Cov}(X, Y) + b^2\mathbf{Var}(Y).$$

bilibinéarité de la covariance

Donc

$$\mathbf{Var}(aX + bY) = \begin{pmatrix} a & b \end{pmatrix} K_Z \begin{pmatrix} a \\ b \end{pmatrix}.$$

Soit $\lambda \in \text{Spec } K_Z$, et notons $X = \begin{pmatrix} a \\ b \end{pmatrix}$ un vecteur propre associé. Alors $K_Z X = \lambda X$.

D'après la question précédente, en choisissant $\begin{pmatrix} a \\ b \end{pmatrix} = X$, on obtient alors

$$\begin{pmatrix} a & b \end{pmatrix} K_Z \begin{pmatrix} a \\ b \end{pmatrix} = \lambda(a^2 + b^2) = \mathbf{Var}(aX + bY) \geq 0,$$

dès lors, on obtient :

$$\lambda \geq 0.$$

Donc : les valeurs propres sont positives.

3. D'après le théorème spectral, il existe $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \in \mathfrak{M}_2(\mathbf{R})$ une matrice diagonale, et P inversible de taille 2×2 telles que :

$$K_Z = P \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} {}^T P.$$

Alors constatons que $\sqrt{\lambda_i}^2 = \lambda_i$ pour $i \in \{1, 2\}$ car les λ_i sont positifs. Notons $D = \begin{pmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \end{pmatrix}$. Alors

$$K_Z = PD^2 {}^T P.$$

C'est ce que l'on voulait.

4.3. Projection orthogonale

Exercice ALG.6.12 | Orthogonal de parties On munit \mathbf{R}^4 de son produit scalaire euclidien, et on considère les sous-espaces vectoriels E_1, E_2 définis ci-dessous. Déterminer pour chacun d'eux une base de E_i^\perp avec $i \in \{1, 2, 3\}$.

- $E_1 = \text{Vect}((3, 2, 0, 4), (1, 0, 0, -2), (1, -1, -1, 1))$.
- $E_2 = \{(x, y, z, t) \in \mathbf{R}^4 / 2x + 3y - t = 0\}$.

Solution (exercice ALG.6.12)

- Une famille génératrice de E_1 est (e_1, e_2, e_3) où $e_1 = (3, 2, 0, 4)$, $e_2 = (1, 0, 0, -2)$ et $e_3 = (1, -1, -1, 1)$. Donc :

$$(x, y, z, t) \in E_1^\perp \iff (x, y, z, t) \perp e_1, e_2, e_3 \iff \begin{cases} 3x + 2y + 4z &= 0 \\ x &- 2t = 0 \\ x - y - z + t &= 0 \end{cases}$$

$$\iff \begin{cases} y + 2z &= -3t \\ y + z &= 3t \\ x &= 2t \end{cases}$$

$$\iff \begin{cases} z &= -6t \\ y &= 9t \\ x &= 2t \end{cases}$$

Donc $E_1^\perp = \text{Vect}(2, 9, -6, 1)$.

- Cette fois-ci, l'espace vectoriel est donné sous forme d'une équation implicite (ou cartésienne). On peut commencer par chercher une famille génératrice.

$$(x, y, z, t) \in E_2 \iff 2x + 3y - t = 0 \iff t = 2x + 3y$$

$$\iff (x, y, z, t) = (x, y, z, 2x + 3y) = x(1, 0, 0, 2) + y(0, 1, 0, 3) + z(0, 0, 1, 0)$$

Donc $E_2 = \text{Vect}((1, 0, 0, 2), (0, 1, 0, 3), (0, 0, 1, 0))$. C'est un espace vectoriel de dimension trois, on s'y attendait puisque E_2 est un noyau de forme linéaire. Ensuite, on

procède comme avant :

$$(x, y, z, t) \in E_1^\perp \iff \begin{cases} x &+ 2t = 0 \\ y &+ 3t = 0 \\ z &= 0 \end{cases}$$

$$\iff \begin{cases} x &= -2t \\ y &= -3t \\ z &= 0 \end{cases}$$

$$\iff (x, y, z, t) = (-2t, -3t, 0, t) = t(-2, -3, 0, 1)$$

Donc $E_2^\perp = \text{Vect}((-2, -3, 0, 1))$.

Exercice ALG.6.13 | Projection orthogonale sur une partie et son orthogonal On munit \mathbf{R}^4 de son produit scalaire euclidien.

- Calculer la matrice de la projection orthogonale p sur $F = \text{Vect}(e_1, e_2)$ dans la base canonique avec $e_1 = (1, 0, 1, 0)$ et $e_2 = (0, 1, 1, 0)$.
- Déterminer F^\perp et écrire la matrice de la projection orthogonale q sur F^\perp .

Solution (exercice ALG.6.13)

- Calculons une expression analytique de la projection à l'aide de la définition. Soit

$(x, y, z, t) \in \mathbf{R}^4$, alors :

$$\begin{aligned}
 (x', y', z', t') = p_F(x, y, z, t) &\iff \begin{cases} (x', y', z', t') \in F \\ (x - x', y - y', z - z', t - t') \perp e_1 \\ (x - x', y - y', z - z', t - t') \perp e_2 \end{cases} \\
 &\iff \begin{cases} \exists \lambda, \mu \in \mathbf{R}, (x', y', z', t') = \lambda(1, 0, 1, 0) + \mu(0, 1, 1, 0), \\ (x - x') + (z - z') = 0, \\ (y - y') + (z - z') = 0, \end{cases} \\
 &\iff \begin{cases} \exists \lambda, \mu \in \mathbf{R}, (x', y', z', t') = (\lambda, \mu, \lambda + \mu, 0) \\ x' + z' = x + z, \\ y' + z' = y + z, \end{cases} \\
 &\iff \begin{cases} \exists \lambda, \mu \in \mathbf{R}, (x', y', z', t') = (\lambda, \mu, \lambda + \mu, 0), \\ 2\lambda + \mu = x + z, \\ \lambda + 2\mu = y + z, \end{cases}
 \end{aligned}$$

Il s'agit ensuite de résoudre le système précédent en λ, μ , on trouve :

$$\begin{cases} \exists \lambda, \mu \in \mathbf{R}, (x', y', z', t') = (\lambda, \mu, \lambda + \mu, 0), \\ 2\lambda + \mu = x + z, \\ \lambda + 2\mu = y + z, \end{cases} \iff \begin{cases} \exists \lambda, \mu \in \mathbf{R}, (x', y', z', t') = (\lambda, \mu, \lambda + \mu, 0), \\ \mu = \frac{1}{3}(-x + 2y + z), \\ \lambda = \frac{1}{3}(2x - y + z) \end{cases}$$

Donc: $(x', y', z', t') = (\lambda, \mu, \lambda + \mu, 0) = \left(\frac{1}{3}(2x - y + z), \frac{1}{3}(-x + 2y + z), \frac{1}{3}(x + y + 2z), 0\right)$.

Donc :

$$p_F(x, y, z) = \left(\frac{1}{3}(2x - y + z), \frac{1}{3}(-x + 2y + z), \frac{1}{3}(x + y + 2z), 0\right).$$

On déduit alors la matrice dans la base canonique :

$$\mathcal{M}_{\mathcal{B}^{\text{can}}}^{\text{at}}(p_F) = \frac{1}{3} \begin{pmatrix} 2 & -1 & 1 & 0 \\ -1 & 2 & 1 & 0 \\ 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

2. On a :

$$\begin{aligned}
 (x, y, z, t) \in F^\perp &\iff \begin{cases} x + z = 0 \\ y + z = 0 \end{cases} \\
 &\iff \begin{cases} x = -z \\ y = -z \end{cases} \iff (x, y, z, t) = (x, x, -x, t) = x(1, 1, -1, 0) + t(0, 0, 0, 1)
 \end{aligned}$$

On déduit $F^\perp = \text{Vect}((1, 1, -1, 0), (0, 0, 0, 1))$. Calculons la projection orthogonale demandée, en orthonormalisant la famille (u, v) où $u = (1, 1, -1, 0), v = (0, 0, 0, 1)$. On constate que $u \perp v$, il suffit donc de normaliser les deux vecteurs, posons donc :

$$f_1 = \frac{1}{\sqrt{3}}(1, 1, -1, 0), \quad f_2 = (0, 0, 0, 1).$$

Donc d'après le cours, pour tout $(x, y, z, t) \in \mathbf{R}^4$,

$$\begin{aligned}
 p_{F^\perp}(x, y, z, t) &= \langle (x, y, z, t) | f_1 \rangle f_1 + \langle (x, y, z, t) | f_2 \rangle f_2 \\
 &= \frac{1}{3}(x + y - z)(1, 1, -1, 0) + t(0, 0, 0, 1) = \frac{1}{3}(x + y - z, x + y - z, z - x - y, 3t)
 \end{aligned}$$

Donc $\mathcal{M}_{\mathcal{B}^{\text{can}}}^{\text{at}}(p_{F^\perp}) = \frac{1}{3} \begin{pmatrix} 1 & 1 & -1 & 0 \\ 1 & 1 & -1 & 0 \\ -1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$. Remarquons une chose : géométrie-ment, on se doute que $p_F + p_{F^\perp} = \text{Id}_{\mathbf{R}^4}$. Cette propriété est hors programme en BCPST, mais elle nous aurait permis d'affirmer directement que $p_{F^\perp}(x, y, z, t) = (x, y, z, t) - p_F(x, y, z, t)$ puis on utiliserait la question précédente.

Exercice ALG.6.14 |

1. On considère dans \mathbf{R}^2 la droite vectorielle $(\mathcal{D}) \quad x + y = 0$. On note p la projection orthogonale sur \mathcal{D} .

1.1) Déterminer la matrice A de p dans la base canonique de \mathbf{R}^2 .

- 1.2) Justifier l'existence d'une base orthonormale \mathcal{C} de \mathbf{R}^2 telle que $\mathcal{M}_{\mathcal{C}}(p)$ soit diagonale, et déterminer une telle base.
- 2. On considère dans \mathbf{R}^3 le plan vectoriel $(\mathcal{P}) \quad x + y - z = 0$. On note p la projection orthogonale sur \mathcal{P} .
 - 2.1) Déterminer la matrice A de p dans la base canonique de \mathbf{R}^3 .
 - 2.2) Justifier l'existence d'une base orthonormale \mathcal{C} de \mathbf{R}^3 telle que $\mathcal{M}_{\mathcal{C}}(p)$ soit diagonale, et déterminer une telle base.

Solution (exercice ALG.6.14)

- 1. 1.1) Une base de \mathcal{D} est $(1, -1)$. Donc une base orthonormée est $(1/\sqrt{2}, -1/\sqrt{2})$. Ainsi, d'après le cours, pour tout $(x, y) \in \mathbf{R}^2$:

$$p_{\mathcal{D}}(x, y) = \frac{1}{2} \langle (x, y) | (1, -1) \rangle (1, -1) = \frac{x-y}{2} (1, -1).$$

Donc
$$\mathcal{M}_{\mathcal{B}^{\text{can}}}(\mathcal{M}_{\mathcal{D}}(p_{\mathcal{D}})) = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

- 1.2) ...
- 2. 2.1) Une base de \mathcal{P} est $((1, 0, 1), (0, 1, 1))$. Soient $(x, y, z), (x', y', z') \in \mathbf{R}^3$, alors :

$$(x', y', z') = p_{\mathcal{P}}(x, y, z)$$

$$\Leftrightarrow \begin{cases} x' + y' - z' = 0 \\ (x - x', y - y', z - z') \in \mathcal{P}^{\perp} \end{cases},$$

$$\Leftrightarrow \begin{cases} x' + y' - z' = 0 \\ x' + z' = x + z \\ y' + z' = y + z \end{cases},$$

$$\Leftrightarrow \begin{cases} x' = \frac{1}{3}(2x - y + z) \\ y' = -\frac{1}{3}(x - 2y - z) \\ z' = \frac{1}{3}(x + y + 2z) \end{cases}.$$

On déduit alors la matrice de la projection dans la base canonique :

$$\mathcal{M}_{\mathcal{P}^{\text{F}}}(\mathcal{M}_{\mathcal{B}^{\text{can}}}(p)) = \frac{1}{3} \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

2.2) ...

Exercice ALG.6.15 | Calcul d'un extrema à l'aide d'une projection orthogonale Soit

$$f \begin{cases} \mathbf{R}^2 \longrightarrow \mathbf{R}, \\ (x, y) \longmapsto (x + y - 2)^2 + (2x + y - 1)^2 + (2x + y - 3)^2 + (3x + y - 2)^2. \end{cases}$$

Déterminer le minimum de f sur \mathbf{R}^2 . *Indication : On pourra commencer par interpréter $f(x, y)$ comme le carré d'une norme euclidienne, pour tous $(x, y) \in \mathbf{R}^2$.*

Solution (exercice ALG.6.15)

Soit $(x, y) \in \mathbf{R}^2$. Alors :

$$f(x, y) = \|(x + y - 2, 2x + y - 1, 2x + y - 3, 3x + y - 2)\|^2 = \|x(1, 2, 2, 3) + y(1, 1, 1, 1) - (2, 1, 3, 2)\|^2$$

Ainsi, puisque f est minorée elle admet une borne inférieure et :

$$\inf_{(x,y) \in \mathbf{R}^2} f(x, y) = \inf_{(x,y) \in \mathbf{R}^2} \|(2, 1, 3, 2) - x(1, 2, 2, 3) + y(1, 1, 1, 1)\|^2 = \inf_{v \in \mathcal{F}} \|(2, 1, 3, 2) - v\|^2$$

où $F = \text{Vect}((1,2,2,3), (1,1,1,1))$. On calcule donc tout d'abord $p_F((2,1,3,2))$. C'est l'unique vecteur $(a, b, c, d) \in \mathbf{R}^4$ tel que :

$$\begin{cases} (2,1,3,2) - (a,b,c,d) \perp (1,2,2,3), (1,1,1,1) \\ (a,b,c,d) \in F. \end{cases}$$

$$\Leftrightarrow \begin{cases} 2 - a + 2(1 - b) + 2(3 - c) + 3(2 - d) = 0, \\ 2 - a + 1 - b + 3 - c + 2 - d = 0, \\ (a, b, c, d) = x(1,2,2,3) + y(1,1,1,1), \quad x, y \in \mathbf{R} \end{cases}$$

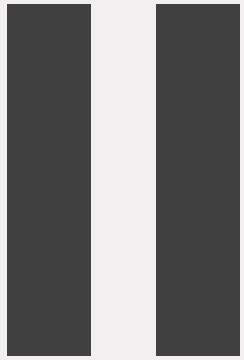
$$\Leftrightarrow \begin{cases} 2 - a + 2(1 - b) + 2(3 - c) + 3(2 - d) = 0, \\ 2 - a + 1 - b + 3 - c + 2 - d = 0, \\ (a, b, c, d) = x(1,2,2,3) + y(1,1,1,1), \quad x, y \in \mathbf{R} \end{cases}$$

$$\Leftrightarrow \begin{cases} 18x + 8y = 16, \\ 8x + 4y = 8, \\ (a, b, c, d) = (x + y, 2x + y, 2x + y, 3x + y), \quad x, y \in \mathbf{R} \end{cases}$$

Il reste ensuite à résoudre le système précédent en (x, y) , on trouve $x = 0, y = 2$ comme solution. Ainsi, d'après le cours, f est minimale au point $(0, 2)$. Si on souhaite obtenir la valeur dudit minimum, on remplace : $f(0, 2) = 2$, ou alors on calcule

$$\|(2, 1, 3, 2) - p_F((2, 1, 3, 2))\|^2 = \|(2, 1, 3, 2) - (0, 1, 2, 3) + 2(1, 1, 1, 1)\|^2,$$

mais c'est plus long.



Deuxième partie

Analyse

Chapitre ANA.7.

Fonctions de la variable réelle

Résumé & Plan

L'objectif de ce chapitre est de rappeler les éléments d'analyse réelle des fonctions d'une variable. Sa vocation n'est donc pas de se substituer à votre cours de première année, il faut le voir comme un complément.

Nous allons commencer par revoir des généralités sur les fonctions (du vocabulaire mais pas que), nous passerons ensuite aux propriétés analytiques : la continuité pour commencer, ensuite la dérivabilité, et enfin les développements limités qui généralisent l'approximation locale d'une fonction par sa tangente, afin de résoudre des problèmes de calculs de limite par exemple.

1	Généralités	2
2	Limites et continuité	6
2.1	Limite d'une fonction en un point	6
2.2	Continuité	11
3	Dérivation	17
3.1	Généralités et premières propriétés	17
3.2	Dérivées d'ordre supérieur	20
3.3	Extrema et théorème de ROLLE	21

4	Développements limités	25
4.1	Généralités	25
4.2	Développement géométrique	28
4.3	Développements obtenus par la formule de TAYLOR-YOUNG	28
4.4	Développements obtenus par primitivation	30
4.5	Développements obtenus par produits	32
4.6	Composition de développements limités	33
5	Exercices	35
5.1	Continuité	35
5.2	Dérivabilité	37
5.3	Fonctions usuelles	41
5.4	Développements limités	43
5.5	Méthode numérique	44

5.6 Du côté de l'Algèbre 46

Douter de tout ou tout croire, ce sont deux solutions également commodes, qui l'une et l'autre nous dispensent de réfléchir.

— **Henri POINCARÉ**

des couples $(x, f(x))$, où x parcourt D . Si on le note \mathcal{C}_f , on a alors, pour tout $(x, y) \in \mathbf{R} \times \mathbf{R}$:

$$(x, y) \in \mathcal{C}_f \iff [x \in D \text{ et } y = f(x)] .$$

Définition ANA.7.1 | Intervalle & Voisinage ouvert

On appelle *intervalle* de \mathbf{R} une partie convexe de \mathbf{R} , c'est-à-dire un ensemble I vérifiant la propriété suivante :

$$\forall (x, y) \in I^2, \quad (x \leq y \implies [x, y] \subset I).$$

Soit $x_0 \in \mathbf{R}$. On appellera *voisinage ouvert* de x_0 tout intervalle de la forme $]x_0 - \varepsilon, x_0 + \varepsilon[$ avec $\varepsilon > 0$.

Attention

Il ne faut pas confondre une fonction f et l'image $f(x)$ de x par f . En particulier, on écrira pas « la fonction $f(x)$ ».

1. GÉNÉRALITÉS

Dans cette première section, les fonctions seront définies sur une partie $D \subset \mathbf{R}$ et à valeurs dans \mathbf{R} .

Rappelons également qu'une fonction est une application, à ce titre elle ne peut posséder plusieurs images associées à un même antécédent. On rappelle également que si $f : D \rightarrow \mathbf{R}$ est une fonction, l'ensemble *image* noté $f(D)$ est :

$$f(D) = \{f(x), \quad x \in D\}.$$

Notation

$\mathcal{F}(D, \mathbf{R})$ (et \mathbf{R}^D) désignent l'ensemble des fonctions de D dans \mathbf{R} à valeurs réelles.

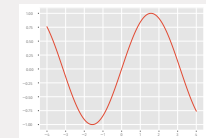
Définition ANA.7.2 | Ensemble de définition et graphe

L'*ensemble de définition* d'une fonction, ou ensemble de départ, est l'ensemble des réels en lesquels elle est définie. Le *graphe* (ou *représentation graphique* de f) d'une fonction $f : D \rightarrow \mathbf{R}$ définie sur une partie D de \mathbf{R} est l'ensemble constitué

TRACÉS DE GRAPHES DE FONCTIONS EN PYTHON. Pour rappel, les tracés de courbe en Python se font toujours par discrétisation : on relie un nuage de points construit à l'aide de Python, avec un espacement suffisamment petit.

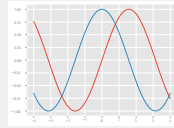
Représentation de la fonction sinus entre -4 et 4

```
import numpy as np
import matplotlib.pyplot as plt
X = np.linspace(-4, 4, 1000) #Maillage de
  1000 points de l'intervalle [-4,4]
Y = np.sin(X) #Fonction sin appliquée
  coordonnée par coordonnée
plt.plot(X, Y)
```



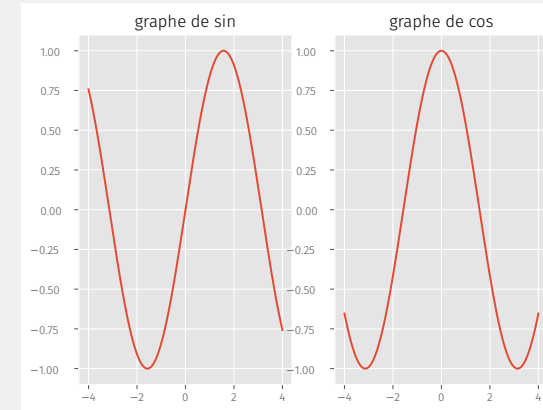
Représentation des fonctions sinus et cosinus entre -4 et 4

```
import numpy as np
import matplotlib.pyplot as plt
X = np.linspace(-4, 4, 1000) #Maillage de
↳ 1000 points de l'intervalle [-4,4]
Y = np.sin(X) #Fonction sin appliquée
↳ coordonnée par coordonnée
Z = np.cos(X) #Fonction cos appliquée
↳ coordonnée par coordonnée
plt.plot(X, Y)
plt.plot(X, Z)
```



Représentation des fonctions sinus et cosinus entre -4 et 4 (en grille)

```
import numpy as np
import matplotlib.pyplot as plt
X = np.linspace(-4, 4, 1000) #maillage de 1000 points de
↳ l'intervalle [-4,4]
Y = np.sin(X)
Z = np.cos(X) #Fonction appliquée coordonnée par coordonnée
plt.subplot(121) #grille de graphes de format 2 L x 1 C, 1er
↳ emplacement pour sin
plt.plot(X, Y)
plt.title('graphe de sin')
plt.subplot(122) #grille de graphes de format 2 L x 1 C, 2ème
↳ emplacement pour cos
plt.plot(X, Z)
plt.title('graphe de cos')
```



Dans les tracés précédents, on peut se passer de la commande `linspace` et utiliser des boucles classiques ou listes par compréhension. Nous en ferons des exemples en TP d'Informatique.

⊗ **Attention**

Vous devez savoir tracer des fonctions en autonomie, aucun rappel sur ce sujet n'est inscrit dans les sujets d'oraux.¹

Définition ANA.73 | Opérations sur les fonctions

Soient $f : D \rightarrow \mathbf{R}$ et $g : D \rightarrow \mathbf{R}$ deux fonctions définies sur D . On appelle *somme* de f et g l'application $f + g : D \rightarrow \mathbf{R}$ définie par :

$$\forall x \in D, \quad (f + g)(x) = f(x) + g(x).$$

De même, on appelle *produit* de f et g l'application $f g : D \rightarrow \mathbf{R}$ définie par :

$$\forall x \in D, \quad (f g)(x) = f(x)g(x).$$

Si g ne s'annule pas, alors on appelle *inverse* de g (resp. *quotient de f par g*) l'application $\frac{1}{g} : D \rightarrow \mathbf{R}$ (resp. $\frac{f}{g} : D \rightarrow \mathbf{R}$) définie par :

$$\forall x \in D, \quad \left(\frac{1}{g}\right)(x) = \frac{1}{g(x)}, \quad \left(\text{resp.} \forall x \in D, \quad \left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}\right)$$

¹à la rigueur, on vous rappellera que la commande `plot` existe

Soit $n \in \mathbf{Z}$. Si $n \geq 0$ ou si f ne s'annule pas, alors on appelle *puissance $n^{\text{ème}}$* de f l'application $f^n : D \rightarrow \mathbf{R}$ définie par :

$$\forall x \in D, \quad (f^n)(x) = (f(x))^n.$$

Attention

À l'inverse des applications linéaires, la puissance pour les fonctions réelles correspond au produit classique des réels.

Remarque 1.1 — On rappelle que

$$(\mathcal{F}(D, \mathbf{R}), +, \times)$$

est un \mathbf{R} -espace vectoriel : consulter le [Chapter ALG.3](#).

Définition ANA.7.4 | Parité/Imparité

Soit $f : D \rightarrow \mathbf{R}$ une fonction. On dit que f est *paire* si pour tout $x \in D$, on a :


$$-x \in D \quad \text{et} \quad f(-x) = f(x).$$

De même, f est dite *impaire* si pour tout $x \in D$, on a :

$$-x \in D \quad \text{et} \quad f(-x) = -f(x).$$

Proposition ANA.7.1 | Structure d'espace vectoriel des fonctions paires/impaires

Soit $D \subset \mathbf{R}$. Alors l'ensemble des fonctions paires et l'ensemble des fonctions impaires de \mathbf{R}^D sont des sous-espaces vectoriels de $E = \mathbf{R}^D$.

Preuve Faisons la preuve pour les fonctions paires. Elle est quasiment identique pour les fonctions impaires. 

Définition ANA.7.5 | Périodicité

Soit $T > 0$. Une fonction $f : D \rightarrow \mathbf{R}$ est dite *T-périodique* si pour tout $x \in D$, on a :


$$x + T \in D \quad \text{et} \quad f(x + T) = f(x).$$

Le réel T est appelé une *période* de la fonction. Une fonction est dite *périodique* s'il existe $T > 0$ tel que f soit T -périodique.

Attention

Il n'y a pas unicité de la période. En effet, une fonction T -périodique est *a fortiori* $2T$ -périodique, $3T$ -périodique, *etc.* Par ailleurs, une période est forcément positive.

Exemple 1 — La fonction \tan est π -périodique, et elle est définie sur $\mathbf{R} \setminus \left\{ \frac{\pi}{2} + \pi\mathbf{Z} \right\}$.

Exemple 2 — Déterminer la plus petite période de $f : x \mapsto \cos(2x) + \sin(3x)$. 

Définition ANA.7.6 | Monotonie pour le cas $\mathbf{R} = \mathbf{R}$

Une fonction $f : D \rightarrow \mathbf{R}$ est dite *croissante* (resp. *strictement croissante*) si :

$$\forall (x, y) \in D^2, \quad x \leq y \implies f(x) \leq f(y)$$

$$\text{(resp. } \forall (x, y) \in D^2, \quad x < y \implies f(x) < f(y)\text{)}.$$

Une fonction $f : D \rightarrow \mathbf{R}$ est dite *décroissante* (resp. *strictement décroissante*) si :

$$\forall (x, y) \in D^2, \quad x \leq y \implies f(x) \geq f(y)$$

$$\text{(resp. } \forall (x, y) \in D^2, \quad x < y \implies f(x) > f(y)\text{)}.$$

Une fonction $f : D \rightarrow \mathbf{R}$ est dite *monotone* (resp. *strictement monotone*) si elle est croissante ou décroissante (resp. strictement croissante ou strictement décroissante). Une fonction *constante* est une fonction croissante et décroissante.

Proposition ANA.7.2 | Structure d'espace vectoriel des fonctions périodiques

Soit $D \subset \mathbf{R}$ et $T > 0$. Alors l'ensemble des fonctions T -périodiques de \mathbf{R}^D est un sous-espace vectoriel de $E = \mathbf{R}^D$.

Preuve

**Attention aux confusions avec les suites!**

- ▶ Si pour tout $x \in D$ tel que $x + 1 \in D$, $f(x + 1) \geq f(x)$, on ne peut en général rien déduire sur la croissance de f .
- ▶ Si pour tout $x \in D$ tel que $x + 1 \in D$, $f(x + 1) \leq f(x)$, on ne peut en général rien déduire sur la décroissance de f .

Par exemple, la fonction f définie sur \mathbf{R}^+ par $f(n) = n$ pour tout $n \in \mathbf{N}$ et $f(x) = 0$ pour tout $x \in \mathbf{R}^+ \setminus \mathbf{N}$ vérifie $f(x + 1) \geq f(x)$ pour tout $x \in \mathbf{R}^+$ et pourtant elle n'est clairement pas croissante.

Proposition ANA.7.3 | Les fonctions monotones ne forment pas un sous-espace vectoriel des fonctions

Soit $D \subset \mathbf{R}$. Alors les ensembles de fonctions monotones (ou croissantes, décroissantes) ne sont pas des espaces vectoriels.

Preuve Pour les fonctions croissantes : il suffit de considérer $x \in D \mapsto x$. Elle est croissante, mais son opposé est décroissante. De même pour les fonctions décroissantes en considérant $x \in D \mapsto -x$. Ces deux ensembles ne sont pas stables par opposé, donc ne sont *a fortiori* pas des espaces vectoriels.

Pour les fonctions monotones, constatons, sur $D = \mathbf{R}^+$, que $f : x \mapsto x^2$ et $g : x \mapsto -x$ sont des fonctions monotones, mais $f + g$ n'est pas monotone sur \mathbf{R}^+ . Cet en-

semble n'est pas stable par addition, donc *a fortiori* n'est pas un espace vectoriel.

Définition ANA.7.7 | Borne

Une fonction $f : D \rightarrow \mathbf{R}$ est dite *majorée* (resp. *minorée*, *bornée*) si l'ensemble $f(D)$ est majoré (resp. minoré, borné). Autrement dit s'il existe $M \in \mathbf{R}$ (resp. $m \in \mathbf{R}$, $(m, M) \in \mathbf{R}^2$) tel que :

$$\begin{aligned} & \forall x \in D, \quad f(x) \leq M \\ \text{(resp. } & \forall x \in D, m \leq f(x), \\ & \forall x \in D, m \leq f(x) \leq M) \end{aligned}$$

Proposition ANA.7.4 | Structure d'espace vectoriel des fonctions bornées

L'ensemble des fonctions bornées forme un sous-espace vectoriel de l'ensemble des fonctions de D dans \mathbf{R} .

Preuve



Définition ANA.7.8 | Adhérence d'un ensemble

Soit I est un intervalle, on dit qu'un réel $x \in \mathbf{R}$ est *adhérent* à I si :

$$\forall \varepsilon > 0, \quad]x - \varepsilon, x + \varepsilon[\cap I \neq \emptyset,$$

alors on appelle *adhérence* de I l'ensemble des réels adhérents à I . On le note en général \bar{I} .

Ce sont donc des points « très proches » de I , mais pas forcément dedans.

Exemple 3 — $\sqrt{2}$ est adhérent à $[1, 2[$, alors que $\frac{5}{2}$ ne l'est pas.

Remarque 2.1 — Si I est un intervalle, alors \bar{I} est l'intervalle obtenu en adjoignant à I sa borne supérieure et sa borne inférieure dans $\bar{\mathbf{R}}$:

$$\bar{I} = I \cup \{\inf I, \sup I\}.$$

Ainsi, on généralise la notation $\bar{\mathbf{R}} = \mathbf{R} \cup \{-\infty, +\infty\}$.

2. LIMITES ET CONTINUITÉ



Cadre

Dans cette section, I désignera toujours un intervalle de \mathbf{R} non vide et non réduit à un point.

2.1. Limite d'une fonction en un point

Définition ANA.7.9 | Limite en un point adhérent

Soient $a \in \bar{I}$, $\ell \in \mathbf{R} \cup \{\pm\infty\}$ et une application f définie sur I ou sur $I \setminus \{a\}$. On note :

- ▶ $f(x) \xrightarrow{x \rightarrow a} \ell$ si :

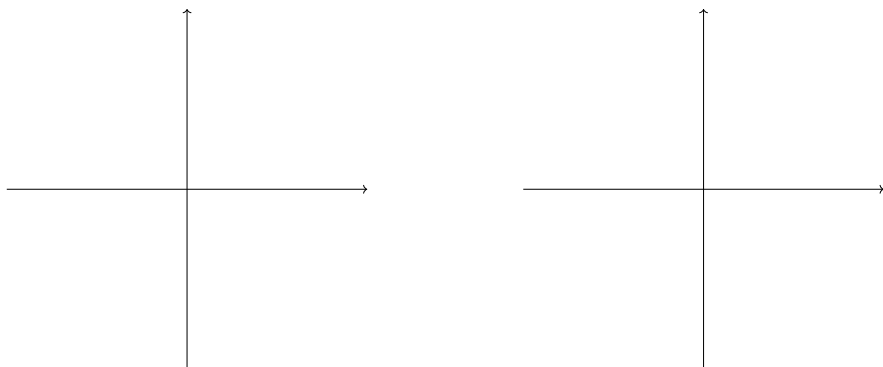
$$\forall \varepsilon > 0, \exists \eta_\varepsilon > 0, \forall x \in I, |x - a| < \eta_\varepsilon \implies |f(x) - \ell| < \varepsilon,$$
- ▶ $f(x) \xrightarrow{x \rightarrow a} +(-)\infty$ si :

$$\forall A \in \mathbf{R}, \exists \eta_A > 0, \forall x \in I, |x - a| < \eta_A \implies f(x) > (<)A.$$
- ▶ $f(x) \xrightarrow{x \rightarrow +(-)\infty} \ell$ si :

$$\forall \varepsilon > 0, \exists B_\varepsilon \in \mathbf{R}, \forall x \in I, x > (<)B_\varepsilon \implies |f(x) - \ell| < \varepsilon.$$
- ▶ $f(x) \xrightarrow{x \rightarrow +(-)\infty} +(-)\infty$ si :

$$\forall A \in \mathbf{R}, \exists B_A \in \mathbf{R}, \forall x \in I, x > (<)B_A \implies f(x) > (<)A.$$

Notez bien que a n'est pas nécessairement dans l'ensemble de définition de f , il est soit dedans, soit au pire très proche (dans \bar{I}). Illustrons deux des définitions sur un dessin.



Remarque 2.2 — Où intervient $a \in \bar{I}$? Si c'est le cas, l'ensemble $\{x \in I, |x - a| < \eta\}$ est non vide pour tout $\eta > 0$. Bien entendu c'est déjà le cas si $a \in I \subset \bar{I}$, mais ça l'est aussi plus généralement pour n'importe quel point de l'adhérence.

Proposition ANA.7.5 | Unicité de la limite

Soient $a \in \bar{I}$, $\ell \in \mathbf{R} \cup \{\pm\infty\}$ et une application f définie sur I ou sur $I \setminus \{a\}$. Si f admet une limite en a , alors cette limite est unique.

Preuve Traitons le cas d'une limite finie par exemple. Supposons que $f \xrightarrow{x \rightarrow a} \ell$ et $f \xrightarrow{x \rightarrow a} \ell'$ avec $\ell \neq \ell'$. Alors pour $\varepsilon = \frac{|\ell - \ell'|}{3}$, la définition de la limite donne : $f(x) \in]\ell - \varepsilon, \ell + \varepsilon[\cap]\ell' - \varepsilon, \ell' + \varepsilon[$ pour x assez proche de a . Comme ces deux intervalles sont disjoints, c'est absurde et $\ell = \ell'$.

Définition ANA.7.10 | Limite d'une fonction

Soient $a \in \bar{I}$ et $\ell \in \bar{\mathbf{R}}$ tels que $f(x) \xrightarrow{x \rightarrow a} \ell$. On dit que ℓ est la *limite* de f en a , ce que l'on note :

$$\ell = \lim_{x \rightarrow a} f(x) \quad \text{ou, plus simplement,} \quad \ell = \lim_a f.$$

Remarque 2.3 — Une fonction peut très bien ne pas avoir de limite (par exemple la fonction sin en $+\infty$). Si la proposition précédente assure l'unicité de la limite, on a toutefois un problème d'existence. On n'écrira donc jamais $\lim_a f$ avant d'en avoir prouvé l'existence.

Attention

Parfois on trouve dans la littérature les mêmes définitions qu'au-dessus mais avec « $x \in I \setminus \{a\}$ » au lieu de « $x \in I$ » (on parlera de *limite épointée*). Les différences entre ces deux notions sont minimes, mais sources de pièges. Par exemple, pour les limites épointées, la Proposition ANA.7.6 ci-dessous est fausse.

Proposition ANA.7.6 |

Soit $f : I \rightarrow \mathbf{R}$, où I est un intervalle, et $a \in \mathbf{R}$. Si f est **définie en** a (i.e. $a \in I$) et possède une limite en a , alors : $\lim_{x \rightarrow a} f(x) = f(a)$.

Preuve Faisons (par exemple) la preuve dans le cas où $\ell = \lim_{x \rightarrow a} f$ est finie. (déf.) Notons que la limite ℓ est nécessairement finie. En effet, si l'on suppose que $\lim_a f = +\infty$ par exemple, alors choisissons $\eta > 0$ tel que si $|x - a| < \eta$, alors $f(x) > f(a) + 1$

(par définition de la limite). En faisant $x = a$ on obtient alors une contradiction. Donc la limite est finie, et toujours par définition de la limite, on a :

$$\forall \varepsilon > 0, \exists \eta > 0, \forall x \in I, |x - a| < \eta \implies |f(x) - \ell| < \varepsilon.$$

Donc en faisant $x = a$: on obtient

$$\forall \varepsilon > 0, |f(a) - \ell| < \varepsilon.$$

Il reste à faire tendre ε vers 0 pour avoir le résultat : $\ell = f(a)$.

Proposition ANA.7.7

Si $f : I \rightarrow \mathbf{R}$ admet une limite finie en $a \in \bar{I}$, alors f est bornée au voisinage de a .

Preuve Faisons (par exemple) la preuve dans le cas où $a \in I$ (en particulier a est fini).

Considérer $\varepsilon = 1$, par exemple, dans la définition de la limite, montre que $\ell - 1 \leq f(x) \leq \ell + 1$ au voisinage de a , où $\ell = \lim_a f$.

On peut aussi affaiblir la définition de limite précédente et faire tendre x vers a que d'un côté. C'est ce que nous voyons dès à présent.

LIMITE À DROITE OU À GAUCHE. Une limite peut être caractérisée par une convergence à droite et à gauche.

Définition ANA.7.11 | Limite à droite/gauche

Soient $f : I \rightarrow \mathbf{R}$, $a \in \bar{I} \cap \mathbf{R}$ et $\ell \in \bar{\mathbf{R}}$. On pose $I_+ = I \cap]a, +\infty[$, et on suppose que $I_+ \neq \emptyset$.

On dit que f admet ℓ pour limite à droite en a si la restriction de f à I_+ admet ℓ pour limite en a . On note alors $\ell = \lim_{x \rightarrow a^+} f(x)$. On définit de même la notion de limite à gauche en a .

Proposition ANA.7.8

Soient $f : I \rightarrow \mathbf{R}$, $a \in \bar{I} \cap \mathbf{R}$. Alors :

$$f \text{ admet une limite en } a \iff \left\{ \begin{array}{l} f \text{ admet deux limites à droite et gauche,} \\ \text{qui sont de plus égales, en } a \end{array} \right.$$

Remarque 2.4 — Dans l'écriture avec des quantificateurs, on remplace $|x - a| < \eta$ par $a - \eta < x < a$ pour la limite à gauche, et $a < x < a + \eta$ pour la notion de limite à droite.

Exemple 4 — Les fonctions $x \mapsto \frac{1}{x^2}$ admet une limite à droite et à gauche en 0, égale à $+\infty$. Donc admet une limite en zéro qui vaut $+\infty$.

Ce n'est pas le cas de $x \mapsto \frac{1}{x}$ qui admet seulement une limite à droite et à gauche.

Exemple 5 — **Partie entière** Rappelons la définition de la partie entière et étudions la continuité de la fonction.

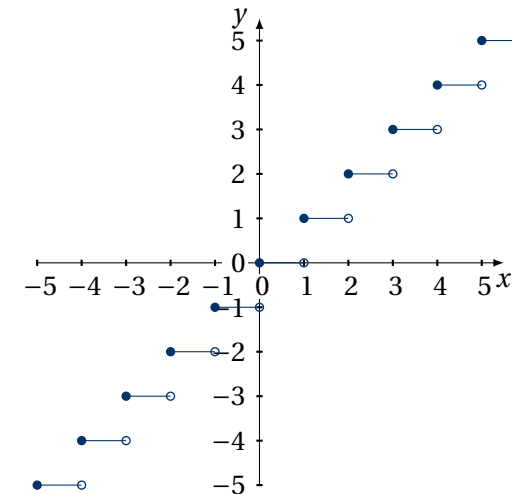


FIG. ANA.7.1. : Graphe de la partie entière



REFORMULATION SÉQUENTIELLE. On souhaite reformuler ici la définition de la limite avec des suites.

Théorème ANA.7.1 | Caractérisation séquentielle de la limite

Soient $a \in \bar{I}$ et $\ell \in \bar{\mathbf{R}}$. Alors :

$$f(x) \xrightarrow{x \rightarrow a} \ell \in \bar{\mathbf{R}} \iff \forall (x_n)_{n \in \mathbf{N}} \in I^{\mathbf{N}}, \quad x_n \xrightarrow{n \rightarrow \infty} a, \text{ on a : } f(x_n) \xrightarrow{n \rightarrow \infty} \ell.$$

Attention Le quantificateur \forall est très important.

Remarque 2.5 — Nous utiliserons largement ce fait dans l'étude des suites récurrentes du type « $x_{n+1} = f(x_n)$ ».

Preuve Procédons par double implication, dans les cas où $a \in \mathbf{R}$ et $\ell \in \mathbf{R}$ (les autres cas sont laissés en exercice).

\Rightarrow Supposons que $f(x) \xrightarrow{x \rightarrow a} \ell$. Soit (x_n) une suite de points de I convergeant vers a . Soit $\varepsilon > 0$. Comme $f(x) \xrightarrow{x \rightarrow a} \ell$, alors il existe $\alpha > 0$ tel que pour $x \in I$:

$$|x - a| \leq \alpha \implies |f(x) - \ell| \leq \varepsilon.$$

Or $x_n \xrightarrow{n \rightarrow \infty} a$, donc il existe $n_0 \in \mathbf{N}$ tel que pour $n \in \mathbf{N}$:

$$n \geq n_0 \implies |x_n - a| \leq \alpha.$$

On en déduit que pour tout $n \in \mathbf{N}$:

$$n \geq n_0 \implies |f(x_n) - \ell| \leq \varepsilon$$

c'est-à-dire $f(x_n) \xrightarrow{n \rightarrow \infty} \ell$.

\Leftarrow Prouvons l'implication réciproque par contraposée. Supposons donc que $f(x)$ ne tend pas vers ℓ lorsque x tend vers a . On en déduit qu'il existe donc $\varepsilon > 0$ tel que :

$$\forall \alpha > 0, \quad \exists x \in I, \quad |x - a| \leq \alpha \quad \text{et} \quad |f(x) - \ell| > \varepsilon.$$

Pour tout $n \in \mathbf{N}$, on peut appliquer ce qui précède à $\alpha = \frac{1}{n+1}$: il existe ainsi $x_n \in I$ tel que $|x_n - a| \leq \frac{1}{n+1}$ et $|f(x_n) - \ell| > \varepsilon$. Ainsi, (x_n) est une suite de points de I convergeant vers a , et pourtant la suite $(f(x_n))$ ne converge pas vers ℓ , puisqu'il existe $\varepsilon > 0$ tel que :

$$\forall n_0 \in \mathbf{N}, \quad \exists n \in \mathbf{N}, \quad n \geq n_0 \quad \text{et} \quad |f(x_n) - \ell| > \varepsilon$$

(tout entier $n \geq n_0$ convient).

Cet énoncé, comme le précise la méthode suivante, est utile pour montrer qu'une fonction $f : I \rightarrow \mathbf{R}$ n'admet pas de limite en $a \in \bar{I}$.

Méthode Nier l'existence d'une limite : fonctions d'une variable
Exhiber deux suites (u_n) et (v_n) d'éléments de I tendant vers a , mais telles que :
 $((f(u_n))_n$ et $(f(v_n))_n$ ne convergent pas vers la même limite.

Exemple 6 — L'application $f : \begin{cases} \mathbf{R}_+^* \rightarrow \mathbf{R} \\ x \mapsto \sin \frac{1}{x} \end{cases}$ n'admet pas de limite en 0^+ ?



PROPRIÉTÉS DE LA LIMITE. Nous rappelons dans ce paragraphe les propriétés (sans démonstration) de la limite à connaître.

Théorème ANA.7.2 | Stabilité des inégalités larges par passage à la limite

Soient I un intervalle, $a \in \bar{I}$ et $\ell \in \mathbf{R}$. Soient deux fonctions $f : I \rightarrow \mathbf{R}$, $g : I \rightarrow \mathbf{R}$ telles que :

- ▶ $f(x) \leq g(x)$ dans un voisinage de a ,
- ▶ $f(x) \xrightarrow{x \rightarrow a} \ell \in \mathbf{R}$,
- ▶ $g(x) \xrightarrow{x \rightarrow a} \ell' \in \mathbf{R}$.

Alors :

$$\ell \leq \ell'.$$

Théorème ANA.7.3 | Opérations sur les limites

Soient I un intervalle, $f : I \rightarrow \mathbf{R}$ et $g : I \rightarrow \mathbf{R}$ deux fonctions admettant *resp.* $\ell \in \bar{\mathbf{R}}$ et $\ell' \in \bar{\mathbf{R}}$ pour limites en $a \in \bar{I}$. Alors :

1. la fonction $|f|$ admet $|\ell|$ pour limite en a ,
2. si $\{\ell, \ell'\} \neq \{-\infty, +\infty\}$, alors la fonction $f + g$ admet $\ell + \ell'$ pour limite en a ,
3. pour tout $\lambda \in \mathbf{R}$, si $(\lambda, \ell) \neq (0, \pm\infty)$, alors la fonction λf admet $\lambda \ell$ pour limite en a ,
4. si $\{|\ell|, |\ell'|\} \neq \{0, +\infty\}$, alors la fonction $f g$ admet $\ell \ell'$ pour limite en a ,
5. si $\ell \neq 0$, alors la fonction $\frac{1}{f}$ admet $\frac{1}{\ell}$ pour limite en a ,
6. si $\ell' \neq 0$ alors la fonction $\frac{f}{g}$ admet $\frac{\ell}{\ell'}$ pour limite en a ,

7. si $\ell = 0$ et $f > 0$ dans un voisinage de a , alors $\frac{1}{f}$ admet $+\infty$ pour limite en a .

Théorème ANA.7.4 | Compositions de limites

Soient I et J deux intervalles, $f : I \rightarrow J$ et $g : J \rightarrow \mathbf{R}$ deux applications, $a \in \bar{I}$, $b \in \bar{J}$ et $\ell \in \bar{\mathbf{R}}$. Alors :

$$\left. \begin{array}{l} f(x) \xrightarrow{x \rightarrow a} b \\ g(y) \xrightarrow{y \rightarrow b} \ell \end{array} \right\} \Rightarrow g(f(x)) \xrightarrow{x \rightarrow a} \ell.$$


Théorème ANA.7.5 | Théorème d'encadrement

Soient I un intervalle, $a \in \bar{I}$ et $\ell \in \mathbf{R}$. On considère trois fonctions $f : I \rightarrow \mathbf{R}$, $g : I \rightarrow \mathbf{R}$ et $h : I \rightarrow \mathbf{R}$ telles que :

1. $f \leq g \leq h$ au voisinage de a ,
2. les deux fonctions f et h admettent ℓ pour limite en a .

Alors :

$$g(x) \xrightarrow{x \rightarrow a} \ell.$$

Exemple 7 — L'application $f : \begin{cases} \mathbf{R}_+^* \rightarrow \mathbf{R} \\ x \mapsto x \sin \frac{1}{x} \end{cases}$ admet-elle une limite en 0^+ ? 

2.2. Continuité

Rappelons les principaux éléments relatifs à la continuité des fonctions.

Définition ANA.7.12 | en un point

Soient $f : I \rightarrow \mathbf{R}$ et $a \in I$. On dit que f est *continue en a* si $\lim_{x \rightarrow a} f(x) = f(a)$, i.e.

$$\forall \varepsilon > 0, \exists \alpha > 0, \forall x \in I, |x - a| < \alpha \implies |f(x) - f(a)| < \varepsilon.$$

Définition ANA.7.13 | sur un intervalle

Soient I un intervalle et $f : I \rightarrow \mathbf{R}$ une fonction. On dit que f est *continue sur I* si elle est continue en tout $a \in I$. Ainsi, f est continue sur I si et seulement si :

$$\forall a \in I, \forall \varepsilon > 0, \exists \alpha > 0, \forall x \in I, |x - a| < \alpha \implies |f(x) - f(a)| < \varepsilon.$$

On note $\mathcal{C}^0(I, \mathbf{R})$, l'ensemble des fonctions continues sur I à valeurs réelles.

Définition ANA.7.14 | Prolongement par continuité

Soit f une fonction définie sur $I \setminus \{a\}$ telle que : f admet une limite finie ℓ en a . On appelle alors *prolongement par continuité de f en a* la fonction \tilde{f} définie sur I en par :

$$\tilde{f}(a) = \ell, \quad \tilde{f} = f \quad \text{sur } I \setminus \{a\}.$$

La fonction \tilde{f} est souvent encore notée $f : I \rightarrow \mathbf{R}$, et elle est continue en a . On dit aussi que l'on a effectué un *prolongement par continuité de f en a* .

Exemple 8 — Prolonger par continuité les deux fonctions ci-dessous aux bornes de leur ensemble de définition.

$$\begin{aligned} 1. f &: \begin{cases} \mathbf{R}^* \rightarrow \mathbf{R}, \\ x \rightarrow \frac{\sin x}{x}, \end{cases} \\ 2. g &: \begin{cases} \mathbf{R}^* \rightarrow \mathbf{R}, \\ x \rightarrow \frac{\arctan x}{x}. \end{cases} \end{aligned}$$

Théorème ANA.7.6 | Théorème de minoration

Soient I un intervalle et $a \in \bar{I}$. On considère deux fonctions $f : I \rightarrow \mathbf{R}$ et $g : I \rightarrow \mathbf{R}$ telles que :

- $f \leq g$ au voisinage de a ,
 - $f(x) \xrightarrow{x \rightarrow a} +\infty$.
- Alors : $g(x) \xrightarrow{x \rightarrow a} +\infty$.

Théorème ANA.7.7 | Théorème de majoration

Soient I un intervalle et $a \in \bar{I}$. On considère deux fonctions $f : I \rightarrow \mathbf{R}$ et $g : I \rightarrow \mathbf{R}$ telles que :

- $f \leq g$ au voisinage de a ,
 - $g(x) \xrightarrow{x \rightarrow a} -\infty$.
- Alors : $f(x) \xrightarrow{x \rightarrow a} -\infty$.

Théorème ANA.7.8 | Théorème de la limite monotone

Soient $(a, b) \in \bar{\mathbf{R}}^2$, tel que $a < b$, et $f :]a, b[\rightarrow \mathbf{R}$ une fonction croissante.

- Si f est majorée, alors f admet une limite finie en b .
- Si f n'est pas majorée, alors : $f(x) \xrightarrow{x \rightarrow b} +\infty$.



REFORMULATION SÉQUENTIELLE. La caractérisation séquentielle de la limite nous livre donc directement celle de la continuité.

Proposition ANA.7.9 | Caractérisation séquentielle de la continuité en un point

Soient I un intervalle, $f : I \rightarrow \mathbf{R}$ et $a \in I$. Alors :

f est continue en $a \iff \forall (x_n)_{n \in \mathbf{N}} \in I^{\mathbf{N}}, x_n \xrightarrow{n \rightarrow \infty} a, \text{ on a : } f(x_n) \xrightarrow{n \rightarrow \infty} f(a).$



Méthode Continuité et permutation de limites

Il faut surtout retenir la caractérisation séquentielle de la manière suivante : si f est continue, et avec les mêmes notations que *supra*,

$$f\left(\lim_{n \rightarrow \infty} (\dots)\right) = \lim_{n \rightarrow \infty} f(\dots).$$

PROPRIÉTÉS LOCALES DES FONCTIONS CONTINUES. De-même que les opérations sur les limites livrent celles sur la continuité.

Proposition ANA.7.10 | Opérations sur les fonctions continues en un point

Soient I un intervalle, $f : I \rightarrow \mathbf{R}$ et $g : I \rightarrow \mathbf{R}$ deux fonctions continues en $a \in I$.

Alors les fonctions $|f|$, $f + g$, λf (où $\lambda \in \mathbf{R}$) et fg sont encore continues en a . En

outre, si $g(a) \neq 0$, alors $\frac{f}{g}$ est définie sur un voisinage de a et est continue en a .

Théorème ANA.7.9 | Compositions de fonctions continues

Soient I et J deux intervalles, f et g deux fonctions composables :

$$\begin{array}{c} I \xrightarrow{f} J \xrightarrow{g} \mathbf{R} \\ \quad \quad \quad \curvearrowright \\ \quad \quad \quad g \circ f \end{array}$$

et $a \in I$. Si f est continue en a et g est continue en $f(a)$, alors $g \circ f$ est continue en a .

PROPRIÉTÉS GLOBALES DES FONCTIONS CONTINUES. On déduit immédiatement de la définition les versions globales des deux énoncés précédents :

- ▶ Soient $f : I \rightarrow \mathbf{R}$ et $g : I \rightarrow \mathbf{R}$ deux fonctions continues sur I . Alors les fonctions $|f|, f+g, \lambda f$ (où $\lambda \in \mathbf{R}$) et fg sont encore continues sur I . En outre, si g ne s'annule pas, alors $\frac{f}{g}$ est continue sur I .
- ▶ Soient I et J deux intervalles, f et g deux fonctions composables :
 $I \xrightarrow{f} J \xrightarrow{g} \mathbf{R}$ Si f est continue sur I et g est continue sur J , alors $g \circ f$ est continue sur I .

IMAGE D'UN INTERVALLE PAR UNE FONCTION CONTINUE, THÉORÈME DES VALEURS INTERMÉDIAIRES.

Théorème ANA.7.10 | Théorème des valeurs intermédiaires

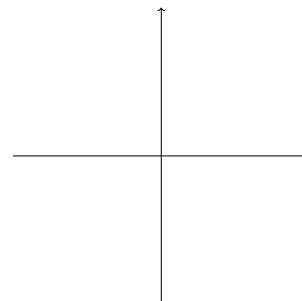
Soient I un intervalle, $(a, b) \in \mathbf{I}^2$ tels que $a \leq b$, $f : I \rightarrow \mathbf{R}$ une application **continue** et c un réel entre $f(a)$ et $f(b)$. Alors :
 il existe $x_0 \in [a, b]$ tel que $f(x_0) = c$.

Nous présentons ici une preuve « constructive » du théorème : nous allons construire x_0 comme la limite de deux suites adjacentes.

Preuve (*Approximation de x_0 par dichotomie*) Supposons pour simplifier que f est croissante $f(a) \leq f(b)$. On va définir deux suites (a_n) et (b_n) telles que les intervalles $[a_n, b_n]$ soient de plus en plus petits, et on montre qu'elles convergent vers une même limite x_0 .

Posons $a_0 = a$ et $b_0 = b$, puis par récurrence :

$$\begin{cases} a_{n+1} = a_n & \text{et } b_{n+1} = \frac{a_n + b_n}{2} & \text{si : } f\left(\frac{a_n + b_n}{2}\right) \geq c, \\ a_{n+1} = \frac{a_n + b_n}{2} & \text{et } b_{n+1} = b_n & \text{si : } f\left(\frac{a_n + b_n}{2}\right) < c. \end{cases}$$




Alors, on montre par récurrence la propriété (\mathcal{P}_n) :

$$a_n, b_n \in [a, b], \quad a = a_0 \leq \dots \leq a_n, \quad b_n \leq \dots \leq b_0 = b,$$

$$b_n - a_n = \frac{b - a}{2^n}, \quad f(a_n) \leq c \leq f(b_n).$$



On est à présent capables de conclure. 

On se rappelle la **Définition ANA.8.1** d'un intervalle vue en début de chapitre, une version alternative du théorème des valeurs intermédiaires consiste à dire que l'image d'un intervalle (un ensemble «sans trou») par une application continue est également un intervalle (également «sans trou»).

Cette dernière preuve nous livre directement un algorithme pour approcher ce type de zéros.

Recherche d'un zéro par dichotomie

```
def dichotomie(a, b, f, prec):
    """
    Retourne une valeur approchée d'un zéro de f entre a et b avec
    ↪ précision prec
    Retourne faux si aucune racine n'existe
    """
    while b - a > prec:
        c = (a + b)/2
        if f(a)*f(c) <= 0:
            # changement de signe sur [a,c]
            b = c
        else:
            # pas de changement de signe sur [a,c]
            a = c
    return (a + b)/2
```


On peut aussi adopter une version récursive.

Recherche d'un zéro par dichotomie, version récursive


```
def dichotomie_rec(a, b, f, prec):
    """
    Retourne une valeur approchée d'un zéro de f entre a et b avec
    ↪ précision prec
    selon un principe dichotomique
    Retourne faux si aucune racine n'existe
    """
    if f(a)*f(b) > 0:
        return False
    if b - a <= prec:
        return (a + b)/2
    else:
        c = (a + b)/2
        if f(a)*f(c) <= 0:
            # changement de signe sur [a,c]
            return dichotomie_rec(a, c, f, prec)
        else:
            # pas de changement de signe sur [a,c]
            return dichotomie_rec(c, b, f, prec)
```

Il existe encore d'autres méthodes plus sophistiquées (méthode de NEWTON, de la sécante, *etc.*...), mais qui ne sont pas à notre programme, nous les aborderons donc uniquement en TP d'Informatique². La démonstration précédente prouve la convergence de ces deux fonctions.

Exemple 9 – Variations autour de la dichotomie Adapter les codes précédents pour résoudre une équation de la forme (on admet l'existence d'une unique solution) :

1. $f(x) = \alpha$ avec $\alpha \in \mathbf{R}$, 

²la méthode de NEWTON a notamment fait l'objet d'une partie du sujet de Modélisation 2020

2. $f(x) = x$. Une telle solution est donc un *point fixe* de f , voir ci-dessous. 

Dans ces deux exemples, on peut aussi très bien utiliser les fonctions *dicho* et *dicho_rec* précédentes que l'on applique aux fonctions $x \mapsto f(x) - \alpha$ et $x \mapsto f(x) - x$.

Exemple 10 — Toute fonction polynomiale réelle de degré impair s'annule au moins une fois sur \mathbf{R} . Nous l'avons montré dans le [Chapter ALG.2](#).

Définition ANA.7.15 | Point fixe

On appelle *point fixe* d'une fonction $f : D \rightarrow \mathbf{R}$ avec $D \subset \mathbf{R}$ tout réel $c \in D$ tel que $f(c) = c$.

Exemple 11 — Soit $f : [0, 1] \rightarrow [0, 1]$ une application continue. Alors, f admet un *point fixe* : il existe $c \in [0, 1]$ tel que $f(c) = c$. 

Le théorème des valeurs intermédiaires est à utiliser lorsque l'on souhaite prouver l'existence d'un réel solution à une équation. Lorsque l'on souhaite montrer l'unicité, on aura recours au théorème de la bijection sur un intervalle bien choisi.

Théorème ANA.7.11 | Théorème de la bijection

Soient I un *intervalle* et $f : I \rightarrow \mathbf{R}$ une fonction continue et strictement monotone. Alors :

1. $f(I)$ est un intervalle.
2. f réalise une bijection de I sur $f(I)$.
3. la bijection réciproque $f^{-1} : f(I) \rightarrow I$ est continue et strictement monotone, de même sens de monotonie que f .

Preuve Voir cours de première année.

Dans la première assertion, on peut préciser l'intervalle $f(I)$, selon la monotonie de f et l'intervalle I . Par exemple, si f est décroissante alors :

$$I =]a, b] \quad \Rightarrow \quad f(I) = [f(b), \lim_a f[.$$

⊗ Attention

La fonction f n'est pas nécessairement globalement³ bijective, mais elle **réalise une bijection de I sur $f(I)$** . Attention à la distinction entre les deux.


🔧 Méthode Utilisation du théorème de la bijection ou du théorème des valeurs intermédiaires


On souhaite justifier l'existence et l'unicité éventuelle d'une solution $x \in \mathbf{R}$ à l'équation $f(x) = \alpha$ avec $\alpha \in \mathbf{R}$.

1. Si l'unicité n'est pas souhaitée : on applique le théorème des valeurs intermédiaires.
2. Si l'unicité est souhaitée : on applique le théorème de la bijection sur un intervalle I tel que $\alpha \in f(I)$.

Notez également que l'équation $f(x) = x$ est équivalente à $g(x) = \alpha$ avec $g = f - \text{Id}$ et $\alpha = 0$.

Exemple 12 — Utilisation du théorème de la bijection dans deux contextes

1. Soit $t \geq 0$. Le polynôme $x \mapsto x^3 + tx - 1$ admet une unique racine réelle $u(t)$. 

2. **(Existence d'une suite implicite)** Soit $n \geq 1$, et $f_n :]0, \infty[\rightarrow x^{n+1} - x^n$. L'équation $f_n(x) = 1$ admet une unique solution $\alpha_n \in [1, \infty[$. 

³c'est-à-dire sur son ensemble de définition tout entier

IMAGE D'UN SEGMENT PAR UNE FONCTION CONTINUE. Le théorème suivant est également fondamental et connu sous le nom de théorème des « bornes atteintes ».


Théorème ANA.7.12 | Bornes atteintes

Soit $f : [a, b] \rightarrow \mathbf{R}$ une application continue. Alors :
 f est bornée sur $[a, b]$ et ses bornes sont atteintes.

Attention

Il n'y a aucune raison pour que les bornes soient atteintes en a et b , donc sur les bords de l'intervalle. Penser à la fonction sinus sur $[0, 2\pi]$ par exemple.

Preuve Admis.

Exemple 13 — Soit $f : \mathbf{R} \rightarrow \mathbf{R}$ une fonction continue, alors $f \circ \sin$ est bornée. 

3. DÉRIVATION

Dans cette partie, I désignera encore et toujours un intervalle de \mathbf{R} *non trivial*, c'est-à-dire non vide et non réduit à un point.

3.1. Généralités et premières propriétés

Définition ANA.7.16 | Dérivabilité

Soient $f : I \rightarrow \mathbf{R}$ une application et $x_0 \in I$. On dit que f est *dérivable* en x_0 si l'application

$$\begin{array}{l} I \setminus \{x_0\} \longrightarrow \mathbf{R} \\ x \longrightarrow \frac{f(x) - f(x_0)}{x - x_0} \end{array}$$

admet une limite finie⁴ en x_0 . La limite est alors appelée le *nombre dérivé de f en x_0* et on le note $f'(x_0)$, ou encore $\frac{df}{dx}(x_0)$.

Le nombre $\frac{f(x) - f(x_0)}{x - x_0}$ est appelé le *taux d'accroissement de f entre x_0 et x* et s'interprète comme la pente de la corde du graphe de f entre les points d'abscisses x_0 et x . Lorsque f est dérivable en x_0 , le nombre $f'(x_0)$ s'interprète alors comme la « pente limite » de ces cordes. Par définition, c'est la pente de la tangente au graphe de f au point d'abscisse x_0 .

⁴ Comme $x_0 \in I$ et que I est un intervalle, alors $x_0 \in \overline{I \setminus \{x_0\}}$ et calculer la limite de ce quotient a bien un sens.

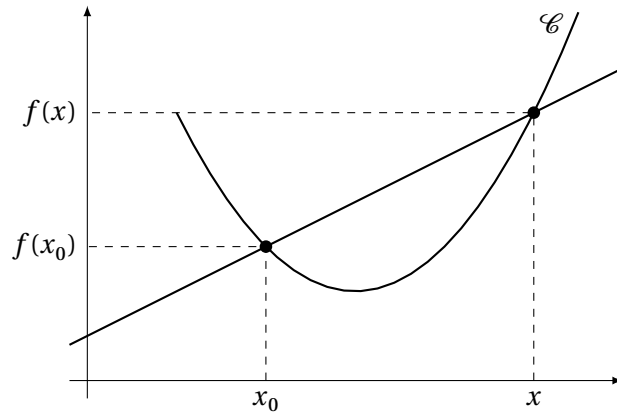


FIG. ANA.7.2. : Nombre dérivé

De même, le cas échéant, $f'_d(x_0)$ (resp. $f'_g(x_0)$) est la pente de la demi-tangente à droite (resp. à gauche) au graphe de f en x_0 .

Définition ANA.7.17 | Dérivabilité à droite ou gauche

On dit que f est *dérivable à droite* (resp. *dérivable à gauche*), si $f_{|I \cap]x_0, +\infty[}$ (resp. $f_{|I \cap]-\infty, x_0]}$) est dérivable en x_0 . Le nombre dérivé de la restriction s'appelle alors la *dérivée à droite* (resp. *dérivée à gauche*) de f en x_0 et se note $f'_d(x_0)$ (resp. $f'_g(x_0)$).

Définition ANA.7.18 | Ensemble $\mathcal{D}^1(I, \mathbf{R})$

Soit $f : I \rightarrow \mathbf{R}$ une application. On dit que f est *dérivable sur* I si f est dérivable en x pour tout $x \in I$. On définit alors la *fonction dérivée* de f ainsi :

$$f' : \begin{cases} I & \rightarrow \mathbf{R}, \\ x & \rightarrow f'(x). \end{cases}$$

On la note f' , ou encore $\frac{df}{dx}$. On note $\mathcal{D}^1(I, \mathbf{R})$ l'ensemble des fonctions dérivables sur I à valeurs réelles.

Notation Dérivée d'une expression

Soit une expression $f(x)$ dépendant de $x \in \mathbf{R}$, avec f une fonction dérivable. On notera dans la suite indifféremment :

- ▶ $\frac{df}{dx}(x)$ la fonction f' évaluée en x ,
- ▶ $\frac{d}{dx}[f(x)]$ la dérivée de l'expression $f(x)$ par rapport à x .

En particulier, on évitera dans la mesure du possible d'écrire $f(x)'$.

Définition ANA.7.19 | Ensemble dans $\mathcal{C}^1(I, \mathbf{R})$

Une fonction $f : I \rightarrow \mathbf{R}$ est dite *de classe \mathcal{C}^1 sur* I si elle est dérivable sur I et si f' est continue sur I . On note $\mathcal{C}^1(I, \mathbf{R}) \subset \mathcal{D}^1(I, \mathbf{R})$ l'ensemble de ces fonctions.

Théorème ANA.7.13 | Dérivabilité et développement limité à l'ordre un

Une fonction $f : I \rightarrow \mathbf{R}$ est dérivable en $x_0 \in I$ si et seulement s'il existe $c \in \mathbf{R}$ tel que :

$$f(x) \underset{x \rightarrow x_0}{=} f(x_0) + c(x - x_0) + o(x - x_0).$$

On a alors $c = f'(x_0)$. En particulier, si f est dérivable en x_0 , alors f est continue en x_0 .

PROPRIÉTÉS DE LA DÉRIVABILITÉ. Comme pour les propriétés sur les limites, celles sur la dérivabilité sont rappelées sans démonstration.

Théorème ANA.7.14 | Opérations sur les fonctions dérivables et les dérivées

Soient f et g deux fonctions de I dans \mathbf{R} dérivables en $x_0 \in I$.

- ▶ $f + g$ est dérivable en x_0 et $(f + g)'(x_0) = f'(x_0) + g'(x_0)$,
- ▶ pour tout $\lambda \in \mathbf{R}$, λf est dérivable en x_0 et $(\lambda f)'(x_0) = \lambda f'(x_0)$,
- ▶ $f g$ est dérivable en x_0 et $(f g)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$,
- ▶ si $g(x_0) \neq 0$, alors il existe un voisinage de x_0 sur lequel g ne s'annule pas, $\frac{1}{g}$ est dérivable en x_0 et :

$$\left(\frac{1}{g}\right)'(x_0) = -\frac{g'(x_0)}{g(x_0)^2}.$$

- ▶ si $g(x_0) \neq 0$, alors il existe un voisinage de x_0 sur lequel g ne s'annule pas, $\frac{f}{g}$ est dérivable en x_0 et :

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}.$$

- ▶ Pour tout $n \in \mathbf{Z}$ (en supposant que $f(x_0) \neq 0$ si $n \leq 0$), alors f^n (puissance $n^{\text{ème}}$) est dérivable en x_0 et :
 $(f^n)'(x_0) = n f^{n-1}(x_0) f'(x_0)$.

Corollaire ANA.7.1 | Reformulation linéaire

- ▶ Les triplets $(\mathcal{D}^1(I, \mathbf{R}), +, \times|_{\mathbf{R} \times \mathcal{D}^1(I, \mathbf{R})})$, $(\mathcal{C}^1(I, \mathbf{R}), +, \times|_{\mathbf{R} \times \mathcal{D}^1(I, \mathbf{R})})$ sont des sous-espaces vectoriels de $\mathcal{F}^1(I, \mathbf{R})$.
- ▶ La dérivation $\frac{d}{dx} : \mathcal{D}^1(I, \mathbf{R}) \rightarrow \mathcal{F}^1(I, \mathbf{R})$ est une application linéaire.

Théorème ANA.7.15 | Dérivation d'une composée

Soient $f : I \rightarrow J$ et $g : J \rightarrow \mathbf{R}$ dérivables en $x_0 \in I$ et $f(x_0) \in J$. Alors $g \circ f$ est dérivable en x_0 et :

$$(g \circ f)'(x_0) = g'(f(x_0)) \cdot f'(x_0).$$

Par conséquent, si $f : I \rightarrow J$ et $g : J \rightarrow \mathbf{R}$ sont deux fonctions dérivables sur I et J . Alors $g \circ f$ est dérivable sur I et :

$$(g \circ f)' = (g' \circ f) \cdot f'.$$

Appliquant ceci à la relation $f \circ f^{-1} = \text{Id}$ lorsque f est bijective d'inverse f^{-1} , nous obtenons le corollaire suivant.

Corollaire ANA.7.2 | Dérivation d'une bijection réciproque

Soit $f : I \rightarrow J$ une application bijective et continue et $y_0 \in J$. Si :

- ▶ f est dérivable en $f^{-1}(y_0) \in I$,
- ▶ et $f'(f^{-1}(y_0)) \neq 0$,

alors f^{-1} est dérivable en y_0 et :

$$(f^{-1})'(y_0) = \frac{1}{f'(f^{-1}(y_0))} \text{ } ^5.$$

En particulier, si $f : I \rightarrow J$ est une bijection, dérivable sur I , alors si f' ne s'annule pas, la fonction f^{-1} est dérivable sur J et :

$$(f^{-1})' = \frac{1}{f' \circ f^{-1}}.$$

L'exemple ci-dessous est à très bien connaître.

Exemple 14 — Existence d'arctan et graphe.

1. Rappeler le principe de construction de arctan, ainsi que son graphe.

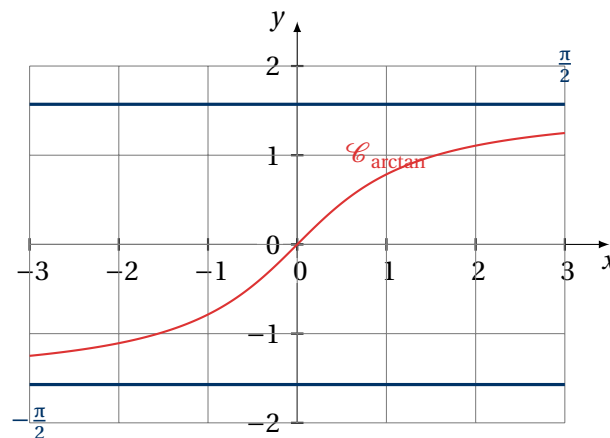




FIG. ANA.7.3. : Graphe de l'arctangente

- 2.

⁵l'hypothèse est très simple à retenir, puisque c'est l'hypothèse garantissant la non-nullité du dénominateur dans cette formule

3. Calculer \arctan' après avoir justifié l'existence sur un ensemble à préciser. 

4. Montrer la formule : $\forall x \in \mathbf{R}^*, \arctan(x) + \arctan\left(\frac{1}{x}\right) = \operatorname{sgn}(x) \frac{\pi}{2}$. 

Cette construction d'un inverse peut être aussi faite sur les fonctions \cos , \sin en restreignant correctement l'ensemble de définition, les fonctions inverses sont alors notées \arccos , \arcsin mais ne sont pas au programme de BCPST. Nous les étudierons donc uniquement en TD.

3.2. Dérivées d'ordre supérieur

Définition ANA.7.20 | Classe \mathcal{D}^n

Soit $f : I \rightarrow \mathbf{R}$ une application. On définit, sous réserve d'existence, les *dérivées successives* de f en posant $f^{(0)} = f$ et pour $n \in \mathbf{N}$:

$$f^{(n+1)} = (f^{(n)})'.$$

Définition ANA.7.21 | Classe \mathcal{C}^n

Soit $n \in \mathbf{N}$. Une fonction $f : I \rightarrow \mathbf{R}$ est dite *de classe \mathcal{C}^n* sur I si :

- ▶ f est n fois dérivable sur I ,
- ▶ et si $f^{(n)}$ est continue sur I .

La fonction est dite *de classe \mathcal{C}^∞* si elle est de classe \mathcal{C}^n pour tout $n \in \mathbf{N}$.

Notation

- ▶ L'application $f^{(n)}$ est aussi notée ou $\frac{d^n f}{dx^n}$. On note encore f'' , f''' , etc, pour $f^{(2)}$, $f^{(3)}$, etc.
- ▶ On note $\mathcal{C}^n(I, \mathbf{R})$ (resp. $\mathcal{C}^\infty(I, \mathbf{R})$) l'ensemble des fonctions de classe \mathcal{C}^n (resp. \mathcal{C}^∞) sur I et à valeurs réelles.

Exemple 15 — La fonction \sin est \mathcal{C}^∞ et pour tout $n \in \mathbf{N}^*$ on a :

$$\forall x \in \mathbf{R}, \quad \sin^{(n)}(x) = \sin\left(x + n\frac{\pi}{2}\right).$$



Proposition ANA.7.11 | Opérations sur les fonctions de classe $\mathcal{C}^n/\mathcal{D}^n$

Soient $n \in \mathbf{N} \cup \{\infty\}$, $(f, g) \in \mathcal{C}^n(I, \mathbf{R})^2$ (resp. \mathcal{D}^n) et $(\lambda, \mu) \in \mathbf{R}^2$. Alors :

- ▶ $\lambda f + \mu g \in \mathcal{C}^n(I, \mathbf{R})$ (resp. \mathcal{D}^n),
- ▶ $f g \in \mathcal{C}^n(I, \mathbf{R})$ (resp. \mathcal{D}^n),
- ▶ et si g ne s'annule pas, alors : $\frac{f}{g} \in \mathcal{C}^n(I, \mathbf{R})$ (resp. \mathcal{D}^n).

Corollaire ANA.7.3

Soit $n \geq 0$. Les triplets

$$\left(\mathcal{D}^n(I, \mathbf{R}), +, \times\right)_{\mathbf{R} \times \mathcal{D}^n(I, \mathbf{R})}, \quad \left(\mathcal{C}^n(I, \mathbf{R}), +, \times\right)_{\mathbf{R} \times \mathcal{C}^n(I, \mathbf{R})}$$

sont des espaces vectoriels.

Théorème ANA.7.16 | Composée de fonctions de classe $\mathcal{C}^n/\mathcal{D}^n$

Soient $n \in \mathbf{N} \cup \{\infty\}$, $f \in \mathcal{C}^n(I, J)$ (resp. $\mathcal{D}^n(J, \mathbf{R})$) et $g \in \mathcal{C}^n(J, \mathbf{R})$ (resp. $\mathcal{D}^n(J, \mathbf{R})$). Alors : $g \circ f \in \mathcal{C}^n(I, \mathbf{R})$ (resp. $\mathcal{D}^n(I, \mathbf{R})$).

Théorème ANA.7.17 | Réciproque d'une fonction de classe $\mathcal{C}^n/\mathcal{D}^n$

Soient $n \in \mathbf{N}^* \cup \{\infty\}$ et $f \in \mathcal{C}^n(I, \mathbf{R})$ (resp. $\mathcal{D}^n(I, \mathbf{R})$). On suppose que f' ne s'annule pas sur I . La fonction f réalise une bijection de I sur $J = f(I)$, et :

$$f^{-1} \in \mathcal{C}^n(J, I), \quad (\text{resp. } \mathcal{D}^n(I, \mathbf{R})).$$

Résumé | Lien entre les notions de régularités pour les fonctions d'une variable



3.3. Extrema et théorème de ROLLE

Définition ANA.7.22 | Extrema

On dit qu'une fonction $f : I \rightarrow \mathbf{R}$ admet en $x_0 \in I$ un *minimum* (resp. *maximum*) si :

$$\forall x \in I, \quad f(x) \geq f(x_0) \\ (\text{resp. } \forall x \in I, \quad f(x) \leq f(x_0)).$$

On dit que f admet en $x_0 \in I$ un *minimum local* (resp. *maximum local*) s'il existe

$\alpha > 0$ tel que :

$$\forall x \in I \cap]x_0 - \alpha, x_0 + \alpha[, \quad f(x) \geq f(x_0)$$

(resp. $\forall x \in I \cap]x_0 - \alpha, x_0 + \alpha[, \quad f(x) \leq f(x_0)$).

On dit que f admet en x_0 un *extremum* (resp. *extremum local*) si f admet en x_0 un minimum ou un maximum (resp. un minimum local ou un maximum local).

Notation

Lorsque f admet un unique minimum (resp. maximum) global x_0 sur I , on note généralement

$$x_0 = \operatorname{argmin}_I f, \quad \text{resp.} \quad x_0 = \operatorname{argmax}_I f.$$

Définition ANA.7.23 | Point critique

On appelle *point critique* d'une fonction dérivable $f : I \rightarrow \mathbf{R}$ tout réel $x \in I$ tel que $f'(x) = 0$.

Théorème ANA.7.18 | Condition nécessaire pour qu'un point intérieur soit un extremum local

Soient $f : I \rightarrow \mathbf{R}$ une application et $c \in I$. On suppose que :

1. c est à l'intérieur de I , i.e. il existe $\varepsilon > 0$ vérifiant : $]c - \varepsilon, c + \varepsilon[\subset I$,
2. f admet en c un extremum local,
3. f est dérivable en c .

Alors

c est un point critique de f .

Preuve Supposons par exemple que f possède un maximum local en c , alors quitte à diminuer ε , nous pouvons supposer qu'il existe $\varepsilon > 0$ tel que : $]c - \varepsilon, c + \varepsilon[\subset I$ et :

$$\forall x \in]c - \varepsilon, c + \varepsilon[, \quad f(x) \leq f(c).$$

Ainsi,

$$\forall x \in]c - \varepsilon, c[, \quad \frac{f(x) - f(c)}{x - c} \leq 0, \quad \forall x \in]c, c + \varepsilon[, \quad \frac{f(x) - f(c)}{x - c} \geq 0.$$

Comme f est dérivable, les limites à gauche et droite en c du taux d'accroissement sont les mêmes, égales au nombre dérivé. Ici, en passant à la limite dans les inégalités précédentes il devrait donc être positif et négatif à la fois, donc nul.

Attention La condition « c est à l'intérieur de I » est indispensable

Si $f : x \in [2, 5] \rightarrow x^2$, elle admet un minimum en 2, $f(2) = 4$, un maximum en 5, $f(5) = 25$ et pourtant 2, 5 ne sont pas des points critiques.

Théorème ANA.7.19 | Théorème de ROLLE

Soit $f : [a, b] \rightarrow \mathbf{R}$ (où $a < b$) une fonction telle que :

- ▶ f est continue sur $[a, b]$,
- ▶ f est dérivable sur $]a, b[$,
- ▶ $f(a) = f(b)$.

Alors : il existe $c \in]a, b[$ tel que $f'(c) = 0$.

L'interprétation du théorème est la suivante : à un moment, la courbe possède une tangente horizontale.

Attention

Il n'y a pas unicité d'un tel c . Prendre par exemple la fonction \sin sur l'intervalle $[0, 2\pi]$.

Remarque 3.1 — Dans le théorème de ROLLE,

- ▶ f n'a pas besoin d'être dérivable aux bornes de l'intervalle (considérer $x \rightarrow \sqrt{1 - x^2}$ sur $[-1, 1]$).
- ▶ La continuité aux bornes est nécessaire (considérer $x \rightarrow x - \lfloor x \rfloor$ sur $[0, 1]$).
- ▶ La dérivabilité est nécessaire : considérer $x \rightarrow |x|$ sur $[-1, 1]$.

Preuve Si f est constante sur $[a, b]$, alors le résultat est évident. Sinon, l'application f étant continue sur le segment $[a, b]$, elle admet un minimum et un maximum distincts sur cet intervalle. Au moins un de ces *extrema* est différent de $f(a) = f(b)$ et correspond donc à un point intérieur à $[a, b]$, c'est-à-dire à un point de $]a, b[$. Le **Théorème ANA.7.18** conclut.

Ce théorème peut être très utile pour prouver l'existence de zéros d'une fonction.

Exemple 16 — Itération de ROLLE et entrelacement des points d'annulation Soit $f : I \rightarrow \mathbf{R}$ une fonction dérivable.

1. Si f s'annule $n + 1$ fois, alors f' s'annule au moins n fois.
2. En appliquant plusieurs fois ce résultat, déduire que si f est n fois dérivable et s'annule $n + 1$ fois, alors $f^{(n)}$, la dérivée $n^{\text{ème}}$ de f , s'annule au moins une fois.



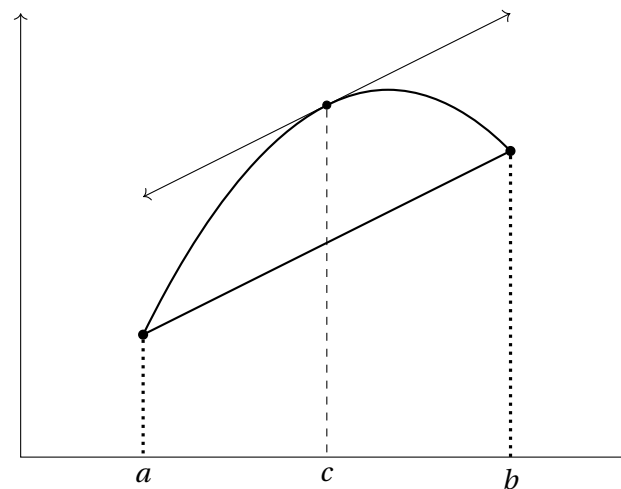
Théorème ANA.7.20 | Égalité des accroissements finis

Soit $f : [a, b] \rightarrow \mathbf{R}$ (où $a < b$) une fonction telle que :

1. f est continue sur $[a, b]$,
2. f est dérivable sur $]a, b[$.

Alors il existe $c \in]a, b[$ tel que : $f'(c) = \frac{f(b) - f(a)}{b - a}$.

L'interprétation du théorème est la suivante : à un moment, la pente de la tangente vaut celle de la sécante entre les points $(a, f(a))$ et $(b, f(b))$.



Preuve



Exemple 17 — Montrons à l'aide de l'égalité des accroissements finis (appliquée à $t \mapsto \ln(1+t)$ sur $[0, x]$ pour tout $x > 0$) que :

$$\forall x > 0, \quad \frac{x}{1+x} < \ln(1+x) < x.$$




4. DÉVELOPPEMENTS LIMITÉS

Il est nécessaire de revoir les relations de comparaison au programme de BCPST (équivalent et « petit o ») **avant** d'aborder ce paragraphe.


Exemple 18 — Manipulations de petit o


▶ $x^2 = o(x^4)$, MAIS : $x^4 \underset{x \rightarrow 0}{=} o(x^2)$, 

▶ $\frac{1}{x^2} \underset{x \rightarrow +\infty}{=} o\left(\frac{1}{x}\right)$, MAIS : $\frac{1}{x} \underset{x \rightarrow 0}{=} o\left(\frac{1}{x^2}\right)$, 

▶ $n^2 \underset{n \rightarrow +\infty}{=} o(n^4)$, $2^n \underset{n \rightarrow \infty}{=} o(3^n)$, $\frac{1}{n^2} \underset{n \rightarrow +\infty}{=} o\left(\frac{1}{n}\right)$. 

Exemple 19 — Manipulations d'équivalents La symbole équivalent est à manipuler avec d'extrêmes précautions : retenez que les seules opérations compatibles sont la multiplication, le quotient et les puissances d'équivalents

▶ $n \underset{n \rightarrow \infty}{\sim} n+1$, 

▶ et pourtant $e^n \underset{n \rightarrow \infty}{\sim} e^{n+1}$. 

▶ En revanche, on peut montrer l'équivalent suivant : $\ln(n+1) \underset{n \rightarrow \infty}{\sim} \ln(n)$. On consultera l'**Exemple 24** ci-après.

Nous rappelons ici uniquement les résultats théoriques classiques qui fondent les développements limités, ainsi que la liste des développements à connaître (ou au moins à savoir retrouver très rapidement).

4.1. Généralités

Définition ANA.724 | Définition d'un développement limité

Soient un intervalle I , un réel $x_0 \in \bar{I} \cap \mathbf{R}$, un entier $n \in \mathbf{N}$ et une application $f : I \rightarrow \mathbf{R}$. On dit que f admet un *développement limité* à l'ordre n au voisinage de x_0 s'il existe des réels a_0, a_1, \dots, a_n tels que :

$$f(x) \underset{x \rightarrow x_0}{=} a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \dots + a_n(x - x_0)^n + o((x - x_0)^n). \quad (\text{DL})$$

Le polynôme

$$R_n(X) \underset{(\text{déf.})}{=} a_0 + a_1(X - x_0) + a_2(X - x_0)^2 + \dots + a_n(X - x_0)^n$$

est appelé *partie régulière* (ou *principale, entière*), du développement limité.

Notation

Nous noterons $DL_n(x_0)$ pour « développement limité à l'ordre n au voisinage de x_0 ».

Remarque 4.1 — Admettre un $DL_n(x_0)$ signifie donc que

$$\lim_{x \rightarrow x_0} \frac{(f - R_n(f))(x)}{(x - x_0)^n} = 0.$$

On rappelle qu'un terme noté $o((x - x_0)^n)$ est un terme qui s'écrit sous la $\varepsilon(x)(x - x_0)^n$ avec $\varepsilon(x) \xrightarrow{x \rightarrow x_0} 0$, et c'est une fonction définie au moins au voisinage de x_0 . De plus, on a la règle de calcul suivante :

$$\frac{o((x - x_0)^n)}{x - x_0} = o((x - x_0)^{n-1}).$$

Ainsi, quand on divise un développement limité, il ne faut pas oublier de diviser les petit o . En cas de doute, mieux vaut toujours se ramener au brouillon à la notation $\varepsilon(x)(x - x_0)^n$ avant de faire des manipulations.

Puisque Eq. (DL) n'est finalement qu'une limite déguisée, et que l'on peut faire des changements de variable dans les limites, nous pouvons donc aussi en faire dans les développements limités.



Méthode Translation x_0 dans un développement limité

Si $x_0 \neq 0$, alors la recherche d'un $DL_n(x_0)$ pour une fonction f se fera en se ramenant au voisinage de 0 par le changement de variable « $h = x - x_0$ ». Plus précisément,

1. considérer $g : h \mapsto f(x_0 + h)$,
2. faire un $DL_n(0)$ de g : on obtient une expression du type $g(h) \underset{h \rightarrow 0}{=} R_n(h) + o(h^n)$ (avec R_n fonction polynomiale de degré n définie au voisinage de zéro),
3. un $DL_n(x_0)$ de f est alors : $f(x) \underset{x \rightarrow x_0}{=} R_n(x - x_0) + o((x - x_0)^n)$.

PROPRIÉTÉS DES DÉVELOPPEMENTS LIMITÉS. Les développements limités satisfont des propriétés très similaires aux polynômes, la plus notable est l'identification coefficient par coefficient appelée plus sobrement *unicité du développement limité* dans la suite.

Proposition ANA.7.12 | Premières propriétés

Soient $n \in \mathbf{N}$ et f une fonction admettant un $DL_n(x_0)$. Alors :

- ▶ la partie régulière du $DL_n(x_0)$ est unique,
- ▶ **(Troncation)** f admet un $DL_p(x_0)$ pour tout $p \leq n$.
- ▶ Si $x_0 = 0$ et f est paire (*resp.* impaire), alors la partie principale du $DL_n(0)$ est un polynôme pair (*resp.* impair).
- ▶ Soient g une fonction admettant un $DL_n(x_0)$ et $(\lambda, \mu) \in \mathbf{R}^2$. Alors $\lambda f + \mu g$ admet un $DL_n(x_0)$ et la partie régulière est obtenue par combinaison linéaire des parties régulières des $DL_n(x_0)$ de f et g .

Preuve

Montrons la propriété de troncation.

Proposition ANA.7.13 | Continuité, dérivabilité et développement limité

Soient $f : I \rightarrow \mathbf{R}$ une fonction et $x_0 \in I$. Alors :

- ▶ f est continue en x_0 si et seulement si elle admet un $DL_0(x_0) : f(x) \underset{x \rightarrow x_0}{=} a + o(1)$. On a alors $a = f(x_0)$.
- ▶ f est dérivable en x_0 si et seulement si elle admet un $DL_1(x_0) : f(x) \underset{x \rightarrow x_0}{=} a + b(x - x_0) + o(x - x_0)$. On a alors $a = f(x_0)$ et $b = f'(x_0)$.

Attention

Cet énoncé ne se généralise **pas** pour des valeurs supérieures de n .

Exemple 20 — Contre-exemple Par exemple, pour $n \geq 2$, l'application

$$f : \begin{cases} \mathbf{R} \rightarrow \mathbf{R} \\ x \mapsto \begin{cases} x^{n+1} \sin\left(\frac{1}{x^n}\right) & \text{si } x \neq 0, \\ 0 & \text{si } x = 0. \end{cases} \end{cases}$$

admet un $DL_n(0)$.

Pourtant, f' n'est pas continue en 0, donc f n'est même pas deux fois dérivable.



de ce point.

Exemple 21 — Étudier l'existence d'un prolongement continu en zéro pour $f : x \mapsto \frac{\sin x}{x}$, étudier la dérivabilité du prolongement.

Proposition ANA.7.14 | Développement limité et prolongement

Soient $x_0 \in I$ et $f : I \setminus \{x_0\} \rightarrow \mathbf{R}$. Si f admet un $DL_1(x_0)^6$,

$$f(x) \underset{x \rightarrow x_0}{=} a + b(x - x_0) + o(x - x_0),$$

alors la fonction f se prolonge par continuité en x_0 , en posant $f(x_0) = a$. De plus, la fonction ainsi prolongée est dérivable en x_0 de dérivée égale à b .

Preuve



Méthode Prolongement



Pour montrer qu'une fonction est prolongeable par continuité et/ou dérivable en un point x_0 , il suffit de faire un développement limité d'ordre 0/1 au voisinage

⁶il n'y a rien d'anormal au fait qu'une fonction *a priori* non définie en un point admette un développement limité en ce point. Car, rappelons-le, un développement limité est un résultat sur une limite, la variable **tend vers** x_0 mais en restant dans l'ensemble de définition.

Proposition ANA.7.15 | Développement limité et équivalent

Soit f une fonction admettant un développement limité écrit sous la forme suivante :

$$f(x) \underset{x \rightarrow x_0}{=} (x - x_0)^p (a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \dots + a_n(x - x_0)^n + o((x - x_0)^n)),$$

avec $a_0 \neq 0$, $a_1, \dots, a_n \in \mathbf{R}$. Alors : $f(x) \underset{x \rightarrow x_0}{\sim} a_0(x - x_0)^p$.

Preuve En mettant en facteur le premier terme, on obtient puisque $a_0 \neq 0$:

$$f(x) \underset{x \rightarrow x_0}{=} a_0(x - x_0)^p \times \left(1 + \frac{a_1}{a_0}(x - x_0) + \frac{a_2}{a_0}(x - x_0)^2 + \dots + \frac{a_n}{a_0}(x - x_0)^n + o((x - x_0)^n) \right).$$

Il reste à faire tendre x vers x_0 et à constater que la parenthèse converge vers 1. De plus, une autre écriture donne, en supposant $a_1 \neq 0$:

$$f(x) - a_0(x - x_0)^p \underset{x \rightarrow x_0}{=} a_1(x - x_0)^{p+1} \left(1 + \frac{a_2}{a_1}(x - x_0) + \frac{a_3}{a_1}(x - x_0)^2 + \dots + \frac{a_n}{a_1}(x - x_0)^{n-1} + o((x - x_0)^{n-1}) \right).$$

La parenthèse est positive pour x assez proche de x_0 (puisqu'elle tend vers 1), donc le signe de $x \mapsto f(x) - a_0(x - x_0)^p$ au voisinage de zéro est donné par celui de $x \mapsto a_1(x - x_1)^{p+1}$. Si $a_1 = 0$ on factorise simplement par le premier coefficient non nul.



Méthode Recherche d'équivalent/Signe local d'une fonction

On lit :

1. un équivalent de fonction au voisinage de x_0 en regardant le **premier terme non nul** du $DL_n(x_0)$ pour un certain entier n assez grand (de sorte que le développement limité ne soit pas nul).
2. le signe de f au voisinage de x_0 grâce à celui de $a_0(x - x_0)^p$. Cela dépend donc de la parité de p notamment.

Nous allons voir à présent comment obtenir concrètement l'expression d'un développement limité. La formule de TAYLOR-YOUNG ci-dessous nous propose une hypothèse (caractère \mathcal{C}^n fois dérivable) pour avoir l'existence d'un développement limité à l'ordre n , en plus les coefficients du développement sont donnés par les dérivées successives de la fonction. Bien entendu, nous n'allons pas nous amuser à calculer ces dérivées successives à chaque fois. Nous allons plutôt calculer les développements des fonctions usuelles, puis effectuer des opérations (somme, produit, primitivation, etc...) pour obtenir celui qui nous intéresse.

4.2. Développement géométrique

La formule de sommation de termes géométriques nous permet aisément de déduire un premier développement limité.

Théorème ANA.7.21 | Développement limité géométrique

$$\begin{aligned} \frac{1}{1-x} & \underset{x \rightarrow 0}{=} \sum_{k=0}^n x^k + o(x^n), \\ & \underset{x \rightarrow 0}{=} 1 + x + x^2 + \dots + x^n + o(x^n). \end{aligned}$$

Preuve D'après la formule sur les sommes de termes géométriques, nous avons pour tout $x \neq 1$:

$$\sum_{k=0}^n x^k = \frac{1-x^{n+1}}{1-x} \iff \frac{1}{1-x} = \sum_{k=0}^n x^k + \frac{x^{n+1}}{1-x}.$$

Mais $\frac{x^{n+1}}{1-x} = \frac{x}{1-x} \xrightarrow{x \rightarrow 0} 0$, on déduit alors la formule.

On en déduit aisément, par compositions des limites, les développements suivants :

$$\begin{aligned} \frac{1}{1+x} & \underset{x \rightarrow 0}{=} \sum_{k=0}^n (-x)^k + o(x^n) \\ & \underset{x \rightarrow 0}{=} 1 - x + x^2 - x^3 + \dots + (-x)^n + o(x^n) \\ \frac{1}{1-x^2} & \underset{x \rightarrow 0}{=} \sum_{k=0}^n x^{2k} + o(x^{2n}) \\ & \underset{x \rightarrow 0}{=} 1 + x^2 + x^4 + x^6 + \dots + x^{2n} + o(x^{2n}) \end{aligned}$$

4.3. Développements obtenus par la formule de TAYLOR-YOUNG

Pour obtenir notamment celui de l'exponentielle, nous allons nous servir de la formule de TAYLOR-YOUNG.

Théorème ANA.7.22 | Formule de TAYLOR-YOUNG

Soit $f \in \mathcal{C}^n(I, \mathbf{R})^7$, où I est un intervalle tel que $x_0 \in I$. Alors f admet un $DL_n(x_0)$:

$$\begin{aligned} f(x) & \underset{x \rightarrow x_0}{=} \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + o((x - x_0)^n), \\ & \underset{x \rightarrow x_0}{=} f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!} (x - x_0)^2 + \dots + \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n \\ & \quad + o((x - x_0)^n). \end{aligned}$$

⁷Ce n'est pas la meilleure hypothèse, mais celle du programme. On pourrait la supposer \mathcal{D}^n sur I si $n \neq 0$ et continue sinon.

Attention

Il existe des fonctions admettant un développement limité à l'ordre n mais qui ne sont pas \mathcal{C}^n (on peut regarder à nouveau l'Exemple 20) précédent. La formule de TAYLOR-YOUNG donne une **condition nécessaire** uniquement pour l'existence d'un développement limité.

Preuve Récurrence sur $n \in \mathbf{N}$, en admettant provisoirement le Lemme ANA.7.1 (de primitivation des développements limités) présenté plus bas.

- ▶ Pour $n = 0$: la formule correspond à la définition de la continuité de f en x_0 .
- ▶ Supposons la formule vérifiée à l'ordre n . Soit f une fonction de classe \mathcal{C}^{n+1} . Alors f' est de classe \mathcal{C}^n , et donc on peut lui appliquer la formule de TAYLOR-YOUNG :

$$f'(x) = \sum_{k=0}^n \frac{f^{(k+1)}(x_0)}{k!} (x - x_0)^k + o((x - x_0)^n).$$

Le résultat de primitivation terme à terme (Lemme ANA.7.1) fournit alors :

$$f(x) = f(x_0) + \sum_{k=0}^n \frac{f^{(k+1)}(x_0)}{(k+1)!} (x - x_0)^{k+1} + o((x - x_0)^{n+1}).$$

Il reste alors à effectuer le changement de variable $\ell = k + 1$ dans la somme, cela nous donne la formule de TAYLOR-YOUNG à l'ordre $n + 1$.


Pour tout $n \in \mathbf{N}$, on en déduit les développements limités suivants au voisinage de 0 :

Proposition ANA.7.16 | Développements limités provenant de TAYLOR-YOUNG

$$\begin{aligned} e^x &\underset{x \rightarrow 0}{=} \sum_{k=0}^n \frac{x^k}{k!} + o(x^n) \\ &= 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + o(x^n) \\ \cos x &\underset{x \rightarrow 0}{=} \sum_{k=0}^n (-1)^k \frac{x^{2k}}{(2k)!} + o(x^{2n+1}) \\ &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \dots + (-1)^n \frac{x^{2n}}{(2n)!} + o(x^{2n+1}), \end{aligned}$$

$$\begin{aligned} \sin x &\underset{x \rightarrow 0}{=} \sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{(2k+1)!} + o(x^{2n+2}) \\ &= x - \frac{x^3}{3!} + \frac{x^5}{5!} + \dots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + o(x^{2n+2}), \\ (1+x)^\alpha &\underset{x \rightarrow 0}{=} \sum_{k=0}^n \left(\prod_{0 \leq q < k} (\alpha - q) \right) \frac{x^k}{k!} + o(x^n) \\ &= 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!} x^2 + \dots + \frac{\alpha(\alpha-1) \dots (\alpha-n+1)}{n!} x^n + o(x^n), \end{aligned}$$

pour tout $\alpha \in \mathbf{R}$.

Preuve Constatons, par exemple, le tout dernier. 

Exemple 22 — On en déduit par exemple :

$$\begin{aligned} \sqrt{1+x} &= (1+x)^{1/2} \underset{x \rightarrow 0}{=} 1 + \frac{x}{2} - \frac{x^2}{8} + o(x^2), \\ \frac{1}{\sqrt{1+x}} &= (1+x)^{-1/2} \underset{x \rightarrow 0}{=} 1 - \frac{x}{2} + \frac{3}{8}x^2 + o(x^2). \end{aligned}$$

Exemple 23 — Déterminer un développement limité à l'ordre $x \mapsto \sqrt{\frac{1-x}{1+x}}$ au voisinage de zéro à l'ordre 2. 

Preuve Soit $x \in I \setminus \{x_0\}$. Notons

$$I_x \quad (\text{resp. } \overline{I_x}) = \begin{cases}]x_0, x[& (\text{resp. } [x_0, x]) \quad \text{si } x_0 < x, \\]x, x_0[& (\text{resp. } [x, x_0]) \quad \text{si } x < x_0. \end{cases}$$

La fonction f est continue sur $\overline{I_x}$ et dérivable sur I_x , donc par l'égalité des accroissements finis, on sait qu'il existe $c_x \in I_x$ tel que :

$$\frac{f(x) - f(x_0)}{x - x_0} = f'(c_x).$$

Comme $c_x \in I_x$, alors $0 < |c_x - x_0| \leq |x - x_0|$. On en déduit :

$$\left| \frac{f(x) - f(x_0)}{(x - x_0)^{n+1}} \right| = \frac{1}{|x - x_0|^n} |f'(c_x)| \leq \left| \frac{f'(c_x)}{(c_x - x_0)^n} \right|.$$

Or par hypothèse $\left| \frac{f'(t)}{(t - x_0)^n} \right| \xrightarrow{t \rightarrow x_0} 0$, et $c_x \xrightarrow{x \rightarrow x_0} x_0$ (car $c_x \in I_x$), d'où par composition de limites :

$$\left| \frac{f'(c_x)}{(c_x - x_0)^n} \right| \xrightarrow{x \rightarrow x_0} 0.$$

Par le théorème de convergence par encadrement, il vient donc $\left| \frac{f(x) - f(x_0)}{(x - x_0)^{n+1}} \right| \xrightarrow{x \rightarrow x_0} 0$, d'où le résultat.

4.4. Développements obtenus par primitivation

Lemme ANA.7.1 | Primitivation des $o(\dots)$

Soient I un intervalle, $f : I \rightarrow \mathbf{R}$ une fonction dérivable sur I , $x_0 \in I$ et $n \in \mathbf{N}$. Si

$$f'(x) \underset{x \rightarrow x_0}{=} o((x - x_0)^n),$$

alors

$$f(x) \underset{x \rightarrow x_0}{=} f(x_0) + o((x - x_0)^{n+1}),$$

ou de manière équivalente⁸

$$\int_{x_0}^x f'(t) dt \underset{x \rightarrow x_0}{=} o\left(\int_{x_0}^x (t - x_0)^n dt\right).$$

⁸Étant donné que la constante $\frac{1}{n+1}$ qui apparaît en calculant l'intégrale de droite peut être omise dans le petit o .

Théorème ANA.7.23 | Primitivation de développement limité

Soient I un intervalle contenant x_0 et $f : I \rightarrow \mathbf{R}$ une fonction dérivable. On suppose que f' admet un $DL_n(x_0)$ de la forme :

$$\begin{aligned} f'(x) &\underset{x \rightarrow x_0}{=} \sum_{k=0}^n a_k (x - x_0)^k + o((x - x_0)^n) \\ &\underset{x \rightarrow x_0}{=} a_0 + a_1(x - x_0) + \dots + a_n(x - x_0)^n + o((x - x_0)^n). \end{aligned}$$

Alors f admet un $DL_{n+1}(x_0)$ dont la partie régulière est obtenue en primitivant

celle de f' et en ajoutant $f(x_0)$:

$$\begin{aligned} f(x) &\underset{x \rightarrow x_0}{=} f(x_0) + \sum_{k=0}^n \frac{a_k}{k+1} (x-x_0)^{k+1} + o((x-x_0)^{n+1}) \\ &\underset{x \rightarrow x_0}{=} f(x_0) + a_0(x-x_0) + \frac{a_1}{2}(x-x_0)^2 + \cdots + \frac{a_n}{n+1}(x-x_0)^{n+1} \\ &\quad + o((x-x_0)^{n+1}). \end{aligned}$$

Preuve Il suffit d'appliquer le lemme précédent à la fonction g définie par :


$$g(x) = f(x) - \sum_{k=0}^n \frac{a_k}{k+1} (x-x_0)^{k+1} \text{ définie pour } x \text{ assez proche de } x_0.$$

Le théorème précédent permet d'obtenir de nouveaux développements limités. Pour tout $n \in \mathbf{N}$ (le cas échéant) :

Proposition ANA.7.17 | Développements limités obtenus par primitivation

$$\begin{aligned} \ln(1+x) &\underset{x \rightarrow 0}{=} \sum_{k=1}^n (-1)^{k+1} \frac{x^k}{k} + o(x^n), \\ &\underset{x \rightarrow 0}{=} x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots + (-1)^{n+1} \frac{x^n}{n} + o(x^n). \\ \arctan x &\underset{x \rightarrow 0}{=} \sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{2k+1} + o(x^{2n+2}), \\ &\underset{x \rightarrow 0}{=} x - \frac{x^3}{3} + \frac{x^5}{5} + \cdots + (-1)^n \frac{x^{2n+1}}{2n+1} + o(x^{2n+2}). \end{aligned}$$

Preuve Il suffit de primitiver les développements limités de $x \mapsto \frac{1}{1+x}$ et $x \mapsto \frac{1}{1+x^2}$ au voisinage de zéro.

Exemple 24 — Montrer que $\ln(n+1) \underset{n \rightarrow \infty}{\sim} \ln(n)$. 

Exemple 25 — En utilisant les développements limités précédents, faire l'étude locale de $f(x) = \frac{\ln(1-2x)}{1+x}$ au voisinage de 0. 

⁹Il y a donc une constante qui apparaît que l'on pourrait appeler « constante de primitivation »

4.5. Développements obtenus par produits

Théorème ANA.7.24 | Produit de développements limités.


Si deux fonction f et g admettent des $DL_n(0)$ de parties régulières F et G respectivement, alors fg admet un $DL_n(0)$ de partie régulière obtenue en ne gardant que les termes de degré inférieur ou égal à n dans le produit FG .

Preuve Multiplier et tronquer le développement limité obtenu.

Exemple 26 — Déterminons le $DL_2(0)$ de $\sqrt{1+x} \cdot \cos x$. 


Remarque 4.2 — Ordre d'un produit de développements limités

- ▶ Les termes d'ordre strictement supérieur à n n'ont pas de signification particulière et sont « absorbés » par le $o(x^n)$. On ne doit surtout pas les garder... dans le cas contraire, l'identité fournie ne sera pas un développement limité.
- ▶ Il est important d'observer que si l'on effectue le produit de deux développements limités dont les premiers coefficients sont nuls, alors on gagne des ordres dans le développement limité produit! Ceci sert pour prévoir les ordres nécessaires dans chacun des facteurs lorsque l'on veut un certain ordre pour le produit.

Exemple 27 — Donner le développement limité d'ordre maximal¹⁰ de $(x^2 - 2x^3 + o(x^3))(5x^3 - x^4 + o(x^4))$ au voisinage de zéro. 

Exemple 28 — Déterminer un développement limité des expressions suivantes :

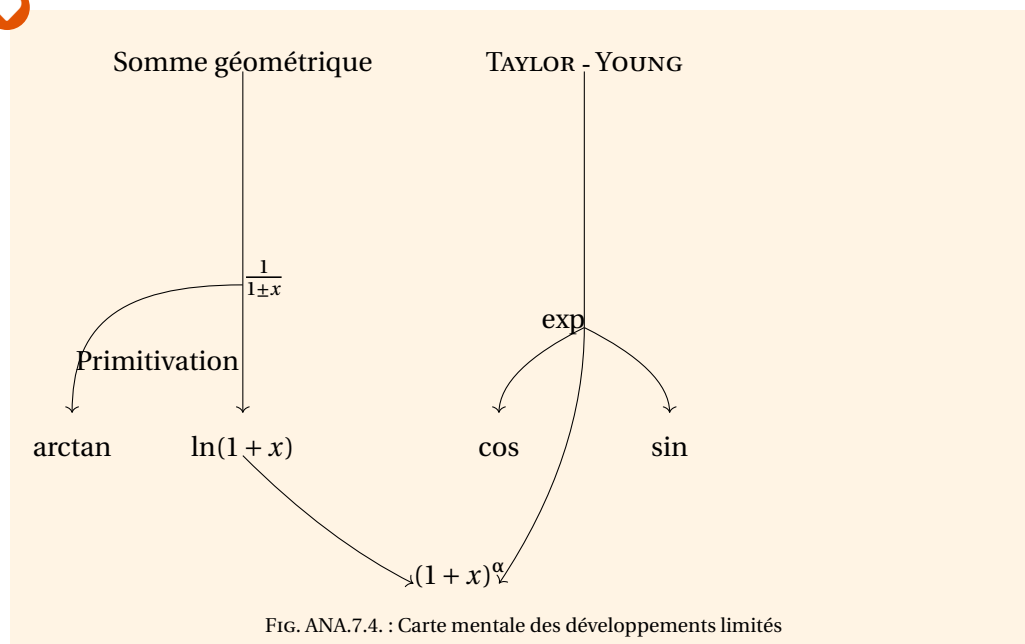
¹⁰Compte-tenu des informations que l'on vous donne.

1. $x \mapsto (\cos x - 1)(\sin x - x)$ à l'ordre sept en zéro.  *Le premier terme de $\cos - 1$ est d'ordre deux, le premier de $x \mapsto \sin(x) - x$ est d'ordre trois. Donc il va falloir pousser le développement limité de $\cos - 1$ à l'ordre 4 (solution de $3 + p = 7$ avec $p \in \mathbf{N}$) et il va falloir pousser le développement limité de $x \mapsto \sin(x) - x$ à l'ordre 3 (solution de $4 + p = 7$ avec $p \in \mathbf{N}$. Ceci étant dit, passons aux calculs.*

$$\begin{aligned} (\cos x - 1)(\sin x - x) &= \left(-\frac{x^2}{2} + \frac{x^4}{4!} + o(x^4) \right) \left(-\frac{x^3}{3!} + \frac{x^5}{5!} + o(x^5) \right) \\ &= \frac{x^5}{6} - x^7 \left(\frac{1}{2 \cdot 5!} + \frac{1}{4! \cdot 3!} \right) + o(x^7). \end{aligned}$$

2. \sin^6 à l'ordre 8, 

Résumé



4.6. Composition de développements limités

Par exemple, supposons connu un $DL_n(0)$ d'une fonction f est connu


$$f(x) \underset{x \rightarrow 0}{=} a_0 + a_1x + a_2x^2 + \dots + a_nx^n + o(x^n).$$


Peut-on écrire

$$f(-x) \underset{x \rightarrow 0}{=} a_0 - a_1x + a_2x^2 + \dots + (-1)^n a_nx^n + o(x^n) \quad ?$$

La réponse est **oui** — par composition de limites (puisque $o(x^n) = o((-x)^n)$). Nous avons déjà utilisé cet argument. Voyons quelques exemples.

Exemple 29 — On note $f : x \mapsto \ln(\cos x)$ pour x dans son domaine de définition.

1. Déterminer le domaine de définition de f et justifier qu'elle admet un développement limité à l'ordre quatre au voisinage de zéro. 

2. Déterminer ce développement limité. 

*** **Fin du chapitre** ***

5. EXERCICES

Exercice ANA.7.1 | Vrai ou Faux?

1. L'inverse d'une fonction bornée strictement positive est borné.
2. Une fonction quotient de fonctions décroissantes est croissante.
3. Une fonction continue bornée atteint ses bornes.
4. La fonction $x \mapsto x|x|$ est de classe \mathcal{C}^1 .
5. Si $f(x) \sim h(x)$, alors $xf(x) \sim 0(x + \sqrt{x})h(x)$.
6. $\left(1 - \frac{1}{n}\right)^{-n} \underset{n \rightarrow \infty}{\sim} 1$.

5.1. Continuité

Exercice ANA.7.2 | Étudier la définition et la continuité éventuelle de $f : x \mapsto (-1)^{\lfloor x \rfloor} \left(x - \lfloor x \rfloor - \frac{1}{2}\right)$.

Solution (exercice ANA.7.2)

La fonction est continue sur $\mathbf{R} \setminus \mathbf{Z}$.
 Soit $k \in 2\mathbf{Z}$ un entier pair. Alors $\lim_{x \rightarrow k, >} f(x) = \lim_{x \rightarrow k, >} \left(x - \lfloor x \rfloor - \frac{1}{2}\right) = -\frac{1}{2}$, même chose à gauche :
 $\lim_{x \rightarrow k, <} f(x) = -\lim_{x \rightarrow k, <} \left(x - (k-1) - \frac{1}{2}\right) = -\frac{1}{2}$, bref f est continue également aux entiers pairs.
 Soit $k \in 2\mathbf{Z} + 1$, $\lim_{x \rightarrow k, >} f(x) = -\lim_{x \rightarrow k, >} \left(x - \lfloor x \rfloor - \frac{1}{2}\right) = \frac{1}{2}$ et $\lim_{x \rightarrow k, <} f(x) = \lim_{x \rightarrow k, <} \left(x - (k-1) - \frac{1}{2}\right) = \frac{1}{2}$, elle l'est aussi aux entiers impairs.
 Donc : f est continue sur \mathbf{R} tout entier.

Exercice ANA.7.3 | Soient f et g définies sur \mathbf{R} par :

$$(x) = \lfloor x \rfloor \sin(\pi x) \quad g(x) = \lfloor x \rfloor \sin(x).$$

Étudier leur continuité sur \mathbf{R} .

Solution (exercice ANA.7.3)

Les fonctions f et g sont trivialement continues sur $\mathbf{R} \setminus \mathbf{Z}$.
 Étudions la continuité aux entiers relatifs $k \in \mathbf{Z}$. On a : $\lim_{x \rightarrow k^+} f(x) = \lim_{x \rightarrow k^+} (k \times \sin(k\pi)) = k \times 0 = 0$. De même, on a : $\lim_{x \rightarrow k^-} f(x) = \lim_{x \rightarrow k^-} ((k-1) \times \sin(k\pi)) = (k-1) \times 0 = 0$.
 En revanche, pour g , nous obtiendrons de la même façon :
 $\lim_{x \rightarrow k^+} g(x) = \lim_{x \rightarrow k^+} (k \times \sin(k)) = k \sin k$ et $\lim_{x \rightarrow k^-} g(x) = \lim_{x \rightarrow k^-} ((k-1) \times \sin(k)) = (k-1) \sin k \neq k \sin k$ dès que $\sin k \neq 0$. En conclusion : f est continue, g ne l'est pas.

Exercice ANA.7.4 | Soit $f : [0, 1] \rightarrow \mathbf{R}$ continue, telle que $f(0) = f(1)$.
 Montrer qu'il existe x_1 et x_2 dans $]0, 1[$ tels que $f(x_1) = f(x_2)$ et $x_1 - x_2 = \frac{1}{2}$.

Solution (exercice ANA.7.4)

Il s'agit de montrer qu'il existe $x_2 \in [0, 1]$ tel que

$$f\left(x_2 + \frac{1}{2}\right) = f(x_2).$$

Pour cela, introduisons la fonction $g \left| \begin{array}{l} [0, 1/2] \rightarrow \mathbf{R} \\ x \mapsto f(x + 1/2) - f(x) \end{array} \right.$. Elle est continue en tant que différence de telles fonctions, et de plus $g(1/2) = f(1) - f(1/2)$, $g(0) = f(1/2) - f(0) = -g(1)$. Donc g change de signe sur $[0, 1/2]$ et d'après le théorème des valeurs intermédiaires il existe $x_2 \in [0, 1/2]$ tel que $g(1/2) = 0$, i.e. $f\left(x_2 + \frac{1}{2}\right) = f(x_2)$. En posant $x_1 = x_2 + \frac{1}{2}$, on a donc montré :

$$\exists x_1, x_2 \in [0, 1], \quad f(x_1) = f(x_2), \quad x_1 - x_2 = \frac{1}{2}.$$

Exercice ANA.7.5 | Agro—Véto Soit $t \geq 0$, pour tout x , on pose $P_t(x) = x^3 + tx - 1$.

1. Montrer que P_t admet une unique racine réelle $u(t)$.
2. **2.1)** Montrer que $u(\mathbf{R}^+) \subset]0, 1[$.
- 2.2) Démontrer que u est strictement décroissante sur \mathbf{R}^+ .

2.3) Calculer $\lim_{t \rightarrow \infty} u(t)$. *Indication : Utiliser l'expression de $P_t(u(t))$.*

2.4) Montrer que u est bijective de \mathbf{R}^+ vers $]0, 1]$, de réciproque :

$$v : \begin{cases}]0; 1] & \rightarrow \mathbf{R}^+, \\ y & \mapsto \frac{1-y^3}{y}. \end{cases}$$

2.5)  Représenter graphiquement cette fonction à l'aide de Python sur $]0, 1]$. En déduire le tracé de la représentation graphique de u .

2.6) Justifier que la fonction u est continue sur \mathbf{R}^+ .

2.7) Démontrer que la fonction u est dérivable sur \mathbf{R}^+ , puis déterminer, pour tout $t \geq 0$, une expression de $u'(t)$ en fonction de t et $u(t)$.

Solution (exercice ANA.7.5)

1. La fonction P_t est strictement croissante puisque pour tout $x \in \mathbf{R}$, $P_t'(x) = 3x^2 + t \geq 0$ (et est nulle uniquement lorsque $x = t = 0$). Elle est de plus continue, car polynomiale. Donc d'après le théorème de la bijection, elle réalise une bijection de \mathbf{R} sur $[-\infty, \infty] \ni 0$. Donc P_t admet une unique racine réelle $u(t)$.

2. 2.1) IL s'agit de montrer que $u(t) \in]0, 1]$ pour tout $t \in \mathbf{R}^+$, donc d'affiner le théorème de la bijection précédent. La fonction P_t est toujours continue et $P_t(0) = -1 < 0$, $P_t(1) = t \geq 0$ donc d'après le théorème des valeurs intermédiaires, $u(t) \in]0, 1]$. Donc $u(\mathbf{R}^+) \subset]0, 1]$.

2.2) Soient $t < s$, montrons que $u(t) \leq u(s)$. Pour cela, calculons $P_t(u(s))$ (ou $P_s(u(t))$) et cherchons son signe.

$$\begin{aligned} P_t(u(s)) &= u(s)^3 + t u(s) - 1 \\ &< u(s)^3 + s u(s) - 1 \quad \left. \begin{array}{l} \\ \end{array} \right\} t \leq s \\ &= P_s(u(s)) = P_t(u(t)). \end{aligned}$$

Or P_t est strictement croissante, donc $u(s) < u(t)$. Donc u est strictement décroissante sur \mathbf{R}^+ .

2.3) On sait que $\lim_{t \rightarrow \infty} u(t)$ existe d'après le théorème de convergence monotone puisque u est décroissante minorée par zéro, et notons $\ell = \lim_{t \rightarrow \infty} u(t) \in [0, 1]$. Supposons donc que $\ell \in]0, 1]$. Alors comme

$$t u(t)^3 + t u(t) = 1,$$

on voit que $\lim_{t \rightarrow \infty} (t u(t)^3 + t u(t)) = \infty$ par règles usuelles sur les limites alors que le membre de droite converge vers 1 — contradiction. Donc


$$\lim_{t \rightarrow \infty} u(t) = 0.$$

2.4) On ne peut se contenter du théorème de la bijection, puisqu'on ne sait pas si u est continue. Soit donc $y \in]0, 1]$, résolvons en $t \in \mathbf{R}^+$,

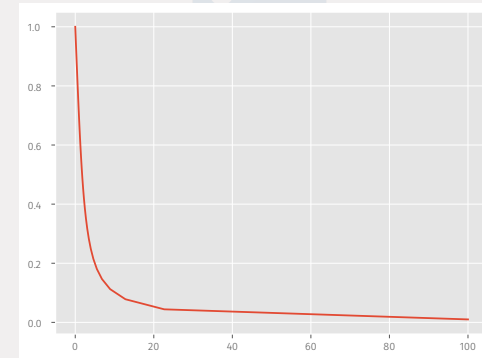
$$u(t) = y \iff P_t(y) = 0 \iff y^3 + t y = 1 \iff t = \frac{1 - y^3}{y}.$$

Donc u est bijective de \mathbf{R}^+ vers $]0, 1]$, de réciproque :

$$v : \begin{cases}]0; 1] & \rightarrow \mathbf{R}^+, \\ y & \mapsto \frac{1-y^3}{y}. \end{cases}$$

2.5)  Représentons graphiquement cette fonction à l'aide de Python sur $]0, 1]$, pour avoir le tracé de u on effectue un symétrie d'axe $y = x$ en inversant les variables dans `plt.plot`.

```
import matplotlib.pyplot as plt
X = np.linspace(0.01, 1, 10*3)
Y = [(1-x**3)/x for x in X]
plt.plot(Y, X)
```



2.6) Puisque v est continue par opérations élémentaires, la fonction u est alors continue sur \mathbf{R}^+ .

2.7) Nous avons

$$u'(t) = (v^{-1})'(t) = \frac{1}{v' \circ u(t)},$$

pour tout t tel que $v' \circ u(t) \neq 0$. Or $v'(y) = \frac{-3y^3 - (1-y^3)}{y^2} = \frac{-(2y^3+1)}{y^2}$ pour tout $y \in]0, 1]$. Donc $v' \circ u(t) \neq 0$ pour tout $t \in \mathbf{R}^+$ puisque $u(t) \geq 0$. Et

$$\forall t \in \mathbf{R}^+, \quad \frac{-u(t)^2}{(2u(t)^3 + 1)}.$$

5.2. Dérivabilité

Exercice ANA.7.6 | Étude complète de fonction Soit g la fonction définie sur $[0, +\infty[$ par :

$$\forall x \in [0, +\infty[, \quad g(x) = e^x - xe^x + 1.$$

- Déterminer la limite de g en $+\infty$.
- Étudier les variations de g et tracer son tableau de variation.
- Montrer que l'équation $g(x) = 0$ admet une unique solution sur \mathbf{R}^+ . On note α cette solution.
- Écrire un programme ValeurApprocheeAlpha qui calcule une valeur approchée de α à $\varepsilon = 10^{-3}$ près.
- Soit A la fonction définie sur $[0, +\infty[$ par : $\forall x \in [0, +\infty[, \quad A(x) = \frac{4x}{e^x+1}$.
 - Étudier la dérivabilité de A , et écrire A' en fonction de g .
 - En déduire les variations de A sur \mathbf{R}^+ .

Solution (exercice ANA.7.6)

Notons $\mathcal{D}_g = [0, \infty[$.

1. Nous avons pour tout $x \in \mathcal{D}_g$.

$$g(x) = e^x (1 - x + e^{-x}),$$

mais $\lim_{x \rightarrow \infty} (1 - x + e^{-x}) = -\infty$. Donc $\lim_{x \rightarrow \infty} g(x) = -\infty$.

2. Pour tout $x \in \mathcal{D}_g$, $g'(x) = e^x - e^x - xe^x = -xe^x \leq 0$. Donc g est décroissante et on a la tableau ci-après.

x	0	α	$+\infty$
$g'(x)$	0	-	
$g(x)$	2	0	$-\infty$

3. On ne peut pas la résoudre, donc appliquons le théorème de la bijection. La fonction g est strictement décroissante, continue et $g(\mathcal{D}_g) =]-\infty, 2] \ni 0$, donc g réalise une bijection de \mathcal{D}_g sur $g(\mathcal{D}_g)$, donc l'équation $g(x) = 0$ admet une unique solution sur \mathbf{R}^+ . On note α cette unique solution.

4. ➤ En tapant dans la console la valeur de g en 2 on trouve que $g(2) \leq 0$.

```
import math as ma

def g(x):
    return ma.exp(x)-x*ma.exp(x)+1

def ValeurApprocheeAlpha(prec):
    """
    Retourne une valeur approchée de alpha à précision prec
    Retourne faux si aucune racine n'existe
    """
    a = 0
    b = 2
    if g(a)*g(b) > 0:
        return False
    while b - a > prec:
        c = (a + b)/2
```

```

if g(a)*g(c) <= 0:
    # changement de signe sur [a,c]
    b = c
else:
    # pas de changement de signe sur [a,c]
    a = c
return (a + b)/2

```

Alors ValeurApprocheeAlpha(10**(-3)) retourne 1.27880859375.

5. Soit A la fonction définie sur $[0, +\infty[$ par : $\forall x \in [0, +\infty[, A(x) = \frac{4x}{e^x+1}$.
 - 5.1) Etudier la dérivabilité de A, et écrire A' en fonction de g.
 - 5.2) En déduire les variations de A sur \mathbf{R}^+ .

Exercice ANA.7.7 | Soit $n \in \mathbf{N}$.

1. Soit $a \in \mathbf{R}$. On note $f_a : x \mapsto \frac{1}{x-a}$. Montrer que f_a est n fois dérivable sur un ensemble à préciser, et calculer sa dérivée n -ième.
2. Calculer, après avoir justifié leur existence, les dérivées d'ordre n de la fonction g définie par $g(x) = \frac{x^2+1}{(x-1)(x-2)(x-3)}$, pour tout x dans un ensemble à préciser.

Solution (exercice ANA.7.7)

1. Soit $a \in \mathbf{R}$,

$$\begin{aligned}
 \forall x \in \mathbf{R} \setminus \{a\}, \quad & \left(\frac{1}{x-a}\right)^{(n)} = ((x-a)^{-1})^{(n)} \\
 & = ((-1)(x-a)^{-2})^{(n-1)} = \dots \\
 & = ((-1)^n n!(x-a)^{-1-n})^{(n-1)}.
 \end{aligned}$$

D'où :

$$\forall x \in \mathbf{R} \setminus \{a\}, \quad \boxed{\left(\frac{1}{x-a}\right)^{(n)} = \frac{(-1)^n n!}{(x-a)^{n+1}}}.$$

2. Nous avons $\mathcal{D}_f = \mathbf{R} \setminus \{1, 2, 3\}$. La fonction g est un quotient de fonctions dérivables n fois, donc g l'est également. Cherchons A, B, C tels que

$$\forall x \in \mathcal{D}_f, \quad g(x) = \frac{A}{x-1} + \frac{B}{x-2} + \frac{C}{x-3}.$$

Pour déterminer les constantes, réduisons au même dénominateur.

$$\begin{aligned}
 \frac{A}{x-1} + \frac{B}{x-2} + \frac{C}{x-3} &= \frac{A(x-2)(x-3) + B(x-1)(x-3) + C(x-1)(x-2)}{(x-1)(x-2)(x-3)} \\
 &= \frac{A(x-2)(x-3) + B(x-1)(x-3) + C(x-1)(x-2)}{(x-1)(x-2)(x-3)} \\
 &= \frac{(A+B+C)x^2 + (-5A-4B-3C)x + (6A+3B+2C)}{(x-1)(x-2)(x-3)}.
 \end{aligned}$$

On identifie coefficient par coefficient le numérateur, on obtient :

$$\begin{cases} A+B+C = 1 \\ -5A-4B-3C = 0 \\ 6A+3B+2C = 1 \end{cases} \iff A=1, B=-5, C=5.$$

Donc

$$\forall x \in \mathcal{D}_f, \quad g(x) = \frac{1}{x-1} - \frac{5}{x-2} + \frac{5}{x-3}.$$

On peut ensuite conclure :

$$\forall x \in \mathcal{D}_f, \quad \boxed{g^{(n)}(x) = \frac{(-1)^n n!}{(x-1)^{n+1}} - \frac{5(-1)^n n!}{(x-2)^{n+1}} + \frac{5(-1)^n n!}{(x-3)^{n+1}}}.$$

Exercice ANA.7.8 | Formule de LEIBNIZ

1. Soient $n \in \mathbf{N}$ et f, g deux fonctions n fois dérivables sur un certain intervalle. Montrer la *formule de LEIBNIZ* :

$$(fg)^{(n)} = \sum_{k=0}^n \binom{n}{k} f^{(k)} g^{(n-k)}.$$

2. (Polynômes de LAGUERRE) Soit $n \in \mathbf{N}$, on note

$$P_n : x \longmapsto \frac{e^x}{n!} (x^n e^{-x})^{(n)}.$$

Montrer que P_n est une fonction polynomiale, préciser son degré et coefficient dominant.

Solution (exercice ANA.7.8)

1. Faisons une récurrence sur n . Pour $n = 0$, nous avons $(fg)^{(0)} = fg$ donc la formule est vérifiée. Supposons-là vraie au rang n et montrons-là au rang $n + 1$. Comme pour la formule du binôme c'est la formule de PASCAL sur les coefficients binomiaux qui sera un intermédiaire technique important.

$$\begin{aligned} & (fg)^{(n+1)} \\ &= (fg)^{(n)'} = \left(\sum_{k=0}^n \binom{n}{k} f^{(k)} g^{(n-k)} \right)' \quad \left. \begin{array}{l} \\ \end{array} \right\} \text{dérivée d'un produit} \\ &= \sum_{k=0}^n \binom{n}{k} f^{(k+1)} g^{(n-k)} + \sum_{k=0}^n \binom{n}{k} f^{(k)} g^{(n+1-k)} \\ &= \sum_{k=1}^{n+1} \binom{n}{k-1} f^{(k)} g^{(n+1-k)} + \sum_{k=0}^n \binom{n}{k} f^{(k)} g^{(n+1-k)} \\ &= \binom{n}{0} f^{(0)} g^{(n+1)} + \sum_{k=1}^n \left(\binom{n}{k-1} + \binom{n}{k} \right) f^{(k)} g^{(n+1-k)} + \binom{n}{n} f^{(n+1)} g^{(0)} \quad \left. \begin{array}{l} \\ \end{array} \right\} \text{formule de PASCAL} \\ &= \binom{n+1}{0} f^{(0)} g^{(n+1)} + \sum_{k=1}^n \binom{n+1}{k} f^{(k)} g^{(n+1-k)} + \binom{n+1}{n+1} f^{(n+1)} Q^{(0)} \\ &= \sum_{k=0}^{n+1} \binom{n+1}{k} f^{(k)} Q^{(n+1-k)}. \end{aligned}$$

D'où la formule par principe de récurrence.

$$(fg)^{(n)} = \sum_{k=0}^n \binom{n}{k} f^{(k)} g^{(n-k)}.$$

2. Nous avons, pour tout $x \in \mathbf{R}$,

$$\begin{aligned} P_n(x) &= \frac{e^x}{n!} (x^n e^{-x})^{(n)} \\ &= \frac{e^x}{n!} \sum_{k=0}^n \binom{n}{k} (x^n)^{(k)} (e^{-x})^{(n-k)} \\ &= \frac{e^x}{n!} \sum_{k=0}^n \binom{n}{k} n(n-1) \dots (n-k+1) x^{n-k} (-1)^{n-k} e^{-x} \\ &= \frac{1}{n!} \sum_{k=0}^n \binom{n}{k} \frac{n!}{(n-k)!} x^{n-k} (-1)^{n-k} \\ &= \sum_{k=0}^n \binom{n}{k} \frac{1}{(n-k)!} x^{n-k} (-1)^{n-k}. \end{aligned}$$

Donc P_n est un polynôme de degré n , et son coefficient dominant vaut :

$$\frac{(-1)^n}{n!}.$$

Exercice ANA.7.9 | En utilisant le théorème des accroissements finis, déterminer

$$\lim_{x \rightarrow \infty} \left((x+1)e^{\frac{1}{x+1}} - xe^{\frac{1}{x}} \right).$$

Solution (exercice ANA.7.9)

Bien entendu, il s'agit d'une forme indéterminée, il nous faut donc procéder autrement.

Notons $f : t > 0 \longmapsto te^{\frac{1}{t}}$. Soit $x > 0$. Alors f est continue sur $[x, x+1]$ et dérivable sur $]x, x+1[$. Donc d'après l'égalité des accroissements finis, il existe $c_x \in]x, x+1[$

$$f(x+1) - f(x) = f'(c_x).$$

Or $f'(t) = e^{\frac{1}{t}} - \frac{1}{t}e^{\frac{1}{t}} = e^{\frac{1}{t}} \left(1 - \frac{1}{t}\right)$ pour tout $t > 0$. Puisque $x < c_x < x + 1$, on a

$$1 - \frac{1}{x} < 1 - \frac{1}{c_x} < 1 - \frac{1}{x+1}$$

donc $\lim_{x \rightarrow \infty} \left(1 - \frac{1}{c_x}\right) = 1$ par théorème d'encadrement. Et donc

$$\lim_{x \rightarrow \infty} f'(c_x) = 1.$$

On déduit alors

$$\lim_{x \rightarrow \infty} \left((x+1)e^{\frac{1}{x+1}} - xe^{\frac{1}{x}} \right) = 1.$$

Exercice ANA.710 | Théorème de la limite de la dérivée Montrer que si f est continue sur $[a, b]$ et dérivable sur $]a, b[$ et que f' admet une limite finie quand $x \rightarrow a$, alors :

$$f \text{ est dérivable en } a, \quad \text{et} \quad f'(a) = \lim_{x \rightarrow a} f'(x).$$

Solution (exercice ANA.710)

Les hypothèses nous invitent clairement à appliquer le théorème des accroissements finis entre x et a pour tout $x \in [a, b]$. Il existe $c_x \in]a, x[$, tel que

$$\frac{f(x) - f(a)}{x - a} = f'(c_x),$$

puisque f est continue sur $[a, b]$ et dérivable sur $]a, b[$. Mais comme f' admet une limite finie quand $x \rightarrow a$ et que $c_x \xrightarrow{x \rightarrow a} a$ (puisque $c_x \in]a, x[$), il vient par composition des limites :

$$\lim_{x \rightarrow a} f'(c_x) = \lim_{x \rightarrow a} f'(x).$$

Donc

$$\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = \lim_{x \rightarrow a} f'(x).$$

Il vient que f est dérivable en a , et $f'(a) = \lim_{x \rightarrow a} f'(x)$.

Exercice ANA.711 | Soit $f \in \mathcal{C}^2(\mathbf{R}^{+*}, \mathbf{R})$. On suppose que $\lim_{x \rightarrow 0} f(x) = 0$ et que la fonction $x \mapsto x^2 f''(x)$ est bornée sur \mathbf{R}^{+*} .

1. (Égalité de TAYLOR-LAGRANGE à l'ordre 2) Montrer que pour tout $x \in \mathbf{R}^{+*}$ et $a \in]0, 1[$, il existe $\xi \in]ax, x[$ tel que :

$$f(ax) = f(x) - (1-a)xf'(x) + \frac{(1-a)^2}{2}x^2f''(\xi).$$

Indication : On pourra considérer la fonction $\varphi : t \mapsto f(t) - f(x) - (x-t)f'(x) + \frac{(x-t)^2}{2}A$ avec $A \in \mathbf{R}$ à choisir correctement.

2. En déduire que $\lim_{x \rightarrow 0} xf'(x) = 0$.

Solution (exercice ANA.711)

1. On cherche à appliquer le théorème de ROLLE. On choisit une constante A telle que

$$f(ax) = f(x) - (1-a)xf'(x) + \frac{(1-a)^2}{2}x^2A,$$

i.e.

$$A = \frac{f(ax) - f(x) + (1-a)xf'(x)}{(1-a)^2x^2}.$$

Avec ce choix, on a $\varphi(x) = 0 = \varphi(ax)$. Et φ est continue sur $[ax, x]$, dérivable sur $]ax, x[$, donc d'après le théorème de ROLLE, il existe $\zeta \in]ax, x[$ tel que

$$\varphi'(\zeta) = 0.$$

Or, $\varphi'(t) = f'(t) - f'(x) + (x-t)A$. Donc φ' s'annule encore en x et ax , et est continue sur $[ax, x]$, dérivable sur $]ax, x[$, donc en réappliquant le théorème de ROLLE, on a l'existence d'un $\xi \in]ax, x[$ tel que

$$\varphi''(\xi) = 0 = f''(\xi) - A.$$

Donc $A = f''(\xi)$, et donc en utilisant l'égalité définissant A, on déduit :

$$A = f''(\xi) = (f(ax) - f(x) + (1-a)xf'(x)) \frac{2}{(1-a)^2 x^2}$$

qui équivaut à

$$f(ax) = f(x) - (1-a)xf'(x) + \frac{(1-a)^2}{2} x^2 f''(\xi).$$

2. Supposons que $\lim_{x \rightarrow 0} f(x) = 0$. En faisant tendre x vers 0 dans la première question, on obtient

$$\lim_{x \rightarrow 0} xf'(x) = \lim_{x \rightarrow 0} \frac{1-a}{2} x^2 f''(\xi) \leq \frac{1-a}{2} \sup_{y \in \mathbb{R}^{++}} |y^2 f''(y)|.$$

En faisant $a \rightarrow 1$, on obtient la nullité de la limite. Donc $\lim_{x \rightarrow 0} f(x) = 0$.

5.3. Fonctions usuelles

Exercice ANA.7.12 | Fonctions trigonométriques inverses

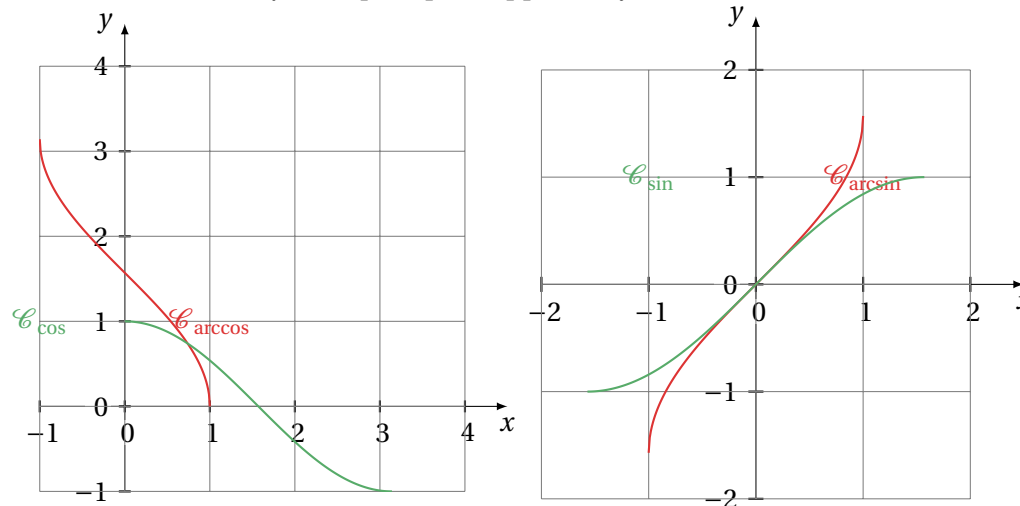
1. Montrer que la fonction $\cos|_{[0,\pi]}$ est bijective. On notera arccos sa bijection réciproque.
2. Calculer arccos' après avoir justifié l'existence sur un ensemble à préciser.
3. Faire la même étude avec $\sin|_{[-\pi/2,\pi/2]}$, on note arcsin la réciproque de la fonction précédente, et calculer arcsin' après avoir justifié l'existence sur un ensemble à préciser.

Solution (exercice ANA.7.12)

Nous traitons les deux fonctions d'un coup. Les fonctions cos, sin sont continues, et

- ▶ $\cos|_{[0,\pi]}$ est strictement décroissante,
- ▶ $\sin|_{[-\pi/2,\pi/2]}$ est strictement croissante.

Donc d'après le théorème de la bijection, elles réalisent une bijection de $[0, \pi]$ pour $\cos|_{[0,\pi]}$, $[-\pi/2, \pi/2]$ pour $\sin|_{[-\pi/2,\pi/2]}$, vers $[-1, 1]$. Nous obtenons les courbes en considérant les symétriques par rapport à $y = x$ des courbes habituelles.



On note alors

$$\arccos = (\cos|_{[0,\pi]})^{-1}, \quad \arcsin = (\sin|_{[-\pi/2,\pi/2]})^{-1}.$$

On s'intéresse à présent aux dérivées. Pour tout $y \in [0, \pi]$ tel que $-\sin(\arccos(y)) \neq 0$, nous avons :

$$\begin{aligned} \arccos'(y) &= \frac{1}{-\sin(\arccos(y))} \\ &= -\frac{1}{|\sin(\arccos(y))|} \quad \left. \vphantom{\arccos'(y)} \right\} \arccos(y) \in [0, \pi] \\ &= -\frac{1}{\sqrt{1 - \cos^2(\arccos(y))}} \\ &= -\frac{1}{1 - y^2}. \end{aligned}$$

Et d'après les calculs précédents $-\sin(\arccos(y)) \neq 0$ si et seulement si $y \neq -1, 1$. En

faisant de-même avec arcsin, on conclut :

$$\forall y \in]-1, 1[, \quad \arccos'(y) = -\frac{1}{1-y^2}, \quad \arcsin'(y) = \frac{1}{1-y^2}.$$

Exercice ANA.713 | Fonctions hyperboliques

1. Faire l'étude complète des fonctions ch et sh définies ci-dessous :

$$\text{ch} : x \in \mathbf{R} \mapsto \frac{e^x + e^{-x}}{2}, \quad \text{sh} : x \in \mathbf{R} \mapsto \frac{e^x - e^{-x}}{2}.$$

2. Montrer que sh est bijective, et déterminer une expression de $\text{sh}^{-1} = \text{argsh}$.

Solution (exercice ANA.713)

Elles sont définies sur \mathbf{R} en tant que somme de telles fonctions. De plus, elles sont dérivables, et pour tout $x \in \mathbf{R}$:

$$\text{sh}' x = \left[\frac{e^x - e^{-x}}{2} \right]' = \frac{e^x - (-e^{-x})}{2} = \frac{e^x + e^{-x}}{2} = \text{ch } x,$$

d'où $\text{sh}' = \text{ch}$, on obtient de-même $\text{ch}' = \text{sh}$.

Mais $\text{ch} \geq 0$ puisque l'exponentielle est positive. Par ailleurs, pour $x \in \mathbf{R}$:

$$\text{sh}(x) \geq 0 \iff e^x \geq e^{-x} \iff e^{2x} \geq 1 \iff x \geq 0.$$

Donc sh est positive sur \mathbf{R}^+ , négative sur \mathbf{R}^- . On déduit alors les variations suivantes en utilisant les dérivées calculées avant :

x	$-\infty$	0	$+\infty$
$\text{ch}'(x) = \text{sh}(x)$	-	0	+
$\text{ch}(x)$	$+\infty$	1	$+\infty$

x	$-\infty$	$+\infty$
$\text{sh}'(x) = \text{ch}(x)$	+	
$\text{sh}(x)$	$-\infty$	$+\infty$

Les limites proviennent de règles usuelles d'opérations sur les limites.

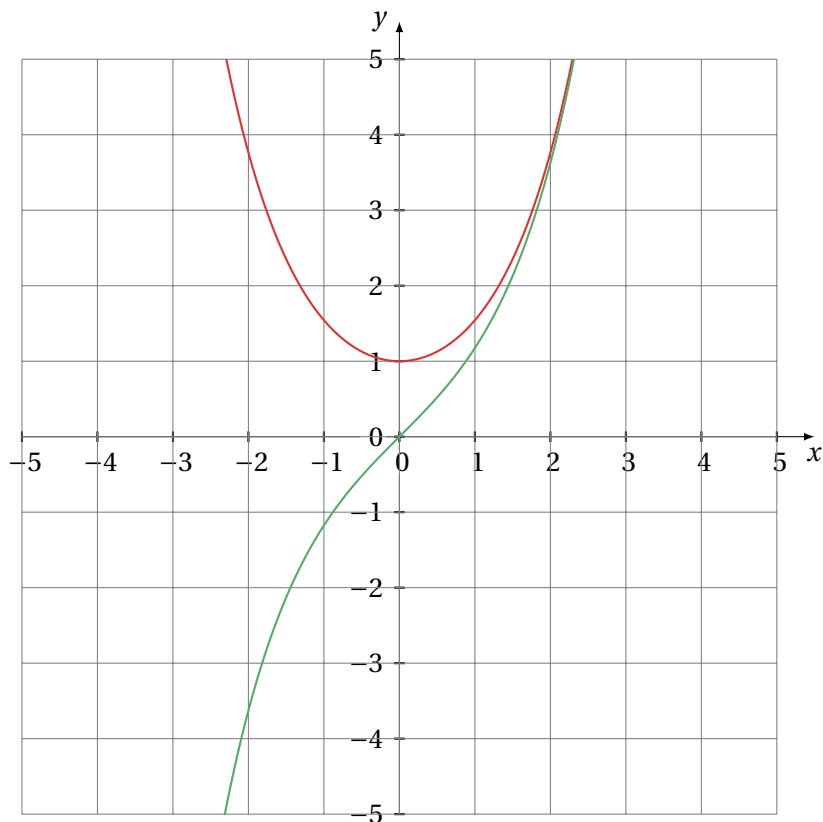


FIG. ANA.7.5. : Graphe de ch, sh

5.4. Développements limités

Exercice ANA.7.14 | Opération nettoyage «Nettoyer» le plus possible les expressions suivantes.

1. $\dots \underset{x \rightarrow 0}{=} \frac{1}{x^2} + o\left(\frac{1}{x^2}\right) + 3 + \frac{1}{x} + o\left(\frac{1}{x}\right),$

2. $\dots \underset{x \rightarrow 0}{=} x + o(x) + x \ln x + o(x^2 \ln x) + x^2 + o(x^2),$

3. $\dots \underset{n \rightarrow \infty}{=} \frac{1}{\sqrt{n}} + \frac{1}{2^n} + o\left(\frac{1}{2^n}\right) + \frac{1}{n} + o\left(\frac{1}{n}\right) + \frac{1}{n^2}.$

Exercice ANA.7.15 | Calculer les développements limités des fonctions suivantes au voisinage de zéro :

1. $x \mapsto \sqrt{1 + \sin x}$ à l'ordre 3,

2. $x \mapsto \frac{\sqrt[3]{\cos x}}{1+x^2} - e^{\cos x}$ à l'ordre 5.

Solution (exercice ANA.7.15)

On a $\sin x \underset{x \rightarrow 0}{=} x - \frac{x^3}{6} + o(x^3)$ et d'autre part $\sqrt{1+u} \underset{u \rightarrow 0}{=} 1 + \frac{u}{2} - \frac{u^2}{8} + \frac{u^3}{16} + o(u^3)$. Comme $\sin x \underset{x \rightarrow 0}{\rightarrow} 0$, on peut composer les développements pour obtenir

$\sqrt{1 + \sin x} \underset{x \rightarrow 0}{=} 1 + \frac{1}{2}\left(x - \frac{x^3}{6}\right) - \frac{1}{8}x^2 + \frac{1}{16}x^3 + o(x^3).$

Le cosinus étant positif au voisinage de zéro, la seconde question a bien un sens et la racine cubique a pour définition associée la forme exponentielle.

Calculons :

$\cos x \underset{x \rightarrow 0}{=} 1 - \frac{x^2}{2} + \frac{x^4}{24} + o(x^5), \quad \frac{1}{1+x^2} \underset{x \rightarrow 0}{=} 1 - x^2 + x^4 + o(x^5),$ Donc après calculs :

$\frac{\sqrt[3]{\cos x}}{1+x^2} - e^{\cos x} \underset{x \rightarrow 0}{=} -\frac{172}{100} + \frac{39}{100}x^2 + \frac{1679}{400}x^4 + o(x^5)$

Exercice ANA.7.16 | Étude d'un prolongement On considère la fonction $f : x \mapsto \frac{\sin x}{x(2-\cos x)}$.

- Déterminer le domaine de définition de f .
- Montrer que f est prolongeable par continuité en 0, on notera encore f ce prolongement.
- Déterminer un développement limité en 0 à l'ordre 3 de f .
- Faire l'étude locale en zéro de la fonction.

Solution (exercice ANA.7.16)

1. Puisque $\cos x \neq 2$ pour tout x , il vient que $\mathcal{D}_f = \mathbf{R}^*$.

2. On peut raisonner ici avec des équivalents :

$$\sin x \underset{x \rightarrow 0}{\sim} x, \quad 2 - \cos x \underset{x \rightarrow 0}{\longrightarrow} 1,$$

donc $f(x) \underset{x \rightarrow 0}{\sim} \frac{x}{2x} = \frac{1}{2}$. Donc f est prolongeable par continuité en zéro, et on pose

$$f(0) = \frac{1}{2}.$$

3. $f(x) \underset{x \rightarrow 0}{\sim} \frac{x - \frac{x^3}{6} + o(x^4)}{x \left(2 - \left(1 - \frac{x^2}{2} + o(x^3) \right) \right)}$

$$\underset{x \rightarrow 0}{\sim} \frac{1 - \frac{x^2}{6} + o(x^3)}{1 + \frac{x^2}{2} + o(x^3)}$$

$$\underset{x \rightarrow 0}{\sim} \left(1 - \frac{x^2}{6} + o(x^3) \right) \left(1 - \frac{x^2}{2} + o(x^3) \right)$$

$$\underset{x \rightarrow 0}{\sim} 1 - \frac{x^2}{2} - \frac{x^2}{6} + o(x^3)$$

$$\underset{x \rightarrow 0}{\sim} \boxed{1 - \frac{2x^2}{3} + o(x^3)}.$$

DL de $\frac{1}{1+X}$ avec $X = \frac{x^2}{2} + o(x^3)$ qui tend vers zéro

Puisque la fonction f est paire, il est parfaitement logique nous n'ayons que des termes pairs.

4. Ainsi, f ainsi prolongée comme en seconde question est dérivable en zéro, et $f'(0) = 0$. Elle admet pour tangente $y = 0$ (horizontale), et la courbe est localement en-dessous de cette tangente.

5.5. Méthode numérique

Exercice ANA.7.17 | Approximation informatique d'un argmin — Modélisation 2021

On considère une fonction $f : I = [a, b] \rightarrow \mathbf{R}$, avec $a < b$ deux réels, telle que : il existe $x_0 \in [a, b]$ vérifiant :

- ▶ f est strictement décroissante sur $[a, x_0]$,
- ▶ f est strictement croissante sur $[x_0, b]$.

Ainsi, f admet un minimum global sur I en x_0 . On souhaite trouver un algorithme convergeant vers $x_0 = \operatorname{argmin}_I f$.

1. Dessiner l'allure de f .
2. L'idée est de construire une suite d'intervalles $[a_k, b_k], k \in \mathbf{N}$ qui vont approcher de plus en plus x_0 , on stoppe l'algorithme une fois une certaine précision $\varepsilon > 0$ atteinte.
 - ▶ On pose $a_0 = a, b_0 = b$.
 - ▶ Supposons que a_k, b_k sont bien définis pour un entier $k \in \mathbf{N}$. On découpe l'intervalle $[a_k, b_k]$ en trois morceaux de taille identique, i.e. on pose

$$a_k < x_g = a_k + \frac{b_k - a_k}{3} < x_d = b_k - \frac{b_k - a_k}{3} < b_k.$$

- Si $f(x^g) < f(x^d)$, alors x_0 est «à gauche» de x_d ¹¹, on pose alors $a_{k+1} = a_k, b_{k+1} = x^d$.
- Si $f(x^g) > f(x^d)$, alors x_0 est «à gauche» de x_g , on pose alors $g_{k+1} = x^g, d_{k+1} = d_k$.
- Si $f(x^g) = f(x^d)$, alors on pose $g_{k+1} = x^g, d_{k+1} = x^d$.

On recommence ce procédé tant que $d - g > \varepsilon$ et on retourne à la fin la meilleure des valeurs parmi $f(g), f(d), f((g + d)/2)$.

- 2.1) Écrire une fonction d'en-tête `argmin3(f, x, y, z)` qui renvoie parmi les trois valeurs x, y et z , celle en laquelle f est minimale.
- 2.2) Compléter le script suivant implémentant cet algorithme.

Algorithme d'approximation d'un minimum

```
def minimum(f, A, B, eps):
    g = A
    d = B
    while d - g > eps:
        # calcul de xg xd
```



¹¹Raisonnement par l'absurde, s'il était à droite de x_d alors x_g, x_d se situeraient dans l'ensemble où f est strictement décroissante, donc on aurait $f(x^g) > f(x^d)$ car $x_g < x_d$ — contradiction.

```

xg = g + (d-g)/3
xd = d - (d-g)/3
if f(xg) < f(xd):
    d = xd
elif f(xg) > f(xd):
    g = -----
else:
    # on est sûr que x 0 est entre g et d
    g = -----
    d = -----
return argmin3(f, g, d, (g+d)/2)

```

3. On considère la fonction $f : x \in [0, 1] \mapsto x^2 - \frac{1}{2}x + 1$. Montrer que la fonction f vérifie bien les conditions de l'énoncé et lui appliquer l'algorithme. Quelle valeur approchée s'attendait-on à trouver? Est-ce cohérent?

Solution (exercice ANA.7.17)

1. Par exemple, une parabole.

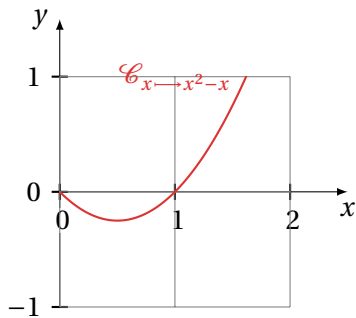


FIG. ANA.7.6. : $a = 0, b = 1, x_0 = \frac{1}{2}$

Algorithme d'approximation d'un minimum

2. `def argmin3(f, x, y, z):`
`mini = x`

```

if f(y) < f(mini):
    mini = y
if f(z) < f(mini):
    mini = z
return mini

def minimum(f, A, B, eps):
    g = A
    d = B
    while d - g > eps:
        # calcul de xg xd
        xg = g + (d-g)/3
        xd = d - (d-g)/3
        if f(xg) < f(xd):
            d = xd
        elif f(xg) > f(xd):
            g = xg
        else:
            # on est sûr que x 0 est entre g et d
            g = xg
            d = xd
    return argmin3(f, g, d, (g+d)/2)

```

3. Faisons un test sur la fonction proposée.

```

def f(x):
    return x**2-1/2*x+1
argmini = minimum(f, 0, 1, 10**(-3))

```

Alors `argmini` vaut 0.2500324034230415. Ce qui est tout à fait cohérent, puisque le minimum est atteint en $-\frac{-\frac{1}{2}}{2} = \frac{1}{4}$.

5.6. Du côté de l'Algèbre

Exercice ANA.718 | Fonctions à variations bornées Soit $D \subset \mathbf{R}$. On considère $\Delta = \{f - g, f, g \text{ croissantes sur } D\}$. On l'appelle l'ensemble des fonctions à variations bornées.

Montrer que Δ est un sous-espace vectoriel de \mathbf{R}^D .

Solution (exercice ANA.718)

- ▶ Nous avons $\Delta \subset \mathbf{R}^D$.
- ▶ De plus, comme $0 = 0 - 0$ et que la fonction nulle est croissante, il vient que $0_{\mathbf{R}^D}$ est dans Δ .
- ▶ Soient donc $\lambda, \mu \in \mathbf{R}$, et $\varphi_1 = f_1 - g_1, \varphi_2 = f_2 - g_2$ deux fonctions de Δ , où les f_i, g_i sont croissantes pour $i \in \llbracket 1, 2 \rrbracket$.

$$\lambda\varphi_1 + \mu\varphi_2 = \lambda(f_1 - g_1) + \mu(f_2 - g_2).$$

Plusieurs cas se présentent.

- ▶ 1. Si $\lambda, \mu \geq 0$. Alors $\lambda\varphi_1 + \mu\varphi_2 = (\lambda f_1 + \mu f_2) - (\lambda g_1 + \mu g_2)$. Et cette écriture fait bien apparaître une différence de fonctions croissantes.
- ▶ 2. Si $\lambda \leq 0, \mu \geq 0$. Alors $\lambda\varphi_1 + \mu\varphi_2 = (\lambda f_1 + \mu f_2) - (\lambda g_1 + \mu g_2) = ((-\lambda)g_1 + \mu f_2) - ((-\lambda)f_1 + \mu g_2)$. Alors comme $-\lambda \geq 0$, cette écriture fait encore bien apparaître une différence de fonctions croissantes.
- ▶ 3. Si $\lambda \geq 0, \mu \leq 0$, on fait la même analyse avec l'écriture $\lambda\varphi_1 + \mu\varphi_2 = (\lambda f_1 + (-\mu)g_2) - (\lambda g_1 + (-\mu)f_2)$.
- ▶ 4. Si $\lambda \leq 0, \mu \leq 0$. Alors $\lambda\varphi_1 + \mu\varphi_2 = ((-\lambda)g_1 + (-\mu)g_2) - ((-\lambda)f_1 + (-\mu)f_2)$.

En conclusion, Δ est un sous-espace vectoriel de \mathbf{R}^D .

Chapitre ANA.8.

Fonctions de plusieurs variables

Résumé & Plan

Dans ce chapitre nous allons étudier les fonctions de n variables avec $n \in \mathbf{N}^*$.¹ Plus précisément, nous allons voir une notion de continuité, dérivabilité directionnelle, et de classe \mathcal{C}^k avec $k \geq 0$. Il existe aussi une notion de dérivée (classique, non partielle) appelée *différentiabilité* mais qui n'est pas au programme de BCPST.

1	Généralités	2
1.1	Premiers exemples	2
1.2	Vocabulaire	3
2	Limite et continuité	4
2.1	Limite	5
2.2	Continuité	6

¹Officiellement, seules les fonctions de deux variables sont au programme, mais les sujets de concours sortent bien souvent de ce cadre. Par ailleurs, la plupart des énoncés ne présentent aucune difficulté à être énoncés pour des fonctions à plus de deux variables.

3	Dérivabilité	8
3.1	Dérivées directionnelles	8
3.2	Gradient	12
3.3	Fonctions \mathcal{C}^1	13
3.4	Fonctions \mathcal{C}^2	17
3.5	Extrema	19
4	Exercices	23
4.1	Généralités	23
4.2	Dérivation & Primitivation	24
4.3	Règle de la chaîne et équation aux dérivées partielles	26
4.4	En Physique	31

4.5 Optimisation 32

Le mathématicien et philosophe LEIBNIZ (1646-1716) fut le premier à démontrer la formule pour la dérivation d'un produit. De ce fait, cette formule porte aussi parfois le nom de règle de LEIBNIZ.

— Le saviez-vous ?

Cadre
 Nous considérerons dans ce chapitre des fonctions de n variables, i.e. des applications $f : \mathbb{R}^n \rightarrow \mathbb{R}$, ou plus généralement $f : P \rightarrow \mathbb{R}$ avec $P \subset \mathbb{R}^n$ si elle est seulement définie sur un sous-ensemble P de \mathbb{R}^n .

Définition ANA.8.1 | Pavé & Pavé ouvert

On appelle *pavé* de \mathbb{R}^n tout ensemble de la forme

$$I_1 \times \dots \times I_n = \{(x_1, \dots, x_n), x_1 \in I_1, \dots, x_n \in I_n\},$$

où I_1, \dots, I_n sont des intervalles de \mathbb{R} . Soit $X = (x_1, \dots, x_n) \in \mathbb{R}^n$. On appellera *voisinage ouvert* de $X = (x_1, \dots, x_n)$ tout ensemble de la forme

$$]x_1 - \varepsilon_1, x_1 + \varepsilon_1[\times \dots \times]x_n - \varepsilon_n, x_n + \varepsilon_n[, \quad \text{avec } \varepsilon_1, \dots, \varepsilon_n > 0.$$

Pour $n = 1$ un voisinage ouvert d'un point est un intervalle ouvert centré en ce point. Pour $n = 2$ c'est un rectangle ouvert de centre de gravité le point considéré.

1. GÉNÉRALITÉS

1.1. Premiers exemples

Citons quelques contextes où interviennent naturellement les fonctions de plusieurs variables.

- ▶ L'aire d'un rectangle de côtés x, y est $S(x, y) = xy$. C'est donc une fonction $S \left| \begin{array}{l} \mathbb{R}^2 \rightarrow \mathbb{R} \\ (x, y) \mapsto xy \end{array} \right.$ de deux variables. On peut se poser la question suivante : peut-on trouver un minimum pour S ? un minimum sous une contrainte de périmètre? i.e. $2(x + y) = p$ avec $p \in \mathbb{R}^{+*}$.
- ▶ **(En mécanique)** Si $\vec{AB} = \begin{pmatrix} a \\ b \end{pmatrix}$ est un vecteur de \mathbb{R}^2 , alors le travail d'une force $\vec{F} = \begin{pmatrix} x \\ y \end{pmatrix}$ le long d'un déplacement entre A et B est une fonction de quatre variables :

$$F(a, b, x, y) = ax + by.$$

Deux variables de position et deux variables de force.

- ▶ **(En géométrie)** Considérons le plan affine $\mathcal{P} : ax + by + cz + d = 0$ avec $(a, b, c, d) \in \mathbb{R}^4$ tel que $(a, b, c) \neq (0, 0, 0)$ et $B = (x_B, y_B, z_B) \in \mathbb{R}^3$. En première année, la distance d'un point B au plan M a été définie comme :

$$d(B, \mathcal{P}) = \inf_{M \in \mathcal{P}} d(B, M) = \inf_{M \in \mathcal{P}} \sqrt{(x_M - x_B)^2 + (y_M - y_B)^2 + (z_M - z_B)^2}.$$

Il s'agit donc ici de minimiser la fonction

$$F \left| \begin{array}{l} \mathcal{P} \subset \mathbb{R}^3 \rightarrow \mathbb{R} \\ (x, y, z) \mapsto \sqrt{(x - x_B)^2 + (y - y_B)^2 + (z - z_B)^2} \end{array} \right.$$

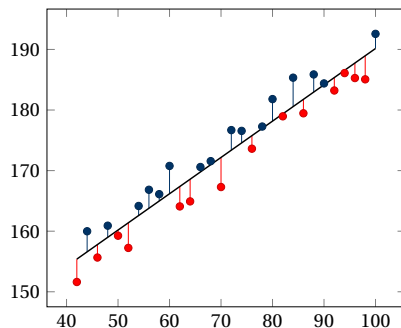
sur l'ensemble des vecteurs $(x, y, z) \in \mathcal{P}$.

- ▶ **(En statistiques : droite de régression)**

Si $(x_i, y_i)_{1 \leq i \leq n}$ avec $n \geq 1$ est un nuage de n points et que l'on se pose la question de l'existence d'une droite passant au plus près de ces points (au sens des « moindres carrés »), cela revient à minimiser cette fonctionnelle de deux variables :

$$F(a, b) = \sum_{i=1}^n (y_i - ax_i - b)^2,$$

où l'éventuel minimum trouvé (a^*, b^*) correspondra au couple coefficient directeur/ordonnée à l'origine de la droite dite des *moindres carrés*. Représentons cette quantité F sur un dessin.



La quantité $F(a, b)$ correspond donc à la somme des longueurs algébriques des traits verticaux bleus et rouges (positive si bleu, négative si rouge). Minimiser F revient donc à minimiser chacune de ces longueurs, et donc *in fine* approcher au plus près le nuage de points considéré.

1.2. Vocabulaire

Si $f : (x, y) \in \mathbf{R}^2 \rightarrow \mathbf{R}$ est une fonction de deux variables, alors on appellera « restrictions partielles » les applications définies pour tout $x_0, y_0 \in \mathbf{R}$ par :

$$x \mapsto f(x, y_0), \quad y \mapsto f(x_0, y).$$

On « fixe » donc une variable et on fait varier l'autre, si une fonction a deux variables, alors les restrictions partielles en auront qu'une. Plus généralement, si une fonction a n variables, alors les restrictions partielles en auront une. Passons à la définition.

Définition ANA.8.2 | Restriction partielle

Soit $f : \mathbf{R}^n \rightarrow \mathbf{R}$ ou plus généralement $f : P \rightarrow \mathbf{R}$ avec $P \subset \mathbf{R}^n$. Alors les applications, pour $i \in \llbracket 1, n \rrbracket$,

$$x \in \mathbf{R} \mapsto f(x_1, \dots, x_{i-1}, \boxed{x}, x_{i+1}, \dots, x_n)$$

sont appelées les applications *restrictions partielles* associées à f , ce sont des fonctions réelles d'une variable réelle.²

Exemple 1 — Les applications partielles de $f : (x_1, x_2) \in \mathbf{R}^2 \mapsto 3x_1^3 + e^{x_2}$ au point $a = (2, 7)$ sont $f_1 : t \mapsto 3t^3 + e^7$, $f_2 : t \mapsto 24 + e^t$.

Rappelons que le graphe d'une fonction d'une variable $f : I \subset \mathbf{R} \rightarrow \mathbf{R}$ a été défini dans le **Chapter ANA.7** comme étant

$$\{(x, f(x)), x \in I\}.$$

On généralise sans difficulté cette définition aux fonctions de **deux** variables uniquement.

Définition ANA.8.3 | Graphe & Ligne de niveau

Soit $f : P \rightarrow \mathbf{R}$ avec $P \subset \mathbf{R}^2$ une fonction de deux variables. On appelle *graphe de f* (ou *surface représentative de f*) l'ensemble

$$\mathcal{S}_f = \{(x, y, f(x, y)), (x, y) \in P\}.$$

Une *ligne de niveau $k \in \mathbf{R}$ fixé*, est le sous-ensemble de \mathbf{R}^2 défini par :

$$\mathcal{L}_k(f) = \{(x, y) \in P, f(x, y) = k\}.$$
³

Ainsi, dans le cas d'une fonction de deux variables, le graphe n'est donc ni plus ni moins qu'une surface dans l'espace, alors que les lignes de niveaux sont des courbes dans le plan. Voyons quelques exemples de lignes de niveaux.

Exemple 2 — Soit $f : (x, y) \in \mathbf{R}^2 \mapsto \ln(1 + x^2 + y^2)$. Déterminer $\mathcal{L}_k(f)$ pour tout $k \in \mathbf{R}$, en distinguant le cas $k = 0$.

²Pour les fonctions d'une variable nous n'avons pas parlé de restriction partielle, puisque ces dernières coïncideraient avec la fonction de départ.

³Une ligne de niveau est donc une courbe tracée sur l'ensemble de définition P

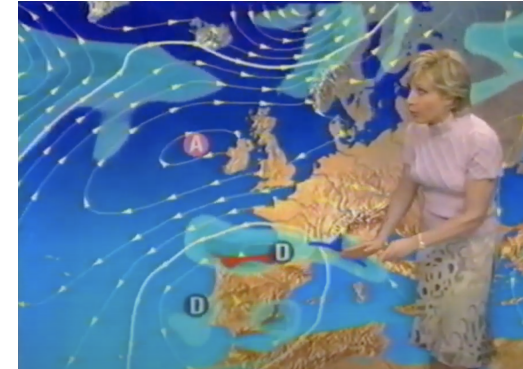


Fig. ANA.8.2. : Une carte météo TF1 (04/12/2003) faisant apparaître les isobares

Exemple 3 — Cartes IGN Le relief est un graphe de fonction de deux variables, et on représente souvent également les lieux de même altitude, qui correspondent donc aux lignes de niveau de \mathcal{S}_f . Si l'on faisait de-même pour des cartes de pression, les lignes de niveau représenteraient ce que l'on appelle plus communément les *isobares*.

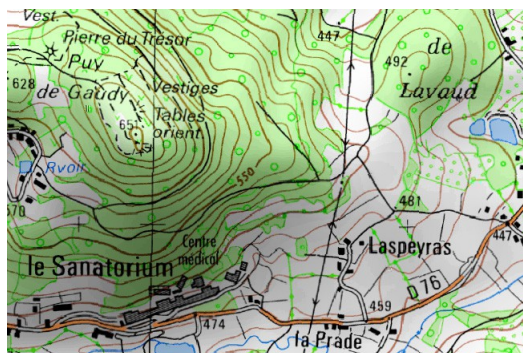


Fig. ANA.8.1. : Carte IGN d'un département Français

Exemple 4 — Cartes Météo Les courbes des lieux de même pression sur une carte météo, appelées *courbes isobares*, sont les lignes de niveau d'un graphe \mathcal{S}_f de fonction de deux variables, celle qui à un lieu géographique associe la pression en ce lieu.

2. LIMITE ET CONTINUITÉ

Réfléchissons à comment étendre toutes les notions d'analyse à des fonctions $f : \mathbf{R}^n \rightarrow \mathbf{R}$? Reprenons la définition de fonction continue $f : I \rightarrow \mathbf{R}$ en $a \in I$:

$$\forall \varepsilon > 0, \exists \eta > 0, \forall x \in I, \quad |x - a| < \eta \implies |f(x) - f(a)| < \varepsilon.$$

Que doit-on changer dans cette définition ? Quasiment rien, les quantités $f(x) - f(a)$ et $x - a$ sont bien définies si $x, a \in \mathbf{R}^n$, puisque l'on peut faire la différence de deux vecteurs de \mathbf{R}^n , mais comment donner un sens à

$|x - a|$? En effet, la valeur absolue d'un vecteur n'existe pas.

Notons $x = (x_1, \dots, x_n), a = (a_1, \dots, a_n) \in \mathbf{R}^n$. Une solution : nous allons remplacer « $|x - a| < \eta$ » dans la définition de la limite par :

$$\ll |x_1 - a_1| < \eta, \quad \dots, \quad |x_n - a_n| < \eta. \gg$$

2.1. Limite

Cadre
 Dans cette section, P désignera toujours un pavé de \mathbf{R}^n non vide et non réduit à un point.

Passons maintenant aux *nouvelles* définitions de limite, pas si nouvelles finalement puisqu'il s'agit simplement de faire la substitution annoncée précédemment, et d'enlever celles qui n'ont aucun sens (vers les infinis notamment). Avant cela, nous avons besoin de la notion d'adhérence.


Définition ANA.8.4 | Adhérence d'un ensemble

Soit P pavé de \mathbf{R}^n , on dit qu'un élément $x = (x_1, \dots, x_n) \in \mathbf{R}^n$ est *adhérent* à P si :

$$\forall \varepsilon > 0, \quad (]x_1 - \varepsilon, x_1 + \varepsilon[\times \dots \times]x_n - \varepsilon, x_n + \varepsilon[) \cap P \neq \emptyset.^4$$

Alors on appelle *adhérence* de P l'ensemble des réels adhérents à P . On le note en général \bar{P} .

Remarque 2.1 — Comme dans \mathbf{R} , ce sont donc des points « très proches » de P , mais pas forcément dedans.

Exemple 5 — Dessinons $P = [-1, 1] \times [0, 2[$. Alors $(0, 2)$ est adhérent à P , alors que $(0, 3)$ ne l'est pas. 

⁴Pour $n = 2$, cet ensemble $]x_1 - \varepsilon, x_1 + \varepsilon[\times]x_2 - \varepsilon, x_2 + \varepsilon[$ n'est autre qu'un carré de longueur de côté 2ε , et centré en (x_1, x_2)

Définition ANA.8.5 | Limite en un point adhérent

Soient $A = (a_1, \dots, a_n) \in \bar{P}$, où P est un sous-ensemble de \mathbf{R}^n , $\ell \in \mathbf{R} \cup \{\pm\infty\}$ et une application f définie sur P ou sur $P \setminus \{A\}$. On note :

▶ $f(X) \xrightarrow{X \rightarrow A} \ell$ si :

$$\forall \varepsilon > 0, \exists \eta_\varepsilon > 0, \forall X = (x_1, \dots, x_n) \in P, \quad |x_1 - a_1| < \eta_\varepsilon, \dots, |x_n - a_n| < \eta_\varepsilon \implies |f(X) - \ell| < \varepsilon,$$

▶ $f(X) \xrightarrow{X \rightarrow A} +(-)\infty$ si :

$$\forall A \in \mathbf{R}, \exists \eta_A > 0, \forall X = (x_1, \dots, x_n) \in P, \quad |x_1 - a_1| < \eta_\varepsilon, \dots, |x_n - a_n| < \eta_\varepsilon \implies f(X) > (<)A.$$

Attention

L'ensemble \mathbf{R}^n pour $n \geq 2$ ne possède pas d'« infini » *a priori*. Les notions de limites en ces points n'existent donc pas.

Remarque 2.2 — Où intervient $A = (a_1, \dots, a_n) \in \bar{P}$? Si c'est le cas, l'ensemble

$$\{X = (x_1, \dots, x_n) \in P, \quad |x_1 - a_1| < \eta, \dots, |x_n - a_n| < \eta\} \neq \emptyset \text{ pour tout } \eta > 0.$$

Bien entendu c'est déjà le cas si $A \in P \subset \bar{P}$, mais ça l'est aussi plus généralement pour n'importe quel point de l'adhérence.

Cette notion de limite possède le même type de propriétés que pour la limite de fonctions de \mathbf{R} dans \mathbf{R} ; unicité, opérations, inégalités, ... nous ne les détaillerons pas à nouveau ici.

Proposition ANA.8.1 |

Soit $f : P \rightarrow \mathbf{R}$, où P est un pavé de \mathbf{R}^n et $A \in \mathbf{R}^n$. Si f est **définie en** A (i.e. $A \in P$) et possède une limite en A , alors :

$$\lim_{X \rightarrow A} f(X) = f(A).$$

Preuve Identique au cas des fonctions d'une variable.

NIER L'EXISTENCE D'UNE LIMITE. Dans le cas des fonctions d'une variable, on utilisait très souvent la caractérisation séquentielle : on exhibe deux suites convergeant vers la même chose, mais dont les suites images par f tendent vers des limites différentes. Traditionnellement, pour les fonctions de plusieurs variables, on nie directement la définition de la limite, comme précisé dans la méthode qui suit.

 **Méthode Nier l'existence d'une limite pour les fonctions de plusieurs variables**

Si $f : P \subset \mathbf{R}^n \rightarrow \mathbf{R}$ et $A \in P$. Alors si f admet une limite $\ell \in \overline{\mathbf{R}}$ en A , pour toute fonction $\varphi : Q \subset \mathbf{R}^n \rightarrow P$ (avec $A \in \overline{Q}$) telle que $\lim_B \varphi = A$ pour $B \in \overline{Q}$, on a :

$$\lim_{X \rightarrow B} f \circ \varphi(X) = \ell.$$

Ainsi, si l'on trouve deux telles fonctions $\varphi_1 : Q_1 \subset \mathbf{R}^n \rightarrow P$ (associée à un point $B_1 \in \overline{Q_1}$) et $\varphi_2 : Q_2 \subset \mathbf{R}^n \rightarrow P$ (associée à un point $B_2 \in \overline{Q_2}$), vérifiant :

$$\lim_{X \rightarrow B_1} f \circ \varphi_1(X) \neq \lim_{X \rightarrow B_2} f \circ \varphi_2(X),$$


la fonction n'admet donc pas de limite en A . On dit que l'on a trouvé des « chemins le long desquels la fonction ne converge pas vers la même limite ».

En résumé, la limite doit être la même peu importe le « chemin emprunté » pour tendre vers A .

Exemple 6 — Considérons les deux fonctions ci-dessous :

1. Soit f_1 la fonction de deux variables définie par :

$$f_1 : \begin{cases} \mathbf{R}^2 \longrightarrow \mathbf{R} \\ (x, y) \longrightarrow \begin{cases} \frac{xy}{x^2+y^2} & \text{si } (x, y) \neq (0, 0), \\ 0 & \text{sinon.} \end{cases} \end{cases} \quad \text{La fonction } f_1 \text{ n'admet pas de li-}$$

mite en $(0, 0)$. 

2. Soit f_2 la fonction de deux variables définie par :

$$f_2 : \begin{cases} \mathbf{R}^2 \longrightarrow \mathbf{R} \\ (x, y) \longrightarrow \begin{cases} \frac{x^2 y^2}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0), \\ 0 & \text{sinon.} \end{cases} \end{cases} \quad \text{La fonction } f_2 \text{ vérifie } \lim_{(x,y) \rightarrow (0,0)} f_2(x, y) =$$

0. 

2.2. Continuité

Nous avons maintenant une notion de limite, on peut donc regarder l'ensemble des fonctions continues que l'on définit dès à présent.

Définition ANA.8.6 | Continuité


Soit $f : P \rightarrow \mathbf{R}$ où P est un pavé de \mathbf{R}^n et $A \in P$. On dit que f est *continue en A* si :

$$f(A) = \lim_{X \rightarrow A} f(X)$$

On dit que f est *continue sur P* si elle l'est en tout point de P .

Attention

Disons-le de suite, la continuité des restrictions partielles de f n'entraîne pas la continuité de f !

Exemple 7 — On reprend les fonctions f_1, f_2 de l'Exemple 6. Étudier la continuité de f_1, f_2 sur \mathbf{R}^2 , ainsi que celle de leurs restrictions partielles. 

Attention

De manière générale, la continuité entraîne celle des restrictions partielles mais la réciproque est fautive.

Exemple 8 — *Autres applications continues usuelles*

1. Les projections $p_i : \begin{cases} \mathbf{R}^n & \rightarrow \mathbf{R} \\ (x_1, \dots, x_n) & \mapsto x_i \end{cases}$ définies pour tout $i \in \llbracket 1, n \rrbracket$ sont continues.
2. Les applications polynomiales sont continues.

3. DÉRIVABILITÉ


Nous avons considéré dans le **Chapter ANA.7** des fonctions $f : I \rightarrow \mathbf{R}$ et $a \in \bar{I}$ où I est un intervalle de \mathbf{R} . Ici nous allons définir le nombre dérivé sur des ensembles dits ouverts \mathcal{O} de \mathbf{R}^n , *i.e.* en des points qui ne sont pas situés sur le *bord* de l'ensemble, afin d'éviter de parler de limite à droite/gauche (qui n'existe pas dans \mathbf{R}^n).


Définition ANA.8.7 | Ensemble ouvert

Soit $\mathcal{O} \subset \mathbf{R}^n$. On dit que \mathcal{O} est un *ouvert de \mathbf{R}^n* si pour tout $x = (x_1, \dots, x_n) \in \mathcal{O}$, il existe $\varepsilon > 0$ tel que :

$$]x_1 - \varepsilon, x_1 + \varepsilon[\times \dots \times]x_n - \varepsilon, x_n + \varepsilon[\subset \mathcal{O}.^5$$

Exemple 9 — Parmi les ensembles suivants, lesquels sont ouverts?

▶ $(n = 1)$ $]1, 2[, [2, 9[,]1, +\infty[,]1, 2[\cup]2, 10[?$ 

▶ $(n = 2)$ $[-1, 1] \times [0, 1]?$ 

⁵Pour $n = 2$, cela signifie qu'il existe un carré ouvert assez petit inclus dans \mathcal{O} autour de x .

3.1. Dérivées directionnelles

Pour les fonctions de plusieurs variables, nous allons définir la dérivation, selon une direction $V \in \mathbf{R}^n$ au point $A \in \mathbf{R}^n$, lorsque les fonctions sont définies sur un ensemble ouvert. Autrement dit, nous allons observer comment $f(X)$ varie, lorsque X se déplace sur la droite passant par A et dirigée par V . Cela revient à considérer le taux d'accroissement de la fonction d'une variable réelle $t \rightarrow f(A + tV)$ lorsque t tend vers zéro.

Définition ANA.8.8 | Dérivée directionnelle

Soit $f : \mathcal{O} \rightarrow \mathbf{R}$, où \mathcal{O} est un sous-ensemble ouvert⁶ de \mathbf{R}^n , $A \in \mathcal{O}$ et V un vecteur de \mathbf{R}^n . On dit que f est *dérivable en A dans la direction V* si :

$$t \mapsto \frac{f(A + tV) - f(A)}{t}$$

admet une limite finie en zéro. Dans ce cas, on note $\frac{\partial f}{\partial V}(A)$ la valeur de la limite.

Remarque 3.1 — Interprétation La dérivée directionnelle $\frac{\partial f}{\partial V}(A)$ peut être interprétée comme le taux d'accroissement infinitésimal de f au point A dans la direction V .


Attention

Le fait que f soit dérivable partiellement, même dans toutes les directions $V \in \mathbf{R}^n$, **n'entraîne pas** la continuité de f au point considéré.

Exemple 10 — Contre-exemple Considérons la fonction $g : \mathbf{R}^2 \rightarrow \mathbf{R}$ définie par $(x, y) \mapsto \begin{cases} \frac{y^2}{x} & \text{si } x \neq 0, \\ y & \text{sinon.} \end{cases}$ Alors g admet une dérivée dans toutes les directions en $(0, 0)$, mais n'est pas continue en $(0, 0)$.

⁶Ainsi, sous cette hypothèse, pour t assez petit, la quantité $A + tV$ est dans \mathcal{O} donc $f(A + tV)$ est bien définie.

► **(Dérivées en (0,0) dans toutes les directions)** soit $V = (v_1, v_2) \in \mathbf{R}^2$, et $t \in \mathbf{R}^*$.

 Si $v_1 \neq 0$ alors :

$$\frac{g(0 + tv_1, 0 + tv_2) - g(0,0)}{t} = \frac{1}{t} \frac{t^2 v_2^2}{tv_1} \xrightarrow{t \rightarrow 0} \frac{v_2^2}{v_1}.$$

Si $v_1 = 0$:

$$\frac{g(0 + tv_1, 0 + tv_2) - g(0,0)}{t} = \frac{1}{t} tv_2 \xrightarrow{t \rightarrow 0} v_2.$$

Dans les deux cas f est dérivable en $(0,0)$ dans la direction (v_1, v_2) .

► Pourtant, cette fonction n'est pas continue en $(0,0)$. En effet, $g(x,0) = x \xrightarrow{x \rightarrow 0} 0$, et pour $x \neq 0$, $g(x^2, x) = 1 \xrightarrow{x \rightarrow 0} 1 \neq 0$.

CAS PARTICULIER : LES DÉRIVÉES PARTIELLES. Lorsque les directions V sont les vecteurs de la base canonique $\mathcal{B}^{\text{can}} = (e_1, \dots, e_n)$ de \mathbf{R}^n , on appelle généralement ces objets les *dérivées partielles* de f .

Définition ANA.8.9 | Dérivée partielle

Soit $f : \mathcal{O} \rightarrow \mathbf{R}$, où \mathcal{O} est un sous-ensemble ouvert de \mathbf{R}^n , et $A = (a_1, \dots, a_n) \in \mathcal{O}$, $i \in \{1, \dots, n\}$. On dit que f possède une *dérivée partielle dans la direction i* si f possède une dérivée directionnelle selon $V = e_i$, i.e. si :

$$t \mapsto \frac{f(A + te_i) - f(A)}{t}$$

$$\text{i.e. si } t \mapsto \frac{f(a_1, \dots, a_{i-1}, \boxed{a_i + t}, a_{i+1}, \dots, a_n) - f(a_1, \dots, a_n)}{t}$$

admet une limite finie en zéro. Dans ce cas, on note $\partial_i f(A)$ (ou encore $\partial_i f(A)$, $\frac{\partial f}{\partial e_i}(A)$) la valeur de la limite.

Remarque 3.2 — On « fixe » donc toutes les variables sauf une, selon laquelle on calcule le taux d'accroissement et on dérive. C'est donc la dérivée en a_i de la i -ème restriction partielle.



Notation Extension pour la Physique

Nous avons défini ici une notion de dérivation par rapport à l'une des variables, mais parfois, dans certains contextes la fonction f n'est pas explicitement donnée. Par exemple, rappelons la loi fondamentale d'un resistor de résistance R :

$$U = R \cdot I,$$

sous les notations du cours de Physique. Que signifie $\frac{\partial U}{\partial I}$ souvent employée par les physiciens? La quantité U s'écrit $U = f(R, I)$ avec $f : (x, y) \mapsto xy$, et

$$\frac{\partial U}{\partial I} \text{ (déf.)} = \frac{\partial f}{\partial y}(R, I) = ((x, y) \mapsto x)(R, I) = R.$$



Notation Cas $n = 2$ — les plus importantes

Notons $A = (a_1, a_2) \in \mathbf{R}^2$ et $f : \mathbf{R}^2 \rightarrow \mathbf{R}$. On rappelle que la base canonique de \mathbf{R}^2 est (e_1, e_2) où $e_1 = (1, 0)$, $e_2 = (0, 1)$. On peut donc dériver suivant deux directions de la base canonique, et on note avec les mêmes notations que dans la définition :

$$\partial_1 f(A) = \frac{\partial f}{\partial x}(A) = \lim_{t \rightarrow 0} \frac{f(a_1 + t, a_2) - f(a_1, a_2)}{t}, \quad \text{(derx)}$$

$$\partial_2 f(A) = \frac{\partial f}{\partial y}(A) = \lim_{t \rightarrow 0} \frac{f(a_1, a_2 + t) - f(a_1, a_2)}{t}. \quad \text{(dery)}$$

L'existence de ces limites implique la continuité des restrictions partielles $a_1 \mapsto f(a_1, a_2)$ pour tout $a_2 \in \mathbf{R}$, et $a_2 \mapsto f(a_1, a_2)$ pour tout $a_1 \in \mathbf{R}$. Les quantités **(derx)** et **(dery)** sont simplement les taux d'accroissement de ces deux fonctions aux points a_1 et a_2 .



Attention Cas $n = 2$ — attention aux notations!

Soit $f : \mathbf{R}^2 \rightarrow \mathbf{R}$. Il faut bien comprendre, que :



► **(dérivée partielle évaluée en ...)**

$$\frac{\partial f}{\partial x}(x, y), \quad \frac{\partial f}{\partial x}(u, v), \quad \frac{\partial f}{\partial x}(s, t), \quad \dots$$

Parfois on peut préférer — et je la préfère!⁷ — la notation $\partial_1 f$, $\partial_2 f$ qui ne fait pas intervenir le nom éventuel donné à chaque variable.

est différent de :

► **(dériver une expression par rapport à ...)**

$$\frac{\partial}{\partial x}(f(x, x^2)) \quad \left(\text{ou } \frac{d}{dx}(f(x, x^2)) \text{ si une seule variable est présente} \right)$$

signifie en revanche que l'on dérive **l'expression** par rapport à x : attention donc à l'emplacement des parenthèses.

Les constats précédents seront importants pour bien comprendre les formules de dérivation d'une composée qui vont suivre.

Les dérivées partielles & directionnelles héritent des mêmes propriétés de dérivabilité que pour les sommes, produits, quotients, composées déjà connues. Nous les citons ici sans démonstration.

Proposition ANA.8.2 | Propriétés des dérivées partielles et directionnelles

Soient $f : \mathcal{O} \rightarrow \mathbf{R}$ et $g : \mathcal{O} \rightarrow \mathbf{R}$, où \mathcal{O} est un sous-ensemble ouvert de \mathbf{R}^n , et $A = (a_1, \dots, a_n) \in \mathcal{O}$, $V \in \mathbf{R}^n$ une direction.

► **(Produit et inverse)** Si f, g possèdent en A une dérivée directionnelle dans la direction V , alors $f \times g$ aussi et :

$$\frac{\partial (f \times g)}{\partial V}(A) = \frac{\partial f}{\partial V}(A)g(A) + f(A)\frac{\partial g}{\partial V}(A).$$

⁷Mais je ne l'utiliserai quasiment pas, car les sujets de concours ne l'utilisent pas

En conséquence, si $f(A) \neq 0$, on a :

$$\frac{\partial \left(\frac{1}{f} \right)}{\partial V}(A) = -\frac{1}{f(A)^2} \frac{\partial f}{\partial V}(A).$$

► **(Composition à gauche par une fonction d'une variable)** Si f possède en A une dérivée directionnelle dans la direction V , et $\varphi : \mathbf{R} \rightarrow \mathbf{R}$ dérivable en $f(A)$, alors $\varphi \circ f$ possède une dérivée directionnelle dans la direction V et :


$$\frac{\partial (\varphi \circ f)}{\partial V}(A) = \frac{\partial f}{\partial V}(A)\varphi'(f(A)).$$


Remarque 3.3 — En particulier, la proposition est vérifiée pour les dérivées partielles.


Preuve Même preuve que pour les fonctions d'une variable.

Exemple 11 — Soit $f : (x, y) \in \mathbf{R}^2 \mapsto xy$. Déterminer si elles existent les dérivées partielles de f , **en utilisant la définition.**

Exemple 12 – Déterminer les ensembles de définition des fonctions suivantes, et déterminer si elles existent leurs dérivées partielles.


1. $g : (x, y) \in \mathbf{R}^2 \mapsto 1 + xy + y^2$, 

2. $h : (x, y) \in \mathbf{R}^2 \mapsto \sqrt{x^2 + y^2}$, 

3. $i : (x, y, z) \in \mathbf{R}^3 \mapsto \ln(x + 2z) + \sqrt{1 + y^2}$. 

De-même que l'on peut dériver partiellement une expression de plusieurs variables, on peut se poser la question de l'existence et du calcul d'une primitive partielle selon l'une des variables. Voyons quelques exemples classiques.

Exemple 13 – *Primitivation partielle première*

1. Déterminer les fonctions f de classe \mathcal{C}^1 telles que $\frac{\partial f}{\partial x} = 0$. 

2. Déterminer les fonctions f de classe \mathcal{C}^1 sur $\mathbf{R}^2 \setminus \{(0, 0)\}$ et vérifiant :

$$\forall (x, y) \in \mathbf{R}^2 \setminus \{(0, 0)\}, \quad \begin{cases} \frac{\partial f}{\partial x}(x, y) = \frac{x}{\sqrt{x^2 + y^2}}, \\ \frac{\partial f}{\partial y}(x, y) = \frac{y}{\sqrt{x^2 + y^2}}. \end{cases}$$



3.2. Gradient

La notion de gradient définie ci-dessous généralise la quantité $f'(a)$ lorsque $f : \mathbf{R} \rightarrow \mathbf{R}$ et $a \in \mathbf{R}$, où f est dérivable en a .

Définition ANA.8.10 | Gradient

Soit $f : \mathcal{O} \rightarrow \mathbf{R}$ où \mathcal{O} est un sous-ensemble ouvert de \mathbf{R}^n , et f une fonction admettant des dérivées partielles dans toutes les directions en un point $A \in \mathcal{O}$. On appelle *gradient de f en $A \in \mathcal{O}$* le vecteur de \mathbf{R}^n :

$$\text{grad } f(A) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(A) \\ \vdots \\ \frac{\partial f}{\partial x_n}(A) \end{pmatrix}.$$



Notation

On le note généralement $\text{grad } f(A)$ ou encore $\nabla_A f$.

Remarque 3.4 – Interprétation La directionnelle $\frac{\partial f}{\partial v}(A)$ peut être interprétée comme le taux d'accroissement infinitesimal de f au point A dans la direction v .

Définition ANA.8.11 | Point critique

On appelle *point critique* d'une fonction admettant des dérivées partielles dans toutes les directions $f : \mathcal{O} \rightarrow \mathbf{R}$ avec \mathcal{O} un ensemble ouvert, tout élément $C \in \mathcal{O}$ tel que $\text{grad } f(C) = 0$.

Remarque 3.5 – Cas $n = 2$ Avec les notations de la définition, si $f : \mathbf{R}^2 \rightarrow \mathbf{R}$, un point C est un point critique si et seulement si :

$$\begin{cases} \partial_1 f(C) = \frac{\partial f}{\partial x}(C) = 0, \\ \partial_2 f(C) = \frac{\partial f}{\partial y}(C) = 0. \end{cases}$$

Exemple 14 – On note $f : \begin{cases} \mathbf{R}^2 & \rightarrow \\ (x, y) & \mapsto \sqrt{1+x^2+y^2} \end{cases} \mathbf{R}$. Montrer que f admet un gradient en tout point de \mathbf{R}^2 , puis calculer $\text{grad } f$ en tout point $A = (a_1, a_2) \in \mathbf{R}^2$, ainsi que ses points critiques.

Exemple 15 – On note $g : \begin{cases} \mathbf{R}^n & \rightarrow \\ (x_1, \dots, x_n) & \mapsto \sum_{i=1}^n x_i^2 \end{cases} \mathbf{R}$. Montrer que f admet un gradient en tout point de \mathbf{R}^n , puis calculer $\text{grad } g$ en tout point $x \in \mathbf{R}^n$, ainsi que ses points critiques.


3.3. Fonctions \mathcal{C}^1

Définition ANA.8.12 | Classe \mathcal{C}^1

Soit $f : \mathcal{O} \rightarrow \mathbf{R}$ où \mathcal{O} est un sous-ensemble ouvert de \mathbf{R}^n . On dit que f est de classe \mathcal{C}^1 sur \mathcal{O} si toutes ses dérivées partielles sont définies et continues sur \mathcal{O} .

Notation

On note $\mathcal{C}^1(\mathcal{O}, \mathbf{R})$ l'ensemble des fonctions de classe \mathcal{C}^1 sur \mathcal{O} .

Exemple 16 — Les fonctions g, h définies dans l'Exemple 12 sont-elles \mathcal{C}^1 ? 

Exemple 17 —

- ▶ Les projections $p_i \left| \begin{array}{l} \mathbf{R}^n \rightarrow \mathbf{R} \\ (x_1, \dots, x_n) \mapsto x_i \end{array} \right.$ définies pour tout $i \in \llbracket 1, n \rrbracket$ sont de classe \mathcal{C}^1 .
- ▶ Les applications polynomiales sont de classe \mathcal{C}^1 .

Proposition ANA.8.3 | Propriétés des fonctions \mathcal{C}^1

Soient $f : \mathcal{O} \rightarrow \mathbf{R}$ et $g : \mathcal{O} \rightarrow \mathbf{R}$, où \mathcal{O} est un sous-ensemble ouvert de \mathbf{R}^n , deux fonctions de classe \mathcal{C}^1 sur \mathcal{O} . Alors :

- ▶ $f + g$ et $f \times g$ sont aussi de classe \mathcal{C}^1 ,
- ▶ si g ne s'annule pas, le quotient $\frac{f}{g}$ est aussi de classe \mathcal{C}^1 .
- ▶ **(Composition à gauche par une fonction d'une variable)** Si $\varphi : \mathbf{R} \rightarrow \mathbf{R}$ est également de classe \mathcal{C}^1 , alors $\varphi \circ f$ est également de classe \mathcal{C}^1 .

Preuve Utiliser les résultats sur les fonctions d'une variable aux restrictions partielles.

Enfin, la formule de TAYLOR-YOUNG reste valable pour des fonctions de n variables de classe \mathcal{C}^1 . Nous l'énonçons à l'ordre un. Rappelons-là déjà pour les fonctions d'une variable : si I est un intervalle, $a \in I$ et $f : I \rightarrow \mathbf{R}$ une fonction de classe \mathcal{C}^1 , alors

$$f(x) \underset{x \rightarrow a}{=} f(a) + f'(a)(x - a) + o(x - a).$$

Le seul terme qui va changer pour les fonctions de plusieurs variables est le terme d'ordre 1 : nous allons remplacer $f'(a)(x - a)$ par son analogue dans \mathbf{R}^n :

$$f'(a)(x - a) \sim \langle \text{grad } f(A) | X - A \rangle,$$

où $\langle \cdot, \cdot \rangle$ est défini ci-après.

Notation Produit scalaire, petit $\circ()$

Dans la suite, pour deux vecteurs $X = (x_1, \dots, x_n), Y = (y_1, \dots, y_n)$ de \mathbf{R}^n , on notera :

$$\langle X | Y \rangle = \sum_{i=1}^n x_i y_i.^8$$

Théorème ANA.8.1 | TAYLOR-YOUNG à l'ordre un [H.P]

Soient $f : \mathcal{O} \rightarrow \mathbf{R}$ de classe \mathcal{C}^1 et $A = (a_1, \dots, a_n) \in \mathcal{O}$ où \mathcal{O} est un sous-ensemble ouvert de \mathbf{R}^n . Alors :

$$f(X) \underset{X \rightarrow A}{=} f(A) + \langle \text{grad } f(A) | X - A \rangle + o(X - A).$$

⁸Ce produit scalaire sera étudié plus en détail dans un prochain chapitre

où $o(X - A)$ est une quantité réelle telle que :

$$o(X - A) = \left(\sum_{i=1}^n |x_i - a_i| \right) \times \varepsilon(X - A)^9,$$

pour une certaine fonction $\varepsilon : \mathbf{R}^n \rightarrow \mathbf{R}$ telle que : $\lim_0 \varepsilon = 0$.

Preuve Admis.

Remarque 3.6 — De manière équivalente, par changement de variable « $H = X - A$ » dans la limite, le résultat peut s'écrire sous la forme suivante :

$$f(A + H) = f(A) + \langle \text{grad } f(A) | H \rangle + o(H),$$

où $o(H) = \left(\sum_{i=1}^n |H_i| \right) \varepsilon(H)$ pour une certaine fonction $\varepsilon : \mathbf{R}^n \rightarrow \mathbf{R}$ telle que : $\lim_0 \varepsilon = 0$.

Corollaire ANA.8.1 | TAYLOR-YOUNG à l'ordre un (pour deux variables)

Soient $f : \mathcal{O} \rightarrow \mathbf{R}$ et $A = (a_1, a_2) \in \mathcal{O}$ où \mathcal{O} est un sous-ensemble ouvert de \mathbf{R}^2 . Alors :

$$f(X) \underset{X \rightarrow A}{=} f(A) + \langle \text{grad } f(A) | X - A \rangle + o(X - A),$$

$$f(x_1, x_2) \underset{(x,y) \rightarrow (a_1, a_2)}{=} f(a_1, a_2) + \frac{\partial f(A)}{\partial x} \cdot (x - a_1) + \frac{\partial f(A)}{\partial y} \cdot (y - a_2) + o(X - A).$$

Où, de manière équivalente,

$$f(a_1 + h, a_2 + k) \underset{(h,k) \rightarrow (0,0)}{=} f(a_1, a_2) + \frac{\partial f(A)}{\partial x_1} h + \frac{\partial f(A)}{\partial x_2} k + o((h, k)).$$

Preuve Conséquence directe du théorème général.


⁹Pour $n = 1$, on retrouve la classique définition du petit $o(\cdot)$ (une fonction tend vers zéro si et seulement si sa valeur absolue tend vers zéro)


Corollaire ANA.8.2 | Une fonction \mathcal{C}^1 admet des dérivées directionnelles

Soient $f : \mathcal{O} \rightarrow \mathbf{R}$ et $A = (a_1, a_2) \in \mathcal{O}$, $V \in \mathbf{R}^n$ où \mathcal{O} est un sous-ensemble ouvert de \mathbf{R}^2 . Alors f admet une dérivée directionnelle dans la direction V , donnée par :

$$\frac{\partial f}{\partial V}(A) = \langle \text{grad } f(A) | V \rangle.$$

Remarque 3.7 — Interprétation géométrique du gradient Parmi les directions V de norme un, on peut montrer que la direction qui maximise $|\langle \text{grad } f(A) | V \rangle|$ est un vecteur V colinéaire à $\text{grad } f(A)$. On résume cela en « le gradient donne la direction dans laquelle f varie le plus vite autour du point A ».

Exemple 18 — Soit $f : (x, y, z) \in \mathbf{R}^3 \mapsto 1 + xy^2 + x^2z$ et $a = (2, 1, 0)$. Écrire la formule de TAYLOR-YOUNG pour f à l'ordre 1 en a . 

Exemple 19 — Application à l'approximation On reprend la fonction $g : (x, y) \in \mathbf{R}^2 \mapsto \sqrt{x^3 + y^3}$. Déterminer une valeur approchée de $\sqrt{1,02^3 + 2,07^3}$ sans calculatrice. 

où f est une fonction $f : \mathbf{R}^n \rightarrow \mathbf{R}$ supposée de classe \mathcal{C}^1 . Supposons que des relevés des x_i peuvent être entachés d'une erreur $\delta x_i \geq 0$ pour tout $i \in \llbracket 1, n \rrbracket$. Alors, on notant $\delta f(x_1, \dots, x_n) = f(x_1 + \delta x_1, \dots, x_n + \delta x_n) - f(x_1, \dots, x_n)$ l'erreur correspondante sur f :

$$\delta f(x_1, \dots, x_n) \approx \sum_{i=1}^n \delta x_i f'(x_1, \dots, x_n)_i,$$

ou encore en valeur absolue, d'après l'inégalité triangulaire :

$$|\delta f(x_1, \dots, x_n)| \leq \sum_{i=1}^n |\delta x_i| |f'(x_1, \dots, x_n)_i|.$$

Cette formule permet donc de contrôler l'erreur maximale sur f commise en fonction des erreurs qui entachent chacune des mesures.

Remarque 3.8 — Interprétation géométrique de la formule de la TAYLOR-YOUNG pour $n = 2$ La formule de TAYLOR-YOUNG, comme pour les fonctions d'une variable¹⁰, possède une interprétation géométrique. En effet, avec les mêmes notations que dans la formule, notons $T_A(f)$ le plan ci-après :

$$T_A(f) : z = f(A) + \frac{\partial f}{\partial x}(A)(x - a_1) + \frac{\partial f}{\partial y}(A)(y - a_2).$$

C'est un plan de \mathbf{R}^3 , dont un vecteur normal est donné par

$$\left(\frac{\partial f}{\partial x}(A), \frac{\partial f}{\partial y}(A), 1 \right)^{11}$$

On l'appelle le *plan tangent à \mathcal{S}_f — surface représentative de f — en A* . La formule de TAYLOR-YOUNG nous dit qu'au voisinage de $A = (a_1, a_2) \in \mathbf{R}^2$, la surface représentative \mathcal{S}_f de f est alors « très proche » du plan tangent $T_A(f)$.

On termine par un corollaire classique, qui se démontre comme pour les fonctions d'une variable : dès qu'une fonction admet un développement limité au voisinage d'un point, alors la fonction est continue en ce point. Ici, l'existence d'un tel développement est garantie par la formule de TAYLOR-YOUNG qui requiert une hypothèse \mathcal{C}^1 dans ce cours.

¹⁰ où l'on sait qu'une courbe s'approche localement par sa tangente

Exemple 20 — Incertitudes de mesures Supposons qu'une quantité dépende de manière \mathcal{C}^1 de plusieurs paramètres notés x_1, \dots, x_n avec la relation

$$y = f(x_1, \dots, x_n)$$

Corollaire ANA.8.3 | Toute fonction \mathcal{C}^1 est continue

Soient $f : \mathcal{O} \rightarrow \mathbf{R}$ et $A = (a_1, \dots, a_n) \in \mathcal{O}$ où \mathcal{O} est un sous-ensemble ouvert de \mathbf{R}^n . Si f est de classe \mathcal{C}^1 , alors elle est continue.

Preuve Il suffit de faire tendre X vers A dans la formule de TAYLOR-YOUNG : en effet, $\langle \text{grad } f(A) | X - A \rangle = \sum_{i=1}^n \frac{\partial f(A)}{\partial x_i} (x_i - a_i) \xrightarrow{X \rightarrow A} 0$. On obtient $f(X) \xrightarrow{X \rightarrow A} f(A)$ puisque $o(X - A) \xrightarrow{X \rightarrow A} 0$ par définition d'un « petit o ».

Résumé Lien entre les notions de régularités pour les fonctions de plusieurs variables

**Théorème ANA.8.2 | Règle de la chaîne (1)**

Soient $f \in \mathcal{C}^1(\mathcal{O}, \mathbf{R})$, où \mathcal{O} est un ouvert de \mathbf{R}^n et $x_1, \dots, x_n : \mathcal{O} \rightarrow \mathbf{R}$ n fonctions de $\mathcal{C}^1(I, \mathbf{R})$ ¹², telles que $(x_1(t), \dots, x_n(t)) \in \mathcal{O}$ pour tout $t \in I$ avec I intervalle réel.

Alors $g \left| \begin{array}{l} I \rightarrow \mathbf{R} \\ t \mapsto f(x_1(t), \dots, x_n(t)) \end{array} \right.$ est de classe \mathcal{C}^1 et

$$\forall t \in I, \quad g'(t) = \sum_{i=1}^n x'_i(t) \partial_i f(x_1(t), \dots, x_n(t)).$$
¹³

Preuve Admis.

Corollaire ANA.8.4 | Règle de la chaîne (1) – Cas $n = 2$

Soient $f \in \mathcal{C}^1(\mathcal{O}, \mathbf{R})$, où \mathcal{O} est un ouvert de \mathbf{R}^2 et $x, y : \mathcal{O} \rightarrow \mathbf{R}$ deux fonctions de $\mathcal{C}^1(I, \mathbf{R})$, telles que $(x(t), y(t)) \in \mathcal{O}$ pour tout $t \in I$ avec I intervalle réel. Alors

$g \left| \begin{array}{l} I \rightarrow \mathbf{R} \\ t \mapsto f(x(t), y(t)) \end{array} \right.$ est de classe \mathcal{C}^1 et

$$\forall t \in I, \quad g'(t) = x'(t) \frac{\partial f}{\partial x}(x(t), y(t)) + y'(t) \frac{\partial f}{\partial y}(x(t), y(t)).$$

Preuve Appliquer la formule précédente dans le cas $n = 2$.

Exemple 21 – Soit $f : \mathbf{R}^2 \rightarrow \mathbf{R}$ une fonction \mathcal{C}^1 . Calculer $\frac{d}{dt} (f(t^2, \sqrt{t}))$ pour tout $t > 0$.

**3.3.1. Dérivation de composées**

¹²Encore une fois, les formules ci-dessous ne requièrent aucunement l'hypothèse \mathcal{C}^1 en vérité, mais nous les énonçons ainsi conformément au programme

¹³Remarquons que pour $n = 1$, on retrouve la formule de dérivation d'une composée pour les fonctions d'une variable.

Théorème ANA.8.3 | Règle de la chaîne (2) – Cas $n = 2$


Soient $\mathcal{O}_1, \mathcal{O}_2$ deux ouverts de \mathbf{R}^2 , et $x : \mathcal{O}_1 \rightarrow \mathbf{R}$, $y : \mathcal{O}_1 \rightarrow \mathbf{R}$, $f : \mathcal{O}_2 \rightarrow \mathbf{R}$, trois applications \mathcal{C}^1 telles que : $\forall (u, v) \in \mathcal{O}_1, (x(u, v), y(u, v)) \in \mathcal{O}_2$.

Alors $g \left| \begin{array}{l} \mathcal{O}_1 \rightarrow \mathbf{R} \\ (u, v) \rightarrow f(x(u, v), y(u, v)) \end{array} \right.$ est de classe \mathcal{C}^1 et

$$\frac{\partial g}{\partial u}(u, v) = \frac{\partial x}{\partial u}(u, v) \frac{\partial f}{\partial x}(x(u, v), y(u, v)) + \frac{\partial y}{\partial u}(u, v) \frac{\partial f}{\partial y}(x(u, v), y(u, v)),$$

$$\frac{\partial g}{\partial v}(u, v) = \frac{\partial x}{\partial v}(u, v) \frac{\partial f}{\partial x}(x(u, v), y(u, v)) + \frac{\partial y}{\partial v}(u, v) \frac{\partial f}{\partial y}(x(u, v), y(u, v)).$$

Preuve Nous admettons cette deuxième formule.

Exemple 22 – Cas du changement polaire Si $f : \mathbf{R}^2 \rightarrow \mathbf{R}$ est de classe \mathcal{C}^1 , montrer que $f^{\text{pol}} : (r, \theta) \in \mathbf{R}^+ \times]0, 2\pi[\rightarrow f(r \cos \theta, r \sin \theta)$ est aussi de classe \mathcal{C}^1 et calculer $\frac{\partial}{\partial r}(f^{\text{pol}}(r, \theta)), \frac{\partial}{\partial \theta}(f^{\text{pol}}(r, \theta))$ en fonction des dérivées partielles de f , pour tous $(r, \theta) \in \mathbf{R}^+ \times]0, 2\pi[$. 

3.4. Fonctions \mathcal{C}^2

On se restreint dans cette partie aux fonctions de deux variables ($n = 2$) pour simplifier, même si l'ensemble des résultats peuvent être étendus aux fonctions de n variables.

Définition ANA.8.13 | Dérivées d'ordre supérieure

Soit $f : \mathcal{O} \rightarrow \mathbf{R}$ où \mathcal{O} est un sous-ensemble ouvert de \mathbf{R}^2 .

1. Supposons que f admette une dérivée partielle par rapport à x sur \mathcal{O} , et soit $A \in \mathcal{O}$.

▶ si $\frac{\partial f}{\partial x}$ possède une dérivée partielle au point A par rapport à sa première variable, on pose :

$$\frac{\partial^2 f}{\partial x^2}(A) = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right) (A),$$

▶ si $\frac{\partial f}{\partial x}$ possède une dérivée partielle par rapport à y en A , on pose :

$$\frac{\partial^2 f}{\partial y \partial x}(A) = \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) (A).$$

2. On définit de la même manière $\frac{\partial^2 f}{\partial x \partial y}(A) = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) (A)$ et $\frac{\partial^2 f}{\partial y^2}(A) = \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial y} \right) (A)$.

Définition ANA.8.14 | Classe \mathcal{C}^2 , cas $n = 2$

Soit $f : \mathcal{O} \rightarrow \mathbf{R}$ où \mathcal{O} est un sous-ensemble ouvert de \mathbf{R}^2 . On dit que f est de classe \mathcal{C}^2 sur \mathcal{O} si toutes ses dérivées partielles secondes sont définies et continues sur \mathcal{O} .

Notation

On note $\mathcal{C}^2(\mathcal{O}, \mathbf{R})$ l'ensemble des fonctions de classe \mathcal{C}^2 sur \mathcal{O} .


Pour finir, un théorème qui sera lui aussi admis (la démonstration est difficile) : on peut permuter l'ordre des dérivées dans les dérivées croisées dès que la fonction est \mathcal{C}^2 .


Théorème ANA.8.4 | SCWHARZ

Soient \mathcal{O} un ouvert de \mathbf{R}^2 et $f \in \mathcal{C}^2(\mathcal{O}, \mathbf{R})$. Alors :


$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}.$$


Preuve Admis.

Exemple 23 — Vérifier le résultat du théorème de SCWHARZ sur la fonction $f : (x, y) \in \mathbf{R}^2 \mapsto x^3 - y^3 x + \sqrt{1 + x^2 + y^2}$ après avoir précisé son ensemble de définition. 

Exemple 24 — *Cas du changement polaire* Si $f : \mathbf{R}^2 \rightarrow \mathbf{R}$ est de classe \mathcal{C}^2 , montrer que f^{pol} est aussi de classe \mathcal{C}^2 et calculer $\frac{\partial^2}{\partial r^2}(f^{\text{pol}}(r, \theta))$ en fonction des dérivées partielles secondes de f , pour tous $(r, \theta) \in \mathbf{R}^+ \times [0, 2\pi[$. *Indication* : On pourra se servir des dérivées partielles d'ordre un, qui ont déjà été calculées. 

Exemple 25 — Primitivation partielle seconde

1. Déterminer les fonctions f de classe \mathcal{C}^2 sur \mathbf{R} telles que $\frac{\partial^2 f}{\partial x^2} = 0$. 

2. Déterminer les fonctions f de classe \mathcal{C}^2 sur \mathbf{R} telles que $\frac{\partial^2 f}{\partial x \partial y} = 0$. 

3.5. Extrema

Pour terminer, nous revenons à l'une des motivations initiales : l'optimisation de fonctions de plusieurs variables, *i.e.* la recherche de minimums et de maximums.

Définition ANA.8.15 | Extrema

► **(Global)** On dit qu'une fonction $f : P \rightarrow \mathbf{R}$ admet en $A \in P$ un *minimum* (*resp. maximum*) si :

$$\forall X \in P, \quad f(X) \geq f(A)$$

(*resp.* $\forall X \in P, \quad f(X) \leq f(A)$),

ou de manière équivalente :

$$\forall H \in \mathbf{R}^n, A + H \in P \quad f(A + H) \geq f(A)$$

(*resp.* $\forall H \in \mathbf{R}^n, A + H \in P, \quad f(A + H) \leq f(A)$), ¹⁴

► **(Local)** On dit que f admet en $A = (a_0, \dots, a_n) \in P$ un *minimum local* (*resp. maximum local*) si l'une des égalités précédentes a lieu uniquement sur un voisinage ouvert de A .¹⁵ On dit que f admet en A un *extremum* (*resp. extremum local*) si f admet en A un minimum ou un maximum (*resp.* un minimum local ou un maximum local).

Pour les fonctions dérivables d'une variable, on sait qu'en tout extremum (même local) situé dans l'intérieur du domaine de définition, la dérivée première est nulle (on avait parlé de point critique). Voici l'analogie pour les fonctions de plusieurs variables.

Théorème ANA.8.5 | Condition nécessaire pour être un extremum local

Soient $f : \mathcal{O} \rightarrow \mathbf{R}$ une application, \mathcal{O} un ouvert, et $C \in \mathcal{O}$. On suppose que :

1. f admet en C un extremum local,
2. f admet des dérivées partielles dans toutes les directions en C .

Alors :

C est un point critique de f .

Preuve Utiliser le résultat déjà montré pour les fonctions d'une variable à chaque dérivée partielle de f .

Attention

Il s'agit d'une condition **nécessaire**. En effet, si nous considérons la fonction $f : (x, y) \in \mathbf{R}^2 \rightarrow x^2 - y^2$. Les points critiques $(x, y) \in \mathbf{R}^2$ sont les solutions du système :

$$\begin{cases} 2x = 0, \\ 2y = 0, \end{cases} \iff (x, y) = (0, 0).$$

Et $(0, 0)$ n'est ni un maximum, ni un minimum local (et *a fortiori* global), en effet : pour tout $\varepsilon > 0$,

¹⁴Il est souvent plus visuel d'utiliser la forme en « $A + H$ » que celle « en X » car on préfère faire tendre des variables vers 0 que vers A .

¹⁵la seule différence entre les deux définitions portent donc sur l'ensemble où l'inégalité est vraie. Dans le second, uniquement sur un petit domaine autour du point.

- ✘
- ▶ $f(\varepsilon, \sqrt{\varepsilon}) - f(0, 0) = \varepsilon^2 - \varepsilon = \varepsilon(\varepsilon - 1)$ est du signe de $-\varepsilon$ lorsque ε tend vers zéro,
 - ▶ $f(\sqrt{\varepsilon}, \varepsilon) - f(0, 0) = \varepsilon - \varepsilon^2 = \varepsilon(1 - \varepsilon)$ est du signe de $+\varepsilon$ lorsque ε tend vers zéro.
- Ceci prouve que $(0, 0)$ n'est pas un *extrema* local. ✎

On peut le constater sur un dessin.

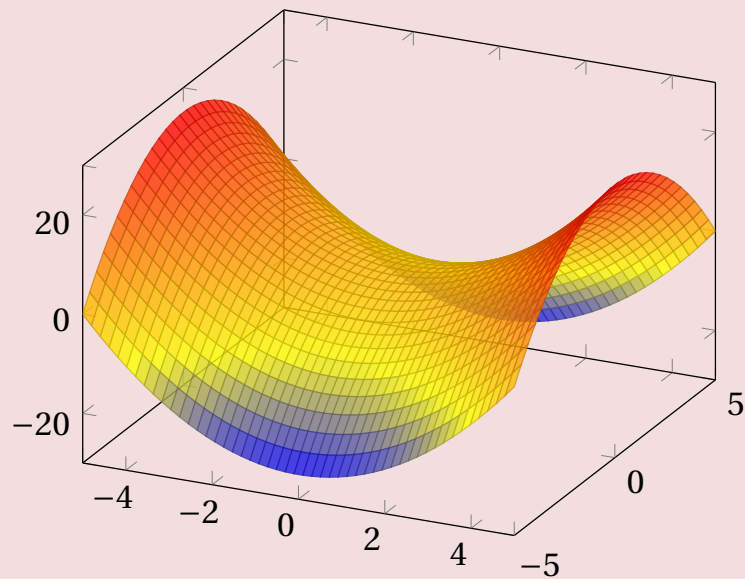


FIG. ANA.8.3. : Exemple de *point selle* en $(0, 0)$: ni maximum, ni minimum en ce point

Preuve Appliquer la preuve du [Chapter ANA.7](#) pour chacune des dérivées partielles.

Remarque 3.9 — Il existe des méthodes pour déterminer la nature d'un point critique mais qui ne sont pas à notre programme. Tout exercice sur le sujet sera donc guidé et consistera en la vérification de la [Définition ANA.8.15](#).

Exemple 26 —

1. Déterminer les points critiques de la fonction $f : (x, y) \mapsto x^3 - 3x + y^2$. ✎

2. Déterminer les points critiques de la fonction $g : (x, y) \mapsto x^2 - 4x + y^3 - 3y$. ✎

Exemple 27 — Retour sur la régression linéaire On considère $(x_i, y_i)_{1 \leq i \leq n}$ avec $n \geq 1$ est un nuage de n points de \mathbf{R}^2 . On rappelle que dans le problème de recherche de droite des moindres carrés, on souhaite minimiser la quantité définie ci-après. Déterminer l'unique point critique (a^*, b^*) de la fonction $F : \mathbf{R}^2 \rightarrow \mathbf{R}$:

$$\forall (a, b) \in \mathbf{R}^2, \quad F(a, b) = \sum_{i=1}^n (y_i - ax_i - b)^2.$$

On rappelle les définitions des quantités statistiques usuelles suivantes :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \overline{x^2} = \frac{1}{n} \sum_{i=1}^n x_i^2, \quad \mathbf{V}_x = \overline{x^2} - \bar{x}^2 = \sigma_x^2, \quad \mathbf{C}_{x,y} = \frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \bar{y}.$$

On montrera que :

$$a^* = \frac{\mathbf{C}_{x,y}}{\sigma_x^2}, \quad b^* = \bar{y} - a^* \bar{x}.$$



Remarque 3.10 — Nous referons appel à ce résultat dans le [Chapter ALEA.17](#), et nous montrerons que (a^*, b^*) est un minimum global de F .

*** **Fin du chapitre** ***

4. EXERCICES

4.1. Généralités

Exercice ANA.8.1 | Soit f la fonction définie par $f(x, y) = \ln(2x + x^2 + y^2)$, pour tout (x, y) dans un ensemble à préciser dans la suite.

- Déterminer l'ensemble de définition de f , et le représenter graphiquement.
- Montrer que les lignes de niveau de f sont des cercles dont on donnera le centre et le rayon.

Solution (exercice ANA.8.1)

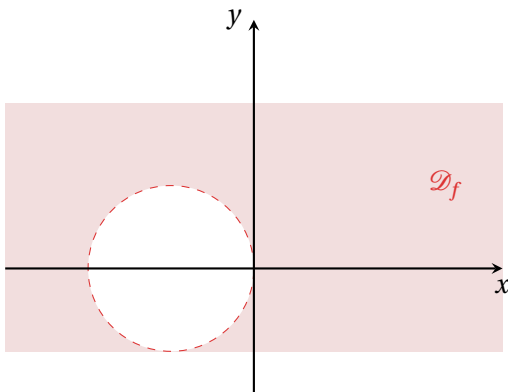
- La fonction f est définie pour tout $(x, y) \in \mathbb{R}^2$ tel que

$$2x + x^2 + y^2 > 0 \iff (x + 1)^2 + y^2 > 1,$$

en mettant le premier polynôme sous forme canonique. Donc l'ensemble de définition de f est

$$\mathcal{D}_f = \{(x, y) \in \mathbb{R}^2, (x + 1)^2 + y^2 > 1\}.$$

L'ensemble $(x + 1)^2 + y^2 > 1$ est le cercle de centre $(-1, 0)$ et de rayon 1, donc \mathcal{D}_f est la partie extérieure strictement à ce cercle (en rouge sur le dessin)



- Soit $k \in \mathbb{R}$, alors

$$\begin{aligned} \ln(2x + x^2 + y^2) &= k, \\ 2x + x^2 + y^2 &= e^k, \\ (x + 1)^2 + y^2 &= e^k + 1. \end{aligned}$$

Donc la ligne de niveau k est le cercle de centre $(-1, 0)$ et de rayon $\sqrt{e^k - 1}$.

Exercice ANA.8.2 | Analyse d'une surface représentative Des espèces d'insectes ont été inventoriées dans plusieurs milieux sur une colline calcaire, selon une échelle de dynamique végétale v depuis la pelouse sèche (valeur 1) jusqu'au sous-bois (valeur 6) et la hauteur h en cm de la végétation herbacée (entre 5 et 40 cm). Le nombre NB d'espèces d'insectes est modélisé par la fonction

$$f(v, h) = -0.466v^2 + 2.960v - 0.00655h^2 + 0.34625h + 1.08725,$$

de la variable $(v, h) \in \mathbb{R}^2$. On peut visualiser sa surface représentative.

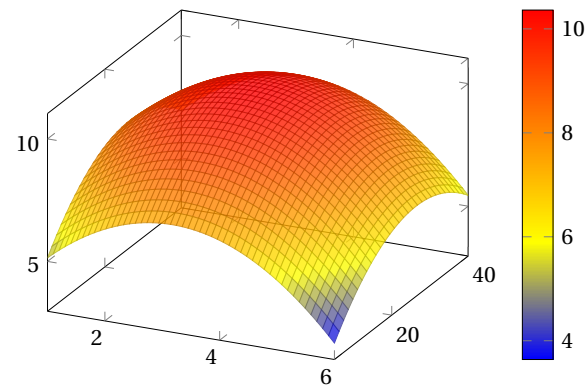


FIG. ANA.8.4. : Graphe de f

- Légender la figure.

2. Quelle est la valeur du milieu où l'on attend la plus grande richesse en espèce d'insectes? Quel est le NB maximal associé?
3. En un point quelconque $M(v, h), (v, h) \in \mathbf{R}^2$, quelle est la direction $V \in \mathbf{R}^2$ dans laquelle la fonction f «varie le plus vite»? (i.e. celle où $|\frac{\partial f}{\partial V}|$ est maximale) Quelle est la valeur $\Delta(v, h)$ de cette variation maximale?

Solution (exercice ANA.8.2)

1. L'axe des abscisses correspond donc à l'indice du milieu (entre 0 et 6), celui des ordonnées à la hauteur de végétation herbacée (entre 20 et 40), et NB se lit sur la cote.
2. L'indice maximal du milieu correspond à NB maximal semble être $v = 3, NB_{max} \approx 12$. Quelle est la valeur du milieu où l'on attend la plus grande richesse en espèce d'insectes? Quel est le NB maximal associé?
3. Soit $M(v, h), (v, h) \in \mathbf{R}^2$. D'après le cours, on sait que la direction cherchée est donnée par le gradient au point M. Or,

$$\frac{\partial f}{\partial v}(M) = -0.932v + 2.960, \quad \frac{\partial f}{\partial h}(M) = -0.0131h + 0.34625.$$

Donc au point M, f varie le plus vite dans la direction $V = (-0.932v + 2.960, -0.0131h + 0.34625)$.

La valeur de la variation maximale est alors

$$\left| \frac{\partial f}{\partial V}(M) \right|.$$

D'après le cours nous savons que cette variation maximale est

$$\|\text{grad} f(M)\| = \sqrt{(-0.932v + 2.960)^2 + (-0.0131h + 0.34625)^2}.$$

4.2. Dérivation & Primitivation

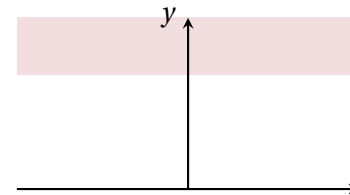
Exercice ANA.8.3 | Dérivabilité partielle Pour chacune des fonctions ci-dessous f :

$\mathbf{R}^2 \rightarrow \mathbf{R}$, déterminer le domaine de définition et le représenter, puis calculer les dérivées partielles d'ordre 1.

1. $f : (x, y) \mapsto \ln(2 - y) + e^x \sin(y) - \sqrt{x^2 + 1}$,
2. $f : (x, y) \mapsto \frac{\ln(1-x)\ln(y-2)}{\sqrt{2x+y-1}}$,
3. $f : (x, y) \mapsto \cos(x) \cos(y) \cos(y - x)$,
4. $f : (x, y) \mapsto \sqrt{2 - x^2 - y^2} + 2x - 2y + \ln(x^2 y)$,
5. $f : (x, y) \mapsto \arctan\left(\frac{x^2}{1+y^2}\right)$.

Solution (exercice ANA.8.3)

1. $\mathcal{D}_f = \{(x, y) \in \mathbf{R}^2, y > 2\}$. C'est donc le demi-plan supérieur à la droite $y =$



2. Calculons à présent les dérivées partielles : $\forall (x, y) \in \mathcal{D}_f,$

$$\frac{\partial f}{\partial x}(x, y) = \sin(y) - \frac{2x}{2\sqrt{x^2 + 1}},$$

$$\frac{\partial f}{\partial y}(x, y) = -\frac{1}{2-y} + e^x.$$

2. $\mathcal{D}_f = \{(x, y) \in \mathbf{R}^2, x < 1, y > 2, y > 1 - 2x\}$. Calculons à présent les dérivées par-

telles : $\forall (x, y) \in \mathcal{D}_f$,

$$\frac{\partial f}{\partial x}(x, y) = \ln(y-2) \frac{-\frac{1}{1-x} \sqrt{2x+y-1} - \ln(1-x) \frac{2}{2\sqrt{2x+y-1}}}{2x+y-1},$$

$$\frac{\partial f}{\partial y}(x, y) = \ln(1-x) \frac{\frac{1}{y-2} \sqrt{2x+y-1} - \ln(y-2) \frac{1}{2\sqrt{2x+y-1}}}{2x+y-1}.$$

3. ...

Exercice ANA.8.4 | Primitivation partielle

1. Déterminer les fonctions f de classe \mathcal{C}^1 sur \mathbf{R}^2 et vérifiant :

$$\forall (x, y) \in \mathbf{R}^2, \quad \begin{cases} \frac{\partial f}{\partial x}(x, y) = \sin y e^{x \sin y}, \\ \frac{\partial f}{\partial y}(x, y) = (x e^{x \sin y} + y) \cos y. \end{cases}$$

2. Déterminer les fonctions f de classe \mathcal{C}^1 sur \mathbf{R}^2 et vérifiant :

$$\forall (x, y) \in \mathbf{R}^2, \quad \begin{cases} \frac{\partial f}{\partial x}(x, y) = 2x \ln(x^2 + 1), \\ \frac{\partial f}{\partial y}(x, y) = -2f(x, y) + 2(x^2 + 1) \ln(x^2 + 1) - 2x^2 + y^2. \end{cases}$$

Solution (exercice ANA.8.4)

Notons à chaque fois (S) le système de dérivées partielles.

1. Nous aurons besoin dans la suite de déterminer une primitive de $y \mapsto y \cos y$.
Faisons donc une intégration par parties. Pour tout $y \in \mathbf{R}$,

$$\begin{aligned} \int_0^y t \cos(t) dt &= - \int_0^y \sin(t) dt + [t \sin t]_0^y \\ &= \cos y - 1 + y \sin y, \end{aligned}$$

puisque les fonctions $t \mapsto t, t \mapsto \sin t$ sont \mathcal{C}^1 . Soit maintenant $(x, y) \in \mathbf{R}^2$.

$$(S) \iff \begin{cases} \frac{\partial f}{\partial x}(x, y) = \sin y e^{x \sin y}, \\ \frac{\partial f}{\partial y}(x, y) = (x e^{x \sin y} + y) \cos y. \end{cases}$$

$$\iff \begin{cases} f(x, y) = e^{x \sin y} + K(y), \\ f(x, y) = e^{x \sin y} + \cos y - 1 + y \sin y + L(x) \end{cases}, \quad K, L \text{ dérivables.}$$

En égalisant les deux lignes, on obtient :

$$\forall (x, y) \in \mathbf{R}^2, \quad K(y) - \cos y + 1 - y \sin y = L(x).$$

Or, une fonction de x ne peut être égale à une fonction de y que si les deux fonctions sont constantes. En effet, si f, g désignent deux fonctions d'une variable telles que pour tout $(x, y) \in \mathbf{R}^2, f(x) = g(y)$. Alors pour tout $x \in \mathbf{R}, f(x) = g(0)$ (faire $y = 0$), et pour tout $y \in \mathbf{R}, f(0) = g(y)$ (faire $x = 0$). Donc f, g sont constantes et nécessairement valent la même constante car $f(0) = g(0)$. Utilisant ceci, nous déduisons l'existence de $M \in \mathbf{R}$ telle que

$$\forall (x, y) \in \mathbf{R}^2, \quad K(y) = M + \cos y - 1 + y \sin y, \quad L(x) = M.$$

Et donc

$$\begin{cases} f(x, y) = e^{x \sin y} + M + \cos y - 1 + y \sin y, \\ f(x, y) = e^{x \sin y} + M + \cos y - 1 + y \sin y. \end{cases}$$

L'ensemble des solutions est donc

$$\left\{ (x, y) \mapsto e^{x \sin y} + M + \cos y - 1 + y \sin y, M \in \mathbf{R} \right\}.$$

2. Commençons par déterminer une primitive de $x \mapsto 2x \ln(x^2 + 1)$. Faisons donc

une intégration par parties. Pour tout $x \in \mathbf{R}$,

$$\begin{aligned} \int_0^x (2t) \ln(t^2 + 1) dt &= -2 \int_0^x \frac{t^2}{2} \frac{2t}{t^2 + 1} dt + 2 \left[\frac{t^2}{2} \ln(t^2 + 1) \right]_0^x \\ &= -2 \int_0^x \frac{t^3}{t^2 + 1} dt + x^2 \ln(x^2 + 1) \\ &= -2 \int_0^x \frac{t(t^2 + 1) - t}{t^2 + 1} dt + x^2 \ln(x^2 + 1) \\ &= -2 \left[\frac{t^2}{2} - \frac{1}{2} \ln(t^2 + 1) \right]_0^x + x^2 \ln(x^2 + 1) \\ &= \ln(x^2 + 1) - x^2 + x^2 \ln(x^2 + 1), \end{aligned}$$

puisque les fonctions $t \mapsto \frac{t^2}{2}, t \mapsto \ln(t^2 + 1)$ sont \mathcal{C}^1 . Soit maintenant $(x, y) \in \mathbf{R}^2$.

$$\begin{aligned} \text{(S)} \iff & \begin{cases} \frac{\partial f}{\partial x}(x, y) &= 2x \ln(x^2 + 1), \\ \frac{\partial f}{\partial y}(x, y) &= -2f(x, y) + 2(x^2 + 1) \ln(x^2 + 1) - 2x^2 + y^2 \end{cases} \\ \iff & \begin{cases} f(x, y) &= \ln(x^2 + 1) - x^2 + x^2 \ln(x^2 + 1) + K(y), \\ \frac{\partial f}{\partial y}(x, y) &= -2f(x, y) + 2(x^2 + 1) \ln(x^2 + 1) - 2x^2 + y^2 \end{cases} \quad \text{K dérivable} \\ \iff & \begin{cases} f(x, y) &= \ln(x^2 + 1) - x^2 + x^2 \ln(x^2 + 1) + K(y), \\ \frac{\partial f}{\partial y}(x, y) &= -2f(x, y) + 2(x^2 + 1) \ln(x^2 + 1) - 2x^2 + y^2 \end{cases} \quad \text{K dérivable.} \end{aligned}$$

En injectant la ligne 1 dans la ligne 2, on trouve

$$\begin{aligned} K'(y) &= -2(\ln(x^2 + 1) - x^2 + x^2 \ln(x^2 + 1) + K(y)) + 2(x^2 + 1) \ln(x^2 + 1) - 2x^2 + y^2 \\ \iff K'(y) &= -2K(y) + y^2. \end{aligned}$$

Il s'agit d'une équation différentielle d'ordre 1, donc $K(y) = \text{Le}^{-2y} + \widehat{K}(y)$, où \widehat{K} est une solution particulière de l'équation avec second membre. On cherche cette solution particulière par variation de la constante, on trouve $\widehat{K}(y) = -\frac{1}{2}(y - 1)y + \frac{1}{4}$. Donc

$$K(y) = \text{Le}^{-2y} - \frac{1}{2}(y - 1)y + \frac{1}{4},$$

on peut enfin conclure sur l'ensemble des solutions de (S) :

$$\left\{ (x, y) \mapsto \ln(x^2 + 1) - x^2 + x^2 \ln(x^2 + 1) + \text{Le}^{-2y} - \frac{1}{2}(y - 1)y + \frac{1}{4}, L \in \mathbf{R} \right\}.$$

4.3. Règle de la chaîne et équation aux dérivées partielles

Exercice ANA.8.5 | Soit V une application de classe \mathcal{C}^1 sur \mathbf{R}^2 . Calculer la dérivée de la fonction Φ définie par

$$\forall t > 0, \quad \Phi(t) = V(t^2 + 1, e^t + \ln t - 1),$$

après avoir justifié la dérivabilité.

Solution (exercice ANA.8.5)

La fonction V est de classe \mathcal{C}^1 et $t \mapsto t^2 + 1, t \mapsto e^t + \ln t - 1$ sont de classe \mathcal{C}^1 sur \mathbf{R}^{++} . La règle de la chaîne nous dit alors que Φ est dérivable. De plus, pour tout $t > 0$,

$$\Phi'(t) = 2t \frac{\partial V}{\partial x}(t^2 + 1, e^t + \ln t - 1) + \left(e^t + \frac{1}{t} \right) \frac{\partial V}{\partial y}(t^2 + 1, e^t + \ln t - 1).$$

Exercice ANA.8.6 | **EDP et changement affine** On cherche à déterminer l'ensemble E des fonctions f de classe \mathcal{C}^1 de \mathbf{R}^2 dans \mathbf{R} , telles que :

$$\forall (x, y) \in \mathbf{R}^2, \quad \frac{\partial f}{\partial x}(x, y) = \frac{\partial f}{\partial y}(x, y). \quad \text{(E)}$$

- On pose $u = x + y$ et $v = x - y$. Exprimer x et y comme fonctions de u et v . Ces deux fonctions seront notées $(u, v) \in \mathbf{R}^2 \mapsto x(u, v)$ et $(u, v) \in \mathbf{R}^2 \mapsto y(u, v)$ dans la suite.
- On pose $g(u, v) = f(x(u, v), y(u, v))$ pour tout $(u, v) \in \mathbf{R}^2$. Exprimer les dérivées partielles premières de f en fonction de celles de g .

- Donner une condition sur la fonction g qui équivaut à (E). Déterminer les fonctions g qui satisfont cette condition.
- En déduire l'ensemble des solutions de (E).

Solution (exercice ANA.8.6)

- On pose $\begin{cases} u = x + y, \\ v = x - y, \end{cases}$ il s'agit d'un système linéaire que l'on résout très facilement :

$$u + v = 2x, \quad u - v = 2y,$$

On déduit alors : $x = \frac{u+v}{2}, \quad y = \frac{u-v}{2}.$

- On pose $g(u, v) = f(x(u, v), y(u, v)) = f\left(\frac{u+v}{2}, \frac{u-v}{2}\right)$. De manière équivalente, nous avons :

$$\forall x, y \in \mathbf{R}, \quad f(x, y) = g(x + y, x - y).$$

Alors, d'après la formule de la chaîne nous avons pour tout $(x, y) \in \mathbf{R}^2$:

$$\frac{\partial f}{\partial x}(x, y) = \frac{\partial g}{\partial u}(x, y) + \frac{\partial g}{\partial v}(x, y),$$

$$\frac{\partial f}{\partial y}(x, y) = \frac{\partial g}{\partial u}(x, y) - \frac{\partial g}{\partial v}(x, y).$$

- Alors f est solution de (E) si et seulement si :

$$\forall (x, y) \in \mathbf{R}^2, \quad \frac{\partial g}{\partial u}(x, y) + \frac{\partial g}{\partial v}(x, y) = \frac{\partial g}{\partial u}(x, y) - \frac{\partial g}{\partial v}(x, y),$$

ceci est encore équivalent à

$$\frac{\partial g}{\partial v} = 0.$$

Cette équation est un problème de primitivation partielle classique : g est alors une fonction de u , il existe une fonction K de classe \mathcal{C}^1 telle que

$$\forall (u, v) \in \mathbf{R}^2, \quad g(u, v) = K(u).$$

- On peut ensuite conclure sur les solutions en f :

$$\forall (x, y) \in \mathbf{R}^2, \quad f(x, y) = g(x + y, x - y) = K(x + y).$$

Exercice ANA.8.7 | EDP et changement polaire Résoudre à l'aide des coordonnées polaires l'équation les équations aux dérivées partielles dans l'ensemble des fonctions \mathcal{C}^1 :

- $x \frac{\partial f}{\partial y} - y \frac{\partial f}{\partial x} = 0$ sur $\{(x, y) \in \mathbf{R}^2, (x, y) \neq (0, 0)\}$.
- $x \frac{\partial f}{\partial x}(x, y) + y \frac{\partial f}{\partial y}(x, y) = x^2 + y^2$, sur $\{(x, y) \in \mathbf{R}^2, (x, y) \neq (0, 0), y > 0\}$.

Indication : On considèrera la fonction $g : (r, \theta) \in [0, 2\pi[\times \mathbf{R}^+ \rightarrow f(r \cos \theta, r \sin \theta)$.

Solution (exercice ANA.8.7)

D'après la règle de la chaîne, la fonction g est de classe \mathcal{C}^1 , et

$$\frac{\partial g}{\partial r}(r, \theta) = \cos \theta \frac{\partial f}{\partial x}(r \cos \theta, r \sin \theta) + \sin \theta \frac{\partial f}{\partial y}(r \cos \theta, r \sin \theta)$$

$$\frac{\partial g}{\partial \theta}(r, \theta) = -r \sin \theta \frac{\partial f}{\partial x}(r \cos \theta, r \sin \theta) + r \cos \theta \frac{\partial f}{\partial y}(r \cos \theta, r \sin \theta).$$

- Nous avons

$$\forall (x, y) \in \mathbf{R}^2, \quad x \frac{\partial f}{\partial y}(x, y) - y \frac{\partial f}{\partial x}(x, y) = 0$$

$$\forall (r, \theta) \in \mathbf{R}^+ \times [0, 2\pi], \quad (r \cos \theta) \frac{\partial f}{\partial y}(r \cos \theta, r \sin \theta) - (r \sin \theta) \frac{\partial f}{\partial x}(r \cos \theta, r \sin \theta) = 0$$

coordonnées polaires bijectives, on remplace $x = r \cos \theta, y = r \sin \theta$

$$- (r \sin \theta) \frac{\partial f}{\partial x}(r \cos \theta, r \sin \theta) = 0$$

calculs précédents

$$\forall (r, \theta) \in \mathbf{R}^+ \times [0, 2\pi], \quad \frac{\partial g}{\partial \theta}(r, \theta) = 0$$

$$\forall (r, \theta) \in \mathbf{R}^+ \times [0, 2\pi], \quad g(r, \theta) = K(r), \quad K \text{ dérivable,}$$

$$\forall (x, y) \in \mathbf{R}^2, (x, y) \neq (0, 0), \quad f(x, y) = K\left(\sqrt{x^2 + y^2}\right), \quad K \text{ dérivable.}$$

$\begin{matrix} x = r \cos \theta, \\ y = r \sin \theta \end{matrix}$ donne $x^2 + y^2 = r^2$

2. Nous avons

$$\forall (x, y) \in \mathbf{R}^2, \quad x \frac{\partial f}{\partial x}(x, y) + y \frac{\partial f}{\partial y}(x, y) = x^2 + y^2$$

$$\forall (r, \theta) \in \mathbf{R}^+ \times [0, 2\pi], \quad r \cos \theta \frac{\partial f}{\partial x}(r \cos \theta, r \sin \theta) + r \sin \theta \frac{\partial f}{\partial y}(x, y) = r^2$$

$$\forall (r, \theta) \in \mathbf{R}^+ \times [0, 2\pi], \quad r \frac{\partial g}{\partial r}(r, \theta) = r^2$$

$$\forall (r, \theta) \in \mathbf{R}^{+\ast} \times [0, 2\pi], \quad \frac{\partial g}{\partial r}(r, \theta) = r$$

$$\forall (r, \theta) \in \mathbf{R}^{+\ast} \times [0, 2\pi], \quad g(r, \theta) = \frac{r^2}{2} + K(\theta), \quad K \text{ dérivable,}$$

$$\forall (x, y) \in \mathbf{R}^2, (x, y) \neq (0, 0), y > 0, \quad \boxed{f(x, y) = \frac{x^2 + y^2}{2} + K\left(\arctan\left(\frac{y}{x}\right)\right)}, \quad K \text{ dérivable.}$$

calculs
ou résout
précédents
sur l'en-
semble des
couples
non nuls,
donc $r \neq 0$

Pour revenir aux coordonnées x, y , nous avons utilisé $x = r \cos \theta, y = r \sin \theta$ qui donne $x^2 + y^2 = r^2, \tan \theta = \frac{y}{x}$ et puisque $y > 0, \theta \in]-\pi/2, \pi/2[$, on peut donc passer à l'arctan pour obtenir $\theta = \arctan\left(\frac{y}{x}\right)$.

Exercice ANA.8.8 | Relation fondamentale de l'analyse pour les fonctions de deux variables Soit $f : \mathbf{R}^2 \rightarrow \mathbf{R}$ une fonction \mathcal{C}^1 . On note $g : t \in [0, 1] \rightarrow f(tx_1, tx_2)$ pour tous $(x_1, x_2) \in \mathbf{R}^2$.

1. Justifier que g est dérivable et que : $g(1) = g(0) + \int_0^1 g'(t) dt$.

2. En déduire que :

$$f(x_1, x_2) = f(0, 0) + x_1 \int_0^1 \frac{\partial f}{\partial x}(tx_1, tx_2) dt + x_2 \int_0^1 \frac{\partial f}{\partial y}(tx_1, tx_2) dt.$$

Solution (exercice ANA.8.8)

1. Puisque $t \mapsto tx_i$ sont de classe \mathcal{C}^1 pour $i \in \llbracket 1, 2 \rrbracket$, et que f est de classe \mathcal{C}^1 , g est également de classe \mathcal{C}^1 d'après la règle de la chaîne (version I). Donc g' est continue et

$$\int_0^1 g'(t) dt = [g]_0^1 = g(1) - g(0).$$

Donc : $g(1) = g(0) + \int_0^1 g'(t) dt$.

2. Par ailleurs, d'après la règle de la chaîne,

$$g'(t) = x_1 \frac{\partial f}{\partial x}(tx_1, tx_2) + x_2 \frac{\partial f}{\partial y}(tx_1, tx_2),$$

pour tout $t \in [0, 1], g(0) = f(0, 0), g(1) = f(x_1, x_2)$. En combinant les deux questions, on déduit :

$$\boxed{f(x_1, x_2) = f(0, 0) + x_1 \int_0^1 \frac{\partial f}{\partial x} f(tx_1, tx_2) dt + x_2 \int_0^1 \frac{\partial f}{\partial y} f(tx_1, tx_2) dt}.$$

Exercice ANA.8.9 | Une équation aux dérivées partielles de transport/diffusion – Extrait ENS 2020 Dans cette partie, on fixe $m \in \mathbf{R}$, et on s'intéresse aux fonctions u de $[0, 1] \times \mathbf{R}^+$ dans \mathbf{R} , de classe \mathcal{C}^2 , et solutions de l'équation aux dérivées partielles suivante :

$$\frac{\partial u}{\partial t}(x, t) = m \frac{\partial u}{\partial x}(x, t) + \frac{1}{2} \frac{\partial^2 u}{\partial x^2}(x, t), \quad (1)$$

avec conditions de bord $\forall t \geq 0, u(0, t) = u(1, t) = 0$ (2). On cherche les solutions u non identiquement nulles (*i.e.* non identiquement nulles) qui peuvent se décomposer en un produit de deux fonctions d'une seule variable, *i.e.* les solutions dites à *variables séparables*.

Soient donc $f : [0, 1] \rightarrow \mathbf{R}$ et $g : \mathbf{R}^+ \rightarrow \mathbf{R}$ deux fonctions de classe \mathcal{C}^2 . On suppose que la fonction $u : [0, 1] \times \mathbf{R}^+ \rightarrow \mathbf{R}$ définie par $u(x, t) = f(x)g(t)$ ne s'annule pas et est solution de (1) et (2).

1. Vérifier que les fonctions f et g ne s'annulent pas non plus.
2. Montrer qu'il existe un réel λ tel que la fonction g vérifie l'équation différentielle : $\forall t \in \mathbf{R}_+, g'(t) = \lambda g(t)$, puis résoudre cette équation différentielle.
3. Montrer que pour cette même valeur de λ , la fonction f vérifie l'équation différentielle

$$\forall x \in [0, 1], \quad \frac{1}{2}f''(x) + mf'(x) = \lambda f(x),$$

avec condition de bord $f(0) = f(1) = 0$.

4. On suppose dans cette question que $\lambda > -\frac{m^2}{2}$.

4.1) Montrer qu'il existe deux réels a et b tels que :

$$\forall x \in [0, 1], \quad f(x) = ae^{r_+x} + be^{r_-x}, \text{ avec } r_+, r_- \in \mathbf{R} \text{ à préciser.}$$

4.2) Montrer que a et b sont nécessairement nuls, et obtenir une contradiction.

5. En procédant comme dans la question précédente, obtenir une contradiction si l'on suppose $\lambda = -\frac{m^2}{2}$.

6. On suppose dans cette question que $\lambda < -\frac{m^2}{2}$.

6.1) Montrer qu'il existe deux réels a et b tels que :

$$\forall x \in [0, 1], \quad f(x) = (a \cos(\pi \ell x) + b \sin(\pi \ell x))e^{-mx}, \text{ avec } \ell = \frac{1}{\pi} \sqrt{-2\lambda - m^2}.$$

6.2) Montrer que l'on a nécessairement $a = 0, b \neq 0$, et $\ell \in \mathbf{N}^*$.

7. Déterminer tous les couples (f, g) de fonctions \mathcal{C}^2 tels que la fonction $u : (x, t) \mapsto u(x, t) = f(x)g(t)$ ne s'annule pas, et soit solution de (1) et (2).

Solution (exercice ANA.8.9)

1. Si l'une des fonctions f et g est identiquement nulle, alors la fonction u est identiquement nulle. Par contraposée, comme u est supposée non identiquement nulle, chacune des deux fonctions f et g est non identiquement nulle.
2. Comme les fonctions f et g sont supposées de classe \mathcal{C}^2 , la fonction u admet des dérivées partielles du premier et du second ordre sur le pavé $[0, 1] \times \mathbf{R}_+$, et on a :

$$\forall (x, t) \in [0, 1] \times \mathbf{R}_+, \quad \frac{\partial u}{\partial t}(x, t) = f(x)g'(t), \quad \frac{\partial u}{\partial x}(x, t) = f'(x)g(t), \quad \frac{\partial^2 u}{\partial x^2}(x, t) = f''(x)g(t).$$

L'équation aux dérivées partielles (1) s'écrit alors :

$$\forall (x, t) \in [0, 1] \times \mathbf{R}_+, \quad f(x)g'(t) = mf'(x)g(t) + \frac{1}{2}f''(x)g(t) \quad (1').$$

Comme la fonction f est non identiquement nulle, il existe un réel $x_0 \in [0, 1]$ vérifiant $f(x_0) \neq 0$. Divisons membre à membre par $f(x_0)$ dans l'identité ci-dessus :

$$\forall t \in \mathbf{R}_+, \quad g'(t) = \frac{1}{f(x_0)} \left(mf'(x_0) + \frac{1}{2}f''(x_0) \right) g(t).$$

Il suffit de poser $\lambda = \frac{1}{f(x_0)} \left(mf'(x_0) + \frac{1}{2}f''(x_0) \right)$, et g vérifie alors $g'(t) = \lambda g(t)$ pour tout $t \in \mathbf{R}_+$. On sait résoudre cette équation différentielle du premier ordre homogène à coefficients constants :

$$\forall t \in \mathbf{R}_+, \quad g(t) = g(0)e^{\lambda t}.$$

3. On injecte ce dernier résultat dans l'équation (1') :

$$\forall x \in [0, 1], \quad f(x)g(0)\lambda e^{\lambda t} = (mf'(x) + \frac{1}{2}f''(x))g(0)e^{\lambda t}.$$

Si le réel $g(0)$ était nul, la fonction g serait identiquement nulle, ce qui serait en contradiction avec le résultat de la question 1. Ainsi, le réel $g(0)$ n'est pas nul. On peut donc diviser membre à membre par $g(0)e^{\lambda t}$, ce qui donne :

$$\forall x \in [0, 1], \quad \frac{1}{2}f''(x) + mf'(x) = \lambda f(x).$$

De plus, on a

$$f(0) = \frac{u(0,0)}{g(0)} = 0 \quad \text{et} \quad f(1) = \frac{u(1,0)}{g(0)} = 0.$$

4. **4.1)** La fonction f est solution d'une équation différentielle linéaire homogène du second ordre à coefficients constants, d'équation caractéristique associée

$$\frac{1}{2}r^2 + mr - \lambda = 0.$$

Le discriminant du polynôme du second degré associé vaut

$$\Delta = m^2 + 2\lambda.$$

Dans cette question, l'énoncé suppose $\Delta > 0$. L'équation caractéristique admet donc deux racines réelles distinctes

$$r_+ = -m + \sqrt{m^2 + 2\lambda} \quad \text{et} \quad r_- = -m - \sqrt{m^2 + 2\lambda}.$$

Par conséquent, il existe un unique couple (a, b) de réels vérifiant :

$$\forall x \in [0, 1], \quad f(x) = ae^{r_+x} + be^{r_-x}.$$

4.2) Les conditions aux limites $f(0) = 0$ et $f(1) = 0$ fournissent le système :

$$\begin{cases} a + b = 0 \\ e^{r_+}a + e^{r_-}b = 0. \end{cases}$$

Il s'agit d'un système linéaire de deux équations à deux inconnues a et b . Le déterminant de ce système vaut

$$e^{r_-} - e^{r_+} \neq 0$$

puisque les réels r_+ et r_- sont distincts et la fonction exponentielle injective sur \mathbf{R} . Ce système linéaire admet donc une unique solution, qui est la solution triviale $(a, b) = (0, 0)$. Ainsi,

la fonction f est identiquement nulle, ce qui contredit le résultat de la question 1

5. Dans cette question, l'énoncé suppose que l'on a $\Delta = 0$ (voir notation précédente). L'équation caractéristique admet donc une unique racine réelle $r = -m$. Par conséquent, il existe un unique couple (a, b) de réels vérifiant :

$$\forall x \in [0, 1], \quad f(x) = (ax + b)e^{rx}.$$

Les conditions aux limites $f(0) = 0$ et $f(1) = 0$ fournissent le système :

$$\begin{cases} b = 0 \\ (a + b)e^r = 0 \end{cases}$$

qui équivaut à $a = b = 0$. Ainsi,

la fonction f est identiquement nulle, ce qui contredit là encore le résultat de la question 1

6. Dans cette question, l'énoncé suppose que l'on a $\Delta < 0$.

6.1) L'équation caractéristique admet donc deux racines complexes non réelles conjuguées

$$r_+ = -m + i\sqrt{-m^2 - 2\lambda} \quad \text{et} \quad r_- = -m - i\sqrt{-m^2 - 2\lambda}.$$

Par conséquent, il existe un unique couple (a, b) de réels vérifiant :

$$\forall x \in [0, 1], \quad f(x) = \left(a \cos(\sqrt{-m^2 - 2\lambda}x) + b \sin(\sqrt{-m^2 - 2\lambda}x) \right) e^{-mx}.$$

c'est-à-dire, avec le réel ℓ introduit par l'énoncé,

$$\forall x \in [0, 1], \quad f(x) = \left(a \cos(\pi\ell x) + b \sin(\pi\ell x) \right) e^{-mx}.$$

6.2) La condition aux limites $f(0) = 0$ montre que l'on a $a = 0$. Comme la fonction f n'est pas identiquement nulle (question 1), on a nécessairement $b \neq 0$. La condition initiale $f(1) = 0$ montre, puisque b et l'exponentielle de s'annulent pas :

$$\sin(\pi\ell) = 0$$

donc ℓ est un entier relatif, non nul puisque f est non identiquement nulle. Sa définition par $\ell = \frac{1}{\pi}\sqrt{\dots}$ montre que l'entier ℓ est positif. Ainsi, on a

$$\ell \in \mathbf{N}^*.$$

7. On a effectué une *analyse* : s'il existe une fonction $u : (x, t) \rightarrow f(x)g(t)$ de classe \mathcal{C}^2 sur $[0, 1] \times \mathbf{R}_+$ non identiquement nulle vérifiant (1) et (2), alors nécessairement, en notant x_0 un élément de $[0, 1]$ vérifiant $f(x_0) \neq 0$ et $\lambda = \frac{1}{f(x_0)} \left(mf'(x_0)g(t) + \frac{1}{2}f''(x_0) \right)$, on a :

- ▶ $-2\lambda - m^2 > 0$
- ▶ le réel $\ell = \frac{1}{\pi}\sqrt{-2\lambda - m^2}$ est un entier naturel non nul
- ▶ il existe $b \in \mathbf{R}^*$ tel que :

$$\forall x \in [0, 1], \quad f(x) = b \sin(\pi\ell x) e^{-mx}.$$

- ▶ il existe un réel c tel que

$$\forall t \in \mathbf{R}_+, \quad g(t) = ce^{\lambda t}.$$

En d'autres termes, en notant \mathcal{S} l'ensemble des fonctions u de $[0, 1] \times \mathbf{R}_+$ dans \mathbf{R} , s'écrivant sous la forme $u : (x, t) \mapsto f(x)g(t)$ où $f : [0, 1] \rightarrow \mathbf{R}$ et $g : \mathbf{R}_+ \rightarrow \mathbf{R}$ sont deux fonctions de classe \mathcal{C}^2 , solutions de l'équation aux dérivées partielles (1) et vérifiant les conditions aux limites (2), on vient de démontrer l'inclusion :

$$\mathcal{S} \subset \left\{ u : [0, 1] \times \mathbf{R}_+ \rightarrow \mathbf{R} \right. \\ \left. (x, t) \mapsto C \sin(\pi \ell x) \exp\left(-mx - \frac{m^2 + \pi^2 \ell^2}{2} t\right), (\ell, C) \in \mathbf{N}^* \times \mathbf{R}_* \right\}$$

Il ne reste plus qu'à effectuer la *synthèse*, c'est-à-dire examiner l'autre inclusion. On fixe un entier naturel non nul ℓ et un réel non nul C . On pose :

$$f \left| \begin{array}{l} [0, 1] \rightarrow \mathbf{R} \\ x \rightarrow \sin(\pi \ell x) e^{-mx}, \end{array} \right. \quad g \left| \begin{array}{l} \mathbf{R}_+ \rightarrow \mathbf{R} \\ t \rightarrow C \exp\left(-\frac{m^2 + \pi^2 \ell^2}{2} t\right) \end{array} \right.$$

et

$$u : [0, 1] \times \mathbf{R}_+ \rightarrow \mathbf{R} \\ (x, t) \mapsto f(x)g(t) = C \sin(\pi \ell x) \exp\left(-mx - \frac{m^2 + \pi^2 \ell^2}{2} t\right).$$

La fonction u ainsi définie est clairement de classe \mathcal{C}^2 sur le pavé $[0, 1] \times \mathbf{R}_+$. Elle vérifie clairement les conditions aux limites (2). Posons $\lambda = -\frac{m^2 + \pi^2 \ell^2}{2} < -\frac{m^2}{2}$. La fonction g est bien solution de l'équation différentielle :

$$\forall t \in \mathbf{R}_+, \quad g'(t) = \lambda g(t)$$

et d'après le résultat de la question 4, la fonction f est solution de l'équation différentielle linéaire homogène d'ordre 2 à coefficients constants (on peut aussi dériver deux fois la fonction f qui est connue explicitement ici) :

$$\forall x \in [0, 1], \quad \frac{1}{2} f''(x) + m f'(x) = \lambda f(x).$$

En multipliant cette dernière équation membre à membre par $g(t)$, on obtient l'équation (1') :

$$\forall (x, t) \in [0, 1] \times \mathbf{R}_+, \quad f(x)g'(t) = m f'(x)g(t) + \frac{1}{2} f''(x)g(t) \quad (1')$$

D'après ce que l'on a écrit au début de la réponse à la question 2, cette dernière identité implique le fait que la fonction u est solution de l'équation aux dérivées partielles (1). L'autre inclusion est donc établie. Conclusion :

$$\mathcal{S} = \left\{ u : [0, 1] \times \mathbf{R}_+ \rightarrow \mathbf{R} \right. \\ \left. (x, t) \mapsto C \sin(\pi \ell x) \exp\left(-mx - \frac{m^2 + \pi^2 \ell^2}{2} t\right), (\ell, C) \in \mathbf{N}^* \times \mathbf{R}_* \right\}$$

4.4. En Physique

Exercice ANA.8.10 | Gaz parfaits La loi des gaz parfaits peut prendre la forme $PV = nRT$ où n est le nombre de molécules de gaz présentes dans un volume V , à la température T , P étant la pression R une constante. Montrer que

$$\frac{\partial V}{\partial T} \times \frac{\partial T}{\partial P} \times \frac{\partial P}{\partial V} = -1.$$

Solution (exercice ANA.8.10)

Nous avons $PV = RnT$, donc V, T, P sont des fonctions de T, P, V et les dérivées partielles existent bien. Et de plus,

$$\frac{\partial V}{\partial T} = \frac{Rn}{P}, \quad \frac{\partial T}{\partial P} = \frac{V}{Rn}, \quad \frac{\partial P}{\partial V} = -\frac{RnT}{V^2},$$

donc en faisant le produit :

$$\frac{\partial V}{\partial T} \times \frac{\partial T}{\partial P} \times \frac{\partial P}{\partial V} = -1.$$

Exercice ANA.8.11 | Dans un fluide quelconque on définit le coefficient de dilatation isobare α et le coefficient de compressibilité isotherme χ par :

$$\alpha = \frac{1}{V} \left(\frac{\partial V}{\partial T} \right), \quad \chi = -\frac{1}{V} \left(\frac{\partial V}{\partial P} \right),$$

avec les notations du cours de Physique. Montrer que :

$$\frac{\partial \chi}{\partial T} = -\frac{\partial \alpha}{\partial P}.$$

Solution (exercice ANA.8.11)

Supposant que le volume dépende de manière \mathcal{C}^2 de T, P i.e. la température et la pression, on peut appliquer le théorème de SCHWARZ. Ainsi

$$\frac{\partial^2 V}{\partial T \partial P} = \frac{\partial^2 V}{\partial P \partial T},$$

donc

$$\frac{\partial \chi}{\partial T} = -\frac{1}{V} \frac{\partial^2 V}{\partial T \partial P} = -\frac{1}{V} \frac{\partial^2 V}{\partial P \partial T} = \frac{\partial \alpha}{\partial P}.$$

Donc

$$\boxed{\frac{\partial \chi}{\partial T} = -\frac{\partial \alpha}{\partial P}.$$

4.5. Optimisation

Exercice ANA.8.12 | Calculs de points critiques Pour chacune des fonctions suivantes, déterminer les points critiques.

1. $f : (x, y) \in \mathbf{R}^2 \mapsto (x + y)e^{-x^2 - y^2}$,
2. $f : (x, y) \in \mathbf{R}^2 \mapsto xe^y + ye^x$,
3. $f : (x, y) \in \mathbf{R}^2 \mapsto x^4 + y^4 - 4(x - y)^2$,

Solution (exercice ANA.8.12)

1. Soit $(x, y) \in \mathbf{R}^2$. Alors

$$\begin{cases} \frac{\partial f}{\partial x}(x, y) = e^{-x^2 - y^2} + (-2x)(x + y)e^{-x^2 - y^2} = 0 \\ \frac{\partial f}{\partial y}(x, y) = e^{-x^2 - y^2} + (-2y)(x + y)e^{-x^2 - y^2} = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} \frac{\partial f}{\partial x}(x, y) = e^{-x^2 - y^2}(1 - 2x(x + y))e^{-x^2 - y^2} = 0 \\ \frac{\partial f}{\partial y}(x, y) = e^{-x^2 - y^2}(1 - 2y(x + y)) = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} 1 - 2x(x + y) = 0 \\ 1 - 2y(x + y) = 0 \end{cases} \quad (\star)$$

En faisant la différence des deux lignes, on obtient :

$$2(y - x)(x + y) = 0.$$

Donc

- ▶ soit $y = x$, dans ce cas

$$1 - 4x^2 = 0 = (1 - 2x)(1 + 2x) \implies x = \pm \frac{1}{2}.$$

D'où les deux points critiques $(\frac{1}{2}, \frac{1}{2})$ et $(-\frac{1}{2}, -\frac{1}{2})$.

- ▶ soit $y \neq x$, et $y = -x$, ceci est incompatible avec le système (\star) ;

On a donc *in fine* deux points critiques : $(\frac{1}{2}, \frac{1}{2})$ et $(-\frac{1}{2}, -\frac{1}{2})$.

2. Soit $(x, y) \in \mathbf{R}^2$. Alors

$$\begin{cases} \frac{\partial f}{\partial x}(x, y) = e^y + ye^x = 0 \\ \frac{\partial f}{\partial y}(x, y) = xe^y + e^x = 0 \end{cases} \quad (\star)$$

Faisons ensuite $L_2 \leftarrow L_2 - xL_1$, on déduit en regardant L_2 :

$$(1 - xy)e^x = 0.$$

Donc $xy = 1$. Donc $x \neq 0, y \neq 0$, et $y = \frac{1}{x}$. En injectant cette relation dans (\star) , on déduit

$$xe^{\frac{1}{x}} + e^x = 0.$$

Notons $g : x \in \mathbf{R}^{+\star} \mapsto xe^{\frac{1}{x}} + e^x$. La fonction g ne peut s'annuler sur $\mathbf{R}^{+\star}$, car somme de deux fonctions strictement positives sur $\mathbf{R}^{+\star}$. Par ailleurs,

$$g'(x) = e^{\frac{1}{x}} \left(1 - \frac{1}{x}\right) + e^x.$$

Si $x \in \mathbf{R}^{+\star}$, alors $g'(x) > 0$ puisque $1 - \frac{1}{x} > 0$. Donc g est continue strictement croissante sur $\mathbf{R}^{+\star}$, $\lim_{x \rightarrow 0^+} g(x) = 1$ par opérations usuelles sur les limites. Donc g ne peut s'annuler qu'une seule fois sur $\mathbf{R}^{+\star}$, car elle réalise une bijection de $\mathbf{R}^{+\star}$ vers $] -\infty, 1[$ d'après le théorème de la bijection. Mais $g(-1) = 0$, donc $(-1, -1)$ est l'unique point critique.

3. Soit $(x, y) \in \mathbf{R}^2$. Alors

$$\begin{cases} \frac{\partial f}{\partial x}(x, y) = 4x^3 - 8(x - y) = 0 \\ \frac{\partial f}{\partial y}(x, y) = 4y^3 + 8(x - y) = 0 \end{cases} \quad (\star)$$

En sommant les deux lignes du système, on trouve

$$x^3 + y^3 = 0 \text{ i.e. } x^3 = -y^3. \quad (\star)$$

Si $x = 0$, alors le système de points critiques donne $y = 0$, donc $(0, 0)$ est point critique. Sinon $\left(\frac{y}{x}\right)^3 = -1$ d'après (\star) , mais $t \mapsto t^3$ est bijective donc $\frac{y}{x} = -1$ et $y = -x$. Remplaçant ceci dans le système définissant les points critiques, on trouve :

$$4x^3 - 8.2x = 0, \quad -4x^3 + 8.2x = 0,$$

soit $x^2 - 4 = 0$ et $x = \pm 2$. Donc $(2, -2)$, $(-2, 2)$ sont des points critiques. En conclusion, les points critiques sont :

$$(2, -2), \quad (-2, 2).$$

Exercice ANA.8.13 | Étudier les *extrema* locaux et globaux de :

$$f : (x, y) \mapsto x^2 + y^2 - x^3.$$

Solution (exercice ANA.8.13)

Soit $(x, y) \in \mathbf{R}^2$. Alors (x, y) est un point critique si et seulement si

$$\begin{cases} 2x - 3x^2 = 0, \\ 2y = 0, \end{cases} \iff y = 0, x(2 - 3x) = 0.$$

On déduit alors les points critiques : $(0, 0), (2/3, 0)$.

- ▶ **(Étude de $(0, 0)$)** On étudie le signe de $f(x, y) - f(0, 0) = x^2 + y^2 - x^3$ pour $(x, y) \in \mathbf{R}^2$. Si $|x| \leq 1, |y| \leq 1$, alors $x^3 \leq x^2 \leq x^2 + y^2$ donc pour (x, y) assez proche de $(0, 0), f(x, y) - f(0, 0) \geq 0$, donc $(0, 0)$ est un minimum local. En revanche, il n'est pas global, puisque : $f(2, 0) = -4 \leq f(0, 0) = 0$.
- ▶ **(Étude de $(2/3, 0)$)** On étudie le signe de $f(x, y) - f(2/3, 0)$ pour (x, y) assez proche de $(2/3, 0)$ ou encore de $f(2/3 + h, k) - f(2/3, 0)$ pour (h, k) assez petit.

$$\begin{aligned} f(2/3 + h, k) - f(2/3, 0) &= \left(\frac{2}{3} + h\right)^2 + k^2 - \left(\frac{2}{3} + h\right)^3 - \frac{4}{9} + \frac{8}{27} \\ &= \frac{4h}{3} + h^2 + k^2 - \left(\frac{8}{27} + 3\frac{4}{9}h + 3\frac{2}{3}h^2 + h^3\right) + \frac{8}{27} \\ &= h^2 + k^2 - 2h^2 - h^3 \\ &= -h^2 + k^2 - h^3. \end{aligned}$$

Donc $f(2/3 + h, 0) - f(2/3, 0)$ est négatif pour (h, k) assez petit, et $f(2/3 + 0, k) - f(2/3, 0) \geq 0$. Donc $(2/3, 0)$ n'est pas un extremum local.

Exercice ANA.8.14 | Soit $f \left| \begin{array}{l} \mathbf{R}^2 \rightarrow \mathbf{R}, \\ (x, y) \mapsto x^4 + y^4 - x^2 + y^2. \end{array} \right.$

1. Vérifier que f est de classe \mathcal{C}^1 sur \mathbf{R}^2 et calculer ses dérivées partielles premières.
2. Montrer que f possède exactement trois points critiques.
3. **3.1)** Calculer $f(0, 0)$ et étudier le signe de $f(x, 0)$ et $f(0, x)$ pour x assez proche de zéro.

3.2) Que peut-on en conclure?

4. 4.1) Calculer $f\left(\frac{1}{\sqrt{2}}, 0\right)$, puis $f\left(\frac{1}{\sqrt{2}} + h, k\right) - f\left(\frac{1}{\sqrt{2}}, 0\right)$ pour tout $(h, k) \in \mathbf{R}^2$.

4.2) Que peut-on en déduire?

Solution (exercice ANA.8.14)

1. La fonction f est une fonction polynomiale donc est de classe \mathcal{C}^1 . Pour tout $(x, y) \in \mathbf{R}^2$,

$$\frac{\partial f}{\partial x}(x, y) = 4x^3 - 2x \quad \frac{\partial f}{\partial y}(x, y) = 4y^3 + 2y.$$

2. Soit $(x, y) \in \mathbf{R}^2$, alors $\text{grad} f(x, y) = 0$ si et seulement si :

$$\begin{cases} \frac{\partial f}{\partial x}(x, y) = 0 \\ \frac{\partial f}{\partial y}(x, y) = 0 \end{cases} \Leftrightarrow \begin{cases} 2x(2x^2 - 1) = 0 \\ 2y(2y^2 + 1) = 0 \end{cases} \Leftrightarrow \begin{cases} x = 0 \text{ ou } x = \frac{1}{\sqrt{2}} \text{ ou } x = -\frac{1}{\sqrt{2}} \\ y = 0 \end{cases}$$

La fonction f admet trois points critiques : $(0, 0)$, $\left(\frac{1}{\sqrt{2}}, 0\right)$ et $\left(-\frac{1}{\sqrt{2}}, 0\right)$.

3. 3.1) $f(0, 0) = 0$ et pour tout x de \mathbf{R} , nous avons $f(x, 0) = x^2(x^2 - 1)$ donc pour $x \in]-1; 0[\cup]0; 1[$, $f(x, 0) < 0$. Enfin, $f(0, x) = x^4 + x^2 > 0$ pour tout $x \in \mathbf{R}^*$.

3.2) D'après la question précédente, $f(x, y) - f(0, 0)$ n'est pas de signe constant au voisinage de $(0, 0)$, ainsi f n'admet d'extremum en $(0, 0)$.

4. 4.1) $f\left(\frac{1}{\sqrt{2}}, 0\right) = \frac{1}{4} - \frac{1}{2} = -\frac{1}{4}$, ensuite fixons $(h, k) \in \mathbf{R}^2$, alors

$$\begin{aligned} & f\left(\frac{1}{\sqrt{2}} + h, k\right) - f\left(\frac{1}{\sqrt{2}}, 0\right) \\ &= \left(\frac{1}{\sqrt{2}} + h\right)^4 + k^4 - \left(\frac{1}{\sqrt{2}} + h\right)^2 + k^2 + \frac{1}{4} \\ &= \frac{1}{4} + \frac{4h}{2\sqrt{2}} + \frac{6h^2}{2} + \frac{4h^3}{\sqrt{2}} + h^4 - \frac{1}{2} - \frac{2h}{\sqrt{2}} - h^2 + k^2 + \frac{1}{4} \\ &= k^4 + \frac{4h^3}{\sqrt{2}} + 2h^2 + k^2 \\ &= h^2(h + \sqrt{2})^2 + k^4 + k^2. \end{aligned}$$

binôme de NEWTON
simplifications
factorisation par h^2

4.2) Ainsi, pour tout $(h, k) \in \mathbf{R}^2$, $f\left(\frac{1}{\sqrt{2}} + h, k\right) - f\left(\frac{1}{\sqrt{2}}, 0\right) \geq 0$.

La fonction f admet donc un minimum global en $\left(\frac{1}{\sqrt{2}}, 0\right)$.

Chapitre ANA.9.

Dynamique : Équations différentielles & Suites numériques

Résumé & Plan

Nous repassons en revue dans ce chapitre les principaux éléments de la théorie des équations différentielles et des suites numériques vue en première année, et généralisons à un contexte plus général. Ces objets mathématiques traduisent la notion de «dynamique» dans un contexte continu d'une part, et discret d'autre part, *i.e.* traduisent une évolution temporelle d'un système (en des temps discrets pour des modèles discrets, ou en tout temps pour des modèles continus).

Une dynamique apparaît naturellement dans beaucoup de disciplines ; les lois de la mécanique projetées dans un repère sont des équations différentielles, les lois régissant l'électricité aussi (mailles pour la tension & noeuds pour l'intensité), ou encore les modèles de dynamique des populations (discrets et continus).

1	Dynamique discrète : suites numériques	2
1.1	Généralités	2
1.2	Limite d'une suite	6
1.3	Théorèmes de convergence par majoration, minoration et encadrement ..	8
1.4	Suites extraites des termes pairs et impairs	9
1.5	Théorèmes de convergence par monotonie	10
1.6	Suites remarquables	13

2	Dynamique continue : équations différentielles	23
2.1	Généralités	23
2.2	Équations différentielles linéaires scalaires	26
2.3	Résolution approchée par la méthode d'EULER	37
2.4	Systèmes différentiels linéaires à coefficients constants	40
3	Modélisation de phénomènes dynamiques	42
3.1	Taux d'évolution et dynamiques linéaires	42
3.2	Modélisation de systèmes présentant une dynamique	43
3.3	Dynamique des populations	45
4	Exercices	50
4.1	Suites numériques	50
4.2	Équations différentielles	60

I say, that the power of population is indefinitely greater than the power in the earth to produce subsistence for man. Population, when unchecked, increases in a geometrical ratio. Subsistence increases only in an arithmetical ratio.

— Thomas. R. MALTHUS¹

Parmi toutes les disciplines mathématiques, la théorie des équations différentielles est la plus importante. Elle fournit l'explication de toutes les manifestations élémentaires de la nature où le temps est impliqué

— Sophus LIE

Définition ANA.9.2

- ▶ On appelle *suite réelle* toute application d'un sous-ensemble \mathcal{N} de \mathbf{N} dans \mathbf{R} , i.e. toute famille de réels indexée par un sous-ensemble de \mathbf{N} notée $(u_n)_{n \in \mathcal{N}}$.
- ▶ Si $\mathcal{N} = [n_0, \infty[\cap \mathbf{N}$ avec $n_0 \in \mathbf{N}$, on dit que $(u_n)_{n \in \mathcal{N}}$ est *définie à partir d'un certain rang*.
- ▶ Pour tout $n \in \mathcal{N}$, on dit que u_n est le *n-ième terme* de la suite, ou le terme de rang n .

La plupart du temps, \mathcal{N} sera égal à \mathbf{N} ou \mathbf{N}^* , on rappelle également que $\mathbf{R}^{\mathbf{N}}$ désigne l'ensemble des suites réelles.

Notation Abus de ...



Lorsque l'ensemble \mathcal{N} importe peu, nous nous autoriserons à noter seulement (u_n) au lieu de $(u_n)_{n \in \mathcal{N}}$. Cela signifiera donc implicitement que l'on considère que \mathcal{N} est l'ensemble de définition de l'expression u_n .



Attention

De même qu'il ne faut pas confondre une fonction f et l'image $f(x)$ de x par f , on prendra garde de bien distinguer la suite (u_n) de son terme général d'ordre n noté lui u_n sans parenthèse.

1. DYNAMIQUE DISCRÈTE : SUITES NUMÉRIQUES

Nous commençons ce chapitre par des révisions de première année sur les suites numériques.

Définition ANA.9.1 | À partir d'un certain rang

Pour tout $n \in \mathbf{N}$, soit \mathcal{P}_n une propriété. On dit qu'elle est *vraie à partir d'un certain rang* s'il existe $n_0 \in \mathbf{N}$, tel que \mathcal{P}_n soit vraie pour tout $n \geq n_0$.

¹à l'origine du premier modèle continu de dynamique des populations

Définition ANA.9.3 | Ensemble de définition et graphe

L'*ensemble de définition* d'une suite, ou ensemble de départ, est l'ensemble des entiers en lesquels elle est définie. Le *graphe* (ou *représentation graphique* de (u_n)) d'une suite (u_n) définie sur une partie \mathcal{N} de \mathbf{N} est l'ensemble constitué des couples (n, u_n) , où n parcourt \mathcal{N} .

On rappelle également de manière informelle qu'une suite peut être définie selon un mode *explicite*, i.e. son terme général en fonction de n , ou un mode par *réurrence*, i.e. u_{n+1} en fonction des termes précédents \dots, u_{n-1}, u_n . C'est ce second mode qui fait que l'étude des suites est indispensable, en plus de l'étude des fonctions.

On peut ainsi représenter une suite par son *graphe*. Voyons comment faire cela avec Python.

Représentation graphique d'une suite définie explicitement avec Python

Un premier exemple, une suite définie explicitement en fonction de n . par exemple (u_n) définie par

$$\forall n \in \mathbf{N}, \quad u_n = \sqrt{n+1}.$$

```
import matplotlib.pyplot as plt
```

```
plt.plot([(k + 1)**0.5 for k in range(11)], 'bo')
```

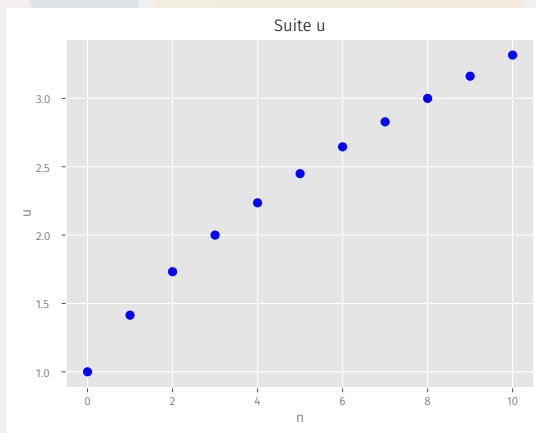
```
#Affichage en bleu et cercles
```

```
plt.title("Suite u")
```

```
plt.xlabel("n")
```

```
plt.ylabel("u")
```

```
plt.show()
```



Représentation graphique d'une suite définie par récurrence avec Python

Un second exemple, une suite définie par une relation de récurrence. par exemple (u_n) définie par

$$u_0 = 1, \quad \forall n \in \mathbf{N}, \quad u_{n+1} = \sqrt{u_n + 1}.$$

```
import matplotlib.pyplot as plt
```

```
def u(n):
```

```
    u = 1
```

```
    for k in range(n):
```

```
        u = (u + 1)**0.5
```

```
    return u
```

```
# version récursive
```

```
def u_bis(n):
```

```
    if n == 0:
```

```
        return 1
```

```
    else:
```

```
        return (u_bis(n-1) + 1)**0.5
```

```
plt.plot([u(k) for k in range(11)], 'ro')
```

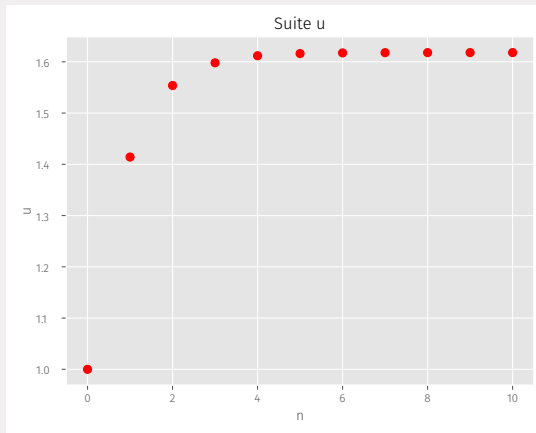
```
#Affichage en bleu et cercles
```

```
plt.title("Suite u")
```

```
plt.xlabel("n")
```

```
plt.ylabel("u")
```

```
plt.show()
```



Dans l'exemple précédent, pour chaque terme on réexécute la fonction `u` pour calculer le n -ième terme. On peut bien entendu améliorer ceci pour le tracé.

```
import matplotlib.pyplot as plt
```

```
def u(n):
    u = 1
    for k in range(n):
        u = (u + 1)**0.5
    return u
```

```
# version récursive
```

```
def liste_u(n):
    L = [1]
    for _ in range(2, n+1):
        dernier = L[-1]
        u = ma.sqrt(dernier+1)
        L.append(u)
    return L
```

Ensuite, on trace la suite comme nous l'avons fait avant, en utilisant cette fois-ci la liste `liste_u(n)`.

Remarque 1.1 – sur les scripts

- ▶ Dans les exemples précédents, on voit que nous ne sommes pas obligés de mentionner un premier argument dans les commandes `plt.plot`. Par défaut, Python placera en abscisses les entiers positifs. Ainsi, si on veut tracer une suite sur \mathbf{N}^* (par exemple), on précisera `list(range(1, ...))` en premier argument.
- ▶ Dans les options de `plt.plot`, `'o'` signifie que l'on trace des points non reliés entre eux, `'r'` et `'b'` correspondent aux couleurs.

Exemple 1 – Proposer quelques commandes à la volée permettant de tracer sur $[[0, 10]]$ la suite (u_n) ci-dessous définie par récurrence :

$$u_0 = 1, \quad \forall n \in \mathbf{N}, \quad u_{n+1} = |u_n| + 1.$$



Cadre

Dans la suite, afin de simplifier la présentation, nous considérons que (u_n) est définie sur \mathbf{N} .

Définition ANA.9.4 | Bornes et opérations sur les suites

Pour tout $(u_n) \in \mathbf{R}^{\mathbf{N}}$, $(v_n) \in \mathbf{R}^{\mathbf{N}}$ et $\lambda \in \mathbf{R}$, on pose :

- ▶ $(u_n) + (v_n) = (u_n + v_n)$,
- ▶ $(u_n) \times (v_n) = (u_n \cdot v_n)$,
- ▶ $\lambda (u_n) = (\lambda u_n)$.

La structure d'espace vectoriel des espaces de suites numériques a déjà été établie dans le **Chapter ALG.3** : $(\mathbf{R}^{\mathbf{N}}, +, \cdot)$ est un espace vectoriel.

Puisqu'une suite est une fonction bien particulière, la notion de suite bornée, monotone, etc. a déjà été définie dans le **Chapter ANA.7**. Nous les rappelons ici reformulées dans le cadre des suites.

Définition ANA.9.5 | Borne

On dit que la suite (u_n) est *majorée* (resp. *minorée* par $m \in \mathbf{R}$) s'il existe $M \in \mathbf{R}$ tel que :

$$\forall n \in \mathbf{N}, \quad u_n \leq M \quad (\text{resp. } \forall n \in \mathbf{N}, u_n \geq m).$$

Elle est dite *bornée* si elle est majorée et minorée.

Proposition ANA.9.1

Soit (u_n) une suite. Alors :

$$(u_n) \text{ est bornée} \iff \exists M \in \mathbf{R}^+, \quad \forall n \in \mathbf{N}, \quad |u_n| \leq M.$$

Preuve

\Rightarrow Supposons la suite (u_n) bornée. Alors il existe $(m, M) \in \mathbf{R}^+$ tel que : $\forall n \in \mathbf{N}, \quad m \leq u_n \leq M$. Alors nous avons immédiatement que : $\forall n \in \mathbf{N}, \quad |u_n| \leq \max(m, M)$.


\Leftarrow Puisque l'égalité $|u_n| \leq M$ pour tout $n \in \mathbf{N}$, $M \in \mathbf{R}^+$ est équivalente à $-M \leq u_n \leq M$ on en déduit que (u_n) est bornée.

Comme pour les fonctions, les ensembles de suites majorées, minorées ne sont pas des espaces vectoriels. En revanche, nous avons le résultat suivant.

Proposition ANA.9.2 | Structure d'espace vectoriel des suites bornées

L'ensemble des suites bornées forme un sous-espace vectoriel de $\mathbf{R}^{\mathbf{N}}$.

Preuve

Preuve identique aux fonctions. 

Définition ANA.9.6 | Monotonie

La suite (u_n) est dite :

- ▶ *croissante* (resp. *strictement croissante*) si :

$$\forall n \geq 0, \quad u_n \leq u_{n+1} \quad (\text{resp. } \forall n \geq 0, u_n < u_{n+1}),$$

- ▶ *décroissante* (resp. *strictement décroissante*) si :

$$\forall n \geq 0, \quad u_n \geq u_{n+1} \quad (\text{resp. } \forall n \geq 0, u_n > u_{n+1}),$$


- ▶ *monotone* (resp. *strictement monotone*) si elle est croissante **ou** décroissante (resp. strictement croissante ou strictement décroissante),
- ▶ **constante** si elle est croissante **et** décroissante.




Méthode Trouver la monotonie d'une suite

1. On étudie le signe de la différence $u_{n+1} - u_n$ pour tout entier n .
2. Pour les suites qui ne s'annulent pas, on peut également comparer le quotient $\frac{u_{n+1}}{u_n}$ à 1 pour tout entier $n \in \mathbf{N}$. Cette seconde version est bien adaptée aux produits.

Exemple 2 –

1. Montrer que $(\sqrt{n})_{n \in \mathbb{N}}$ est minorée par cinq à partir d'un certain rang. Expliciter un tel rang. 

2. Montrer qu'une suite majorée à partir d'un certain rang est majorée. 

1.2. Limite d'une suite

Définition ANA.9.8 | Convergence

Soient $\ell \in \mathbb{R}$ et (u_n) une suite. On dit que (u_n) converge vers ℓ si :

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{N}, \forall n \in \mathbb{N}, n \geq n_0 \implies |u_n - \ell| < \varepsilon.$$

On dit alors que la suite est *convergente* et le réel ℓ , appelé *limite* de la suite (u_n) , est encore noté $\ell = \lim_{n \rightarrow \infty} u_n$, ou encore $u_n \xrightarrow{n \rightarrow \infty} \ell$.

Cela signifie que l'on est « aussi proche que l'on veut » de la limite ℓ , pourvu que n soit assez grand.

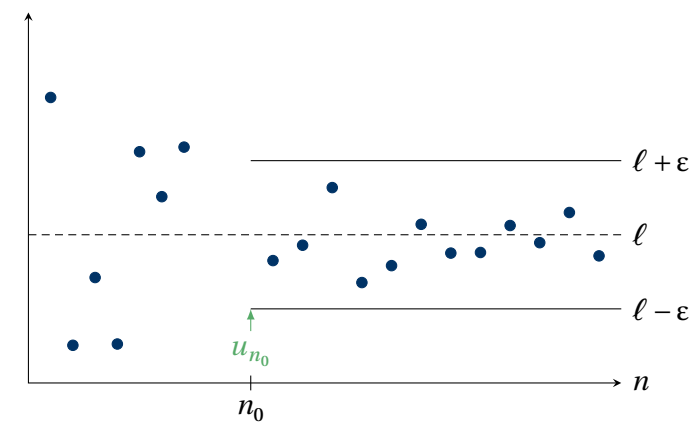



FIG. ANA.9.1. : Convergence d'une suite

Définition ANA.9.7 | Stationnaire

On dit qu'une suite (u_n) est *stationnaire* si elle est constante à partir d'un certain rang, *i.e.* si

$$\exists n_0 \in \mathbb{N}, \forall n \geq n_0, u_n = u_{n_0}.$$

Exemple 3 — La suite $\left(3 + \left[\frac{4}{2^n}\right]\right)_{n \in \mathbb{N}}$ est stationnaire. Expliciter l'entier n_0 de la définition précédente. 

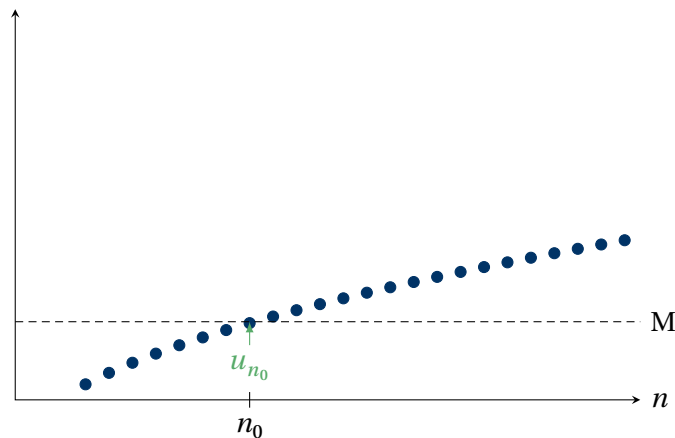


FIG. ANA.9.2. : Divergence d'une suite

Définition ANA.9.9 | Divergence

On dit qu'une suite (u_n) *diverge vers* $+\infty$ (resp. vers $-\infty$), et on note $u_n \xrightarrow{n \rightarrow \infty} +\infty$ (resp. $u_n \xrightarrow{n \rightarrow \infty} -\infty$), si :

$$\forall A \in \mathbf{R}, \exists n_0 \in \mathbf{N}, \forall n \in \mathbf{N}, n \geq n_0 \implies u_n \geq A$$

$$\text{(resp. : } \forall A \in \mathbf{R}, \exists n_0 \in \mathbf{N}, \forall n \in \mathbf{N}, n \geq n_0 \implies u_n \leq A)$$

Une suite est dite *divergente* si elle n'est pas convergente.

Cela signifie que l'on est «aussi grand que l'on veut», pourvu que n soit assez grand.

Définition ANA.9.10 | Nature

Déterminer la nature d'une suite c'est déterminer si elle converge ou diverge.

⊗ Attention Diverger ne signifie pas tendre vers $\pm\infty$

Il existe des suites qui ne convergent pas et qui ne divergent pas vers $\pm\infty$: ce sont celles n'ayant pas de limite. Par exemple, la suite (u_n) définie par $u_n = (-1)^n n^2$ pour tout $n \in \mathbf{N}$.

Théorème ANA.9.1 | Unicité de la limite

La limite d'une suite, si elle existe, est unique.

Preuve Faisons la preuve dans le cas d'une limite finie. Raisonnons par l'absurde et supposons que $u_n \xrightarrow{n \rightarrow \infty} \ell$, $u_n \xrightarrow{n \rightarrow \infty} \ell'$ et $\ell < \ell'$. Posons $\varepsilon = \frac{\ell' - \ell}{3} > 0$. Alors :

$$\ell - \varepsilon \leq u_n \leq \ell + \varepsilon$$

à partir d'un certain rang n_1 et

$$\ell' - \varepsilon \leq u_n \leq \ell' + \varepsilon$$

à partir d'un certain rang n_2 . Or, ε a été choisi de sorte que $\ell + \varepsilon < \ell' - \varepsilon$. Pour $n \geq n_0 = \max\{n_1, n_2\}$, on a donc :

$$u_n \leq \ell + \varepsilon < \ell' - \varepsilon \leq u_n$$

ce qui est absurde.

Théorème ANA.9.2

Toute suite convergente est bornée.

Preuve Soient $\ell = \lim_{n \rightarrow \infty} u_n$ et $\varepsilon = 1 > 0$. Alors en appliquant la définition de la convergence de (u_n) , on a :

$$\ell - 1 \leq u_n \leq \ell + 1$$

à partir d'un certain rang n_0 . Posons alors :

$$M = \max\{|u_0|, |u_1|, |u_2|, \dots, |u_{n_0-1}|, -(\ell - 1), \ell + 1\}.$$

Pour tout $n \geq n_0$, on a : $u_n \leq \ell + 1 \leq M$ et $-u_n \leq -(\ell - 1) \leq M$. Ainsi, pour $n \geq n_0$, on a : $|u_n| \leq M$. Comme la majoration est immédiate pour $n < n_0$, alors on a prouvé : $\forall n \in \mathbf{N}, |u_n| \leq M$. La suite (u_n) est bornée.

Théorème ANA.9.3

Soit (u_n) une suite.

1. Si $u_n \xrightarrow{n \rightarrow \infty} \ell > 0$, alors : $u_n > 0$ **à partir d'un certain rang.**
2. Plus généralement, si $u_n \xrightarrow{n \rightarrow \infty} \ell$ et $\ell' < \ell < \ell''$, alors, on a $\ell' < u_n < \ell''$ à partir

d'un certain rang.

⊗ **Attention**

La réciproque de **1**) est fautive : considérer par exemple la suite $(1/n)_{n \geq 1}$: elle est strictement positive, et pourtant elle converge vers zéro.

Preuve Soit $\ell = \lim_{n \rightarrow \infty} u_n > 0$. Posons $\varepsilon = \frac{\ell}{2} > 0$. Alors $\ell - \varepsilon \leq u_n \leq \ell + \varepsilon$ à partir d'un certain rang. Comme $\ell - \varepsilon = \frac{\ell}{2} > 0$, alors les termes u_n sont bien strictement positifs à partir d'un certain rang. Pour le cas général, appliquer le cas particulier à la suite $(u_n - \ell')$, qui converge vers $\ell - \ell' > 0$ et à la suite $(\ell'' - u_n)$, qui converge vers $\ell'' - \ell > 0$. On en déduit bien que $u_n - \ell' > 0$ et $\ell'' - u_n > 0$ à partir d'un certain rang.

⊗ **Attention**

Notez bien que l'encadrement de ℓ nécessite des inégalités strictes. L'idée étant que si les inégalités sont larges on peut converger en «se tassant» sur la limite ℓ' (mais en restant en-dessous par exemple). En revanche, si c'est strict, on peut choisir ε assez petit dans la définition de sorte que la suite pour n assez grand se retrouve strictement au-dessus de ℓ' .

Théorème ANA.9.4 | Passage à la limite dans les inégalités larges

Soient (u_n) et (v_n) deux suites réelles telles que :

1. $u_n \leq v_n$ à partir d'un certain rang,
2. $u_n \xrightarrow{n \rightarrow \infty} \ell \in \mathbf{R}$,
3. $v_n \xrightarrow{n \rightarrow \infty} \ell' \in \mathbf{R}$.

Alors : $\ell \leq \ell'$.

Preuve Procédons par l'absurde et supposons que $\ell > \ell'$. Alors $u_n - v_n \xrightarrow{n \rightarrow \infty} \ell - \ell' > 0$, et par le théorème précédent, $u_n - v_n > 0$ à partir d'un certain rang. C'est absurde.

⊗ **Attention**

On ne peut avoir mieux qu'une inégalité large $\ell \leq \ell'$ dans la conclusion, même



si l'on a $u_n < v_n$. Considérer $u_n = -\frac{1}{n}$ et $v_n = \frac{1}{n}$ par exemple.

Proposition ANA.9.3 | Propriétés de la limite

Soient (u_n) et (v_n) deux suites convergeant respectivement vers $\ell \in \overline{\mathbf{R}}$ et $\ell' \in \overline{\mathbf{R}}$. Alors :

- ▶ la suite $(|u_n|)$ converge vers $|\ell|$.
- ▶ Si $\{\ell, \ell'\} \neq \{-\infty, +\infty\}$, alors la suite $(u_n + v_n)$ converge vers $\ell + \ell'$.
- ▶ Pour tout $\lambda \in \mathbf{R}^*$, la suite (λu_n) converge vers $\lambda \ell$.
- ▶ Si $\{|\ell|, |\ell'|\} \neq \{0, +\infty\}$, alors la suite $(u_n v_n)$ converge vers $\ell \ell'$.
- ▶ Si $\ell \neq 0$, alors $u_n \neq 0$ à partir d'un certain rang et la suite $\left(\frac{1}{u_n}\right)$ converge vers $\frac{1}{\ell}$.
- ▶ Si $\ell = 0$ et $u_n > 0$ à partir d'un certain rang, alors $\left(\frac{1}{u_n}\right)$ diverge vers $+\infty$.

1.3. Théorèmes de convergence par majoration, minoration et encadrement

Théorème ANA.9.5 | Théorème de convergence par encadrement

On considère trois suites (u_n) , (v_n) et (w_n) telles que :

1. $u_n \leq v_n \leq w_n$ au moins à partir d'un certain rang,
2. les deux suites (u_n) et (w_n) convergent vers une même limite $\ell \in \mathbf{R}$.

Alors : $v_n \xrightarrow{n \rightarrow \infty} \ell$.

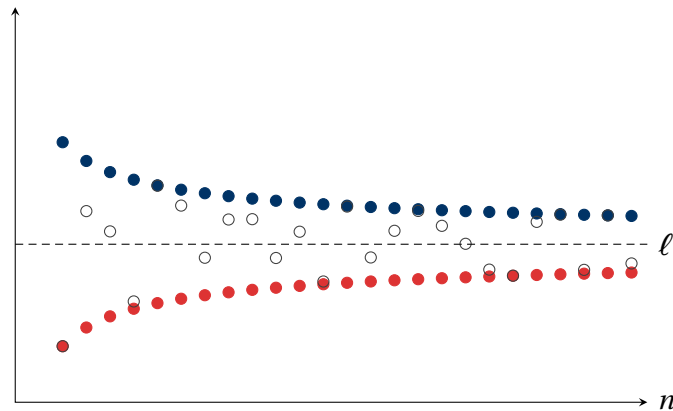


FIG. ANA.9.3. : Théorème d'encadrement : la suite représentée au milieu est nécessairement convergente

Preuve Soit $\varepsilon > 0$. Par hypothèse sur (u_n) et (w_n) , il existe $n_1 \in \mathbf{N}$ et $n_2 \in \mathbf{N}$ tels que pour tout $n \in \mathbf{N}$:

$$n \geq n_1 \implies \ell - \varepsilon \leq u_n \leq \ell + \varepsilon \quad \text{et} \quad n \geq n_2 \implies \ell - \varepsilon \leq w_n \leq \ell + \varepsilon.$$

Par ailleurs, il existe $n_3 \in \mathbf{N}$ tel que pour $n \geq n_3 : u_n \leq v_n \leq w_n$. Ainsi, en posant $n_0 = \max\{n_1, n_2, n_3\}$, on a pour tout $n \geq n_0 : \ell - \varepsilon \leq u_n \leq v_n \leq w_n \leq \ell + \varepsilon$. Ainsi, (v_n) converge vers ℓ .

Corollaire ANA.9.1

Le produit d'une suite bornée et d'une suite convergeant vers zéro est une suite convergeant vers zéro.

Preuve Soit (u_n) une suite bornée disons par $M \in \mathbf{R}^+$, et (v_n) convergeant vers zéro. Alors pour tout n ,

$$0 \leq |u_n v_n| \leq M |v_n|,$$

or $\lim_{n \rightarrow \infty} v_n = 0$, donc par théorème d'encadrement $\lim_{n \rightarrow \infty} (u_n v_n) = 0$.

Théorème ANA.9.6 | Théorème de divergence par minoration

On considère deux suites (u_n) et (v_n) telles que :

1. $u_n \leq v_n$ à partir d'un certain rang,

2. $u_n \xrightarrow{n \rightarrow \infty} +\infty$.

Alors : $v_n \xrightarrow{n \rightarrow \infty} +\infty$.

Preuve Soit $A \in \mathbf{R}$. Alors $u_n \geq A$ à partir d'un certain rang n_1 . Comme $u_n \leq v_n$ à partir d'un certain rang n_2 , alors $A \leq v_n$ à partir du rang $\max\{n_1, n_2\}$, d'où la conclusion.

Corollaire ANA.9.2 | Théorème de divergence par majoration

On considère deux suites (u_n) et (v_n) telles que :

1. $u_n \leq v_n$ au moins à partir d'un certain rang,

2. $v_n \xrightarrow{n \rightarrow \infty} -\infty$.

Alors : $u_n \xrightarrow{n \rightarrow \infty} -\infty$.

Preuve Appliquer le théorème de divergence par minoration aux suites $(-v_n)$ et $(-u_n)$.

1.4. Suites extraites des termes pairs et impairs

Théorème ANA.9.7 | Convergence des suites des termes pairs et impairs

Soit (u_n) une suite. On suppose que les deux suites extraites (u_{2n}) et (u_{2n+1}) convergent vers une même limite $\ell \in \overline{\mathbf{R}}$. Alors (u_n) converge vers ℓ .

Remarque 1.2 — Par contraposée, si les deux sous-suites des termes pairs et impairs ne convergent pas vers la même limite, alors la suite ne converge pas.


Preuve Supposons que $\ell \in \mathbf{R}$ (le cas où $\ell = \pm\infty$ est laissé en exercice). Soit $\varepsilon > 0$. Par convergence des deux suites extraites, il existe deux entiers n_1 et n_2 tels que $|u_{2n} - \ell| \leq \varepsilon$ pour $n \geq n_1$ et $|u_{2n+1} - \ell| \leq \varepsilon$ pour $n \geq n_2$. Posons alors $n_0 = \max\{2n_1, 2n_2 + 1\}$. Pour tout $n \geq n_0$, on a alors $|u_n - \ell| \leq \varepsilon$, que n soit pair ou impair : en effet si $n = 2p$, alors $2p \geq 2n_1$, donc $p \geq n_1$ d'où $|u_n - \ell| = |u_{2p} - \ell| \leq \varepsilon$, et si $n = 2p + 1$, alors $2p + 1 \geq 2n_2 + 1$, donc $p \geq n_2$ d'où $|u_n - \ell| = |u_{2p+1} - \ell| \leq \varepsilon$.

Ainsi, (u_n) converge vers ℓ .

Il existe des résultats faisant intervenir des suites autres que celles des termes impairs/pairs $((2n)$ et $(2n + 1))$, mais qui ne sont pas à notre programme.

Exemple 4 — Si une suite (u_n) vérifie

$$\forall n, p \in \mathbf{N}^*, \quad 0 \leq u_{n+p} \leq \frac{n+p}{np},$$

alors elle converge. 

Preuve Admise.

Remarque 1.3 —

- ▶ On peut préciser l'énoncé : une suite croissante (u_n) converge vers $\sup_{n \in \mathbf{N}} u_n$ (la limite étant finie ou non, selon que (u_n) est majorée ou non), *i.e.* vers le plus petit $M \in \mathbf{R}$ tel que : $\forall n \in \mathbf{N}, u_n \leq M$.
- ▶ Noter que la convergence de la suite est dans la conclusion du théorème.
- ▶ Cet énoncé permet de conclure quant à la convergence des suites croissantes, même si on a aucune idée de la limite!

ADJACENCE DE SUITES.

Définition ANA.9.11 | Suites adjacentes

Deux suites (u_n) et (v_n) sont dites *adjacentes* si :

1. elles sont monotones de sens contraires,
2. **(la différence tend vers zéro)** $u_n - v_n \xrightarrow{n \rightarrow \infty} 0$.

1.5. Théorèmes de convergence par monotonie

Théorème ANA.9.8 | Théorème de la limite monotone

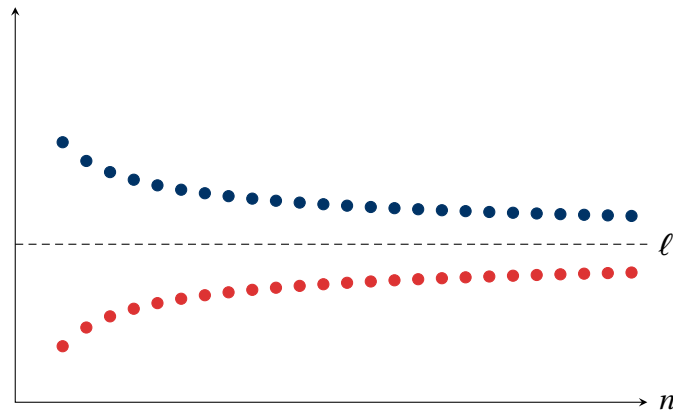
1. Toute suite réelle croissante et majorée (ou décroissante minorée) converge vers une limite finie.
2. Toute suite réelle croissante non majorée (*resp.* décroissante non minorée) diverge vers $+\infty$ (*resp.* $-\infty$).

Théorème ANA.9.9 | Convergence des suites adjacentes

Deux suites adjacentes convergent, et vers une même limite finie.

⊗ Attention

La convergence est dans **la conclusion**, et non dans la définition de l'adjacence.

FIG. ANA.9.4. : Deux suites adjacentes : elles convergent toutes les deux vers ℓ


Preuve On peut supposer sans perte de généralité que c'est (u_n) qui est croissante et (v_n) décroissante.

Ainsi, $(v_n - u_n)$ est décroissante et converge vers 0 : on a donc nécessairement $v_n - u_n \geq 0$ pour tout $n \in \mathbf{N}$. On en déduit que pour tout $n \in \mathbf{N}$: $u_0 \leq u_n \leq v_n \leq v_0$. Ainsi, la suite (u_n) est croissante et majorée par v_0 , elle converge donc vers un réel u d'après le théorème de la limite monotone. De même, (v_n) est décroissante et minorée par u_0 : elle converge vers un réel v . Enfin, $(v_n - u_n)$ converge vers 0 par hypothèse et vers $v - u$ par les théorèmes généraux. Par unicité de la limite, on a donc $u = v$.

Exemple 5 – Convergence d'une série alternée Soit la suite (u_n) définie par : $\forall n \in$


$$\mathbf{N}, u_n = \sum_{k=0}^n \frac{(-1)^k}{1+k}. \text{ Montrons que cette suite converge.}^2 \quad \square$$

²Cet exemple est à très bien connaître

3. Montrer que que $(v_n), (w_n)$ sont adjacentes. On appelle alors *constante d' EULER* notée γ , qui vaut $0,577$ à 10^{-3} près, la limite commune de ces deux suites. 

Exemple 6 — Existence de la constante d'EULER Soit $n \in \mathbf{N}^*$, on note

$$H_n = \sum_{k=1}^n \frac{1}{k}, \quad v_n = S_{n-1} - \ln n, \quad w_n = S_n - \ln n.$$

1. Justifier l'inégalité $x - \frac{x^2}{2} \leq \ln(1+x) \leq x$ pour tout $x \in [0, 1]$. 

2. Dédire que pour tout $n \in \mathbf{N}^*$, $\frac{1}{n+1} \leq \ln\left(1 + \frac{1}{n}\right) \leq \frac{1}{n}$. 

4.  Écrire un script Python permettant d'avoir une valeur approchée de γ à 10^{-3} près. 

1.6. Suites remarquables

1.6.1. Suites arithmétiques et géométriques.

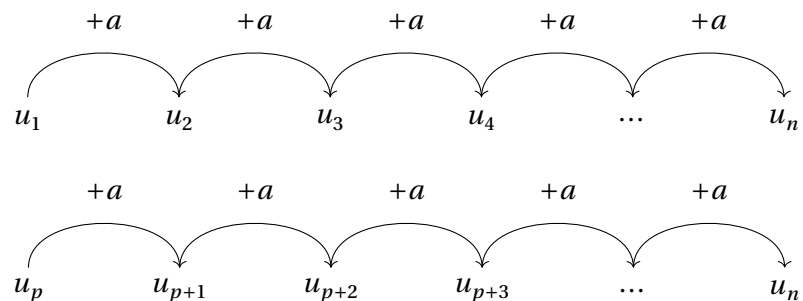
Définition ANA.9.12

On appelle *suite arithmétique* de raison $a \in \mathbf{R}$ toute suite $(u_n) \in \mathbf{R}^{\mathbf{N}}$ telle que :

$$\forall n \in \mathbf{N}, u_{n+1} = u_n + a.$$

On rappelle que l'on a alors :

$$\forall n \in \mathbf{N}, u_n = u_0 + na.$$



Il est clair qu'aucune suite arithmétique de raison non nulle ne converge. Nous pouvons cependant en expliciter leur somme.

5. Montrer que :

$$H_n \stackrel{\text{(d  f.)}}{=} \sum_{k=1}^n \frac{1}{k} \stackrel{n \rightarrow \infty}{\approx} \ln(n) + \gamma + o(1).$$



Théorème ANA.9.10 | Somme des termes d'une suite arithmétique.

Soit $a \in \mathbf{R}$. Pour tout $n \in \mathbf{N}$, on pose $S_n = \sum_{k=0}^n ka$. Alors :

$$S_n = a \frac{n(n+1)}{2}.$$

Preuve Voir votre cours de 1ère année.

Nous pouvons retenir une formule générale de la manière suivante :

$$\sum \text{suite arithmétique} = \text{nb termes} \times \frac{\text{premier terme} + \text{dernier terme}}{2}.$$

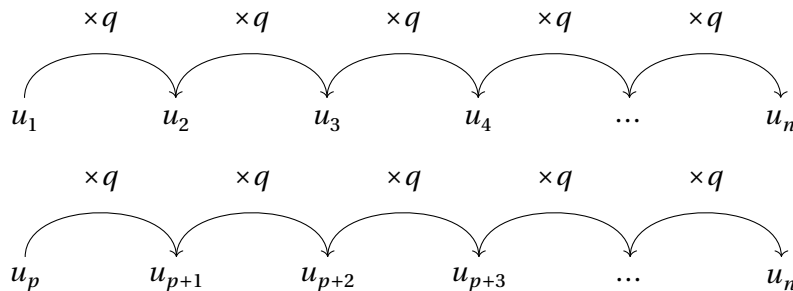
Définition ANA.9.13 | Suite géométrique

On appelle *suite géométrique* de raison $q \in \mathbf{R}$ toute suite (u_n) telle que :

$$\forall n \in \mathbf{N}, u_{n+1} = qu_n.$$

On rappelle que l'on a alors :

$$\forall n \in \mathbf{N}, u_n = u_0q^n.$$



Théorème ANA.9.11 | Convergence des suites géométriques.

Soit $q \in \mathbf{R}$. Alors la suite (q^n) converge si et seulement si $|q| < 1$ ou $q = 1$. Elle converge vers 0 si $|q| < 1$ et vers 1 si $q = 1$.

Théorème ANA.9.12 | Somme géométrique.

Soit $q \in \mathbf{R}$. Pour tout $n \in \mathbf{N}$, on pose $S_n = \sum_{k=0}^n q^k$. Alors :

$$S_n = \begin{cases} \frac{1 - q^{n+1}}{1 - q} & \text{si } q \neq 1, \\ n + 1 & \text{si } q = 1. \end{cases}$$

Preuve Si $q = 1$, alors il est clair que $S_n = n + 1$. Supposons donc que $q \neq 1$.

Pour tout $n \in \mathbf{N}$, on a : $S_n = \sum_{k=0}^n q^k$. Ainsi, pour tout $n \in \mathbf{N}$:

$$(1 - q)S_n = S_n - qS_n = \sum_{k=0}^n q^k - \sum_{k=0}^n q^{k+1} = 1 - q^{n+1}$$

et comme $1 - q \neq 0$, on obtient l'expression de S_n annoncée en divisant par $1 - q$.

Nous pouvons retenir une formule générale de la manière suivante :

$$\sum \text{suite géométrique} = \text{premier terme} \times \frac{1 - \text{raison}^{\text{nb termes}}}{1 - \text{raison}}.$$

1.6.2. Suites arithmético-géométriques

On rappelle également ici la méthode permettant d'étudier les mélanges des deux suites précédentes : les *suites arithmético-géométriques*.

Définition ANA.9.14

On appelle *suite arithmético-géométrique* toute suite (u_n) pour laquelle il existe $q \in \mathbf{R}$ et $a \in \mathbf{R}$ tels que :

$$\forall n \in \mathbf{N}, u_{n+1} = qu_n + a.$$

Proposition ANA.9.4

Soit (u_n) une suite telle que pour tout $n \in \mathbf{N}$ on ait $u_{n+1} = qu_n + a$, où $(q, a) \in \mathbf{R}^2$ et $q \neq 1$. Alors,


- ▶ il existe une constante $\alpha \in \mathbf{R}$ telle que $(u_n - \alpha)$ est une suite géométrique de raison q .
- ▶ **(Terme général en fonction de n)** $\forall n \in \mathbf{N}, u_n = \alpha + (u_0 - \alpha)q^n$.³

Preuve Commençons par déterminer une suite constante (α) vérifiant la même relation :

$$\alpha = q\alpha + a \iff \alpha = \frac{a}{1-q} \quad (\text{car } q \neq 1).$$

Posons $\alpha = \frac{a}{1-q}$. On a donc $\alpha = q\alpha + a$ et pour tout $n \in \mathbf{N}$: $u_{n+1} = qu_n + a$. Ainsi, par différence membre à membre, on a : $u_{n+1} - \alpha = q(u_n - \alpha)$. La suite $(u_n - \alpha)$ est donc une suite géométrique de raison q . On sait alors que $u_n - \alpha = (u_0 - \alpha)q^n$ (pour tout $n \in \mathbf{N}$).

Exemple 7 — Soit (u_n) la suite définie par $u_0 = 1$ et pour tout $n \in \mathbf{N}$: $u_{n+1} = 3u_n + 2$.

Déterminer une expression explicite en fonction de n de : $\sum_{k=0}^n u_k$. 

³Cette expression ne doit pas être donnée telle quelle mais être retrouvée par le calcul.

Définition ANA.9.15

On appelle *suite récurrente linéaire homogène d'ordre deux à coefficients constants* sur \mathbf{R} toute suite (u_n) pour laquelle il existe $a, b, c \in \mathbf{R}$ tels que :

$$\forall n \in \mathbf{N}, au_{n+2} + bu_{n+1} + cu_n = 0.$$

On appelle *équation caractéristique associée à (u_n)* l'équation

$$ax^2 + bx + c = 0. \quad (\text{EC})$$

Notation

On note \mathcal{S} l'ensemble des suites vérifiant la relation récurrente linéaire d'ordre deux précédente.

Théorème ANA.9.13 | Expression explicite

Soient (u_n) une suite récurrente linéaire homogène d'ordre deux à coefficients constants et Eq. (EC) son équation caractéristique. On suppose que $a \neq 0, c \neq 0$. On note Δ le discriminant de Eq. (EC).

1. (Si $\Delta > 0$) alors Eq. (EC) admet deux racines simples réelles distinctes α et β , et :

$$\mathcal{S} = \text{Vect}((\alpha^n), (\beta^n)).$$

2. (Si $\Delta = 0$) alors Eq. (EC) admet une racine double (nécessairement réelle) $\alpha \neq 0$, et :

$$\mathcal{S} = \text{Vect}((\alpha^n), (n\alpha^n)) = \{((\lambda + \mu n)\alpha^n), \lambda, \mu \in \mathbf{R}\}.$$

3. (Si $\Delta < 0$) alors Eq. (EC) admet deux racines complexes conjuguées α et $\bar{\alpha}$. On pose $\rho = |\alpha| > 0$ et $\theta \equiv \text{Arg}(\alpha) \in]-\pi, \pi[$ ($\theta \neq 0, \pm\pi$ car $\alpha \notin \mathbf{R}$), si bien que $\alpha = \rho e^{i\theta}$. Alors :

$$\mathcal{S} = \text{Vect}((\rho^n \cos(n\theta)), (\rho^n \sin(n\theta))).$$


Attention


aux confusions avec le résultat analogue sur les équations différentielles dans le cas $\Delta < 0$: il fait appel à la **forme algébrique des racines** pour les équations différentielles, et la **forme exponentielle** pour les suites.


La preuve ci-dessous doit être considérée comme un exercice d'Algèbre linéaire.


Preuve Faisons-là dans le cas $\Delta > 0$ i.e. où (EC) admet deux racines réelles distinctes.

1. Notons $\Phi \begin{cases} \mathcal{S} & \longrightarrow & \mathbf{R}^2, \\ (u_n) & \longmapsto & (u_0, u_1). \end{cases}$ Alors Φ est un isomorphisme linéaire, en particulier $\dim \mathcal{S} = 2$. En effet,

- ▶ Φ est linéaire. 

- ▶ Φ est injective. 


► Φ est surjective. 

4. Donc c'est une base puisque \mathcal{S} est de dimension deux. 

C'est terminé.

Exemple 8 — Suite de FIBONACCI à l'aide de l'équation caractéristique La suite de Fibonacci est définie par $u_0 = 0$, $u_1 = 1$, et :

$$\forall n \in \mathbf{N}, u_{n+2} = u_{n+1} + u_n.$$

1. Déterminer une expression explicite du n -ième terme en fonction de n . En déduire que la suite $\left(\frac{u_{n+1}}{u_n}\right)$ converge vers une limite à déterminer. 

2. Puisque α et β sont deux racines de (EC), alors $(\alpha^n), (\beta^n) \in \mathcal{S}$. 

3. On montre ensuite que la famille $((\alpha^n), (\beta^n))$ est libre dans \mathcal{S} . 

2. ➤ À l'aide Python, déterminer une valeur approchée de $\alpha = \frac{1+\sqrt{5}}{2}$.



Ainsi, en définissant

$$X_n = \begin{pmatrix} u_n \\ u_{n+1} \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix},$$

nous obtenons la récurrence vectorielle ci-après :

$$\forall n \in \mathbf{N}, X_{n+1} = AX_n, \quad X_0 = \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Trouver u_n en fonction de $n \in \mathbf{N}$ revient donc à trouver X_n en fonction de n , puis à regarder la première coordonnée de X_n .

► On déduit alors par récurrence que : $\forall n \in \mathbf{N}, X_n = A^n X_0$.



TRAITEMENT MATRICIEL D'UNE RÉCURRENCE LINÉAIRE D'ORDRE 2. Il est possible de reformuler une récurrence linéaire d'ordre deux ou plus à l'aide de matrices, mais pour simplifier nous ferons la présentation uniquement pour l'ordre deux.

Exemple 9 — Suite de FIBONACCI à l'aide de matrices On considère à nouveau la suite de Fibonacci est définie par $u_0 = 0$, $u_1 = 1$, et :

$$\forall n \in \mathbf{N}, u_{n+2} = u_{n+1} + u_n.$$

Cette dernière égalité est équivalente à la suivante :

$$\forall n \in \mathbf{N}, \begin{pmatrix} u_{n+1} \\ u_{n+2} \end{pmatrix} = \begin{pmatrix} u_{n+1} \\ u_{n+1} + u_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} u_n \\ u_{n+1} \end{pmatrix}.$$

► On calcule ensuite A^n en diagonalisant la matrice A . On montrera que $A = PDP^{-1}$

$$\text{avec } P = \begin{pmatrix} -\sqrt{5}-1 & \sqrt{5}-1 \\ 2 & 2 \end{pmatrix} \text{ et } D = \begin{pmatrix} \frac{1-\sqrt{5}}{2} & 0 \\ 0 & \frac{1+\sqrt{5}}{2} \end{pmatrix}.$$



1.6.4. Système de suites récurrentes linéaires à coefficients constants


On s'intéresse ici à un système de récurrences linéaires d'ordre 1 « imbriquées ». C'est l'analogie discret des systèmes différentiels linéaires d'ordre 1 que nous verrons dans la prochaine section. Traitons un exemple.

Exemple 10 — On considère trois suites $(u_n), (v_n), (w_n)$ telles que

$$\forall n \in \mathbf{N}, \begin{cases} u_{n+1} = u_n + v_n + w_n, \\ v_{n+1} = 2v_n + 2w_n, \\ w_{n+1} = 3w_n, \end{cases} \quad \text{et} \quad u_0 = 1, v_0 = 2, w_0 = 3.$$

Exprimons $(u_n), (v_n), (w_n)$ en fonction de $n \in \mathbf{N}$.

1. Écrire le système précédent à l'aide d'une matrice A et de $X_n = \begin{pmatrix} u_n \\ v_n \\ w_n \end{pmatrix}$ pour tout


$n \in \mathbf{N}$. 

► On peut conclure. 


2. On rappelle que dans l'**Exemple 12** du **Chapter ALG.5**, nous avons diagonalisé la

matrice $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix}$, et avons montré que :

$$A = P \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \cdot P^{-1}, \quad \text{avec} \quad P = \begin{pmatrix} 1 & 1 & \frac{3}{2} \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

3. Notons $Y_n = P^{-1}X_n$ pour tout $n \in \mathbf{N}$. Montrer que (X_n) est solution du système si et seulement si (Y_n) est solution de $Y_{n+1} = DY_n$ pour tout $n \in \mathbf{N}$. 

Remarque 1.4 — Les calculs précédents s'étendent naturellement aux récurrences d'ordre supérieur, le vecteur colonne X_n aura simplement une dimension plus grande que 2.

4. En déduire les solutions du système initial. 

⊗ Attention

Un point important est que le calcul effectif de P^{-1} n'est pas nécessaire dans la démarche.

🔧 Méthode Exprimer les solutions d'un système de suites récurrentes linéaires – Cas diagonalisable

On considère trois suites $(u_n), (v_n), (w_n)$ telles que

$$\forall n \in \mathbf{N}, \begin{pmatrix} u_{n+1} \\ v_{n+1} \\ w_{n+1} \end{pmatrix} = A \begin{pmatrix} u_n \\ v_n \\ w_n \end{pmatrix}, \quad \text{avec } A \in \mathfrak{M}_{3,3}(\mathbf{R}) \text{ supposée diagonalisable.}$$

Alors pour exprimer les suites $(u_n), (v_n), (w_n)$ en fonction de n , on :

1. diagonalise A , *i.e.* on cherche P inversible, D diagonale telle que $A = PDP^{-1}$,
2. on pose $X_n = \begin{pmatrix} u_n \\ v_n \\ w_n \end{pmatrix}$, et $Y_n = P^{-1}X_n$ pour tout $n \in \mathbf{N}$. On montre que $Y_{n+1} = DY_n$ pour tout $n \in \mathbf{N}$.
3. On peut alors exprimer Y_n en fonction de $n \in \mathbf{N}$.
4. Puis on conclut à l'aide de P : $X_n = PY_n$.

1.6.5. Suites récurrentes générales d'ordre 1

Cette fois-ci on ne suppose plus linéaire la relation de récurrence, mais seulement d'ordre 1 (*i.e.* elle fait intervenir un terme et le suivant).

Définition/Proposition ANA.9.1

On appelle *suite récurrente d'ordre un* toute suite (u_n) vérifiant une relation de récurrence de la forme :

$$\forall n \in \mathbf{N}, \quad u_{n+1} = f(u_n), \quad u_0 \in D$$

où $f : D \rightarrow \mathbf{R}$ est une fonction définie sur un domaine $D \subset \mathbf{R}$ vérifiant

$$f(D) \subset D.$$

On peut résumer cette technique dans une méthode.

Preuve Si $f(D) \subset D$, montrons par récurrence \mathcal{P}_n « u_n est bien défini, et $u_n \in D$ », pour $n \in \mathbf{N}$. Pour $n = 0$, $u_0 \in D$. Si la propriété est vraie au rang n , alors $u_n \in D$ puis $u_{n+1} = f(u_n) \in f(D) \subset D$ est alors aussi bien définie. Donc la suite (u_n) est bien définie.

Constatons que dans cette partie nous généralisons certaines suites remarquables déjà étudiées.

Exemple 11 — Précisons les récurrences obtenues dans quelques cas particuliers.

- ▶ $f(x) = x + a$, où $a \in \mathbf{R}$: la suite (u_n) associée est alors arithmétique.
- ▶ $f(x) = qx$, où $q \in \mathbf{R}$: la suite (u_n) associée est alors géométrique.
- ▶ $f(x) = qx + a$, où $(q, a) \in \mathbf{R}^2$: la suite (u_n) associée est alors arithmético-arithmétique.

La suite de cette partie est axée autour de deux problèmes :

- ▶ celui de la définition de la suite.
- ▶ Puis, celui de l'existence éventuelle d'une limite.

NATURE DE LA SUITE. Le plus souvent, dans les exercices, on vous fera :

1. analyser la monotonie de la suite. Cela se fait en comparant u_0 et u_1 puis en utilisant la monotonie de f . Si par exemple $u_1 \leq u_0$ et f est croissante, alors, en appliquant n fois f à l'inégalité précédente, on obtient pour tout $n \in \mathbf{N}$, $u_{n+1} \leq u_n$.
2. Dédire la convergence/divergence en utilisant les théorèmes de convergence/divergence monotone.

Attention En règle générale, la monotonie de f ne donne pas directement la monotonie de (u_n) .

Par exemple, si la suite est croissante (*resp.* décroissante), et que D est majorée (*resp.* minorée), alors on peut établir la convergence. Enfin, pour trouver la valeur de la limite, on passe à la recherche de points fixes pour la fonction f .

Définition ANA.9.16 | Point fixe

On appelle *point fixe* de $f : D \rightarrow \mathbf{R}$ tout réel $x \in D$ tel que $f(x) = x$.

Théorème ANA.9.14

Soient $f : D \rightarrow \mathbf{R}$, où $D \subset \mathbf{R}$, et (u_n) une suite réelle telle que :


$$\forall n \in \mathbf{N}, \quad u_{n+1} = f(u_n).$$


Si (u_n) converge vers ℓ et f est définie et **continue** en ℓ , alors ℓ est un point fixe de f .

Attention Ne pas oublier la continuité!

Preuve Pour tout $n \in \mathbf{N}$, on a $u_{n+1} = f(u_n)$. Le membre de gauche tend vers ℓ lorsque n tend vers $+\infty$ et celui de droite vers $f(\ell)$ par continuité de f . Ainsi, $f(\ell) = \ell$.

Exemple 12 — Étudions les suites récurrentes suivantes :

1. $u_0 \geq 1$ et $u_{n+1} = \sqrt{u_n}$. 

2. $u_0 \in [0, 1]$ et $u_{n+1} = u_n^2$. 

Dans les exemples précédents, la fonction f associée était toujours croissante. Des exemples où la fonction f est décroissante seront vus en TD.

1.6.6. Suites implicites

Définition ANA.9.17

On appelle *suite implicite* toute suite (x_n) dont le terme général x_n est donné comme solution (en général unique) d'une équation dépendant d'un paramètre $n \in \mathbf{N}$, i.e. vérifiant une égalité du type

$$f_n(x_n) = 0$$


avec f_n une fonction pour tout entier $n \in \mathbf{N}$.


Il n'y a pas de résultat général au programme, mais leur étude s'appuie souvent sur un schéma proche de l'exemple ci-après. La difficulté est qu'*a priori* on ne connaît pas l'expression générale d'une suite implicite, on utilisera le théorème de convergence monotone pour établir la convergence.

Exemple 13 — Étude d'une suite implicite Pour tout $n \in \mathbf{N}$, on considère l'équation :


$$f_n(x) = 0 \quad (E_n)$$


d'inconnue $x \in \mathbf{R}_+^*$, où : $f_n(x) = nx + \ln(x)$.

1. Pour tout $n \in \mathbf{N}$, l'équation (E_n) admet une unique solution sur \mathbf{R}_+^* . On la note désormais x_n .  La fonction f_n est continue sur \mathbf{R}^{+*} , et $\lim_{x \rightarrow \infty} f_n(x) = \infty$ par croissances comparées, et $\lim_{x \rightarrow 0} f_n(x) = -\infty$. De plus, en calculant la dérivée, on constate facilement que la fonction est même strictement croissante. Donc d'après le théorème de la bijection, la fonction f_n réalise une bijection de \mathbf{R}^{+*} vers $f(\mathbf{R}^{+*}) = \mathbf{R}$ (d'après le calcul de limites et la monotonie de f). Comme $0 \in \mathbf{R}$, il existe un unique $x_n \in \mathbf{R}^{+*}$ comme prétendu dans l'énoncé.

2. Pour tout $n \in \mathbf{N}$, on a $x_n \in]0, 1]$.  Comme $f_n(1) = n > 0$, on peut même affirmer que $x_n \in]0, 1]$.

3. La suite (x_n) décroît. Indication : On cherchera à comparer $0 = f_n(x_n)$ et $f_n(x_{n+1})$.

 Nous avons $f_n(x_{n+1}) = nx_{n+1} + \ln(x_{n+1}) = (n+1)x_{n+1} + \ln(x_{n+1}) - x_{n+1} = f_{n+1}(x_{n+1}) - x_{n+1} = -x_{n+1} \leq 0$ car $f_{n+1}(x_{n+1}) = 0$. Ainsi, puisque $0 = f_n(x_n) \geq f_n(x_{n+1})$, et que la fonction f_n est croissante, nous avons immédiatement $x_n \geq x_{n+1}$ pour tout entier $n \in \mathbf{N}$.

4. La suite (x_n) converge vers une limite $\ell \in [0, 1]$ que l'on peut déterminer.  La suite est d'après ce qui précède décroissante minorée par zéro, donc converge vers une limite finie.

Supposons que $\ell \in]0, 1]$. Alors puisque $nx_n = -\ln(x_n)$, nous aurions en passant à la limite : $-\ln(\ell) = \infty$ ce qui est clairement une contradiction.

Méthode Plan d'étude d'une suite implicite

1. Établir l'existence et l'unicité de la suite grâce au théorème de la bijection.
2. Chercher la monotonie en comparant $f_n(x_{n+1})$ à $f_n(x_n) = 0$ (ou $f_{n+1}(x_n)$ à $f_n(x_n) = 0$).
3. Trouver la valeur de la limite : en général on raisonne par l'absurde dans l'identité $f_n(x_n) = 0$.

2. DYNAMIQUE CONTINUE : ÉQUATIONS DIFFÉRENTIELLES

Cadre

Dans toute cette section,

- ▶ I désignera un intervalle réel, qui sera le domaine de définition de l'équation différentielle.
- ▶ L'entier n désignera l'ordre de l'équation différentielle.
- ▶ L'entier p désignera le nombre de coordonnée des solutions, qui seront des matrices colonnes de $\mathfrak{M}_{p,1}(\mathbf{R})$.

2.1. Généralités

Définition ANA.9.18 | Équation différentielle

Soient $n, p \in \mathbf{N}^*$. On appelle *équation différentielle d'ordre n sur $\mathfrak{M}_{p,1}(\mathbf{R})$* toute équation⁴ de la forme

$$y^{(n)} = f(t, y, \dots, y^{(n-1)})$$

où :

- ▶ l'intervalle I est appelé le *domaine de définition* de l'équation différentielle.
- ▶ $f : I \times (\mathfrak{M}_{p,1}(\mathbf{R}))^n \rightarrow \mathfrak{M}_{p,1}(\mathbf{R})$ est appelée *générateur de l'équation différentielle*.
- ▶ La fonction $y \in \mathcal{D}^n(I, \mathfrak{M}_{p,1}(\mathbf{R}))$ est appelée *inconnue* de l'équation. On appelle *solution* toute fonction $y : I \rightarrow \mathfrak{M}_{p,1}(\mathbf{R})$ dérivable⁵ telle que :

$$\forall t \in I, \quad y^{(n)}(t) = f(t, y'(t), \dots, y^{(n-1)}(t)).$$

Si $p = 1$, la courbe représentative d'une solution est appelée une *courbe intégrale* de l'équation différentielle et on parle d'équation différentielle *scalaire*.

Ainsi, on appelle toute équation différentielle toute relation entre une fonction et ses dérivées.

Remarque 2.1 — On omet souvent la variable dans les fonctions inconnues. Mais, en toute rigueur, il faudrait considérer comme équation différentielle une équation de la forme

$$\forall t \in I, \quad y^{(n)}(t) = f(t, y'(t), \dots, y^{(n-1)}(t)).$$

Définition ANA.9.19 | Système différentiel

On appelle *système différentiel d'ordre 1* (ou simplement *système différentiel*) toute équation différentielle où :

- ▶ $n = 1$,⁶
- ▶ il existe $B \in \mathfrak{M}_{p,1}(\mathbf{R})$, et pour tout $t \in I$ une matrice $A(t) \in I$, telles que :

$$\forall x \in \mathfrak{M}_{p,1}(\mathbf{R}), \quad f(t, x) = A(t)x + B.$$

Le système est dit à *coefficients constants* si la fonction $t \in I \rightarrow A(t)$ est constante.

⁴sous-entendue en une fonction y

⁵i.e. que chaque coordonnée de y est dérivable


⁶l'équation différentielle ne fait alors apparaître que l'inconnue y et sa dérivée première y'


De manière plus explicite, un système différentiel en l'inconnue $Y(t) = \begin{pmatrix} y_1(t) \\ \vdots \\ y_p(t) \end{pmatrix}$, est une équation différentielle de la forme


$$Y'(t) = \begin{pmatrix} y_1'(t) \\ \vdots \\ y_p'(t) \end{pmatrix} = A(t) \begin{pmatrix} y_1(t) \\ \vdots \\ y_p(t) \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix},$$


où $b_i \in \mathbf{R}$ pour tout $i \in \llbracket 1, p \rrbracket$. Nous verrons une méthode afin de résoudre un tel système différentiel linéaire lorsque A est diagonalisable.

Exemple 14 — *Écriture d'une équation différentielle sous la forme* $y^{(n)} = f(t, y', \dots, y^{(n-1)})$ Dans chacun des exemples ci-dessous, écrire l'équation sous la forme $y^{(n)} = f(t, y', \dots, y^{(n-1)})$ de la définition. On précisera les entiers n, p associés et la fonction f .⁷

1. $y''' = 4ty''$, 

2. $y'' = 4xy' - 2y$, 

3. $z'' = 4z' - 2z$, 

4. **(Système de Lokta-Volterra)** $\begin{cases} P_1' = \left(1 - \frac{P_2}{3}\right)P_1, \\ P_2' = \left(1 - \frac{P_1}{2}\right)P_2. \end{cases}$ Est-ce un système différentiel linéaire? 

⁷Les notations de variable, fonction inconnue sont volontairement diverses dans cet exemple, vous devez savoir vous adapter

5. (Système de chauffe)
$$\begin{cases} m_1 C_1 \frac{dT_1}{dt} = h_1 S_1 (T_2 - T_1), \\ m_2 C_2 \frac{dT_2}{dt} = h_1 S_1 (T_1 - T_2) + h_2 S_2 (T_\infty - T_2), \end{cases} \text{ avec}$$

 $m_1, C_1, m_2, C_2, h_1, h_2, T_\infty \in \mathbf{R}$. C'est même un système différentiel à coefficients constants. Préciser $t \mapsto A(t)$ et B comme définis dans la Définition ANA.9.19.



Attention Non-unicité du générateur et des « entiers n, p »

Il n'y a aucune unicité des objets précédents. Par exemple, $y'' = 4xy' - 2y$ peut également être vue comme :

$$\frac{d}{dx} \begin{pmatrix} y \\ y' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 4x & -2 \end{pmatrix} \begin{pmatrix} y \\ y' \end{pmatrix},$$

donc comme une équation différentielle dont la fonction inconnue $\begin{pmatrix} y \\ y' \end{pmatrix}$ est à valeurs dans $\mathfrak{M}_{2,1}(\mathbf{R})$ ($p = 2$), d'ordre $n = 1$ et de fonction associée

$$f \left| \begin{array}{l} \mathbf{R} \times \mathfrak{M}_{2,1}(\mathbf{R}) \longrightarrow \mathfrak{M}_{2,1}(\mathbf{R}) \\ (x, Y) \longrightarrow \begin{pmatrix} 0 & 1 \\ -2 & 4x \end{pmatrix} Y \end{array} \right.$$

✗ donc un système différentiel *in fine*.

Plaçons-nous sans plus tarder dans le cas linéaire scalaire.

2.2. Équations différentielles linéaires scalaires

2.2.1. Généralités

Définition ANA.9.20

Soit $n \in \mathbf{N}^*$. On appelle *équation différentielle linéaire d'ordre n sur \mathbf{R}* toute équation de la forme

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + \dots + a_1(t)y' + a_0(t)y = b(t) \tag{E_n}$$

où $a_i \in \mathcal{C}^0(I, \mathbf{R})$ pour tout $i \in \llbracket 1, n-1 \rrbracket$. La fonction $y \in \mathcal{D}^n(I, \mathbf{R})$ est appelée *inconnue* de (E_n) . On appelle *solution* de (E_n) toute fonction $y : I \rightarrow \mathbf{R}$ dérivable telle que :

$$\forall t \in I, \quad y^{(n)}(t) + a_{n-1}(t)y^{(n-1)}(t) + \dots + a_1(t)y'(t) + a_0(t)y(t) = b(t).$$

Elle est dite à coefficients constants si les fonctions a_{n-1}, \dots, a_1 sont constantes.

Remarque 2.2 — Les définitions de courbe intégrale et d'intervalle de définition restent bien entendu en vigueur.

Définition ANA.9.21 | Homogène

- ▶ L'équation (E_n) est dite *homogène*, ou *sans second membre*, si b est la fonction nulle.
- ▶ On appelle *équation homogène associée* à (E_n) ou encore *équation sans second membre associée* à (E_n) l'équation suivante :

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + \dots + a_1(t)y' + a_0(t)y = 0. \tag{H_n}$$

Σ Notation

Dans la suite, nous noterons \mathcal{S} l'ensemble des solutions de (E_n) , et \mathcal{S}^0 l'ensemble des solutions de (H_n) .

Remarque 2.3 — Forme normalisée \longleftrightarrow Forme générale Une équation de la forme

$$a_n(t)y^{(n)} + a_{n-1}(t)y^{(n-1)} + \dots + a_1(t)y' + a_0(t)y = c(t)$$


est encore appelée une équation différentielle linéaire d'ordre n . La forme faisant intervenir un coefficient 1 devant la dérivée s'appelle la *forme normalisée* de l'équation différentielle, elle s'obtient en divisant les deux membres par la fonction a_n sur tout intervalle J où a_n ne s'annule pas,


$$\forall t \in I \cap J, \quad y^{(n)} + \frac{a_{n-1}(t)}{a_n(t)}y^{(n-1)} + \dots + \frac{a_1(t)}{a_n(t)}y' + \frac{a_0(t)}{a_n(t)}y = \frac{b(t)}{a_n(t)}.$$

Dans la suite tous les résultats seront énoncés pour la forme normalisée, *i.e.* celle des équations (E_n) et (H_n) .

Remarque 2.4 — Caractère \mathcal{C}^n des solutions Si y est une solution de (E_n) , alors $y^{(n)}$ est même continue sur I (*i.e.* $y \in \mathcal{C}^n(I, \mathbf{R})$), puisque $y^{(n)} = b - a_{n-1}y^{(n-1)} - \dots - a_0y$ est alors une somme de fonctions continues.

Exemple 15 — Justifier que les équations différentielles suivantes sont linéaires, et déterminer leur homogène associée. Préciser leur domaine de définition.

1. $2 - xz' = x^2z''$. 

2. $\sum_{i=1}^n iy^{(i)} = \pi.$ 

► Puisque \mathcal{S}_0 est un noyau,^a c'est en particulier un espace vectoriel.

Attention

Ce n'est pas le cas de l'ensemble \mathcal{S} ! En effet, dès que le second membre est non nul, la solution nulle n'est plus solution, \mathcal{S} n'est donc *a fortiori* pas un espace vectoriel.

Définition/Proposition ANA.9.2 | Endomorphisme associé à une équation différentielle linéaire

► On appelle *endomorphisme associé à (E_n)* l'application linéaire suivante :


$$\Phi_n(a_0, \dots, a_{n-1}) \left| \begin{array}{l} \mathcal{D}^n \longrightarrow \mathbf{R}^1, \\ y \longmapsto y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1y' + a_0y. \end{array} \right.$$

► $\mathcal{S}_0 = \ker(\Phi_n(a_0, \dots, a_{n-1}))$. Par conséquent, \mathcal{S}_0 est un espace vectoriel.⁸

STRUCTURE DE L'ENSEMBLE DES SOLUTIONS DE (E_n) . Une équation différentielle linéaire est une «équation linéaire» sur $\mathcal{C}^n(I, \mathbf{R})$, donc l'ensemble des solutions possède une structure classique.

Dans la suite on notera plus simplement Φ_n au lieu de $\Phi_n(a_0, \dots, a_{n-1})$.

Preuve

► L'application Φ_n est linéaire. 

Théorème ANA.9.15 | Structure des solutions de l'équation complète

Si $y^{\text{par}} : I \rightarrow \mathbf{R}$ est **une** solution particulière de l'équation complète (E_n) , alors les solutions de (E_n) sont toutes les fonctions d'expression :

$$y = y_H + y^{\text{par}},$$

où y_H est une solution de (H_n) . Autrement dit l'ensemble \mathcal{S} des solutions sur \mathbf{R} de (E_n) est :

$$\mathcal{S} = \mathcal{S}_0 + y^{\text{par}} \stackrel{\text{(déf.)}}{=} \{y_H + y^{\text{par}}, y_H \in \mathcal{S}_0\},$$

où \mathcal{S}_0 est l'ensemble des solutions de (H_n) .

⁸On peut montrer également qu'il est de dimension finie égale à n . Nous n'avons pas les outils pour le montrer de manière générale, nous le ferons donc seulement pour $n = 1, 2$.

^aOn aurait pu utiliser aussi la définition d'un sous-espace vectoriel et montrer notamment la stabilité par combinaison linéaire

Attention La preuve ci-dessous exploite très largement la linéarité de l'équation, ce résultat est faux dans le cas contraire.

Preuve (Point clef — *Faire la différence entre l'ED avec la solution particulière et une solution générale*)



PRINCIPE DE SUPERPOSITION. Passons à une autre conséquence de la linéarité : le principe de superposition.

Méthode Principe de superposition

Lorsque le second membre b de (E_n) s'écrit comme une somme fonctions, par exemple $b = \lambda b_1 + \mu b_2$ avec b_i continue pour tout $i \in \llbracket 1, 2 \rrbracket$, $\lambda, \mu \in \mathbf{R}$, on peut appliquer le *principe de superposition*. Il s'agit de considérer les deux équations différentielles linéaires :

$$y^{(n)} + a_{n-1}^1(t)y^{(n-1)} + \dots + a_1^1(t)y' + a_0^1(t)y = b_1(t) \quad (E_n^1)$$

$$y^{(n)} + a_{n-1}^2(t)y^{(n-1)} + \dots + a_1^2(t)y' + a_0^2(t)y = b_2(t) \quad (E_n^2)$$

Si l'on a :

déterminé une solution particulière y_1^{par} de (E_n^1) ,
 et une solution particulière y_2^{par} de (E_n^2) .
 La somme $\lambda y_1^{\text{par}} + \mu y_2^{\text{par}}$ est alors une solution particulière de (E_n) . Le principe reste naturellement vrai si l'on considère une somme de n fonctions dans le second membre.

Preuve Sommer simplement les équations (E_n^1) et (E_n^2) .

Des exemples d'application seront vues dans les prochaines sections. Maintenant que le résultat général de structure des ensembles de solution est établi, nous allons l'appliquer à l'ordre 1 et 2. Tout l'enjeu est alors de savoir :

- calculer explicitement l'ensemble \mathcal{S}_0 des solutions de l'homogène,
- déterminer une solution particulière y^{par} . Pour l'ordre 1, nous aurons une méthode systématique appelée *variation de la constante*.

2.2.2. Équations différentielles linéaires du 1^{er} ordre

Définition ANA.9.22 | Définition pour $n = 1$
 On appelle *équation différentielle linéaire du premier ordre sur \mathbf{R}* toute équation de la forme (E_1) .

C'est donc une équation du type

$$y' + a(t)y = b(t), \quad (E_1)$$

où $a, b \in \mathcal{C}^0(I, \mathbf{R})$. Rappelons que nous avons également défini la notion d' *équation homogène associée* ou encore *équation sans second membre associée* à (E_1) l'équation suivante :

$$y' + a(t)y = 0. \quad (H_1)$$

Remarque 2.5 — Caractère \mathcal{C}^1 des solutions Si y est une solution de (E_1) , alors y' est continue sur I , puisque $y' = b - ay$. La fonction y est donc un élément de $\mathcal{C}^1(I, \mathbf{R})$.

RÉSOLUTION DE L'ÉQUATION HOMOGÈNE On connaît par un calcul direct l'ensemble des solutions de l'équation (H_1) , donné dans le théorème suivant. Nous savions déjà en revanche que c'est un espace vectoriel!

Théorème ANA.9.16 | Résolution de l'équation homogène

L'ensemble \mathcal{S}_0 des solutions de (H_1) est :

$$\begin{aligned}\mathcal{S}_0 &= \text{Vect} \left(t \mapsto e^{-A(t)} \right), \text{ où } A : I \rightarrow \mathbf{R} \text{ est une primitive de } a \\ &= \text{Vect} \left(t \mapsto \exp \left(- \int_{t_0}^t a(s) ds \right) \right), \text{ où } t_0 \in I.\end{aligned}$$

Ainsi, \mathcal{S}_0 est un espace vectoriel de dimension 1.

Preuve La fonction a étant continue sur I , elle admet une primitive A sur cet intervalle. Une fonction dérivable $y : I \rightarrow \mathbf{R}$ est alors solution de (H_1) si et seulement si pour tout $t \in I$:

$$y'(t)e^{A(t)} + a(t)y(t)e^{A(t)} = 0$$

(la fonction $t \mapsto e^{A(t)}$ ne s'annule pas sur I). On reconnaît la dérivée du produit $y(t)e^{A(t)}$. Ainsi, y est solution si et seulement si $t \mapsto y(t)e^{A(t)}$ est constante sur l'intervalle I . Les solutions sont donc bien les fonctions $y : I \rightarrow \mathbf{R}$ telles que : $\forall t \in \mathbf{R}, y(t)e^{A(t)} = C$, où $C \in \mathbf{R}$, d'où le résultat.

Exemple 16 — Résoudre $(1 + t^2)y' + 4ty = 0$. 



Cadre

Dans la suite de cette sous-section, on se fixe une équation différentielle $y' + ay = b$, avec $a, b : I \rightarrow \mathbf{R}$ continues, et A désigne une primitive de a .

RÉSOLUTION DE L'ÉQUATION COMPLÈTE On applique simplement le théorème déjà démontré sur le sujet : toute solution est obtenue en sommant les solutions de l'homogène et une solution particulière.

Théorème ANA.9.17 | Résolution de l'équation complète

L'ensemble \mathcal{S} des solutions sur \mathbf{R} de (E_1) est :

$$\mathcal{S} = \left\{ t \in I \mapsto Ce^{-A(t)} + y^{\text{par}}(t), \quad C \in \mathbf{R} \right\} \stackrel{\text{notation}}{=} \mathcal{S}_0 + y^{\text{par}},$$

où $A : I \rightarrow \mathbf{R}$ est une primitive de a .

Pour résoudre complètement l'équation différentielle (E_1) , il reste donc à déterminer une solution particulière y^{par} de (E_1) .

VARIATION DE LA CONSTANCE. Nous commençons par une méthode qui fonctionne toujours dès que le second membre est continu : la *méthode de variation de la constante*. Il s'agit de chercher une solution de la forme des solutions de (H_1) , où la *constante* C est remplacée par une **fonction** dérivable $t \in I \rightarrow C(t)$. Nous faisons donc *varier* la constante C au sens propre du terme.




Méthode Variation de la constante

Chercher y^{par} sous la forme $t \in I \rightarrow C(t)e^{-A(t)}$, où la fonction $C : I \rightarrow \mathbf{R}$ est dérivable et est à déterminer.

Preuve Si l'on pose $y^{\text{par}}(t) = C(t)e^{-A(t)}$, pour tout $t \in I$, où C est une fonction dérivable sur I , alors :

$$\begin{aligned} y^{\text{par}} \text{ solution de } (E_1) &\iff (y^{\text{par}})' + ay^{\text{par}} = b \\ &\iff (Ce^{-A})' + aCe^{-A} = b \\ &\iff C'e^{-A} - CA'e^{-A} + aCe^{-A} = b \\ &\iff C'e^{-A} - Ca'e^{-A} + aCe^{-A} = b \\ &\iff C'e^{-A} = b \\ &\iff C' = be^A \\ &\iff C \text{ est une primitive de } be^A \text{ sur } I \end{aligned}$$

On est donc ramené à un calcul de primitive. Une fois C déterminé (à une constante additive près!), une solution particulière est donnée sur I par $y^{\text{par}}(t) = C(t)e^{-A(t)}$. Donc la méthode de variation de la constante fonctionne si et seulement s'il existe une primitive à la fonction be^A sur I . Cette dernière existe dès que la $t \rightarrow b(t)e^{A(t)}$ est continue, ce qui est le cas.

Exemple 17 — Résoudre : $y' + 3x^2y = e^{x-x^3}$. 

Nous savons donc à présent résoudre complètement une équation différentielle linéaire d'ordre 1. Lorsque l'on ajoute en plus une *condition initiale*, alors il existe une unique solution.

Théorème ANA.9.18 | Résolution avec condition initiale

Soient $t_0 \in I$ et $y_0 \in \mathbf{R}$. Il existe une et une seule solution au «problème de CAU-

CHY » :

$$\begin{cases} y' + a(t)y = b(t), \\ y(t_0) = y_0 \in \mathbf{R}. \end{cases}$$

De plus,

- ▶ l'unique solution a pour expression

$$\forall t \in I, \quad y(t) = e^{A(t_0)-A(t)} y_0 + e^{-A(t)} \int_{t_0}^t e^{A(u)} b(u) du.$$

- ▶ En particulier, toute solution qui s'annule est identiquement nulle.

En résumé, sans condition initiale on a une infinité de solutions. Avec une condition initiale il y a généralement unicité.

Preuve



UTILISATION DU PRINCIPE DE SUPERPOSITION. La méthode a déjà été présentée dans la section de généralités. La voici appliquée dans le cas de l'ordre 1.

Exemple 18 — Déterminer une solution particulière sur \mathbf{R} de l'équation différentielle :

$$y' - 2y = 3e^t + e^{2t}.$$

On cherche donc une solution particulière de ces deux équations différentielles :

$$y' - 2y = 3e^t, \quad y' - 2y = e^{2t}.$$

- ▶ Pour la première, on trouve par variation de la constante : pour tout $t \in \mathbf{R}$, $y_1(t) = -3e^t$.
- ▶ Dans le second cas $y_2(t) = te^{2t}$.


Par principe de superposition, la somme est alors une solution particulière de l'équation différentielle de départ :

$$y^{\text{par}}(t) = te^{2t} - 3e^t, \quad \forall t \in \mathbf{R}.$$

2.2.3. Équations différentielles linéaires du 2nd ordre à coefficients constants

Définition ANA.9.23 | Définition pour $n = 2$

On appelle *équation différentielle linéaire du second ordre sur \mathbf{R}* toute équation de la forme (E_2) .

 **Cadre**
 Dans toute la suite de cette sous-section, nous considèrerons des équations différentielles à coefficients constants.⁹

C'est donc une équation du type

$$ay'' + by' + cy = d(t) \quad (\text{E}_2)$$

où $(a, b, c) \in \mathbf{R}^* \times \mathbf{R} \times \mathbf{R}$, et $b \in \mathcal{C}^0(I, \mathbf{R})$. Rappelons que nous avons également défini la notion d'équation homogène associée ou encore équation sans second membre associée à (E_2) l'équation suivante :


$$ay'' + by' + cy = b. \quad (\text{H}_2)$$

Définition ANA.9.24 | Équation caractéristique

On introduit également l'équation caractéristique de (E_2) :

$$ar^2 + br + c = 0, \quad \text{d'inconnue } r \in \mathbf{R}. \quad (\text{EC})$$

Remarque 2.6 — Caractère \mathcal{C}^2 des solutions Si y est une solution de (E_2) , alors y'' est continue sur I , puisque $y'' = \frac{1}{a}(f - by' - cy)$. La fonction y est donc un élément de $\mathcal{C}^2(I, \mathbf{R})$.

 **Cadre**
 Dans la suite de cette sous-section, on se fixe une équation différentielle $ay'' + by' + cy = d(t)$, avec $d : I \rightarrow \mathbf{R}$ continue, et $a, b, c \in \mathbf{R}$.

RÉSOLUTION DE L'ÉQUATION HOMOGÈNE Nous savons là encore déterminer facilement l'ensemble des solutions de l'équation homogène.

⁹Ce sont les seules au programme.

Théorème ANA.9.19 | Résolution de l'équation homogène

On suppose ici $(a, b, c) \in \mathbf{R}^3$.

1. Si (EC) possède deux racines réelles distinctes $\alpha \in \mathbf{R}$ et $\beta \in \mathbf{R}$, alors

$$\mathcal{S}_0 = \text{Vect} \left(t \mapsto e^{\alpha t}, t \mapsto e^{\beta t} \right).$$

2. Si (EC) possède une racine double $\alpha \in \mathbf{R}$, alors

$$\mathcal{S}_0 = \text{Vect} \left(t \mapsto e^{\alpha t}, t \mapsto te^{\alpha t} \right) = \left\{ t \mapsto (A + Bt)e^{\alpha t}, (A, B) \in \mathbf{R}^2 \right\}.$$


3. Si (EC) possède deux racines complexes conjuguées $\alpha + i\beta$ et $\alpha - i\beta$ avec $(\alpha, \beta) \in \mathbf{R} \times \mathbf{R}^*$, alors


$$\begin{aligned} \mathcal{S}_0 &= \text{Vect} \left(t \mapsto e^{\alpha t} \cos(\beta t), t \mapsto e^{\alpha t} \sin(\beta t) \right), \\ &= \left\{ t \mapsto e^{\alpha t} (A \cos(\beta t) + B \sin(\beta t)), A, B \in \mathbf{R} \right\} \end{aligned}$$


Ainsi, \mathcal{S}_0 est un espace vectoriel de dimension 2.

Preuve Voir cours de première année. On vérifie également sans difficulté que les familles de solutions proposées sont libres, et donc forment des bases de l'ensemble des solutions de l'homogène.

Exemple 19 — Résoudre

1. $y'' - \omega^2 y = 0$ et $y'' + \omega^2 y = 0$ (où ω est un réel non nul). 

2. $y'' - 4y' + 13y = 0.$ 

3. $y'' - 4y' + 4y = 0.$ 

RÉSOLUTION DE L'ÉQUATION COMPLÈTE On applique encore une fois le théorème déjà démontré sur le sujet : toute solution est obtenue en sommant les solutions de l'homogène et **une** solution particulière.

Théorème ANA.9.20

L'ensemble \mathcal{S} des solutions sur \mathbf{R} de (E_2) est :

$$\mathcal{S} = \{y + y^{\text{par}}, y \in \mathcal{S}_0\} \stackrel{\text{(nota.)}}{=} \mathcal{S}_0 + y^{\text{par}}.$$

Preuve Ce théorème est une conséquence directe du **Théorème ANA.9.15** pour $n = 2$ et de la résolution explicite de l'homogène faite précédemment.

DÉTERMINATION DE y^{par} Pour résoudre complètement l'équation différentielle (E_2) , il reste donc à déterminer *une* solution particulière y^{par} de (E_1) . On ne sait le faire que pour certaines formes du second membre $d(t)$.

Remarque 2.7 — Il existe une méthode générale, comme pour le premier ordre, appelée *méthode de variations des constantes*, mais elle n'est pas à notre programme.

Admettant l'existence d'une solution particulière avec un second membre continu d , on peut démontrer l'existence et l'unicité ci-après.


Théorème ANA.9.21 | Résolution avec condition initiale

Soient $t_0 \in I$, $(y_0, y'_0) \in \mathbf{R}^2$. Il existe une et une seule solution au problème de Cau-

chy :

$$\begin{cases} ay'' + by' + cy = d(t), \\ y(t_0) = y_0, \\ y'(t_0) = y'_0. \end{cases}$$

En particulier, toute solution qui s'annule est identiquement nulle.

Preuve Nous admettons l'existence d'une solution particulière. Faisons la preuve dans le cas « $\Delta > 0$ », les autres étant analogues. 

où P est une fonction polynomiale à coefficients dans \mathbf{R} et $m \in \mathbf{R}$. Alors il existe une solution particulière de (E_2) d'expression :

$$\forall t \in \mathbf{R}, \quad y^{\text{par}}(t) = Q(t)e^{mt},$$

où Q est une fonction polynomiale à coefficients dans \mathbf{R} et :

- ▶ **si m n'est pas racine de (EC)** : $\deg(Q) = \deg(P)$,
- ▶ **si m est racine simple de (EC)** : $\deg(Q) = \deg(P) + 1$,
- ▶ **si m est racine double de (EC)** : $\deg(Q) = \deg(P) + 2$.

Preuve Cherchons à quelle condition (E_1) admet une solution d'expression

$$\forall t \in \mathbf{R}, \quad y^{\text{par}}(t) = Q(t)e^{mt}$$

où Q est une fonction polynomiale. On calcule, pour tout $t \in \mathbf{R}$:

$$\begin{aligned} (y^{\text{par}})'(t) &= (Q'(t) + mQ(t))e^{mt}, \\ (y^{\text{par}})''(t) &= (Q''(t) + 2mQ'(t) + m^2Q(t))e^{mt}. \end{aligned}$$

Cette fonction y^{par} est donc solution de (E_1) si et seulement si pour tout $t \in \mathbf{R}$:

$$a(y^{\text{par}})''(t) + b(y^{\text{par}})'(t) + cy^{\text{par}}(t) = P(t)e^{mt}$$

ce qui, après simplification par e^{mt} et regroupement des termes, équivaut à :

$$aQ''(t) + (2am + b)Q'(t) + (am^2 + bm + c)Q(t) = P(t).$$

Il suffit alors de résoudre cette équation polynomiale en injectant $Q(t) = \sum_{k=0}^n a_k t^k$ dans l'identité avec les a_k des éléments de \mathbf{R} à déterminer [...].


CAS PARTICULIER : SECOND MEMBRE DE LA FORME « POLYNÔME \times EXPONENTIELLE ». Notez que le résultat ci-après, en prenant $a = 0$, fournit aussi une solution particulière lorsque l'équation différentielle est d'ordre 1.


Théorème ANA.9.22 | Solution particulière pour $a \in \mathbf{R}$ une constante, et $b(t) = P(t)e^{mt}$, $P \in \mathbf{R}[X]$, $m \in \mathbf{R}$


On suppose que le second membre de (E_2) est de la forme


$$\forall t \in \mathbf{R}, \quad b(t) = P(t)e^{mt},$$

Exemple 20 — Déterminer une solution particulière réelle des équations différentielles linéaires d'ordre 2 à coefficients constants suivantes :

1. $y'' - 4y = (8t + 6)e^{2t}$. 

2. $y'' - y' - 2y = 3e^t + 1.$ 


3. $y'' - 6y' + 9y = e^{3x}.$ 

4. $y'' - 4y' + 4y = (t^2 + 1)e^{2t}.$ 

PRINCIPE DE SUPERPOSITION. Le principe de superposition s'applique encore pour les équations différentielles linéaires du second ordre à coefficients constants admettant un second membre somme de plusieurs fonctions simples.

2.2.4. Technique du changement de fonction inconnue


On peut aussi parfois réaliser des changements de fonction inconnue dans les équations différentielles pour se ramener à un type que l'on sait résoudre (exemple typique : passer d'une équation différentielle linéaire du second ordre à coefficients non constants à une version à coefficients constants, ou encore, comme dans l'exemple ci-dessous, passer d'une équation différentielle non linéaire à une version linéaire). Voyons cela sur un exemple avant de systématiser dans une méthode.

Exemple 21 — Non linéaire à linéaire Résoudre l'équation différentielle (E) $y' = y \ln y$. Indication : On pourra réaliser le changement de fonction inconnue $y(t) = e^{z(t)}$ pour tout t . 

Résumé

\Rightarrow Soit y une solution de (E). Alors posons $z = \ln \circ y$. On a vérifié que z est solution d'une équation différentielle (E') résoluble.

\Leftarrow Soit z une solution de (E'), alors $y = \exp \circ z$ est une solution de (E). En d'autres termes, il y a une correspondance bijective entre les solutions de (E) et (E') — il suffit donc de résoudre l'une ou l'autre pour toutes les avoir.

Exemple 22 — Non constants à constants Résoudre $(1 + x^2)^2 y'' + 2x(1 + x^2)y' + 4y = 0$ sur $\bigcup_{k \in \mathbb{Z}}]-\frac{\pi}{2} + k\pi, \frac{\pi}{2} + k\pi[$. Indication : On pourra réaliser le changement de fonction inconnue $z(x) = y(\tan x)$ pour tout t . 



3. Évaluer (E) en $\varphi(t)$ pour tout $t \in \mathbf{R}$.
4. Combiner 2) et 3) pour trouver une équation différentielle en z .

2.3. Résolution approchée par la méthode d'EULER

Rappelons que l'on appelle équation différentielle toute équation de la forme

$$y^{(n)} = f(t, y', \dots, y^{(n-1)})$$

où I est un intervalle réel, $f : I \times (\mathfrak{M}_{p,1}(\mathbf{R}))^n \rightarrow \mathfrak{M}_{p,1}(\mathbf{R})$, avec $n, p \in \mathbf{N}^*$. Nous allons dans la suite nous placer dans le cadre $n = 1$, *i.e.* considérer des équations de la forme

$$y' = f(t, y),$$

sur un intervalle réel I , en revanche on autorise y à avoir p coordonnées, ainsi $y = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix}$. On retrouve le cas scalaire pour $p = 1$. On cherche ainsi à connaître une solution approchée d'un problème de Cauchy de la forme

$$\begin{cases} y' = f(t, y), \\ y(t_0) = y_0 \in \mathfrak{M}_{p,1}(\mathbf{R}), \end{cases} \quad I = [0, \tau], \quad f : I \rightarrow \mathbf{R} \text{ continue,}$$

avec $\tau > 0$. Notez bien que cette méthode permet de résoudre, au moins de manière approchée, tout problème de Cauchy du moment que f est une fonction continue (pour que les intégrales ci-dessous aient un sens). En ce sens approché, on fait donc beaucoup mieux que le seul cas linéaire proposé dans les deux précédentes parties.

Méthode Résolution par changement de fonction inconnue



Soit (E) une équation différentielle en une fonction y que l'on ne sait pas résoudre *a priori*.

1. Soit une fonction z dépendant de y donnée par l'exercice (généralement « de la forme $z(t) = y \circ \varphi(t)$ »).
2. Calculer les dérivées successives z, z', z'', \dots (en fonction de l'ordre de l'équation différentielle en y).

PRINCIPE DE LA MÉTHODE. Soit $0 = t_0 < \dots < t_N = \tau$ une subdivision de $[0, \tau]$. L'idée de la méthode est la suivante de construire un nuage de points (t_i, y_i) de proche en proche (donc grâce à une relation de récurrence), tels que ces points y_i soient très proches des $y(t_i)$ (où y désigne l'unique solution au problème de Cauchy). Le nuage de points est construit par « analyse-synthèse » (ou heuristique ici).

**Méthode d'EULER**

1. On commence par subdiviser l'intervalle $[0, \tau]$ de manière uniforme à l'aide de $N + 1$ points espacés d'un pas de $h = \frac{\tau}{N}$. Plus précisément, on pose :

$$t_0 = 0, \quad t_1 = h, \quad t_2 = 2h, \quad \dots, \quad t_k = kh, \quad \dots, \quad t_N = \tau.$$

2. **(Heuristique)** On considère que pour h petit :

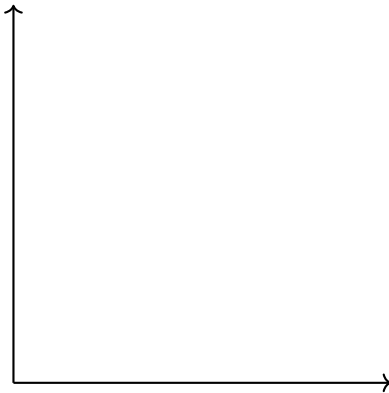
$$\text{Si } y \text{ est une solution, alors } y'(t) = f(t, y(t)) \approx \frac{y(t+h) - y(t)}{h}.$$

Donc que pour toute solution y :

$$y(t+h) \approx y(t) + hf(t, y(t)) \quad (2.1)$$

3. On définit donc une suite de points y_i satisfaisant le plus possible l'approximation précédente, en posant :

$$\forall k \in \llbracket 0, N-1 \rrbracket, \quad y_{k+1} = y_k + h \times f(t_k, y_k).$$



>_🐍 IMPLÉMENTATION EN PYTHON. Nous commençons par donner la fonction en dimension 1.

```
"""
f : fonction, y0 : valeur en zéro, tau : borne max de
↳ l'intervalle,
N : nombre de points de la subdivision->
retourne un couple (abs, solution approchée)
"""
h = tau/N
T = np.linspace(0, tau, N+1)
Y = np.zeros(N+1)
Y[0] = y0
for k in range(N):
    Y[k+1] = Y[k] + h * f(T[k], Y[k])
return T, Y
```

Si l'inconnue y possède k coordonnées, alors c'est pareil, on approche simplement la dérivée de chaque coordonnée.

```
import numpy as np
def euler_vec(f, y0, tau, N):
    """
    f : fonction, y0 : vecteur en zéro, tau : borne max de
    ↳ l'intervalle,
    N : nombre de points de la subdivision->
    retourne un couple (abs, solution approchée)
    solution approchée est une matrice (en chaque point de la
    ↳ subdivision, k
    coordonnées)
    """
    h = tau/N
    T = np.linspace(0, tau, N+1)
    Y = np.zeros((N+1, len(y0)))
    Y[0] = y0
```

```

for k in range(N):
    Y[k+1] = Y[k] + h * f(T[k], Y[k])
return T, Y

```

Exemple 23 – Fonction exponentielle Testons avec la fonction exponentielle, *i.e.* l'équation différentielle $y' = y$.

```

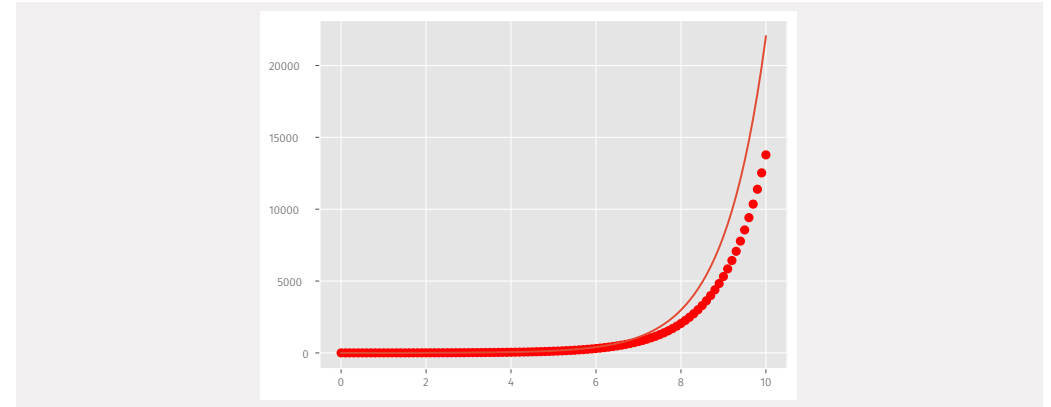
def f(t,x):
    return x

import numpy as np
import matplotlib.pyplot as plt

def euler(f, y0, tau, N):
    """
    f : fonction, y0 : valeur en zéro, tau : borne max de
    l'intervalle,
    N : nombre de points de la subdivision->
    retourne un couple (abs, solution approchée)
    """
    h = tau/N
    T = np.linspace(0, tau, N+1)
    Y = np.zeros(N+1)
    Y[0] = y0
    for k in range(N):
        Y[k+1] = Y[k] + h * f(T[k], Y[k])
    return T, Y

X, sol_app = euler(f, 1, 10, 100)
plt.plot(X, sol_app, 'ro') #sol. approchée
plt.plot(X, np.exp(X)) #solution exacte

```



Exemple 24 – Système proie/prédateur de Lotka & Volterra Reprenons un système déjà introduit dans un exemple précédent, correspondant à une compétition entre deux espèces :

$$\begin{cases} y_1' = (2 - 4y_2)y_1, \\ y_2' = (-2 - 2y_1)y_2. \end{cases}$$

D'après ce que l'on a vu dans la section précédente, si par exemple y_1 correspond à une population de lapins, et y_2 une population de renards, alors nous avons ici :

- ▶ un fort taux de prédation des renards sur les lapins, un faible taux de prédation des lapins sur les renards,
- ▶ des taux de natalité comparables.

On s'attend alors en toute logique à observer une extinction rapide de la population de lapins.

```

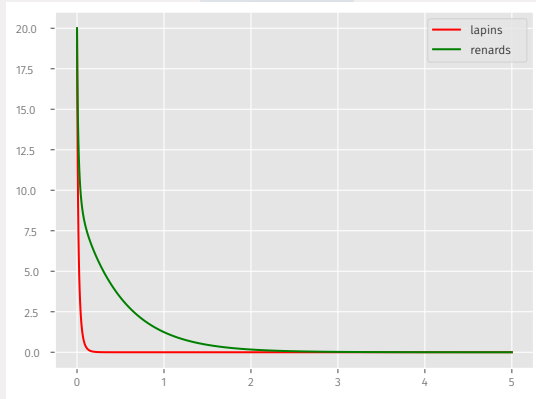
def f(t, x):
    return np.array([(2-4*x[1])*x[0], (-2-2*x[0])*x[1]])

import numpy as np
import matplotlib.pyplot as plt

X, sol_app = euler_vec(f, [20, 20], 5, 10000)

```

```
plt.plot(X, [sol_app[i][0] for i in range(len(sol_app))], 'r',
         label = 'lapins')
plt.plot(X, [sol_app[i][1] for i in range(len(sol_app))],
         'g', label = 'renards')
plt.legend()
```



Commenter le graphe obtenu. 

2.4. Systèmes différentiels linéaires à coefficients constants

Soit $p \geq 1$ un entier. On rappelle qu'un système différentiel en l'inconnue $t \mapsto Y(t) = \begin{pmatrix} y_1(t) \\ \vdots \\ y_p(t) \end{pmatrix}$, est une équation différentielle de la forme

$$Y'(t) = \begin{pmatrix} y_1'(t) \\ \vdots \\ y_p'(t) \end{pmatrix} = A(t) \begin{pmatrix} y_1(t) \\ \vdots \\ y_p(t) \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix},$$

où $b_i \in \mathbf{R}$ pour tout $i \in \llbracket 1, p \rrbracket$. Nous allons voir une méthode afin de résoudre un tel système différentiel linéaire lorsque A est diagonalisable. Mais commençons par un lemme, souvent admis dans les sujets de concours.

Lemme ANA.9.1 |

Soit $t \mapsto X(t)$ un vecteur de fonctions dérivables de $\mathfrak{M}_{p,1}(\mathbf{R})$, et $B \in \mathfrak{M}_{p,p}(\mathbf{R})$ une matrice constante. Alors

$$(B \cdot X)' = B \cdot X'.$$

Preuve Notons $X = \begin{pmatrix} x_1(t) \\ \vdots \\ x_p(t) \end{pmatrix}$ pour tout $t \in \mathbf{R}$, avec x_1, \dots, x_p dérivables. Notons $B = (b_{i,j})_{1 \leq i,j \leq p}$. Alors pour tout t , on a

$$B \cdot X(t) = \begin{pmatrix} \sum_{j=1}^p b_{1,j} x_j(t) \\ \vdots \\ \sum_{j=1}^p b_{p,j} x_j(t) \end{pmatrix} \implies (B \cdot X(t))' = \begin{pmatrix} \sum_{j=1}^p b_{1,j} x_j'(t) \\ \vdots \\ \sum_{j=1}^p b_{p,j} x_j'(t) \end{pmatrix},$$

ce qui prouve

$$(B \cdot X)'(t) = B \cdot X'(t).$$

**Méthode Résolution d'un système différentiel**

Notons $Y'(t) = AY(t)$ le système avec $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix} \in \mathfrak{M}_{p,1}(\mathbf{K})$ un vecteur de fonctions

dérivables, et $A \in \mathfrak{M}_{p,p}(\mathbf{K})$ une matrice diagonalisable à coefficients constants. Pour résoudre le système, on commence par

1. diagonaliser la matrice A . On obtient $A = PDP^{-1}$ avec D diagonale et P inversible.
2. Considérer la nouvelle fonction inconnue $Z = P^{-1}Y$.
3. Établir un système en Z : puisque P est une matrice constante, on a $Z' = P^{-1}Y' = P^{-1}PDP^{-1}Y \implies Z' = DZ$.
4. Résoudre $Z' = DZ$, c'est un système d'équations différentielles indépendantes.
5. On déduit $Y = PZ$.

Exemple 25 — en dimension 3 Soit le système différentiel suivant (en les fonctions inconnues y_1, y_2, y_3 :

$$\begin{cases} y_1' = y_1 + y_2 + y_3, \\ y_2' = 2y_2 + 2y_3, \\ y_3' = 3y_3, \end{cases} \quad y_1, y_2, y_3 \text{ de classe } \mathcal{C}^1.$$

1. Écrire le système précédent à l'aide d'une matrice A et de $Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$.


2. On rappelle que dans l'Exemple 12 du Chapter ALG.5, nous avons diagonalisé la

matrice $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix}$, et avons montré que :

$$A = P \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \cdot P^{-1}, \quad \text{avec } P = \begin{pmatrix} 1 & 1 & \frac{3}{2} \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

3. Notons $Z = P^{-1}X$. Montrer que Y est solution de $Y' = AY$ si et seulement si Z est solution de $Z' = DZ$.

4. En déduire les solutions du système initial.

Attention  Là encore, dans la méthode présentée, il n'y a pas besoin de connaître la matrice P^{-1} , mais uniquement P .

3. MODÉLISATION DE PHÉNOMÈNES DYNAMIQUES

L'interprétation physique du nombre dérivé est une vitesse instantanée. Il est donc possible de décrire des phénomènes d'évolution continue à l'aide d'équations différentielles ou des suites récurrentes pour des évolutions discrètes. Nous nous intéresserons également dans la suite aux modèles décrivant des dynamiques de population.

3.1. Taux d'évolution et dynamiques linéaires

Lorsqu'une grandeur discrète $U = (u_n)$ (une suite) ou continue y (une fonction dérivable) varie, on cherche à quantifier souvent la variation de cette grandeur entre deux instants. Pour le signe, *i.e.* savoir si elle croît ou décroît, on dispose déjà d'outils :

$$u_{n+1} - u_n, n \in \mathbf{N} \quad \text{pour les suites, et } y' \quad \text{pour les fonctions.}$$

On pourrait d'autre part s'intéresser plus précisément à l'amplitude de la variation en un temps donné.

- ▶ **(Discret)** dans le cas discret, si n, p sont deux entiers, alors la variation de (u_n) entre les temps n et $n + p$ est définie comme

$$\Delta_n^{n+p}(U) = u_{n+p} - u_n.$$

- ▶ **(Continu)** dans le cas continu, si t, h sont deux réels positifs, alors la variation de y entre les temps t et $t + h$ est définie comme

$$\Delta_t^{t+h}(y) = y(t+h) - y(t).$$

Plutôt que de regarder des variations comme *supra*, on peut aussi se demander quel est le pourcentage ou taux d'augmentation/diminution par rapport à une valeur antérieure — c'est le principe par exemple des livrets bancaires.

- ▶ **(Discret)** dans le cas discret, si n, p sont deux entiers, alors intuitivement le taux d'évolution de (u_n) entre les temps n et $n + p$ est définie comme le réel $\tau_{n,p}(U)$ tel que

$$u_{n+p} = u_n + \tau_{n,p} u_n$$

donc de manière équivalente

$$\tau_{n,p}(U) = \frac{u_{n+p} - u_n}{u_n}.$$

- ▶ **(Continu)** dans le cas continu, si t, h sont deux réels positifs, alors on définit de même le taux d'évolution de y entre les temps t et $t + h$:

$$\tau_{t,t+h}(y) = \frac{y(t+h) - y(t)}{y(t)}.$$

On peut préférer aussi les versions *instantanées* (*cf.* la définition ci-après).

Définition ANA.9.25 | Taux d'évolution

- ▶ **(Discret)** Soit $U = (u_n)$ une suite. Si n, p sont deux entiers, alors on définit :
 - le *taux d'évolution entre n et p* noté $\tau_{n,p}(U)$ par :

$$\tau_{n,p}(U) = \frac{u_{n+p} - u_n}{u_n}.$$

- le *taux d'évolution en n* noté $\tau_n(U)$ par :

$$\tau_n(U) = \frac{u_{n+1} - u_n}{u_n}.$$

- ▶ **(Continu)** Soit y une fonction dérivable. Si t, h sont deux réels positifs, alors on définit :

— le *taux d'évolution* entre t et $t + h$ noté $\tau_{t,t+h}(y)$ par :

$$\tau_{t,t+h}(y) = \frac{y(t+h) - y(t)}{y(t)}.$$

— le *taux d'évolution en t* noté $\tau_t(y)$ par :


$$\tau_t(y) = \frac{y'(t)}{y(t)} = \lim_{h \rightarrow 0} \frac{\tau_{t,t+h}(y)}{h} \text{ }^{10}.$$


Lorsque y décrit une population d'individus, on parle de *taux de reproduction* (resp. *mortalité*) en t si $\tau(t) \geq 0$ (resp. ≤ 0).

► **(Évolution linéaire)** Dans chacun des cas, on dit que U (ou y) suit une *évolution linéaire* lorsque le taux d'évolution est constant au cours du temps.

Dans la suite de cette section, nous n'utiliserons presque plus que le taux d'évolution en un temps donné. L'avantage de cette quantité est qu'elle possède une interprétation très intuitive, et permet donc de formaliser mathématiquement des situations très diverses d'évolution. Voyons quelques exemples.

Exemple 26 — Dans cet exemple, on considère $U = (u_n)_{n \in \mathbb{N}}$ une suite numérique, $y : \mathbf{R} \rightarrow \mathbf{R}$ une fonction dérivable.

1. Quelles sont les suites U de taux d'évolution constant égal à $\tau \in \mathbf{R}$? Analyser la convergence en temps long, commentez. 

2. Même question pour les fonctions y , on analysera la convergence en temps long, commentez. 

Remarque 3.1 — C'est d'ailleurs la démarche de l'exemple précédent qui poussa à introduire la fonction exponentielle. On dit aussi de manière plus vague que y possède une « croissance exponentielle ».

3.2. Modélisation de systèmes présentant une dynamique


L'utilisation de suites numériques d'une part, et de fonctions d'autre part permet de modéliser de manière très fidèle des systèmes présentant une dynamique simple. De manière formelle, si une grandeur V — discrète ou continue et éventuellement vectorielle — décrit une évolution, alors schématiquement :


- **(Discret)** bilan de V entre n et $n + 1 \implies$ relation de récurrence sur (v_n) .
- **(Continu)** bilan de V entre t et $t + h \implies$ équation différentielle sur V .


¹⁰Par définition du nombre dérivé

Faisons ce travail de bilan de grandeur, dans plusieurs contextes.

Exemple 27 — Modélisation Modéliser les situations ci-après à l'aide d'une suite ou d'une fonction. Commentez.

1. (Reproduction cellulaire très simplifiée) On considère une population de cellules, qui se reproduisent tous les ans. On suppose que chaque cellule donne alors lieu à deux cellules à l'année suivante. 


2. (Cinétique chimique d'ordre 1) On considère une réaction chimique notée $A \rightarrow B$, on suppose que le réactif A disparaît de manière proportionnelle au temps et à la concentration $[A]$ en réactif A. 

3. (Évolution radioactive) On considère une population d'atomes de carbone 14. On suppose qu'en tout temps $t \in \mathbf{R}^+$, une proportion $\rho_t \in \mathbf{R}^+$ d'atomes se désintègrent et une proportion $t\rho_b \in \mathbf{R}^+$ se crée. 

4. (Deux populations à évolution liée) On considère une population de tortues, composée de bébés (âgés d'un an, population décrite¹¹ par une suite (b_n)) d'une part et d'adultes (âgés de deux ans ou plus, associés à une suite (a_n)) d'autre part.

On suppose que

- ▶ les tortues deviennent adultes à 2 ans, et que seules 20% parviennent à cet âge,
- ▶ 40% des tortues adultes de l'année n meurent avant la fin de l'année,
- ▶ les femelles composent la moitié de la population et donnent naissance à 4 bébés chaque année, de l'âge de 2 ans jusqu'à la fin de leur vie.

On déterminera a_{n+2} en fonction de a_{n+1} et b_n , ainsi que b_{n+1} en fonction de a_n 

Nous verrons d'autres contextes en TD.

3.3. Dynamique des populations

3.3.1. Modèle malthusien : évolution libre

Il est largement considéré que c'est au crédit de Thomas. R. MALTHUS que l'on accorde la paternité de l'un des premiers modèles mathématiques de croissance de population, qui publie (anonymement, dans sa première version) son célèbre *Essai sur le principe de population* en 1798.

¹¹Je n'emploie pas le mot cardinal car b_n n'est pas forcément un entier

Selon MALTHUS, cf. la citation présentée en début de chapitre, la croissance d'une population aurait un ratio géométrique, *a contrario* de la croissance des ressources qui serait arithmétique. MALTHUS ne vérifia pas rigoureusement cette théorie de rapport de croissances, prenant comme argument d'autorité l'évidence de ses propos. MALTHUS estime que la population mondiale double toutes les 25 années - fatalement, la population tendra rapidement à s'accroître au-delà des moyens de subsistance, engendrant selon lui pléthores de conséquences plutôt dévastatrices (guerres, famines, épidémies, ...).

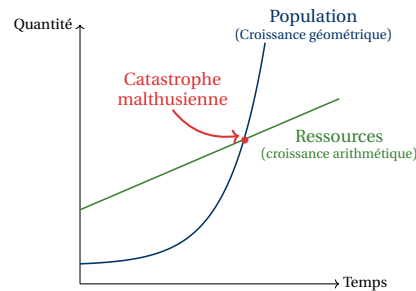


FIG. ANA.9.5. : Représentation du modèle de MALTHUS. Lorsque la quantité de population (augmentation géométrique) dépasse celle des ressources (augmentation arithmétique), la catastrophe malthusienne s'enclenche.

L'appréhension d'une telle catastrophe démographique associée à une préconisation de la limitation du nombre de naissances porte désormais le nom de *malthusianisme*. Notons que, comme le souligne l'anthropologue C. Meillassoux, cette peur d'une croissance excessive au delà des moyens de subsistances est complètement irréaliste : en réalité, comment une population pourrait-elle continuer à croître exponentiellement en ayant épuisé les ressources nécessaires à son développement ? MALTHUS le reconnaît d'ailleurs lui-même (traduction en Français) : « Je sais bien, que les millions d'habitants en excès dont j'ai parlé n'existeront jamais ». Mais cette pensée malthusienne eut tout de même des conséquences importantes, comme par exemple la politique de l'enfant unique en Chine. Présentons à présent une description mathématique.

On suppose que la population grandit en se multipliant par un nombre fixe β appelé *taux de reproduction* ou *taux de fertilité*, et meurt selon un *taux de mortalité* $\mu > 0$ supposés constant ici. Ainsi, si P désigne le cardinal¹² de la population.

¹²abus de vocabulaire, puisque P n'a aucune raison d'être un entier positif

BILAN DE POPULATION MALTHUSIEN

► (modèle discret)



► (modèle continu)



On débouche alors sur la définition ci-après.

Définition ANA.9.26 | Modèle de MALTHUS

- ▶ On dit qu'une suite (p_n) suit un *modèle malthusien de taux β, μ , condition initiale $P_0 \in \mathbf{R}$* si (p_n) vérifie :

$$p_{n+1} = p_n + (\beta - \mu)p_n \quad (n \in \mathbf{N}), \quad p_0 = P_0. \quad (\text{Malthus, disc})$$

- ▶ On dit qu'une fonction P suit un *modèle malthusien de taux β, μ , condition initiale $P_0 \in \mathbf{R}$* si P est dérivable en t et vérifie :

$$P'(t) = (\beta - \mu)P(t) \quad (t \geq 0), \quad P(0) = P_0. \quad (\text{Malthus})$$

On appelle β le taux de natalité, μ le taux de mortalité. Le taux d'évolution est alors $\beta - \mu$.¹³

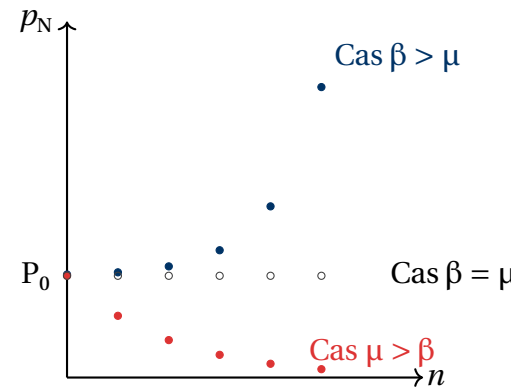


FIG. ANA.9.6. : Modèle malthusien discret

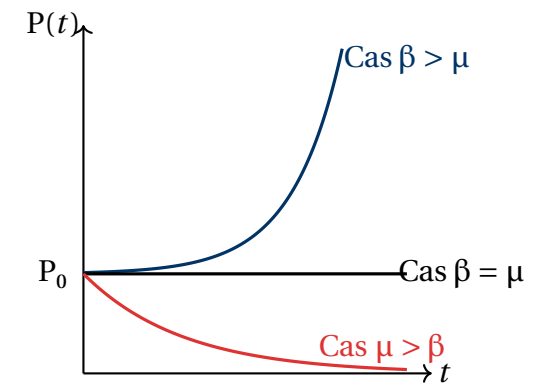


FIG. ANA.9.7. : Modèle malthusien continu

Ce modèle très simpliste met en évidence un point très important : l'évolution d'une population est dictée par la balance entre taux de fertilité et taux de mortalité, que ce soit dans le cas discret ou le cas continu.

De manière équivalente, cela signifie que :

- ▶ **(Cas discret)** $p_n = (1 + \beta - \mu)^n P_0$ pour tout $n \in \mathbf{N}$,
- ▶ **(Cas continu)** $P(t) = P_0 e^{(\beta - \mu)t}$ pour tout $t \geq 0$.

Ici nous ne supposons donc pas l'existence de prédateurs, et que les ressources naturelles sont en quantité illimitée. Ainsi la population a donc la possibilité de se développer indéfiniment.

¹³Conséquence directe de la définition

3.3.2. Modèle logistique de VERHULST : évolution sous capacité de milieu

En 1838, Pierre-François VERHULST répond à MALTHUS en proposant un modèle *logistique* de dynamique de population. À l'instar du modèle de MALTHUS, il suppose qu'une population sans limitation de ressources croît de manière exponentielle mais que la croissance de la population est freinée par sa propre dynamique et par la limitation des ressources du milieu.

Dans sa note, VERHULST suppose la résistance à la croissance d'une population proportionnelle au carré de la vitesse avec laquelle la population tend à croître (à l'instar d'un mobile en chute libre traversant un milieu résistant — cette intuition fut d'ailleurs fournie à VERHULST par le physicien QUETELET). Le bilan de population est quant à lui le même que pour MALTHUS, avec simplement l'ajout d'un terme proportionnel à $-P(t)^2$ (ou $-p_n^2$). À renommage de variables près, on arrive directement sur la définition ci-après.

Définition ANA.9.27 | Modèle de VERHULST

- ▶ On dit qu'une suite (p_n) suit un *modèle de VERHULST de taux* β, μ, κ , *condition initiale* $P_0 \in \mathbf{R}$ si (p_n) vérifie :

$$p_{n+1} = p_n + (\beta - \mu) \left(p_n - \frac{p_n^2}{\kappa} \right) \quad (n \in \mathbf{N}), \quad p_0 = P_0. \quad (\text{Verhulst, disc})$$

- ▶ On dit qu'une fonction P suit un *modèle de VERHULST de taux* β, μ, κ , *condition initiale* $P_0 \in \mathbf{R}$ si P est dérivable en t et vérifie :

$$P'(t) = (\beta - \mu) \left(P(t) - \frac{P(t)^2}{\kappa} \right) \quad (t \geq 0), \quad P(0) = P_0. \quad (\text{Verhulst})$$

On appelle β le taux de natalité, μ le taux de mortalité, et κ la capacité du milieu¹⁴.

Pour le modèle de VERHULST, il n'est pas du tout aisé d'obtenir une expression de (p_n) en fonction de n . En revanche dans le cas continu, si $P_0 > 0$, nous pouvons établir (cf. TD) que :

$$P(t) = \frac{\kappa}{1 + e^{-(\beta-\mu)t} \left(\frac{\kappa}{P_0} - 1 \right)}.$$

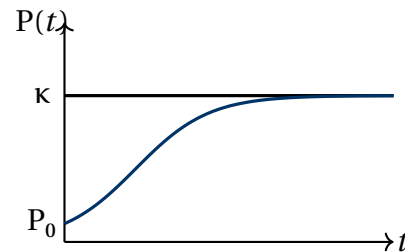


FIG. ANA.9.8 : Modèle de VERHULST continu

Définition ANA.9.28 | Modèle de GOMPERTZ

- ▶ On dit qu'une suite (p_n) suit un *modèle de GOMPERTZ de taux* β, μ, κ , *condition initiale* $P_0 \in \mathbf{R}$ si (p_n) vérifie :

$$p_{n+1} = p_n + (\beta - \mu) \ln \left(\frac{\kappa}{p_n} \right) p_n \quad (n \in \mathbf{N}), \quad p_0 = P_0. \quad (\text{Gompertz, disc})$$

- ▶ On dit qu'une fonction P suit un *modèle de GOMPERTZ de taux* β, μ, κ , *condition initiale* $P_0 \in \mathbf{R}$ si P est dérivable en t et vérifie :

$$P'(t) = (\beta - \mu) \ln \left(\frac{\kappa}{P(t)} \right) P(t) \quad (t \geq 0), \quad P(0) = P_0. \quad (\text{Gompertz})$$

On appelle β le taux de natalité, μ le taux de mortalité, et κ la capacité du milieu.

Là encore, on ne peut pas obtenir d'expression explicite pour le modèle discret. En revanche dans le cas continu, si $P_0 > 0$, nous pouvons établir (cf. TD) que :

$$P(t) = \kappa e^{\ln \left(\frac{P_0}{\kappa} \right) e^{-(\beta-\mu)t}}.$$

Nous ferons la résolution explicite en TD.

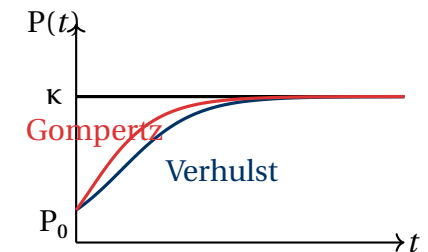


FIG. ANA.9.9 : Modèle de GOMPERTZ continu

3.3.3. Modèle de GOMPERTZ

Un modèle ressemblant au précédent, seule la vitesse de convergence vers la valeur limite est modifiée ainsi que la pente initiale.

¹⁴C'est donc ce terme supplémentaire qui vient freiner la croissance de la population si $\beta - \mu > 0$, et accélérer la croissance si $\beta - \mu < 0$

3.3.4. Modèle proies-prédateurs de LOTKA-VOLTERRA : compétition entre deux populations

Si deux espèces dont les populations sont représentées par P_1 et P_2 se partagent le milieu, on peut adapter le modèle de VERHULST pour tenir compte de cette compétition.

Définition ANA.9.29 | Modèle de Lotka-Volterra

On dit qu'une fonction P suit un *modèle de Lotka-Volterra de taux* β, μ, κ si P est dérivable en t et vérifie :

$$\begin{cases} P_1'(t) = ((\beta_1 - \mu_1) - \pi_1 P_2(t)) P_1(t), \\ P_2'(t) = ((\beta_2 - \mu_2) - \pi_2 P_1(t)) P_2(t). \end{cases} \quad (\text{LoktaVolt})$$

On appelle β_1 le taux de natalité, μ_1 le taux de mortalité pour la première (avec β_2, μ_2 pour la seconde), et π_1, π_2 les taux de prédation.

Remarque 3.2 — Interprétation du système Nous annotons chacun des termes présents dans le système. 

$$\begin{cases} P_1'(t) = ((\beta_1 - \mu_1) - \pi_1 P_2(t)) P_1(t), \\ P_2'(t) = ((\beta_2 - \mu_2) - \pi_2 P_1(t)) P_2(t). \end{cases} \quad (3.1)$$

En plus de la non-linéarité, nous avons à travailler ici sur une équation différentielle dans $\mathfrak{M}_{2,1}(\mathbf{R})$. Nous étudierons plus en détail ce système en séance d'Informatique en approchant les solutions grâce à la méthode d'EULER.

*** **Fin du chapitre** ***

4. EXERCICES

Exercice ANA.9.1 | Vrai ou Faux?

- $\left(1 + \frac{3}{n}\right)^n \xrightarrow{n \rightarrow \infty} 1$
- Une suite bornée converge.
- La partie entière d'une suite réelle convergente est convergente.
- Une suite géométrique bornée converge.
- Soit f une fonction solution de $y' = 3y$. Alors pour tout x , il existe $C \in \mathbf{R}$ telle que $f(x) = Ce^{2x}$.
- Il y a une unique solution de $y' = xy$ avec $y(0) = 1$.
- Les solutions de $y'' - 2y'(x) + 2y = 0$ sont toutes bornées.
- La suite provenant de la méthode d'EULER appliquée à une équation différentielle linéaire d'ordre 1 est arithmético-géométrique.

4.1. Suites numériques

Exercice ANA.9.2 | Étudier la convergence des suites (u_n) définies par :

- $u_n = \frac{n^2 + n - 1}{n^2 - n \cos n}$ pour $n \geq 1$,
- $u_n = \frac{n^2(\ln(n+1) - \ln n)}{\sqrt{n^2 + 1}}$ pour $n \geq 1$,
- $u_n = n - \sqrt{n^2 + (-1)^n}$,
- $u_n = \frac{2^n - 3^n}{2^n - 3^{n+1}}$,
- $u_n = n \left(2 + \cos\left(\frac{n\pi}{12}\right)\right)$,
- $u_n = \left(\cos\left(\frac{1}{3n}\right)\right)^{3n^2 + 1}$ pour $n \geq 1$.

Solution (exercice ANA.9.2)

$$1. u_n = \frac{n^2 \left(1 + \frac{1}{n} - \frac{1}{n^2}\right)}{n^2 \left(1 - \frac{1}{\cos n}\right)} \xrightarrow{n \rightarrow \infty} \boxed{1}.$$

2. Pour tout $n \geq 1$,

$$u_n = \frac{n^2}{\sqrt{n^2 + 1}} \ln\left(1 + \frac{1}{n}\right) \underset{n \rightarrow \infty}{\sim} \frac{n^2}{\sqrt{n^2 + 1}} \frac{1}{n} \underset{n \rightarrow \infty}{\sim} \frac{n^2}{n^2} \xrightarrow{n \rightarrow \infty} \boxed{1}.$$

- $u_n = n - \sqrt{n^2 + (-1)^n} = n \sqrt{1 + \frac{(-1)^n}{n^2}} \xrightarrow{n \rightarrow \infty} \boxed{\infty}$ par règles usuelles sur les limites.
- On met en facteur les termes «les plus lourds» au numérateur et dénominateur.

$$u_n = \frac{3^n \left(\frac{2}{3}\right)^n - 1}{3^{n+1} \frac{1}{3} \left(\frac{2}{3}\right)^n - 1} \xrightarrow{n \rightarrow \infty} \boxed{\frac{1}{3}}.$$

5. On sait que le cosinus est borné par 1, donc

$$\forall n \in \mathbf{N}, \quad 1 \leq 2 + \cos\left(\frac{n\pi}{12}\right) \leq 3,$$

d'où

$$\forall n \in \mathbf{N}, \quad n \leq u_n.$$

Il vient $\boxed{u_n \xrightarrow{n \rightarrow \infty} \infty}$ par théorème de divergence par minoration.

- Faisons un développement limité. Constatons que la suite est bien définie puisque $0 \leq \frac{1}{3n} \leq 1$ et donc $\cos\left(\frac{1}{3n}\right) > 0$ pour tout entier $n \in \mathbf{N}$.

$$\begin{aligned} u_n &= \left(\cos\left(\frac{1}{3n}\right)\right)^{3n^2 + 1} \\ &= \exp\left((3n^2 + 1) \ln \cos\left(\frac{1}{3n}\right)\right) \\ &= \exp\left((3n^2 + 1) \ln\left(1 - \frac{1}{18n^2} + o\left(\frac{1}{n^2}\right)\right)\right) \\ &\xrightarrow{n \rightarrow \infty} \boxed{e^{-\frac{1}{6}}}. \end{aligned}$$

Exercice ANA.9.3 | Suite récurrentes linéaires d'ordre deux Exprimer u_n en fonction de n pour tout n et déterminer la limite de :

1. $u_0 = 1, u_1 = 9, u_{n+2} = u_{n+1} - \frac{1}{4}u_n$. ➤ Proposer un programme Python traçant les 100 premiers termes.
2. $u_0 = -1, u_1 = 5, u_{n+2} + u_n = u_{n+1}$,
3. $u_0 = 1, u_1 = 2, u_n = \frac{2}{\frac{1}{u_{n-1}} + \frac{1}{u_{n-2}}}$, pour tout $n \geq 2$,
4. $u_0 = 1, u_1 = 2, 6u_n u_{n+1} - 5u_n u_{n+2} + u_{n+1} u_{n+2} = 0$.

Exercice ANA.9.4 | Deux récurrences non classiques

1. Soit E l'ensemble des suites réelles $(x_n)_{n \in \mathbf{N}}$ vérifiant : $\forall n \in \mathbf{N}, x_{n+1} + x_n = n$.
 - 1.1) Trouver une suite arithmétique élément de E. On la notera (y_n) .
 - 1.2) Trouver toutes les suites de E, on pourra former une relation de récurrence sur la suite $(x_n - y_n)_n$, si $(x_n) \in E$.
 - 1.3) Donner un équivalent d'une suite (x_n) de E.
2. On considère la suite (u_n) définie par

$$u_0 = \frac{2}{3}, \quad \forall n \in \mathbf{N}, \quad u_{n+1} = \frac{u_n}{2} + \frac{n}{2\sqrt{2}} + \frac{1}{\sqrt{2}}.$$
 - 2.1) En étudiant (v_n) définie par $v_n = u_n \times \sqrt{2} - n$ pour tout n , donner l'expression de u_n en fonction de n .
 - 2.2) Calculer $S_n = \sum_{i=0}^n u_i$ en fonction de n .

Solution (exercice ANA.9.4)

1. Soit E l'ensemble des suites réelles $(x_n)_{n \in \mathbf{N}}$ vérifiant : $\forall n \in \mathbf{N}, x_{n+1} + x_n = n$.
 - 1.1) On cherche (y_n) sous la forme $y_n = y_0 + nr$ avec $r \in \mathbf{R}$. En remplaçant, on obtient :

$$\forall n \in \mathbf{N}, y_0 + (n+1)r + y_0 + nr = n \iff (2y_0 + r) + 2nr = n.$$

Les membres étant des polynômes en n , on déduit :

$$r = \frac{1}{2}, \quad y_0 = -\frac{1}{4}.$$

La seule suite arithmétique de (E) est donc

$$(y_n) = \left(-\frac{1}{4} + \frac{n}{2} \right).$$

- 1.2) Soit $n \in \mathbf{N}$. On a

$$\begin{cases} x_{n+1} = -x_n + n, \\ y_{n+1} = -y_n + n. \end{cases}$$

Donc en faisant la différence

$$x_{n+1} - y_{n+1} = -(x_n - y_n),$$

donc la suite $(x_n - y_n)$ est géométrique de raison -1 , et

$$x_n - y_n = (-1)^n (x_0 - y_0) = (-1)^n \left(x_0 + \frac{1}{4} \right),$$

cela implique

$$x_n = (-1)^n \left(x_0 + \frac{1}{4} \right) + -\frac{1}{4} + \frac{n}{2}.$$

- 1.3) En mettant $\frac{n}{2}$, on montre sans difficulté que $x_n \underset{n \rightarrow \infty}{\sim} \frac{n}{2}$ pour toute suite $(x_n) \in E$.
2. Attention, (u_n) n'est pas du tout arithmético-géométrique.
 - 2.1) Soit $n \in \mathbf{N}$. Alors :

$$\begin{aligned} v_{n+1} &= \sqrt{2} \cdot u_{n+1} - (n+1) \\ &= \sqrt{2} \left(\frac{u_n}{2} + \frac{n}{2\sqrt{2}} + \frac{1}{\sqrt{2}} \right) - (n+1) \\ &= \frac{u_n}{\sqrt{2}} + \frac{n}{2} + 1 - n - 1 \\ &= \frac{v_n + n}{\sqrt{2}} + \frac{n}{2} - n \\ &= \frac{v_n}{2} + n - n = \frac{1}{2}v_n. \end{aligned}$$

Donc (v_n) est géométrique de raison $\frac{1}{2}$, et donc $v_n = \frac{v_0}{2^n} = \frac{2\sqrt{2}}{3} = \frac{\sqrt{2}}{3 \cdot 2^{n-1}}$ pour tout $n \in \mathbf{N}$. Ainsi, on déduit que

$$\forall n \in \mathbf{N}, u_n = \frac{1}{3 \cdot 2^{n-1}} + \frac{n}{\sqrt{2}}.$$

$$\begin{aligned}
 2.2) \quad S_n &= \sum_{i=0}^n \left(\frac{1}{3 \cdot 2^{i-1}} + \frac{i}{\sqrt{2}} \right) \\
 &= \frac{2}{3} \sum_{i=0}^n \frac{1}{2^i} + \frac{1}{\sqrt{2}} \frac{n(n+1)}{2} \\
 &= \frac{2}{3} \frac{1 - \frac{1}{2^{n+1}}}{\frac{1}{2}} + \frac{1}{\sqrt{2}} \frac{n(n+1)}{2} \\
 &= \boxed{\frac{4}{3} \left(1 - \frac{1}{2^{n+1}} \right) + \frac{1}{\sqrt{2}} \frac{n(n+1)}{2}}.
 \end{aligned}$$

Exercice ANA.9.5 | Approximation de l'exponentielle Soit $x \in \mathbf{R}$. On souhaite trouver un moyen algorithmique permettant d'obtenir une valeur approchée de e^x . Pour cela, introduisons une suite $(u_n(x))$ définie par :

$$\forall n > -x, \quad u_n(x) = \left(1 + \frac{x}{n} \right)^n.$$


1. Déterminer un développement asymptotique de $(u_n(x))$ à l'ordre 1, puis montrer que $u_n(x) \xrightarrow{n \rightarrow \infty} e^x$. *Ce résultat pourra être utilisé dans la suite.*
2. On introduit une seconde suite $(v_n(x))$ définie par :

$$\forall n \in \mathcal{N}, \quad v_n(x) = \left(1 - \frac{x}{n} \right)^{-n},$$

où \mathcal{N} est un ensemble à préciser.



- 2.1) Justifier que les suites $(u_n(x)), (v_n(x))$ sont définies pour $n > |x|$.
- 2.2) Montrer par récurrence, que pour tout $n \in \mathbf{N}, y > -1$,

$$(1 + y)^n \geq 1 + ny.$$

- 2.3)  Conjecturer les variations des suites à l'aide de Python.
- 2.4) En montrant que pour $n \in \mathbf{N}, n > |x|$,

$$u_{n+1}(x) = \left(1 + \frac{x}{n} \right)^{n+1} \left(1 - \frac{x}{n(n+1) \left(1 + \frac{x}{n} \right)} \right)^{n+1},$$

établir que $(u_n(x))_{n > |x|}$ est croissante.

- 2.5) En faisant le lien entre $v_n(x)$ et $u_n(-x)$ pour tout n, x vérifiant $n > |x|$, donner la monotonie de $(v_n(x))$.
- 2.6) Conclure quant à la convergence des suites $(u_n(x))$ et $(v_n(x))$.
- 2.7)  En déduire une fonction Python d'en-tête `approx_expo(x, prec)` qui fournit une valeur approchée de e^x à précision `prec`.
- 2.8)  Vérifier le programme précédent à l'aide du module `math` ou `numpy`.

Solution (exercice ANA.9.5)

1. On utilise les formules de développements limités usuels.

$$\begin{aligned}
 u_n(x) &= \left(1 + \frac{x}{n} \right)^n \\
 &= \exp \left(n \ln \left(1 + \frac{x}{n} \right) \right) \\
 &\underset{n \rightarrow \infty}{=} \exp \left(n \left(\frac{x}{n} - \frac{x^2}{n^2} + o \left(\frac{1}{n^2} \right) \right) \right) \\
 &\underset{n \rightarrow \infty}{=} \exp \left(x - \frac{x^2}{n} + o \left(\frac{1}{n} \right) \right) \\
 &\underset{n \rightarrow \infty}{=} e^x \exp \left(-\frac{x^2}{n} + o \left(\frac{1}{n} \right) \right) \\
 &\underset{n \rightarrow \infty}{=} \boxed{e^x \left(-\frac{x^2}{n} + o \left(\frac{1}{n} \right) \right)}.
 \end{aligned}$$

$\left. \begin{array}{l} \exp \left(-\frac{x^2}{n} + o \left(\frac{1}{n} \right) \right) \\ \exp \left(-\frac{x^2}{n} + o \left(\frac{1}{n} \right) \right) \end{array} \right\} \xrightarrow{n \rightarrow \infty} 0$

2. On introduit une seconde suite $(v_n(x))$ définie par :

$$\forall n \in \mathbf{N}, \quad v_n(x) = \left(1 - \frac{x}{n} \right)^{-n}.$$

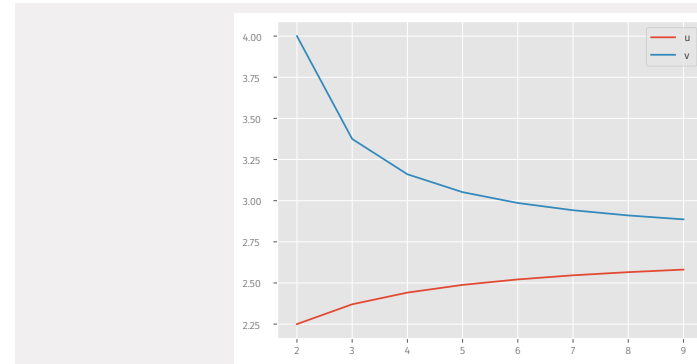
- 2.1) La première est définie dès que $n > -x$. La seconde est définie pour les n tels que $1 - \frac{x}{n} > 0$ donc si et seulement si $n > x$. Or, $n > |x|$ signifie que $-n < x < n$, donc $(u_n(x)), (v_n(x))$ sont définies sur le domaine commun $\boxed{n > |x|}$.

2.2) Soit $y > -1$. La propriété est vraie pour $n = 0$ car $1 \geq 1$. Supposons-la vraie au rang n , alors puisque $1 + y > 0$ on a par hypothèse de récurrence :

$$\begin{aligned} (1 + y)^{n+1} &\geq (1 + ny)(1 + y) \\ &= 1 + ny + x + ny^2 \\ &\geq 1 + (n + 1)y. \end{aligned}$$

```
import math as ma
def calcul_trace(n, x):
    """
    retourne deux listes correspondant à u,v
    """
    n_0 = int(abs(x))+1
    L_u = []
    L_v = []
    for k in range(n_0, n):
        L_u.append((1+x/k)**k)
        L_v.append((1-x/k)**(-k))
    return L_u, L_v
```

```
import matplotlib.pyplot as plt
L_u, L_v = calcul_trace(10, 1)
plt.plot(range(2, 10), L_u, label='u')
plt.plot(range(2, 10), L_v, label='v')
plt.legend()
```



On conjecture donc que $(u_n(x))$ semble croissante et (v_n) décroissante.

2.3) On a en suivant l'indication

$$\begin{aligned} u_{n+1}(x) &= \left(1 + \frac{x}{n+1}\right)^{n+1} \\ &= \left(1 + \frac{x}{n}\right)^{n+1} \left(1 - \frac{x}{n(n+1)\left(1 + \frac{x}{n}\right)}\right)^{n+1} \\ &\geq \left(1 + \frac{x}{n}\right)^{n+1} \left(1 - \frac{x}{n\left(1 + \frac{x}{n}\right)}\right)^{n+1} \quad \left. \begin{array}{l} \text{par une question précédente} \\ \curvearrowright \end{array} \right. \\ &= u_n(x). \end{aligned}$$

Donc $(u_n(x))_{n>|x|}$ est croissante.

2.4) On a $v_n(x) = \frac{1}{u_n(-x)}$, donc $(v_n(x))$ est décroissante comme inverse d'une suite croissante. Il reste enfin à montrer la convergence vers zéro, c'est immédiat car il s'agit d'une différence de suites convergent vers e^x .


2.5) >_☞

```
import math as ma
def approx_expo(x, prec):
    """
    retourne deux listes correspondant à u,v
    """
```

```

k = int(abs(x))+1
u = (1+x/k)**k
v = (1-x/k)**(-k)
while v - u > prec:
    k += 1
    u = (1+x/k)**k
    v = (1-x/k)**(-k)
return ((1+x/k)**k+(1-x/k)**(-k))/2

```

2.6)  On peut alors comparer

```

>>> import math as ma
>>> ma.exp(1)
2.718281828459045
>>> approx_expo(1, 10**(-3))
2.7182819969815766

```

Exercice ANA.9.6 | Une suite récurrente avec « f » décroissante On considère la fonction f définie sur \mathbf{R}_+^* par : $\forall x \in \mathbf{R}_+^*, f(x) = \frac{e^{-x}}{x}$. On considère également la suite $(u_n)_{n \in \mathbf{N}}$ définie par $u_0 = 1$ et par la relation $u_{n+1} = f(u_n)$, valable pour tout entier naturel n .

1. **1.1)** Dresser le tableau de variation de f , limites comprises.
- 1.2) Vérifier que chaque terme de la suite $(u_n)_{n \in \mathbf{N}}$ est parfaitement défini et strictement positif.
2. Les scripts suivants renvoient, 5 pour programme1 et 6 pour programme2. Que sait-on de u_5 et u_6 ? Quelle conjecture peut-on émettre sur le comportement de la suite $(u_n)_{n \in \mathbf{N}}$?

```

import math as ma
def programme1():
    u = 1

```

```


n = 0
while u > 0.00001:
    u = ma.exp(-u)/u
    n += 1
return n

import math as ma
def programme2():
    u = 1
    n = 0
    while u < 100000:
        u = ma.exp(-u)/u
        n += 1
    return n

```

3. **3.1)** étudier les variations de la fonction g définie sur \mathbf{R}_+ par :

$$\forall x \in \mathbf{R}_+, g(x) = e^{-x} - x^2$$

- 3.2) En déduire que l'équation $f(x) = x$, d'inconnue x , possède une seule solution, que l'on notera α , sur \mathbf{R}_+^* .
- 3.3)  Montrer que $\frac{1}{e} < \alpha < 1$. Écrire un programme qui calcule une valeur approchée de α , en utilisant l'algorithme de dichotomie.
4. **4.1)** établir les deux inégalités : $u_2 > u_0$ et $u_3 < u_1$.
- 4.2) En déduire les variations des suites $(u_{2n})_{n \in \mathbf{N}}$ et $(u_{2n+1})_{n \in \mathbf{N}}$.
5. On pose : $h(x) = \begin{cases} (f \circ f)(x) & \text{si } x > 0, \\ 0 & \text{si } x = 0. \end{cases}$
 - 5.1) Déterminer $h(x)$ pour tout réel x strictement positif et vérifier que h est continue en 0.
 - 5.2) Résoudre l'équation $h(x) = x$, d'inconnue x élément de \mathbf{R}_+ .
 - 5.3) En déduire la limite de la suite $(u_{2n+1})_{n \in \mathbf{N}}$.
 - 5.4) Montrer par l'absurde que la suite $(u_{2n})_{n \in \mathbf{N}}$ diverge puis donner $\lim_{n \rightarrow +\infty} u_{2n}$.

Solution (exercice ANA.9.6)

1. 1.1) La fonction f est définie dérivable (en tant que quotient de telles fonctions dont le dénominateur ne s'annule pas) sur \mathbf{R}^{++} . De plus, pour tout $x \in \mathbf{R}^{++}$, $f'(x) = \frac{-e^{-x}x - e^{-x}}{x^2} = -\frac{e^{-x}(x+1)}{x^2}$. On déduit alors le tableau de variations de f à l'aide du signe de $x \mapsto -(x+1) = -x-1$.

x	$-\infty$	-1	0	$+\infty$
$f(x)$	$-\infty$	$-e$	$+\infty$	0

Les limites s'obtiennent par opérations classiques sur les limites et par croissances comparées. En effet,

$$\lim_{x \rightarrow -\infty} f(x) = -\lim_{y \rightarrow \infty} \frac{e^y}{y} = -\infty,$$

par croissances comparées. Les limites en zéro découlent d'opérations usuelles sur les limites, celle en $+\infty$ de $\lim_{x \rightarrow \infty} e^{-x} = 0$.

- 1.2) Il s'agit de vérifier que la suite ne s'annule jamais, et donc que l'on reste dans \mathcal{D}_f à chaque itération, i.e. $f(\mathbf{R}^{++}) \subset \mathbf{R}^{++}$. Mais comme f est continue, le théorème des valeurs intermédiaires garantit que $f(\mathbf{R}^{++}) =]\lim_{x \rightarrow -\infty} f(x), \lim_{x \rightarrow 0} f(x)[=$

\mathbf{R}^{++} . Donc : la suite est bien définie.

2. Les boucles while s'arrêtent dès que u_n est très grand ou très petit. Pour programme1, le résultat donne $u_5 \leq 0.00001$. Pour programme2, nous avons $u_6 \geq 100000$. On peut conjecturer le comportement des sous-suites extraites des termes pairs et impairs : les pairs convergent peut-être vers 0, et les impairs divergent peut-être vers $+\infty$.

3. 3.1) La fonction g est dérivable sur \mathbf{R}^+ en tant que différence de telles fonctions. Soit $x \in \mathbf{R}_+^*$. Alors

$$g'(x) = -e^{-x} - 2x = -xf(x) - 2x = -x(f(x) + 2).$$

Puisque f est strictement positive sur \mathbf{R}^{++} (donc $f + 2$ aussi) d'après le tableau de variations, la fonction g' est quant à elle négative sur \mathbf{R}^{++} . Donc :


g est une fonction strictement décroissante sur \mathbf{R}^+ .

- 3.2) Déduisons que l'équation $f(x) = x$, d'inconnue x , possède une seule solution, que l'on notera α , sur \mathbf{R}_+^* . On souhaite ici montrer l'existence et l'unicité, il faut donc chercher à appliquer le théorème de la bijection sur un intervalle bien choisi.

En effet, remarquons déjà que pour tout $x \in \mathbf{R}^{++}$, $f(x) = x \iff e^{-x} = x^2 \iff g(x) = 0$. Or, g est une fonction continue sur \mathbf{R}^{++} , elle est de plus strictement décroissante et $\lim_{x \rightarrow 0} g(x) = 1$, $\lim_{x \rightarrow \infty} g(x) = \lim_{x \rightarrow \infty} g(x) = -\infty$.

Ainsi, g réalise une bijection de \mathbf{R}^{++} vers $] -\infty, 1]$, et comme $0 \in] -\infty, 1]$,

il existe un unique $\alpha \in \mathbf{R}^{++}$ tel que $g(\alpha) = 0$ donc tel que $f(\alpha) = \alpha$.

- 3.3)  Montrer que $\frac{1}{e} < \alpha < 1$, i.e. que :

$$g\left(\frac{1}{e}\right) > 0, \quad g(1) < 0.$$

Nous avons $f(1) = \frac{1}{e} < 1$ donc $g(1) < 0$. Et $g\left(\frac{1}{e}\right) = f\left(\frac{1}{e}\right) - \frac{1}{e} = e^{1-\frac{1}{e}} - e^{-1} = e^{-1} \left(e^{2-\frac{1}{e}} - 1 \right) > 0$ car $2 - \frac{1}{e} > 1$. Donc : $\frac{1}{e} < \alpha < 1$. Cette inégalité va nous permettre d'initialiser l'algorithme de Dichotomie ci-après.

```
import math as ma
def g(x):
    return ma.exp(-x)-x**2
def dichotomie(prec):
    """
    retourne une valeur approchée de alpha à précision eps
    via l'algorithme de dichotomie
    """
    a = 1/ma.exp(1)
    b = 1
    while b-a > prec:
        c = (a+b)/2
        if g(a)*g(c) <= 0:
            b = c
        else:
            a = c
```

`return (a+b)/2`

Par exemple pour `prec = 10*(-3)`, on obtient 0.7033848354324981.

4. 4.1) Nous avons $u_2 = \frac{e^{-e^{-1}}}{e^{-1}} = e^{1-e^{-1}}$, il s'agit donc de vérifier que $1 - e^{-1} > 1$, ce qui est équivalent à $1 > e^{-1}$. L'égalité est vérifiée. Pour la seconde, $u_3 = \frac{e^{-e^{1-e^{-1}}}}{e^{1-e^{-1}}}$ il s'agit de montrer que $\frac{e^{-e^{1-e^{-1}}}}{e^{1-e^{-1}}} < e^{-1}$, ou encore en passant au logarithme :

$$e^{-1} - 1 - e^{1-e^{-1}} < -1 \iff e^{-1} < e^{1-e^{-1}}.$$

En passant à nouveau au logarithme, on obtient :

$$-2 < -e^{-1} \text{ égalité bien entendu vérifiée.}$$

D'où en conclusion : $u_2 > u_0$ et $u_3 < u_1$.

- 4.2) La fonction f est décroissante sur \mathbf{R}^{+*} , donc $f \circ f$ est quant à elle croissante sur \mathbf{R}^{+*} . Soit $n \in \mathbf{N}$, appliquons n fois la fonction $f \circ f$ aux inégalités établies dans la question précédente. On obtient :

$$u_{2n+2} > u_{2n}, \quad u_{2n+3} < u_{2n+1}.$$

C'est exactement dire que

(u_{2n}) (resp. (u_{2n})) est strictement croissante (resp. strictement décroissante).

5. On pose : $h(x) = \begin{cases} (f \circ f)(x) & \text{si } x > 0, \\ 0 & \text{si } x = 0. \end{cases}$. On va établir d'abord qu'il s'agit du prolongement par continuité en zéro de la fonction f^2 .

- 5.1) Soit x un réel strictement positif. Par définition, pour $x > 0$,

$$h(x) = \frac{e^{-\frac{e^{-x}}{x}}}{\frac{e^{-x}}{x}} = xe^{x-\frac{e^{-x}}{x}}.$$

Par opérations sur les limites, on a : $x - \frac{e^{-x}}{x} \xrightarrow{x \rightarrow 0} -\infty$ donc par produit $h(x) \xrightarrow{x \rightarrow 0} 0$. Ainsi, la fonction h est continue en zéro.

- 5.2) L'équation $h(x) = x$ admet déjà zéro comme solution, résolvons-là à présent sur \mathbf{R}_{+*} . On a

$$h(x) = x$$

$$\begin{aligned} \iff xe^{x-\frac{e^{-x}}{x}} &= x \\ \iff x - \frac{e^{-x}}{x} &= 0 \quad \left. \begin{array}{l} \\ \end{array} \right\} \text{simplification par } x \neq 0 \text{ puis passage au logarithme} \\ \iff x^2 = e^{-x} \\ \iff x = f(x). \quad \left. \begin{array}{l} \\ \end{array} \right\} x \neq 0 \end{aligned}$$

Les solutions de l'équation sont donc les éléments de $\{0, \alpha\}$.

- 5.3) Constatons d'abord que $(u_{2n+1})_{n \in \mathbf{N}}$ vérifie une relation de récurrence à un pas. En effet, soit $n \in \mathbf{N}$, alors

$$u_{2(n+1)+1} = u_{2n+3} = f(u_{2n+2}) = f \circ f(u_{2n+1}) = h(u_{2n+1}) \quad (*)$$

car les termes de la suite sont strictement positifs d'après une question précédente. La suite $(u_{2n+1})_{n \in \mathbf{N}}$ est décroissante (déjà vu), et est minorée par zéro, donc converge vers une limite finie que nous notons $\ell \in \mathbf{R}$. Passons à la limite dans $(*)$:

$$\ell = \lim_{n \rightarrow \infty} h(u_{2n+1}) = h\left(\lim_{n \rightarrow \infty} u_{2n+1}\right) = h(\ell),$$

car h est continue sur \mathbf{R}^+ (étape!). Donc ℓ est une solution de l'équation $h(x) = x$, donc $\ell = 0$ ou α . La seconde limite est impossible, puisque $(u_{2n+1})_{n \in \mathbf{N}}$ est décroissante et que $u_{2n+1} \leq u_1 = \frac{1}{e} < \alpha$ d'après une question précédente. Donc si $(u_{2n+1})_{n \in \mathbf{N}}$ convergeait vers α on aurait par passage à la limite $\alpha < \alpha$ — contradiction.

Donc, en conclusion : $u_{2n+1} \xrightarrow{n \rightarrow \infty} 0$.

- 5.4) La suite des termes pairs est quant à elle croissante, et donc pour tout $n \in \mathbf{N}$, $u_{2n} \geq u_0 = 1$. On montre de-même qu'avant que, comme la fonction h est continue, la suite (u_{2n}) ne peut converger que vers une solution de l'équation $x = h(x)$. Comme $\alpha < 1$, et que $u_{2n} \geq 1 > 0$ pour tout $n \in \mathbf{N}$, elle ne peut converger ni vers α , ni vers zéro. Donc elle diverge et c'est forcément vers $+\infty$ car elle est positive. Donc $u_{2n} \xrightarrow{n \rightarrow \infty} \infty$.

Exercice ANA.9.7 | Étude d'une suite récurrente d'ordre 1. Approximation d'une racine par une suite On considère un réel $a > 0$ et une suite (u_n) définie par son premier terme u_0 tel que $u_0 > \sqrt{a}$, et

$$\forall n \in \mathbf{N}, \quad u_{n+1} = \frac{1}{2} \left(u_n + \frac{a}{u_n} \right).$$

- Étudier brièvement sur \mathbf{R}^{++} la fonction $f : x \mapsto \frac{1}{2} \left(x + \frac{a}{x} \right)$ et le signe de $f(x) - x$.
- _🔗 Écrire une fonction en Python qu'on appellera $U(u_0, a, n)$ qui prend en argument le premier terme de la suite, le réel a et l'entier $n \geq 1$, et renvoie la liste des termes $[u_0, \dots, u_n]$.
- En utilisant la bibliothèque `matplotlib.pyplot` tracer le graphique qui donne les 100 premiers termes de (u_n) pour $a = 2$ et $u_0 = 3,2$.
- Montrer que pour tout $n \in \mathbf{N} : \sqrt{a} < u_{n+1} < u_n$. En déduire que (u_n) converge et calculer sa limite.
- On pose $v_n = \frac{u_n - a}{u_n + \sqrt{a}}$.
 - Montrer que $v_{n+1} = v_n^2$ et en déduire $v_n = v_0^{2^n} \forall n$.
 - Montrer que pour tout $n, \quad 0 \leq u_n - \sqrt{a} \leq 2u_0 v_0^{2^n}$ et retrouver la limite de (u_n) .
- Dans cette question $a = 2$ et $u_0 = 2$. On veut déterminer le plus petit entier n tel que $2u_0 v_0^{2^n} < 10^{-100}$.
 - _🔗 Le déterminer en écrivant une fonction python `NMin`.
 - Que peut-on dire de u_n avec le n trouvé ci-dessus?

Solution (exercice ANA.9.7)

- La fonction f est dérivable sur \mathbf{R}^{++} en tant que somme de fonctions dérivables. De plus, $f'(x) = \frac{1}{2} \left(1 - \frac{a}{x^2} \right)$ pour tout $x \in \mathbf{R}^{++}$.

La fonction f est donc décroissante sur $[0, \sqrt{a}]$ et croissante sur $[\sqrt{a}, \infty[$.

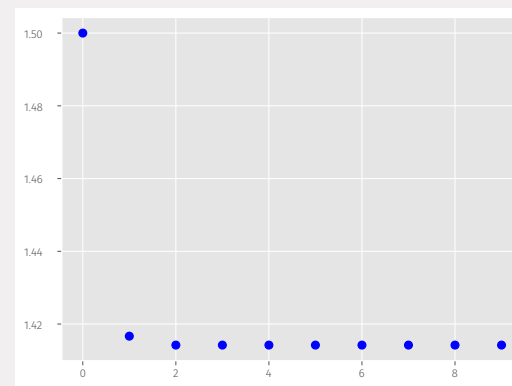
De plus, si l'on pose $g(x) = f(x) - x = \frac{1}{2} \left(\frac{a}{x} - x \right)$, on constate immédiatement que :

g est positive si $x \leq \sqrt{a}$ et négative si $x \geq \sqrt{a}$.

```
2. def U(u0, a, n):
    """
    Renvoie la liste des termes jusqu'au rang n
    """
    L = [u0]
    u = u0
    for _ in range(2, n+1):
        u = (1/2)*(u+a/u)
        L.append(u)
    return L
```

- On peut donc ensuite tracer ces termes.

```
import matplotlib.pyplot as plt
plt.plot(U(3/2,2,10), 'bo')
```



- Constatons que $f(\sqrt{a}) = a$, c'est un point fixe de la fonction. Alors rappelons que $u_0 > \sqrt{a}$. Donc en appliquant f qui est strictement croissante sur ce domaine, on trouve $u_1 > \sqrt{a}$. Mais comme g est strictement négative sur $]\sqrt{a}, \infty[$, on a alors pour tout $x \in]\sqrt{a}, \infty[$, $f(x) < x$ donc en faisant $x = u_0$ on trouve $u_1 < u_0$. En combinant les deux encadrements on trouve $\sqrt{a} < u_1 < u_0$, i.e. la propriété au rang $n = 0$. Supposons-là vraie au rang n , alors appliquons f dans l'encadrement, on a alors :

$\sqrt{a} < u_{n+2} < u_{n+1}$ puisque f est strictement croissante sur $[\sqrt{a}, \infty[$. La propriété est donc vraie pour tout n par récurrence.

Ainsi la suite est décroissante minorée donc (u_n) converge.

5. 5.1) Soit $n \in \mathbf{N}$. Alors :

$$v_{n+1} = \frac{\frac{1}{2}\left(u_n + \frac{a}{u_n}\right) - \sqrt{a}}{\frac{1}{2}\left(u_n + \frac{a}{u_n}\right) + \sqrt{a}} = \frac{(u_n^2 + a) - 2u_n\sqrt{a}}{(u_n^2 + a) + 2u_n\sqrt{a}} = v_n^2.$$

On vérifie ensuite la formule $v_n = v_0^{2^n}$ par récurrence sur $n \in \mathbf{N}$.

5.2) On a déjà établi que pour tout entier $n \in \mathbf{N}$, $u_n \geq \sqrt{a}$. Et pour l'autre partie de l'encadrement, on écrit que $u_n - \sqrt{a} = v_n(u_n + \sqrt{a}) \leq v_n 2u_0$. Puis on utilise l'expression explicite de (v_n) trouvée précédemment. Puisque $u_0 > \sqrt{a}$, on vérifie par exemple par étude de fonction que $v_0 \in]-1, 1[$, donc que : $\lim_{n \rightarrow \infty} v_0^{2^n} = 0$.

Donc : $\lim_{n \rightarrow \infty} u_n = 0$.

```
6. 6.1) import math as ma
def NMin(u0, a):
    """
    Cherche le plus petit entier n tel que le majorant de
    u-sqrt(a) soit assez petit
    """
    n = 0
    v = (u0-ma.sqrt(a))/(u0+ma.sqrt(a))
    while 2*(u0)*v >= 10**(-100):
        n += 1
        v = v**2
    return n
```

Par exemple pour les paramètres données dans l'énoncé, on trouve : 8.

6.2) Pour ce n on est sûr que u_n est une approximation de \sqrt{a} avec précision 10^{-100} .

Exercice ANA.9.8 | Agro—Véto, 2018. Étude d'une suite implicite Pour $n \geq 1$, on définit la suite de fonctions (f_n) par :

$$\forall x \in \mathbf{R} \quad f_n(x) = nx^3 + n^2x - 2.$$

1. 1.1) Montrer pour n fixé que l'équation $f_n(x) = 0$ admet une unique solution sur \mathbf{R} notée a_n . Montrer que $a_n > 0$.
- 1.2) ✎ Écrire une fonction `cho(n)` qui calcule une valeur approchée de a_n pour un n donné, à précision 10^{-2} , en utilisant le principe de dichotomie. La tester pour $n = 2$.
- 1.3) Montrer que la suite $(a_n)_{n \geq 1}$ est convergente et déterminer sa limite.
2. On pose $u_n = n^2 a_n$ pour tout $n \geq 1$.
 - 2.1) ✎ Écrire un programme qui trace les termes u_n pour $n \in [10, 40]$. En déduire une conjecture sur la limite de la suite (u_n) .
 - 2.2) Démontrer cette conjecture. En déduire un équivalent de la suite (a_n) .
3. Soit g définie sur $[0, 1]$ par $g(x) = \frac{2x^3+1}{3x^2+2}$ pour tout $x \in [0, 1]$.
 - 3.1) Montrer que g est croissante sur $[a_2, 1]$.
 - 3.2) On définit la suite $(x_n)_{n \in \mathbf{N}}$ par $x_0 = 1$ et $x_{n+1} = g(x_n)$ pour tout entier $n \in \mathbf{N}$. Montrer que la suite (x_n) converge et déterminer sa limite.

Solution (exercice ANA.9.8)

1. Pour n fixé f_n est dérivable sur \mathbf{R} (c'est un polynôme) et $f_n'(x) = 3nx^2 + n^2 > 0$ pour tout x , donc f_n strictement croissante sur \mathbf{R} . La limite de f_n en $-\infty$ est $-\infty$, et en $+\infty$ c'est $+\infty$. De plus f_n est continue sur \mathbf{R} , donc d'après le théorème de la bijection, f_n s'annule une unique fois : Il existe un unique réel a_n tel que $f_n(a_n) = 0$. De plus $f_n(0) = -2 < 0 \Rightarrow 0 < a_n$ car f_n strictement croissante et $f_n(a_n) = 0$.
 - 1.1) Pour démarrer le principe de dichotomie, on a besoin aussi d'un majorant de a_n , or $f_n(1) = n + n^2 - 2 \geq 1 + 1 - 2 = 0$ car $n \geq 1$, donc $1 \geq a_n$.

```
def f(n, x):
    return n*x**3 + n*n*x - 2
```

```
def dichotomie(n):
    a = 0
    b = 1
    while b - a > 10**(-3):
        c = (a+b)/2
        if f(n, a)*f(n, c) <= 0:
            b = c
        else:
            a = c
    return (a+b)/2
```

Par exemple :

```
>>> dichotomie(10)
0.02001953125
```

1.2) Pour n fixé, $f_{n+1}(a_n) = (n+1)a_n^3 + (n+1)^2 a_n - 2 = \underbrace{na_n^3 + n^2 a_n - 2}_{=f_n(a_n)=0} + \underbrace{a_n^3 + (2n+1)a_n}_{>0 \text{ car } a_n > 0} > 0 = f_{n+1}(a_{n+1})$.

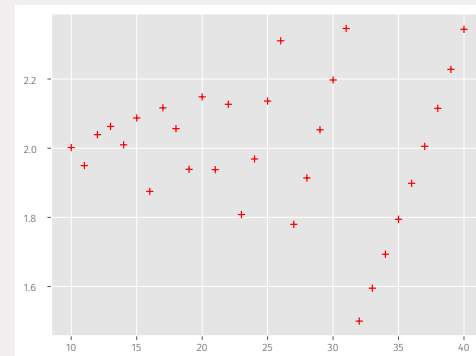
Donc $a_n > a_{n+1}$ puisque la fonction f_{n+1} est strictement croissante sur \mathbf{R} , et ceci pour tout $n \geq 1$. On en déduit que la suite $(a_n)_{n \geq 1}$ est décroissante, et minorée par 0, donc convergente.

De plus $f_n(a_n) = 0 \iff na_n^3 + n^2 a_n - 2 = 0 \iff a_n = \frac{2}{n} \left(\frac{1}{a_n^2 + n} \right)$.
Or quand $n \rightarrow +\infty$, $a_n \rightarrow \ell \in \mathbf{R}$ donc $a_n^2 + n \rightarrow +\infty \Rightarrow \frac{1}{a_n^2 + n} \rightarrow 0$ et $a_n \rightarrow 0$ donc $\ell = 0$.

2. 2.1) Si l'on garde la précision de 10^{-3} utilisé au départ dans le programme de dichotomie, on obtient une représentation graphique incohérente des u_n , on va donc demander une précision eps supérieure.

```
import matplotlib.pyplot as plt

X = range(10, 41)
Y = [n*dichotomie(n) for n in X]
plt.plot(X, Y, '+r')
```



Les valeurs de la suite sont proches de deux, mais certes très anarchiques. La méthode de dichotomie fonctionne donc assez mal sur cet exemple.

On « conjecture » que la suite $(u_n)_{n \geq 1}$ converge vers 2.

2.2) $\forall n \geq 1 \quad u_n = n^2 a_n$ donc $f_n(a_n) = 0 \iff \frac{u_n}{n} a_n^2 + u_n - 2 = 0 \iff u_n \left(1 + \frac{a_n^2}{n} \right) = 2$. Comme $\frac{a_n^2}{n} \xrightarrow{n \rightarrow \infty} 0$, nous obtenons : $u_n \xrightarrow{n \rightarrow \infty} 2$. Ainsi, nous obtenons l'équivalent : $u_n \underset{n \rightarrow \infty}{\sim} \frac{2}{n}$.

3. 3.1) Soit $g(x) = \frac{2x^3+1}{3x^2+2}$. On peut remarquer qu'elle est définie et dérivable sur \mathbf{R} . De plus, pour tout $x \in \mathbf{R}$:

$$g'(x) = \frac{6x^2(3x^2+2) - 6x(2x^3+1)}{(3x^2+2)^2} = \frac{6x}{(3x^2+2)^2} \underbrace{(x^3+2x-1)}_{h(x)}$$

La fonction h est donc dérivable sur \mathbf{R} et $h'(x) = 3x^2 + 2 > 0$ donc h est strictement croissante sur \mathbf{R} , donc sur $[0, 1]$. Or $a_2 \in [0, 1]$ et $h(a_2) = a_2^3 + 2a_2 - 1 = \frac{1}{2} (2a_2^3 + 4a_2 - 2)$, donc g' du signe de h sur $[0, 1]$ et est donc négative avant $=0$ par définition de a_2

a_2 et g' est positive sur $[a_2, 1]$, donc g est croissante sur $[a_2, 1]$.

3.2) On remarque de plus que $g(1) = \frac{3}{5} < 1$, $g(a_2) - a_2 = \frac{2a_2^3+1}{3a_2^2+2} - a_2 = \frac{-a_2^3-2a_2+1}{3a_2^2+2} = 0$ donc $g(a_2) = a_2$. Donc $g([a_2, 1]) \subset [a_2, 1]$ et g est croissante et continue sur cet intervalle. Nous obtenons successivement :

— $x_n \in [a_2, 1]$ pour tout entier $n \in \mathbf{N}$ par stabilité de $[a_2, 1]$ par g , — $x_1 =$

$g(x_0) = \frac{3}{5} < x_0$, donc par récurrence (en utilisant la croissance de g), on peut montrer que :

$\forall n, x_{n+1} < x_n$, c'est-à-dire que (x_n) est décroissante. De plus cette suite est bornée, donc elle converge, vers x_ℓ vérifiant $g(x_\ell) = x_\ell$ par continuité de g . Or l'unique solution dans $[0, 1]$ de cette équation est a_2 . En conclusion : (x_n) converge vers a_2

4.2. Équations différentielles

Exercice ANA.9.9 | Équation «quasi-»différentielle Déterminer les fonctions $f \in \mathcal{D}^1(\mathbf{R}, \mathbf{R})$ vérifiant :

$$\forall x \in \mathbf{R}, f'(x) = 2f(-x) + x.$$

Solution (exercice ANA.9.9)

► Pour éliminer la variable $-x$ qui nous gêne, on peut essayer de dériver comme somme et composée de fonctions dérivables, donc f vérifie l'équation différentielle

$$\forall x \in \mathbf{R}, f''(x) = -2f'(-x) + 1 = -2(2f(x) - x) + 1 \iff f''(x) + 4f(x) = 2x + 1.$$

Les solutions de l'homogène sont les fonctions $x \mapsto A \sin(2x) + B \cos(2x)$ avec $A, B \in \mathbf{R}$. En cherchant une solution particulière sous forme polynomiale (ou par variation de la constante), on trouve l'ensemble des solutions qui est :

$$\left\{ x \mapsto A \sin(2x) + B \cos(2x) + \frac{1}{2}x + \frac{1}{4}, \quad A, B \in \mathbf{R} \right\}.$$

► Inversement, prenons f de la forme $f(x) = A \sin(2x) + B \cos(2x) + \frac{1}{2}x + \frac{1}{4}, x \in \mathbf{R}$ et regarder si elle est solution de l'équation de départ. Alors f est solution de l'équation de départ si et seulement si, pour tout $x \in \mathbf{R}$,

$$2A \cos(2x) - 2B \sin(2x) + \frac{1}{2} = -2A \sin(2x) + 2B \cos(2x) - x + \frac{1}{2} + x.$$

Ce qui permet d'écrire, en rassemblant les termes, que :

$$\forall x \in \mathbf{R}, 2(A - B)(\cos(2x) + \sin(2x)) = 0 \implies A = B.$$

Il faut donc restreindre l'ensemble initial avec des constantes A et B égales. L'ensemble solution est

$$\left\{ x \mapsto A(\sin(2x) + \cos(2x)) + \frac{1}{2}x + \frac{1}{4}, \quad A \in \mathbf{R} \right\}.$$

Exercice ANA.9.10 |

1. Diagonaliser la matrice $A = \begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix}$.
2. Résoudre, par deux méthodes, le système différentiel suivant sur \mathbf{R} :

$$\begin{cases} y' - y = z, \\ z' + z = 3y. \end{cases}$$

Solution (exercice ANA.9.10)

1. Soit $\lambda \in \mathbf{C}$, alors

$$A - \lambda I_2 = \begin{pmatrix} 1 - \lambda & 1 \\ 3 & -1 - \lambda \end{pmatrix} \begin{array}{l} \left. \begin{array}{l} \phantom{\tilde{L}} \\ \phantom{\tilde{L}} \end{array} \right\} L_1 \leftrightarrow L_3 \\ \left. \begin{array}{l} \tilde{L} \begin{pmatrix} 3 & -1 - \lambda \\ 1 - \lambda & 1 \end{pmatrix} \\ \tilde{L} \begin{pmatrix} 3 & -1 - \lambda \\ 0 & 2 - \lambda^2 \end{pmatrix} \end{array} \right\} L_2 \leftrightarrow 3L_2 - (1 - \lambda)L_1 \end{array}$$

Donc les valeurs propres sont ± 2 . Pour les espaces propres, on résout le système associé.

- ▶ $X = \begin{pmatrix} x \\ y \end{pmatrix} \in E_2(A)$ si et seulement si $3x - 3y = 0$, d'où $E_2(A) = \text{Vect} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$,
- ▶ $X = \begin{pmatrix} x \\ y \end{pmatrix} \in E_{-2}(A)$ si et seulement si $3x + y = 0$, d'où $E_{-2}(A) = \text{Vect} \begin{pmatrix} -\frac{1}{3} \\ 1 \end{pmatrix}$.

Donc

$$A = PDP^{-1}, \quad D = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & -\frac{1}{3} \\ 1 & 1 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ -\frac{3}{4} & \frac{3}{4} \end{pmatrix}.$$

2. ▶ **(Première méthode : matricielle)** Notons $X = \begin{pmatrix} y \\ z \end{pmatrix}$. Alors le système se réécrit

$$X' = \begin{pmatrix} y + z \\ 3y - z \end{pmatrix} = AX.$$

Donc introduisons $Y = P^{-1}X$. Alors $Y' = P^{-1}X' = P^{-1}PDP^{-1}X = DY$.

Notons à présent $Y = \begin{pmatrix} F \\ G \end{pmatrix}$, avec F, G deux fonctions dérivables. Alors, d'après le calcul précédent, elles sont solutions de

$$F' = 2F, \quad G' = -2G,$$

donc :

$$\exists K, L \in \mathbf{R}, \quad \forall x \in \mathbf{R}, \quad F(x) = Ke^{2x}, \quad G(x) = Le^{-2x}.$$

On revient ensuite aux variables y, z , pour tout $x \in \mathbf{R}$

$$X(x) = PY(x) = \begin{pmatrix} 1 & -\frac{1}{3} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} Ke^{2x} \\ Le^{-2x} \end{pmatrix} = \begin{pmatrix} Ke^{2x} - \frac{1}{3}Le^{-2x} \\ Ke^{2x} + Le^{-2x} \end{pmatrix}.$$

- ▶ **(Deuxième méthode : artisanale)** Les deux fonctions y, z sont *a priori* dérivables une fois. Mais comme $y' = z + y$, y est deux fois dérivable et on obtient

en redérivant : $y'' - 4y = 0$ par substitution. Donc y est nécessairement de la forme $y(x) = Ke^{2x} + Le^{-2x}$ pour tout $x \in \mathbf{R}$.

En injectant dans la première équation, on trouve $z(x) = Ke^{2x} - 3Le^{-2x}$ pour tout $x \in \mathbf{R}$.

Inversement, on vérifie que les deux fonctions précédentes sont effectivement solutions. Et en renommant L en $\frac{1}{3}L$, on retrouve bien les solutions obtenues par la première méthode.

Exercice ANA.9.11 | Changement de fonction inconnue Résoudre

$$x^2 y'' + 3xy' + y = (x + 1)^2$$

sur \mathbf{R}^{+*} . *Indication :* On pourra réaliser le changement de fonction inconnue $z(t) = y(e^t)$ pour tout $t \in \mathbf{R}$.

Solution (exercice ANA.9.11)

On commence par calculer les dérivées de z . Pour tout $t \in \mathbf{R}$,

$$\begin{aligned} z'(t) &= e^t y'(e^t), \\ &= e^t y'(e^t) + (e^t)^2 y''(e^t) \\ &= z'(t) + (e^t)^2 y''(e^t). \end{aligned}$$

Or, pour tout $t \in \mathbf{R}$, on a en utilisant l'équation différentielle de départ :

$$(e^t)^2 y''(e^t) = -3e^t y'(e^t) - y(e^t) - (e^t + 1)^2,$$

donc

$$(e^t)^2 y''(e^t) = -3z'(t) - z(t) - (e^t + 1)^2.$$

Donc z vérifie :

$$\forall t \in \mathbf{R}, \quad z''(t) = z'(t) - 3z'(t) - z(t) - (e^t + 1)^2,$$

donc

$$\forall t \in \mathbf{R}, \quad z''(t) + 2z'(t) + z(t) = (e^t + 1)^2.$$

C'est cette fois-ci une équation différentielle linéaire à coefficients constants, l'équation caractéristique est $x^2 + 2x + 1 = 0 = (x + 1)^2$. Donc

$$\exists A, B \in \mathbf{R}, \forall t \in \mathbf{R}, \quad z(t) = (At + B)e^{-t}.$$

Constatons que

$$\forall t \in \mathbf{R}, \quad z(t) = y(e^t) \iff \forall x > 0, \quad y(x) = z(\ln x).$$

Donc

$$\boxed{\forall x > 0, \quad y(x) = \frac{A \ln x + B}{x}}.$$

Exercice ANA.9.12 | On note E l'ensemble des applications de \mathbf{R} dans \mathbf{R} de classe \mathcal{C}^2 vérifiant

$$\forall x \in \mathbf{R}, \quad x^2 f''(x) - 4x f'(x) + 6f(x) = 0.$$

1. Vérifier que E est un espace vectoriel.
2. Déterminer l'ensemble des fonctions polynomiales sur \mathbf{R} de E.
3. Soit $f \in E$. On définit alors : $g \begin{cases} \mathbf{R} \setminus \{0\} & \longrightarrow \mathbf{R} \\ x & \longrightarrow \frac{f(x)}{x^2} \end{cases}$. Montrer que g est de classe \mathcal{C}^2 sur \mathbf{R}^* et que g'' est nulle. En déduire la dimension et une base de E.
4. Retrouver les solutions sur \mathbf{R}^{++} en faisant le changement de variable $t = \ln x$.

Solution (exercice ANA.9.12)

1. Justifions que E est un sous-espace vectoriel de $\mathcal{C}^2(\mathbf{R}, \mathbf{R})$. Puisque l'équation différentielle est linéaire homogène, la fonction nulle est solution. Soient $f, g \in E$ et $\lambda, \mu \in \mathbf{R}$. Alors, pour tout $x \in \mathbf{R}$,

$$\begin{aligned} & x^2(\lambda f + \mu g)''(x) - 4x(\lambda f + \mu g)'(x) + 6(\lambda f + \mu g)(x) \\ &= x^2(\lambda f''(x) + \mu g''(x)) - 4x(\lambda f'(x) + \mu g'(x)) + 6(\lambda f(x) + \mu g(x)) \\ &= \lambda 0 + \mu 0 \\ &= 0. \end{aligned} \quad \left. \begin{array}{l} \text{linéarité de la dé-} \\ \text{rivation} \\ f, g \in E \end{array} \right\}$$

Donc E est un sous-espace vectoriel de $\mathcal{C}^2(\mathbf{R}, \mathbf{R})$, donc **E est un espace vectoriel.**

2. Soit $f = a_n X^n + a_{n-1} X^{n-1} + \dots + a_0$ avec $a_n \neq 0$ et $a_1, \dots, a_n \in \mathbf{R}$. Alors commençons par chercher s'il y a une condition sur le degré n , en analysant le coefficient dominant. Le terme d'ordre n de $X^2 f''(X) - 4X f'(X) + 6f(X)$ est :

$$a_n n(n-1)X^n - 4n a_n X^n + 6a_n X^n = a_n(n^2 - 5n + 6)X^n.$$

Ainsi, si $f \in E$ alors tous les coefficients de $X^2 f''(x) - 4X f'(x) + 6f(x)$ sont nuls, en particulier $n^2 - 5n + 6 = (n-3)(n-2) = 0$ puisque $a_n \neq 0$. Donc : $\deg f \in \{3, 2\}$. Supposons que $f = a_3 X^3 + \dots + a_0$. Alors f est solution si et seulement si

$$\begin{aligned} X^2(6a_3 X + 2a_2) - 4X(3a_3 X^2 + 2a_2 X + a_1) + 6(a_3 X^3 + \dots + a_0) &= 0, \\ 0X^3 + 0X^2 + (-4a_1 + 6a_1)X - 4a_0 &= 0. \end{aligned}$$

En identifiant, on obtient alors comme conditions $a_1 = 0, a_0 = 0$.

Donc f est solution si et seulement si $f \in \text{Vect}(X^2, X^3)$.

En résumé, **les solutions polynomiales sont les fonctions de $\text{Vect}(X^2, X^3)$.**

3. La fonction **g est de classe \mathcal{C}^2 sur \mathbf{R}^*** en tant que quotient de telles fonctions, dont le dénominateur ne s'annule pas. De plus, pour $x \in \mathbf{R}^*$,

$$g'(x) = \frac{f'(x)x^2 - 2xf(x)}{x^4} = \frac{f'(x)}{x^2} - 2\frac{f(x)}{x^3}.$$

Puis on déduit la dérivée seconde :

$$g''(x) = \frac{f''(x)x^2 - 2xf'(x)}{x^4} - 2\frac{f'(x)x^3 - 3x^2 f(x)}{x^6}.$$

Ainsi, en mettant les fractions au même dénominateur, on obtient

$$g''(x) = \frac{x^2 f''(x) - 2xf'(x) - 2xf'(x) + 6f'(x)}{x^4} = \frac{0}{x^4}.$$

Donc : $g'' = 0$ sur \mathbf{R}^* .

On déduit alors qu'il existe $A, B \in \mathbf{R}$ tels que $f(x) = Ax^3 + Bx^2$ pour tout $x \in \mathbf{R}^*$. Par continuité de f sur \mathbf{R} , on a même :

$$\forall x \in \mathbf{R}, \quad f(x) = Ax^3 + Bx^2.$$

Ainsi, $(x \mapsto x^2, x \mapsto x^3)$ est une famille génératrice de E , elle est de plus libre en tant que sous-famille d'une famille libre (la base canonique). Donc

$(x \mapsto x^2, x \mapsto x^3)$ est une base de E et $\dim E = 2$.

4. Définissons la fonction $h : x \mapsto f(e^x)$. Alors en tant que composée de telles fonctions, la fonction h est deux fois dérivable, et pour tout $x \in \mathbf{R}^{+\ast}$,

$$h'(x) = e^x f'(e^x)$$

$$h''(x) = e^x f'(e^x) + (e^x)^2 f''(e^x).$$

Mais d'après l'équation différentielle, nous avons

$$(e^x)^2 f''(e^x) - (4e^x) f'(e^x) + 6f(e^x) = 0.$$

Donc en utilisant h , on trouve :

$$(h''(x) - h'(x)) - 4h'(x) + 6h(x) = 0 \iff h''(x) - 5h'(x) + 6h(x) = 0.$$

L'équation caractéristique est alors $x^2 - 5x + 6 = 0 = (x - 2)(x - 3)$. Donc il existe $A, B \in \mathbf{R}$ telles que pour tout $x \in \mathbf{R}$, $f(e^x) = h(x) = Ae^{2x} + Be^{3x}$. De manière équivalente,

$$\forall x \in \mathbf{R}^{+\ast}, \quad f(x) = Ax^2 + Bx^3.$$

On retombe bien sur les solutions trouvées par l'autre méthode.

Exercice ANA.9.13 | WRONSKIEN pour l'ordre 2 Soient f et g deux solutions de l'équation différentielle ci-dessous sur \mathbf{R}^*

$$\forall x \in \mathbf{R}^*, \quad y'' + \left(1 + \frac{2}{x}\right)y' + \left(1 + \frac{1}{x}\right)y = 0.$$

On admet l'existence de ces deux solutions. On considère par ailleurs, pour tout $x \in \mathbf{R}^*$ la matrice

$$M(x) = \begin{pmatrix} f(x) & g(x) \\ f'(x) & g'(x) \end{pmatrix}.$$

On dit que (f, g) est un *système fondamental de solutions* si et seulement si

$$\forall x \in \mathbf{R}^{+\ast}, \quad W(x) \neq 0.$$

1. Montrer que $W = \det M$ — c'est une fonction de \mathbf{R} dans \mathbf{R} ! — vérifie une équation différentielle simple du premier ordre.
2. En déduire que (f, g) est un système fondamental de solution si et seulement si $W(1) \neq 0$.

Solution (exercice ANA.9.13)

1. Par définition du déterminant, on a $W(x) = fg' - f'g$. En utilisant la formule de

dérivation d'un produit, on a pour tout $x \in \mathbf{R}^*$,

$$\begin{aligned}
 W'(x) &= (fg' - f'g)'(x) \\
 &= f'(x)g'(x) + f(x)g''(x) - f''(x)g(x) - f'(x)g'(x) \\
 &= f'(x)g'(x) + f(x)\left(-\left(1 + \frac{2}{x}\right)g'(x) - \left(1 + \frac{1}{x}\right)g(x)\right) \\
 &\quad - \left(-\left(1 + \frac{2}{x}\right)f'(x) - \left(1 + \frac{1}{x}\right)f(x)\right)g(x) - f'(x)g'(x) \\
 &= \cancel{f'(x)g'(x)} - f(x)g'(x) - \cancel{f(x)g'(x)} + f'(x)g(x) + \cancel{f(x)g'(x)} - \cancel{f'(x)g'(x)} \quad \left. \begin{array}{l} \text{dévelop-} \\ \text{pement} \end{array} \right\} \\
 &\quad - \frac{2}{x}f(x)g'(x) - \frac{1}{x}f(x)g(x) + \frac{2}{x}f'(x)g(x) + \frac{1}{x}f(x)g(x) \\
 &= -W(x) - \frac{2}{x}W(x) = -\left(1 + \frac{2}{x}\right)W(x).
 \end{aligned}$$

Donc : W est solution de $y' = -\left(1 + \frac{2}{x}\right)y$.

2. On déduit alors que pour tout $x \in \mathbf{R}^*$,

$$W(x) = Ke^{-x-2\ln|x|} = \frac{Ke^{-x}}{x^2}.$$

En faisant $x = 1$, on obtient $W(1) = Ke^{-1}$, donc $K = W(1)e$ et

$$W(x) = W(1)\frac{e^{1-x}}{x^2}.$$

Donc, puisque'une exponentielle n'est jamais nulle, (f, g) est un système fondamental de solution si et seulement si $W(1) \neq 0$.

Exercice ANA.9.14 | Système différentiel non linéaire On considère le système différentiel (\mathcal{S}) suivant :

$$(\mathcal{S}) \quad \begin{cases} x' &= y^2, \\ y' &= \sin x, \end{cases}$$

où les inconnues sont deux fonctions dérivables x et y . On admet que \mathcal{S} est non vide.

1. **(Points d'équilibres)** Déterminer les solutions constantes de (\mathcal{S}) . Ces solutions sont appelées *points d'équilibre du système*.

2. Soit V une fonction de classe \mathcal{C}^1 sur \mathbf{R}^2 .

2.1) Donner une condition sur $\frac{\partial V}{\partial x}$ et $\frac{\partial V}{\partial y}$ pour que la fonction

$$\phi : t \longrightarrow V(x(t), y(t))$$

soit constante pour toute solution (x, y) de (\mathcal{S}) .

2.2) Vérifier que la fonction

$$V : (x, y) \in \mathbf{R}^2 \longrightarrow \cos x + y^3/3$$

satisfait la condition précédente.

3. On admet qu'il existe une unique solution (x_0, y_0) de (\mathcal{S}) vérifiant $x_0(0) = 0$ et $y_0(0) = -\sqrt[3]{6}$.

3.1) Quelle relation existe-t-il entre x_0 et y_0 ?

3.2) Exprimer y_0 en fonction de x_0 puis tracer le graphe correspondant.

4.3. Modélisation discrète et continue

Exercice ANA.9.15 | Toutes les heures, on injecte à un sujet, une même dose de 1,8 unités, d'une substance médicamenteuse dans le sang. Les injections sont faites par piqûre intraveineuse. On suppose que la substance se répartit instantanément dans le sang et qu'elle est ensuite progressivement éliminée. En l'espace d'une heure, la quantité de cette substance présente dans le sang diminue de 30%. La première injection se fait à $t = 0$.

Pour $n \in \mathbf{N}^*$, on note Q_n la quantité de substance présente dans le sang à l'instant $t = n$ (en heures), dès que la nouvelle injection est faite.

1. Calculer pour $n \in \mathbf{N}^*$ le terme Q_n . Donner une approximation au dixième près de la quantité de substance présente dans le sang à l'instant à $t = 5$.

2. Déterminer la limite de (Q_n) .

3. ➤ Proposer un programme prenant en argument $\varepsilon > 0$ et permettant de connaître le plus petit entier n tel que $|Q_n - 6| < \varepsilon$.

Solution (exercice ANA.9.15)

1. Nous avons donc la récurrence suivante après lecture de l'énoncé : pour tout entier $n \in \mathbf{N}$, $Q_{n+1} = 1.8 + 0.7Q_n$. Donc on constate que l'unique point fixe associé est 6, donc $(Q_n - 6)$ est géométrique de raison 0.7. On obtient alors pour tout entier $n \in \mathbf{N}$:

$$Q_n - 6 = (0.7)^n(Q_0 - 6) \iff Q_n = 6 + (0.7)^n(1.8 - 6).$$

2. $Q_n \xrightarrow{n \rightarrow \infty} 6$ puisque $|0.7| < 1$.
 3. On déduit le script Python ci-dessous :

```
def approx6(eps):
    n = 0
    Q = 1.8
    while abs(Q-6) >= eps:
        n += 1
        Q = 1.8 + 0.7*Q
    return n
```

Une exécution pour $\varepsilon = 10^{-2}$ renvoie 17.

Exercice ANA.9.16 | Cinétique chimique d'ordre 2 On considère une réaction chimique notée $A \rightarrow B$, on suppose que le réactif A disparaît avec un taux proportionnel (on note $k \in \mathbf{R}$ le coefficient associé) à la concentration de réactif. Déterminer la concentration $[A]$ en réactif en fonction du temps.

Solution (exercice ANA.9.16)

Le taux d'évolution est négatif, puisqu'il y a disparition comme l'indique l'énoncé, et

donc

$$\frac{[A]'(t)}{[A](t)} = -k \cdot [A](t).$$

On peut aussi reformuler sans utiliser la définition du taux, en faisant un bilan de concentration entre t et $t + h$, $t, h \geq 0$:

$$[A](t + h) = [A](t) - k([A](t) \cdot [A](t)),$$

donc en faisant $h \rightarrow 0$, on retrouve que $[A]$ est solution de

$$y' = -ky^2.$$

Supposons que la concentration ne s'annule pas, on obtient :

$$\begin{aligned} \forall t \geq 0, \int_0^t -\frac{[A]'(u)}{[A]^2(u)} du &= \int_0^t k du \\ \iff \left[\frac{1}{[A](u)} \right]_0^t &= k[u]_0^t \iff \frac{1}{[A](t)} - \frac{1}{[A](0)} \\ &= kt \iff [A](t) = \frac{1}{kt + \frac{1}{[A](0)}}. \end{aligned}$$

Exercice ANA.9.17 | Modèle simplifié de croissance tumorale & contrôlabilité du système à zéro Une première étape de la maladie du cancer consiste en une division frénétique et anarchique d'un certain nombre de cellules cancéreuses, qui forment un agrégat de cellules appelé tumeur. Durant cette phase, le nombre de cellules cancéreuses augmente entre deux temps avec un taux de reproduction égal à deux. Une modélisation simplifiée de ce processus en une équation différentielle vérifiée par le nombre y de cellules cancéreuses au temps t , la fonction y est solution de l'équation homogène

$$\begin{cases} y' = f(y), \\ y(0) = 1000 \end{cases} \quad f : \mathbf{R} \rightarrow \mathbf{R} \text{ une fonction.} \quad (\text{Cancer})$$

Dans toute la suite, le nombre de cellules y est une variable du temps t que l'on exprime en jour.

1. Proposer une fonction f en accord avec l'énoncé, puis résoudre l'équation différentielle.
2. Trois biologistes en compétition souhaitent éradiquer la maladie au bout de 10 jours *via* une chimiothérapie, qui a pour effet de tuer un certain nombre de cellules cancéreuses au cours du temps. La nouvelle équation

$$\begin{cases} y' = f(y) + c, \\ y(0) = 1000, \end{cases} \quad c \text{ une fonction } \mathbf{R} \longrightarrow \mathbf{R}. \quad (\text{Chimio})$$

où $c(t)$ désigne l'effet de la chimiothérapie à l'instant t pour tout $t \in \mathbf{R}$.

- 2.1) Donner une formule explicite pour y solution de (Chimio).
- 2.2) En déduire l'existence d'une fonction c répondant aux exigences des médecins.

Exercice ANA.9.18 | Hauteur d'une baignoire On considère une baignoire de forme parallélépipédique dont la base est de dimensions ab que l'on remplit avec un débit constant noté d . On note $z(t)$ la hauteur d'eau dans la baignoire à l'instant t et $V(t)$ son volume. On suppose que $V(0) = 0$. La baignoire a une fissure au fond qui laisse s'échapper plus ou moins d'eau en fonction de la pression exercée par l'eau sur celle-ci. On rappelle que la pression au fond de la baignoire est égale à $p(z) = \rho g z$. Le débit de la fuite est $d_f = \alpha p$ avec $\alpha > 0$ et p la pression qui s'y exerce. Toutes les grandeurs sont en unité SI.

1. Si on suppose la baignoire suffisamment haute, montrer que le volume d'eau tend vers un volume à l'équilibre V_{eq} que l'on déterminera.
2. Si la baignoire a un volume V (que l'on suppose inférieur au volume d'équilibre), au bout de combien de temps sera-t-elle pleine?

Exercice ANA.9.19 | Modèle de VERHULST continu On dit qu'une fonction P suit un modèle de Verhulst de taux β, μ, κ si P est dérivable en t et vérifie :

$$P'(t) = (\beta - \mu) \left(P(t) - \frac{P(t)^2}{\kappa} \right) \quad (t \geq 0), \quad P(0) = P_0. \quad (\text{Verhulst})$$

Dans la suite on note $r = \beta - \mu$. On cherche résoudre l'équation (Verhulst) pour $P(0) = P_0 \in]0, \kappa[$.

1. **(Positivité)** On cherche déjà à prouver que pour tout $t \geq 0$, $P(t) \in]0, \kappa[$. On suppose que P est une solution, et on note $f(t) = rP(t) \left(1 - \frac{P(t)}{\kappa} \right)$.

- 1.1) Montrer que f est dérivable sur \mathbf{R}^+ puis justifier que f est solution d'une équation différentielle linéaire homogène du premier ordre à coefficients continus.
- 1.2) Montrer que $f(0) \neq 0$ et en déduire que pour tout $t \geq 0$, $f(t) \neq 0$.
- 1.3) Montrer que pour tout $t \geq 0$, $P(t) \in]0, \kappa[$.

2. (Résolution)

- 2.1) Montrer qu'il existe deux constantes $a, b \in \mathbf{R}^2$ telles que : $\forall N \in \mathbf{R} \setminus \{0, \kappa\}, \frac{\kappa}{N(\kappa-N)} = \frac{a}{N} + \frac{b}{\kappa-N}$.
- 2.2) Résoudre (Verhulst).

Solution (exercice ANA.9.19)

1. **(Positivité)** On cherche déjà à prouver que pour tout $t \geq 0$, $P(t) \in]0, \kappa[$. On suppose que P est une solution, et on note $f(t) = rP(t) \left(1 - \frac{P(t)}{\kappa} \right)$.

- 1.1) la fonction f est dérivable sur \mathbf{R}^+ en tant que somme et produit de fonctions dérivables. De plus, pour tout $t \geq 0$,

$$\begin{aligned} f'(t) &= rP'(t) \left(1 - \frac{P(t)}{\kappa} \right) + rP(t) \left(-\frac{1}{\kappa} P'(t) \right) \\ &= r^2 P(t) \left(1 - \frac{P(t)}{\kappa} \right)^2 - \frac{r^2}{\kappa} P(t) \left(1 - \frac{P(t)}{\kappa} \right) \\ &= rP(t) \left(1 - \frac{P(t)}{\kappa} \right) \left[r \left(1 - \frac{P(t)}{\kappa} \right) - \frac{r}{\kappa} \right] \\ &= f(t)g(t), \end{aligned}$$

où $g(t) = r \left(1 - \frac{P(t)}{\kappa} \right) - \frac{r}{\kappa}$ pour tout $t \geq 0$. Constatons que g est une fonction continue, donc f est solution de $y' = g(t)y$, une équation différentielle linéaire homogène du premier ordre à coefficients continus.

- 1.2) On a

$$f(0) = rP_0 \left(1 - \frac{P_0}{\kappa} \right) > 0$$

car $\frac{P_0}{\kappa} \in]0, 1[$ par hypothèse. Donc $f(0) \neq 0$, par ailleurs, d'après la question précédente, si on note G la primitive de g qui s'annule en zéro, alors

$$f(t) = f(0) \cdot e^{G(t)},$$

or $e^{G(t)} \neq 0$ pour tout t donc $f(t) \neq 0$ pour tout $t \geq 0$.

1.3) Notons $Q = rX \left(1 - \frac{X}{\kappa}\right)$, c'est un polynôme de degré deux, de graphe une parabole orientée vers le bas, et de racines $0, \kappa$.

$$f = Q \circ P.$$

On sait que $P(0) \in]0, \kappa[$, donc la valeur initiale de P se situe strictement entre les racines de Q et $f(0) > 0$ (le polynôme Q est positif entre ses racines). Si en un certain t , $P(t) \notin]0, \kappa[$ alors $f(t) < 0$. Puisque f est continue, d'après le théorème des valeurs intermédiaires, il existerait t tel que $f(t) = 0$ — contradiction. Donc $\boxed{\text{pour tout } t \geq 0, P(t) \in]0, \kappa[}$.

2. (Résolution)

2.1) Après calculs, on trouve :

$$\forall N \in \mathbf{R} \setminus \{0, \kappa\}, \quad \boxed{\frac{\kappa}{N(\kappa - N)} = \frac{1}{N} + \frac{1}{\kappa - N}}.$$

2.2) On sait que f ne s'annule pas, donc on peut commencer par diviser de chaque côté l'équation différentielle par f . On résout alors :

$$\frac{P'(t)}{rP(t) \left(1 - \frac{P(t)}{\kappa}\right)} = \frac{P'(t)}{r} \cdot \frac{\kappa}{P(t)(\kappa - P(t))} = 1.$$

D'après la question précédente, elle est équivalente à

$$\frac{P'(t)}{r} \left(\frac{1}{P(t)} + \frac{1}{\kappa - P(t)} \right) = 1,$$

ou encore

$$\frac{1}{r} \left(\frac{P'(t)}{P(t)} - \frac{-P'(t)}{\kappa - P(t)} \right) = 1.$$

' En primitivant de chaque côté, on déduit l'existence d'une constante $K \in \mathbf{R}$ satisfaisant

$$\forall t \geq 0, \quad \frac{1}{r} (\ln |P(t)| - \ln |\kappa - P(t)|) = t + K.$$

Par propriété du log, en passant à l'exponentielle, et en utilisant le fait que $P(t) \in]0, \kappa[$, on obtient finalement

$$\forall t \geq 0, \quad \frac{P(t)}{\kappa - P(t)} = e^{r(t+K)} \iff \boxed{P(t) = \frac{\kappa e^{r(t+K)}}{1 + e^{r(t+K)}}}.$$

Exercice ANA.9.20 | Modèle de GOMBERTZ continu On dit qu'une fonction P suit un modèle de Gompertz de taux β, μ, κ si P est dérivable en t et vérifie :

$$P'(t) = (\beta - \mu) \ln \left(\frac{\kappa}{P(t)} \right) P(t) \quad (t \geq 0), \quad P(0) = P_0 > 0.$$

On appelle β le taux de natalité, μ le taux de mortalité, et κ la capacité du milieu. **On admet que P ne s'annule pas.**

Exprimer $P(t)$ en fonction de t . *Indication : On pourra commencer par diviser par P .*

Chapitre ANA.10.

Séries Numériques

Résumé & Plan

L'objectif principal de ce chapitre est de donner un sens à la notion de somme infinie, que nous allons définir comme une limite, en cas d'existence.

Nous allons donc étudier des suites particulières, que l'on appelle «série», *i.e.* les suites dont le terme général d'indice $n \in \mathbb{N}$ est la somme de $n + 1$ termes d'une suite numérique. Ces suites apparaissent naturellement dans bon nombre de problèmes en Mathématiques; de la définition de l'intégrale d'une fonction continue sur un segment, en passant par les calculs d'espérances de variable aléatoire discrètes à support dénombrable. Bref, une théorie générale s'impose.

1	Généralités	2
1.1	Définitions	2
1.2	Propriétés	5
2	Séries usuelles	6
2.1	Géométrique	6
2.2	Exponentielle	8
2.3	de RIEMANN	10

3	Séries de signe constant & Convergence absolue	11
3.1	Séries de signe constant	11
3.2	Séries à termes quelconques & Convergence absolue	17
3.3	Résumé : plan d'étude d'une série numérique	19
4	Séries doubles	19
5	Exercices	22
5.1	Séries simples	22
5.2	Séries doubles	28

$$\sum_{k=1}^{\infty} \frac{1}{k^4} = \frac{\pi^2}{90}$$

— Le saviez-vous ?



Cadre
 Dans toute la suite, nous considèrerons des suites définies sur \mathbb{N} : il convient d'adapter les notions pour des suites définies à partir d'un certain rang.

1. GÉNÉRALITÉS

1.1. Définitions

Définition ANA.10.1 | Somme partielle

Soit $(u_n) \in \mathbf{R}^{\mathbf{N}}$. On lui associe une nouvelle **suite** (S_n) en posant :

$$\forall n \geq 0, \quad S_n = \sum_{k=0}^n u_k.$$

- ▶ La **suite** (S_n) s'appelle la *série de terme général* u_n , et on la note généralement $(\sum u_n)$.
- ▶ Soit $n \in \mathbf{N}$. Alors le **réel** S_n est appelé la *somme partielle d'ordre* n .

Définition ANA.10.2 | Série convergente/divergente

- ▶ On dit qu'une série $(\sum u_n)$ *converge*, ou encore que $\sum_{n=0}^{\infty} u_n$ *converge*, si la suite des sommes partielles (S_n) converge, *i.e.* si et seulement s'il existe $S \in \mathbf{R}$ tel que :

$$\sum_{k=0}^n u_k \xrightarrow{n \rightarrow \infty} S.$$

- ▶ La limite S est alors appelée la *somme de la série*, et on la note $S = \sum_{n=0}^{+\infty} u_n$. On dit qu'une série *diverge* si elle ne converge pas.
- ▶ *Déterminer la nature d'une série* c'est déterminer si elle converge ou diverge.

Attention

Il faut prendre du recul sur l'expression « $\sum_{n=0}^{\infty} u_n$ converge». Car en écrivant $\sum_{n=0}^{\infty} u_n$ on suppose déjà l'existence d'une limite (et donc la convergence). C'est donc un abus de langage, mais cependant très utilisé y compris pour les intégrales généralisées que nous verrons dans le [Chapter ANA.11](#) : je l'utiliserai donc aussi.

Conjecture de la nature d'une série par tracé

Généralement, pour conjecturer la nature d'une série avec Python, on trace par exemple la suite des sommes partielles en fonction de $n \in \mathbf{N}$. On constate, avec les notations de la définition précédente, que (S_n) vérifie une relation de récurrence simple :

$$\forall n \in \mathbf{N}, \quad S_{n+1} = S_n + u_{n+1}.$$

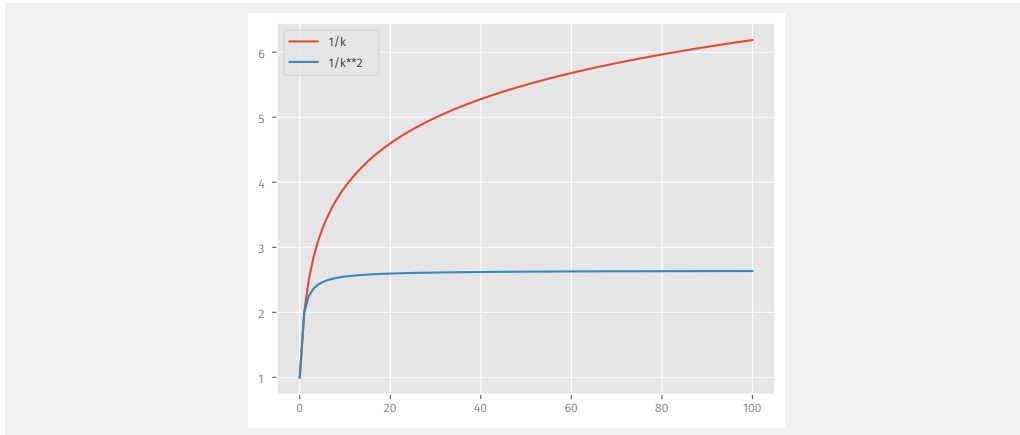
On peut donc construire facilement la liste $[S_0, \dots, S_n]$. Voyons deux exemples, les séries

$$\sum_{n=1}^{\infty} \frac{1}{n}, \quad \text{et} \quad \sum_{n=1}^{\infty} \frac{1}{n^2}.$$

```
import matplotlib.pyplot as plt
```

```
n = 100
S_1 = 1
L_1 = [S_1]
S_2 = 1
L_2 = [S_2]
for k in range(1, n+1):
    S_1 += 1/k
    S_2 += 1/k**2
    L_1.append(S_1)
    L_2.append(S_2)
plt.plot(L_1, label='1/k')
plt.plot(L_2, label='1/k**2')
plt.legend()
```

On peut ensuite tracer.



Proposition ANA.10.2 | Condition nécessaire de convergence sur le reste

Si $\sum_{n=0}^{\infty} u_n$ converge, on a pour tout entier $n \geq 0$, $S = S_n + R_n$. En particulier,

$$R_n \xrightarrow{n \rightarrow \infty} 0.$$

Preuve



Définition ANA.10.3 | Reste d'ordre n

Si $\sum_{n=0}^{\infty} u_n$ converge, on appelle *suite des restes partiels*, la suite (R_n) définie par :

$$\forall n \geq 0, R_n = S - S_n = \sum_{k=n+1}^{+\infty} u_k \quad \left(= \lim_{N \rightarrow \infty} \sum_{k=n+1}^N u_k \right).$$

Proposition ANA.10.1

Soit $(u_n)_{n \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ et $n_0 \in \mathbb{N}$. Alors :

$$\sum_{n=0}^{\infty} u_n \text{ converge} \iff \sum_{n=n_0}^{\infty} u_n \text{ converge}.$$

Remarque 1.1 — La nature d'une série ne dépend donc que du comportement de u_n lorsque $n \rightarrow \infty$.

Preuve Précisons les deux membres de l'équivalence.

- ▶ $\sum_{n=0}^{\infty} u_n$ converge si et seulement si $\lim_{n \rightarrow \infty} (u_0 + \dots + u_n)$ converge.
- ▶ $\sum_{n=n_0}^{\infty} u_n$ converge si et seulement si $\lim_{n \rightarrow \infty} (u_{n_0} + \dots + u_n)$ converge.



SÉRIE DE FORME TÉLÉSCOPIQUE. On s'intéresse ici aux séries dont la somme partielle est une somme télescopique, *i.e.* avec un terme général de la forme $u_{n+1} - u_n, n \in \mathbb{N}$.

Proposition ANA.10.3 | Série télescopique

Soit $(u_n) \in \mathbb{R}^{\mathbb{N}}$. Alors :

$$(u_n) \text{ converge} \iff \sum_{n=0}^{\infty} (u_{n+1} - u_n) \text{ converge}.$$

Preuve



**Méthode** Série quasi-téléscopique

Lorsque le terme général d'une série est de la forme $u_{n+p} - u_n$, avec $p \in \mathbf{N}$ fixé, alors on se ramène à une série téléscopique en écrivant

$$\forall n \in \mathbf{N}, \quad u_{n+p} - u_n = (u_{n+p} - u_{n+p-1}) + \cdots + (u_{n+1} - u_n).$$

Chacune des séries intervenant est alors téléscopique.

Exemple 1 — Montrer que $\sum_{n=1}^{\infty} \frac{1}{n(n+2)}$ converge, et calculer sa somme.

DIVERGENCE GROSSIÈRE. Lorsque la série converge, nous allons montrer que son terme général tend vers zéro, il s'agit donc d'une condition nécessaire. Lorsque ladite condition ne sera pas satisfaite, on parlera de *divergence grossière*.

Proposition ANA.10.4 |

Si $(\sum u_n)$ est une série convergente, **alors** :

$$u_n \xrightarrow{n \rightarrow \infty} 0.$$

Preuve

En effet, $u_n = S_n - S_{n-1}$ pour tout entier $n \geq 1$. Si $\sum_{n=0}^{\infty} u_n$ converge, choisissons $S \in \mathbf{R}$ de sorte que $S_n \xrightarrow{n \rightarrow \infty} S$. En passant à la limite, on récupère $u_n = S_n - S_{n-1} \xrightarrow{n \rightarrow \infty} S - S = 0$.

**Attention**

La réciproque est fautive! Nous montrerons que $(\sum \frac{1}{n})_{n \geq 1}$ diverge et pourtant $\frac{1}{n} \xrightarrow{n \rightarrow \infty} 0$. Ce constat est **le** point clef du chapitre à bien assimiler.


S'il suffisait que le terme général converge vers zéro pour assurer la convergence, ce chapitre sur les séries serait totalement inutile. Le caractère convergeant d'une série est **beaucoup plus fort** que la seule convergence de son terme général vers zéro. En revanche, si vous parvenez à montrer que le terme général ne converge pas vers zéro alors *a fortiori* la série ne converge pas. Nous parlerons *infra* de *divergence grossière*.

Définition/Proposition ANA.10.1 | Divergence grossière

On dit que $\sum_{n=0}^{\infty} u_n$ *diverge grossièrement* si :

$$u_n \not\xrightarrow{n \rightarrow \infty} 0.$$

Dans ce cas, $\sum_{n=0}^{\infty} u_n$ diverge.

Exemple 2 — La série $(\sum \cos(n))$ diverge grossièrement.  En effet, si l'on avait $\cos(n) \xrightarrow{n \rightarrow \infty} 0$, alors l'égalité $\cos(2n) = 2 \cos^2(n) - 1$ donnerait une absurdité par passage à la limite. Ou alors, nous avons déjà vu que cette suite n'admet carrément pas de limite.¹

1.2. Propriétés

Tous les théorèmes connus pour la convergence des suites peuvent être utilisés pour la convergence de la suite des sommes partielles, donc pour la convergence de la série.

Proposition ANA.10.5 | Linéarité de la somme, structure d'espace vectoriel des séries convergentes

Soient $\sum_{n=0}^{\infty} u_n$ et $\sum_{n=0}^{\infty} v_n$ deux séries convergentes à valeurs dans \mathbf{R} avec $(\lambda, \mu) \in \mathbf{R}^2$. Alors :

- ▶ $\sum_{n=0}^{\infty} (\lambda u_n + \mu v_n)$ est convergente, et
- ▶ $\sum_{n=0}^{\infty} (\lambda u_n + \mu v_n) = \lambda \sum_{n=0}^{\infty} u_n + \mu \sum_{n=0}^{\infty} v_n$.

En particulier,

- ▶ l'ensemble des séries convergentes est un sous-espace vectoriel de l'espace vectoriel des suites à valeurs dans \mathbf{R} .
- ▶ Une somme de séries convergentes est convergente.

Preuve Montrons que l'ensemble des suites convergentes est un sous-espace vectoriel de l'ensemble des suites.

- ▶ L'ensemble des séries convergentes est bien inclus dans l'ensemble des suites convergentes.
- ▶ La suite nulle $(0)_n = (0 + \dots + 0)_n$ est bien une série convergente.
- ▶ Soient $(u_n), (v_n)$ deux suites telles que $(\sum u_n), (\sum v_n)$ convergent, et $\lambda, \mu \in \mathbf{R}$. Alors la somme partielle de $(\sum \lambda u_n + \mu v_n)$ est définie pour tout $n \in \mathbf{N}$ par :

$$\sum_{k=0}^n (\lambda u_k + \mu v_k) = \lambda \sum_{k=0}^n u_k + \mu \sum_{k=0}^n v_k.$$

¹ En effet, \cos n'a pas de limite en $+\infty$, $\cos(2n\pi) \xrightarrow{n \rightarrow \infty} 1$ et $\cos(2n\pi + \pi/2) \xrightarrow{n \rightarrow \infty} 0$ alors que $2n\pi \xrightarrow{n \rightarrow \infty} \infty$ et $\pi/2 + 2n\pi \xrightarrow{n \rightarrow \infty} \infty$.

Par passage à la limite, on obtient alors l'égalité annoncée, et le fait que $\sum_{n=0}^{\infty} (\lambda u_n + \mu v_n)$ converge.

Attention

On veillera donc, avant d'invoquer la linéarité d'une somme infinie, à justifier la convergence de toutes les séries apparaissant dans le calcul.

Exemple 3 — d'égalité illicite La somme $\sum_{n=1}^{+\infty} \frac{1}{n(n+1)}$ est convergente mais on **ne peut pas écrire** :

$$\sum_{n=1}^{+\infty} \frac{1}{n(n+1)} = \sum_{n=1}^{+\infty} \left(\frac{1}{n} - \frac{1}{n+1} \right) = \sum_{n=1}^{+\infty} \frac{1}{n} - \sum_{n=1}^{+\infty} \frac{1}{n+1}$$

FAUX!

car les deux sommes du membre de droite sont divergentes.

Proposition ANA.10.6 | « Convergent + Divergent = Divergent »

Soient $\sum_{n=0}^{\infty} u_n$ une série convergente et $\sum_{n=0}^{\infty} v_n$ une série divergente. Alors :

$$\sum_{n=0}^{\infty} (u_n + v_n) \text{ est divergente.}$$

Preuve Notons (S_n) la somme partielle de $\sum_{n=0}^{\infty} (u_n + v_n)$. Soit $n \in \mathbf{N}$:

$$S_n = \sum_{k=0}^n (u_k + v_k) = \sum_{k=0}^n u_k + \sum_{k=0}^n v_k.$$

Donc (S_n) est une somme de deux suites, dont l'une converge (la suite $(\sum_{k=0}^n u_k)_n$), et l'autre diverge (la suite $(\sum_{k=0}^n v_k)_n$). Donc (S_n) diverge.

Afin d'achever ce paragraphe de propriété, énonçons une propriété déjà bien connue dans le cas des sommes finies, et qui reste vraie aussi pour les sommes infinies.

Proposition ANA.10.7 | Somme infinie nulle

Soit $(\sum u_n)$ une série convergente. Alors :

$$\sum_{n=0}^{\infty} |u_n| = 0 \implies \forall n \in \mathbf{N}, \quad u_n = 0.$$

En particulier, si (u_n) est une **suite positive** ou **négative**,

$$\sum_{n=0}^{\infty} u_n = 0 \implies \forall n \in \mathbf{N}, \quad u_n = 0.$$

Ainsi, la propriété déjà bien connue pour les sommes finies reste vraie pour les sommes infinies. L'hypothèse de signe constant (positif ou négatif) sur les termes reste cruciale.

Preuve Notons $S_n = \sum_{k=0}^n |u_k|$ pour tout entier $n \in \mathbf{N}$, alors :

$$0 \leq S_n \leq \sum_{k=0}^{\infty} |u_k| = 0,$$

puisque $|u_n| \geq 0$ pour tout n . Donc pour tout entier n , $S_n = 0$. Or, il s'agit d'une somme finie de termes positifs qui est nulle donc chaque terme est nul :

$$\forall k \in \llbracket 0, n \rrbracket, \quad |u_k| = 0.$$

Ceci implique

$$\forall k \in \llbracket 0, n \rrbracket, \quad u_k = 0.$$

Ceci étant vrai pour tout entier $n \in \mathbf{N}$, on récupère alors :

$$\forall n \in \mathbf{N}, \quad u_n = 0.$$

Remarque 1.2 — Cette proposition nous servira notamment dans le **Chapitre AEA.13** afin de montrer la propriété suivante : si une variable aléatoire réelle discrète X positive ou négative est d'espérance nulle, alors X est nulle.

2. SÉRIES USUELLES

Plusieurs séries usuelles sont au programme :

- ▶ les séries géométriques, dont nous allons établir la convergence à l'aide de résultats sur les sommes de termes géométriques,
- ▶ deux séries du type $\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$ avec $\alpha = 1$ ou 2 — appelées séries de RIEMANN,
- ▶ et la série dite « exponentielle », puisqu'en sommant nous retrouverons la fonction exp.

2.1. Géométrie**Définition/Proposition ANA.10.2 | Série géométrique**

Soit $q \in \mathbf{R}$ et $n \in \mathbf{N}$. On appelle *série géométrique de raison q* la série

$$(\sum q^n)_{n \geq 0}.$$

Elle converge si et seulement si $|q| < 1$. De plus, dans ce cas

$$\sum_{n=0}^{\infty} q^n = \frac{1}{1-q} \quad (\text{Geo})$$

et son reste à l'ordre n est donné par :

$$R_n = \frac{q^{n+1}}{1-q}.$$

Comme pour les sommes finies géométriques, la formule de la somme partielle s'adapte lorsque le premier terme de la somme est différent de 1 — c'est dans ce cas aussi la formule mentionnée sur le reste R_n . Nous pouvons donc retenir une formule générale (pour la somme partielle et le reste) de la manière suivante :

$$\sum_{n=0}^{\infty} \text{suite géométrique} = \text{premier terme} \times \frac{1}{1 - \text{raison}}.$$

Preuve*(Point clef — Formule de sommation de termes géométriques)*

$$\blacktriangleright \sum_{n=2}^{\infty} n(n-1)q^{n-2} = \left(\frac{1}{1-q}\right)^{(2)} = \frac{2}{(1-q)^3}.$$

Remarque 2.1 — La formule Eq. (GeoDer) signifie que :

$$\sum_{n=k}^{\infty} \frac{n(n-1)\dots(n-k+1)q^{n-k}}{(q^n)^{(k)}} = \left(\frac{1}{1-q}\right)^{(k)} = \left(\sum_{n=0}^{\infty} q^n\right)^{(k)}.$$

donc que,

Dans le cas des séries géométriques, on peut échanger dérivée en q et somme **infinie** en n .²

Preuve Nous faisons la preuve pour $k = 1$. Le cas général d'une dérivée d'ordre k se traite de la même manière.

Définition/Proposition ANA.10.3 | Série géométrique et géométrique dérivée (réelle)

Soit $q \in \mathbf{R}$ et $n \in \mathbf{N}$. On appelle *série géométrique dérivée* $k \geq 0$ fois de raison q la série

$$\left(\sum (q^n)^{(k)}\right)_{n \geq k} = \left(\sum n(n-1)\dots(n-k+1)q^{n-k}\right)_{n \geq k}.$$


Elle converge si et seulement si $|q| < 1$. De plus, dans ce cas :

$$\sum_{n=k}^{\infty} \frac{n(n-1)\dots(n-k+1)q^{n-k}}{(q^n)^{(k)}} = \left(\frac{1}{1-q}\right)^{(k)} = \frac{k!}{(1-q)^{k+1}} \quad (\text{GeoDer})$$

En particulier, si $k = 2$ et $k = 3$, nous avons :

$$\blacktriangleright \sum_{n=1}^{\infty} nq^{n-1} = \left(\frac{1}{1-q}\right)^{(2)} = \frac{1}{(1-q)^2},$$

²Ce qui n'est *a priori* pas évident, seul le cas des sommes finies est évident d'après les résultats généraux sur la dérivation

Exemple 4 — Déterminer la nature et somme éventuelle de $\sum_{n=0}^{\infty} \frac{n^2+1}{3^n}$. 



Méthode Calcul d'une somme de série polynôme fois géométrique

Une série du type $\left(\sum_{n \geq 0} n^k q^n\right)$, avec $|q| < 1, k \geq 0$, peut s'exprimer en fonction de séries géométriques et géométriques dérivées. En effet,

- ▶ pour $k = 1$, on écrit

$$n^1 q^n = q(nq^{n-1}).$$

- ▶ Pour $k = 2$, on écrit

$$n^2 q^n = q^2(n(n-1)q^{n-2}) + q(nq^{n-1}).$$

2.2. Exponentielle

Définition/Proposition ANA.10.4 | Série exponentielle

La série $\left(\sum \frac{x^n}{n!}\right)$, appelée *série exponentielle de x* , converge pour tout $x \in \mathbf{R}$, et :

$$\sum_{n=0}^{+\infty} \frac{x^n}{n!} = e^x.$$

Cette série peut d'ailleurs servir de définition de l'exponentielle — mais bien entendu inaccessible en Terminale ! Elle vient s'ajouter aux autres :

- ▶ l'unique solution du problème $y' = y, y(0) = 1$.
- ▶ La limite de $\left(1 + \frac{x}{n}\right)^n$ lorsque $n \rightarrow \infty$ définit e^x pour tout $x \in \mathbf{R}$.

Commençons par montrer le lemme ci-après : le résultat est trivial si $|x| < 1$ (le numérateur tend vers zéro !), mais pas dans le cas général.

Lemme ANA.10.1

Soit $x \in \mathbf{R}$, alors $\lim_{n \rightarrow \infty} \frac{x^n}{n!} = 0$.

Systematisons les calculs précédents dans une méthode.

Preuve (du lemme) L'idée est la suivante : il y a autant de facteurs au numérateur qu'au dénominateur, mais les facteurs du dénominateur ne cessent de grandir alors que ceux du numérateur sont toujours « des x ». Cela nous incite donc à découper le dénominateur en « deux morceaux », dont le second sera composé de facteurs strictement supérieurs à x .

$$0 \leq \left| \frac{x^n}{n!} \right| = \frac{|x|^{|x|}}{|x|!} \times \frac{|x|^{n-|x|}}{(|x|+1) \times \dots \times n} \leq \frac{|x|^{|x|}}{|x|!} \left(\frac{|x|}{|x|+1} \right)^{n-|x|}.$$

Puisque $\left| \frac{|x|}{|x|+1} \right| < 1$, le majorant converge vers zéro — c'est le terme général d'une suite géométrique. Donc $\lim_{n \rightarrow \infty} \frac{x^n}{n!} = 0$ par théorème d'encadrement.

Preuve On admettra dans la preuve que pour tout $x \in \mathbf{R}$:

$$\lim_{n \rightarrow \infty} \frac{x^n}{n!} = 0.$$

Montrons la formule suivante par récurrence sur $n \in \mathbf{N}$ — appelée égalité de TAYLOR-LAPLACE pour l'exponentielle :

$$e^x = \sum_{k=0}^n \frac{x^k}{k!} + \int_0^x \frac{(x-t)^n}{n!} e^t dt.$$



Ainsi, il suffit d'établir que $\lim_{n \rightarrow \infty} \int_0^x \frac{(x-t)^n}{n!} e^t dt = 0$. Pour ceci, on majore.

Exemple 5 — Déterminer la nature et somme éventuelle de $\sum_{n=0}^{\infty} \frac{n^2}{n!}$.

Systématisons les calculs précédents dans une méthode.

Méthode Calcul d'une somme de série polynôme fois exponentielle

Une série du type $\left(\sum_{n \geq 0} \frac{n^k}{n!}\right)$, avec $k \in \mathbf{N}$, peut s'exprimer en fonction de séries exponentielles après changement d'indice. En effet,

- ▶ pour $k = 1$, on écrit pour $n \geq 1$

$$\frac{n^1}{n!} = \frac{1}{(n-1)!}.$$

- ▶ Pour $k = 2$, on écrit

$$\begin{aligned} \frac{n^2}{n!} &= \frac{n}{(n-1)!} \quad \text{si } n \geq 1, \\ &= \frac{n-1+1}{(n-1)!} = \frac{1}{(n-2)!} + \frac{1}{(n-1)!}. \end{aligned} \quad \left. \vphantom{\frac{n^2}{n!}} \right) \text{ si } n \geq 2$$

Il apparaît parfois aussi des sommes exponentielles mais ne faisant intervenir que les termes pairs et impairs, elles s'expriment toutes les deux à l'aide de séries exponentielles comme nous allons le voir.

Exemple 6 — Exponentielles des termes pairs/impairs Soit $x \in \mathbf{R}$, alors :

$$\frac{e^x + e^{-x}}{2} = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!}, \quad \frac{e^x - e^{-x}}{2} = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!}. \quad 3$$



2.3. de RIEMANN


Définition/Proposition ANA.10.5 | Série harmonique, constante d'EULER

- ▶ La série harmonique $\sum_{n \geq 1} \frac{1}{n}$ diverge, mais ne diverge pas grossièrement.⁴
- ▶ De plus, il existe une constante γ , appelée *constante d'EULER*, qui vaut 0,577 à 10^{-3} près, telle que :

$$H_n \stackrel{\text{(déf.)}}{=} \sum_{k=1}^n \frac{1}{k} \underset{n \rightarrow \infty}{=} \ln(n) + \gamma + o(1).$$

Preuve

1. (Divergence.)

- ▶ **(Méthode 1 : par l'absurde et minoration)**  Notons $H_n = \sum_{k=1}^n \frac{1}{k}$ la somme partielle de la série d'ordre n . On peut montrer que pour tout entier n , $H_{2n} - H_n \geq \frac{1}{2}$. Si (H_n) convergeait disons vers $\ell \in \mathbf{R}$, alors en passant à la limite dans l'inégalité précédente, nous aurions $\ell - \ell = 0 \geq \frac{1}{2}$. C'est une

⁴De multiples proverbes interprètent ce résultat : «Les petits ruisseaux font les grands fleuves», «Les grains de sable amassés forment une montagne»

contradiction.

- ▶ **(Méthode 2 : par comparaison série intégrale)** Soient $k \geq 1$ et $n \in \mathbf{N}^*$. La fonction $t \mapsto \frac{1}{t}$ étant décroissante sur $[k, k+1]$, nous avons

$$\Rightarrow \int_k^{k+1} \frac{1}{k+1} dt \leq \int_k^{k+1} \frac{1}{t} dt \leq \int_k^{k+1} \frac{1}{k} dt \quad \left. \begin{array}{l} \forall t \in [k, k+1] \\ \int \end{array} \right\} \text{intégration sur } [k, k+1]$$

$$\frac{1}{k+1} \leq \ln(k+1) - \ln(k) \leq \frac{1}{k}.$$

En considérant le membre de droite, puis en sommant de $k = 1$ à n , on obtient :

$$\sum_{k=1}^n (\ln(k+1) - \ln(k)) = \ln(n+1) - 0 \leq \sum_{k=1}^n \frac{1}{k} = H_n. \quad \left. \begin{array}{l} \int \\ \end{array} \right\} \text{téléscopage}$$

Comme $\ln(n+1) \xrightarrow{n \rightarrow \infty} \infty$, on conclut en utilisant le théorème de divergence par minoration.

- 2. **(Existence de la constante d'EULER)** Nous avons déjà vu une première méthode dans l'Exemple 6 du Chapter ANA.9.

La technique mise en jeu dans la preuve précédente, appelée *comparaison série/intégrale*, sera étudiée un peu plus tard dans le chapitre.

Exemple 7 – Autres exemples

- 1. La série $\sum_{n \geq 1} \frac{1}{n^2}$ converge et : $\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$. Ce résultat est classique, mais non immédiat, nous prouverons la convergence un peu plus tard.
- 2. La série $\sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n}$ converge et : $\sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n} = \ln(2)$. Nous en verrons une démonstration en exercice.

3. SÉRIES DE SIGNE CONSTANT & CONVERGENCE ABSOLUE

3.1. Séries de signe constant

Définition ANA.10.4 | Séries à termes particuliers

- ▶ Une série $(\sum u_n)$ est à *termes positifs* (resp. *négatifs*) si $u_n \geq 0$ pour tout n (resp. $u_n \leq 0$ pour tout n). Si les inégalités précédentes sont vraies pour n assez grand, nous parlerons de séries à *termes positifs* (resp. *négatifs*) pour n assez grand.
- ▶ Une série $(\sum u_n)$ est à *termes de signes constants* si elle est à termes positifs ou négatifs.



Cadre

Nous nous intéressons dans cette section aux séries associées à des suites à termes positifs. Les résultats analogues s'appliquent :

- ▶ aux séries à termes négatifs en considérant $(\sum (-u_n))$.
- ▶ Aux séries à termes positifs pour n assez grand — extension triviale dans les démonstrations!

L'étude spéciale des séries à termes positifs est motivée par le constat ci-après, qui se trouve être central dans la démarche.

Soit (u_n) une suite de \mathbf{R}^+ , alors pour tout entier n :

$$S_{n+1} - S_n = (u_1 + \dots + u_n + u_{n+1}) - (u_1 + \dots + u_n) = u_{n+1} \geq 0.$$

La suite (S_n) est donc croissante, et sa convergence (la convergence de la série donc) se réduit à son éventuel caractère majoré (d'après le théorème de convergence monotone, toute suite monotone converge vers une limite finie si et seulement si elle est majorée ou minorée, cf. Section 1 sur les suites numériques).


En résumé :

les sommes partielles d'une série de signe constant sont monotones!

Proposition ANA.10.8 | Convergence des séries positives

Soit $(\sum u_n)$ une série à termes positifs, de somme partielle S_n pour tout $n \in \mathbf{N}$.
Alors :


$$\sum_{n=0}^{\infty} \text{converge} \iff (S_n) \text{ est majorée.}$$


Preuve  Soit $(\sum u_n)$ une suite à termes positifs, et soit (S_n) la suite de ses sommes partielles. Comme pour tout $n \in \mathbf{N}$ on a

$$S_{n+1} - S_n = u_{n+1} \geq 0,$$

la suite (S_n) est croissante. D'après le théorème de la limite monotone, elle est donc convergente dans $\overline{\mathbf{R}} = \mathbf{R} \cup \{\pm\infty\}$. De plus, elle converge si et seulement si elle est majorée, et elle diverge vers $+\infty$ sinon. La conclusion en découle.

Exemple 8 –

1. La série $\sum_{n=1}^{\infty} \frac{1}{n!}$ est convergente. Indication : Minorer $k!$, on n'utilisera pas de résultat sur les séries exponentielles. 

2. La série $\sum_{n=1}^{\infty} \frac{1}{n^2}$ est convergente. Indication : Minorer k^2 par $k(k-1)$ (pour $k \geq 2$). 

COMPARAISON DE SÉRIES À TERMES POSITIFS. Nous systématisons la technique mise en oeuvre dans les exemples précédents : nous avons majoré le terme général d'une série à termes positifs par celui d'une série à termes positifs convergente, nous en avons déduit la convergence de la série à termes positifs de départ. Plus généralement, nous avons le résultat qui suit.

Théorème ANA.10.1 | Comparaison de séries à termes positifs

Soient (u_n) et (v_n) deux suites à termes positifs. On suppose que l'on a

$$\forall n \in \mathbf{N}, \quad 0 \leq u_n \leq v_n,$$

ou au moins pour n assez grand. Alors :

- ▶ $\sum_{n=0}^{\infty} v_n$ converge $\implies \sum_{n=0}^{\infty} u_n$ converge,
- ▶ $\sum_{n=0}^{\infty} u_n$ diverge $\implies \sum_{n=0}^{\infty} v_n$ diverge.

Preuve



- ▶ On contrapose simplement la première partie.

L'utilisation du théorème de comparaison se fera en général au travers de deux moyens :

- ▶ soit on vous a fait trouver un encadrement (ou alors elle est évidente comme dans les exemples précédents),
- ▶ soit on utilise la formulation avec des équivalents ci-après. Elle est officiellement [H.P] mais les sujets la nécessitant l'admettent en préambule d'exercice.

Corollaire ANA.10.1 | Comparaison de séries à termes positifs avec \sim [H.P]

Soient (u_n) et (v_n) deux suites à termes positifs. On suppose que l'on a

$$u_n \underset{n \rightarrow \infty}{\sim} v_n.$$

Alors : $\sum_{n=0}^{\infty} u_n$ et $\sum_{n=0}^{\infty} v_n$ sont de même nature.

On utilisera donc ce résultat **uniquement si l'énoncé vous le permet**.



Attention Faux si les séries ne sont pas à termes positifs

- ▶ La série $\sum_{n=1}^{\infty} \frac{(-1)^n}{\sqrt{n}}$ converge; nous serons capable de le justifier sans difficulté après avoir vu les séries *alternées* en TD.
- ▶ La série $\sum_{n=1}^{\infty} \left(\frac{(-1)^n}{\sqrt{n}} + \frac{1}{n} \right)$ diverge en tant que somme d'une série convergente et d'une série divergente,
- ▶ alors que $\frac{(-1)^n}{\sqrt{n}} \underset{n \rightarrow \infty}{\sim} \frac{(-1)^n}{\sqrt{n}} + \frac{1}{n}$.

Preuve (Point clef — Traduire l'équivalent à l'aide d'une inégalité)

On sait que $u_n \underset{n \rightarrow \infty}{\sim} v_n$. Cela signifie que v_n est non nulle pour n assez grand et que $\frac{u_n}{v_n} \xrightarrow{n \rightarrow \infty} 1$, i.e.

$$\exists n_0 \in \mathbf{N}, \quad \forall n \geq n_0, \quad \left| \frac{u_n}{v_n} - 1 \right| < \frac{1}{2}.^a$$

En choisissant $\varepsilon = \frac{1}{2}$, on constate que pour n assez grand $\frac{u_n}{v_n}$ est entre $\frac{1}{2}$ et $\frac{3}{2}$:

$$\exists n_0 \in \mathbf{N}, \quad \forall n \geq n_0, \quad -\frac{1}{2} < \frac{u_n}{v_n} - 1 < \frac{1}{2} \iff \frac{1}{2} < \frac{u_n}{v_n} < \frac{3}{2}.$$

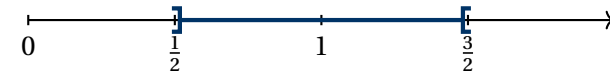


FIG. ANA.10.1 : Le quotient u_n/v_n est dans cet intervalle, au moins pour n assez grand

Donc pour n assez grand, on a finalement l'encadrement ci-après :

$$0 \leq \frac{v_n}{2} \leq u_n \leq \frac{3v_n}{2}. \quad (\star)$$

On peut à présent appliquer le théorème de comparaison.


- ▶ Supposons que $\sum_{n=0}^{\infty} u_n$ converge, alors puisque $0 \leq \frac{v_n}{2} \leq u_n$ pour n assez grand — partie gauche de (\star) — le théorème de comparaison livre que $\sum_{n=0}^{\infty} \frac{v_n}{2}$ converge et donc la convergence de $\sum_{n=0}^{\infty} v_n$.
- ▶ Supposons que $\sum_{n=0}^{\infty} u_n$ diverge, alors puisque $0 \leq u_n \leq \frac{3v_n}{2}$ pour n assez grand — partie droite de (\star) — le théorème de comparaison livre que $\sum_{n=0}^{\infty} \frac{3v_n}{2}$ diverge et donc la divergence de $\sum_{n=0}^{\infty} v_n$.

On a donc montré que les deux séries ont la même nature.

Exemple 9 — Déterminer la nature des séries ci-après, en appliquant le théorème de comparaison si une inégalité est évidente, ou le critère sur les équivalents sinon.

1. Les séries $\left(\sum \frac{\ln n}{n} \right)_{n \geq 1}$ et $\left(\sum \frac{1}{\ln n} \right)_{n \geq 2}$ divergent.

^aDéfinition de la limite avec « $\varepsilon = \frac{1}{2}$ »

2. La série $\left(\sum e^{-\sqrt{n}}\right)$ converge. 

3. La série $\left(\sum \frac{1}{2n+1}\right)$ diverge.

▶ (Inégalité) 

▶ (Équivalents) 

Le deuxième exemple peut se généraliser sans problème à des exponentielles plus générales.

 **Méthode** Convergence des séries exponentielles convergeant vers zéro

Soit (u_n) une suite telle que $u_n \xrightarrow{n \rightarrow \infty} \infty$. Alors :

1. $n^2 e^{-u_n} \xrightarrow{n \rightarrow \infty} 0$ par croissances comparées,
2. donc pour n assez grand, $n^2 e^{-u_n} \leq 1 \implies 0 \leq e^{-u_n} \leq \frac{1}{n^2}$.
3. Comme $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converge, le théorème de comparaison donne la convergence de $\sum_{n=0}^{\infty} e^{-u_n}$.


SÉRIES DE TYPE RIEMANN ET COMPARAISON SÉRIE—INTÉGRALE. On appelle série de RIEMANN toute série de la forme


$$\sum_{n=1}^{\infty} \frac{1}{n^{\alpha}}, \quad \alpha \in \mathbf{R}.$$

Il existe une technique générale permettant d'étudier la convergence de ce type de série, très communes dans la pratique mais commençons par un exemple. Rappelons avant tout que nous avons déjà croisé deux types de série de RIEMANN :

- ▶ $(\alpha = 2)$ $\sum_{n=1}^{\infty} \frac{1}{n^2}$ qui converge, nous l'avons admis,
- ▶ $(\alpha = 1)$ et $\sum_{n=1}^{\infty} \frac{1}{n}$ qui diverge, nous l'avons démontré.

Exemple 10 — Cas $\alpha = 1/2$ La série $\left(\sum \frac{1}{\sqrt{n}}\right)_{n \geq 1}$ est divergente.


- ▶ **(Par comparaison)** En comparant \sqrt{n} et n pour tout $n \geq 1$, établir la divergence. 

- ▶ **(Par comparaison série/intégrale)** Montrer la divergence par comparaison série-intégrale. 


- ▶ ou alors réaliser une comparaison série-intégrale.

De manière plus générale, nous avons le résultat suivant.

Exemple 11 – Cas général Montrer que : $\sum_{n \geq 1} \frac{1}{n^\alpha}$ est convergente $\iff \alpha > 1$.

- ▶ **(Par comparaison)** En comparant $\frac{1}{n^\alpha}$ à $\frac{1}{n}$ ou $\frac{1}{n^2}$ en fonction de α . Notons pour commencer que la série diverge grossièrement si $\alpha \leq 0$. 

- ▶ **(Par comparaison série/intégrale)** En effectuant une comparaison série-intégrale. Notons pour commencer que la série diverge grossièrement si $\alpha \leq 0$.

 Supposons donc $\alpha > 0$ et posons $S_n = \sum_{k=1}^n \frac{1}{k^\alpha}$ pour $n \geq 1$ la somme partielle

de la série. Par décroissance de $t \mapsto \frac{1}{t^\alpha}$, on a pour $k \geq 1$:

$$\frac{1}{(k+1)^\alpha} \leq \int_k^{k+1} \frac{dt}{t^\alpha} \leq \frac{1}{k^\alpha}$$

Deux techniques semblent donc possibles pour étudier une série de RIEMANN :

- ▶ comparer $\frac{1}{n^\alpha}$ à $\frac{1}{n}$ et $\frac{1}{n^2}$,

puis pour $n \geq 1$, nous obtenons avec la relation de CHASLES en sommant entre 1 et $n - 1$:

$$S_{n+1} - 1 \leq \int_1^{n+1} \frac{dt}{t^\alpha} \leq S_n$$

d'où, en mettant S_n au milieu de l'encadrement et en calculant l'intégrale :

$$\frac{(n+1)^{-\alpha+1} - 1}{1-\alpha} \leq S_n \leq 1 + \frac{n^{-\alpha+1} - 1}{1-\alpha}.$$

- Si $\alpha = 1$, nous avons déjà montré la divergence de la série.
- Si $\alpha > 1$: alors $n^{-\alpha+1} \xrightarrow{n \rightarrow \infty} 0$, donc (S_n) est majorée. Puisqu'elle est de plus croissante, elle converge. Donc la série converge.
- si $\alpha \in]0, 1[$: alors $(n+1)^{-\alpha+1} \xrightarrow{n \rightarrow \infty} \infty$, donc par divergence par minoration on déduit que (S_n) diverge. Donc la série diverge.

► En déduire que si $\alpha > 1$, on a :

$$\sum_{n=1}^{+\infty} \frac{1}{n^\alpha} \leq \frac{\alpha}{\alpha-1}.$$



Remarque 3.1 — De cet exemple on retient :

- le cas « limite » pour les natures des séries de RIEMANN est donc $\left(\sum \frac{1}{n}\right)$.
- Tout ce qui décroît plus lentement vers zéro, comme par exemple $\left(\sum \frac{1}{\sqrt{n}}\right)$ et $\left(\sum \frac{1}{\sqrt[3]{n}}\right)$, sont divergentes.
- Tout ce qui décroît plus rapidement vers zéro comme par exemple $\left(\sum \frac{1}{n^2}\right)$ et $\left(\sum \frac{1}{n^{2022}}\right)$, sont convergentes.

La démarche utilisée dans la preuve précédente, qui consiste à comparer les sommes partielles à des intégrales, est appelée *comparaison série—intégrale*. Elle fonctionne pour toutes les séries de la forme $\sum f(n)$ où f est une fonction continue, positive et monotone, dont une primitive est connue. Voyons un exemple qui sort du contexte des séries de RIEMANN.

Exemple 12 — *Comparaison série-intégrale pour un terme général décroissant* Étudier la convergence de la série $\left(\sum \frac{1}{n \ln n}\right)_{n \geq 2}$.

Théorème ANA.10.2 | Convergence absolue \Rightarrow Convergence

Soit $(\sum u_n)$ une série. Alors :

$$\left(\sum |u_n|\right) \text{ converge} \implies \left(\sum u_n\right) \text{ converge.}$$

Autrement dit, toute série absolument convergente est convergente.

Définition ANA.10.6 | Série semi-convergente

On dit qu'une série $(\sum u_n)$ est *semi-convergente* si la série $\sum_{n=0}^{\infty} u_n$ converge mais pas $(\sum |u_n|)$.

Remarque 3.2 — La comparaison série—intégrale permet souvent de donner en plus un équivalent des restes partiels R_n lorsque la série converge (ils tendent vers zéro), ou un équivalent des sommes partielles S_n lorsque la série diverge (elles tendent vers $+\infty$).

3.2. Séries à termes quelconques & Convergence absolue

Nous considérons de nouveau dans cette section des suites dont les termes peuvent être des réels de signe quelconque. Nous allons regarder une notion de convergence plus forte que la convergence des sommes partielles : il s'agit de la *convergence absolue*.

Définition ANA.10.5 | Série absolument convergente

On dit qu'une série $(\sum u_n)$ *converge absolument* si la série $(\sum |u_n|)$ converge.

Remarque 3.3 — La série $(\sum |u_n|)$ est une série à termes positifs : tous les critères vus dans la [Section 3](#) peuvent donc s'appliquer et c'est cela le gros avantage de la notion de convergence absolue.

Attention Réciproque fautive

La réciproque du théorème précédent est fautive, précisément pour toutes les séries semi-convergentes. En effet, on peut montrer que la série $(\sum \frac{(-1)^n}{n})$ converge, mais $(\sum |\frac{(-1)^n}{n}|) = (\sum \frac{1}{n})$ ne converge pas (nous l'avons déjà vu).

Preuve Nous avons, pour tout $n \in \mathbf{N}$:

$$-|u_n| \leq u_n \leq |u_n| \implies 0 \leq u_n + |u_n| \leq 2|u_n|.$$

Ainsi la série $(\sum (u_n + |u_n|))$ est une série à termes positifs, tout comme $(\sum |u_n|)$ qui converge par hypothèse, donc en appliquant le théorème de comparaison, nous déduisons que :

$$(\sum (u_n + |u_n|)) \text{ converge.}$$

Mais comme $u_n = (u_n + |u_n|) - |u_n|$, la série $(\sum u_n)$ s'exprime alors comme la différence de deux séries convergentes, elle est donc elle aussi convergente.

Nous croiserons rarement des séries à termes quelconques semi-convergentes (donc telle que $(\sum |u_n|)$ ne converge pas), tout simplement parce que les techniques existantes permettant de les traiter (transformation d'ABEL, critère des séries alternées, etc.) ne sont pas au programme de BCPST.

Exemple 13 — La série $\sum_{n \geq 1} \frac{\cos n}{n^2}$ est absolument convergente, donc convergente.



Proposition ANA.10.9 | Structure d'espace vectoriel des séries absolument convergentes

Soient $\sum_{n=0}^{\infty} u_n$ et $\sum_{n=0}^{\infty} v_n$ deux séries absolument convergentes à valeurs dans \mathbf{R} avec $(\lambda, \mu) \in \mathbf{R}^2$. Alors :

- ▶ $\sum_{n=0}^{\infty} (\lambda u_n + \mu v_n)$ est absolument convergente, et
- ▶ En particulier, l'ensemble des séries convergentes est donc un sous-espace vectoriel de l'espace vectoriel des suites convergentes.

Preuve



Proposition ANA.10.10 | Inégalité triangulaire généralisée

Soit $(\sum u_n)$ une série absolument convergente. Alors :

$$\left| \sum_{n=0}^{+\infty} u_n \right| \leq \sum_{n=0}^{\infty} |u_n|.$$

Preuve



Pour tout $n \in \mathbf{N}$, on a d'après l'inégalité triangulaire usuelle :

$$\left| \sum_{k=0}^n u_k \right| \leq \sum_{k=0}^n |u_k|.$$

Le résultat s'obtient alors par passage à la limite, la fonction $x \mapsto |x|$ étant continue.

ORDRE DE SOMMATION. On admet le théorème suivant qui sera utile en probabilités et qui nous autorise, s'il y a convergence absolue, à sommer dans n'importe quel ordre les éléments d'une série. Attention, cela n'est pas du tout évident ! Nous pouvons toujours le faire pour une somme finie, mais, pour une somme quelconque, cela n'est pas clair du tout. Considérons par exemple pour tout $n \in \mathbf{N}$ la suite $u_n = \frac{(-1)^n}{n+1}$.

Alors on peut démontrer que $(\sum u_n)$ converge (on note $\ell \in \mathbf{R}$ la somme de la série), mais pas absolument (c'est la série harmonique). En effet nous ne pouvons pas permuter n'importe comment l'ordre des termes, puisque si cela était possible nous

aurions :

$$\begin{aligned}
 \ell &= \boxed{1} - \frac{1}{2} + \boxed{\frac{1}{3}} - \frac{1}{4} + \boxed{\frac{1}{5}} - \frac{1}{6} + \frac{1}{7} - \dots \\
 &= \left(\boxed{1} - \frac{1}{2} - \frac{1}{4} \right) + \left(\boxed{\frac{1}{3}} - \frac{1}{6} - \frac{1}{8} \right) + \dots + \left(\boxed{\frac{1}{2k-1}} - \frac{1}{4k-2} - \frac{1}{4k} \right) + \dots \\
 &= \left(\frac{1}{2} - \frac{1}{4} \right) + \left(\frac{1}{6} - \frac{1}{8} \right) + \left(\frac{1}{10} - \frac{1}{12} \right) + \dots + \left(\frac{1}{4k-2} - \frac{1}{4k} \right) + \dots \\
 &= \frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \frac{1}{10} - \frac{1}{12} + \dots + \frac{1}{4k-2} - \frac{1}{4k} + \dots \\
 &= \frac{1}{2} \left(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots \right) = \frac{\ell}{2}.
 \end{aligned}$$

Théorème ANA.10.3 | Permutation des termes d'une série absolument convergente

Soit $\sigma : \mathbf{N} \rightarrow \mathbf{N}$ une bijection, et $(\sum u_n)_{n \geq 0}$ une série absolument convergente.

Alors : $(\sum u_{\sigma(n)})$ converge absolument, et $\sum_{n=0}^{\infty} u_n = \sum_{n=0}^{\infty} u_{\sigma(n)}$.

Preuve Admis.

Remarque 3.4 — Si la convergence absolue est en défaut, on ne peut rien dire
 Lorsque $(\sum u_n)_{n \geq 0}$ ne converge pas absolument mais converge (i.e. $(\sum u_n)_{n \geq 0}$ est semi-convergente), alors on peut même construire une bijection $\sigma : \mathbf{N} \rightarrow \mathbf{N}$ telle que $(\sum u_{\sigma(n)})_{n \geq 0}$ diverge. L'hypothèse de convergence absolue est donc cruciale.

3.3. Résumé : plan d'étude d'une série numérique



4.

SÉRIES DOUBLES

Nous aurons besoin en probabilités de permuter parfois des sommes infinies (pour des calculs d'espérance de couples de variables aléatoires réelles notamment). Nous donnons ici sans démonstration le résultat principal : il s'agit du théorème de FUBINI, et plus précisément ici d'un cas particulier. Là encore, ce qui paraît évident pour des sommes finies ne l'est pas forcément pour des sommes quelconques : nous ne pourrions pas toujours permuter deux sommes infinies.

Définition ANA.10.7 | Suite double

On appelle *suite double* toute application d'un sous-ensemble \mathcal{N} de \mathbf{N}^2 dans \mathbf{R} , i.e. toute famille de scalaires de \mathbf{R} indexée par un sous-ensemble de \mathbf{N}^2 , elle est

notée $(u_{n,p})_{(n,p) \in \mathcal{N}}$.

Pour tout $(n, p) \in \mathcal{N}$, on dit que $u_{n,p}$ est le (n, p) -ième *terme* de la suite, ou le terme de *rang* (n, p) . Si $\mathcal{N} = ([n_0, \infty[\cap \mathbf{N}) \times ([n_1, \infty[\cap \mathbf{N})$ avec $n_0, n_1 \in \mathbf{N}$, on dit que $(u_{n,p})_{(n,p) \in \mathcal{N}}$ est définie à partir d'un certain rang.

Exemple 14 — Pour tout $i, j \geq 0$, si $u_{i,j} = \frac{1}{2^{i+j}}$, alors $(u_{i,j})$ est une suite double.

Cadre
 Dans toute la suite, nous considérerons des suites doubles définies sur \mathbf{N}^2 : il convient d'adapter les notions pour des suites définies à partir d'un certain rang.

Notation Abus de ...
 Comme d'habitude si le contexte n'impose pas de le préciser, nous notons $(u_{n,p})$ au lieu de $(u_{n,p})_{(n,p) \in \mathcal{N}}$. On note $\mathbf{R}^{\mathbf{N} \times \mathbf{N}}$ les suites doubles indexées par $\mathbf{N} \times \mathbf{N}$.

Théorème ANA.10.4 | Permutation de sommes infinies si $u_{n,p}$ est positive

Soit $(u_{n,p})$ une suite double réelle **positive**. On suppose que :

1. pour tout n , $(\sum_{p \geq 0} u_{n,p})_{n \geq 0}$ converge,
2. $(\sum_{p=0}^{\infty} \sum_{n \geq 0} u_{n,p})$ converge.

Alors : pour tout p , $(\sum_{n \geq 0} u_{n,p})_{p \geq 0}$ converge, et la série $(\sum_{n=0}^{\infty} \sum_{p \geq 0} u_{n,p})_{p \geq 0}$ converge, et on a :

$$\sum_{n=0}^{\infty} \left(\sum_{p=0}^{\infty} u_{n,p} \right) = \sum_{p=0}^{\infty} \left(\sum_{n=0}^{\infty} u_{n,p} \right). \tag{4.1}$$

Preuve Admis.

Remarque 4.1 — Version allégée Parfois dans la littérature on peut trouver comme seule hypothèse : « la suite $u_{n,p}$ est positive » au lieu des deux hypothèses de convergence *supra*.

En fait, si 1) n'est pas vérifiée, alors puisque $u_{n,p}$ est positive, la somme partielle diverge forcément vers $+\infty$:

$$\sum_{n=0}^{\infty} u_{n,p} = \infty \quad \text{pour tout entier } p \in \mathbf{N}.$$

Ainsi la série $(\sum_{p=0}^{\infty} \sum_{n=0}^{\infty} u_{n,p})$ de 2) diverge vers $+\infty$ elle aussi. On peut justifier dans ce cas que l'égalité Eq. (4.1) devient :

$$+\infty = +\infty.$$

Elle est donc vérifiée là encore, mais le programme ne souhaite pas que vous écriviez de telles choses.

Lorsque les séries sont à termes quelconques, là il est nécessaire de rajouter de la convergence absolue dans les hypothèses.

Théorème ANA.10.5 | Permutation de sommes infinies dans le cas général

Soit $(u_{n,p})$ une suite double réelle. On suppose que :

1. pour tout n , $(\sum |u_{n,p}|)$ converge,
2. $(\sum_{p=0}^{\infty} \sum_{n=0}^{\infty} |u_{n,p}|)$ converge.

Alors pour tout p , $(\sum |u_{n,p}|)$ converge, la série $(\sum_{n=0}^{\infty} \sum_{p=0}^{\infty} |u_{n,p}|)$ converge et on a :

$$\sum_{n=0}^{\infty} \left(\sum_{p=0}^{\infty} u_{n,p} \right) = \sum_{p=0}^{\infty} \left(\sum_{n=0}^{\infty} u_{n,p} \right). \tag{4.2}$$

Vous ne pouvez retenir que ce résultat, qui est plus général que celui sur les séries à termes positifs.

Preuve Admis.

Définition ANA.10.8 | Série double convergente

On dira que $(\sum u_{n,p})_{n,p \in \mathbb{N}}$ converge absolument si les hypothèses du théorème de FUBINI sont vérifiées, *i.e.*

1. pour tout n , $(\sum_{p \in \mathbb{N}} |u_{n,p}|)$ converge,

2. et $(\sum_{n \in \mathbb{N}} \underbrace{\sum_{p=0}^{\infty} |u_{n,p}|}_{\text{une série en } n})$ converge.

Dans ce cas, on notera $\sum_{(n,p) \in \mathbb{N}^2} u_{n,p}$ la valeur de la somme $\sum_{n=0}^{\infty} \sum_{p=0}^{\infty} |u_{n,p}|$.

Remarque 4.2 — D'après le **Théorème ANA.10.5**, nous pourrions aussi prendre comme définition de série double convergente : pour tout p , $(\sum_{n \in \mathbb{N}} |u_{n,p}|)$ converge, la série $(\sum_{n=0}^{\infty} \sum_{p \in \mathbb{N}} |u_{n,p}|)$ converge. L'une impliquant l'autre d'après le théorème de FUBINI.

Exemple 15 — Étudier la convergence absolue de la série double $(\sum_{(i,j) \in \mathbb{N}^2} \frac{1}{2^{i+j}})$.

Calculer sa somme le cas échéant. 

*** Fin du chapitre ***

5. EXERCICES

5.1. Séries simples

Exercice ANA.10.1 | Études de convergences Étudier la convergence de $(\sum u_n)$, et calculer la somme si cela vous semble possible, lorsque :

- ▶ $u_n = \frac{n+2}{2^n}$ pour $n \geq 0$,
- ▶ $u_n = \frac{n^2}{n!}$ pour $n \geq 0$,
- ▶ $u_n = \frac{n(-1)^n}{3^{n-2}}$ pour tout $n \geq 0$,
- ▶ pour $k \geq 0$, $u_n = \frac{\binom{n}{k}}{n!}$, pour tout $n \geq k$,
- ▶ $u_n = \frac{1}{n^2-1}$ pour $n \geq 2$,
- ▶ $u_n = ne^{-n}$,
- ▶ $u_n = \sin(2^{-n})$,
- ▶ $u_n = \ln\left(1 + \frac{1}{n+1}\right)$,
- ▶ $u_n = \ln\left(1 - \frac{1}{n^2}\right)$,
- ▶ $u_n = \frac{\ln(n)}{n}$ pour $n \geq 1$,
- ▶ $u_n = \frac{(-1)^n \cos(n)}{n^2}$ pour $n \geq 1$.

Dans tout l'exercice, on pourra utiliser, uniquement en cas de besoin, le résultat suivant : si $(u_n), (v_n)$ sont deux suites positives telles que $u_n \sim_{n \rightarrow \infty} v_n$, alors $(\sum u_n), (\sum v_n)$ sont de même nature.

Solution (exercice ANA.10.1)

- ▶ Soit $n \geq 0$. On reconnaît la somme d'une série géométrique et géométrique dérivée.

$$\begin{aligned} \sum_{k=0}^n u_k &= \sum_{k=0}^n \frac{k+2}{2^k} \\ &= \frac{1}{2} \sum_{k=1}^n k \left(\frac{1}{2}\right)^{k-1} + \sum_{k=0}^n \left(\frac{1}{2}\right)^k \\ &\xrightarrow{n \rightarrow \infty} \frac{1}{2} \frac{1}{\left(1 - \frac{1}{2}\right)^2} + \frac{1}{1 - \frac{1}{2}} \end{aligned} \quad \left. \vphantom{\sum_{k=0}^n} \right\} \left| \frac{1}{2} \right|$$

$$= \boxed{4.}$$

- ▶ Soit $n \geq 0$. On reconnaît quasiment une série exponentielle.

$$\begin{aligned} \sum_{k=0}^n u_k &= \sum_{k=0}^n \frac{k^2}{k!} \\ &= \sum_{k=1}^n \frac{k^2}{k!} \\ &= \sum_{k=1}^n \frac{(k-1)}{(k-1)!} + \sum_{k=1}^n \frac{1}{(k-1)!} \\ &= \sum_{k=2}^n \frac{1}{(k-2)!} + \sum_{k=1}^n \frac{1}{(k-1)!} \\ &\xrightarrow{n \rightarrow \infty} e + e \end{aligned} \quad \left. \vphantom{\sum_{k=0}^n} \right\} \text{changement indice}$$

$$= \boxed{2e.}$$

- ▶ Soit $n \geq 0$. On reconnaît quasiment une série géométrique.

$$\begin{aligned} \sum_{k=0}^n u_k &= \sum_{k=0}^n \frac{k(-1)^k}{3^{k-2}} \\ &= 9 \sum_{k=0}^n k \left(-\frac{1}{3}\right)^k \\ &= -3 \sum_{k=1}^n k \left(-\frac{1}{3}\right)^{k-1} \\ &\xrightarrow{n \rightarrow \infty} -3 \frac{1}{\left(1 + \frac{1}{3}\right)^2} \end{aligned} \quad \left. \vphantom{\sum_{k=0}^n} \right\} \left| -\frac{1}{3} \right| < 1$$

$$= \boxed{2e.}$$

- ▶ Pour tout $n \geq k$, on a

$$u_n = \frac{\binom{n}{k}}{n!} = \frac{1}{k!(n-k)!}.$$

Donc pour tout $N \geq n$,

$$\begin{aligned} \sum_{n=k}^N u_n &= \sum_{n=k}^N \frac{1}{k!(n-k)!} \\ &= \frac{1}{k!} \sum_{n=k}^N \frac{1}{(n-k)!} \\ &= \frac{1}{k!} \sum_{n=0}^{N-k} \frac{1}{n!} \quad \left. \begin{array}{l} \text{changement indice} \end{array} \right\} \\ &\xrightarrow{N \rightarrow \infty} \boxed{\frac{e}{k!}} \end{aligned}$$

► Pour tout $n \geq 2$,

$$\begin{aligned} \sum_{k=2}^n u_k &= \sum_{k=2}^n \frac{1}{k^2-1} = \sum_{k=2}^n \frac{1}{(k-1)(k+1)} \\ &= \frac{1}{2} \sum_{k=2}^n \left(\frac{1}{k-1} - \frac{1}{k+1} \right) \quad \left. \begin{array}{l} \text{changement d'indice} \\ \text{téléscopage} \end{array} \right\} \\ &= \frac{1}{2} \left(\sum_{k=1}^{n-1} \frac{1}{k} - \sum_{k=3}^{n+1} \frac{1}{k} \right) \\ &= \frac{1}{2} \left(1 + \frac{1}{2} - \frac{1}{n} - \frac{1}{n+1} \right) \\ &\xrightarrow{N \rightarrow \infty} \boxed{\frac{3}{4}} \end{aligned}$$

► Soit $n \geq 0$. On reconnaît quasiment une série géométrique dérivée.

$$\begin{aligned} \sum_{k=0}^n u_k &= \sum_{k=0}^n k \left(\frac{1}{e} \right)^k \\ &= \frac{1}{e} \sum_{k=1}^n k \left(\frac{1}{e} \right)^{k-1} \quad \left. \begin{array}{l} \left| \frac{1}{e} \right| < 1 \end{array} \right\} \\ &\xrightarrow{N \rightarrow \infty} \frac{1}{e} \frac{1}{\left(1 - \frac{1}{e}\right)^2} \\ &= \boxed{\frac{e}{(1-e)^2}} \end{aligned}$$

► Par étude de fonction, on montre sans difficulté que pour tout $x \geq 0$, $\sin x \leq x$.
Donc pour tout $n \geq 0$:

$$0 \leq u_n \leq 2^{-n}.$$

Or, $(\sum 2^{-n})$ converge car $|\frac{1}{2}| < 1$. Donc d'après le théorème de comparaison pour les séries à termes positifs,

$$\boxed{(\sum u_n) \text{ converge.}}$$

► Pour tout $n \geq 0$,

$$\begin{aligned} \sum_{k=0}^n u_k &= \sum_{k=0}^n \ln \left(1 + \frac{1}{k+1} \right) = \sum_{k=0}^n \ln \left(\frac{k+2}{k+1} \right) \\ &= \sum_{k=0}^n \ln(k+2) - \sum_{k=0}^n \ln(k+1) \quad \left. \begin{array}{l} \text{changement d'indice} \\ \text{téléscopage} \end{array} \right\} \\ &= \sum_{k=2}^{n+2} \ln(k) - \sum_{k=1}^{n+1} \ln(k) \\ &= \ln(n+2) - \ln(1) \\ &\xrightarrow{N \rightarrow \infty} \boxed{\infty}. \end{aligned}$$

Donc la série diverge, résultat que l'on aurait pu retrouver en regardant un équivalent, et en utilisant la divergence de la série harmonique.

► Pour tout $n \geq 2$,

$$\begin{aligned} \sum_{k=2}^n u_k &= \sum_{k=2}^n \ln \left(\frac{(k+1)(k-1)}{k \cdot k} \right) = \sum_{k=2}^n \ln \left(\frac{k+1}{k} \right) - \sum_{k=2}^n \ln \left(\frac{k}{k-1} \right) \quad \left. \begin{array}{l} \text{téléscopage} \end{array} \right\} \\ &= \ln \left(\frac{n+1}{n} \right) - \ln 2 \\ &\xrightarrow{N \rightarrow \infty} \boxed{-\ln 2}. \end{aligned}$$

► Pour tout $n \geq 3$, $u_n = \frac{\ln(n)}{n} \geq \frac{1}{n} \geq 0$. Or, $(\sum \frac{1}{n})$ diverge, donc d'après le théorème de comparaison pour les séries à termes positifs,

$$\boxed{(\sum u_n) \text{ diverge.}}$$

- Ce n'est pas une série à termes positifs, et ce n'est pas non plus une série usuelle. On regarde donc la convergence absolue,

$$\forall n \in \mathbf{N}^*, \quad 0 \leq |u_n| \leq \frac{1}{n^2}.$$

Or, $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converge, donc par théorème de comparaison pour les séries positives, on déduit que $\sum_{n=1}^{\infty} u_n$ converge (et même absolument!).

Exercice ANA.10.2 | On considère la série $(\sum u_n)_{n \geq 2}$ où $u_n = \frac{2n^2}{n^3-1}$ pour tout $n \geq 2$.

1. Montrer que pour tout $n \geq 2$, $u_n \geq \frac{2}{n}$.
2. En déduire la nature de $(\sum u_n)_{n \geq 2}$.
3. ➤_🔗 Écrire une fonction en Python qui étant donné un entier n renvoie une liste contenant les valeurs des sommes partielles d'indice k pour $k \in \llbracket 2, n \rrbracket$. Représenter alors graphiquement les 49 premières valeurs de ces sommes partielles.

Solution (exercice ANA.10.2)

1. Soit $n \geq 2$. Alors $u_n \geq \frac{2}{n}$ est équivalent à $\frac{2n^2}{n^3-1} \geq \frac{2}{n}$, soit $2n^3 \geq 2(n^3-1)$ soit $0 \geq -2$. Cette inégalité étant vérifiée, nous déduisons que $u_n \geq \frac{2}{n}$.
2. D'après le cours, la série $(\sum \frac{2}{n})_{n \geq 1}$ diverge, les séries étant positives, d'après le théorème de comparaison, nous déduisons que :

$$(\sum u_n)_{n \geq 2} \text{ diverge.}$$

3. ➤_🔗

```
import matplotlib.pyplot as plt
```

```
def liste_S():
```

```
    """
```

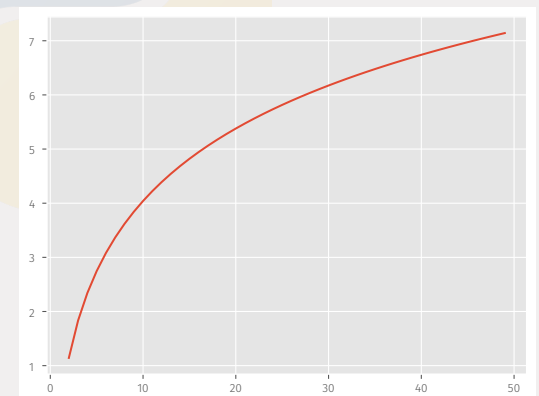
```
    retourne une liste contenant S2, ..., S49
```



```
    """
    liste_somme_part = [8/7]
    S = liste_somme_part[0]
    for i in range(3, 50):
        S += (2*i**2)/(i**3-1)
        liste_somme_part.append(S)
    return liste_somme_part
```

```
# ON TRACE ENSUITE CES TERMES
```

```
plt.plot(list(range(2, 50)), liste_S())
```



Puis on l'affiche.

Exercice ANA.10.3 | Avec des croissances comparées

1. Justifier que pour n assez grand dans \mathbf{N} , on a : $n \ln ne^{-2n} \leq \frac{1}{n^2}$.
2. Déterminer la nature de $(\sum (-1)^n n \ln ne^{-2n})_{n \geq 1}$.

Solution (exercice ANA.10.3)

1. Par croissances comparées, $n^2 \cdot n \ln ne^{-2n} \xrightarrow{n \rightarrow \infty} 0$, donc pour n assez grand, cette

quantité est majorée par 1. Ainsi, il existe $n_0 \in \mathbf{N}$ tel que pour tout $n \geq n_0$:

$$n \ln ne^{-2n} \leq \frac{1}{n^2}.$$

2. Il ne s'agit pas d'une série à termes positifs, donc étudions la convergence absolue. Pour tout $n \geq n_0$,

$$0 \leq |(-1)^n n \ln ne^{-2n}| \leq \frac{1}{n^2}.$$

Or $\left(\sum \frac{1}{n^2}\right)$ converge, donc par théorème de comparaison, on déduit que

$$\left(\sum (-1)^n n \ln ne^{-2n}\right)_{n \geq 1} \text{ converge absolument, donc converge.}$$

Exercice ANA.10.4 |

1. Montrer qu'il existe $a, b \in \mathbf{R}$ tels que :

$$\forall n \in \mathbf{N}, \quad \frac{4n}{n^4 + 2n^2 + 9} = \frac{a}{(n-1)^2 + 2} + \frac{b}{(n+1)^2 + 2}.$$

2. Étudier la nature de la série de terme général $u_n = \frac{4n}{n^4 + 2n^2 + 9}$, $n \geq 0$. Calculer sa somme en cas de convergence.

Solution (exercice ANA.10.4)

1. On cherche $a, b \in \mathbf{R}$ de sorte que

$$\forall n \in \mathbf{N}, \quad \frac{4n}{n^4 + 2n^2 + 9} = \frac{a}{(n-1)^2 + 2} + \frac{b}{(n+1)^2 + 2}.$$

Ceci est équivalent à

$$\frac{4n}{n^4 + 2n^2 + 9} = \frac{a((n+1)^2 + 2) + b((n-1)^2 + 2)}{n^4 + 2n^2 + 9}.$$

En développant le numérateur et en regroupant chaque terme suivant les puissances de n , on obtient la condition équivalente ci-après :

$$\forall n \in \mathbf{N}, \quad n^2(a+b) + 2n(a-b) + (a+2+b+2) = 4n.$$

Deux polynômes sont égaux si et seulement si ses coefficients le sont, donc

$$a+b=0, \quad 2(a-b)=4, \quad a+b=-4 \iff a=1, \quad b=-1.$$

Donc

$$\forall n \in \mathbf{N}, \quad \frac{4n}{n^4 + 2n^2 + 9} = \frac{1}{(n-1)^2 + 2} - \frac{1}{(n+1)^2 + 2}.$$

2. On constate alors une série de forme télescopique. Soit $N \geq 0$,

$$\begin{aligned} S_N &= \sum_{n=0}^N u_n = \sum_{n=0}^N \left(\frac{1}{(n-1)^2 + 2} - \frac{1}{(n+1)^2 + 2} \right) \\ &= \sum_{n=0}^N \left(\frac{1}{(n-1)^2 + 2} - \frac{1}{n^2 + 2} \right) + \sum_{n=0}^N \left(\frac{1}{n^2 + 2} - \frac{1}{(n+1)^2 + 2} \right) \quad \left. \vphantom{S_N} \right\} \text{télescopage} \\ &= \frac{1}{3} - \frac{1}{N^2 + 2} + \frac{1}{2} - \frac{1}{(N+1)^2 + 2} \\ &\xrightarrow{N \rightarrow \infty} \frac{1}{3} + \frac{1}{2} = \frac{5}{6}. \end{aligned}$$

Ce qui prouve donc la convergence de la série associée.

Exercice ANA.10.5 | Soit $u_0 \in]0, 1[$ et (u_n) la suite définie par : $\forall n \in \mathbf{N}, \quad u_{n+1} = u_n - u_n^2$.

1. Montrer que $u_n \in]0, 1[$ pour tout entier $n \in \mathbf{N}$, puis que $u_n \xrightarrow{n \rightarrow \infty} 0$.
2. Montrer que $(\sum u_n^2)$ est convergente et calculer la somme.

Solution (exercice ANA.10.5)

1. La suite (u_n) vérifie une relation de la forme $u_{n+1} = f(u_n)$ avec $f : x \in \mathbf{R} \mapsto x - x^2$, et $u_0 \in]0, 1[$. Or, le graphe de f est une parabole orientée vers le bas, de racines 0 et 1, le maximum étant atteint en $-\frac{1}{(-2)} = \frac{1}{2}$, la valeur du maximum est alors $f(1/2) = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$. Donc $f(]0, 1[) =]0, 1/4[\subset]0, 1[$, donc $]0, 1[$ est un intervalle stable par f . Ainsi, par récurrence immédiate on déduit que $u_n \in]0, 1[$ pour tout entier $n \in \mathbf{N}$.

Essayons d'appliquer le théorème de convergence monotone pour montrer la convergence de la suite. Soit $n \in \mathbf{N}$, $u_{n+1} - u_n = -u_n^2 \leq 0$ donc la suite (u_n) est décroissante, elle est de plus minorée par zéro donc converge vers une limite finie $\ell \in [0, 1]$. Puisque f est continue, $\ell = \ell - \ell^2$ soit $\ell^2 = 0$, ceci implique que $\ell = 0$. En conclusion :

$$\lim_{n \rightarrow \infty} u_n = 0.$$

2. Montrons que $(\sum u_n^2)$ est convergente et calculons sa somme. Soit $n \in \mathbf{N}$, alors

$$S_n = \sum_{k=0}^n u_k^2 = \sum_{k=0}^n (u_k - u_{k+1}),$$

la dernière somme étant télescopique, nous déduisons

$$S_n = u_0 - u_{n+1} \xrightarrow{n \rightarrow \infty} u_0.$$

Donc la série $(\sum u_n^2)$ est convergente, et $\sum_{k=0}^{\infty} u_k^2 = u_0$.

Exercice ANA.10.6 | On pose pour tout $n \in \mathbf{N}$, $u_n = \arctan\left(\frac{2}{(n+1)^2}\right)$.

1. Montrer la formule suivante : $\forall a, b \in \mathbf{R} \setminus \left\{\frac{\pi}{2} + k\pi, k \in \mathbf{Z}\right\}$:

$$\tan(a + b) = \frac{\tan a + \tan b}{1 - \tan a \tan b}.$$

2. Montrer que : $\forall n \in \mathbf{N}^*$, $u_n = \arctan\frac{1}{n} - \arctan\frac{1}{n+2}$.
 3. Étudier la convergence et calculer la somme de $(\sum u_n)_{n \geq 0}$.

Solution (exercice ANA.10.6)

1. Soient $a, b \in \mathbf{R} \setminus \left\{\frac{\pi}{2} + k\pi, k \in \mathbf{Z}\right\}$. Alors

$$\begin{aligned} \tan(a + b) &= \frac{\sin(a + b)}{\cos(a + b)} \\ &= \frac{\sin a \cos b + \sin b \cos a}{\cos a \cos b - \sin a \sin b} \end{aligned} \quad \left. \vphantom{\frac{\sin(a + b)}{\cos(a + b)}} \right\} \text{division par } \cos a \cos b$$

$$\frac{\tan a + \tan b}{1 - \tan a \tan b}.$$

On obtient par imparité aussi :

$$\tan(a - b) = \frac{\tan a - \tan b}{1 + \tan a \tan b}.$$

2. Notons $v_n = \arctan\frac{1}{n} - \arctan\frac{1}{n+2}$ pour tout $n \in \mathbf{N}^*$. Alors, en utilisant la question précédente, on a :

$$\tan v_n = \tan\left(\arctan\frac{1}{n} - \arctan\frac{1}{n+2}\right) = \frac{\frac{1}{n} - \frac{1}{n+2}}{1 + \frac{1}{n} \frac{1}{n+2}},$$

donc $\tan v_n = \frac{(n+2)-n}{n(n+2)+1} = \frac{2}{n(n+2)+1} = \frac{2}{(n+1)^2}$. Or, on sait que

$$\forall x \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[, \quad \arctan(\tan x) = x.$$

Donc comme $v_n \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ (c'est une différence de deux éléments dans $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$), on obtient :

$$v_n = \arctan\left(\frac{2}{(n+1)^2}\right) = \arctan\frac{1}{n} - \arctan\frac{1}{n+2}.$$

3. La série apparaît alors comme une série télescopique. Soit $n \in \mathbf{N}^*$,

$$\begin{aligned}
 S_n &= \sum_{k=0}^n u_k \\
 &= u_0 + \sum_{k=1}^n \left(\arctan \frac{1}{k} - \arctan \frac{1}{k+2} \right) \quad \left. \begin{array}{l} \text{question précédente pour } k \geq 1 \\ \text{té-} \\ \text{lé-} \\ \text{sco-} \\ \text{page} \end{array} \right\} \\
 &= \arctan(2) + \sum_{k=1}^n \left(\arctan \frac{1}{k} - \arctan \frac{1}{k+1} \right) + \sum_{k=1}^n \left(\arctan \frac{1}{k+1} - \arctan \frac{1}{k+2} \right) \\
 &= \arctan(2) + \arctan(1) - \arctan \left(\frac{1}{n+1} \right) + \arctan(2) - \arctan \left(\frac{1}{n+2} \right) \\
 &\xrightarrow{n \rightarrow \infty} \boxed{2 \arctan 2 + \frac{\pi}{4}}.
 \end{aligned}$$

Ce qui prouve donc la convergence de $\sum_{n=0}^{\infty} u_n$.

Exercice ANA.10.7 | Critère général des séries alternées Soit une suite $(v_n)_{n \in \mathbf{N}}$ définie pour tout entier $n \in \mathbf{N}$ par : $v_n = (-1)^n u_n$ avec (u_n) une suite réelle positive décroissant vers zéro.

1. On note (S_n) la suite des sommes partielles de (v_n) . Montrer que $(S_{2n})_{n \in \mathbf{N}}$ et $(S_{2n+1})_{n \in \mathbf{N}}$ sont deux suites adjacentes. Qu'en déduit-on pour $(\sum v_n)$?

2. (Application) Étudier les séries $\left(\sum \frac{(-1)^n}{n!} \right)_{n \geq 0}$, et la série $\left(\sum \sin \left(\pi \frac{n^2 + n + 1}{n + 1} \right) \right)_{n \geq 0}$.

Solution (exercice ANA.10.7)

1. On a, pour tout $n \in \mathbf{N}$: $S_n = \sum_{k=0}^n v_k$.

- ▶ $S_{2(n+1)} - S_{2n} = v_{2n+2} + v_{2n+1} = u_{2n+2} - u_{2n+1} \leq 0$ puisque (u_n) est décroissante,
- ▶ $S_{2(n+1)+1} - S_{2n+1} = v_{2n+3} + v_{2n+2} = -u_{2n+3} + u_{2n+2} \geq 0$ puisque (u_n) est décroissante,

▶ $S_{2n+1} - S_{2n} \xrightarrow{n \rightarrow \infty} 0$.

Donc $(S_{2n})_{n \in \mathbf{N}}$ et $(S_{2n+1})_{n \in \mathbf{N}}$ sont deux suites adjacentes, convergent donc *de facto* vers la même limite, donc (S_n) converge aussi vers cette limite par propriété sur les suites extraites. Ainsi, $\left(\sum v_n \right)$ converge.

2. (Application)

▶ Pour la première, notons $u_n = \frac{1}{n!}$ pour tout $n \geq 0$. Alors (u_n) décroît comme inverse de suite croissante, et tend vers 0 par règles usuelles. Donc

$\left(\sum \frac{(-1)^n}{n!} \right)_{n \geq 0}$ converge d'après la première question.

▶ Pour la seconde, constatons que

$$\sin \left(\pi \frac{n^2 + n + 1}{n + 1} \right) = \sin \left(\pi \frac{n(n + 1) + 1}{n + 1} \right) = \sin \left(n\pi + \frac{\pi}{n + 1} \right).$$

Rappelons que

$$\forall x \in \mathbf{R}, n \in \mathbf{N}, \quad (-1)^n \sin x = \sin(x + n\pi).$$

Donc

$$\sin \left(\pi \frac{n^2 + n + 1}{n + 1} \right) = (-1)^n \sin \left(\frac{\pi}{n + 1} \right).$$

Notons à présent $u_n = \sin \left(\frac{\pi}{n+1} \right)$. Cette suite décroît (puisque $(\pi/(n+1))$ décroît aussi, et $u_n \xrightarrow{n \rightarrow \infty} 0$). Donc

$\left(\sum \sin \left(\pi \frac{n^2 + n + 1}{n + 1} \right) \right)_{n \geq 0}$ converge.

Exercice ANA.10.8 | Transformation d'ABEL

1. Soient (u_n) et $(v_n)_{n \in \mathbf{N}}$ deux suites de $\mathbf{R}^{\mathbf{N}}$. On note $(V_n)_{n \in \mathbf{N}}$ la somme partielle de v . On suppose que

- ▶ $u_n \xrightarrow{n \rightarrow \infty} 0$, (u_n) est décroissante,
- ▶ et que (V_n) est bornée.

L'objectif de cette question est d'établir que $(\sum u_n v_n)_{n \geq 1}$ converge.

1.1) (Formule d'intégration par parties discrète) Montrer que pour tout $n \in \mathbf{N}$,

$$\sum_{k=0}^n u_k v_k = u_n V_n - \sum_{k=0}^{n-1} (u_{k+1} - u_k) V_k.$$

1.2) En déduire que la série $\sum_{n=0}^{\infty} u_n v_n$ converge.

2. Montrer que $\sum_{n=1}^{\infty} \frac{\cos n}{n}$ converge.

Solution (exercice ANA.10.8)

1. 1.1) Pour la démontrer, dans le cas d'une intégrale, on intègre simplement la relation $(fg)' = f'g + fg'$, dans le cas de sommes on réindexe à l'aide d'un changement de variable. Pour tout $n \in \mathbf{N}$,

$$\begin{aligned} & \sum_{k=0}^n u_k v_k \\ &= \sum_{k=1}^n u_k v_k + u_0 v_0 = \sum_{k=1}^n u_k (V_k - V_{k-1}) + u_0 v_0 \\ &= \sum_{k=1}^n u_k V_k - \sum_{k=1}^n u_k V_{k-1} + u_0 v_0 \\ &= \sum_{k=1}^n u_k V_k - \sum_{k=0}^{n-1} u_{k+1} V_k + u_0 v_0 \quad \left. \begin{array}{l} \\ \end{array} \right\} \text{changement d'indices} \\ &= \sum_{k=1}^{n-1} (u_k - u_{k+1}) V_k + u_0 v_0 - u_1 v_0 + u_n v_n \\ &= \sum_{k=0}^{n-1} (u_k - u_{k+1}) V_k + u_n v_n. \end{aligned}$$

1.2) Il suffit de démontrer que le membre de droite admet une limite lorsque $n \rightarrow \infty$. Comme V est une suite bornée, u converge vers 0, la suite $(u_n V_n)$ converge donc vers 0. Le premier terme est une somme partielle de série absolument convergente : en effet, puisque $u_k - u_{k+1} \geq 0$ pour tout entier k et que (V_k) est bornée, il existe un constant M positive telle que pour tout n :

$$\sum_{k=0}^{n-1} |(u_k - u_{k+1}) V_k| \leq M \sum_{k=0}^{n-1} (u_k - u_{k+1}) = M(u_0 - u_n),$$

comme (u_n) converge vers zéro, la suite $(\sum_{k=0}^{n-1} |(u_k - u_{k+1}) V_k|)_{n \in \mathbf{N}}$ est majorée, et c'est une somme partielle de série positive donc elle est croissante, donc $\sum_{n=0}^{\infty} u_n v_n$ converge (et même absolument).

2. Il n'est pas clair qu'il y ait convergence absolue, puisque le cosinus est borné par 1. Mais, la suite $(\cos 1 + \dots + \cos n)_{n \in \mathbf{N}}$ est bornée. En effet,

$$\begin{aligned} \sum_{k=1}^n \cos k &= \operatorname{Re} \left(\sum_{k=1}^n e^{ik} \right) \\ &= \operatorname{Re} \left(e^i \cdot \frac{1 - (e^i)^n}{1 - e^i} \right) \\ &= \operatorname{Re} \left(e^i \frac{e^{\frac{ni}{2}} \sin(n/2)}{e^{\frac{i}{2}} \sin(1/2)} \right) \quad \left. \begin{array}{l} \\ \end{array} \right\} \text{angle moitié} \\ &= \frac{\sin(n/2 + 1) \sin(n/2)}{\sin(1/2)}. \end{aligned}$$

Donc $|\sum_{k=1}^n \cos k| \leq \frac{1}{|\sin(1/2)|}$ pour tout entier $n \in \mathbf{N}$, ce qui prouve que $(\cos 1 + \dots + V_n)$ est bornée. Enfin, la suite $(1/n)_n$ décroît vers zéro, donc d'après la question précédente, $\sum_{n=1}^{\infty} \frac{\cos n}{n}$ converge.

5.2. Séries doubles

Exercice ANA.10.9 | Étudier la nature des séries doubles ci-dessous, calculer leur somme lorsque cela est possible.

1. $\left(\sum_{i,j \geq 0} \frac{(i+j)\lambda^{i+j}}{i!j!} \right)_{i,j \geq 0}$,
2. $\left(\sum_{i,j \geq 0} \frac{a}{2^{i+j}} \right)_{i,j \geq 0}$, pour tout $a \in \mathbf{R}$. Déterminer a de sorte que la somme soit égale à un.

Solution (exercice ANA.10.9)

1. ▶ Fixons $i \in \mathbf{N}$, étudions la convergence absolue de la série en $j \in \mathbf{N}$. Soit $n \in \mathbf{N}$, alors

$$\begin{aligned} & \sum_{j=0}^n \frac{(i+j)|\lambda|^{i+j}}{i!j!} \\ &= \frac{|\lambda|^i}{i!} \left(\sum_{j=0}^n \frac{i|\lambda|^j}{j!} + \sum_{j=0}^n \frac{j|\lambda|^j}{j!} \right) \\ &= \frac{|\lambda|^i}{i!} \left(i \sum_{j=0}^n \frac{|\lambda|^j}{j!} + |\lambda| \sum_{j=1}^n \frac{|\lambda|^{j-1}}{(j-1)!} \right) \quad \left. \begin{array}{l} \\ \end{array} \right\} \text{changement d'indice + série exponentielle} \\ &\xrightarrow{n \rightarrow \infty} \frac{|\lambda|^i}{i!} (ie^{|\lambda|} + |\lambda|e^{|\lambda|}) \\ &= \frac{|\lambda|^i e^{|\lambda|}}{i!} (i + |\lambda|). \end{aligned}$$

- ▶ Étudions à présent la convergence absolue de la série en $i \in \mathbf{N}$ de terme général $\frac{|\lambda|^i e^{|\lambda|}}{i!} (i + |\lambda|)$. Soit $n \in \mathbf{N}$, alors

$$\begin{aligned} & \sum_{i=0}^n \left(\frac{|\lambda|^i e^{|\lambda|}}{i!} (i + |\lambda|) \right) \\ &= e^{|\lambda|} \left(\sum_{i=0}^n \frac{i|\lambda|^i}{i!} + \sum_{i=0}^n \frac{|\lambda|^{i+1}}{i!} \right) \\ &= e^{|\lambda|} \left(\sum_{i=1}^n |\lambda| \frac{|\lambda|^{i-1}}{(i-1)!} + |\lambda| \sum_{i=0}^n \frac{|\lambda|^i}{i!} \right) \quad \left. \begin{array}{l} \\ \end{array} \right\} \text{changement d'indice + série exponentielle} \\ &\xrightarrow{n \rightarrow \infty} e^{|\lambda|} |\lambda| (e^{|\lambda|} + e^{|\lambda|}) = 2e^{2|\lambda|} |\lambda|. \end{aligned}$$

De tous ces calculs, on déduit la convergence de la série double. En les refaisant sans valeur absolue, on obtient immédiatement :

$$\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{(i+j)\lambda^{i+j}}{i!j!} = \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \frac{(i+j)\lambda^{i+j}}{i!j!} = 2\lambda e^{2\lambda}.$$

2. ▶ Fixons $i \in \mathbf{N}$, étudions la convergence absolue de la série en $j \in \mathbf{N}$. Soit $n \in \mathbf{N}$, alors

$$\begin{aligned} & \sum_{j=0}^n \frac{|a|}{2^{i+j}} \\ &= \frac{|a|}{2^i} \sum_{j=0}^n \frac{1}{2^j} \quad \left. \begin{array}{l} \\ \end{array} \right\} \left| \frac{1}{2} \right| < 1 \\ &\xrightarrow{n \rightarrow \infty} \frac{|a|}{2^{i-1}}. \end{aligned}$$

- ▶ Étudions à présent la convergence absolue de la série en $i \in \mathbf{N}$ de terme général $\frac{|a|}{2^{i-1}}$. Soit $n \in \mathbf{N}$, alors

$$\begin{aligned} & \sum_{i=0}^n \frac{|a|}{2^{i-1}} \\ &= 2|a| \sum_{i=0}^n \frac{1}{2^i} \quad \left. \begin{array}{l} \\ \end{array} \right\} \left| \frac{1}{2} \right| < 1 \\ &\xrightarrow{n \rightarrow \infty} 4|a|. \end{aligned}$$

De tous ces calculs, on déduit la convergence de la série double. En les refaisant sans valeur absolue, on obtient immédiatement :

$$\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{a}{2^{i+j}} = \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \frac{a}{2^{i+j}} = 4a.$$

Donc la somme double est égale à 1 si et seulement si $a = \frac{1}{4}$.

Chapitre ANA.11.

Intégration

Résumé & Plan

Ce chapitre se découpe en deux grandes parties : une première de révisions à propos des calculs de primitives et d'intégrales un segment, une seconde ayant pour but de d'étendre l'intégrale à des intervalles plus généraux que les segments. Nous en aurons besoin plus tard pour définir la notion de variable aléatoire réelle à densité. Comme pour les séries, tout sera une affaire de passage à la limite.

1	Primitives & Intégration sur un segment	2
1.1	Primitives	2
1.2	Intégrale sur un segment	2
1.3	Propriétés	5
1.4	Lien entre primitive et intégrale	7
1.5	Calculs d'intégrales	9
1.6	Sommes de RIEMANN & Intégration Numérique	14
2	Intégration sur un intervalle quelconque	17
2.1	Généralités	18
2.2	Propriétés des intégrales convergentes	25
2.3	Calculs d'intégrales	26

2.4	Intégrales de fonctions de signe constant	29
2.5	Fonctions de signe quelconque & Convergence absolue	32
2.6	Plan d'étude d'une intégrale	34
3	Exercices	35
3.1	Intégrales sur un segment	35
3.2	Intégrales impropres	44

Le nombre d'or φ — i.e. la plus grande des solutions de $x^2 - x - 1 = 0$ — et π vérifient les identités ci-après :

$$\bullet \pi\varphi = \int_0^{\infty} \frac{5}{1+x^{10}} dx,$$

$$\bullet \frac{\pi}{\varphi} = \int_0^{\infty} \frac{5x^2}{1+x^{10}} dx.$$

Eh oui!

— Le saviez-vous ?

1. PRIMITIVES & INTÉGRATION SUR UN SEGMENT

Commençons ce chapitre par des révisions d'intégration de première année. On définit l'intégrale d'une fonction continue sur un segment à l'aide d'une primitive de cette fonction, l'existence d'une telle primitive sera quant à elle admise.

1.1. Primitives

Définition ANA.11.1

Soit $f : I \rightarrow \mathbf{R}$ une application définie sur un *intervalle* I de \mathbf{R} . On appelle *primitive* de f sur I toute application $F : I \rightarrow \mathbf{R}$ dérivable telle que $F' = f$.

Proposition ANA.11.1 | Ensemble des primitives

Soit $f : I \rightarrow \mathbf{R}$ une fonction définie sur un intervalle I . Si $F : I \rightarrow \mathbf{R}$ est une primitive de f sur l'intervalle I , alors les primitives de f sur I sont les fonctions de la forme $F + c$, où $c \in \mathbf{R}$.

On retiendra notamment que si f admet une primitive, alors elle en admet même une infinité¹ ! Il n'est donc pas question de parler de *la* primitive de f .

Preuve



Théorème ANA.11.1 | Existence de primitives pour les fonctions continues

Soit $f : I \rightarrow \mathbf{R}$ une fonction continue sur un intervalle I . Alors f possède une primitive sur I .

Comme déjà annoncé, ce théorème est admis.

¹Puisque si F est une primitive, toutes les fonctions $F + c$ avec c une constante en sont aussi



Méthode Justifier l'existence d'une primitive

Il suffit de montrer la continuité de la fonction.

1.2. Intégrale sur un segment



Cadre

Dans toute cette sous-section, la notation $[a, b]$ désignera toujours un segment, avec $a, b \in \mathbf{R}$.

Nous allons commencer par les fonctions continues sur un segment puis nous généraliserons aux fonctions continues par morceaux, *i.e.* continues sauf en quelques points en lesquels elle est prolongeable par continuité.

Définition/Proposition ANA.11.1 | Intégrale d'une fonction continue ou prolongeable par continuité sur un segment

Soit $f : [a, b] \rightarrow \mathbf{R}$ une fonction continue, **ou** continue sur $[a, b[$ (*resp.* $]a, b]$, *resp.* $]a, b)$ et prolongeable par continuité en a (*resp.* b , *resp.* a et b).

- ▶ **(Si f est continue sur $[a, b]$)** On appelle *intégrale de f sur le segment $[a, b]$* le réel noté $\int_a^b f$ (ou encore $\int_a^b f(x) dx$, $\int_{[a,b]} f(x) dx$) défini par :

$$\int_a^b f(x) dx \stackrel{(\text{déf.})}{=} [F(x)]_a^b \stackrel{(\text{déf.})}{=} F(b) - F(a),$$


où F désigne une primitive de f .

On appelle *intégrande de $\int_a^b f$* la fonction f .

- ▶ **(Si f est prolongeable par continuité sur $[a, b]$)** Si f est seulement continue sur $[a, b[,]a]b$ ou $]a, b]$ et **prolongeable par continuité aux bornes**, on appelle *intégrale de f sur le segment $[a, b]$* l'intégrale de son prolongement continu.


Remarquez que si $a = b$, alors avec les notations de la définition précédente, on a :

$$\int_a^a f = [F]_a^a = F(a) - F(a) = 0.$$

Preuve La quantité $\int_a^b f(x) dx$ ne dépend pas de la primitive choisie. 

La définition de l'intégrale est donc bien posée.

Nous verrons dans la suite de ce chapitre ce que l'on peut faire si f n'est ni continue, ni prolongeable par continuité.

Exemple 1 — L'intégrale $\int_0^\pi \frac{\sin t}{t} dt$ existe. 

EXTENSION AUX FONCTIONS CONTINUES PAR MORCEAUX.

Définition ANA.11.2 | Fonction continue/constante par morceaux sur un segment

Une fonction $f : [a, b] \rightarrow \mathbf{R}$ est dite *continue par morceaux* s'il existe une subdivision de $[a, b]$, i.e. une famille de points $x_0 < \dots < x_N$ avec N un entier, vérifiant $a = x_0 < x_1 < x_2 < \dots < x_{N-1} < x_N = b$, et telle que, pour tout $k \in \llbracket 0, n-1 \rrbracket$:

- ▶ f est continue sur $]x_k, x_{k+1}[$,
- ▶ f est prolongeable par continuité en x_k et en x_{k+1} .

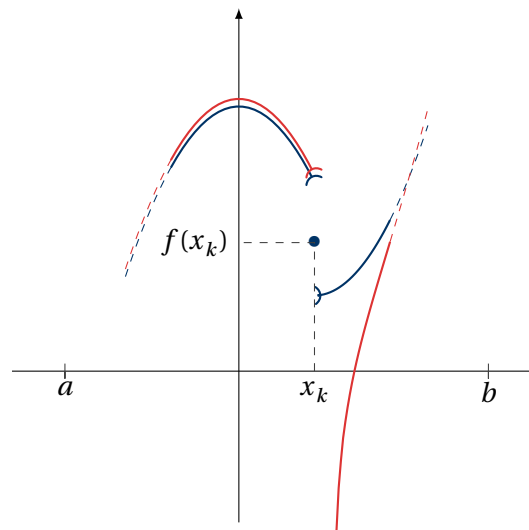
Si de plus f est constante sur $]x_k, x_{k+1}[$ pour tout $k \in \llbracket 0, n-1 \rrbracket$ alors elle est dite *constante par morceaux*.

Σ Notation

On note $\mathcal{C}_m^0([a, b], \mathbf{R})$ l'ensemble des fonctions continues par morceaux sur $[a, b]$.

⊗ Attention

Nous n'avons pas défini la notion de fonction continue par morceaux sur \mathbf{R} ! Seulement sur un segment.



La courbe bleue correspond à une fonction continue par morceaux. La rouge en revanche n'admet pas de limite finie en x_k et donc n'est pas continue par morceaux.

FIG. ANA.11.1. : Fonction continue (et non) par morceaux

Définition ANA.11.3 | Intégrale sur un segment d'une fonction continue par morceaux

Avec les notations de la définition précédente, on définit alors l'*intégrale de la fonction continue par morceaux* f comme étant le réel noté $\int_a^b f$ (ou encore $\int_a^b f(x) dx$, $\int_{[a,b]} f(x) dx$) égal à :

$$\int_a^b f(x) dx = \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} f|_{]x_k, x_{k+1}[}$$


On peut vérifier que ce nombre est indépendant de la subdivision choisie.

Méthode Justifier que l'intégrale d'une fonction définie sur un segment existe

Montrer que la fonction est :

- ▶ continue sur le segment,
- ▶ ou que l'on peut la prolonger en une fonction continue,
- ▶ ou encore qu'elle est continue par morceaux.

Remarque 1.1 — Changer une fonction en un nombre fini de points ne change pas l'intégrale. On observe que si deux fonctions continues par morceaux sur $[a, b]$ sont égales sauf en un nombre fini de points — autrement dit elles ne sont pas discontinues forcément aux mêmes points — alors elles ont même intégrale sur $[a, b]$: en effet, les prolongements aux bornes seront alors uniques, donc les intégrales associées seront égales.

Exemple 2 — Justifier l'existence et calculer $\int_0^n e^{\lfloor t \rfloor} dt$. 

VALEUR MOYENNE. Les éléments qui suivent seront le coeur de la définition de l'espérance d'une variable aléatoire à densité. Nous savons déjà «moyenner» des valeurs discrètes, à l'aide de séries, dans le cas continu on remplace essentiellement le symbole \sum par \int .

Définition ANA.11.4 | Valeur moyenne & Valeur moyenne pondérée

Soit $f : [a, b] \rightarrow \mathbf{R}$ une fonction continue par morceaux.

- ▶ **(Moyenne)** On appelle *valeur moyenne de f sur le segment $[a, b]$* le réel :

$$\frac{1}{b-a} \int_a^b f.$$

- ▶ **(Moyenne pondérée)** Si $p : [a, b] \rightarrow \mathbf{R}^+$ est une fonction continue par morceaux positive telle que :

$$\int_a^b p = 1.$$


Alors on appelle *valeur moyenne de f , pondérée par p* le réel

$$\int_a^b p \cdot f.$$

Exemple 3 — Par exemple, cela permet de donner un sens à la valeur moyenne d'une intensité électrique i sur un intervalle de temps $[0, T]$: $\frac{1}{T} \int_0^T i(t) dt$, ou de toute autre grandeur physique sur un intervalle de temps borné.

Exemple 4 —

- ▶ Déterminer la valeur moyenne de \sin et de \sin^2 sur $[0, 2\pi]$. 

- ▶ Déterminer la valeur moyenne de $t \mapsto t$ sur $[0, 2\pi]$ pondérée par $p : t \in [0, 2\pi] \mapsto \frac{1}{2\pi}$. 

1.3. Propriétés

Proposition ANA.11.2 | Propriétés de l'intégrale.

Soient I un intervalle et $(a, b) \in I^2$. Alors :

1. **(Linéarité)** L'application $\psi : \begin{cases} \mathcal{C}_m^0(I, \mathbf{R}) & \longrightarrow \mathbf{R} \\ f & \longmapsto \int_a^b f \end{cases}$ est une forme linéaire, i.e. pour tout $(f, g) \in (\mathcal{C}_m^0(I, \mathbf{R}))^2$ et $(\lambda, \mu) \in \mathbf{R}^2$, on a :

$$\int_a^b (\lambda f + \mu g) = \lambda \int_a^b f + \mu \int_a^b g.$$

2. **(Positivité)** Si $f \in \mathcal{C}_m^0(I, \mathbf{R})$ et $a \leq b$, alors :

$$f \geq 0 \implies \int_a^b f \geq 0.$$

3. **(Croissance)** Si $(f, g) \in (\mathcal{C}_m^0(I, \mathbf{R}))^2$ et $a \leq b$, alors :

$$f \leq g \implies \int_a^b f \leq \int_a^b g.$$


4. **(Relation de CHASLES)** Soient $f \in \mathcal{C}_m^0(I, \mathbf{R})$ et $c \in I$. Alors :

$$\int_a^c f = \int_a^b f + \int_b^c f.$$

Preuve

- Constater que si F et G sont des primitives de f et g , alors $\lambda F + \mu G$ est une primitive de $\lambda f + \mu g$.
- Si F est une primitive de f , alors l'hypothèse nous donne $F' \geq 0$, donc que F est croissante. On obtient immédiatement $\int_a^b f = F(b) - F(a) \geq 0$ puisque $a \leq b$.
- Il suffit d'appliquer le résultat précédent à la fonction $g - f$ et d'utiliser la linéarité.
- Si F est une primitive, alors $\int_a^c f = F(c) - F(a) = F(b) - F(a) + F(c) - F(b) = \int_a^b f + \int_b^c f$.

Exemple 5 — Intégrale à paramètre Pour tout $x \in \mathbf{R}_+$, posons : $A(x) = \int_0^1 \frac{e^t}{1+xt} dt$.

La fonction A est bien définie sur \mathbf{R}_+ et elle est décroissante. 

Théorème ANA.11.2 | Inégalité triangulaire & Majoration de l'intégrale d'un produit

Soit $(a, b) \in \mathbf{R}^2$ tel que $a \leq b$ et deux fonctions $f \in \mathcal{C}_m^0([a, b], \mathbf{R})$, $g \in \mathcal{C}_m^0([a, b], \mathbf{R})$ telle que g est bornée. Alors :

- ▶ $\left| \int_a^b f \right| \leq \int_a^b |f|$,
- ▶ $\left| \int_a^b fg \right| \leq \sup_{x \in [a, b]} |g(x)| \cdot \int_a^b |f|$.²

Preuve


²On rappelle que la borne supérieure d'une fonction bornée désigne le plus petit des majorants.


▶ On a $-|f| \leq f \leq |f|$. Donc en intégrant entre a et b , on déduit

$$-\int_a^b |f| \leq \int_a^b f \leq \int_a^b |f| \iff \left| \int_a^b f \right| \leq \int_a^b |f|.$$

▶ Pour la deuxième, on pose $M = \sup_{x \in [a,b]} |g(x)|$, puis on écrit : $\left| \int_a^b fg \right| \leq \int_a^b |fg| \leq \int_a^b M \cdot |f| = M \cdot \int_a^b |f|$.

Exemple 6 –

1. Pour tout $n \in \mathbf{N}$, posons $I_n = \int_0^1 \frac{x^n}{1+x^2} dx$ et $J_n = \int_0^1 x^n \arctan(1-nx) dx$. Montrons que les suites (nI_n) et (nJ_n) sont bornées. 

2. Donnons un équivalent en 0^+ de la fonction F définie par $F(x) = \int_x^{x^2} \frac{e^t}{t} dt$.
Indication : On pourra, par exemple, chercher à encadrer le numérateur de l'intégrande.  Il s'agit d'encadrer F pour x assez proche de zéro. Pour $x \in [0, 1]$, $x^2 \leq x$, les bornes de l'intégrale ne sont donc pas dans le bon sens. Notons plutôt $G(x) = -F(x) = \int_{x^2}^x \frac{e^t}{t} dt$ et cherchons un équivalent de G . Puisque \exp est croissante, on a

$$\forall x \in]0, 1[, \quad e^{x^2} \int_{x^2}^x \frac{1}{t} dt \leq G(x) \leq e^x \int_{x^2}^x \frac{1}{t} dt$$

ou de manière équivalente :

$$-\ln x \cdot e^{x^2} \leq G(x) \leq -\ln x \cdot e^x.$$

On déduit alors un équivalent, puisque pour $x \in]0, 1[$:

$$1 \leq \frac{G(x)}{-\ln x \cdot e^{x^2}} \leq \frac{-\ln x \cdot e^x}{-\ln x \cdot e^{x^2}} = e^{x-x^2} \xrightarrow{x \rightarrow 0^+} 1.$$

En conclusion :
$$F(x) \underset{x \rightarrow 0^+}{\sim} \ln x \cdot e^{x^2}.$$

Théorème ANA.11.3 | Fonction positive d'intégrale nulle

Soit $f : [a, b] \rightarrow \mathbf{R}$. On suppose que f est **continue** et de signe constant. Alors :

- $\int_a^b f = 0 \iff f$ est la fonction nulle.
- $f > 0$ (resp. > 0) $\implies \int_a^b f > 0$ (resp. < 0).

Attention

La continuité est indispensable, continue par morceaux ne suffit pas.³

Pour une fonction continue par morceaux, on peut utiliser ce résultat au prolongement continu de la fonction, sur chaque sous-intervalle.

Preuve

1. Seule l'implication directe n'est pas triviale. On la montre pour $f \geq 0$ (quitte à changer f en $-f$) en raisonnant par l'absurde. Supposons qu'il existe $x_0 \in [a, b]$ tel que $f(x_0) > 0$. Alors, en utilisant la définition de la continuité, on obtient l'existence de $\eta > 0$ tel que pour tout $x \in]x_0 - \eta, x_0 + \eta[$,

$$f(x) \in \left] \frac{f(x_0)}{2}, \frac{3f(x_0)}{2} \right[.$$

En notant $g = \frac{3f(x_0)}{2} \mathbb{1}_{]x_0 - \eta, x_0 + \eta[}$, on a alors $f \geq g$ puisque f est positive, et donc

$$\int_a^b f \geq \int_a^b g = 2\eta \frac{3f(x_0)}{2} = 3\eta f(x_0) > 0.$$

³ Il faudrait sinon remplacer « f est la fonction nulle » par « f est nulle sauf en un nombre fini de points » dans la conclusion.

Donc $\int_a^b f > 0$ — contradiction.

2. 

Preuve



1.4. Lien entre primitive et intégrale

Par définition de l'intégrale, il est nécessaire de connaître une primitive pour la calculer, il existe donc un fort lien entre les deux notions. Voyons lequel.

Théorème ANA.11.4 | Lien primitive / intégrale.

Soient I un intervalle, $a \in I$ et $f : I \rightarrow \mathbf{R}$ une fonction continue. Alors l'application

$$F : \begin{cases} I & \longrightarrow & \mathbf{R} \\ x & \longmapsto & \int_a^x f \end{cases}$$

est **l'unique primitive de f sur I qui s'annule en a** et est de classe \mathcal{C}^1 . Par conséquent, si f est de classe \mathcal{C}^1 sur I , alors elle est égale à l'intégrale de sa dérivée :

$$\forall x \in I, \quad f(x) - f(a) = \int_a^x f'.^4 \quad (\text{RFA})$$

Méthode Primitiver une fonction en utilisant une intégrale

Lorsque vous avez besoin d'une technique d'intégration (intégration par parties ou changement de variable par exemple) pour primitiver une fonction $f : I \rightarrow \mathbf{R}$, choisir $a \in I$, puis calculer $\int_a^x f$ pour tout $x \in I$.

⁴Cette égalité est appelée assez pompeusement «relation fondamentale de l'analyse». Pour notre définition de l'intégrale, elle est évidente. Mais pour d'autres définitions, il faut travailler pour l'établir.

Remarque 1.2 — L'écriture d'une fonction sous la forme Eq. (RFA) peut donner de précieux renseignements sur f si on en connaît certains sur f' . Voir l'exemple ci-après.

Exemple 7 — Inégalité des accroissements finis Soit $f \in \mathcal{C}^1([a, b], \mathbf{R})$ avec $a < b$ et telle que f' soit bornée par $M \in \mathbf{R}^+$. Montrer, en écrivant f comme intégrale de sa dérivée,⁵ que :

$$|f(b) - f(a)| \leq M|b - a|.$$



LES INTÉGRALES À BORNE(S) VARIABLE(S). Dans ce paragraphe, nous nous intéressons aux intégrales de fonctions dont une ou plusieurs bornes de l'intégrales dépendant d'une variable. Le théorème ci-dessous est clairement hors-programme, il faut donc uniquement en connaître la démarche de la preuve associée.

⁵Mais on peut aussi appliquer l'égalité des accroissements finis

Théorème ANA.11.5 | Intégrale à deux bornes variables [H.P]

Soient I un intervalle, a un point de I et $f : I \rightarrow \mathbf{R}$ une fonction continue. Soient par ailleurs $u, v : J \rightarrow I$ deux fonctions dérivables sur J où J est un intervalle réel. Alors la fonction

- ▶ $g : x \mapsto \int_{u(x)}^{v(x)} f$ est dérivable sur J et sa dérivée est donnée par :
- ▶ $\forall x \in J, \quad g'(x) = \frac{d}{dx} \left(\int_{u(x)}^{v(x)} f \right) = f(v(x)) \cdot v'(x) - f(u(x)) \cdot u'(x).$

Preuve (Point clef — **Introduire une primitive de l'intégrande**)

Puisque f est continue, choisissons une primitive de f notée F . 

Méthode Calcul d'une intégrale à bornes variables


Soient I un intervalle, a un point de I et $f : I \rightarrow \mathbf{R}$ une fonction continue. Soient par ailleurs $u, v : J \rightarrow I$ deux fonctions définies sur J où J est un intervalle réel.


Pour étudier la dérivabilité de $x \in J \mapsto \int_{u(x)}^{v(x)} f$, on :


1. introduit une primitive de f , notée F .
2. Alors : $\int_{u(x)}^{v(x)} f = F \circ v(x) - F \circ u(x).$
3. Justifier la dérivabilité et dériver à l'aide la formule de dérivation d'une composée.
4. On obtient *in fine* $\frac{d}{dx} \left(\int_{u(x)}^{v(x)} f \right) = f \circ v(x) v'(x) - f \circ u(x) u'(x)$ — cette formule ne doit pas être apprise par coeur, il faut savoir la retrouver en dérivant une composée.

Un point important est que le résultat ne dépend pas de F ; inutile donc de chercher à calculer F explicitement.

Exemple 8 — Fonction d'une variable dépendant d'une intégrale Justifier que f, g sont dérivables, où f, g sont définies ci-dessous, puis calculer leur dérivée.

1. $f : x \in \mathbf{R} \mapsto \int_0^x \sqrt{1+t^2} dt$. En déduire les variations de f . 

2. $g : x \in \mathbf{R} \mapsto \int_{-x}^{x^2} \arctan(t^2) dt$. En déduire les variations de g sur \mathbf{R}^+ . 

Exemple 9 — Fonction de deux variables dépendant d'une intégrale La fonction $i : (x, y) \in \mathbf{R}^2 \mapsto \int_{-x}^{y^2} \frac{1}{t^2 + t + 1} dt$ est bien définie, de classe \mathcal{C}^1 , et déterminer son gradient en tout point $(x, y) \in \mathbf{R}^2$. 

1.5. Calculs d'intégrales

Nous avons vu précédemment que calculer une primitive revient à un calcul d'intégrale. Pour ces dernières nous disposons de deux techniques principales de calcul : l'intégration par parties et le changement de variable.

Notez bien que les deux théorèmes qui suivent sont vraies pour des fonctions intégrées **continues** : pour les continues par morceaux, il faut les appliquer entre deux points d'une subdivision.

1.5.1. Intégration par parties

Théorème ANA.11.6 | Intégration par parties

Soient $u, v : [a, b] \rightarrow \mathbf{R}$ deux fonctions de classe \mathcal{C}^1 . Alors :

$$\int_a^b u \cdot v' = [u \cdot v]_a^b - \int_a^b u' \cdot v.$$



Attention

Toute intégration par parties doit être justifiée.

Preuve (Point clef — Intégrer la formule de dérivation d'un produit)

Notons que les hypothèses assurent que les fonctions intégrées (uv' et $u'v$) sont continues, donc intégrables, sur $[a, b]$. En regroupant les deux intégrales, on fait apparaître la dérivée d'un produit :

$$\int_a^b u' \cdot v' + \int_a^b u' \cdot v = \int_a^b (u' \cdot v' + u' \cdot v) = \int_a^b (uv)'(t) = [u \cdot v]_a^b.$$

d'où le résultat annoncé.





Méthode Quand utiliser l'intégration par parties ?


Pour intégrer un *produit* de deux fonctions, dont l'une est facile à *primitive* et l'autre est facile à *dériver*. Exemple : une exponentielle multipliée par un polynôme.


Exemple 10 — Calculer les intégrales suivantes (où $x \in \mathbf{R}$).

1. $\int_0^x \arctan t \, dt$, 

2. $\int_0^x t \ln(t^2 + 1) \, dt$. 

3. $\int_0^x (t^2 - t + 3)e^{2t} dt.$ 

4. $\int_0^x e^t \sin t dt.$ 

5. Calculer une primitive de $t \mapsto \frac{\ln t}{t^2}$ sur un domaine à préciser. 

1.5.2. Changement de variable

Théorème ANA.11.7 | Changement de variables.

Soient $f : I \rightarrow \mathbf{R}$ une fonction définie et continue sur un intervalle I , et $\varphi : [\alpha, \beta] \rightarrow I$ une fonction de classe \mathcal{C}^1 appelée *changement de variables*. Alors :

$$\int_{\varphi(\alpha)}^{\varphi(\beta)} f(u) du = \int_{\alpha}^{\beta} f(\varphi(t))\varphi'(t) dt.$$

« On pose $u = \varphi(t)$ »

⊗ Attention

Tout changement de variable doit être justifié.

Dans la pratique, on réalise assez peu souvent un changement de variable en essayant de « coller » à cette formule. On utilise plutôt les calculs formels ci-après, qui correspondent à la formule de changement de variable non-intégrée⁶ :

$$\ll f(u) du = f(\varphi(t)) d(\varphi(t)) = f(\varphi(t))\varphi'(t) dt. \gg$$

⁶Et avec des gros guillemets, car cette version sans intégrale n'a aucun sens mathématique.

Ainsi, pour réaliser le changement $u = \varphi(t)$, on commence par écrire formellement :

« $du = d(\varphi(t)) = \varphi'(t) dt$ » — on calcule donc ni plus ni moins la dérivée de φ !



Méthode Mise en place d'un changement de variable

En pratique, on écrit les calculs formels ci-après au brouillon⁷ :

$$\left\{ \begin{array}{l} u = \varphi(t) \\ du = \varphi'(t) dt \\ x = \varphi(\alpha) \iff t = \alpha \\ x = \varphi(\beta) \iff t = \beta \end{array} \right.$$

en **justifiant que φ est de classe \mathcal{C}^1** ⁸. Ce n'est qu'après que l'on pourra écrire l'égalité :

$$\int_{\varphi(\alpha)}^{\varphi(\beta)} f(u) du = \int_{\alpha}^{\beta} f(\varphi(t))\varphi'(t) dt.$$

Preuve (Point clef — *Intégrer la formule de dérivation d'une composée.*)

Notons que f et $f \circ \varphi \cdot \varphi'$ sont continues sur I et sur $[\alpha, \beta]$ respectivement, ce qui assure l'existence des intégrales. Introduisons une primitive F de f sur I (il en existe puisque f est continue). Alors $F \circ \varphi$ est dérivable de dérivée $F' \circ \varphi \cdot \varphi' = f \circ \varphi \cdot \varphi'$. Ainsi :

$$\int_{\varphi(\alpha)}^{\varphi(\beta)} f(x) dx = [F(x)]_{\varphi(\alpha)}^{\varphi(\beta)} = F(\varphi(\beta)) - F(\varphi(\alpha))$$

et

$$\int_{\alpha}^{\beta} f(\varphi(t)) \cdot \varphi'(t) dt = [F \circ \varphi(t)]_{\alpha}^{\beta} = F(\varphi(\beta)) - F(\varphi(\alpha))$$

d'où le résultat.

Remarque 1.3 — On peut lire la formule de changement de variable dans les deux sens

⁷Ils n'ont pas vocation à servir de preuve de la formule de changement de variable

⁸Le changement de variable vous sera donné dans la pratique

- ▶ Soit on réalise un changement « implicite » en l'ancienne variable u , on pose « $u = \varphi(t)$ » (de gauche à droite dans la formule), et parfois on peut avoir besoin d'explicitier t en fonction de u , ou au moins de savoir exprimer les bornes initiales comme image par φ de deux bornes α, β :

$$\int_{\alpha}^{\beta} f(u) du = \int_{\varphi^{-1}(\alpha)}^{\varphi^{-1}(\beta)} f(\varphi(t)) \cdot \varphi'(t) dt.$$

- ▶ Soit on réalise un changement « explicite » en l'ancienne variable t , on pose « $u = \varphi(t)$ » (de droite à gauche dans la formule). La formule de changement de variable est donc ici assez peu utilisée, puisque l'on sait primitiver l'intégrande $f \circ \varphi \times \varphi'$ — c'est une dérivée de composée.

Exemple 11 — Formule de changement de variable dans le « sens \rightarrow »

1. Soit $f : [-a, a] \rightarrow \mathbf{R}$ une fonction continue (où $a > 0$). Montrer que si f est paire, alors $\int_{-a}^a f(x) dx = 2 \int_0^a f(x) dx$. Que peut-on dire si f est impaire?

2. Calculer $\int_{-1}^1 \sqrt{1-x^2} dx$.

3. Soit $I = \int_0^{\frac{\pi}{2}} \cos^2 x \, dx$. Montrer que $I = \int_0^{\frac{\pi}{2}} \sin^2 x \, dx$, puis déterminer la valeur de



I.

Exemple 12 — Formule de changement de variable dans le sens « sens ← » Comme déjà précisé dans une remarque, la formule de changement de variable n'est pas indispensable dans ce cas, puisque l'on reconnaît à chaque fois des dérivées de composées usuelles. Cependant on va tout de même rédiger ces deux exemples en utilisant un changement de variable.

1. Calculer $\int_1^4 \frac{e^{1+\sqrt{t}}}{\sqrt{t}} \, dt$.

2. Calculer $\int_0^{\pi/2} \sin^3 t \cos^2 t \, dt$. Faisons le changement $u = \cos^3(t)$, $du = -3 \cos^2(t) \sin(t) \, dt$, et la fonction \cos^3 est de classe \mathcal{C}^1 . Par ailleurs, $\sin^2 t = 1 - \cos^2(t) = 1 - u^{2/3}$, donc

$$\int_0^{\pi/2} \sin^3 t \cos^2 t \, dt = -\frac{1}{3} \int_1^0 (1 - u^{2/3}) \, du = \left[u - \frac{3}{5} u^{5/3} \right]_0^1 = \boxed{\frac{2}{5}}.$$

1.5.3. Primitives usuelles


Dans les tableaux suivants, pour chaque fonction f définie sur un *intervalle* I précisé, on donne *une* primitive F . Les primitives suivantes doivent être connues par cœur, ou *a minima* être retrouvées rapidement.


$f(x) = \dots$	$F(x) = \dots$	$x \in I \subset \dots$	Condition
$(x - a)^\alpha$	$\frac{(x - a)^{\alpha+1}}{\alpha + 1}$	$]a, \infty[$	$\alpha \neq -1$
$\frac{1}{x - a}$	$\ln x - a $	$\mathbf{R} \setminus \{a\}$	$a \in \mathbf{R}$
e^{ax}	$\frac{e^{ax}}{a}$	\mathbf{R}	$a \in \mathbf{R}^*$
$\ln x $	$x \ln x - x$	\mathbf{R}^*	
$\sin(ax)$	$\frac{-\cos(ax)}{a}$		
$\cos(ax)$	$\frac{\sin(ax)}{a}$	\mathbf{R}	$a \in \mathbf{R}^*$
$\frac{1}{x^2 + 1}$	$\arctan x$	\mathbf{R}	




Méthode Essayer de se ramener à une notation sous forme de fonction puissance

Beaucoup d'expressions peuvent se mettre sous la forme « $(x - a)^\alpha$ » : la formule de primitivation de cette expression est donc centrale.

Exemple 13 — Déterminer, sur un ensemble à préciser, une primitive des fonctions suivantes. 

1. $f : x \mapsto \frac{1}{\sqrt{1-3x}}$. 

2. $g : x \mapsto x(\sqrt{1+x^2})^3$. 



Méthode Primitives de fractions rationnelles

On sait déterminer une primitive des fonctions de la forme $x \mapsto \frac{1}{ax^2 + bx + c}$ où a, b et c sont des constantes réelles et $a \neq 0$. Il suffit de discuter selon la valeur du discriminant Δ :

1. si $\Delta > 0$, alors on factorise le dénominateur pour se ramener à $x \mapsto \frac{1}{(x - \alpha)(x - \beta)}$, puis on écrit la fraction comme somme de deux autres qui se primitivent avec un logarithme.
2. Si $\Delta = 0$, alors on factorise le dénominateur pour se ramener à $x \mapsto \frac{1}{(x - \alpha)^2}$,
3. si $\Delta < 0$, alors on met le dénominateur **sous forme canonique** et on effectue un changement de variable pour se ramener à $u \mapsto \frac{1}{u^2 + 1}$.


Exemple 14 — Déterminer une primitive des fonctions suivantes :

$$x \mapsto \frac{1}{2x^2 + x - 1}, \quad x \mapsto \frac{1}{4x^2 - 4x + 1} \quad \text{et} \quad x \mapsto \frac{1}{x^2 + x + 1}.$$

Et enfin, nous rappelons une méthode usuelle concernant les fonctions trigonométriques.

 **Méthode** Calcul d'une primitive avec des fonctions trigonométriques

1. Commencer par linéariser l'expression, à l'aide de nombres complexes si besoin.
2. Primitiver avec les formules usuelles.

Exemple 15 — Déterminer, *via* deux méthodes, une primitive de $x \mapsto \cos^3 x \sin x$ sur un ensemble à préciser. 

1.6. Sommes de RIEMANN & Intégration Numérique

La motivation première de l'introduction du calcul intégral fut celle du calcul d'aires, et de volumes (pour les intégrales doubles). Pour le moment nous n'avons pas encore réalisé cette interprétation, c'est l'objectif de cette sous-section. Commençons avec un premier exemple : celui d'une fonction constante.

Si $f : [a, b] \rightarrow \mathbf{R}$ est une fonction constante égale à $k \in \mathbf{R}$, alors : $\int_a^b f = (b-a)k$.

Maintenant si f est constante par morceaux, alors avec les notations de la **Définition ANA.11.3**, l'intégrale de f vaut :

$$\sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} f|_{]x_k, x_{k+1}[} = \sum_{k=0}^{n-1} (x_{k+1} - x_k) \alpha_k$$

où α_k désigne la valeur que prend le prolongement continue de f sur $[x_k, x_{k+1}]$. Cette quantité est donc une somme d'aires de rectangles. Maintenant, dans le cas général, nous avons le théorème suivant.

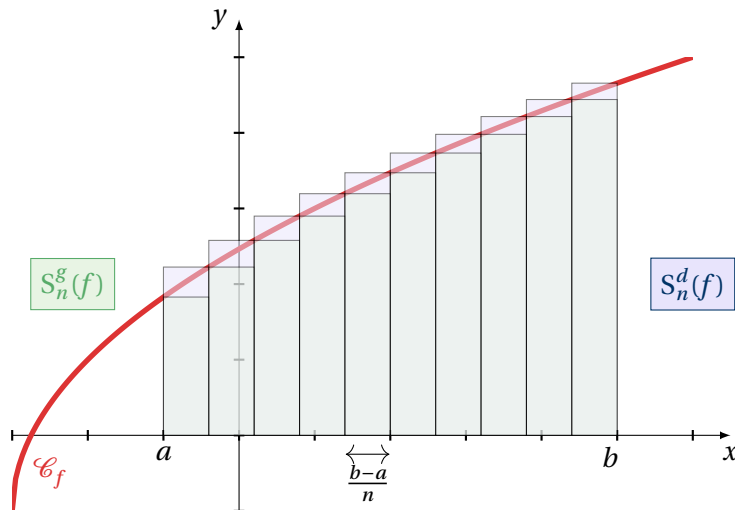
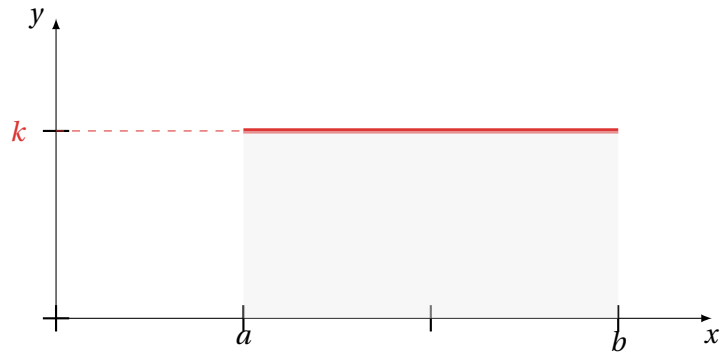


FIG. ANA.11.2. : Méthode des rectangles.

Alors :

$$S_n^g(f) \xrightarrow{n \rightarrow \infty} \int_a^b f, \quad S_n^d(f) \xrightarrow{n \rightarrow \infty} \int_a^b f.$$

2. (Estimation de l'erreur dans le cas \mathcal{C}^1) Si on suppose en outre que la fonction f est de classe \mathcal{C}^1 , alors on a la majoration de l'erreur suivante :

$$\left| \int_a^b f - S_n(f) \right| \leq \frac{(b-a)^2}{2n} \sup_{[a,b]} |f'|,$$

où $S_n(f)$ désigne l'une ou l'autre des sommes $S_n^g(f), S_n^d(f)$.

Les quantités $S_n^g(f)$ (resp. $S_n^d(f)$) correspondent à la somme des aires des rectangles verts (resp. bleus) sur le dessin précédent. Les points $a = x_0, \dots, x_n = b$ sont espacés d'un pas $\frac{b-a}{n}$ et «découpe» donc l'intervalle $[a, b]$ en sous-intervalles de même largeur.

Par ailleurs, connaissant une expression de f , il est alors très facile d'en déduire une valeur approchée de son intégrale sur $[a, b]$ à l'aide d'un outil informatique.

Preuve (Point clef — **Relation de CHASLES, majoration d'intégrales, inégalité des accroissements finis**)

La convergence dans le cas où f est de classe \mathcal{C}^0 est admise. Nous allons montrer la majoration de l'erreur dans le cas \mathcal{C}^1 et, par exemple, pour la somme des rectangles à gauche, ce qui suffira à prouver la convergence par théorème d'encadrement : en effet,

$$\lim_{n \rightarrow \infty} \frac{(b-a)^2}{2n} \sup_{[a,b]} |f'| = 0.$$

En notant $x_k = a + k \frac{b-a}{n}$ pour $k \in \llbracket 0, n \rrbracket$, et $M_1 = \sup_{[a,b]} |f'|$. Cette borne supérieure

Définition/Proposition ANA.11.2 | Convergence des sommes de RIEMANN

1. Soit $f : [a, b] \rightarrow \mathbf{R}$ une fonction continue. Pour tout $n \in \mathbf{N}^*$, on appelle *somme de RIEMANN gauche* (resp. *droite*) ou *somme des rectangles gauche* (resp. *droite*) les quantités :

$$S_n^g(f) = \frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + k \frac{b-a}{n}\right), \quad S_n^d(f) = \frac{b-a}{n} \sum_{k=1}^n f\left(a + k \frac{b-a}{n}\right).$$

existe puisque f' est continue sur le segment $[a, b]$, elle est donc bornée. On a alors :

$$\begin{aligned}
 0 &\leq \left| \int_a^b f - S_n^g(f) \right| \\
 &= \left| \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} (f(t) - f(x_k)) dt \right| \\
 &\leq \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} |f(t) - f(x_k)| dt \\
 &\leq \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} M_1 |t - x_k| dt \\
 &\leq \sum_{k=0}^{n-1} M_1 \left[\frac{(t - x_k)^2}{2} \right]_{x_k}^{x_{k+1}} \\
 &\leq M_1 \frac{(b-a)^2}{2n}.
 \end{aligned}$$

inégalités triangulaires pour les sommes/intégrales

Donc par théorème d'encadrement, nous obtenons :

$$\lim_{n \rightarrow \infty} \left| \int_a^b f - S_n^g(f) \right| = 0.$$

Preuve (Dans le cas monotone) (Point clef — **Comparaison série-intégrale**)

On propose une preuve dans un contexte simplifié, en supposant que f est monotone. Supposons par exemple que f est croissante, alors soit $k \in \llbracket 0, n-1 \rrbracket$ alors par croissance :

$$\forall x \in \left[a + k \frac{b-a}{n}, a + (k+1) \frac{b-a}{n} \right], \quad f\left(a + k \frac{b-a}{n}\right) \leq f(x) \leq f\left(a + (k+1) \frac{b-a}{n}\right).$$



Dans la fin de la preuve précédente, on a aussi finalement montré que si f est croissante; les rectangles à gauche sont «en-dessous» des rectangles à droite, ce qui se constate aisément sur un dessin.

Exemple 16 — Sommes de RIEMANN Identifier les sommes ci-dessous comme des sommes de RIEMANN, et en déduire les valeurs données des limites.

1. $\sum_{k=1}^n \frac{1}{n+k} \xrightarrow{n \rightarrow \infty} \ln 2.$

2. $\sum_{k=0}^{n-1} \frac{\sqrt{n^2 - k^2}}{n^2} \xrightarrow{n \rightarrow \infty} \frac{\pi}{4}.$

INTÉGRATION NUMÉRIQUE. Lorsque l'on ne sait pas calculer explicitement une intégrale, nous pouvons l'approcher à l'aide d'une somme de RIEMANN comme nous venons le voir. Avec les notations de **Définition/Proposition ANA.11.2**, la quantité $S_n(f)$ pour n assez grand peut donc servir d'approximation de $\int_a^b f$. On en déduit alors le script Python suivant.

APPROXIMATION D'INTÉGRALES PAR LA MÉTHODE DES RECTANGLES

Méthode des rectangles (gauche)

```
def rectangle_RG(f, a, b, n):
    """
    Calcule la somme des rectangles gauche associée à f
    """
    S = 0
    h = (b-a)/n
    for i in range(n):
        S += f(a+h*i)
    return S*h
```

Méthode des rectangles (droite)

```
def rectangle_RD(f, a, b, n):
    """
    Calcule la somme des rectangles droite associée à f
    """
    S = 0
    h = (b-a)/n
    for i in range(1,n+1):
        S += f(a+h*i)
    return S*h
```

Par exemple,

```
>>> rectangle_RG(lambda x:x**2, 0, 1, 10**3)
0.33283349999999995
>>> rectangle_RD(lambda x:x**2, 0, 1, 10**3)
0.33383349999999995
```

Ce que l'on peut retrouver par le calcul.

$$\int_0^1 t^2 dt = \frac{1}{3}.$$

Il existe beaucoup d'autres méthodes : celle du point milieu (méthode des rectangles où l'on choisit le milieu des intervalles comme hauteur), celle des trapèzes (les rectangles sont remplacés par des trapèzes), de SIMPSON (les rectangles sont des branches de paraboles). Une méthode sera d'autant meilleure qu'elle converge rapidement vers la bonne valeur théorique inconnue.

2. INTÉGRATION SUR UN INTERVALLE QUELCONQUE

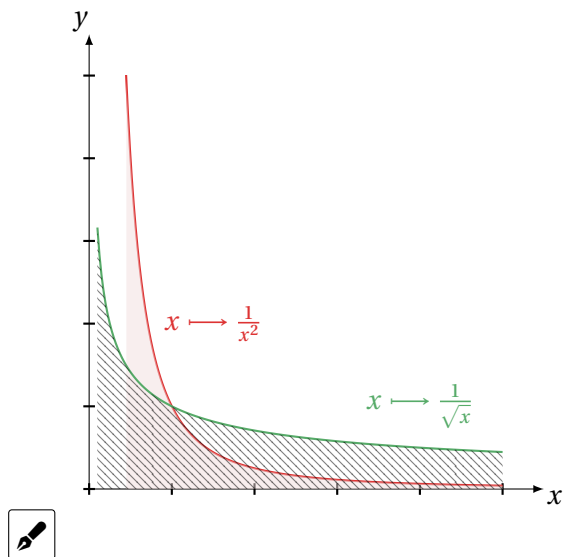
En 1ère année, et dans la section précédente, nous avons donné un sens $\int_a^b f$ avec a, b deux réels, pour une fonction continue ou continue par morceaux sur $[a, b]$. On souhaiterait étendre la notion pour intégrer des fonctions continues définies sur I un intervalle quelconque. Par exemple $I = [a, +\infty[$ ou $I = [a, b[, I =]a, b[$. Disons-le de suite, comme pour les sommes infinies de nombres réels, cela ne sera pas toujours possible. Regardons un premier exemple.

Exemple 17 — *Est-il possible de définir $\int_0^1 \frac{1}{\sqrt{t}} dt$ et $\int_0^1 \frac{1}{t^2} dt$?*

Les deux fonctions $t \mapsto \frac{1}{\sqrt{t}}$ et $t \mapsto \frac{1}{t^2}$ ne sont ni continues sur $[0, 1]$ ni prolongeables par continuité sur ce même intervalle, il n'est donc pas possible de définir leur intégrale avec la [Section 1](#).

1. Une idée est donc d'analyser l'existence éventuelle des deux limites suivantes :

$$\lim_{\varepsilon \rightarrow 0} \int_{\varepsilon}^1 \frac{1}{\sqrt{t}} dt, \quad \lim_{\varepsilon \rightarrow 0} \int_{\varepsilon}^1 \frac{1}{t^2} dt.$$



On constate que l'aire rouge semble être de plus en plus grande lorsque l'on se rapproche de l'origine, alors que la verte semble converger vers une limite finie. En langage sur les fonctions, $x \mapsto \frac{1}{x^2}$ diverge beaucoup plus vite en zéro que $x \mapsto \frac{1}{\sqrt{x}}$. Plus précisément, nous avons pour tout $\varepsilon > 0$:

Tout l'enjeu est donc, comme nous venons de le voir, de justifier l'existence des limites précédentes. Et, comme pour les séries, les difficultés se présenteront lorsque nous ne saurons pas primitiver les intégrandes.

Toute fonction du type $\varepsilon > 0 \mapsto \int_{\varepsilon}^a f$ (resp. $A > 0 \mapsto \int_a^A f$) avec f fonction continue sur $[\varepsilon, a]$ pour tout ε et a (resp. pour tout a et A) est **monotone** si f est **de signe constant**. Ainsi, d'après le théorème de la limite monotone, il nous suffira de montrer qu'elles sont majorées pour justifier l'existence de leur limite quand $\varepsilon \rightarrow 0$ (resp. $A \rightarrow \infty$).

Ainsi, nous aurons le même type de résultats dans la suite que dans le [Chapter ANA.10](#) sur les suites et séries :

- ▶ des résultats sur les fonctions positives incluant un théorème de comparaison (l'hypothèse sera vérifiée ou non grâce à des équivalents, développements limités, etc.) comme pour les séries.
- ▶ Puis nous passerons à l'extension aux fonctions à valeurs quelconques. La convergence absolue des séries sera remplacée par l'existence de l'intégrale de la valeur absolue, on parlera toujours de «convergence absolue».

Ainsi, par ce biais, nous pouvons définir $\int_0^1 \frac{1}{\sqrt{t}} dt$ même si $t \mapsto \frac{1}{\sqrt{t}}$ n'est pas continue ou prolongeable en zéro. En revanche, ce n'est pas le cas de $t \mapsto \frac{1}{t^2}$.

2. Peut-on, et comment, définir $\int_1^{\infty} \frac{1}{\sqrt{t}} dt$ et $\int_1^{\infty} \frac{1}{t^2} dt$?

2.1. Généralités

2.1.1. Intégrale impropre sur $[a, b[$



Cadre

Dans cette sous-section, les fonctions seront définies sur un intervalle $[a, b[$ avec a un réel et $b \in \mathbf{R} \cup \{+\infty\}$. Par exemple : $[0, 1[$ ou $[0, +\infty[$.

Définition ANA.11.5 | Intégrale partielle

Soit $f \in \mathcal{C}^0([a, b[, \mathbf{R})$. Alors on appelle *intégrale partielle de f en b* la fonction

$$\begin{cases} [a, b[& \longrightarrow & \mathbf{R}, \\ B & \longrightarrow & \int_a^B f. \end{cases}$$

Définition ANA.11.6 | Intégrale convergente/divergente

▶ On dit que l'intégrale $\int_a^b f$ converge si $\lim_{B \rightarrow b^-} \int_a^B f$ existe. On pose alors :

$$\int_a^b f = \lim_{B \rightarrow b^-} \int_a^B f.$$

On appelle alors *intégrande de $\int_a^b f$* la fonction f .

- ▶ La limite est alors appelée l'intégrale de f entre a et b .
- ▶ Déterminer la nature de l'intégrale c'est déterminer si elle converge ou diverge.

La notion d'intégrale partielle est analogue à la notion de somme partielle, la terminologie « $\int_a^b \dots$ converge » est analogue à « $\sum_{n=0}^\infty \dots$ converge » mais en revanche il n'y a pas de vocabulaire analogue de la notion de « série » pour les intégrales (*i.e.* la **suite** des sommes partielles).⁹

Définition/Proposition ANA.11.3 | Intégrale faussement impropre

Soit $f \in \mathcal{C}^0([a, b[, \mathbf{R})$ et supposons que f est prolongeable par continuité en b , de prolongement \tilde{f} . Alors :

$$\text{l'intégrale } \int_a^b f \text{ est convergente, et : } \int_a^b f = \int_a^b \tilde{f}$$


On dit alors que l'intégrale est *faussement impropre* en b .

Rappelons que $\int_a^b \tilde{f}$ était notre définition de l'intégrale dans la **Section 1**.

⁹Mais ce serait la fonction intégrale partielle.

Preuve En effet, notons F une primitive de \tilde{f} ainsi que $B \in [a, b[$. Alors, puisque $B < b$,

$$\begin{aligned} \int_a^B f &= \int_a^B \tilde{f}(t) dt = F(B) - F(a) \\ &\xrightarrow{B \rightarrow b^-} F(b) - F(a) \\ &= \int_a^b \tilde{f} dt. \end{aligned} \quad \left. \begin{array}{l} \text{car } F \text{ est continue} \\ \text{par définition de l'intégrale} \end{array} \right\}$$

Exemple 18 — L'intégrale $\int_{1/2}^1 \frac{\ln t}{t-1} dt$ est faussement impropre en 1. 

Définition ANA.11.7

Soit $f \in \mathcal{C}^0([a, b[, \mathbf{R})$. L'intégrale $\int_a^b f$ est dite *impropre* (ou *généralisée*) en b si :

- ▶ $\int_a^b f$ converge, **mais que**
- ▶ f n'est **pas** prolongeable par continuité en b .

L'adjectif « impropre » signifie qu'il y a un passage à la limite derrière, et donc que l'on ne peut définir l'intégrale qu'avec la seule **Section 1**.

Définition ANA.11.8 | Reste d'une intégrale impropre

Soit $f \in \mathcal{C}^0([a, b[, \mathbf{R})$ telle que $\int_a^b f$ converge, on appelle *reste en a de $\int_a^b f$* , la fonction

$$\begin{cases}]a, b[& \longrightarrow & \mathbf{R}, \\ B & \longrightarrow & \boxed{\int_a^B f} - \boxed{\int_a^B f} = \int_B^b f. \end{cases}$$

intégrale totale intégrale partielle

La définition est analogue à celle du reste ($R_n = S - S_n$) pour les séries (voir le [Chapitre ANA.10](#)).

Proposition ANA.11.3 | Le reste d'une intégrale convergente tend vers zéro (au point généralisé)

Soit $f \in \mathcal{C}^0([a, b[, \mathbf{R})$ telle que $\int_a^b f$ converge. Alors son reste tend vers zéro au point où l'intégrale est impropre, *i.e.*

$$\lim_{x \rightarrow b^-} \int_x^b f = 0.$$

Preuve



La propriété précédente est donc identique à celle déjà connue pour les séries.

Proposition ANA.11.4 | Si généralisée en b , la borne du bas n'a pas d'importance

Soit $f \in \mathcal{C}^0([a, b[, \mathbf{R})$ et $c \in [a, b[$. Alors :


$$\int_a^b f \text{ converge} \iff \int_c^b f \text{ converge.}$$


La nature d'une intégrale impropre ne dépend donc que du comportement de f au voisinage de la borne impropre. De même que si l'on enlève un nombre fini de termes dans l'étude d'une série, cela ne change pas sa nature.

Preuve



Exemple 19 — Étudier la nature des intégrales suivantes et déterminer leur valeur en cas de convergence.

1. $\int_0^{+\infty} \frac{1}{1+t^2} dt.$ 


2. $\int_0^{\frac{\pi}{2}} \tan t dt.$ 

2.1.2. Intégrale impropre sur $]a, b]$


Dans le cas d'une fonction f continue sur $]a, b]$ avec b un réel et $a \in \mathbf{R} \cup \{-\infty\}$, on définit de la même manière l'intégrale de f comme la limite éventuelle de $\int_A^b f$ lorsque A tend vers a à droite. En particulier on appellera :


1. *intégrale partielle de f en a* la fonction $\left| \begin{array}{l}]a, b] \rightarrow \mathbf{R}, \\ A \rightarrow \int_A^b f. \end{array} \right.$ L'intégrande d'une intégrale désigne alors encore une fois la fonction f .
2. Et *reste en a de $\int_a^b f$* , lorsque $\int_a^b f$ converge, la fonction

$$\left| \begin{array}{l}]a, b] \rightarrow \mathbf{R}, \\ A \rightarrow \int_a^b f - \int_A^b f = \int_a^A f. \end{array} \right.$$

Exemple 20 — L'intégrale $\int_0^{\frac{\pi}{2}} \frac{\sin(t)}{t} dt$ est faussement impropre en $\frac{\pi}{2}$. 

Exemple 21 — Étudier la nature des intégrales suivantes et déterminer leur valeur en cas de convergence.

1. $\int_0^1 t \ln(t) dt$, 

2. $\int_0^1 \ln(t) dt$, 

2.1.3. Intégrale impropre sur $]a, b[$

On traite dans ce paragraphe le cas d'une fonction définie sur $]a, b[$, avec $a \in \mathbf{R} \cup \{-\infty\}$ et $b \in \mathbf{R} \cup \{+\infty\}$.

Définition/Proposition ANA.11.4 | Intégrale doublement impropre

Soit $f \in \mathcal{C}^0(]a, b[, \mathbf{R})$ ($-\infty \leq a < b \leq +\infty$). On dit que l'intégrale $\int_a^b f$ converge si :

$$\exists c \in]a, b[, \quad \int_a^c f, \quad \text{et} \quad \int_c^b f \quad \text{convergent.}$$

On définit dans ce cas : $\int_a^b f = \int_a^c f + \int_c^b f$. De plus, la définition — du point de vue de la convergence et de la valeur de l'intégrale — ne dépend pas du choix de c .

Preuve Montrons que la valeur de l'intégrale ne dépend pas du choix de c . En effet, si l'on choisit $c' \in \mathbf{R}$ tel que $c' \geq c$ par exemple et que les intégrales convergent, alors :

$$\begin{aligned} \int_a^{c'} f + \int_{c'}^b f &= \int_a^c f + \int_c^{c'} f + \int_{c'}^b f \\ &= \int_a^c f + \int_c^b f. \end{aligned}$$

On montre de même que la convergence ne dépend pas du c considéré.

Remarque 2.1 — Puisque la définition vous autorise n'importe quel c , choisissez-en un qui vous arrange. Par exemple, pour l'étude de $\int_{-\infty}^{\infty} e^{-|t|} dt$, il est plus intéressant d'étudier

$$\int_0^{\infty} e^{-|t|} dt \quad \text{et} \quad \int_{-\infty}^0 e^{-|t|} dt,$$

que d'étudier

$$\int_{\frac{17}{3}}^{\infty} e^{-|t|} dt \quad \text{et} \quad \int_{-\infty}^{\frac{17}{3}} e^{-|t|} dt.$$

⊗ Attention Deux limites distinctes

Si $a = -\infty$ et $b = +\infty$, il ne suffit pas que $\lim_{A \rightarrow \infty} \int_{-A}^A f$ existe et soit finie pour garantir l'existence de l'intégrale, mais plutôt que :

$$\lim_{A \rightarrow \infty} \int_{-A}^c f, \quad \text{et} \quad \lim_{B \rightarrow \infty} \int_c^B f$$


existent et sont finies pour un certain $c \in \mathbf{R}$.

Un contre-exemple très simple existe : la fonction $t \mapsto t$. Pour tout $A \in \mathbf{R}^+$, $\int_{-A}^A t dt = 0$ et pourtant $\int_{-\infty}^{\infty} t dt$ diverge puisque $\int_0^{\infty} t dt$ diverge.

Exemple 22 — L'intégrale $\int_{-\infty}^{+\infty} \frac{1}{1+t^2} dt$ est convergente, et de plus :

$$\int_{-\infty}^{+\infty} \frac{1}{1+t^2} dt = \pi.$$



Exemple 23 — Existence et calcul de $\int_{-\infty}^{+\infty} e^{-|t-1|} dt$. 

Il y aura beaucoup d'analogies avec le cours sur les séries dans ce chapitre. Cependant, voici une différence notable : il n'existe pas de divergence grossière pour les intégrales.

Attention **Différence avec les séries**

Soit f une fonction continue sur l'intervalle $[a, \infty[$. Alors :

$$\int_a^\infty f \text{ converge} \not\Rightarrow \lim_{t \rightarrow \infty} f(t) = 0.$$

Il n'existe donc pas de notion de « divergence grossière » pour les intégrales.

Exemple 24 – Contre-exemple Considérons une fonction « triangulaire par morceaux » dont les triangles sont d'aires $\frac{1}{n^2}$, de hauteur 1, centrés sur le milieu du segment $[n, n + 1]$, i.e.

$$f(x) = \begin{cases} n^2 \left(x - n - \frac{1}{2}\right) + 1 & \text{si } x \in \left[n + \frac{1}{2} - \frac{1}{n^2}, n + \frac{1}{2}\right], \\ -n^2 \left(x - n - \frac{1}{2}\right) - 1 & \text{si } x \in \left[n + \frac{1}{2}, n + \frac{1}{2} + \frac{1}{n^2}\right], \\ 0 & \text{sinon.} \end{cases}$$

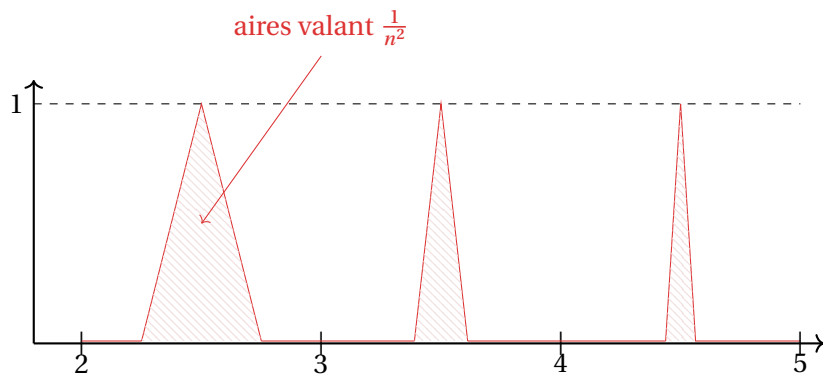


FIG. ANA.11.3. : Graphe de la fonction f

Rappelons que la série $\left(\sum_{k \geq 1} \frac{1}{k^2}\right)$ est convergente. Soit $x \geq 2$, alors grâce à la relation de CHASLES, nous avons :

$$\forall x \geq 2, \int_2^x f \leq \sum_{k=2}^{\lfloor x \rfloor + 1} \frac{1}{k^2} \leq \sum_{k=2}^{+\infty} \frac{1}{k^2}.$$

Donc les intégrales partielles sont majorées, la fonction f est positive, donc f est d'intégrale convergente. Et pourtant, il est clair qu'elle ne converge pas vers zéro.

2.1.4. Intégrale impropre en un nombre fini de points

Définition ANA.11.9 | Intégrale d'une fonction continue sauf en un nombre fini de points

Soit $f :]a, b[\rightarrow \mathbf{R}$ ($-\infty \leq a < b \leq +\infty$). On suppose qu'il existe une subdivision $a = x_0 < x_1 < x_2 < \dots < x_{n-1} < x_n = b$, et telle que, pour tout $k \in \llbracket 0, n-1 \rrbracket$, f est définie et continue sur chaque intervalle $]x_k, x_{k+1}[$. On dit que f est continue sur $]a, b[$ sauf en un nombre fini de points.

- ▶ On dit que l'intégrale $\int_a^b f$ converge si, pour tout $k \in \llbracket 0, n-1 \rrbracket$, $\int_{x_k}^{x_{k+1}} f$ converge.
- ▶ En cas de convergence, on pose

$$\int_a^b f = \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} f.$$

Remarque 2.2 –

- ▶ L'hypothèse ne signifie pas que f est continue par morceaux : on ne suppose pas que f est prolongeable par continuité aux points x_k (mais certaines peuvent l'être), les intégrales intervenant dans la somme ne sont donc pas « faussement impropres ».
- ▶ Le résultat ne dépend pas de la subdivision choisie.
- ▶ Il faut vérifier la convergence de toutes les intégrales aux bornes où elle est généralisée.

Exemple 25 – Que faut-il étudier comme intégrales pour étudier la convergence de $\int_0^{+\infty} \frac{\ln(t)}{t^2 - 1} dt$? Nous réaliserons l'étude complète plus tard.

VALEUR MOYENNE GÉNÉRALISÉE. On généralise ici sans peine la notion de valeur moyenne sur un intervalle quelconque, en cas de convergence.

Définition ANA.11.10 | Valeur moyenne & Valeur moyenne pondérée


Soit f une fonction continue sauf éventuellement en un nombre fini de points sur un intervalle I de la forme $]a, b]$, $[a, b[$ ou $]a, b[$ avec $(-\infty \leq a < b \leq +\infty)$, ainsi que $p : [a, b] \rightarrow \mathbf{R}$ une fonction continue sauf éventuellement en un nombre fini de points telle que :

$$\int_a^b p = 1.$$

En cas de convergence de l'intégrale, on appelle *valeur moyenne de f sur I pondérée par p* le réel :

$$\int_a^b p \cdot f.$$

Cette définition sera la base de la définition de l'espérance pour les variables aléatoires à densité dans le [Chapter ALEA.14](#).

Exemple 26 — Déterminer la valeur moyenne de $t \mapsto t$ sur \mathbf{R}^+ pondérée par $p : t \in \mathbf{R}^+ \mapsto \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$. (On admettra que p est d'intégrale 1) 

2.1.5. Intégrales de référence

Tous les résultats, sauf le tout dernier sur l'intégrale de GAUSS, qui suivent doivent être vus comme des exercices, aucun résultat sur les intégrales usuelles n'est clairement au programme de BCPST.

Proposition ANA.11.5 | Intégrande exponentielle décroissant vers zéro [H.P]

Soit $a \in \mathbf{R}$. Alors $\int_0^{+\infty} e^{-at} dt$ converge $\iff a > 0$.

Preuve



Le résultat ci-dessous précise la convergence des intégrales associées à des fonctions du type $t \mapsto \frac{1}{t^\alpha}$ avec $\alpha \in \mathbf{R}$. Rappelons-nous que nous avons établi en introduction les natures suivantes :

- ▶ $\int_0^1 \frac{1}{t^2} dt$ diverge
- ▶ $\int_1^\infty \frac{1}{t^2} dt$ converge,
- ▶ $\int_0^1 \frac{1}{\sqrt{t}} dt$ converge,
- ▶ $\int_0^\infty \frac{1}{t^2} dt$ diverge.

Il semble donc se dessiner deux critères suivant que l'on se place autour de zéro, ou au voisinage de $+\infty$. Précisons cela.

Proposition ANA.11.6 | Intégrales de RIEMANN [H.P.]

Soit $\alpha \in \mathbf{R}$.

1. (Caractère impropre en $+\infty$) $\int_1^{+\infty} \frac{1}{t^\alpha} dt$ converge $\iff \alpha > 1$.
2. (Caractère impropre en zéro) $\int_0^1 \frac{1}{t^\alpha} dt$ converge $\iff \alpha < 1$.



Attention

Vous devez savoir refaire la preuve de cette proposition pour la valeur de α considérée.

Attention

Le critère au voisinage de $+\infty$ est le même que pour les séries. En revanche, en zéro, la convergence a lieu pour les valeurs strictement inférieures à 1.

Preuve

1.  Soit $A > 1$. Calculons : $\int_1^A \frac{1}{t^\alpha} dt = \int_1^A t^{-\alpha} dt = \begin{cases} \left[\frac{t^{1-\alpha}}{1-\alpha} \right]_1^A = \frac{1}{1-\alpha} (A^{1-\alpha} - 1) & \text{si } \alpha \neq 1, \\ [\ln|t|]_1^A = \ln A & \text{si } \alpha = 1 \end{cases}$. On constate alors que, lorsque $A \rightarrow \infty$, $\int_1^A \frac{1}{t^\alpha} dt$ converge vers une limite finie si et seulement si $1 - \alpha < 0$ i.e. $\alpha < 1$.
2.  Soit $\varepsilon > 0$. Calculons : $\int_\varepsilon^1 \frac{1}{t^\alpha} dt = \int_\varepsilon^1 t^{-\alpha} dt = \begin{cases} \left[\frac{t^{1-\alpha}}{1-\alpha} \right]_\varepsilon^1 = \frac{1}{1-\alpha} (1 - \varepsilon^{1-\alpha}) & \text{si } \alpha \neq 1, \\ [\ln|t|]_\varepsilon^1 = -\ln \varepsilon & \text{si } \alpha = 1 \end{cases}$. On constate alors que, lorsque $\varepsilon \rightarrow 0+$, $\int_\varepsilon^1 \frac{1}{t^\alpha} dt$ converge vers une limite finie si et seulement si $1 - \alpha > 0$ i.e. $\alpha < 1$.

Le résultat suivant est admis, cette intégrale interviendra dans un chapitre ultérieur.

Théorème ANA.11.8 | Intégrale de GAUSS

L'intégrale $\int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt$ est convergente, et $\int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt = \sqrt{2\pi}$.

Preuve La valeur est admise, la convergence est provisoirement admise.

2.2. Propriétés des intégrales convergentes

Proposition ANA.11.7 | Propriétés de l'intégrale.

Soient $a, b \in \mathbf{R}$ ($-\infty \leq a < b \leq +\infty$) tels que $a \neq b$. Alors :

1. (**Linéarité**) L'ensemble des fonctions f des fonctions définies et continues (ou éventuellement sauf en un nombre fini de points) sur $]a, b[$ ($[a, b[$ ou $]a, b]$) telles que $\int_a^b f$ converge est un \mathbf{R} -espace vectoriel sur \mathbf{R} , et l'application $f \mapsto \int_a^b f$ est une forme linéaire sur cet espace, i.e. si $\int_a^b f$ et $\int_a^b g$ convergent avec f, g continues (ou éventuellement sauf en un nombre fini de points) sur $]a, b[$ et $\lambda, \mu \in \mathbf{R}$, alors $\int_a^b (\lambda f + \mu g)$ converge aussi et

$$\int_a^b (\lambda f + \mu g) = \lambda \int_a^b f + \mu \int_a^b g.$$

2. Avec les mêmes notations que précédemment,
 - ▶ si $\int_a^b f$ converge et $\int_a^b g$ diverge, alors $\int_a^b (f + g)$ diverge.
 - ▶ Si $\int_a^b f$ et $\int_a^b g$ divergent, **on ne peut pas conclure** pour la nature de $\int_a^b (f + g)$.
3. (**Positivité**) Soit f définie et continue sur $[a, b[$ ($[a, b[$ ou $]a, b]$) ou éventuellement sauf en un nombre fini de points. Si $\int_a^b f$ converge, et si f est positive, alors on a $\int_a^b f \geq 0$.
4. (**Croissance**) Avec les hypothèses précédentes, si, pour tout $t \in]a, b[$ ($[a, b[$ ou $]a, b]$), $f(t) \leq g(t)$, et les intégrales $\int_a^b f$ et $\int_a^b g$ convergent, alors

$$\int_a^b f \leq \int_a^b g.$$

5. (**Relation de CHASLES**) Soit $c \in]a, b[$. Si $\int_a^b f$ converge, il en est de même

de $\int_a^c f$ et $\int_c^b f$, et on a :

$$\int_a^b f = \int_a^c f + \int_c^b f.$$

Preuve Nous admettons toutes ces propriétés qui découlent de simples passages à la limite dans celles que nous connaissons déjà pour l'intégrale sur un segment.

⊗ Attention

On veillera donc, avant d'invoquer la linéarité de l'intégrale pour des intégrales généralisées, à *justifier la convergence de toutes les intégrales* apparaissant dans le calcul.

Exemple 27 — d'égalité illicite L'intégrale $\int_1^{+\infty} \frac{1}{t(t+1)} dt$ est convergente mais on **ne peut pas écrire** :

$$\int_1^{+\infty} \frac{1}{t(t+1)} dt = \int_1^{+\infty} \left(\frac{1}{t} - \frac{1}{t+1} \right) dt = \underbrace{\int_1^{+\infty} \frac{dt}{t} - \int_1^{+\infty} \frac{dt}{t+1}}_{\text{interdit}}$$

car les deux intégrales du membre de droite sont divergentes.

Enfin, nous pouvons généraliser sans peine la relation fondamentale de l'analyse dans le cas où l'une des bornes est infinie. Cela nous servira notamment dans le **Chapitre ALEA.14**.

Proposition ANA.11.8 | Intégrale généralisée à une borne variable

Soit $f \in \mathcal{C}^0(\mathbf{R}, \mathbf{R})$ telle que $\int_{-\infty}^{\infty} f$ converge. Alors :

▶ la fonction

$$g : x \in \mathbf{R} \mapsto \int_{-\infty}^x f \text{ est bien définie.}$$

▶ g est dérivable sur \mathbf{R} et sa dérivée est donnée par :

$$\forall x \in \mathbf{I}, \quad g'(x) = \frac{d}{dx} \left(\int_{-\infty}^x f \right) = f(x).$$

Preuve



2.3. Calculs d'intégrales

Nous cette sous-section, nous voyons comment intégrer par parties et réaliser des changements de variable dans les intégrales généralisées.

2.3.1. Technique d'intégration par parties

En pratique, les intégrations par parties se feront *toujours* en revenant à des intégrales sur un segment, et en passant à la limite dans les bornes ensuite (ce qui montrera alors la convergences des intégrales généralisées qui apparaissent dans le calcul). La formule déjà connue reste donc vraie.



Méthode intégration par parties pour les intégrales généralisées

1. Revenir à une intégrale partielle.
2. Utiliser la formule déjà connue sur le segment.
3. Chercher à passer à la limite.

⊗ Attention

Toute intégration par parties doit être justifiée!

Exemple 28 — Calculer $\int_0^{+\infty} te^{-t} dt$.

2.3.2. Changement de variables

Pour simplifier, on énonce la formule de changement de variable dans le cas d'intégrales généralisées en deux points. Puisqu'une intégrale sur un intervalle du type $]a, b[$ ou $[a, b[$ peut être vue comme généralisée sur $]a, b[$, la formule est donc vraie pour tout type d'intégrale généralisée.

Théorème ANA.11.9 | Changement de variables

Soit $(\alpha, \beta) \in \mathbf{R} \cup \{\pm\infty\}$ tel que $\alpha < \beta$, et φ une fonction, appelée *changement de variables*, de classe \mathcal{C}^1 strictement monotone sur $]a, b[$ de limites :

$$a = \lim_{\alpha^+} \varphi, \quad b = \lim_{\beta^+} \varphi.$$

Soit de plus $f \in \mathcal{C}^0(]a, b[, \mathbf{R})$ (si $a < b$) ou $f \in \mathcal{C}^0(]b, a[, \mathbf{R})$ (si $a > b$). Alors les intégrales

$$\int_{\alpha}^{\beta} f(\varphi(t)) \cdot \varphi'(t) dt \quad \text{et} \quad \int_a^b f(u) du \quad \text{sont de même nature,}$$

et en cas de convergence :

$$\int_a^b f(u) du = \int_{\alpha}^{\beta} f(\varphi(t)) \cdot \varphi'(t) dt. \quad \text{«On pose } u = \varphi(t)\text{»}$$

Attention

- ▶ Tout changement de variable doit être justifié!
- ▶ Un changement de variable peut transformer une intégrale simple en intégrale généralisée et vice-versa. Dans ce cas, à condition de le remarquer, la convergence des deux intégrales est immédiate.

Preuve Écrivons tout d'abord la formule de changement de variable déjà connue sur les segments. Soient alors $x \in]\alpha, \beta[$ et $y \in [x, \beta[$, alors puisque φ est de classe \mathcal{C}^1 :

$$\int_{\varphi(x)}^{\varphi(y)} f(u) du = \int_x^y f(\varphi(t)) \cdot \varphi'(t) dt. \quad (\text{ChgtVarSeg})$$

Constatons également que par hypothèse φ réalise une bijection de $]a, b[$ vers $]a, b[$.

⇐ Supposons que $\int_a^b f(u) du$ converge, i.e. pour tout $c \in]a, b[$, les intégrales

$$\int_a^c f(u) du \quad \text{et} \quad \int_c^b f(u) du \quad \text{convergent.}$$

Ainsi, par hypothèse la fonction $y \mapsto \int_{\varphi(x)}^{\varphi(y)} f(u) du$ admet une limite lorsque $y \rightarrow \beta^-$, donc c'est le cas aussi de $y \mapsto \int_x^y f(\varphi(t)) \cdot \varphi'(t) dt$ d'après (ChgtVarSeg). Ceci étant vrai pour tout x , et comme φ est bijective donc tout réel $c \in]\alpha, \beta[$ s'écrit sous la forme $\varphi(x)$ pour un certain x , on obtient que :

$$\forall c \in]\alpha, \beta[, \quad \int_c^{\beta} f(\varphi(t)) \cdot \varphi'(t) dt \quad \text{converge.}$$

En faisant de même pour $x \mapsto \int_{\varphi(x)}^{\varphi(y)} f(u) du$ lorsque $x \rightarrow \alpha^+$, on obtient que

$$\forall c \in]\alpha, \beta[, \quad \int_{\alpha}^c f(\varphi(t)) \cdot \varphi'(t) dt \quad \text{converge.}$$

⇒ On montre exactement de la même façon, i.e. en exploitant (ChgtVarSeg) que si $\int_a^b f(u) du$ converge, alors $\int_{\alpha}^{\beta} f(\varphi(t)) \varphi'(t) dt$ converge aussi. Enfin, reprenons (ChgtVarSeg) : pour $x \in]\alpha, \beta[$ et $y \in [x, \beta[$, nous avons

$$\int_{\varphi(x)}^{\varphi(y)} f(u) du = \int_x^y f(\varphi(t)) \cdot \varphi'(t) dt.$$

Donc en faisant $x \rightarrow \alpha$, on déduit

$$\int_a^{\varphi(y)} f(u) du = \int_a^y f(\varphi(t)) \cdot \varphi'(t) dt,$$

puis le résultat en faisant $y \rightarrow \beta$.

On notera que, étant donné que φ est une bijection de $]a, b[$ sur $]a, b[$, nous avons $\lim_{t \rightarrow \alpha} \varphi(t) = a$ et $\lim_{t \rightarrow \beta} \varphi(t) = b$ dans le cas où φ est strictement croissante (et l'inverse sinon). C'est ce point qui est utilisé de manière centrale dans la preuve et qui n'a pas

lieu d'être sur un segment car on ne passe pas à la limite.

Attention

L'hypothèse de monotonie est cruciale. En effet, si on réalise le changement de variable $t = u^2$ dans


$$I_1 = \int_{-1}^1 \frac{1}{u} du,$$


qui diverge (car par exemple $\int_{-1}^0 \frac{1}{u} du$ diverge), alors on obtiendrait que I_1 a même nature que

$$I_2 = \int_1^1 \frac{1}{2t} dt = 0$$

qui elle, converge.

Exemple 29 — Déterminer la nature des intégrales suivantes à l'aide du changement de variable indiqué. Les calculer en cas de convergence.

1. $I = \int_0^{+\infty} \frac{1}{(1+u^2)^2} du$. *Indication* : On posera $u = \tan(t)$. 

2. $I = \int_3^{+\infty} \frac{1}{t(\ln(t))} dt$. *Indication* : On posera $t = e^v$. 

Corollaire ANA.11.1 | Conséquence à l'aide de la parité


1. Soit f une fonction continue sur un intervalle $] -a, a[$ avec $a \in]0, +\infty[$. Si f est paire alors $\int_{-a}^a f$, $\int_0^a f$ sont de même nature et, en cas de convergence :


$$\int_{-a}^a f = 2 \int_0^a f.$$

2. Soit f une fonction continue sur un intervalle $] -a, a[$ avec $a \in]0, +\infty[$. Si f est impaire alors $\int_{-a}^a f$, $\int_0^a f$ sont de même nature et, en cas de convergence :

$$\int_{-a}^a f = 0.$$

Preuve

Nous faisons la preuve dans le cas pair, l'autre est identique. 

Exemple 30 — On a $\int_0^{+\infty} e^{-\frac{t^2}{2}} dt = \sqrt{\frac{\pi}{2}}$ et $\int_0^{+\infty} e^{-t^2} dt = \frac{\sqrt{\pi}}{2}$.  Puisque


la fonction $t \mapsto e^{-\frac{t^2}{2}}$ est une fonction paire, et que $\int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt$ converge,¹⁰ alors $\int_0^{+\infty} e^{-\frac{t^2}{2}} dt$ converge également. De plus,

$$\int_0^{+\infty} e^{-\frac{t^2}{2}} dt = \frac{1}{2} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt = \frac{1}{2} \sqrt{2\pi} = \boxed{\sqrt{\frac{\pi}{2}}}.$$

Faisons le changement de variable « $u = \frac{t}{\sqrt{2}}$ » dans $\int_0^{+\infty} e^{-\frac{t^2}{2}} dt$, ce changement de variable est possible puisque la fonction $t \mapsto \frac{t}{\sqrt{2}}$ est \mathcal{C}^1 strictement croissante. Ainsi,

$$\int_0^{+\infty} e^{-\frac{t^2}{2}} dt = \sqrt{2} \int_0^{+\infty} e^{-u^2} du \implies \boxed{\int_0^{+\infty} e^{-u^2} du = \frac{\sqrt{\pi}}{2}}.$$

2.4. Intégrales de fonctions de signe constant

 **Cadre**
 Nous nous intéressons dans cette section aux intégrales de fonctions positives, et continues sauf éventuellement aux bornes.¹¹ Les résultats analogues s'appliquent aux fonctions négatives en considérant $\int(-f)$.

Comme pour les séries, l'étude spéciale des intégrandes positives est motivée par le fait suivant; l'intégrale partielle d'une intégrale de fonction positive est monotone. Soit donc f une fonction continue et positive sur $[a, b[$, alors pour tous $x \leq y$ tels que $x, y \in [a, b[$:

$$\int_a^y f - \int_a^x f = \int_x^y f \geq 0 \implies \int_a^y f \geq \int_a^x f.$$

¹⁰Résultat toujours admis pour le moment

¹¹Pour les fonctions continues sauf en un nombre fini de points, il suffit d'appliquer ces résultats à plusieurs intégrales

La fonction $x \in [a, b[\mapsto \int_a^x f$ est donc croissante, et sa convergence (la convergence de l'intégrale donc) se réduit à son éventuel caractère majoré (d'après le théorème de convergence monotone pour les fonctions). Le même constat peut être fait pour une intégrale généralisée en b par exemple, ou même généralisée en les deux bornes.

En résumé :

les intégrales partielles d'une intégrale de signe constant sont monotones!

On déduit alors immédiatement le théorème suivant.

Proposition ANA.11.9 | Convergence des intégrales de fonctions positives

Soit f une fonction continue et positive sur $[a, b[$ (resp. $]a, b]$). Alors :

$$\int_a^b f \text{ converge} \iff \begin{cases} \text{la fonction } x \mapsto \int_a^x f \text{ est majorée sur } [a, b[\\ \text{resp. } x \mapsto \int_x^b f \text{ sur }]a, b]. \end{cases}$$

COMPARAISON D'INTÉGRALES DE FONCTIONS POSITIVES.

Théorème ANA.11.10 | Comparaison des intégrales de fonctions positives

Soient $f, g \in \mathcal{C}^0([a, b])$ ou $]a, b[$ ou $]a, b]$, **positives**, telles que pour tout $t \in [a, b[$ (ou $]a, b]$) ou $]a, b[$,

$$0 \leq f(t) \leq g(t).$$

- ▶ Si l'intégrale $\int_a^b g$ converge, alors l'intégrale $\int_a^b f$ converge.
- ▶ Si l'intégrale $\int_a^b f$ diverge, alors l'intégrale $\int_a^b g$ diverge.

Preuve



Commençons par traiter le cas où $f, g \in \mathcal{C}^0([a, b])$.

- ▶ On contrapose simplement la première partie. Le cas $f, g \in \mathcal{C}^0([a, b])$ se traite de la même manière. Enfin, supposons que $f, g \in \mathcal{C}^0(]a, b[)$.
- ▶ Si $\int_a^b g$ converge, alors pour tout $c \in]a, b[$, $\int_a^c g, \int_c^b g$ convergent. Donc en appliquant les deux cas précédents, on déduit que pour tout $c \in]a, b[$, $\int_a^c f, \int_c^b f$ convergent aussi. Ainsi, $\int_a^b f$ converge.
- ▶ On contrapose simplement la première partie.

L'utilisation du théorème de comparaison se fera en général au travers de deux moyens :

- ▶ soit on vous a fait trouver un encadrement (ou alors elle est évidente comme dans les exemples précédents),
- ▶ soit on utilise la formulation avec des équivalents ci-après. Elle est officiellement [H.P], comme pour les séries, mais les sujets la nécessitant l'admettent en préambule d'exercice.

Corollaire ANA.11.2 | Comparaison des intégrales de fonctions positives avec \sim [H.P]

Soient $f, g \in \mathcal{C}^0(]a, b[)$ (resp. $]a, b[$) **positives**, et telles que

$$f(t) \underset{t \rightarrow b}{\sim} g(t), \quad (\text{resp. } f(t) \underset{t \rightarrow a}{\sim} g(t)).$$

Alors : $\int_a^b f$ et $\int_a^b g$ sont de même nature.

On utilisera donc ce résultat **uniquement si l'énoncé vous le permet**.

Attention Faux si les fonctions ne sont pas positives

Pour les fonctions qui ne sont pas de signe constant, les théorèmes de comparaison ne s'appliquent pas : considérer sur $[1, +\infty[$ les fonctions f, g définies par :

$$\forall t \in [1, +\infty[, \quad f(t) = \frac{\sin(t)}{\sqrt{t}}, \quad g(t) = \frac{\sin(t)}{\sqrt{t}} + \frac{|\sin(t)|}{t}.$$

Elles sont équivalentes en $+\infty$ mais leurs intégrales ne sont pas de même nature.

Preuve (Point clef — Traduire l'équivalent à l'aide d'une inégalité)

Faisons par exemple le cas où f, g sont définies sur $[a, b]$, supposons que $f(t) \underset{t \rightarrow b}{\sim} g(t)$. Cela signifie que $g(t)$ est non nulle pour t assez proche de b et que $\frac{f(t)}{g(t)} \underset{t \rightarrow b}{\rightarrow} 1$, i.e.

$$\exists \eta > 0, \quad \forall t \in [b - \eta, b + \eta], \quad \left| \frac{f(t)}{g(t)} - 1 \right| < \frac{1}{2}.$$

En choisissant $\varepsilon = \frac{1}{2}$, on constate que pour t assez proche de b , $\frac{f(t)}{g(t)}$ est entre $\frac{1}{2}$ et $\frac{3}{2}$:

$$\exists \eta > 0, \quad \forall t \in [b - \eta, b + \eta], \quad -\frac{1}{2} < \frac{f(t)}{g(t)} - 1 < \frac{1}{2} \iff \frac{1}{2} < \frac{f(t)}{g(t)} < \frac{3}{2}.$$

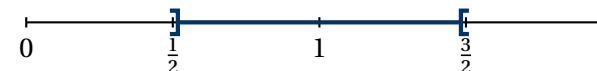


FIG. ANA.11.4. : Le quotient $\frac{f(t)}{g(t)}$ est dans cet intervalle, au moins pour t assez proche de t_0

Donc pour t assez proche de b , on a finalement l'encadrement ci-après :


$$0 \leq \frac{g(t)}{2} \leq f(t) \leq \frac{3g(t)}{2}. \quad (\star)$$


On peut à présent appliquer le théorème de comparaison.

- ▶ Supposons que $\int_a^b f$ converge, alors puisque $0 \leq \frac{g(t)}{2} \leq f(t)$ pour t assez proche de b — partie gauche de (\star) — le théorème de comparaison livre que $\int_a^b \frac{g}{2}$ converge et donc la convergence de $\int_a^b g$.

- Supposons que $\int_a^b f$ diverge, alors puisque $0 \leq f(t) \leq \frac{3g(t)}{2}$ pour t assez proche de b — partie droite de (\star) — le théorème de comparaison livre que $\int_a^b \frac{3g}{2}$ diverge et donc $\int_a^b g$ diverge.
- On a donc montré que les deux intégrales ont la même nature.

Exemple 31 — Déterminer la nature des intégrales ci-après, en appliquant le théorème de comparaison si une inégalité est évidente, ou le critère sur les équivalents sinon.

1. L'intégrale $\int_1^{\infty} \frac{\ln t}{t} dt$ diverge. 

2. L'intégrale $\int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt$ est convergente. 

3. Retour sur l'Exemple 25. Existence de $\int_0^{+\infty} \frac{\ln(t)}{t^2 - 1} dt$? 


^aDéfinition de la limite avec « $\varepsilon = \frac{1}{2}$ »



Méthode Convergence d'intégrale à intégrande exponentielle décroissante

Soit $f : \mathbf{R} \rightarrow \mathbf{R}$ une fonction telle que $f(t) \xrightarrow{t \rightarrow \infty} \infty$. Alors :

1. $t^2 e^{-f(t)} \xrightarrow{t \rightarrow \infty} 0$ par croissances comparées,
2. donc pour t assez grand, $t^2 e^{-f(t)} \leq 1 \implies 0 \leq e^{-f(t)} \leq \frac{1}{t^2}$.

 3. Comme $\int_1^{\infty} \frac{1}{t^2} dt$ converge, le théorème de comparaison donne la convergence de $\int_1^{\infty} e^{-f(t)} dt$.¹²

2.5. Fonctions de signe quelconque & Convergence absolue

Nous considérons de nouveau dans cette section des fonctions non forcément positives. Nous allons regarder une notion de convergence plus forte que la convergence des intégrales partielles : il s'agit de la *convergence absolue*.

Définition ANA.11.11 | Intégrale absolument convergente

Soit f une fonction continue définie sur $]a, b[$, $[a, b[$ ou $]a, b]$. On dit que l'intégrale $\int_a^b f$ est *absolument convergente* (ou que f est *intégrable*) si $\int_a^b |f(t)| dt$ est convergente.

Remarque 2.3 — L'intégrale $\int_a^b |f(t)| dt$ est d'intégrande positive, tous les critères vus précédemment s'appliquent donc.

Définition ANA.11.12 | Intégrale semi-convergente

Soit f une fonction continue définie sur $]a, b[$, $[a, b[$ ou $]a, b]$. On dit que l'intégrale $\int_a^b f$ est *semi-convergente* si $\int_a^b f$ converge mais pas $\int_a^b |f(t)| dt$.

Théorème ANA.11.11 | La convergence absolue implique la convergence

Soit f une fonction continue définie sur $]a, b[$, $[a, b[$ ou $]a, b]$.

$$\int_a^b f \text{ est absolument convergente} \implies \int_a^b f \text{ est convergente.}$$

Autrement dit, toute intégrale absolument convergente est convergente.

¹²Attention à bien démarrer l'intégrale à une borne strictement différente de zéro, 1 par exemple.


Preuve Nous avons :


$$-|f| \leq f \leq |f| \implies 0 \leq f + |f| \leq 2|f|.$$

Ainsi l'intégrale $\int_a^b (f + |f|)$ est d'intégrande positive, tout comme $\int_a^b |f|$ qui converge par hypothèse, donc en appliquant le théorème de comparaison, nous déduisons que :

$$\int_a^b (f + |f|) \text{ converge.}$$

Mais comme $f = (f + |f|) - |f|$, l'intégrale $\int_a^b f$ s'exprime alors comme la différence de deux intégrales convergentes, elle est donc elle aussi convergente.

Exemple 32 — L'intégrale $\int_0^1 \sin\left(\frac{1}{t}\right) dt$ est absolument convergente. 

Exemple 33 — Montrer, à l'aide d'une intégration par parties, que l'intégrale $\int_0^{+\infty} \frac{\sin(t)}{t} dt$ est convergente. 

d'intégrales absolument convergentes est un espace vectoriel.

Exemple 34 — Déterminer la nature et la valeur en cas de convergence de

$$\int_{-\infty}^{+\infty} \frac{\sin(t^4)}{t^3} dt.$$

Preuve



Proposition ANA.11.11 | Inégalité triangulaire généralisée

Soit f une fonction continue définie sur $]a, b[$, $[a, b[$ ou $]a, b]$. Si $\int_a^b f$ est absolument convergente, alors :

$$\left| \int_a^b f \right| \leq \int_a^b |f(t)| dt.$$

Preuve



Proposition ANA.11.10 | Structure d'espace vectoriel

Soient f et g deux fonctions définies sur $]a, b[$, $[a, b[$ ou $]a, b]$ d'intégrales $\int_a^b f$ et

$\int_a^b g$ absolument convergentes et $(\lambda, \mu) \in \mathbf{R}^2$. Alors :

- ▶ $\int_a^b (\lambda f + \mu g)$ est absolument convergente.
- ▶ En particulier, l'ensemble des fonctions définies sur $]a, b[$, $[a, b[$ ou $]a, b]$

2.6. Plan d'étude d'une intégrale



Méthode Plan d'étude d'une intégrale

Soit f continue ou continue sauf en un nombre fini de points sur $]a, b[$ ou $[a, b[$ ou $]a, b]$ ($-\infty \leq a < b \leq +\infty$). Il s'agit de se poser les questions dans l'ordre suivant afin d'analyser l'existence de l'intégrale.

1. Suis-je capable de calculer l'intégrale $\int_a^A f$ (ou $\int_A^a f$ en fonction du cas), ou les deux dans le cas de $]a, b[$) explicitement pour A dans l'intervalle d'intégration ?
Si oui, on la calcule et on analyse l'existence d'une limite en A .
2. Sinon, et ce sera l'immense majorité des cas, on se demande si :
 - ▶ elle est positive, dans ce cas on essaie de la majorer ou minorer par une fonction simple dont on connaît la nature de l'intégrale. On utilise éventuellement des développements limités et relations de comparaisons pour cela.
 - ▶ elle n'est pas positive, on étudie la convergence absolue.

*** Fin du chapitre ***

3. EXERCICES

Exercice ANA.11.1 | Vrai ou Faux?

1. Une primitive de $x \mapsto \cos(-x + 1)$ est $x \mapsto \sin(x - 1)$.
2. Soit $n \in \mathbb{N}$. Une primitive de $x \mapsto \frac{x}{(a^2+x^2)^n}$ est $x \mapsto \frac{1}{2(n-1)} \frac{1}{(a^2+x^2)^{n-1}}$.
3. $\int_1^2 x \, dx = \frac{5}{2}$,
4. Si f est continue sur $[0, 1]$, alors $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n f\left(\frac{k}{n}\right) = \int_0^1 f(t) \, dt$.
5. L'intégrale $\int_0^\infty \frac{1}{x^{5/2}} \, dx$ converge.

3.1. Intégrales sur un segment

Exercice ANA.11.2 | Avec décomposition d'une fraction rationnelle On considère la fonction f définie sur $] -3, 2[$ par $f(x) = \frac{3x^2+4x-25}{x^2+x-6}$ pour tout $x \in] -3, 2[$.

1. Montrer qu'il existe trois réels a, b, c que l'on déterminera, tels que

$$\forall x \in] -3, 2[, \quad f(x) = a + \frac{b}{x+3} + \frac{c}{x-2}.$$

2. En déduire l'existence et la valeur de $\int_0^1 f(x) \, dx$.

Solution (exercice ANA.11.2)

1. On cherche a, b, c de sorte que :

$$\forall x \in] -3, 2[, \quad f(x) = a + \frac{b}{x+3} + \frac{c}{x-2} = \frac{a(x+3)(x-2) + b(x-2) + c(x+3)}{(x+3)(x-2)}.$$

En développant puis en factorisant numérateur suivant les degrés, on déduit la condition :

$$\forall x \in] -3, 2[, \quad a \cdot x^2 + x(-2a + 3a + b + c) + (-6a - 2b + 3c) = 3x^2 + 4x - 25.$$

Soit le système de conditions

$$a = 3, \quad b + c = 1, \quad -2b + 3c = -7.$$

Après résolution : $5c = -5, c = -1$ puis $b = 2$. Ainsi,

$$\forall x \in] -3, 2[, \quad f(x) = 3 + \frac{2}{x+3} - \frac{1}{x-2}.$$

2. La fonction f est continue sur $[0, 1]$, en tant que quotient de telles fonctions dont le dénominateur ne s'annule pas, donc l'intégrale existe bien. Et par linéarité de l'intégrale :

$$\begin{aligned} \int_0^1 f &= \int_0^1 3 \, dx + 2 \int_0^1 \frac{dx}{x+3} - \int_0^1 \frac{dx}{x-2} \\ &= 3 + 2 [\ln|x+3|]_0^1 - [\ln|x-2|]_0^1 \\ &= 3 + 2 \ln 4 - 2 \ln 3 + \ln 2 \\ &= \boxed{3 + 5 \ln 2 - 2 \ln 3}. \end{aligned}$$

Exercice ANA.11.3 | Changements de variables

1. Montrer que $\int_0^{\pi/2} (\cos t)^n \, dt = \int_0^{\pi/2} (\sin t)^n \, dt$.
2. Calculer les intégrales suivantes par changement de variable.

2.1) $\int_1^e \frac{\ln t}{t^2} \, dt,$

2.2) $\int_0^1 \frac{dx}{(1+x^2)^{3/2}},$ *Indication : Poser $x = \tan t$.*

Solution (exercice ANA.11.3)

1. Faisons le changement de variable $u = \frac{\pi}{2} - t$, la fonction $x \mapsto \frac{\pi}{2} - x$ étant \mathcal{C}^1 , on obtient :

$$\int_0^{\pi/2} (\cos t)^n \, dt = \int_{\pi/2}^0 \left(\cos \left(\frac{\pi}{2} - u \right) \right) (-du) = \int_0^{\pi/2} (\sin t)^n \, dt.$$

Donc $\int_0^{\pi/2} (\cos t)^n dt = \int_0^{\pi/2} (\sin t)^n dt.$

2. 2.1) Posons $u = \frac{1}{t}$, $du = -\frac{dt}{t^2}$. Donc comme $t \in [1, e] \rightarrow \frac{1}{t}$ est de classe \mathcal{C}^1 , on déduit que :

$$\int_1^e \frac{\ln t}{t^2} dt = \int_1^{1/e} (-\ln u)(-du) = \int_1^{1/e} \ln u du.$$

Puis en utilisant une primitive du logarithme, on obtient :

$$\int_1^e \frac{\ln t}{t^2} dt = [u \ln u - u]_1^{1/e} = -\frac{1}{e} - \frac{1}{e} + 1 = \boxed{1 - \frac{2}{e}}.$$

2.2) $\int_0^1 \frac{dx}{(1+x^2)^{3/2}}$. Faisons le changement $x = \tan t$, $dx = (1 + \tan^2(t)) dt$. Donc comme \tan est de classe \mathcal{C}^1 sur $[0, \frac{\pi}{4}]$, on déduit :

$$\int_0^1 \frac{dx}{(1+x^2)^{3/2}} = \int_0^{\pi/4} \frac{(1 + \tan^2(t)) dt}{(1 + \tan^2(t))^{3/2}} = \int_0^{\pi/4} \frac{dt}{\sqrt{1 + \tan^2(t)}}.$$

Mais comme $1 + \tan^2 = \frac{1}{\cos^2}$, on a $\sqrt{1 + \tan^2} = \frac{1}{|\cos t|}$. Donc puisque \cos est positive sur $[0, \pi/4]$, on a :

$$\int_0^1 \frac{dx}{(1+x^2)^{3/2}} = \int_0^{\pi/4} \cos t dt = [\sin t]_0^{\pi/4} = \boxed{\frac{\sqrt{2}}{2}}.$$

Exercice ANA.11.4 | Intégrale au service d'un équivalent

1. Montrer que : $\forall n \in \mathbf{N}^*$, $\int_1^n \ln(x) dx \leq \ln(n!) \leq \int_1^{n+1} \ln(x) dx.$
2. En déduire que : $\ln(n!) \underset{n \rightarrow \infty}{\sim} n \ln(n).$

Solution (exercice ANA.11.4)

1. Soit $k \in \mathbf{N}^*$. Alors :

$$\forall x \in [k, k+1], \quad \ln(k) \leq \ln(x) \leq \ln(k+1).$$

Donc en intégrant entre k et $k+1$:

$$\ln(k) \leq \int_k^{k+1} \ln(x) dx \leq \ln(k+1).$$

Sommons entre $k = 1$ et $k = n$:

$$\sum_{k=1}^n \ln(k) \leq \sum_{k=1}^n \int_k^{k+1} \ln(x) dx \leq \sum_{k=1}^n \ln(k+1).$$

Puis par propriété du logarithme et relation de CHASLES,

$$\ln(n!) \leq \int_1^{n+1} \ln(x) dx \leq \ln((n+1)!).$$

On peut ensuite remettre le terme souhaité au milieu :

$$\boxed{\int_1^n \ln(x) dx \leq \ln(n!) \leq \int_1^{n+1} \ln(x) dx.}$$

2. On sait qu'une primitive du logarithme est $x \mapsto x \ln(x) - x$ sur \mathbf{R}^{++} . Donc

$$[x \ln(x) - x]_1^n \leq \ln(n!) \leq [x \ln(x) - x]_1^{n+1}.$$

D'où, après calculs :

$$n \ln(n) - n + 1 \leq \ln(n!) \leq (n+1) \ln(n+1) - (n+1) + 1.$$

Et enfin, en mettant $n \ln n$ en facteur :

$$1 - \frac{1}{\ln n} + \frac{1}{n \ln n} \leq \frac{\ln(n!)}{n \ln n} \leq \frac{(n+1) \ln(n+1)}{n \ln n} - \frac{n}{\ln n}. \quad (*)$$

Mais

$$\frac{(n+1) \ln(n+1)}{n \ln n} \underset{n \rightarrow \infty}{\sim} \frac{\ln(n+1)}{\ln(n)} = \frac{\ln(n) + \ln(1 + \frac{1}{n})}{\ln n} = 1 + \frac{\ln(1 + \frac{1}{n})}{\ln n} \xrightarrow{n \rightarrow \infty} 0.$$

Le théorème d'encadrement dans (*) permet alors de conclure :

$$\boxed{\ln(n!) \underset{n \rightarrow \infty}{\sim} n \ln(n).}$$

Exercice ANA.11.5 | Déterminer la limite quand $n \rightarrow +\infty$ des suites suivantes :

$$u_n = \sum_{k=1}^n \frac{n}{n^2 + k^2}, \quad v_n = \frac{1}{n} \left[\prod_{k=1}^n \left(1 + \frac{k}{n} \right) \right]^{\frac{1}{n}}.$$

Solution (exercice ANA.11.5)

$$\begin{aligned} u_n &= \sum_{k=1}^n \frac{n}{n^2 + k^2} \\ &= \frac{1}{n} \sum_{k=1}^n \frac{1}{1 + \left(\frac{k}{n}\right)^2} \\ &\xrightarrow{n \rightarrow \infty} \int_0^1 \frac{1}{1+t^2} dt = [\arctan]_0^1 = \frac{\pi}{4}. \end{aligned}$$

car $t \mapsto \frac{1}{1+t^2}$ est continue

Pour la seconde, on va plutôt passer au logarithme pour faire apparaître des sommes.

$$\ln v_n = -\ln n + \frac{1}{n} \sum_{k=1}^n \ln \left(1 + \frac{k}{n} \right).$$

Or, puisque $x \in [0, 1] \mapsto \ln(1+x)$ est continue, il vient

$$\frac{1}{n} \sum_{k=1}^n \ln \left(1 + \frac{k}{n} \right) \xrightarrow{n \rightarrow \infty} \int_0^1 \ln(1+t) dt = \frac{1}{2} - 1 = -\frac{1}{2}.$$

Donc $\ln v_n \xrightarrow{n \rightarrow \infty} -\infty$ par opération sur les limites. Donc par composition :

$$\boxed{v_n \xrightarrow{n \rightarrow \infty} 0.}$$

Exercice ANA.11.6 | **Méthode des trapèzes sur un exemple** On pose pour tout $t \in [0, 1]$,

$$h(t) = \frac{e^t}{1+t}, \quad J = \int_0^1 h \quad \text{et} \quad \forall n \geq 1,$$

$$U_n = \frac{1}{n} \sum_{k=0}^{n-1} h\left(\frac{k}{n}\right), \quad V_n = \frac{1}{n} \sum_{k=0}^{n-1} h\left(\frac{k+1}{n}\right).$$

1. Vérifier que h est croissante.
2. Montrer que pour tout $k \in \{0, \dots, n-1\}$,

$$\frac{1}{n} h\left(\frac{k}{n}\right) \leq \int_{k/n}^{(k+1)/n} h(t) dt \leq \frac{1}{n} h\left(\frac{k+1}{n}\right).$$

3. En déduire que : $\forall n \in \mathbf{N}^*, \quad U_n \leq J \leq V_n.$
4. Vérifier que pour tout $n \geq 1$:

$$\left| J - \frac{U_n + V_n}{2} \right| \leq \frac{h(1) - h(0)}{2n}.$$

5. Écrire une fonction qui prend en paramètre ε et qui retourne une valeur approchée de J à ε près.
6. Que représente géométriquement la quantité $\frac{U_n + V_n}{2}$?

Solution (exercice ANA.11.6)

1. Soit $t \in [0, 1]$. Alors

$$h'(t) = \frac{e^t(1+t) - e^t}{(t+1)^2} = \frac{t \cdot e^t}{(t+1)^2} \geq 0.$$

Donc h est une fonction croissante.

2. Soit $k \in \{0, \dots, n-1\}$, alors pour tout $t \in \left[\frac{k}{n}, \frac{k+1}{n}\right]$, on a :

$$\begin{aligned} h\left(\frac{k}{n}\right) &\leq h(t) \leq h\left(\frac{k+1}{n}\right) \\ \int_{k/n}^{(k+1)/n} h\left(\frac{k}{n}\right) dt &\leq \int_{k/n}^{(k+1)/n} h(t) dt \leq \int_{k/n}^{(k+1)/n} h\left(\frac{k+1}{n}\right) dt. \end{aligned}$$

intégration

Puis en calculant la différence des deux bornes, on déduit l'encadrement souhaité :

$$\boxed{\frac{1}{n} h\left(\frac{k}{n}\right) \leq \int_{k/n}^{(k+1)/n} h(t) dt \leq \frac{1}{n} h\left(\frac{k+1}{n}\right).}$$

3. Soit $n \in \mathbf{N}^*$. Alors d'après la relation de CHASLES, on a :

$$\frac{1}{n} \sum_{k=0}^{n-1} h\left(\frac{k}{n}\right) \leq \int_0^1 h(t) dt \leq \frac{1}{n} \sum_{k=0}^{n-1} h\left(\frac{k+1}{n}\right).$$

On reconnaît alors :

$$\boxed{U_n \leq J \leq V_n}.$$

4. Soit $n \geq 1$, la majoration est équivalente à :

$$J - \frac{U_n + V_n}{2} \leq \frac{h(1) - h(0)}{2n}.$$

Mais comme $J \leq V_n$, on a

$$\begin{aligned} J - \frac{U_n + V_n}{2} &\leq V_n - \frac{U_n + V_n}{2} \\ &= \frac{V_n - U_n}{2} \\ &= \frac{1}{n} \sum_{k=0}^{n-1} \left(h\left(\frac{k+1}{n}\right) - h\left(\frac{k}{n}\right) \right) \\ &= \frac{h(1) - h(0)}{2n}. \end{aligned} \quad \left. \begin{array}{l} \\ \\ \\ \end{array} \right\} \text{téléscopage}$$

De-même pour le minorant :

$$\begin{aligned} J - \frac{U_n + V_n}{2} &\geq U_n - \frac{U_n + V_n}{2} \\ &= \frac{U_n - V_n}{2} \\ &= \frac{h(0) - h(1)}{2n}. \end{aligned} \quad \left. \begin{array}{l} \\ \\ \end{array} \right\} \text{téléscopage}$$

Donc :

$$\boxed{\left| J - \frac{U_n + V_n}{2} \right| \leq \frac{h(1) - h(0)}{2n}}.$$

5. D'après la question précédente, la moyenne des sommes est une bonne approximation de J .

```
import math as ma
```

```
def h(x):
    return ma.exp(x)/(1+x)
```

```
def trapeze(h, eps):
    n = 1
    while (h(1)-h(0))/(2*n) > eps:
        n += 1
    # on a calculé le bon n, puis on calcule le T associe
    T = 0
    for k in range(0, n):
        T += h(k/n)+h((k+1)/n)
    return T/(2*n)
```

On peut comparer par exemple avec les méthodes des rectangles du cours.

```
>>> trapeze(h, 10**(-3))
1.125387830946576
>>> rectangle_RD(h, 0, 1, 10**3)
1.1255657101712513
>>> rectangle_RD(h, 0, 1, 10**3)
1.1255657101712513
```

6. Que représente géométriquement la quantité $\frac{U_n + V_n}{2}$? La somme des aires des trapèzes de base les segments $\left[\frac{k}{n}, \frac{k+1}{n}\right]$.

Exercice ANA.11.7 | Pour $n \in \mathbf{N}$, on pose :

$$I_n = \frac{1}{2^{n+1}n!} \int_0^1 (1-t)^n e^{\frac{t}{2}} dt.$$

1. Calculer I_0 .

2. Trouver une relation entre I_n et I_{n+1} . Que vaut $\lim_{n \rightarrow +\infty} I_n$?
3. On pose $S_n = \sum_{k=0}^n \frac{1}{2^k k!}$. Dédurre des résultats précédents la convergence de la suite $(S_n)_{n \in \mathbb{N}}$ et la valeur de sa limite.

Solution (exercice ANA.11.7)

1. $I_0 = \frac{1}{2} \int_0^1 e^{\frac{t}{2}} dt = \left[e^{\frac{t}{2}} \right]_0^1 = \boxed{\sqrt{e} - 1}$.

2. Soit $n \in \mathbb{N}$. Alors, en faisant une intégration par parties, on déduit :

$$\begin{aligned} \int_0^1 (1-t)^{n+1} e^{\frac{t}{2}} dt &= 2 \int_0^1 (1-t)^{n+1} \left(\frac{1}{2} e^{\frac{t}{2}} \right) dt \\ &= 2 \int_0^1 (n+1)(1-t)^n e^{\frac{t}{2}} dt + 2 \left[(1-t)^{n+1} e^{\frac{t}{2}} \right]_0^1 \end{aligned}$$

$\left. \begin{array}{l} t \mapsto e^{\frac{t}{2}}, t \mapsto \\ (1-t)^{n+1} \text{ sont } \mathcal{C}^1 \end{array} \right\}$

$$= 2(n+1)2^{n+1} n! I_n - 2.$$

Donc finalement

$$2^{n+2}(n+1)I_{n+1} = 2(n+1)2^{n+1} n! I_n - 2.$$

Ou autrement dit :

$$\boxed{I_{n+1} = I_n - \frac{1}{2^{n+1}(n+1)!}}$$

Cette relation de récurrence n'est pas standard. En revanche, on peut majorer l'intégrale :

$$|I_n| \leq \frac{1}{2^{n+1} n!} \int_0^1 |(1-t)^n e^{\frac{t}{2}}| dt.$$

Or, en utilisant la croissance de l'exponentielle

$$\forall t \in [0, 1], \quad |(1-t)^n e^{\frac{t}{2}}| \leq 1\sqrt{e}.$$

Donc :

$$0 \leq |I_n| \leq \frac{\sqrt{e}}{2^{n+1} n!} \xrightarrow{n \rightarrow \infty} 0.$$

Ainsi, par théorème d'encadrement : $\boxed{\lim_{n \rightarrow +\infty} I_n = 0}$.

3. Nous avons établi

$$I_{k+1} = I_k - \frac{1}{2^{k+1}(k+1)!}.$$

Donc en sommant entre 0 et n , on déduit

$$\sum_{k=0}^n (I_k - I_{k+1}) = \sum_{k=0}^n \frac{1}{2^{k+1}(k+1)!}.$$

Ce qui, par télescopage, donne :

$$I_0 - I_{n+1} = S_{n+1} - \frac{1}{2^{0} 0!}.$$

Donc en faisant $n \rightarrow \infty$, on déduit :

$$S_n \xrightarrow{n \rightarrow \infty} I_0 + 1 = \boxed{\sqrt{e}}.$$

En d'autres termes, la série associée converge et

$$\sum_{n=0}^{\infty} \frac{1}{2^k k!} = \sum_{n=0}^{\infty} \frac{(1/2)^k}{k!} = \sqrt{e}.$$

Ce que l'on savait déjà grâce au cours sur les séries.

Exercice ANA.11.8 | On pose $f(t) = \frac{\exp(t)-1}{t^2} - \frac{1}{t}$ pour tout $t \in \mathbb{R}^*$.

1. Montrer que f est prolongeable par continuité en 0.
2. On suppose $0 < a < b$. On note

$$\forall x \in \mathbb{R}^{+*}, \quad g(x) = \int_{ax}^{bx} \frac{\exp(t) - 1}{t^2} dt.$$

Déterminer la limite de $g(x)$ quand x tend vers 0^+ .

Solution (exercice ANA.11.8)

1. Faisons un développement limité au voisinage de 0 pour la fonction f . On a :

$$f(t) = \frac{\exp(t) - 1}{t^2} - \frac{1}{t}$$

$$\underset{t \rightarrow 0}{\sim} \frac{t + t^2 + o(t^2)}{t^2} - \frac{1}{t}$$

$$\underset{t \rightarrow 0}{\sim} 1 + o(1) \xrightarrow{t \rightarrow 0} 1.$$

Donc f est prolongeable par continuité en zéro, en posant $f(0) = 1$.

2. On suppose que $0 < a < b$. Faisons intervenir la fonction f en écrivant :

$$g(x) = \int_{ax}^{bx} f + \int_{ax}^{bx} \frac{dt}{t}.$$

Choisissons une primitive de f sur \mathbf{R}^+ , que l'on note F . Alors

$$g(x) = F(bx) - F(ax) + \ln\left(\frac{a}{b}\right).$$

Or, F est continue car dérivable sur \mathbf{R}^+ , donc

$$F(bx) - F(ax) \xrightarrow{x \rightarrow 0^+} F(0) - F(0) = 0.$$

Ainsi, $\lim_{x \rightarrow 0^+} g(x) = \ln\left(\frac{a}{b}\right)$.

Exercice ANA.11.9 | Intégrale à bornes variables Soit $F(x) = \int_0^x e^{-t^2} dt$ et $G(x) = \int_0^1 e^{-(xu)^2} du$ pour tout $x \in \mathbf{R}$.

1. Montrer à l'aide d'un changement de variable que : $\forall x \in \mathbf{R}, F(x) = xG(x)$.

2. Montrer que G est dérivable sur \mathbf{R}^* et que :

$$\forall x \neq 0 \quad G'(x) = \frac{xf(x) - F(x)}{x^2}, \quad \text{où } f(x) = e^{-x^2}.$$

Est-elle dérivable en zéro ?

3. En déduire les variations de G .

Solution (exercice ANA.11.9)

1. On a envie de poser $t = xu$ dans la seconde intégrale. Le changement est licite puisque $u \mapsto xu$ est une fonction \mathcal{C}^1 pour tout $x \neq 0$ — on ne peut faire ce changement si $x = 0$ puisqu'on ici on a besoin de diviser par la dérivée, qui est nulle. Donc,

$$\int_0^1 e^{-(xu)^2} du = \int_0^1 e^{-t^2} \frac{dt}{x}.$$

Donc en multipliant par x , on déduit :

$$\forall x \neq 0 \quad F(x) = xG(x).$$

Notez que la formule est vraie aussi pour $x = 0$ puisque $F(0) = 0$ et $0G(0) = 0$. Finalement

$$\forall x \in \mathbf{R}, F(x) = xG(x).$$

2. Soit $x \neq 0$, on a

$$G(x) = \frac{F(x)}{x}.$$

Donc G est dérivable sur \mathbf{R}^* car F l'est (c'est l'unique primitive de $t \mapsto e^{-t^2}$ s'annulant en zéro). Et :

$$\forall x \in \mathbf{R}^*, G'(x) = \frac{f(x) \cdot x - F(x)}{x^2}.$$

Montrons que G est dérivable en zéro. Et en fait elle n'est pas continue en zéro ! Puisque $G(0) = 1$, alors que :

$$\forall x \neq 0, G(x) = \frac{F(x)}{x} = \frac{F(x) - F(0)}{x - 0} \xrightarrow{x \rightarrow 0} f(0) = 1 = G(0).$$

Donc : G est continue en zéro. De plus, pour tout $x \neq 0$,

$$\frac{G(x) - G(0)}{x - 0} = \frac{\frac{F(x)}{x} - 1}{x} = \frac{F(x) - x}{x^2} = \frac{1}{x} \frac{(F(x) - x) - (F(0) - 0)}{x}.$$

Or, puisque $x \mapsto F(x) - x$ est dérivable en zéro,

$$\frac{(F(x) - x) - (F(0) - 0)}{x} \xrightarrow{x \rightarrow 0} f(0) - 0 = 1,$$

et donc

$$\frac{G(x) - G(0)}{x - 0} \xrightarrow{x \rightarrow 0^\pm} \pm\infty.$$

Ainsi, G n'est pas dérivable en zéro.

3. La meilleure option est de ne pas utiliser la dérivée. Soient $x \leq y$ deux réels. Alors

$$\forall u \in [0, 1], \quad e^{-(yu)^2} \leq e^{-(xu)^2}.$$

Donc en intégrant entre 0 et 1, on déduit : $G(y) \leq G(x)$. Ainsi, G est une fonction décroissante.

Exercice ANA.11.10 | Intégrale sur une période d'une fonction continue périodique

Soit $f : \mathbf{R} \rightarrow \mathbf{R}$ une fonction continue, supposée de plus T-périodique avec $T > 0$.

1. Montrer que :

$$\forall a \in \mathbf{R}, \quad \int_T^{a+T} f(t) dt = \int_0^a f(t) dt.$$

2. En déduire :

$$\forall a \in \mathbf{R}, \quad \int_0^T f = \int_a^{a+T} f.$$

Solution (exercice ANA.11.10)

1. Soit $a \in \mathbf{R}$. Alors faisons le changement $u = t - T$, qui est \mathcal{C}^1 , donc puis f est T-périodique, on déduit que :

$$\int_T^{a+T} f(t) dt = \int_0^a f(u + T) du = \int_0^a f(u) du.$$

2. De plus, par la relation de CHASLES : $\forall a \in \mathbf{R}$,

$$\begin{aligned} \int_a^{a+T} f &= \int_a^0 f + \int_0^T f + \int_T^{a+T} f \\ &= \int_a^0 f + \int_0^T f + \int_0^a f(u) du. \end{aligned}$$

Les termes 1 et 3 s'annulant, il vient le résultat.

Exercice ANA.11.11 | Inégalité de CAUCHY-SCHWARZ Soient $a < b$ deux réels, et f, g deux fonctions continues sur $[a, b]$.

1. Montrer que :

$$\left| \int_a^b f \cdot g \right| \leq \sqrt{\int_a^b f^2} \cdot \sqrt{\int_a^b g^2}.$$

On pourra introduire

$$P : \begin{cases} \mathbf{R} \rightarrow \mathbf{R}, \\ \lambda \mapsto \int_a^b (\lambda f + g)^2. \end{cases}$$

2. Dans le cas où f et g sont continues, on a égalité si et seulement si (f, g) est une famille liée.

Solution (exercice ANA.11.11)

1. Soit $\lambda \in \mathbf{R}$. Alors

$$P(\lambda) = \int_a^b (\lambda^2 f^2 + g^2 + 2\lambda f g).$$

Donc par linéarité de l'intégrale, on déduit :

$$P(\lambda) = \lambda^2 \int_a^b f^2 + 2\lambda \int_a^b f g + \int_a^b g^2 = \int_a^b (\lambda f + g)^2 \geqslant .$$

Donc P est un trinôme en λ , qui est de plus positif et coefficient dominant positif aussi. Ainsi, le discriminant de P est négatif, de sorte que

$$4 \left(\int_a^b f g \right)^2 - 4 \cdot \int_a^b f^2 \cdot \int_a^b g^2 \leqslant 0.$$

En passant à la racine, on déduit alors

$$\left| \int_a^b f \cdot g \right| \leqslant \sqrt{\int_a^b f^2} \cdot \sqrt{\int_a^b g^2}.$$

2. On a égalité dans dans l'inégalité de CAUCHY-SCHWARZ si

$$\begin{aligned} \Delta = 0 &\iff \exists \lambda \in \mathbf{R}, \quad P(\lambda) = 0 \\ &\iff \exists \lambda \in \mathbf{R}, \quad \int_a^b (\lambda f + g)^2 = 0 \\ &\iff \lambda f + g = 0 \\ &\iff \boxed{(f, g) \text{ liée.}} \end{aligned} \quad \left. \vphantom{\int_a^b (\lambda f + g)^2 = 0} \right\} \textit{intégrande positive}$$

Exercice ANA.11.12 | Intégrale à paramètre Soit f la fonction définie sur \mathbf{R} par : $\forall x \in$

$$\mathbf{R}, \quad f(x) = \int_0^1 \frac{\sin(tx)}{1+t^2} dt.$$

1. Montrer que f est bien définie et calculer $f(0)$.
2. ➤_🔧 À l'aide d'une méthode des rectangles, proposer un script permettant de réaliser son tracé sur l'intervalle $[0, 10]$.

3. Montrer que pour tous $x, y \in \mathbf{R}$, $|f(x) - f(y)| \leqslant \frac{\ln 2}{2} |x - y|$.
4. Montrer que f est continue bornée sur \mathbf{R} .

Solution (exercice ANA.11.12)

1. Puisque pour tout $x \in \mathbf{R}$, $t \mapsto \frac{\sin(tx)}{1+t^2}$ est continue, la fonction f est bien définie. De plus,

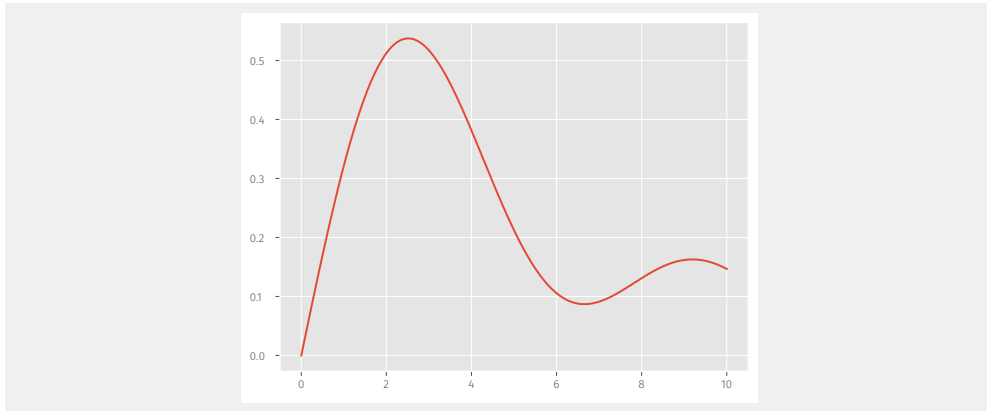
$$f(0) = \int_0^1 0 = \boxed{0}.$$

```
2. def g(t, x):
    return ma.sin(t*x)/(1+t**2)

def f(x):
    """
    retourne une approximation de f(x) suivant la méthode des
    rectangles à gauche
    """
    a = 0
    b = 1
    n = 10**3
    S = 0
    h = (b-a)/n
    for i in range(n):
        S += g(a+h*i, x)
    return S*h
```

Ensuite, on trace.

```
X = np.linspace(0, 10, 10**3)
Y = [f(x) for x in X]
plt.plot(X, Y)
```



3. Constatons déjà que pour tout $a, b \in \mathbf{R}$, on a l'existence d'un $c \in]a, b[$ (puisque \sin est continue sur $[a, b]$, dérivable sur $]a, b[$) vérifiant :

$$|\sin(a) - \sin(b)| = |\cos(c)| |b - a|,$$

donc finalement on a montré que :

$$\forall a, b \in \mathbf{R}, \quad |\sin(a) - \sin(b)| \leq |b - a|. \quad (\star)$$

Soient x, y deux réels. Alors

$$\begin{aligned} |f(x) - f(y)| &\leq \int_0^1 \frac{1}{1+t^2} |\sin(tx) - \sin(ty)| dt \\ &\leq \int_0^1 \frac{1}{1+t^2} |tx - ty| dt \\ &\leq |x - y| \int_0^1 \frac{t}{1+t^2} dt = \frac{1}{2} |x - y| \left[\ln(1+t^2) \right]_0^1. \end{aligned} \quad (\star)$$

Finalement, on a établi :

$$\boxed{\forall x, y \in \mathbf{R}, \quad |f(x) - f(y)| \leq \frac{\ln 2}{2} |x - y|}.$$

4. Vérifions par exemple la caractérisation séquentielle de la continuité. Soit $a \in \mathbf{R}$ et (x_n) telle que $x_n \xrightarrow{n \rightarrow \infty} a$. Montrons que $f(x_n) \xrightarrow{n \rightarrow \infty} f(a)$. D'après la question précédente, on a :

$$\forall n \in \mathbf{N}, \quad 0 \leq |f(x_n) - f(a)| \leq \frac{\ln 2}{2} |x_n - a|,$$

puisque $|x_n - a| \xrightarrow{n \rightarrow \infty} 0$, le théorème d'encadrement permet alors de conclure : $f(x_n) \xrightarrow{n \rightarrow \infty} f(a)$. Donc f est continue en a pour tout a , donc est continue sur \mathbf{R} . De plus, pour $x \in \mathbf{R}$,

$$|f(x)| \leq \int_0^1 |\sin(tx)| \frac{dt}{1+t^2} \leq \int_0^1 \frac{dt}{1+t^2} = \frac{\pi}{4}.$$

Donc f est bornée.

Exercice ANA.11.13 | Application linéaire définie par une intégrale Soit E le \mathbf{R} -espace vectoriel des fonctions de classe \mathcal{C}^1 sur $[0, 1]$. Pour $n \in \mathbf{N}$, on définit

$$L_n \begin{cases} E & \longrightarrow & F, \\ f & \longmapsto & L_n(f) : \left(x \in [0, 1] \longrightarrow \int_0^x \frac{(x-t)^n}{n!} f(t) dt \right). \end{cases}$$

1. Soit $n \in \mathbf{N}^*$. Montrer que $L_n(f)$ est dérivable et que $(L_n(f))' = L_{n-1}(f)$.
2. Montrer que $L_n(f)$ est de classe \mathcal{C}^{n+2} sur $[0, 1]$ pour tout $n \in \mathbf{N}$.
3. Montrer que L_n est un endomorphisme de E et déterminer son noyau.

Solution (exercice ANA.11.13)

1. Le problème majeur est que la variable x est présente aussi dans l'intégrale. Nous allons développer l'intégrande à l'aide de la formule du binôme. Pour tout $x \in \mathbf{R}$,

$$\begin{aligned} \int_0^x \frac{(x-t)^n}{n!} f(t) dt &= \int_0^x \sum_{k=0}^n \binom{n}{k} \frac{x^k (-1)^{n-k} t^{n-k}}{n!} f(t) dt \\ &= \sum_{k=0}^n \frac{(-1)^{n-k}}{k!(n-k)!} x^k \int_0^x t^{n-k} f(t) dt. \end{aligned} \quad \left. \begin{array}{l} \\ \\ \end{array} \right\} \text{linéarité de l'intégrale}$$

Rappelons que $x \longmapsto \int_0^x t^{n-k} f(t) dt$ est l'unique primitive de $t \longmapsto t^{n-k} f(t)$ s'annulant en zéro. On peut alors dériver en utilisant la dérivée d'un produit. Pour tout

$x \in \mathbf{R}$,

$$\begin{aligned} L_n(f)'(x) &= \sum_{k=0}^n \frac{(-1)^{n-k}}{k!(n-k)!} \left(kx^{k-1} \cdot \int_0^x t^{n-k} f(t) dt + x^k \cdot x^{n-k} f(x) \right) \\ &= \sum_{k=0}^n \frac{(-1)^{n-k}}{k!(n-k)!} \left(kx^{k-1} \cdot \int_0^x t^{n-k} f(t) dt + x^n f(x) \right) \\ &= \sum_{k=1}^n \frac{(-1)^{n-k}}{(k-1)!(n-k)!} x^{k-1} \cdot \int_0^x t^{n-k} f(t) dt + \left(\sum_{k=0}^n \frac{(-1)^{n-k}}{k!(n-k)!} \right) \cdot x^n f(x) \\ &= \sum_{k=0}^{n-1} \frac{(-1)^{n-k-1}}{k!(n-k-1)!} x^k \cdot \int_0^x t^{n-k-1} f(t) dt + n!(1-1)^n \cdot x^n f(x) \\ &= \frac{1}{(n-1)!} \int_0^x \sum_{k=0}^{n-1} \binom{n-1}{k} x^k (-t)^{n-1-k} f(t) dt + 0 \\ &= \int_0^x \frac{(x-t)^{n-1}}{(n-1)!} f(t) dt. \end{aligned}$$

changement d'indice et binôme linéarité de l'intégrale
binôme

Donc : $(L_n(f))' = L_{n-1}(f)$ pour tout $n \geq 1$.

2. Par récurrence immédiate à partir de la première question, on a que $L_n(f)$ est n fois dérivable, et que $L_n(f)^{(n)} = L_0(f)$. Or

$$L_0(f) : x \in [0, 1] \mapsto \int_0^x f$$

est l'unique primitive de f s'annulant en zéro, elle est donc aussi dérivable et

$$L_n(f)^{(n+1)} = L_0(f)' = f.$$

Mais f est \mathcal{C}^1 donc $L_n(f)$ est $n+2$ fois dérivable de dérivée $(n+2)$ -ième continue.

Ainsi $L_n(f)$ est de classe \mathcal{C}^{n+2} .

3. Si $f \in \mathbb{E}$, alors $L_n(f) \in \mathbb{E}$ d'après la question précédente. De plus, soient $f, g \in \mathbb{E}, \lambda, \mu \in \mathbf{R}$ et $x \in [0, 1]$:

$$\begin{aligned} L_n(\lambda f + \mu g)(x) &= \int_0^x \frac{(x-t)^n}{n!} (\lambda f + \mu g)(t) dt \\ &= \lambda \int_0^x \frac{(x-t)^n}{n!} f(t) dt + \mu \int_0^x \frac{(x-t)^n}{n!} g(t) dt. \end{aligned}$$

linéarité de l'intégrale

Donc : L_n est un endomorphisme de \mathbb{E} . Soit de plus $f \in \ker L_n$ i.e. telle que $L_n(f) = 0$. Alors en dérivant $n-1$ fois, on obtient que : $L_0(f)(x) = 0 = \int_0^x f$ pour tout $x \in \mathbf{R}$. Ainsi, en redérivant, on déduit que f est nulle, donc $\ker L_n = \{0\}$.

3.2. Intégrales impropres

Exercice ANA.11.14 | Déterminer la nature des intégrales ci-dessous.

- ▶ $\int_0^{+\infty} \frac{t^2}{e^t - 1} dt,$
- ▶ $\int_2^{+\infty} \sin\left(\frac{1}{(3t+1)^2}\right) dt,$
- ▶ $\int_3^{+\infty} \frac{dt}{t^2 - 4t + 3},$
- ▶ $\int_0^5 \frac{dt}{\sqrt{t(5-t)}},$
- ▶ $\int_0^1 \frac{x}{1-x^2} dx,$
- ▶ $\int_3^{+\infty} \frac{\sqrt{x^3}}{x^2 - 1} dx.$
- ▶ $\int_0^{+\infty} e^{-t^2} dt,$
- ▶ $\int_0^{+\infty} x \sin x e^{-x} dx,$
- ▶ $\int_0^{+\infty} \ln t e^{-t} dt,$
- ▶ $\int_0^1 \frac{dt}{(1-t)\sqrt{t}},$
- ▶ $\int_0^1 \cos^2\left(\frac{1}{t}\right) dt.$

Dans tout l'exercice, on pourra utiliser, uniquement en cas de besoin, le résultat sur les équivalents.

Solution (exercice ANA.11.14)

- ▶ L'intégrale est généralisée en 0 et $+\infty$, on étudie donc $\int_0^1 \frac{t^2}{e^t - 1} dt$ et $\int_1^{+\infty} \frac{t^2}{e^t - 1} dt$. Mais $\frac{t^2}{e^t - 1} \underset{t \rightarrow 0}{\sim} \frac{t^2}{t} \xrightarrow{t \rightarrow \infty} 0$ donc la fonction est prolongeable en zéro par continuité donc la première intégrale converge. De plus, $t^2 \frac{t^2}{e^t - 1} \xrightarrow{t \rightarrow \infty} 0$ par

croissances comparées, donc pour t assez grand

$$0 \leq \frac{t^2}{e^t - 1} \leq \frac{1}{t^2}.$$

Or, $\int_1^\infty \frac{1}{t^2} dt$ converge, donc par théorème de comparaison, on déduit que

$$\int_1^\infty \frac{t^2}{e^t - 1} dt \text{ converge. Ainsi, } \boxed{\int_0^{+\infty} \frac{t^2}{e^t - 1} dt \text{ converge.}}$$

- L'intégrande est bien une fonction positive, l'intégrale est généralisée en $+\infty$. De plus,

$$\forall t \in [2, \infty[, \quad 0 \leq \left| \sin\left(\frac{1}{(3t+1)^2}\right) \right| \leq \frac{1}{(3t+1)^2} \leq \frac{1}{9t^2}$$

puisque pour tout $y \in \mathbf{R}^+$, $\sin(y) \leq y$ — à montrer par étude de fonction.

Puisque $\int_2^\infty \frac{1}{t^2} dt$ converge, on obtient par théorème de comparaison que

$$\boxed{\int_2^{+\infty} \sin\left(\frac{1}{(3t+1)^2}\right) dt \text{ converge.}}$$

- Remarquons que 3 est l'unique racine du dénominateur sur $[3, \infty[$, l'intégrale est donc généralisée en $3, \infty$. On étudie $\int_3^4 \frac{dt}{t^2 - 4t + 3}$ et $\int_4^\infty \frac{dt}{t^2 - 4t + 3}$. De plus, $t^2 - 4t + 3 = (t - 3)(t - 1)$ pour tout t . Cherchons a, b deux réels de sorte que

$$\forall t > 3, \quad \frac{1}{t^2 - 4t + 3} = \frac{a}{t - 3} + \frac{b}{t - 1} = \frac{a(t - 1) + b(t - 3)}{t^2 - 4t + 3}.$$

Par identification, on déduit que $a + b = 0$, $-a - 3b = 1$ soit $b = -\frac{1}{2}$, $a = \frac{1}{2}$. On est donc capables de calculer les intégrales partielles.

— Soit $A \in]3, 4[$:

$$\begin{aligned} \int_A^4 \frac{dt}{t^2 - 4t + 3} &= \frac{1}{2} \left([\ln|t - 3|]_A^4 - [\ln|t - 1|]_A^4 \right) \\ &= \frac{1}{2} \left(-\ln|A - 3| - \ln 3 + \ln(A - 1) \right) \\ &\xrightarrow{A \rightarrow 4} \frac{\ln(2/3)}{2}. \end{aligned}$$

— On obtient de même sur $[4, \infty[$, pour $B \in]4, \infty[$:

$$\begin{aligned} \int_4^B \frac{dt}{t^2 - 4t + 3} &= \frac{1}{2} \left([\ln|t - 3|]_4^B - [\ln|t - 1|]_4^B \right) \\ &= \frac{1}{2} \left(\ln|B - 3| + \ln 3 - \ln(B - 1) \right) \\ &= \frac{1}{2} \left(\ln\left(\frac{B - 3}{B - 1}\right) + \ln 3 \right) \\ &\xrightarrow{B \rightarrow \infty} \frac{\ln 3}{2}. \end{aligned}$$

L'intégrale $\boxed{\int_3^{+\infty} \frac{dt}{t^2 - 4t + 3}}$ est donc convergente.

- On va procéder par équivalents. L'intégrale est généralisée en 0 et 5, on étudie donc $\int_0^{5/2} \frac{dt}{\sqrt{t(5-t)}}$ et $\int_{5/2}^5 \frac{dt}{\sqrt{t(5-t)}}$. La fonction intégrée est bien positive. De plus,

$$\frac{1}{\sqrt{t(5-t)}} \underset{t \rightarrow 0}{\sim} \frac{1}{\sqrt{5}} \frac{1}{\sqrt{t}}, \quad \frac{1}{\sqrt{t(5-t)}} \underset{t \rightarrow 5}{\sim} \frac{1}{\sqrt{5}} \frac{1}{\sqrt{5-t}}.$$

Mais, $\int_0^{5/2} \frac{1}{\sqrt{t}} dt$ converge et $\int_{5/2}^5 \frac{1}{\sqrt{5-t}} dt$. En effet, montrons la convergence de la deuxième : soit $B \in [5/2, 5[$,

$$\int_{5/2}^B \frac{dt}{\sqrt{5-t}} = \left[-2\sqrt{5-t} \right]_{5/2}^B = 2 \left(\sqrt{5 - \frac{5}{2}} - \sqrt{5-B} \right)$$

qui converge lorsque B tend vers 5. D'après le critère de comparaison avec équivalents, on a donc prouvé que

$$\int_0^5 \frac{dt}{\sqrt{t(5-t)}} \text{ converge.}$$

- Un calcul explicite de primitive est possible. L'intégrale est généralisée en 1, soit donc $A \in [0, 1[$, alors

$$\int_0^A \frac{x}{1-x^2} dx = -\frac{1}{2} \ln(1-A^2) \xrightarrow{A \rightarrow 1} +\infty.$$

Donc $\int_0^1 \frac{x}{1-x^2} dx$ diverge.

- ▶ Le résultat sur les équivalents nous convainc assez vite de la divergence. Pour le montrer, constatons que :

$$\forall x \geq 3, \quad \frac{\sqrt{x^3}}{x^2-1} \geq \frac{\sqrt{x^3}}{x^2} = \frac{1}{\sqrt{x}}.$$

Or, pour tout $A > 3$,

$$\int_3^A \frac{dx}{\sqrt{x}} = \left[\frac{x^{-1/2+1}}{-1/2+1} \right]_3^A \xrightarrow{A \rightarrow \infty} \infty.$$

Le théorème de comparaison pour les fonctions positives permet alors de conclure :

$\int_3^{+\infty} \frac{\sqrt{x^3}}{x^2-1} dx$ diverge.

- ▶ Il suffit d'établir la convergence de $\int_1^{+\infty} e^{-t^2} dt$, l'intégrande étant continue sur $[0, 1]$. Mais, pour tout $t \geq 1$,

$$0 \leq e^{-t^2} \leq e^{-t^2/2}.$$

Or, d'après le cours $\int_1^{\infty} e^{-t^2/2} dt$ converge donc par théorème de comparaison

pour les fonctions positives $\int_1^{\infty} e^{-t^2} dt$ puis $\int_0^{\infty} e^{-t^2/2} dt$ converge.

- ▶ La fonction n'est clairement pas positive, on majore donc la valeur absolue :

$$\forall x \in \mathbf{R}^+, \quad 0 \leq |x \sin x e^{-x}| \leq x e^{-x}.$$

Or, par intégration par parties (cf. cours), $\int_0^{\infty} x e^{-x} dx$ converge, donc par théo-

rème de comparaison $\int_0^{\infty} x \sin x e^{-x} dx$ converge car converge absolument.

- ▶ Pour les dernières intégrales, on s'autorise le critère sur les équivalents. On a : $\ln(t)e^{-t} \underset{t \rightarrow 0}{\sim} \ln(t)$, mais $\int_0^1 \ln(t) dt$ converge (cf. cours, avec une intégration par

parties) et \ln est de signe constant sur $]0, 1]$ donc par théorème sur les équivalents, $\int_0^1 \ln(t)e^{-t} dt$ converge. De plus, par croissances comparées :

$$t^2 \ln(t)e^{-t} \xrightarrow{t \rightarrow \infty} 0,$$

donc pour t assez grand, $\ln(t)e^{-t} \leq \frac{1}{t^2}$. Mais $\int_1^{\infty} \frac{dt}{t^2}$ converge, d'où la conver-

gence par théorème de comparaison. Ainsi, $\int_0^1 \ln(t)e^{-t} dt$ converge.

- ▶ Là encore, on s'autorise le critère sur les équivalents. On a $\frac{1}{(1-t)\sqrt{t}} \underset{t \rightarrow 0}{\sim} \frac{1}{\sqrt{t}}$, $\frac{1}{(1-t)\sqrt{t}} \underset{t \rightarrow 1}{\sim} \frac{1}{1-t}$, les fonctions mises en jeu sont positives, et en calculant les intégrales partielles on montre facilement que $\int_{1/2}^1 \frac{1}{1-t} dt$ diverge. Donc

$\int_0^1 \frac{dt}{(1-t)\sqrt{t}}$ diverge.

- ▶ On a la majoration évidente suivante $|\cos^2(\frac{1}{t})| \leq 1$ pour tout $t \in]0, 1]$, donc par théorème de comparaison, puisque $\int_0^1 1 dt$ converge, on déduit que

$\int_0^1 \cos^2(\frac{1}{t}) dt$ converge aussi.

Exercice ANA.11.15 | On considère l'intégrale c-dessous, en cas de convergence

$$I = \int_{-\infty}^{+\infty} \frac{1}{(1+e^t)(1+e^{-t})} dt.$$

1. Montrer que $\frac{1}{(1+e^t)(1+e^{-t})} = \frac{e^t}{(1+e^t)^2}$ pour tout $t \in \mathbf{R}$.
2. Étudier la convergence de I , et calculer la valeur de I le cas échéant.

Solution (exercice ANA.11.15)

1. Soit $t \in \mathbf{R}$. Alors :

$$\frac{1}{(1+e^t)(1+e^{-t})} = \frac{e^t}{(1+e^t)(e^t+1)} = \frac{e^t}{(1+e^t)^2}$$

2. Cette précédente formule nous permet de calculer l'intégrale partielle. Constatons déjà que l'intégrande est paire, donc il suffit d'étudier la nature de $I' = \int_0^{+\infty} \frac{1}{(1+e^t)(1+e^{-t})} dt$. Soit $A > 0$, alors

$$\int_0^A \frac{1}{(1+e^t)(1+e^{-t})} dt = \left[-\frac{e^t}{(1+e^t)^2} \right]_0^A = \frac{1}{2} - \frac{e^A}{(1+e^A)^2} \xrightarrow{A \rightarrow \infty} \frac{1}{2}$$

Donc l'intégrale converge et $I = \frac{1}{2} + \frac{1}{2} = \boxed{1}$.

Exercice ANA.11.16 | Factorielles sous forme d'intégrale On définit une suite d'intégrales (I_n) par :

$$\forall n \in \mathbf{N}, \quad I_n = \int_0^{+\infty} t^n e^{-t} dt.$$

1. Montrer que I_n est définie pour tout $n \in \mathbf{N}$.
2. Déterminer l'expression de I_n en fonction de $n \in \mathbf{N}$.

Solution (exercice ANA.11.16)

1. Montrer la convergence par récurrence.

■ **Initialisation.** Soit $A > 0$, alors $\int_0^A 1e^{-t} dt = 1 - e^{-A} \xrightarrow{A \rightarrow \infty} 1$. Donc I_0 converge et vers 1. Donc

■ **Hérédité.** Supposons que I_n converge, alors la fonction $A \mapsto \int_0^A t^n e^{-t} dt$ admet une limite quand $A \rightarrow \infty$. Alors, puisque les fonctions $t \mapsto t^{n+1}$, $t \mapsto -e^{-t}$ sont de classe \mathcal{C}^1 , on obtient par intégration par parties :

$$\int_0^A t^{n+1} e^{-t} dt = -(n+1) \int_0^A t^n (-e^{-t}) dt + [-t^{n+1} e^{-t}]_0^A.$$

Mais $[-t^{n+1} e^{-t}]_0^A \xrightarrow{A \rightarrow \infty} 0$ par croissances comparées, et par hypothèse de récurrence, $-(n+1) \int_0^A t^n (-e^{-t}) dt \xrightarrow{A \rightarrow \infty} (n+1)I_n$. Donc I_{n+1} converge.

Finalement, par principe de récurrence, on a montré que I_n converge pour tout $n \in \mathbf{N}$ et $I_{n+1} = (n+1)I_n$.

2. Par récurrence évidente, on a $I_n = (n-1)!I_0 = \boxed{(n-1)!}$.

Exercice ANA.11.17 | Déterminer la nature des intégrales ci-dessous en effectuant le changement de variable donné.

1. $\int_0^{+\infty} \frac{dt}{\sqrt{1+e^t}}$ *Indication : On pourra poser $u = \sqrt{1+e^t}$.*
2. $\int_0^{+\infty} \frac{t^2-1}{(1+t^2)\sqrt{1+t^4}} dt$. *Indication : On pourra poser $t = \frac{1}{u}$.*
3. $\int_0^1 \sqrt{\frac{1+u}{1-u}} du$. *Indication : On pourra poser $u = \cos t$.*

Solution (exercice ANA.11.17)

1. On a $du = \frac{e^t}{2\sqrt{1+e^t}} dt$, ou de manière équivalente

$$\frac{dt}{\sqrt{1+e^t}} = \frac{2 du}{u^2-1}$$

Donc, puisque $t \mapsto \sqrt{e^t + 1}$ est strictement croissante et de classe \mathcal{C}^1 , $\int_0^{+\infty} \frac{dt}{\sqrt{1+e^t}}$ a même nature que

$$2 \int_{\sqrt{2}}^{+\infty} \frac{du}{u^2 - 1}.$$

Mais pour tout $A \in [\sqrt{2}, \infty[$, on a :

$$\begin{aligned} 2 \int_{\sqrt{2}}^A \frac{du}{u^2 - 1} &= \int_{\sqrt{2}}^A \left(\frac{1}{u-1} - \frac{1}{u+1} \right) du \\ &= [\ln|u-1|]_{\sqrt{2}}^A - [\ln|u+1|]_{\sqrt{2}}^A \\ &= \ln \left| \frac{u-1}{u+1} \right| + \ln \left(\frac{\sqrt{2}+1}{\sqrt{2}-1} \right) \\ &\xrightarrow{A \rightarrow \infty} \boxed{\ln \left(\frac{\sqrt{2}+1}{\sqrt{2}-1} \right)}. \end{aligned}$$

2. On a $dt = -\frac{du}{u^2}$, puisque la fonction inverse est de classe \mathcal{C}^1 et est strictement décroissante sur $]0, \infty[$, l'intégrale $\int_0^{+\infty} \frac{t^2 - 1}{(1+t^2)\sqrt{1+t^4}} dt$ a même nature que :

$$\int_0^{+\infty} \frac{t^2 - 1}{(1+t^2)\sqrt{1+t^4}} dt = - \int_{+\infty}^0 \frac{\left(\frac{1}{u}\right)^2 - 1}{\left(1 + \left(\frac{1}{u}\right)^2\right)\sqrt{1 + \left(\frac{1}{u}\right)^4}} \frac{du}{u^2}.$$

Après simplifications, on observe que $I = \int_0^{+\infty} \frac{t^2 - 1}{(1+t^2)\sqrt{1+t^4}} dt$ a même nature que $-I$ (mais on le savait déjà!), et en cas de convergence sont égales. Donc, en cas de convergence, $I = -I$ donc $\boxed{I = 0}$. Reste à montrer la convergence. On a pour tout $t \geq 0$:

$$0 \leq \frac{t^2 - 1}{(1+t^2)\sqrt{1+t^4}} \leq \frac{1}{\sqrt{1+t^4}} \leq \frac{1}{t^2}.$$

Or, $\int_1^{\infty} \frac{dt}{t^2}$ converge en calculant l'intégrale partielle, donc par théorème de comparaison $\int_1^{\infty} \frac{t^2-1}{(1+t^2)\sqrt{1+t^4}} dt$ converge. De plus, $\int_0^1 \frac{t^2-1}{(1+t^2)\sqrt{1+t^4}} dt$ converge puisque l'intégrande est continue sur cet intervalle, donc I converge bien.

3. $\int_0^1 \sqrt{\frac{1+u}{1-u}} du$. On a $du = -\sin t dt$, et $0 = \cos(\pi/2)$, $1 = \cos(0)$, or $\cos : \left[0, \frac{\pi}{2}\right] \rightarrow [0, 1]$ est \mathcal{C}^1 strictement monotone donc $\int_0^1 \sqrt{\frac{1+u}{1-u}} du$ a même nature que

$$\int_{\frac{\pi}{2}}^0 \sqrt{\frac{1+\cos t}{1-\cos t}} (-\sin t) dt.$$

Mais

$$1 + \cos t = 2 \cos^2 \left(\frac{t}{2} \right), \quad 1 - \cos t = 2 \sin^2 \left(\frac{t}{2} \right).$$

Donc on obtient comme nouvelle intégrale à étudier :

$$\int_{\frac{\pi}{2}}^0 \left| \frac{\cos\left(\frac{t}{2}\right)}{\sin\left(\frac{t}{2}\right)} \right| (-\sin t) dt = 2 \int_0^{\frac{\pi}{2}} \cos^2 \left(\frac{t}{2} \right) dt$$

où à la dernière étape nous avons utilisé $\sin(t) = 2 \sin\left(\frac{t}{2}\right) \cdot \cos\left(\frac{t}{2}\right)$. Puis enfin :

$$2 \int_0^{\frac{\pi}{2}} \cos^2 \left(\frac{t}{2} \right) dt = \int_0^{\frac{\pi}{2}} (\cos(t) + 1) dt,$$

qui donne après calculs de l'intégrale :

$$[\sin(t) + t]_0^{\frac{\pi}{2}} = \boxed{1 + \frac{\pi}{2}}.$$

On a donc obtenu la convergence et la valeur associée.

Exercice ANA.11.18 | Intégrales couplées On pose

$$I = \int_0^{+\infty} \frac{1}{1+t^4} dt \quad \text{et} \quad J = \int_0^{+\infty} \frac{t^2}{1+t^4} dt.$$

1. Montrer que ces intégrales convergent et que $I = J$.
2. En effectuant le changement de variable $x = t - \frac{1}{t}$ dans l'intégrale $I + J$, calculer la valeur commune de I et J .

Solution (exercice ANA.11.18)

1. L'intégrale I est généralisée en $+\infty$, et pour tout $t \in [1; +\infty[$, $\frac{1}{1+t^4} \leq \frac{1}{t^4}$. Or, $\int_1^{+\infty} \frac{1}{t^4} dt$ converge (refaire la démonstration avec l'intégrale partielle). Donc $I = \int_0^{+\infty} \frac{1}{1+t^4} dt$ converge par critère de comparaison pour les fonctions positives. On effectue ensuite un changement de variable, qui montrera à la fois l'égalité entre les deux intégrales, et la convergence de J . Plus précisément, posons $u = \frac{1}{t} \Leftrightarrow t = \frac{1}{u}$. Puisque la fonction $t \mapsto \frac{1}{t}$ est strictement monotone \mathcal{C}^1 de $[0, +\infty[$ vers $[0, +\infty[$, on a alors, sous réserve de convergence :

$$\begin{aligned} \int_0^{+\infty} \frac{dt}{1+t^4} &= \int_{+\infty}^0 -\frac{1}{u^2} \frac{du}{1+\frac{1}{u^4}} \\ &= \int_0^{+\infty} \frac{du}{u^2 + \frac{1}{u^2}} \\ &= J. \end{aligned}$$

Ainsi J est aussi convergente et $I = J$.

2. En tant que somme de deux intégrales convergentes, $I + J = \int_0^{+\infty} \frac{1+t^2}{1+t^4} dt$ converge. Faisons, comme suggéré par l'énoncé, le changement de variable $x = t - \frac{1}{t} \implies dx = \left(1 + \frac{1}{t^2}\right) dt = \frac{t^2+1}{t^2} dt$. Le changement est licite puisque la fonction $t \mapsto t - \frac{1}{t}$ est strictement monotone \mathcal{C}^1 de $[0, \infty[$ vers $]-\infty, 0]$. De plus $\frac{1+t^2}{1+t^4} = \frac{1+t^2}{t^2} \times \frac{1}{\frac{1}{t^2}+t^2}$ et $x^2 = \left(t - \frac{1}{t}\right)^2 = t^2 + \frac{1}{t^2} - 2 \Leftrightarrow t^2 + \frac{1}{t^2} = x^2 + 2$. Donc $\int_0^{+\infty} \frac{1+t^2}{1+t^4} dt = \int_{-\infty}^0 \frac{dx}{x^2+2} = \frac{1}{2} \int_{-\infty}^0 \frac{dx}{\left(\frac{x}{\sqrt{2}}\right)^2 + 1} = \frac{1}{2} \left[\sqrt{2} \arctan\left(\frac{x}{\sqrt{2}}\right) \right]_{-\infty}^0 = \frac{\pi}{\sqrt{2}}$. D'où $I + J = \frac{\pi}{\sqrt{2}}$. Or $I = J$, donc $I = J = \frac{\pi}{2\sqrt{2}}$.

Exercice ANA.11.19 | Soit f une fonction définie sur \mathbf{R}^+ et à valeurs dans \mathbf{R} , continue décroissante et telle que

$$\int_0^{+\infty} f(t) dt \text{ converge.}$$

1. Montrer que pour tout $x \geq 0$:

$$\int_x^{2x} f(t) dt \leq xf(x) \leq 2 \int_{x/2}^x f(t) dt.$$

2. Montrer que :

$$\lim_{x \rightarrow +\infty} \int_x^{2x} f(t) dt = 0.$$

3. En déduire que $\lim_{x \rightarrow +\infty} xf(x) = 0$.

Solution (exercice ANA.11.19)

1. Soit $x \geq 0$, alors puisque f est décroissante :

$$\forall t \in [x, 2x], \quad f(t) \leq f(x),$$

donc en intégrant :

$$\int_x^{2x} f(t) dt \geq (2x - x) \cdot f(x) = xf(x).$$

Pour le majorant, on écrit :

$$\forall t \in \left[\frac{x}{2}, x\right], \quad f(x) \leq f(t),$$

donc en intégrant :

$$\int_{x/2}^x f(x) dt = \frac{x}{2} f(x) \leq \int_{x/2}^x f(t) dt.$$

D'où en réunissant les deux inégalités :

$$\int_x^{2x} f(t) dt \leq xf(x) \leq 2 \int_{x/2}^x f(t) dt.$$

2. Il suffit d'écrire :

$$\int_x^{2x} f(t) dt = \int_0^{2x} f(t) dt - \int_0^x f(t) dt \xrightarrow{x \rightarrow \infty} \int_0^{\infty} f(t) dt - \int_0^{\infty} f(t) dt = 0.$$

3. On a d'après la question précédente et par composition des limites la convergence suivante :

$$\int_{x/2}^x f(t) dt \xrightarrow{x \rightarrow \infty} 0.$$

Donc $\lim_{x \rightarrow +\infty} x f(x) = 0$ d'après le théorème d'encadrement appliqué à la première question.

Exercice ANA.11.20 | Lemme de RIEMANN-LEBESGUE

1. **(Sur un segment)** Soient $a < b$ et $f \in \mathcal{C}^1([a, b], \mathbf{R})$. Montrer que :

$$\lim_{\lambda \rightarrow \infty} \int_a^b f(t) \cos(\lambda t) dt = 0.$$

2. **(Sur \mathbf{R}^+)** Soit $f \in \mathcal{C}^1(\mathbf{R}^+, \mathbf{R})$ telle que

$$\int_0^{\infty} |f'(t)| dt \text{ converge, } f \text{ est bornée.}$$

Montrer que :

$$\lim_{\lambda \rightarrow \infty} \int_0^{\infty} f(t) \cos(\lambda t) dt = 0.$$

3. Les résultats précédents vous semblent-ils maintenus si cos est remplacé par sin ?



Troisième partie

Aléatoire

Chapitre ALEA.12.

Dénombrement, Espaces probabilisés, Variables aléatoires

Résumé & Plan

L'objectif de ce chapitre est de lever certaines restrictions du programme de première année, à savoir la finitude des univers d'expériences aléatoires. En effet, il vous était impossible en première année de traduire l'expérience du lancer infini de pièce à l'aide d'un espace probabilisé.

Nous allons donc maintenant définir un cadre général pour les expériences aléatoires, tel que décrit par Andreï NIKOLAIEVITCH KOLMOGOROV (largement adopté par la communauté mathématique à partir de 1950) :

- ▶ un univers Ω contenant les données brutes d'une expérience aléatoire,
- ▶ une tribu \mathcal{F} qui contient des parties de Ω dont on calculera la probabilité et qui sont les évènements liés à Ω ,
- ▶ et enfin une probabilité \mathbf{P} définie sur \mathcal{F} , qui doit donner une indication sur la vraisemblance d'un évènement.

Afin de disposer de tous les outils permettant de réaliser des calculs probabilistes, nous commençons par une section de révision de dénombrement.

1	Dénombrement	2
1.1	Cardinal d'un ensemble fini	4
1.2	Listes, Permutations, Combinaisons	7
2	Axiomatique des probabilités	11
2.1	Espace probabilisé	12

2.2	Univers & Espace probabilisable	13
2.3	Espace probabilisé	15
2.4	Résultat d'existence de probabilités	18
2.5	Conditionnement & Indépendance d'évènements	20
3	Variables aléatoires	26
3.1	Généralités	27
3.2	Fonction de répartition & Loi	28
4	Exercices	33
4.1	Dénombrement	33
4.2	Espaces probabilisés	36

Un être humain possède environ 150 000 cheveux (en tout cas, moins d'un million). Comme la ville de Paris compte 2,141 millions d'habitants, d'après le principe des tiroirs de DIRICHLET, il existe au moins deux personnes à Paris qui ont exactement le même nombre de cheveux.

— Le saviez-vous ?

On estime à environ 10^{120} le nombre de parties différentes possibles au jeu d'échecs. Ce nombre s'appelle le nombre de SHANNON.

— Le saviez-vous? ▶

Quels que soient les progrès des connaissances humaines, il y aura toujours place pour l'ignorance et par suite pour le hasard et la probabilité.

— Emile BOREL

$\frac{1}{6} \times 6 = 1$ euro tandis que les gains moyens sur les autres faces sont plus faibles : on vient de prouver que jouer le 6 est la meilleure stratégie. Le joueur vient de faire des probabilités.

En revanche, tout le raisonnement précédent est basé sur le caractère non pipé du dé. Les *statistiques* permettent de répondre à la question suivante : en se basant sur un grand nombre d'observations, peut-on en conclure, avec forte probabilité, que le dé est non pipé? Pour le vérifier, le joueur peut lancer avant de commencer à parier 100 fois le dé et remarque que le 6 ne sort pas une seule fois et que le 4 sort une fois sur deux... il se dit donc naturellement que le dé n'est pas équilibré et que le 4 a une chance sur deux de sortir. Le joueur vient de faire des statistiques, et cela change donc la stratégie à adopter.

Avant de formaliser les expériences aléatoires à l'aide d'espaces probabilisés, nous allons avoir besoin de revoir des éléments de dénombrement.

1. DÉNOMBREMENT

Il est indispensable, avant d'entamer la lecture de cette partie, de revoir votre chapitre de théorie des ensembles de première année (ensemble, application, réunion, intersection, propriétés, *etc.*...).

RAPPELS SUR LA NOTATION INDICATRICE.

Σ Notation Fonction indicatrice d'un ensemble.

Dans tout ce chapitre, nous nous servirons de la notation $\mathbb{1}_A$ où A est sous-ensemble d'un ensemble E . Il s'agit de l'application

$$\mathbb{1}_A : \begin{cases} E & \longrightarrow & \{0, 1\} \\ x & \longmapsto & \begin{cases} 1 & \text{si } x \in A, \\ 0 & \text{sinon.} \end{cases} \end{cases}$$

On rappelle également une propriété importante :

Définition ALEA.12.1

On appelle *expérience aléatoire* une expérience renouvelable, et qui, renouvelée dans des conditions identiques – pour autant que l'observateur puisse s'en assurer – ne donne pas forcément le même résultat à chaque renouvellement. L'ensemble des issues possibles d'une expérience aléatoire est appelé *univers*, on le notera en général Ω dans la suite.

PROBABILITÉS OU STATISTIQUES ? C'est le premier chapitre de Probabilités & Statistiques de l'année, précisions brièvement de quoi on parle au travers d'un exemple.

Un joueur parie sur le résultat du lancé d'un dé à 6 faces. S'il annonce le bon résultat, il gagne la valeur de la face en euros, s'il perd il ne gagne rien du tout.

- ▶ Les *probabilités* permettent de répondre à la question suivante : quelle stratégie le joueur doit-il adopter pour maximiser ses gains? Il s'agit de faire la remarque suivante : quelle que soit la face qu'il annonce, il a toujours une chance sur six de gagner. Par contre, les gains en cas de bonne réponse dépendent de la face : s'il annonce 1 il a une chance sur six de gagner 1 euros, s'il annonce 2 il a une chance sur six de gagner 2 euros *etc.*... On devine ainsi que le joueur a tout intérêt à parier sur le 6 : il gagnera aussi souvent qu'avec une autre face mais gagnera plus. Mathématiquement cela s'écrit en terme de gain moyen, c'est-à-dire les chances de gagner multipliées par le gain, cela représente ce que le joueur gagne en moyenne par partie. On voit alors que le gain moyen en pariant sur le 6 est de

Proposition ALEA.12.1 | Indicatrice d'une réunion et d'une intersection

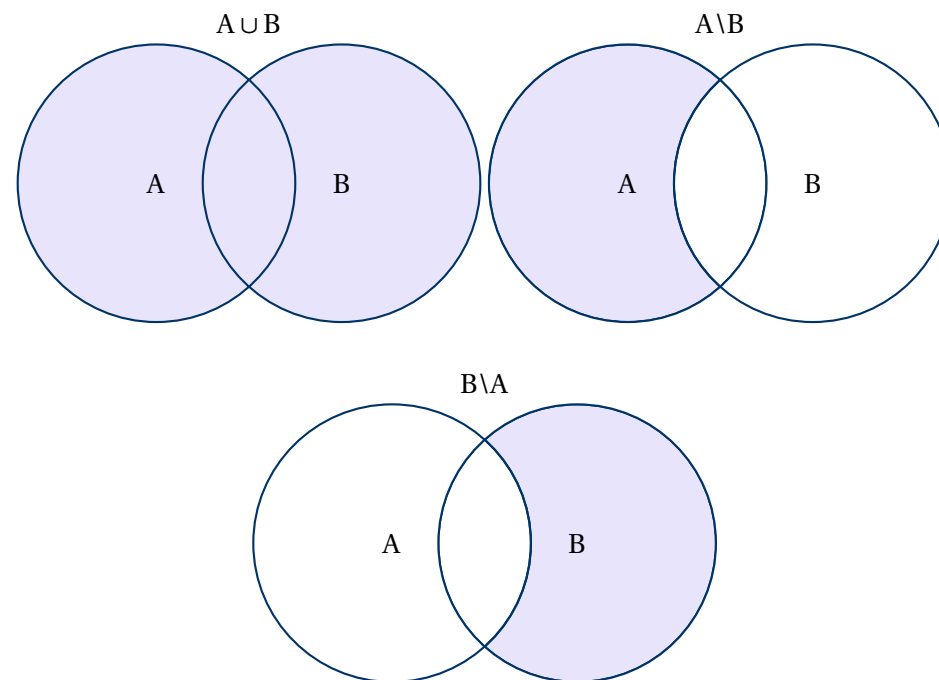
Soient $A, B \in \mathcal{P}(E)$ deux sous-ensembles d'un ensemble E . Alors :

1. $\mathbb{1}_{A \cap B} = \mathbb{1}_A \mathbb{1}_B$,
2. $\mathbb{1}_{A \cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_{A \cap B}$.

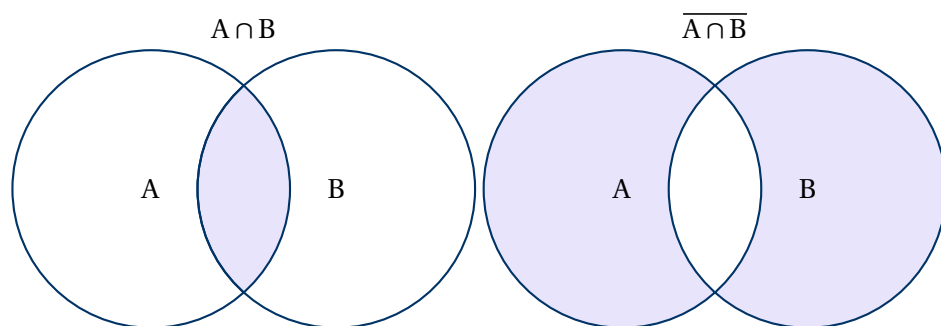
Preuve

1. 

2. Soit $x \in E$, alors on montre que $\mathbb{1}_{A \cup B}(x) = \mathbb{1}_A(x) + \mathbb{1}_B(x) - \mathbb{1}_{A \cap B}(x)$ en distinguant les cas $x \in A \setminus B$, $x \in B \setminus A$ puis $x \in A \cap B$.



RAPPELS SUR LA THÉORIE DES ENSEMBLES Soient A, B deux sous-ensembles d'un ensemble E . On rappelle les différentes opérations ensemblistes définies en première année au travers des diagrammes ci-dessous.



Ainsi, pour $x \in E$,

- ▶ $x \in A \cap B$ si : $x \in A$ **et** $x \in B$,
- ▶ $x \notin A \cap B$ si : $x \in A \setminus B$ **ou** $x \in B \setminus A$,
- ▶ $x \in A \cup B$ si : $x \in A$ **ou** $x \in B$,
- ▶ $x \in A \setminus B$ si : $x \in A$ **et** $x \notin B$. De-même pour $x \in B \setminus A$.

1.1. Cardinal d'un ensemble fini

Définition/Proposition ALEA.12.1

- ▶ On dit qu'un ensemble E est *fini* s'il existe une bijection de E sur $\llbracket 1, n \rrbracket$.
- ▶ L'entier n est unique, et est appelé le *cardinal* de E , noté $\# E$. Par convention, l'ensemble vide \emptyset est un ensemble fini de cardinal zéro : $\# \emptyset = 0$.

Plus simplement, le cardinal d'un ensemble fini est donc son nombre d'éléments. Intuitivement, les éléments d'un ensemble de cardinal n peuvent être *numérotés* de 1 à n . En pratique, on pourra, si besoin, écrire

$$E = \{x_1, \dots, x_n\}.$$

S'il existe une bijection entre deux ensembles E et F , alors E est un ensemble fini si et seulement si F est un ensemble fini, et dans ce cas, ils ont le même cardinal.

Dans la pratique, nous allons rarement compter les éléments d'un ensemble «à la main». On pourra souvent représenter le contexte à l'aide d'une «situation typique» du cours (des uplets, des applications injectives/surjectives, des choix d'éléments dans un ensemble avec ou sans ordre, distincts ou non *etc...*). Nous allons donc étudier toutes ces situations dans la suite du cours. Avantage : une fois qu'une telle situation est identifiée, il n'y aura plus rien à faire!¹

Exemple 1 –

- $\# \{5, 8, 12\} = 3$.
- Soit $(n, p) \in \mathbf{N}^2$ tel que $n \leq p$. Alors : $\# (\llbracket n, p \rrbracket) = p - n + 1$. C'est donc la différence des deux $\boxed{+1}$.

¹Si on connaît son cours, bien entendu

Proposition ALEA.12.2 | Cardinal d'une partie

- Soit E un ensemble fini et $F \subset E$. Alors :
- ▶ F est un ensemble fini, et : $\# F \leq \# E$.
 - ▶ De plus, $F = E \iff \# F = \# E$.

On remarquera l'analogie de cette proposition avec celle faisant intervenir des espaces vectoriels et la dimension.

Proposition ALEA.12.3 | Cardinal du complémentaire

- Soient E un ensemble fini et A une partie de E . Alors :
- ▶ $E \setminus A$ est un ensemble fini, et
 - ▶ $\# (E \setminus A) = \# E - \# A$.

PARTITIONNEMENT. On formalise ici la notion «découpage» d'ensemble en sous-parties disjointes.

Définition ALEA.12.2 | Partition

Soit E un ensemble. Alors une famille de parties $(A_n)_{n \in \mathbf{N}} \in \mathcal{P}(E)^{\mathbf{N}}$ est appelée *partition de E* si :

- ▶ pour tout $n \neq m$, $A_n \cap A_m = \emptyset$,
- ▶ $\bigcup_{n=0}^{\infty} A_n = E$.

On dit aussi que E est la *réunion disjointe des $A_n, n \in \mathbf{N}$* .



Notation Réunion disjointe

Lorsque les hypothèses précédentes sont satisfaites, on note $E = \bigsqcup_{n \in \mathbf{N}} A_n$.

Si E est un ensemble fini, alors tous les sous-ensembles $A_n, n \in \mathbf{N}$, qui précèdent sont finis. Et dans le cas fini, on obtient la formule de passage au cardinal ci-dessous.

Proposition ALEA.12.4 | Cardinal d'une réunion disjointe

Soient E un ensemble fini et $(A_n)_{1 \leq n \leq p}$, $p \in \mathbf{N}^*$, une partition de E . Alors :


$$\# E = \sum_{n=1}^p \# A_n.$$


Remarque 1.1 — de vocabulaire Lorsque $E = \Omega$ est un univers probabiliste, alors une partition est simplement un système complet d'évènements. Nous le reverrons bientôt.

Preuve (*Point clef — Récurrence sur p , le nombre d'éléments qui constituent la partie*)

Pour $p = 2$, on utilise la [Proposition ALEA.12.3](#) puisque $A_2 = E \setminus A_1$. Supposons la propriété vraie au rang p . 

Remarque 1.2 — Cette proposition est d'une importance cruciale pour donner une consistance à vos raisonnements de dénombrement. On écrit, lorsque c'est possible, notre ensemble de départ comme une réunion disjointe d'ensembles plus simples. Le passage au cardinal nous permet ensuite de conclure.

Exemple 2 — Soit $n \in \mathbf{N}^*$. Déterminer, en utilisant un cardinal de réunion, le nombre de couples $(x, y) \in \llbracket 1, n \rrbracket^2$ tels que $x + y = n$. Lister les couples possibles et retrouver le résultat précédent. 

Exemple 3 — Listes binaires sans termes consécutifs égaux Soit n un entier non nul. On désigne par u_n le nombre de listes — on tient donc compte de l'ordre — de n termes, chaque terme étant 0 ou 1, et n'ayant pas deux termes 1 consécutifs. Établir une relation de récurrence sur les termes de (u_n) . Exprimer alors u_n en fonction de $n \in \mathbf{N}^*$. 

En particulier, si E est un ensemble fini et $p \in \mathbf{N}^*$, alors :

$$\#(E^p) = \#E^p.$$

Preuve Pour $p = 0$, on suppose donc que $E = \emptyset$. Alors $\#(\emptyset \times F) = \#\emptyset = 0$.
Soit $p \in \mathbf{N}^*$. Alors notons $E = \{x_1, \dots, x_p\}$, on a : $E \times F = \bigcup_{i=1}^p \{e_i\} \times F$. C'est une partition du produit donc on a :

$$\#(E \times F) = \sum_{i=1}^p \#(\{e_i\} \times F) = \sum_{i=1}^p \#F = p\#F.$$

La formule est donc démontrée.

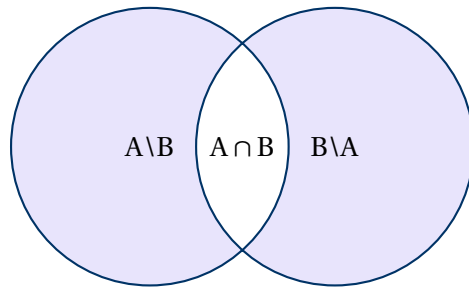
Exemple 4 — On suppose que l'on peut représenter une cellule par une suite de $N \in \mathbf{N}$ entiers entre -5 et 5 , et un entier qui vaut soit 0 soit 3 : le 0 correspondant à une cellule cancéreuse, et le 3 à une cellule saine. Combien y-a-t'il de configurations possibles?



Corollaire ALEA.12.1 | Cardinal d'une réunion

Soient E un ensemble fini et A et B deux parties de E . Alors :

$$\#(A \cup B) = \#A + \#B - \#(A \cap B).$$



Preuve La famille de parties $(A, (A \cup B) \setminus A)$ est une partition de $A \cup B$, de-même $(A \cap B, (A \cup B) \setminus A)$ est une permutation de B . En passant ensuite au cardinal, on obtient :

$$\#(A \cup B) = \#A + \#(A \cup B) \setminus A - \#A = \#(A \cap B) + \#(A \cup B) \setminus A.$$

Proposition ALEA.12.5 | Cardinal d'un produit cartésien.

Soient E et F deux ensembles finis. Alors, $E \times F$ est fini et :

$$\#(E \times F) = \#E \times \#F.$$

Plus généralement, soit $(E_k)_{1 \leq k \leq p}$ une famille finie d'ensembles finis. Alors,


- ▶ $E_1 \times \dots \times E_p$ est fini, et
- ▶ $\#(E_1 \times \dots \times E_p) = \prod_{k=1}^p \#E_k$.

Proposition ALEA.12.6 | Cardinal de l'ensemble des applications de E dans F .

Soient E et F deux ensembles finis, de cardinaux non nuls. Alors l'ensemble $F^E = \mathcal{F}(E, F)$ des applications de E dans F est fini et :

$$\# \left(F^E \right) = \# F^{\# E}.$$


Preuve Notons n le cardinal de E et nommons ses éléments : $E = \{x_1, \dots, x_n\}$. Une application $f : E \rightarrow F$ est entièrement déterminée par la famille $(f(x_1), \dots, f(x_n))$ d'éléments de F . Il y a donc autant d'applications $f : E \rightarrow F$ que d'éléments $(y_1, \dots, y_n) \in F^n$. Ainsi, il existe une bijection entre les ensembles F^E et F^n . D'après la ??, on a donc : $\# \left(F^E \right) = \# \left(F^n \right) = \# F^n = (\# F)^{\# E}$.

Exemple 5 – Rangement de boules discernables dans des tiroirs On considère 5 boules discernables que l'on veut placer dans 3 tiroirs distincts, chaque tiroir pouvant contenir de 0 à 5 boules. Donner le nombre de répartitions possibles.  Il s'agit d'affecter un tiroir pour chaque boule, soient 3^5 possibilités en tout. C'est aussi le nombre d'applications d'un ensemble à 5 éléments vers un ensemble à 3 éléments.

Proposition ALEA.12.7 | Cardinal de l'ensemble des parties.

Soit E un ensemble fini. Alors l'ensemble $\mathcal{P}(E)$ des parties de E est fini et :

$$\# \mathcal{P}(E) = 2^{\# E}.$$

Preuve Se donner une partie A de E c'est se donner une application de E dans $\{0, 1\}$ où $x \in E$ est associé par exemple à 1 si $x \in A$ et 0 sinon. 

1.2. Listes, Permutations, Combinaisons

La plupart des exercices de dénombrement peuvent se ramener au cas de tirages de p éléments parmi les n éléments d'un ensemble E . Il y a alors essentiellement quatre façons différentes de tirer p éléments parmi n :

- ▶ avec ordre et répétition (les p -listes d'éléments distincts dans la suite),
- ▶ avec ordre et sans répétition (les p -listes d'éléments quelconques dans la suite),
- ▶ sans ordre et sans répétition (les p -combinaisons d'éléments distincts dans la suite),
- ▶ sans ordre et avec répétition (les p -combinaisons d'éléments quelconques dans la suite).

La présence d'un ordre ou pas sera fixée par le choix de l'objet mathématique que l'on compte : des uplets pour les éléments ordonnés $((1, 2) \neq (2, 1))$, et des ensembles pour les éléments non ordonnés $(\{1, 2\} = \{2, 1\})$.

1.2.1. Nombre de p -listes d'un ensemble fini : avec ordre

Définition ALEA.12.3 | p -listes

Soient E un ensemble et $p \in \mathbb{N}^*$. On appelle p -uplet ou p -liste d'éléments de E tout élément (x_1, \dots, x_p) de E^p . On parle aussi de *couples* pour $p = 2$, de *triplets* pour $p = 3$, de *quadruplets* pour $p = 4$.

Attention

Par exemple pour $p = 2$, $(x_1, x_2) \neq (x_2, x_1)$ dès que $x_1 \neq x_2$. Ainsi, on tient compte de l'ordre des éléments pour les p -listes.

Proposition ALEA.12.8 | Nombre de p -listes

Soient E un ensemble fini de cardinal $n \in \mathbb{N}^*$ et $p \in \mathbb{N}^*$. Le nombre de p -listes d'éléments de E est n^p .

En probabilités, ce type de cardinal interviendra dans des expériences de tirages effectués **avec remise**.

Preuve



Exemple 6 — En ne supposant aucune contrainte sur les séries de chiffres, combien de numéros de cartes bancaires existe-t-il? Et de codes d'identification?

On peut maintenant également se poser la question de la recherche du cardinal de ces mêmes listes, mais lorsque tous les éléments sont distincts.

Proposition ALEA.12.9 | Nombre de p -listes d'éléments distincts

Soient E un ensemble fini de cardinal n et $p \in \mathbf{N}^*$. Le nombre de p -listes d'éléments **distincts** de E est :

$$\begin{cases} 0 & \text{si } p > n, \\ \frac{n!}{(n-p)!} & \text{si } p \leq n. \end{cases}$$

En probabilités, ce type de cardinal interviendra dans des expériences de tirages effectués **sans remise**.

Preuve Le résultat est évident dans le premier cas. Supposons $p \leq n$. La construction d'un p -uplet (x_1, \dots, x_p) d'éléments distincts de E peut se faire ainsi : choisir x_1 parmi les n éléments de E , puis x_2 parmi les $n-1$ éléments de $E \setminus \{x_1\}$, puis x_3 parmi les $n-2$ éléments de $E \setminus \{x_1, x_2\}$, etc. Le nombre de p -uplets d'éléments distincts est alors bien :

$$n(n-1) \dots (n-p+1) = \frac{n!}{(n-p)!}.$$

Corollaire ALEA.12.2 | Nombre d'injections

Soient E un ensemble fini de cardinal p et F un ensemble fini de cardinal n . Le nombre d'injections (*i.e.* d'applications injectives) de E dans F est :

$$\begin{cases} 0 & \text{si } p > n, \\ \frac{n!}{(n-p)!} & \text{si } p \leq n. \end{cases}$$

Preuve L'image d'une application injective est une p -liste d'éléments distincts dans un ensemble à n éléments, et ces deux ensembles ont même cardinal.

CAS PARTICULIER : LE NOMBRE DE PERMUTATIONS D'UN ENSEMBLE FINI. Rappelons tout d'abord la définition d'une permutation.

Définition ALEA.12.4

On appelle *permutation* d'un ensemble fini E toute bijection de E dans E .

Σ Notation

Si E est un ensemble fini, on note \mathfrak{S}_E l'ensemble des permutations de E . Si $E = \llbracket 1, n \rrbracket$ avec $n \in \mathbf{N}^*$, on note plus simplement \mathfrak{S}_n cet ensemble.

Si $E = \llbracket 1, n \rrbracket$, une application $\sigma \in \mathfrak{S}_n$ peut être écrite comme une application clas-

sique

$$\sigma \begin{cases} E \longrightarrow E, \\ n \longrightarrow \sigma(n), \end{cases}$$

ou encore sous la forme d'une liste $(\sigma(1), \dots, \sigma(n))$ — c'est simplement la liste des images de chaque entier. On dit qu'une permutation *s'identifie* à une liste.² De plus, si σ est bijective, les $\sigma(i)$, $i \in \llbracket 1, n \rrbracket$, sont forcément tous distincts.

Corollaire ALEA.12.3 | Nombre de permutations

Soit E un ensemble fini de cardinal n . Le nombre de permutations de E est

$$\mathfrak{S}_E = \# \mathfrak{S}_n = n!.$$

Preuve Puisque E est fini, $\sigma : E \rightarrow E$ est une permutation si et seulement si elle est injective. Or, d'après le corollaire précédent, le nombre d'applications injectives de E dans E est $\frac{n!}{0!} = n!$.

On peut aussi refaire le calcul à la main. 

Exemple 7 —

- ▶ De combien de façons peut-on asseoir n personnes sur un banc rectiligne? 

²Mais attention, ce sont deux objets de natures très différentes.

- ▶ Autour d'une table ronde? 

Exemple 8 — Si une classe est constituée de 48 étudiants et si la salle de classe comporte exactement 48 places assises, alors il y a $48!$ dispositions différentes possibles, soit environ $12 \cdot 10^{60}$.

1.2.2. Nombre de p -combinaisons d'un ensemble fini : sans ordre

Définition ALEA.12.5

Soient E un ensemble et $p \in \mathbf{N}$. On appelle p -combinaison (ou p -ensemble) d'éléments de E , toute partie $\{x_1, \dots, x_p\}$ de E de cardinal p .

⊗ Attention

Par exemple pour $p = 2$, $\{x_1, x_2\} = \{x_2, x_1\}$. Ainsi, on ne tient pas compte de l'ordre des éléments pour les p -combinaisons.

ÉLÉMENTS DISTINCTS.

Proposition ALEA.12.10 | Nombre de p -combinaisons

Soit E un ensemble fini de cardinal n . Le nombre de p -combinaisons d'éléments **distincts** de E est :

$$\binom{n}{p} \underset{\text{(nota.)}}{=} \begin{cases} 0 & \text{si } p > n \text{ ou } p < 0, \\ \frac{n!}{p!(n-p)!} & \text{si } 0 \leq p \leq n. \end{cases}$$

Preuve



Exemple 9 — Loto Remplir une grille de loto consiste à cocher 5 cases parmi 49. Le nombre de combinaisons possibles est donc $\binom{49}{5} = 1\,906\,884$, dont une seule est gagnante. Au loto, on ne tient pas compte de l'ordre des éléments.

Une partie de l'exemple qui suit, très classique, utilise plutôt des permutations.

Exemple 10 — Anagrammes d'un mot On appelle *anagramme d'un mot* toute permutation de l'ensemble des lettres. Dans la pratique, il faut bien distinguer deux cas, puisque certaines permutations peuvent donner *in fine* le même mot, il ne faut donc pas le compter deux fois!

1. (Cas de lettres distinctes) quel est le nombre d'anagrammes du mot CHEVAL?



2. (Cas de lettres répétées) quel est le nombre d'anagrammes du mot ANANAS?



On essaye de remplir les 6 emplacements lettres par lettres (comme au jeu du pendu). Commençons par choisir la place des A : on doit choisir 3 emplacements parmi 6, et ce sans ordre (car l'ordre des A n'a pas d'importance, ce sont les mêmes lettres), et sans répétition car on ne peut pas placer deux lettres sur le même emplacement. On a donc des combinaisons de 3 lettres parmi 6 emplacements, soit $\binom{6}{3}$ possibilités. On doit ensuite placer les 2 N parmi les 3 emplacements restants, on a donc $\binom{3}{2}$ possibilités. Enfin, il ne reste qu'une possibilité pour la place du S. Ainsi, on a $\binom{6}{3} \times \binom{3}{2} = 60$ possibilités.

Le nombre de p -combinaisons est inférieur au nombre de p -listes d'un ensemble à n éléments.³

Proposition ALEA.12.11 | Propriétés des coefficients binomiaux

Soient $n \in \mathbf{N}$ et $p \in \mathbf{Z}$. Les coefficients binomiaux vérifient les propriétés suivantes :

1. (Symétrie) $\binom{n}{p} = \binom{n}{n-p}$.

2. (Formule de PASCAL) $\binom{n}{p} + \binom{n}{p+1} = \binom{n+1}{p+1}$.

3. (Formule du binôme de NEWTON) Soient $(a, b) \in \mathbf{C}^2$ et $n \in \mathbf{N}$. Alors :

$$(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}.$$

Preuve Démontrons ces propriétés de manière combinatoire.

1. Par passage au complémentaire, il est clair que le nombre de p -combinaisons d'un ensemble de cardinal n est égal au nombre de $(n-p)$ -combinaisons de cet ensemble. Le résultat découle alors.


2. Considérons le cas non trivial où $0 \leq p \leq n-1$ — pour le reste, c'est une simple vérification. Soit E un ensemble fini de cardinal $n+1$. Fixons un élément $a \in E$. Les $(p+1)$ -combinaisons de E se classent en deux parties : celles contenant a (il

³Ceci est tout à fait cohérent avec l'intuition, si on tient compte de l'ordre, il y a plus de possibilités.

y en $\binom{n}{p}$ et celles ne contenant pas a (il y en a $\binom{n}{p+1}$). Comme ces deux classes de parties sont d'intersection vide, on somme leurs cardinaux pour trouver 2).

3. Les termes issus du développement du produit $(a + b)^n$ sont obtenus en choisissant, un certain nombre de puissances de a et de puissances de b tel que la somme des puissances soit égale à n . Plus précisément, notons $k \in \llbracket 0, n \rrbracket$ la puissance de a d'un des termes, celle de b est alors $n - k$. Il y a par ailleurs $\binom{n}{k}$ -façons d'obtenir le terme $a^k b^{n-k}$ — c'est le nombre de façons de choisir k fois a dans le produit $(a + b) \dots (a + b)$. D'où la formule.

ELÉMENTS QUELCONQUES. Ce cas là est plus rare mais il apparaît parfois. Rappelons que nous avons traité l'exemple ci-dessous mais avec des boules discernables.

Exemple 11 — Rangement de boules indiscernables dans des tiroirs On considère 5 boules indiscernables⁴ que l'on veut placer dans 3 tiroirs distincts, chaque tiroir pouvant contenir de 0 à 5 boules. Donner le nombre de répartitions possibles. Dessinons une configuration. 

Cela revient donc à compter le nombre de façons de poser les 5 (boules) + 2 (cloisons internes) = 7 objets. On imagine alors qu'il y a 7 emplacements à remplir, donc le nombre de configurations est le nombre de choix pour les cloisons (cf. des boules,

⁴Le cas de boules indiscernables est plus complexe que le cas discernable, car on peut avoir trois boules dans le 1er tiroir *via* plusieurs configurations (mettre les deux premières que l'on choisit, ou la 1ère et la dernière que l'on choisit par exemple...).

cela mènera au même résultat), soient $\binom{7}{2}$ configurations possibles. En effet, il faut faire ces choix de manière non répétée (on ne met pas deux objets au même endroit) et puisque les cloisons (cf. boules) sont identiques, on ne tient pas compte de l'ordre des emplacements.

Les rappels de dénombrement étant terminés, passons à l'axiomatique probabiliste.

Résumé

On essaie toujours de représenter un exercice à l'aide d'un objet de dénombrement du cours (application, liste, ensemble, *etc.*). Dans un raisonnement dénombrement, la phrase

- ▶ «soit ça soit ça» donnera une addition de cardinal — une partition de l'ensemble est cachée,
- ▶ «ça puis ça» donnera une multiplication de cardinal.

2.

AXIOMATIQUE DES PROBABILITÉS

Certaines réalités doivent être modélisées par des objets mathématiques que nous qualifierons d'*aléatoires*, qui dépendent en fait de la réalisation d'un contexte précis. Voyons pourquoi.

DE L'ALÉATOIRE, POURQUOI? L'aléatoire est présent dans toute expérience scientifique. Les deux grandes explications en sont :

- ▶ d'une part l'aléatoire «intrinsèque» lié à la complexité des individus et des phénomènes étudiés et au manque d'information dans le domaine. Exemple : vous observez une particule, qui se balade de manière anarchique sur la droite réelle,

on essaye d'observer et de trouver une loi de déplacement, l'aléatoire est donc dû au manque d'informations sur cette loi *a priori* inconnue.

- ▶ D'autre part, de l'aléatoire peut intervenir « expérimentalement », par mesures entachées d'erreur, ou encore lorsque les moyens pour relever sont limités. Exemple 1 : un relevé expérimental en physique, il y a une erreur/incertitude due à la machine elle-même. Exemple 2 : vous voulez savoir si un vaccin contre la COVID-19 est efficace, vous pouvez le tester sur un échantillon, mais aussi grand soit-il, un aléatoire sera toujours présent (à moins de le tester à l'échelle de la planète, ce qui est bien entendu irréalisable dans la pratique).

La première source d'aléatoire est donc le manque d'informations, ou l'ignorance, la seconde est matérielle et peut donc être contrôlée et donnée elle-même lieu à une étude plus poussée. Il n'y a pas de raison de penser qu'une source d'aléatoire peut complètement être supprimée. En Biologie (ainsi qu'en Médecine), étant donné l'extrême complexité des systèmes étudiés, l'aléatoire est très présent.

2.1. Espace probabilisé

Les premières formalisations de la notion de hasard, au XVII^{ème} siècle, répondaient pour l'essentiel à diverses questions issues de la théorie des jeux. Depuis la publication en 1933 des Fondements de la théorie des probabilités d'Andreï KOLMOGOROV, les probabilités sont solidement ancrées sur la formalisation proposée par KOLMOGOROV.

Afin de donner un sens intuitif au concept d'espace probabilisé (les variables aléatoires seront vues plus tard), considérons par exemple les expériences aléatoires ci-dessous.

1. **(Expérience 1)** L'expérience d'un lancer de dé à 6 faces peut conduire à 6 résultats selon la face obtenue.
2. **(Expérience 2)** L'expérience d'un nombre infini de lancers d'une pièce amenant à pile «P» ou face «F».

- ▶ **(L'univers Ω)** Un ensemble qui décrit les résultats bruts de l'expérience. Dans l'expérience 1, on a $\Omega = \{1, \dots, 6\}$. Les résultats de l'expérience 2 peuvent être considérés comme des suites

$$(u_n) = (u_1, \dots, u_n \dots), \quad \forall i, \quad u_i \in \{P, F\}.$$

Donc $\Omega = \{P, F\}^{\mathbb{N}}$.

- ▶ **(Une tribu \mathcal{F} sur Ω)** C'est un sous-ensemble de parties de Ω qui correspond à ce que l'on considère comme observable. Dans l'expérience 1, tout est observable, donc $\mathcal{F} = \mathcal{P}(\Omega)$ décrit bien la situation. En revanche, dans l'expérience 2, un observateur humain ne peut relever qu'un nombre fini de lancers pour des raisons évidentes. Donc \mathcal{F} contiendra au moins des événements de la forme

$$A_i \text{ «le lancer } i \text{ a donné P», } \quad i \geq 0,$$

ainsi que leur complémentaire réunion intersection. Les parties de Ω qui sont dans \mathcal{F} seront appelées les *événements*.

- ▶ **(Une probabilité P sur (Ω, \mathcal{F}))** Une application qui donne un indicateur dans $[0, 1]$ de la vraisemblance d'un événement (*i.e.* les chances qu'à une observation de survenir). Dans l'expérience 1, puisque le dé est supposé non truqué, l'application

$$P : \omega \in \Omega \longmapsto \frac{1}{6}$$

décrit bien la situation. Ou plus généralement

$$P : A \in \mathcal{F} = \mathcal{P}(\Omega) \longmapsto \frac{\# A}{6}.$$

Dans l'exemple 2, on poserait si la pièce est non truquée :

$$P : A_i \in \mathcal{F} \longmapsto \frac{1}{2}.$$

On admet l'existence d'une probabilité, qui est non triviale dans l'Exemple 2.

Remarque 2.1 – Complément sur les tribus La notion de tribu ne paraît pas ici indispensable, et pourtant elle joue un rôle fondamental dans plusieurs contextes importants en Mathématiques :

1. la construction de l'intégrale de LEBESGUE : qui unifie l'intégrale de RIEMANN, les séries et d'autres choses, et qui fournit une définition générale de l'espérance.
2. Les phénomènes d'évolution, où l'information disponible peut dépendre du temps et donc on est amené à considérer des suites ou familles de tribus (on parle dans ce cas de *filtration*).

Bref, à notre niveau, il convient simplement de voir le concept de tribu comme de l'information disponible et accessible par observation, comme cela a été exposé précédemment.

En 1ère année les univers étaient toujours finis, ce qui était bien sûr restrictif (l'expérience 2 ne rentre pas dans ce cadre).

2.2. Univers & Espace probabilisable

Nous allons maintenant devoir imposer certaines propriétés sur les tribus et probabilités afin d'obtenir un cadre raisonnable permettant de faire des calculs. Voyons lesquels.

Définition ALEA.12.6 | Tribu sur un univers Ω

Soit Ω un ensemble appelé *univers*. On appelle *tribu* (ou σ -*algèbre*) sur Ω un sous-ensemble \mathcal{F} des parties de Ω vérifiant :

1. $\emptyset \in \mathcal{F}$.
2. (Stabilité par passage au complémentaire) $A \in \mathcal{F} \implies {}^cA \in \mathcal{F}$,
3. (stabilité par réunion dénombrable) si $(A_n)_{n \in \mathbb{N}} \in \mathcal{F}^{\mathbb{N}} \implies \bigcup_{n=0}^{\infty} A_n \in \mathcal{F}$.⁵

Le couple (Ω, \mathcal{F}) est alors appelé *espace probabilisable*. Les éléments de Ω sont appelés les *résultats* (ou *issues*, *éventualités*), notés généralement ω . Les éléments de \mathcal{F} sont appelés les *événements*, l'ensemble \emptyset est appelé *événement impossible* et Ω *événement quasi-certain*.

⁵Notez que l'on aurait pu (en utilisant 3)) remplacer l'axiome 1) par $\Omega \in \mathcal{F}$. À partir des trois axiomes

Attachons certaines propriétés à la notion d'univers.

Définition ALEA.12.7 | Ensemble dénombrable

Un univers Ω d'un $(\Omega, \mathcal{F}, \mathbf{P})$ est dit :

1. *fini* si l'ensemble Ω est un ensemble fini,
2. *dénombrable* s'il est en bijection avec \mathbb{N} ,
3. *au plus dénombrable* s'il est fini ou dénombrable.

Un univers dénombrable est donc en particulier de cardinal non fini. Tout ensemble dénombrable Ω peut s'écrire donc sous la forme « numérotée » suivante :

$$\Omega = \{\omega_i, i \in \mathbb{N}\}.$$

Remarque 2.2 — L'axiome 3) est légitimé en considérant n'importe quelle expérience aléatoire mettant en jeu une série infinie d'actions. Imaginons qu'un joueur lance une pièce de monnaie et qu'il gagne s'il obtient trois fois de suite pile, si A_n « le joueur obtient, pour la première fois, trois fois de suite pile au n -ième lancer », alors l'évènement « le jour gagne » est $\bigcup_{n \in \mathbb{N}} A_n$.

Donnons sans plus tarder quelques exemples d'espaces probabilisables.

Exemple 12 — *Exemples d'univers* Le résultat d'une expérience aléatoire peut prendre des formes variées :

1. le tirage de trois cartes dans un jeu de 32. 

principaux, on peut en déduire d'autres, qui auraient pu être pris d'ailleurs comme définition.

2. Lancer d'une pièce jusqu'à obtenir pile : $\Omega = \{P, FP, FFP, FFFP, \dots\}$.
3. Lancer infini d'une pièce : $\Omega = \{P, F\}^{\mathbb{N}}$ (non dénombrable cette fois-ci),
4. Durée de vie d'une ampoule : $\Omega = \mathbf{R}^{++}$ (non dénombrable),
5. Jeu de fléchettes : $\Omega = \{(x, y) \in \mathbf{R}^2, x^2 + y^2 \leq R^2\}$ (non dénombrable), avec $R \in \mathbf{R}^*$.

Exemple 13 — Exemples de tribus On peut munir un univers Ω des tribus ci-dessous.

- ▶ $\{\emptyset, \Omega\}$ appelée *tribu grossière*,
- ▶ l'ensemble $\mathcal{P}(\Omega)$ des parties de Ω , appelée *tribu discrète*,
- ▶ si $A \in \mathcal{P}(\Omega)$ alors la *tribu engendrée par A* notée $\sigma(A)$ est $\{\emptyset, A, {}^cA, \Omega\}$, on montre sans difficulté qu'il s'agit de la plus petite tribu contenant A.⁶

Proposition ALEA.12.12

Soit (Ω, \mathcal{T}) un espace probabilisable. Alors

1. $\Omega \in \mathcal{T}$,
2. (**\mathcal{T} est stable par réunion finie**) c'est-à-dire que pour toute famille finie $A_1, \dots, A_N \in \mathcal{T}$, $N \geq 1$, on a : $\bigcup_{i=1}^N A_i \in \mathcal{T}$.
3. (**\mathcal{T} est stable par intersection dénombrable**) pour toute suite $(A_n)_{n \in \mathbb{N}} \in \mathcal{T}^{\mathbb{N}}$, on a : $\bigcap_{n \in \mathbb{N}} A_n \in \mathcal{T}$. En particulier, elle est aussi stable par intersection finie.
4. (**\mathcal{T} est stable par différence**) Si $(A, B) \in \mathcal{T}^2$, alors : $A \setminus B = A \cap {}^cB \in \mathcal{T}$.

Preuve

1. Comme $\emptyset \in \Omega$, vu que \mathcal{T} est stable par passage au complémentaire, ${}^c\emptyset = \Omega \in \mathcal{T}$.
2. On sait que \mathcal{T} est stable par réunion dénombrable. L'idée est qu'une réunion finie est en particulier dénombrable en la prolongeant par le vide. Considérons $A_1, \dots, A_N \in \mathcal{T}$. Posant $A_k = \emptyset$ pour tout $k \geq N + 1$, on a

$$\bigcup_{k=1}^{\infty} A_k = \left(\bigcup_{k=1}^N A_k \right) \cup \emptyset = \bigcup_{k=1}^N A_k \in \mathcal{T}.$$

⁶De manière plus générale, étant donnée une partie \mathcal{P} de $\mathcal{P}(\Omega)$, on peut même montrer qu'il existe toujours une plus petite tribu sur Ω contenant \mathcal{P} .

3. Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'éléments de \mathcal{T} . Alors $({}^cA_n)_{n \in \mathbb{N}}$ est aussi une suite d'éléments de \mathcal{T} . D'après 2), on obtient :

$${}^c\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \bigcap_{n \in \mathbb{N}} {}^cA_n = \bigcap_{n \in \mathbb{N}} A_n \in \mathcal{T}.$$

4. Si $A, B \in \mathcal{T}$, alors cB est aussi dans \mathcal{T} , donc d'après 3), on obtient : $A \cap {}^cB \in \mathcal{T}$.

Définition ALEA.12.8 | Système complet d'évènements

Soient Ω un ensemble et $(A_n)_{n \in \mathbb{N}}$ une famille de parties de Ω . On dit que $(A_n)_{n \in \mathbb{N}}$ est un *système complet d'évènements* si $(A_n)_{n \in \mathbb{N}}$ est une partition de Ω , i.e. si :

- ▶ $\bigcup_{n=0}^{\infty} A_n = \Omega$,
- ▶ les A_n sont deux à deux disjoints, i.e. pour tous $i \neq j$ deux entiers, on a : $A_i \cap A_j = \emptyset$.

La notion de système complet d'évènements est en fait une notion purement ensembliste (appelée partition dans le début du chapitre). Dans un contexte plutôt probabiliste (i.e. lorsque l'on partitionne un univers Ω provenant d'une expérience aléatoire), on parle plutôt de système complet d'évènements. C'est juste une question de vocabulaire.

Autrement dit, la famille $(A_n)_{n \in \mathbb{N}}$ est une partition de Ω si pour tout $\omega \in \Omega$, il existe un et un seul $n \in \mathbb{N}$ tel que $\omega \in A_n$.

Exemple 14 —

- ▶ Pour tout $A \subset \Omega$, l'ensemble $\{A, {}^cA\}$ est un système complet d'évènements.
- ▶ Si $\Omega = \{1, 2, \dots, 6\}$, alors (A, B) où $A = \langle \text{faire pair} \rangle = \{2, 4, 6\}$, $B = \langle \text{faire impair} \rangle = \{1, 3, 5\}$ est un système complet d'évènements.

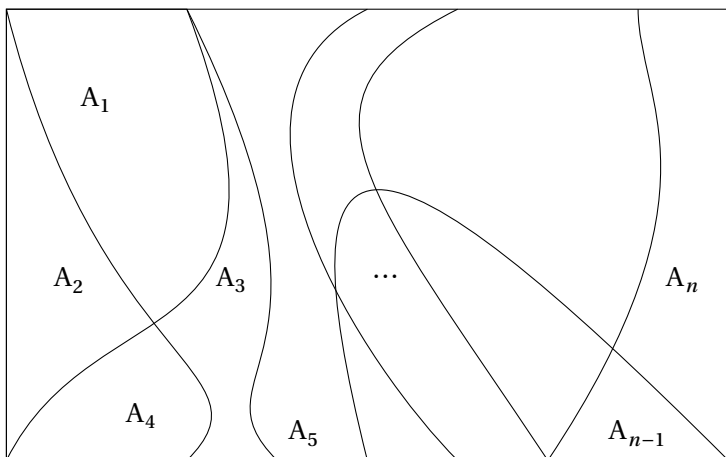


Fig. ALEA.12.1. : Représentation d'un système complet d'évènements

2.3. Espace probabilisé

Définition ALEA.12.9 | Probabilité

Soient un ensemble Ω et une tribu \mathcal{T} sur Ω , on appelle *probabilité* (ou plus simplement *probabilité*) sur (Ω, \mathcal{T}) une application $\mathbf{P} : \mathcal{T} \rightarrow [0, 1]$ telle que les propriétés ci-dessous soient vérifiées :

- ▶ **(Additivité dénombrable ou σ -additivité)** pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'éléments de \mathcal{T} deux à deux disjoints (*i.e.* $A_n \cap A_m = \emptyset$ pour tout $n \neq m$)

$$\sum_{n=0}^{\infty} \mathbf{P}(A_n) \text{ converge, } \mathbf{P}\left(\bigcup_{n=0}^{\infty} A_n\right) = \sum_{n=0}^{\infty} \mathbf{P}(A_n).$$

- ▶ $\mathbf{P}(\Omega) = 1$.

On appelle *espace probabilisé* un triplet $(\Omega, \mathcal{T}, \mathbf{P})$ où \mathcal{T} est une tribu sur Ω , et \mathbf{P} une probabilité sur (Ω, \mathcal{T}) . En particulier, si $(A_n)_{n \in \mathbb{N}}$ est un système complet d'évènements, alors : $\sum_{n=0}^{\infty} \mathbf{P}(A_n) = 1$.

Remarque 2.3 – La définition est bien posée Il y a une petite subtilité dans l'égalité

$$\mathbf{P}\left(\bigcup_{n=0}^{\infty} A_n\right) = \sum_{n=0}^{\infty} \mathbf{P}(A_n)$$

de la **Définition ALEA.12.9**. Étant donné que $\bigcup_{n=0}^{\infty} A_n$ est toujours invariante par permutation des A_n (vérification facile avec la définition d'une réunion), il faudrait justifier que c'est le cas aussi de la somme $\sum_{n=0}^{\infty} \mathbf{P}(A_n)$. Dans le cas contraire, on pourrait avoir plusieurs valeurs en fonction de l'ordre en lequel on somme, et l'hypothèse n'est pas bien posée.

Pour justifier ceci, il nous suffit d'appliquer le **Théorème ANA.10.3** du **Chapter ANA.10** et donc de justifier que $(\sum_{n=0}^N \mathbf{P}(A_n))$ converge absolument. En effet, c'est bien le cas, car pour tout entier N , on a :

$$\sum_{n=0}^N |\mathbf{P}(A_n)| = \sum_{n=0}^N \mathbf{P}(A_n) = \mathbf{P}\left(\bigcup_{n=0}^N A_n\right) \leq 1.$$

La série précédente étant une série à termes positifs de somme partielle majorée, elle converge absolument.

Définition ALEA.12.10 | Certitude & Négligence d'un évènement

Soit $(\Omega, \mathcal{T}, \mathbf{P})$ un espace probabilisé.

- ▶ Un évènement A tel que $\mathbf{P}(A) = 1$ est dit **\mathbf{P} -quasi-certain**, on dit aussi que A a lieu **\mathbf{P} -presque sûrement**, « p.s. » en abrégé.
- ▶ Un évènement A vérifiant $\mathbf{P}(A) = 0$ est dit **\mathbf{P} -négligeable** (ou **\mathbf{P} -quasi-impossible** parfois).

Définition ALEA.12.11 | Système quasi-complet

Soient Ω un ensemble. Considérons $(A_n)_{n \in \mathbb{N}}$ une famille de parties de Ω . On dit que $(A_n)_{n \in \mathbb{N}}$ est un *système quasi-complet d'évènements* de Ω si les deux propriétés suivantes sont vérifiées :

- ▶ $\sum_{n=0}^{\infty} \mathbf{P}(A_n) = 1,$ ⁷
- ▶ les A_n sont deux à deux disjoints, *i.e.* pour tous $i \neq j$ deux entiers, on a :
 $A_i \cap A_j = \emptyset.$

Proposition ALEA.12.13

Tout système complet d'évènements est un système quasi-complet d'évènements.

Preuve

**Remarque 2.4 — Structure de l'ensemble des probabilités**

- ▶ L'ensemble des probabilités sur un espace probabilisable (Ω, \mathcal{F}) n'est **pas** un espace vectoriel, car ce n'est pas un sous-espace vectoriel de $[0, 1]^\Omega$.
- ▶ **(Stabilité par combinaison convexe)** L'ensemble des probabilités sur un espace probabilisable (Ω, \mathcal{F}) est stable par combinaison convexe : *i.e.* si \mathbf{P}_1 et \mathbf{P}_2 sont deux probabilités sur (Ω, \mathcal{F}) alors $\lambda\mathbf{P}_1 + \mu\mathbf{P}_2$ est aussi une probabilité avec $\lambda + \mu = 1$.

⁷On a remplacé la réunion dans la définition de système complet d'évènements par une égalité en probabilité.

PROPRIÉTÉS ÉLÉMENTAIRES D'UNE PROBABILITÉ.**Proposition ALEA.12.14 |**

Soit un espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$. Alors pour tout $(A, B) \in \mathcal{F}^2$, on a :

1. **(Différence)** $\mathbf{P}(B \setminus A) = \mathbf{P}(B) - \mathbf{P}(A \cap B)$. En particulier si $A \subset B$ alors $\mathbf{P}(B \setminus A) = \mathbf{P}(B) - \mathbf{P}(A)$.
2. **(Formule d'inclusion/exclusion)** $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$,
3. **(Additivité)** si $A \cap B = \emptyset$ alors $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$. En particulier, $\mathbf{P}(\emptyset) = 0$ et $\mathbf{P}(^c A) = 1 - \mathbf{P}(A)$,
4. **(Monotonie pour \subset)** si $A \subset B$ alors $\mathbf{P}(A) \leq \mathbf{P}(B)$.

Preuve



Si la suite $(A_n)_{n \in \mathbb{N}}$ dans la **Définition ALEA.12.9** n'est pas monotone, on a en revanche qu'une inégalité.

Proposition ALEA.12.15 | Sous-additivité dénombrable

Soit $(\Omega, \mathcal{F}, \mathbf{P})$ un espace de probabilité et soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements de \mathcal{F} . Alors,

$$\text{si } \sum_{n=0}^{\infty} \mathbf{P}(A_n) \text{ converge, alors } \mathbf{P}\left(\bigcup_{n=0}^{\infty} A_n\right) \leq \sum_{n=0}^{\infty} \mathbf{P}(A_n).$$

Remarque 2.5 — Même si la série majorante diverge, c'est forcément vers $+\infty$ puisqu'elle est à termes positifs. L'inégalité reste donc vraie également dans ce cas.


Preuve Nous admettons cette preuve. L'idée principale étant d'écrire la réunion *a priori* non disjointe $\bigcup_{n=0}^{\infty} A_n$ en une réunion disjointe $\bigcup_{n=0}^{\infty} \widetilde{A}_n$ avec $\widetilde{A}_n \in \mathcal{F}$ pour tout $n \in \mathbb{N}$. L'axiome d'additivité dénombrable d'une probabilité nous permet ensuite de conclure en faisant le lien entre les $\mathbf{P}(A_n)$ et les $\mathbf{P}(\widetilde{A}_n)$.

Passons à présent à des exemples classiques d'espaces probabilisables.

Exemple 15 — Un jeu non truqué de pile ou face fini. Le pile ou face en une étape est décrit par l'espace $(\Omega, \mathcal{F}, \mathbf{P})$ avec $\Omega = \{P, F\} = \{P\} \cup \{F\}$,

$$\mathcal{F} = \mathcal{P}(\Omega) = \{\emptyset, \{P\}, \{F\}, \{P, F\}\},$$

où \mathbf{P} est la probabilité uniforme sur Ω donnée par $\mathbf{P}(A) = \frac{\#A}{\#\Omega} = \frac{\#A}{2}$ pour tout $A \in \mathcal{F}$.

Vérifions que $(\Omega, \mathcal{F}, \mathbf{P})$ est un espace probabilisé. 

Si l'on veut tenir compte d'une succession de N jeux, on prend plutôt $\Omega = \{P, F\}^N$, $\mathcal{F} = \mathcal{P}(\Omega)$ et \mathbf{P} la probabilité uniforme sur Ω donnée par $\mathbf{P}(A) = \frac{\#A}{\#\Omega} = \frac{\#A}{2^N}$ pour tout $A \in \mathcal{F}$.

Exemple 16 — Main de Poker La sélection d'une main au poker peut être décrite par l'espace $(\Omega, \mathcal{F}, \mathbf{P})$ avec $\Omega = \mathcal{P}_5(\{cartes\})$ l'ensemble des parties à cinq éléments de l'ensemble des cartes, $\mathcal{F} = \mathcal{P}(\Omega)$, le nombre de cartes en main dépendant de la manche, et \mathbf{P} la probabilité uniforme (sur Ω) donnée par $\mathbf{P}(A) = \frac{\#A}{\#\Omega} = \frac{\#A}{\binom{52}{5}}$ pour tout $A \in \mathcal{F}$.


On vérifie de la même manière que $(\Omega, \mathcal{F}, \mathbf{P})$ est un espace probabilisé.

Exemple 17 — Généralisation : Probabilité uniforme discrète, issues équiprobables

Soit Ω un ensemble fini non vide. Pour tout $A \in \mathcal{P}(\Omega)$, posons

$$\mathbf{P}_{\Omega}^{\#}(A) = \frac{\# A}{\# \Omega}.$$

Alors on montre que $\mathbf{P}_{\Omega}^{\#}$ définit une probabilité sur $(\Omega, \mathcal{P}(\Omega))$ appelée *probabilité uniforme* sur Ω . Quelles expériences aléatoires pourraient correspondre à ce triplet $(\Omega, \mathcal{T}, \mathbf{P}_{\Omega}^{\#})$?

 *N'importe quelle expérience aléatoire entrant dans un contexte d'équiprobabilité. Par exemple $\Omega = \llbracket 1, 6 \rrbracket$ correspondant au lancer d'un dé équilibré. Chaque issue est alors de probabilité 1/6.*

Exemple 18 — Jeu infini de pile ou face Il est naturel d'introduire comme espace des issues l'ensemble Ω des suites $(\omega_1, \omega_2, \dots)$ où les ω_i valent P ou F (et non plus des N-uplets finis). On a ainsi


$$\Omega = \{P, F\}^{\mathbf{N}^*},$$

l'ensemble des suites à valeurs dans $\{0, 1\}$. **Cet univers n'est pas au plus dénombrable** : on peut montrer qu'il est en bijection avec \mathbf{R} .

La construction d'une tribu est beaucoup moins évidente que dans les autres cas. On suppose qu'il existe $\mathcal{T} \subset \mathcal{P}(\Omega)$ telle que pour tout $n \in \mathbf{N}^*$, \mathcal{T} contient A_n «le n -ième lancer a donné face» (et donc *a fortiori* toute intersection ou réunion de tels ensembles ou de leur complémentaire — ce qui signifie, en terme d'information disponible, que l'on peut observer uniquement un nombre **fini** de lancers). Nous admettons que sur cet espace probabilisable (Ω, \mathcal{T}) ainsi obtenu on peut définir une probabilité vérifiant la propriété suivante : pour tout $n \in \mathbf{N}^*$,

$$\mathbf{P}(A_n) = \frac{1}{2}.$$

Un joueur lance une pièce équilibré jusqu'à ce qu'il obtienne pile. Montrer que l'évènement F «le joueur obtient face à tous les lancers» est négligeable. *Indication* : Introduire l'évènement B_n «le joueur obtient face au cours des n premiers lancers» pour tout $n \in \mathbf{N}^*$.

 *Notons F «le joueur obtient face à tous les lancers». C'est un évènement car*

$$F = \bigcap_{n=0}^{\infty} A_n,$$


c'est donc une réunion dénombrable d'évènements (par hypothèse les A_n sont dans \mathcal{T}).

*Mêmes si les A_n sont indépendants on ne peut dire que la probabilité de l'intersection est le produit infini des probabilités.*⁸

Mais, $F = \bigcap_{k=1}^{\infty} A_k \subset \bigcap_{k=1}^n A_k = B_n$. Or, $\mathbf{P}(B_n) = \left(\frac{1}{2}\right)^n \xrightarrow{n \rightarrow \infty} 0$. Donc par théorème de majoration, $\lim_{n \rightarrow \infty} \mathbf{P}\left(\bigcap_{k=1}^n A_k\right) = 0$. L'évènement F est donc bien négligeable.

Exemple 19 — Masse de Dirac, issues certaines Soit Ω un ensemble quelconque et $\omega \in \Omega$. On définit une probabilité sur $(\Omega, \mathcal{P}(\Omega))$ en posant pour tout $A \in \mathcal{P}(\Omega)$:

$$\delta_{\omega}(A) = \begin{cases} 1 & \text{si } \omega \in A, \\ 0 & \text{sinon.} \end{cases}$$

On vérifie facilement que pour tout $\omega \in \Omega$, δ_{ω} est une probabilité sur $(\Omega, \mathcal{P}(\Omega))$, elle est appelée *masse de DIRAC* en a . Pour cette probabilité, les évènements sont toujours certains ou négligeables. Quelle expérience aléatoire pourrait correspondre à ce triplet $(\Omega, \mathcal{P}(\Omega), \delta_{\omega})$?  *Par exemple, un lancer de dé qui ne comporterait que des 6, avec ω l'évènement «faire 6 au lancer».*

2.4. Résultat d'existence de probabilités

On rappelle que le qualificatif *au plus dénombrable* signifie dénombrable ou fini.

Vous avez déjà rencontré le théorème suivant en première année dans le cas où Ω est fini. Si on note $\Omega = \{\omega_1, \dots, \omega_n\}$, avec $n \in \mathbf{N}$ alors si P est une probabilité sur $(\Omega, \mathcal{P}(\Omega))$, on a pour tout $A \in \mathcal{P}(\Omega)$:

$$\mathbf{P}(A) = \mathbf{P}\left(\bigoplus_{\omega \in A} \{\omega\}\right) = \sum_{\omega \in A} \mathbf{P}(\{\omega\}).$$

⁸Mais on peut utiliser le théorème dit de «continuité le long des suites croissantes/décroissantes», mais hors programme en BCPST.

En d'autres termes, il faut définir $\mathbf{P}(\{\omega\})$, $\omega \in \Omega$ pour construire entièrement \mathbf{P} .⁹ Mais peut-on prendre n'importe quoi pour $\mathbf{P}(\{\omega\})$? La réponse est non, en effet, puisque $\Omega = \bigsqcup_{\omega \in \Omega} \{\omega\}$ et que $\mathbf{P}(\Omega) = 1$, l'axiome d'additivité dénombrable impose

$$\mathbf{P}(\Omega) = 1 = \sum_{i=1}^n \mathbf{P}(\{\omega_i\}).$$

Ces faits peuvent être généralisés sans peine aux univers au plus dénombrable, c'est ce qui est précisé dans le prochain théorème.

Théorème ALEA.12.1 | Existence d'une probabilité sur un univers au plus dénombrable

Soit Ω un univers non vide, fini ou dénombrable. Pour toute famille $(p_\omega)_{\omega \in \Omega}$ de \mathbf{R}^+ telle que $\sum_{\omega \in \Omega} p_\omega = 1$, il existe une unique probabilité \mathbf{P} sur $(\Omega, \mathcal{P}(\Omega))$ telle que :

- ▶ $\mathbf{P}(\{\omega\}) = p_\omega$ pour tout $\omega \in \Omega$.
- ▶ Plus précisément, elle est définie par : $\forall A \in \mathcal{P}(\Omega), \mathbf{P}(A) = \sum_{\omega \in A} p_\omega$.


Remarque 2.6 – Notation $\sum_{\omega \in \Omega} p_\omega$ Notez que $\sum_{\omega \in \Omega} p_\omega$ a du sens puisque Ω est supposé au plus dénombrable : en effet, on peut écrire Ω sous forme « numérotée » $\Omega = \{\omega_n, n \in \mathbf{N}\}$ et on pose

$$\sum_{\omega \in \Omega} p_\omega = \sum_{n=0}^{\infty} p_{\omega_n}.$$

De plus, cette somme existe et ne dépend pas de la numérotation puisque la série converge absolument. De-même pour tout $A \in \mathcal{P}(\Omega)$, on définit ainsi $\sum_{\omega \in A} p_\omega$.


Preuve ■ **Analyse** – (ou unicité). Si \mathbf{P} est une probabilité sur Ω telle que pour tout $\omega \in \Omega$ on ait $\mathbf{P}(\{\omega\}) = p_\omega$, alors, si $A \in \Omega$, on a en utilisant la propriété d'additivité

⁹De la même manière que, pour définir une application linéaire, il suffit de la définir sur une base


dénombrable d'une probabilité : 

$$\mathbf{P}(A) = \mathbf{P}\left(\bigsqcup_{\omega \in A} \{\omega\}\right) = \sum_{\omega \in A} \mathbf{P}(\{\omega\}).$$

On obtient la formule du théorème qui définit \mathbf{P} , donc si \mathbf{P} existe, elle est forcément définie par cette formule.

■ **Synthèse.** Définissons alors \mathbf{P} par : $\forall A \in \mathcal{P}(\Omega), \mathbf{P}(A) = \sum_{\omega \in A} \mathbf{P}(\{\omega\})$. Vérifions que \mathbf{P} est une probabilité sur $(\Omega, \mathcal{P}(\Omega))$. 

fions que \mathbf{P} est une probabilité sur $(\Omega, \mathcal{P}(\Omega))$. 

Exemple 20 – La suite $\left(\frac{1}{2^n}\right)_{n \geq 1}$ définit une probabilité sur \mathbf{N}^* , mais aussi sur $\{(1), (2, 2), (3, 3, 3), \dots, (n, \dots, n), \dots\}$ munis de leur tribu des parties. 

2.5. Conditionnement & Indépendance d'évènements

Un énoncé de probabilité conditionnelle est un énoncé du type : « si B se produit alors la probabilité que A se produise est p ».

- ▶ Par exemple, considérons les évènements « il neige » et « le bus est en retard ». Si je sais qu'il neige, la probabilité que le bus soit en retard devrait être augmentée.
- ▶ Autre exemple, si l'on considère le jeu ci-après, trois portes, derrière l'une des trois se cache un trésor. J'ai le droit de regarder successivement derrière chacune des portes et d'en choisir une à chaque tour. Sachant que j'ai déjà regardé une porte, j'ai plus de chance au tour suivant de tomber sur la bonne.

On souhaite donc définir une nouvelle notion probabiliste qui tienne compte de cet apport d'information.

2.5.1. Conditionnement par rapport à un évènement

Mathématiquement, nous allons considérer la définition suivante.

Définition ALEA.12.12 |

Soit $(\Omega, \mathcal{F}, \mathbf{P})$ un espace probabilisé. Soit B un évènement non-négligeable *i.e.* tel que $\mathbf{P}(B) > 0$. Pour tout évènement $A \in \mathcal{F}$ on pose

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}.$$

Cette quantité est appelée *probabilité de A sachant B* ou *probabilité conditionnellement à B*.¹⁰

On aurait pu penser à introduire l'application


$$A \in \mathcal{F} \longrightarrow \mathbf{P}(A \cap B)$$

pour décrire l'apport d'information. Seulement cette application n'est pas une probabilité sur (Ω, \mathcal{F}) , puisque $\mathbf{P}(\Omega \cap B) = \mathbf{P}(B) \neq 1$ *a priori*, il faut donc diviser par $\mathbf{P}(B)$ pour former une probabilité.

Attention L'évènement $\{A | B\}$ n'existe pas!

- ▶ Nous n'avons pas défini l'évènement $\{A | B\}$, **seulement la probabilité conditionnelle sachant B**.
- ▶ À ne pas confondre avec l'évènement $A \setminus B = A \cap {}^c B$. De plus, avant de parler de conditionnement **il est nécessaire de préciser que B n'est pas négligeable**.

Exemple 21 — Une famille a deux enfants. Quelle est la probabilité que les deux soient des garçons conditionnellement au fait qu'au moins l'un des deux est un garçon, que les deux soient des garçons? Préciser un espace probabilisé associé à l'expérience.

 Avec des notations évidentes, l'espace probabilisé est $\Omega = \{GG, GF, FG, FF\}$ muni de la probabilité uniforme. La probabilité cherchée est

$$\mathbf{P}(GG|\{GG, GF, FG\}) = \frac{\mathbf{P}(\{GG\})}{\mathbf{P}(\{GG, GF, FG\})} = \frac{1/4}{3/4} = \frac{1}{3}.$$

En revanche, $\mathbf{P}(GG) = \frac{1}{4} < \frac{1}{3}$.

¹⁰Parfois aussi notée $\mathbf{P}_B(A)$.

2.5.2. Formules probabilistes

On déduit directement de la définition une formule de permutation de conditionnement appelée formule de BAYES.

 **Attention**

Un raisonnement probabiliste doit s'appuyer sur l'une des formules ci-après. Vous pouvez éventuellement vous appuyer sur un arbre **au brouillon**, mais l'utilisation d'un seul arbre sur une copie n'est pas une rédaction suffisamment convaincante.

Proposition ALEA.12.16 | Formule de BAYES OU DE RÉCIPROCITÉ CONDITIONNELLE

1. Soient A et B des événements non négligeables, alors

$$\mathbf{P(A|B)} = \mathbf{P(B|A)} \frac{\mathbf{P(A)}}{\mathbf{P(B)}}.$$

2. Soit $(A_n)_{n \in \mathbb{N}}$ un système complet d'événements non négligeables. Alors :

$$\mathbf{P(A|B)} = \mathbf{P(B|A)} \frac{\mathbf{P(A)}}{\sum_{n=0}^{\infty} \mathbf{P(B|A_n)} \mathbf{P(A_n)}}.$$

Preuve



Exemple 22 – Intérêt des QCM pour les examens Considérons des questions où m réponses possibles sont proposées et supposons qu'un candidat a une probabilité p de connaître la réponse à une question prise au hasard parmi un ensemble fini de questions. Sachant que le candidat a répondu correctement à la question, quelle est la probabilité qu'il sache effectivement la réponse? On suppose qu'un candidat ne sachant pas la réponse répond « au hasard », et donc que chacune des m réponses possibles sont équiprobables.

Soit RC l'événement « le candidat répond correctement » et Co l'événement « le candidat connaît la réponse ». Appliquons la règle de BAYES,

$$\begin{aligned} \mathbf{P(Co | RC)} &= \frac{\mathbf{P(RC | Co)} \mathbf{P(Co)}}{\mathbf{P(RC)}} \\ &= \frac{\mathbf{P(RC | Co)} \mathbf{P(Co)}}{\mathbf{P(RC \cap Co)} + \mathbf{P(RC \cap {}^cCo)}} \\ &= \frac{\mathbf{P(RC | Co)} \mathbf{P(Co)}}{\mathbf{P(RC | Co)} \mathbf{P(Co)} + \mathbf{P(RC | {}^cCo)} \mathbf{P}({}^cCo)} = \frac{1 \cdot p}{1 \cdot p + \frac{1}{m}(1 - p)} \\ &= \frac{mp}{mp + 1 - p}. \end{aligned}$$

Donc plus m est grand, plus $\mathbf{P(Co|RC)}$ est grand. C'est assez intuitif car il est probable que le candidat connaisse la réponse s'il a donné une bonne réponse parmi de nombreuses proposées. Remarquons que pour $m = 3$ et $p = 1/2$, $\mathbf{P(Co | RC)} = 3/4$. Ce qui est somme toute assez grand! On conçoit donc qu'un questionnaire d'une trentaine de questions, chacune à trois ou quatre réponses possibles, soit à même de rendre compte du savoir d'un étudiant!

La formule des probabilités totales est la première grande formule à connaître en probabilités.

Proposition ALEA.12.17 | Formule des probabilités totales

Soit $(\Omega, \mathcal{F}, \mathbf{P})$ un espace probabilisé et $(A_n)_{n \in \mathbf{N}} \in \mathcal{F}^{\mathbf{N}}$ un système complet d'événements de Ω . Alors, pour tout événement $B \in \mathcal{F}$, on a :

$$\mathbf{P}(B) = \sum_{n=0}^{\infty} \mathbf{P}(B \cap A_n).$$

En particulier, si pour tout $n \in \mathbf{N}$, $\mathbf{P}(A_n) \neq 0$, alors :

$$\mathbf{P}(B) = \sum_{n=0}^{\infty} \mathbf{P}(B|A_n)\mathbf{P}(A_n).$$

Remarque 2.7 — Parfois, les énoncés ne font pas figurer la condition «pour tout $n \in \mathbf{N}$, $\mathbf{P}(A_n) \neq 0$ » mais décrètent que **par convention** :

$$\mathbf{P}(B|A_n)\mathbf{P}(A_n) = 0 \quad \text{si } A_n \text{ est négligeable.}$$

Méthode Quand utiliser la formule des probabilités totales

Pour calculer la probabilité d'un événement pour lequel on a besoin de faire une disjonction de cas. Exemple typique : deux urnes dont le tirage se fait dans l'une ou l'autre en fonction du résultat du tirage précédent, on utilise alors le résultat du tirage précédent comme système complet d'événements puis on applique la formule des probabilités totales.

Preuve

- (1er cas : $(A_n)_n$ est un système complet d'événements)

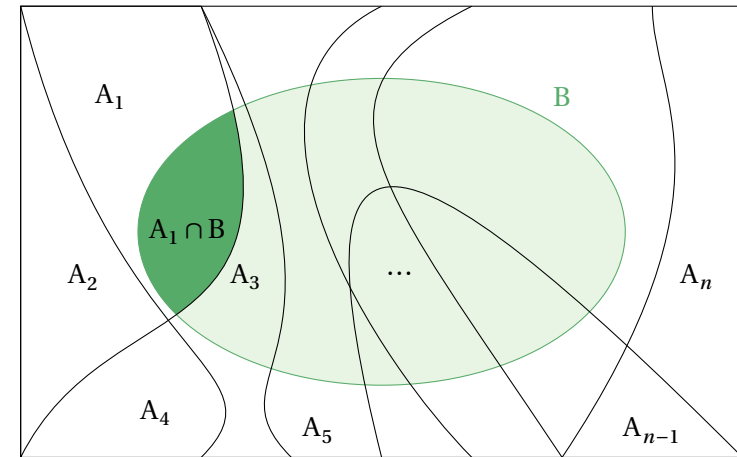


FIG. ALEA.12.2. : Écriture d'un événement B avec un système complet d'événements (A_n)

- (2ème cas : $(A_n)_n$ est un système quasi-complet d'événements) Nous avons donc ici

$$\forall n \neq m, \quad A_n \cap A_m = \emptyset, \quad \sum_{k=0}^{\infty} \mathbf{P}(A_k) = 1.$$

De manière équivalente, c'est supposer que

$$\forall n \neq m, A_n \cap A_m = \emptyset, \quad \mathbf{P}\left(\bigsqcup_{n \in \mathbb{N}} A_n\right) = 1.$$

Notons dans la suite $A = \bigsqcup_{n \in \mathbb{N}} A_n$. On a

$$\begin{aligned} B &= (B \cap A) \bigsqcup (B \cap {}^c A) \\ &= \left(\bigsqcup_{n \in \mathbb{N}} B \cap A_n\right) \bigsqcup (B \cap {}^c A), \end{aligned} \quad \left. \begin{array}{l} \\ \\ \end{array} \right\} \text{en passant aux proba}$$

$$\mathbf{P}(B) = \sum_{n=0}^{\infty} \mathbf{P}(B \cap A_n) + \mathbf{P}(B \cap {}^c A).$$

Or, par hypothèse $\mathbf{P}(A) = 1$, donc $\mathbf{P}({}^c A) = 0$. Mais comme $(B \cap {}^c A) \subset {}^c A$, on obtient $0 \leq \mathbf{P}(B \cap {}^c A) \leq \mathbf{P}({}^c A) = 0$, donc $\mathbf{P}(B \cap {}^c A) = 0$ et la formule est établie.

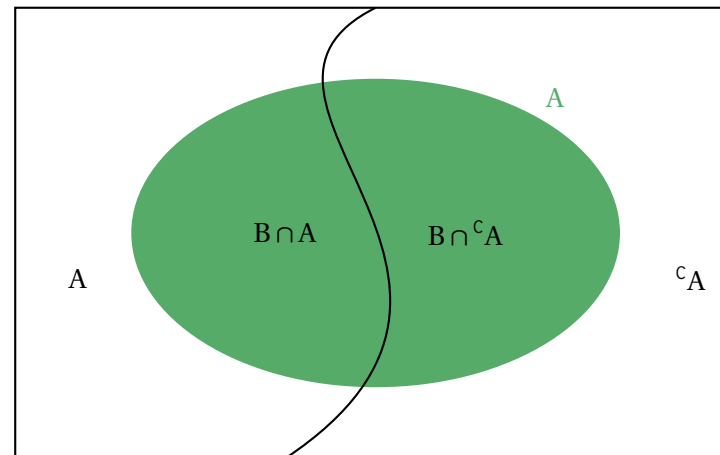


FIG. ALEA.12.3. : système complet d'évènements avec un évènement et son complémentaire


Puisqu'un système complet d'évènements est un système quasi-complet d'évènements nous aurions pu nous contenter du second cas.

Remarque 2.8 — Un cas particulier qui revient souvent est la formule

$$\mathbf{P}(B) = \mathbf{P}(B|A) \mathbf{P}(A) + \mathbf{P}(B|{}^c A) \mathbf{P}({}^c A)$$

valable dès que $0 < \mathbf{P}(A) \leq 1$ qui provient directement du système complet d'évènements $(A, {}^c A)$. Alors la visualisation ensembliste est la suivante.

Exemple 23 — Test et faux positifs. Une maladie affecte une personne sur 1000. Le test de dépistage n'est pas parfait : le résultat est toujours positif pour une personne malade et pour une personne saine il est positif (donc erroné) 2 fois sur 100. Quelle est la probabilité qu'une personne ayant un résultat positif au test soit effectivement malade ?

 Soit T l'évènement « le test est positif » et M l'évènement « la personne est malade ». On cherche $\mathbf{P}(M|T)$. On écrit

$$\mathbf{P}(M|T) = \mathbf{P}(T|M) \frac{\mathbf{P}(M)}{\mathbf{P}(T)}.$$

D'après les données du problème $\mathbf{P}(T|M) = 1$ et $\mathbf{P}(M) = 0,001$. De plus

$$\mathbf{P}(T) = \mathbf{P}(T|M) \mathbf{P}(M) + \mathbf{P}(T|M^c) \mathbf{P}(M^c) = 1 \times 0,001 + 0,02 \times 0,999.$$

En regroupant tout, on trouve que $\mathbf{P}(M|T)$ est de l'ordre de 5%. Le test est probablement erroné.

Nous avons par définition, que pour tout couple d'évènements $(A, B) \in \mathcal{F}^2$, tels que B ne soit pas négligeable :

$$\mathbf{P}(A \cap B) = \mathbf{P}(A|B) \mathbf{P}(B),$$

on peut généraliser sans trop de difficulté à une intersection de n évènements.

Théorème ALEA.12.2 | Formule des probabilités composées

Soient $(\Omega, \mathcal{F}, \mathbf{P})$, $n \geq 2$, et $A_1, \dots, A_n \in \mathcal{F}$ tels que : $\mathbf{P}(A_1 \cap \dots \cap A_{n-1}) \neq 0$. Alors :

$$\mathbf{P}(A_1 \cap \dots \cap A_n) = \mathbf{P}(A_1) \mathbf{P}(A_2|A_1) \mathbf{P}(A_3|A_1 \cap A_2) \cdots \mathbf{P}(A_n|A_1 \cap \dots \cap A_{n-1}).$$

**Méthode** Quand utiliser la formule des probabilités composées

Pour calculer la probabilité d'un évènement qui est une intersection d'évènements non indépendants.¹¹ Exemple typique : une urne dont on change les proportions de boules de chaque type étape par étape, piocher pour la première fois une boule d'un type au rang n revient à piocher que des boules des autres types jusqu'au rang $n - 1$ puis une boule du bon type au rang n .

Remarque 2.9 — L'hypothèse $\mathbf{P}(A_1 \cap \dots \cap A_{n-1}) \neq 0$ garantit que toutes les probabilités conditionnelles existent. C'est un point clef de la démonstration ci-dessous.

Preuve



Remarque 2.10 — On l'utilise en général pour toute répétition d'expériences, dont la réalisation de la n -ième dépend de la $n - 1$ -ième.

Exercice ALEA.12.1 | Une histoire de princesse. Une princesse est retenue prisonnière dans un chateau. Un prince charmant se met en tête de la délivrer. Lorsqu'il arrive à

¹¹ sinon c'est plus facile, on utilise l'indépendance.

l'entrée du chateau, il se trouve devant 3 portes. Il en ouvre une au hasard (équiprobable).

- ▶ S'il ouvre la 1ère porte, il délivre la princesse.
- ▶ S'il ouvre la deuxième porte, un dragon apparait et le dévore.
- ▶ S'il ouvre la troisième porte, une sorcière lui fait boire un filtre, il oublie tout ce qu'il a vu et est mis à la porte du chateau.

Le prince renouvelle ses tentatives jusqu'à ce qu'il meure ou qu'il délivre la princesse.

1. Calculer la probabilité de l'évènement D_k : « il délivre la princesse au k -ème essai ».



2. Calculer la probabilité de l'évènement D : « il délivre la princesse ».



2.5.3. Indépendance d'évènements

De façon intuitive, on dit que A est indépendant de B si savoir B ne change pas la probabilité de A. C'est-à-dire si

$$\mathbf{P}(A|B) = \mathbf{P}(A).$$

Pour que cette formule ait un sens on est obligé de supposer que $\mathbf{P}(B) > 0$, ce qui n'est pas le cas dans la définition suivante. Mais si c'est le cas l'égalité précédente signifie simplement

$$\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B).$$

C'est cette définition que nous allons utiliser.

Définition ALEA.12.13

Soit $(\Omega, \mathcal{F}, \mathbf{P})$ un espace probabilisé et $(A_n)_{n \in \mathbf{N}} \in \mathcal{F}^{\mathbf{N}}$.


1. **(Indépendance)** On dit que les A_n sont *mutuellement indépendants* (on dit parfois seulement *indépendants*) si

$$\mathbf{P}\left(\bigcap_{n \in \mathcal{N}} A_n\right) = \prod_{n \in \mathcal{N}} \mathbf{P}(A_n)$$

pour toute partie \mathcal{N} finie non vide incluse dans \mathbf{N} . Pour deux évènements A, B, on notera $A \perp\!\!\!\perp B$ lorsqu'ils sont indépendants.

2. **(Indépendance deux à deux)** On dit que les A_n sont *indépendants deux à deux* si pour tous $i, j \in \mathbf{N}$, $A_i \perp\!\!\!\perp A_j$.

En particulier, deux évènements A et B sont donc *indépendants* si $\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$.

Exemple 24 — On tire une carte dans un paquet de 52. L'évènement A : «tirer un roi» est indépendant de l'évènement B : «tirer un pique». Pourquoi?  En effet $\mathbf{P}(A \cap B)$ est la probabilité de tirer le roi de pique, soit $1/52$, qui est bien égal à $\mathbf{P}(A)\mathbf{P}(B) = (4/52) \times (13/52) = 1/52$.

Proposition ALEA.12.18 | Propriétés de l'indépendance

Soient $A, B \in \mathcal{F}$. Alors :

1. $A \perp\!\!\!\perp A \implies \mathbf{P}(A) \in \{0, 1\}$.
2. $A \perp\!\!\!\perp B \iff {}^c A \perp\!\!\!\perp {}^c B$.
3. $A \perp\!\!\!\perp B \implies A \perp\!\!\!\perp {}^c B, B \perp\!\!\!\perp {}^c A$.
4. **(Généralisation)** Soit $(A_n)_{n \in \mathbf{N}}$ une famille d'évènements deux à deux indépendants (*resp.* mutuellement indépendants), et $(B_n)_{n \in \mathbf{N}}$ une autre famille d'évènements telle que $B_n = A_n$ ou ${}^c A_n$ pour tout $n \in \mathbf{N}$. Alors : $(B_n)_{n \in \mathbf{N}}$ une famille d'évènements deux à deux indépendants (*resp.* mutuellement indépendants).


Preuve

1. 

2. 

3. 

4. Récurrence sur le nombre d'évènements après avoir choisi une partie finie.

Exemple 25 — On jette deux pièces. Les événements «la première pièce tombe sur pile», «la deuxième pièce tombe sur pile» et «les deux pièces donnent le même résultat» sont deux à deux indépendants, mais pas mutuellement indépendants. 

3. VARIABLES ALÉATOIRES

Reprenons notre expérience 1 introductive de lancer de dé. Nous l'avons décrite au moyen d'un espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$. Un autre point de vue est celui des variables aléatoires : *i.e.* de voir les valeurs d'un lancer de dé comme les valeurs d'une fonction sur un ensemble Ω .

On ne sera donc pas toujours intéressé par le résultat complet d'une expérience aléatoire, *i.e.* les éléments de Ω , mais plutôt par une fonction de ces derniers. Une telle fonction sera appelée *variable aléatoire* si la quantité d'intérêt est à valeurs dans \mathbf{R} ou *vecteur aléatoire* si la quantité est à valeurs dans \mathbf{R}^d , $d \geq 2$.

3.1. Généralités

En règle générale, puisqu'en 2^{ème} année les évènements sont les éléments de \mathcal{F} — et pas n'importe quelle partie — la définition de variable aléatoire sera très légèrement modifiée par rapport à celle de première année.

Définition ALEA.12.14 | Variable aléatoire réelle

Une application $X : \Omega \rightarrow \mathbf{R}$ est appelée *variable aléatoire réelle* si

$$\forall a \in \mathbf{R}, \quad \{X \leq a\} \stackrel{\text{(déf.)}}{=} \{\omega \in \Omega, \quad X(\omega) \leq a\} \in \mathcal{F}. \quad (\text{VaProp})$$

On appelle *support* de X (ou *univers image* de X) l'ensemble $X(\Omega)$.

Dans Eq. (VaProp) on peut en fait remplacer le symbole « $\leq a$ » par n'importe quel intervalle « $\in I$ », comme nous le voyons dans la proposition qui suit.

Proposition ALEA.12.19

Une application $X : \Omega \rightarrow \mathbf{R}$ est une variable aléatoire réelle si pour tout intervalle $I \subset \mathbf{R}$:

$$\{X \in I\} = \{\omega, X(\omega) \in I\} \in \mathcal{F}.$$

En particulier : pour tout $a \in I$, $\{X = a\} = \{\omega, X(\omega) = a\} \in \mathcal{F}$.

Σ Notation
Généralement, on note plus simplement $\mathbf{P}(X \in I)$ plutôt, qu'en toute rigueur, $\mathbf{P}(\{X \in I\})$.

L'assertion $\{X \leq a\} \in \mathcal{F}$ signifie que les probabilités

$$\mathbf{P}(X \leq a) \stackrel{\text{(nota.)}}{=} \mathbf{P}(\{X \leq a\})$$


sont bien définies pour tout $a \in \mathbf{R}$ (on rappelle que \mathbf{P} est définie sur \mathcal{F} uniquement).

Malgré la terminologie, une variable aléatoire est avant tout une **application** et pas une «variable» au sens propre du terme.

Remarque 3.1 —

- ▶ Notez bien que la notion de variable aléatoire réelle dépend de la tribu sous-jacente \mathcal{F} .
- ▶ Lorsque $\mathcal{F} = \mathcal{P}(\Omega)$ (cadre de première année) alors il est immédiat que Eq. (VaProp) est satisfaite pour tout $a \in \mathbf{R}$. En revanche si \mathcal{F} ne contient pas toutes les parties de Ω , ce n'est plus forcément le cas.

Exemple 26 — Cas discret

- ▶ On lance simultanément deux dés discernables et on choisit comme univers $\Omega = \llbracket 1, 6 \rrbracket^2$ (cf. Chapter ALEA.12), comme tribu celle des parties $\mathcal{F} = \mathcal{P}(\Omega)$ et que l'on munit de la probabilité uniforme. Cette probabilité décrit l'expérience puisque les dés ne sont pas truqués. Notons alors :
 - X la somme des valeurs des dés,
 - Y le maximum des deux valeurs.
 Que valent $X(\Omega)$ et $Y(\Omega)$? 

Écrire explicitement les applications X et Y mises en jeu. 

- ▶ Les applications X, Y sont bien des variables aléatoires puisque $\mathcal{F} = \mathcal{P}(\Omega)$.
- ▶ On tire à pile ou face avec une pièce jusqu'à obtenir pile et on note X le nombre de tirs effectués,

$$\Omega = \{P, FP, FFP, \dots\} \cup \underbrace{\{FFFFF \dots\}}_{\omega_0},$$

auquel on associe la tribu $\mathcal{F} = \mathcal{P}(\Omega)$. Mais $X(\omega_0)$ n'aura pas de sens dans ce cas

puisque X est à valeurs dans \mathbf{R} (et pas dans $\overline{\mathbf{R}}$!).

Exemple 27 — Cas réel

- ▶ On considère une ampoule électrique et on note X sa durée de vie.
- ▶ On observe deux bactéries et on s'intéresse à la durée de vie T de celle qui disparaîtra la première. On choisit comme univers $\Omega = [0, \infty[\times [0, \infty[$. On admet que

$$T \begin{cases} \Omega & \longrightarrow \mathbf{R} \\ (\omega_1, \omega_2) & \longrightarrow \min(\omega_1, \omega_2) \end{cases}$$

est une variable aléatoire relativement à une tribu sur Ω .¹²


Proposition ALEA.12.20 | Structure d'espace vectoriel, opérations

- ▶ Soient X, Y deux variables aléatoires réelles sur $(\Omega, \mathcal{F}, \mathbf{P})$ et $(\lambda, \mu) \in \mathbf{R}^2$, alors :
 - $\lambda X + \mu Y$ est une variable aléatoire réelle,
 - XY est une variable aléatoire réelle.
 De plus, l'ensemble des variables aléatoires réelles muni de l'addition et de la multiplication externe des applications est un \mathbf{R} -espace vectoriel.
- ▶ **(Minimum/maximum)** Soient X_1, \dots, X_n une collection de n variables aléatoires réelles, alors :

$$\min(X_1, \dots, X_n) \quad \text{et} \quad \max(X_1, \dots, X_n)$$

sont des variables aléatoires réelles.

La preuve qui suit est à bien connaître car nous utiliserons des techniques similaires dans des chapitres ultérieurs.

Preuve Nous admettons 1). Nous montrons 2). 

¹²que l'on ne précisera pas, il s'agit de la plus petite tribu contenant tous les produits d'intervalles ouverts de $[0, \infty[$

3.2. Fonction de répartition & Loi

On reprend dans cette section les précédentes notations, on se donne $X : \Omega \longrightarrow \mathbf{R}$ une variable aléatoire réelle.

Définition ALEA.12.15 | Loi

On appelle *loi* de la variable aléatoire réelle X , la fonction \mathbf{P}_X qui à un intervalle ouvert I de \mathbf{R} associe $\mathbf{P}_X(I)$ définie par

$$\mathbf{P}_X(I) = \mathbf{P}(X \in I).$$

Nous aurons plutôt recours à la fonction ci-après pour étudier une variable aléatoire réelle.

Définition ALEA.12.16 | Fonction de répartition

On appelle *fonction de répartition*¹³ de la variable aléatoire réelle X , la fonction $F_X : \mathbf{R} \rightarrow [0, 1]$ donnée, pour $x \in \mathbf{R}$, par :

$$F_X(x) = \mathbf{P}_X(]-\infty, x]) = \mathbf{P}(X \in]-\infty, x]) = \mathbf{P}(X \leq x),$$

Déterminer la loi d'une variable aléatoire c'est donc trouver avec quelle probabilité elle arrive dans un certain intervalle. Trouver la fonction de répartition c'est déterminer avec quelle probabilité elle est plus petite qu'une certaine valeur.

Remarque 3.2 — Notion de loi dans le cas discret. Dans le cas discret, que nous verrons plus tard, on appellera plutôt loi (ou *fonction de masse*) l'application

$$\mathbf{P}_X : x \in X(\Omega) \mapsto \mathbf{P}(X = x)$$

Nous verrons que la donnée de la loi, ou de la fonction précédente, sont équivalentes.

PROPRIÉTÉS DE LA FONCTION DE RÉPARTITION.

Proposition ALEA.12.21 | Propriétés analytiques

Soit $X : \Omega \rightarrow \mathbf{R}$ une variable aléatoire et F_X sa fonction de répartition. Alors

1. $0 \leq F_X \leq 1$,
2. F_X est croissante,
3. F_X est c.à.d.l.a.g. : *i.e.* elle est continue à droite en tout $x \in \mathbf{R}$, *i.e.*

$$F_X(x) = \lim_{\substack{y \rightarrow x \\ y > x}} F_X(y),$$

et possède une limite à gauche notée $F_X(x-) = \lim_{\substack{y \rightarrow x \\ y < x}} F_X(y)$ en tout $x \in \mathbf{R}$.

¹³ *cumulative distribution function* en Anglais

4. Pour tout $x \in \mathbf{R}$,

$$\lim_{x \rightarrow -\infty} F_X(x) = 0, \quad \lim_{x \rightarrow +\infty} F_X(x) = 1.$$

Preuve

1. Soit $x \in \mathbf{R}$, alors $\{X \leq x\} \subset \Omega$, alors par propriété de monotonie de probabilité :

$$F_X(x) \leq \mathbf{P}(\Omega) = 1.$$

2. Soient $x \leq y$, alors $\{X \leq x\} \subset \{X \leq y\}$ donc $\mathbf{P}(X \leq x) \leq \mathbf{P}(X \leq y)$, ce qui montre que F_X est croissante.
3. La fonction F_X possède nécessairement une limite finie à droite et à gauche puisqu'elle est croissante et à valeurs dans $[0, 1]$. Nous admettons la continuité à droite.
4. Admise.

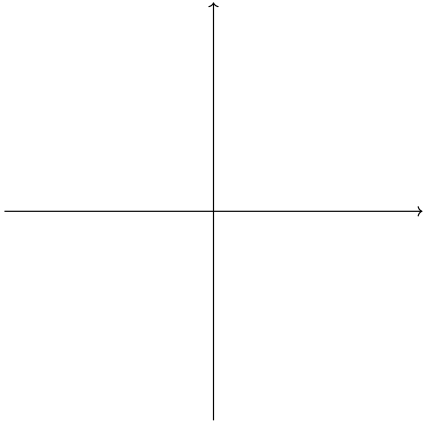
Remarque 3.3 — Ces propriétés caractérisent la notion de fonction de répartition au sens suivant : si F_X est une fonction vérifiant les trois propriétés précédentes alors on peut trouver un espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$ et une variable aléatoire $X : \Omega \rightarrow \mathbf{R}$ telle que $\mathbf{P}(X \leq x) = F_X(x)$ pour tout $x \in \mathbf{R}$. Plus précisément, si on a seulement la continuité à droite, cela suffit, comme le mentionne le théorème suivant.

Théorème ALEA.12.3 | Existence d'une variable aléatoire de fonction de répartition fixée

Soit F une fonction définie sur \mathbf{R} à valeurs réelles telle que :

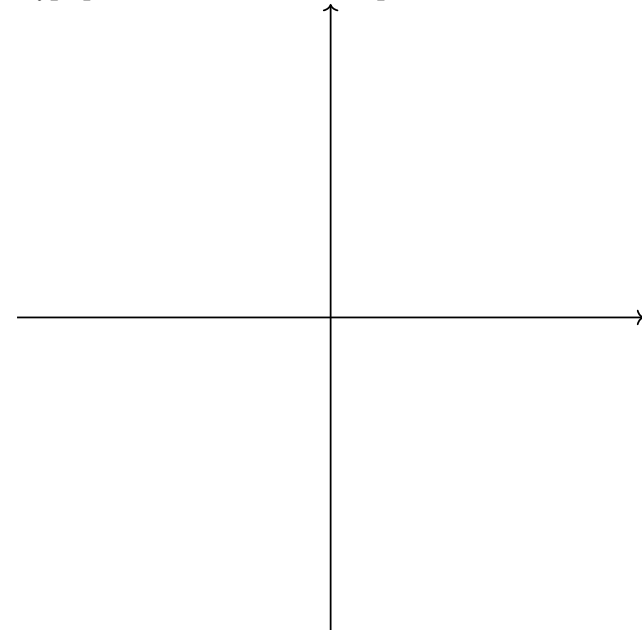
1. F est croissante sur \mathbf{R} ,
 2. $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$,
 3. F est continue à droite en tout point,
- alors il existe un espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$ et une variable aléatoire X définie sur cet espace tels que $F_X = F$.

Exemple 28 — On considère une variable aléatoire X à valeurs dans $\{1, \dots, 6\}$ donnant le résultat du lancé d'un dé à 6 faces. Calculer sa fonction de répartition et la tracer.



1. $\mathbf{P}(X > x) = 1 - F_X(x)$, la fonction $1 - F_X$ est en général appelée *fonction d'anti-répartition*,
 2. $\mathbf{P}(x < X \leq y) = F_X(y) - F_X(x)$,
 3. $\mathbf{P}(X < x) = F_X(x-)$,
 4. $\mathbf{P}(X = x) = F_X(x) - F_X(x-)$. En particulier, F_X est continue en $x \iff \mathbf{P}(X = x) = 0$.
- Lorsque $\mathbf{P}(X = x) = 0$ pour tout $x \in \mathbf{R}$, on dit que X est *sans atome*.

Si X n'a pas d'atome, alors la fonction F_X est continue en chaque point. Esquissons à présent le dessin typique d'une fonction de répartition. :



Preuve

1. 

Proposition ALEA.12.22 | Propriétés probabilistes

Soit $X : \Omega \rightarrow \mathbf{R}$ une variable aléatoire et F_X sa fonction de répartition. Soient $x \leq y$ des réels, on a

2.  Découlent directement de la définition de F_X , en écrivant que pour tous $x, y \in \mathbf{R}$ tels que $x \leq y$,

$$\{x < X \leq y\} = \{X \leq y\} \setminus \{X \leq x\},$$

on applique ensuite \mathbf{P} de chaque côté.

3. Constatons que pour tout $x \in \mathbf{R}$:

$$]-\infty; x[=]-\infty; x-1[\uplus \left(\bigcup_{n=1}^{+\infty} \left] x - \frac{1}{n}; x - \frac{1}{n+1} \right[\right).$$

Cette égalité se prouve par double inclusion. Puis on déduit :

$$\{X \in]-\infty; x[\} = \{X \in]-\infty; x-1[\} \uplus \bigcup_{n=1}^{+\infty} \left\{ X \in \left] x - \frac{1}{n}; x - \frac{1}{n+1} \right[\right\}.$$

En passant aux probabilités, puisque les réunions sont disjointes, on obtient :

$$\begin{aligned} \mathbf{P}(X < x) &= F_X(x-1) + \sum_{n=1}^{\infty} \mathbf{P}\left(X \in \left] x - \frac{1}{n}; x - \frac{1}{n+1} \right[\right), \\ &= F_X(x-1) + \sum_{n=1}^{\infty} \left(F_X\left(\frac{1}{n+1}\right) - F_X\left(\frac{1}{n}\right) \right) \\ &= F_X(x-1) + \lim_{n \rightarrow \infty} F_X\left(x - \frac{1}{n+1}\right) - F_X(x-1) = F_X(x-). \end{aligned} \quad \left. \begin{array}{l} \curvearrowright \\ \text{téléscopage} \end{array} \right\}$$

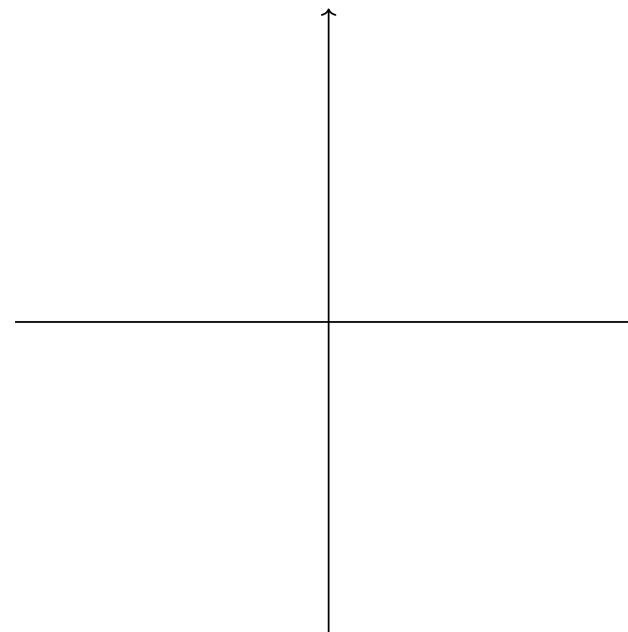
4. 

REPRÉSENTATION D'UNE VARIABLE ALÉATOIRE PAR HISTOGRAMME.

Définition ALEA.12.17

Soit $X : \Omega \rightarrow \mathbf{R}$ une variable aléatoire. On appelle *histogramme* de X de pas $h > 0$ la courbe représentative de la fonction H constante par morceaux, satisfaisant pour tout $i \in \mathbf{Z}$:

$$H : x \in [ih, (i+1)h[\longrightarrow \mathbf{P}(X \in [ih, (i+1)h[).$$



Nous verrons plus tard comment tracer un tel histogramme avec Python, en effectuant des simulations.

ABSENCE DE MÉMOIRE. Pour termine, présentons une propriété qui caractérise deux lois que nous rencontrerons plus tard.

Définition ALEA.12.18 | Absence de mémoire

Une variable aléatoire réelle X est dite *sans mémoire* si :

1. elle est positive ou nulle,
2. pour tout couple $(t, s) \in (\mathbf{R}^{++})^2$, on a :
- 3.

$$\mathbf{P}(X > t + s) = \mathbf{P}(X > s)\mathbf{P}(X > t). \quad (\text{Abs,Mém})$$

Avec les mêmes notations que dans la définition, l'Eq. (Abs,Mém) signifie aussi, de manière équivalente, que la fonction de répartition F vérifie la proposition suivante.

Proposition ALEA.12.23 | Reformulation de l'absence de mémoire

Une variable aléatoire réelle X est dite *sans mémoire* si et seulement si elle est positive, et pour tout couple $(t, s) \in (\mathbf{R}^{++})^2$, on a :

$$F(t + s) = F(t) + F(s) - F(t)F(s) \iff \mathbf{P}(X > t + s | X > s) = \mathbf{P}(X > t).$$

Preuve



Exemple 29 — Dans des chapitres ultérieures, nous montrerons que les variables ci-après sont à absence de mémoire :

- ▶ **(Loi géométrique)** X telle que $X(\Omega) = \mathbf{N}^*$ et pour tout $k \geq 1$, $\mathbf{P}(X = k) = p \cdot (1 - p)^{k-1}$ avec $p \in [0, 1]$. C'est même la seule loi discrète vérifiant cette propriété.

- ▶ **(Loi exponentielle)** X telle que pour tout $a < b$ réels, $\mathbf{P}(X \in [a, b]) = \int_a^b \lambda e^{-\lambda x} dx$ avec $\lambda \geq 0$. C'est même la seule loi à densité vérifiant cette propriété.

*** Fin du chapitre ***

4. EXERCICES

4.1. Dénombrement

Exercice ALEA.12.2 | Combien le mot BCPST possède-t-il d'anagrammes distinctes? Le mot CONFINEMENT?

Solution (exercice ALEA.12.2)

Dans le mot BCPST, il y a 5 lettres distinctes. Ainsi, un anagramme s'identifie à une permutation d'un ensemble à 5 éléments, il y a donc $5!$ anagrammes.

Pour le mot CONFINEMENT, on a des lettres identiques donc on place les lettres une par une. On a $\binom{11}{3}$ emplacements pour les N, puis $\binom{8}{2}$ emplacements pour les E (on choisit des combinaisons puisque l'ordre n'importe pas pour ces lettres identiques et on ne peut pas mettre deux lettres dans le même emplacement) et enfin $6!$ possibilités pour les lettres distinctes restantes, soit au total : $\binom{11}{3} \times \binom{8}{2} \times 6!$ anagrammes.

Exercice ALEA.12.3 | À partir d'un alphabet de p lettres, combien de mots de n lettres peut-on former

1. ne contenant que des lettres distinctes?
2. ne contenant jamais deux lettres identiques consécutives?

Solution (exercice ALEA.12.3)

1. Cela revient à compter le nombre de n -listes d'éléments distincts dans un ensemble (l'alphabet) de p éléments. Le cours donne alors le cardinal cherché :

$$\begin{cases} \frac{(p-n)!}{n!} & \text{si } n \leq p, \\ 0 & \text{sinon.} \end{cases}$$

Tout est logique, on peut former de mot avec que des lettres distinctes s'il y a moins de lettres dans l'alphabet que d'emplacements dans le mot.

2. Pour la première lettre, on peut choisir n'importe quelle lettre de l'alphabet (p possibilités). Ensuite, pour la suivante, on en a $p - 1$ car on ne peut reprendre la même. C'est ensuite le cas pour toutes les autres, on a donc au total $p(p - 1)^{n-1}$ mots possibles.

Exercice ALEA.12.4 | Combien existe-t-il de mots de 9 lettres contenant le mot « MERCI »?

Solution (exercice ALEA.12.4)

Le mot « MERCI » contient 5 lettres. On peut donc placer le « m » en position 1,2,3 ou 4. Une fois « MERCI » placé, il nous reste 4 lettres possibles, avec 26 choix pour chaque donc 26^4 possibilités.

Notons \mathcal{M}_k l'ensemble des mots contenant « MERCI » en position $k \in \llbracket 1, 4 \rrbracket$, et \mathcal{M} l'ensemble des mots. Alors

$$\mathcal{M} = \bigsqcup_{k=1}^4 \mathcal{M}_k,$$

donc en passant au cardinal :

$$\# \mathcal{M} = \sum_{k=1}^4 \# \mathcal{M}_k = \sum_{k=1}^4 26^4,$$

ainsi, $\# \mathcal{M} = 4 \cdot 26^4$.

Exercice ALEA.12.5 | Soient n couples qui se rencontrent et se saluent. Chaque personne serre (une fois) la main de chacune des autres (sauf celle de son conjoint). Combien y aura-t-il de poignées de mains échangées?

Solution (exercice ALEA.12.5)

Il y a ici $2n$ personnes et n couples.

- **(1ère Méthode)** commençons par compter les poignées de main pour le couple 1 : on en a $2n - 2$ pour chaque membre du couple donc $2(2n - 2)$ pour le couple 1. Ensuite pour le couple deux, on ne reconsidère pas les poignées déjà comptées, on a donc $2(2n - 4)$ — c'est simplement le résultat précédent où l'on a remplacé le nombre de couples n par $n - 1$ (le couple 1 n'étant plus compté). Donc on obtient finalement

$$2(2n-2)+2(2n-4)+\dots+2\times 2=2^2((n-1)+(n-2)+\dots+1)=2^2\frac{n(n-1)}{2}=\boxed{2n(n-1)}.$$

- **(2ème Méthode)** on ne se préoccupe pas des poignées déjà comptées. On a alors $2(2n - 2)$ poignées par couple, donc pour n couples au total $2n(2n - 2)$ poignées de main. Si l'on ne souhaite pas compter deux fois chaque poignées, on divise simplement le résultat précédent par deux, d'où : $n(2n-2) = \boxed{2n(n-1)}$ au total. On retrouve le précédent.

Exercice ALEA.12.6 | À l'issu d'un concours, 160 candidats sont admis et classés dont 70 garçons. Déterminer le nombre de classements possibles des 10 premiers admis qui contiennent autant de filles que de garçons.

Solution (exercice ALEA.12.6)

On choisit d'abord la composition du classement : 5 filles parmi les 90, soient $\binom{90}{5}$ et 5 garçons parmi les 70, soient $\binom{70}{5}$, ensuite il faut considérer toutes façons de permuter les 10 candidats sélectionnés, donc les permutations d'un ensemble à 10 éléments. Donc au total

$$\boxed{\binom{90}{5} \cdot \binom{70}{5} \cdot 10!}.$$

Exercice ALEA.12.7 | Soit E un ensemble fini de cardinal n et

$$F = \{(A, B) \in \mathcal{P}(E)^2, \#(A \cap B) = 1\}.$$

- Calculer $\# F$ de deux manières différentes.
- En déduire la valeur de $\sum_{k=0}^n \binom{n}{k} k 2^{n-k}$.

Solution (exercice ALEA.12.7)

- On peut commencer par découper F suivant l'élément que constitue l'intersection, *i.e.* en écrivant que

$$F = \bigcup_{x \in E} F_x$$

où $F_x = \{(A, B) \in \mathcal{P}(E)^2, A \cap B = \{x\}\}$, et chercher $\# F_x$ pour tout $x \in E$. Pour ce calcul, on ne cherche pas à remplir les deux parties A, B mais plutôt on va parcourir les éléments de E sauf x et regarder combien il y a de possibilités d'appartenance. Notons $E = \{x_1, \dots, x_{n-1}\} \cup \{x\}$, avec $x_i \neq x$ pour tout $i \in \llbracket 1, n-1 \rrbracket$.

Le nombre de couples de parties $(A, B) \in F_x$ est de manière équivalente le nombre de façons de répartir les éléments x entre $A \setminus A \cap B, B \setminus A \cap B$ et $E \setminus A \cup B$, il y a donc 3 choix possibles pour chaque x_i , donc 3^{n-1} choix possibles au total. Ainsi

$$\# F = \sum_{x \in E} 3^{n-1} = \boxed{n 3^{n-1}}.$$

- On peut aussi découper F selon le nombre d'éléments dans A (par exemple).

$$F = \bigcup_{k=1}^n A_k$$

où $A_k = \{(A, B) \in \mathcal{P}(E)^2, \#(A \cap B) = 1, \# A = k\}$. Pour compter les éléments de $A_k, k \in \llbracket 0, n \rrbracket$, on commence par choisir les k éléments de A au total $\binom{n}{k}$ possibilités, puis l'élément de $A \cap B$ donc k choix possibles, et enfin pour les $n - k$

restants dans E qui ne sont pas dans A cela revient à compter le nombre de parties possibles dans un ensemble à 2^{n-k} éléments donc 2^{n-k} possibilités.

Ainsi

$$\# A_k = \binom{n}{k} \cdot k \cdot 2^{n-k}.$$

Ainsi

$$\# F = \sum_{k=1}^n \binom{n}{k} \cdot k \cdot 2^{n-k}.$$

2. On a donc établi

$$\sum_{k=1}^n \binom{n}{k} \cdot k \cdot 2^{n-k} = n3^{n-1}.$$

Exercice ALEA.12.8 | Formule de VANDERMONDE

1. Soient $(a, b, c) \in \mathbb{N}^3$. Donner deux démonstrations de l'égalité

$$\sum_{k=0}^c \binom{a}{k} \binom{b}{c-k} = \binom{a+b}{c},$$

- 1.1) en développant de deux manières $(1+X)^a(1+X)^b$,
- 1.2) en calculant de deux manières le nombre de parties de cardinal c dans $E \cup F$, où E et F sont des ensembles disjoints de cardinaux a et b respectivement.

2. (Applications) Calculer $\sum_{p=0}^n \binom{n}{p}^2$.

Solution (exercice ALEA.12.8)

1. 1.1) ▶ D'une part, on peut utiliser les règles usuelles sur les puissances.

$$(1+X)^a(1+X)^b = (1+X)^{a+b} = \sum_{c=0}^{a+b} \binom{a+b}{c} X^c. \quad \text{formule du binôme}$$

▶ Deuxième calcul :

$$(1+X)^a(1+X)^b = \left(\sum_{k=0}^a \binom{a}{k} X^k \right) \cdot \left(\sum_{j=0}^b \binom{b}{j} X^j \right) = \sum_{k=0}^a \sum_{j=0}^b \binom{a}{k} \binom{b}{j} X^{k+j} = \sum_{c=0}^{a+b} \left[\sum_{k=0}^c \binom{a}{k} \binom{b}{c-k} \right] X^c. \quad \text{réécriture de la somme}$$

Il reste à identifier les coefficients de ces deux polynômes, les coefficients devant X^c donnent le résultat :

$$\sum_{k=0}^c \binom{a}{k} \binom{b}{c-k} = \binom{a+b}{c}.$$

1.2) Découpons les parties à c éléments de $E \uplus F$ de cardinal c , noté $\mathcal{P}_c(E \uplus F)$, en fonction du nombre d'éléments dans E. Plus précisément, on pose

$$\mathcal{P}_c(E \uplus F) = \bigsqcup_{k=0}^c E_k, \quad E_c = \{A \in \mathcal{P}_c(E \uplus F), \#(A \cap E) = k\}.$$

Pour définir $A \in E_k$, on choisit donc k éléments dans E puis $c - k$ dans F, donc au total

$$\# E_c = \binom{a}{k} \binom{b}{c-k}.$$

Donc en passant au cardinal on obtient

$$\binom{a+b}{c} = \#(\mathcal{P}_c(E \uplus F)) = \sum_{k=0}^c \# E_k = \sum_{k=0}^c \binom{a}{k} \binom{b}{c-k}.$$

On a donc établi de nouveau que

$$\sum_{k=0}^c \binom{a}{k} \binom{b}{c-k} = \binom{a+b}{c}.$$

2. Il suffit de choisir $a = n, b = n, c = n$. On obtient

$$\sum_{p=0}^n \binom{n}{p}^2 = \binom{2n}{n}.$$

4.2. Espaces probabilisés

4.2.1. Généralités

Exercice ALEA.12.9 | Deux probabilités sur $\{1, \dots, n\}$ Soit $\Omega = \{1, 2, \dots, n\}$.

- Déterminer une probabilité \mathbf{P} sur Ω , tel que, pour tout $k \in \llbracket 1, n \rrbracket$, $\mathbf{P}(\{k\})$ soit proportionnel à k .
- Déterminer une probabilité \mathbf{P} sur Ω , tel que, pour tout $k \in \llbracket 1, n \rrbracket$, $\mathbf{P}(\{1, 2, \dots, k\})$ soit proportionnel à k^2 .

Solution (exercice ALEA.12.9)

1. Une condition nécessaire et suffisante est la suivante :

$$\sum_{k=1}^n \mathbf{P}(\{k\}) = \mathbf{P}(\Omega) = 1.$$

Notons $\lambda \in \mathbf{R}^+$ de sorte que $\mathbf{P}(\{k\}) = \lambda \cdot k$. Alors

$$\lambda \sum_{k=1}^n k = 1 \implies \lambda = \frac{2}{n(n+1)}.$$

Donc

$$\forall k \in \llbracket 1, n \rrbracket, \quad \mathbf{P}(\{k\}) = \frac{2k}{n(n+1)}.$$

2. Une condition nécessaire et suffisante est la suivante :

$$\sum_{k=1}^n \mathbf{P}(\{k\}) = \mathbf{P}(\Omega) = 1.$$

Or, pour tout $k \geq 2$,

$$\mathbf{P}(\{k\}) = \mathbf{P}(\{1, 2, \dots, k\}) - \mathbf{P}(\{1, 2, \dots, k-1\}).$$

Notons $\lambda \in \mathbf{R}^+$ de sorte que $\mathbf{P}(\{1, 2, \dots, k\}) = \lambda \cdot k$. Alors

$$\lambda \sum_{k=2}^n \lambda(k - (k-1)) + \lambda 1 = 1 \implies \lambda = \frac{1}{n}.$$

Donc

$$\forall k \in \llbracket 1, n \rrbracket, \quad \mathbf{P}(\{k\}) = \frac{k}{n}.$$

Exercice ALEA.12.10 | L'équiprobabilité sur \mathbf{N} existe-t-elle? Peut-on choisir un entier naturel au hasard de manière équiprobable?

Solution (exercice ALEA.12.10)

Si une telle probabilité existait, alors $\mathbf{P}(\{n\}) = p$ pour tout $n \in \mathbf{N}$ avec $p \in [0, 1]$. Donc par additivité dénombrable, on aurait $\mathbf{P}(\mathbf{N}) = \sum_{n \in \mathbf{N}} p = 1$. Si $p = 0$, alors c'est déjà une contradiction. Sinon $p \neq 0$, et par le cours sur les séries, on sait bien que $\sum_{n \in \mathbf{N}} p$ diverge — c'est là encore une contradiction.

Il n'existe donc pas d'équiprobabilité sur \mathbf{N} .

Exercice ALEA.12.11 |

1. Montrer que : $\forall x \in \mathbf{R}, e^x \geq 1 + x$.
2. Soient A_1, \dots, A_n des évènements indépendants d'un espace probabilisé (Ω, \mathbf{P}) . Montrer que la probabilité qu'aucun des A_i ne soit réalisé est inférieure à $\exp\left(-\sum_{i=1}^n \mathbf{P}(A_i)\right)$.

Solution (exercice ALEA.12.11)

1. Simple étude de fonction pour $x \mapsto e^x - 1 - x$.
2. On souhaite donc majorer la probabilité $\mathbf{P}(A_1^c \cap \dots \cap A_n^c)$. Comme A_1, \dots, A_n sont indépendants, les évènements A_1^c, \dots, A_n^c le sont aussi (pour le voir : commencer avec $n = 2$ et cela s'étend ensuite par récurrence). Donc : $\mathbf{P}(A_1^c \cap \dots \cap A_n^c) = \prod_{k=1}^n (1 - \mathbf{P}(A_k))$. Or, pour tout $x \in \mathbf{R}, 1 - x \leq e^{-x}$ d'après la question précédente. L'inégalité découle alors en faisant le produit puisque tous les termes sont positifs.

4.2.2. Expériences aléatoires

Exercice ALEA.12.12 | Roulette russe Un revolver à 6 coups contient une seule balle mais on ne sait pas à quel endroit du barillet. Le premier joueur place le revolver sur sa tempe et presse la gâchette. S'il survit le deuxième joueur fait de même. Vaut-il mieux jouer en premier ou en second? La taille du barillet importe-t-elle?

Solution (exercice ALEA.12.12)

Notons B_k l'évènement « on se prend la balle en jouant à la partie k ». Alors :

$$\begin{aligned} \mathbf{P}(B_1) &= \frac{1}{6}, \\ \mathbf{P}(B_2) &= \mathbf{P}(B_2 \cap {}^c B_1) \\ &= \mathbf{P}(B_2 | {}^c B_1) \mathbf{P}({}^c B_1) \\ &= \frac{1}{5} \cdot \frac{5}{6} = \frac{1}{6}. \end{aligned}$$

Il n'y a donc pas de raison de jouer en premier ou en deuxième. Cela est intuitivement clair, puisque certes avec un barillet entièrement vierge on a qu'une chance sur 6 de mourir (1 chance sur 5 après un essai), mais en jouant en deuxième on est avantagé par le fait que le premier joueur prend un risque avant nous. *In fine*, en probabilité, le jeu est équilibré.

Le barillet n'a aucune incidence sur la conclusions, s'il comporte $N \geq 1$ balles, nous aurions

$$\begin{aligned} \mathbf{P}(B_1) &= \frac{1}{N}, \\ \mathbf{P}(B_2) &= \mathbf{P}(B_2 \cap {}^c B_1) \\ &= \mathbf{P}(B_2 | {}^c B_1) \mathbf{P}({}^c B_1) \\ &= \frac{1}{N-1} \cdot \frac{N-1}{N} = \frac{1}{N}. \end{aligned}$$

La conclusion est toujours la même :

$$\mathbf{P}(B_1) = \mathbf{P}(B_2).$$

Exercice ALEA.12.13 | Indice de coïncidence d'un texte L'indice de coïncidence I_c d'un texte est la probabilité pour que, si on prenne deux lettres au hasard dans ce texte, ce soient les mêmes. On note n_A, \dots, n_Z le nombre de A, \dots, Z dans le texte, et n le nombre total de lettres. Montrer que :

$$I_c = \frac{n_A \times (n_A - 1)}{n \times (n - 1)} + \dots + \frac{n_Z \times (n_Z - 1)}{n \times (n - 1)}.$$

Cet indice est très utilisé en cryptographie, notamment dans celui de VIGENÈRE qui effectue des décalages par blocs de lettres. Le calcul de l'indice permet alors de trouver assez facilement la longueur des blocs ayant permis le codage en testant les possibilités, et en comparant à l'indice de coïncidence de la langue française.

Solution (exercice ALEA.12.13)

L'expérience aléatoire consiste ici à piocher deux lettres dans le texte, au hasard. L'univers Ω est l'ensemble des parties à deux éléments formées par les n lettres du texte. On munit Ω de la probabilité uniforme, sachant que $\#\Omega = \binom{n}{2} = \frac{n(n-1)}{2}$.

Notons $\mathcal{A} = \{A, B, \dots, Z\}$ l'alphabet, E «deux lettres choisies au hasard dans le texte sont les mêmes» et, pour chaque lettre $a \in \mathcal{A}$, E_a «deux lettres choisies au hasard dans le texte sont les mêmes et sont des a ». Alors $(E_a)_{a \in \mathcal{A}}$ forme un système complet d'évènements, et donc

$$I_c = \mathbf{P}(E) = \sum_{a \in \mathcal{A}} \mathbf{P}(E_a) = \sum_{a \in \mathcal{A}} \frac{\#E_a}{\#\Omega} = \sum_{a \in \mathcal{A}} \frac{\binom{n_a}{2}}{\binom{n}{2}} = \sum_{a \in \mathcal{A}} \frac{n_a(n_a-1)}{n(n-1)}.$$

On obtient alors :

$$I_c = \frac{n_A \times (n_A - 1)}{n \times (n - 1)} + \frac{n_B \times (n_B - 1)}{n \times (n - 1)} + \frac{n_C \times (n_C - 1)}{n \times (n - 1)} + \dots + \frac{n_Z \times (n_Z - 1)}{n \times (n - 1)}.$$

Exercice ALEA.12.14 | Relectures d'un livre pour la détection d'erreurs Un livre contient 4 erreurs. On le confie à N lecteurs différents, indépendants entre eux, pour détecter ces erreurs; pour chaque lecteur, chaque erreur est détectée, indépendamment des autres erreurs, avec une probabilité $1/3$.

1. On note p_N la probabilité qu'il ne subsiste aucune erreur après N lectures. Calculer p_N en fonction de N .
2. Déterminer, à l'aide de commandes Python, le nombre de relectures nécessaires pour que la probabilité qu'elles soient toutes corrigées soit supérieure à 0,90.
3. On suppose que le nombre x d'erreurs est réparti uniformément sur $\{0, 1, 2, 3, 4\}$.

- 3.1) Calculer à nouveau p_N en fonction de N .
- 3.2) Calculer la limite de cette probabilité lorsque $N \rightarrow \infty$. Cela est-il cohérent?

Solution (exercice ALEA.12.14)

1. Notons par exemple A, B, C, D les quatre erreurs et A_n, B_n, C_n, D_n les évènements «l'erreur A, B, C ou D n'est pas corrigée au bout de n lectures». On a : $\mathbf{P}(A_n) = \left(\frac{2}{3}\right)^n$ puisque les lectures réalisées de manière indépendantes. Si l'on ne souhaite pas passer à l'évènement complémentaire, on peut aussi regarder la probabilité de «l'erreur A est corrigée en n lectures» qui est

$$\sum_{k=1}^n \frac{2^{k-1}}{3} \frac{1}{3} = \frac{1}{3} \frac{1 - (2/3)^n}{1 - 2/3} = 1 - (2/3)^n.$$

On retrouve bien l'expression précédente, la somme est obtenue en écrivant Ω comme réunion des évènements «l'erreur A est corrigée à la première lecture», «l'erreur A est corrigée à la seconde lecture», ..., «l'erreur A est corrigée à la n -ième lecture». Comme les lectures se font de manière indépendante, la probabilité demandée est :

$$\begin{aligned} \mathbf{P}({}^cA_n \cap {}^cB_n \cap {}^cC_n \cap {}^cD_n) &= \mathbf{P}({}^cA_n) \times \mathbf{P}({}^cB_n) \times \mathbf{P}({}^cC_n) \times \mathbf{P}({}^cD_n) \\ &= (1 - (2/3)^n)^4. \end{aligned}$$

2. Dans ce cas on ne connaît plus le nombre d'erreurs, mais la question précédente s'étend sans peine à k erreurs avec $k \in \mathbf{N}$:

$$\mathbf{P}(E_1 \cap E_2 \cap \dots \cap E_k) = (1 - (2/3)^n)^k,$$

où E_i est l'évènement «l'erreur i n'est pas corrigée en n lectures». Ainsi, en condi-

tionnant suivant le nombre d'erreurs entre 0 et 4 on a la probabilité cherchée :

$$\begin{aligned} & \mathbf{P}(\{\text{tout est corrigée en } n \text{ lectures}\}) \\ &= \sum_{k=0}^4 \mathbf{P}(\{\text{tout est corrigée en } n \text{ lectures}\} | N_k) \mathbf{P}(N_k) \\ &= \sum_{k=0}^4 (1 - (2/3)^n)^k \frac{1}{5} \\ &= \frac{5^n}{5 \cdot 2^n} [1 - (1 - (2/3)^n)^5] \end{aligned}$$

où N_k est l'évènement « il y a k erreurs », qui est bien non négligeable pour tout k , et en utilisant la formule donnant la somme de premiers termes d'une suite géométrique.

Exercice ALEA.12.15 | Une urne contient six boules dont 4 blanches et 2 noires. On extrait une boule de l'urne, on note sa couleur, puis on la remet dans l'urne. On effectue ensuite des tirages avec remise jusqu'à l'obtention d'une boule de la même couleur que précédemment. Soit $k \in \mathbf{N}^*$. Déterminer la probabilité que le nombre de tirage après remise de la boule initialement tirée soit k .

Solution (exercice ALEA.12.15)

Ici, on a besoin de faire une disjonction de cas suivant le résultat du premier lancer. La formule des probabilités totales est donc bien adaptée.

Notons N_k l'évènement « le tirage k mène à une boule noire » pour $k \geq 0$. Le tirage 0 étant considéré comme le tirage initial sur lequel est basée l'expérience aléatoire. Notons T_n l'évènement « il y a $n \geq 1$ tirages après le tirage 0 ». Alors, puisque $(N_0, {}^cN_0)$

est un système complet d'évènements :

$$\begin{aligned} \mathbf{P}(T_n) &= \mathbf{P}(T_n | N_0) \mathbf{P}(N_0) + \mathbf{P}(T_n | {}^cN_0) \mathbf{P}({}^cN_0) \\ &= \left(\frac{2}{6}\right)^{n-1} \left(\frac{4}{6}\right) \cdot \frac{2}{6} + \left(\frac{4}{6}\right)^{n-1} \left(\frac{2}{6}\right) \cdot \frac{4}{6} \\ &= \frac{2^{n+2} + 2^{n+1}}{6^{n+1}} \\ &= \frac{3 \cdot 2^{n+1}}{6^{n+1}}. \end{aligned}$$

Exercice ALEA.12.16 | Urne à composition variable Soient $n, b \in \mathbf{N}^*$. On dispose d'une urne contenant initialement deux boules blanches et deux boules noires dans laquelle on effectue des tirages successifs d'une boule selon le protocole suivant : si la boule tirée est noire, on arrête les tirages. Si la boule tirée est blanche, elle est remise dans l'urne avec en plus b autres boues blanches. On note N_k « obtenir une boule noire lors du $k^{\text{ème}}$ tirage », A_n l'évènement « une boule blanche apparaît à chacun des n premiers tirages » et F « ne jamais obtenir de boules noires ». L'objectif de l'exercice est de montrer que l'évènement F est négligeable.

1. Calculer $\mathbf{P}(N_n)$ et montrer que $\mathbf{P}(A_n) = \prod_{k=0}^{n-1} \frac{kb+2}{kb+4}$ pour tout $n \geq 1$.
2. Supposons ici que $b = 1$. Calculer $\sum_{n=1}^{+\infty} \mathbf{P}(N_n)$. Conclure.
3. Nous allons généraliser ce résultat à b quelconque dans la suite.
 - 3.1) Établir que $\prod_{k=1}^n (1 + x_k) \geq \sum_{k=1}^n x_k$ pour toute famille x_1, \dots, x_n de réels positifs.
 - 3.2) Établir une minoration de $\frac{1}{\mathbf{P}(A_n)}$ pour tout $n \in \mathbf{N}^*$, et montrer que $\lim_{n \rightarrow \infty} \mathbf{P}(A_n) = 0$.
 - 3.3) En utilisant le fait que pour tout $m \in \mathbf{N}^*$, $\bigcap_{n=1}^{+\infty} A_n \subset A_m$, calculer

$$\mathbf{P}\left(\bigcap_{n=1}^{+\infty} A_n\right).$$

- 3.4) Conclure.

Solution (exercice ALEA.12.16)

1. On utilisera deux fois la formule des probabilités composées.

$$\begin{aligned} \mathbf{P}(N_n) &= \mathbf{P}(N_n \cap {}^c N_{n-1} \cap \dots \cap N_1) \\ &= \mathbf{P}(N_n | {}^c N_{n-1} \cap \dots \cap {}^c N_1) \times \mathbf{P}({}^c N_{n-1} | {}^c N_{n-2} \cap \dots \cap {}^c N_1) \times \dots \times \mathbf{P}({}^c N_2 | {}^c N_1) \mathbf{P}({}^c N_1) \\ &= \frac{2}{4+(n-1)b} \cdot \frac{2+(n-2)b}{4+(n-2)b} \cdot \dots \cdot \frac{2+b}{4+b} \frac{2}{4} \\ &= \frac{2}{4+(n-1)b} \cdot \prod_{k=0}^{n-2} \frac{2+kb}{4+kb} \end{aligned}$$

De-même :

$$\begin{aligned} \mathbf{P}(A_n) &= \frac{2+(n-1)b}{4+(n-1)b} \cdot \dots \cdot \frac{2+b}{4+b} \frac{2}{4} \\ &= \prod_{k=0}^{n-1} \frac{kb+2}{kb+4} \end{aligned}$$

2. Supposons ici que $b = 1$. On a dans ce cas

$$\mathbf{P}(N_n) = \frac{2}{n-3} \cdot \prod_{k=0}^{n-2} \frac{2+k}{4+k} = \frac{12n!}{(n+3)!} = \frac{12}{(n+1)(n+2)(n+3)}$$

On cherche ensuite a, b, c trois réels tels que

$$\forall n \in \mathbf{N}, \frac{1}{(n+1)(n+2)(n+3)} = \frac{a}{n+1} + \frac{b}{n+2} + \frac{c}{n+3}$$

Après calculs on trouve $a = 1/2, b = -1, c = 1/2$, donc

$$\begin{aligned} \forall N \in \mathbf{N}^*, \sum_{n=1}^N \mathbf{P}(N_n) &= 6 \sum_{n=1}^N \left(\frac{1}{n+1} - \frac{2}{n+2} + \frac{1}{n+3} \right) \\ &= 6 \left(\sum_{n=1}^N \left(\frac{1}{n+1} - \frac{1}{n+2} \right) + \sum_{n=1}^N \left(\frac{1}{n+3} - \frac{1}{n+2} \right) \right) \\ &\xrightarrow{N \rightarrow \infty} 1. \end{aligned}$$

Donc $\sum_{n=1}^{+\infty} \mathbf{P}(N_n) = 1$.

Donc $\bigcup_{n=1}^{\infty} N_n$ est un évènement quasi-certain, de complémentaire F, donc

$\mathbf{P}(F) = 0$.

3. Nous allons généraliser ce résultat à b quelconque dans la suite.

3.1) Faisons une récurrence sur $n \in \mathbf{N}^*$. Pour $n = 1$, on a bien $1 + x_1 \geq x_1$. Supposons la propriété vraie au rang n , alors

$$\begin{aligned} &\prod_{k=1}^{n+1} (1+x_k) \\ &= \prod_{k=1}^n (1+x_k)(1+x_{n+1}) \\ &= \prod_{k=1}^n (1+x_k) + x_{n+1} \cdot \prod_{k=1}^n (1+x_k) \\ &\geq \sum_{k=1}^n x_k + x_{n+1} \\ &= \sum_{k=1}^{n+1} x_k \end{aligned}$$

hypothèse de récurrence
HR + produit de réels ≥ 1

La propriété est donc établie par récurrence. On a donc

$$\prod_{k=1}^n (1+x_k) \geq \sum_{k=1}^n x_k$$

pour toute famille x_1, \dots, x_n de réels positifs.

3.2)

$$\begin{aligned} \frac{1}{\mathbf{P}(A_n)} &= \prod_{k=0}^{n-1} \frac{kb+4}{kb+2} \\ &= \prod_{k=0}^{n-1} \left(1 + \frac{2}{kb+2} \right) \\ &\geq \sum_{k=0}^{n-1} \frac{2}{kb+2} \end{aligned}$$

De plus, $\lim_{n \rightarrow \infty} \sum_{k=0}^{n-1} \frac{2}{kb+2} = \infty$ puisque la série harmonique est divergente.

Donc $\mathbf{P}(A_n) \xrightarrow{n \rightarrow \infty} 0$.

3.3) On a pour tout $m \in \mathbf{N}^*$, $\bigcap_{n=1}^{+\infty} A_n \subset A_m$, donc par monotonie d'une probabilité

$$0 \leq \mathbf{P}\left(\bigcap_{n=1}^{+\infty} A_n\right) \leq \mathbf{P}(A_m) \xrightarrow{m \rightarrow \infty} 0.$$

Donc par théorème d'encadrement, et puisque le minorant ne dépendant pas de m , on déduit

$$\mathbf{P}\left(\bigcap_{n=1}^{+\infty} A_n\right) = 0.$$

3.4) Nous avons

$$F = \bigcap_{n=1}^{+\infty} A_n.$$

On déduit alors $\mathbf{P}(F) = 0$. En d'autres termes, l'évènement « ne jamais obtenir de noir » est négligeable.

Exercice ALEA.12.17 | Propagation d'un message On considère n « menteurs » I_1, I_2, \dots, I_n . Le premier menteur I_1 reçoit une information sous la forme de « oui » ou « non ». Il transmet l'information à I_2 , ainsi de suite jusqu'à I_n qui l'annonce au monde. Chacun des menteurs transmet ce qu'il a entendu avec la probabilité p et le contraire avec la probabilité $1 - p$ où $0 < p < 1$. De plus, les réponses des n individus sont indépendantes.

1. Soient $n \in \mathbf{N}$ et p_n la probabilité que l'information possédée par le menteur n soit correcte *i.e.* celle d'origine. Déterminer une relation liant p_n et p_{n+1} .
2. En déduire la valeur de p_n et sa limite lorsque n tend vers l'infini.

Solution (exercice ALEA.12.17)

1. Notons C_n l'évènement « l'information possédée par le menteur n est correcte », *i.e.* $\mathbf{P}(C_n) = p_n$ pour tout $n \in \mathbf{N}$. Alors, puisque $\{C_n, {}^c C_n\}$ est un système complet d'évènements,

$$\begin{aligned} \mathbf{P}(C_{n+1}) &= \mathbf{P}(C_{n+1}|C_n) \mathbf{P}(C_n) + \mathbf{P}(C_{n+1}|{}^c C_n) \mathbf{P}({}^c C_n) \\ &= p p_n + (1 - p)(1 - p_n) \\ &= p_n(2p - 1) + (1 - p). \end{aligned}$$

2. La suite (p_n) est une suite arithmético-géométrique. Deux cas se présentent.

- ▶ Si $p = 1$, alors $p_{n+1} = p_n = p_1 = 1$ pour tout $n \in \mathbf{N}$. Donc $p_n \xrightarrow{n \rightarrow \infty} 1$. L'information est donc presque-sûrement correctement transmise.
- ▶ Si $p \neq 1$, cherchons $C \in \mathbf{R}$ de sorte que

$$C = (2p - 1)C + (1 - p) \iff C = \frac{1 - p}{2(1 - p)} = \frac{1}{2}.$$

Alors pour tout $n \in \mathbf{N}$, on a

$$p_n - \frac{1}{2} = (2p - 1)^{n-1} \left(p_1 - \frac{1}{2} \right) = (2p - 1)^{n-1} \left(1 - \frac{1}{2} \right) = \frac{1}{2} (2p - 1)^{n-1}.$$

Donc :

$$p_n = \frac{1}{2} + \frac{1}{2} (2p - 1)^{n-1}.$$

Puisque $p \neq 1$, $(2p - 1)^{n-1} \xrightarrow{n \rightarrow \infty} 0$ et donc $p_n \xrightarrow{n \rightarrow \infty} \frac{1}{2}$. Il y a donc une chance sur deux qu'elle soit bien transmise.

Exercice ALEA.12.18 | Relation de récurrence Un fumeur cherche à arrêter de fumer chaque jour. On note p_n la probabilité qu'il fume le jour n .

- ▶ S'il a réussi à ne pas fumer un jour, il ne fume pas le lendemain avec une probabilité $\frac{1}{2}$.

► S'il fume un jour, il ne fume pas le lendemain avec une probabilité $\frac{1}{4}$.

1. Chercher une relation de récurrence sur la suite (p_n) .
2. Exprimer p_n en fonction de $n \in \mathbf{N}$ pour tout $n \in \mathbf{N}$.
3. Déterminer $\lim_{n \rightarrow \infty} p_n$.

Solution (exercice ALEA.12.18)

1. Notons F_n l'évènement «fume le jour n ». Alors puisque $(F_n, {}^cF_n)$ est un système complet d'évènements pour tout $n \in \mathbf{N}$, nous avons

$$\begin{aligned} \mathbf{P}(F_{n+1}) &= \mathbf{P}(F_{n+1}|F_n)\mathbf{P}(F_n) + \mathbf{P}(F_{n+1}|{}^cF_n)\mathbf{P}({}^cF_n) \\ &= \left(1 - \frac{1}{4}\right)p_n + \frac{1}{2}(1 - p_n) \\ &= \frac{3}{4}p_n - \frac{1}{2}p_n + \frac{1}{2} \\ &= \boxed{\frac{1}{4}p_n + \frac{1}{2}}. \end{aligned}$$

2. Cherchons $C \in \mathbf{R}$ de sorte que

$$C = \frac{1}{4}C + \frac{1}{2} \iff C = \frac{2}{3}.$$

Alors $(p_n - C)$ est géométrique de raison $\frac{1}{4}$, et donc

$$\forall n \in \mathbf{N}, \quad p_n - \frac{2}{3} = \frac{1}{4^{n-1}} \left(p_1 - \frac{2}{3}\right),$$

d'où

$$\forall n \in \mathbf{N}, \quad \boxed{p_n = \frac{2}{3} + \frac{1}{4^{n-1}} \left(p_1 - \frac{2}{3}\right)}.$$

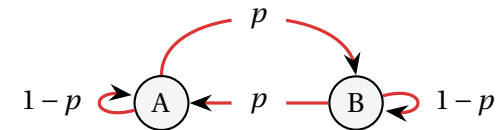
3. On déduit $\boxed{p_n \xrightarrow{n \rightarrow \infty} \frac{2}{3}}$. Il y a deux chances sur 3 qu'il continue à fumer lorsque le nombre de jours grandit.

Exercice ALEA.12.19 | Relation vectorielle en dimension 2 Une abeille va chaque jour sur l'une des deux fleurs A et B. Au jour 0, elle va à la fleur A. À chaque nouvelle journée, il y a une probabilité $p \in]0, 1[$ qu'elle aille sur la même fleur que la veille. Pour tout entier n , on note A_n l'évènement «l'abeille est sur la fleur A le jour n » et B_n l'évènement «l'abeille est sur la fleur B le jour n ». On pose de plus $a_n = \mathbf{P}(A_n)$ et $b_n = \mathbf{P}(B_n)$ pour tout $n \in \mathbf{N}$.

1. Pour tout entier n , exprimer a_{n+1} et b_{n+1} en fonction de a_n et b_n .
2. Déterminer les expressions explicites de a_n et b_n en fonction de $n \in \mathbf{N}$.
3. Vers quoi tendent les deux suites? Interpréter.

Solution (exercice ALEA.12.19)

On peut représenter les transitions sur un graphe de sommets les noms de fleurs.



1. Soit $n \in \mathbf{N}$. Alors constatons que (A_n, B_n) forme un système complet d'évènements. Donc par formule des probabilités totales,

$$\begin{aligned} a_{n+1} &= \mathbf{P}(A_{n+1}) \\ &= \mathbf{P}(A_{n+1}|A_n)\mathbf{P}(A_n) + \mathbf{P}(A_{n+1}|B_n)\mathbf{P}(B_n) \\ &= \boxed{(1 - p)\mathbf{P}(A_n) + p\mathbf{P}(B_n)} \end{aligned}$$

$$\begin{aligned} b_{n+1} &= \mathbf{P}(B_{n+1}) \\ &= \mathbf{P}(B_{n+1}|A_n)\mathbf{P}(A_n) + \mathbf{P}(B_{n+1}|B_n)\mathbf{P}(B_n) \\ &= \boxed{p\mathbf{P}(A_n) + (1 - p)\mathbf{P}(B_n)}. \end{aligned}$$

2. Il s'agit ici de deux suites récurrentes linéaires imbriquées, on passe donc par des matrices. Notons $X_n = \begin{pmatrix} a_n \\ b_n \end{pmatrix}$ pour tout $n \in \mathbf{N}$, de sorte que

$$X_{n+1} = AX_n, \quad \text{avec} \quad A = \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix}.$$

Diagonalisons la matrice A. Puisque la somme sur chaque ligne est constante égale à 1, on sait que $1 \in \text{Spec} A$ et $\lambda = 1$ est une valeur propre dont $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ est un vecteur propre associé. Soit $\lambda \in \mathbf{C}$, alors

$$A - \lambda I_2 = \begin{pmatrix} 1-p-\lambda & p \\ p & 1-p-\lambda \end{pmatrix} \begin{array}{l} \left. \begin{array}{l} \phantom{\tilde{L}} \\ \phantom{\tilde{L}} \end{array} \right\} L_1 \leftrightarrow L_2 \\ \tilde{L} \begin{pmatrix} p & 1-p-\lambda \\ 1-p-\lambda & p \end{pmatrix} \\ \left. \begin{array}{l} \phantom{\tilde{L}} \\ \phantom{\tilde{L}} \end{array} \right\} L_2 \leftarrow pL_2 - (1-p-\lambda)L_1 \\ \tilde{L} \begin{pmatrix} p & 1-p-\lambda \\ 0 & p^2 - (1-p-\lambda)^2 \end{pmatrix} \end{array}$$

De plus, $p^2 - (1-p-\lambda)^2 = (p-1+p+\lambda)(1-\lambda) = (2p-1+\lambda)(1-\lambda)$. Donc $\text{Spec} A = \{1, 1-2p\}$. Mais $1-2p = 1 \iff p = 1$ ce qui est exclu, donc A possède donc bien deux valeurs propres distinctes et est donc diagonalisable. De plus, les deux espaces propres sont de dimension 1. Et $X = \begin{pmatrix} x \\ y \end{pmatrix} \in E_{1-2p}(A)$ si et seulement si

$$px + (1-p-1+2p)y = 0 \iff x + y = 0,$$

donc $E_{1-2p}(A) = \text{Vect} \left(\begin{pmatrix} 1 \\ -1 \end{pmatrix} \right)$. Ainsi

$$A = P \cdot D \cdot P^{-1}, \quad P = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1-2p & 0 \\ 0 & 1 \end{pmatrix}.$$

Ensuite, notons $Y_n = P^{-1}X_n$. Alors $Y_{n+1} = P^{-1}X_{n+1} = P^{-1}PDP^{-1}X_n$. Donc $Y_{n+1} = DY_n$. Si on note $Y_n = \begin{pmatrix} a'_n \\ b'_n \end{pmatrix}$, on a alors

$$\forall n \in \mathbf{N}, \quad a'_{n+1} = (1-2p)a'_n, \quad b'_{n+1} = b'_n.$$

Et donc

$$\forall n \in \mathbf{N}, \quad a'_n = (1-2p)^n a'_0, \quad b'_n = b'_0.$$

Mais $a_0 = 1, b_0 = 0$ donc

$$Y_0 = P^{-1} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}.$$

On déduit que

$$X_n = PY_n = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{2} \left((1-2p)^n \right) \\ \frac{1}{2} \left(1 \right) \end{pmatrix} = \frac{1}{2} \begin{pmatrix} (1-2p)^n + 1 \\ -(1-2p)^n + 1 \end{pmatrix}.$$

Or $|1-2p| < 1$ donc $\lim_{n \rightarrow \infty} a_n = \frac{1}{2}, \lim_{n \rightarrow \infty} b_n = \frac{1}{2}$.

Finalement, en temps long, cela revient à considérer que l'abeille choisit chaque fleur de manière équiprobable! Ceci n'est pas surprenant puisque les transitions $A \rightarrow B$ et $B \rightarrow A$ ne privilégie pas une fleur plutôt que l'autre.

Exercice ALEA.12.20 | Relation vectorielle en dimension 3 Dans tout l'énoncé p est un réel de l'intervalle $]0, 1[$ et $q = 1 - p$. Sur une table sont placées deux boules noires (étape 0), puis une des deux boules est choisie au hasard et éliminée de la table. Ensuite on repose sur la table (étape 1) :

- ▶ soit une boule blanche, avec la probabilité p ,
- ▶ soit une boule noire, avec la probabilité q .

Cette action est répétée ainsi indéfiniment, de sorte qu'à la $k \in \mathbf{N}$ -ième itération de l'expérience, deux boules sont sur la table :

- ▶ soit deux noires (événement noté NN_k),
- ▶ soit une noire et une blanche (événement noté NB_k),

► soit deux blanches (événement noté BB_k).

Pour $k \in \mathbf{N}$, on note également $a_k = \mathbf{P}(NN_k)$, $b_k = \mathbf{P}(NB_k)$, $c_k = \mathbf{P}(BB_k)$, et on définit les matrices suivantes :

$$M = \begin{pmatrix} q & q/2 & 0 \\ p & 1/2 & q \\ 0 & p/2 & p \end{pmatrix}, \quad D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & q & q^2 \\ -2 & p-q & 2pq \\ 1 & -p & p^2 \end{pmatrix},$$

$$U_k = \begin{pmatrix} a_k \\ b_k \\ c_k \end{pmatrix}, \quad U_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

1. **1.1)** Calculer les produits PD , MP . Que dire de $MP - PD$?
1. **1.2)** En déduire que M est diagonalisable.
2. **2.1)** Donner a_0, b_0, c_0 . Justifier $a_1 = q$, $b_1 = p$ et $c_1 = 0$.
2. **2.2)** Montrer que pour tout entier naturel k non nul : $U_{k+1} = MU_k$.
3. En déduire que pour tout $k \in \mathbf{N}^*$:

$$U_k = PD^k \begin{pmatrix} p^2 \\ 2p \\ 1 \end{pmatrix}.$$

4. En déduire pour tout entier naturel $k \in \mathbf{N}^*$, a_k, b_k, c_k en fonction de k et montrer que :

$$\lim_{k \rightarrow +\infty} a_k = q^2, \quad \lim_{k \rightarrow +\infty} b_k = 2pq, \quad \lim_{k \rightarrow +\infty} c_k = p^2.$$

5. Interpréter.

Solution (exercice ALEA.12.20)

1. **1.1)** On a

$$P \cdot D = \begin{pmatrix} 0 & \frac{q}{2} & q^2 \\ 0 & \frac{-p+q}{2} & 2pq \\ 0 & \frac{-p}{2} & p^2 \end{pmatrix}, \quad M \cdot P = \begin{pmatrix} 0 & \frac{q}{2}(q+p) & q^2(q+p) \\ p+q-1 & \frac{-p+q}{2} & pq(q+p+1) \\ 0 & -\frac{p}{2}(q+p) & p^2(p+q) \end{pmatrix}.$$

En utilisant la relation $p+q=1$, on déduit que $P \cdot D = M \cdot P$. On a donc obtenu que $MP - PD = 0$.

1. **1.2)** Donc puisque P est inversible, on a obtenu :

$$M = PDP^{-1}.$$

Donc M est diagonalisable, et $\text{Spec}(M) = \{0, \frac{1}{2}, 1\}$.

2. **2.1)** On a $a_0 = \mathbf{P}(NN_0) = 1$ car $NN_0 = \Omega$, et $b_0 = c_0 = 0$. Pour a_1 , étant donné qu'il y a forcément deux boules noires au temps zéro, la probabilité d'avoir deux noires au temps 1 est celle de l'ajout d'une boule noire à la première étape, donc q .

$$\begin{aligned} a_1 &= \mathbf{P}(NN_1) = q \\ c_1 &= \mathbf{P}(BB_1) = 0, \quad (BB_1 = \emptyset) \\ b_1 &= 1 - a_1 - c_1 \\ &= 1 - q = p. \end{aligned}$$

2. **2.2)** Soit k un entier non nul. Le système (NN_k, NB_k, BB_k) est un système complet d'évènements. Donc d'après la formule des probabilités totales, on a (on donne les explications pour la première ligne uniquement pour alléger

la rédaction) :

$$\begin{aligned}
 & \mathbf{P}(\text{NN}_{k+1}) \\
 &= \mathbf{P}(\text{NN}_{k+1} | \text{NN}_k) \mathbf{P}(\text{NN}_k) + \mathbf{P}(\text{NN}_{k+1} | \text{NB}_k) \mathbf{P}(\text{NB}_k) + \mathbf{P}(\text{NN}_{k+1} | \text{BB}_k) \mathbf{P}(\text{BB}_k) \\
 &= \boxed{q} \mathbf{P}(\text{NN}_k) + \boxed{\frac{1}{2} \cdot q} + 0 \cdot \mathbf{P}(\text{BB}_k) \\
 & \quad \text{pioche une N puis rempl. par N} \qquad \text{pioche B puis rempl. par N} \\
 & \mathbf{P}(\text{NB}_{k+1}) \\
 &= \mathbf{P}(\text{NB}_{k+1} | \text{NN}_k) \mathbf{P}(\text{NN}_k) + \mathbf{P}(\text{NB}_{k+1} | \text{NB}_k) \mathbf{P}(\text{NB}_k) + \mathbf{P}(\text{NB}_{k+1} | \text{BB}_k) \mathbf{P}(\text{BB}_k) \\
 &= p \cdot \mathbf{P}(\text{NN}_k) + \boxed{\frac{1}{2} \cdot q + \frac{1}{2} \cdot p} \mathbf{P}(\text{NB}_k) \\
 & \quad \text{pioche B puis rempl. par B ou pioche N puis rempl. par B} \\
 & + q \cdot \mathbf{P}(\text{BB}_k) \\
 & \mathbf{P}(\text{BB}_{k+1}) \\
 &= \mathbf{P}(\text{BB}_{k+1} | \text{NN}_k) \mathbf{P}(\text{NN}_k) + \mathbf{P}(\text{BB}_{k+1} | \text{NB}_k) \mathbf{P}(\text{NB}_k) + \mathbf{P}(\text{BB}_{k+1} | \text{BB}_k) \mathbf{P}(\text{BB}_k) \\
 &= 0 \cdot \mathbf{P}(\text{NN}_k) + \frac{p}{2} \cdot \mathbf{P}(\text{NB}_k) + p \cdot \mathbf{P}(\text{BB}_k).
 \end{aligned}$$

Les trois relations forment alors la relation matricielle

$$\boxed{\mathbf{U}_{k+1} = \mathbf{M} \cdot \mathbf{U}_k}$$

3. Pour $k = 1$, on a $\mathbf{U}_1 = \begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix} = \begin{pmatrix} q \\ p \\ 0 \end{pmatrix}$. D'autre part

$$\text{PD} \begin{pmatrix} p^2 \\ 2p \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & q & q^2 \\ -2 & p-q & 2pq \\ 1 & -p & p^2 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} p^2 \\ 2p \\ 1 \end{pmatrix} = \begin{pmatrix} q \\ p \\ 0 \end{pmatrix}.$$

La relation est donc initialisée pour $k = 1$. Supposons-là vraie au rang k , alors

$$\begin{aligned}
 \mathbf{U}_{k+1} &= \text{PDP}^{-1} \mathbf{U}_k \\
 &= \text{PDP}^{-1} \text{PD}^k \begin{pmatrix} p^2 \\ 2p \\ 1 \end{pmatrix} \quad \text{hypothèse de récurrence} \\
 &= \text{PD}^{k+1} \begin{pmatrix} p^2 \\ 2p \\ 1 \end{pmatrix}.
 \end{aligned}$$

Par principe de récurrence, on déduit que pour tout $k \in \mathbf{N}^*$:

$$\boxed{\mathbf{U}_k = \text{PD}^k \begin{pmatrix} p^2 \\ 2p \\ 1 \end{pmatrix}}$$

4. Ainsi, pour tout $k \in \mathbf{N}^*$,

$$\mathbf{U}_k = \begin{pmatrix} 1 & q & q^2 \\ -2 & p-q & 2pq \\ 1 & -p & p^2 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1/2^k & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} p^2 \\ 2p \\ 1 \end{pmatrix} = \begin{pmatrix} q \left(\frac{p}{2^{k-1}} \right) \\ 2p \left(\frac{p-q}{2^k} + q \right) \\ p^2 \left(1 - \frac{1}{2^{k-1}} \right) \end{pmatrix}.$$

Donc en faisant $k \xrightarrow{k \rightarrow \infty} \infty$ dans chaque coordonnée, on obtient

$$\boxed{\lim_{k \rightarrow +\infty} a_k = q^2, \quad \lim_{k \rightarrow +\infty} b_k = 2pq, \quad \lim_{k \rightarrow +\infty} c_k = p^2}$$

5. Quand on attend suffisamment longtemps, on a donc une probabilité q^2 d'avoir deux boules noires, $2pq$ d'avoir une noire et une boule blanche et p^2 d'avoir deux boules blanches. Ces probabilités correspondent à celle d'obtenir les configurations BB, NB, NN lorsque l'on fait deux tirages successifs dans un ensemble où la probabilité d'obtenir une noire est q et celle d'obtenir une blanche p .

Chapitre ALEA.13.

Variables aléatoires discrètes


Résumé & Plan

Nous avons étudié dans le [Chapter ALEA.12](#) les espaces probabilisés, cadre pour définir sur ces espaces des variables et vecteurs aléatoires. Dans ce chapitre, nous allons étudier ces objets à valeurs dans un ensemble au plus dénombrable. Dans ce cas précis, nous allons donner une expression de l'espérance en terme de somme de série, qui généralisera la définition vue en première année.

1	Généralités	2	3	Lois discrètes usuelles	22
1.1	Définitions	2	3.1	Loi uniforme discrète	24
1.2	Loi & fonction de répartition	5	3.2	Loi de BERNOULLI & binomiale	26
1.3	Propriétés des variables aléatoires réelles discrètes	9	3.3	Loi Hypergéométrique	30
1.4	Indépendance	12	3.4	Loi Géométrique & Absence de mémoire	32
2	Espérance, Variance, Moments	14	3.5	Loi de POISSON	36
2.1	Espérance	14	3.6	Bilan des lois discrètes	38
2.2	Moments d'ordre supérieur	19	4	Exercices	39
			4.1	Généralités	39
			4.2	Expérience aléatoire	45
			4.3	Matrices aléatoires	54

La notion d'espérance a été initialement introduite par HUYGENS en 1657, dans son traité De Ratiociniis in Aleae Ludo («de la logique du jeu de dé»). Le nom d'espérance y apparaît en latin sous le nom de expectatio, avec l'interprétation d'être «le juste prix auquel un joueur accepterait de céder sa place dans une partie».

— Le saviez-vous ?

 **Cadre**
 Dans tout le chapitre, et même lorsque cela n'est pas précisé, on se fixe un espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$.

Nous avons déjà motivé l'introduction de variables aléatoires, dans le chapitre précédent.

Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$. Si Ω était fini — *i.e.* le contexte de première année — on pouvait donc l'écrire sous la forme $\Omega = \{\omega_1, \dots, \omega_n\}$ avec $n = \#\Omega$. Ainsi

$$X(\Omega) = \{X(\omega_1), \dots, X(\omega_n)\} \quad \text{est aussi un ensemble fini.}$$

Puisque nos univers sont désormais des ensembles quelconques, le support $X(\Omega)$ devient lui aussi un ensemble quelconque. Nous étudions dans la suite le cas particulier où il est au plus dénombrable, c'est l'objet précisément de ce chapitre.

1. GÉNÉRALITÉS

On ne rappelle pas la définition de variable aléatoire vue précédemment : consulter pour cela le [Chapter ALEA.12](#) si besoin.

1.1. Définitions

Définition ALEA.13.1 | Variable aléatoire discrète

Soit X une variable aléatoire.

- ▶ On dit que X est *discrète* si $X(\Omega)$ est un ensemble au plus dénombrable, *i.e.* en bijection avec \mathbf{N} **ou** fini.
- ▶ Si $X(\Omega)$ est un ensemble fini, on parle de *variable aléatoire discrète finie*.

Pour rappel, l'hypothèse « $X(\Omega)$ est au plus dénombrable» signifie que cet ensemble s'écrit sous la forme :

- ▶ $X(\Omega) = \{x_n, n \in \mathbf{N}\}$ s'il est dénombrable, les $x_n, n \in \mathbf{N}$ étant deux à deux distincts,
- ▶ $X(\Omega) = \{x_n, n \in \llbracket 0, N \rrbracket\}$ avec $N \in \mathbf{N}$ s'il est fini.

Notation Somme infinie sur un ensemble dénombrable

Considérons E un ensemble au plus dénombrable et $f : E \rightarrow \mathbf{R}$, on souhaite donner un sens à la quantité $\sum_{x \in E} f(x)$.

- ▶ Si E est fini, alors $E = \{x_n, n \in \llbracket 0, N \rrbracket\}$ avec $N \in \mathbf{N}$. Alors on pose :

$$\sum_{x \in E} f(x) \stackrel{\text{(défi.)}}{=} \sum_{i=1}^N f(x_i).$$

C'est une somme finie de première année, rien de plus à dire puisqu'elle ne dépend pas de la numérotation des éléments de E .

- ▶ Si E est dénombrable, alors $E = \{x_n, n \in \mathbf{N}\}$. Alors, **si** $(\sum f(x_n))_{n \geq 0}$ converge absolument, on pose :

$$\sum_{x \in E} f(x) \stackrel{\text{(défi.)}}{=} \sum_{n=0}^{\infty} f(x_n).$$

En effet, on sait depuis le [Chapter ANA.10](#), qu'en cas de convergence absolue, la somme ainsi définie ne dépendra pas de la numérotation (x_n) choisie pour E . Ainsi, pour utiliser cette notation, il faudra d'abord justifier la convergence absolue.

Lorsque le support est au plus dénombrable, la condition **VaProp** du **Chapter ALEA.12** (définition d'une variable aléatoire) se réécrit comme ci-dessous.

Proposition ALEA.13.1 | Réécriture de l'hypothèse « $\{X \leq a\} \in \mathcal{F}$ »

Soit $X : \Omega \rightarrow \mathbf{R}$ avec $X(\Omega)$ au plus dénombrable. Alors :

$$X \text{ est une variable aléatoire discrète} \iff \forall x \in X(\Omega), \{X = x\} \in \mathcal{F}.$$

Preuve

\Rightarrow Notons $X(\Omega) = \{x_n, n \in \mathbf{N}\}$ avec (x_n) supposée croissante, et soit $x = x_n \in X(\Omega)$ avec $n \in \mathbf{N}$. Alors

$$\{X = x_n\} = \{X \leq x_n\} \setminus \{X \leq x_{n-1}\} \in \mathcal{F}$$

par propriété de stabilité par différence d'une tribu.

\Leftarrow Supposons que pour tout $x \in X(\Omega)$, $\{X = x\} \in \mathcal{F}$. Alors

$$\forall a \in \mathbf{R}, \{X \leq a\} = \bigcup_{\substack{x \leq a \\ x \in X(\Omega)}} \{X = x\} \in \mathcal{F}$$

par propriété de stabilité par réunion dénombrable d'une tribu.

Proposition ALEA.13.2 | Variable aléatoire définie sur un univers au plus dénombrable

Soit $X : \Omega \rightarrow \mathbf{R}$ une variable aléatoire avec Ω au plus dénombrable. Alors :

X est une variable aléatoire discrète.


Preuve




Remarque 1.1 — Ainsi, sur les univers finis de première année, il ne pouvait exister que des variables aléatoires discrètes.

Exemple 1 — Cas fini : lancé d'un dé et gain associé Un jeu de hasard consiste à lancer un dé équilibré à 6 faces. Le lanceur :

1. gagne le double de la valeur de la face obtenue si celle-ci est paire.
2. Sinon, il perd le double de la valeur indiquée par le dé.

Décrire l'expérience en précisant un triplet $(\Omega, \mathcal{F}, \mathbf{P})$ associé ainsi qu'une variable aléatoire X décrivant le gain, c'est une variable aléatoire discrète. 

Exemple 2 – Cas dénombrable On lance une infinité de fois une pièce non truquée. On note X la variable aléatoire qui donne le nombre de lancers nécessaires jusqu'à obtention du premier pile. Décrire l'expérience en précisant un triplet $(\Omega, \mathcal{F}, \mathbf{P})$ associé ainsi qu'une variable aléatoire X . On admettra qu'il existe une tribu \mathcal{F} telle que X soit bien une variable aléatoire. 

Définition/Proposition ALEA.13.1 | Système complet associé à une variable aléatoire discrète


Si X est une variable aléatoire réelle discrète, alors :

- ▶ $\{X = x\}_{x \in X(\Omega)}$ est un système complet d'évènements.
- ▶ En particulier, il est quasi-complet :

$$\sum_{x \in X(\Omega)} \mathbf{P}(X = x) = 1, \quad \{X = x, X = x'\} = \emptyset, \quad \forall x \neq x' \in X(\Omega).$$

On l'appelle le *système complet associé à X* .

Preuve

1. Existence de $\sum_{x \in X(\Omega)} \mathbf{P}(X = x)$. 

2. On a clairement $\{X = x\} \cap \{X = x'\} = \emptyset$ dès que $x \neq x'$ sont deux éléments de $X(\Omega)$, par définition même d'une application (un élément de l'espace de départ ne peut avoir deux images!). Reste à montrer que :

$$\bigcup_{x \in X(\Omega)} \{X = x\} = \Omega.$$

On a

$$\bigcup_{x \in X(\Omega)} \{X = x\} \subset \Omega,$$

puisque'une réunion de parties de Ω est encore une partie de Ω . Inversement, soit $\omega \in \Omega$, alors $X(\omega) \in X(\Omega)$ donc ω est dans l'un des $\{X = x\}$ pour $x \in X(\Omega)$, ce qui prouve

$$\Omega \subset \bigcup_{x \in X(\Omega)} \{X = x\}.$$

SYSTÈME COMPLET ASSOCIÉ.

1.2. Loi & fonction de répartition

Rappelons que nous avons défini la loi de X comme l'application qui à tout intervalle réel I associe

$$\mathbf{P}_X(I) = \mathbf{P}(X \in I).$$

Comment simplifier cette définition dans le cas de variables aléatoires discrètes?

$$\{X \in I\} = \{X \in I \cap X(\Omega)\}, \quad \text{donc} \quad \mathbf{P}(X \in I) = \mathbf{P}(X \in I \cap X(\Omega)).$$

Or,

$$\{X \in I \cap X(\Omega)\} = \bigsqcup_{\substack{x \in X(\Omega) \\ x \in I}} \{X = x\} \implies \mathbf{P}(X \in I \cap X(\Omega)) = \sum_{\substack{x \in X(\Omega) \\ x \in I}} \mathbf{P}(X = x).$$

Ainsi, pour obtenir la loi, il suffit de connaître :

- ▶ d'une part l'univers-image $X(\Omega)$,
- ▶ et d'autre part tous les $\mathbf{P}(X = x)$ pour tout $x \in X(\Omega)$.

Ce constat nous mène tout droit à la définition suivante.

Définition ALEA.13.2 | Loi d'une variable aléatoire discrète

Soit X une variable aléatoire réelle discrète.

- ▶ **(Loi)** On appelle *loi* de la variable aléatoire réelle discrète X — par abus de langage — ou parfois *fonction de masse* la fonction encore notée \mathbf{P}_X et définie par :

$$\mathbf{P}_X : x \in X(\Omega) \longrightarrow \mathbf{P}(X = x).$$

- ▶ **(Déterminer la loi)** Déterminer la loi d'une variable aléatoire discrète c'est calculer $X(\Omega)$ et $\mathbf{P}(X = x)$ pour tout $x \in X(\Omega)$.
- ▶ **(Avoir même loi que)** Soit Y une autre variable aléatoire discrète définie sur $(\Omega, \mathcal{F}, \mathbf{P})$. On dit que X, Y ont même loi si :

$$\forall z \in X(\Omega) \cup Y(\Omega)^1, \quad \mathbf{P}(X = z) = \mathbf{P}(Y = z).$$


¹la plupart du temps, dans des contextes d'expériences aléatoires, les ensembles $X(\Omega), Y(\Omega)$ seront

Notation


- ▶ Lorsque X, Y ont même loi, on note $X \sim Y$.
- ▶ Si Y suit une loi usuelle \mathcal{L} (une BERNOULLI, binomiale, etc.), on note $X \hookrightarrow \mathcal{L}$.

Exemple 3 — Une pièce amène pile avec la probabilité p et face avec la probabilité $1 - p, 0 < p < 1$. On la lance n fois de suite. Soit X le nombre de fois où pile apparaît au cours de ces lancers. Chercher la loi de X sans utiliser de résultat sur les lois usuelles.



Exemple 4 — Variable aléatoire constante Soit $C \in \mathbf{R}$. Alors C peut être vu comme une variable aléatoire constante, et on peut déterminer sa loi. 

Exemple 5 — Variable aléatoire indicatrice Soit $(\Omega, \mathcal{F}, \mathbf{P})$ un espace probabilisé et $A \in \mathcal{F}$ un évènement. Alors l'application $\mathbb{1}_A$ est une variable aléatoire sur $(\Omega, \mathcal{F}, \mathbf{P})$, et égaux

on peut déterminer sa loi. 


Proposition ALEA.13.3 | La loi ne dit rien sur l'égalité de deux variables aléatoires

Soient X, Y deux variables aléatoires définies sur $(\Omega, \mathcal{F}, \mathbf{P})$.

1. $X = Y \implies X \sim Y$.
2. En revanche, la réciproque est fautive : si $X \sim Y$ alors X n'est pas forcément égale à Y .

Preuve

1. 


2. **est faux.** Par exemple si $(\Omega, \mathcal{F}, \mathbf{P})$ est l'espace probabilisé associé au lancer d'une pièce **équilibrée**. On note X la variable aléatoire égale à 1 si on obtient pile, égale à 0 si on obtient face. On note Y la variable aléatoire égale à 1 si on obtient face, égale à 0 si on obtient pile. 

Définition ALEA.13.3 | Loi conditionnelle

Soit X une variable aléatoire réelle discrète et $A \in \mathcal{F}$ tel que $\mathbf{P}(A) \neq 0$. On appelle *loi conditionnelle sachant A* de X la fonction

$$\mathbf{P}_X(\cdot | A) : x \in X(\Omega) \longrightarrow \mathbf{P}(X = x | A) = \frac{\mathbf{P}(\{X = x\} \cap A)}{\mathbf{P}(A)}.$$

Notation
Si la loi conditionnelle précédente est une loi usuelle \mathcal{L} , on notera $X \hookrightarrow \big|_A \mathcal{L}$.

Exemple 6 — Lancé d'un dé et gain associé conditionnel On reconsidère l'**Exemple 1**. Déterminer la loi conditionnelle de X sachant A «le résultat du dé est pair», même question sachant B «le résultat du dé est un multiple de 3». 

Résumons les points précédents dans une méthode.

Méthode Répondre à la question «déterminer la loi d'une variable aléatoire réelle discrète X »

1. Commencer par déterminer son support $X(\Omega)$ s'il n'est pas déjà donné *i.e.* l'ensemble de départ de \mathbf{P}_X .
2. Calculer les $\mathbf{P}(X = x)$ pour tout $x \in X(\Omega)$. Si $X(\Omega)$ est fini, il n'y a donc qu'un nombre fini de probabilités à déterminer, on peut alors les représenter sous forme d'un tableau comme ceci :



$X = k$	k_1	k_2	...
$\mathbf{P}(X = k)$	$\mathbf{P}(X = k_1)$	$\mathbf{P}(X = k_2)$...

LIEN ENTRE LOI & FONCTION DE RÉPARTITION.

Proposition ALEA.13.4 | Lien avec la fonction de répartition

Soit X est une variable aléatoire discrète.

1. Si $X(\Omega)$ est fini, notons $X(\Omega) = \{x_0, \dots, x_N\}$ où $(x_n)_{0 \leq n \leq N}$ est supposée croissante. Alors pour tout $n \in \llbracket 1, N \rrbracket$:

$$\mathbf{P}_X(x_n) = F_X(x_n) - F_X(x_{n-1}).$$

2. Si $X(\Omega)$ est dénombrable, notons $X(\Omega) = \{x_n, n \in \mathbf{N}\}$ où (x_n) est supposée croissante, alors pour tout $n \in \mathbf{N}^*$:

$$\mathbf{P}_X(x_n) = F_X(x_n) - F_X(x_{n-1}).$$

En particulier, si $X(\Omega) = \mathbf{Z}$ ou \mathbf{N} , on a :

$$\forall k \in \mathbf{Z} \text{ ou } \mathbf{N}, \quad \mathbf{P}_X(k) = F_X(k) - F_X(k-1).²$$

Preuve Faisons, pour simplifier, la preuve dans le cas $X(\Omega) = \mathbf{N}$.

Corollaire ALEA.13.1 | La fonction de répartition caractérise la loi

Soient X, Y deux variables aléatoires discrètes. Alors :

$$X \sim Y \iff F_X = F_Y.$$

On dit aussi que la fonction de répartition *caractérise la loi* au sens suivant : deux variables aléatoires discrètes ont même loi si et seulement si elles ont même fonction de répartition.

Preuve Faisons, pour simplifier, la preuve dans le cas $X(\Omega) = \mathbf{N}$.



EXISTENCE D'UNE VARIABLE ALÉATOIRE RÉELLE DISCRÈTE DE LOI FIXÉE. Un certain nombre d'énoncés de probabilité commencent par la phrase suivante :

«soit X une variable aléatoire telle que $\mathbf{P}(X = x_i) = p_i$ » avec $p_i \in [0, 1]$, i entier et (x_i) une famille.

Ces énoncés supposent l'existence de X . Mais cela ne définit pas X en tant qu'application, c'est-à-dire $X(\omega)$ pour tout $\omega \in \Omega$, mais existe-t-elle vraiment? On aimerait donc au moins savoir si une telle variable aléatoire existe sur un certain $(\Omega, \mathcal{F}, \mathbf{P})$ à trouver : la réponse est oui dès que la somme des p_i supposés positifs vaut un, comme le

²Cette égalité est en fait une reformulation dans le cas discret de l'égalité $\mathbf{P}(X = x) = F_X(x) - F_X(x-)$ vue dans la [Section 3](#) du [Chapter ALEA.12](#) pour les variables aléatoires générales.

précise le théorème qui suit.

Théorème ALEA.13.1 | Variable aléatoire associé à une famille de somme un

Soit $\mathbf{X} = \{x_i, i \in \mathbf{N}\}$ (resp. $\mathbf{X} = \{x_i, i \in \llbracket 0, \mathbf{N} \rrbracket\}$ avec $\mathbf{N} \in \mathbf{N}$) une partie au plus dénombrable de \mathbf{R} , et $(p_i)_{i \in \mathbf{N}}$ (resp. $(p_i)_{i \in \llbracket 1, \mathbf{N} \rrbracket}$) une famille de réels tels que :

$$\sum_{i \in \mathbf{N}} p_i = 1 \quad \left(\text{resp.} \sum_{i=1}^{\mathbf{N}} p_i = 1 \right) \quad \text{et} \quad p_i \geq 0.$$

Alors il existe un espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$ et une variable aléatoire réelle $X : \Omega \rightarrow \mathbf{R}$ discrète tels que :

- ▶ $X(\Omega) = \mathbf{X}$.
- ▶ Pour tout i , $\mathbf{P}(X = x_i) = p_i$.

Dans la plupart des situations que nous étudierons en pratique, le travail commencera par la donnée d'une ou plusieurs variables aléatoires de lois prescrites que nous étudierons sans jamais préciser l'espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$ sur lequel elles sont définies.

Celui-ci est voué à rester caché et ne présente de toute façon aucun caractère d'unicité, et de nombreux choix d'espace probabilisé sont possibles pour la description d'une même situation (en voici un dans la preuve ci-après).

Preuve Faisons la preuve par exemple dans le cas dénombrable $\mathbf{X} = \{x_i, i \in \mathbf{N}\}$.

1. On cherche un exemple, rien qu'un, d'espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$ et de variable aléatoire qui répondent au problème posé. Posons alors :


$$\begin{aligned} (\Omega, \mathcal{F}) &= (\mathbf{X}, \mathcal{P}(\mathbf{X})), \\ X &= \text{Id}_\Omega = \text{Id}_\mathbf{X}, \end{aligned}$$


puis enfin pour \mathbf{P} l'unique probabilité telle que : $\mathbf{P}(\{x_i\}) = p_i$ pour tout $i \in \mathbf{N}$ — i.e. celle définie dans la [Section 2.4](#) du [Chapter ALEA.12](#). Alors ce choix convient, puisque

- ▶ $X(\Omega) = \text{Id}_\Omega(\Omega) = \Omega$,


- ▶ et si $i \in \mathbf{N}$ alors

$$\begin{aligned} \mathbf{P}(X = x_i) &= \mathbf{P}(\text{Id}_\Omega = x_i) \\ &= \mathbf{P}(\{\omega \in \Omega, \omega = x_i\}) \\ &= \mathbf{P}(x_i) = p_i. \end{aligned}$$

Exemple 7 — Soit X une variable aléatoire réelle discrète telle que : $X(\Omega) = \mathbf{N}^*$ et pour tout $n \geq 1$, $\mathbf{P}(X = n) = \frac{1}{n(n+1)}$. Justifier l'existence de X et calculer F_X . 

Exemple 8 — Existe-t-il une variable aléatoire réelle discrète X de support $X(\Omega) = \mathbf{N}^*$ de loi donnée par les $\left(\frac{1}{2^n}\right)_{n \geq 1}$? 

1.3. Propriétés des variables aléatoires réelles discrètes

Exemple 9 — Soit $\lambda > 0$. À quelle condition sur $a \in \mathbf{R}$ existe-t-il une variable aléatoire réelle discrète X de support $X(\Omega) = \mathbf{N}$ de loi donnée par les $(p_n)_{n \geq 0} = \left(a \frac{\lambda^n}{n!}\right)_{n \geq 0}$? 

Dans la **Section 3** nous avons vu qu'une combinaison linéaire de variables aléatoires réelles est une variable aléatoire réelle. Tout ceci reste vrai pour les variables aléatoires réelles discrètes et en plus toutes les variables aléatoires obtenues par ces opérations sont encore discrètes.

Proposition ALEA.13.5 | Structure d'espace vectoriel, opérations sur les variables aléatoires réelles discrètes

1. L'ensemble des variables aléatoires discrètes, muni de l'addition et de la multiplication de réels, est un \mathbf{R} -espace vectoriel, *i.e.* si X, Y sont discrètes, $(\lambda, \mu) \in \mathbf{R}^2$, alors : $\lambda X + \mu Y$ est aussi une variable aléatoire réelle discrète.
2. Si X, Y sont deux variables aléatoires réelles, alors : XY est une variable aléatoire réelle discrète.
3. **(Minimum/maximum)** si X_1, \dots, X_n sont n variables aléatoires réelles discrètes, alors :

$$\min(X_1, \dots, X_n) \quad \text{et} \quad \max(X_1, \dots, X_n)$$

sont des variables aléatoires réelles discrètes.

4. **(Image d'une variable aléatoire discrète par une application)** Si X est discrète et $f : X(\Omega) \rightarrow \mathbf{R}$, alors : $f(X) \left| \begin{array}{l} \Omega \rightarrow \mathbf{R} \\ \omega \mapsto f(X(\omega)) \end{array} \right.$ est une variable aléatoire réelle discrète. Son univers-image est $f(X)(\Omega) = f[X(\Omega)]$ et sa loi est donnée par :

$$\forall y \in f(X(\Omega)), \quad \mathbf{P}(f(X) = y) = \sum_{\substack{x \in X(\Omega) \\ x = f(y)}} \mathbf{P}(X = y).$$

Preuve Nous avons déjà vu que toutes ces applications sont des variables aléatoires. On montre sans difficulté que les supports associés sont bien au plus dénombrables. Faisons la preuve dans le cas où $X(\Omega) = \{x_n, n \in \mathbf{N}\}$ est dénombrable.

1. 

2. 

3. Évident.


4. 

Pour la loi, écrivons que pour tout $y \in f(X(\Omega))$,

$$\begin{aligned} \{f(X) = y\} &= \{f(X) = y\} \cap \left(\bigsqcup_{x \in X(\Omega)} \{X = x\} \right) \\ &= \bigsqcup_{x \in X(\Omega)} \{f(X) = y\} \cap \{X = x\} \\ &= \bigsqcup_{x \in X(\Omega)} \{f(x) = y\} \cap \{X = x\} \\ &= \bigsqcup_{\substack{x \in X(\Omega) \\ x=f(y)}} \{X = x\}. \end{aligned}$$

Le résultat s'en suit en passant aux probabilités.

Pour les calculs de loi de $f(X)$, il faut surtout savoir les effectuer sur des exemples de fonction f , voyons cela.

Exemple 10 — On lance un dé équilibré, et soit X la variable aléatoire égale au résultat du lancer. Déterminer les lois de X , $|X|$ et X^2 . 

PROPRIÉTÉS DE LA FONCTION DE RÉPARTITION D'UNE LOI DISCRÈTE. La plupart des propriétés ci-dessous ont déjà été vues dans la [Section 3](#) du [Chapter ALEA.12](#).

Proposition ALEA.13.6 | Fonction de répartition des variables aléatoires réelles discrètes

Soit X est une variable aléatoire discrète.

1. F_X est croissante,
2. F_X est c.à.d.l.a.g. : elle est continue à droite, et possède une limite à gauche
 $F(x-) = \lim_{y \rightarrow x^-} F_X(y)$,
 (nota.) $\lim_{y < x}$
3. $\lim_{x \rightarrow -\infty} F_X(x) = 0$ et $\lim_{x \rightarrow +\infty} F_X(x) = 1$.
4. $\mathbf{P}(X > x) = 1 - F_X(x)$,
5. $\mathbf{P}(x < X \leq y) = F_X(y) - F_X(x)$,
6. $\mathbf{P}(X < x) = F_X(x-)$,
7. $\mathbf{P}(X = x) = F_X(x) - F_X(x-)$.
8. La fonction F_X est constante par morceaux.³

Exemple 11 — Calculer et tracer la fonction de répartition de X donnée par l'Exemple 8.

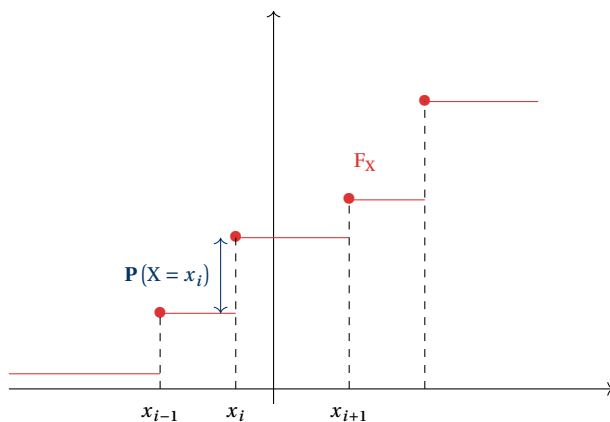


FIG. ALEA.13.1. : Allure d'une fonction de répartition dans le cas discret

Preuve


Montrons 8 comme annoncé, la seule nouvelle propriété découverte.




³Propriété caractéristique d'une loi discrète

Enfin, et c'est un exemple **très important**, on peut remarquer que beaucoup de probabilités peuvent s'exprimer en fonction des $\mathbf{P}(X = x)$ pour $x \in X(\Omega)$ et de F_X .

Exemple 12 — *Exprimer des évènements en fonction de la fonction de répartition et de la loi* Soit X une variable aléatoire discrète réelle telle que $X(\Omega) = \mathbf{N}$. Calculer les probabilités ci-dessous en fonction de $\mathbf{P}(X = n)$ pour tout $n \in X(\Omega)$ et de F_X lorsque cela vous semble possible.

1. Pour $N \in \mathbf{N}$, $\mathbf{P}(X \geq N)$, $\mathbf{P}(X < N)$. 

2. Pour $x \in \mathbf{R}$, $\mathbf{P}(X \leq x)$, $\mathbf{P}(X > x)$. 

3. $\mathbf{P}(X \text{ pair})$, $\mathbf{P}(X \text{ est le carré d'un entier positif})$. 

1.4. Indépendance

Définition ALEA.13.4 | Indépendance de variables aléatoires discrètes

Soit X_1, \dots, X_n une collection de n variables aléatoires discrètes.

► **(Indépendance)** Les variables aléatoires X_1, \dots, X_n sont dites *indépendantes* si pour tout $(x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega)$,

$$\mathbf{P}(X_1 = x_1, \dots, X_n = x_n) = \mathbf{P}(X_1 = x_1) \times \dots \times \mathbf{P}(X_n = x_n).$$

► **(Indépendance deux à deux)** Les variables aléatoires X_1, \dots, X_n sont dites *indépendantes deux à deux* si pour tous $i \neq j$, X_i et X_j sont indépendantes. Plus généralement, si (X_n) est une famille quelconque de variables aléatoires, elles sont dites *indépendantes* (*resp. indépendantes deux à deux*) si toute sous-famille finie est indépendante (*resp. deux à deux indépendantes*).

Σ Notation

On notera $X_1 \perp\!\!\!\perp X_2$ pour signifier que les variables aléatoires X_1 et X_2 sont indépendantes.

Définition ALEA.13.5 | Suite i.i.d.

On dit qu'une suite (X_n) de variables aléatoires discrètes est i.i.d. (on dit *indépendantes et identiquement distribuées*) si elles sont indépendantes et de même loi.

Proposition ALEA.13.7 | Indépendance version fonction de répartition

Soit X_1, \dots, X_n une collection n variables aléatoires discrètes. Alors X_1, \dots, X_n sont *indépendantes* si et seulement si pour tout $(x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega)$,

$$\mathbf{P}(X_1 \leq x_1, \dots, X_n \leq x_n) = \mathbf{P}(X_1 \leq x_1) \times \dots \times \mathbf{P}(X_n \leq x_n).$$

FONCTIONS DE VARIABLES ALÉATOIRES INDÉPENDANTES. Nous admettons également les deux résultats qui suivent.

Théorème ALEA.13.2

Soit (X_1, \dots, X_n) une famille de n variables aléatoires mutuellement indépendantes. Alors pour toutes fonctions f_1, \dots, f_n où pour tout $f_i : X_i(\Omega) \rightarrow \mathbf{R}$, les variables aléatoires réelles $f_1(X_1), \dots, f_n(X_n)$ sont des variables aléatoires réelles indépendantes.

Exemple 13 — Si X_1, \dots, X_3 sont indépendantes et X_3 ne s'annule pas, alors $X_1^2, X_2^2, 1/X_3$ sont indépendantes.

Théorème ALEA.13.3 | Lemme des coalitions ou Indépendance par paquets

Soit (X_1, \dots, X_n) une famille de n variables aléatoires mutuellement indépendantes. Alors pour toutes fonctions $\varphi_1, \dots, \varphi_k$, et n_1, \dots, n_k , k des entiers tels que $\sum_{i=1}^k n_i = n$, où pour tout $\varphi_i : \mathbf{R}^{n_i} \rightarrow \mathbf{R}$, les variables aléatoires réelles

$$\varphi_1(X_1, \dots, X_{n_1}), \dots, \varphi_k(X_{n_1+\dots+n_{k-1}}, \dots, X_n)$$

sont des variables aléatoires réelles mutuellement indépendantes.

Exemple 14 — Si X_1, X_2, X_3, X_4, X_5 sont indépendantes alors $X_1^2 + X_2^2, X_3X_5, X_4$ sont indépendantes.

La proposition suivante est indiquée comme hors-programme dans la mesure où la

formule obtenue n'est pas à connaître par coeur. En revanche, comme dans le cas discret, la méthode mise en jeu dans la démonstration doit être maîtrisée.

Proposition ALEA.13.8 | Minimum & Maximum de variables aléatoires discrètes

[H.P]

Soient X_1, \dots, X_n discrètes et mutuellement indépendantes. Alors pour tout $x \in \mathbf{R}$,

$$\mathbf{P}(\min\{X_1, \dots, X_n\} \geq x) = \prod_{i=1}^n \mathbf{P}(X_i \geq x),$$

$$\mathbf{P}(\max\{X_1, \dots, X_n\} \leq x) = \prod_{i=1}^n \mathbf{P}(X_i \leq x).$$

Si de plus les variables aléatoires sont i.i.d. de même loi qu'une variable aléatoire X de fonction de répartition F_X , alors :

$$\mathbf{P}(\min\{X_1, \dots, X_n\} \geq x) = \mathbf{P}(X \geq x)^n,$$

$$\mathbf{P}(\max\{X_1, \dots, X_n\} \leq x) = F_X(x)^n.$$

Notez que lorsque $n = 2$, on a

$$\max(X_1, X_2) + \min(X_1, X_2) = X_1 + X_2, \quad \max(X_1, X_2) \cdot \min(X_1, X_2) = X_1 \cdot X_2.$$

La première relation permet par exemple de calculer l'espérance de l'un connaissant l'espérance de l'autre (voir plus bas pour la définition de l'espérance).

Preuve




Attention

Les égalités

$$\mathbf{P}(\min \{X_1, \dots, X_n\} \leq x) = \mathbf{P}\left(\bigcup_{i=1}^n \{X_i \leq x\}\right),$$

$$\mathbf{P}(\max \{X_1, \dots, X_n\} \geq x) = \mathbf{P}\left(\bigcup_{i=1}^n \{X_i \geq x\}\right)$$

sont vraies. Mais en revanche, les réunions précédentes ne sont **pas disjointes**.**Méthode** Trouver la loi d'un min ou max de variables aléatoires discrètes indépendantesPour le max $X = \max(X_1, \dots, X_n)$, si X_1, \dots, X_n sont par exemple à valeurs dans \mathbf{N} .1. On calcule la fonction de répartition : $\mathbf{P}(X \leq k) = \mathbf{P}(X_1 \leq k) \dots \mathbf{P}(X_n \leq k)$ pour tout $k \in \mathbf{N}$. On invoque l'indépendance au moment adéquat.2. On calcule ensuite $\mathbf{P}(X = k)$ en fonction de $\mathbf{P}(X \leq k)$ et $\mathbf{P}(X \leq k - 1)$, i.e. $\mathbf{P}(X = k) = \mathbf{P}(X \leq k) - \mathbf{P}(X \leq k - 1)$ pour tout entier k .Pour $X = \min(X_1, \dots, X_n)$, remplacer dans 1) \leq par \geq , puis en déduire la fonction de répartition. Étape 2) inchangée, mais utiliser la relation $\mathbf{P}(X = k) = \mathbf{P}(X \geq k) - \mathbf{P}(X \geq k + 1)$.

Exemple 15 — Soient U_1, U_2 les résultats de 2 lancers de dés à 6 faces et non pipés, supposés indépendants. Déterminer la loi de U défini comme le plus grand des lancers.  Il est évident que $U(\Omega) = \llbracket 1, 6 \rrbracket$. Soit donc $k \in U(\Omega)$, calculons $\mathbf{P}(U = k)$. Pour étudier un max, on utilise la fonction de répartition.

$$\mathbf{P}(U \leq k) = \mathbf{P}(U_1 \leq k, U_2 \leq k) = \mathbf{P}(U_1 \leq k) \cdot \mathbf{P}(U_2 \leq k),$$

par indépendance. Or,

$$\begin{aligned} \mathbf{P}(U_1 \leq k) &= \sum_{\ell=1}^k \mathbf{P}(U_1 = \ell) = \frac{k-1+1}{6} \\ &= \frac{k}{6}. \end{aligned}$$

Donc finalement :

$$\forall k \in U(\Omega), \quad \mathbf{P}(U \leq k) = \frac{k^2}{36}.$$

Puis on récupère la loi, en écrivant que

$$\mathbf{P}(U = k) = \mathbf{P}(U \leq k) - \mathbf{P}(U \leq k - 1)$$

$$= \begin{cases} \frac{k^2}{36} - \frac{(k-1)^2}{36} & \text{si } k \in \llbracket 2, 6 \rrbracket, \\ \frac{1}{36} - 0 & \text{si } k = 1, \end{cases}$$

$$= \frac{k^2}{36} - \frac{(k-1)^2}{36}.$$

les formules se réunissent en une seule

Donc la loi de U est la suivante, en utilisant une identité remarquable :

$$U(\Omega) = \llbracket 1, 6 \rrbracket, \quad \mathbf{P}(U = k) = \frac{2k-1}{36}.$$

2.

ESPÉRANCE, VARIANCE, MOMENTS

2.1.

Espérance

On en vient au coeur du chapitre, la définition de l'espérance pour une variable aléatoire réelle discrète à support non forcément fini mais, plus généralement, au plus dénombrable. Reprenons le cours de première année : si $X(\Omega) = \{x_1, \dots, x_n\}$ est fini, vous aviez appelé *espérance de X* la quantité suivante :

$$\mathbf{E}(X) = \sum_{x \in X(\Omega)} x \mathbf{P}(X = x) = \sum_{i=1}^n x_i \mathbf{P}(X = x_i),$$

formule qui était directement inspirée du cas de la loi uniforme :

$$\mathbf{E}(X) = \frac{1}{n} \sum_{i=1}^n x_i \quad (\text{moyenne d'une série statistique})$$

Dans ce cadre, l'espérance était une somme finie donc existait toujours. À présent, étant donné que le support n'est plus fini, nous allons nous intéresser plutôt à la convergence d'une série, *i.e.* à une moyenne sur un « nombre infini de termes ».

Définition ALEA.13.6 | Espérance d'une variable aléatoire réelle discrète

Soit X une variable aléatoire réelle discrète.

1. (Admettre une espérance) On dit que X admet une espérance si :

$$\left(\sum x \mathbf{P}(X = x)\right)_{x \in X(\Omega)} \text{ converge } \underline{\text{absolument}}.$$

2. (Valeur de l'espérance) Si X admet une espérance, alors on appelle *espérance de X* la quantité

$$\mathbf{E}(X) = \sum_{x \in X(\Omega)} x \mathbf{P}(X = x).^4$$

Une variable aléatoire d'espérance nulle est dite *centrée*.

Remarque 2.1 — Pourquoi supposer une convergence absolue? Nous avons vu dans le [Chapter ANA.10](#) qu'en cas de convergence absolue, l'ordre de sommation des termes d'une série n'a aucune importance. Ainsi l'espérance définie précédemment ne dépend pas de la numérotation choisie pour les éléments de $X(\Omega) = \{x_n, n \in \mathbf{N}\}$ (ou $\{x_n, n \in \llbracket 0, N \rrbracket\}$).

On vérifie immédiatement que toute variable aléatoire à support fini possède une espérance.

Proposition ALEA.13.9 | Cas où $X(\Omega)$ est fini

Si $X(\Omega) = \{x_1, \dots, x_N\}$ avec $x_1, \dots, x_N \in \mathbf{R}$ et $N \in \mathbf{N}$, alors X admet toujours une


espérance définie par :


$$\mathbf{E}(X) = \sum_{n=1}^N x_n \mathbf{P}(X = x_n).$$

Proposition ALEA.13.10 | Espérance d'une indicatrice

Soit $A \in \mathcal{F}$, alors : $\mathbf{E}(\mathbb{1}_A) = \mathbf{P}(A)$.

La proposition précédente paraît anecdotique, mais elle serait d'importance capitale dans de futurs chapitres, afin notamment d'interpréter les probabilités comme des moyennes et pouvoir les approcher par simulation.


Preuve Nous avons déjà établi que $\mathbb{1}_A$ est une variable aléatoire discrète. Calculons son espérance. 

Exemple 16 — Soit X une variable aléatoire décrivant le lancer d'un dé. Calculer l'espérance de X^2 , en utilisant la définition de l'espérance. 

⁴On rappelle encore une fois la signification du symbole somme dans cette formule : si par exemple

$$X(\Omega) = \{x_i, i \in \mathbf{N}\} \text{ est dénombrable, on note } \sum_{x \in X(\Omega)} x \mathbf{P}(X = x) = \sum_{i=0}^{\infty} x_i \mathbf{P}(X = x_i)$$

Exemple 17 — Soit X une variable aléatoire réelle discrète telle que $X(\Omega) = \mathbf{N}^*$ et pour tout $n \geq 1$: $\mathbf{P}(X = n) = \frac{1}{2^n}$.

La variable X admet une espérance, calculons-la. 

Exemple 18 — Soit X une variable aléatoire réelle discrète telle que $X(\Omega) = \mathbf{N}^*$ et pour tout $n \geq 1$: $\mathbf{P}(X = n) = \frac{1}{n(n+1)}$. La variable X n'admet pas d'espérance, pourquoi?



PROPRIÉTÉS DE L'ESPÉRANCE.

Théorème ALEA.13.4

Soient X et Y deux variables aléatoires discrètes admettant une espérance. Alors :

1. (Linéarité de l'espérance) Soient $\lambda, \mu \in \mathbf{R}$, $\lambda X + \mu Y$ admet une espérance et :

$$\mathbf{E}(\lambda X + \mu Y) = \lambda \mathbf{E}(X) + \mu \mathbf{E}(Y).$$

2. (Positivité de l'espérance) $X \geq 0 \implies \mathbf{E}(X) \geq 0$, et :

$$X \geq 0, \quad \mathbf{E}(X) = 0 \iff X = 0 \text{ p.s. } \left(i.e. \mathbf{P}(X = 0) = 1 \right).$$

Le résultat subsiste si on a seulement $\mathbf{P}(X \geq 0) = 1$ en hypothèse, *i.e.*

$$\mathbf{P}(X \geq 0) = 1 \implies \mathbf{E}(X) \geq 0,$$

$$\mathbf{P}(X \geq 0) = 1, \quad \mathbf{E}(X) = 0 \iff X = 0 \text{ p.s. } \left(i.e. \mathbf{P}(X = 0) = 1 \right).$$

3. (Croissance de l'espérance)

$$X \leq Y \implies \mathbf{E}(X) \leq \mathbf{E}(Y).$$

Le résultat subsiste si on a seulement $\mathbf{P}(X \leq Y) = 1$ en hypothèse, *i.e.*

$$\mathbf{P}(X \leq Y) = 1 \implies \mathbf{E}(X) \leq \mathbf{E}(Y).$$

Preuve

1. Nous admettons provisoirement la linéarité de l'espérance : la preuve nécessitant le théorème de transfert pour les couples aléatoires, que nous verrons plus tard.
2. Nous admettons provisoirement ce fait, qui nécessite l'inégalité de BIENAYMÉ-

TCHEBYCHEV, et qui sera vue dans un prochaine chapitre.

3. 

FORMULE DE TRANSFERT POUR UNE VARIABLE ALÉATOIRE DISCRÈTE. L'objectif est ici d'obtenir une formule pour calculer des espérances de fonctions de variables aléatoires $f(X)$ sans avoir à trouver la loi de $f(X)$ (ce qui peut se révéler compliqué). Le théorème de transfert répond à ce problème.

Remarque 2.2 — On peut donc, avec ce théorème, calculer l'espérance de $f(X)$ en connaissant seulement la loi de X

Théorème ALEA.13.5 | Transfert pour les variables aléatoires discrètes

Soit X une variable aléatoire discrète et $f : X(\Omega) \rightarrow \mathbf{R}$. Alors :

$f(X)$ possède une espérance $\iff (\sum_{x \in X(\Omega)} f(x) \mathbf{P}(X = x))$ converge **absolument**.

Dans ce cas, nous avons :

$$\mathbf{E}(f(X)) = \sum_{x \in X(\Omega)} f(x) \mathbf{P}(X = x).$$

Preuve C'est un calcul direct, en utilisant la loi de $f(X)$ calculée dans la **Proposition ALEA.15.8** :

$$\begin{aligned} \mathbf{E}(f(X)) &= \sum_{x \in f(X)(\Omega)} x \mathbf{P}(f(X) = x) \\ &= \sum_{x \in f(X)(\Omega)} x \sum_{\substack{y \in X(\Omega) \\ x=f(y)}} \mathbf{P}(X = y) \\ &= \sum_{x \in f(X)(\Omega)} \sum_{\substack{y \in X(\Omega) \\ x=f(y)}} x \mathbf{P}(X = y) = \sum_{x \in f(X)(\Omega)} \sum_{\substack{y \in X(\Omega) \\ x=f(y)}} y \mathbf{P}(X = y) \\ &= \sum_{y \in Y(\Omega)} y \mathbf{P}(Y = y). \end{aligned}$$

Nous admettons la dernière égalité.

Remarque 2.3 — On note donc que, dans la formule de transfert, nous avons $\mathbf{P}(X = x)$ avec $x \in X(\Omega)$ qui est largement préférable à $\mathbf{P}(f(X) = y)$ pour tout $y \in f(X)(\Omega)$.

Corollaire ALEA.13.2 | Inégalité triangulaire pour l'espérance

Soit X une variable aléatoire réelle discrète. Alors :

X admet une espérance $\iff |X|$ admet une espérance.

Et dans ce cas :

$$|\mathbf{E}(X)| \leq \mathbf{E}(|X|).$$

Preuve (Point clef — *Théorème du transfert & Inégalité triangulaire pour les séries*)



Exemple 19 – Cas fini On considère la variable aléatoire donnée par le tableau suivant :

$X = k$	-3	-1	0	1	2	3
$P(X = k)$	2/10	1/10	1/10	2/10	3/10	1/10

et $Y = 3X + 2$. Calculer de deux manières son espérance. 

Corollaire ALEA.13.3 | Transfert pour les fonctions affines

Soit X une variable aléatoire discrète et $a, b \in \mathbf{R}$. Alors :

1. X possède une espérance $\implies aX + b$ possède une espérance.
2. Supposons que $a \neq 0$. Alors : $aX + b$ possède une espérance $\implies X$ possède une espérance.

De plus, si X et $aX + b$ possèdent une espérance, nous avons :

$$\mathbf{E}(aX + b) = a\mathbf{E}(X) + b.$$

Preuve




Corollaire ALEA.13.4 | Centrage

Soit X une variable aléatoire discrète possédant une espérance, alors

$X - \mathbf{E}(X)$ est une variable aléatoire centrée.

Preuve Appliquer le résultat précédent avec $a = 1$ et $b = -\mathbf{E}(X)$. Alors $X - \mathbf{E}(X)$ possède donc une espérance, et

$$\mathbf{E}(X - \mathbf{E}(X)) = \mathbf{E}(X) - \mathbf{E}(X) = 0.$$

Exemple 20 – Cas dénombrable Soit $Y = \frac{1+(-1)^X}{2}$ où X est définie par $X(\Omega) = \mathbf{N}^*$ et $\mathbf{P}(X = k) = \frac{1}{2^k}$ pour tout $k \in \mathbf{N}^*$. Calculer de deux manières son espérance. 

2.2. Moments d'ordre supérieur

À l'aide du théorème de transfert, nous pouvons donc affirmer les points suivants.

Définition/Proposition ALEA.13.2 | Variance, écart-type, moments, version discrète

- ▶ **(Moments d'ordre k)** On dit que X admet un moment d'ordre $k \in \mathbf{N}$ si l'équivalence suivante est réalisée :

$$\mathbf{E}(|X|^k) < \infty \iff \left(\sum_{x \in X(\Omega)} |x|^k \mathbf{P}(X = x) \right) \text{ converge.}$$

On appelle alors *moment d'ordre k* : $\mathbf{E}(X^k) = \sum_{x \in X(\Omega)} x^k \mathbf{P}(X = x)$.

- ▶ **(Moments d'ordre 2)** On dit que X admet un moment d'ordre deux si l'équivalence suivante est réalisée :

$$\mathbf{E}(|X|^2) < \infty \iff \left(\sum_{x \in X(\Omega)} |x|^2 \mathbf{P}(X = x) \right) \text{ converge.}$$

On appelle alors *moment d'ordre 2* : $\mathbf{E}(X^2) = \sum_{x \in X(\Omega)} x^2 \mathbf{P}(X = x)$.

- ▶ **(Variance)** Si X admet un moment d'ordre deux, alors X admet un moment d'ordre un (*i.e.* une espérance), et on appelle *variance de X* la quantité notée $\mathbf{Var}(X)$ et définie par : $\mathbf{Var}(X) = \mathbf{E}((X - \mathbf{E}(X))^2)$. La variable aléatoire

X possède une variance si et seulement si :


$$\left(\sum_{x \in X(\Omega)} (x - \mathbf{E}(X))^2 \mathbf{P}(X = x) \right) \text{ converge}$$

$$\left(\sum_{x \in X(\Omega)} x^2 \mathbf{P}(X = x) \right) \text{ converge}$$

Dans ce cas,

$$\mathbf{Var}(X) = \sum_{x \in X(\Omega)} (x - \mathbf{E}(X))^2 \mathbf{P}(X = x).$$


On appelle *écart-type* de X, la quantité notée σ_X et définie par $\sigma_X := \sqrt{\mathbf{Var}(X)}$. Une variable aléatoire de variance un est dite *réduite*.

Preuve Montrons que : si X admet un moment d'ordre deux, alors X admet un moment d'ordre un. Le reste provient de simples applications du théorème de transfert. 

Remarque 2.4 — On retiendra également que :

1. si X n'a pas d'espérance, alors elle n'a pas de variance.
2. Le moment d'ordre 1 correspond donc à l'espérance.

3. La variance d'une variable aléatoire mesure l'écart «quadratique» (au carré) moyen entre X et sa valeur moyenne $\mathbf{E}(X)$, et en plus elle réalise le minimum parmi tous les écarts au carré (voir l'exemple ci-après).

Exemple 21 — L'espérance minimise l'écart quadratique à X Soit X une variable aléatoire admettant un moment d'ordre deux. Démontrer que $a \mapsto \mathbf{E}((X - a)^2)$ est minimale pour $a = \mathbf{E}(X)$. 

 **Méthode Étudier l'existence d'une variance dans le cas discret**

Cela revient à étudier la convergence de :

$$\sum_{x \in X(\Omega)} |x|^2 \mathbf{P}(X = x).$$


Proposition ALEA.13.11 | Cas d'une variable aléatoire bornée

Soit X une variable aléatoire discrète.

- ▶ Si X est bornée, alors X admet une variance.
- ▶ Le résultat est encore vrai si X est presque-sûrement bornée : s'il existe $M \in \mathbf{R}^+$ tel que $\mathbf{P}(|X| \leq M) = 1$ alors X admet une variance.

Preuve Montrons que

$$\sum_{x \in X(\Omega)} x^2 \mathbf{P}(X = x) \quad \text{converge (et donc absolument aussi car positive).}$$

Faisons le cas général : on suppose qu'il existe $M \in \mathbf{R}^+$ tel que $\mathbf{P}(|X| \leq M) = 1$. 

OPÉRATION DE CENTRAGE/RÉDUCTION. La proposition ci-dessous paraît anecdotique mais elle sera d'un intérêt majeur plus tard dans l'année.

Définition/Proposition ALEA.13.3

Soit X une variable aléatoire discrète ayant une variance, alors

$$X^* \stackrel{\text{(déf.)}}{=} \frac{X - \mathbf{E}(X)}{\sigma_X}$$

est une variable aléatoire réelle centrée réduite. On l'appelle *la centrée/réduite de* X .

Proposition ALEA.13.12 | Propriétés de la variance/covariance

Soient X, Y admettent une variance, et $\lambda, \mu \in \mathbf{R}$.

1. **(Variance nulle)** $\mathbf{Var}(X) = 0 \iff X = \mathbf{E}(X)$ p.s.
2. **(Variance d'une expression affine)** $\mathbf{Var}(\lambda X + \mu) = \lambda^2 \mathbf{Var}(X)$.
3. **(Formule de KÖNIG-HUYGENS)** $\mathbf{Var}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2$.

Preuve



Preuve



Exemple 22 — Soit $X \hookrightarrow \mathcal{B}\left(n, \frac{1}{2}\right)$. Dans quel ensemble fini X^* prend-elle ses valeurs?

Calculer explicitement X^* dans ce cas. 

CAS DE VARIABLES ALÉATOIRES INDÉPENDANTES. On présente sans démonstration le résultat suivant, qui sera démontré et étudié dans le [Chapter ALEA.16](#) sur les couples aléatoires discrets.

Proposition ALEA.13.13

Si X_1, \dots, X_n sont n variables aléatoires discrètes mutuellement indépendantes. Alors :

1. (**Espérance d'un produit**) si les X_i admettent une espérance,

$$\mathbf{E}(X_1 \dots X_n) = \mathbf{E}(X_1) \dots \mathbf{E}(X_n).$$

2. (**Variance d'une somme**) si les X_i admettent une variance, alors

$$\mathbf{Var}(X_1 + \dots + X_n) = \mathbf{Var}(X_1) + \dots + \mathbf{Var}(X_n).$$

Preuve Provisoirement admis.

3. LOIS DISCRÈTES USUELLES

Pour chacune des lois ci-dessous, il est important de connaître :

- ▶ sa loi — ce qui inclut $X(\Omega)$! — et une idée de son histogramme,
- ▶ son espérance/variance,
- ▶ le type d'expérience dans lequel elle intervient,
- ▶ \gg et comment la simuler à l'aide de Python.

Dans la pratique, on essaiera autant que possible de traduire l'énoncé avec des lois usuelles. Nous supposons dans la suite avoir effectué les importations suivantes.

Pour l'aspect informatique, nous aurons besoin des importations suivantes, que l'on suppose donc réalisées dans toute la suite.

```
import random as rd # pour les simulations
import numpy as np # pour les fonctions classiques et/ou la
    - simulation
import matplotlib.pyplot as plt # pour les représentations
    - graphiques
```

GÉNÉRALITÉS À PROPOS DE LA SIMULATION DE VARIABLES ALÉATOIRES RÉELLES. Commençons par définir ce que l'on appelle *simulation de variable aléatoire* en Mathématiques.

Définition ALEA.13.7 | Simulation

Soit X une variable aléatoire. Alors on appelle *simulation de la variable aléatoire* X toute procédure permettant de renvoyer un $X(\omega)$ pour $\omega \in \Omega$, de sorte que si l'on effectue n simulations $X(\omega_1), \dots, X(\omega_n)$ avec $n \in \mathbf{N}$ selon cette même procédure, on ait :

$$\forall x \in \mathbf{R}, \quad \frac{\sum_{i=1}^n \mathbb{1}_{\{X(\omega_i) \leq x\}}}{n} \xrightarrow{n \rightarrow \infty} \mathbf{P}(X \leq x).$$

Autrement dit, l'histogramme associé à la suite de simulations est proche du véritable histogramme de la loi, ce qui est un comportement naturel attendu.

Remarque 3.1 — Loi uniforme continue On dira qu'une variable aléatoire U à valeurs dans \mathbf{R} suit une *loi uniforme sur* $[0, 1]$ si pour tout intervalle I ,


$$\mathbf{P}(U \in I) = \text{Long}(I \cap [0, 1]).$$

Autrement dit la probabilité d'être dans un certain intervalle est la longueur dudit intervalle. Nous étudierons plus en détail cette variable aléatoire non discrète plus tard dans l'année ([Chapter ALEA.14](#)).

Proposition ALEA.13.14

Soit X une variable aléatoire réelle telle que F_X soit bijective. Alors :

$$F_X^{-1}(U) \sim X, \quad \text{où } U \hookrightarrow \mathcal{U}([0, 1]).$$

Preuve Les variables aléatoires $F_X^{-1}(U)$ et X ont même loi si et seulement si elles ont même fonction de répartition. 

Remarque 3.2 — Et si F_X n'est pas bijective? En fait, même si F_X n'est pas bijective on peut construire une fonction G_X telle que $G_X(U)$ ait pour loi X .⁵

Tous les constats précédents se résument en l'idée suivante :

 **Résumé**
la simulation d'une loi quelconque se ramène à la simulation d'un réel aléatoire dans $[0, 1[$.

C'est ce que nous allons observer dans tous les scripts de simulation qui suivent. Même pour l'un des plus simples, celui de la loi de BERNOULLI, on utilisera une simulation de $\mathcal{U}[0, 1]$. Nous verrons en TP avec quelle méthode Python simule ce réel dans $[0, 1[$.⁶

⁵En l'occurrence, la fonction G_X ci-après convient : $\forall u \in \mathbf{R}, G_X(u) = \inf\{t \in \mathbf{R}, F_X(t) \geq u\}$. On l'appelle l'*inverse généralisé* de F_X . On peut démontrer que $G_X \circ F_X(u) = u$ pour tout $u \in \mathbf{R}$, et cela permet de justifier que $F_X^{-1}(U)$ est une variable aléatoire de loi celle de X .

⁶Il n'existe pas de «vrai» aléatoire en informatique, Python fournit en fait les valeurs successives de certaines suites récurrentes bien choisies

Simulation d'un réel entre 0 et 1

```
>>> import random as rd
>>> rd.random()
0.3125923664052699
```



Le module `random` sait aussi simuler beaucoup de lois usuelles, mais vous devez aussi savoir la simuler «à la main». Dans la suite nous donnerons systématiquement les deux.

COMMENT TRACER UNE FONCTION DE RÉPARTITION EN PYTHON? Pour tracer une fonction de répartition d'une loi discrète, on utilise le fait déjà établi suivant : c'est une fonction constante par morceaux, et chaque saut est aux éléments du support de X , l'amplitude d'un saut valant $\mathbf{P}(X = k)$ si $k \in X(\Omega)$. On en déduit alors la fonction générale suivante.

```
def trace_fdr(Support, Loi):
    """
    trace la fonction de répartition de loi loi donné dans la liste
    ← Loi, et de support support
    """
    Fdr = [0 for _ in range(len(Support))]
    for k in range(0, len(Support)):
        Fdr[k] = sum([Loi[i] for i in range(0, k+1)])
    plt.step(Support, Fdr)
```

3.1. Loi uniforme discrète


Preuve

**Définition/Proposition ALEA.13.4 | Loi uniforme sur un sous-ensemble fini de \mathbf{Z}**

Soit E un sous-ensemble **fini** de \mathbf{Z} . On dit qu'une variable aléatoire suit une *loi uniforme sur E* (on note $X \hookrightarrow \mathcal{U}(E)$), si :

$$X(\Omega) = E, \quad \forall k \in E, \quad \mathbf{P}(X = k) = \frac{1}{\#E}.$$

Remarque 3.3 — Modélisation. Toute expérience aléatoire dont les issues sont des nombres entiers en nombre fini, apparaissant de manière équiprobable.

Preuve Il existe un espace probabilisé et une variable aléatoire vérifiant ces conditions. En effet, 

Exemple 23 — Cas d'un intervalle d'entiers consécutifs. En particulier, si $E = \llbracket a, b \rrbracket$ avec $(a, b) \in \mathbf{Z}^2$ deux entiers tels que $a < b$, $X \hookrightarrow \mathcal{U}(\llbracket a, b \rrbracket)$, si :

$$\forall k \in \llbracket a, b \rrbracket, \quad \mathbf{P}(X = k) = \frac{1}{b - a + 1}.$$

La quantité $b - a + 1$ est simplement $\# \llbracket a, b \rrbracket$.


Proposition ALEA.13.15 | Espérance, variance

Si $X \hookrightarrow \mathcal{U}(\llbracket a, b \rrbracket)$ avec $(a, b) \in \mathbf{Z}^2$, alors X possède une espérance et une variance, et :

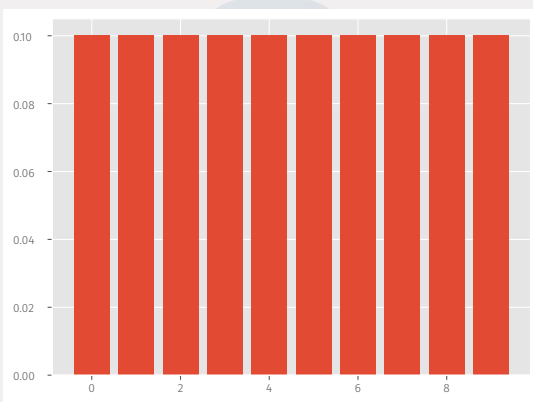
$$\mathbf{E}(X) = \frac{a + b}{2}, \quad \mathbf{Var}(X) = \frac{(b - a) \cdot (b - a + 1)}{12}.$$

HISTOGRAMME, FONCTION DE RÉPARTITION.**Histogramme de la loi uniforme**

```
a = 0
b = 9
Support = range(a, b+1)
Loi = [0 for _ in range(len(Support))]
```

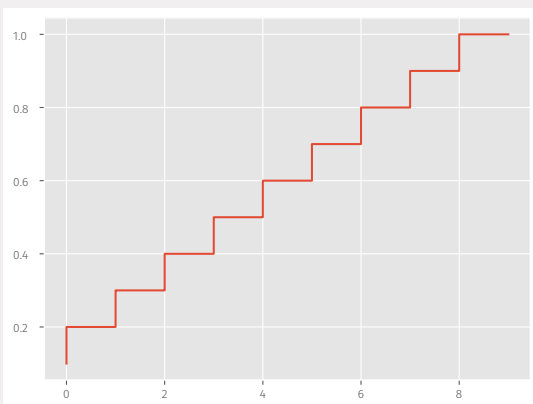


```
for k in range(b-a+1):
    Loi[k] = 1/(b-a+1)
plt.bar(Support, Loi)
```



Fonction de répartition de la loi uniforme

```
trace_fdr(Support, Loi)
```



Résumé


`plt.bar` permet de tracer des diagrammes en bâtons et `plt.step` trace des fonctions en reliant les points **par morceaux**

SIMULATION. On va pouvoir se ramener à une loi uniforme sur $[0, 1]$ comme nous l'avions démontré de manière générale au début de cette sous-section.

Proposition ALEA.13.16 | Simulation d'une loi uniforme discrète sur $[[a, b]]$

Soient a, b deux entiers distincts et $U \hookrightarrow \mathcal{U}([0, 1])$. Alors :

$$a + \lfloor U(b - a + 1) \rfloor \hookrightarrow \mathcal{U}([a, b]).$$

Preuve La variable aléatoire est à valeurs entières supérieures ou égales à a et inférieures à b . Soit donc $k \geq a$, calculons $\mathbf{P}(a + \lfloor U(b - a + 1) \rfloor = k)$. 

On en déduit alors le script suivant de simulation.

Simulation de la loi uniforme sur $[[a, b]]$

```
def uniforme(a, b):
    return a + int(rd.random()*(b-a+1))
```

Ou bien on utilise la fonction existante du module `random`.

```
rd.randint(a, b)
```

3.2. Loi de BERNOULLI & binomiale

Définition/Proposition ALEA.13.5 | Loi de BERNOULLI/Rademacher de paramètre p

Soit $p \in [0, 1]$. On dit qu'une variable aléatoire suit une :

- ▶ **(loi de BERNOULLI de paramètre p)** (on note $X \hookrightarrow \mathcal{B}(p)$), si :


$$X(\Omega) = \{0, 1\}, \quad \mathbf{P}(X = 1) = p, \quad \mathbf{P}(X = 0) = 1 - p.$$

- ▶ **(loi de RADEMACHER de paramètre p)** (on note $X \hookrightarrow \mathcal{R}(p)$), si :

$$X(\Omega) = \{-1, 1\}, \quad \mathbf{P}(X = 1) = p, \quad \mathbf{P}(X = -1) = 1 - p.$$

Remarque 3.4 — Modélisation. Toute expérience aléatoire dont les issues sont au nombre de deux, dont l'une apparaît avec probabilité p .

Remarque 3.5 — Si $p = 1$ (cf. $p = 0$) alors $\mathbf{P}(X = 1) = 1$ (cf. $= 0$) et $\mathbf{P}(X = 0) = 0$ (cf. $= 1$) donc X est constante égale à 1 (cf. 0) presque-sûrement.

Preuve Il existe un espace probabilisé et une variable aléatoire vérifiant ces conditions pour l'une ou l'autre des lois. En effet, 

Proposition ALEA.13.17 | Exemples typiques, Obtenir le paramètre

- ▶ **(Indicatrice)** Soit $A \in \mathcal{F}$, alors $\mathbb{1}_A \hookrightarrow \mathcal{B}(\mathbf{P}(A))$.
- ▶ **(Valeurs 0, 1)** Si X est une variable aléatoire telle que $X(\Omega) = \{0, 1\}$ alors

$$X \hookrightarrow \mathcal{B}(\mathbf{P}(X = 1)) = \mathcal{B}(\mathbf{E}(X)).$$

En particulier, si X suit une certaine loi de BERNOULLI, son paramètre est donné par $\mathbf{P}(X = 1) = \mathbf{E}(X)$.⁷

Cette propriété nous informe donc que toute expérience aléatoire à deux issues peut être modélisée par une variable aléatoire de BERNOULLI, après numérotation des deux issues par 0, 1.

Preuve



Proposition ALEA.13.18 | Lien entre BERNOULLI et RADEMACHER

Soit $p \in [0, 1]$, alors :

$$X \hookrightarrow \mathcal{B}(p) \iff 2X - 1 \hookrightarrow \mathcal{R}(p).$$

Preuve




⁷Plus tard dans l'année, nous aurons des outils pour estimer l'espérance ou une probabilité, et donc pour estimer un paramètre de BERNOULLI *a priori* inconnu.

Définition/Proposition ALEA.13.6 | Loi binomiale de paramètres p et n

Soient $p \in [0, 1]$ et $n \in \mathbf{N}$. On dit qu'une variable aléatoire suit une *loi binomiale de paramètres p et n* (on note $X \hookrightarrow \mathcal{B}(n, p)$), si :


$$X(\Omega) = \llbracket 0, n \rrbracket, \quad \forall k \in \llbracket 0, n \rrbracket, \quad \mathbf{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}.$$

Remarque 3.6 — Modélisation. Toute épreuve constituée de n épreuves aléatoires dont les résultats sont **indépendants**, chacune ayant deux issues appelées succès (de probabilité p) et échec (de probabilité $1 - p$). La variable aléatoire X est le nombre total de succès dans ces n épreuves. Un raisonnement de dénombrement conduit à la formule mentionnée *supra*.

Preuve Il existe un espace probabilisé et une variable aléatoire vérifiant ces conditions. En effet, 

Exemple 24 — Stabilité par retournement Justifions les deux faits ci-après, tout à fait conformes à l'intuition. Soient $p \in [0, 1]$ et $n \geq 1$.

▶ Si $X \hookrightarrow \mathcal{B}(p)$, alors $1 - X \hookrightarrow \mathcal{B}(1 - p)$. 

▶ Si $X \hookrightarrow \mathcal{B}(n, p)$, alors $n - X \hookrightarrow \mathcal{B}(n, 1 - p)$. 

La propriété qui suit nous servira notamment pour interpréter la loi binomiale en terme de variables aléatoires suivant une loi de BERNOULLI et donc de pouvoir la si-

muler. Nous n'avons pas encore les outils pour la démontrer, elle est donc pour le moment admise.

Proposition ALEA.13.19 | Somme de binomiales indépendantes

Soient $p \in [0, 1]$, $X_1, \dots, X_n, n \geq 1$ des variables aléatoires réelles **indépendantes** telles que pour tout $i \in \llbracket 1, n \rrbracket$, $X_i \hookrightarrow \mathcal{B}(p)$. Alors :

$$X_1 + \dots + X_n \hookrightarrow \mathcal{B}(n, p).$$

Preuve Provisoirement admise.

Proposition ALEA.13.20 | Espérance, variance

▶ Si $X \hookrightarrow \mathcal{B}(p)$ avec $p \in [0, 1]$, alors

$$\mathbf{E}(X) = p, \quad \mathbf{Var}(X) = p(1 - p).$$

▶ Si $X \hookrightarrow \mathcal{B}(n, p)$ avec $p \in [0, 1]$ et $n \in \mathbf{N}$, alors X possède une espérance et une variance, et :

$$\mathbf{E}(X) = np, \quad \mathbf{Var}(X) = np(1 - p).$$




Preuve (formules de l'espérance et la variance)

▶ 

▶ 

Exemple 25 — Marche aléatoire simple sur Z Soient (R_n) une suite de variables aléatoires indépendantes de même loi $\mathcal{R}(p)$ avec $p \in [0, 1]$. On pose :

$$S_n = R_1 + \dots + R_n.$$

1. Pour tout $i \in \mathbf{N}$, on note $X_i = \frac{R_i+1}{2}$. Quelle est la loi de X_i ? 
2. Dédurre la loi de $\frac{S_n+n}{2}$ pour tout $n \in \mathbf{N}$, puis l'espérance et la variance de S_n . 
3. Un(e) élève de BCPST titube en rentrant d'un bar boulevard des Pyrénées, situé à une abscisse 0. On suppose que cet élève avance de 1 avec probabilité p , recule de 1 avec probabilité $1-p$. En moyenne, a-t-il une chance d'atteindre son appartement situé à l'abscisse $x = 10$? Si oui, préciser environ au bout de combien de temps. 

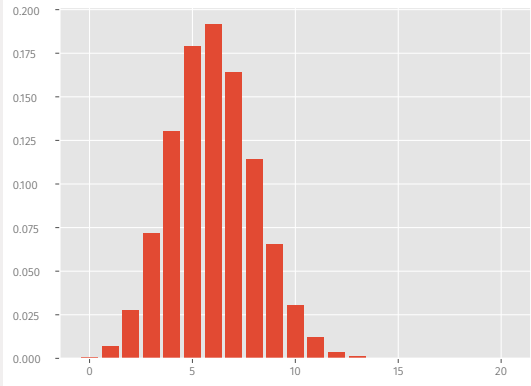
HISTOGRAMME, FONCTION DE RÉPARTITION. Même principe que pour la loi uniforme, on reprend les scripts précédents en modifiant le support et la loi. Notez que, étant donné le choix de paramètres, le support de la loi se concentre dans la partie gauche du graphique (l'espérance vaut $\frac{20}{3}$).

Histogramme de la loi binomiale

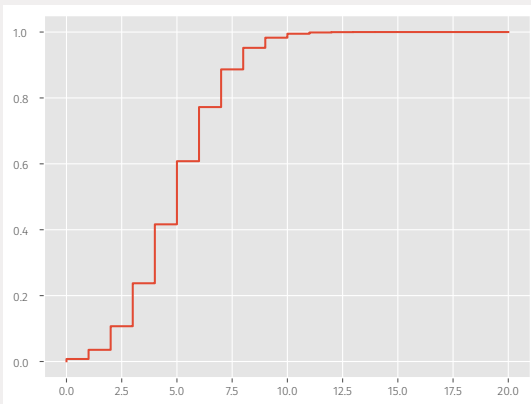
```
import scipy.special

def binom(n, k):
    """
    (n : int, k : int) -> renvoie le coefficient binomial k parmi
    n
    """
    if k > n:
        return 0
    else:
        return scipy.special.binom(n, k) #commande toute faite
        ← pour le coefficient binomial

n = 20
p = 0.3
Support = range(0, n+1)
Loi = [0 for _ in range(len(Support))]
for k in range(len(Support)):
    Loi[k] = binom(n, k)*(p**k)*((1-p)**(n-k))
plt.bar(Support, Loi)
```


Fonction de répartition de la loi $\mathcal{B}(20, 0.3)$

trace_fdr(Support, Loi)



SIMULATION. Le point de départ pour ces deux lois est la simulation d'une BERNOULLI, qui se fait en regardant dans quelle portion de l'intervalle $[0, 1[$ ($[0, p]$ ou $[p, 1 - p]$) se trouve `rd.random()`. Ensuite pour déduire la binomiale un certain nombre de résultats, d'où l'importance de bien connaître l'interprétation de ces lois en terme

d'expérience aléatoire. Nous ferons de-même pour la loi géométrique plus tard.

Simulation de la loi de BERNOULLI et de la binomiale

```
import random as rd
def bernoulli(p):
    """
    simule une bernoulli
    """
    if rd.random() < p:
        return 1
    else:
        return 0

def binomiale(n,p):
    """
    simule une binomiale
    """
    S = 0
    for i in range(n):
        if rd.random() < p:
            S += 1
    return S
```

Le module `random` ne sait pas simuler directement les lois de BERNOULLI et binomiale. On fait donc plutôt appel pour cela à la sous bibliothèque `random` de `numpy`.

```
np.random.binomial(n,p)
np.random.binomial(n,p,nb_simu) # Si l'on souhaite un tableau
↳ numpy de simulations
```

3.3. Loi Hypergéométrique

Définition/Proposition ALEA.13.7 | Loi hypergéométrique de paramètres p et

$n, N \in \mathbf{N}$

Soient $p \in [0, 1]$ et $n, N \in \mathbf{N}$ tels que : $1 \leq n \leq N$, $Np \in \mathbf{N}$. On pose $q = 1 - p$. On dit qu'une variable aléatoire suit une *loi hypergéométrique de paramètres p et n, N* (on note $X \hookrightarrow \mathcal{H}(N, n, p)$), si :

$$X(\Omega) = \llbracket 0, n \rrbracket, \quad \forall k \in \llbracket 0, n \rrbracket, \quad \mathbf{P}(X = k) = \frac{\binom{Np}{k} \binom{Nq}{n-k}}{\binom{N}{n}}.$$

Remarque 3.7 — Modélisation. Décrit une série de tirages sans remise (ou simultanés) de n éléments dans un ensemble contenant Np éléments de type 1 et Nq éléments de type 2. La variable aléatoire X représente le nombre d'éléments de type 1 obtenus dans ce paquet de n éléments.

Expliquons cette remarque à l'aide d'un raisonnement de dénombrement. 

Preuve Il existe un espace probabilisé et une variable aléatoire vérifiant ces conditions. C'est une conséquence, notamment, de la formule de VANDERMONDE. Nous admettons le reste.

Remarque 3.8 — Lorsque N devient grand, nous montrerons dans le **Chapitre ALEA.16** que cette loi est proche d'une $\mathcal{B}(n, p)$. Et c'est bien logique si l'on analyse l'expression aléatoire type qu'elle décrit. En effet, si le nombre de boules dans l'urne est très grand, la non-remise devient « équivalente » à la remise.

Remarque 3.9 — À propos du support. On peut même construire l'espace probabilisé sous-jacent de sorte que

$$X(\Omega) = \llbracket \max(0, n - Nq), \min(n, Np) \rrbracket.$$

Pour constater cela, analyser la nullité des coefficients binomiaux en fonction des paramètres. Pour simplifier, nous conserverons le support de la définition précédente.

Proposition ALEA.13.21 | Espérance, variance

Si $X \hookrightarrow \mathcal{H}(N, n, p)$ avec $p \in [0, 1]$ et $n, N \in \mathbf{N}$, alors X possède une espérance et une variance, et :

$$\mathbf{E}(X) = np, \quad \mathbf{Var}(X) = npq \frac{N-n}{N-1}.$$

Preuve Numérotons les éléments de type 1 entre 1 et Np .

On peut écrire X de la manière suivante : $X = \sum_{k=1}^{Np} \mathbb{1}_{E_k}$ où E_k est l'évènement « nous avons tiré l'élément k dans notre paquet de n éléments ». Alors pour tout en-

tier $k \in \llbracket 1, Np \rrbracket$,

$$\mathbf{P}(E_k) = 1 - \mathbf{P}^c(E_k) = 1 - \frac{\binom{N-1}{n}}{\binom{N}{n}} = 1 - \frac{N-n}{N} = \frac{n}{N}.$$

Ainsi, par linéarité de l'espérance, on a :

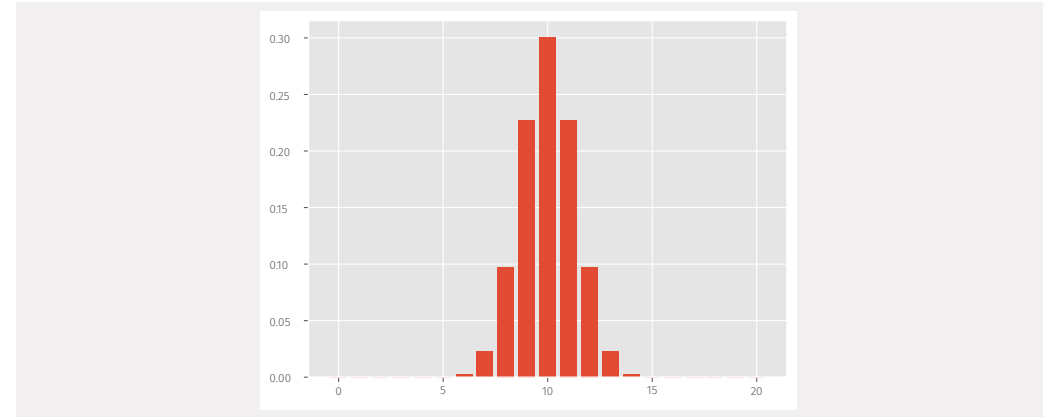
$$\mathbf{E}(X) = \sum_{k=1}^{Np} \mathbf{P}(E_k) = \sum_{k=1}^{Np} \frac{n}{N} = Np \frac{n}{N} = Np \frac{n}{Np + Nq} = p \frac{n}{p+q} = np$$

car $p + q = 1$. Nous admettons la formule de la variance.

HISTOGRAMME, FONCTION DE RÉPARTITION. Même principe que pour la loi uniforme, on reprend les scripts précédents en modifiant le support et la loi. Notez que, étant donnée le choix de paramètres, le support de la loi se concentre dans la partie gauche du graphique (l'espérance vaut $\frac{20}{3}$).

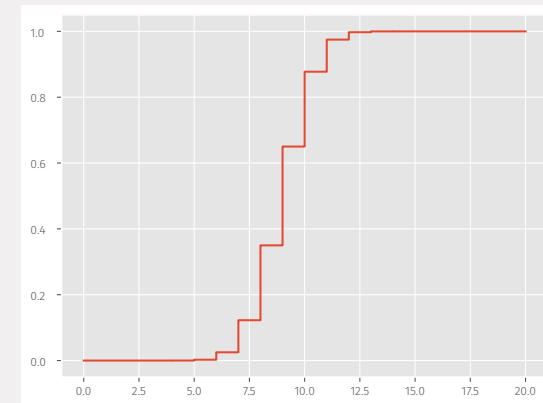
Histogramme de la loi hypergéométrique

```
N = 30
n = 20
p = 0.5
q = 1-p
Support = range(0, n+1)
Loi = [0 for _ in range(len(Support))]
for k in range(len(Support)):
    Loi[k] = binom(N*p, k)*binom(N*q, n-k)/binom(N, n)
plt.bar(Support, Loi)
```



Fonction de répartition de la loi $\mathcal{H}(30, 20, 0.5)$

trace_fdr(Support, Loi)



SIMULATION. On crée deux variables correspondant aux nombres d'objets de type 1 et 2, on fait des tirages selon les proportions calculées et on actualise les proportions après chaque tirage.⁸

⁸Intuitivement, il s'agit de simuler une succession de tirages sans remise dans une urne.

Simulation de la loi hypergéométrique

```

import random as rd
def hypergeometrique(N, n, p):
    """
    simule une hypergeometrique
    N : nombre total d'éléments
    n : nombre d'éléments piochés
    p : proportion éléments type 1
    """
    Nb_type1 = N * p
    Nb_type2 = N - Nb_type1
    S = 0
    for i in range(n):
        if rd.random() < p:
            Nb_type1 -= 1
            S += 1
        else:
            Nb_type2 -= 1
    p = Nb_type1 / (Nb_type1 + Nb_type2) # proportion boules de
    ↪ type 1
    return S

```

Le module random ne sait pas simuler directement la loi hypergéométrique. On fait donc plutôt appel pour cela à la sous bibliothèque random de numpy.

```

np.random.hypergeometric(Np, Nq, n) # Attention à l'ordre des
↪ paramètres, et aux paramètres eux-mêmes

```

Nous passons aux lois propres au programme de seconde année : les lois discrètes dont le support n'est pas fini mais plus généralement dénombrable.

3.4. Loi Géométrique & Absence de mémoire

Définition/Proposition ALEA.13.8 | Loi géométrique de paramètre p

Soit $p \in]0, 1]$, on pose $q = 1 - p$. On dit qu'une variable aléatoire suit une *loi géométrique de paramètres p* (on note $X \hookrightarrow \mathcal{G}(p)$), si :

$$X(\Omega) = \mathbf{N}^*, \quad \forall k \in \mathbf{N}^*, \quad \mathbf{P}(X = k) = pq^{k-1}.$$


Attention

Le support de la loi géométrique est \mathbf{N}^* et non \mathbf{N} : le premier succès ne peut arriver sans commencer l'expérience.

Remarque 3.10 — Modélisation. Temps d'apparition du premier succès dans la répétition, de manière indépendante, d'une expérience aléatoire de BERNOULLI (succès $p \in]0, 1]$, échec $q = 1 - p$).

Remarque 3.11 —

- ▶ Si $p = 1$, alors le premier succès arrive presque-sûrement dès le premier essai donc X devrait être égale à 1 presque-sûrement. On retrouve ce fait en analysant l'expression de $\mathbf{P}(X = k)$ pour tout $k \geq 1$ — en effet, $q = 0$ et donc $\mathbf{P}(X = 1) = 1$ mais $\mathbf{P}(X = k) = 0$ dès que $k \geq 2$.
- ▶ En revanche, on exclut $p = 0$ dans la définition, car dans ce cas on aurait « $X = \infty$ » presque-sûrement. De manière plus pragmatique, la somme des probabilités ne serait pas égale à 1.

Preuve Il existe un espace probabilisé et une variable aléatoire vérifiant ces conditions. En effet, 

Le support de cette loi est **non borné**, donc **l'existence d'une espérance n'est ici plus automatique** et se ramène à l'étude de la convergence d'une série.

Proposition ALEA.13.22 | Espérance, variance

Si $X \hookrightarrow \mathcal{G}(p)$ avec $p \in]0, 1]$, alors X possède une espérance et une variance, et :


$$\mathbf{E}(X) = \frac{1}{p}, \quad \mathbf{Var}(X) = \frac{q}{p^2}.$$

Preuve



Comme pour la loi de BERNOULLI, on peut très facilement obtenir le paramètre connaissant l'espérance.

Preuve Immédiat d'après la formule de l'espérance, et celle de la fonction d'anti-répartition, puisque nous avons montré que pour tout $k \in \mathbf{N}^*$, $\mathbf{P}(X^k) = (1-p)^k$.

Exemple 26 – Probabilité conditionnelle selon une loi géométrique Soient $X \hookrightarrow \mathcal{G}(p)$ avec $p \in]0, 1]$ et A un évènement tel que $\mathbf{P}(A|X = k) = \left(\frac{1}{3}\right)^k$ pour tout entier $k \in \mathbf{N}^*$. Calculer $\mathbf{P}(A)$, puis la loi conditionnelle de X sachant A . 

Corollaire ALEA.13.5 | Obtenir le paramètre

Si $X \hookrightarrow \mathcal{G}(p)$ avec $p \in]0, 1]$ alors :

$$p = 1 - \mathbf{P}(X > 1) = \frac{1}{\mathbf{E}(X)}.$$

Pour les lois précédentes, nous n'étions pas capables de donner une formule explicite pour la fonction de répartition, *i.e.* sans faire intervenir de somme. Pour la loi géométrique, c'est possible.

Proposition ALEA.13.23

Soit $X \hookrightarrow \mathcal{G}(p)$ avec $p \in]0, 1]$. Alors :

- ▶ **(Antirépartition)** pour tout $k \in \mathbf{N}^*$, $1 - F_X(k) = \mathbf{P}(X > k) = q^k$,⁹
- ▶ **(Répartition)** pour tout $k \in \mathbf{N}^*$, $F_X(k) = 1 - q^k$, et pour tout $x \in \mathbf{R}$,

$$F_X(x) = \mathbf{P}(X \leq \lfloor x \rfloor) = \begin{cases} 1 - q^{\lfloor x \rfloor} & \text{si } x \geq 1, \\ 0 & \text{sinon.} \end{cases}$$

Preuve



ABSENCE DE MÉMOIRE. Avant de passer aux autres propriétés de la loi géométrique, définissons de manière générale la notion d'absence de mémoire.

⁹Attention, cette expression est fautive si $k = 0$ par exemple car $\mathbf{P}(X > 0) = \mathbf{P}(X \geq 1) = 1$

Définition ALEA.13.8 | Absence de mémoire

Une variable aléatoire réelle X est dite *sans mémoire* si :

1. elle est positive ou nulle,
2. pour tout couple $(x, y) \in (\mathbf{R}^{++})^2$, on a :

$$\mathbf{P}(X > t + s) = \mathbf{P}(X > s)\mathbf{P}(X > t). \quad (\text{Abs,Mém})$$

Avec les mêmes notations que dans la définition, l'Eq. (Abs,Mém) signifie aussi, de manière équivalente, que la fonction de répartition F vérifie :

$$F(t + s) = F(t) + F(s) - F(t)F(s).$$

Ou encore, en utilisant la formule d'une probabilité conditionnelle, que si $\mathbf{P}(X > s) > 0$:

$$\mathbf{P}(X > t + s | X > s) = \mathbf{P}(X > t).$$

C'est cette dernière formule qui provient de l'intuition.

Remarque 3.12 — Cas où X est discrète telle $X(\Omega) = \mathbf{N}^*$ Prenons le cas où X est discrète telle que $X(\Omega) = \mathbf{N}^*$. Alors la propriété d'absence de mémoire est équivalente à :

$$\forall (k, \ell) \in (\mathbf{N}^{++})^2, \quad \mathbf{P}(X > k + \ell) = \mathbf{P}(X > k)\mathbf{P}(X > \ell).$$

Nous allons à présent établir que la loi géométrique est à absence de mémoire. En fait, on peut même démontrer que c'est la seule loi discrète à l'être. On commence par des propriétés sur la fonction de répartition de la loi géométrique, donc on peut établir une expression.

Proposition ALEA.13.24

Soit $X \hookrightarrow \mathcal{G}(p)$ avec $p \in]0, 1]$. Alors X est à absence de mémoire.

Preuve Soit $k \in \mathbb{N}^{+*}$, rappelons que

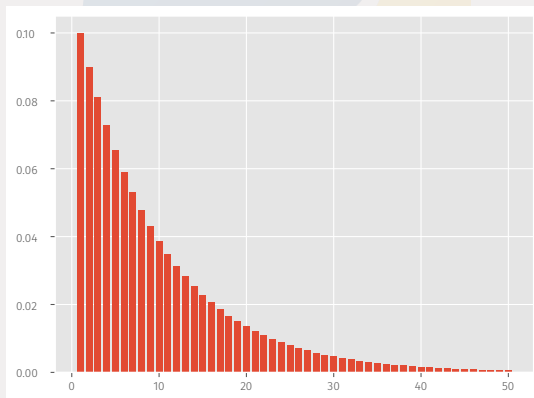
$$P(X > k) = q^k.$$



HISTOGRAMME, FONCTION DE RÉPARTITION.

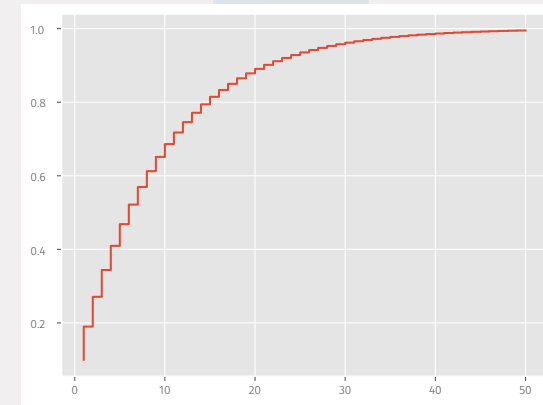
Histogramme de la loi géométrique

```
N = 50 # Troncature du support
n = 20
p = 0.1
q = 1-p
Support = range(1, N+1)
Loi = [0 for _ in range(len(Support))]
for k in range(len(Support)):
    Loi[k] = p*q**k # k démarre à zéro, on décale donc de 1
plt.bar(Support, Loi)
```



Fonction de répartition de la loi $\mathcal{G}(0.1)$

```
# Construction des probabilités cumulées :
Fdr = np.zeros(len(Support))
for k in range(0, len(Support)):
    Fdr[k] = np.sum([Loi[i] for i in range(0, k+1)])
trace_fdr(Support, Loi)
```



SIMULATION. La simulation est encore une fois basée sur un schéma de BERNOULLI puis : on crée deux variables correspondant aux nombres d'objets de type 1 et 2, on fait des tirages selon les proportions calculées et on actualise les proportions après chaque tirage.

Simulation de la loi géométrique

```
import random as rd
def geometrique(p):
    """
    simule une geometrique
    """
    S = 1
```



```
while rd.random() > p:
    S += 1
return S
```

Le module random ne sait pas simuler directement la loi géométrique. On fait donc plutôt appel pour cela à la sous bibliothèque random de numpy.

```
np.random.geometric(p) # Attention à l'ordre des paramètres, et
↳ aux paramètres eux-mêmes
```


3.5. Loi de POISSON

Définition/Proposition ALEA.13.9 | Loi de POISSON de paramètre λ

Soit $\lambda \in \mathbf{R}^{++}$. On dit qu'une variable aléatoire suit une *loi de Poisson de paramètre* λ (on note $X \hookrightarrow \mathcal{P}(\lambda)$), si :

$$X(\Omega) = \mathbf{N}, \quad \forall k \in \mathbf{N}, \quad \mathbf{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Remarque 3.13 — Nous montrerons dans le [Chapter ALEA.16](#) que cette loi peut être approchée par des lois binomiales.

Preuve Il existe un espace probabilisé et une variable aléatoire vérifiant ces conditions. En effet, 

Remarque 3.14 — **Modélisation.** Nombre d'événements se produisant dans un intervalle de temps fixé, si ces événements se produisent avec une fréquence moyenne ou espérance connue et indépendamment du temps écoulé depuis l'événement pré-

cedent.

Proposition ALEA.13.25 | Espérance, variance

Si $X \hookrightarrow \mathcal{P}(\lambda)$ avec $\lambda \in \mathbf{R}^{++}$, alors X possède une espérance et une variance, et :

$$\mathbf{E}(X) = \lambda, \quad \mathbf{Var}(X) = \lambda.$$

Preuve



Corollaire ALEA.13.6 | Obtenir le paramètre

On peut constater que si $X \hookrightarrow \mathcal{P}(\lambda)$ avec $\lambda \geq 0$, alors :

$$\lambda = \mathbf{E}(X).$$

Exemple 27 — **Probabilité conditionnelle selon une loi de Poisson** Soient $X \hookrightarrow \mathcal{P}(\lambda)$ et A un événement tel que $\mathbf{P}(A|X = k) = \left(\frac{1}{3}\right)^k$ pour tout entier $k \in \mathbf{N}$. Calculer $\mathbf{P}(A)$,

puis la loi conditionnelle de X sachant A. 

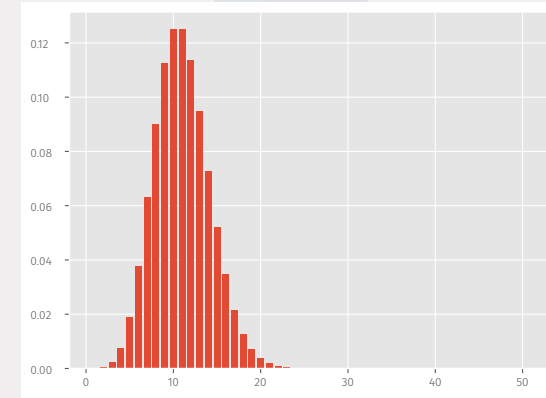
HISTOGRAMME, FONCTION DE RÉPARTITION.

Histogramme de la loi de Poisson

```
N = 50 # Troncature du support
lambda = 10
```

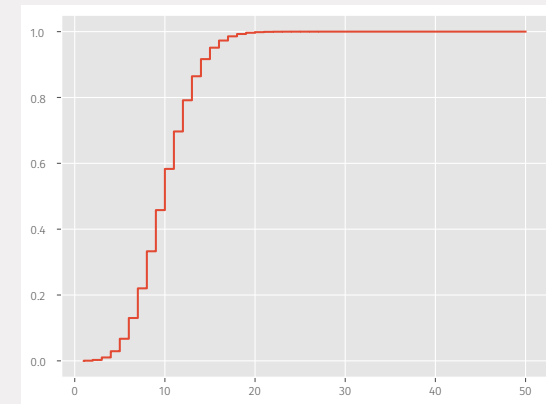


```
Support = range(1, N+1)
Loi = [0 for _ in range(len(Support))]
for k in range(len(Support)):
    Loi[k] = lambda**k/ma.factorial(k)*np.exp(-lambda)
plt.bar(Support, Loi)
```



Fonction de répartition de la loi $\mathcal{P}(10)$

```
trace_fdr(Support, Loi)
```




SIMULATION. Commençons par la propriété qui va nous permettre de simuler la loi de POISSON à partir de l'uniforme sur $[0, 1]$.

Proposition ALEA.13.26 | Simulation de la loi de POISSON

Soit $U \hookrightarrow \mathcal{U}([0, 1])$ et $\lambda > 0$. Alors :

$$X = \min \left\{ n \in \mathbb{N}, \sum_{k=0}^n \frac{\lambda^k e^{-\lambda}}{k!} \geq U \right\} \hookrightarrow \mathcal{P}(\lambda).$$

Preuve

 Il est clair que $X(\Omega) = N$. De plus, soit $k \in N$, alors :

$$\begin{aligned} P(X = k) &= P\left(\sum_{k=0}^{n-1} \frac{\lambda^k e^{-\lambda}}{k!} < U \leq \sum_{k=0}^n \frac{\lambda^k e^{-\lambda}}{k!}\right) \\ &= \text{Long}\left(\left[\sum_{k=0}^{n-1} \frac{\lambda^k e^{-\lambda}}{k!}, \sum_{k=0}^n \frac{\lambda^k e^{-\lambda}}{k!}\right]\right), \quad \text{définition de la loi uniforme} \\ &= \sum_{k=0}^n \frac{\lambda^k e^{-\lambda}}{k!} - \sum_{k=0}^{n-1} \frac{\lambda^k e^{-\lambda}}{k!} \\ &= \frac{\lambda^n e^{-\lambda}}{n!}. \quad \text{téléscopage} \end{aligned}$$

Donc $X \hookrightarrow \mathcal{P}(\lambda)$.

Simulation de la loi de POISSON

On déduit alors à l'aide d'une simple boucle **while** un script permettant de simuler la loi de POISSON.

```
import random as rd
def poisson(lamb):
    """
    Simule une loi de Poisson par inversion
    """
    U = rd.random()
```

```
F = 0
i = 0
while F < U:
    i += 1
    F += np.exp(-lamb)*lamb**i/ma.factorial(i)
return i-1
```

Le module random ne sait pas simuler directement la loi de POISSON. On fait donc plutôt appel pour cela à la sous bibliothèque random de numpy.

```
np.random.poisson(1/lamb) # Attention au paramètre : c'est
↳ l'inverse du paramètre mathématique
```

3.6. Bilan des lois discrètes

Le tableau suivant rassemble quelques lois discrètes usuelles.

Nom	Paramètre(s)	Notation	Support $X(\Omega)$	$P(X = k)$
BERNOULLI	$p \in [0, 1]$	$\mathcal{B}(p)$	$\{0, 1\}$	$p\mathbb{1}_1(k) + (1 - p)\mathbb{1}_0(k)$
BINOMIALE	$(n, p) \in \mathbb{N}^* \times [0, 1]$	$\mathcal{B}(n, p)$	$\{0, \dots, n\}$	$\binom{n}{k} p^k (1 - p)^{n-k}$
HYPER-GÉOMÉTRIQUE	$p \in [0, 1], N \in \mathbb{N}, n \in \{0, \dots, N\}, Np \in \mathbb{N}^*$	$\mathcal{H}(N, n, p)$	$[[0, n]]$	$\frac{\binom{Np}{k} \binom{N(1-p)}{n-k}}{\binom{N}{n}}$
GÉOMÉTRIQUE	$p \in [0, 1]$	$\mathcal{G}(p)$	\mathbb{N}^*	$p(1 - p)^{k-1}$
POISSON	$\lambda \in]0, \infty[$	$\mathcal{P}(\lambda)$	\mathbb{N}	$e^{-\lambda} \frac{\lambda^k}{k!}$

*** **Fin du chapitre** ***

4. EXERCICES

4.1. Généralités

Exercice ALEA.13.1 | Soit X une variable aléatoire discrète à valeurs dans \mathbf{N} , telle que :

$$\mathbf{P}(X = k) = \frac{p^k}{(1+p)^{k+1}}, \quad \text{pour tout } k \in \mathbf{N}, \quad p > 0.$$

1. Justifier l'existence de X .
2. La variable aléatoire X admet-elle une espérance et une variance? Les calculer en cas d'existence.

Solution (exercice ALEA.13.1)

1. Pour tout $k \in \mathbf{N}$, on a

$$\frac{p^k}{(1+p)^{k+1}} \geq 0.$$

De plus, puisque $\left| \frac{p}{1+p} \right| < 1$, on a :

$$\sum_{k=0}^{\infty} \frac{p^k}{(1+p)^k} = \frac{1}{1 - \frac{p}{1+p}} = 1 + p.$$

Donc

$$\sum_{k=0}^{\infty} \frac{p^k}{(1+p)^{k+1}} = 1.$$

Le cours livre l'existence de X telle que

$$\boxed{X(\Omega) = \mathbf{N}, \quad \mathbf{P}(X = k) = \frac{p^k}{(1+p)^{k+1}}.}$$

2. Toujours car $\left| \frac{p}{1+p} \right| < 1$, on sait que

$$\sum_{k=1}^{\infty} k \frac{p^{k-1}}{(1+p)^{k-1}} = \frac{1}{\left(1 - \frac{p}{1+p}\right)^2} = (1+p)^2,$$

et

$$\sum_{k=2}^{\infty} k(k-1) \frac{p^{k-2}}{(1+p)^{k-2}} = \frac{2}{\left(1 - \frac{p}{1+p}\right)^3} = 2 \cdot (1+p)^3.$$

En multipliant la seconde égalité par $\frac{p^2}{(1+p)^3}$, on déduit

$$\sum_{k=2}^{\infty} k(k-1)\mathbf{P}(X = k) = 2p^2 = \sum_{k=0}^{\infty} k\mathbf{P}(X = k).$$

D'après le théorème de transfert, cela signifie que $X(X-1)$ admet une espérance égale à $2p^2$. En multipliant la première égalité par $\frac{p}{(1+p)^2}$, on déduit

$$\sum_{k=1}^{\infty} k\mathbf{P}(X = k) = p = \sum_{k=0}^{\infty} k\mathbf{P}(X = k).$$

D'après le théorème de transfert, cela signifie que X admet une espérance égale à p . Ainsi, $X(X-1) + X = X^2$ admet aussi une espérance et donc X admet une variance et une espérance. On a de plus :

$$\begin{aligned} \mathbf{E}(X^2) &= \mathbf{E}(X(X-1)) + \mathbf{E}(X) \\ &= 2p^2 + p, \end{aligned}$$

$$\begin{aligned} \mathbf{Var}(X^2) &= \mathbf{E}(X^2) - \mathbf{E}(X)^2 \\ &= 2p^2 + p - p^2 = \boxed{p^2 + p} \end{aligned}$$

$$\mathbf{E}(X) = \boxed{p}.$$

Exercice ALEA.13.2 | Soit X une variable aléatoire à valeurs dans \mathbf{N}^* telle que :

$$\forall n \in \mathbf{N}^*, \quad \mathbf{P}(X = n) = \frac{1}{n^2} - \frac{1}{(n+1)^2}.$$

1. Montrer que l'on définit bien ainsi une variable aléatoire.
2. Montrer que X possède une espérance et que $E(X) = \sum_{n=1}^{\infty} \frac{1}{n^2}$. Possède-t-elle une variance?
3. Montrer que $E(X) \leq 2$.

Exercice ALEA.13.3 | Deux applications du théorème de transfert

1. Soient $n \in \mathbb{N}^*$ et $p \in]0, 1[$. On considère une variable aléatoire $X \hookrightarrow \mathcal{B}(n, p)$. Calculer $E\left(\frac{1}{X+1}\right)$ après avoir justifié l'existence.
2. Soient $p \in]0, 1[$ et X une variable aléatoire suivant une loi $\mathcal{G}(p)$. On pose $Y = X^2$. Calculer $E(Y)$ après avoir justifié l'existence.

Exercice ALEA.13.4 | Soit $X \hookrightarrow \mathcal{G}(p)$ avec $p \in]0, 1[$, on note

$$Y = \frac{1}{2} (1 + (-1)^X).$$

Déterminer la loi de Y, son espérance et sa variance.

Solution (exercice ALEA.13.4)

On note $q = p - 1$. Constatons d'abord que $Y(\Omega) = \left\{ \frac{1}{2}(1 + (-1)), \frac{1}{2}(1 + 0) \right\} = \{0, 1\}$. Par ailleurs,

$$\begin{aligned} P(Y = 1) &= P(X \text{ pair}) = P\left(\bigcup_{k=1}^{\infty} \{X = 2k\}\right) \\ &= \sum_{k=1}^{\infty} p q^{2k-1} \\ &= \frac{p}{q} \sum_{k=1}^{\infty} (q^2)^k \\ &= \frac{p}{q} \frac{q^2}{1 - q^2} \quad \left. \begin{array}{l} \\ \end{array} \right\} |q^2| < 1 \\ &= \frac{qp}{p(1+q)} \\ &= \frac{q}{1+q}. \end{aligned}$$

Donc : $Y \hookrightarrow \mathcal{B}\left(\frac{q}{1+q}\right)$

Exercice ALEA.13.5 | Identité de WALD – Espérance d'une somme aléatoire de variables aléatoires discrètes Soit (X_n) une suite de variables aléatoires réelles discrètes i.i.d. de même loi $\mathcal{B}(p)$ avec $p \in]0, 1[$. Soit $N \hookrightarrow \mathcal{P}(\lambda)$ avec $\lambda \in \mathbb{R}$. On définit

$$\text{alors } Y \text{ par : } Y = \begin{cases} 0 & \text{si } N = 0, \\ \sum_{k=1}^N X_k & \text{si } N \neq 0. \end{cases}$$

1. Soit $n \in \mathbb{N}^*$. Quelle est la loi de $S_n = \sum_{k=1}^n X_k$?
2. Déterminer $P(Y = 0)$. Déterminer alors $P(Y = r)$ pour tout $r \in \mathbb{N}^*$.
3. Déterminer $E(Y)$ si elle existe.

Solution (exercice ALEA.13.5)

La difficulté de l'exercice est la suivante : puisque N est une variable aléatoire, on ne peut pas utiliser la linéarité de l'espérance pour calculer l'espérance de Y! En particulier, il est **faux** d'écrire

$$E(Y) = E(X_1) + \dots + E(X_N) = pN \text{ -- une espérance ne peut être aléatoire!}$$

1. Soit $n \in \mathbb{N}^*$. Alors par indépendance des X_k , on sait d'après le cours :

$$S_n = \sum_{k=1}^n X_k \hookrightarrow \mathcal{B}(n, p).$$

2. On a d'après la formule des probabilités totales au système complet d'évènements

$$\{N = 0, N = n, n \in \mathbf{N}^*\},$$

$$\begin{aligned} \mathbf{P}(Y = 0) &= \mathbf{P}(Y = 0|N = 0)\mathbf{P}(N = 0) + \sum_{n=1}^{\infty} \mathbf{P}(Y = 0|N = n)\mathbf{P}(N = n) \\ &= 1 \cdot e^{-\lambda} + \sum_{n=1}^{\infty} p^0(1-p)^{n-0}e^{-\lambda} \frac{\lambda^n}{n!} \\ &= e^{-\lambda} + e^{-\lambda} \sum_{n=1}^{\infty} \frac{(\lambda(1-p))^n}{n!} \\ &= e^{-\lambda} + e^{-\lambda} (e^{\lambda(1-p)} - 1) \\ &= \boxed{e^{-\lambda p}}. \end{aligned}$$

} question précédente

3. Soit $r \in \mathbf{N}^*$, on a d'après la formule des probabilités totales appliquée au système complet d'évènements $\{N = n\}_{n \in \mathbf{N}}$

$$\begin{aligned} \mathbf{P}(Y = r) &= \mathbf{P}(Y = r, N = 0) + \sum_{n=1}^{\infty} \mathbf{P}(Y = r|N = n)\mathbf{P}(N = n) \\ &= 0 + \sum_{n=1}^{\infty} \binom{n}{r} p^r(1-p)^{n-r} e^{-\lambda} \frac{\lambda^n}{n!} \\ &= e^{-\lambda} \sum_{n=r}^{\infty} \frac{1}{r!(n-r)!} p^r(1-p)^{n-r} \lambda^n \\ &= e^{-\lambda} \sum_{n=r}^{\infty} \frac{1}{r!(n-r)!} (\lambda p)^r (\lambda(1-p))^{n-r} \\ &= e^{-\lambda} \frac{(\lambda p)^r}{r!} \sum_{n=0}^{\infty} \frac{1}{n!} (\lambda(1-p))^n \\ &= \frac{(\lambda p)^r}{r!} e^{-\lambda} e^{\lambda(1-p)} \\ &= \frac{(\lambda p)^r}{r!} e^{-\lambda p}. \end{aligned}$$

} $r \geq 1$
} $\lambda^n = \lambda^r \lambda^{n-r}$
} changement d'indice

Puisque de plus $\mathbf{P}(Y = 0) = e^{-\lambda p}$, on déduit que $Y \hookrightarrow \mathcal{P}(\lambda p)$.

4. Donc Y admet une espérance et

$$\boxed{\mathbf{E}(Y) = \lambda p = \mathbf{E}(N) \cdot \mathbf{E}(X_1)}.$$

Cette dernière identité s'appelle l'identité de WALD et est vraie pour des lois plus générales que celles prises dans l'exercice.

Exercice ALEA.13.6 | Étude de la série génératrice Dans tout l'exercice, en cas d'existence, on notera pour X une variable aléatoire discrète, $G_X(t) = \mathbf{E}(t^X)$ pour $t \in \mathbf{R}$. On l'appelle la *série génératrice* de X.

1. Soient X, Y deux variables aléatoires discrètes indépendantes et $t \in \mathbf{R}$ tel que $G_X(t)$ et $G_Y(t)$ convergent. Montrer que : $G_{X+Y}(t) = G_X(t)G_Y(t)$.
2. **(Cas d'un support non borné)** Soit X une variable aléatoire réelle discrète de support $X(\Omega) = \mathbf{N}$.
 - 2.1) Justifier l'existence de G_X sur $] -1, 1[$.
 - 2.2) Étudier l'existence et calculer $G_X(1)$.
3. **(Cas d'un support fini)** On suppose dans cette question que le support de X est fini et composé d'entiers : $X(\Omega) = \{k_0, \dots, k_N\}$ avec $k_i \in \mathbf{N}$ pour tout $i \in \llbracket 0, N \rrbracket$. On suppose de plus que la suite $(k_i)_i$ est croissante.
 - 3.1) Montrer que G_X est une fonction polynomiale, précisez son degré ainsi que ses coefficients. En déduire que deux variables aléatoires finies ont même loi si et seulement si elles ont même série génératrice.
 - 3.2) En déduire que : $\mathbf{P}(X = k_j) = \frac{G_X^{(k_j)}(0)}{k_j!}$ pour tout $j \in \llbracket 0, N \rrbracket$.
 - 3.3) Exprimer espérance et variance de X en fonction de G_X .
 - 3.4) Soit $n \in \mathbf{N}$, et $X_1, \dots, X_n \hookrightarrow \mathcal{B}(p)$ indépendantes avec $p \in]0, 1[$. À l'aide des questions précédentes, établir les faits suivants :
 - ▶ $X_1 + \dots + X_n \hookrightarrow \mathcal{B}(n, p)$,
 - ▶ retrouver les formules de l'espérance et de la variance.

Solution (exercice ALEA.13.6)

1. Soit $t \in \mathbf{R}$, alors par théorème du cours, t^X, t^Y sont indépendantes, donc

$$\mathbf{E}(t^{X+Y}) = \mathbf{E}(t^X \cdot t^Y) = \mathbf{E}(t^X) \cdot \mathbf{E}(t^Y) = G_X(t) \cdot G_Y(t).$$

Donc : $\boxed{G_{X+Y}(t) = G_X(t)G_Y(t)}.$

2. (Cas d'un support non borné) Soit X une variable aléatoire réelle discrète de support $X(\Omega) = \mathbf{N}$.

2.1) Cette fois-ci on a, d'après le théorème de transfert, en cas de convergence absolue,

$$G_X(t) = \sum_{k=0}^{\infty} t^k \mathbf{P}(X = k).$$

On étudie donc la nature de

$$\sum_{k=0}^{\infty} |t|^k \mathbf{P}(X = k).$$

Puisqu'une probabilité est majorée par 1, on a

$$\forall t \in \mathbf{R}, \quad 0 \leq |t|^k \mathbf{P}(X = k) \leq |t|^k,$$

or le majorant est le terme général d'une série convergente si et seulement si $|t| < 1$. Donc si $t \in]-1, 1[$, $\sum_{k=0}^{\infty} |t|^k \mathbf{P}(X = k)$ converge d'après le théorème de comparaison pour les séries à termes positifs. Ainsi,

G_X est définie sur $] -1, 1[$.

2.2) On étudie $\sum_{k=0}^{\infty} 1 \mathbf{P}(X = k)$, cette série converge et est de somme 1, puisque $\{X = k, k \in \mathbf{N}\}$ est un système complet d'évènements. Donc $G_X(1) = 1$.

3. (Cas d'un support fini)

3.1) Puisque $X(\Omega)$ est supposé fini, on a d'après le théorème de transfert :

$$\forall t \in \mathbf{R}, \quad G_X(t) = \sum_{i=0}^N t^{k_i} \mathbf{P}(X = k_i).$$

Donc G_X est une fonction polynomiale de degré k_N puisque la suite des k_i est supposée croissante. De plus, les coefficients sont

$$\mathbf{P}(X = k_0), \dots, \mathbf{P}(X = k_N).$$

Puisque deux polynômes sont égaux si et seulement si leurs coefficients sont égaux, on déduit que

deux variables aléatoires finies ont même loi si et seulement si elles ont même série génératrice.

3.2) Soit $j \in \llbracket 0, N \rrbracket$,

$$\begin{aligned} G_X^{k_j}(t) &= \left(\sum_{i=0}^N t^{k_i} \mathbf{P}(X = k_i) \right)^{k_j} \\ &= \sum_{i=0}^N k_i \cdot (k_i - 1) \cdots (k_i - k_j + 1) t^{k_i - k_j} \mathbf{P}(X = k_i) \\ G_X^{k_j}(0) &= k_j \cdot (k_j - 1) \cdots (k_j - k_j + 1) \mathbf{P}(X = k_j) + 0 \end{aligned}$$

Donc

$$\frac{G_X^{k_j}(0)}{k_j!} = \mathbf{P}(X = k_j).$$

3.3) Soit $t \in \mathbf{R}$. Alors

$$\begin{aligned} G_X'(t) &= \sum_{i=1}^N k_i t^{k_i-1} \mathbf{P}(X = k_i) \\ G_X''(t) &= \sum_{i=2}^N k_i(k_i - 1) t^{k_i-2} \mathbf{P}(X = k_i) \\ G_X'(1) &= \sum_{i=1}^N k_i 1^{k_i-1} \mathbf{P}(X = k_i) = \mathbf{E}(X) \\ G_X''(1) &= \sum_{i=2}^N k_i(k_i - 1) 1^{k_i-2} \mathbf{P}(X = k_i) = \mathbf{E}(X(X - 1)). \end{aligned}$$

Donc $\mathbf{E}(X) = G_X'(1)$, et puisque $\mathbf{E}(X(X - 1)) = G_X''(1)$, on déduit que

$$\mathbf{E}(X^2) = \mathbf{E}(X(X - 1)) + \mathbf{E}(X) = G_X''(1) + G_X'(1),$$

donc par KÖNIG-HUYGENS :

$$\mathbf{Var}(X) = G_X''(1) + G_X'(1) - G_X'(1)^2.$$

3.4) Soit $n \in \mathbf{N}$, et $X_1, \dots, X_n \hookrightarrow \mathcal{B}(p)$ indépendantes avec $p \in]0, 1[$. À l'aide des questions précédentes, établir les faits suivants :

- ▶ Notons $S_n = X_1 + \dots + X_n$. Alors montrons que S_n a même série génératrice qu'une $\mathcal{B}(n, p)$, on commence par calculer la série génératrice de S_n . Puisque $X_1 + \dots + X_n, X_1, \dots, X_n$ sont des variables aléatoires finies, elles admettent une fonction génératrice définie sur \mathbf{R} . Et par indépendance

$$\forall t \in \mathbf{R}, \quad G_{X_1+\dots+X_n}(t) = G_{X_1(t)} \cdots G_{X_n}(t).$$

Or, pour tout $i \in \llbracket 1, n \rrbracket$,

$$G_{X_i}(t) = t^0(1-p) + t^1 p = 1 - p + tp.$$

Donc

$$\forall t \in \mathbf{R}, \quad G_{X_1+\dots+X_n}(t) = (1 - p + tp)^n.$$

Par ailleurs, notons $X \hookrightarrow \mathcal{B}(n, p)$. Alors pour tout $t \in \mathbf{R}$:

$$\begin{aligned} G_X(t) &= \sum_{k=0}^n t^k \mathbf{P}(X = k) \\ &= \sum_{k=0}^n t^k \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=0}^n \binom{n}{k} (pt)^k (1-p)^{n-k} \\ &= (pt + 1 - p)^n = G_{S_n(t)}. \end{aligned} \quad \left. \begin{array}{l} \\ \\ \end{array} \right\} \text{binôme}$$

Donc $S_n \hookrightarrow \mathcal{B}(n, p)$.

- ▶ Nous avons montré que

$$\forall t \in \mathbf{R}, \quad G_{S_n}(t) = (pt + 1 - p)^n.$$

Donc

$$\begin{aligned} \mathbf{E}(S_n) &= G'_{S_n}(1) = \left. \frac{d[(pt + 1 - p)^n]}{dt} \right|_{t=1} \\ &= pn(p + 1 - p)^{n-1} \\ &= np, \end{aligned}$$

$\mathbf{Var}(S_n)$

$$\begin{aligned} &= G''_{S_n}(1) + G'_{S_n}(1) - G'_{S_n}(1)^2 \\ &= p^2 n(n-1)(p + 1 - p)^{n-2} + pn(p + 1 - p)^{n-1} - (pn(p + 1 - p)^{n-1})^2 \\ &= np(1-p). \end{aligned}$$

Exercice ALEA.13.7 | Formule des cumulants, cas discret Soit X une variable aléatoire réelle discrète telle que $X(\Omega) = \mathbf{N}^*$. On souhaite démontrer et utiliser que X possède une espérance si et seulement si $\sum_{k=1}^{\infty} \mathbf{P}(X \geq k)$ converge, et que dans ce cas :

$$\mathbf{E}(X) = \sum_{k=1}^{\infty} \mathbf{P}(X \geq k). \quad (\star)$$

1. (Application) Soient X, Y deux variables aléatoires indépendantes de même loi $\mathcal{G}\left(\frac{1}{3}\right)$. En utilisant (\star) , montrer que $Z = \min(X, Y)$ possède une espérance et calculer $\mathbf{E}(Z)$.

2. On souhaite montrer la propriété admise en introduction.

2.1) Montrer que pour tout $N \in \mathbf{N}^*$:

$$\begin{aligned} S_N &= \sum_{k=1}^N k \mathbf{P}(X = k) \\ &= \sum_{k=1}^{N+1} \mathbf{P}(X \geq k) - (N+1) \mathbf{P}(X \geq N+1). \end{aligned}$$

2.2) Conclure.

3. Retrouver le résultat de la première question sans utiliser (\star) .

Solution (exercice ALEA.13.7)

1. (Application) Notons $Z = \min(X, Y)$, et soit $k \in \mathbf{N}^*$. Alors

$$\begin{aligned} \mathbf{P}(Z \geq k) &= \mathbf{P}(X \geq k, Y \geq k) \\ &= \mathbf{P}(X \geq k) \cdot \mathbf{P}(Y \geq k) \quad \left. \begin{array}{l} \\ \end{array} \right\} \text{indépendance} \\ &= \left(\sum_{\ell=k}^{\infty} \left(\frac{1}{3} \frac{2}{3} \right)^{\ell-1} \right)^2 \\ &= \left(\frac{2}{3} \right)^{2(k-1)}. \end{aligned}$$

Donc comme $|\frac{4}{9}| < 1$, la somme $\sum_{k=1}^{\infty} \mathbf{P}(Z \geq k)$ converge, et

$$\sum_{k=1}^{\infty} \mathbf{P}(Z \geq k) = \sum_{k=1}^{\infty} \left(\frac{4}{9} \right)^{k-1} = \frac{1}{1 - \frac{4}{9}} = \frac{9}{5}.$$

Donc $\min(X, Y)$ admet une espérance, et

$$\mathbf{E}(Z) = \frac{9}{5}.$$

2. On souhaite montrer la propriété admise en introduction.

2.1) Soit $N \in \mathbf{N}^*$:

$$\begin{aligned} S_N &= \sum_{k=1}^N k (\mathbf{P}(X \geq k) - \mathbf{P}(X \geq k+1)) \\ &= \sum_{k=1}^N (k\mathbf{P}(X \geq k) - k\mathbf{P}(X \geq k+1)) \\ &= \sum_{k=1}^N (k\mathbf{P}(X \geq k) - (k+1)\mathbf{P}(X \geq k+1)) + \sum_{k=1}^N \mathbf{P}(X \geq k+1) \quad \left. \begin{array}{l} k = (k+1) - 1 \\ \end{array} \right\} \\ &= \mathbf{P}(X \geq 1) - (N+1)\mathbf{P}(X \geq N+1) + \sum_{k=1}^N \mathbf{P}(X \geq k+1) \quad \left. \begin{array}{l} \\ \end{array} \right\} \text{téléscopage} \\ &= \sum_{k=1}^{N+1} \mathbf{P}(X \geq k) - (N+1)\mathbf{P}(X \geq N+1). \quad (\star) \end{aligned}$$

2.2) \Rightarrow Supposons dans un premier temps que X possède une espérance. Alors la suite $(S_N)_{N \in \mathbf{N}^*}$ converge vers $S \in \mathbf{R}$. On a envie, vue la question, de montrer que

$$(N+1)\mathbf{P}(X \geq N+1) \xrightarrow{N \rightarrow \infty} 0.$$

Et en effet, c'est le cas :

$$\begin{aligned} (N+1)\mathbf{P}(X \geq N+1) &= (N+1) \sum_{k=N+1}^{\infty} \mathbf{P}(X = k) = \sum_{k=N+1}^{\infty} (N+1)\mathbf{P}(X = k) \\ &\leq \sum_{k=N+1}^{\infty} k\mathbf{P}(X = k). \end{aligned}$$

Or $\sum_{k=N+1}^{\infty} k\mathbf{P}(X = k) \xrightarrow{N \rightarrow \infty} 0$ puisque c'est le reste d'une série convergente, donc par théorème d'encadrement on déduit le résultat. Donc en faisant $N \rightarrow \infty$ dans (\star) , on déduit que $\sum_{k=1}^{\infty} \mathbf{P}(X \geq k)$ converge, et que :

$$\mathbf{E}(X) = \sum_{k=1}^{\infty} \mathbf{P}(X \geq k).$$

\Rightarrow Supposons dans un second temps que $\sum_{k=1}^{\infty} \mathbf{P}(X \geq k)$ converge et montrons que $(S_N)_N$ converge. D'après (\star) , on a

$$S_N + (N+1)\mathbf{P}(X \geq N+1) = \sum_{k=1}^{N+1} \mathbf{P}(X \geq k).$$

Donc : $S_N \leq \sum_{k=1}^{N+1} \mathbf{P}(X \geq k)$. Ainsi, puisque $\sum_{k=1}^{\infty} \mathbf{P}(X \geq k)$ converge, la suite $\left(\sum_{k=1}^{N+1} \mathbf{P}(X \geq k) \right)_N$ est majorée (car somme partielle d'une somme positive), donc $(S_N)_N$ est également majorée. Mais comme c'est une suite croissante (puisque la série associée à l'espérance est positive comme $X(\Omega) = \mathbf{N}^*$), elle converge donc. Et donc X admet une espérance. Enfin, le sens précédent livre à nouveau que

$$(N+1)\mathbf{P}(X \geq N+1) \xrightarrow{N \rightarrow \infty} 0,$$

donc

$$\mathbf{E}(X) = \sum_{k=1}^{\infty} \mathbf{P}(X \geq k).$$

On a donc bien prouvé l'équivalence de la question.

3. On peut utiliser le calcul de la loi de $Z = \max(X, Y)$. On a clairement $Z(\Omega) = \mathbf{N}^*$. Pour la loi d'un minimum on peut calculer l'antirépartition. On a déjà montré que :

$$\mathbf{P}(Z \geq k) = \left(\frac{2}{3}\right)^{2(k-1)}.$$

Donc pour tout $k \in \mathbf{N}^*$:

$$\begin{aligned} \mathbf{P}(Z = k) &= \mathbf{P}(Z \geq k) - \mathbf{P}(Z \geq k + 1) \\ &= \left(\frac{2}{3}\right)^{2(k-1)} - \left(\frac{2}{3}\right)^{2k} \\ &= \left(\frac{2}{3}\right)^{2(k-1)} \left(1 - \left(\frac{2}{3}\right)^2\right) \\ &= \frac{5}{9} \cdot \left(\frac{2}{3}\right)^{2(k-1)} \\ &= \frac{5}{4} \cdot \left(\frac{4}{9}\right)^k. \end{aligned}$$

On retrouve ensuite l'espérance de Z en utilisant des formules de sommes géométriques dérivées.

$$\begin{aligned} \mathbf{E}(Z) &= \frac{5}{4} \sum_{k=1}^{\infty} k \left(\frac{4}{9}\right)^k \\ &= \frac{5}{4} \frac{4}{9} \sum_{k=1}^{\infty} k \left(\frac{4}{9}\right)^{k-1} \\ &= \frac{5}{9} \frac{1}{\left(1 - \frac{4}{9}\right)^2} \quad \left. \begin{array}{l} \left. \right) \left| \frac{4}{9} \right| < 1 \end{array} \right\} \\ &= \boxed{\frac{9}{5}}. \end{aligned}$$

4.2. Expérience aléatoire

Exercice ALEA.13.8 | Une rame de tram/bus circule sur une ligne de 4 stations numérotées de 0 à 3. Quand il arrive à la station 3 il fait demi tour, de même à la station 0. On suppose qu'il passe 1 min à chaque station avec un temps de trajet négligeable entre deux stations.

1. Un(e) étudiant(e) de BCPST s'endort après à la station 0 après une soirée trop festive. On suppose qu'à chaque arrivée en station, il se réveille avec probabilité $p \in]0, 1[$ et que les réveils/poursuites de siestes sont indépendants. On note X le numéro de station à laquelle il se réveille. Déterminer la loi de X . Calculer son espérance et sa variance.
2. On note Y le nombre d'aller-retours effectués. Déterminer la loi de Y .

Solution (exercice ALEA.13.8)

$X(\Omega) = \{0, 1, 2, 3\}$. On note $q = 1 - p$. $\mathbf{P}(X = 0) = \mathbf{P}(T \in 6\mathbf{N}^*) = \sum_{k=1}^{+\infty} \mathbf{P}(X = 6k) = \sum_{k=1}^{+\infty} q^{6k-1} p = \frac{pq^5}{1 - q^6}$. $\mathbf{P}(X = 1) = \mathbf{P}(T \in (6\mathbf{N} + 1) \cup (6\mathbf{N} + 5)) = \sum_{k=0}^{+\infty} \mathbf{P}(X = 6k + 1) + \sum_{k=0}^{+\infty} \mathbf{P}(X = 6k + 5) = \frac{p + pq^4}{1 - q^6}$. $\mathbf{P}(X = 2) = \sum_{k=0}^{+\infty} \mathbf{P}(X = 6k + 2) + \sum_{k=0}^{+\infty} \mathbf{P}(X = 6k + 4) = \frac{\mathbf{P}(q^3 + q)}{1 - q^6}$. $\mathbf{P}(X = 3) = \sum_{k=0}^{+\infty} \mathbf{P}(X = 6k + 3) = \frac{pq^2}{1 - q^6}$. $Y(\Omega) \in \mathbf{N}$. Pour tout $k \in \mathbf{N}$, $Y = k \Leftrightarrow T = 6k + r$ avec $r \in \{0, 1, 2, 3, 4, 5\}$. Si $k \in \mathbf{N}^*$, $\mathbf{P}(Y = k) = p \sum_{j=0}^5 q^{6k-1+j} = q^{6k-1} (1 - q^6)$. Si $k = 0$, $\mathbf{P}(Y = 0) = \mathbf{P}(T \in \{1, 2, 3, 4, 5\}) = \sum_{j=0}^5 pq^{6k-1+j} = 1 - q^5$.

Exercice ALEA.13.9 | **Paradoxe de Saint-Pétersbourg** Un casino propose à ses clients le jeu suivant. Le joueur commence par verser à la banque une mise de un million d'euros. On lance alors une pièce de monnaie (non truquée). Si pile tombe, la banque verse au joueur un euro et le jeu s'arrête. Si c'est face qui tombe, on relance la pièce de monnaie et, si c'est pile qui tombe à ce deuxième lancer, la banque verse au joueur

deux euros (et le jeu s'arrête). On relance ainsi la pièce de monnaie jusqu'à ce que pile tombe (ce qui arrive presque-sûrement). À chaque lancer de la pièce, le montant qui sera versé au joueur en cas de sortie de pile est doublé. Le jeu est-il équitable, favorable à la banque ou bien favorable au joueur?

Solution (exercice ALEA.13.9)

On note X le gain du joueur (c'est-à-dire la somme d'argent qui lui sera versée par la banque quand pile sortira). La variable aléatoire X prend les valeurs 1, 2, 4, 8, 16, 32, etc.. Pour tout $k \in \mathbf{N}^*$, on a $\mathbf{P}(X = 2^k) = \frac{1}{2^{k+1}}$ car c'est la probabilité de $k + 1$ évènements indépendants de même probabilité $\frac{1}{2}$ (celle de faire pile au lancer de la pièce). Il s'agit de se poser la question de l'existence d'une espérance pour la variable aléatoire X , i.e. sur la convergence absolue de $\left(\sum_{k \geq 0} \frac{2^k}{2^{k+1}}\right)$. Or, cette série diverge grossièrement, donc X n'admet pas d'espérance, plus précisément la somme partielle de la série diverge vers $+\infty$. Le jeu est donc plus favorable au joueur. Ceci paraît un peu délirant dans la mesure où la mise initiale d'entrée dans le jeu est très élevée; et en fait peu importe cette mise, le jeu sera toujours plus favorable au joueur.

4.2.1. Autour de la loi uniforme

Exercice ALEA.13.10 | Étude d'un min / max On lance deux dés à 6 faces honnêtes. On note alors X le plus grand des numéros obtenus et Y le plus petit.

- Déterminer les lois de X et de Y .
- Calculer $\mathbf{E}(X)$ et $\mathbf{E}(Y)$. Comparer les espérances et commenter.

Solution (exercice ALEA.13.10)

- On a $X(\Omega) = Y(\Omega) = \llbracket 1, 6 \rrbracket$. De plus, introduisons $U_1, U_2 \hookrightarrow \mathcal{U} \llbracket 1, 6 \rrbracket$ deux variables aléatoires indépendantes de sorte que $X = \max(X, Y), Y = \min(X, Y)$. Soit donc $k \in$

$\llbracket 1, 6 \rrbracket$,

$$\begin{aligned} \mathbf{P}(X \leq k) &= \mathbf{P}(U_1 \leq k, U_2 \leq k) \\ &= \mathbf{P}(U_1 \leq k) \mathbf{P}(U_2 \leq k) \quad \left. \vphantom{\mathbf{P}(X \leq k)} \right\} \text{indépendance} \\ &= \left(\frac{k}{6}\right)^2. \end{aligned}$$

Donc pour tout $k \geq 1$,

$$\mathbf{P}(X = k) = \mathbf{P}(X \leq k) - \mathbf{P}(X \leq k - 1) = \begin{cases} \left(\frac{k}{6}\right)^2 - \left(\frac{k-1}{6}\right)^2 & \text{si } k \geq 2, \\ \left(\frac{1}{6}\right)^2 & \text{si } k = 1. \end{cases}$$

Les deux formules se réunissent en une seule :

$$\forall k \in \llbracket 1, 6 \rrbracket, \quad \mathbf{P}(X = k) = \left(\frac{k}{6}\right)^2 - \left(\frac{k-1}{6}\right)^2.$$

Pour le minimum, on utilise plutôt l'antirépartition.

$$\begin{aligned} \mathbf{P}(Y \geq k) &= \mathbf{P}(U_1 \geq k, U_2 \geq k) \\ &= \mathbf{P}(U_1 \geq k) \mathbf{P}(U_2 \geq k) \quad \left. \vphantom{\mathbf{P}(Y \geq k)} \right\} \text{indépendance} \\ &= \left(\frac{6-k+1}{6}\right)^2. \end{aligned}$$

Donc pour tout $k \geq 1$,

$$\mathbf{P}(Y = k) = \mathbf{P}(Y \geq k) - \mathbf{P}(Y \geq k + 1) = \begin{cases} \left(\frac{6-k+1}{6}\right)^2 - \left(\frac{6-k}{6}\right)^2 & \text{si } k \leq 5, \\ \left(\frac{1}{6}\right)^2 & \text{si } k = 6. \end{cases}$$

Les deux formules se réunissent en une seule :

$$\forall k \in \llbracket 1, 6 \rrbracket, \quad \mathbf{P}(Y = k) = \left(\frac{6-k+1}{6}\right)^2 - \left(\frac{6-k}{6}\right)^2.$$

- Calculons $\mathbf{E}(X)$ et $\mathbf{E}(Y)$. Pour Y , on peut constater la relation ci-après afin de gagner du temps :

$$U_1 + U_2 = X + Y.$$

La variable aléatoire X est finie, donc

$$\begin{aligned}
 \mathbf{E}(X) &= \sum_{k=1}^6 k \left(\left(\frac{k}{6}\right)^2 - \left(\frac{k-1}{6}\right)^2 \right) \\
 &= \sum_{k=1}^6 \left(k \left(\frac{k}{6}\right)^2 - (k-1) \left(\frac{k-1}{6}\right)^2 \right) - \sum_{k=1}^6 \left(\frac{k-1}{6}\right)^2 \\
 &= 6 \left(\frac{6}{6}\right)^2 - 0 - \sum_{k=1}^6 \left(\frac{k-1}{6}\right)^2 \quad \left. \begin{array}{l} \text{téléscopage} \\ \text{changement d'indice} \end{array} \right\} \\
 &= 6 \left(\frac{6}{6}\right)^2 - \frac{1}{36} \sum_{k=0}^5 k^2 \\
 &= \boxed{\frac{91}{36}}.
 \end{aligned}$$

Puisque $\mathbf{E}(U_1) = \mathbf{E}(U_2) = \frac{1+6}{2} = \frac{7}{2}$, on déduit que

$$\mathbf{E}(Y) = 7 - \frac{91}{36} = \boxed{\frac{161}{36}}.$$

On observe bien

$$\mathbf{E}(X) = \mathbf{E}(\min(U_1, U_2)) \leq \mathbf{E}(\min(U_1, U_2)) = \mathbf{E}(Y).$$

4.2.2. Autour des lois binomiales et hypergéométriques

Exercice ALEA.13.11 | Équilibrage d'une stratégie de jeu Soit N un entier naturel ≥ 2 . On définit la fonction f sur \mathbf{R}^2 par

$$\forall (x, y) \in \mathbf{R}^2, \quad f(x, y) = y(N(N-1) - x(x-1)) - Nx(x-1).$$

1. (Résolution informatique d'une équation dans $\llbracket 1, 10 \rrbracket^2$)

- 1.1) ➤_🔗 Écrire une fonction Python d'en-tête $f(x, y, N=10)$ qui retourne la valeur de $f(x, y)$ avec 10 pour valeur de N par défaut.
- 1.2) À l'aide de la fonction précédente, dans le cas $N = 10$, donner tous les couples (x, y) d'entiers de $\llbracket 1, 10 \rrbracket$ tels que $f(x, y) = 0$.
2. Une urne A contient $N = 10$ tickets dont x sont gagnants, une urne B contient 10 tickets dont y gagnants.
 - ▶ Un joueur E tire 2 tickets dans l'urne A.
 - ▶ Si les 2 tickets tirés sont gagnants, le joueur E a gagné,
 - ▶ sinon le joueur F tire un ticket dans l'urne B et est déclaré gagnant s'il tire un ticket gagnant.
 - ▶ S'il n'y pas de gagnant, la partie est déclarée nulle.
- 2.1) Calculer la probabilité que E soit gagnant, et que F soit gagnant.
- 2.2) Déterminer x et y tels que la partie soit équitable.
- 2.3) ➤_🔗 Écrire une fonction Python Gagnant $(x, y, N=10)$ simulant ce jeu et retournant +1 si E gagne, 0 si la partie est nulle et -1 si F gagne.
3. Vérifier, à l'aide de Python, que pour le(s) couple(s) (x, y) trouvés en précédemment, on a effectivement un jeu équitable.

Solution (exercice ALEA.13.11)

1. 1.1)

```

1.2) def f(x, y, N=10):
    return y*(N*(N-1) - x*(x-1)) - N*x*(x-1)
def couplesf(N=10):
    L = []
    for k in range(1, N+1):
        for l in range(1, N+1):
            if f(k, l) == 0:
                L.append([k, l])
    return L
    
```

La fonction couplesf() renvoie $\llbracket 6, 5 \rrbracket$.

2. 2.1) Notons G_E l'évènement «le joueur E est gagnant», on tire simultanément

deux tickets dans une urne qui contient x gagnants, et $N - x$ perdants : c'est donc un contexte de loi hypergéométrique. On a donc : $\mathbf{P}(G_E) = \frac{\binom{2}{2} \binom{N-x}{0}}{\binom{N}{2}} =$

$$\frac{\frac{x(x-1)}{2} \cdot 1}{\frac{N(N-1)}{2}} = \frac{x(x-1)}{N(N-1)}.$$

Notons G_F l'évènement «le joueur F est gagnant». Nous avons $G_F = (G_F \cap G_E) \cup (G_F \cap {}^c G_E) = G_F \cap {}^c G_E$ puisque le premier évènement est vide. Donc :

$$\mathbf{P}(G_F) = \mathbf{P}(G_F | {}^c G_E) \mathbf{P}({}^c G_E) = \frac{y}{N} \left(1 - \frac{x(x-1)}{N(N-1)} \right).$$

- 2.2) On cherche donc les couples (x, y) tels que $\frac{y}{N} \left(1 - \frac{x(x-1)}{N(N-1)} \right) = \frac{x(x-1)}{N(N-1)}$. En manipulant l'équation (multiplication par $N(N-1)$) on trouve comme condition : $f(x, y) = 0$, et d'après python : $x=6, y=5$. La partie est donc équitable pour cet unique couple.

```
3. def Gagnant(x, y, N=10):
    #Tour de E
    if rd.random() < x/N:
        if rd.random() < (x-1)/(N-1):
            return 1
    #Tour de F
    if rd.random() < y/N:
        return -1
    #Personne n'a gagné
    return 0
```


4. Pour cela il suffit de simuler un grand nombre de fois l'expérience et de compter les proportions de victoires pour chaque joueur. Par exemple de cette manière :

```
def Equitable(x, y, Nb_Simu, N=10):
    """
    Proportion de parties gagnantes pour chaque jour sur Nb_Simu
    """
    Gagne_E, Gagne_F = 0, 0
```

```
for _ in range(Nb_Simu):
    Res = Gagnant(x, y, N)
    if Res == 1:
        Gagne_E += 1
    elif Res == -1:
        Gagne_F += 1
return Gagne_E/Nb_Simu, Gagne_F/Nb_Simu
```

On constate que les proportions sont très proches pour $N=10000$: `Equitable(4, 6, 100)` renvoie $(0.17, 0.48)$.

Exercice ALEA.13.12 | On dispose d'une infinité de boules blanches, numérotées et représentées par leur indice $k \in \mathbf{N}^*$. L'urne comporte initialement a boules noires non numérotées. Si l'on tire une boule blanche, on la remet dans l'urne avec une autre boule blanche de même numéro. Si on tire une boule noire, on la remet et on ajoute une boule blanche portant un numéro non présent dans l'urne (peu importe lequel, les variables aléatoires définies *infra* ne dépendront pas de ce choix). On appelle Y_k la variable aléatoire égale à 1 si une boule noire est tirée au k -ième tirage et 0 sinon. On note X_n la variable aléatoire égale au nombre de numéros différents présents dans l'urne avant le $(n+1)$ -ième tirage.

- Quelle est la probabilité de tirer une boule noire au n -ième tirage ?
- Soit $n \in \mathbf{N}$. Exprimer X_n en fonction des Y_k pour k et n entiers, montrer que $\mathbf{E}(X_n) = a \sum_{k=a}^{a+n-1} \frac{1}{k}$. Trouver une formule similaire pour $\mathbf{Var}(X_n)$.
- 
 - Écrire une fonction `SimulX(a, n)` qui simule une variable de même loi que X_n .
 - Écrire une fonction qui permet d'estimer $\mathbf{E}(X_n)$, conjecturer la limite $\frac{\mathbf{E}(X_n)}{n}$ quand n tend vers l'infini.
- Déterminer un équivalent de $\mathbf{E}(X_n)$.

Solution (exercice ALEA.13.12)

1. À chaque étape nous avons rajouté une boule soit noire soit blanche, ainsi avant le tirage n , nous avons

$$a + n - 1 \text{ boules dans l'urne.}$$

2. On a alors $\mathbf{P}(Y_n = 1) = \frac{a}{a+n-1}$ la proportion de boules noires au n -ième tirage.

3. La variable X_n est égale au nombre de tirages de boules noires, *i.e.* le nombre de fois que l'on a ajouté un numéro différent. Ainsi, $X_n = \sum_{k=1}^n Y_k$.

En utilisant la linéarité de l'espérance, on en déduit que : $\mathbf{E}(X_n) = \sum_{k=1}^n \mathbf{E}(Y_k) = \sum_{k=1}^n \mathbf{P}(Y_k = 1) = \sum_{k=1}^n \frac{a}{a+k-1}$. Il suffit alors de faire le changement de variable affine $\ell = a + k - 1$, fournissant ainsi l'égalité demandée :

$$\mathbf{E}(X_n) = a \sum_{\ell=a}^{a+n-1} \frac{1}{\ell}.$$

Pour la variance, étant donné que les variables aléatoires Y_k sont indépendantes (en effet, Y_k ne dépend pas du tirage précédent, mais uniquement du rang du tirage, *i.e.* l'entier k , puisque la proportion de boules noires ne dépend que de k). Ainsi,

$$\mathbf{Var}(X_n) = \sum_{k=1}^n \mathbf{Var}(Y_k) = \sum_{k=1}^n \left(1^2 \frac{a}{a+k-1} - \left(\frac{a}{a+k-1} \right)^2 \right) = \sum_{k=1}^n \frac{a}{a+k-1} \frac{a+k-1-a}{a+k-1} =$$

$$a \sum_{k=1}^n \frac{k-1}{(a+k-1)^2}.$$

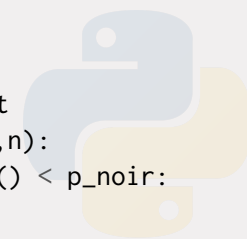
De la même manière, nous pouvons faire ensuite le changement de variable $\ell = a + k - 1$, qui donne

$$\mathbf{Var}(X_n) = a \sum_{\ell=a}^{a+n-1} \frac{\ell-a}{\ell^2} = a \left(\sum_{\ell=a}^{a+n-1} \frac{1}{\ell} - a \sum_{\ell=a}^{a+n-1} \frac{1}{\ell^2} \right).$$

4. 4.1) `import random as rd`
`def SimulX(a, n):`



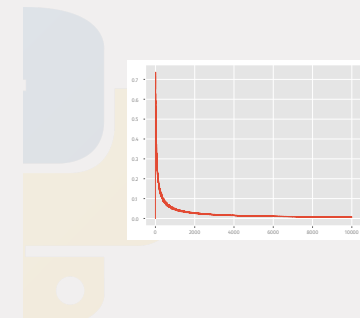
```
(a : pro boules noires, n : nb tirages)->nombre de
boules de numéros différents
"""
S = 0
Nb_tot = a
p_noir = a/Nb_tot
for _ in range(1,n):
    if rd.random() < p_noir:
        S += 1
        Nb_tot += 1
        p_noir = a/Nb_tot
return S
```



Par exemple, `SimulX(10, 50)` renvoie 19.

4.2) En effectuant un grand nombre de simulations N , la moyenne des réalisations de X_n converge vers l'espérance.

```
import random as rd
def EspX(a, n, N):
    """
    (a, n, N)-> simule Xn
    N fois, renvoie la
    moyenne
    """
    return
    sum([SimulX(a,n)
        for _ in
        range(N)])/N
import matplotlib.pyplot
as plt
plt.plot([EspX(10, n,
10)/n for n in
range(1,10000)])
```



Il nous reste ensuite à regarder la suite donnée dans l'énoncé en fonction de n : on trouve $\text{Esp}X(5, 100, 1000)$ environ égale à 0.27127. La suite semble converger vers zéro, mais la convergence est très lente (on ne la constate même pas clairement si on trace jusque $n = 100$). Cette lenteur s'explique par l'équivalent trouvé ci-dessous.

5. On réalise une comparaison série—intégrale. Il suffit de dire que la fonction $t \mapsto \frac{1}{t}$ est continue, strictement décroissante sur $[1, \infty[$.

On somme l'encadrement entre a et $a + n - 1$ pour obtenir :

$$\int_a^{a+n} \frac{1}{t} dt \leq \mathbf{E}(X_n) \leq \int_{a-1}^{a+n-1} \frac{1}{t} dt.$$

En calculant explicitement explicitement les intégrales, on trouve pour $a \geq 2$:

$$a \ln \left| \frac{a+n}{a} \right| \leq \mathbf{E}(X_n) \leq a \ln \left| \frac{a+n-1}{a-1} \right|.$$

À l'aide de cet encadrement, on déduit que : $\lim_{n \rightarrow \infty} \frac{\mathbf{E}(X_n)}{a \ln(a+n)} = 1$ soit :

$\mathbf{E}(X_n) \sim_{n \rightarrow \infty} a \ln(a+n)$. Pour $a = 1$, on prouve de la même manière (mais on ne somme qu'à partir de $k = 2$ et on ajoute à la fin le terme $k = 1$), que : $\mathbf{E}(X_n) \sim_{n \rightarrow \infty} \ln(n)$. C'est l'équivalent de la série harmonique calculé en cours.

4.2.3. Autour de la loi géométrique

Exercice ALEA.13.13 | On dispose d'une pièce truquée telle que la probabilité d'obtenir pile est α , $\alpha \in]0, 1[$. On lance deux fois la pièce :

- ▶ si on obtient (F, P) on a gagné;
- ▶ si on obtient (P, F) on a perdu;
- ▶ sinon on recommence.

Déterminer le nombre moyen X de (couples de) lancers effectués.

Solution (exercice ALEA.13.13)


Notons X le nombre de lancers. Le nombre de parties correspond au nombre de succès (faire (F, P) ou (P, F)) dans une répétition d'expériences de BERNOULLI (lancer deux fois la pièce) indépendantes, puisque le jeu s'arrête une fois un succès obtenu. Par ailleurs, la probabilité de succès est

$$\mathbf{P}((P, F) \cup (F, P)) = \alpha(1 - \alpha) + (1 - \alpha)\alpha = 2\alpha(1 - \alpha).$$

Donc $X \hookrightarrow \mathcal{G}(2\alpha(1 - \alpha))$, et le nombre moyen de lancers est donc

$$\mathbf{E}(X) = \frac{1}{2\alpha(1 - \alpha)}.$$

Exercice ALEA.13.14 | Chance de capture d'un Pokemon Blue joue à Pokemon rouge. Il se retrouve face à MEWTWO dans la caverne azurée. La probabilité de capture de MEWTWO est très faible : $\frac{1}{n}$ pour un entier $n \in \mathbf{N}^*$ très grand. Blue, équipé d'un stock de n pokéballs, s'apprête à toutes les lancer une à une.

1. Calculer la probabilité p_n que Blue capture MEWTWO.
2. Déterminer $\lim_{n \rightarrow \infty} p_n$.
3. On suppose que suite à la découverte d'un bug dans le jeu, qu'on dispose d'une infinité de pokéballs pour tenter notre capture. Déterminer le temps d'attente moyen jusqu'à capture de MEWTWO.
4.  Proposer un programme Python donnant une simulation du temps de capture.

Solution (exercice ALEA.13.14)

1. Le plus simple est de passer par le complémentaire. Notons C_n «il capture en n

essais» et pour tout $i \in \llbracket 1, n \rrbracket$, E_i «l'essai i échoue». Alors

$$\begin{aligned} 1 - p_n &= \mathbf{P}({}^c C_n) \\ &= \mathbf{P}(E_1 \cap \dots \cap E_n) \\ &= \mathbf{P}(E_1) \mathbf{P}(E_2 | E_1) \dots \mathbf{P}(E_n | E_1 \cap \dots \cap E_{n-1}) \quad \left. \vphantom{\mathbf{P}(E_1)} \right\} \text{probabilités composées} \\ &= \left(1 - \frac{1}{n}\right)^n. \end{aligned}$$

Donc : $p_n = 1 - \left(1 - \frac{1}{n}\right)^n$.

2. Pour la limite, il s'agit de revenir à la forme exponentielle puis de faire un développement limité.

$$\begin{aligned} p_n &= 1 - \exp\left(n \ln\left(1 - \frac{1}{n}\right)\right) \\ &= 1 - \exp\left(n \left(-\frac{1}{n} + o\left(\frac{1}{n}\right)\right)\right) \\ &\xrightarrow{n \rightarrow \infty} 1 - e^{-1}. \end{aligned}$$

3. Cette fois-ci, il s'agit du temps d'attente d'un succès (la capture, de probabilité $\frac{1}{n}$ dans une répétition d'expériences de BERNOULLI (capturer ou non) indépendantes. Donc $T \hookrightarrow \mathcal{G}(1/n)$, et le nombre moyen de balls lancées est $\mathbf{E}(T) = n$.

```
4. import random as rd
def geometrique(n):
    """
    simule une geometrique de paramètre 1/n
    """
    balls = 1
    while rd.random() > 1/n:
        balls += 1
    return balls
```

Exercice ALEA.13.15 | Délivrer la princesse Une princesse est retenue prisonnière dans un château. Un prince charmant se met en tête de la délivrer. Lorsqu'il arrive à l'entrée du château, il se trouve devant 3 portes. Il en ouvre une au hasard (équiprobable).

- ▶ S'il ouvre la 1ère porte, il délivre la princesse.
- ▶ S'il ouvre la deuxième porte, un dragon apparait et le dévore.
- ▶ S'il ouvre la troisième porte, une sorcière lui fait boire un filtre, il oublie tout ce qu'il a vu et est mis à la porte du château.

Le prince renouvelle ses tentatives jusqu'à ce qu'il meure ou qu'il délivre la princesse.

1. Calculer la probabilité de l'événement D_k : «il délivre la princesse au k -ème essai».
2. Calculer la probabilité de l'événement D : «il délivre la princesse».
3. On note T le nombre de tentatives du prince. Donner la loi de T ainsi que son espérance.
4. Si le prince échoue dans sa tâche, le syndicat des princes envoie immédiatement un autre prince, jusqu'à ce que la princesse soit délivrée. On suppose que le prince suivant, après avoir vu le précédent prince se faire dévorer, sera tout aussi performant. Calculer le nombre moyen de princes utilisés pour délivrer la princesse.

Solution (exercice ALEA.13.15)

1. Notons S_k l'évènement «il choisit la sorcière au k -ième essai», avec $k \in \mathbf{N}^*$. Alors

$$\begin{aligned} \mathbf{P}(D_k) &= \mathbf{P}(S_1 \cap \dots \cap S_{k-1} \cap D_k) \\ &= \mathbf{P}(S_1) \mathbf{P}(S_2 | S_1) \dots \mathbf{P}(S_{k-1} | S_1 \cap \dots \cap S_{k-2}) \mathbf{P}(D_k | S_1 \cap \dots \cap S_{k-1}) \\ &= \frac{1}{3^k}. \end{aligned}$$

2. On a $D = \bigcup_{k=1}^{\infty} D_k$. Ainsi, par additivité dénombrable d'une probabilité :

$$\begin{aligned} \mathbf{P}(D) &= \sum_{k=1}^{\infty} \mathbf{P}(D_k) \\ &= \sum_{k=1}^{\infty} \frac{1}{3^k} \\ &= \frac{1}{3} \frac{1}{1 - \frac{1}{3}} \\ &= \boxed{\frac{1}{2}}. \end{aligned}$$

3. La variable aléatoire T compte le temps d'attente d'un succès (choisir la bonne porte ou le dragon, car dans ces deux cas l'expérience s'arrête) dans une répétition d'expériences de BERNOULLI (le choix de la bonne porte/dragon ou la sorcière) indépendantes. Par ailleurs, la probabilité de choisir la la porte de la princesse ou du dragon est $\frac{2}{3}$. Donc

$$T \hookrightarrow \mathcal{G}(2/3).$$

Ainsi $\mathbf{E}(T) = \boxed{\frac{3}{2}}$.

4. Dans ce cas, l'expérience continue tant que le prince est dévoré. Cette fois-ci l'expérience de BERNOULLI associée est la suivante : on envoie un prince au casse-pipe, soit il délivre la princesse (succès, de probabilité $\mathbf{P}(D) = \frac{1}{2}$), soit il ne la délivre pas (échec de probabilité $\frac{1}{2}$ aussi). Donc le nombre de prince envoyés, puisque les envois sont supposés indépendantes, suit une géométrique de paramètre $\frac{1}{2}$. Le nombre moyen de princes envoyés est donc de 2.

4.2.4. Atour de la loi de Poisson

Exercice ALEA.13.16 |

1. Justifier que pour tout $x \in \mathbf{R}$, on a :

$$\frac{e^x + e^{-x}}{2} = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!}, \quad \frac{e^x - e^{-x}}{2} = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!}.$$

Dans la suite, on pourra noter :

$$\forall x \in \mathbf{R}, \quad \text{ch } x = \frac{e^x + e^{-x}}{2}, \quad \text{sh } x = \frac{e^x - e^{-x}}{2}.$$

2. On tire un nombre entier naturel X au hasard, et on suppose que $X \hookrightarrow \mathcal{P}(a)$ avec $a > 0$.

- ▶ Si X est impair, PIERRE gagne et reçoit X euros de PAUL.
- ▶ Si X est pair non nul, PAUL gagne et reçoit X euros de PIERRE.
- ▶ Si X = 0, la partie est nulle.

On note p la probabilité que PIERRE gagne et q la probabilité que PAUL gagne.

- 2.1) Déterminer la valeur de p et de q.
- 2.2) Détermine l'espérance des gains de chacun.

Solution (exercice ALEA.13.16)

1. On sait que $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$ pour tout x réel. Donc

$$\begin{aligned} \frac{e^x + e^{-x}}{2} &= \sum_{k=0}^{\infty} \frac{x^k}{k!} + \sum_{k=0}^{\infty} \frac{(-1)^k x^k}{k!} \\ &= \sum_{k=0}^{\infty} \frac{1 + (-1)^k}{2} \frac{x^k}{k!} \\ &= \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!} = \text{ch } x. \end{aligned}$$

linéarité de la somme
1 + (-1)^k = 0 si k impair, 2 sinon

On obtient de-même :

$$\forall x \in \mathbf{R}, \quad \text{sh } x = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!}.$$

2. 2.1) Nous avons

$$\begin{aligned}
 p &= \mathbf{P}(\text{PIERRE gagne}) = \mathbf{P}(X \text{ impair}) = \mathbf{P}\left(\biguplus_{k=0}^{\infty} \{X = 2k + 1\}\right) \\
 &= \sum_{k=0}^{\infty} \mathbf{P}(X = 2k + 1) \\
 &= \sum_{k=0}^{\infty} \frac{a^{2k+1}}{(2k + 1)!} e^{-a} \\
 &= e^{-a} \cdot \text{sh}(a) = \frac{1 - e^{-2a}}{2}.
 \end{aligned}$$

De-même,

$$\begin{aligned}
 p &= \mathbf{P}(\text{PAUL gagne}) = \mathbf{P}(X \text{ pair non nul}) = \mathbf{P}\left(\biguplus_{k=1}^{\infty} \{X = 2k\}\right) \\
 &= \sum_{k=1}^{\infty} \mathbf{P}(X = 2k) \\
 &= \sum_{k=1}^{\infty} \frac{a^{2k}}{(2k)!} e^{-a} \\
 &= e^{-a} \cdot \text{ch}(a) = \frac{1 + e^{-2a}}{2}.
 \end{aligned}$$

2.2) Notons G_a le gain de PAUL et G_i le gain de PIERRE. Alors

► $G_i(\Omega) = \{2k + 1, k \in \mathbf{N}\} \uplus \{-2k, k \in \mathbf{N}^*\}$. De plus,

$$\begin{aligned}
 \forall k \in \mathbf{N}, \quad \mathbf{P}(G_i = 2k + 1) &= \mathbf{P}(X = 2k + 1) \\
 &= e^{-a} \frac{a^{2k+1}}{(2k + 1)!}. \\
 \forall k \in \mathbf{N}^*, \quad \mathbf{P}(G_i = -2k) &= \mathbf{P}(X = 2k) \\
 &= e^{-a} \frac{a^{2k}}{(2k)!}.
 \end{aligned}$$

Donc pour l'espérance, on étudie la nature de

$$\sum_{k=0}^{\infty} (2k + 1) \left(e^{-a} \frac{a^{2k+1}}{(2k + 1)!} \right), \quad \sum_{k=0}^{\infty} |(-2k)| \left(e^{-a} \frac{a^{2k}}{(2k)!} \right).$$

Les deux séries convergent car, après avoir simplifié les factorielles, on obtient des séries exponentielles convergentes d'après la première question. De plus,

$$\begin{aligned}
 \sum_{k=0}^{\infty} (2k + 1) \left(e^{-a} \frac{a^{2k+1}}{(2k + 1)!} \right) &= e^{-a} a \sum_{k=0}^{\infty} \frac{a^{2k}}{(2k)!} = e^{-a} a \cdot \text{ch}(a) \\
 &= a \frac{1 + e^{-2a}}{2}, \\
 \sum_{k=1}^{\infty} (2k) \left(e^{-a} \frac{a^{2k}}{(2k)!} \right) &= e^{-a} a \sum_{k=0}^{\infty} \frac{a^{2k-1}}{(2k - 1)!} = e^{-a} a \cdot \text{sh}(a) \\
 &= a \frac{1 - e^{-2a}}{2}.
 \end{aligned}$$

Donc G_i possède une espérance, et

$$\mathbf{E}(G_i) = a \left(\frac{1 - e^{-2a}}{2} + \frac{1 + e^{-2a}}{2} \right) = a.$$

► $G_a(\Omega) = \{-(2k + 1), k \in \mathbf{N}\} \uplus \{2k, k \in \mathbf{N}^*\}$. De plus,

$$\begin{aligned}
 \forall k \in \mathbf{N}, \quad \mathbf{P}(G_a = -(2k + 1)) &= \mathbf{P}(X = 2k + 1) \\
 &= e^{-a} \frac{a^{2k+1}}{(2k + 1)!}. \\
 \forall k \in \mathbf{N}^*, \quad \mathbf{P}(G_a = 2k) &= \mathbf{P}(X = 2k) \\
 &= e^{-a} \frac{a^{2k}}{(2k)!}.
 \end{aligned}$$

Donc pour l'espérance, on étudie la nature de

$$\sum_{k=0}^{\infty} |-(2k + 1)| \left(e^{-a} \frac{a^{2k+1}}{(2k + 1)!} \right), \quad \sum_{k=0}^{\infty} (2k) \left(e^{-a} \frac{a^{2k}}{(2k)!} \right).$$

Les deux séries convergent car, après avoir simplifié les factorielles, on obtient des séries exponentielles convergentes d'après la première

question. De plus,

$$\begin{aligned}\sum_{k=0}^{\infty} (-2k+1) \left(e^{-a} \frac{a^{2k+1}}{(2k+1)!} \right) &= -e^{-a} a \sum_{k=0}^{\infty} \frac{a^{2k}}{(2k)!} = -e^{-a} a \cdot \text{ch}(a) \\ &= -a \frac{1+e^{-2a}}{2}, \\ \sum_{k=1}^{\infty} (2k) \left(e^{-a} \frac{a^{2k}}{(2k)!} \right) &= e^{-a} a \sum_{k=1}^{\infty} \frac{a^{2k-1}}{(2k-1)!} = e^{-a} a \cdot \text{sh}(a) \\ &= a \frac{1-e^{-2a}}{2}.\end{aligned}$$

Donc G_i possède une espérance, et

$$\mathbf{E}(G_i) = a \left(-\frac{1+e^{-2a}}{2} + \frac{1-e^{-2a}}{2} \right) = -ae^{-2a}.$$

4.3. Matrices aléatoires

Exercice ALEA.13.17 | Matrice aléatoire binomiale Soit $(\Omega, \mathcal{F}, \mathbf{P})$ un espace probabilisé, et A et B deux variables aléatoires indépendantes de même loi $\mathcal{B}(n, p)$ définies sur cet espace. Soit M la variable aléatoire, à valeurs dans $\mathcal{M}_2(\mathbf{R})$, définie pour tout $\omega \in \Omega$ par

$$M(\omega) = \begin{pmatrix} A(\omega) & B(\omega) \\ A(\omega) & B(\omega) \end{pmatrix}.$$

- Proposer une fonction Python qui simule M . On renverra le résultat sous forme de tableau numpy.
- Donner la loi, l'espérance et la variance du nombre de valeurs propres de M . Quelle est la probabilité que M soit diagonalisable?

- Donner la loi de la plus grande des valeurs propres.
- Quelle est la probabilité que $M^2 = 0$? vérifie $M^2 = M$? Proposer une commande Python permettant de retrouver ces résultats par simulation, à l'aide de la première question.

Solution (exercice ALEA.13.17)

- M est inversible si et seulement si $AB - BA \neq 0$. Donc l'évènement « M inversible » est finalement l'ensemble vide, donc de probabilité nulle. Calculons à présent M^2 .
Nous avons $M^2(\omega) = \begin{pmatrix} A(\omega) & B(\omega) \\ A(\omega) & B(\omega) \end{pmatrix} \begin{pmatrix} A(\omega) & B(\omega) \\ A(\omega) & B(\omega) \end{pmatrix} = \begin{pmatrix} (A^2 + AB)(\omega) & (AB + B^2)(\omega) \\ (A^2 + AB)(\omega) & (AB + B^2)(\omega) \end{pmatrix}$. La matrice précédente est nulle si et seulement si $A + B = 0$ ou bien $A = 0$ et $B = 0$. Il reste à écrire cela avec des réunions et intersections.

$$\mathbf{P}(M \text{ nilpotente}) = \mathbf{P}(\{A = 0\} \cap \{B = 0\}) \cup \{A + B = 0\}).$$

Attention, les deux évènements précédents ne sont pas disjoints, on ne peut écrire qu'il s'agit de la somme des deux probabilités. Mais, $\{A + B = 0\} = \{A = 0\} \cap \{B = 0\}$ puisque A, B sont positives. On déduit, par indépendance de A et B :

$$\mathbf{P}(M \text{ nilpotente}) = \mathbf{P}(A = 0)\mathbf{P}(B = 0) = (1-p)^{2n}.$$

Nous avons par ailleurs,

$$\begin{aligned}\mathbf{P}(M^2 = M) &= \mathbf{P}(A^2 + AB = A, AB + B^2 = B) = \mathbf{P}(A(A+B-1) = 0, B(B+A-1) = 0) \\ &= \mathbf{P}(A+B=1) + \mathbf{P}(A=0)\mathbf{P}(B=0) \\ &= \mathbf{P}(A=0, B=1) + \mathbf{P}(A=1, B=0) + \mathbf{P}(A=0)\mathbf{P}(B=0) \\ &= \mathbf{P}(A=0)\mathbf{P}(B=1) + \mathbf{P}(A=1)\mathbf{P}(B=0) + \mathbf{P}(A=0)\mathbf{P}(B=0) \\ &= 2n(1-p)^n p(1-p)^{n-1} + (1-p)^{2n}\end{aligned}$$

Dans ce cas, en revanche, les évènements $\{A+B-1=0\}$ et $\{A=0, B=0\}$ sont bien disjoints, ce qui légitime le calcul précédent.

2. Notons N la variable aléatoire donnant le nombre de valeurs propres. Alors $\mathbf{N}(\Omega) = \{1, 2\}$ puisque M est de format 2×2 . De plus, par calcul simple de déterminant, on déduit que : $\text{Spec}(M) = \{0, A+B\}$. Donc $\mathbf{P}(N = 1) = \mathbf{P}(A + B = 0) = \mathbf{P}(A = 0)\mathbf{P}(B = 0) = (1 - p)^{2n}$. Et $\mathbf{P}(N = 2 = 1 - (1 - p)^{2n})$. L'espérance vaut alors $\mathbf{E}(N) = 1 - (1 - p)^{2n}$ et $\mathbf{Var}(N) = 1 - (1 - p)^{2n}$. Utilisons le système complet $\{N = 1\}, \{N = 2\}$.

$$\begin{aligned} \mathbf{P}(M \text{ diagonalisable}) &= \mathbf{P}(M \text{ diagonalisable} \cap \{N = 1\}) + \mathbf{P}(M \text{ diagonalisable} \cap \{N = 2\}) \\ &= \mathbf{P}(M \text{ diagonalisable} \cap \{N = 1\}) + \mathbf{P}(\{N = 2\}) \end{aligned}$$

où à la dernière ligne nous avons utilisé le fait que toute matrice ayant deux valeurs propres distinctes est diagonalisable. De plus,

$$\mathbf{P}(M \text{ diagonalisable} \cap \{N = 1\}) = \mathbf{P}(A = 0, B = 0, A = B = 0) = \mathbf{P}(A = 0)\mathbf{P}(B = 0) = (1 - p)^{2n}.$$

Car toute matrice diagonalisable ayant une seule valeur propre est du type λI avec $\lambda \in \mathbf{R}$. Donc

$$\mathbf{P}(M \text{ diagonalisable}) = (1 - p)^{2n} + 1 - (1 - p)^{2n}.$$

3. La plus grande valeur propre au sens large est $A + B$, qui suit, par indépendance, une $\mathcal{B}(2n, p)$.

4.4. Récurrence & Chaîne de MARKOV

Exercice ALEA.13.18 | Marche aléatoire vers l'avant avec retour en zéro Un mobile se déplace sur un axe de la façon suivante :

- à l'instant 0, il est au point d'abscisse 0;
- si, à l'instant n , le mobile est au point d'abscisse $k \in \mathbf{N}$, alors, à l'instant $n + 1$, soit il sera au point d'abscisse $k + 1$ avec une probabilité $p \in]0, 1[$, soit il retournera au point 0 avec une probabilité $1 - p$.

On note X_n la variable aléatoire égale à l'abscisse du mobile à l'instant n .

1. >_☛

- Créer une fonction d'en-tête `Simu_X(n, p)` qui simule une réalisation de X_n .
- En déduire une valeur approchée de $\mathbf{E}(X_n)$ à l'aide de la fonction précédente.

2. Déterminer la loi de X_1 . Déterminer l'ensemble des valeurs prises par X_n .

3. Soient $n \in \mathbf{N}^*$ et $k \in \llbracket 1, n \rrbracket$. Trouver une relation entre $\mathbf{P}(X_n = k)$ et $\mathbf{P}(X_{n-1} = k - 1)$.

4. En déduire une relation de récurrence entre $\mathbf{E}(X_n)$ et $\mathbf{E}(X_{n-1})$ pour $n \geq 1$.

5. En déduire une expression de $\mathbf{E}(X_n)$ en fonction de $n \in \mathbf{N}$ et p .

Solution (exercice ALEA.13.18)

1. >_☛

```
1.1) import random as rd
def Simu_X(n, p):
    X = 0
    for _ in range(1, n+1):
        if rd.random() < p:
            X += 1
        else:
            X = 0
    return X
```

```
>>> Simu_X(30, 0.5)
0
```

1.2) Une valeur approchée de $\mathbf{E}(X_n)$ est donnée par la loi des grands nombres, on fait une moyenne de beaucoup de simulations de X_n .

```
N = 10**3 # arbitrairement grand
S = 0
for _ in range(N):
```

```
S += Simu_X(100, 0.5)

>>> S/N      # approximation de l'espérance de X10 pour proba
_ 1/2
1.028
```

- 2. On a $X_1 \in \mathcal{B}(p)$. De plus, $X_n(\Omega) = \{0, \dots, n\}$, puisque dans le pire des cas nous aurons avancé de 1 à chaque fois.
- 3. Soient $n \in \mathbf{N}^*$ et $k \in \llbracket 1, n \rrbracket$. Si on se trouve à la position $k \geq 1$ au temps n , c'est qu'avant on se trouvait à la position $k - 1$. Donc

$$\mathbf{P}(X_n = k) = \mathbf{P}(X_n = k, X_{n-1} = k - 1) = \mathbf{P}(X_n = k | X_{n-1} = k - 1) \mathbf{P}(X_{n-1} = k - 1),$$

donc

$$\mathbf{P}(X_n = k) = p \mathbf{P}(X_{n-1} = k - 1).$$

- 4. Constatons tout d'abord que X_n est, pour tout $n \in \mathbf{N}^*$, une variable aléatoire finie. Donc

$$\mathbf{E}(X_n) = \sum_{k=0}^n k \mathbf{P}(X_n = k) = \sum_{k=1}^n k \mathbf{P}(X_n = k),$$

donc d'après la question précédente

$$\mathbf{E}(X_n) = \sum_{k=1}^n k p \mathbf{P}(X_{n-1} = k - 1) = p \sum_{k=1}^n k \mathbf{P}(X_{n-1} = k - 1) = p \sum_{\ell=0}^{n-1} (\ell + 1) \mathbf{P}(X_{n-1} = \ell).$$

On reconnaît alors l'espérance de X_{n-1} puisque $X_{n-1}(\Omega) = \llbracket 0, n - 1 \rrbracket$. D'où

$$\forall n \in \mathbf{N}^*, \quad \mathbf{E}(X_n) = p \mathbf{E}(X_{n-1}).$$

- 5. Puisqu'on a une suite géométrique, on déduit l'expression en fonction de n :

$$\forall n \in \mathbf{N}^*, \quad \mathbf{E}(X_n) = p^{n-1} \mathbf{E}(X_1) = p^{n-1} p = p^n.$$

Donc

$$\mathbf{E}(X_n) = \begin{cases} p^n & \text{si } n \geq 1, \\ 0 & \text{sinon.} \end{cases}$$

Exercice ALEA.13.19 | Après la pluie viendra le beau temps, enfin presque-sûrement, deux états. Dans un certain pays, il fait rarement beau deux jours de suite.

- ▶ Si un jour il fait beau, le lendemain il fait moche avec probabilité $\frac{2}{3}$.
- ▶ Si un jour il fait moche, il y a une chance sur deux qu'il y ait changement de temps le lendemain.

On note M «il fait moche», et B «il fait beau», et X_n l'état du temps au jour n pour tout n , ainsi que pour tout $n \geq 1$:

$$Z_n = \begin{pmatrix} \mathbf{P}(X_n = B) \\ \mathbf{P}(X_n = M) \end{pmatrix}.$$

- 1. Déterminer une matrice $M \in \mathcal{M}_{2,2}(\mathbf{R})$ telle que pour tout $n \in \mathbf{N}^*$,

$$Z_{n+1} = M \cdot Z_n.$$

- 2. Déterminer une matrice P inversible et une matrice D diagonale, de format 2×2 , telles que $M = P \cdot D \cdot P^{-1}$.
- 3. ➤ Proposer une fonction `simu_temps(n, temps_initial)` et qui retourne une simulation de X_n , où le temps initial à $n = 1$ est donné par `temps_initial`. Conjecturer alors une comparaison entre les deux limites ci-dessous, en admettant l'existence :

$$\lim_{n \rightarrow \infty} \mathbf{P}(X_n = M), \quad \lim_{n \rightarrow \infty} \mathbf{P}(X_n = B).$$

Le temps initial influe-t-il sur ces limites ?

- 4. Démontrer cette conjecture. Est-ce conforme à l'intuition? *On pourra commencer par introduire $Y_n = P^{-1} \cdot X_n$ pour tout $n \in \mathbf{N}^*$, et chercher une récurrence simple sur Y_n*

Solution (exercice ALEA.13.19)

1. Soit $n \in \mathbf{N}$. Appliquons la formule des probabilités totales au système complet d'évènements $(\{X_n = B\}, \{X_n = M\})$.

$$\begin{aligned} \mathbf{P}(X_{n+1} = B) &= \mathbf{P}(X_{n+1} = B | X_n = B) \mathbf{P}(X_n = B) + \mathbf{P}(X_{n+1} = B | X_n = M) \mathbf{P}(X_n = M) \\ &= \left(1 - \frac{2}{3}\right) \cdot \mathbf{P}(X_n = B) + \frac{1}{2} \cdot \mathbf{P}(X_n = M), \end{aligned}$$

$$\begin{aligned} \mathbf{P}(X_{n+1} = M) &= \mathbf{P}(X_{n+1} = M | X_n = B) \mathbf{P}(X_n = B) + \mathbf{P}(X_{n+1} = M | X_n = M) \mathbf{P}(X_n = M) \\ &= \frac{2}{3} \cdot \mathbf{P}(X_n = B) + \frac{1}{2} \cdot \mathbf{P}(X_n = M). \end{aligned}$$

Donc :

$$\forall n \in \mathbf{N}, \quad Z_{n+1} = M \cdot Z_n \text{ avec } M = \begin{pmatrix} \frac{1}{3} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{2} \end{pmatrix}.$$

2. Soit $\lambda \in \mathbf{C}$, alors

$$\begin{aligned} M - \lambda I_2 &= \begin{pmatrix} \frac{1}{3} - \lambda & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{2} - \lambda \end{pmatrix} \\ \tilde{L} \begin{pmatrix} \frac{2}{3} & \frac{1}{2} - \lambda \\ \frac{1}{3} - \lambda & \frac{1}{2} \end{pmatrix} &\leftarrow L_1 \leftrightarrow L_2 \\ \tilde{L} \begin{pmatrix} \frac{2}{3} & \frac{1}{2} - \lambda \\ 0 & P(\lambda) \end{pmatrix} &\leftarrow L_2 \leftarrow \frac{2}{3}L_2 - (1/3 - \lambda)L_1 \end{aligned}$$

où $P(\lambda) = \frac{1}{2} \cdot \frac{2}{3} - \left(\frac{1}{3} - \lambda\right) \left(\frac{1}{2} - \lambda\right) = -\left(\lambda + \frac{1}{6}\right) (\lambda - 1)$. Donc $\text{Spec}(M) = \left\{-\frac{1}{6}, 1\right\}$.

De plus,

► $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_1(M)$ si et seulement si :

$$\frac{2}{3}x - \frac{1}{2}y = 0 \iff X \in \text{Vect} \begin{pmatrix} 3 \\ 4 \\ 1 \end{pmatrix}.$$

► $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_{-\frac{1}{6}}(M)$ si et seulement si :

$$\frac{2}{3}x - \frac{2}{3}y = 0 \iff X \in \text{Vect} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}.$$

Donc :

$$M = P \cdot D \cdot P^{-1}, \quad D = \begin{pmatrix} -1/6 & 0 \\ 0 & 1 \end{pmatrix}, \quad P = \begin{pmatrix} -1 & 3 \\ 1 & 4 \end{pmatrix}, \quad P^{-1} = \frac{1}{7} \begin{pmatrix} 4 & -3 \\ -1 & -1 \end{pmatrix}.$$

3. `import random as rd`

```
def simu_temps(n, temps_initial):
    """
    1 : beau temps
    0 : mauvais temps
    retourne une simulation du temps au bout de n jours
    """
    temps = temps_initial
    for _ in range(2, n+1):
        if temps == 1:
            if rd.random() < 2/3:
                temps = 0
        else:
            # il faisait donc moche
            if rd.random() < 1/2:
                temps = 1
    return temps
```

Affichons par exemple 10 essais pour le jour 200.

```
>>> L = [simu_temps(200, 1) for _ in range(100)]
>>> len([temps for temps in L if temps == 1]) # nb beau
42
```

```
>>> len([temps for temps in L if temps == 0]) # nb moche
```

58

On constate une plus forte probabilité de mauvais temps, on conjecture alors :

$$\lim_{n \rightarrow \infty} \mathbf{P}(X_n = B) \leq \lim_{n \rightarrow \infty} \mathbf{P}(X_n = M).$$

4. Est-ce conforme à l'intuition? Oui, puisque la probabilité conditionnelle $2/3$ est élevée. Posons

$$Y_n = P^{-1}Z_n.$$

Alors

$$\begin{aligned} Y_{n+1} &= P^{-1}Z_{n+1} \\ &= P^{-1}PDP^{-1}Z_n \\ &= DY_n. \end{aligned}$$

Donc notons $Y_n = \begin{pmatrix} a_n \\ b_n \end{pmatrix}$ avec $(a_n), (b_n)$ deux suites réelles. Alors

$$\forall n \in \mathbf{N}^*, \quad a_{n+1} = -\frac{1}{6}a_n, \quad b_{n+1} = b_n,$$

donc comme $|\frac{1}{6}| < 1$, on a

$$\lim_{n \rightarrow \infty} a_n = 0, \quad \lim_{n \rightarrow \infty} b_n = b_1.$$

Or,

$$Z_n = P \cdot Y_n = \begin{pmatrix} -1 & 3 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} a_n \\ b_n \end{pmatrix} \xrightarrow{n \rightarrow \infty} \begin{pmatrix} -1 & 3 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 0 \\ b_1 \end{pmatrix} = \begin{pmatrix} 3b_1 \\ 4b_1 \end{pmatrix}.$$

Or $Y_1 = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} = P^{-1} \begin{pmatrix} \mathbf{P}(X_1 = B) \\ \mathbf{P}(X_1 = M) \end{pmatrix}$, donc

$$\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} = \frac{1}{7} \begin{pmatrix} 4 & -3 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} \mathbf{P}(X_1 = B) \\ \mathbf{P}(X_1 = M) \end{pmatrix},$$

donc en analysant la deuxième ligne du produit matriciel, on obtient :

$$b_1 = \frac{1}{7} (\mathbf{P}(X_1 = B) + \mathbf{P}(X_1 = M)) = \frac{1}{7}.$$

D'où

$$Z_n \xrightarrow{n \rightarrow \infty} \begin{pmatrix} \frac{3}{7} \\ \frac{4}{7} \end{pmatrix}.$$

Et

$$\lim_{n \rightarrow \infty} \mathbf{P}(X_n = B) = \frac{3}{7}, \quad \lim_{n \rightarrow \infty} \mathbf{P}(X_n = M) = \frac{4}{7}.$$

Exercice ALEA.13.20 | Souris dans un tunnel, trois états. Des chercheurs font des expériences sur des souris, ils disposent de trois tunnels A, B et C. Les deux premiers sont des cul-de-sac et le troisième, le C, permet à la souris de sortir. On constate que :

- ▶ la première fois, elle choisit au hasard l'un des trois tunnels.
- ▶ Lorsque la souris se trompe (donc aboutit à un cul-de-sac), la fois d'après, elle choisit au hasard l'un des deux autres tunnels.
- ▶ Lorsqu'elle réussit à sortir, la fois d'après, elle reprend le même tunnel.

On note X_n le tunnel choisi (donc A,B,C) à la $n \in \mathbf{N}^*$ -ième tentative de sortie, et on note pour tout entier $n \in \mathbf{N}^*$,

$$Z_n = \begin{pmatrix} \mathbf{P}(X_n = A) \\ \mathbf{P}(X_n = B) \\ \mathbf{P}(X_n = C) \end{pmatrix}.$$

1. Déterminer Z_1 .
2. Déterminer une matrice $M \in \mathfrak{M}_{3,3}(\mathbf{R})$ telle que

$$\forall n \in \mathbf{N}^*, \quad Z_{n+1} = MZ_n.$$

- Déterminer une matrice P inversible et une matrice D diagonale, de format 3×3 , telles que $M = P \cdot D \cdot P^{-1}$. Il ne sera pas nécessaire de calculer P^{-1} .
- On note à présent $Y_n = P^{-1} \cdot X_n$, montrer que $Y_{n+1} = D \cdot Y_n$ pour tout $n \in \mathbf{N}^*$.
- En déduire l'existence et la valeur des limites

$$\lim_{n \rightarrow \infty} \mathbf{P}(X_n = A), \quad \lim_{n \rightarrow \infty} \mathbf{P}(X_n = B), \quad \lim_{n \rightarrow \infty} \mathbf{P}(X_n = C).$$

Est-ce conforme à l'intuition?

Solution (exercice ALEA.13.20)

- Puisque la première fois il choisit un tunnel au hasard X_1 suit une loi uniforme sur

l'ensemble $\{A, B, C\}$ donc $Z_1 = \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

- Soit $n \in \mathbf{N}$. Alors puisque $(X_n = A, X_n = B, X_n = C)$ est un système complet d'évènements, on déduit par la formule des probabilités totales :

$$\begin{aligned} \mathbf{P}(X_{n+1} = A) &= \mathbf{P}(X_{n+1} = A | X_n = A) \mathbf{P}(X_n = A) + \mathbf{P}(X_{n+1} = A | X_n = B) \mathbf{P}(X_n = B) \\ &\quad + \mathbf{P}(X_{n+1} = A | X_n = C) \mathbf{P}(X_n = C) \\ &= 0 \cdot \mathbf{P}(X_n = A) + \frac{1}{2} \cdot \mathbf{P}(X_n = B) + 0 \cdot \mathbf{P}(X_n = C), \end{aligned}$$

$$\begin{aligned} \mathbf{P}(X_{n+1} = B) &= \mathbf{P}(X_{n+1} = B | X_n = A) \mathbf{P}(X_n = A) + \mathbf{P}(X_{n+1} = B | X_n = B) \mathbf{P}(X_n = B) \\ &\quad + \mathbf{P}(X_{n+1} = B | X_n = C) \mathbf{P}(X_n = C) \\ &= \frac{1}{2} \cdot \mathbf{P}(X_n = A) + 0 \cdot \mathbf{P}(X_n = B) + 0 \cdot \mathbf{P}(X_n = C), \end{aligned}$$

$$\begin{aligned} \mathbf{P}(X_{n+1} = C) &= \mathbf{P}(X_{n+1} = C | X_n = A) \mathbf{P}(X_n = A) + \mathbf{P}(X_{n+1} = C | X_n = B) \mathbf{P}(X_n = B) \\ &\quad + \mathbf{P}(X_{n+1} = C | X_n = C) \mathbf{P}(X_n = C) \\ &= \frac{1}{2} \cdot \mathbf{P}(X_n = A) + \frac{1}{2} \cdot \mathbf{P}(X_n = B) + 1 \cdot \mathbf{P}(X_n = C). \end{aligned}$$

Donc :

$$\forall n \in \mathbf{N}^*, \quad Z_{n+1} = MZ_n, \quad M = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix}.$$

- Soit $\lambda \in \mathbf{C}$. Alors

$$\begin{aligned} M - \lambda I_3 &= \begin{pmatrix} -\lambda & \frac{1}{2} & 0 \\ \frac{1}{2} & -\lambda & 0 \\ \frac{1}{2} & \frac{1}{2} & 1 - \lambda \end{pmatrix} \\ &\stackrel{\sim}{\mathbf{L}} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 1 - \lambda \\ \frac{1}{2} & -\lambda & 0 \\ -\lambda & \frac{1}{2} & 0 \end{pmatrix} && \begin{matrix} L_1 \leftrightarrow L_3 \\ \\ \\ \end{matrix} \\ &\stackrel{\sim}{\mathbf{L}} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 1 - \lambda \\ 0 & -(\lambda + \frac{1}{2}) & \lambda - 1 \\ 0 & \frac{1}{2}(\lambda + \frac{1}{2}) & \lambda(1 - \lambda) \end{pmatrix} && \begin{matrix} L_2 \leftarrow L_2 - L_1 \\ L_3 \leftarrow \frac{1}{2}L_3 + \lambda L_1 \\ \\ \end{matrix} \\ &\stackrel{\sim}{\mathbf{L}} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 1 - \lambda \\ 0 & -(\lambda + \frac{1}{2}) & \lambda - 1 \\ 0 & 0 & P(\lambda) \end{pmatrix}, && \begin{matrix} \\ \\ L_3 \leftarrow L_3 + \frac{1}{2}L_2 \end{matrix} \end{aligned}$$

avec $P(\lambda) = \lambda(1 - \lambda) + \frac{1}{2}(\lambda - 1) = (1 - \lambda) \left(\lambda - \frac{1}{2} \right)$. Donc :

$$\text{Spec}(M) = \left\{ -\frac{1}{2}, \frac{1}{2}, 1 \right\}. \text{ Et,}$$

- par ailleurs,
- $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_1(M)$ si et seulement si :

$$\frac{x}{2} + \frac{y}{2} = 0, \quad y = 0 \iff x = y = 0.$$

$$\text{Donc } E_1(M) = \text{Vect} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

► $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_{1/2}(M)$ si et seulement si :

$$x + y + z = 0, -y - \frac{z}{2} = 0 \iff z = -2y, x = y.$$

$$\text{Donc } E_{1/2}(M) = \text{Vect} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}.$$

► $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_{-1/2}(M)$ si et seulement si :

$$x + y + 3z = 0, z = 0 \iff z = 0, x = -y, z = 0.$$

$$\text{Donc } E_{-1/2}(M) = \text{Vect} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

Donc

$$M = P \cdot D \cdot P^{-1}, \quad P = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} -1/2 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

4. Soit $n \in \mathbf{N}^*$, on note $Y_n = P^{-1} \cdot X_n$. Alors

$$Y_{n+1} = P^{-1} \cdot P \cdot D \cdot P^{-1} \cdot X_n = D(P^{-1}X_n) = D \cdot Y_n.$$

$$\text{Donc } Y_{n+1} = D \cdot Y_n.$$

5. Notons $Y_n = \begin{pmatrix} a_n \\ b_n \\ c_n \end{pmatrix}$ pour tout $n \in \mathbf{N}^*$, alors la question précédente livre que (a_n)

(resp. $(b_n), (c_n)$) est géométrique de raison $-\frac{1}{2}$ (resp. $\frac{1}{2}, 1$). Donc :

$$\lim_{n \rightarrow \infty} a_n = 0 = \lim_{n \rightarrow \infty} b_n, \quad c_n \xrightarrow{n \rightarrow \infty} c_1.$$

Dans la suite, le symbole «tend vers» pour des matrices correspondra à des limites coefficient par coefficient. On a

$$Y_n \xrightarrow{n \rightarrow \infty} \begin{pmatrix} 0 \\ 0 \\ c_1 \end{pmatrix}.$$

Donc puisque P est une matrice constante,

$$X_n = P Y_n \xrightarrow{n \rightarrow \infty} P \begin{pmatrix} 0 \\ 0 \\ c_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ c_1 \end{pmatrix}.$$

Pour calculer c_1 il nous faut maintenant calculer P^{-1} . Après calculs, on trouve

$$P^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{-1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

Donc

$$\begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix} = Y_1 = \begin{pmatrix} \frac{1}{2} & \frac{-1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

donc $c_1 = 1$. Ainsi

$$\lim_{n \rightarrow \infty} \mathbf{P}(X_n = A) = 0, \quad \lim_{n \rightarrow \infty} \mathbf{P}(X_n = B) = 0, \quad \lim_{n \rightarrow \infty} \mathbf{P}(X_n = C) = 1.$$

Est-ce conforme à l'intuition? Oui, puisqu'une fois le bon tunnel choisi (le C), on ne se trompe plus. *En langage chaîne de MARKOV, on parle de tunnel «absorbant»*

Chapitre ALEA.14.

Variables aléatoires à densité

Résumé & Plan

La définition d'intégrale généralisée du [Chapter ANA.11](#) va permettre l'étude d'une nouvelle famille de variables aléatoires que l'on dit *à densité*; elles sont réelles, et la probabilité que cette variable soit dans un certain ensemble se mesure comme l'intégrale d'une certaine fonction.

1	Généralités	2	3	Lois usuelles	24
1.1	Définitions	2	3.1	Loi uniforme sur $[a, b]$	24
1.2	Loi & fonction de répartition	4	3.2	Loi exponentielle	27
1.3	Fonctions de variables aléatoires à densité	11	3.3	La loi normale	30
1.4	Indépendance	14	3.4	Bilan des lois continues usuelles	35
2	Espérance, Variance, Moments	17	4	Exercices	38
2.1	Espérance d'une variable à densité	17	4.1	Généralités	38
2.2	Moments d'ordre supérieur	22	4.2	Études de lois	40
			4.3	Fonctions de variables aléatoires	41
			4.4	Discret vs. Continu	41

4.5 Autour des min et max 45

L'origine de la loi normale remonte aux travaux de Jacques BERNOULLI sur son théorème d'or, appelé aujourd'hui loi des grands nombres, publié dans son oeuvre Ars Conjectandi en 1713. Il y calcule des probabilités liées à des paris sur des jeux de pile ou face à 2, notamment le calcul de la probabilité que la moyenne du nombre de pile soit proche de la moyenne théorique 1/2, calcul dans lequel apparaît le calcul des factorielles.

— Le saviez-vous ?

Cadre
 Dans tout le chapitre, et même lorsque cela n'est pas précisé, on se fixe un espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$.

1. GÉNÉRALITÉS

1.1. Définitions

Définition ALEA.14.1 | Densité de probabilité

Soit $f : \mathbf{R} \rightarrow \mathbf{R}$. On dit que f est une *densité de probabilité* si :

- ▶ f est positive ou nulle, continue sur \mathbf{R} sauf éventuellement en un nombre fini de points,
- ▶ l'intégrale $\int_{-\infty}^{+\infty} f$ est convergente et :

$$\int_{-\infty}^{+\infty} f = 1.$$

L'intégrale $\int_{-\infty}^{+\infty} f$ est donc définie au sens de la **Définition ANA.11.9** du **Chapitre ANA.11** (intégrale d'une fonction continue sauf en un nombre fini de points).

Exemple 1 — Exemples de densité

1. **(Densité de LAPLACE)** Soit f définie sur \mathbf{R} par $f(x) = \frac{1}{2}e^{-|x|}$ pour tout $x \in \mathbf{R}$. C'est une densité de probabilité. 

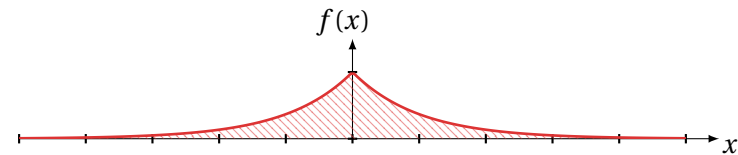




FIG. ALEA.14.1. : Graphe de la densité de LAPLACE

2. Soit g la fonction définie pour tout $x \in \mathbf{R}$ par : $g(x) = \begin{cases} 0 & \text{si } x < 1, \\ \frac{b}{2^x} & \text{si } x \geq 1, \end{cases}$ où $b \in \mathbf{R}$. Il existe b tel que g soit une densité de probabilité. 

3. (Densité normale) Soit h la fonction définie pour tout $x \in \mathbf{R}$ par : $x \in \mathbf{R}$ par :
 $h(x) = ce^{-\frac{x^2}{2}}$ où $c \in \mathbf{R}$. Il existe c tel que h soit une densité de probabilité. 

Avant de passer aux définitions qui suivent, constatons un premier lemme de convergence d'intégrale.

Lemme ALEA.14.1

Soit $f : \mathbf{R} \rightarrow \mathbf{R}$ une densité de probabilité. Alors pour tout $x \in \mathbf{R}$, $\int_{-\infty}^x f$ converge.

Preuve Rappelons que $\int_{-\infty}^{\infty} f = 1$. Cette intégrale étant généralisée en $\pm\infty$, cela signifie que pour tout $x \in \mathbf{R}$,

$$\int_{-\infty}^x f, \int_x^{\infty} f \text{ convergent.}$$

D'où le résultat.

Définition/Proposition ALEA.14.1 | Variable aléatoire à densité & continue

Soit X une variable aléatoire réelle. On dit que :

► X est une *variable aléatoire réelle à densité* s'il existe une densité de proba-

bilité notée f_X telle que, pour tout $x \in \mathbf{R}$,

$$F_X(x) = \int_{-\infty}^x f_X.$$

Une telle fonction f_X est appelée *une* densité de X .

► X est une *variable aléatoire réelle continue* si elle est à densité, et qu'il existe une densité de probabilité f_X continue. Il n'existe alors qu'une seule densité de probabilité, c'est f_X .

Σ Notation

Si X est une variable à densité, on notera f_X une densité de X .

⊗ Attention

On prendra garde à parler **d'une** densité, et **pas de la** densité, sauf dans le cas continu.

Preuve (de l'unicité dans le cas continu) Soit X une variable aléatoire continue, choisissons f_X une densité de probabilité continue. Alors :

$$\forall x \in \mathbf{R}, F_X(x) = \int_{-\infty}^x f_X.$$

Comme f_X est continue, un résultat du **Chapter ANA.11** nous donne alors :

$$\forall x \in \mathbf{R}, F'_X(x) = f_X(x).$$

D'où l'unicité : l'unique densité qui convient est la dérivée de F_X .

En revanche, pour une variable aléatoire seulement à densité, l'égalité a lieu sauf en un nombre fini de points.

Proposition ALEA.14.1 | Non-unicité d'une densité

Soit X une variable aléatoire réelle dont une densité est notée f_X et g_X une autre densité, alors $f_X = g_X$ **sauf en un nombre fini de points.**

Preuve Notons f_1, \dots, f_d et $g_1, \dots, g_{d'}$ une collection de $d + d'$, $(d, d') \in (\mathbf{N}^*)^2$ nombre réels tels que f_X (resp. g_X) est continue sur $\mathbf{R} \setminus \{f_1, \dots, f_d\}$ (resp. sur $\mathbf{R} \setminus \{g_1, \dots, g_{d'}\}$). Alors en notant $\mathcal{D} = \{f_1, \dots, f_d\} \cup \{g_1, \dots, g_{d'}\}$, nous avons, d'après le cours d'intégration :

- ▶ $x \in \mathbf{R} \setminus \mathcal{D} \longrightarrow \int_{-\infty}^x f_X$ et $x \in \mathbf{R} \setminus \mathcal{D} \longrightarrow \int_{-\infty}^x g_X$ sont dérivables,
 - ▶ et par définition d'une densité, $\forall x \in \mathbf{R}, \int_{-\infty}^x f_X = \int_{-\infty}^x g_X = \mathbf{P}(X \leq x)$ (*).
- Ainsi, pour tout $x \in \mathbf{R} \setminus \mathcal{D}$, on obtient en dérivant (*): $f_X = g_X$. En d'autres termes, $f_X = g_X$ sauf en un nombre fini de points.

On achève ce paragraphe par un résultat provisoirement admis.

Proposition ALEA.14.2

Soit X une variable à densité de densité f_X . Alors pour tout intervalle I de la forme $[a, b], [a, b[,]a, b]$ ou $]a, b[, a < b, (a, b) \in (\mathbf{R} \cup \{\pm\infty\})^2$ on a :

$$\mathbf{P}(X \in I) = \int_I f_X.$$

Preuve Faisons le cas $I =]a, b]$, les autres sont provisoirement admis.

$$\begin{aligned} \mathbf{P}(X \in I) &= \mathbf{P}(\{X \leq b\} \setminus \{X \leq a\}) \\ &= \mathbf{P}(X \leq b) - \mathbf{P}(X \leq a) \\ &= \int_{-\infty}^b f_X - \int_{-\infty}^a f_X \\ &= \int_a^b f_X = \int_I f_X. \end{aligned}$$

1.2. Loi & fonction de répartition

Rappelons que nous avons défini la loi d'une variable aléatoire réelle X comme l'application qui à tout intervalle réel I associe

$$\mathbf{P}_X(I) = \mathbf{P}(X \in I).$$

Nous avons déjà vu qu'il est aisé d'exprimer $\mathbf{P}(X \in I)$ en fonction d'une densité. Plus précisément

$$\mathbf{P}(X \in I) = \int_I f_X.$$

Cela nous invite donc à assimiler la notion de loi à celle de la donnée d'une fonction densité f_X .

Définition/Proposition ALEA.14.2 | Loi d'une variable aléatoire à densité

Soit X une variable aléatoire réelle à densité de densité f_X .

- ▶ On appelle *loi* — par abus de langage — de la variable aléatoire réelle à densité X la donnée d'une densité f_X de X .
- ▶ **(avoir même loi que)** Soit Y une autre variable aléatoire à densité de densité f_Y . On dit que X, Y ont même loi (on note $X \sim Y$) si $\mathbf{P}_X = \mathbf{P}_Y$ i.e. si pour tout intervalle I ,

$$\mathbf{P}(X \in I) = \mathbf{P}(Y \in I).$$

On a de plus :

$$X \sim Y \iff f_X = f_Y \quad \text{sauf en un nombre fini de points.}$$

- ▶ **(Déterminer la loi)** Déterminer la loi d'une variable aléatoire réelle à densité c'est calculer une densité f_X .

Preuve Montrons

$$X \sim Y \iff f_X = f_Y \text{ sauf en un nombre fini de points.}$$

\Rightarrow Supposons que $X \sim Y$. Alors pour tout intervalle I ,

$$\begin{aligned} \mathbf{P}(X \in I) &= \int_I f_X \\ &= \int_I f_Y = \mathbf{P}(Y \in I). \end{aligned} \quad \left. \vphantom{\int_I} \right\} \textit{on ne change pas la valeur d'une intégrale en modifiant les valeurs des fonctions en un nombre fini de points}$$

Donc $\mathbf{P}_X = \mathbf{P}_Y$.

\Leftarrow Supposons que pour tout intervalle I , $\mathbf{P}(X \in I) = \mathbf{P}(Y \in I)$. Alors, en choisissant $I =]-\infty, x]$ pour tout $x \in \mathbf{R}$, on obtient :

$$\int_{-\infty}^x f_X = \int_{-\infty}^x f_Y.$$

Ainsi, f_Y est aussi une densité de probabilité de X , tout comme f_X donc par un résultat précédent, on obtient :

$$f_X = f_Y \text{ sauf en un nombre fini de points.}$$

Σ Notation
Lorsque X, Y ont même loi, on note $X \sim Y$. Si Y suit une loi usuelle \mathcal{L} (une loi normale, exponentielle, etc.), on note $X \hookrightarrow \mathcal{L}$.

Comme dans le cas discret, la fonction de répartition caractérise la loi.

Proposition ALEA.14.3 | La fonction de répartition caractérise la loi
Soient X, Y deux variables aléatoires réelles à densité. Alors :

$$X \sim Y \iff F_X = F_Y.$$

Deux variables aléatoires à densité ont donc même loi si et seulement si elles ont même fonction de répartition.

Preuve



Le mot «loi» est finalement assez peu souvent utilisé pour les variables aléatoires à densité, puisque finalement calculer la loi ou une densité c'est la même chose.

LIEN ENTRE DENSITÉ ET FONCTION DE RÉPARTITION. En pratique pour montrer qu'une variable aléatoire est à densité, on utilise le théorème fondamental suivant.

Théorème ALEA.14.1 | Loi à densité & fonction de répartition

Soit X une variable aléatoire. Alors :

$$X \text{ est à densité} \iff \begin{cases} \text{(i)} & F_X \text{ est continue sur } \mathbf{R}, \\ \text{(ii)} & F_X \text{ est de classe } \mathcal{C}^1 \text{ sauf éventuellement en un nombre fini de points.} \end{cases}$$

Dans ce cas, si f est une fonction positive ou nulle telle que


$$F'_X(x) = f(x) \text{ en tout point } x \text{ où } F_X \text{ est dérivable,}$$


alors f est une densité de X . En particulier,

$$f : x \in \mathbf{R} \longrightarrow \begin{cases} F'_X(x) & \text{si } F_X \text{ dérivable en } x, \\ 0 & \text{sinon.} \end{cases}$$

est une densité de X.

Remarque 1.1 — Pour les lois continues, on remplacerait « \mathcal{C}^1 sauf en un nombre fini de points » par « \mathcal{C}^1 ». Une densité s’obtiendrait en dérivant la fonction de répartition sur \mathbf{R} tout entier.



 **Attention** Attention aux mélanges entre la définition d’une densité, et le théorème précédent.

 **Méthode** Pour montrer qu’une variable aléatoire est à densité

Deux méthodes sont possibles :

1. on revient à la définition en devinant une densité,
2. **ou**¹ on montre que la fonction de répartition F_X est continue et de classe \mathcal{C}^1 sauf éventuellement en un nombre fini points. Une densité est alors obtenue en dérivant F_X là où elle est dérivable. On met en général la valeur zéro² pour une densité là où F_X n’est pas dérivable.

Preuve

-  Admis.
-  Notons f une densité de X, et $d_1 < \dots < d_n$ les points où elle n’est pas continue.
- ▶ **(i)** La fonction F_X est continue en dehors de $\{d_1, \dots, d_n\}$, et continue à droite en les $d_i, i \in \llbracket 1, n \rrbracket$ car c’est une fonction de répartition (voir le [Chapter ALEA.12](#)). Il suffit donc de montrer la continuité à gauche de

$$F_X : x \in \mathbf{R} \longrightarrow \int_{-\infty}^x f,$$

en chaque $d_i, i \in \llbracket 1, n \rrbracket$. On calcule alors :

$$\begin{aligned} F_X(d_i) &= \int_{-\infty}^{d_i} f \\ &= \lim_{x \rightarrow d_i^-} \int_{-\infty}^x f \quad \left. \vphantom{\int_{-\infty}^x f} \right\} \text{définition de l'intégrale} \\ &= \lim_{x \rightarrow d_i^-} F_X(x). \end{aligned}$$

Ce qui prouve la continuité à gauche et donc la continuité.

- ▶ **(ii)** Soit $b \in \mathbf{R} \setminus \{d_1; \dots; d_n\}$. Puisque les d_i sont en nombre fini, on peut trouver $a \in \mathbf{R} \setminus \{d_1; \dots; d_n\}$ de sorte que $a < b$ et que f soit continue sur $[a, b]$ intervalle ne contenant aucun des d_i . Par la relation de CHASLES, on a alors :

$$\forall x \in [a, b], \quad F_X(x) = F_X(a) + \int_a^x f.$$

Or f est continue sur $[a, b]$ donc d’après le cours d’intégration, $x \in [a, b] \longmapsto \int_a^x f$ est dérivable avec $F'_X(x) = 0 + f(x)$ pour tout $x \in [a, b]$. Comme f est continue sur $\mathbf{R} \setminus \{d_1; \dots; d_n\}$, F_X l’est aussi.

On rappelle le théorème ci-dessous, et énoncé dans le [Chapter ALEA.12](#), permettant de montrer qu’une fonction est une fonction de répartition d’une certaine variable aléatoire.

Théorème ALEA.14.2 | Existence d’une variable aléatoire de fonction de répartition fixée [Rappel]

Soit F une fonction définie sur \mathbf{R} à valeurs réelles telle que :

1. F est croissante sur \mathbf{R} ,
2. $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$,
3. F est continue à droite en tout point,

alors il existe un espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$ et une variable aléatoire X définie sur cet espace tels que $F_X = F$.


Les deux résultats s’utilisent parfois conjointement, en montrant d’abord l’existence de X (et d’un $(\Omega, \mathcal{F}, \mathbf{P})$ sur lequel elle est définie) puis en montrant qu’elle est à den-

¹C’est cette méthode que nous utiliserons le plus souvent
²valeur arbitraire

sité. Voyons cela sur un exemple.

Exemple 2 — Soit F la fonction définie sur \mathbf{R} par

$$\forall x \in \mathbf{R}, \quad F(x) = \begin{cases} \frac{1-\sqrt{-x}}{2} & \text{si } x \in [-1, 0[, \\ \frac{1+\sqrt{x}}{2} & \text{si } x \in [0, 1[, \\ 1 & \text{si } x \geq 1, \\ 0 & \text{sinon.} \end{cases}$$

1. Justifier que F est la fonction de répartition d'une variable aléatoire à densité X , dont on donnera une densité f . 

2. Tracer les courbes respectives de F et f . 

DIFFÉRENCES MAJEURES AVEC LES VARIABLES ALÉATOIRES DISCRÈTES & CONSÉQUENCES. Rappelons que dans le [Chapter ALEA.12](#) nous avons défini la notion d'atome d'une variable aléatoire réelle X comme des points $x \in \mathbf{R}$ où $\mathbf{P}(X = x) \neq 0$. Nous avons vu que si X est discrète, alors tous les $x \in X(\Omega)$ sont des atomes. En revanche, pour les lois continues ou à densité il n'y en a aucun, comme nous allons le voir dans le corollaire qui suit.

Corollaire ALEA.14.1 | Une variable aléatoire à densité n'a pas d'atome

Soit X une variable à densité. Pour tout réel x , on a :

$$\mathbf{P}(X = x) = 0.$$

Preuve Soit $x \in \mathbf{R}$. Alors pour tout entier $n \in \mathbf{N}^*$, on a

$$0 \leq \mathbf{P}(X = x) \leq \mathbf{P}\left(X \in \left]x - \frac{1}{n}, x\right]\right) = F_X(x) - F_X\left(x - \frac{1}{n}\right) \quad (\star).$$

Donc puisque F_X est continue en x , elle est en particulier continue à gauche en x , donc


$$\lim_{n \rightarrow \infty} F_X\left(x - \frac{1}{n}\right) = F_X(x).$$

On conclut ensuite par théorème d'encadrement dans (\star) , d'où l'on tire :

$$\lim_{n \rightarrow \infty} \mathbf{P}(X = x) = \mathbf{P}(X = x) = 0.$$

Corollaire ALEA.14.2 | Une variable aléatoire à densité n'est pas à support dénombrable

Soit X une variable à densité. Alors $X(\Omega)$ n'est pas au plus dénombrable, *i.e.* il n'est ni fini, ni dénombrable.

Preuve Supposons par l'absurde que $X(\Omega)$ est au plus dénombrable. 

2. 

Puisque X est sans atome, on peut établir facilement le corollaire qui suit et qui nous dit la chose suivante : pour obtenir la probabilité qu'une variable aléatoire à densité soit dans un certain intervalle, on intègre la densité sur cet intervalle. De plus, l'ouverture ou non des bornes n'a aucune importance.

3. 

Corollaire ALEA.14.3 | Bornes des intervalles & loi à densité

Soit X une variable à densité de fonction de répartition F_X et on note f_X une densité. Pour tous réels a et b tels que $a < b$, on a :

$$1. F_X(b) = \mathbf{P}(X \leq b) = \mathbf{P}(X < b) = \int_{-\infty}^b f_X,$$

$$2. 1 - F_X(a) = \mathbf{P}(X \geq a) = \mathbf{P}(X > a) = \int_a^{+\infty} f_X,$$

$$3. \mathbf{P}(a < X < b) = \mathbf{P}(a \leq X \leq b) = \mathbf{P}(a < X \leq b) = \mathbf{P}(a \leq X < b) = \int_a^b f_X.$$

VALEURS PRISES PAR UNE VARIABLE ALÉATOIRE À DENSITÉ. On se pose dans ce paragraphe la question suivante : que dire de l'univers-image ? Comment obtenir des informations sur les valeurs prises par X ? Autant les variables aléatoires discrètes sont définies à l'aide de $X(\Omega)$, pour les variables aléatoires à densité, il convient de réfléchir un peu. Nous allons pouvoir déterminer des intervalles réels I tels que

$$\mathbf{P}(X \in I) = 1,$$

mais pas directement $X(\Omega)$. Autrement dit, X prendra ses valeurs dans I avec probabilité 1.

- ▶ **(À l'aide de l'expression de X)** Si par exemple $X = h(\tilde{X})$ avec \tilde{X} une variable aléatoire et $h : \mathbf{R} \rightarrow \mathbf{R}$ une fonction, alors $X(\Omega) = h(\tilde{X}(\Omega)) \subset h(\mathbf{R})$. Par exemple, si $X = \tilde{X}^2$ avec on a clairement $X(\Omega) \subset \mathbf{R}^+$.
- ▶ **(À l'aide d'une densité f_X)** Voyons quelques exemples.
 1. On considère $f : x \in \mathbf{R} \mapsto \begin{cases} e^{-x} & \text{si } x \geq 0, \\ 0 & \text{sinon.} \end{cases}$ et X une variable aléatoire de densité f . Alors $\mathbf{P}(X \in \mathbf{R}^+) = 1$. Ce sera, plus tard, une densité usuelle, celle de la loi

Attention


En règle générale,

$$\mathbf{P}(a < X < b) = \mathbf{P}(a \leq X \leq b) = \mathbf{P}(a < X \leq b) = \mathbf{P}(a \leq X < b)$$


est faux pour des variables aléatoires discrètes.

Preuve

1. 

exponentielle de paramètre 1. On admet pour le moment que c'est bien une densité. 

2. On considère $f : x \in \mathbf{R} \mapsto \frac{1}{2}e^{-|x|}$ la densité de LAPLACE déjà étudiée et X une variable aléatoire de densité f . Alors $\mathbf{P}(X \in \mathbf{R}) = 1$ par définition d'une densité, et on ne peut pas faire mieux.

3. On considère $f : x \in \mathbf{R} \mapsto \begin{cases} 1 & \text{si } x \in [0, 1], \\ 0 & \text{sinon.} \end{cases}$ et X une variable aléatoire de densité f . Alors $\mathbf{P}(X \in [0, 1]) = 1$. Que vaut $\mathbf{P}(X \in [-1, 1])$? $\mathbf{P}(X \in [-10^{23}, 1])$?
Ce sera, plus tard, une densité usuelle, celle de la loi uniforme sur $[0, 1]$. 

 **Attention**

Du dernier exemple, on constate qu'il n'y a pas unicité de I tel que

$$\mathbf{P}(X \in I) = 1.$$

Tous les $J \supset I$ vérifient aussi $\mathbf{P}(X \in J) = 1$.

En résumé, les intervalles où f_X est non nulle donnent un intervalle I tel que $\mathbf{P}(X \in I) = 1$.

▶ **(À l'aide de la fonction de répartition)** Notons $a < b$ deux réels, avec éventuellement $a = -\infty$ et $b = \infty$, de sorte que

$$F_X(x) = \begin{cases} 0 & \text{si } x \leq (\text{ou } <) a, \\ 1 & \text{si } x \geq (\text{ou } >) b. \end{cases}$$

Alors en dérivant là où F_X est dérivable, on obtient une densité qui est nulle en dehors de $[a, b]$. Donc $\mathbf{P}(X \in [a, b]) = 1$.

En résumé, les intervalles où F_X est non nulle et différente de 1 donnent un intervalle I tel que $\mathbf{P}(X \in I) = 1$.

 **Méthode** Obtenir des renseignements sur $X(\Omega)$ lorsque X est à densité

Au choix :

- ▶ on peut déduire des informations si X est une fonction de variable aléatoire (par exemple un carré, une valeur absolue, une racine carrée, etc.)
- ▶ on regarde une densité puis on analyse les points où elle est non nulle.
- ▶ On regarde là où la fonction de répartition est différente de zéro et un.

EXISTENCE D'UNE VARIABLE À DENSITÉ DE LOI FIXÉE. Un certain nombre d'énoncés probabilistes commencent par la phrase suivante :

«soit X une variable aléatoire à densité de densité f » avec f une densité de probabilité.

Ces énoncés supposent l'existence de X . Mais cela ne définit pas X en tant qu'application, c'est-à-dire $X(\omega)$ pour tout $\omega \in \Omega$, mais existe-t-elle vraiment? On aimerait donc au moins savoir si une telle variable aléatoire existe sur un certain $(\Omega, \mathcal{F}, \mathbf{P})$ à trouver : la réponse est oui, comme dans le cas discret, dès que f est bien une densité de probabilité.

Théorème ALEA.14.3 | Existence d'un espace probabilisé associé à une densité de probabilité

Soit f une densité de probabilités. Alors il existe un espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$ et une variable aléatoire X de densité f définie sur $(\Omega, \mathcal{F}, \mathbf{P})$.


Dans la plupart des situations que nous étudierons en pratique, le travail commencera par la donnée d'une ou plusieurs variables aléatoires de lois prescrites que nous étudierons sans jamais préciser là encore l'espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$ sur lequel elles sont définies.


Celui-ci est voué à rester caché comme dans le cas discret.

Preuve Considérons

$$F \begin{cases} \mathbf{R} & \longrightarrow [0, 1], \\ x & \longrightarrow \int_{-\infty}^x f. \end{cases}$$

Montrons que F est la fonction de répartition d'une variable aléatoire à densité en appliquant le **Théorème ALEA.12.3** du **Chapter ALEA.12**.

▶ F est croissante sur \mathbf{R} . 

▶ $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$. 

▶ F est continue à droite en tout point et même continue comme nous l'avons déjà vérifié au début du chapitre.

Donc il existe un espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$ et une variable aléatoire X définie sur cet espace tels que $F_X = F$. Par construction, X est alors à densité puisque pour tout


$x \in \mathbf{R}$,^a


$$F_X(x) = \int_{-\infty}^x f_X.$$

Remarque 1.2 — L'hypothèse de densité remplace donc l'hypothèse « $\sum_{x \in X(\Omega)} p_x = 1$ » dans le cas discret.

Exemple 3 — On considère la fonction f définie par :

$$f(t) = \begin{cases} \frac{1}{2\sqrt{t-1}} & \text{si } t \in]1, 2], \\ 0 & \text{sinon.} \end{cases}$$

1. Il existe une variable aléatoire réelle X de densité f . 

2. Déterminer de plus la fonction de répartition de X . 

^aOn vérifie ici la définition d'une variable aléatoire à densité

1.3. Fonctions de variables aléatoires à densité


Traisons quelques exemples simples d'images de lois à densité. De manière générale, toute image d'une variable aléatoire à densité par une fonction n'est pas à densité. En effet, l'image par la fonction nulle d'une variable aléatoire à densité est la variable aléatoire nulle, qui est clairement discrète et n'est donc pas à densité.


Attention Les calculs qui suivent sont extrêmement classiques en pratique, à savoir refaire absolument.


Méthode Montrer que l'image d'une variable aléatoire réelle à densité est à densité
Soit X une variable à densité et $Y = g(X)$ avec g une fonction au moins définie sur $X(\Omega)$ (valeur absolue, logarithme, etc.).


1. Deviner un ensemble contenant $Y(\Omega)$ au moyen de l'image de la fonction g .
2. Calculer la fonction de répartition de la variable aléatoire Y .
3. Calculer la fonction de répartition de Y i.e. $\mathbf{P}(Y \leq y)$ pour tout $y \in \mathbf{R}$ et vérifier le **Théorème ALEA.14.1**. Pour effectuer ledit calcul, faire des disjonctions de cas en y en utilisant l'ensemble trouvé en 1.

Exemple 4 — Valeur absolue & Carré Soit X une variable à densité de densité f . Alors :

- $Y = |X|$ est à densité, déterminer une densité. 

- $Z = X^2$ est à densité, déterminer une densité. 

Exemple 5 — Carré dans un cas particulier Soit X une variable à densité de densité $t \in \mathbf{R} \mapsto \begin{cases} \frac{1}{3} & \text{si } t \in [-1, 2], \\ 0 & \text{sinon.} \end{cases}$ Montrer que $Y = X^2$ est une variable aléatoire à densité, déterminer une densité. 

Exemple 6 – Exponentielle Soit X une variable à densité de densité f , et $Y = e^X$. Montrer que Y est une variable aléatoire à densité, déterminer une densité. 

Exemple 7 – Logarithme Soit X une variable à densité telle que $X(\Omega) \subset]0, +\infty[$, et $Y = \ln(X)$. Montrer que Y est une variable aléatoire à densité, déterminer une densité.



- ▶ (Seconde méthode : en utilisant la définition.) 

Exemple 8 – Fonction affine Soit $(a, b) \in \mathbf{R}^2$ tel que $a \neq 0$. Alors $Y = aX + b$ est une variable aléatoire réelle admettant une densité g définie par

$$\forall y \in \mathbf{R}, \quad g(y) = \frac{1}{|a|} f\left(\frac{y-b}{a}\right).$$

- ▶ (Première méthode : en utilisant le **Théorème ALEA.14.1.**) 

1.4. Indépendance

C'est la même définition que dans le cas discret, mais on remplace les égalités dans les évènements par des symboles « \leq ».

Définition ALEA.14.2 | Indépendance de variables aléatoires à densité

Soient X_1, \dots, X_n une collection de n variables aléatoires à densité.

- ▶ **(Indépendance)** Les variables aléatoires X_1, \dots, X_n sont dites *indépendantes* si :

$$\forall (x_1, \dots, x_n) \in \mathbf{R}^n, \quad \mathbf{P}(X_1 \leq x_1, \dots, X_n \leq x_n) = \mathbf{P}(X_1 \leq x_1) \times \dots \times \mathbf{P}(X_n \leq x_n).$$

- ▶ **(Indépendance deux à deux)** Les variables aléatoires X_1, \dots, X_n sont dites *indépendantes deux à deux* si pour tous $i \neq j$, X_i et X_j sont indépendantes. Plus généralement, si (X_n) est une famille infinie de variables aléatoires, elles sont indépendantes (*resp.* indépendantes deux à deux) si et seulement si toute sous-famille finie est indépendante (*resp.* deux à deux indépendantes).

Σ Notation

Dans ce cas on notera, comme dans le cas discret, $X_1 \perp\!\!\!\perp X_2$ pour $n = 2$, et $(X_n) \perp\!\!\!\perp$ en cas d'indépendance mutuelle.

Proposition ALEA.14.4 | Indépendance & Intervalles

Soient X_1, \dots, X_n des variables aléatoires à densité indépendantes, et $I_1 \dots I_n$ une collection de n intervalles réels. Alors :

$$\mathbf{P}(X_1 \in I_1, \dots, X_n \in I_n) = \mathbf{P}(X_1 \in I_1) \times \dots \times \mathbf{P}(X_n \in I_n).$$

Définition ALEA.14.3 | Suite i.i.d.

On dit qu'une suite (X_n) de variables aléatoires à densité est i.i.d. (on dit *indépendantes et identiquement distribuées*) si elles sont indépendantes et de même loi.

FONCTIONS DE VARIABLES ALÉATOIRES INDÉPENDANTES. Nous admettons les résultats qui suivent, qui sont les mêmes que dans le cas discret.

Théorème ALEA.14.4

Soient X_1, \dots, X_n une collection de n variables aléatoires mutuellement indépendantes. Alors pour toutes fonctions f_1, \dots, f_n où pour tout $f_i : X_i(\Omega) \rightarrow \mathbf{R}$, les variables aléatoires réelles $f_1(X_1), \dots, f_n(X_n)$ sont des variables aléatoires réelles indépendantes.

Exemple 9 — Si X_1, \dots, X_3 sont indépendantes et X_3 ne s'annule pas, alors $X_1^2, X_2^2, 1/X_3$ sont indépendantes.

Théorème ALEA.14.5 | Lemme des coalitions ou Indépendance par paquets

Soient X_1, \dots, X_n une collection de n variables aléatoires mutuellement indépendantes. Alors pour toutes fonctions $\varphi_1, \dots, \varphi_k$, et n_1, \dots, n_k , k des entiers tels que $\sum_{i=1}^k n_i = n$, où pour tout $\varphi_i : \mathbf{R}^{n_i} \rightarrow \mathbf{R}$, les variables aléatoires réelles

$$\varphi_1(X_1, \dots, X_{n_1}), \dots, \varphi_k(X_{n_1+\dots+n_{k-1}}, \dots, X_n)$$

sont des variables aléatoires réelles mutuellement indépendantes.

Exemple 10 — Si X_1, X_2, X_3, X_4, X_5 sont indépendantes alors $X_1^2 + X_2^2, X_3 X_5, X_4$ sont indépendantes.

La proposition suivante est indiquée comme hors-programme dans la mesure où la formule obtenue n'est pas à connaître par coeur. En revanche, comme dans le cas discret, la méthode mise en jeu dans la démonstration doit être maîtrisée.

Proposition ALEA.14.5 | Minimum & Maximum de variables aléatoires à densité [H.P]

Soient X_1, \dots, X_n à densité de fonctions de répartition respectives F_1, \dots, F_n et mutuellement indépendantes. Alors pour tout $x \in \mathbf{R}$,

$$\mathbf{P}(\min \{X_1, \dots, X_n\} > x) = \prod_{i=1}^n (1 - F_i(x)),$$

$$\mathbf{P}(\max \{X_1, \dots, X_n\} \leq x) = \prod_{i=1}^n F_i(x).$$

Si de plus les variables aléatoires sont i.i.d. de même loi qu'une variable aléatoire X de fonction de répartition F_X , alors :

$$\mathbf{P}(\min \{X_1, \dots, X_n\} > x) = (1 - F_X(x))^n,$$

$$\mathbf{P}(\max \{X_1, \dots, X_n\} \leq x) = F_X(x)^n.$$

Par conséquent, les variables aléatoires $\min \{X_1, \dots, X_n\}$ et $\max \{X_1, \dots, X_n\}$ sont des variables aléatoires réelles à densité.

Notez que lorsque $n = 2$, on a

$$\max(X_1, X_2) + \min(X_1, X_2) = X_1 + X_2, \quad \max(X_1, X_2) \cdot \min(X_1, X_2) = X_1 \cdot X_2.$$

La première relation permet par exemple de calculer l'espérance de l'un connaissant l'espérance de l'autre (voir plus bas pour la définition de l'espérance).

Preuve



Attention

Les égalités

$$\mathbf{P}(\min \{X_1, \dots, X_n\} \leq x) = \mathbf{P}\left(\bigcup_{i=1}^n \{X_i \leq x\}\right),$$

$$\mathbf{P}(\max \{X_1, \dots, X_n\} \geq x) = \mathbf{P}\left(\bigcup_{i=1}^n \{X_i \geq x\}\right)$$


sont vraies. Mais en revanche, les réunions précédentes ne sont **pas disjointes**.



Méthode Trouver la loi d'un max de variables aléatoires à densité indépendantes

Pour le $\max X = \max(X_1, \dots, X_n)$.

1. on calcule la fonction de répartition : $\mathbf{P}(X \leq x) = \mathbf{P}(X_1 \leq x) \dots \mathbf{P}(X_n \leq x)$ pour tout $x \in \mathbf{R}$. On invoque l'indépendance au moment adéquat.
2. On utilise ensuite l'expression de la fonction de répartition à l'aide de la densité donnée.
3. On vérifie le **Théorème ALEA.14.1** i.e. que $x \mapsto \mathbf{P}(X \leq x)$ est bien continue et


 de classe \mathcal{C}^1 sauf en un nombre fini de points. On déduit alors également une densité.

 **Méthode Trouver la loi d'un min de variables aléatoires à densité indépendantes**

Pour le min $X = \min(X_1, \dots, X_n)$.

1. on calcule la fonction d'antirépartition : $\mathbf{P}(X > x) = \mathbf{P}(X_1 > x) \dots \mathbf{P}(X_n > x)$ pour tout $x \in \mathbf{R}$. On invoque l'indépendance au moment adéquat.
2. On utilise ensuite l'expression de la fonction de répartition à l'aide de la densité donnée.
3. On vérifie le **Théorème ALEA.14.1** i.e. que $x \mapsto 1 - \mathbf{P}(X > x) = \mathbf{P}(X \leq x)$ est bien continue et de classe \mathcal{C}^1 sauf en un nombre fini de points. On déduit alors également une densité.

Exemple 11 — Cas d'uniformes Soit (X_1, \dots, X_n) des variables aléatoires mutuellement indépendantes de loi $\mathcal{U}[0, 1]$ i.e. de densité $f : x \mapsto \begin{cases} 1 & \text{si } x \in [0, 1], \\ 0 & \text{sinon.} \end{cases}$

Montrer que $S = \max(X_1, \dots, X_n)$ est une variable aléatoire à densité, déterminer une densité. 

2. ESPÉRANCE, VARIANCE, MOMENTS

2.1. Espérance d'une variable à densité

Rappelons que la notion d'espérance dans le cas discret est une moyenne (éventuellement une somme infinie) des valeurs de la variable aléatoire en question, et pondérées par sa loi. Dans le [Chapter ANA.11](#), nous avons défini la notion de moyenne pondérée pour une fonction : en particulier la *valeur moyenne de $t \mapsto t$ pondérée par f_X* une densité de probabilité³ est :

$$\int_{-\infty}^{+\infty} t f_X(t) dt.$$

Puisque f_X ne s'annule pas là où X prend ses valeurs, on est bien en train de « moyenner » les valeurs de X , pondérées par f_X . Ce sera notre définition de l'espérance.

Définition ALEA.14.4 | Espérance d'une variable aléatoire à densité

Soit X une variable aléatoire à densité, de densité f_X .

► **(Admettre une espérance)** On dit que X *admet une espérance* ou que X est *intégrable* si :

$$\int_{-\infty}^{+\infty} t f_X(t) dt \quad \text{converge absolument.}$$

► **(Valeur de l'espérance)** Si X admet une espérance, alors on appelle *espérance de X* la quantité

$$E(X) = \int_{-\infty}^{+\infty} t f_X(t) dt.$$

Une variable aléatoire d'espérance nulle est dite *centrée*.

³Qui est donc bien d'intégrale 1

Attention

Comme dans le cas discret, l'existence d'une espérance n'est pas automatique!

Attention

Il convient d'être vigilant, comme dans le cas discret, sur le vocabulaire. En effet :

1. admettre une espérance signifie que : $\int_{-\infty}^{+\infty} |t f_X(t)| dt$ converge.
2. La valeur de l'espérance est quant à elle : $E(X) = \int_{-\infty}^{+\infty} t f_X(t) dt$.

Remarque 2.1 — Pourquoi supposer de la convergence absolue? La convergence absolue n'est en réalité pas nécessaire dans le cas à densité pour que la notion d'espérance soit bien posée (mais écrite dans le programme).

Remarque 2.2 — L'espérance ne dépend pas du choix d'une densité Puisque deux densités coïncident sauf en un nombre fini de points, et que l'on ne change pas la valeur d'une intégrale en modifiant la fonction en un nombre fini de points, on justifie sans peine que la définition ne dépend pas du choix d'une densité.

Proposition ALEA.14.6 | Cas d'une densité paire


Soit X une variable aléatoire à densité admettant une espérance de densité f paire, alors :


$$E(X) = 0.$$


Preuve



Exemple 12 — Montrer que la fonction $f : x \mapsto \begin{cases} \frac{3}{x^4} & \text{si } x \geq 1, \\ 0 & \text{sinon,} \end{cases}$ est une densité d'une

variable aléatoire X , et que X admet une espérance qui vaut $\mathbf{E}(X) = \frac{3}{2}$. 

Exemple 14 — Montrer que la fonction $f : x \mapsto \begin{cases} 6x(1-x) & \text{si } x \in [0, 1], \\ 0 & \text{sinon,} \end{cases}$ est une densité d'une variable aléatoire X , et calculer son espérance le cas échéant. 

Exemple 13 — *Loi de CAUCHY* La fonction $f : x \in \mathbf{R} \mapsto \frac{1}{\pi(1+x^2)}$ est une densité d'une variable aléatoire X qui n'admet pas d'espérance. 

Exemple 15 — *Loi de LAPLACE* Soit X une variable aléatoire de densité $f : x \mapsto \frac{1}{2}e^{-|x|}$. La variable aléatoire X est alors centrée. 

PROPRIÉTÉS DE L'ESPÉRANCE.**Théorème ALEA.14.6 |**

Soient X et Y deux variables aléatoires à densité admettant une espérance. Alors :

- 1. (Linéarité de l'espérance)** Soient $\lambda, \mu \in \mathbf{R}$, $\lambda X + \mu Y$ admet une espérance et :

$$\mathbf{E}(\lambda X + \mu Y) = \lambda \mathbf{E}(X) + \mu \mathbf{E}(Y).$$

- 2. (Positivité de l'espérance)** $X \geq 0 \implies \mathbf{E}(X) \geq 0$, et :

$$X \geq 0, \quad \mathbf{E}(X) = 0 \iff X = 0 \text{ p.s. (i.e. } \mathbf{P}(X = 0) = 1).$$

Le résultat subsiste si on a seulement $\mathbf{P}(X \geq 0) = 1$ en hypothèse, *i.e.*

$$\mathbf{P}(X \geq 0) = 1 \implies \mathbf{E}(X) \geq 0,$$

$$\mathbf{P}(X \geq 0) = 1, \quad \mathbf{E}(X) = 0 \iff X = 0 \text{ p.s. (i.e. } \mathbf{P}(X = 0) = 1).$$

3. (Croissance de l'espérance)

$$X \leq Y \implies \mathbf{E}(X) \leq \mathbf{E}(Y).$$

Le résultat subsiste si on a seulement $\mathbf{P}(X \leq Y) = 1$ en hypothèse, *i.e.*

$$\mathbf{P}(X \leq Y) = 1 \implies \mathbf{E}(X) \leq \mathbf{E}(Y).$$

Preuve

- 1.** Nous admettons la linéarité de l'espérance.

- 2.** Supposons que $\mathbf{P}(X \geq 0) = 1$. Alors $F_X(0) = 0 = \int_{-\infty}^0 f_X$. Or, la fonction f_X est positive et continue sauf en un nombre fini de points, donc

$$f_X|_{\mathbf{R}^-} = 0 \quad \text{sauf en un nombre fini de points.}$$

Ainsi,

$$\mathbf{E}(X) = \int_{-\infty}^0 t f_X(t) dt + \int_0^{\infty} t f_X(t) dt = 0 + \int_0^{\infty} \underbrace{t f_X(t)}_{\geq 0} dt \geq 0.$$

Nous admettons l'équivalence.

- 3.** Appliquer **2** avec $X \leftarrow X - Y$ et utiliser la linéarité de l'espérance.

FORMULE DE TRANSFERT POUR UNE VARIABLE ALÉATOIRE À DENSITÉ. Le théorème suivant permet de calculer l'espérance de $g(X)$ lorsque X est une variable à densité et g une fonction continue. Là encore, le point clef est que nous n'avons pas besoin de connaître la loi de $g(X)$ (ou une densité de $g(X)$) pour calculer son espérance!

Théorème ALEA.14.7 | Transfert pour les variables aléatoires à densité

Soit X une variable aléatoire réelle de densité f_X telle que $X(\Omega) \subset I$ où I est un intervalle réel (éventuellement non borné), et $g : I \rightarrow \mathbf{R}$ continue sauf éventuellement en un nombre fini de points de sorte que $g(X)$ est discrète ou à densité.

Alors :

$Y = g(X)$ admet une espérance

$$\iff \int_I g(t) f_X(t) dt \text{ est absolument convergente.}$$

Dans ce cas, nous avons :

$$\mathbf{E}(g(X)) = \int_I g(t) f_X(t) dt.$$

Preuve (Point clef — Dans le cas étudié ci-dessous, la preuve est un changement de variable)

Montrons dans le cas où g est de classe \mathcal{C}^1 avec g' strictement positive sur $I =]a, b[$, avec $a < b$ réels ou égaux à $-\infty$ ou $+\infty$. Par le théorème de la bijection, la réciproque g^{-1} existe bien.

D'après la formule de changement de variable du [Chapter ANA.11](#), les intégrales

$$\int_a^b g(t) f_X(t) dt \text{ et}$$

$$\int_{g(a)}^{g(b)} u \times \frac{f_X(g^{-1}(u))}{g'(g^{-1}(u))} du = \int_{g(a)}^{g(b)} u \times (F_X \circ g^{-1})'(u) du$$

sont de même nature en faisant formellement le changement de variable « $u = g(t) \iff g^{-1}(u) = t$ ». Or, toute densité de $Y = g(X)$ est nulle en dehors de $]g(a), g(b)[$ puisque $Y(\Omega) \subset g(I) \subset]g(a), g(b)[$ et sur cet intervalle,

$$F_Y(y) = P(g(X) \leq y) = P(X \leq g^{-1}(y)) = (F_X \circ g^{-1})(y).$$

Donc, en cas de convergence,


$$\int_{g(a)}^{g(b)} u \times (F_X \circ g^{-1})'(u) du = \int_{-\infty}^{+\infty} u \times F_Y'(u) du = \int_{-\infty}^{+\infty} u f_Y(u) du.$$

Finalement, on a montré que les intégrales $\int_a^b g(t) f_X(t) dt$ et $\int_{-\infty}^{+\infty} u f_Y(u) du$ sont de même nature et sont égales en cas de convergence. Or, cette dernière intégrale converge absolument si et seulement si Y admet une espérance. Et en cas de convergence elles sont égales, le théorème de transfert est donc démontré dans ce cas.


⊗ Attention

La variable aléatoire $g(X)$ n'est pas forcément une variable à densité comme nous l'avons déjà vu. Mais on sait néanmoins dire si elle admet une espérance et la calculer le cas échéant.


Exemple 16 —

- ▶ Soit $b \in \mathbf{R}$, retrouver $\mathbf{E}(b)$ à l'aide du théorème du transfert. Comment pourrait-on aussi la calculer? 

- ▶ Soit X une variable aléatoire de densité $f : x \mapsto \begin{cases} 1 & \text{si } x \in [0, 1], \\ 0 & \text{sinon.} \end{cases}$

Montrer que \sqrt{X} admet une espérance, et la calculer. 

- ▶ Soit X une variable aléatoire de densité $f : x \mapsto \begin{cases} e^{-x} & \text{si } x \geq 0, \\ 0 & \text{sinon.} \end{cases}$

Montrer que X^2 admet une espérance, et la calculer. 

- ▶ On reconsidère l'Exemple 13. Montrer que $\sin(X)$ admet une espérance, et la calculer.

Corollaire ALEA.14.4 | Inégalité triangulaire pour l'espérance

Soit X une variable aléatoire réelle à densité. Alors :

X admet une espérance $\iff |X|$ admet une espérance.

Et dans ce cas :

$$|E(X)| \leq E(|X|).$$

Preuve (Point clef — *Théorème du transfert & Inégalité triangulaire pour les intégrales généralisées*)



Corollaire ALEA.14.5 | Transfert pour les fonctions affines.


Soit X une variable aléatoire à densité et $a, b \in \mathbf{R}$. Alors :

1. X possède une espérance $\implies aX + b$ possède une espérance.
2. Supposons que $a \neq 0$. Alors : $aX + b$ possède une espérance $\implies X$ possède une espérance.

De plus, si X et $aX + b$ possèdent une espérance, nous avons :

$$E(aX + b) = aE(X) + b.$$

Preuve

1. Le résultat est évident pour $a = 0$: en effet, $0X + b = b$ dans ce cas et la variable aléatoire discrète b admet bien une espérance qui vaut $b = 0E(X) + b$. Considérons le cas $a \neq 0$. 

2. 

Remarque 2.3 — Notons que, comme déjà constaté dans un exemple, $aX + b$ est une variable aléatoire réelle à densité à densité si $a \neq 0$. L'espérance précédente est donc donnée sous forme d'une intégrale dans ce cas, et si $a = 0$, nous obtenons l'espérance d'une constante qui est une variable aléatoire réelle discrète (donc définie au sens du [Chapter ALEA.13](#)).

Corollaire ALEA.14.6 | Centrage

Soit X une variable aléatoire à densité possédant une espérance, alors

$X - E(X)$ est une variable aléatoire centrée.

Preuve Appliquer le résultat précédent avec $a = 1$ et $b = -\mathbf{E}(X)$. Alors $X - \mathbf{E}(X)$ possède donc une espérance, et

$$\mathbf{E}(X - \mathbf{E}(X)) = \mathbf{E}(X) - \mathbf{E}(X) = 0.$$

2.2. Moments d'ordre supérieur

À l'aide du théorème de transfert, nous pouvons donc affirmer les points suivants.

Définition/Proposition ALEA.14.3 | Variance, écart-type, moments, version continue

Soit X une variable aléatoire à densité f_X .

- ▶ **(Moments d'ordre k)** X admet un moment d'ordre $k \in \mathbf{N}$ si et seulement si l'équivalence suivante est réalisée :

$$\mathbf{E}(|X|^k) < \infty \iff \int_{-\infty}^{\infty} |t|^k f_X(t) dt \text{ converge.}$$

On appelle alors *moment d'ordre k* : $\mathbf{E}(X^k) = \int_{-\infty}^{\infty} t^k f_X(t) dt$.

- ▶ **(Moment d'ordre 2 et variance)** On dit que X admet un moment d'ordre deux si et seulement si l'équivalence suivante est réalisée :

$$\mathbf{E}(|X|^2) < \infty \iff \int_{-\infty}^{\infty} |t|^2 f_X(t) dt \text{ converge.}$$


On appelle alors *moment d'ordre 2* : $\mathbf{E}(X^2) = \int_{-\infty}^{\infty} |t|^2 f_X(t) dt$. Si X admet un moment d'ordre deux, alors elle admet un moment d'ordre un (*i.e.* une espérance), et on appelle *variance de X* la quantité notée $\mathbf{Var}(X)$ et définie par : $\mathbf{Var}(X) = \mathbf{E}((X - \mathbf{E}(X))^2)$. La variable aléatoire X possède une variance si et seulement si :

$$\int_{-\infty}^{\infty} (t - \mathbf{E}(X))^2 f_X(t) dt \text{ converge.}$$

Dans ce cas,

$$\mathbf{Var}(X) = \int_{-\infty}^{\infty} (t - \mathbf{E}(X))^2 f_X(t) dt.$$

On appelle *écart-type de X* , la quantité notée σ_X , et définie par $\sigma_X := \sqrt{\mathbf{Var}(X)}$. Une variable aléatoire de variance un est dite *réduite*.

Preuve Montrons que : si X admet un moment d'ordre deux, alors X admet un moment d'ordre un. Le reste provient de simples applications du théorème de transfert. 

Remarque 2.4 —

1. On pourra retenir également que : si X n'a pas d'espérance, alors elle n'a pas de variance.
2. Le moment d'ordre 1 correspond donc à l'espérance.
3. La variance d'une variable aléatoire mesure l'écart quadratique moyen entre X et sa valeur moyenne $\mathbf{E}(X)$.



Méthode Étudier l'existence d'une variance dans le cas à densité

On étudie l'existence d'un moment d'ordre deux, *i.e.* la convergence de

$$\int_{-\infty}^{\infty} |t|^2 f_X(t) dt.$$

Proposition ALEA.14.7 | Cas d'une variable aléatoire bornée

Soit X une variable aléatoire à densité f_X .

- ▶ Si X est bornée, alors X admet une variance, et donc une espérance.
- ▶ Le résultat est encore vrai si X est presque-sûrement bornée *i.e.* s'il existe $M \in \mathbf{R}^+$ tel que $\mathbf{P}(|X| \leq M) = 1$.

Preuve On montre que

$$\int_{-\infty}^{\infty} |t|^2 f_X(t) dt \text{ converge (et donc aussi absolument car l'intégrande est positive).}$$



Enfin toutes les propriétés en vigueur dans le **Chapter ALEA.12** restent donc valables dans le cas à densité.

Proposition ALEA.14.8 | Propriétés de la variance/covariance

Soient X, Y deux variables aléatoires à densité admettant une variance, et $\lambda, \mu \in \mathbf{R}$.

1. (**Variance nulle**) $\text{Var}(X) = 0 \iff X = \mathbf{E}(X)$ p.s.
2. (**Variance d'une expression affine**) $\text{Var}(\lambda X + \mu) = \lambda^2 \text{Var}(X)$.
3. (**Formule de KÖNIG-HUYGENS**) $\text{Var}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2$.

Preuve Identiques au cas discret.

OPÉRATION DE CENTRAGE/RÉDUCTION. La proposition ci-dessous paraît anecdotique mais elle sera d'un intérêt majeur plus tard dans l'année.

Définition/Proposition ALEA.14.4

Soit X une variable aléatoire à densité ayant une variance, alors

$$X^* \stackrel{\text{(déf.)}}{=} \frac{X - \mathbf{E}(X)}{\sigma_X}$$

est une variable aléatoire réelle centrée réduite. On l'appelle *la centrée/réduite de X*. Elle est de plus à densité.

Preuve

- ▶ par linéarité de l'espérance,

$$\mathbf{E}(X^*) = \frac{\mathbf{E}(X) - \mathbf{E}(\mathbf{E}(X))}{\sigma_X} = \frac{\mathbf{E}(X) - \mathbf{E}(X)}{\sigma_X} = 0.$$

- ▶ D'après l'expression de la variance d'une expression affine, on a :

$$\text{Var}(X^*) = \frac{\text{Var}(X - \mathbf{E}(X))}{\sigma_X^2} = \frac{\text{Var}(X)}{\sigma_X^2} = \frac{\sigma_X^2}{\sigma_X^2} = 1.$$

CAS DE VARIABLES ALÉATOIRES INDÉPENDANTES. On présente sans démonstration le résultat suivant.

Proposition ALEA.14.9

Si X_1, \dots, X_n sont n variables aléatoires à densité mutuellement indépendantes. Alors :

1. (**Espérance d'un produit**) si les X_i admettent une espérance,

$$\mathbf{E}(X_1 \dots X_n) = \mathbf{E}(X_1) \dots \mathbf{E}(X_n).$$

2. (**Variance d'une somme**) si les X_i admettent une variance, alors

$$\mathbf{Var}(X_1 + \dots + X_n) = \mathbf{Var}(X_1) + \dots + \mathbf{Var}(X_n).$$

Preuve Admis.^a

3. LOIS USUELLES

Commençons par définir la loi uniforme sur des ensembles non discrets. Précisons un peu de vocabulaire pour certaines commandes Python :

- ▶ cdf signifie « cumulative density function », cet acronyme sera donc réservé aux fonctions de répartition,
- ▶ pdf signifie « probability density function », cet acronyme sera donc réservé aux densités.

Pour la simulation, le principe est similaire au cas discret déjà vu : la plupart des scripts dépendront de la simulation d'un réel aléatoire entre 0 et 1. On rappelle un fait déjà constaté dans le [Chapter ALEA.13](#) : toute loi de fonction de répartition bijective peut être simulée en cherchant F_X^{-1} , car $F_X^{-1}(U)$ a même loi que X si $U \hookrightarrow \mathcal{U}[0, 1]$. Ce sera parfois le cas dans les lois usuelles qui vont suivre.

Comme pour les lois discrètes usuelles, on justifiera à chaque fois l'existence d'une telle variable aléatoire en montrant qu'une certaine fonction est une densité.

^aLes vecteurs aléatoires à densité étant hors programme, nous n'aurons pas les outils pour prouver ces propriétés contrairement au cas discret.

3.1. Loi uniforme sur $[a, b]$

Définition ALEA.14.5 | Loi uniforme continue

Soient a et b deux réels tels que $a < b$. On dit qu'une variable aléatoire X suit la loi uniforme sur $[a, b]$, $]a, b[$ ou $]a, b[$ si elle admet pour densité

$$f_{a,b} : x \mapsto \begin{cases} \frac{1}{b-a} & \text{si } x \in [a, b], \\ 0 & \text{sinon.} \end{cases}$$

Remarque 3.1 — Modélisation Toute expérience aléatoire dont les issues sont des réels d'un intervalle $[a, b]$ et apparaissant de manière équiprobable.

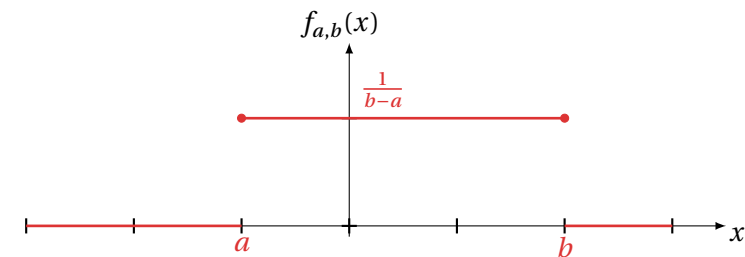




FIG. ALEA.14.2. : Graphe de la densité de l'uniforme

Remarque 3.2 — Valeurs prises En particulier,

$$\mathbf{P}(X \in [a, b]) = \mathbf{P}(X \in]a, b]) = \mathbf{P}(X \in [a, b[) = \mathbf{P}(X \in]a, b[) = 1.$$

Preuve Il existe un espace probabilisé et une variable aléatoire vérifiant ces conditions. En effet, 

Exemple 17 — Soient $X \hookrightarrow \mathcal{U}[-1, 3], Y \hookrightarrow \mathcal{U}[0, 2]$. Donner l'expression de leur densité et les représenter graphiquement. 

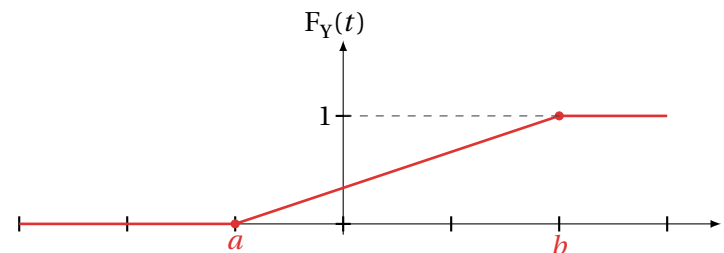


FIG. ALEA.14.3. : Graphe de la fonction de répartition de l'uniforme sur $[a, b]$

Proposition ALEA.14.10 | Espérance, variance, propriétés

Soient a et b deux réels tels que $a < b$ et X, Y deux variables aléatoires réelles.

1. Supposons que $X \hookrightarrow \mathcal{U}[0, 1]$ et $Y \hookrightarrow \mathcal{U}[a, b]$. Alors pour tout $t \in \mathbf{R}$,


$$F_X(t) = \begin{cases} 0 & \text{si } t < 0, \\ t & \text{si } t \in [0, 1], \\ 1 & \text{sinon,} \end{cases} \text{ et plus généralement : } F_Y(t) = \begin{cases} 0 & \text{si } t < a, \\ \frac{t-a}{b-a} & \text{si } t \in [a, b], \\ 1 & \text{sinon.} \end{cases}$$



2. **(Stabilité)** $X \hookrightarrow \mathcal{U}[0, 1] \iff Y = a + (b - a)X \hookrightarrow \mathcal{U}[a, b]$.



3. **(Espérance / Variance)** Les variables aléatoires X et Y admettent une espérance et une variance, données par :

$$\mathbf{E}(Y) = \frac{a + b}{2}, \quad \mathbf{Var}(Y) = \frac{(b - a)^2}{12}.$$

Preuve


1. Il suffit de faire les calculs pour Y . 

2.  

3. Faisons les calculs pour $X \hookrightarrow \mathcal{U}([0, 1])$, on déduira alors ceux pour $Y \hookrightarrow \mathcal{U}([a, b])$.



Exemple 18 — Fonction de loi uniforme Soit $X \hookrightarrow \mathcal{U}\left(\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]\right)$. Montrer que $Y = \tan X$ est à densité, en déterminer une. 

SIMULATION. Nous avons déjà établi le lien entre l'uniforme $\mathcal{U}[0, 1]$ et $\mathcal{U}[a, b]$ pour $a < b$. Il nous sert donc à simuler l'uniforme générale.

Simulation de la loi uniforme continue

```
import random as rd
def uniforme(a,b):
    """
    (a,b)->une simulation de l'uniforme sur [a,b]
    """
    return a + (b - a)*rd.random()
```

Le module `np.random` ne sait pas simuler directement la loi uniforme continue. On fait donc plutôt appel pour cela à la sous-bibliothèque `random` de `numpy`.

```
>>> np.random.uniform(-1, 2)
0.7509610442953076
```

3.2. Loi exponentielle

Définition ALEA.14.6 | Loi exponentielle

Soit $\lambda > 0$. On dit qu'une variable aléatoire X suit une *loi exponentielle* $\mathcal{E}(\lambda)$ si elle admet pour densité

$$f_\lambda : x \mapsto \begin{cases} \lambda e^{-\lambda x} & \text{si } x \in [0, +\infty[, \\ 0 & \text{sinon.} \end{cases}$$

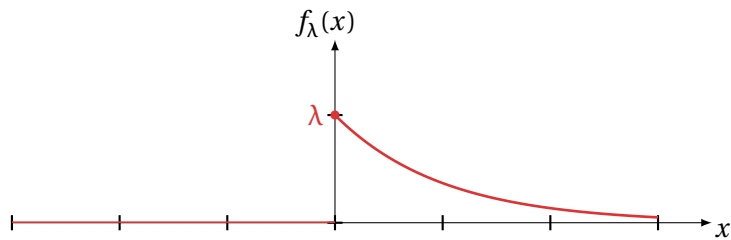



FIG. ALEA.14.4. : Graphe de la densité de l'exponentielle

Remarque 3.3 — Valeurs prises En particulier,

$$P(X \in \mathbf{R}^+) = 1.$$

Preuve Il existe un espace probabilisé et une variable aléatoire vérifiant ces conditions. En effet, 

Proposition ALEA.14.11 | Espérance, variance, propriétés

Soit $\lambda > 0$ et X une variable aléatoire réelle.

1. Supposons que $X \leftrightarrow \mathcal{E}(\lambda)$. Alors pour tout $t \in \mathbf{R}$,

$$F_X(t) = \begin{cases} 0 & \text{si } t < 0, \\ 1 - e^{-\lambda t} & \text{sinon.} \end{cases}$$

2. **(Stabilité)** $X \leftrightarrow \mathcal{E}(\lambda) \iff \lambda X \leftrightarrow \mathcal{E}(1)$.

3. **(Espérance / Variance)** Supposons que $X \leftrightarrow \mathcal{E}(\lambda)$. Alors

$$E(X) = \frac{1}{\lambda}, \quad \text{Var}(X) = \frac{1}{\lambda^2}.$$

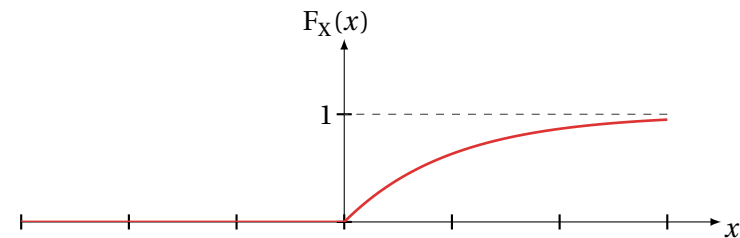



FIG. ALEA.14.5. : Graphe de la fonction de répartition de l'exponentielle

Preuve


1. 


2. \implies 



3. Faisons les calculs pour $X \hookrightarrow \mathcal{E}(1)$, on déduira alors ceux pour $Y = \lambda X \hookrightarrow \mathcal{E}(\lambda)$.



Exemple 20 – Minimum de lois exponentielles Soient $n \geq 1$ et $X_1, \dots, X_n \hookrightarrow \mathcal{E}(\lambda)$ avec $\lambda > 0$. On suppose de plus ces variables aléatoires indépendantes. Montrer que $Y = \min(X_1, \dots, X_n) \hookrightarrow \mathcal{E}(n\lambda)$. 

Exemple 19 – Fonction de loi exponentielle Soit $X \hookrightarrow \mathcal{E}(1)$. Montrer que $Y = \ln X$ est à densité, en déterminer une. 

PROPRIÉTÉ D'ABSENCE DE MÉMOIRE. Une loi exponentielle modélise la durée de vie d'un phénomène continu sans mémoire, ou sans vieillissement, ou sans usure. Nous avons déjà croisé, pour rappel, une loi de probabilité sans mémoire discrète : la loi géométrique.

Proposition ALEA.14.12 | Absence de mémoire

Soit $X \hookrightarrow \mathcal{E}(\lambda)$ avec $\lambda > 0$. Alors X est sans mémoire.

Remarque 3.4 — Comme pour la loi géométrique, la réciproque est vraie aussi : la seule loi à densité qui est à absence de mémoire est la loi exponentielle.

Preuve



SIMULATION. Commençons par la propriété qui va nous permettre de simuler la loi exponentielle à partir de l'uniforme sur $[0, 1]$.

Proposition ALEA.14.13 | Simulation de l'exponentielle

Soit $U \hookrightarrow \mathcal{U}([0, 1])$ et $\lambda > 0$. Alors :

$$X = -\frac{1}{\lambda} \ln(1 - U) \hookrightarrow \mathcal{E}(\lambda).$$

Attention

La preuve qui suit est à maîtriser parfaitement, très classique dans la pratique.

Preuve (*Point clef — Montrer que la fonction de répartition de $-\frac{1}{\lambda} \ln(1 - U) \hookrightarrow \mathcal{E}(\lambda)$ est celle d'une uniforme*)

► La variable aléatoire X est bien définie. 

► 

Simulation de la loi exponentielle

```
import random as rd
from math import log
def expo(lamba):
    """
    lamba -> une simulation de l'exponentielle de paramètre lamba
    """
    return -(1/lamba)*log(rd.random())
```

Le module random ne sait pas simuler directement la loi exponentielle. On fait donc plutôt appel pour cela à la sous bibliothèque random de numpy.

```
>>> np.random.exponential(1/3) # Attention au paramètre : c'est
↳ l'inverse du paramètre mathématique qu'il faut indiquer
0.026001255610887404
```

3.3. La loi normale

Cette loi de probabilité, aussi appelée loi de Gauß ou gaussienne, joue un rôle essentiel en probabilités et en statistiques. On verra dans le [Chapter ALEA.16](#) son caractère universel dans les situations où l'on observe une somme de variables aléatoires indépendantes et centrées : par exemple, on peut envisager les fluctuations d'un cours de bourse comme le résultat d'une addition de phénomènes aléatoires indépendants et de faible ampleur relative.

3.3.1. Loi normale centrée réduite

Définition ALEA.14.7 | Loi normale centrée réduite

On dit qu'une variable aléatoire X suit la loi normale centrée réduite ou gaussienne standard si elle admet pour densité

$$\varphi : x \in \mathbf{R} \mapsto \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

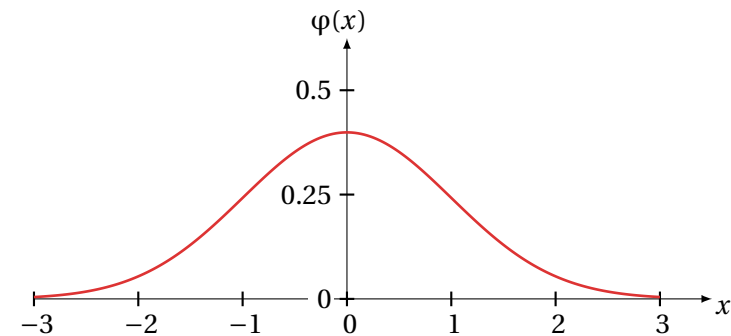



FIG. ALEA.14.6. : Graphe de la densité de la normale centrée réduite

On appelle aussi parfois ce graphe la « courbe en cloche ».

Preuve Il existe un espace probabilisé et une variable aléatoire vérifiant ces conditions. En effet, 

Proposition ALEA.14.14 | Espérance, variance, propriétés

Soit $X \hookrightarrow \mathcal{N}(0, 1)$. Alors X admet une espérance et une variance données par :

$$\mathbf{E}(X) = 0, \quad \mathbf{Var}(X) = 1.$$

Preuve



Exemple 21 — Fonction de loi normale Soit $X \hookrightarrow \mathcal{N}(0, 1)$. Montrer que $Y = |X|$ est à densité, en déterminer une.

Notation Fonction de répartition

Σ

On notera dans la suite Φ la fonction de répartition de la loi $\mathcal{N}(0, 1)$:

$$\forall x \in \mathbf{R}, \quad \Phi(x) = \mathbf{P}(X \leq x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt.$$

Notons qu'elle n'est pas calculable explicitement contrairement aux lois exponentielle et uniforme. En revanche, il est possible d'établir l'équivalent suivant.

Proposition ALEA.14.15 | Propriétés de Φ

1. (Symétrie) Pour tout $x \in \mathbf{R}$, on a :

$$\Phi(-x) = 1 - \Phi(x).$$

2. La fonction $\Phi : \mathbf{R} \rightarrow]0, 1[$ est bijective.

La première propriété peut être également comprise par un dessin.

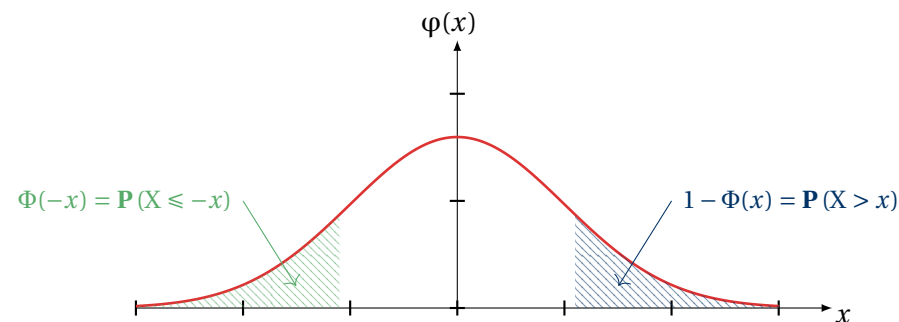


FIG. ALEA.14.7 : Illustration de la propriété de symétrie sur Φ

Preuve

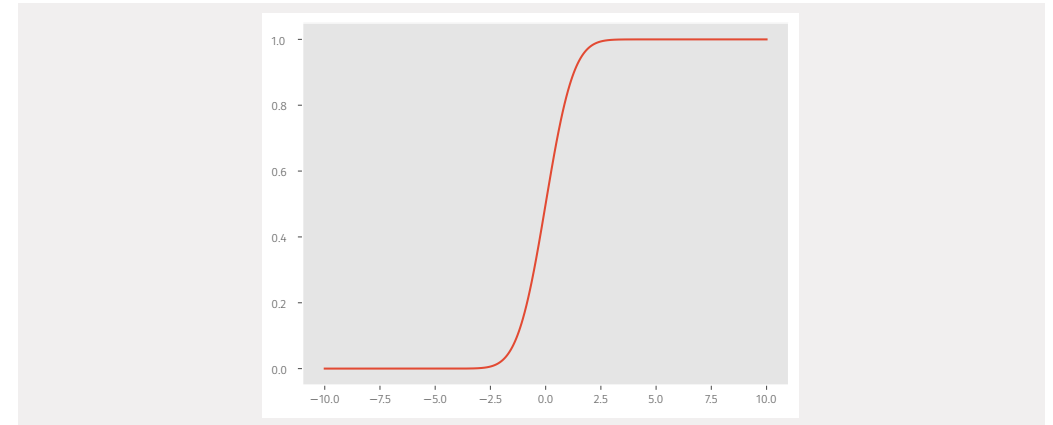
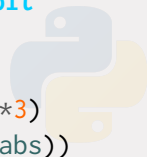
1. 2. 

On en dispose pas d'expression de Φ . On peut cependant utiliser Python pour esquisser son graphe.

Fonction de répartition de la normale centrée/réduite

```
import numpy as np
import matplotlib.pyplot as plt
import scipy.stats as stat

abs = np.linspace(-10,10,10**3)
plt.plot(abs, stat.norm.cdf(abs))
plt.grid()
```



On connaît cependant des tables de valeurs de Φ (voir l'annexe en fin de chapitre).

SIMULATION. Puisque la fonction de répartition est bijective dans le cas de la loi normale, nous pouvons donc facilement déduire une méthode de simulation.

Proposition ALEA.14.16 | Simulation de la normale centrée réduite

Soit $U \hookrightarrow \mathcal{U}([0, 1])$. Alors :

$$\Phi^{-1}(U) \hookrightarrow \mathcal{N}(0, 1).$$

Preuve



La fonction Φ^{-1} ne se calcule pas non plus facilement, on va utiliser le sous-module `scipy.stats` pour obtenir des valeurs approchées.

Simulation de la loi normale centrée réduite

```
import random as rd
import numpy as np
# import scipy.stats as stat

def norm():
    """
    ->une simulation de la N(0,1)
    """
    return stat.norm.ppf(rd.random())
```

Le module `random` ne sait pas simuler directement la loi normale. On fait donc plutôt appel pour cela à la sous-bibliothèque `random` de `numpy`.

```
>>> np.random.normal()
1.2747097167695032
```

3.3.2. Loi normale générale

Définition ALEA.14.8 | Loi normale

Soient $\mu \in \mathbf{R}$ et $\sigma > 0$. On dit qu'une variable aléatoire X suit la loi normale (ou gaussienne) $\mathcal{N}(\mu, \sigma^2)$ si elle admet pour densité

$$\varphi_{\mu, \sigma^2} : x \mapsto \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

On le note $X \hookrightarrow \mathcal{N}(\mu, \sigma^2)$.

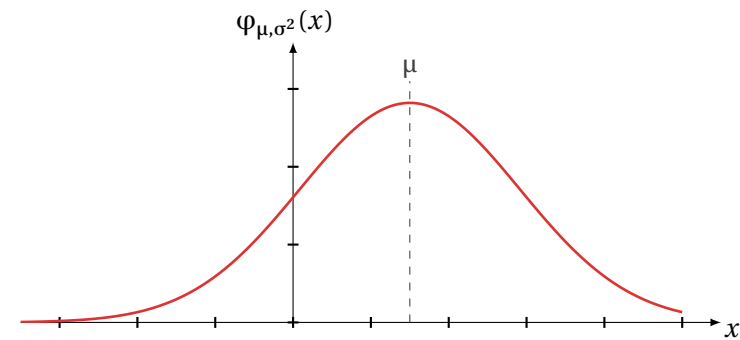



FIG. ALEA.14.8 : Graphe de la densité de la normale générale

Preuve Il existe un espace probabilisé et une variable aléatoire vérifiant ces conditions. En effet, 

Proposition ALEA.14.17 | Espérance et variance

Soit $X \hookrightarrow \mathcal{N}(\mu, \sigma^2)$ alors X admet une espérance et une variance données par :

$$\mathbf{E}(X) = \mu, \quad \mathbf{Var}(X) = \sigma^2.$$

Autrement dit,

$$X \hookrightarrow \mathcal{N}(\mu, \mathbf{Var}(X)).$$

Proposition ALEA.14.18 | Stabilité

► **(Lien avec la gaussienne standard : centrage / réduction)** Soient $\mu \in \mathbf{R}$ et $\sigma > 0$.


$$X \hookrightarrow \mathcal{N}(\mu, \sigma^2) \iff X^* = \frac{X - \mu}{\sigma} \hookrightarrow \mathcal{N}(0, 1).$$

En d'autres termes, si une variable aléatoire X suit une loi $\mathcal{N}(\mu, \sigma^2)$ alors sa centrée réduite associée suit une loi normale centrée réduite, et inversement.

- ▶ **(Fonction affine d'une loi normale)** Plus généralement, soient $\mu \in \mathbf{R}$, $\sigma > 0$ et $X \hookrightarrow \mathcal{N}(\mu, \sigma^2)$. Alors pour tout $(a, b) \in \mathbf{R}^2$ de sorte que $a \neq 0$,


$$aX + b \hookrightarrow \mathcal{N}(a\mu + b, (a\sigma)^2).$$

En d'autres termes, toute variable aléatoire affine non constante d'une loi normale suit encore une loi normale, de paramètres son espérance et sa variance.

Preuve Commençons par prouver le second point. 

On déduit alors facilement le premier.



Preuve On peut maintenant démontrer la [Proposition ALEA.14.17](#). 

Proposition ALEA.14.19 | Somme de lois normales indépendantes

Soient X_1, \dots, X_n une collection de n variables aléatoires indépendantes telles que :

$$\forall i \in \llbracket 1, n \rrbracket, \quad X_i \hookrightarrow \mathcal{N}(\mu_i, \sigma_i^2), \quad \mu_i \in \mathbf{R}, \quad \sigma_i > 0.$$

- ▶ Alors :

$$\sum_{i=1}^n X_i \hookrightarrow \mathcal{N}\left(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2\right) = \mathcal{N}\left(\mathbf{E}\left(\sum_{i=1}^n X_i\right), \mathbf{Var}\left(\sum_{i=1}^n X_i\right)\right).$$

- ▶ Et plus généralement : pour toute famille $(\lambda_1, \dots, \lambda_n) \in \mathbf{R}^n \setminus \{(0, \dots, 0)\}$, on a :

$$\begin{aligned} \sum_{i=1}^n \lambda_i X_i &\hookrightarrow \mathcal{N}\left(\sum_{i=1}^n \lambda_i \cdot \mu_i, \sum_{i=1}^n \lambda_i \cdot \sigma_i^2\right) \\ &= \mathcal{N}\left(\mathbf{E}\left(\sum_{i=1}^n \lambda_i \cdot X_i\right), \mathbf{Var}\left(\sum_{i=1}^n \lambda_i \cdot X_i\right)\right). \end{aligned}$$

En résumé, toute combinaison linéaire de lois normales indépendantes est encore une loi normale. On obtient alors ses paramètres en calculant son espérance et sa variance.

Preuve Provisoirement admis.

Exemple 22 — Soient $n \geq 1$, X_1, \dots, X_n une collection de n variables aléatoires indépendantes, telles que $X_i \hookrightarrow \mathcal{N}\left(0, \frac{1}{i^2}\right)$ pour tout $i \in \llbracket 1, n \rrbracket$. Quelle est la loi de $\sum_{i=1}^n i \cdot X_i$?



Nom	Para- mètre(s)	Notation	Densité	Espérance	Variance
Uniforme	$a < b$ deux réels	$\mathcal{U}_{[a,b]}$	$\frac{1}{b-a} \mathbb{1}_{[a,b]}(x)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponentielle	$\lambda \in]0, +\infty[$	$\mathcal{E}(\lambda)$	$\lambda e^{-\lambda x} \mathbb{1}_{[0, +\infty[}(x)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Laplace	$\lambda \in]0, \infty[$	$\mathcal{L}(\lambda)$	$\frac{\lambda}{2} e^{-\lambda x }$	0	2
Normale (dans \mathbf{R})	$(m, \sigma^2) \in \mathbf{R} \times]0, +\infty[$	$\mathcal{N}(m, \sigma^2)$	$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}$	μ	σ^2

DENSITÉ, FONCTION DE RÉPARTITION, SIMULATION. Techniques similaires à la loi $\mathcal{N}(0, 1)$, on ne les détaille donc pas comme précédemment.

- ▶ On peut rajouter des paramètres loc, et scale pour obtenir les graphes des fonctions de répartition et densité de la loi normale générale.
- ▶ Pour la simulation, on utilise celle de la $\mathcal{N}(0, 1)$, puis on retourne $\mu + \sigma \cdot \mathcal{N}(0, 1)$.
- ▶ Pour la simulation, on peut aussi utiliser les fonctions existantes de `scipy.stats` en précisant directement les bons paramètres.

3.4. Bilan des lois continues usuelles

Le tableau suivant rassemble quelques lois continues usuelles.

ANNEXE – TABLE DE VALEURS DE LA $\mathcal{N}(0, 1)$

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.090
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545

**Méthode Lecture d'une table**

Si l'on souhaite avoir, par exemple, $\Phi(0,96)$, on :

1. se place sur la ligne «0.9»,
2. se place ensuite sur la colonne «0.06».
3. On obtient alors la valeur désirée.

1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

$$\Phi(x) = \mathbf{P}(X \leq x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \quad \text{et} \quad \Phi(-x) = 1 - \Phi(x).$$

On retiendra en particulier la valeur typique $\Phi(1.96) = 0.975$, de sorte que :

$$\mathbf{P}(X \in [-1.96; 1.96]) = 2\Phi(1.96) - 1 = 0.95$$

*** **Fin du chapitre** ***

4. EXERCICES

4.1. Généralités

Exercice ALEA.14.1 | Soit X une variable aléatoire réelle à densité de densité f continue et telle que $E(X^2)$ existe.

1. Montrer que si $x > 0$, alors :

$$0 \leq x^2 \mathbf{P}(|X| \geq x) \leq \int_{-\infty}^{-x} t^2 f(t) dt + \int_x^{\infty} t^2 f(t) dt.$$

En déduire que $\lim_{x \rightarrow \infty} [x^2 \mathbf{P}(|X| \geq x)] = 0$.

2. Soit $x \in \mathbf{R}^+$. Montrer que :

$$\int_0^x t \mathbf{P}(|X| \geq t) dt = \frac{x^2}{2} \mathbf{P}(|X| \geq x) + \frac{1}{2} \int_{-x}^x t^2 f(t) dt.$$

Indication : On pourra utiliser la fonction de répartition F de X et une intégration par parties.

Solution (exercice ALEA.14.1)

1. Soit $x > 0$, alors :

$$\begin{aligned} 0 \leq x^2 \mathbf{P}(|X| \geq x) &= x^2 \mathbf{P}(X \leq -x) + x^2 \mathbf{P}(X \geq x) \\ &= \int_{-\infty}^{-x} x^2 f(t) dt + \int_x^{\infty} x^2 f(t) dt \\ &\leq \int_{-\infty}^{-x} t^2 f(t) dt + \int_x^{\infty} t^2 f(t) dt. \end{aligned}$$

Puisque $E(X^2)$ existe, l'intégrale

$$\int_{-\infty}^{\infty} t^2 f(t) dt$$

converge, donc

$$\int_x^{\infty} t^2 f(t) dt = \int_{-\infty}^{\infty} t^2 f(t) dt - \int_{-\infty}^x t^2 f(t) dt \xrightarrow{x \rightarrow \infty} 0$$

par définition de l'intégrale. On montre de-même que

$$\int_{-\infty}^{-x} t^2 f(t) dt \xrightarrow{x \rightarrow \infty} 0,$$

puis par théorème d'encadrement que :

$$\lim_{x \rightarrow \infty} [x^2 \mathbf{P}(|X| \geq x)] = 0.$$

2. Soit $x \in \mathbf{R}^+$. Puisque $|X|$ est à densité, on a :

$$\forall t \in \mathbf{R}, \quad \mathbf{P}(|X| \geq t) = 1 - \mathbf{P}(|X| \leq t) = 1 - (F(t) - F(-t)).$$

Alors :

$$\begin{aligned} \int_0^x t \mathbf{P}(|X| \geq t) dt &= \int_0^x t(1 - (F(t) - F(-t))) dt \\ &= - \int_0^x \frac{t^2}{2} (-f(t) + f(-t)) dt + \left[(1 - (F(t) - F(-t))) \frac{t^2}{2} \right]_0^x \\ &= \frac{1}{2} \int_0^x t^2 f(t) dt - \frac{1}{2} \int_0^x t^2 f(-t) dt + \left[\mathbf{P}(|X| \geq t) \frac{t^2}{2} \right]_0^x \\ &= \frac{x^2}{2} \mathbf{P}(|X| \geq x) + \frac{1}{2} \int_{-x}^x t^2 f(t) dt, \end{aligned}$$

$t \rightarrow \frac{t^2}{2}, t \rightarrow 1 - (F(t) - F(-t))$
sont \mathcal{C}^1 car f est continue

où à la dernière étape nous avons réalisé le changement de variable $u = -t$ dans la dernière intégrale, de classe \mathcal{C}^1 , puis utilisé la relation de CHASLES.

Exercice ALEA.14.2 | Formule des cumulants, cas continu Soit X une variable aléatoire positive admettant une densité positive f . On notera F sa fonction de répartition.

On souhaite montrer que X admet une espérance si et seulement si $\int_0^\infty \mathbf{P}(X > t) dt$ converge, et qu'en cas de convergence :

$$\mathbf{E}(X) = \int_0^\infty \mathbf{P}(X > t) dt.$$

1. En admettant provisoirement le résultat de l'exercice, montrer que si X, Y sont deux variables aléatoires indépendantes suivant une $\mathcal{E}(\lambda)$ avec $\lambda > 0$, alors $Z = \min(X, Y)$ admet une espérance et la calculer.
2. On souhaite maintenant démontrer la proposition dans le cas où f est continue, pour simplifier.

2.1) Soit $A > 0$. Montrer la formule suivante :

$$\int_0^A tf(t) dt = \int_0^A \mathbf{P}(X > t) dt + A(F(A) - 1).$$

2.2) Conclure.

Solution (exercice ALEA.14.2)

1. On a $Z = \min(X, Y)$. Puisque $X, Y \hookrightarrow \mathcal{E}(\lambda)$, on a pour tout $t \geq 0$

$$\begin{aligned} \mathbf{P}(Z > t) &= \mathbf{P}(X > t, Y > t) \\ &= \mathbf{P}(X > t) \mathbf{P}(Y > t) \quad \left. \begin{array}{l} \\ \end{array} \right\} \text{indépendance} \\ &= (e^{-\lambda t})^2 \\ &= e^{-2\lambda t}. \end{aligned}$$

Par ailleurs, $\mathbf{P}(Z > t) = 1$ si $t < 0$. Donc :

$$\forall t \in \mathbf{R}, \quad F_Z(t) = \begin{cases} 1 - e^{-2\lambda t} & \text{si } t \geq 0, \\ 0 & \text{sinon.} \end{cases}$$

Donc F_Z est bien continue, puisque $1 - e^{-2\lambda t} \xrightarrow{t \rightarrow 0} 0$, et elle est \mathcal{C}^1 sauf peut-être en 0. Donc Z est à densité, et une densité est donnée par :

$$\forall t \in \mathbf{R}, \quad f_Z(t) = \begin{cases} 2\lambda e^{-2\lambda t} & \text{si } t \geq 0, \\ 0 & \text{sinon.} \end{cases}$$

Autrement dit : $Z \hookrightarrow \mathcal{E}(2\lambda)$. Elle possède donc une espérance, et elle vaut

$$\mathbf{E}(Z) = \frac{1}{2\lambda}.$$

- 2.1) Soit $A > 0$. Puisque $G : x \mapsto F(x) - 1$ est une primitive de f , et que f est continue, G est de classe \mathcal{C}^1 . Par ailleurs, $x \mapsto x$ est également de classe \mathcal{C}^1 . Donc par intégration par parties :

$$\int_0^A tf(t) dt = - \int_0^A (F(t) - 1) dt + A(F(A) - 1) - 0(F(0) - 1).$$

Ainsi, puisque $F(t) - 1 = -(1 - F(t)) = -\mathbf{P}(X > t)$ pour tout $t \in \mathbf{R}$, on déduit que :

$$\int_0^A tf(t) dt = \int_0^A \mathbf{P}(X > t) dt + A(F(A) - 1).$$

- 2.2) Procédons par double implication.

\Leftarrow Supposons que $\int_0^\infty \mathbf{P}(X > t) dt$ converge. Alors l'intégrale de droite converge vers une limite finie ou de manière équivalente (on intègre des fonctions positives) il existe $M \in \mathbf{R}^+$ tel que

$$\forall A \in \mathbf{R}^+, \quad \int_0^A \mathbf{P}(X > t) dt \leq M.$$

Et puisque $A(F(A) - 1) \leq 0$ pour tout $A \in \mathbf{R}^+$, on déduit

$$0 \leq \int_0^A tf(t) dt \leq \int_0^A \mathbf{P}(X > t) dt \leq M.$$

Ainsi, la fonction $A \mapsto \int_0^A tf(t) dt$ est majorée, croissante car l'intégrande est positive, donc converge. Ainsi, X possède une espérance.

\Rightarrow Réciproquement, supposons que X possède une espérance. Alors

$$0 \leq A(1 - F(A)) = A \int_A^\infty f(t) dt \leq \int_A^\infty tf(t) dt \xrightarrow{A \rightarrow \infty} 0$$

car il s'agit du reste d'une intégrale convergente. Donc par théorème d'encadrement, on déduit que si X admet une espérance, alors

$$A(1 - F(A)) \xrightarrow{A \rightarrow \infty} 0,$$

et donc $A(F(A) - 1) \xrightarrow{A \rightarrow \infty} 0$. D'après la question précédente, vient alors :

$$\int_0^\infty \mathbf{P}(X > t) dt \text{ converge.}$$


4.2. Études de lois

Exercice ALEA.14.3 | Soit g la fonction définie sur \mathbf{R} par :

$$\forall t \in \mathbf{R}, \quad g(t) = \begin{cases} 0 & \text{si } t < 1, \\ \frac{b}{2^t} & \text{si } t \geq 1, \end{cases} \quad b \in \mathbf{R}.$$


1. Déterminer b pour que g soit la densité d'une variable aléatoire à densité. Déterminer la loi de $X - 1$.
2. En déduire que X possède une espérance et une variance.

Exercice ALEA.14.4 | **Loi de CAUCHY** Soit $\Psi : x \mapsto \frac{1}{\pi(1+x^2)}$.

1. Montrer que Ψ est une densité de probabilité. Dans la suite on considère X de densité Ψ . On dit que X suit une loi de CAUCHY.
2. Déterminer la fonction de répartition F associée à X .
3. La variable aléatoire X admet-elle une espérance? Une variance?
4. **(Simulation)** Soit U une variable aléatoire réelle de loi uniforme sur $]0, 1[$.
 - 4.1) Après avoir déterminé F^{-1} , montrer que X et $F^{-1}(U)$ ont même loi, *i.e.* même fonction de répartition.
 - 4.2)  En déduire une méthode de simulation de la variable aléatoire X .

Exercice ALEA.14.5 |

1. **(Existence de arccos)**

- 1.1) Justifier que $\cos|_{]0, \pi]}$ réalise une bijection de $]0, \pi[$ vers $]-1, 1[$. On notera arccos la fonction réciproque de cette bijection.
- 1.2) Justifier que arccos est dérivable sur $]-1, 1[$ et calculer arccos' sur cet intervalle. Quelle est sa monotonie?
2. Soit X une variable aléatoire uniforme sur $[0, 1]$.
 - 2.1) Déterminer la loi de $Y = \cos(\pi X)$.
 - 2.2)  Écrire une fonction `simulY()` qui simule la variable Y .
 - 2.3) Calculer, si elles existent, l'espérance et la variance de Y .

Solution (exercice ALEA.14.5)

- 1.
2. 2.1) Déterminons la loi de $Y = \cos(\pi X)$, c'est-à-dire une densité dans ce contexte. Constatons que $\mathbf{P}(Y \in [-1, 1]) = 1$. Soit $t \in \mathbf{R}$, alors puisque arccos est décroissante et que $\arccos(\cos(\pi x)) = \pi x$ pour tout $x \in [0, 1]$,

$$\begin{aligned} \mathbf{P}(\cos(\pi X) \leq t) &= \begin{cases} 0 & \text{si } t \leq -1 \\ \mathbf{P}(\cos(\pi X) \leq t) = \mathbf{P}(\pi X \geq \arccos t) & \text{si } t \in]0, 1[\\ 1 & \text{si } t \geq 1 \end{cases} \\ &= \begin{cases} 0 & \text{si } t \leq -1 \\ \mathbf{P}\left(X \geq \frac{\arccos t}{\pi}\right) & \text{si } t \in]0, 1[\\ 1 & \text{si } t \geq 1. \end{cases} \end{aligned}$$


Lorsque $t \in]0, 1[$, puisque $X \hookrightarrow \mathcal{U}([0, 1])$, nous avons $\mathbf{P}(\cos(\pi X) \leq t) = 1 - \frac{\arccos t}{\pi}$. On déduit alors la fonction de répartition de Y :

$$\forall t \in \mathbf{R}, \quad F_Y(t) = \begin{cases} 0 & \text{si } t \leq -1 \\ 1 - \frac{\arccos t}{\pi} & \text{si } t \in]-1, 1[\\ 1 & \text{si } t \geq 1. \end{cases}$$

Puisque $1 - \frac{\arccos t}{\pi} \xrightarrow{t \rightarrow -1} = 1 - \frac{\pi}{\pi} = 0$, et $1 - \frac{\arccos t}{\pi} \xrightarrow{t \rightarrow 1} = 1 - \frac{0}{\pi} = 1$, la fonction F_Y est continue. De plus, elle est d'après la première question \mathcal{C}^1 sauf peut-être

en $-1, 1$. Donc Y est à densité et une densité est donnée par :

$$t \mapsto f_Y(t) = \begin{cases} \frac{1}{\pi\sqrt{1-t^2}} & \text{si } t \in]-1, 1[\\ 0 & \text{sinon.} \end{cases}$$

2.2)  Écrivons une fonction `simuY()` qui simule la variable Y .

```
from math import pi, cos
import random as rd
def simuY():
    return cos(pi*rd.random())
```

2.3) On applique le théorème du transfert, puisque $\mathbf{P}(Y \in [-1, 1]) = 1$ et elle admet donc une espérance donnée par :

$$\mathbf{E}(X) = \int_0^1 x \cos(\pi x) dx, \quad \mathbf{Var}(X) = \int_0^1 x^2 \cos(\pi x) dx - \mathbf{E}(X)^2$$

Les deux intégrales se calculant ensuite par intégration par parties. En effet,

$$\begin{aligned} \mathbf{E}(X) &= \int_0^1 x \cos(\pi x) dx \\ &= -\frac{1}{\pi} \int_0^1 \sin(\pi x) dx + \left[\frac{x}{\pi} \sin(\pi x) \right]_0^1 \\ &= \frac{1}{\pi^2} [\cos(\pi x)]_0^1 \\ &= \frac{1}{\pi^2} (-1 - 1) = \frac{-2}{\pi^2}. \end{aligned}$$

$x \mapsto x, x \mapsto \frac{1}{\pi} \sin(\pi x)$ sont \mathcal{C}^1

On procède de-même pour la variance.

Exercice ALEA.14.6 | Moments d'ordre n de lois usuelles. Après en avoir justifié l'existence, calculer $\mathbf{E}(X^n)$ dans les cas suivants :

1. $X \mapsto \mathcal{U}[a, b]$,
2. $X \mapsto \mathcal{E}(\lambda)$,
3. $X \mapsto \mathcal{N}(0, 1)$.

4.3. Fonctions de variables aléatoires

Exercice ALEA.14.7 | Carré lois classiques Déterminer la densité de $Y = X^2$ dans les cas suivants :

1. $X \mapsto \mathcal{U}([-1, 1])$,
2. $X \mapsto \mathcal{U}([-1, 2])$.
3. $X \mapsto \mathcal{N}(0, 1)$.

Exercice ALEA.14.8 | Loi log-normale Soit $a > 0$ et h la fonction définie par

$$\forall x \in \mathbf{R}, \quad h(x) = \begin{cases} \frac{1}{ax\sqrt{2\pi}} e^{-\frac{(\ln x)^2}{2a^2}} & \text{si } x > 0, \\ 0 & \text{sinon.} \end{cases}$$

1. Montrer que h est une densité de probabilité.
2. Soit X une variable aléatoire de densité h . Montrer que X admet une espérance, calculer $\mathbf{E}(X)$.
3. Déterminer la loi de $Y = \frac{\ln X}{a}$.

4.4. Discret vs. Continu

Exercice ALEA.14.9 | Soit $\lambda \in [0, 1]$, et f_λ la fonction définie sur \mathbf{R} par

$$\forall x \in \mathbf{R}, \quad f_\lambda(x) = \begin{cases} \lambda & \text{si } x \in]-1, 0], \\ 1 - \lambda & \text{si } x \in]0, 1], \\ 0 & \text{sinon.} \end{cases}$$

1. Montrer que f_λ est une densité de probabilité. Soit X admettant f_λ pour densité.
2. Montrer que X admet une espérance et une variance. Les calculer.
3. Soient $\alpha > 0$ et $N \hookrightarrow \mathcal{P}(\alpha)$. Soit de plus Y une variable aléatoire telle que, pour tout $k \geq 0$, la loi conditionnelle sachant $\{N = k\}$ de Y soit à densité de densité $f_{\frac{1}{k+1}}$, i.e. pour tout intervalle $I \subset \mathbf{R}$,

$$\mathbf{P}(Y \in I | N = k) = \int_I f_{\frac{1}{k+1}}.$$

- 3.1) Déterminer la fonction de répartition de Y .
- 3.2) En déduire que Y est à densité et donner une densité.
- 3.3) Montrer qu'il existe une unique valeur de α telle que $\mathbf{E}(Y) = 0$.
- 3.4) ➤_🔗 Écrire une fonction Python qui donne une valeur approchée de α à eps près, le faire tourner et donner une valeur approchée de α à 10^{-3} près.

Solution (exercice ALEA.14.9)

1. La fonction f_λ est continue sauf éventuellement en $-1, 0, 1$, positive car $\lambda \in [0, 1]$, et son intégrale converge car f_λ est nulle en dehors de $]0, 1]$. De plus,

$$\int_{-1}^1 f_\lambda = \int_{-1}^0 \lambda dx + \int_0^1 (1 - \lambda) dx = \lambda + 1 - \lambda = 1.$$

Ainsi f_λ est une densité de probabilités.

2. Puisque $\mathbf{P}(X \in]-1, 1]) = 1$, elle possède une espérance et une variance. Et

$$\begin{aligned} \mathbf{E}(X) &= \int_{-1}^1 x f_\lambda(x) dx \\ &= \lambda \left[\frac{x^2}{2} \right]_{-1}^0 + (1 - \lambda) \left[\frac{x^2}{2} \right]_0^1 \\ &= -\frac{\lambda}{2} + \frac{1}{2}(1 - \lambda) = \boxed{\frac{1}{2} - \lambda}, \end{aligned}$$

$$\begin{aligned} \mathbf{E}(X^2) &= \int_{-1}^1 x^2 f_\lambda(x) dx \\ &= \lambda \left[\frac{x^3}{3} \right]_{-1}^0 + (1 - \lambda) \left[\frac{x^3}{3} \right]_0^1 \\ &= \frac{\lambda}{3} + (1 - \lambda) \frac{1}{3} \\ &= \frac{1}{3}, \end{aligned}$$

$$\mathbf{Var}(X) = \boxed{\frac{1}{3} - \left(\frac{1}{2} - \lambda\right)^2}.$$

3. 3.1) Pour tout $t \in \mathbf{R}, k \in \mathbf{N}$, commençons par calculer

$$\mathbf{P}(Y \leq t | N = k) = \int_{-\infty}^t f_{\frac{1}{k+1}}(x) dx.$$

En faisant un dessin, on se convainc des sous-cas à considérer.

- ▶ Si $t \leq -1$, alors $\mathbf{P}(Y \leq t | N = k) = 0$.
- ▶ Si $t \in]-1, 0]$, alors

$$\mathbf{P}(Y \leq t | N = k) = \int_{-1}^t \frac{dx}{k+1} = \frac{t+1}{k+1}.$$

- ▶ Si $t \in]0, 1]$, alors

$$\mathbf{P}(Y \leq t | N = k) = \int_{-1}^0 \frac{dx}{k+1} + \int_0^t \frac{k}{k+1} dx = \frac{t k + 1}{k + 1}.$$

- ▶ Si $t > 1$, alors $\mathbf{P}(Y \leq t | N = k) = \int_{-\infty}^{\infty} f_{\frac{1}{k+1}} = 1$.

En résumé,

$$\forall t \in \mathbf{R}, \quad \mathbf{P}(Y \leq t | N = k) = \begin{cases} 0 & \text{si } t \leq -1, \\ \frac{t+1}{k+1} & \text{si } t \in]-1, 0], \\ \frac{tk+1}{k+1} & \text{si } t \in]0, 1], \\ 1 & \text{sinon.} \end{cases}$$

Donc d'après la formule des probabilités totales appliquée au système complet d'évènements $\{N = k\}_{k \in \mathbf{N}}$: En résumé,

$$\forall t \in \mathbf{R}, \quad \mathbf{P}(Y \leq t) = \begin{cases} \sum_{k=0}^{\infty} 0 \mathbf{P}(N = k) = 0 & \text{si } t \leq -1, \\ \sum_{k=0}^{\infty} \frac{t+1}{k+1} \mathbf{P}(N = k) & \text{si } t \in]-1, 0], \\ \sum_{k=0}^{\infty} \frac{tk+1}{k+1} \mathbf{P}(N = k) & \text{si } t \in]0, 1], \\ \sum_{k=0}^{\infty} 1 \mathbf{P}(N = k) = 1 & \text{sinon.} \end{cases}$$

Il ne reste donc plus que deux cas.

► Si $t \in]-1, 0]$, alors :

$$\begin{aligned} F_Y(t) &= \sum_{k=0}^{\infty} \frac{t+1}{k+1} e^{-\alpha} \frac{\alpha^k}{k!} \\ &= e^{-\alpha} (t+1) \sum_{k=0}^{\infty} \frac{\alpha^k}{(k+1)!} \\ &= \frac{e^{-\alpha}}{\alpha} (t+1) \sum_{\ell=1}^{\infty} \frac{\alpha^\ell}{\ell!} \\ &= \frac{e^{-\alpha}}{\alpha} (e^\alpha - 1) \\ &= \frac{1 - e^{-\alpha}}{\alpha} (t+1). \end{aligned}$$

► Si $t \in]0, 1]$, alors :

$$\begin{aligned} F_Y(t) &= \sum_{k=0}^{\infty} \frac{tk+1}{k+1} e^{-\alpha} \frac{\alpha^k}{k!} \\ &= e^{-\alpha} \sum_{k=0}^{\infty} \left(t + \frac{1-t}{k+1} \right) \frac{\alpha^k}{k!} \quad \left. \begin{array}{l} \frac{tk+1}{k+1} = \frac{t(k+1)+(1-t)}{k+1} = t + \frac{1-t}{k+1} \end{array} \right\} \\ &= e^{-\alpha} \left(te^\alpha + (1-t) \frac{1}{\alpha} \sum_{k=0}^{\infty} \frac{\alpha^{k+1}}{(k+1)!} \right) \\ &= e^{-\alpha} \left(te^\alpha + (1-t) \frac{1}{\alpha} (e^\alpha - 1) \right) \\ &= t + (1-t) \frac{1}{\alpha} (1 - e^{-\alpha}). \end{aligned}$$

On a donc enfin obtenu la fonction de répartition cherchée :

$$\forall t \in \mathbf{R}, \quad \mathbf{P}(Y \leq t) = \begin{cases} 0 & \text{si } t \leq -1, \\ \frac{1-e^{-\alpha}}{\alpha} (t+1) & \text{si } t \in]-1, 0], \\ t + (1-t) \frac{1}{\alpha} (1 - e^{-\alpha}) & \text{si } t \in]0, 1], \\ 1 & \text{sinon.} \end{cases}$$

3.2) On vérifie ensuite que F_Y est bien continue, et elle est \mathcal{C}^1 en tant que somme composée de telles fonctions, sauf éventuellement en $-1, 0, 1$. Donc Y est à densité et une densité est donnée par :

$$f_Y : t \longrightarrow \begin{cases} \frac{1-e^{-\alpha}}{\alpha} & \text{si } t \in]-1, 0], \\ 1 - \frac{1-e^{-\alpha}}{\alpha} & \text{si } t \in]0, 1], \\ 0 & \text{sinon.} \end{cases}$$

On observe que Y a pour densité f_λ avec $\lambda = \frac{1-e^{-\alpha}}{\alpha}$.

3.3) D'après la première question, on déduit que :

$$\mathbf{E}(Y) = \frac{1}{2} - \frac{1 - e^{-\alpha}}{\alpha}$$

Ainsi Y est centrée si et seulement si

$$\alpha - 2 + 2e^{-\alpha} = 0.$$

Notons $f : x \mapsto x - 2 + 2e^{-x}$. Elle est dérivable et $f'(x) = 1 - 2e^{-x} = 0$ si et seulement si $x = \ln 2$. En résolvant les inéquations associées, on trouve que $f'(x) \geq 0$ si et seulement si $x \geq \ln(2)$ donc f est croissante strictement sur $[\ln 2, \infty[$ et décroissante strictement sur $]0, \ln 2[$. Mais comme $f(0) = 0$, f ne peut s'annuler sur $]0, \ln 2[$, $f(\ln 2) = \ln(2) - 2 + 1 = \ln(2) - 1 < 0$, $\lim_{x \rightarrow \infty} f(x) = \infty$ par règles usuelles sur les limites et f est continue strictement croissante sur $[\ln 2, \infty[$ donc réalise une bijection vers $[\ln 2 - 1, \infty[\ni 0$.

Il existe donc un unique $\alpha > 0$ qui centre Y .

3.4) >_☞

```
import math as ma
def f(x):
    return x - 2 + 2*ma.exp(-x)

def dichot(a, b, f, prec):
    """
    Retourne une valeur approchée d'un zéro de f entre a et
    - b avec précision prec
    Retourne faux si aucune racine n'existe
    """
    while b - a > prec:
        c = (a + b)/2
        if f(a)*f(c) <= 0:
            # changement de signe sur [a,c]
            b = c
        else:
            # pas de changement de signe sur [a,c]
            a = c
    return (a + b)/2
```

Pour savoir comment initialiser la borne de droite, on peut calculer plusieurs valeurs de f sur $[\ln 2, \infty[$ et prendre la première qui dépasse 0.

```
>>> dichot(ma.log(2), 2, f, 10**(-3))
1.593841884973831
```

Exercice ALEA.14.10 | Discret / Continu.

1. Soit X une variable aléatoire de loi exponentielle $\mathcal{E}(\lambda)$. Déterminer la loi de $Y = \lfloor X \rfloor$.
2. Soit $Z \mapsto \mathcal{B}\left(n, \frac{1}{2}\right)$ telle que X et Z sont indépendantes — on ne demande pas dans la suite de justifier qu'un tel Z existe.
Montrer que $T = \frac{X}{Z+1}$ est une variable aléatoire à densité que l'on déterminera.

Exercice ALEA.14.11 | Loi normale alternée Soit $X \mapsto \mathcal{N}(0, 1)$ et ε une variable aléatoire discrète indépendante de X telle que $\varepsilon(\Omega) = \{\pm 1\}$ avec

$$\mathbf{P}(\varepsilon = 1) = \frac{1}{2} = \mathbf{P}(\varepsilon = -1).$$

1. Déterminer la loi de $Y = \varepsilon X$.
2. On pose $Z = Y - 2X$. Justifier que Z admet une espérance et une variance. Les déterminer.

Solution (exercice ALEA.14.11)

1. Tout d'abord, $\mathbf{P}(Y \in \mathbf{R}) = 1$ et *a priori* on ne peut pas réduire l'ensemble des valeurs prises. Pour calculer la fonction de répartition, appliquons la formule des probabilités totales à $\{\varepsilon = \pm 1\}$ qui est un système complet d'événements. Soit $t \in \mathbf{R}$,

$$\begin{aligned} \mathbf{P}(Y \leq t) &= \mathbf{P}(1.X \leq t | \varepsilon = 1) \mathbf{P}(\varepsilon = 1) + \mathbf{P}(-1.X \leq t | \varepsilon = -1) \mathbf{P}(\varepsilon = -1) \\ &= \Phi(t) \frac{1}{2} + (1 - \Phi(-t)) \frac{1}{2} \\ &= \Phi(t). \end{aligned} \quad \left. \vphantom{\begin{aligned} \mathbf{P}(Y \leq t) &= \mathbf{P}(1.X \leq t | \varepsilon = 1) \mathbf{P}(\varepsilon = 1) + \mathbf{P}(-1.X \leq t | \varepsilon = -1) \mathbf{P}(\varepsilon = -1) \\ &= \Phi(t) \frac{1}{2} + (1 - \Phi(-t)) \frac{1}{2} \\ &= \Phi(t). \end{aligned}} \right\} \text{d'après le cours } \Phi(-t) = 1 - \Phi(t)$$

Donc Y a même loi que X, i.e. $Y \hookrightarrow \mathcal{N}(0, 1)$.

2. La variable aléatoire $Z = Y - 2X$ admet une espérance car Y et X en admette une, et

$$\mathbf{E}(Z) = \mathbf{E}(Y) - 2\mathbf{E}(X) = \boxed{0}.$$

Pour la variance, puisque Y, X ne sont clairement pas indépendantes, écrivons plutôt :

$$Z = (\varepsilon - 2)X \implies Z^2 = (\varepsilon - 2)^2 X^2.$$

Ainsi, par indépendance, on déduit

$$\mathbf{E}(Z^2) = \mathbf{E}((\varepsilon - 2)^2) \mathbf{E}(X^2),$$

mais $X \hookrightarrow \mathcal{N}(0, 1)$, donc $\mathbf{E}(X^2) = \mathbf{Var}(X) + \mathbf{E}(X)^2 = 1$. Alors que

$$\mathbf{E}((\varepsilon - 2)^2) = \mathbf{Var}(\varepsilon - 2) + \mathbf{E}(\varepsilon - 2)^2 = \mathbf{Var}(\varepsilon) + (-2)^2 = \frac{2 \times 4}{12} = \frac{2}{3}.$$

Donc Z admet une variance, et

$$\mathbf{Var}(Z) = \frac{2}{3} - 0^2 = \boxed{\frac{2}{3}}.$$

4.5. Autour des min et max

Exercice ALEA.14.12 | Soient X_1, X_2 deux variables aléatoires i.i.d. de lois respectives $\mathcal{E}(\lambda_1)$ et $\mathcal{E}(\lambda_2)$. On pose $Y = \min(X_1, X_2)$.


1. Montrer que Y est à densité, et en déterminer une. Reconnait-on une loi classique?
2. Deux guichets sont ouverts dans une banque, le temps de service au premier guichet (resp. au deuxième) suit une loi exponentielle de moyenne 20 min. (resp. 30 min.). Deux clients rentrent simultanément, l'un choisit le guichet 1, et l'autre le 2. En moyenne, après combien de temps sort le premier?

3. Même question mais avec le dernier.

Exercice ALEA.14.13 | **Loi de GUMBEL. Max de lois exponentielles.** Soit F la fonction définie sur \mathbf{R} par $F(x) = e^{-e^{-x}}$ pour tout $x \in \mathbf{R}$.

1. Montrer que F est la fonction de répartition d'une variable aléatoire à densité X. On dira que X suit une loi de GUMBEL.
2. Déterminer une densité f de X.
3. Soit (X_n) une suite de variables aléatoires mutuellement indépendantes de loi $\mathcal{E}(1)$. On note $M_n = \max(X_1, \dots, X_n)$.
 - 3.1) Étudier pour $x \in \mathbf{R}$,

$$\lim_{n \rightarrow \infty} \mathbf{P}(M_n - \ln(n) \leq x).$$

- 3.2)  En déduire une fonction Python permettant de renvoyer une simulation de X. On pourra se servir ici de la commande `np.random.exponential(1)` qui permet de simuler une exponentielle de paramètre 1.

Solution (exercice ALEA.14.13)

1. Il y a deux choses à montrer ici : l'existence de la variable aléatoire, puis le caractère à densité. Pour l'existence, constatons que F est continue (donc continue à droite), $\lim_{x \rightarrow \infty} F(x) = 1$, $\lim_{x \rightarrow -\infty} F(x) = 0$ puisque $e^{-x} \xrightarrow{x \rightarrow \infty} 0$, $e^{-x} \xrightarrow{x \rightarrow -\infty} \infty$. Enfin, elle est croissante puisque $x \mapsto -e^{-x}$ l'est et que l'exponentielle aussi. Donc il existe X et $(\Omega, \mathcal{F}, \mathbf{P})$, $x : \Omega \rightarrow \mathbf{R}$ de sorte que $F_X = F$. Enfin, X est à densité puisque F est continue comme déjà dit, et F est \mathcal{C}^1 , X est donc à densité.
2. Et une densité est donnée par la dérivée de la fonction F :

$$f_X : x \mapsto e^{-x} e^{-e^{-x}} = e^{-x - e^{-x}}.$$

3. 3.1) On calcule la probabilité demandée.


$$\begin{aligned}
 \mathbf{P}(M_n - \ln(n) \leq x) &= \mathbf{P}(M_n \leq x + \ln n) \\
 &= \mathbf{P}(X_1 \leq x + \ln n) \cdots \mathbf{P}(X_n \leq x + \ln n) \quad \left. \begin{array}{l} \text{indépendance} \\ x + \ln n \geq 0 \text{ au} \\ \text{moins pour } n \text{ as-} \\ \text{sez grand} \end{array} \right\} \\
 &= \left(1 - e^{-x - \ln n}\right)^n \\
 &= \left(1 - \frac{e^{-x}}{n}\right)^n
 \end{aligned}$$

Il s'agit ensuite de trouver la limite de cette suite réelle. On commence déjà par passer à la forme exponentielle.

$$\begin{aligned}
 \left(1 - \frac{e^{-x}}{n}\right)^n &= \exp\left(n \ln\left(1 - \frac{e^{-x}}{n}\right)\right) \\
 &\underset{n \rightarrow \infty}{=} \exp\left(n \left(-\frac{e^{-x}}{n} + o\left(\frac{1}{n}\right)\right)\right) \\
 &\underset{n \rightarrow \infty}{\rightarrow} e^{-e^{-x}}.
 \end{aligned}$$

Donc :

$$\forall x \in \mathbf{R}, \quad \mathbf{P}(M_n - \ln(n) \leq x) \xrightarrow{n \rightarrow \infty} F(x).$$

3.2)  Ainsi, pour simuler X , on peut simuler le maximum précédent et lui retrancher $\ln(n)$.

```

import numpy as np

def simu_gumbel():
    n = 10**3
    Y = [np.random.exponential(1) for _ in range(n)]
    # Recherche du max
    maxi = Y[0]
    for i in range(1, n):
        if Y[i] > maxi:
            maxi = Y[i]
    return maxi - np.log(n)

```

Voici une simulation.

```

>>> simu_gumbel()
-1.4678252490298291

```



Exercice ALEA.14.14 | Max d'un nombre aléatoire de variables Soit N une variable aléatoire suivant la loi binomiale de paramètres n et p avec $n \in \mathbf{N}^*$, $p \in]0, 1[$. Soient X_0, X_1, \dots, X_n une collection de $n + 1$ variables aléatoires de loi $\mathcal{U}([0, 1])$. On suppose que X_0, X_1, \dots, X_n, N sont mutuellement indépendantes.

- Soit $k \in \{0, \dots, n\}$. Déterminer la fonction de répartition de $T_k = \min(X_0, X_1, \dots, X_k)$.
- On pose $T = \min(X_0, X_1, \dots, X_N)$.
 - Montrer que T est bien une variable aléatoire sur le même espace probabilisé sur lequel sont définies X_0, X_1, \dots, X_n, N .
 - Montrer que T est une variable à densité et déterminer une densité de T .

Exercice ALEA.14.15 | Max de lois normales On considère deux variables aléatoires X et Y indépendantes et suivant toutes deux la loi normale centrée réduite (de densité notée φ et de fonction de répartition notée Φ). On pose

$$Z = \max(X, Y)$$

et l'on se propose de déterminer la loi de Z , ainsi que son espérance et sa variance.

- Montrer que Z est une variable aléatoire à densité. On exprimera sa densité f en fonction de φ et Φ .
- En remarquant que, pour tout $x \in \mathbf{R}$, $\varphi'(x) = -x\varphi(x)$, montrer que Z a une espérance et donner sa valeur. *Indication : On pourra effectuer une intégration par parties.*

Solution (exercice ALEA.14.15)

1. Soit $x \in \mathbf{R}$.

$$\begin{aligned} \mathbf{P}(Z \leq x) &= \mathbf{P}(X \leq x, Y \leq y) \\ &= \Phi(x)^2. \end{aligned} \quad \left. \vphantom{\mathbf{P}(Z \leq x)} \right\} \text{indépendance}$$

Donc par composition, Φ^2 est $\mathcal{C}^0, \mathcal{C}^1$ donc Z est à densité et une densité est donnée par :

$$f : x \in \mathbf{R} \longrightarrow 2\Phi(x)\varphi(x).$$

2. ▶ Soit $x \in \mathbf{R}$, alors

$$\begin{aligned} \varphi'(x) &= \frac{d^2}{dx^2} \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \right) \\ &= \frac{-x}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \\ &= -x\varphi(x). \end{aligned}$$

▶ La variable aléatoire Z admet une espérance si et seulement si :

$$\int_{-\infty}^{\infty} |x| (2\Phi(x)\varphi(x)) dx \quad \text{converge.}$$

Commençons par étudier l'intégrale sur \mathbf{R}^+ , en effectuant une intégration par parties. Soit $A > 0$, alors puisque $-\varphi, \Phi$ sont \mathcal{C}^1 :

$$\begin{aligned} 2 \int_0^A \Phi(x) (x\varphi(x)) dx &= \int_0^A (2\varphi(x)\varphi(x)) dx + 2[-\varphi(x)\Phi(x)]_0^A \\ &= 2 \int_0^A \varphi^2(x) dx + 2A\varphi(A)\Phi(A) \quad (\star) \end{aligned}$$

Or $\varphi^2(x) = \frac{1}{2\pi} e^{-x^2}$ dx. Par changement de variable $x = \frac{u}{\sqrt{2}}$ (licite car \mathcal{C}^1 strictement croissant), $\int_0^{\infty} e^{-x^2} dx$ a même nature que

$$\sqrt{2} \int_0^{\infty} e^{-u^2/2} du = \sqrt{2} \frac{\sqrt{2\pi}}{2} = \sqrt{\pi}.$$

Donc en en faisant $A \rightarrow \infty$ dans (\star) (la limite du deuxième terme venant des croissances comparées), puisque Φ est une fonction bornée, on obtient alors :

$$2 \int_0^{\infty} \Phi(x) (x\varphi(x)) dx = 2 \frac{1}{2\pi} \sqrt{\pi} + 0 = \frac{1}{\sqrt{\pi}}.$$

De manière analogue : soit $B < 0$, alors

$$\begin{aligned} 2 \int_B^0 \Phi(x) (x\varphi(x)) dx &= \int_B^0 (2\varphi(x)\varphi(x)) dx + 2[-\varphi(x)\Phi(x)]_B^0 \\ &= 2 \int_B^0 \varphi^2(x) dx + 2B\varphi(B)\Phi(B) \\ &\xrightarrow{B \rightarrow -\infty} 2 \int_{-\infty}^0 \varphi^2(x) dx = \frac{1}{\sqrt{\pi}}. \end{aligned}$$

Donc Z admet une espérance, et

$$\mathbf{E}(Z) = \frac{2}{\sqrt{\pi}}.$$

Chapitre ALEA.15.

Couples aléatoires

Résumé & Plan


L'objectif de ce chapitre est d'étudier les couples aléatoires, *i.e.* les couples de deux variables aléatoires, principalement discrètes d'une part et à densité d'autre part. Nous verrons alors comment calculer la loi de sommes de variables aléatoires indépendantes.

1	Couples aléatoires	2	3	Sommes de variables aléatoires indépendantes & convolution	21
1.1	Généralités	2	3.1	Cas discret	21
1.2	Fonction de répartition & Loi	3	3.2	Cas à densité	24
2	Couples aléatoires discrets	4	4	Exercices	28
2.1	Système complet associé	4	4.1	Calculs déterministes	28
2.2	Loi marginale, conjointe, conditionnelle	5	4.2	Généralités	28
2.3	Espérance, Covariance	14	4.3	Étude de lois	30
			4.4	Somme de variables aléatoires & Stabilité de lois	33

Un être humain possède environ 150 000 cheveux (en tout cas, moins d'un million). Comme la ville de Paris compte 2,141 millions d'habitants, d'après le principe des tiroirs de DIRICHLET, il existe au moins deux personnes à Paris qui ont exactement le même nombre de cheveux.

— Le saviez-vous ?

 **Cadre**
 Dans tout le chapitre, et même lorsque cela n'est pas précisé, on se fixe un espace probabilisé $(\Omega, \mathcal{T}, \mathbf{P})$.

$X(b) = 1, X(c) = 2$ et $Y(a) = Y(c) = 3, Y(b) = 1$. On note $Z = (X, Y)$. Déterminer $X(\Omega)$, $X(\Omega) \times Y(\Omega)$ et $Z(\Omega)$. Le couple Z est un couple aléatoire discret. 

1. COUPLES ALÉATOIRES

1.1. Généralités

Définition ALEA.15.1 | Couple aléatoire

On appelle *couple aléatoire* toute application $Z : \Omega \rightarrow \mathbf{R}^2$ de la forme

$$Z : \Omega \rightarrow (X(\omega), Y(\omega)),$$

où X, Y sont deux variables aléatoires réelles. On note $Z = (X, Y)$ cette application.

On dit que :

- ▶ Z est un *couple aléatoire discret* si X, Y sont discrètes,
- ▶ Z est un *couple aléatoire à densité* si X, Y sont à densité.

Notez bien que la notion de couple aléatoire dépend, comme pour la notion de variable aléatoire, de la tribu sous-jacente \mathcal{T} .

Proposition ALEA.15.1 | À propos de l'univers-image

Soit $Z = (X, Y)$ un couple aléatoire. Alors

$$Z(\Omega) = \{(X(\omega), Y(\omega)), \omega \in \Omega\} \subset \{(X(\omega_1), Y(\omega_2)), (\omega_1, \omega_2) \in \Omega^2 \times \Omega\} = X(\Omega) \times Y(\Omega).$$

Voyons deux exemples dans le cas discret puis un exemple à densité.

Exemple 1 — Notons $\Omega = \{a, b, c\}$, et X, Y deux variables aléatoires telles que $X(a) =$

Exemple 2 — On lance deux pièces. On appelle X le nombre de «pile» obtenus et on pose $Y = 1$ si on obtient au moins une fois face, et 0 sinon. Les quantités X et Y sont des variables aléatoires et le couple (X, Y) forme un vecteur aléatoire. De manière formelle, on considère

$$\Omega = \{PP, PF, FP, FF\},$$

muni de la probabilité uniforme si l'expérience se déroule avec des pièces équilibrées,

et X et Y sont définies par

$$X(PP) = 2, \quad Y(PP) = 0, \quad X(PF) = 1, \quad Y(PF) = 1,$$

$$X(FP) = 1, \quad Y(FP) = 1, \quad X(FF) = 0, \quad Y(FF) = 1.$$

Le vecteur (X, Y) est lui défini par

$$(X, Y)(PP) = (X(PP), Y(PP)) = (2, 0), \quad (X, Y)(PF) = (1, 1),$$

$$(X, Y)(FP) = (1, 1), \quad (X, Y)(FF) = (0, 1).$$

Exemple 3 – à densité On observe deux bactéries et on s'intéresse à leurs durées de vie X, Y qui suivent une loi exponentielle de paramètre $\lambda > 0$, supposées indépendantes. Alors (X, Y) est un couple aléatoire à densité, mais c'est le cas aussi de $(X, \max(X, Y))$ puisque nous avons montré dans le [Chapter ALEA.14](#) que $\max(X, Y)$ est à densité.

1.2. Fonction de répartition & Loi

Définition ALEA.15.2 | Fonction de répartition

On appelle *fonction de répartition* du couple aléatoire $Z = (X, Y)$ l'application $F_Z : \mathbf{R}^2 \rightarrow [0, 1]$ donnée, pour $(x, y) \in \mathbf{R}^2$, par

$$F_Z(x, y) = \mathbf{P}(X \leq x, Y \leq y).$$

La fonction de répartition d'un couple aléatoire est donc une fonction de deux variables.

Définition/Proposition ALEA.15.1 | Loi

On appelle *loi* du couple aléatoire $Z = (X, Y)$, la fonction \mathbf{P}_Z qui à un pavé $I \times J \subset \mathbf{R}^2$

(I, J sont des intervalles) associe $\mathbf{P}_Z(I \times J)$ définie par :

$$\mathbf{P}_Z(I \times J) = \mathbf{P}(X \in I, Y \in J).$$

Proposition ALEA.15.2 | La loi ne dit rien sur l'égalité

Soient Z, Z' deux couples aléatoires.

1. $Z = Z' \implies \mathbf{P}_Z = \mathbf{P}_{Z'}$.

2. La réciproque est fautive : si $\mathbf{P}_Z = \mathbf{P}_{Z'}$ alors Z n'est pas forcément égale à Z' .

Preuve

1. L'hypothèse nous dit que $X = X', Y = Y'$. Soient I, J deux intervalles réels. Alors

$$\begin{aligned} \mathbf{P}_Z(I \times J) &= \mathbf{P}(Z \in I \times J) \\ &= \mathbf{P}(X \in I, Y \in J) \\ &= \mathbf{P}(X' \in I, Y' \in J) = \mathbf{P}(Z' \in I \times J) \quad \downarrow X = X', Y = Y' \\ &= \mathbf{P}_{Z'}(I \times J). \end{aligned}$$

2. **est faux.** On lance une pièce équilibrée. On note X la variable aléatoire égale à 1 si on obtient pile, égale à 0 si on obtient face. On note Y la variable aléatoire égale à 1 si on obtient face, égale à 0 si on obtient pile. Alors (X, Y) et (Y, X) ont même loi et pourtant $(X, Y) \neq (Y, X)$ en tant qu'applications.

Lorsque l'univers est discret, il n'existe que des couples aléatoires discrets.

Proposition ALEA.15.3 | Vecteur aléatoire défini sur un univers au plus dénombrable

Soit $Z = (X, Y)$ un couple aléatoire avec Ω au plus dénombrable. Alors :

Z est un couple aléatoire discret.

Preuve Puisque Ω est dénombrable, d'après le [Chapter ALEA.13](#), X et Y sont discrètes. Donc Z est un couple aléatoire discret.

2.1. Système complet associé

Proposition ALEA.15.4 | Structure d'espace vectoriel, opérations

1. L'ensemble des couples aléatoires réels muni de l'addition et de la multiplication externe des applications est un \mathbf{R} -espace vectoriel, i.e. : si Z_1, Z_2 sont deux couples aléatoires et $(\lambda, \mu) \in \mathbf{R}^2$, alors :

$$\lambda Z_1 + \mu Z_2 \text{ est un vecteur aléatoire réel.}$$

2. Si $Z_1 = (X_1, Y_1), Z_2 = (X_2, Y_2)$ sont deux couples aléatoires, alors :

$$\langle Z_1 | Z_2 \rangle = X_1 X_2 + Y_1 Y_2$$

est une variable aléatoire réelle.

Attention

Dans la proposition précédente, il s'agit d'additions dans \mathbf{R}^2 , donc de couples.

Preuve

- ▶ $\lambda Z_1 + \mu Z_2 = (\lambda X_1 + \mu X_2, \lambda Y_1 + \mu Y_2)$ et $\lambda X_1 + \mu X_2, \lambda Y_1 + \mu Y_2$ sont des variables aléatoires réelles d'après le [Chapter ALEA.12](#).
- ▶ Car l'ensemble des variables aléatoires réelles est stable par somme et produit.

2. COUPLES ALÉATOIRES DISCRETS

Dans cette section nous ne considérerons que des couples aléatoires discrets du type $Z = (X, Y)$. On rappelle que cela signifie que $X(\Omega), Y(\Omega)$ sont des ensembles au plus dénombrables. Du fait de la présence de plusieurs variables aléatoires, de nouvelles notions de loi apparaissent, en plus de la notion classique : l'ancienne notion a son analogue pour les couples : la «loi conjointe».

Cadre

On se fixe donc dans toute la suite (X, Y) un couple aléatoire discret.

Notation Somme double infinie sur un ensemble dénombrable

Considérons E, F deux ensembles au plus dénombrable et $f : E \times F \rightarrow \mathbf{R}$, on souhaite donner un sens à la quantité $\sum_{(x,y) \in E \times F} f(x, y)$.

- ▶ Si E, F sont finis, alors $E = \{x_n, n \in \llbracket 0, N_1 \rrbracket\}$, $F = \{y_n, n \in \llbracket 0, N_2 \rrbracket\}$ avec $N_1, N_2 \in \mathbf{N}$. Alors on pose :

$$\sum_{(x,y) \in E \times F} f(x, y) \stackrel{\text{(défi.)}}{=} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} f(x_i, y_j) = \sum_{j=1}^{N_2} \sum_{i=1}^{N_1} f(x_i, y_j).$$

C'est une somme double finie de première année, rien de plus à dire puisqu'elle ne dépend pas de la numérotation des éléments de E, F .

- ▶ Si E, F sont dénombrables, alors $E = \{x_n, n \in \mathbf{N}\}$, $F = \{y_n, n \in \mathbf{N}\}$. Alors, si les hypothèses de FUBINI sont vérifiées :

— pour tout $m \in \mathbf{N}$, $(\sum f(x_n, y_m))_{n \geq 0}$ converge absolument, et

— si $(\sum_{n=0}^{\infty} \sum_{m \geq 0} f(x_n, y_m))_{m \geq 0}$ converge absolument, alors

on pose :

$$\sum_{(x,y) \in E \times F} f(x, y) \stackrel{\text{(défi.)}}{=} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} f(x_n, y_m) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} f(x_n, y_m).$$


En effet, on sait depuis le [Chapter ANA.10](#), que les sommes infinies précédentes ne dépendront pas de la numérotation des éléments de E et F , et qu'en plus les hypothèses du théorème de FUBINI seront vérifiées donc que l'on peut échanger l'ordre des sommes infinies.

Définition/Proposition ALEA.15.2 | Système d'évènements associé à un couple aléatoire discret

Si $Z = (X, Y)$ est un couple aléatoire discret, alors :

- ▶ $\{X = x, Y = y\}_{(x,y) \in X(\Omega) \times Y(\Omega)}$ est un système complet d'évènements.
- ▶ En particulier, il est quasi-complet :
 - $\sum_{(x,y) \in X(\Omega) \times Y(\Omega)} \mathbf{P}(X = x, Y = y) = 1,$
 - $\{(X, Y) = (x, y), (X, Y) = (x', y')\} = \emptyset, \quad \forall (x, y) \neq (x', y').$

On l'appelle le *système complet associé* à (X, Y) .

Preuve Ne nous fatiguons par trop et utilisons le résultat analogue vu dans le **Chapter ALEA.13.** 

2.2. Loi marginale, conjointe, conditionnelle

Rappelons que nous avons défini la loi de (X, Y) comme l'application qui à tout tout pavé réel $I \times J$ associe

$$\mathbf{P}_X(I \times J) = \mathbf{P}(X \in I, Y \in J).$$

Comment simplifier cette définition dans le cas de couples aléatoires discrets? C'est la même démarche que pour les variables aléatoires discrètes.

$$\{X \in I, Y \in J\} = \{(X, Y) \in I \times J, (X, Y) \in (X, Y)(\Omega)\} \quad \text{donc,}$$

$$\mathbf{P}(X \in I, Y \in J) = \mathbf{P}((X, Y) \in I \times J, (X, Y) \in (X, Y)(\Omega)).$$

Or,

$$\begin{aligned} \{(X, Y) \in I \times J, (X, Y) \in (X, Y)(\Omega)\} &= \bigsqcup_{\substack{(x,y) \in (X,Y)(\Omega) \\ x \in I, y \in J}} \{X = x, Y = y\}, \\ \implies \mathbf{P}((X, Y) \in I \times J) &= \sum_{\substack{(x,y) \in (X,Y)(\Omega) \\ x \in I, y \in J}} \mathbf{P}(X = x, Y = y). \end{aligned}$$

Ainsi, pour obtenir la loi, il suffit de connaître :

- ▶ d'une part l'univers-image $(X, Y)(\Omega)$,
- ▶ et d'autre part tous les $\mathbf{P}(X = x, Y = y)$ pour tout $(x, y) \in (X, Y)(\Omega)$.

Ce constat nous mène tout droit à la définition suivante.

Définition ALEA.15.3 | Loi conjointe. Lois marginales.

Soit $Z = (X, Y)$ un couple aléatoire discret.

- ▶ **(Loi conjointe)** On appelle *loi conjointe* du couple aléatoire discret (X, Y) ou *fonction de masse* la fonction encore notée $\mathbf{P}_{(X,Y)}$ et définie par :

$$\mathbf{P}_{(X,Y)} : (x, y) \in (X, Y)(\Omega) \longrightarrow \mathbf{P}(X = x, Y = y).$$

- ▶ **(Déterminer la loi conjointe)** Déterminer la loi conjointe d'un couple aléatoire discret c'est calculer $(X, Y)(\Omega)$ et $\mathbf{P}(X = x, Y = y)$ pour tout $x \in X(\Omega), y \in Y(\Omega)$.
- ▶ **(Marginales)**
 - La loi $\mathbf{P}_X : x \in X(\Omega) \longrightarrow \mathbf{P}(X = x)$ de X est appelée *première loi marginale*.
 - La loi $\mathbf{P}_Y : y \in Y(\Omega) \longrightarrow \mathbf{P}(Y = y)$ de Y est appelée *seconde loi marginale*.
- ▶ **(Avoir même loi que)** Soit $Z = (X', Y')$ un couple aléatoire discret définie sur $(\Omega, \mathcal{F}, \mathbf{P})$. On dit que $Z = Z'$ ont même loi si :


$$\forall (a, b) \in Z(\Omega) \cup Z'(\Omega)^1, \quad \mathbf{P}(Z = (a, b)) = \mathbf{P}(Z' = (a, b)).$$

¹la plupart du temps, dans des contextes d'expériences aléatoires, les ensembles $Z(\Omega), Z'(\Omega)$ seront égaux


 **Notation**


Lorsque $Z = (X, Y), Z' = (X', Y')$ ont même loi, on note $(X, Y) \sim (X', Y')$ ou $Z \sim Z'$.


Méthode Répondre à la question « déterminer la loi conjointe du vecteur aléatoire X, Y »

-  1. Commencer par déterminer son support $(X, Y)(\Omega)$ s'il n'est pas déjà donné.
2. Calculer les $\mathbf{P}(X = x, Y = y)$ pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$. Si $X(\Omega), Y(\Omega)$ sont finis, il n'y a donc qu'un nombre fini de probabilités à déterminer, on peut les présenter sous forme d'un tableau comme ceci :

$Y = \ell / X = k$	k_1	k_2	...
ℓ_1	$\mathbf{P}(Y = \ell_1, X = k_1)$	$\mathbf{P}(Y = \ell_1, X = k_2)$...
\vdots			\vdots

Exemple 4 — Min / Max et lancer de dés On lance deux dés. On note D_1, D_2 les valeurs obtenues et on pose $X = \min(D_1, D_2), Y = \max(D_1, D_2)$. Déterminer la loi conjointe de (X, Y) . 

Exemple 5 — On considère une urne avec quatre boules numérotées de 1 à 4 et les tirages de deux boules successifs avec remise. On note X le numéro de la boule au 1er tirage et Y le numéro de la boule au 2ème tirage. On pose enfin $M = \max(X, Y)$ et $Z = (X, M)$. Déterminer la loi conjointe de Z . 

Exemple 6 — Égalité de variables aléatoires discrètes Soient X, Y deux variables aléatoires discrètes indépendantes de loi $\mathcal{G}(p)$ avec $p \in]0, 1[$. Exprimer $\mathbf{P}(X = Y)^2$ à l'aide de la loi conjointe de (X, Y) , puis calculer la somme. 

Généralisons les calculs faits dans le dernier exemple, dans une méthode.

Méthode Déterminer des probabilités d'évènements du type $\{X = Y\}, \dots, \{X_1 = \dots = X_n\}$



Par exemple, si X, Y sont deux variables aléatoires discrètes, on souhaite calculer $\mathbf{P}(X = Y)$.

1. Introduire le système complet associé à X (ou Y) dans l'évènement $\{X = Y\}$:

$$\{X = Y\} = \bigcup_{k \in X(\Omega)} \{X = k, Y = k\}.$$

2. On passe ensuite aux probabilités :

$$\mathbf{P}(X = Y) = \sum_{k \in X(\Omega)} \mathbf{P}(X = k, Y = k).$$

En cas d'indépendance, on écrit alors

$$\mathbf{P}(X = Y) = \sum_{k \in X(\Omega)} \mathbf{P}(X = k) \mathbf{P}(Y = k).$$

De manière générale, pour $\mathbf{P}(X_1 = \dots = X_n)$, on introduit le système complet d'évènements associé $n - 1$ variables aléatoires parmi les n .

Nous avons défini la notion de loi conditionnelle d'une variable aléatoire discrète X sachant un évènement non négligeable A comme l'application

$$x \in X(\Omega) \longmapsto \mathbf{P}(X = x|A).$$

²Au passage, les calculs faits ici prouvent que $\{X = Y\}$ est un évènement : on l'a écrit comme une réunion dénombrable d'évènements.

Nous pouvons en particulier regarder la loi conditionnelle de X sachant l'évènement $\{Y = y\}$ pour $y \in Y(\Omega)$ (ou l'inverse) si (X, Y) un couple aléatoire discret.

Définition ALEA.15.4 | Loi conditionnelle

1. Soit $y \in Y(\Omega)$ tel que $\mathbf{P}(Y = y) \neq 0$. Alors on appelle *loi conditionnelle de X sachant $\{Y = y\}$* l'application

$$\mathbf{P}(\cdot | Y = y) \left| \begin{array}{l} X(\Omega) \longrightarrow [0, 1], \\ x \longrightarrow \mathbf{P}(X = x | Y = y). \end{array} \right.$$

2. Soit $x \in X(\Omega)$ tel que $\mathbf{P}(X = x) \neq 0$. Alors on appelle *loi conditionnelle de Y sachant $\{X = x\}$* l'application

$$\mathbf{P}(\cdot | X = x) \left| \begin{array}{l} Y(\Omega) \longrightarrow [0, 1], \\ y \longrightarrow \mathbf{P}(Y = y | X = x). \end{array} \right.$$

LIENS ENTRE LES NOTIONS DE LOI. Les formules des probabilités totales et celle de BAYES permettent de faire le lien entre ces différentes notions de lois, voyons comment. Il n'y a donc, dans ce paragraphe, rien de neuf à apprendre.

Proposition ALEA.15.5 | Loi conjointe \rightarrow Loi marginale

Les lois marginales de (X, Y) sont données, en fonction de la loi conjointe, par les formules suivantes :

$$\forall x \in X(\Omega), \quad \mathbf{P}_X(x) = \mathbf{P}(X = x) = \sum_{y \in Y(\Omega)} \mathbf{P}(X = x, Y = y),$$


$$\forall y \in Y(\Omega), \quad \mathbf{P}_Y(y) = \mathbf{P}(Y = y) = \sum_{x \in X(\Omega)} \mathbf{P}(X = x, Y = y).$$

Preuve (Point clef — Formule des probabilités totales)




Attention La réciproque est fautive. En effet, si X et Y sont telles que $X(\Omega) = Y(\Omega) = \{0, 1\}$, on peut définir deux lois de couple différentes³ par :

$$\forall (x, y) \in \{0, 1\}^2, \quad \mathbf{P}(X = x, Y = y) = \frac{1}{4}.$$

On constate que $X, Y \hookrightarrow \mathcal{U}\{0, 1\}$. 

Et si

$$\forall (x, y) \in \{0, 1\}^2, \quad \mathbf{P}(X = x, Y = y) = \begin{cases} \frac{1}{2} & \text{si } (x, y) = (0, 0) \text{ ou } (1, 1), \\ 0 & \text{sinon,} \end{cases}$$

on constate là encore que $X, Y \hookrightarrow \mathcal{U}\{0, 1\}$. 

Remarque 2.1 — Si la loi conjointe est donnée sous forme d'un tableau, la formule se comprend de la manière suivante : pour calculer une loi marginale, il suffit de sommer les probabilités sur les lignes ou les colonnes du tableau.

³leur existence est garantie par le **Théorème ALEA.15.1** puisque $\frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = 1$ et $\frac{1}{2} + \frac{1}{2} = 1$

Proposition ALEA.15.6 | Lois maringales + indépendance → Loi conjointe

Si X et Y sont indépendantes, alors


$$\forall (x, y) \in (X, Y)(\Omega), \quad \mathbf{P}_Z(x, y) = \mathbf{P}_X(x) \cdot \mathbf{P}_Y(y).$$

Preuve



Exemple 7 — Cas fini Soit (X, Y) un couple aléatoire discret dont la loi marginale est donnée par le tableau ci-après :


X/Y	1	2	3	4
1	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
2	0	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$
3	0	0	$\frac{3}{16}$	$\frac{1}{16}$
4	0	0	0	$\frac{1}{4}$

Déterminer la les lois marginales de (X, Y) . 

Exemple 8 — Cas « semi- » dénombrable Soit $p \in]0, 1[$. Soient X et Y deux variables

aléatoires à valeurs dans \mathbf{N} telles que :

$$\forall (x, y) \in \mathbf{N}^2, \quad \mathbf{P}(X = x \cap Y = y) = \begin{cases} e^{-p} - p + pe^{-p} & \text{si } (x, y) = (0, 0), \\ p - pe^{-p} & \text{si } (x, y) = (0, 1), \\ pe^{-p} & \text{si } (x, y) = (1, 1), \\ \frac{p^n}{n!} e^{-p} & \text{si } (x, y) = (n, 0), \text{ et } n \geq 2, \\ 0 & \text{sinon.} \end{cases}$$

Déterminer les lois marginales de (X, Y) , on montrera que $X \hookrightarrow \mathcal{P}(p)$ et que $Y \hookrightarrow \mathcal{B}(p)$. 

Proposition ALEA.15.7 | Lien loi conditionnelle / marginale / conjointe

Soit (X, Y) un couple aléatoire discret. Alors pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$:

$$\mathbf{P}(X = x, Y = y) = \mathbf{P}(X = x | Y = y) \mathbf{P}(Y = y) = \mathbf{P}(Y = y | X = x) \mathbf{P}(X = x),$$

et pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$:

$$\mathbf{P}(X = x) = \sum_{y \in Y(\Omega)} \mathbf{P}(X = x | Y = y) \mathbf{P}(Y = y),$$

$$\mathbf{P}(Y = y) = \sum_{x \in X(\Omega)} \mathbf{P}(Y = y | X = x) \mathbf{P}(X = x).$$

Remarque 2.2 — Rappel d'une convention On rappelle la convention ci-après déjà précisée dans le [Chapter ALEA.12](#) : pour $(x, y) \in X(\Omega) \times Y(\Omega)$,


$$\mathbf{P}(X = x | Y = y) \mathbf{P}(Y = y) = 0, \quad \text{si } \mathbf{P}(Y = y) = 0.$$


Preuve (Point clef — *Formule des probabilités totales, définition d'une probabilité conditionnelle*)




Exemple 9 – Urnes choisies aléatoirement On considère $n + 1$ urnes numérotées $0, \dots, n$. On suppose que l'urne $i \in \llbracket 0, n \rrbracket$ contient i boules noires et $n + 1 - i$ boules blanches. On procède à l'expérience suivante : on choisit une urne suivant une loi $\mathcal{B}\left(n, \frac{1}{2}\right)$ puis on effectue le tirage d'une boule. On note N le numéro de l'urne choisie à la première étape, et X la variable qui vaut 1 si on a obtenu une boule noire, et 0 sinon.

- ▶ Déterminer la loi conditionnelle de X sachant $\{N = i\}$ pour tout $i \in \llbracket 0, n \rrbracket$. 

- ▶ Déterminer la loi conjointe du couple (N, X) . 

- ▶ Déterminer la loi de X . Quelle est la probabilité que la boule tirée soit noire? 

Exemple 10 – Pile ou Face jusqu'à deux piles Dans une succession de piles ou faces pour laquelle la probabilité d'obtenir pile est $p \in]0, 1[$ et d'obtenir face est $q = 1 - p$, on note X le rang d'apparition du premier pile et Y celui du second. Déterminer la loi conjointe de (X, Y) puis vérifier que le système quasi-complet d'événements associé à (X, Y) en est bien un.  On a

$$(X, Y)(\Omega) = \{(k, \ell) \in \mathbf{N}^* \times \mathbf{N}^*, \ell > k\}.$$

Nous avons clairement $X \hookrightarrow \mathcal{G}(p)$ et par ailleurs, la loi conditionnelle de Y sachant $\{X = k\}$ pour $k \in \mathbf{N}^*$ est donnée par l'expression suivante, puisque les lancers sont indépendants :


$$\forall \ell \in \llbracket k + 1, \infty \rrbracket, \quad \mathbf{P}(Y = \ell | X = k) = pq^{\ell - k - 1}.$$

On obtient alors pour tout couple $(k, \ell) \in (\mathbf{N}^*)^2$ tel que $\ell > k$:

$$\mathbf{P}(Y = \ell, X = k) = \mathbf{P}(Y = \ell | X = k) \mathbf{P}(X = k) = pq^{\ell - k - 1} \times pq^{k-1} = p^2 q^{\ell - 2}.$$

On vérifie ensuite que la somme des ces probabilités vaut bien un :

$$\sum_{k=1}^{\infty} \sum_{\ell=k+1}^{\infty} \mathbf{P}(Y = \ell, X = k) = \sum_{k=1}^{\infty} \sum_{\ell=k+1}^{\infty} p^2 q^{\ell - 2} = \left(\frac{p}{q}\right)^2 \sum_{k=1}^{\infty} \frac{q^{k+1}}{1 - q} = \frac{p}{q^2} q^2 \frac{1}{1 - q} = \boxed{1}.$$

Exemple 11 – Conditionnement POISSON / binomiale Soient $n \in \mathbf{N}^*$, $p \in]0, 1[$, X une variable aléatoire telle que $X \hookrightarrow \mathcal{P}(\lambda)$, et Y une variable aléatoire telle que la loi conditionnelle sachant $\{X = n\}$ est une loi $\mathcal{B}(n, p)$. Déterminer la loi de Y . 

PROPRIÉTÉS DES COUPLES ALÉATOIRES DISCRETS. Précisons, en plus de celles mentionnées dans la Proposition ALEA.15.4, quelles sont les propriétés vérifiées par les couples aléatoires discrets.

Proposition ALEA.15.8 | Structure d'espace vectoriel, opérations sur les variables aléatoires réelles discrètes

1. L'ensemble des couples aléatoires discrets, muni de l'addition et de la multiplication de réels, est un \mathbf{R} -espace vectoriel, *i.e.* si Z et Z' sont deux couples aléatoires discrets, $(\lambda, \mu) \in \mathbf{R}^2$, alors :

$$\lambda Z + \mu Z'$$

est un couple aléatoire discret.

2. **(Produit scalaire)** Si $Z = (X, Y), Z' = (X', Y')$ sont deux couples discrets, alors :

$$\langle (X, Y) | (X', Y') \rangle = XX' + YY'$$

est une variable aléatoire discrète.

3. **(Image d'un couple aléatoire par une application)** Si (X, Y) est un couple aléatoire discret et $f : (X, Y)(\Omega) \rightarrow \mathbf{R}$, alors

$f(X, Y) \begin{cases} \Omega \longrightarrow \mathbf{R} \\ \omega \longmapsto f(X(\omega), Y(\omega)) \end{cases}$ est une variable aléatoire réelle discrète.

Son univers-image est

$$f(X, Y)(\Omega) = f((X, Y)(\Omega))$$

et sa loi est donnée par :

$$\forall z \in f(X, Y)(\Omega), \quad \mathbf{P}(f(X, Y) = z) = \sum_{\substack{(x,y) \in X(\Omega) \times Y(\Omega) \\ z=f(x,y)}} \mathbf{P}(X = x, Y = y). \text{ [H.P]}$$

Dans la dernière formule, on utilise donc la loi conjointe de (X, Y) pour déterminer celle de $f(X, Y)$.

Preuve

1. 

2. Même preuve qu'en 1.

3. 

EXISTENCE D'UN COUPLE ALÉATOIRE DISCRET DE LOI CONJOINTE FIXÉE. On généralise sans difficulté le théorème d'existence de variable aléatoire discrète de loi fixée, vu dans le **Chapter ALEA.13**.

Théorème ALEA.15.1 | Existence d'un espace probabilisé associé à une suite double de somme un

Soit $\mathbf{X} = \{(x_i, y_j), i, j \in \mathbf{N}\}$ (resp. $\{(x_i, y_j), i \in \llbracket 0, N_1 \rrbracket, j \in \llbracket 0, N_2 \rrbracket\}$ avec $N_1, N_2 \in \mathbf{N}$) une partie au plus dénombrable de \mathbf{R}^2 , et $(p_{i,j})_{(i,j) \in \mathbf{N}^2}$ (resp. $(p_{i,j})_{(i,j) \in \llbracket 1, N_1 \rrbracket \times \llbracket 1, N_2 \rrbracket}$) une famille de réels telle que :

$$\sum_{i,j \in \mathbf{N}} p_{i,j} = 1 \quad \left(\text{resp.} \sum_{\substack{0 \leq i \leq N_1 \\ 0 \leq j \leq N_2}} p_{i,j} = 1 \right) \quad \text{et} \quad p_{i,j} \geq 0.$$

Alors il existe un espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$ et un couple aléatoire discret (X, Y) tel que :

- ▶ $(X, Y)(\Omega) = \mathbf{X}$.
- ▶ Pour tout (i, j) , $\mathbf{P}(X = x_i, Y = y_j) = p_{i,j}$.


La preuve est identique au cas des variables aléatoires discrètes.

Preuve

- ▶ On cherche un exemple, rien qu'un, d'espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$ et de couple aléatoire qui répondent au problème posé. Posons alors :

$$\begin{aligned} (\Omega, \mathcal{F}) &= (\mathbf{X}, \mathcal{P}(\mathbf{X})), \\ (X, Y) &= \text{Id}_{\Omega \times \Omega} = \text{Id}_{\mathbf{X} \times \mathbf{X}}, \end{aligned}$$

puis enfin pour \mathbf{P} l'unique probabilité telle que : $\mathbf{P}(\{(x_i, y_j)\}) = p_{i,j}$ pour tout $i, j \in \mathbf{N}$ (ou $\llbracket 1, N_1 \rrbracket \times \llbracket 1, N_2 \rrbracket$) — i.e. celle définie dans la **Section 2.4 du Chapter ALEA.12**. On vérifie sans difficulté que ce choix convient.

Exemple 12 — Soit $a \in \mathbf{R}$. Déterminer a pour qu'il existe un couple aléatoire discret (X, Y) tel que $\mathbf{P}(X = i, Y = j) = \frac{a}{2^{i+j}}$ pour tout $(i, j) \in (\mathbf{N}^*)^2$. 

- ▶ Soit $i \in \mathbb{N}^*$, alors étudions la convergence de $\left(\sum_{j \geq 1} \frac{1}{2^{i+j}}\right)$. Puisque $|1/2| < 1$, nous avons

$$\sum_{j=1}^{\infty} \frac{1}{2^{i+j}} = \frac{1}{2^i} \sum_{j=1}^{\infty} \frac{1}{2^j} = \frac{1}{2^i} \frac{1}{1-1/2} = \frac{1}{2^i}.$$

- ▶ De-même $\sum_{i=1}^{\infty} \frac{1}{2^i} = \frac{1}{2} \frac{1}{1-1/2} = 1$.

Les hypothèses du théorème de FUBINI sont donc vérifiées, et nous avons

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \frac{1}{2^{i+j}} = 1.$$

Donc l'unique valeur de a telle que

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \mathbf{P}(X = i, Y = j) = 1$$

est $\boxed{a = 1}$. De plus, pour tout $(i, j) \in (\mathbb{N}^*)^2$,

$$\frac{a}{2^{i+j}} \geq 0$$

d'où l'existence d'un couple aléatoire (X, Y) de loi conjointe donnée précédemment.

2.3. Espérance, Covariance

2.3.1. Espérance & Transfert

Définition ALEA.15.5 | Espérance d'un couple aléatoire discret

Soit $Z = (X, Y)$ un couple aléatoire discret.

- ▶ **(Admettre une espérance)** On dit que $Z = (X, Y)$ admet une espérance si X et Y admettent une espérance.
- ▶ Si X, Y admettent une espérance, alors on appelle *espérance de (X, Y)* le vec-

teur

$$\mathbf{E}(Z) = (\mathbf{E}(X), \mathbf{E}(Y)) \in \mathbf{R}^2.$$

L'espérance d'un vecteur aléatoire est donc aussi un vecteur : c'est le couple formé par les espérances.

FORMULE DE TRANSFERT POUR LES COUPLES ALÉATOIRES DISCRETS. L'objectif est ici d'obtenir une formule pour calculer des espérances de fonctions de variables aléatoires $f(X, Y)$ sans avoir à trouver la loi. Nous allons voir également que la formule de transfert pour les couples aléatoires va nous permettre de démontrer certaines propriétés sur les variables aléatoires réelles discrètes admises jusque là. Notamment, en l'appliquant à $f : (x, y) \rightarrow \lambda x + \mu y$ pour $(\lambda, \mu) \in \mathbf{R}^2$, nous pourrions montrer la linéarité de l'espérance.

Théorème ALEA.15.2 | Transfert pour les couples aléatoires discrets

Soit (X, Y) un couple aléatoire discret et $g : (X, Y)(\Omega) \rightarrow \mathbf{R}$. Alors :

$g(X, Y)$ possède une espérance

$$\iff \left(\sum_{(x,y) \in X(\Omega) \times Y(\Omega)} g(x, y) \mathbf{P}(X = x, Y = y)\right) \text{ converge } \underline{\text{absolument}}.$$

Dans ce cas, nous avons :

$$\begin{aligned} \mathbf{E}(g(X, Y)) &= \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} g(x, y) \times \mathbf{P}(X = x, Y = y) \\ &= \sum_{y \in Y(\Omega)} \sum_{x \in X(\Omega)} g(x, y) \times \mathbf{P}(X = x, Y = y). \end{aligned}$$

Preuve La permutation des sommes est justifiée par la convergence absolue et le **Théorème ANA.10.5**. Nous admettons le reste.

On peut donc enfin montrer la linéarité de l'espérance, dans le cas discret.

Corollaire ALEA.15.1 | Linéarité de l'espérance

Soient X, Y deux variables aléatoires discrètes admettant une espérance, et $\lambda, \mu \in \mathbf{R}$. Alors $\lambda X + \mu Y$ admet une espérance et :

$$\mathbf{E}(\lambda X + \mu Y) = \lambda \mathbf{E}(X) + \mu \mathbf{E}(Y).$$

Preuve



Corollaire ALEA.15.2 | Espérance d'un produit

▶ Si X_1, \dots, X_n est une collection de n variables aléatoires discrètes **indépendantes** admettant une espérance. Alors :

$$\mathbf{E}(X_1 \times \dots \times X_n) = \mathbf{E}(X_1) \times \dots \times \mathbf{E}(X_n).$$

▶ Soient X, Y deux variables aléatoires indépendantes discrètes ayant une espérance. Alors :

$$\mathbf{E}(X \cdot Y) = \mathbf{E}(X) \cdot \mathbf{E}(Y).$$


Attention La réciproque est fautive

✗ Soit $X \hookrightarrow \mathcal{U}(\{-1, 0, 1\})$ et $Y = X^2$. Alors $\mathbf{E}(X) = (-1) \times 31 + 0 \times 13 + 1 \times 13 = 0$ et de même, avec le théorème de transfert, $\mathbf{E}(XY) = \mathbf{E}(X^3) = 0$. Ainsi $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y) = 0$ et pourtant $Y = X^2$ n'est pas indépendante de X .

Preuve Nous démontrons le résultat pour $n = 2$: il s'étend alors automatiquement pour n quelconque par récurrence, notamment car $X_1 \times \dots \times X_{n-1} \perp\!\!\!\perp X_n$.

Exemple 13 – Cas fini On considère une urne contenant n boules numérotées de 1 à n . On tire successivement deux boules avec remise. On note X le numéro de la 1ère boule, Y le numéro de la 2ème. Soit $Z = \max(X, Y)$. Quelle est l'espérance de Z ?

$$\begin{aligned} \mathbf{E}(Z) &= \sum_{1 \leq i, j \leq n} \max(i, j) \mathbf{P}\left((X = x_i) \cap (Y = y_j)\right) = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \max(i, j) \\ &= \frac{1}{n^2} \sum_{i=1}^{n-1} \left[\sum_{j=1}^i \max(i, j) + \sum_{j=i+1}^n \max(i, j) \right] + \frac{1}{n^2} n^2 \quad \left. \begin{array}{l} \text{on sort de le terme} \\ i = n + \text{CHASLES} \end{array} \right\} \\ &= \frac{1}{n^2} \left(\sum_{i=1}^{n-1} \left[\sum_{j=1}^i i + \sum_{j=i+1}^n j \right] + n^2 \right) \\ &= \frac{1}{n^2} \sum_{i=1}^n \left[i^2 + \frac{n(n+1)}{2} - \frac{i(i+1)}{2} \right] \quad \left. \begin{array}{l} n^2 \text{ est le terme } i = n \text{ de la somme} \end{array} \right\} \\ &= \frac{1}{2n^2} \sum_{i=1}^n \left[n(n+1) + i^2 - i \right] \\ &= \frac{1}{2n^2} \left(n^2(n+1) + \frac{n(n+1)(2n+1)}{6} - \frac{n(n+1)}{2} \right) \\ &= \frac{(n+1)(4n-1)}{6n}. \end{aligned}$$

Exemple 14 — Cas «semi-»dénombrable Soient X, Y deux variables aléatoires indépendantes telles que $X \hookrightarrow \mathcal{G}(p)$, avec $p \in]0, 1[$ et $Y \hookrightarrow \mathcal{B}(p)$. Déterminer si $Z = \min(X, Y + 1)$ admet une espérance, la calculer le cas échéant.  Il s'agit de calculer $E(g(X, Y))$ si $g : (x, y) \in \mathbf{R}^2 \longrightarrow \min(x, y + 1)$. Étudions donc la convergence absolue de

$$\left(\sum_{k \geq 1, \ell \in \{0, 1\}} \min(k, \ell + 1) \mathbf{P}(X = k, Y = \ell) \right)$$

Soit $\ell \in \{0, 1\}$. Il s'agit donc d'étudier la convergence absolue de la série en k , pour ces deux valeurs de ℓ .

► Supposons que $\ell = 0$. Soit $k \geq 1$. Alors par indépendance

$$\min(k, \ell + 1) \mathbf{P}(X = k, Y = \ell) = q \min(k, 1) \mathbf{P}(X = k) = q \mathbf{P}(X = k)$$

car $k \geq 1$. Comme $\sum_{k=1}^{\infty} \mathbf{P}(X = k) = 1$, nous déduisons la convergence souhaitée.

► Supposons que $\ell = 1$. Alors par indépendance

$$\min(k, \ell + 1) \mathbf{P}(X = k, Y = \ell) = p \min(k, 2) (p q^{k-1}) = p^2 \min(k, 2) q^{k-1}.$$

Comme $|q| < 1$, on a alors

$$\sum_{k=1}^{\infty} p^2 \min(k, 2) q^{k-1} = p^2 + p^2 \sum_{k=2}^{\infty} k q^{k-1} = p^2 + p^2 \left(\frac{1}{(1-q)^2} - 1 \right) = 1$$

car $1 - q = p$. D'où la convergence de la série double, et

$$E(Z) = E(\min(X, Y + 1)) = \boxed{q + 1}.$$

2.3.2. Covariance

La covariance de deux variables est un objet qui mesure intuitivement la dépendance d'une des deux variables aléatoires par rapport à l'autre. Un cas limite étant celui de variables aléatoires indépendantes, qui correspondra à une covariance nulle. Voyons la définition.

Définition/Proposition ALEA.15.3 | Covariance

Soit (X, Y) un couple aléatoire discret, tel que X, Y admettent un moment d'ordre deux. On appelle *covariance de X et Y* le réel noté $\mathbf{Cov}(X, Y)$ défini par :

$$\mathbf{Cov}(X, Y) = \mathbf{E}([X - \mathbf{E}(X)] [Y - \mathbf{E}(Y)]).$$

Les variables aléatoires X et Y sont dites *non corrélées* si :

$$\mathbf{Cov}(X, Y) = 0.$$

Preuve (de l'existence de la covariance) Soit $(x, y) \in X(\Omega) \times Y(\Omega)$. En appliquant le théorème de transfert à $g(X, Y)$ avec $g : (x, y) \in \mathbf{R}^2 \mapsto (x - \mathbf{E}(X))(y - \mathbf{E}(Y))$, on voit qu'il suffit de vérifier la convergence de la série double

$$\left(\sum_{(x,y) \in X(\Omega) \times Y(\Omega)} (x - \mathbf{E}(X))(y - \mathbf{E}(Y)) \mathbf{P}(X = x, Y = y) \right)$$

Or,

$$\begin{aligned} & |(x - \mathbf{E}(X))(y - \mathbf{E}(Y))| \mathbf{P}(X = x, Y = y) \\ & \leq \frac{1}{2} (x - \mathbf{E}(X))^2 \mathbf{P}(X = x, Y = y) + \frac{1}{2} (y - \mathbf{E}(Y))^2 \mathbf{P}(X = x, Y = y).^a \end{aligned}$$

► Pour le premier terme, constatons que d'après la formule des probabilités totales, nous avons

$$\sum_{y \in Y(\Omega)} (x - \mathbf{E}(X))^2 \mathbf{P}(X = x, Y = y) = (x - \mathbf{E}(X))^2 \mathbf{P}(X = x),$$

puis comme X admet une variance,

$$\sum_{x \in X(\Omega)} (x - \mathbf{E}(X))^2 \mathbf{P}(X = x) = \mathbf{Var}(X).$$

D'où la convergence de

$$\left(\sum_{(x,y) \in X(\Omega) \times Y(\Omega)} \frac{1}{2} (x - \mathbf{E}(X))^2 \mathbf{P}(X = x, Y = y) \right).$$

► La série

$$\left(\sum_{(x,y) \in X(\Omega) \times Y(\Omega)} \frac{1}{2} (y - \mathbf{E}(Y))^2 \mathbf{P}(X = x, Y = y) \right).$$

converge aussi par symétrie des rôles entre x, y .

Donc d'après le théorème de comparaison pour les séries, la série double initiale converge.

On retiendra aussi de la démonstration l'écriture en série double de la covariance, qui est résumée par la proposition ci-après.

Proposition ALEA.15.9 | Covariance en somme double

Soit (X, Y) un couple aléatoire discret, tel que X, Y admettent un moment d'ordre

^aOn rappelle $|ab| \leq \frac{a^2}{2} + \frac{b^2}{2}$ pour tout a, b réels.

deux, i.e.

$$\left(\sum_{x \in X(\Omega)} x^2 \mathbf{P}(X = x) \right), \quad \left(\sum_{y \in Y(\Omega)} y^2 \mathbf{P}(Y = y) \right) \text{ convergent.}$$

$$\begin{aligned} \text{Alors : } \mathbf{Cov}(X, Y) &= \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} (x - \mathbf{E}(X))(y - \mathbf{E}(Y)) \times \mathbf{P}(X = x, Y = y) \\ &= \sum_{y \in Y(\Omega)} \sum_{x \in X(\Omega)} (x - \mathbf{E}(X))(y - \mathbf{E}(Y)) \times \mathbf{P}(X = x, Y = y). \end{aligned}$$

Proposition ALEA.15.10 | Propriétés de la covariance

Soit (X, Y) un couple aléatoire discret, tel que X, Y admettent un moment d'ordre deux.

1. **(Lien covariance/variance)** $\mathbf{Cov}(X, X) = \mathbf{Var}(X)$.
2. **(Formule de KÖNIG-HUYGENS)** $\mathbf{Cov}(X, Y) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y)$.
3. **(Symétrie)** $\mathbf{Cov}(X, Y) = \mathbf{Cov}(Y, X)$.
4. **(Constante)** $\mathbf{Cov}(X, c) = \mathbf{Cov}(c, X) = 0$ pour toute constante $c \in \mathbf{R}$.
5. **(Covariance et somme)** Les applications $\mathbf{Cov}(\cdot, Y) : Z \mapsto \mathbf{Cov}(Z, Y)$ et $\mathbf{Cov}(X, \cdot) : Z \mapsto \mathbf{Cov}(X, Z)$ sont des applications linéaires sur l'espace vectoriel des variables aléatoires admettant un moment d'ordre deux, i.e. si $(\lambda, \mu) \in \mathbf{R}^2$ et X, Y, Z sont trois variables aléatoires admettant un moment d'ordre deux, alors

$$\mathbf{Cov}(\lambda X + \mu Y, Z) = \lambda \mathbf{Cov}(X, Z) + \mu \mathbf{Cov}(Y, Z),$$

$$\mathbf{Cov}(Z, \lambda X + \mu Y) = \lambda \mathbf{Cov}(Z, X) + \mu \mathbf{Cov}(Z, Y).$$

6. **(Variance d'une somme (cas $n = 2$))**

$$\mathbf{Var}(X + Y) = \mathbf{Var}(X) + \mathbf{Var}(Y) + 2\mathbf{Cov}(X, Y).$$

En particulier, si X et Y sont non corrélées, nous avons :

$$\mathbf{Var}(X + Y) = \mathbf{Var}(X) + \mathbf{Var}(Y).$$


Preuve

1. 2. 3. 4. 5. 6. 

La propriété de bilinéarité doit être comprise comme une règle de développement classique sur les nombres réels du collège. Ainsi, si je décide de noter « \cdot » l'opérateur de covariance, l'assertion⁴

$$\mathbf{Cov}(Z, \lambda X + \mu Y) = \lambda \mathbf{Cov}(Z, X) + \mu \mathbf{Cov}(Z, Y) \quad \text{devient simplement}$$

$$Z \cdot (\lambda X + \mu Y) = \lambda Z \cdot X + \mu Z \cdot Y.$$

Exemple 15 — Soient X, Y deux variables aléatoires admettant un moment d'ordre deux. Développer $\mathbf{Cov}(X - Y, X + Y)$ et $\mathbf{Cov}(2X + 3Y, X - 4Y)$. 

⁴Très simple à retenir donc.

L'expression type KÖNIG-HUYGENS de la covariance nous permet d'établir facilement le corollaire qui suit.

Corollaire ALEA.15.3 | Indépendantes \Rightarrow non-corrélées

Soient X, Y deux variables aléatoires indépendantes, alors X, Y sont non-corrélées.

Preuve D'après le [Corollaire ALEA.15.2](#), $E(XY) - E(X)E(Y) = \mathbf{Cov}(X, Y) = 0$, c'est justement dire que les variables aléatoires X, Y sont non-corrélées.


⊗ Attention

La non corrélation n'implique pas l'indépendance.] Soit $X \hookrightarrow \mathcal{U}(\{-1, 0, 1\})$ et $Y = X^2$. Alors on a déjà montré que $E(XY) = E(X)E(Y) = 0$ donc $\mathbf{Cov}(X, Y) = 0$, et pourtant $Y = X^2$ n'est pas indépendante de X .

🔧 Méthode Développement d'une variance de somme

Soit la quantité $\mathbf{Var}(X + Y)$ avec X, Y deux variables aléatoires discrètes ayant une variance.

1. Écrire la quantité en fonction de la covariance : $\mathbf{Var}(X + Y) = \mathbf{Cov}(X + Y, X + Y)$.
2. Développer en utilisant la bilinéarité de la covariance.

Exemple 16 — Développer $\mathbf{Var}(X - 2Y + 3Z)$ lorsque X, Y, Z sont trois variables aléatoires discrètes ayant une variance. 

Proposition ALEA.15.11 | Variance d'une somme, cas général

Soient X, \dots, X_n une famille de n variables aléatoires possédant un moment d'ordre deux. Alors :

$$\begin{aligned} \mathbf{Var}\left(\sum_{i=1}^n X_i\right) &= \sum_{i=1}^n \mathbf{Var}(X_i) + 2 \sum_{\substack{i < j \\ (i,j) \in \llbracket 1, n \rrbracket^2}} \mathbf{Cov}(X_i, X_j) \\ &= \sum_{i=1}^n \mathbf{Var}(X_i) + \sum_{\substack{i \neq j \\ (i,j) \in \llbracket 1, n \rrbracket^2}} \mathbf{Cov}(X_i, X_j). \quad \text{[H.P]} \end{aligned}$$

En particulier, si elles sont **deux à deux non corrélées** : $\mathbf{Cov}(X_i, X_j) = 0$ pour tout $(i, j) \in \{1, \dots, n\}^2$, $i \neq j$, ou mieux deux à deux indépendantes, alors :

$$\mathbf{Var}(X + \dots + X_n) = \mathbf{Var}(X) + \dots + \mathbf{Var}(X_n).$$

Cette formule n'est pas au programme, elle est cependant très classique. Il faut donc en connaître la démonstration mais pas l'énoncé par coeur.

Preuve



COEFFICIENT DE CORRÉLATION. Un nouvel objet peut être défini à l'aide de la covariance : il s'agit du coefficient de corrélation, qui est un outil de détection de relations affines entre variables aléatoires.

Définition/Proposition ALEA.15.4 | Inégalité de CAUCHY-SCHWARZ/ Coefficient de corrélation.

Soit (X, Y) un couple aléatoire discret, tel que X, Y admettent un un moment d'ordre deux. Alors :

1. (Inégalité de CAUCHY-SCHWARZ)

$$|\mathbf{Cov}(X, Y)| \leq \sigma_X \cdot \sigma_Y.$$

Si X, Y sont d'écart-type non nul. On appelle *coefficient de corrélation entre X et Y* la quantité


$$\rho(X, Y) = \frac{\mathbf{Cov}(X, Y)}{\sigma_X \cdot \sigma_Y} \in [-1, 1].$$

2. (Cas d'égalité) $\rho(X, Y) = \pm 1 \iff \exists a, b \in \mathbf{R}, \mathbf{P}(Y = aX + b) = 1.$

Remarque 2.3 – Interprétation du coefficient de corrélation Ainsi, pour savoir si une variable aléatoire dépend de manière affine de l'autre (avec probabilité 1), il suffit de calculer leur coefficient de corrélation.

- ▶ s'il est proche de 1 ou -1 , on envisage une relation affine entre les deux variables aléatoires.
- ▶ En revanche, plus il est proche de zéro, plus l'existence d'une telle relation devient hautement improbable.

Preuve

1.  Soit $P : t \in \mathbf{R} \longrightarrow \mathbf{V}(tX + Y) = \mathbf{Cov}(tX + Y, tX + Y)$, cette fonction est bien définie puisque X, Y admettent une variance donc $tX + Y$ aussi pour tout t . Alors, par bilinéarité de la covariance :

$$\forall t \in \mathbf{R}, P(t) = \mathbf{V}(X)t^2 + 2t\mathbf{Cov}(X, Y) + \mathbf{V}(Y).$$

Comme $\mathbf{Var}(X) = \sigma_X^2 \neq 0$ par hypothèse, P est une fonction polynomiale de degré 2, positive (par définition et car la variance est positive), ainsi son discriminant est négatif, soit

$$4\mathbf{Cov}(X, Y)^2 - 4\mathbf{V}(X)\mathbf{V}(Y) \leq 0 \iff \mathbf{Cov}(X, Y)^2 \leq \mathbf{V}(X)\mathbf{V}(Y).$$

Il reste à composer par la fonction $\sqrt{\cdot}$, qui est bien croissante, de chaque côté et de l'inégalité. Puisque $\sigma_X \sigma_Y > 0$ par hypothèse, il reste ensuite à diviser de chaque côté par $\sigma_X \sigma_Y$. On obtient bien

$$|\rho(X, Y)| \leq 1.$$

2. 

3. SOMMES DE VARIABLES ALÉATOIRES INDÉPENDANTES & CONVOLUTION

Nous allons voir dans cette partie comment analyser la loi de sommes de variables aléatoires indépendantes, discrètes ou à densité. Dans le cas discret, nous verrons que la formule fait intervenir la loi conjointe du couple. Dans le cas à densité, nous aurons une formule. On applique ensuite ces résultats aux lois usuelles étudiées.

3.1. Cas discret

Proposition ALEA.15.12 | Produit de convolution discret

- ▶ Soit (X, Y) un couple aléatoire discret. Alors $X + Y$ est une variable aléatoire réelle discrète et pour tout $k \in (X + Y)(\Omega)$, on a :

$$\mathbf{P}(X + Y = k) = \sum_{\ell \in X(\Omega)} \mathbf{P}(X = \ell, Y = k - \ell), \text{ ou de manière équivalente :}$$

$$\mathbf{P}_{X+Y}(k) = \sum_{\ell \in X(\Omega)} \mathbf{P}_{(X,Y)}(\ell, k - \ell).$$

- ▶ En particulier, si X et Y sont **indépendantes**, alors : pour tout $k \in (X + Y)(\Omega)$, on a :

$$\mathbf{P}(X + Y = k) = \sum_{\ell \in X(\Omega)} \mathbf{P}(X = \ell) \cdot \mathbf{P}(Y = k - \ell), \text{ ou de manière équivalente :}$$

$$\mathbf{P}_{X+Y}(k) = \sum_{\ell \in X(\Omega)} \mathbf{P}_X(\ell) \cdot \mathbf{P}_Y(k - \ell).$$

Là encore, cette formule n'est pas à retenir par coeur : mémoriser surtout qu'il s'agit

simplement d'une formule des probabilités totales où l'on introduit le système complet d'évènements associé à l'une ou l'autre des variables aléatoires. On retiendra que :

- ▶ la loi de la somme s'exprime donc en fonction de la loi conjointe du couple.
- ▶ En cas d'indépendance, la loi de la somme s'exprime donc en fonction des lois marginales.

Méthode Déterminer la loi de la somme de deux variables aléatoires discrètes indépendantes

1. Introduire le système complet associé à X ou Y dans l'évènement $\{X + Y = k\}$ pour tout $k \in (X + Y)(\Omega)$.
2. Utiliser la formule des probabilités totales pour avoir

$$\mathbf{P}(X + Y = k) = \sum_{\ell \in X(\Omega)} \mathbf{P}(X = \ell) \cdot \mathbf{P}(Y = k - \ell).$$


3. Pour poursuivre le calcul, regarder pour quels $\ell \in X(\Omega)$, on a :


$$\mathbf{P}(X = \ell) \cdot \mathbf{P}(Y = k - \ell) \neq 0.$$

Puis conclure en finissant le calcul de la somme.


Preuve



Exemple 17 — Cas fini Soient X, Y deux variables aléatoires indépendantes suivant une loi $\mathcal{B}(p)$, avec $p \in]0, 1[$. Déterminer la loi de $X + Y$, que remarquer? 

2. Pour $n \geq 2$ fixé, déterminer la loi conditionnelle de X sachant $\{X + Y = n\}$. 

Exemple 18 — Cas dénombrable Soient X, Y deux variables aléatoires indépendantes suivant une loi $\mathcal{G}(p)$, avec $p \in]0, 1[$.

1. Déterminer la loi de $X + Y$. Étudier l'existence d'une espérance, la calculer le cas échéant. 

On peut enfin démontrer certains résultats du [Chapter ALEA.13](#) qui ont été admis.

Théorème ALEA.15.3 | Stabilité des lois binomiale et de POISSON

Soient X, \dots, X_d des variables aléatoires réelles indépendantes, $d \in \mathbf{N}^*$. Alors :


1. si X, \dots, X_d sont d variables aléatoires réelles mutuellement indépendantes de lois $\mathcal{B}(n_1, p), \dots, \mathcal{B}(n_d, p)$, alors

$$X + \dots + X_d \leftrightarrow \mathcal{B}(n_1 + \dots + n_d, p).$$

2. Si X, \dots, X_d sont d variables aléatoires réelles mutuellement indépendantes de loi $\mathcal{P}(\lambda_1), \dots, \mathcal{P}(\lambda_d)$, avec $\lambda_1, \dots, \lambda_d \in \mathbf{R}^+$ alors :

$$X + \dots + X_d \leftrightarrow \mathcal{P}(\lambda_1 + \dots + \lambda_d).$$

La propriété pour la binomiale se comprend aisément puisque si l'on réalise n_1 puis n_2 etc. puis n_d expériences de BERNOULLI identiques indépendantes et qu'on somme les nombres de succès obtenus, cela revient à effectuer $n_1 + \dots + n_d$ cette même expérience et compter les succès.

 **Attention** Pour la loi binomiale, on ne peut en général pas sommer par rapport au paramètre p , et même en cas d'indépendance.

Preuve Faisons les preuves pour $d = 2$, le résultat général s'en déduit alors par récurrence évident sur d .

1. Rappelons la formule de VANDERMONDE vue dans le TD de dénombrement : pour tout $k \in \mathbf{Z}$ tel que $k \leq n_1, n_2$, nous avons


$$\binom{n_1 + n_2}{k} = \sum_{\ell=0}^k \binom{n_1}{\ell} \binom{n_2}{k-\ell}.$$




2. 

3.2. Cas à densité

Exemple 19 — Soient X et Y deux variables aléatoires indépendantes suivant des lois de Poisson de paramètres λ et μ .

1. Donner la loi de $X + Y$. Quelle est son espérance? 

2. Pour $n \geq 0$ fixé, déterminer la loi conditionnelle de X sachant $\{X + Y = n\}$. 

Le théorème suivant est fondamental et sera admis. L'idée est la suivante : on connaît une formule pour la densité de la somme de deux variables aléatoires indépendantes à densité.

Théorème ALEA.15.4 | Produit de convolution à densité

- Soient X et Y deux variables aléatoires **indépendantes**⁵ de densités f_X et f_Y . Alors l'application notée $f_X \star f_Y$ définie par :

$$\forall z \in \mathbf{R}, \quad (f_X \star f_Y)(z) = \begin{cases} \int_{-\infty}^{+\infty} f_X(x)f_Y(z-x) dx & \text{en cas de convergence,} \\ 0 & \text{sinon.} \end{cases}$$

est une densité de $X + Y$. Autrement dit,

$$f_{X+Y} = f_X \star f_Y.$$

- Plus généralement, si X_1, \dots, X_n sont à densité **indépendantes**, alors $X_1 + \dots + X_n$ est également à densité.

Notez l'analogie avec la convolution discrète : le symbole somme est remplacé par une intégrale, les lois sont remplacées par les densités.

Preuve Admis.

Attention

L'hypothèse d'indépendance est cruciale. En effet, soit X est une variable aléatoire à densité. Alors X et $-X$ ne sont pas indépendantes, et

$X + (-X) = 0$ n'est même pas une variable aléatoire à densité (car discrète).

⁵Il existe aussi une formule dans le cas non-indépendant, elle fait intervenir la notion de «densité conjointe» du couple, mais qui n'est pas au programme.

Remarque 3.1 —

- ▶ Comme cela est précisé dans le programme officiel, la condition « $f_X \star f_Y$ est définie et continue sur \mathbf{R} sauf éventuellement en un nombre fini de points» est souvent admise, il convient alors uniquement de justifier que $X \perp\!\!\!\perp Y$ — l'indépendance est quant à elle attendue dans votre rédaction!
- ▶ Il est facile de voir que, dès que l'une des deux densités au moins est bornée, l'intégrale $f_X \star f_Y$ converge. Et c'est bien le cas de toutes les lois usuelles vues dans ce chapitre.
- ▶ On a, par changement de variable :

$$\int_{-\infty}^{+\infty} f_X(x)f_Y(z-x) dx = \int_{-\infty}^{+\infty} f_X(z-y)f_Y(y) dy.$$

Méthode Déterminer une densité de la somme de deux variables aléatoires indépendantes à densité

1. Rappeler l'hypothèse d'indépendance puis écrire l'expression de la convolée des densités :


$$\forall z \in \mathbf{R}, \quad (f_X \star f_Y)(z) = \begin{cases} \int_{-\infty}^{+\infty} f_X(x)f_Y(z-x) dx & \text{en cas de convergence,} \\ 0 & \text{sinon.} \end{cases}$$

2. Pour poursuivre le calcul, regarder pour quels $x \in \mathbf{R}$, on a :

$$f_X(x)f_Y(z-x) \neq 0.$$

Puis conclure en réduisant l'intervalle d'intégration.

Exemple 20 — Calculer une densité de $X + Y$ lorsque X et Y sont deux variables aléatoires indépendantes qui suivent les lois indiquées.

1. $X, Y \hookrightarrow \mathcal{U}[0, 1]$.  Puisque X, Y sont indépendantes, alors une densité de f_{X+Y} est

$f_X \star f_Y$. Ainsi, pour tout $z \in \mathbf{R}$,

$$f_{X+Y}(z) = \int_{-\infty}^{+\infty} f_X(x)f_Y(z-x) dz,$$

or : $f_X(x)f_Y(z-x) \neq 0 \iff 0 \leq x \leq 1, 0 \leq z-x \leq 1$, i.e. :

$$0 \leq x \leq 1, -1+z \leq x \leq z.$$

Dessignons ces deux intervalles en x .





Plusieurs cas de figure se présentent alors.

- ▶ si $z \leq 0$ ou $-1+z \geq 1$ alors on intègre la fonction nulle et $f_{X+Y}(z) = 0$.
- ▶ Si $0 < z \leq 1$, alors $f_{X+Y}(z) = \int_0^z 1 dt = z$.
- ▶ Si $0 < -1+z \leq 1$, alors $f_{X+Y}(z) = \int_{-1+z}^1 1 dt = 2-z$.

On déduit alors une densité de $X + Y$:

$$f_{X+Y}(z) = \begin{cases} 0 & \text{si } z \leq 0, \text{ ou } z \geq 2, \\ z & \text{si } 0 < z \leq 1, \\ 2-z & \text{si } 1 < z < 2. \end{cases}$$

2. $X \hookrightarrow \mathcal{U}[-1, 1], Y \hookrightarrow \mathcal{U}[0, 1]$. 

3. $X \hookrightarrow \mathcal{E}(1), Y \hookrightarrow \mathcal{E}(2)$. 

Proposition ALEA.15.13 | Somme de lois normales indépendantes

Soient X_1, \dots, X_n une collection de n variables aléatoires indépendantes telles que :

$$\forall i \in \llbracket 1, n \rrbracket, \quad X_i \hookrightarrow \mathcal{N}(\mu_i, \sigma_i^2), \quad \mu_i \in \mathbf{R}, \quad \sigma_i > 0.$$

▶ Alors :

$$\sum_{i=1}^n X_i \hookrightarrow \mathcal{N}\left(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2\right) = \mathcal{N}\left(\mathbf{E}\left(\sum_{i=1}^n X_i\right), \mathbf{Var}\left(\sum_{i=1}^n X_i\right)\right).$$


▶ Et plus généralement : pour toute famille $(\lambda_1, \dots, \lambda_n) \in \mathbf{R}^n \setminus \{(0, \dots, 0)\}$, on a :

$$\begin{aligned} \sum_{i=1}^n \lambda_i X_i &\hookrightarrow \mathcal{N}\left(\sum_{i=1}^n \lambda_i \cdot \mu_i, \sum_{i=1}^n \lambda_i^2 \cdot \sigma_i^2\right) \\ &= \mathcal{N}\left(\mathbf{E}\left(\sum_{i=1}^n \lambda_i \cdot X_i\right), \mathbf{Var}\left(\sum_{i=1}^n \lambda_i \cdot X_i\right)\right). \end{aligned}$$

Preuve Nous allons nous contenter de montrer le cas $n = 2, \mu_1 = \mu_2, \sigma_1^2 = \sigma_2^2 = 1$ et $\lambda_1 = \lambda_2 = 1$, *i.e.* montrons que $X_1 + X_2 \hookrightarrow \mathcal{N}(0, 2)$.

▶ Commençons par justifier que $-\frac{1}{2}t^2 - \frac{1}{2}(x-t)^2 = -\frac{1}{4}x^2 - \frac{1}{2}\left(\frac{t-x/2}{1/\sqrt{2}}\right)^2$ pour tout

$$t, x \in \mathbf{R}. \quad \img alt="pencil icon" data-bbox="608 578 627 605"/>$$

▶ On peut alors déduire que $X_1 + X_2 \hookrightarrow \mathcal{N}(0, 2)$. 

On peut enfin démontrer le résultat de stabilité de la loi normale du [Chapter ALEA.14](#) qui avait été admis et dont on rappelle l'énoncé.

Le cas général se démontre en utilisant le produit de convolution mais les calculs sont plus compliqués.

Corollaire ALEA.15.4 | Application rigolote

Soient X et Y deux variables aléatoires indépendantes à densité. Alors :

$$\mathbf{P}(X = Y) = 0.$$

Preuve



*** **Fin du chapitre** ***

4. EXERCICES

4.1. Calculs déterministes

Exercice ALEA.15.1 | Sommes doubles finies Soit $n \in \mathbb{N}^*$. Calculer les sommes suivantes :

$$\begin{array}{ll} \blacktriangleright \sum_{i=0}^n \sum_{j=1}^n (2i+j), & \blacktriangleright \sum_{1 \leq i, j \leq n} |i-j|, \\ \blacktriangleright \sum_{i=1}^{n-1} \sum_{j=i+1}^n 2^{i+j}, & \blacktriangleright \sum_{1 \leq i < j \leq n} (i+j). \\ \blacktriangleright \sum_{i=1}^n \sum_{j=i}^n \frac{i}{j}, & \end{array}$$

Exercice ALEA.15.2 | Séries doubles Étudier la convergence des séries doubles ci-après. Calculer leur somme le cas échéant.

$$\left(\sum_{i,j \geq 0} \frac{ij^2}{3^{i+j}} \right), \quad \left(\sum_{i,j \geq 0} \frac{i^2-1}{i!j!} \right).$$

Exercice ALEA.15.3 | Intégration de fonctions constantes par morceaux Calculer les intégrales suivantes. *On admettra leur existence*

$$\begin{array}{l} 1. \text{ Pour tout } x \in \mathbf{R}, \int_{-\infty}^{2x} f \text{ et } \int_{-3x}^{3x} f \text{ où : } f : t \longmapsto \begin{cases} \frac{1}{4} & \text{si } t \in [-1, 3] \\ 0 & \text{sinon.} \end{cases} \\ 2. \text{ Pour tout } x \in \mathbf{R}, \int_{-\infty}^{\infty} g(t)g(x-t) dt \text{ où : } g : t \longmapsto \begin{cases} e^{-t} & \text{si } t \geq 0, \\ 0 & \text{sinon.} \end{cases} \end{array}$$

4.2. Généralités

Exercice ALEA.15.4 | Coefficient de corrélation de combinaisons linéaires Soient X et Y deux variables aléatoires réelles admettant une variance non nulle, et $a, b, c, d \in \mathbf{R}$ tels que $a \neq 0, c \neq 0$. Montrer que :

$$\rho(aX+b, cY+d) = \text{sgn}(ac)\rho(X, Y), \quad \text{où } \text{sgn}(x) = \begin{cases} 1 & \text{si } x \geq 0, \\ -1 & \text{si } x < 0. \end{cases}$$

Solution (exercice ALEA.15.4)

Tout d'abord, précisons que l'on suppose $a \neq 0, c \neq 0$ pour que $\sigma_{aX+b} \neq 0$ et $\sigma_{cY+d} \neq 0$: en effet,

$$\sigma_{aX+b}^2 = \text{Var}(aX+b) = a^2 \text{Var}(X) \neq 0$$

dès que $a \neq 0$, de-même pour σ_{cY+d} .

Avec les notations de l'énoncé, nous avons

$$\rho(aX+b, cY+d) = \frac{\text{Cov}(aX+b, cY+d)}{\sqrt{\text{Var}(aX+b)}\sqrt{\text{Var}(cY+d)}},$$

or d'après les propriétés classiques sur la variance,

$$\text{Var}(aX+b) = a^2 \text{Var}(X) = a^2 \sigma_X^2, \quad \text{Var}(cY+d) = c^2 \text{Var}(Y) = c^2 \sigma_Y^2.$$

Puis, par bilinéarité de la covariance :

$$\text{Cov}(aX+b, cY+d) = ac \text{Cov}(X, Y) + a \text{Cov}(X, d) + c \text{Cov}(b, Y) + \text{Cov}(b, d).$$

Or la covariance d'une variable aléatoire avec n'importe quelle constante est nulle, donc on trouve finalement

$$\text{Cov}(aX+b, cY+d) = ac \text{Cov}(X, Y).$$

Ainsi,

$$\rho(aX + b, cY + d) = \frac{ac \mathbf{Cov}(X, Y)}{|a| |c| \sigma_X \sigma_Y} = \frac{ac}{|a| |c|} \rho(X, Y).$$

Or, $\frac{ac}{|a| |c|} = \text{sgn}(ac)$, on obtient alors la formule

$$\boxed{\rho(aX + b, cY + d) = \text{sgn}(ac) \rho(X, Y)}.$$

Exercice ALEA.15.5 | Équivalence entre indépendance et non-corrélation pour les couples de BERNOULLI Soient X et Y deux variables aléatoires de BERNOULLI de paramètre $p \in [0, 1]$.

1. Montrer que

$$|\mathbf{Cov}(X, Y)| \leq \frac{1}{4}.$$

2. Montrer que X, Y sont indépendantes, si et seulement si $\mathbf{Cov}(X, Y) = 0$. *Propriété également symptomatique des vecteurs aléatoires de lois normales, appelés «vecteurs gaussiens» : l'indépendance est équivalente à la non-corrélation deux à deux.*

Solution (exercice ALEA.15.5)

1. D'après le cours, $\mathbf{Var}(X) = p(1-p) = \mathbf{Var}(Y)$. Or,

$$|\rho(X, Y)| = \left| \frac{\mathbf{Cov}(X, Y)}{p(1-p)} \right| \leq 1,$$

dès lors, nous obtenons

$$|\mathbf{Cov}(X, Y)| \leq p(1-p).$$

Or, par étude de la fonction $f : x \mapsto x(1-x)$ sur $[0, 1]$, nous déduisons que pour tout $x \in [0, 1]$,

$$x(1-x) \leq \frac{1}{4}.$$

Ainsi, $\boxed{|\mathbf{Cov}(X, Y)| \leq \frac{1}{4}}$.

2. Montrons que X, Y sont indépendantes, si et seulement si $\mathbf{Cov}(X, Y) = 0$.

\Rightarrow Conséquence directe du cours : deux variables indépendantes sont non-corrélées.

\Leftarrow Supposons que $\mathbf{Cov}(X, Y) = 0$. D'après le théorème de transfert, applicable car on travaille sur des univers-images finis, nous avons

$$\mathbf{Cov}(X, Y) = \mathbf{E}(XY) - \mathbf{E}(X) \mathbf{E}(Y) = 0 + 0 + 0 + 1 \times 1 \mathbf{P}(X = 1, Y = 1) - p^2,$$

donc l'hypothèse donne

$$\mathbf{P}(X = 1, Y = 1) - p^2 = 0 \iff \mathbf{P}(X = 1, Y = 1) = \mathbf{P}(X = 1) \mathbf{P}(Y = 1).$$

Maintenant, il faut magouiller avec les évènements pour obtenir le reste de la définition d'indépendance.

$$\mathbf{P}(X = 1, Y = 0) + \mathbf{P}(X = 1, Y = 1) = \mathbf{P}(X = 1) = p,$$

donc on obtient

$$\mathbf{P}(X = 1, Y = 0) = p - p^2 = p(1-p) = \mathbf{P}(X = 1) \mathbf{P}(Y = 0),$$

on obtient par le même procédé

$$\mathbf{P}(X = 0, Y = 1) = p - p^2 = (1-p)p = \mathbf{P}(X = 0) \mathbf{P}(Y = 1),$$

puis enfin

$$\begin{aligned} \mathbf{P}(X = 0, Y = 0) &= \mathbf{P}(X = 0) - \mathbf{P}(X = 0, Y = 1) = 1 - p - (1-p)p \\ &= (1-p)(1-p) = \mathbf{P}(X = 0) \mathbf{P}(Y = 0). \end{aligned}$$

4.3. Étude de lois

Exercice ALEA.15.6 | Soient X et Y deux variable aléatoire prenant leurs valeurs dans $\{0, 1, 2\}$. On suppose que la loi du couple (X, Y) est donnée par le tableau :

$Y \backslash X$	0	1	2
0	p	$\frac{p}{2}$	$\frac{p}{4}$
1	$2p$	p	$\frac{p}{2}$
2	$4p$	$2p$	p

- Déterminer une condition sur p pour que ce tableau représente effectivement la loi conjointe d'un couple aléatoire.
- Déterminer alors les lois marginales de ce couple.
- Calculer le coefficient de corrélation entre X et Y .
- Les variables X et Y sont-elles indépendantes ?

Solution (exercice ALEA.15.6)

- D'après le cours, le tableau représente la loi conjointe d'un couple (X, Y) si et seulement si tous les termes du tableau sont positifs, et si leur somme vaut 1. Et en effet, leur somme est

$$3p + 8p + p + \frac{p}{4} = \frac{49p}{4}.$$

Donc la seule valeur possible de p est $\frac{4}{49}$. Par ailleurs pour cette valeur de p tous les termes du tableau sont bien positifs. Donc $p = \frac{4}{49}$.

- D'après le tableau, on a $X(\Omega) = Y(\Omega) = \{0, 1, 2\}$. D'après la formule des probabili-

tés totales, on a :

$$\mathbf{P}(X = 0) = \sum_{\ell=0}^2 \mathbf{P}(X = 0, Y = \ell) = \boxed{7p = \frac{4}{7}},$$

$$\mathbf{P}(X = 1) = \sum_{\ell=0}^2 \mathbf{P}(X = 1, Y = \ell) = \boxed{\frac{7p}{2} = \frac{2}{7}},$$

$$\mathbf{P}(X = 2) = \sum_{\ell=0}^2 \mathbf{P}(X = 2, Y = \ell) = \boxed{\frac{7p}{4} = \frac{1}{7}}.$$

En faisant de-même pour Y , on trouve

$$\boxed{\mathbf{P}(Y = 0) = \frac{7p}{4} = \frac{1}{7}, \quad \mathbf{P}(Y = 1) = \frac{7p}{2} = \frac{2}{7}, \quad \mathbf{P}(Y = 2) = 7p = \frac{4}{7}}.$$

- On commence par calculer la covariance, *i.e.* $\mathbf{Cov}(X, Y) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y)$, qui existe puisque X, Y sont des variables aléatoires finies. Tout d'abord,

$$\mathbf{E}(X) = \frac{14}{49} + \frac{2}{7} = \frac{4}{7},$$

$$\mathbf{E}(Y) = \frac{14}{49} + \frac{8}{7} = \frac{10}{7}.$$

D'après le théorème de transfert, on a :

$$\begin{aligned} \mathbf{E}(XY) &= 0 + p + 2\frac{p}{2} + 2(2p) + 4p \\ &= 10p = \frac{40}{49}. \end{aligned}$$

Donc

$$\mathbf{Cov}(X, Y) = \frac{40}{49} - \frac{4}{7} \frac{10}{7} = 0.$$

Donc $\rho(X, Y) = \frac{\mathbf{Cov}(X, Y)}{\sigma_X \sigma_Y} = 0$.

- Oui, vérifier tous les cas possibles.

Exercice ALEA.15.7 | Loi multinomiale – D'après concours G2E Soient p_1, p_2 et p_3 des réels strictement positifs. Pour tout $n \in \mathbf{N}$, on considère $D_n = \{(i, j) \in \mathbf{N}^2 : i + j \leq n\}$ et la fonction de deux variable f définie sur D_n par

$$f(i, j) = \frac{n!}{i!j!(n-i-j)!} p_1^i p_2^j p_3^{n-i-j}.$$

- Déterminer une condition nécessaire et suffisante d'existence de (X, Y) de loi conjointe f .
- Déterminer les lois marginales des variables aléatoires X et Y , ainsi que leur espérance et variance.

Solution (exercice ALEA.15.7)

1. Il s'agit de calculer la somme double. On a

$$\begin{aligned} \sum_{(i,j) \in D_n} f(i, j) &= \sum_{i=0}^n \sum_{j=0}^{n-i} \frac{n!}{i!j!(n-i-j)!} p_1^i p_2^j p_3^{n-i-j} \\ &= \sum_{i=0}^n \frac{n! p_1^i}{i!} \sum_{j=0}^{n-i} \frac{1}{j!(n-i-j)!} p_2^j p_3^{n-i-j} \\ &= \sum_{i=0}^n \frac{n! p_1^i}{i!(n-i)!} \sum_{j=0}^{n-i} \binom{n-i}{j} p_2^j p_3^{n-i-j} \\ &= \sum_{i=0}^n \frac{n! p_1^i}{i!(n-i)!} (p_2 + p_3)^{n-i} \\ &= (p_1 + p_2 + p_3)^n. \end{aligned}$$

binôme
binôme

Donc

$$\sum_{(i,j) \in D_n} f(i, j) = 1 \iff (p_1 + p_2 + p_3)^n = 1 \iff p_1 + p_2 + p_3 = 1.$$

Par ailleurs, tous les $f(i, j), (i, j) \in D_n$ sont bien entendu positifs, donc f définit une loi conjointe si et seulement si

$$p_1 + p_2 + p_3 = 1.$$

2. On a $X(\Omega) = \llbracket 0, n \rrbracket = Y(\Omega)$. De plus, soit $i \in \llbracket 0, n \rrbracket$, alors d'après la formule des probabilités totales :

$$\begin{aligned} \mathbf{P}(X = i) &= \sum_{j=0}^n f(i, j) \\ &= \sum_{j=0}^{n-i} \frac{n!}{i!j!(n-i-j)!} p_1^i p_2^j p_3^{n-i-j} \\ &= \frac{n!}{i!(n-i)!} p_1^i \sum_{j=0}^{n-i} \frac{(n-i)!}{j!(n-i-j)!} p_2^j p_3^{n-i-j} \\ &= \binom{n}{i} p_1^i (p_2 + p_3)^{n-i} \\ &= \binom{n}{i} p_1^i (1 - p_1)^{n-i}. \end{aligned}$$

binôme
 $p_2 + p_3 = 1 - p_1$

Donc $X \hookrightarrow \mathcal{B}(n, p_1)$. De-même, en échangeant les rôles de i et j , on trouve

$Y \hookrightarrow \mathcal{B}(n, p_2)$. Donc

$$\mathbf{E}(X) = np_1, \mathbf{Var}(X) = np_1(1 - p_1), \mathbf{E}(Y) = np_2, \mathbf{Var}(Y) = np_2(1 - p_2).$$

Exercice ALEA.15.8 | Soient X et Y deux variables aléatoires à valeurs dans \mathbf{N}^* telles que :

$$\mathbf{P}(X = i, Y = j) = \frac{a}{2^{i+j}} \quad \text{avec } (i, j) \in (\mathbf{N}^*)^2 \text{ et } a \in \mathbf{R}.$$

1. Calculer a pour que la famille $(\mathbf{P}(X = i, Y = j))_{i,j}$ soit bien une loi conjointe d'un couple aléatoire.

2. Déterminer les lois marginales de X et Y.
3. Les variables aléatoires X et Y sont-elles indépendantes?

Solution (exercice ALEA.15.8)

1. — Soit $i \in \mathbf{N}^*$, alors étudions la convergence de $\left(\sum_{j \geq 1} \frac{1}{2^{i+j}}\right)$. Puisque $|1/2| < 1$, nous avons

$$\sum_{j=1}^{\infty} \frac{1}{2^{i+j}} = \frac{1}{2^i} \sum_{j=1}^{\infty} \frac{1}{2^j} = \frac{1}{2^i} \frac{1}{1 - 1/2} = \frac{1}{2^i}.$$

— De même $\sum_{i=1}^{\infty} \frac{1}{2^i} = \frac{1}{2} \frac{1}{1 - 1/2} = 1$.

Les hypothèses du théorème de FUBINI sont donc vérifiées, et nous avons

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \frac{1}{2^{i+j}} = 1.$$

Donc l'unique valeur de a telle que

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \mathbf{P}(X = i, Y = j) = 1$$

est $a = 1$. De plus, pour tout $(i, j) \in (\mathbf{N}^*)^2$,

$$\frac{a}{2^{i+j}} \geq 0$$

d'où l'existence d'un couple aléatoire (X, Y) de loi conjointe donnée précédemment.

2. Nous avons clairement $X(\Omega) = Y(\Omega) = \mathbf{N}^*$. De plus, pour tout $i \geq 1$, nous avons d'après le cours :

$$\mathbf{P}(X = i) = \sum_{j=1}^{\infty} \mathbf{P}(X = i, Y = j) = \frac{1}{2^i} \sum_{j=1}^{\infty} \frac{1}{2^j} = \frac{1}{2^i} = \frac{1}{2} \frac{1}{2^{i-1}},$$

de même, par symétrie des indices, on a

$$\mathbf{P}(Y = j) = \frac{1}{2^j} = \frac{1}{2} \frac{1}{2^{j-1}}.$$

On reconnaît ainsi des lois géométriques : $X, Y \hookrightarrow \mathcal{G}(1/2)$.

3. Les variables aléatoires X, Y sont indépendantes puisque pour tout $(i, j) \in (\mathbf{N}^*)^2$,

$$\mathbf{P}(X = i) \mathbf{P}(Y = j) = \frac{1}{2^i} \frac{1}{2^j} = \frac{1}{2^{i+j}} = \mathbf{P}(X = i, Y = j)$$

donc $X \perp\!\!\!\perp Y$.

Exercice ALEA.15.9 | Une urne contient 3 jetons numérotés 1, 2 et 3, indiscernables au toucher. On effectue une suite de tirages d'un jeton de cette urne, en remplaçant à chaque fois le jeton obtenu avant le tirage suivant.

1. On note Y le nombre de tirage nécessaires pour obtenir, pour la première fois, deux numéros différents. Déterminer la loi de Y et son espérance.
2. On note Z le nombre de tirages nécessaires pour obtenir, pour la première fois, les trois numéros. Déterminer la loi du couple (Z, Y) .
3. Déterminer la loi de Z et son espérance.

Solution (exercice ALEA.15.9)

1. La loi en question ressemble à une géométrique, sachant le résultat du premier tirage. Notons T_i l'évènement «le jeton i sort au tirage 1» pour tout $i \in \llbracket 1, 3 \rrbracket$. Alors sachant l'un de ces évènements, Y compte le temps d'attente de l'un des deux autres jetons, de probabilité $\frac{2}{3}$. Donc $Y(\Omega) = [2, \infty[\cap \mathbf{N}$, et de plus, pour tout $k \geq 2$,

$$\begin{aligned} \mathbf{P}(Y = k) &= \sum_{i=1}^3 \mathbf{P}(Y = k | T_i) \mathbf{P}(T_i) \\ &= \sum_{i=1}^3 \frac{2}{3} \left(\frac{1}{3}\right)^{(k-1)-1} \times \frac{1}{3} \\ &= \frac{2}{3} \left(\frac{1}{3}\right)^{k-2}. \end{aligned}$$

Il faut bien faire attention à la dernière étape au décalage de 1 dans la puissance : c'est seulement une fois le premier lancer réalisé que l'on commence à attendre un succès. Afin de minimiser nos efforts, on peut constater que $Y-1 \hookrightarrow \mathcal{G}\left(\frac{2}{3}\right)$ puisque $(Y-1)(\Omega) = \mathbf{N}^*$ et que pour tout $k \in \mathbf{N}^*$, $\mathbf{P}(Y-1 = k) = \mathbf{P}(Y = k+1) = \frac{2}{3} \left(\frac{1}{3}\right)^{k-1}$. Donc $Y-1$ admet une espérance et

$$\mathbf{E}(Y-1) = \frac{3}{2} \iff \boxed{\mathbf{E}(Y) = \frac{5}{2}}.$$

2. On note Z le nombre de tirages nécessaires pour obtenir, pour la première fois, les trois numéros. Nécessairement, il faut plus de temps pour obtenir les 3 numéros que les 2. Ainsi

$$\boxed{(Z, Y)(\Omega) = \left\{ (k, \ell) \in ([2, \infty[\cap \mathbf{N})^2, k > \ell \right\}}.$$

Soient $k, \ell \geq 2$ de sorte que $k > \ell$. Alors si on a obtenu deux numéros distincts au tirage ℓ , on attendra un dernier succès, de probabilité $\frac{1}{3}$, donc nous aurons une loi géométrique à nouveau.

$$\begin{aligned} \mathbf{P}(Z = k, Y = \ell) &= \mathbf{P}(Z = k | Y = \ell) \mathbf{P}(Y = \ell) \\ &= \frac{1}{3} \left(\frac{2}{3}\right)^{k-1-\ell} \times \frac{2}{3} \left(\frac{1}{3}\right)^{\ell-2} \\ &= \boxed{\left(\frac{1}{3}\right)^{\ell-1} \cdot \left(\frac{2}{3}\right)^{k-\ell}}. \end{aligned}$$

3. On a $\boxed{Z(\Omega) = [3, \infty[\cap \mathbf{N}}$ puisqu'il faut attendre au moins trois lancers pour avoir

tous les numéros. Soit donc $k \geq 3$, alors

$$\begin{aligned} \mathbf{P}(Z = k) &= \sum_{\ell=2}^{k-1} \left(\frac{1}{3}\right)^{\ell-1} \cdot \left(\frac{2}{3}\right)^{k-\ell} \\ &= 3 \left(\frac{2}{3}\right)^k \sum_{\ell=2}^{k-1} \left(\frac{1}{2}\right)^\ell \\ &= 3 \left(\frac{2}{3}\right)^k \left(\frac{1}{2}\right)^2 \frac{1 - \left(\frac{1}{2}\right)^{k-2}}{\frac{1}{2}} \\ &= \boxed{\frac{3}{2} \left(\frac{2}{3}\right)^k \left(1 - \left(\frac{1}{2}\right)^{k-2}\right)}. \end{aligned}$$

4.4. Somme de variables aléatoires & Stabilité de lois

4.4.1. Cas discret

Exercice ALEA.15.10 | On fait deux biopsies à un patient. Dans la première n cellules sont étudiées et on désigne par X le nombre de cellules malignes. Dans la seconde m cellules sont étudiées et on note Y le nombre de cellules malignes. La probabilité qu'une cellule soit maligne est notée p .

1. Par quelle loi peut-on modéliser les variables aléatoires X et Y ?
2. Que représente $X + Y$? Déterminer la loi de $X + Y$.
3. Le laborantin a mélangé par inadvertance les deux éprouvettes. Quelle est alors la loi conditionnelle de X sachant $\{X + Y = k\}$, $k \in [0, n + m]$?

Solution (exercice ALEA.15.10)

- La variable aléatoire X compte le nombre de succès (avoir une cellule maligne, de probabilité p) dans une répétition de n expériences de BERNOULLI (considérer une cellule et analyser si elle est maligne ou pas) indépendantes. Donc $X \hookrightarrow \mathcal{B}(n, p)$, de même $Y \hookrightarrow \mathcal{B}(m, p)$.
- La variable aléatoire $X + Y$ compte alors le nombre total de cellules malignes sur les deux prélèvements, et par indépendance, d'après le cours nous avons $X + Y \hookrightarrow \mathcal{B}(n + m, p)$.
- Le laborantin a mélangé par inadvertance les deux éprouvettes. Soit $\ell \in \llbracket 0, n \rrbracket$, alors

$$\begin{aligned}
 \mathbf{P}(X = \ell | X + Y = k) &= \frac{\mathbf{P}(X = \ell, X + Y = k)}{\mathbf{P}(X + Y = k)} \\
 &= \frac{\mathbf{P}(X = \ell, Y = k - \ell)}{\mathbf{P}(X + Y = k)} \quad \left. \vphantom{\frac{\mathbf{P}(X = \ell, Y = k - \ell)}{\mathbf{P}(X + Y = k)}}} \right) \text{indépendance} \\
 &= \frac{\mathbf{P}(X = \ell) \mathbf{P}(Y = k - \ell)}{\mathbf{P}(X + Y = k)} \\
 &= \frac{\binom{n}{\ell} p^\ell (1-p)^{n-\ell} \cdot \binom{m}{k-\ell} p^{k-\ell} (1-p)^{m-k+\ell}}{\binom{n+m}{k} p^k (1-p)^{n+m-k}} \quad \left. \vphantom{\frac{\binom{n}{\ell} p^\ell (1-p)^{n-\ell} \cdot \binom{m}{k-\ell} p^{k-\ell} (1-p)^{m-k+\ell}}{\binom{n+m}{k} p^k (1-p)^{n+m-k}}} \right) \text{simplifications} \\
 &= \frac{\binom{n}{\ell} \cdot \binom{m}{k-\ell}}{\binom{n+m}{k}}.
 \end{aligned}$$

On reconnaît alors une loi $\mathcal{H}\left(n + m, k, \frac{n}{n+m}\right)$.

Exercice ALEA.15.11 | Sommer exceptionnellement en le paramètre p Une urne contient des jetons numérotés de 1 à $k \geq 2$. Chaque jeton $j \in \llbracket 1, k \rrbracket$ est en proportion $p_j \in]0, 1[$. On effectue $n \geq 1$ tirages avec remise.

- Déterminer la loi de N_i définie comme le nombre de jetons numéros i tirés.
- Soient i, j deux entiers de $\llbracket 1, k \rrbracket$.
 - Déterminer la loi de $N_i + N_j$ si $i \neq j$, son espérance, sa variance.

2.2) Calculer $\mathbf{Cov}(N_i, N_j)$.

Solution (exercice ALEA.15.11)

- On répète n fois de façon indépendante et identique la même expérience de BERNOULLI qui consiste à tirer un jeton dans l'urne. Le succès est de tirer un jeton numéroté i et sa probabilité est p_i . Comme N_i compte le nombre de succès, on a $N_i \hookrightarrow \mathcal{B}(n, p_i)$.
- Il s'agit de traiter globalement la variable $N_i + N_j$:
 - De nouveau comme dans la question précédente, on répète n fois de façon indépendante et identique la même expérience de BERNOULLI qui consiste à tirer un jeton dans l'urne. Mais cette fois, le succès est de tirer un jeton numéroté i ou j et sa probabilité est donc $p_i + p_j$ (proportion de jetons portant le numéro i ou j).
Comme $N_i + N_j$ compte le nombre de succès, on a $N_i + N_j \hookrightarrow \mathcal{B}(n, p_i + p_j)$.
D'où $\mathbf{E}(N_i + N_j) = n(p_i + p_j)$ et $\mathbf{V}(N_i + N_j) = n(p_i + p_j)(1 - (p_i + p_j))$.
 - Or on sait aussi que $\mathbf{Var}(N_i + N_j) = \mathbf{Var}(N_i) + \mathbf{Var}(N_j) + 2\mathbf{Cov}(N_i, N_j)$. On en déduit

$$\begin{aligned}
 \mathbf{Cov}(N_i, N_j) &= \frac{1}{2} \left[\mathbf{Var}(N_i + N_j) - \mathbf{Var}(N_i) + \mathbf{Var}(N_j) \right] \\
 &= \frac{1}{2} \left[n(p_i + p_j)(1 - p_i - p_j) - np_i(1 - p_i) - np_j(1 - p_j) \right] \\
 &= \frac{1}{2} \left[np_i + np_j - np_i^2 - np_j^2 - 2np_i p_j - np_i + np_i^2 - np_j + np_j^2 \right] \\
 &= -np_i p_j
 \end{aligned}$$

En conclusion : $\mathbf{Cov}(N_i, N_j) = -np_i p_j$.

Exercice ALEA.15.12 | Soient $p \in]0, 1[$ et $q = 1 - p$, ainsi que X et Y deux variables

aléatoires indépendantes à valeurs dans \mathbf{N} telle que :

$$\forall k \in \mathbf{N}, \quad \mathbf{P}(X = k) = \mathbf{P}(Y = k) = pq^k.$$

On définit alors $U = \max(X, Y)$ et $V = \min(X, Y)$.

1. Donner la loi conjointe de (U, V) , puis les marginales de U et V .
2. Les variables aléatoires U, V sont-elles indépendantes?

Solution (exercice ALEA.15.12)

1. Nous avons $(U, V)(\Omega) = \{(k, \ell) \in \mathbf{N}^2, k \geq \ell\}$. Soit donc $(k, \ell) \in (U, V)(\Omega)$, alors

$$\mathbf{P}(U = k, V = \ell) = \mathbf{P}(\{X = k, Y = \ell\} \cup \{X = \ell, Y = k\}).$$

Les deux évènements apparaissant dans la réunion sont disjoints uniquement si $k \neq \ell$.

- ▶ Donc si $k \neq \ell$, nous avons $\mathbf{P}(U = k, V = \ell) = \mathbf{P}(X = k, Y = \ell) + \mathbf{P}(X = \ell, Y = k) = 2(pq^k)(pq^\ell) = 2p^2q^{k+\ell}$ puisque $X \perp\!\!\!\perp Y$ et que X, Y ont même loi.
- ▶ Si $k = \ell$, alors $\mathbf{P}(U = k, V = k) = \mathbf{P}(X = k, Y = k) = (pq^k)^2$.

Donc :

$$\forall (k, \ell) \in (U, V)(\Omega), \quad \mathbf{P}(U = k, V = \ell) = \begin{cases} 2p^2q^{k+\ell} & \text{si } \ell < k, \\ p^2q^{2k} & \text{si } k = \ell. \end{cases}$$

Passons à présent aux lois marginales.

- ▶ $U(\Omega) = V(\Omega) = \mathbf{N}$.
- ▶ Soit $k \in U(\Omega)$, alors d'après le cours,

$$\mathbf{P}(U = k) = \sum_{\ell=0}^k \mathbf{P}(U = k, V = \ell).$$

Faisons deux cas. Si $k \geq 1$:

$$\begin{aligned} \mathbf{P}(U = k) &= \sum_{\ell=0}^{k-1} (2p^2q^{k+\ell}) + (pq^k)^2 \\ &= 2p^2q^k \frac{1-q^k}{p} + (pq^k)^2. \end{aligned}$$

Par ailleurs, si $k = 0$:

$$\mathbf{P}(U = 0) = \mathbf{P}(X = 0) \mathbf{P}(Y = 0) = p^2.$$

Donc :

$$\mathbf{P}(U = k) = \begin{cases} 2pq^k + p(p-2)q^{2k} & \text{si } k \neq 0, \\ p^2 & \text{si } k = 0. \end{cases} = \boxed{2pq^k + p(p-2)q^{2k}}.$$

De-même, pour tout $\ell \in \mathbf{N}$,

$$\mathbf{P}(V = \ell) = \sum_{k=\ell}^{\infty} \mathbf{P}(U = k, V = \ell) = \sum_{k=\ell+1}^{\infty} (2p^2q^{k+\ell}) + (pq^\ell)^2.$$

Ainsi,

$$\mathbf{P}(V = \ell) = 2p^2q^\ell q^{\ell+1} \frac{1}{p} + p^2q^{2\ell} = \boxed{2pq^{2\ell+1} + p^2q^{2\ell}}.$$

2. **Non**, car

$$\mathbf{P}(U = 0) \mathbf{P}(V = 0) = p^2(2pq + p^2) = p^3(2q + p) = p^3(2 - p),$$

et

$$\mathbf{P}(U = 0, V = 0) = p^2,$$

mais $p^2 = p^3(2 - p) \iff 1 = p(2 - p) \iff p = 1$ (seule solution dans $[0, 1]$), ce qui est exclu.

Exercice ALEA.15.13 | Loi binomiale négative Une équipe de géologues travaille en zone sismique et installe un sismomètre au sommet d'un volcan. Chaque jour, si ce sismomètre détecte une onde sismique, ce dernier envoie une alerte par satellite à un camp de base situé au pied du volcan. On note X_1 la variable aléatoire réelle correspondant au nombre de jours sans alerte avant une première alerte, par exemple si cette première alerte survient au troisième jour après l'installation du sismomètre alors X_1 prend la valeur 2.

On admet que ce sismomètre envoie au plus une alerte par jour, et que la probabilité que survienne une alerte est constante et égale à p , indépendante des résultats obtenus les jours précédents.

1. **1.1)** Déterminer l'ensemble des valeurs prises par X_1 noté $X_1(\Omega)$, et pour tout $j \in X_1(\Omega)$, calculer $\mathbf{P}(X_1 = j)$.
- 1.2) En déduire que $1 + X_1$ suit la loi géométrique de paramètre p ce qu'on notera dorénavant $X_1 \hookrightarrow \mathcal{G}^*(p)$.
2. Ce sismomètre étant extrêmement sensible, les géologues décident de mener des analyses complémentaires à partir d'une seconde alerte. Aussi on note X_2 la variable aléatoire réelle correspondant au nombre de jours sans alerte avant cette seconde alerte, depuis l'installation du sismomètre. Par exemple, si une première alerte survient au troisième jour après l'installation de ce sismomètre et une seconde au septième jour alors X_2 prend la valeur 5
 - 2.1) Déterminer $(X_1, X_2)(\Omega)$.
 - 2.2) Soit $j \in \mathbf{N}$. Déterminer la loi conditionnelle de X_2 sachant $\{X_1 = j\}$.
 - 2.3) En déduire que :

$$\forall k \in \mathbf{N}, \quad \mathbf{P}(X_2 = k) = (k + 1)q^k p^2.$$

3. On généralise l'étude précédente et pour tout $n \in \mathbf{N}^*$, on note X_n la variable aléatoire réelle correspondant au nombre de jours sans alerte avant une n -ième alerte.

- 3.1) Démontrer par récurrence sur $k \in \mathbf{N}$ que :

$$\forall k \in \mathbf{N}, \quad \binom{n+k}{k} = \sum_{j=0}^k \binom{n+j-1}{j}.$$

- 3.2) Démontrer par récurrence sur $n \in \mathbf{N}^*$ que :

$$\forall k \in \mathbf{N}, \quad \mathbf{P}(X_n = k) = \binom{n+k-1}{k} q^k p^n.$$

Solution (exercice ALEA.15.13)

1. **1.1)** La première alerte peut survenir dès le premier jours, donc $X_1(\Omega) = \mathbf{N}$. Soit $j \in X_1(\Omega)$, alors $\{X_1 = j\}$ signifie que nous avons eu j jours sans alerte puis une alerte au jour $j + 1$, donc :

$$\mathbf{P}(X_1 = j) = q^j p.$$

- 1.2) On a clairement $(X_1 + 1)(\Omega) = \mathbf{N}^*$, donc pour tout $\ell \geq 1$, on a

$$\mathbf{P}(X_1 + 1 = \ell) = \mathbf{P}(X_1 = \ell - 1) = q^{\ell-1} p,$$

$$\text{donc } 1 + X_1 \hookrightarrow \mathcal{G}(p).$$

2. **2.1)** Nécessairement $X_2 \geq X_1$, et on peut éventuellement avoir égalité si deux alertes surviennent l'une après l'autre. Donc

$$(X_1, X_2)(\Omega) = \{(j, k) \in \mathbf{N}^2, j \leq k\}.$$

- 2.2) Soit $j \in \mathbf{N}$. Soit $k \geq j$, alors

$$\mathbf{P}(X_2 = k | X_1 = j) = q^{k-j} p,$$

qui correspond à $k - j$ jours sans alerte puis une alerte.

2.3) On a $X_2(\Omega) = \mathbf{N}$, le cas zéro étant obtenu lorsque deux alertes consécutives ont lieu les deux premiers jours. D'après la formule des probabilités totales, on déduit alors :

$$\begin{aligned} \mathbf{P}(X_2 = k) &= \sum_{j=0}^{\infty} \mathbf{P}(X_2 = k, X_1 = j) \\ &= \sum_{j=0}^k \mathbf{P}(X_2 = k | X_1 = j) \mathbf{P}(X_1 = j) \\ &= \sum_{j=0}^k q^{k-j} p q^j p \\ &= \boxed{\mathbf{P}(X_2 = k) = (k + 1) q^k p^2.} \end{aligned}$$

3. On généralise l'étude précédente et pour tout $n \in \mathbf{N}^*$, on note X_n la variable aléatoire réelle correspondant au nombre de jours sans alerte avant une n -ième alerte.

3.1) Posons $H_k : \binom{n+k}{k} = \sum_{j=0}^k \binom{n+j-1}{j}$ pour tout $k \in \mathbf{N}$.

■ **Initialisation.** $\sum_{j=0}^0 \binom{n+j-1}{j} = \binom{n-1}{0} = 1 = \binom{n}{0}$ donc H_0 est vraie.

■ **Hérédité.** Soit k fixé tel que H_k soit vraie. Alors :

$$\begin{aligned} \sum_{j=0}^{k+1} \binom{n+j-1}{j} &= \sum_{j=0}^k \binom{n+j-1}{j} + \binom{n+k}{k+1} \\ &= \binom{n+k}{k} + \binom{n+k}{k+1}. \end{aligned}$$

De plus, la formule de PASCAL donne $\binom{n+k}{k} + \binom{n+k}{k+1} = \binom{n+k+1}{k+1}$. Donc H_{k+1} est vraie. Donc par principe de récurrence :

$$\boxed{\forall k \in \mathbf{N}, \quad \binom{n+k}{k} = \sum_{j=0}^k \binom{n+j-1}{j}.}$$

3.2) Démontrons par récurrence sur $n \in \mathbf{N}^*$ que :

$$\forall k \in \mathbf{N}, \quad \mathbf{P}(X_n = k) = \binom{n+k-1}{k} q^k p^n.$$

Posons, pour $n \geq 1$: $H_n \ll \mathbf{P}(X_n = k) = \binom{n+k-1}{k} q^k p^n$ pour tout $k \in \mathbf{N}$.

■ **Initialisation.** Pour $n = 1$, nous avons bien

$$\mathbf{P}(X_1 = k) = \binom{k-1}{k} q^k = q^k p,$$

car cela a déjà été prouvé dans une question précédente.

■ **Hérédité.** Supposons la propriété prouvée au rang n , c'est-à-dire :

$$\forall k \in \mathbf{N}, \quad \mathbf{P}(X_n = k) = \binom{n+k-1}{k} q^k p^n.$$

Alors cherchons la loi de X_{n+1} . Soit donc $k \in \mathbf{N}$, nous avons encore d'après la formule des probabilités totales :

$$\begin{aligned} \mathbf{P}(X_{n+1} = k) &= \sum_{j=0}^{\infty} \mathbf{P}(X_{n+1} = k | X_n = j) \mathbf{P}(X_n = j) \\ &= \sum_{j=0}^k q^{k-j} p \mathbf{P}(X_n = j) \\ &= \sum_{j=0}^k q^{k-j} p \binom{n+j-1}{j} q^j p^n \\ &= \binom{n+k}{k} q^k p^{n+1}. \end{aligned}$$

la loi condi. se calcule comme dans une question précédente, $k - j$ sans alerte puis une alerte
hypothèse de récurrence
question précédente

Donc par principe de récurrence, nous avons pour tout $n \in \mathbf{N}^*$:

$$\boxed{\mathbf{P}(X_n = k) = \binom{n+k-1}{k} q^k p^n.}$$

Exercice ALEA.15.14 | Soient X, Y, Z trois variables aléatoires indépendantes suivant toutes la loi de POISSON $\mathcal{P}(\lambda)$ avec $\lambda > 0$. Montrer que $\frac{X+Y}{1+Z}$ admet une espérance, et la calculer.

Solution (exercice ALEA.15.14)

On peut voir $\frac{X+Y}{1+Z}$ comme une fonction de trois variables en X, Y, Z mais nous ne disposons pas de théorème de transfert pour ce cadre. En revanche, notons $T = X + Y$ et $g : (x, y) \in \mathbf{R}^2 \mapsto \frac{x}{1+y}$. Alors puisque $X \perp\!\!\!\perp Y$, $T \hookrightarrow \mathcal{P}(2\lambda)$. Et $\frac{X+Y}{1+Z} = h(T, Z)$. Donc $\frac{X+Y}{1+Z}$ admet une espérance si et seulement si

$$\sum_{k, \ell=0}^{\infty} \frac{k}{1+\ell} \mathbf{P}(T = k, Z = \ell) \quad \text{converge absolument.}$$

Ou encore par indépendance et positivité :

$$\sum_{k, \ell=0}^{\infty} \frac{k}{1+\ell} \mathbf{P}(T = k) \mathbf{P}(Z = \ell).$$

Fixons $\ell \in \mathbf{N}$ et étudions

$$\sum_{k=0}^{\infty} \frac{k}{1+\ell} \frac{e^{-2\lambda} (2\lambda)^k}{k!} \mathbf{P}(Z = \ell).$$

Soit $N \in \mathbf{N}$, alors

$$\begin{aligned} \sum_{k=0}^N \frac{k}{1+\ell} \frac{e^{-2\lambda} (2\lambda)^k}{k!} \mathbf{P}(Z = \ell) &= \frac{\mathbf{P}(Z = \ell)}{1+\ell} \sum_{k=0}^N k \frac{e^{-2\lambda} (2\lambda)^k}{k!} \\ &= \frac{\mathbf{P}(Z = \ell)}{1+\ell} e^{-2\lambda} (2\lambda) \sum_{k=1}^N \frac{(2\lambda)^{k-1}}{(k-1)!} \\ &\xrightarrow{N \rightarrow \infty} \frac{\mathbf{P}(Z = \ell)}{1+\ell} e^{-2\lambda} (2\lambda) e^{2\lambda} \\ &= (2\lambda) \frac{e^{-\lambda} \lambda^{\ell}}{(1+\ell)!} \\ &= 2e^{-\lambda} \frac{\lambda^{\ell+1}}{(1+\ell)!}. \end{aligned}$$

Maintenant, on somme en l'autre variable. On sait qu'elle converge aussi d'après le cours sur la série exponentielle, et de plus par changement d'indice :

$$\sum_{\ell=0}^{\infty} 2e^{-\lambda} \frac{\lambda^{\ell+1}}{(1+\ell)!} = 2e^{-\lambda} (e^{\lambda} - 1).$$

Ainsi, la série double étudiée est absolument convergente et par théorème de transfert, $\frac{X+Y}{1+Z}$ admet une espérance égale à :

$$\mathbf{E} \left(\frac{X+Y}{1+Z} \right) = 2e^{-\lambda} (1 - e^{-\lambda}).$$

Exercice ALEA.15.15 | On considère un concessionnaire de voitures. Soit X la variable aléatoire égale au nombre de voitures vendues par jour, on suppose que X suit une loi de POISSON de paramètre $\lambda > 0$. On considère qu'un acheteur de voiture fait un prêt pour se procurer sa voiture avec une probabilité $p \in]0, 1[$. Soit Y la variable aléatoire égale au nombre d'acheteurs par jour de voitures qui font un prêt.

- Déterminer la loi conjointe de (X, Y) .
- Déterminer la loi de Y , puis donner $\mathbf{E}(Y)$ et $\mathbf{Var}(Y)$.
- On pose $Z = X - Y$. Quelle est la loi de Z ? Les variables aléatoires Y, Z sont-elles indépendantes? Calculer $\mathbf{Cov}(X, Y)$.

Solution (exercice ALEA.15.15)

- Il y a forcément moins d'acheteurs faisant un prêt que d'acheteurs. Donc

$$(X, Y)(\Omega) = \{(k, \ell) \in \mathbf{N}^2, k \geq \ell\}.$$

Soit donc $(k, \ell) \in \mathbf{N}^2$ vérifiant $k \geq \ell$. Ici, d'après l'énoncé, nous connaissons la loi du nombre d'acheteurs faisant un prêt conditionnellement à $\{X = k\}$, qui est une $\mathcal{B}(k, p)$: en effet, il s'agit de compter le nombre de succès (contraction d'un prêt) dans une répétition de k expériences de BERNOUILLI indépendantes. Donc

$$\mathbf{P}(Y = \ell | X = k) = \binom{k}{\ell} p^{\ell} (1-p)^{k-\ell}.$$

Donc, d'après la définition d'une probabilité conditionnelle :

$$\begin{aligned} \mathbf{P}(Y = \ell, X = k) &= \mathbf{P}(Y = \ell | X = k) \mathbf{P}(X = k) \\ &= \binom{k}{\ell} p^\ell (1-p)^{k-\ell} e^{-\lambda} \frac{\lambda^k}{k!} \\ &= \frac{e^{-\lambda}}{\ell!(k-\ell)!} (\lambda p)^\ell (\lambda(1-p))^{k-\ell}. \end{aligned}$$

2. On a $Y(\Omega) = \mathbf{N}$ et d'après la formule des probabilités totales, nous déduisons pour tout $\ell \in \mathbf{N}$

$$\begin{aligned} \mathbf{P}(Y = \ell) &= \sum_{k=0}^{\infty} \mathbf{P}(Y = \ell, X = k) \\ &= \sum_{k=\ell}^{\infty} \frac{e^{-\lambda}}{\ell!(k-\ell)!} (\lambda p)^\ell (\lambda(1-p))^{k-\ell} \\ &= e^{-\lambda} \frac{(\lambda p)^\ell}{\ell!} \sum_{k=\ell}^{\infty} \frac{(\lambda(1-p))^{k-\ell}}{(k-\ell)!} \\ &= e^{-\lambda} \frac{(\lambda p)^\ell}{\ell!} \sum_{k=0}^{\infty} \frac{(\lambda(1-p))^k}{k!} \quad \left. \vphantom{\sum_{k=0}^{\infty}} \right\} \text{changement d'indice} \\ &= e^{-\lambda} \frac{(\lambda p)^\ell}{\ell!} e^{-\lambda(1-p)} \\ &= e^{-\lambda p} \frac{(\lambda p)^\ell}{\ell!}. \end{aligned}$$

Ainsi, $Y \hookrightarrow \mathcal{P}(\lambda p)$.

3. On pose $Z = X - Y$. Il s'agit dans notre expérience du nombre d'acheteurs n'ayant pas contracté de prêt. On peut donc utiliser par analogie les résultats précédents en remplaçant p par $1-p$ (ce qui revient dans le compte de succès à inverser succès et échec). Ainsi, $Z \hookrightarrow \mathcal{P}(\lambda(1-p))$.

Les variables aléatoires Y, Z ne sont clairement pas indépendantes, car il y a une «balance» entre les acheteurs empruntant et les autres. Pour le voir, on peut par exemple écrire :

$$\mathbf{P}(Z = 1, Y = 0) = \mathbf{P}(X = 1, Y = 0) = e^{-\lambda} \lambda (1-p),$$

alors que

$$\mathbf{P}(Z = 1) \mathbf{P}(Y = 0) = e^{-\lambda(1-p)} e^{-\lambda p}.$$

La covariance existe puisque X, Y admettent une variance. On a au choix deux méthodes, la première est à privilégier.

► **(Méthode 1 : utiliser la loi de Z)** On a par bilinéarité de la covariance :

$$\begin{aligned} \mathbf{Var}(X - Y) &= \mathbf{Cov}(X - Y, X - Y) \\ &= \mathbf{Var}(X) - \mathbf{Cov}(X, Y) - \mathbf{Cov}(Y, X) + \mathbf{Var}(Y) \\ &= \mathbf{Var}(X) + \mathbf{Var}(Y) - 2\mathbf{Cov}(X, Y). \end{aligned}$$

Ainsi,

$$\begin{aligned} \mathbf{Cov}(X, Y) &= \frac{1}{2} (\mathbf{Var}(X) + \mathbf{Var}(Y) - \mathbf{Var}(Z)) \\ &= \frac{1}{2} (\lambda + \lambda p - \lambda(1-p)) \\ &= \lambda p. \end{aligned}$$

► **(Méthode 2 : KÖNIG-HUYGENS)** On peut sinon revenir à KÖNIG-HUYGENS,

mais les calculs sont beaucoup plus longs. On $E(X) = \lambda$, $E(Y) = \lambda p$. Et

$$\begin{aligned}
 E(XY) &= \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} k\ell \mathbf{P}(X=k, Y=\ell) \\
 &= \sum_{k=0}^{\infty} \sum_{\ell=0}^k k\ell \frac{e^{-\lambda}}{\ell!(k-\ell)!} (\lambda p)^\ell (\lambda(1-p))^{k-\ell} \\
 &= e^{-\lambda} \sum_{k=0}^{\infty} k \sum_{\ell=1}^k \left(\frac{1}{(\ell-1)!(k-\ell)!} (\lambda p)^\ell (\lambda(1-p))^{k-\ell} \right) \\
 &= \lambda p e^{-\lambda} \sum_{k=1}^{\infty} \frac{k}{(k-1)!} \sum_{\ell=1}^k \binom{k-1}{\ell-1} (\lambda p)^{\ell-1} (\lambda(1-p))^{(k-1)-(\ell-1)} \quad \left. \begin{array}{l} \text{chan-} \\ \text{gement} \\ \text{d'indice} \end{array} \right\} \\
 &= \lambda p e^{-\lambda} \sum_{k=1}^{\infty} \frac{k}{(k-1)!} \sum_{\ell=0}^{k-1} \binom{k-1}{\ell} (\lambda p)^\ell (\lambda(1-p))^{(k-1)-\ell} \quad \left. \begin{array}{l} \text{somme} \\ \text{interne} \\ \text{binôme} \end{array} \right\} \\
 &= \lambda p e^{-\lambda} \sum_{k=1}^{\infty} \frac{k}{(k-1)!} (\lambda p + \lambda(1-p))^{k-1} \\
 &= \lambda p e^{-\lambda} \sum_{k=1}^{\infty} \frac{k-1+1}{(k-1)!} \lambda^{k-1} \\
 &= \lambda p e^{-\lambda} \left(\lambda \sum_{k=2}^{\infty} \frac{1}{(k-2)!} \lambda^{k-2} + \sum_{k=1}^{\infty} \frac{1}{(k-1)!} \lambda^{k-1} \right) \quad \left. \begin{array}{l} \text{changement d'indice +} \\ \text{série exp.} \end{array} \right\} \\
 &= \lambda p e^{-\lambda} (\lambda e^{-\lambda} + e^{-\lambda}) \\
 &= \lambda p (\lambda + 1).
 \end{aligned}$$

Donc par la formule de KÖNIG-HUYGENS :

$$\mathbf{Cov}(X, Y) = \lambda p (\lambda + 1) - \lambda^2 p = \lambda p.$$

Exercice ALEA.15.16 | Produits de lois RADEMACHER Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires indépendantes telles que $\mathbf{P}(X_n = -1) = p$ et $\mathbf{P}(X_n = 1) = 1 - p$ avec $p \in]0, 1[$. On pose $Z_n = \prod_{i=1}^n X_i$.

1. Calculer l'espérance de Z_n .

2. Dédurre de la question précédente la loi de Z_n .

3. Donner une condition sur p pour que Z_1 et Z_2 soient indépendantes.

Solution (exercice ALEA.15.16)

1. Par mutuelle indépendance des X_i , on a $E(\prod_{i=1}^n X_i) = \prod_{i=1}^n E(X_i) \iff E(Z_n) = (E(X_1))^n$. Or $E(X_1) = \mathbf{P}(X_1 = 1) - \mathbf{P}(X_1 = -1) = 1 - p - p = 1 - 2p$, d'où

$$E(Z_n) = (1 - 2p)^n.$$

2. On a deux égalités : $\mathbf{P}(Z_n = 1) - \mathbf{P}(Z_n = -1) = E(Z_n)$ et $\mathbf{P}(Z_n = 1) + \mathbf{P}(Z_n = -1) = 1$, d'où un système d'équations que l'on peut résoudre :

$$\begin{cases} \mathbf{P}(Z_n = 1) - \mathbf{P}(Z_n = -1) = (1 - 2p)^n \\ \mathbf{P}(Z_n = 1) + \mathbf{P}(Z_n = -1) = 1 \end{cases} \iff \begin{cases} 2\mathbf{P}(Z_n = 1) = (1 - 2p)^n + 1 \\ 2\mathbf{P}(Z_n = -1) = 1 - (1 - 2p)^n \end{cases} \iff \begin{cases} \mathbf{P}(Z_n = 1) = \frac{1+(1-2p)^n}{2} \\ \mathbf{P}(Z_n = -1) = \frac{1-(1-2p)^n}{2} \end{cases}$$

Ainsi, $Z_n(\Omega) = \{-1, 1\}$ et $\mathbf{P}(Z_n = 1) = \frac{1+(1-2p)^n}{2}$, $\mathbf{P}(Z_n = -1) = \frac{1-(1-2p)^n}{2}$.

3. On peut commencer par calculer sous quelle condition la covariance est nulle. Nous avons.

$$\begin{aligned}
 \mathbf{Cov}(Z_1, Z_2) &= \mathbf{Cov}(X_1, X_1 X_2) \\
 &= E(X_1^2 X_2) - E(X_1) E(X_1 X_2) \\
 &= E(X_1^2) E(X_2) - E(X_1)^3 \\
 &= E(X_2) - E(X_1)^3 = 1 - 2p - (1 - 2p)^3 \\
 &= (1 - 2p)4p(1 - p).
 \end{aligned}$$

*indépendance et même loi pour les X_i
 $X_1^2 = 1$*

Or, comme $p \in]0, 1[$, Z_1, Z_2 sont donc non corrélées uniquement pour $p = \frac{1}{2}$. Inversement, montrons que pour $p = \frac{1}{2}$, les variables aléatoires Z_1, Z_2 sont bien

indépendantes. On calcule.

$$\begin{aligned} \mathbf{P}(Z_1 = 1, Z_2 = 1) &= \mathbf{P}(X_1 = 1, X_1 X_2 = 1) = \mathbf{P}(X_1 = 1, X_2 = 1) = \mathbf{P}(X_1 = 1) \mathbf{P}(X_2 = 1) = \frac{1}{4} \\ \mathbf{P}(Z_1 = 1) \mathbf{P}(Z_2 = 1) &= \mathbf{P}(X_1 = 1) \mathbf{P}(X_1 X_2 = 1) \quad \left. \begin{array}{l} \{X_1 X_2 = 1\} \\ \{X_1 = 1, X_2 = 1\} \cup \{X_1 = -1, X_2 = -1\} \end{array} \right\} = \\ &= \frac{1}{2} \left(\frac{1}{4} + \frac{1}{4} \right) \\ &= \frac{1}{4}. \end{aligned}$$

On passe aux autres vérifications.

$$\begin{aligned} \mathbf{P}(Z_1 = -1, Z_2 = 1) &= \mathbf{P}(X_1 = -1, X_1 X_2 = 1) = \mathbf{P}(X_1 = -1, X_2 = -1) = \frac{1}{4} \\ \mathbf{P}(Z_1 = -1) \mathbf{P}(Z_2 = 1) &= \mathbf{P}(X_1 = -1) \mathbf{P}(X_1 X_2 = 1) \quad \left. \begin{array}{l} \{X_1 X_2 = 1\} \\ \{X_1 = 1, X_2 = 1\} \cup \{X_1 = -1, X_2 = -1\} \end{array} \right\} = \\ &= \frac{1}{2} \left(\frac{1}{4} + \frac{1}{4} \right) \\ &= \frac{1}{4}. \end{aligned}$$

$$\begin{aligned} \mathbf{P}(Z_1 = 1, Z_2 = -1) &= \mathbf{P}(X_1 = 1, X_1 X_2 = -1) = \mathbf{P}(X_1 = 1, X_2 = -1) = \frac{1}{4} \\ \mathbf{P}(Z_1 = 1) \mathbf{P}(Z_2 = -1) &= \mathbf{P}(X_1 = 1) \mathbf{P}(X_1 X_2 = -1) \quad \left. \begin{array}{l} \{X_1 X_2 = -1\} \\ \{X_1 = 1, X_2 = -1\} \cup \{X_1 = -1, X_2 = 1\} \end{array} \right\} = \\ &= \frac{1}{2} \left(\frac{1}{4} + \frac{1}{4} \right) \\ &= \frac{1}{4}. \end{aligned}$$

Et enfin :

$$\begin{aligned} \mathbf{P}(Z_1 = -1, Z_2 = -1) &= \mathbf{P}(X_1 = -1, X_2 = 1) = \frac{1}{4} \\ \mathbf{P}(Z_1 = -1) \mathbf{P}(Z_2 = -1) &= \mathbf{P}(X_1 = -1) \mathbf{P}(X_1 X_2 = -1) \quad \left. \begin{array}{l} \{X_1 X_2 = -1\} \\ \{X_1 = 1, X_2 = -1\} \cup \{X_1 = -1, X_2 = 1\} \end{array} \right\} = \\ &= \frac{1}{2} \left(\frac{1}{4} + \frac{1}{4} \right) \\ &= \frac{1}{4}. \end{aligned}$$

Donc Z_1, Z_2 sont indépendantes si et seulement si $p = \frac{1}{2}$.

Exercice ALEA.15.17 | Quand l'indépendance n'est pas loin Soit (X_i) une suite de variables de BERNOULLI de même paramètre p , indépendantes. Pour tout $i \in \mathbf{N}$, on pose $Y_i = X_i X_{i+1}$.

1. Quelle est la loi de Y_i pour tout $i \in \mathbf{N}$?
2. Soit $n \in \mathbf{N}$, on pose : $S_n = \sum_{i=1}^n Y_i$.

2.1) Démontrer que :

$$\mathbf{Var}(S_n) = \sum_{i=1}^n \mathbf{Var}(Y_i) + 2 \sum_{1 \leq i < j \leq n} \mathbf{Cov}(Y_i, Y_j).$$

2.2) Déterminer $\mathbf{E}(S_n)$ et $\mathbf{Var}(S_n)$.

Solution (exercice ALEA.15.17)

1. On voit que le support de Y_i est $\{0, 1\}$ pour tout i . De plus, $\mathbf{P}(Y_i = 1) = \mathbf{P}(X_i = 1) \mathbf{P}(X_{i+1} = 1) = p^2$ par indépendance. Ainsi, $\mathbf{P}(Y_i = 0) = 1 - p^2$ et : $Y_i \hookrightarrow \mathcal{B}(p^2)$.
2. L'espérance est linéaire, on en déduit immédiatement que : $\mathbf{E}(S_n) = np^2$.
En revanche, étant donné que les variables aléatoires Y_i ne sont pas indépendantes, la variance se développe avec des termes de covariance :

$$\begin{aligned} \mathbf{Var}(S_n) &= \sum_{i=1}^n \mathbf{Var}(Y_i) + 2 \sum_{1 \leq i < j \leq n} \mathbf{Cov}(Y_i, Y_j) \\ &= np^2(1 - p^2) + 2 \sum_{1 \leq i < j \leq n} \mathbf{Cov}(Y_i, Y_j). \end{aligned}$$

Intéressons-nous aux termes de covariance. On remarque que dès que $j > i + 1$, alors les variables $Y_i = X_i X_{i+1}$ et $Y_j = X_j X_{j+1}$ sont indépendantes (elles n'ont aucun facteur X_ℓ commun, et on peut alors appliquer le lemme des coalitions).

Ainsi,

$$\sum_{1 \leq i < j \leq n} \mathbf{Cov}(Y_i, Y_j) = \sum_{1 \leq i < n} \mathbf{Cov}(Y_i, Y_{i+1}) = \sum_{1 \leq i < n} (\mathbf{E}(Y_i Y_{i+1}) - p^2).$$

Et $\mathbf{E}(Y_i Y_{i+1}) = 1 \times \mathbf{P}(Y_i Y_{i+1} = 1) = \mathbf{P}(Y_i = 1, Y_{i+1} = 1) = \mathbf{P}(X_i X_{i+1} = 1, X_{i+1} X_{i+2} = 1) = \mathbf{P}(X_i = 1, X_{i+1} = 1, X_{i+2} = 1) = p^3$.

D'où finalement : $\mathbf{Var}(S_n) = np^2(1-p^2) + 2(n-1)p^3(1-p)$.

4.4.2. Cas continu

Exercice ALEA.15.18 | On considère U et V deux variables indépendantes suivant la loi uniformes sur [0, 1]

- Déterminer la loi de $X = -\ln(U)$ et de $Y = -\ln(V)$.
- Déterminer la loi de $Z = X + Y$.
- Déterminer la loi de $T = e^Z$, et en déduire la loi de $\frac{1}{UV}$.

Exercice ALEA.15.19 | Autour de la loi Gamma Soit $n \in \mathbf{N}^*$. On pose

$$\Gamma(n) = \int_0^{+\infty} t^{n-1} e^{-t} dt.$$

- Montrer par récurrence que pour tout $n \in \mathbf{N}^*$, $\Gamma(n)$ est convergente.
- Calculer $\Gamma(n)$ en fonction de n .
- On pose $f_n(t) = \begin{cases} \frac{1}{\Gamma(n)} t^{n-1} e^{-t} & \text{si } t \geq 0 \\ 0 & \text{sinon.} \end{cases}$ Montrer que f_n est une densité.
- Soit X_n une variable aléatoire de densité f_{n+1} . Montrer que pour tout $x > 0$, si Y_x suit une loi de Poisson de paramètre x , on a :

$$\mathbf{P}(X_n \leq x) = 1 - \mathbf{P}(Y_x \leq n).$$

Indication : On pourra chercher à dériver cette inégalité.

Solution (exercice ALEA.15.19)

- Montrer la convergence par récurrence.

■ **Initialisation.** Soit $A > 0$, alors $\int_0^A 1e^{-t} dt = 1 - e^{-A} \xrightarrow{A \rightarrow \infty} 1$. Donc $\Gamma(1)$ converge et vers 1.

■ **Hérédité.** Supposons que $\Gamma(n)$ converge, alors la fonction $A \mapsto \int_0^A t^{n-1} e^{-t} dt$ admet une limite quand $A \rightarrow \infty$. Alors, puisque les fonctions $t \mapsto t^n, t \mapsto -e^{-t}$ sont de classe \mathcal{C}^1 , on obtient par intégration par parties :

$$\int_0^A t^n e^{-t} dt = -n \int_0^A t^{n-1} (-e^{-t}) dt + [-t^n e^{-t}]_0^A.$$

Mais $[-t^n e^{-t}]_0^A \xrightarrow{A \rightarrow \infty} 0$ par croissances comparées, et par hypothèse de récurrence, $-n \int_0^A t^{n-1} (-e^{-t}) dt \xrightarrow{A \rightarrow \infty} n\Gamma(n)$. Donc $\Gamma(n+1)$ converge.

Finalement, par principe de récurrence, on a montré que $\Gamma(n)$ converge pour tout $n \in \mathbf{N}^*$ et $\Gamma(n) = (n-1)\Gamma(n-1)$.

- Dans la question précédente, nous avons montré que

$$\forall n \in \mathbf{N}^*, \quad \Gamma(n+1) = n\Gamma(n).$$

On établit alors facilement par récurrence que $\Gamma(n) = (n-1)!\Gamma(1)$. Mais $\Gamma(1) = 1$ comme nous l'avons montré dans la question précédente. Donc

$$\forall n \in \mathbf{N}^*, \quad \Gamma(n) = (n-1)!.$$

- La fonction f_n est continue sauf peut-être en zéro, et f_n est positive puisque Γ paraît comme l'intégrale d'une fonction positive pour tout n donc est positive. Et enfin, d'après le résultat des questions précédentes, $\int_0^1 t^{n-1} e^{-t} dt$ converge et vaut $\Gamma(n)$. Donc l'intégrale totale vaut par linéarité de l'intégrale $\frac{\Gamma(n)}{\Gamma(n)} = 1$. Donc f_n est bien une densité de probabilité.
- L'inégalité dérivée est la suivante, par linéarité de la dérivation et puisque X_n est à densité f_{n+1} :

$$\forall x > 0, \quad f_n(x) = - \sum_{k=0}^n \frac{d\left(e^{-x} \frac{x^k}{k!}\right)}{dx}.$$

Soit en utilisant la dérivée d'un produit :

$$\forall x > 0, \quad f_n(x) = - \sum_{k=0}^n e^{-x} \left(\frac{kx^{k-1}}{k!} - \frac{x^k}{k!} \right) = -e^{-x} \left(\sum_{k=1}^n \left(\frac{x^{k-1}}{(k-1)!} - \frac{x^k}{k!} \right) + (0-1) \right).$$

Donc par télescopage et puisque $\Gamma(n + 1) = n!$, on déduit l'égalité suivante à démontrer :

$$\forall x > 0, \quad f_n(x) = e^{-x} \frac{x^n}{\Gamma(n + 1)}.$$

On retrouve bien l'expression de f_n sur $\mathbf{R}^{+\ast}$. Il ne reste plus qu'à primitiver pour revenir à l'égalité initiale, on déduit l'existence d'une constante $C \in \mathbf{R}$ de sorte que :

$$\forall x > 0, \quad \mathbf{P}(X_n \leq x) = 1 - \mathbf{P}(Y_x \leq n) + C.$$

Faisons $x \xrightarrow{x \rightarrow 0}$: on obtient par continuité de la fonction de répartition,

$$\mathbf{P}(X_n \leq 0) = 0 = 1 - \lim_{x \rightarrow 0} \mathbf{P}(Y_x \leq n) + C.$$

Mais

$$\mathbf{P}(Y_x \leq n) = \sum_{k=0}^n e^{-x} \frac{x^k}{k!},$$

donc

$$\lim_{x \rightarrow 0} \mathbf{P}(Y_x \leq n) = 1 + 0.$$

On trouve alors $C = 0$. La formule est donc établie :

$$\boxed{\forall x > 0, \quad \mathbf{P}(X_n \leq x) = 1 - \mathbf{P}(Y_x \leq n)}.$$

Exercice ALEA.15.20 | Loi du χ^2 de degré 2 Soient X_1, X_2 deux variables aléatoires mutuellement indépendantes, suivant toutes la loi normale centrée réduite. On note

$$Z = \frac{1}{2}(X_1^2 + X_2^2).$$

1. Montrer que $\frac{1}{2}X_1^2$ suit une loi usuelle que l'on déterminera.

2. En déduire une densité de Z . Le résultat fera intervenir une intégrale qu'on ne cherchera pas à calculer.

Solution (exercice ALEA.15.20)

1. Commençons par déterminer la loi de X_1^2 . Soit $x \in \mathbf{R}^+$, alors

$$\begin{aligned} \mathbf{P}\left(\frac{1}{2}X_1^2 \leq x\right) &= \mathbf{P}\left(|X_1| \leq \sqrt{2x}\right) \\ &= \Phi\left(\sqrt{2x}\right) - \Phi\left(-\sqrt{2x}\right) \\ &= 2\Phi\left(\sqrt{2x}\right) - 1. \end{aligned}$$

Et donc :

$$F_{X_1^2/2}(x) = \begin{cases} 2\Phi\left(\sqrt{2x}\right) - 1 & \text{si } x > 0, \\ 0 & \text{sinon.} \end{cases}$$

Donc comme Φ est continue, on a $2\Phi\left(\sqrt{2x}\right) - 1 \xrightarrow{x \rightarrow 0} 2\Phi(0) - 1 = 1 - 1 = 0$, et Φ est \mathcal{C}^1 donc par composition, $F_{X_1^2/2}$ est continue et \mathcal{C}^1 sauf peut-être en zéro. Ainsi, $\frac{1}{2}X_1^2$ est à densité dont une densité est

$$f_{X_1^2/2} : x \mapsto \begin{cases} 2\varphi(\sqrt{2x}) \frac{d(\sqrt{2x})}{dx} & \text{si } x > 0, \\ 0 & \text{sinon.} \end{cases}$$

Ce qui livre en utilisant l'expression de la densité de la $\mathcal{N}(0, 1)$:

$$\boxed{f_{X_1^2/2} : x \mapsto \begin{cases} \frac{1}{\sqrt{\pi x}} e^{-x} & \text{si } x > 0, \\ 0 & \text{sinon.} \end{cases}}$$

2. Pour simplifier on notera $f = f_{X_1^2/2} = f_{X_2^2}$ puisque $\frac{X_1^2}{2}$ et $\frac{X_2^2}{2}$ ont même loi. Par lemme des coalitions, $\frac{1}{2}X_1^2 \perp\!\!\!\perp \frac{1}{2}X_2^2$. Donc Z admet une densité, donnée par

$$\forall z \in \mathbf{R}, \quad f_Z(z) = \int_{-\infty}^{\infty} f(x)f(z-x) dx.$$

Par ailleurs, pour tout x réel, $f(x)f(z-x) \neq 0$ si et seulement si $x > 0$ et $z-x > 0$, soit encore si et seulement si : $x > 0, x < z$. Deux cas se présentent alors. Si $z < 0$, on a $f_Z(z) = 0$, reste ensuite les cas $z \geq 0$:

$$f_Z(z) = \int_0^z \frac{1}{\sqrt{\pi x}} e^{-x} \frac{1}{\sqrt{\pi(z-x)}} e^{-(z-x)} dx$$

$$= \frac{e^{-z}}{\pi} \int_0^z \frac{1}{\sqrt{x(x-z)}} dx$$

En résumé :

$$\forall z \in \mathbf{R}, f_Z(z) = \begin{cases} \frac{e^{-z}}{\pi} \int_0^z \frac{1}{\sqrt{x(x-z)}} & \text{si } z > 0, \\ 0 & \text{sinon.} \end{cases}$$

Exercice ALEA.15.21 | Matrice aléatoire uniforme Soient X, Y deux variables aléatoires réelles à densité de loi $\mathcal{U}([0, 1])$ et indépendantes. On note F_X et F_Y les fonctions de répartitions associées.

1. Montrer que $Z = X^2 - Y$ admet la fonction h ci-dessous pour densité :

$$\forall x \in \mathbf{R}, h(x) = \begin{cases} \sqrt{x+1} & \text{si } -1 \leq x < 0, \\ 1 - \sqrt{x} & \text{si } 0 \leq x \leq 1, \\ 0 & \text{sinon.} \end{cases}$$

2. Déterminer la probabilité pour que la matrice aléatoire $M = \begin{pmatrix} 0 & -1 \\ Y & 2X \end{pmatrix}$ soit diagonalisable dans $\mathfrak{M}_2(\mathbf{R})$.

Solution (exercice ALEA.15.21)

1. On a par stabilité de la loi uniforme, que $-Y \leftrightarrow \mathcal{U}[-1, 0]$. Cherchons à présent une densité de X^2 . On a clairement $\mathbf{P}(X^2 \in [0, 1]) = 1$. Soit donc $x \in [0, 1]$,

$$\mathbf{P}(X^2 \leq x) = \mathbf{P}(-\sqrt{x} \leq X \leq \sqrt{x})$$

$$= F_X(\sqrt{x}) - F_X(-\sqrt{x})$$

$$= F_X(\sqrt{x}) = \sqrt{x} \quad \left. \vphantom{\mathbf{P}(X^2 \leq x)} \right\} \text{car } F_X \text{ est nulle sur } \mathbf{R}^-$$

En résumé,

$$\forall x \in \mathbf{R}, F_{X^2}(x) = \begin{cases} 1 & \text{si } x > 1, \\ \sqrt{x} & \text{si } x \in [0, 1], \\ 0 & \text{sinon.} \end{cases}$$

Or $\sqrt{x} \xrightarrow{x \rightarrow 0} 0, \sqrt{x} \xrightarrow{x \rightarrow 1} 1$, la fonction F_{X^2} est continue. De plus, elle est \mathcal{C}^1 sauf peut-être en 0 et 1. Donc X^2 est à densité donnée par

$$\forall x \in \mathbf{R}, f_{X^2}(x) = \begin{cases} \frac{1}{2\sqrt{x}} & \text{si } x \in]0, 1], \\ 0 & \text{sinon.} \end{cases}$$

Puisque $X \perp\!\!\!\perp Y$ on a par coalitions : $X^2 \perp\!\!\!\perp -Y$. Donc $X^2 - Y$ est à densité, et une densité est

$$\forall z \in \mathbf{R}, f_Z(z) = \int_{-\infty}^{\infty} f_{X^2}(x) f_{-Y}(z-x) dx$$

Par ailleurs,

$$f_{X^2}(x) f_{-Y}(z-x) \neq 0 \iff x \in [0, 1], z-x \in [-1, 0],$$

ou de manière équivalente $0 \leq x \leq 1, z \leq x \leq z+1$. Plusieurs cas se présentent alors.

$$\forall z \in \mathbf{R}, f_Z(z) = \begin{cases} \int_0^{z+1} \frac{1}{2\sqrt{x}} dx & \text{si } z+1 \in [0, 1] \\ \int_z^1 \frac{1}{2\sqrt{x}} dx & \text{si } z \in [0, 1] \\ 0 & \text{sinon.} \end{cases}$$

Ce qui livre après calculs des intégrales :

$$\forall z \in \mathbf{R}, f_Z(z) = h(z) = \begin{cases} \sqrt{z+1} & \text{si } -1 \leq z < 0, \\ 1 - \sqrt{z} & \text{si } 0 \leq z \leq 1, \\ 0 & \text{sinon.} \end{cases}$$

2. Commençons par déterminer les valeurs propres de $M = \begin{pmatrix} 0 & -1 \\ Y & 2X \end{pmatrix}$. Soit $\lambda \in \mathbf{R}$, alors :

$$\det \begin{pmatrix} -\lambda & -1 \\ Y & 2X - \lambda \end{pmatrix} = -\lambda(2X - \lambda) + Y = \lambda^2 - 2X\lambda + Y.$$

Donc notons $\Delta = 4X^2 - 4Y = 4(X^2 - Y)$. Alors on a l'égalité d'évènements :

$$\{M \text{ diagonalisable sur } \mathfrak{M}_{2,2}(\mathbf{R})\} = \{\Delta > 0\},$$

puisque si $\Delta < 0$ elle n'a même pas de valeur propre réelle, et si $\Delta = 0$ elle aurait une unique valeur propre et donc serait de la forme XI_2 ce qui est exclu. Mais

$$\mathbf{P}(\Delta > 0) = \mathbf{P}(X^2 - Y > 0) = \int_0^1 (1 - \sqrt{z}) = \left[z - \frac{z^{3/2}}{3/2} \right]_0^1 = \boxed{\frac{1}{3}}.$$

Exercice ALEA.15.22 | On considère deux variables aléatoires réelles X et Y définies sur le même espace probabilisé, indépendantes, et de même loi $\mathcal{E}(1)$.

1. Soit $t > 0$. Montrer que $Y - tX$ admet pour densité h définie par :

$$\forall x \in \mathbf{R}, h(x) = \begin{cases} \frac{e^{-x}}{t+1} & \text{si } x > 0, \\ \frac{e^{x/t}}{t+1} & \text{si } x \leq 0. \end{cases}$$

2. En déduire une densité de $Z = \frac{Y}{X}$.

Solution (exercice ALEA.15.22)

1. Soit $t > 0$. Alors $\mathbf{P}(-tX \leq 0) = 1$, donc pour tout $x \in \mathbf{R}^{-*}$,

$$\begin{aligned} \mathbf{P}(-tX \leq x) &= \mathbf{P}\left(X \geq -\frac{x}{t}\right) = 1 - F_X\left(-\frac{x}{t}\right) \\ &= \begin{cases} 1 - (1 - e^{\frac{x}{t}}) & \text{si } -\frac{x}{t} \geq 0, \\ 1 - 0 & \text{sinon,} \end{cases} \\ &= \begin{cases} e^{\frac{x}{t}} & \text{si } x \leq 0, \\ 1 & \text{sinon.} \end{cases} \end{aligned}$$

Puisque $e^{\frac{x}{t}} \xrightarrow{x \rightarrow 0} 1$, la fonction F_{-tX} est continue par composition, de plus elle est \mathcal{C}^1 sauf peut-être en zéro. Donc $-tX$ est à densité et une densité est donnée par :

$$\forall x \in \mathbf{R}, f_{-tX}(x) = \begin{cases} \frac{1}{t} e^{\frac{x}{t}} & \text{si } x \leq 0, \\ 0 & \text{sinon.} \end{cases}$$

Par lemme des coalitions, $Y \perp\!\!\!\perp -tX$. Donc $Y - tX$ est à densité et admet pour densité h définie par :

$$\forall x \in \mathbf{R}, h(x) = \int_{-\infty}^{\infty} f_Y(z) f_{-tX}(x - z) dz.$$

Mais $f_Y(z) f_{-tX}(x - z) \neq 0$ si et seulement si $z \geq 0, x - z \leq 0$ soit $z \geq 0, x \leq z$. Deux cas se présentent alors.

$$\begin{aligned} h(x) &= \begin{cases} \int_0^{\infty} e^{-z} \frac{1}{t} e^{\frac{x-z}{t}} dt & \text{si } x > 0, \\ \int_0^x e^{-z} \frac{1}{t} e^{\frac{x-z}{t}} dz & \text{si } x \leq 0. \end{cases} \\ &= \boxed{\begin{cases} \frac{e^{-x}}{t+1} & \text{si } x > 0, \\ \frac{e^{x/t}}{t+1} & \text{si } x \leq 0. \end{cases}} \end{aligned}$$

2. Calculons les fonctions de répartition des deux variables aléatoires. Soit $t \in \mathbf{R}^{+\star}$, alors puisque X prend des valeurs positives, on a :

$$\begin{aligned} \mathbf{P}(Z \leq t) &= \mathbf{P}(Y \leq tX) \\ &= \mathbf{P}(Y - tX \leq 0) \\ &= \int_{-\infty}^0 h(x) dx = \int_{-\infty}^0 \frac{e^{x/t}}{t+1} dx \\ &= \frac{t}{t+1} \left[e^{x/t} \right]_{-\infty}^0 \\ &= \frac{t}{t+1}. \end{aligned}$$

Par ailleurs, pour $t \leq 0$, on a évidemment $\mathbf{P}(Z \leq t) = 0$ puisque Z prend ses valeurs dans \mathbf{R}^+ . De plus, puisque $\frac{t}{t+1} \xrightarrow{t \rightarrow 0} 0$, la fonction F_Z est donc continue sur \mathbf{R} et \mathcal{C}^1 sauf éventuellement en zéro. Donc Z est à densité, et une densité est donnée par

$$f_Z : t \mapsto \begin{cases} \frac{1}{(t+1)^2} & \text{si } t \geq 0, \\ 0 & \text{sinon.} \end{cases}$$

Chapitre ALEA.16.

Théorèmes limites probabilistes

Résumé & Plan

Les théorèmes limites en probabilités nous seront utiles dans plusieurs contextes. Le premier est le calcul approché de fonctions de répartition de lois classiques, le second est celui de la Statistique inférentielle.

1	Inégalités de concentration	2
2	Théorèmes limites	5
2.1	Moyenne et variance empirique	5
2.2	Loi faible des grands nombres	6
2.3	Loi limite de la centrée/réduite : le théorème central limite	12
3	Approximations de lois	17
3.1	Conséquences du théorème central limite	17
3.2	Conséquences de calculs directs	18

4	Exercices	24
4.1	Inégalités de concentration	24
4.2	Théorèmes limites & Approximations	24

La méthode de MONTE-CARLO est une appellation générique pour désigner les méthodes algorithmiques permettant de faire des calculs numériques en utilisant le hasard. Le nom de ces méthodes fait allusion aux jeux de hasard pratiqués à MONTE-CARLO, ce quartier de MONACO célèbre pour son casino.

— Le saviez-vous?



Cadre

Dans tout le chapitre, et même lorsque cela n'est pas précisé, on se fixe un espace probabilisé (Ω, \mathcal{F}, P) .

1. INÉGALITÉS DE CONCENTRATION

En probabilités, les *inégalités de concentration* sont des estimations de la probabilité qu'une variable aléatoire s'écarte de sa moyenne, ou par complémentaire de la probabilité qu'elle se concentre autour. Il en existe beaucoup, mais il vous est demandé de connaître l'inégalité de BIENAYMÉ-TCHEBYCHEV uniquement, conséquence elle-même de l'inégalité de MARKOV, au programme elle aussi. Nous les appliquerons souvent à une somme de variables aléatoires indépendantes afin d'estimer l'écart entre la moyenne de n réalisations d'une même variable aléatoire et son espérance : ce qui nous fournira une preuve de la *loi des grands nombres*.

Proposition ALEA.16.1 | Inégalité de MARKOV

Si X est une variable aléatoire discrète (*resp.* à densité) **positive** possédant une espérance, *i.e.*

$$\left(\sum |x| \mathbf{P}(X = x) \right)_{x \in X(\Omega)} \quad \left(\text{resp.} \quad \int_{-\infty}^{\infty} |t| f_X(t) dt \right) \quad \text{converge.}$$

Alors, pour tout $\varepsilon > 0$, on a :

$$\mathbf{P}(X \geq \varepsilon) \leq \frac{\mathbf{E}(X)}{\varepsilon} \quad (\text{Markov})$$

⊗ Attention

Ne pas oublier l'hypothèse de positivité. Par exemple, on ne peut appliquer cette inégalité à $X \hookrightarrow \mathcal{N}(0, 1)$, en revanche, on peut pour $X^2, |X|, X^4, \text{etc.}$

Preuve

1. (Cas discret)

2. (Cas à densité)

Exemple 1 — Soit $X \hookrightarrow \mathcal{E}(\lambda)$ avec $\lambda > 0$, et soit $\varepsilon > 0$.

1. Calculer précisément $\mathbf{P}(X \geq \varepsilon)$.
2. Majorer $\mathbf{P}(X \geq \varepsilon)$ à l'aide de l'inégalité de MARKOV. Commenter la précision de cette inégalité.



Remarque 1.1 — L'inégalité de MARKOV est donc parfois peu précise. Son intérêt est d'être générale et de s'appliquer pour des variables aléatoires d'espérance connue où $\mathbf{P}(X \geq a)$ peut être difficile à calculer précisément.

Exemple 2 — Soit (Y_n) une suite de variables aléatoires réelles i.i.d. suivant une $\mathcal{E}(\lambda)$ avec $\lambda > 1$. Pour tout $n \geq 1$ on pose $X_n = \prod_{k=1}^n Y_k$. Montrer que :

$$\forall \varepsilon > 0, \quad \mathbf{P}(|X_n| \geq \varepsilon) \xrightarrow{n \rightarrow \infty} 0.$$



Corollaire ALEA.16.1 | Inégalité de MARKOV généralisée

Si X est une variable aléatoire discrète (*resp.* à densité) possédant un moment d'ordre k , *i.e.*

$$\left(\sum |x|^k \mathbf{P}(X = x) \right)_{x \in X(\Omega)} \quad \left(\text{resp.} \quad \int_{-\infty}^{\infty} |t|^k f_X(t) dt \right) \quad \text{converge.}$$

Alors, pour tout $\varepsilon > 0$, on a :

$$\mathbf{P}(|X| \geq \varepsilon) \leq \frac{\mathbf{E}(|X|^k)}{\varepsilon^k}. \quad (\text{Markov, G n })$$

Du fait de la pr sence de valeurs absolues, il n'y a plus besoin d'hypoth se de positivit  dans ce corollaire.

Preuve



Enfin, nous pouvons dans l'Eq. (Markov, G n ) pour $k = 2$, faire la substitution suivante :

$$X \longleftarrow X - \mathbf{E}(X).$$

En adaptant l'hypoth se nous obtenons imm diatement la proposition ci-dessous.

Proposition ALEA.16.2 | BIENAYM -TCHEBYCHEV

Si X est une variable al atoire discr te (*resp.*   densit ) poss dant un moment d'ordre 2, *i.e.*

$$\left(\sum |x|^2 \mathbf{P}(X = x) \right)_{x \in \mathbf{X}(\Omega)} \quad \left(\text{resp.} \quad \int_{-\infty}^{\infty} |t|^2 f_X(t) dt \right) \quad \text{converge.}$$


Alors, pour tout $\varepsilon > 0$, on a :

$$\mathbf{P}(|X - \mathbf{E}(X)| \geq \varepsilon) \leq \frac{\mathbf{Var}(X)}{\varepsilon^2}. \quad (\text{BTCH})$$

Preuve Cons quence directe de la d finition de l'esp rance : $\mathbf{E}((X - \mathbf{E}(X))^2) = \mathbf{Var}(X)$ et de Eq. (Markov, G n ).

Remarque 1.2 — Notez que, comme $\{|X - \mathbf{E}(X)| > \varepsilon\} \subset \{|X - \mathbf{E}(X)| \geq \varepsilon\}$ — puisqu'une in galit  stricte est large — on a aussi :

$$\mathbf{P}(|X - \mathbf{E}(X)| > \varepsilon) \leq \mathbf{P}(|X - \mathbf{E}(X)| \geq \varepsilon) \leq \frac{\mathbf{Var}(X)}{\varepsilon^2}.$$

Exemple 3 — Montrer qu'il suffit de lancer un d    six faces  quilibr  27 778 fois pour garantir avec moins de 5% d'erreur que la fr quence d'apparition du 1 sera de $\frac{1}{6}$   1% pr s. 

Exemple 4 – Version « asymptotique » Soit (X_n) une suite de variables aléatoires réelles positives possédant toutes une variance. On suppose que :

$$\lim_{n \rightarrow \infty} \mathbf{E}(X_n) = \mu \in \mathbf{R}, \quad \text{et} \quad \lim_{n \rightarrow \infty} \mathbf{Var}(X_n) = 0.$$

Montrer que :

$$\forall \varepsilon > 0, \quad \mathbf{P}\left(|X_n - \mu| \geq \varepsilon\right) \xrightarrow{n \rightarrow \infty} 0.$$



2. THÉORÈMES LIMITES

Commençons par introduire quelques notations. Ces suites de variables aléatoires seront étudiées plus en détail dans le chapitre de statistiques.

2.1. Moyenne et variance empirique

Définition ALEA.16.1 | Moyenne/Variance empirique

Soient X_1, \dots, X_n une famille de $n \in \mathbf{N}^*$ variables aléatoires réelles. On appelle :

- ▶ *moyenne empirique des X_i (resp. variance empirique)* les variables aléatoires réelles

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \quad \left(\text{resp.} \quad \sigma_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2 \right).$$

On appelle *écart-type empirique* la variable aléatoire

$$\sigma_n \stackrel{\text{(déf.)}}{=} \sqrt{\sigma_n^2}.$$

- ▶ On appelle *fonction de répartition empirique des X_i* la fonction aléatoire

réelle \overline{F}_n définie pour tout $t \in \mathbf{R}$ par :

$$\overline{F}_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{X_i \leq t\}}.$$

« Empirique » signifie « qui ne s’appuie que sur l’expérience, l’observation ». Les quantités dans leur version « empirique » sont donc des réels formés à partir d’observations, et proches de la vraie quantité. Le lien entre les deux est donné par la loi faible des grands nombres que nous allons voir un peu plus tard.

Toutes les quantités probabilistes existent dans une version empirique : la fonction de répartition, l’anti fonction de répartition, les moments d’ordre supérieur, *etc.* Nous en donnons donc seulement quelques unes dans la définition précédente, leur interprétation est la suivante :

- ▶ \overline{X}_n correspondant à la moyenne statistique des observations X_1, \dots, X_n , σ_n^2, σ_n à leur variance et écart-type.
- ▶ Par ailleurs, pour tout $t \in \mathbf{R}$, $\overline{F}_n(t)$ est la proportion des X_1, \dots, X_n qui sont inférieures ou égales à t .

Proposition ALEA.16.3 | Version KÖNIG-HUYGENS

Soient X_1, \dots, X_n une famille de $n \in \mathbf{N}^*$ variables aléatoires réelles. Alors :

$$\sigma_n^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \overline{X}_n^2.$$

Preuve (Point clef — Développer le carré (comme pour toute formule type KÖNIG-HUYGENS))

$$\begin{aligned} \sigma_n^2 &= \frac{1}{n} \sum_{i=1}^n (X_i - \overline{X}_n)^2 \\ &= \frac{1}{n} \sum_{i=1}^n (X_i^2 + \overline{X}_n^2 - 2X_i\overline{X}_n) \\ &= \frac{1}{n} \sum_{i=1}^n X_i^2 + \frac{1}{n} \sum_{i=1}^n \overline{X}_n^2 - 2\frac{\overline{X}_n}{n} \sum_{i=1}^n X_i \\ &= \frac{1}{n} \sum_{i=1}^n X_i^2 + \overline{X}_n^2 - 2\overline{X}_n^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \overline{X}_n^2. \end{aligned}$$

} développement du carré
} linéarité de la somme

Σ Notation

Nous noterons dans la suite aussi parfois $S_n = X_1 + \dots + X_n$ la somme partielle pour tout $n \in \mathbf{N}^*$.

2.2. Loi faible des grands nombres

La loi faible des grands nombres donne une réponse générale à la question suivante : est-ce que la moyenne empirique est proche de l’espérance de la loi commune, si elle existe ? La réponse est oui. Nous énonçons, conformément au programme, la loi faible des grands nombres qui requiert l’existence d’une variance mais on peut s’en passer et exiger seulement l’existence d’une espérance (on parle de « loi forte », en revanche la démonstration est beaucoup plus difficile).

C’est à Jacques BERNOULLI que l’on doit le premier énoncé de la loi des grands nombres ; il apparaît dans son ouvrage *Ars Conjectandi* publié en 1713, huit ans après sa mort. Il avait pour cadre le jeu du pile ou face (schéma de BERNOULLI). Le terme de « loi des grands nombres » est lui dû à POISSON. Ce terme juridique est à mettre en rapport avec le titre de l’ouvrage dans lequel il l’introduit, *Recherches sur les probabilités des jugements*, paru en 1837. De nombreux mathématiciens ont ensuite généralisé les énoncés de BERNOULLI et POISSON, citons notamment KOLMOGOROV et TCHEBYCHEV. L’idée de moyenniser des résultats et de constater que cette moyenne est proche d’une valeur commune lorsque l’on augmente le nombre de simulations est quant à elle très ancienne, avant même l’apparition du formalisme des probabilités actuellement

utilisé — le chemin fut long à parcourir pour arriver aux énoncés ci-après.

Théorème ALEA.16.1 | Loi faible des grands nombres, version espérance

Soit (X_n) une suite de variables aléatoires réelles

- ▶ i.i.d.,¹
- ▶ et possédant toutes une variance commune $\sigma^2 = \text{Var}(X_1)$. On note alors $\mu = \mathbf{E}(X_1)$.

Alors :

$$\forall \varepsilon > 0, \quad \mathbf{P}\left(|\overline{X}_n - \mu| \geq \varepsilon\right) \xrightarrow{n \rightarrow \infty} 0.$$

On dit que la suite (\overline{X}_n) converge en probabilité vers μ .

Preuve



Corollaire ALEA.16.2 | Loi faible des grands nombres, version probabilité

Soit $A \in \mathcal{F}$ un évènement, et $(A_n)_n$ une suite d'évènements de \mathcal{F} telle que les variables aléatoires $(\mathbb{1}_{A_n})$ soient i.i.d. et aient même loi que $\mathbb{1}_A$. Alors :

$$\forall \varepsilon > 0, \quad \mathbf{P}\left(\left|\frac{\sum_{i=1}^n \mathbb{1}_{A_i}}{n} - \mathbf{P}(A)\right| \geq \varepsilon\right) \xrightarrow{n \rightarrow \infty} 0.$$


Remarque 2.1 — Interprétation Pour estimer par simulation une probabilité $\mathbf{P}(A)$, on simule un grand nombre n de fois l'expérience aléatoire associée, et la proportion de fois sur ces n simulations où cet évènement s'est réalisé, *i.e.*

$\frac{\sum_{i=1}^n \mathbb{1}_{A_i}}{n}$ est alors une bonne approximation de $\mathbf{P}(A)$.

Preuve



¹On rappelle que cela signifie qu'elles sont indépendantes, et qu'elles suivent la même loi

Exemple 5 — Soit un dé à six faces équilibré. On le lance 10 fois et on note E l'évènement «on a obtenu au moins 7 entiers pairs sur les 10 lancers». Créer un script Python permettant d'estimer $\mathbf{P}(E)$. 

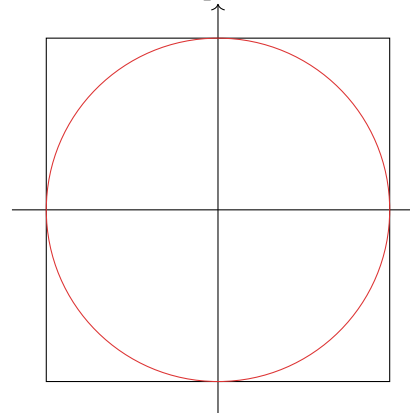
Exemple 6 — *Approximation de π* Soient deux variables aléatoires $X_1, X_2 \hookrightarrow \mathcal{U}[0, 1]$ i.i.d.. On pose $R = X_1^2 + X_2^2$. On admet la convergence et la valeur de

$$\forall z \in]0, 1[, \quad I_z = \int_0^z \frac{1}{\sqrt{x(z-x)}} dx = \pi.$$

1. Estimer, à l'aide de Python, la valeur de la probabilité $\mathbf{P}(R \leq 1) = \mathbf{P}(X_1^2 + X_2^2 \leq 1)$.



2. À votre avis, que vaut $\mathbf{P}(R \leq 1)$? Le montrer, et l'expliquer sur un dessin.



Corollaire ALEA.16.3 | Loi faible des grands nombres, version fonction de répartition

Soit X une variable aléatoire réelle et (X_n) une suite i.i.d. de variables aléatoires de même loi que X . Alors :

$$\forall x \in \mathbf{R}, \forall \varepsilon > 0, \quad \mathbf{P}\left(\left|F_n(x) - F_X(x)\right| \geq \varepsilon\right) \xrightarrow{n \rightarrow \infty} 0.^2$$


Remarque 2.2 – Interprétation Pour estimer par simulation la valeur d'une fonction de répartition en un point x i.e. $F_X(x)$, on simule un grand nombre n de fois la variable aléatoire X , et la proportion de fois sur ces n simulations où $X_i \leq x$, i.e.

$$\frac{\sum_{i=1}^n \mathbb{1}_{\{X_i \leq x\}}}{n}$$
 est alors une bonne approximation de $F_X(x)$.

Preuve



On peut appliquer le corollaire précédent à l'évènement $A = \{X \leq x\}$ pour toute variable aléatoire X et $x \in \mathbf{R}$, on en déduit alors une valeur approchée de la fonction de répartition de X notée F_X , à l'aide de la fonction de répartition empirique :

Exemple 7 – Une usine fabrique 1000 moteurs par jour, la probabilité qu'un moteur donné sortant de l'usine soit défectueux est 0.01, les pannes étant indépendantes. Créer un script Python permettant d'estimer la probabilité qu'au plus 50 moteurs soient en panne en une journée. 

²On peut faire mieux que cette convergence, c'est le résultat du théorème de GLIVENKO-CANTELLI, qui est hors-programme.

**Méthode** >  **Approcher une espérance (ou une probabilité) par simulation**

Soient X, X_1, \dots, X_n une suite i.i.d. d'espérance μ , et admettant une variance.

Alors :

1. $\overline{X}_n \approx \mu$ pour n assez grand, et \overline{X}_n s'obtient en simulant un grand nombre de fois la loi commune aux X_i , $i \in \llbracket 1, n \rrbracket$.

2. **(Pour l'espérance)**

$$\mathbf{E}(X) \approx \overline{X}_n = \frac{\sum_{i=1}^n X_i}{n}, \quad n \text{ assez grand.}$$

On forme la moyenne de simulations X_i . Si cette moyenne ne semble pas converger, alors X n'a probablement pas d'espérance.

3. **(Pour une probabilité/fonction de répartition)** Soit I un intervalle de \mathbf{R} . Alors comme $\mathbf{P}(X \in I) = \mathbf{E}(\mathbb{1}_{\{X \in I\}})$, on peut utiliser **1)** pour approcher la probabilité :

$$\mathbf{P}(X \in I) \approx \frac{\sum_{i=1}^n \mathbb{1}_{\{X_i \in I\}}}{n}, \quad n \text{ assez grand.}$$

On compte le nombre de simulations X_i dans I et on renvoie la moyenne. Pour obtenir une approximation de $F_X(x)$, on compte le nombre de simulations « $\leq x$ ».

Constater la convergence de la loi faible des grands nombres

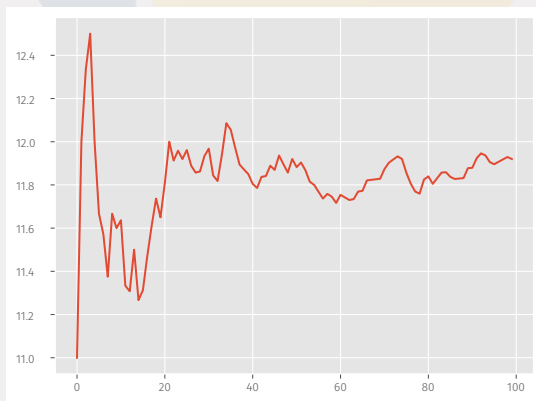
Rappelons que $\overline{X}_n = \frac{1}{n}(X_1 + \dots + X_n)$. Cherchons une relation de récurrence sur \overline{X}_n :

$$\begin{aligned} \overline{X}_n &= \frac{1}{n}(X_1 + \dots + X_n) = \frac{1}{n}(X_1 + \dots + X_{n-1}) + \frac{X_n}{n} \\ &= \frac{n-1}{n} \frac{1}{n-1}(X_1 + \dots + X_{n-1}) + \frac{X_n}{n} \\ &= \frac{n-1}{n} \underbrace{\overline{X}_{n-1}}_{\text{déjà simulé}} + \underbrace{\frac{X_n}{n}}_{\text{à simuler}}. \end{aligned}$$

On résume tout ceci dans la méthode suivante, très importante en pratique, car elle permet d'obtenir des valeurs approchées de quantités probabilistes à l'aide de simulation.

Cette relation permet donc de conserver la trajectoire déjà simulée pour ajouter un nouveau point sur le graphe. Voyons cela sur un exemple.

```
def LFGN_Binomiale(N, n, p):
    """
    Trace la moyenne empirique
    N : nb de simulations
    n,p : paramètres de la binomiale
    """
    X_bar = np.random.binomial(n, p) # simu d'une B(n,p)
    L = [X_bar]
    for k in range(2, N+1):
        X_bar = ((k-1)/k)*X_bar + np.random.binomial(n, p)/k
        L.append(X_bar)
    plt.plot(L)
    return L
LFGN_Binomiale(100, 20, 0.6)
```



On constate bien une convergence vers $0,6 \times 20 = 12$.

MÉTHODE DE MONTE-CARLO POUR L'APPROXIMATION D'INTÉGRALES La loi des grands nombres, dans le cas des variables à densité, nous donne un moyen d'estimer une

intégrale de fonction continue. Voyons comment.

Proposition ALEA.16.4 | Méthode de MONTE-CARLO

Soit $f : [0, 1] \rightarrow \mathbf{R}$ une fonction continue sauf en un nombre fini de points telle que :

$$I = \int_0^1 f(t) dt \text{ et } \int_0^1 f(t)^2 dt \text{ convergent absolument.}$$

Notons $X_i = f(U_i)$ pour tout $i \in \mathbf{N}$, avec $(U_i)_i$ suite i.i.d. de variables aléatoires de loi commune $\mathcal{U}[0, 1]$. On note $\bar{X}_n = \frac{\sum_{i=1}^n X_i}{n} = \frac{\sum_{i=1}^n f(U_i)}{n}$ pour tout $n \in \mathbf{N}^*$. Alors :

1. les X_i admettent toutes une espérance égale à $I = \int_0^1 f(t) dt$ et elles sont indépendantes.
2. Pour tout $\varepsilon > 0$, $\mathbf{P}(|\bar{X}_n - I| \geq \varepsilon) \xrightarrow{n \rightarrow \infty} 0$. Par conséquent, \bar{X}_n est une bonne approximation de I .

Cette méthode probabiliste d'approximation d'intégrale vient donc s'ajouter à celles déjà connues : rectangle gauche ou droite, trapeàzes (méthodes analytiques). La méthode peut s'adapter à d'autres types d'intégrales sur d'autres segments que $[0, 1]$.

Preuve

1. 

2. 

Exemple 8 – On peut déterminer une valeur approchée de $I = \int_0^1 t^2 dt$ à l'aide de Python, on fixera un grand nombre $N = 10^3$ simulations. Bien sûr, dans ce cas on sait même calculer l'intégrale qui vaut $I = \frac{1}{3}$.

```
def f(t):
    return t**2

# Monté-Carlo
def monte_Carlo(f, N):
    X_bar = 0
    for _ in range(N):
        X_bar += f(rd.random())
    return X_bar/N
```

On peut ensuite tester et comparer avec par exemple la méthode des rectangles à gauche.

```
>>> monte_Carlo(f, 10**3)
0.34559465201341527
>>> rectangle_RG(f, 0, 1, 10**3)
0.33283349999999995
```

2.3. Loi limite de la centrée/réduite : le théorème central limite

Le théorème qui suit est dû, dans sa première démonstration pour les lois de BERNOULLI, à Pierre-Simon DE LAPLACE, dans son premier énoncé à Abraham de MOIVRE, et est probablement le résultat le plus remarquable du cours de cette année. En effet, il explique que beaucoup de distributions (*i.e.* d'histogrammes) croisés dans la nature ressemble à une courbe «en cloche», l'aspect remarquable est son hypothèse universelle. En effet, peu importe la loi donnée en entrée (discrète, à densité ou même ni l'un ni l'autre), la centrée réduite d'une moyenne empirique se rapprochera toujours

d'une gaussienne standard. Voyons cela plus en détail.

Soit (X_n) une suite de variables aléatoires réelles i.i.d., et possédant toutes une variance commune $\sigma^2 = \mathbf{Var}(X_1)$. On note alors $\mu = \mathbf{E}(X_1)$. Rappelons que la version centrée/réduite de $\bar{X}_n = \frac{X_1 + \dots + X_n}{n}$, est par définition

$$\begin{aligned} \bar{X}_n^* &= \frac{\bar{X}_n - \mathbf{E}(\bar{X}_n)}{\sqrt{\mathbf{Var}(\bar{X}_n)}} \\ &= \frac{\bar{X}_n - \mu}{\sqrt{n\sigma^2/n^2}} \quad \left. \begin{array}{l} \text{linéarité de l'espérance, et indépendance de } X_1, \dots, X_n \end{array} \right\} \\ &= \frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} = \sqrt{n} \frac{\bar{X}_n - \mu}{\sigma}. \end{aligned}$$

Alors que la version centrée/réduite de $S_n = X_1 + \dots + X_n$, est par définition :



Alors, le théorème central limite, que nous allons énoncer, prétend que la loi de \overline{X}_n^* est «très proche» d'une $\mathcal{N}(0, 1)$ lorsque $n \rightarrow \infty$. Énonçons le résultat complet.

Théorème ALEA.16.2 | Théorème central limite, version convergence en loi

Soit (X_n) une suite de variables aléatoires réelles

i.i.d., et possédant toutes une variance commune $\sigma^2 = \mathbf{Var}(X_1)$.

On note alors $\mu = \mathbf{E}(X_1)$. Alors :

$$\overline{X}_n^* = \frac{\overline{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{S_n - n\mu}{\sqrt{n}\sigma} = S_n^*$$

converge en loi vers une loi normale centrée réduite, i.e. pour tout $(a, b) \in (\mathbf{R} \cup \{\pm\infty\})^2, a \neq b$,

$$\mathbf{P}\left(a \leq \sqrt{n} \frac{\overline{X}_n - \mu}{\sigma} \leq b\right) = \mathbf{P}\left(a \leq \frac{S_n - n\mu}{\sqrt{n}\sigma} \leq b\right) \\ \xrightarrow{n \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{t^2}{2}} dt = \Phi(b) - \Phi(a).$$

Le résultat est admis. Vous pouvez, au choix, appliquer ce théorème à la centrée-réduite de la moyenne empirique, ou la somme partielle. L'essentiel étant de savoir calculer sans erreur l'une ou l'autre des ces centrée-réduites.

Remarque 2.3 —

1. La convergence en loi précisée plus haut est en fait équivalente à la convergence des fonctions de répartition en tout point x de \mathbf{R} :

$$\forall x \in \mathbf{R} \quad F_{\overline{X}_n^*}(x) \xrightarrow{n \rightarrow \infty} F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt.$$

2. Le point remarquable de ce théorème est son universalité : peu importe la loi en entrée de (X_n) — continue ou même discrète! — on a convergence en loi vers une $\mathcal{N}(0, 1)$. Ce fait explique que dans beaucoup de contextes nous obtenons des histogrammes en forme de «cloche».

Constater la convergence du théorème central limite

Étant donné que le résultat du théorème central limite est une convergence en loi, i.e. un «rapprochement des histogrammes» lorsque n devient grand, on :

1. commence par créer un tableau contenant des simulations de la centrée réduite de la loi étudiée.
2. On regroupe ces simulations en classes puis on trace l'histogramme : cela est fait automatiquement avec la commande `plt.hist` où l'on spécifie notamment le nombre de classes souhaité dans l'histogramme.
3. On superpose ce graphique avec celui de la densité de la $\mathcal{N}(0, 1)$, les deux devant être assez proches d'après le théorème central limite.

On réalise les simulations pour l'uniforme sur $[0, 1]$ d'espérance $\frac{1}{2}$ et variance $\frac{1}{12}$.

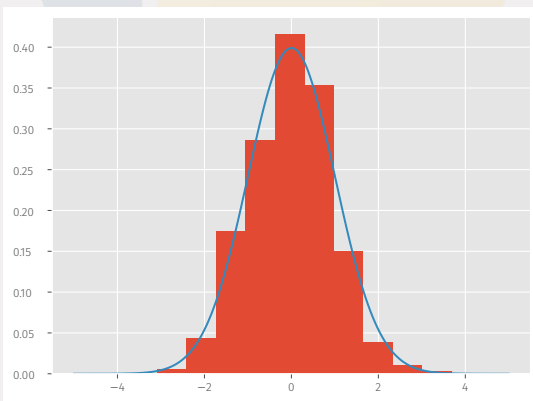
```
import random as rd
import matplotlib.pyplot as plt
import numpy as np
```

```
def unif_star(n):
    """
    retourne une simulation de la centrée réduite de Sn pour
    l'uniforme
    """
    S = 0
```

```
for _ in range(n):
    S += rd.random()
return np.sqrt(n)*(S/n-1/2)/np.sqrt(1/12)
```

```
def TCL_Unif(c, N, n):
    X_star = np.zeros(N)
    for i in range(N):
        X_star[i] = unif_star(n)
    plt.hist(X_star, bins=c, density=True)
    x = np.linspace(-5, 5, 10**3)
    y = np.exp(-x**2/2)/np.sqrt(2*np.pi)
    plt.plot(x, y)
```

```
TCL_Unif(10, 1000, 100)
```



Autrement dit, pour n assez grand,
 ▶ la loi de \bar{X}_n est « proche de » $\mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$.
 ▶ la loi de S_n est « proche de » $\mathcal{N}\left(n\mu, n\sigma^2\right)$.

Attention

Cela n'a pas de sens de dire directement que \bar{X}_n converge en loi vers une $\mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$, tout simplement car les paramètres de la loi normale dépendent de n . On se contentera d'une notation vague « \approx »³ et de préciser sous quelle condition elle est valide.



Méthode Pour retenir l'approximation de \bar{X}_n par une loi normale

Penser aux paramètres : $\mathbf{E}(\bar{X}_n) = \mu$, $\mathbf{Var}(\bar{X}_n) = \frac{\sigma^2}{n}$, « à la limite » on obtient une loi $\mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$ de mêmes paramètres.

Preuve (Point clef — Stabilité de la loi normale)

Nous avons pour tout $n \in \mathbf{N}^*$:

$$\bar{X}_n^* = \sqrt{n} \frac{\bar{X}_n - \mu}{\sigma} \iff \bar{X}_n = \frac{\sigma}{\sqrt{n}} \bar{X}_n^* + \mu.$$



Nous avons pour tout $n \in \mathbf{N}^*$:

$$S_n^* = \frac{S_n - n\mu}{\sigma\sqrt{n}} \iff S_n = \sigma\sqrt{n}S_n^* + n\mu.$$





³Bien sûr, ce symbole n'a aucune légitimité mathématique.


Corollaire ALEA.16.4 | Théorème central limite, version « approximation »

Soit (X_n) une suite de variables aléatoires réelles i.i.d., et possédant toutes une variance commune $\sigma^2 = \mathbf{Var}(X_1)$. On note alors $\mu = \mathbf{E}(X_1)$. Alors pour tout $(a, b) \in (\mathbf{R} \cup \{\pm\infty\})^2$, $a \neq b$,

$$\mathbf{P}\left(a \leq \bar{X}_n \leq b\right) \underset{n \rightarrow \infty}{\approx} \int_a^b \varphi_{\mu, \frac{\sigma^2}{n}}, \quad \mathbf{P}\left(a \leq S_n \leq b\right) \underset{n \rightarrow \infty}{\approx} \int_a^b \varphi_{n\mu, n\sigma^2}.$$

Exemple 9 — Chaque année, M. Durand effectue 2 fois par jour, 5 jours par semaine et pendant 46 semaines, un trajet en voiture dont la durée est une variable aléatoire réelle X qui suit une loi d'espérance 45 min et d'écart type 10 min. On suppose que les durées des trajets sont mutuellement indépendantes. Estimer la probabilité pour que M. Durand passe au moins 350 heures dans sa voiture au cours de l'année. 

Exemple 10 — Une montre fait une erreur relative d'au plus une demi-minute par jour, et que l'erreur commise est uniforme. Déterminer la probabilité que l'erreur commise au bout d'une année non bissextile soit inférieure ou égale à un quart d'heure. 

Exemple 11 — Soit $k \in \mathbf{N}^*$, et $X_k \hookrightarrow \mathcal{U}[0, 1]$ telle que $(X_k)_{k \in \mathbf{N}^*}$ soit une suite i.i.d.. Quelle est la limite de $\mathbf{P}\left(\frac{n}{2} - \sqrt{n} < S_n < \frac{n}{2} + \sqrt{n}\right)$ quand $n \rightarrow +\infty$? 

La seconde forme du théorème central limite consiste en la modification suivante : on peut remplacer σ par sa version empirique, la convergence est alors maintenue. Cette substitution paraît pour le moment anecdotique, mais elle sera d'importance capitale en statistiques. On admet ce résultat également.

Théorème ALEA.16.3 | Théorème central limite en approchant la variance

Soit (X_n) une suite de variables aléatoires réelles

i.i.d., et possédant toutes une variance commune $\sigma^2 = \mathbf{Var}(X_1)$.

On note alors $\mu = \mathbf{E}(X_1)$. Alors $\frac{\bar{X}_n - \mu}{\frac{\sigma_n}{\sqrt{n}}}$ converge en loi vers une loi normale centrée réduite, i.e. pour tous $(a, b) \in (\mathbf{R} \cup \{\pm\infty\})^2$, $a \neq b$,

$$\mathbf{P}\left(a \leq \sqrt{n} \frac{\bar{X}_n - \mu}{\sigma_n} \leq b\right) = \mathbf{P}\left(a \leq \frac{S_n - n\mu}{\sqrt{n}\sigma_n} \leq b\right) \xrightarrow{n \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{t^2}{2}} dt = \Phi(b) - \Phi(a).$$

Remarque 2.4 — Cette substitution par σ_n est en fait une conséquence d'une propriété plus générale de la convergence en loi, appelée « lemme de SLUTSKY ».

3. APPROXIMATIONS DE LOIS

Nous allons à présent appliquer les théorèmes limites vus précédemment pour approcher des lois usuelles par d'autres lois.

3.1. Conséquences du théorème central limite

Commençons avec deux applications directes du théorème central limite.

Proposition ALEA.16.5 | Approximation $\mathcal{B}(n, p) \approx \mathcal{N}(np, np(1-p))$ (MOIVRE-LAPLACE)

Si X_1, \dots, X_n sont i.i.d. de loi $\mathcal{B}(p)$, alors :

- $S_n = \sum_{k=1}^n X_k \hookrightarrow \mathcal{B}(n, p)$, et
- $S_n^* = \frac{S_n - np}{\sqrt{np(1-p)}} \xrightarrow[n \rightarrow \infty]{\text{loi}} \mathcal{N}(0, 1)$. Autrement dit, pour n assez grand, la loi de S_n est « proche de » $\mathcal{N}(np, np(1-p))$.

⊗ Attention

Cela n'a pas de sens de dire directement que S_n^* converge en loi vers une $\mathcal{N}(np, np(1-p))$, tout simplement car les paramètres de la loi normale dépendent de n .

🔧 Méthode Pour retenir l'approximation de la loi binomiale par la normale

Penser aux paramètres : $E(S_n) = np$, $\text{Var}(S_n) = np(1-p)$, « à la limite » on obtient une loi $\mathcal{N}(np, np(1-p))$ qui a même espérance/variance que S_n .

Remarque 3.1 — Validité On considère en pratique cette approximation comme satisfaisante dès que

$$n \geq 30, \quad np \geq 15, \quad np(1-p) \geq 5.$$

Preuve



Exemple 12 — Cas de $\mathcal{B}(40, 0.5)$ et « correction de continuité » Soit $X \hookrightarrow \mathcal{B}(40, 0.5)$.

- Par quelle loi normale peut-on approcher la loi de X ? Nous sommes dans les conditions d'approximation de MOIVRE-LAPLACE : X est une somme de $n = 40 \geq 30$ variables de BERNOULLI et $\frac{n}{2} = 20 \geq 5$. On peut donc approcher la loi de X par une $\mathcal{N}(20, 10)$.
- En déduire une valeur approchée de $\mathbf{P}(17 \leq X < 25)$. On obtient ainsi $\mathbf{P}(17 \leq X < 25) = \mathbf{P}(17 \leq X \leq 24) = \mathbf{P}\left(\frac{-3}{\sqrt{10}} \leq \frac{X-20}{\sqrt{10}} \leq \frac{4}{\sqrt{10}}\right) \approx {}^4 \Phi(1, 26) - \Phi(-0, 95) \approx 0.72$.
- Comment utiliser cette approximation pour estimer $\mathbf{P}(X = 20)$? Si on approche $\mathbf{P}(X = 20)$ par la même probabilité mais pour une loi normale, on trouve zéro (on rappelle que les lois à densité n'ont pas d'atome). On modifie donc un petit peu la probabilité de départ : notons

$$\begin{aligned} \mathbf{P}(X = 20) &\approx \mathbf{P}(19,5 \leq X \leq 20,5) \approx \mathbf{P}(-0,5/\sqrt{10} \leq X - 20 \leq 0,5/\sqrt{10}) \\ &= \Phi\left(0,5/\sqrt{10}\right) - \Phi\left(-0,5/\sqrt{10}\right) \approx 0.126. \end{aligned}$$

Cette opération que l'on vient de réaliser est généralement appelée « correction de continuité ».

⁴Vous constaterez qu'ici on s'est directement ramené au théorème central limite en centrant et réduisant. Il n'est donc pas forcément utilisé d'apprendre par coeur la proposition précédente

Faisons de-même mais cette fois-ci pour la loi de POISSON.

Proposition ALEA.16.6 | Approximation $\mathcal{P}(\lambda) \approx \mathcal{N}(n\lambda, n\lambda)$

Si X_1, \dots, X_n i.i.d. de loi $\mathcal{P}(\lambda)$ avec $\lambda > 0$, alors :

1. $S_n = \sum_{k=1}^n X_k \hookrightarrow \mathcal{P}(n\lambda)$, et
2. $S_n^* = \frac{S_n - n\lambda}{\sqrt{n\lambda}} \xrightarrow[n \rightarrow \infty]{\text{loi}} \mathcal{N}(0, 1)$. Autrement dit, pour n assez grand, la loi de S_n est «proche de» $\mathcal{N}(n\lambda, n\lambda)$.

Remarque 3.2 – Validité On considère en pratique cette approximation comme satisfaisante dès que

$$n\lambda \geq 15.$$

Remarque 3.3 – Cas $n = 1$ Le résultat précédent permet donc d'affirmer que $\mathcal{P}(\lambda)$ est proche de $\mathcal{N}(\lambda, \lambda)$ dès lors que

$$\lambda \geq 15.$$

Preuve



Attention

Cela n'a pas de sens de dire directement que S_n^* converge en loi vers une $\mathcal{N}(n\lambda, n\lambda)$, tout simplement car les paramètres de la loi normale dépendent de n .

Méthode Pour retenir l'approximation de la loi de POISSON par la normale

Penser aux paramètres : $E(S_n) = n\lambda$, $\text{Var}(S_n) = n\lambda$, «à la limite» on obtient une loi $\mathcal{N}(n\lambda, n\lambda)$ de même espérance/variance que S_n .

3.2. Conséquences de calculs directs

Pour terminer ce chapitre, nous établissons deux autres résultats de convergence totalement indépendants des théorèmes limites.

Proposition ALEA.16.7 | Approximation $\mathcal{H}(N, n, p) \approx \mathcal{B}(n, p)$

Soient $p \in]0, 1[$ et $n \in \mathbf{N}$. On pose $q = 1 - p$. Alors pour $N \geq n$, tel que $Np \in \mathbf{N}$, soit $X_N \hookrightarrow \mathcal{H}(N, n, p)$. Alors pour tout $k \in \llbracket 0, n \rrbracket$,

$$\lim_{\substack{N \rightarrow +\infty \\ N \in \mathbf{N}, Np \in \mathbf{N}}} \mathbf{P}(X_N = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

En particulier, la loi $\mathcal{H}(N, n, p)$ est «proche de» la loi $\mathcal{B}(n, p)$.


Remarque 3.4 – Validité On considère en pratique cette approximation comme satisfaisante dès que

$$N \geq 10n.$$

Remarque 3.5 – Comment expliquer ce résultat sans calculs? Si le nombre

Preuve Soit $i \in \mathbf{N}$. Rappelons que $\mathbf{P}(X_i = k) = \frac{\binom{Np}{k} \binom{Nq}{n-k}}{\binom{N}{n}}$ pour tout $k \in X_i(\Omega)$. Alors :


$$\begin{aligned} & \frac{\binom{Np}{k} \binom{Nq}{n-k}}{\binom{N}{n}} \\ &= \frac{\frac{(Np)!}{k!(Np-k)!} \frac{(Nq)!}{(n-k)!(Nq-n+k)!}}{\frac{N!}{n!(N-n)!}} = \frac{n!}{k!(n-k)!} \frac{(Np)!}{(Np-k)!} \frac{(Nq)!}{(Nq-n+k)!} \frac{(N-n)!}{N!} \\ &= \binom{n}{k} \frac{(Np) \dots (Np-k+1)}{1} \frac{(Nq) \dots (Nq-n+k+1)}{1} \frac{1}{(N) \dots (N-n+1)}. \end{aligned}$$


La deuxième fraction apparaît comme un quotient de deux polynômes en N : le numérateur est un polynôme de degré n de coefficient dominant $p^k q^{n-k} N^n$, le second de degré n de coefficient dominant N^n . 

n est très grand, le résultat précédent nous apprend que

$$\mathcal{B}\left(n, \frac{\lambda}{n}\right) \text{ est proche de } \mathcal{P}(\lambda).$$

Mais le paramètre $\frac{\lambda}{n}$ est alors très petit, donc $\mathcal{P}(\lambda)$ correspondrait alors approximativement à un nombre de succès survenus dans un intervalle très grand $\llbracket 0, n \rrbracket$, mais avec une probabilité de succès $\frac{\lambda}{n}$ extrêmement faible. C'est pour cette raison que l'on qualifie parfois la loi de POISSON de «loi des évènements rares».

Preuve Montrons tout d'abord que : $\lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda}$. 

On peut ensuite conclure : 

Proposition ALEA.16.8 | Approximation $\mathcal{P}(\lambda) \approx \mathcal{B}\left(n, \frac{\lambda}{n}\right)$

Soit $\lambda \in \mathbf{R}^{+*}$, et $Y_n \hookrightarrow \mathcal{B}\left(n, \frac{\lambda}{n}\right)$. Alors :

$$\forall k \in \mathbf{N}, \quad \lim_{n \rightarrow \infty} \mathbf{P}(Y_n = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Remarque 3.6 — Validité On considère en pratique cette approximation comme satisfaisante dès que

$$n \geq 30, \quad \frac{\lambda}{n} \leq 0,01.$$

Remarque 3.7 — Interprétation de la loi de POISSON : évènements rares Lorsque

Approximation de la loi de POISSON

```
import scipy.special

def binom(n, k):
    """
```



```

(n : int, k : int) -> renvoie le coefficient binomial k parmi
↳ n
"""
if k > n:
    return 0
else:
    return scipy.special.binom(n, k)

lamba = 10
def rep_Hist_Bin(n):
    """
    trace l'histo de B(n,lamba/n) à n fixé
    """
    p = lamba/n
    Support = np.arange(0,n+1,1)
    Loi = np.zeros(len(Support))
    for k in range(len(Support)):
        Loi[k] = binom(n, k)*(p**k)*((1-p)**(n-k))
    plt.bar(Support, Loi , color='r')

def rep_Hist_Poiss():
    """
    trace l'histo de P(lamba) tronquée à 50
    """
    N = 50 # Troncature du support
    Support = np.arange(1,N+1,1)
    Loi = np.zeros(len(Support))
    for k in range(len(Support)):
        Loi[k] = lamba**k/ma.factorial(k)*np.exp(-lamba)
    plt.bar(Support, Loi)

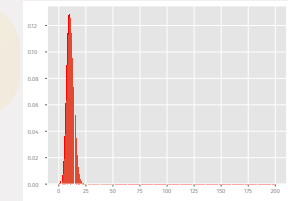
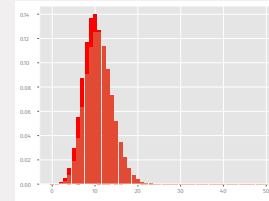
```

```

n = 50
rep_Hist_Bin(n)
rep_Hist_Poiss()


n = 200
rep_Hist_Bin(n)
rep_Hist_Poiss()

```




Pour $n = 200$, nous voyons que les distributions sont déjà extrêmement proches.


Exemple 13 — Sur une autoroute la proportion de camions par rapport à l'ensemble des véhicules est 0.07.

- Soit X le nombre de camions parmi 100 véhicules choisis au hasard. Calculer $\mathbf{P}(X \geq 5)$.  Il s'agit ici de calculer la probabilité $\mathbf{P}(X \geq 5)$ avec $X \hookrightarrow \mathcal{B}(100, 0, 07)$. Nous sommes dans les conditions d'application de la proposition précédente : $n = 100 \geq 30$, $p = 0,07 \leq 0,1$ et $np = 7 \leq 15$. Ainsi on approche cette loi par $\mathcal{P}(100 \times 0,07) = \mathcal{P}(7)$. Donc :

$$\begin{aligned}
 \mathbf{P}(X \geq 5) &\approx \sum_{k=5}^{\infty} e^{-7} \frac{7^k}{k!} \\
 &= 1 - e^{-7} \sum_{k=0}^4 \frac{7^k}{k!} \approx \boxed{0,827}.
 \end{aligned}$$

- Soit Y le nombre de camions parmi 1000 véhicules. Calculer $\mathbf{P}(65 \leq Y \leq 75)$.  Cette fois-ci on souhaite approcher une $\mathcal{B}(1000, 0, 07)$. Comme $np \geq 15$, la bonne approximation à utiliser est plutôt celle de MOIVRE-LAPLACE (vérifier pour cela l'ensemble des conditions : $1000 \times 0,07 = 70 \geq 15$, $70 \times 0,93 = 64$, et $1 > 4$). On approche donc par la loi $\mathcal{N}(70, 65,1)$. Ainsi :

$$\begin{aligned}
 \mathbf{P}(65 \leq Y \leq 75) &= \mathbf{P}\left(\frac{-5}{\sqrt{65,1}} \leq \frac{Y - 70}{\sqrt{65,1}} \leq \frac{5}{\sqrt{65,1}}\right) \\
 &= \Phi\left(\frac{5}{\sqrt{65,1}}\right) - \Phi\left(\frac{-5}{\sqrt{65,1}}\right) \approx \boxed{0,5}.
 \end{aligned}$$

3. On choisit n véhicules au hasard. Déterminer pour quelles valeurs de n on peut affirmer que la proportion des camions parmi ces véhicules est comprise entre 0.06 et 0.08 avec un risque d'erreur inférieur à 0.05.  Cette fois-ci le nombre de camions est S_n avec $S_n \hookrightarrow \mathcal{B}(n, 0,07)$, la proportion des camions est alors $\frac{S_n}{n}$. On cherche donc n tel que

$$P\left(\left|\frac{S_n}{n} - 0,07\right| \geq 0,01\right) = 0,05.$$

Sous les conditions d'application de l'approximation de MOIVRE-LAPLACE, i.e. $n \geq 30$, $0,07n \geq 15$ et $0,07 \times 0,93 \times n > 5$, i.e. $n \geq 215$, nous pouvons approcher la loi de $\frac{S_n}{n}$ par $\mathcal{N}\left(0,07, \frac{0,651}{n}\right)$. La loi de $\frac{S_n}{n} - 0,07$ est, d'après les propriétés de stabilité de la loi normale une $\mathcal{N}\left(0, \frac{0,651}{n}\right)$. Ainsi, puisque $\text{Var}\left(\frac{S_n}{n}\right) = \frac{0,651}{n}$:

$$\begin{aligned} P\left(\left|\frac{S_n}{n} - 0,07\right| \geq 0,01\right) &= P\left(\left|\sqrt{\frac{n}{0,651}} \left(\frac{S_n}{n} - 0,07\right)\right| \geq \sqrt{\frac{n}{0,651}} 0,01\right) \\ &\approx 2\left(1 - \Phi\left(\frac{\sqrt{n}}{\sqrt{651}}\right)\right) \\ &\approx 0,05 \end{aligned}$$

si et seulement si $\Phi\left(\frac{\sqrt{n}}{\sqrt{651}}\right) \approx 0,975$. On trouve avec la table : $\frac{\sqrt{n}}{\sqrt{651}} \approx 1,96$, soit $n \approx 2501 \geq 90$, qui est confirmé donc l'approximation de MOIVRE-LAPLACE réalisée plus haut. Rappelons également que pour trouver n on aurait pu utiliser la commande `norm.ppf(0.975)` (fonction inverse de Φ) qui renvoie 1.959963984540054).

 Rassurez-vous, les critères de validité sont généralement rappelés s'il y en a besoin dans un sujet.

Résumé Bilan des approximations de lois

Loi	Proche de	Validité
Loi $\mathcal{B}(n, p)$	$\mathcal{N}(np, np(1-p))$	$n \geq 30, np \geq 15, np(1-p) \geq 5,$
Loi $\mathcal{P}(\lambda)$	$\mathcal{N}(n\lambda, \sqrt{n\lambda})$	$n\lambda \geq 15,$
Loi $\mathcal{H}(N, n, p)$	$\mathcal{B}(n, p)$	$N \geq 10n,$
Loi $\mathcal{B}(n, \lambda/n)$	$\mathcal{P}(\lambda)$	$n \geq 30, \frac{\lambda}{n} \leq 0,01.$

ANNEXE – TABLE DE VALEURS DE LA $\mathcal{N}(0, 1)$

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.090
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545

**Méthode Lecture d'une table**

Si l'on souhaite avoir, par exemple, $\Phi(0,96)$, on :

1. se place sur la ligne «0.9»,
2. se place ensuite sur la colonne «0.06».
3. On obtient alors la valeur désirée.

1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

$$\Phi(x) = \mathbf{P}(X \leq x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \quad \text{et} \quad \Phi(-x) = 1 - \Phi(x).$$

On retiendra en particulier la valeur typique $\Phi(1.96) = 0.975$, de sorte que :

$$\mathbf{P}(X \in [-1.96; 1.96]) = 2\Phi(1.96) - 1 = 0.95$$

*** **Fin du chapitre** ***

4. EXERCICES

4.1. Inégalités de concentration

Exercice ALEA.16.1 | Soient X une variable aléatoire et $g : \mathbf{R}^{+*} \rightarrow \mathbf{R}^{+*}$ une fonction croissante, telle $g(|X|)$ possède une espérance. Montrer que :

$$\forall a > 0, \quad \mathbf{P}(|X| \geq a) \leq \frac{\mathbf{E}(g(|X|))}{g(a)}.$$

Exercice ALEA.16.2 | On note Φ la fonction de répartition de la loi $\mathcal{N}(0, 1)$.


1. Montrer que : $\forall x \in \mathbf{R}^{+*}, \quad 1 - \Phi(x) \leq \frac{1}{2x^2}$.
2. En déduire l'existence de $\int_0^\infty (1 - \Phi(t)) dt$. Calculer sa valeur.

Exercice ALEA.16.3 | Soit $X \hookrightarrow \mathcal{N}(0, 1)$ et $x > 0$. Montrer que :

$$\int_0^x e^{-\frac{t^2}{2}} dt \geq \sqrt{\frac{\pi}{2}} - \frac{1}{x}.$$

Indication : On pourra essayer d'appliquer l'inégalité de MARKOV.

Exercice ALEA.16.4 | On réalise 400 fois de façon indépendante la même expérience dont la probabilité de succès est 0,8. On note N_{400} la variable aléatoire discrète donnant le nombre de succès dans cette série de 400 expériences.

1. Montrer que : $\mathbf{P}(300 < N < 340) \geq 0,84$.
2.  En simulant un grand nombre de fois la variable aléatoire N , constater le résultat précédent par simulation.

Exercice ALEA.16.5 | Une urne contient 2 boules blanches et 4 boules noires. On y effectue n tirages avec remise et on note F_n la fréquence des boules blanches, obtenues

sur n tirages. Déterminer n tel que

$$\mathbf{P}\left(\left|F_n - \frac{1}{3}\right| \geq 0,02\right) \leq 0,01,$$

en utilisant :

1. en utilisant l'inégalité de BIENAYMÉ-TCHEBYCHEV,
2. en utilisant le théorème central limite.

Solution (exercice ALEA.16.5)

On introduit dans tout l'exercice une suite $(X_n)_n$ i.i.d. de même loi $\mathcal{B}(2/6) = \mathcal{B}(1/3)$. Alors $F_n = \overline{X_n}$.

1. D'après l'inégalité de BIENAYMÉ-TCHEBYCHEV, puisque les X_n admet un moment d'ordre deux, nous avons pour tout $\varepsilon > 0$:

$$\mathbf{P}\left(\left|F_n - \mathbf{E}(F_n)\right| \geq \varepsilon\right) = \mathbf{P}\left(\left|F_n - \frac{1}{3}\right| \geq \varepsilon\right) \leq \frac{\frac{1}{3} \times \frac{2}{3}}{n\varepsilon^2}.$$

Donc

$$\mathbf{P}\left(\left|F_n - \mathbf{E}(F_n)\right| \geq 0,02\right) \leq \frac{2}{9n(0,02)^2} = \frac{2 \cdot 10^4}{36n} = \frac{10^4}{18n}.$$

On cherche donc n tel que $\frac{10^4}{18n} \leq 0,01 \iff n \geq 55555$.

2. ...

4.2. Théorèmes limites & Approximations

Exercice ALEA.16.6 | Soit (X_n) une suite de variables aléatoires indépendantes de loi $\mathcal{B}(p_n)$ avec $p_n \in]0, 1[$ pour tout entier $n \in \mathbf{N}$. On pose $S_n = X_1 + \dots + X_n$ pour tout

$n \geq 1$. On souhaite montrer que pour tout $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\left| \frac{S_n}{n} - \frac{1}{n} \sum_{k=1}^n p_k \right| \geq \varepsilon \right) = 0.$$

1. Peut-on appliquer la loi faible des grands nombres?
2. Calculer l'espérance et la variance de S_n , puis démontrer que $\mathbf{Var}(S_n) \leq n$.
3. En déduire le résultat.

Exercice ALEA.16.7 | Soit $(X_n)_{n \in \mathbf{N}^*}$ une suite de variables aléatoires réelles i.i.d. de même loi $\mathcal{B}(p)$ avec $p \in]0, 1[$. On pose pour tout $n \in \mathbf{N}^*$, $Y_n = X_n + X_{n+1}$ et $\overline{Y}_n = \frac{1}{n} \sum_{k=1}^n Y_k$.

Montrer que la suite $(\overline{Y}_n)_{n \in \mathbf{N}^*}$ converge en probabilité vers la variable aléatoire constante $2p$.

Exercice ALEA.16.8 | Déterminer :

$$\lim_{n \rightarrow \infty} \frac{1}{3^n} \sum_{k=0}^{\lfloor \frac{n}{3} \rfloor} \binom{n}{k} 2^{n-k}.$$

Indication : Appliquer le théorème central limite à une suite de variables aléatoires de BERNOULLI bien choisie.

Exercice ALEA.16.9 | Un commando de soldats doit traverser une rivière infestée de crocodiles. La probabilité qu'un soldat soit dévoré au cours de la traversée est $\frac{1}{10}$. Combien faut-il envoyer de soldats pour que la probabilité d'avoir au moins 300 rescapés soit supérieure à 95 %?

Exercice ALEA.16.10 | Une entreprise compte 300 employés. Chacun d'eux téléphone en moyenne 6 minutes par heure. Combien de lignes téléphoniques doit faire installer l'entreprise afin que la probabilité que toutes les lignes soient occupées en même temps sur un intervalle de temps d'une heure soit inférieure ou égale à 0.025.

Solution (exercice ALEA.16.10)

Dans cet exercice, il faut comprendre le fonctionnement suivant : chaque employé

n'est pas associé à une ligne précise, en revanche, lorsqu'un appel arrive sur l'une des lignes, on souhaite souvent avoir un employé libre pour décrocher. Notons X le nombre d'employés occupés sur un intervalle de temps d'une heure. Puisque les occupations peuvent être supposées indépendantes, et que la probabilité qu'un employé fixé soit occupé est $\frac{6}{60} = \frac{1}{10}$ et qu'il y a 300 employés, on a que $X \hookrightarrow \mathcal{B} \left(300, \frac{1}{10} \right)$

(comptage de succès dans une répétition de 300 expériences de BERNOULLI indépendantes, un succès étant une occupation d'un employé). On souhaite ainsi résoudre $\mathbf{P}(X \geq N) \leq 0,025$, N étant un entier à trouver correspondant au nombre de lignes à installer, qui correspond à la probabilité qu'un appel tombe dans le vide *i.e.* tous les employés sont occupés. On utilise le théorème central limite (ou l'approximation de la binomiale par la normale), puisque X est une somme de variables aléatoires i.i.d. admettant une variance.

$$\begin{aligned} \mathbf{P}(X \geq N) &\leq 0,025 \\ \Leftrightarrow \mathbf{P} \left(X^* \geq \frac{N-30}{\sqrt{27}} \right) &\leq 0,025, \\ \Leftrightarrow 1 - \Phi \left(\frac{N-30}{\sqrt{27}} \right) &\leq 0,025, \\ \Leftrightarrow \Phi \left(\frac{N-30}{\sqrt{27}} \right) &\geq 0,975 = \Phi(1,96), \\ \Leftrightarrow \frac{N-30}{\sqrt{27}} &\geq 1,96 \\ \Leftrightarrow N &\geq 41. \end{aligned}$$

Il faut donc 41 lignes au moins.

Exercice ALEA.16.11 | On lance un dé équilibré à 6 faces 9000 fois. On note A « on obtient la face 6 entre 1400 et 1600 fois ».

1. Donner l'expression exacte de $\mathbf{P}(A)$. On n'essayera pas de calculer une certaine somme.

- Donner une minoration de $\mathbf{P}(A)$ en utilisant l'inégalité de BIENAYMÉ-TCHEBYCHEV.
- En utilisant le théorème central limite, donner une estimation de $\mathbf{P}(A)$. Comparer avec la question précédente.
- Pour utiliser le théorème central limite, on centre et on réduit (on peut utiliser aussi l'approximation de la binomiale par une normale, cela revient au même). Soit $N \hookrightarrow \mathcal{N}(0, 1)$, et on note Φ sa fonction de répartition.

Solution (exercice ALEA.16.11)

- Le nombre de 6 apparus sur les 9000 lancers suit une $\mathcal{B}\left(9000, \frac{1}{6}\right)$, par indépendance des lancers et puisque l'on compte le nombre de succès (faire un 6, de probabilité $1/6$) dans cette répétition de 9000 lancers. Notons X une variable aléatoire de telle loi. Alors

$$\begin{aligned} \mathbf{P}(A) &= \mathbf{P}(X \in [1400, 1600]) \\ &= \sum_{k=1400}^{1600} \mathbf{P}(X = k) \\ &= \sum_{k=1400}^{1600} \binom{9000}{k} \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{9000-k} \\ &= \sum_{k=1400}^{1600} \binom{9000}{k} \frac{5^{9000-k}}{6^{9000}}. \end{aligned}$$

Cette somme est bien entendu non calculable.

- Rappelons que $\mathbf{P}(A) = \mathbf{P}(X \in [1400, 1600])$. Alors pour utiliser BIENAYMÉ-TCHEBYCHEV, on centre la variable aléatoire. On a $\mathbf{E}(X) = \frac{9000}{6} = 1500$. Donc

$$\begin{aligned} \mathbf{P}(X \in [1400, 1600]) &= \mathbf{P}(X - 1500 \in [-100, 100]) \\ &= \mathbf{P}(|X - 1500| \leq 100). \end{aligned}$$

En passant au complémentaire, on déduit

$$\begin{aligned} \mathbf{P}(X \in [1400, 1600]) &= 1 - \mathbf{P}(|X - 1500| > 100) \\ &= 1 - \mathbf{P}(|X - 1500| \geq 101) \\ &\geq 1 - \frac{\mathbf{Var}(X)}{101^2} \\ &= 1 - \frac{1250}{101^2} \approx \boxed{0.877}. \end{aligned}$$

$\left. \begin{array}{l} |X - 1500| \text{ entier} \\ \text{BIENAYMÉ-TCHEBYCHEV car } \mathbf{E}(X) = 1500 \end{array} \right\}$

$$\begin{aligned} \mathbf{P}(X \in [1400, 1600]) &= \mathbf{P}(X - 1500 \in [-100, 100]) \\ &= \mathbf{P}(|X - 1500| \leq 100) \\ &= \mathbf{P}\left(\left|\frac{X - 1500}{\sqrt{\mathbf{Var}(X)}}\right| \leq \frac{100}{\sqrt{\mathbf{Var}(X)}}\right) \\ &\approx \mathbf{P}\left(|N| \leq \frac{100}{\sqrt{1250}}\right) \\ &= \Phi\left(\frac{100}{\sqrt{1250}}\right) - \Phi\left(-\frac{100}{\sqrt{1250}}\right) \\ &= 2\Phi\left(\frac{100}{\sqrt{1250}}\right) - 1 \\ &\approx 2\Phi(2.83) - 1. \end{aligned}$$

En utilisant une table de la loi normale on trouve cette fois $\mathbf{P}(A) \approx 0.995$. La minoration observée à la question précédente est donc correcte, mais pas très précise.

Exercice ALEA.16.12 | Surbooker ou pas? Une compagnie aérienne souhaite optimiser l'occupation en voyageurs de sa flotte d'avions. Pour cela, elle souhaite tenir compte des voyageurs ne se présentant pas à l'embarquement, et proposer ces places à d'autres voyageurs, en prenant en amont plus de réservations que de places disponibles.

On suppose que chaque client se présente à l'embarquement avec une probabilité $p \in]0, 1[$. On souhaite trouver le nombre de réservations à réaliser $n \geq N$, $N \in \mathbf{N}^*$ étant le nombre de places dans l'avion, afin que la probabilité de surbooker effectivement l'avion soit inférieure ou égale à 4%. On note X le nombre de clients se présentant

à l'embarquement sur ces n réservations. On donne la valeur $0,96 = \Phi(1,75)$ où Φ désigne la fonction de répartition de la loi $\mathcal{N}(0, 1)$.

- Déterminer la loi de X .
- On suppose dans la suite que les paramètres nous permettent d'approcher X par une loi normale, préciser de quelle loi normale il s'agit.
- Justifier que n satisfaisant la condition de l'énoncé satisfait une inéquation de la forme :

$$an + b\sqrt{n} + c \leq 0,$$

où a, b, c sont des paramètres réels à préciser.

- Démontrer que cette inéquation admet des solutions et exprimer la meilleure solution du problème en fonction du a, b, c .

Solution (exercice ALEA.16.12)

- On reconnaît un schéma binomial, puisque chaque passager embarque avec la même probabilité p , et indépendamment les uns des autres. On a donc $X \hookrightarrow B(n, p)$ et par conséquent

$$E(X) = np, \quad \text{Var}(X) = np(1-p).$$

- On suppose que les conditions d'approximation du cours sont satisfaites. On peut alors approcher $\mathcal{B}(n, p)$ par $\mathcal{N}(np, np(1-p))$, donc une loi normale de mêmes paramètres.
- On souhaite $P(X > N) \leq 0,04$. Donc en notant $X^* = \frac{X-np}{\sqrt{np(1-p)}}$ et en passant au complémentaire :

$$P(X \leq N) = P\left(X^* \leq \frac{N-np}{\sqrt{np(1-p)}}\right) \geq 0,96$$

donc n vérifie

$$\frac{N-np}{\sqrt{np(1-p)}} \geq \Phi^{-1}(0,96) = 1,75$$

et en réarrangeant :

$$np + 1,75\sqrt{p(1-p)} \cdot \sqrt{n} - N \leq 0.$$

Cette inéquation est bien de la forme :

$$an + b\sqrt{n} + c \leq 0$$

avec $a = p, b = 1,75\sqrt{p(1-p)}$ et $c = -N$.

- Cette équation admet des solutions puisque par exemple $n = 0$ est évidemment solution. On reconnaît une équation du second degré en la variable \sqrt{n} . Le discriminant est $\Delta = 1,75^2 p(1-p) + 4Np > 0$ et

$$n \text{ est solution} \iff \sqrt{n} \in \left[\frac{-b - \sqrt{\Delta}}{2a}, \frac{-b + \sqrt{\Delta}}{2a} \right]$$

et la meilleure solution n_{\max} est donc le plus grand entier tel que $\sqrt{n} \leq \frac{-b + \sqrt{\Delta}}{2a}$, c'est-à-dire :

$$n_{\max} = \left\lfloor \left(\frac{-b + \sqrt{\Delta}}{2a} \right)^2 \right\rfloor.$$

Exercice ALEA.16.13 | D'après Agro—Véto 2016 On considère une population de $n \in \mathbf{N}^*$ insectes. Le nombre (aléatoire) N d'oeufs pondus par un insecte donné suit une loi de POISSON de paramètre $\lambda > 0$. Chaque oeuf peut éclore avec une probabilité $p \in]0, 1[$. On note X la variable aléatoire égale au nombre d'oeufs éclos pour un insecte donné et Y la variable aléatoire égale au nombre d'oeufs non éclos.

- On utilisera les fonctions ci-dessous en prenant comme valeurs des paramètres : $n = 10, p = 0,25$ et $\lambda = 40$.

- En approchant la loi de N par une loi binomiale $\mathcal{B}\left(1000, \frac{\lambda}{1000}\right)$, écrire fonction python d'en-tête def $f1(\lambda)$ qui prend en entrée une valeur de λ et qui simule la variable aléatoire N .

- 1.2) Écrire une fonction python d'en tête `def f2(1, p)` qui prend en entrée une valeur pour λ et une valeur pour p et donne en sortie un couple de valeurs pour (X, Y) .
- 1.3) Écrire une fonction python d'en-tête `def f3(n, l, p)` : qui prend en entrée une valeur pour le nombre d'insectes n , une valeur pour λ , une valeur pour p et donne en sortie le nombre total d'oeufs éclos.
2. 2.1) Donner les lois de X et Y .
- 2.2) Les variables aléatoires X, Y sont-elles indépendantes? Calculer $\mathbf{Cov}(X, Y)$.
- 2.3) On note N_{10} le nombre d'oeufs pondus par 10 insectes, et Z_{10} le nombre d'oeufs éclos pour 10 insectes. Donner les lois de N_{10} et Z_{10} .

Exercice ALEA.16.14 | Caractérisation de la loi normale Soit X une variable aléatoire réelle centrée de carré intégrable, de variance $\sigma^2 > 0$. On suppose que si X_1, X_2 sont indépendantes de même loi que X alors $\frac{X_1+X_2}{\sqrt{2}}$ a même loi que X .

1. Montrer que : $\frac{X_1+\dots+X_4}{\sqrt{2^2}}$ a même loi que X .
2. **(Généralisation)** Soit $(X_n)_n$ une suite i.i.d. de même loi que X , on note pour tout $n \in \mathbf{N}^*$, $S_n = X_1 + \dots + X_n$. Montrer que : $\frac{S_{2^n}}{\sqrt{2^n}}$ a même loi que X pour tout $n \in \mathbf{N}^*$.
3. En déduire, en utilisant le théorème central limite, que $X \rightsquigarrow \mathcal{N}(0, \sigma^2)$.

Chapitre ALEA.17.

Statistiques Descriptives & Inférentielles

Résumé & Plan

Les statistiques sont présentes dans beaucoup de domaines en Sciences, notamment dans l'exploitation de grosses quantités de données. Il existe plusieurs types de statistiques : nous en étudierons deux, les statistiques descriptives d'une part dont le but est d'étudier des séries de données, d'autre part les statistiques inférentielles dont l'objectif est de savoir si des données semblent provenir ou non de réalisations d'une certaine variable aléatoire.

1	La Statistique : position du problème	2
1.1	Où apparaît l'aléatoire?	2
1.2	Statistiques descriptive	2
1.3	Statistique inférentielle	3
2	Statistiques descriptives	3
2.1	Univariées	4
2.2	Bivariées	15

3	Statistiques inférentielles	24
3.1	Estimation ponctuelle	24
3.2	Estimation par Intervalle de confiance	30
3.3	Test de conformité à la moyenne	36
4	Exercices	42
4.1	Descriptives	42
4.2	Estimateurs	43
4.3	Intervalles de confiance	46
4.4	Tests	50

Il existe trois types de mensonges : les mensonges simples, les sacrés mensonges et les statistiques.

— **Mark Twain**

**Cadre**

Dans tout ce chapitre, (Ω, \mathcal{F}, P) désignera un espace probabilisé sur lequel seront définis les objets aléatoires le cas échéant.

1.**LA STATISTIQUE : POSITION DU PROBLÈME****1.1. Où apparaît l'aléatoire ?**

L'aléatoire est présent dans toute expérience scientifique. Les deux grandes explications en sont :

- ▶ d'une part l'aléatoire « intrinsèque » lié à la complexité des individus et des phénomènes étudiés et au manque d'information dans le domaine.
- ▶ D'autre part, de l'aléatoire peut intervenir « expérimentalement », par mesures entachées d'erreur,¹ ou encore lorsque les moyens pour relever sont limités (*i.e.* par exemple un recensement dans une grande population).

La première source d'aléatoire est donc le manque d'informations, ou l'ignorance, la seconde est matérielle et peut donc être contrôlée et donnée elle-même lieu à une étude plus poussée. Il n'y a pas de raison de penser qu'une source d'aléatoire peut complètement être supprimée. En Biologie (ainsi qu'en Médecine), étant donné l'extrême complexité des systèmes étudiés, l'aléatoire est très présent.

Exemple 1 — On étudie l'influence d'un champignon sur une population d'hêtres dans deux parcelles de forêt. Plusieurs sources d'aléatoires vont être prise en compte.

- ▶ Plusieurs individus de la même espèce, de la même origine et du même âge ont

¹Rappelons que nous avons vu dans le [Chapter ANA.8](#) une borne sur l'erreur commise dans une fonction de mesures chacune entachées d'erreurs

- des biomasses différentes : aléatoire « intrinsèque » à l'expérience);
- ▶ les parcelles ne sont pas identiques (aléatoire « extrasèque »);
- ▶ les échantillons de population ne possèdent pas le même nombre d'individus, les calculs vont donc être source d'aléatoire : aléatoire « expérimental », lié au protocole;
- ▶ la grandeur mesurée ne peut l'être qu'avec une précision finie.

Cette analogie aléatoire pousse à interpréter les mesures effectuées comme des réalisations d'une variables aléatoire X « cachée ». Les Statistiques visent alors à étudier cet aléatoire, ou le prédire (Statistiques inférentielles) ou encore à le représenter (Statistiques descriptives)².

1.2.**Statistiques descriptive**

La statistique descriptive est la branche des statistiques qui regroupe les nombreuses techniques utilisées pour décrire un ensemble relativement important de données. Cette description peut être réalisée au moyen de calculs de grandeurs (moyenne, variance, écart-type, mode, *etc.*) ou au moyen de descriptions visuelles (graphiques par exemple). Nous nous poserons également une seconde question : on peut représenter deux séries de données sous forme d'un nuage de points, on peut alors se demander s'il est possible d'évaluer la corrélation entre les deux séries, et plus précisément la faculté de l'une à dépendre de l'autre.

²Mais il existe d'autres types de statistiques!

1.3. Statistique inférentielle

Inférence statistique, ensemble des méthodes permettant de formuler en termes probabilistes un jugement sur une population à partir des résultats observés sur un échantillon extrait au hasard de cette population.

— Larousse

Commençons par expliquer ce que nous souhaitons faire en statistiques inférentielles.

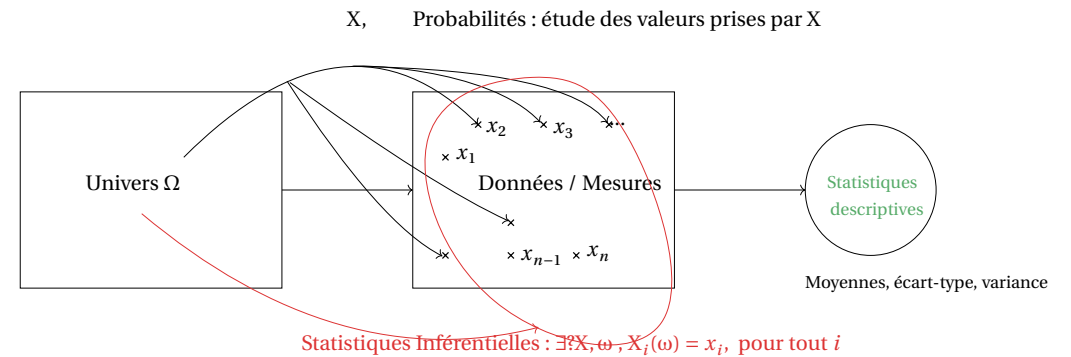
1. Les probabilités ont pour objectif notamment l'étude de réalisations de ce que nous appelons *variables aléatoires*. Nous savons avec quelles probabilités ces variables prennent une certaine valeur (donnée de la loi) et nous étudions ensuite des paramètres : l'espérance, la variance, la covariance, intervalles de fluctuation *etc.*.
2. La *statistique inférentielle* souhaite effectuer la démarche inverse : *i.e.* savoir que si une série de données x_1, \dots, x_n avec n un entier, peut être vue comme n réalisations d'une variable aléatoire suivant une certaine loi, autrement dit si le caractère mesuré semble suivre une certaine loi de probabilité. Bien évidemment, nous ne pourrons pas dire OUI ou NON avec certitude, nous pourrons juste dire OUI ou NON avec une bonne (si possible) probabilité, lorsque n est suffisamment grand.

L'objectif est donc de savoir s'il existe (X_1, \dots, X_n) une famille de n variables aléatoires réelles **de même loi** telles que pour un certain $\omega \in \Omega$, $(x_1, \dots, x_n) = (X_1(\omega), \dots, X_n(\omega))$ ³. Nous les chercherons également indépendantes, ce qui semble être une hypothèse raisonnable (il n'y a pas de raisons que les relevés x_1, \dots, x_n influent les uns sur les autres).

Pour des raisons techniques (application de la loi des grands nombres et du théorème

³Ce point de vue est équivalent à l'existence de $\omega_1, \dots, \omega_n \in \Omega$ tels que $(x_1, \dots, x_n) = (X(\omega_1), \dots, X(\omega_n))$

central limite), nous supposons que la loi commune des X_i possède une espérance μ et une variance σ^2 .⁴



2. STATISTIQUES DESCRIPTIVES

Dans tout ce qui suit, nous verrons que toutes les notions définies ont une analogie dans le monde des probabilités (échantillon, espérance, variance, écart-type, covariance, *etc.*). Cette analogie sera explicitement précisée à chaque fois pour que les définitions deviennent naturelles. Pour plus de précisions, vous pourrez consulter la **Remarque 2** ci-dessous.



Cadre

Dans cette section, les notations n, p, i, j désigneront des entiers même si cela n'est pas précisé.

⁴Ce qui, rappelons-le, n'est pas automatique en deuxième année pour les univers non finis

2.1. Univariées

2.1.1. Série statistique

Définition ALEA.17.1 | Population & Échantillon

Une *population* est un ensemble fini dont les éléments sont appelés des *individus*. Le nombre d'individus d'une population est appelé sa *taille*. Un sous-ensemble d'une population est appelé un *échantillon* de cette population.

Remarque 2.1 — Population ne signifie pas que l'on considère des personnes. Si vous réalisez plusieurs titrages pour votre T.I.P.E. on parlera alors de population de concentrations.

Très souvent, on ne mènera pas notre étude sur la population entière mais sur une sous-partie que l'on espère représentative.

Définition ALEA.17.2 | Caractère, Série statistique

1. Un *caractère* x de la population est une donnée *qualitative* ou *quantitative* attachée à chaque individu de la population. On notera x_i la valeur du caractère x pour un individu i .
2. Une *série statistique* de taille n , est une famille (x_1, \dots, x_n) à n éléments.
3. La donnée des valeurs d'un caractère pour les individus d'un échantillon de taille n est
 - ▶ un caractère est dit *quantitatif* s'il prend des valeurs quantifiables, souvent des réels mais éventuellement des p -uplets ou des matrices.
 - ▶ Un caractère est dit *qualitatif* s'il correspond à une propriété qui ne se quantifie pas (c'est-à-dire couleur des cheveux, opinion politique, *etc.*)

Définition ALEA.17.3 | Modalités

On appelle *modalités* d'un caractère les valeurs possibles qu'il peut prendre.

Exemple 2 — *Série des poids* Le tableau ci-dessous regroupe des données de poids d'individus numérotés entre 1 et n .

Individu	1	2	...	n
Masse	62	80	...	74

Exemple 3 — *Diamètres de pièces* Le tableau ci-dessous regroupe les diamètres en cm de 48 pièces prélevées dans la production d'une machine.

19	26	23	20	22	24	20	24
22	20	21	19	21	22	19	20
21	21	22	21	23	22	21	24
25	23	22	19	20	26	24	25
23	26	25	25	21	22	25	24
23	22	24	24	25	23	25	22

Il s'agit donc d'un échantillon statistique de taille 48 dans la population des pièces fabriquées par la machine. Le caractère étudié est le diamètre de la pièce en centimètre.

Les modalités sont : 1.19, 1.20, 1.21, 1.22, 1.23, 1.24, 1.25 et 1.26.

2.1.2. Effectifs & Regroupements en classes

Une série statistique peut être donnée sous plusieurs formes.

- ▶ Soit on décide de fournir toutes les données de manière exhaustive,
- ▶ soit on décide de les regrouper en modalités et/ou classes, et on doit alors fournir

un tableau d'effectifs associés.

DONNÉES AFFICHÉES DE MANIÈRE EXHAUSTIVE. On peut simplement donner la valeur du caractère pour chaque individu. Voir les **Exemples 2** et **3**.

DONNÉES REGROUPÉES PAR MODALITÉ. Si la taille de l'échantillon est trop grande on préférera donner les nombres d'individus associés à chaque modalité. On appelle cela l'*effectif* associé à ladite modalité.

Exemple 4 – Âges d'enfants Dans cet exemple, une série d'âges sur des enfants entre 0 et 5 ans. On fournit alors les effectifs associés à chaque âge.

Nombre d'enfants	0	1	...	5
Effectif	20	31	...	3

DONNÉES REGROUPÉES EN CLASSES : LORSQUE LE NOMBRE DE MODALITÉS EST TROP IMPORTANT. Parfois le nombre de modalités est trop grand, voire infini pour des modalités dites continues (c'est-à-dire à valeurs réelles). Il est alors pertinent de regrouper les modalités en classes disjointes, le plus souvent en des intervalles qui ne sont pas forcément de tailles égales. Inversement les modalités non regroupées en classe sont dites *ponctuelles*. Il peut également arriver que nos modalités comportent des classes et des modalités ponctuelles.

Exemple 5 – Par exemple, lorsque l'on étudie la démographie urbaine française on va regrouper les communes en classes selon leur nombre d'habitants :

- ▶ Hameaux de 1 à 99 habitants,
- ▶ Village de 100 à 1999 habitants,
- ▶ Ville de 2000 à 99999 habitants,

- ▶ Agglomération à partir de 100000 habitants.

EFFECTIFS & EFFECTIFS CUMULÉS CROISSANTS.

Définition ALEA.17.4

Soit une série statistique de taille n admettant un nombre fini de modalités ou à défaut de classes notées a_1, \dots, a_p .

1. Pour $j \in \llbracket 1, p \rrbracket$, on définit l'*effectif* n_j associé à la valeur a_j comme étant le nombre d'individus i pour lesquels $x_i = a_j$.
2. Pour $j \in \llbracket 1, p \rrbracket$, on définit la *fréquence de* f_j associée à la valeur a_j comme étant la proportion d'individus pour lesquels $x_i = a_j$, c'est-à-dire

$$f_j = \frac{n_j}{n}.$$

Proposition ALEA.17.1 | Formule des fréquences totales / effectifs totaux

Soit une série statistique de taille n admettant un nombre fini de modalités ou à défaut de classes notées a_1, \dots, a_p . Alors

$$\sum_{j=1}^p n_j = n \quad \text{et} \quad \sum_{j=1}^p f_j = 1.$$

Remarque 2.2 – Analogie avec les probabilités Reprenons les notations précédentes, *i.e.* une série statistique de taille n admettant un nombre fini de modalités ou à défaut de classes notées a_1, \dots, a_p . On note n_1, \dots, n_p les effectifs correspondant. Alors si l'on considère une variable aléatoire réelle discrète X telle que⁵ :

$$X(\Omega) = \{a_1, \dots, a_p\}, \quad \mathbf{P}(X = a_j) = \frac{n_j}{n}, \quad j \in \llbracket 1, p \rrbracket.$$

⁵on voit donc la série statistique comme la réalisation d'une variable réelle discrète, de loi la fréquence de ladite observation.

On vérifie sans peine que :

$$\sum_{j=1}^p \frac{n_j}{n} = 1, \quad \frac{n_j}{n} \geq 0$$

pour tout $j \in \llbracket 1, p \rrbracket$, ce qui garantit l'existence de X . Alors la formule précédente est simplement dire que

$$\{X = a_j\}_{j \in \llbracket 1, p \rrbracket}$$

est un système complet d'évènements. Toutes les notions que l'on va définir ci-après (espérance, variance, écart-type), correspondent en fait à :

$$\mathbf{E}(X), \mathbf{Var}(X), \sigma_X, \dots$$

Dans le cas où les caractères étudiés sont des réels (ils peuvent donc être ordonnés), on va introduire les effectifs cumulés croissants et les fréquences cumulées croissantes.

Définition ALEA.17.5 | Effectifs cumulés croissants

On suppose ici les modalités a_1, \dots, a_p rangées dans l'ordre croissant ($a_1 < a_2 < \dots < a_p$). Pour des intervalles $]a, b]$ et $]c, d]$ cela correspond à $a < b \leq c < d$. Soit $j \in \llbracket 1, p \rrbracket$.

1. On définit l'*effectif cumulé croissant* associé à la modalité a_j comme le nombre d'observations $x_i \leq a_j$, c'est-à-dire

$$n_j^c = \# \{i \in \llbracket 1, n \rrbracket, x_i \leq a_j\}.$$

2. On définit la *fréquence cumulée (croissante)* associée à la modalité a_j comme la proportion d'observations $x_i \leq a_j$, c'est-à-dire

$$f_j^c = \frac{n_j^c}{n}.$$

Remarque 2.3 — Analogie probabiliste : la fonction de répartition.

Proposition ALEA.17.2 | Lien effectifs / effectifs cumulés

On suppose ici les modalités a_1, \dots, a_p rangées dans l'ordre croissant ($a_1 < a_2 < \dots < a_p$). Pour des intervalles $]a, b]$ et $]c, d]$ cela correspond à $a < b \leq c < d$.

1. $\forall j \in \llbracket 1, p \rrbracket$, $n_j^c = \sum_{k=1}^j n_k$, et $f_j^c = \sum_{k=1}^j f_k$,
2. (Les effectifs cumulés sont croissants)

$$n_1^c \leq n_2^c \leq \dots \leq n_p^c = n,$$

$$f_1^c \leq f_2^c \leq \dots \leq f_p^c = 1.$$

Preuve

1. On écrit d'abord, avec $a_0 = -\infty$ pour convention, le découpage en tranches^a suivant :

$$\{i \in \llbracket 1, n \rrbracket, x_i \leq a_j\} = \bigsqcup_{k=1}^j \{i \in \llbracket 1, n \rrbracket, a_{k-1} < x_i \leq a_k\}.$$

En passant au cardinal dans cette réunion disjointe, il vient : $n_j^c = \sum_{k=1}^j n_k$. Il suffit de diviser alors par n pour obtenir la version avec fréquences.

2. Il suffit de constater que pour $i, j \in \llbracket 1, p \rrbracket$ tels que $1 \leq i < j \leq p$, nous avons l'inclusion $\{k \in \llbracket 1, n \rrbracket, x_k \leq a_i\} \subset \{k \in \llbracket 1, n \rrbracket, x_k \leq a_j\}$ puisque les classes sont ordonnées. Il suffit ensuite de passer au cardinal. On divise par n pour obtenir la version avec fréquences.

2.1.3. Représentation graphique

Les séries de données statistiques peuvent être représentés de plusieurs manières, en lieu et place de tableaux comme présentés *supra*.

DIAGRAMME EN BÂTONS & HISTOGRAMMES. Le diagramme en bâtons est adapté pour représenter des données ayant un nombre fini de modalités. Pour les données conti-

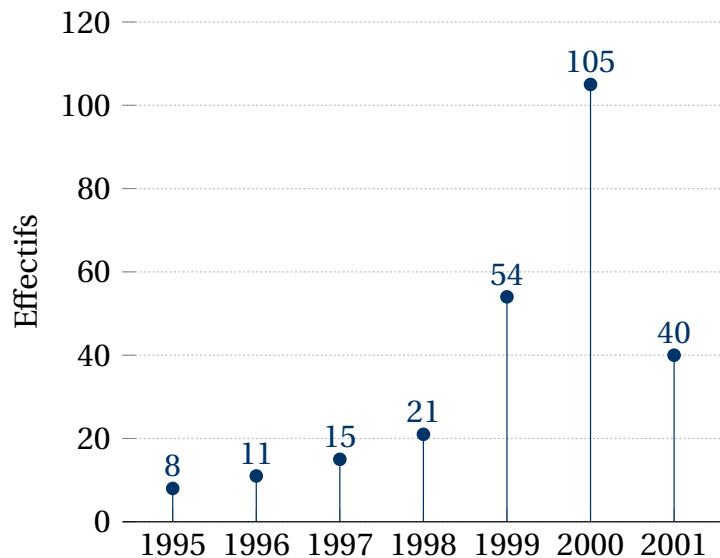
^aLes tranches deviennent des égalités si aucun regroupement en classes n'a été fait.

nues regroupées en classes, on tracera plutôt ce que l'on appelle un *histogramme*.

Pour établir un diagramme en bâtons on va tracer pour chaque modalité un bâton (un rectangle long et fin) centré en a_j et de hauteur f_j ou n_j . Ce type de graphique est adapté aux données ponctuelles et aux données qualitatives.

Exemple 6 – Diagramme en bâtons. Par exemple, ici nous avons représenté le tableau d'effectifs d'âges suivant :

Âge	1994	1995	1996	1997	1998	1999	2000
Effectif	8	11	15	21	54	105	40



Lorsque le nombre de modalités est trop grand, et même infini, nous avons que l'on regroupait généralement les données en classes.

Pour chaque classe on va alors tracer un rectangle dont la largeur vaut l'amplitude de l'intervalle et dont l'aire est proportionnelle à la fréquence ou à l'effectif de la classe.

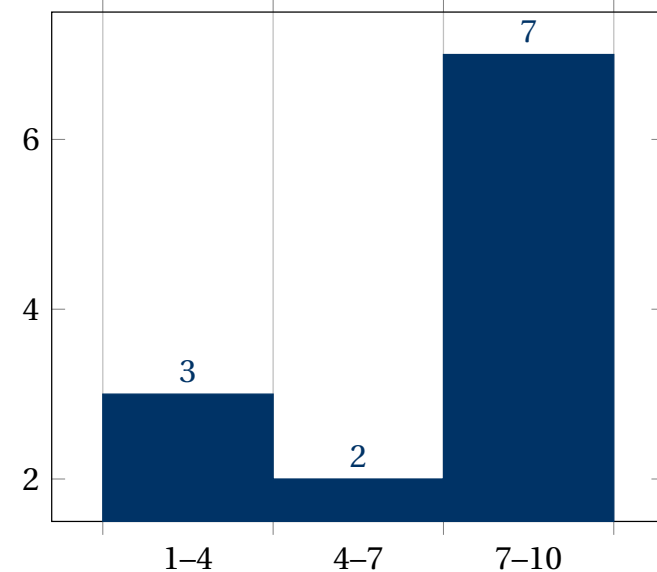
Attention

C'est l'aire du rectangle qui est importante, pas sa hauteur.

Exemple 7 – Histogramme. Par exemple, ici nous avons représenté le tableau d'effectifs d'âges suivant :

Données en classes	[1, 4[[4, 7[[7, 10[
Effectifs	3	2	7

Sur cet exemple les classes ont toutes la même longueur, mais ceci n'est aucunement

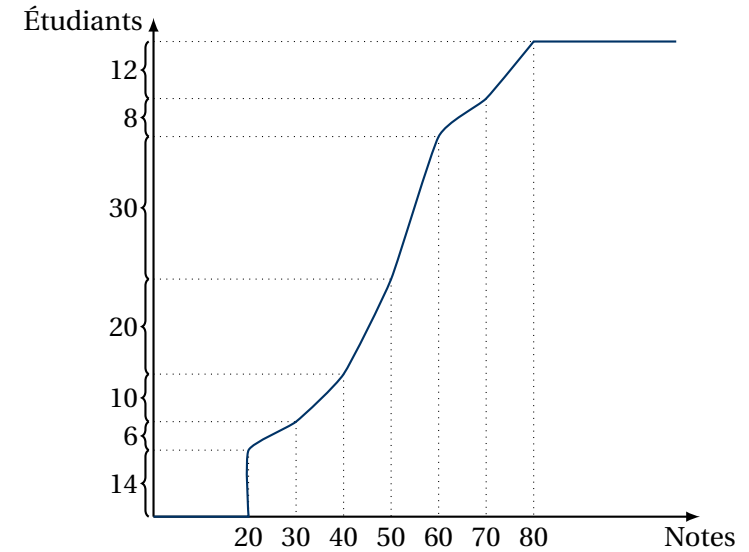


obligatoire.

POLYgone DES FRÉquences CUMULÉES CroISSANTES. La procédure est différente pour des modalités ponctuelles ou regroupées par classes.

- ▶ Pour des modalités ponctuelles, on place les points $A_i(a_i, f_i^c)$. On va alors relier les points $A_i(a_i, f_i^c)$ à $B_i(a_{i+1}, f_i^c)$ puis B_i à $A_{i+1}(a_{i+1}, f_{i+1}^c)$. Schématiquement on trace un trait horizontal puis un trait vertical. *La courbe obtenue ressemble à celle d'une fonction de répartition de loi discrète.*
- ▶ Pour des modalités regroupées en classe $]a_i, a_{i+1}]$ (resp. $[a_i, a_{i+1}[$) on place les points A_i de coordonnées (a_{i+1}, f_i^c) (resp. (a_i, f_i^c)), à la droite (resp. gauche) de l'intervalle donc. On relie ensuite simplement les points A_i par une ligne brisée. *La courbe obtenue ressemble alors à celle d'une fonction de répartition de loi à densité.*

Remarque 2.4 — Cette représentation est pertinente si les individus sont répartis uniformément au sein de la classe. Si vous avez des raisons de penser que ce n'est pas le cas alors il faut séparer votre classe en plusieurs classes.



Par exemple, ici nous avons représenté le tableau d'effectifs d'âges suivant :

Notes en classe	[30, 40[[40, 50[[50, 60[[60, 70[[70, 80[[80, 90[[90, 100[
Effectifs	14	6	10	20	30	8	12

Lire graphiquement le nombre approximatif de notes inférieures à 30, et 56.

Exemple 8 — Notes de DS Par exemple, on peut essayer de représenter le polygone des fréquences cumulées.

2.1.4. Caractéristiques de position

Les caractéristiques de positions ont leur analogue en probabilités : l'espérance et la médiane d'une variable aléatoire⁶. Ce sont des grandeurs dont la vocation est de mesurer la position des données qui constituent la série statistique. En revanche, le mode, défini ci-après, est très peu souvent défini en probabilités : mais ce serait alors, pour une variable aléatoire discrète X , un élément $x \in X(\Omega)$ qui maximise $P(X = x)$.⁷

MODE & CLASSE MODALE.

Définition ALEA.17.6 | Mode & Classe modale

1. On appelle *mode* d'une série statistique x toute modalité de x dont l'effectif est maximal parmi les effectifs de toutes les modalités.
2. Lorsque les modes correspondent à des classes, on appelle alors *classe modale* la classe dont l'effectif est maximal.

Attention Si vos classes sont de tailles différentes alors la classe modale n'est pas forcément la classe qui a le « plus haut » rectangle dans l'histogramme.

Remarque 2.5 — Il est possible qu'une série statistique admette plusieurs modes ou classes modales. D'un point de vue informatique, rechercher un mode revient à rechercher un maximum dans une liste.

MOYENNE.

⁶Mais l'étude de l'existence d'une médiane, pour une variable aléatoire, n'est pas au programme.

⁷Rien ne garantit son existence en revanche.

Définition ALEA.17.7 | Moyenne pour des données ponctuelles

Soit $x = (x_1, \dots, x_n)$ une série statistique quantitative. La *moyenne* de la série, notée \bar{x} , est définie par

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i.$$

Remarque 2.6 — Analogie L'espérance en probabilités.

Proposition ALEA.17.3

Soit $x = (x_1, \dots, x_n)$ une série statistique de modalités (a_1, \dots, a_p) ponctuelles, alors :

$$\bar{x} = \frac{1}{n} \sum_{j=1}^p n_j a_j = \sum_{j=1}^p f_j a_j.$$

On dit alors que \bar{x} est la *moyenne pondérée* des a_j par les fréquences f_j .

Preuve Constater simplement que dans $\sum_{i=1}^n x_i$, chaque x_i apparaît un nombre n_j de fois où n_j désigne l'effectif de x_i .

Quand on travaille avec des données regroupées par classes cette définition n'est pas utilisable. Dans cette situation on va alors considérer que les valeurs sont uniformément réparties dans les intervalles et prendre pour moyenne de la série la moyenne des milieux des intervalles pondérés par les effectifs.

Définition ALEA.17.8 | Moyenne pour des données regroupées en classes

Soit $x = (x_1, \dots, x_n)$ une série statistique quantitative. Supposons que les modalités $(a_j)_{j \in [1, p]}$ de cette série correspondent à des intervalles $[b_j, c_j]$. On définit alors la *moyenne* par :

$$\bar{x} = \frac{1}{n} \sum_{j=1}^p n_j \left(\frac{b_j + c_j}{2} \right) = \sum_{j=1}^p f_j \left(\frac{b_j + c_j}{2} \right).$$

Proposition ALEA.17.4 | Propriétés de la moyenne

Soient x et y deux séries statistiques quantitatives dont les modalités sont à valeurs dans le même ensemble, et d'effectifs respectifs n et m .

1. **(Moyenne d'une série affine)** Soit $(a, b) \in \mathbf{R}^2$ et soit u la série statistique définie par

$$\forall i \in \llbracket 1, n \rrbracket, \quad u_i = ax_i + b, \quad \text{alors : } \bar{u} = a\bar{x} + b.$$

2. **(Moyenne d'un mélange)** Soit z la série statistique obtenue en « concaténant » les séries x et y , *i.e.*

$$z = (x_1, \dots, x_n, y_1, \dots, y_m), \quad \text{alors : } \bar{z} = \frac{n\bar{x} + m\bar{y}}{n + m}.$$

Preuve1. 2. 

MÉDIANE. On suppose ici que nos modalités sont des réels. La moyenne est fortement influencée par les valeurs extrêmes, donc dans ce cas la donnée de la moyenne est

assez peu instructive, et on privilégie une autre quantité appelée la médiane.

Remarque 2.7 — Cela signifie qu'il y a au moins autant d'éléments plus grands que d'éléments plus petits.

Définition ALEA.17.9 | Médiane pour des données ponctuelles

On appelle *médiane* d'une série statistique x de taille n tout réel m , « partageant la série d'observations en deux », *i.e.* tel que :

$$\# (\{i \in \llbracket 1, n \rrbracket, x_i \leq m\}) \geq \frac{n}{2} \quad \text{et} \quad \# (\{i \in \llbracket 1, n \rrbracket, x_i \geq m\}) \geq \frac{n}{2}.$$

En pratique on prend souvent comme médiane la valeur d'une modalité. Dans ce cas, un individu dont le caractère correspond à la médiane est dit être un *individu médian*.

Remarque 2.8 — Les symboles \leq, \geq sont justifiés par le fait que $\frac{n}{2}$ n'est pas toujours un entier. Le signe d'égalité n'aurait dans ce cas aucun sens, puisqu'un cardinal est toujours entier.

Proposition ALEA.17.5 | Cas de modalités ordonnées

Soit x une série statistique de taille n dont les modalités sont données dans l'ordre croissant $a_1 < a_2 < \dots < a_n$.

1. Si n est impair, alors : $a_{\frac{n+1}{2}}$ est une médiane.
2. Si n est pair, alors : tout nombre de l'intervalle $\left[a_{\frac{n}{2}}, a_{\frac{n}{2}+1} \right]$ est une médiane. En particulier,

$$\frac{a_{\frac{n}{2}} + a_{\frac{n}{2}+1}}{2} \text{ est } \underline{\text{une}} \text{ médiane.}^8$$

Attention

Il y a en général plusieurs médianes.

⁸C'est souvent celle-ci qui, par abus de langage, nous appelons parfois *la* médiane.

Remarque 2.9 —

- ▶ La médiane a , par rapport à la moyenne, l'avantage d'être peu influencée par les valeurs extrêmes. Elle est alors plus représentative que la moyenne lorsque la série comporte des valeurs très grandes ou très petites.
Par exemple, en France en 2014 le salaire moyen mensuel était de 1934 euros pour les femmes et 2389 euros pour les hommes tandis que le salaire médian mensuel était de 1619 euros pour les femmes et 1882 euros pour les hommes. La différence s'explique par le fait que les très hauts salaires, même s'ils sont peu nombreux, tirent la moyenne vers le haut.
- ▶ Sur le polygone des fréquences cumulées croissantes on lit une médiane assez simplement. Il suffit de chercher l'abscisse du point de la courbe d'ordonnée $\frac{1}{2}$.

Définition ALEA.17.10 | Médiane pour des données regroupées en classes

Soit x une série statistique dont les modalités sont regroupées en classes. On définit la *médiane* de x comme l'abscisse du point de la courbe des fréquences cumulées croissantes d'ordonnée $\frac{1}{2}$.

2.1.5. Caractéristiques de dispersion

L'idée des caractéristiques de dispersion est de donner une idée de la répartition de la série autour de sa moyenne ou de sa médiane. Les valeurs sont elles relativement proches de la moyenne ou existe-t-il des valeurs très grandes et très petites? L'analogue probabilité est donc évident : l'espérance, la variance, l'écart-type ...

VALEURS EXTRÊMES ET ÉTENDUE.**Définition ALEA.17.11**

Soit x une série statistique de taille n à valeurs réelles.

1. Si les modalités sont en nombre fini, on appelle *valeurs extrêmes* de la série x

les nombres :

$$\min_{i \in \llbracket 1, n \rrbracket} x_i, \quad \text{et} \quad \max_{i \in \llbracket 1, n \rrbracket} x_i.$$

Il s'agit donc des modalités maximales et minimales.

2. Si les modalités sont regroupées en classes, on appelle *valeurs extrêmes* de la série la borne supérieure de la classe maximale et la borne inférieure de la classe minimale.

On appelle *étendue* de la série statistique la différence entre la valeur maximale et la valeur minimale.

Remarque 2.10 — L'étendue est facile à déterminer mais ne délivre que très peu d'informations car elle est très fortement affectée par les valeurs extrêmes. Par exemple en France le revenu annuel se situe entre 0 euros et environ 7 millions, ce qui ne nous donne pas vraiment une idée de la répartition des salaires dans la population.

QUANTILES. Plutôt que d'introduire des quantités qui découpent en deux une série statistique, on peut également partager en trois quatre *etc.*, de manière générale si on souhaite partager en parties de taille $n \cdot \alpha$ avec $\alpha \in [0, 1]$, on parle alors de quantile d'ordre α .

Définition ALEA.17.12 | Quartile - $\alpha = \frac{1}{4}$

Soit x une série statistique de taille n à valeurs réelles. On appelle premier quartile de la série x tout réel Q_1 tel que

$$\# (\{i \in \llbracket 1, n \rrbracket, x_i \leq Q_1\}) \geq \frac{n}{4} \quad \text{et} \quad \# (\{i \in \llbracket 1, n \rrbracket, x_i \geq Q_1\}) \geq \frac{3n}{4}.$$

De-même on appelle troisième quartile de la série x tout réel Q_3 tel que

$$\# (\{i \in \llbracket 1, n \rrbracket, x_i \leq Q_3\}) \geq \frac{3n}{4} \quad \text{et} \quad \# (\{i \in \llbracket 1, n \rrbracket, x_i \geq Q_3\}) \geq \frac{n}{4}.$$

Remarque 2.11 — Il faut comprendre ces définitions de la même manière que pour la médiane : la première signifie qu'il y a au moins 1/4 des observations qui sont inférieures ou égales à Q_1 , de-même pour les autres.

Remarque 2.12 — La médiane est donc aussi parfois appelée *deuxième quartile* ou *quantile d'ordre 2*.

On peut définir la notion de décile de manière similaire.

Définition ALEA.17.13 | Décile - $\alpha = \frac{1}{10}$

] Soit x une série statistique de taille n à valeurs réelles. Pour $j \in \llbracket 1, 9 \rrbracket$, on appelle *j-ème décile* de la série x tout réel d_j tel que :

$$\# \left(\left\{ i \in \llbracket 1, n \rrbracket, x_i \leq d_j \right\} \right) \geq \frac{jn}{10} \quad \text{et} \quad \# \left(\left\{ i \in \llbracket 1, n \rrbracket, x_i \geq d_j \right\} \right) \geq \frac{(10-j)n}{10}.$$

Ces deux notions se généralisent avec la notion de quantiles.

Définition ALEA.17.14

Soit x une série statistique de taille n à valeurs réelles. Soit $t \in [0, 1]$, on appelle α -*quantile* de la série x tout réel q_α tel que

$$\# \left(\left\{ i \in \llbracket 1, n \rrbracket, x_i \leq q_\alpha \right\} \right) \geq nt \quad \text{et} \quad \# \left(\left\{ i \in \llbracket 1, n \rrbracket, x_i \geq q_\alpha \right\} \right) \geq (1-\alpha)n.$$

Comme pour les quartiles et les déciles il n'y pas unicité du α -quantile.

Remarque 2.13 —

- ▶ Le premier quartile est alors un $\frac{1}{4}$ -quantile, *etc.*
- ▶ Si la série statistique est donnée via un regroupements en classes alors on déterminera un quantile d'ordre α en prenant l'abscisse d'un point du polygone des fréquences cumulées croissantes d'ordonnée α .

Définition ALEA.17.15 | Écarts

Soit x une série statistique de taille n à valeurs réelles. On appelle

1. *écart interquartile* la différence $Q_3 - Q_1$ noté parfois «IQR».
2. *intervalle interquartile* l'intervalle $[Q_1, Q_3]$.

De-même on appelle

1. *écart interdécile* la différence $d_9 - d_1$.
2. *intervalle interdécile* l'intervalle $[d_1, d_9]$

La moitié au moins de la population se trouve dans l'intervalle interquartile.

VISUALISATION GRAPHIQUE DES QUANTILES : LE DIAGRAMME DE TUKEY (OU « BOÎTE À MOUSTACHE »).

On peut représenter de manière graphique l'étendue, les quartiles et la médiane en dessinant un diagramme dit *diagramme de TUKEY* conçu de la manière suivante :

- ▶ au centre une boîte allant du premier au troisième quartile, séparée en deux par la médiane;
- ▶ de chaque côté une moustache allant du minimum au premier quartile pour l'une, et du troisième quartile au maximum pour l'autre.

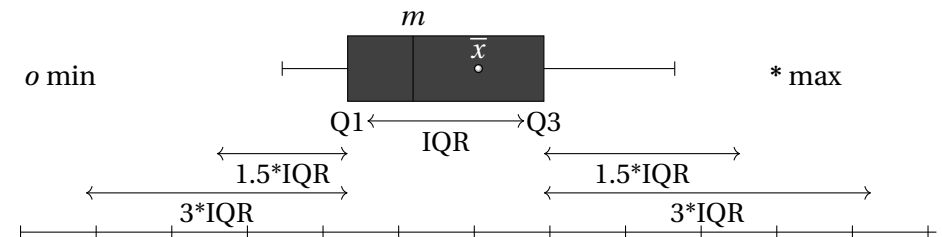


FIG. ALEA.17.1. : Représentation en boîte à moustache d'une série statistique

VARIANCE ET ÉCART-TYPE.

Définition ALEA.17.16 | Variance & Écart-Type

1. La *variance* d'une série statistique quantitative à valeurs réelles $x = (x_1, x_2, \dots, x_n)$ de nombre de modalités finies a_1, \dots, a_p , est le nombre V_x défini par :

$$V_x = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n} \sum_{j=1}^p n_j (a_j - \bar{x})^2 = \sum_{j=1}^p f_j (a_j - \bar{x})^2.$$

2. L'*écart-type* d'une telle série, noté σ_x , est défini par :

$$\sigma_x = \sqrt{V_x}.$$

Dans le cas d'une série regroupée en classes on prendra pour valeurs a_j les centres des classes.

V_x est la moyenne des carrés des écarts à la moyenne donc est toujours positive, d'où la bonne définition de l'écart-type.

Proposition ALEA.17.6 | Variance nulle

Soit une série statistique quantitative à valeurs réelles $x = (x_1, x_2, \dots, x_n)$ de nombre de modalités finies a_1, \dots, a_p , alors :

$$V_x = 0 \iff \text{toutes les données de la série sont égales.}$$

Preuve On raisonne par exemple avec l'expression en fréquences de la variance.

$$\begin{aligned} \sum_{j=1}^p f_j (a_j - \bar{x})^2 &= 0 \\ \iff \forall i \in [1, p], f_j (a_j - \bar{x})^2 &= 0, & \left. \begin{array}{l} \text{somme de termes positifs} \\ f_j \neq 0 \end{array} \right\} \\ \iff \forall i \in [1, p], (a_j - \bar{x})^2 &= 0, \\ \iff \forall i \in [1, p], a_j &= \bar{x}. \end{aligned}$$

C'est ce qu'on voulait.

Proposition ALEA.17.7 | KÖNIG-HUYGENS

Soit x une série statistique quantitative à valeurs réelles. Alors :

$$V_x = \overline{x^2} - \bar{x}^2.$$

Preuve C'est un calcul direct, comme pour les variables aléatoires.

$$\begin{aligned} V_x &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^n (x_i^2 + \bar{x}^2 - 2x_i\bar{x}) = \frac{1}{n} \sum_{i=1}^n x_i^2 + \frac{1}{n} \sum_{i=1}^n \bar{x}^2 - 2\frac{\bar{x}}{n} \sum_{i=1}^n x_i \\ &= \overline{x^2} + \bar{x}^2 - 2\bar{x}^2 = \overline{x^2} - \bar{x}^2. \end{aligned}$$

Remarque 2.14 – Interprétation

- ▶ Plus la variance est grande, plus la série s'éloigne de sa moyenne, et plus la série est donc «étalée». Inversement, plus la variance est proche de zéro et plus la série est concentrée autour de sa moyenne.
- ▶ La variance ne donne pas d'informations sur une éventuelle asymétrie de la série.

Remarque 2.15 – Homogénéité L'intérêt de l'écart-type par rapport à la variance est que l'écart-type s'exprime dans les mêmes unités que les modalités de la série. On pourra alors faire des calculs faisant intervenir modalités, moyenne et écart-type (par exemple dans des situations d'estimation de paramètres ou de test statistique d'hypothèses).

Proposition ALEA.17.8 | Propriétés de la variance

Soit x une série statistique quantitative réelle, $(a, b) \in \mathbf{R}^2$ et y la série statistique $y = ax + b$. On a alors

$$V_y = a^2 V_x \quad \text{et donc} \quad \sigma_y = |a| \sigma_x.$$

Preuve On a

$$\begin{aligned} V_y &= \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 = \frac{1}{n} \sum_{i=1}^n (ax_i + b - (a\bar{x} + b))^2 \\ &= \frac{1}{n} \sum_{i=1}^n a^2(x_i - \bar{x})^2 \\ &= a^2 V_x. \end{aligned}$$

Pour la deuxième, prendre simplement la racine dans l'expression.

2.1.6. En Python : calcul des grandeurs univariées

On suppose dans la suite que toutes les données d'une série à nombre de modalités fini sont contenues dans une liste L donnée en paramètre.

Espérance

```
def moyenne(L):
    """
    Renvoie l'espérance
    """
    S = 0
    for x in L:
        S += x
    return S/len(L)
```

Variance

```
from moyenne import *
def variance(L):
    """
    Renvoie la variance, version KH
    """
```

```
esp = moyenne(L)
V = 0
for x in L:
    V += x**2
return V/len(L) - esp**2
```

On peut également créer une fonction qui retourne une médiane, après avoir triée la liste. Celle-ci peut bien entendu être adaptée à tous les quantiles. On peut également, après recherche du minimum et du maximum, renvoyer l'étendue de la série.

Médiane

```
def mediane(L):
    """
    Cherche la médiane d'une liste, après tri rapide des
    observations
    """
    L = tri_rapide_rec(L)
    n = len(L)
    if n % 2 == 1:
        # Nombre impair d'observations, on prend le milieu
        return L[n//2]
    else:
        # Nombre pair d'observations, on prend la moyenne des deux
        ↪ termes du milieu
        return (L[n//2-1] + L[n//2])/2
```

On peut ensuite tester si la quantité retournée est bien une médiane, en contrôlant la définition.

```
def mediane_verif(L, m):
    """
    Renvoie True si m est bien une médiane de L
    """
    nb_inf = 0
```



```

nb_sup = 0
for x in L:
    if x >= m:
        nb_sup += 1
    if x <= m:
        nb_inf += 1
return nb_sup >= len(L)/2 and nb_inf >= len(L)/2

```

Étendue d'une série

```

def etendue(L):
    """
    Cherche l'étendue de L, sans la trier
    """
    min = L[0]
    max = L[0]
    for i in range(1, len(L)):
        if L[i] < min:
            min = L[i]
        elif L[i] > max:
            max = L[i]
    return max - min

```

Test sur une série de données

Par exemple, si l'on considère des notes de DS.

```

>>> Notes_DS = [5.2, 7.4, 10.1, 12.8, 7.4, 10.0, 14.5, 6.5, 4.4,
↳ 3.3, 11.8, 11.6, 8.0, 8.5, 10.2, 11.9, 8.3, 9.6, 10.0, 9.6,
↳ 9.4, 10.2, 10.3, 15.1, 15.6, 13.2, 9.9, 9.6, 5.9, 14.5, 6.4,
↳ 10.5, 6.7, 6.3, 10.9, 10.1, 7.5]
>>> moy = moyenne(Notes_DS)
>>> moy
9.545945945945943
>>> var = variance(Notes_DS)

```

```

>>> var
8.503564645726854
>>> med = mediane(Notes_DS)
>>> med
9.9
>>> mediane_verif(Notes_DS, med)
True

```

2.2. Bivariées

On a vu dans les sous-sections précédentes diverses manières d'extraire de l'information d'un échantillon statistique. Lorsque l'on ne dispose plus d'un seul mais de plusieurs échantillons statistiques, on peut, au delà de la simple étude des échantillons, étudier les éventuels liens entre eux, c'est l'objet de cette dernière sous-section. Pour des raisons de simplicité on se limitera à deux échantillons. Cette sous-section est essentiellement une reformulation dans le cadre déterministe des notions aléatoires déjà vues dans le [Chapter ALEA.15](#) (covariance, coefficient de corrélation, *etc.*).

2.2.1. Série statistique

On va s'intéresser ici à deux caractères quantitatifs d'une même population. On notera n la taille de l'échantillon étudié et (x, y) les deux caractères étudiés. L'observation des valeurs des caractères se traduit par un échantillon d'éléments de \mathbf{R}^2 , que l'on note :

$$((x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)).$$

Définition ALEA.17.17 | Série statistique bivariée

1. Un caractère (x, y) de la population est une donnée *qualitative* ou *quantitative* attachée à chaque individu de la population. On notera (x_i, y_i) la valeur du caractère (x, y) pour un individu i .
2. Soient x et y deux séries statistiques de taille n . Alors on appelle *série statistique bivariée* une famille de \mathbf{R}^2 du type

$$((x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)).$$

3. On appelle nuage point associé à l'échantillon (x, y) le tracé de tous les points de coordonnées $M(x_k, y_k)$ pour $k \in \llbracket 1, n \rrbracket$.

Définition ALEA.17.18 | Modalités

On appelle *modalités* d'un caractère bivarié les valeurs possibles qu'il peut prendre.

$x \backslash y$	b_1	...	b_j	...	b_q	Totaux
a_1	$n_{1,1}$...	$n_{1,j}$...	$n_{1,q}$	$n_{1,\bullet}$
\vdots	\vdots		\vdots		\vdots	\vdots
a_i	$n_{i,1}$...	$n_{i,j}$...	$n_{i,q}$	$n_{i,\bullet}$
\vdots	\vdots		\vdots		\vdots	\vdots
a_p	$n_{p,1}$...	$n_{p,j}$...	$n_{p,q}$	$n_{p,\bullet}$
Totaux	$n_{\bullet,1}$...	$n_{\bullet,j}$...	$n_{\bullet,q}$	n

$n_{i,j}$ est le cardinal de l'ensemble des individus présentant à la fois les modalités a_i et b_j . Pour $(i, j) \in \llbracket 1, p \rrbracket \times \llbracket 1, q \rrbracket$, on pose

$$n_{i,\bullet} = \sum_{j=1}^q n_{i,j} \quad \text{et} \quad n_{\bullet,j} = \sum_{i=1}^p n_{i,j}.$$

On a alors

$$n = \sum_{i=1}^p \sum_{j=1}^q n_{i,j} = \sum_{i=1}^p n_{i,\bullet} = \sum_{j=1}^q n_{\bullet,j}.$$

2.2.2. Effectifs

Notons (a_1, \dots, a_p) les modalités du caractère x (éventuellement des classes) et (b_1, \dots, b_q) les modalités du caractère y . Le plus souvent on regroupe les individus par modalités, on obtient alors le tableau d'effectifs suivant :

Définition ALEA.17.19 | Fréquence marginale & Effectif marginale

Pour $(i, j) \in \llbracket 1, p \rrbracket \times \llbracket 1, q \rrbracket$, on définit

- ▶ l'*effectif conjoint* (resp. *fréquence conjointe*) des modalités (a_i, b_j) la suite double : $n_{i,j} = n_{i,j}$ (resp. $f_{i,j} = \frac{n_{i,j}}{n}$).
- ▶ L'*effectif marginal de la modalité* a_i (resp. *fréquence marginale de la modalité* a_i) par : $n_{i,\bullet}$ (resp. $f_{i,\bullet} = \frac{n_{i,\bullet}}{n}$).
- ▶ L'*effectif marginal de la modalité* a_j (resp. *fréquence marginale de la modalité* a_j) par : $n_{\bullet,j}$ (resp. $f_{\bullet,j} = \frac{n_{\bullet,j}}{n}$).

Proposition ALEA.17.9 | Propriétés des fréquences/effectifs marginaux

Soit une série (x, y) statistique de taille n admettant un nombre fini de modalités ou à défaut de classes.

- $\sum_{i=1}^p \sum_{j=1}^q f_{i,j} = \sum_{i=1}^p f_{i,\bullet} = \sum_{j=1}^q f_{\bullet,j} = 1,$
- $f_{i,\bullet} = \sum_{j=1}^q f_{i,j}$ et $f_{\bullet,j} = \sum_{i=1}^p f_{i,j}.$

Preuve Diviser par n les égalités précédemment établies sur les effectifs.

Remarque 2.16 — Analogie avec les probabilités Reprenons les notations précédentes, *i.e.* une série statistique bivariée de taille $p \times q$ admettant un nombre fini de modalités ou à défaut de classes notées $(a_i, b_j)_{(i,j) \in \llbracket 1, p \rrbracket \times \llbracket 1, q \rrbracket}$. On note $(n_{i,j})_{(i,j) \in \llbracket 1, p \rrbracket \times \llbracket 1, q \rrbracket}$ les effectifs correspondant. Alors si l'on considère un vecteur aléatoire discret (X, Y) tel que⁹ :

$$(X, Y)(\Omega) = \{(a_i, b_j), (i, j) \in \llbracket 1, p \rrbracket \times \llbracket 1, q \rrbracket\},$$

$$\mathbf{P}(X = a_i, Y = b_j) = \frac{n_{i,j}}{n}, \quad (i, j) \in \llbracket 1, p \rrbracket \times \llbracket 1, q \rrbracket.$$

On vérifie sans peine que :

$$\sum_{i=1}^p \sum_{j=1}^q \frac{n_{i,j}}{n} = 1, \quad \frac{n_{i,j}}{n} \geq 0$$

pour tout $(i, j) \in \llbracket 1, p \rrbracket \times \llbracket 1, q \rrbracket$, ce qui garantit l'existence de (X, Y) . Alors les formules précédentes traduisent simplement que

$$\left\{ X = a_i, Y = b_j \right\}_{\llbracket 1, p \rrbracket \times \llbracket 1, q \rrbracket}$$

est un système complet d'évènements, et d'autre part le lien entre loi marginale et loi conjointe. La covariance d'une série statistique que nous allons définir ci-après sera simplement $\mathbf{Cov}(X, Y)$.

⁹on voit donc la série statistique bivariée comme la réalisation d'un couple aléatoire discret, de loi la fréquence conjointe de ladite modalité.

2.2.3. Caractéristiques de position & dispersion

On définit les moyennes et variances de manière similaire au cas univarié, pour chacune des séries x et y . On peut les exprimer en fonction des notations propres aux séries bivariées.

Proposition ALEA.17.10 | Expression de la moyenne

- (Nombre fini de modalités)** Soit une série (x, y) statistique de taille n admettant un nombre fini de modalités ou à défaut de classes.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^p n_{i,\bullet} a_i = \sum_{i=1}^p f_{i,\bullet} a_i, \quad \bar{y} = \frac{1}{n} \sum_{j=1}^q n_{\bullet,j} b_j = \sum_{j=1}^q f_{\bullet,j} b_j.$$

- (Regroupement en classes)** On remplace les a_i dans les définitions des moyennes par les milieux des classes, **mais les formules précédentes ne sont plus vraies.**

Preuve Montrons **1** pour x . Les modalités de la série statistique univariée x sont les $a_i, i \in \llbracket 1, p \rrbracket$. Chaque modalité apparaît avec une fréquence $\sum_{j=1}^q f_{i,j}$, qui est la proportion des modalités $(a_i, y_1), \dots, (a_i, y_q)$ dans la série statistique bivariée de départ. Ainsi,

$$\bar{x} = \sum_{i=1}^p \left(\sum_{j=1}^q f_{i,j} \right) a_i,$$

on reconnaît alors les fréquences partielles $\bar{x} = \sum_{i=1}^p f_{i,\bullet} a_i$. La formule est démontrée.

Définition ALEA.17.20 | Point moyen du nuage

Le point de coordonnées (\bar{x}, \bar{y}) est appelé *point moyen du nuage* associé à la série bivariée (x, y) .

Proposition ALEA.17.11 | Expression de la variance

1. **(Nombre fini de modalités)** Soit une série (x, y) statistique de taille n admettant un nombre fini de modalités ou à défaut de classes.

$$\begin{aligned} V_x &= \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^p n_{i,\bullet} (a_i - \bar{x})^2 = \sum_{i=1}^p f_{i,\bullet} (a_i - \bar{x})^2, \\ V_y &= \frac{1}{n} \sum_{k=1}^n (y_k - \bar{y})^2 = \frac{1}{n} \sum_{j=1}^q n_{\bullet,j} (b_j - \bar{y})^2 = \sum_{j=1}^q f_{\bullet,j} (b_j - \bar{y})^2. \end{aligned}$$

2. **(Regroupement en classes)** On remplace les a_i dans les définitions des moyennes par les milieux des classes, **mais les formules précédentes ne sont plus vraies.**

Remarque 2.17 — Là aussi, si les données sont regroupées en classes on prend pour a_i et b_j les centres des classes

Preuve Même preuve que pour la moyenne.

Définition ALEA.17.21 | Covariance

Soit une série (x, y) statistique de taille n admettant un nombre fini de modalités ou à défaut de classes. On appelle *covariance de x et de y* , notée $C_{x,y}$, la quantité

$$\begin{aligned} C_{x,y} &= \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})(y_k - \bar{y}) \quad \text{ou de manière équivalente} \\ C_{x,y} &= \frac{1}{n} \sum_{i=1}^p \sum_{j=1}^q n_{i,j} (a_i - \bar{x})(b_j - \bar{y}). \end{aligned}$$

Remarque 2.18 — Interprétation

- ▶ Si $C_{x,y} < 0$ alors x et y ont tendance à varier dans des sens opposés (quand l'un augmente l'autre diminue), si $C_{x,y} > 0$ alors ils ont tendance à varier dans le même sens.
- ▶ On a $\mathbf{Cov}(x, x) = V_x$.

Proposition ALEA.17.12 | KÖNIG-HUYGENS

Soit une série (x, y) statistique de taille n admettant un nombre fini de modalités ou à défaut de classes. Alors :

$$C_{x,y} = \bar{xy} - \bar{x} \cdot \bar{y}.$$

Remarque 2.19 — C'est la moyenne des produits moins le produits des moyennes

Preuve



$$\begin{aligned} C_{x,y} &= \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})(y_k - \bar{y}) \\ &= \frac{1}{n} \sum_{k=1}^n (x_k y_k - \bar{x} y_k - \bar{y} x_k + \bar{x} \cdot \bar{y}) \\ &= \frac{1}{n} \sum_{k=1}^n x_k y_k - \frac{1}{n} \sum_{k=1}^n \bar{x} y_k - \frac{1}{n} \sum_{k=1}^n \bar{y} x_k + \frac{1}{n} \sum_{k=1}^n \bar{x} \cdot \bar{y} \\ &= \bar{xy} - \bar{x} \cdot \bar{y} - \bar{x} \cdot \bar{y} + \bar{x} \cdot \bar{y} \\ &= \bar{xy} - \bar{x} \cdot \bar{y} \end{aligned}$$

développement du produit
linéarité de la somme

Proposition ALEA.17.13 | Variance d'une somme

Soit une série (x, y) statistique de taille n admettant un nombre fini de modalités ou à défaut de classes. Alors :

$$V_{x+y} = V_x + 2C_{x,y} + V_y$$

Preuve

$$\begin{aligned}
 V_{x+y} &= \overline{(x+y)^2} - \overline{x+y}^2 \\
 &= \overline{x^2 + 2xy + y^2} - (\overline{x} + \overline{y})^2 \\
 &= \overline{x^2} + 2\overline{xy} + \overline{y^2} - \overline{x^2} - \overline{y^2} - 2\overline{x} \cdot \overline{y} \quad \left. \vphantom{\overline{x^2}} \right\} \text{linéarité de l'espérance} \\
 &= \overline{x^2} - \overline{x}^2 + \overline{y^2} - \overline{y}^2 + 2(\overline{xy} - \overline{x}\overline{y}) \quad \left. \vphantom{\overline{x^2}} \right\} \text{formule de KÖNIG-HUYGENS} \\
 &= V_x + 2C_{x,y} + V_y.
 \end{aligned}$$

On a également une « inégalité de CAUCHY-SCHWARZ » pour la covariance.

Définition/Proposition ALEA.17.1 | Inégalité de CAUCHY-SCHWARZ/ Coefficient de corrélation.

Soit une série (x, y) statistique de taille n admettant un nombre fini de modalités ou à défaut de classes.

1. (Inégalité de CAUCHY-SCHWARZ)

$$|C_{x,y}| \leq \sqrt{V_x} \sqrt{V_y} \quad \text{ou encore} \quad |C_{x,y}| \leq \sigma_x \sigma_y.$$

Si x, y sont d'écart-type non nul. On appelle *coefficient de corrélation entre x et y* la quantité $\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_x \sigma_y} \in [-1, 1]$.

2. (Cas d'égalité) $\rho(x, y) = \pm 1 \iff \exists a, b \in \mathbf{R}, y = ax + b$, i.e. la série y dépend de x et de manière affine.

Preuve

 On définit l'application

$$P \begin{cases} \mathbf{R} & \longrightarrow & \mathbf{R}, \\ t & \longrightarrow & V_{x+ty}. \end{cases}$$

En développant la variance, d'après une proposition précédente, il vient :

$$\forall t \in \mathbf{R}, \quad P(t) = \sigma_x^2 + 2tC_{x,y} + t^2\sigma_y^2.$$

L'application P est donc une application polynomiale de degré 2. De plus, d'après les propriétés de la variance on a

$$\forall t \in \mathbf{R}, \quad P(t) \geq 0.$$

Comme P est de signe constant il ne peut pas admettre deux racines réelles distinctes, son discriminant est négatif ou nul, c'est-à-dire

$$\Delta = 4C_{x,y}^2 - 4\sigma_x^2\sigma_y^2 \leq 0.$$

D'où

$$C_{x,y}^2 \leq \sigma_x^2\sigma_y^2.$$

Par croissance de la fonction racine carrée on a alors

$$|C_{x,y}| \leq \sigma_x \sigma_y.$$

Pour le cas d'égalité, on constate qu'il est obtenu lorsque $\Delta = 0 \iff P$ possède une racine réelle (double a fortiori), i.e. si et seulement si il existe $t \in \mathbf{R}, P(t) = V_{x+ty} = 0$ donc si et seulement si la série $x + ty$ est constante. C'est ce qu'il fallait montrer.

2.2.4. Ajustement affine

Il est courant, en physique-chimie, en sciences industrielles, ou plus généralement dans toute discipline expérimentale (biologie, chimie, économie, ...), d'avoir à comparer des données expérimentales et de conjecturer une éventuelle dépendance linéaire entre deux paramètres donnés. Vous pourriez avoir ce besoin lors de vos TIPE. Notez qu'il est aussi possible d'étudier les dépendances polynomiales entre deux paramètres pour un degré quelconque, nous n'aborderons pas ce point ici.

L'idée de l'ajustement affine est la suivante : on dispose de deux séries statistiques (souvent expérimentales) x et y et on soupçonne qu'il existe une relation les liant de la forme $y = ax + b$: par exemple après avoir trouvé un coefficient de corrélation proche de un. On veut alors chercher la droite d'équation $y = ax + b$ qui passe « le mieux » par notre nuage de points. Parfois on sait que la relation existe et on veut déterminer a et b .

Plus précisément, soit $(x_i, y_i)_{1 \leq i \leq n}$ avec $n \geq 1$ est un nuage de n points provenant de séries statistiques x, y . En regardant un dessin, nous voyons que si l'on approche le

nuage par la droite $y = ax + b$ avec $(a, b) \in \mathbf{R}^2$, alors l'écart entre cette droite et le nuage au point $x_i, i \in \llbracket 1, n \rrbracket$, est donné par : $y_i - ax_i - b$.

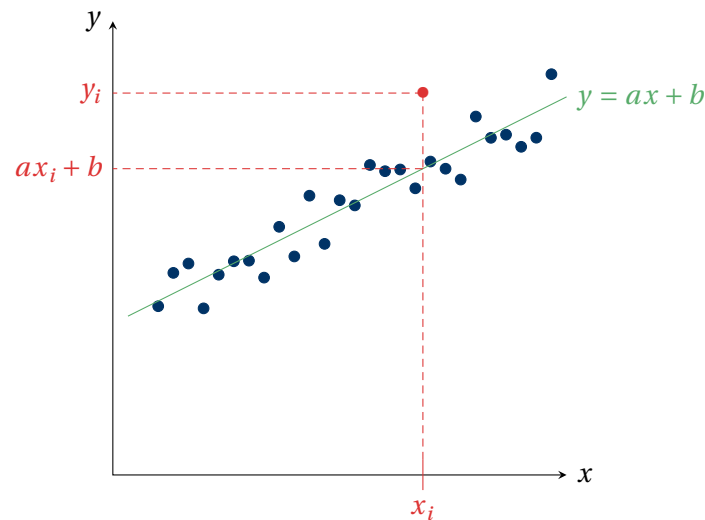


FIG. ALEA.17.2. : Problème de régression linéaire

Ainsi, on pourrait se poser par exemple la question de la minimisation en a, b des quantités suivantes :

$$\max_{1 \leq i \leq n} |y_i - ax_i - b|, \quad \sum_{i=1}^n |y_i - ax_i - b|, \quad F(a, b) = \sum_{i=1}^n (y_i - ax_i - b)^2,$$

de sorte que quand l'une ou l'autre de ces trois quantités est petite, tous les écarts à la droite sont également petits. Dans le dernier cas, on parle de *minimisation au sens des moindres carrés* et c'est cette minimisation que nous allons essayer de réaliser car c'est pour celle-ci que les calculs sont les plus simples. Nous pouvons résoudre ce problème de deux manières. Il s'agira donc de minimiser la fonction de deux variables

$$F \begin{cases} \mathbf{R}^2 & \longrightarrow \mathbf{R} \\ (a, b) \in \mathbf{R}^2 & \longrightarrow \sum_{i=1}^n (y_i - ax_i - b)^2, \end{cases}$$

i.e. de déterminer

$$\inf_{(a,b) \in \mathbf{R}^2} \sum_{i=1}^n (y_i - ax_i - b)^2 = \inf_{(a,b) \in \mathbf{R}^2} F(a, b).$$

HEURISTIQUE. Faisons une heuristique. Si y est réellement affine en x , *i.e.* de la forme $y = ax + b$ avec $a, b \in \mathbf{R}$, alors d'après les propriétés de la covariance et de l'espérance déjà établie, nous avons :

$$\bar{y} = \overline{ax + b} = a\bar{x} + b \implies \text{la droite passe par le point moyen, } b = \bar{y} - a\bar{x}$$

$$\text{Cov}(y, x) = a\text{Cov}(x, x) + 0 \implies a = \frac{\text{Cov}(y, x)}{\sigma_x^2}.$$

Il s'avère que le couple (a, b) obtenu précédemment, noté (a^*, b^*) dans la suite, sera également la solution obtenue au sens des moindres carrés. C'est ce que nous montrons dès à présent.

Théorème ALEA.171 | Existence de la droite des moindres carrés

Soit (x, y) une série statistique double constituée d'une suite de couples $((x_k, y_k))_{1 \leq k \leq n}$. Alors (a^*, b^*) défini ci-dessous est l'unique minimum global de F :

$$a^* = \frac{C_{x,y}}{\sigma_x^2}, \quad b^* = \bar{y} - a^*\bar{x}.$$

La droite de régression par la méthode des moindres carrés de y en x a donc pour équation :

$$y = \frac{C_{x,y}}{\sigma_x^2}(x - \bar{x}) + \bar{y}.$$

PREMIÈRE DÉMONSTRATION : EN OPTIMISANT DES TRINÔMES. Soit $(a, b) \in \mathbf{R}^2$. Comme



$$\frac{\partial F}{\partial b}(a, b) = -2 \sum_{k=1}^n (ax_k + b - y_k) = -2 \left[a \sum_{k=1}^n x_k + b \underbrace{\sum_{k=1}^n 1}_n - \sum_{k=1}^n y_k \right] = 0$$

$$\iff b = \frac{\sum_{k=1}^n y_k}{n} - a \frac{\sum_{k=1}^n x_k}{n} = \bar{y} - a\bar{x},$$

et que pour tout $a \in \mathbf{R}$, le graphe de $b \mapsto F(a, b)$ est une parabole orientée vers le haut, nous avons :

$$\forall a, b \in \mathbf{R}, \quad F(a, b) \geq F(a, \bar{y} - a\bar{x}).$$

On considère ensuite :

$$f : a \in \mathbf{R} \mapsto F(a, \bar{y} - a\bar{x}).$$

Puisque

$$\begin{aligned} f(a) &= \sum_{k=1}^n [a(x_k - \bar{x}) - (y_k - \bar{y})]^2 \\ &= a^2 \sum_{k=1}^n (x_k - \bar{x})^2 - 2a \sum_{k=1}^n (x_k - \bar{x})(y_k - \bar{y}) + \sum_{k=1}^n (y_k - \bar{y})^2 \end{aligned}$$

identité remarquable

$$= a^2 n\sigma_x^2 - a(2n\mathbf{C}_{x,y}) + n\sigma_y^2.$$

polynôme de degré 2 en a

$$f'(a) = 2an\sigma_x^2 - 2n\mathbf{C}_{x,y}.$$

Comme f est encore un trinôme de graphe une parabole orientée vers le haut, elle est minimale là où f' s'annule, i.e. en $a = \frac{\mathbf{C}_{x,y}}{\sigma_x^2}$. En résumé, nous avons montré :

$$\forall a, b \in \mathbf{R}, \quad F(a, b) \geq F(a, \bar{y} - a\bar{x}) \geq F\left(\frac{\mathbf{C}_{x,y}}{\sigma_x^2}, \bar{y} - \frac{\mathbf{C}_{x,y}}{\sigma_x^2}\bar{x}\right).$$

Cette inégalité prouve que $\left(\frac{\mathbf{C}_{x,y}}{\sigma_x^2}, \bar{y} - \frac{\mathbf{C}_{x,y}}{\sigma_x^2}\bar{x}\right)$ est un minimum global de F . Un calcul de points critiques montre ensuite qu'il s'agit du seul minimum possible :

SECONDE DÉMONSTRATION : EN UTILISANT UNE PROJECTION ORTHOGONALE. Nous notons $E = \mathbf{R}^2$. Interprétons autrement la quantité

$$\inf_{(a,b) \in \mathbf{R}^2} \sum_{i=1}^n (y_i - ax_i - b)^2 = \inf_{(a,b) \in \mathbf{R}^2} F(a, b).$$

On constate que :

$$\begin{aligned} \inf_{(a,b) \in \mathbf{R}^2} \sum_{i=1}^n (y_i - ax_i - b)^2 &= \inf_{(a,b) \in \mathbf{R}^2} \|(y_1, \dots, y_n) - a(x_1, \dots, x_n) - b(1, \dots, 1)\|^2 \\ &= \inf_{(a,b) \in \mathbf{R}^2} \|Y - aX - b\mathbb{1}\|^2 = \inf_{Z \in F} \|Y - Z\|^2 = d(Y, F)^2 \end{aligned}$$

où $X = (x_1, \dots, x_n)$, $Y = (y_1, \dots, y_n)$, $\mathbb{1} = (1, \dots, 1)$ et $F = \{aX + b, (a, b) \in \mathbf{R}^2\} = \text{Vect}(X, \mathbb{1})$.

Nous avons interprété le problème initial comme une distance minimale à une partie, qui elle-même s'exprime en fonction de la projection orthogonale de Y sur F :

$$\inf_{(a,b) \in \mathbf{R}^2} \sum_{i=1}^n (y_i - ax_i - b)^2 = \|Y - p_F(Y)\|^2.$$

Pour terminer, il reste à trouver $Y' = p_F(Y)$. En utilisant la définition d'une projection, on résout alors :

$$\begin{aligned} \begin{cases} Y' \in F \\ Y' - Z \in F^\perp \end{cases} &\iff \begin{cases} Y' = aX + b, & a, b \in \mathbf{R} \\ Y - Y' \perp (1, \dots, 1) \\ Y - Y' \perp X \end{cases} \\ &\iff \begin{cases} Y' = aX + b, & a, b \in \mathbf{R} \\ Y - aX - (1, \dots, 1) \perp (1, \dots, 1) \\ Y - aX - (1, \dots, 1) \perp X. \end{cases} \end{aligned}$$

En conclusion : $\left(\frac{C_{x,y}}{\sigma_x^2}, \bar{y} - \frac{C_{x,y}}{\sigma_x^2} \bar{x}\right)$ est l'unique minimum (global) de F sur \mathbf{R}^2 .

On aboutit alors au système sur $(a, b) \in \mathbf{R}^2$:

$$\begin{cases} \sum_{i=1}^n (y_i - ax_i - b) = 0 & \left(\iff \frac{\partial F(a,b)}{\partial b} = 0\right) \\ \sum_{i=1}^n (y_i - ax_i - b)x_i = 0 & \left(\iff \frac{\partial F(a,b)}{\partial a} = 0\right), \end{cases}$$

qui a pour unique solution le couple (a^*, b^*) trouvée précédemment (en cherchant les points critiques). Seulement, la théorie sur la projection orthogonale déroulée plus haut garantit que la solution obtenue cette fois-ci est un minimum : inutile donc de chercher à l'établir. On récupère de suite l'unicité par unicité de la projection orthogonale sur un sous-espace vectoriel de dimension finie.

Remarque 2.20 — À propos du vocabulaire Pourquoi parle-t-on de «régression linéaire»? La réponse est une erreur de traduction. Le mathématicien anglais Sir GALTON étudiait les tailles des fils (y_j) en fonction de la taille de leur père (x_j); et a constaté un «retour à la moyenne». En effet, les grands individus ont en moyenne des enfants plus petits qu'eux et les petits individus ont des enfants plus grand qu'eux. En Anglais le terme «retour à la moyenne» est «régression to the mean», ce terme a ensuite été mal transposé au Français.

QUALITÉ D'UNE RÉGRESSION LINÉAIRE. Comment évaluer la «justesse» d'un ajustement? Pour y répondre on définit un nouvel indicateur statistique : le coefficient de détermination, plutôt que le seul coefficient de corrélation.

Définition/Proposition ALEA.17.2 | Coefficient de détermination d'une régression

Soit (x, y) une série statistique double constituée d'une suite de couples $((x_k, y_k))_{1 \leq k \leq n}$. On appelle *coefficient de détermination de x et y* , noté $r^2(x, y)$, la quantité définie par :

$$r^2(x, y) = \rho(x, y)^2 = \frac{C_{x,y}^2}{V_x V_y} \in [0, 1].$$

⊗ Attention


Ce n'est donc pas le coefficient de corrélation, mais son carré.

Preuve Puisque $\rho(x, y) \in [-1, 1]$, son carré est bien dans $[0, 1]$.

Remarque 2.21 — Interprétation Ainsi $r^2(x, y) = 1$ correspond à une adéquation parfaite tandis que $r^2(x, y)$ proche de 0, équivalent à $\rho(x, y)$ proche de 0, indique une faible liaison linéaire ce qui peut signifier qu'il n'y a pas de lien entre x et y ou bien que x et y sont liés par une relation non-affine. En général, on considère une régression linéaire comme «satisfaisante» lorsque

$$r^2(x, y) \geq 0.9.$$

AUTRES AJUSTEMENTS SE RAMENANT À UNE RÉGRESSION LINÉAIRE. On peut penser à beaucoup d'ajustements. Par exemple :

1. si l'on souhaite tester la relation $y = \lambda e^{\alpha x}$, avec $(\alpha, \lambda) \in \mathbf{R} \times \mathbf{R}^{+*}$ avec y série statistique strictement positive, on peut constater qu'elle est équivalente à $(\ln y = \ln \lambda + \alpha x)$: on fait alors la régression sur $(x, \ln y)$. Comment retrouver α, λ à partir de a^*, b^* ? 
2. si l'on souhaite tester la relation $y = a \ln x + b$, avec $(a, b) \in \mathbf{R}$, on peut faire alors la régression sur $(\ln x, y)$.

2.2.5. En Python : calcul des grandeurs bivariées

On suppose dans la suite que toutes les données d'une série à nombre de modalités fini sont contenues dans une liste L donnée en paramètre.

Covariance

```
def covariance(L, M):
    """
    Renvoie la covariance des deux séries
    """
    Prod = [L[i]*M[i] for i in range(len(M))]
    return esperance(Prod) - esperance(L)*esperance(M)
```

Coefficient de corrélation

```
import math as ma
def coeff_cor(L, M):
    """
    Renvoie le coefficient de corrélation des deux séries
    """
    return covariance(L,
        ↪ M)/(ma.sqrt(variance(L))*ma.sqrt(variance(M)))
```

3.

STATISTIQUES INFÉRENTIELLES

Rappelons que l'inférence statistique (*cf.* introduction) consiste à savoir si une série statistique x peut être vue comme plusieurs réalisations d'une même variable aléatoire, et estimer les valeurs des paramètres de cette loi (espérance, variance, *etc.*). Consulter l'introduction pour plus de détails.

3.1. Estimation ponctuelle

Dans cette première sous-section, l'inférence va se situer dans l'aspect suivant : on essaie d'estimer les paramètres d'une loi à l'aide de réalisations X_1, \dots, X_n (des variables

aléatoires formant ce qu'on appellera un *échantillon* dans la suite) de cette loi.

Définition ALEA.17.22 | n -échantillon et estimateur

1. Un n -échantillon est un n -uplet (X_1, \dots, X_n) de variables aléatoires indépendantes et de même loi (i.i.d.). Une *observation* (ou *réalisation*) de (X_1, \dots, X_n) est $(X_1(\omega), \dots, X_n(\omega))$ pour un certain $\omega \in \Omega$. La loi commune s'appelle la *loi parente* ou la *loi mère*. On parle d'*échantillon gaussien* lorsque tous les $X_i, i \in \llbracket 1, n \rrbracket$ suivent une loi normale.
2. Un *estimateur d'un paramètre* θ inconnu est une suite de variables aléatoires $(\widehat{\theta}_n)$, où $\widehat{\theta}_n$ est une fonction de (X_1, \dots, X_n) , i.e. $\widehat{\theta}_n = \varphi_n(X_1, \dots, X_n)$ avec $\varphi_n : \mathbf{R}^n \rightarrow \mathbf{R}$.


Le plus souvent le paramètre θ sera l'espérance ou la variance de la loi de X , une médiane, ou toute autre valeur caractéristique d'une loi de probabilité.

⊗ Attention à la terminologie « estimateur »

On s'attend à ce qu'un estimateur « estime » θ au sens propre du terme *i.e.* qu'il en soit très proche. La définition précédente ne garantit pas cela : les propriétés vraiment intéressantes d'un estimateur seront énoncées *infra*.

⚙️ Cadre

Dans toute la suite de cette section, on se fixe un n -échantillon (X_1, \dots, X_n) de même loi qu'une variable aléatoire réelle X . On suppose de plus que cette loi commune dépend d'un paramètre $\theta \in \mathbf{R}$.

Exemple 9 – Uniforme en zéro Par exemple, (X_1, \dots, X_n) et X où X_1, \dots, X_n sont i.i.d. de loi $\mathcal{U}[0, \theta]$, et X de même loi indépendante des X_i . C'est un n -échantillon, et $\widehat{\theta}_n = X_1 + \dots + X_n$ est un estimateur de θ . 

Exemple 10 — Élection présidentielle À l'approche du second tour d'une élection présidentielle, on interroge une personne au hasard et on note $X = 1$ si elle se prononce pour le candidat A et $X = 0$ si c'est pour le candidat B. Alors X suit une loi de BERNOULLI de paramètre $\theta \in [0, 1]$ inconnu¹⁰ qui correspond à la proportion de français qui votent pour A.

On questionne cinq individus sur leurs intentions de vote et on obtient les résultats suivants (en notant 1 ou 0 selon que le choix se porte sur le candidat A ou B) : 1, 0, 0, 1, 0. Ces résultats observés correspondent, pour tout $\theta \in [0, 1]$, à la réalisation d'un 5 - échantillon (X_1, \dots, X_5) de loi mère $\mathcal{B}(\theta)$.

¹⁰Sauf bien entendu si on se dit capable de poser la question à l'ensemble de la population, ce qui n'est pas réalisable pour un institut de sondage

Définition ALEA.17.23 | Qualité d'un estimateur

Soit $(\widehat{\theta}_n)$ un estimateur de θ .

1. On appelle *erreur d'estimation* la **variable aléatoire** $\widehat{\theta}_n - \theta$.
2. Supposons que $\widehat{\theta}_n$ admet un moment d'ordre un. On appelle *biais*, l'espérance de l'erreur d'estimation :

$$b(\theta_n) = \mathbf{E}(\widehat{\theta}_n - \theta) = \mathbf{E}(\widehat{\theta}_n) - \theta.¹¹$$

On dit qu'un estimateur est :

- ▶ *sans biais* (ou qu'il estime θ de manière non biaisée) si $b(\theta_n) = 0$ pour tout $n \in \mathbf{N}$, il est dit *biaisé* dans le cas contraire.
 - ▶ Il est dit *asymptotiquement sans biais* si $b(\theta_n) \xrightarrow{n \rightarrow \infty} 0$.
3. Supposons que $\widehat{\theta}_n$ admet un moment d'ordre deux. On appelle *risque quadratique* la quantité :

$$r_n(\theta) = \mathbf{E}((\widehat{\theta}_n - \theta)^2).$$

On dit qu'un estimateur est *plus efficace qu'un autre* si son risque quadratique est moins élevé.

4. Un estimateur $(\widehat{\theta}_n)$ est dit *convergeant vers θ* si :

$$\forall \varepsilon > 0, \quad \mathbf{P}(|\widehat{\theta}_n - \theta| \geq \varepsilon) \xrightarrow{n \rightarrow \infty} 0,$$

i.e. s'il converge en probabilité vers le paramètre θ .

Proposition ALEA.17.14 | Décomposition biais/variance

Soit $(\widehat{\theta}_n)$ un estimateur du paramètre θ admettant un moment d'ordre deux. Alors pour tout entier $n \in \mathbf{N}$:

$$r_n(\theta) = \mathbf{Var}(\widehat{\theta}_n) + (b(\theta_n))^2.$$

¹¹par linéarité de l'espérance

¹²Donc plus on ajoute d'observation, plus la probabilité de s'écarter de θ est faible

Au même titre que la formule de KÖNIG-HUYGENS s'utilise le plus souvent pour calculer une variance lorsque l'on a déjà calculé une espérance, la décomposition biais/variance s'utilise pour calculer un risque lorsque l'on a déjà calculé le biais.

Preuve*(Point clef — Développer le carré)*

Proposition ALEA.17.15 | Condition suffisante de convergence

Soit $(\widehat{\theta}_n)$ un estimateur de θ admettant un moment d'ordre deux, et tel que : $\lim_{n \rightarrow \infty} r_n(\theta) = 0$. Alors :

$(\widehat{\theta}_n)_{n \in \mathbf{N}}$ est convergent vers $\theta \in \mathbf{R}$.

Preuve*(Point clef — On applique l'inégalité de Markov pour les moments d'ordre deux)*

ÉTUDE DES ESTIMATEURS DE LA MOYENNE/VARIANCE/ÉCART-TYPE EMPIRIQUES. On va essentiellement s'intéresser dans la suite aux estimateurs de l'espérance et de la variance, appelés moyenne et variance empirique. Rappelons l'expression de ces estimateurs vue dans le ??.

Définition ALEA.17.24 | Moyenne/Variance empirique

Soient X_1, \dots, X_n une famille de $n \in \mathbf{N}^*$ variables aléatoires réelles. On appelle *moyenne empirique des X_i* (resp. *variance empirique*) les variables aléatoires réelles

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \quad (\text{resp.} \quad \sigma_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \overline{X}_n)^2).$$

On appelle *écart-type empirique* la variable aléatoire

$$\sigma_n \stackrel{\text{(défi.)}}{=} \sqrt{\sigma_n^2}.$$

Nous avons également établi une version KÖNIG-HUYGENS en développant le carré *via* une identité remarquable.

Proposition ALEA.17.16 | Version KÖNIG-HUYGENS

Soient X_1, \dots, X_n une famille de $n \in \mathbf{N}^*$ variables aléatoires réelles. Alors :

$$\sigma_n^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \overline{X}_n^2.$$

Corollaire ALEA.17.1 | Lien entre moyenne/variance empirique pour des BERNOULLI

Soient X_1, \dots, X_n une famille de $n \in \mathbf{N}^*$ variables aléatoires réelles de loi $\mathcal{B}(p)$ avec $p \in [0, 1]$ et $n \in \mathbf{N}^*$. Alors

$$\sigma_n^2 = \overline{X_n} - \overline{X_n}^2.$$

Il n'y a en règle général pas de lien entre espérance et variance empirique (ce qui semble logique). C'est ici un cas très particulier qui provient du fait suivant : la somme des carrés des X_i est égale à la somme, puisqu'une BERNOULLI est à valeurs dans $\{0, 1\}$.

Preuve



Passons maintenant aux qualités de ces estimateurs.

Proposition ALEA.17.17 | Qualité de la moyenne/variance empirique

On suppose que X_1, \dots, X_n admettent une espérance μ et une variance $\sigma^2 > 0$, $n \in \mathbf{N}^*$.

1. (Biais de la moyenne empirique)

$$\mathbf{E}(\overline{X_n}) = \mu, \quad \text{et} \quad \mathbf{Var}(\overline{X_n}) = \frac{\sigma^2}{n}.$$

En particulier, la moyenne empirique $(\overline{X_n})$ est un estimateur sans biais de l'espérance.

2. (Biais de la variance empirique)

$$\mathbf{E}(\sigma_n^2) = \frac{n-1}{n} \sigma^2, \quad \text{et donc :} \quad \mathbf{E}\left(\frac{n}{n-1} \sigma_n^2\right) = \sigma^2.$$

En particulier, la variance empirique (σ_n^2) est un estimateur biaisé de la variance.

Preuve Il s'agit ici de calculer les biais de $\overline{X_n}$ et σ_n^2 .

1. Soit $n \in \mathbf{N}^*$. Nous avons d'une part par linéarité de l'espérance :

$$\mathbf{E}(\overline{X_n}) = \frac{1}{n} \sum_{i=1}^n \mathbf{E}(X_i) = \frac{1}{n} (n \mu) = \mu,$$

et d'autre part par argument d'indépendance des X_i :

$$\mathbf{Var}(\overline{X_n}) = \frac{1}{n^2} \sum_{i=1}^n \sigma^2 = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}.$$

2. Nous allons utiliser la version KÖNIG-HUYGENS de l'estimateur σ_n^2 , i.e.

$$\sigma_n^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \overline{X_n}^2.$$



$$\begin{aligned} \mathbf{E}(\sigma_n^2) &= \mathbf{E}\left(\frac{1}{n} \sum_{i=1}^n X_i^2 - \overline{X_n}^2\right), \\ &= \frac{1}{n} \sum_{i=1}^n \mathbf{E}(X_i^2) - \mathbf{E}(\overline{X_n}^2), \\ &= \frac{1}{n} \sum_{i=1}^n (\sigma^2 + \mu^2) - (\mathbf{Var}(\overline{X_n}) + \mathbf{E}(\overline{X_n})^2), \\ &= \sigma^2 + \mu^2 - \left(\frac{\sigma^2}{n} + \mu^2\right) \\ &= \frac{n-1}{n} \sigma^2. \end{aligned}$$

linéarité de l'espérance
lien moment d'ordre deux et variance
d'après 1

Nous déduisons, par linéarité de l'espérance une version dite « corrigée » de l'estimateur de la variance, *i.e.* un estimateur sans biais appelé *variance corrigée*.

Définition ALEA.17.25 | Variance empirique corrigée

On appelle *variance empirique corrigée* des X_i (*resp.* *écart-type corrigé*) les variables aléatoires définies pour tout $n \geq 2$ par :

$$\sigma_n^{2,\text{cor}} = \frac{n}{n-1} \sigma_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2 \quad (\text{resp.} \quad \sigma_n^{\text{cor}} = \sqrt{\sigma_n^{2,\text{cor}}}).$$

Proposition ALEA.17.18 | Qualité de la variance corrigée

On suppose que X_1, \dots, X_n admettent une espérance μ et une variance $\sigma^2 > 0$. Alors pour tout $n \geq 2$:

$$\mathbf{E}(\sigma_n^{2,\text{cor}}) = \sigma^2.$$

Ainsi, la variance empirique corrigée $\sigma_n^{2,\text{cor}}$ est un estimateur sans biais de la variance.

Preuve



Proposition ALEA.17.19 | Convergence

On suppose que X_1, \dots, X_n admettent une espérance μ et une variance $\sigma^2 > 0$, $n \in \mathbf{N}^*$. Alors :

L'estimateur $(\bar{X}_n)_{n \in \mathbf{N}^*}$ (*resp.* $(\sigma_n^2)_{n \geq 2}$) converge vers μ (*resp.* σ^2).

Preuve


(Point clef — Les risques quadratiques convergent vers zéro)



Nous admettons la convergence de σ_n^2 .


Exemple 11 — Reprenons l'Exemple 10, on note (X_1, \dots, X_n) un $n \in \mathbf{N}^*$ -échantillon associé. Un estimateur naturel pour θ est \bar{X}_n . On peut envisager d'autres estimateurs, par exemple

$$A_n = \frac{2}{n(n+1)} \sum_{k=1}^n kX_k, \quad \text{ou encore} \quad B_n = p.$$

Lequel des trois est le meilleur du point de vue du risque quadratique? On commencera par calculer leur biais, puis leur risque quadratique. Commentez les résultats obtenus. 

1. Qu'estime de manière non biaisée $2\bar{X}_n$? 


2. Montrer que $\widehat{\theta}_n$ est une variable aléatoire à densité pour tout $n \geq 1$. 

3. L'estimateur $(\widehat{\theta}_n)$ est-il biaisé? Donner alors un estimateur non biaisé de θ . 

Notez que des observations de ces trois estimateurs ont été donnés dans l'énoncé :

- ▶ $\bar{X}_5 = \frac{1}{5}(1 + 0 + 0 + 1 + 0) = \frac{2}{5}$,
- ▶ $A_5 = \frac{1}{3}$,
- ▶ $B_5 = 0$.

Exemple 12 — Uniforme en zéro On prolonge l'Exemple 9. Si (X_1, \dots, X_n) est un $n \in \mathbf{N}^*$ -échantillon de loi mère $\mathcal{U}[0, \theta]$ avec $\theta \in \mathbf{R}$. On note $\widehat{\theta}_n = \max(X_1, \dots, X_n)$.

4. Entre l'estimateur précédent et $2\overline{X}_n$, lequel choisir? 

3.2. Estimation par Intervalle de confiance

Une fois le calcul d'un estimateur effectué, on ne peut pas se contenter d'une valeur estimée : il faut mesurer l'erreur commise entre la valeur inconnue et l'estimation. En effet, même avec un risque quadratique faible, on n'est jamais à l'abri de tomber sur un « mauvais » échantillon qui nous donnerait une mauvaise estimation du paramètre.

Cette estimation d'erreur est précisément la vocation de l'estimation par intervalle de confiance qui est plus précise que la seule donnée d'un estimateur : nous allons don-

ner des intervalles, contenant le paramètre à estimer, avec très forte probabilité.¹³

Définition ALEA.17.26 | Intervalle de confiance

Soit $\alpha \in]0, 1[$ et (X_1, \dots, X_n) un échantillon de variables aléatoires indépendantes et de même loi que X , loi dépendant d'un paramètre θ .

1. On appelle *intervalle de confiance de niveau $1 - \alpha$* (on dit aussi de *probabilité de confiance $1 - \alpha$*) pour le paramètre θ , tout intervalle aléatoire noté I_{X_1, \dots, X_n} dépendant des X_1, \dots, X_n tel que :

$$\mathbf{P}(\theta \in I_{X_1, \dots, X_n}) \geq 1 - \alpha.$$

2. On appelle *intervalle de confiance asymptotique de niveau $1 - \alpha$ pour le paramètre θ* (on dit aussi de *probabilité de confiance $1 - \alpha$*) toute suite $(I_{X_1, \dots, X_n})_n$ d'intervalles aléatoires telle que :

$$\lim_{n \rightarrow \infty} \mathbf{P}(\theta \in I_{X_1, \dots, X_n}) \geq 1 - \alpha.$$

Remarque 3.1 —

- ▶ Très souvent, on recherche un intervalle de confiance de θ sous la forme d'un intervalle centré en une estimation ponctuelle de θ . Par exemple, pour l'intervalle de confiance de la moyenne μ d'un échantillon, il sera centré en \overline{X}_n le plus souvent.
- ▶ La plupart du temps, c'est ce niveau de risque de 0.05 qui est utilisé, et qui est communément accepté par exemple en sciences humaines. Mais dans des domaines plus sensibles où l'on n'a pas vraiment de droit à l'erreur (aérospatiale, physique nucléaire, etc), on travaille avec des niveaux de risque de 0.01, voir moins.

¹³En passant au complémentaire, la probabilité que le paramètre soit en dehors de cet intervalle sera donc très petite.

Proposition ALEA.17.20 | Stabilité par élargissement


Soit $\alpha \in]0, 1[$ et (X_1, \dots, X_n) un échantillon de variables aléatoires indépendantes et de même loi que X , loi dépendant d'un paramètre θ . Alors :

si I_{X_1, \dots, X_n} est un intervalle de confiance de niveau $1 - \alpha$ pour θ , alors pour tout intervalle $J_{X_1, \dots, X_n} \supset I_{X_1, \dots, X_n}$, J_{X_1, \dots, X_n} est encore un intervalle de confiance de niveau $1 - \alpha$ pour θ .

Preuve



On commence par deux exemples de recherche d'un intervalle de confiance non-asymptotique.

Exemple 13 — Pesons-nous avec CRUELLA. Intervalle de confiance non asymptotique obtenu par propriétés des lois normales. Pamela est un mannequin célèbre dont le poids est strictement surveillé par Cruella. Cette charmante dame a investi un jour dans l'achat d'une balance Harmonia afin de connaître précisément le poids de sa protégée. Horreur : elle a constaté sur l'emballage de la balance que les fabricants (d'honnêtes artisans suisses) admettaient que leur outil de mesure (nul n'est parfait) pouvait commettre des erreurs de mesure dont l'écart-type valait 0,1 kg, néanmoins l'étiquette précise que les mesures (X_1, \dots, X_n) avec $n \in \mathbf{N}^*$ sont gaussiennes. En effet les pièces détachées ne sont pas toutes exactement identiques, leur montage n'est jamais parfait et le transport à travers les Alpes endommage parfois les balances. Ne faisant ni une ni deux, Cruella, a, dès le lendemain, dévalisé le magasin en investissant dans l'achat de 99 nouvelles balances Harmonia et a forcé Pamela à sauter sur les 100 balances pendant que Cruella relevait scrupuleusement les 100 mesures. Résultat moyen des pesées : 55,4 kg. Donner à Cruella un intervalle de confiance pour le poids moyen de Pamela sur ce type de balance, de probabilité de confiance 0,95. 

Remarque 3.2 — Nous n'avons pas eu besoin d'utiliser le théorème central limite ici, car l'échantillon de départ était déjà gaussien. Sinon, de manière générale, on utilisera la **Proposition ALEA.17.21** ci-après.

Exemple 14 — Intervalle de confiance non asymptotique obtenu via la loi faible des grands nombres On considère une pièce dont on souhaite savoir si elle est truquée ou non. Pour cela, on peut la lancer autant de fois que l'on veut. Mathématiquement, étant donné $n \in \mathbf{N}^*$, on observe la réalisation d'un échantillon (X_1, \dots, X_n) de variables aléatoires indépendantes et de même loi $\mathcal{B}(p)$, où $p \in]0; 1[$ est **connu**.

On cherche alors un intervalle $[a, b]$, dont les bornes dépendent des observations, mais pas de p , et tel que la probabilité que le paramètre inconnu p appartienne à cet intervalle soit égale à 0,95. On note $\widehat{p}_n = \frac{1}{n} \sum_{i=1}^n X_i$.

1. À l'aide de l'inégalité de BIENAYMÉ-TCHEBYCHEV, justifier que pour tout $\varepsilon > 0$:

$$\mathbf{P}\left(|\widehat{p}_n - p| > \varepsilon\right) \leq \frac{1}{4n\varepsilon^2}.$$

2. Déterminer une valeur de ε telle que $\mathbf{P}\left(|\widehat{p}_n - p| > \varepsilon\right) \leq \alpha$.

3. En déduire que $\mathbf{P}\left(|\widehat{p}_n - p| \leq \frac{1}{2\sqrt{\alpha}\sqrt{n}}\right) \geq 1 - \alpha$ et donner un intervalle de confiance de niveau $\alpha = 0,05$.

4. On effectue 2000 lancers de la pièce et on trouve $\widehat{p}_n = 0,57$. Que peut-on conclure?



INTERVALLE DE CONFIANCE ASYMPTOTIQUE OBTENU PAR LE THÉORÈME CENTRAL LIMITE On cherche à présent un intervalle de confiance asymptotique pour μ à l'aide du théorème central limite. Puisque (X_n) est une suite de variables aléatoires i.i.d. possédant une variance, on peut lui appliquer le théorème central limite : pour tout $(a, b) \in \mathbf{R}^2$,

$$\lim_{n \rightarrow +\infty} \mathbf{P} \left(a \leq \sqrt{n} \frac{\overline{X}_n - \mu}{\sigma_n} \leq b \right) = \Phi(b) - \Phi(a).$$

On souhaite obtenir un intervalle de confiance asymptotique au niveau de confiance $1 - \alpha$ avec $\alpha \in]0, 1[$. Pour cela, on doit donc choisir a, b de sorte que

$$\Phi(b) - \Phi(a) \geq 1 - \alpha.$$

Il y a une infinité de façon de choisir a et b . On choisit couramment $a = -b$ (intervalle symétrique), et donc

$$\Phi(b) - \Phi(a) = \Phi(b) - \Phi(-b) = \Phi(b) - (1 - \Phi(b)) = 2\Phi(b) - 1.$$

On cherche donc b de sorte que :

$$2\Phi(b) - 1 = 1 - \alpha \iff \Phi(b) = 1 - \frac{\alpha}{2} \iff b = \Phi^{-1} \left(1 - \frac{\alpha}{2} \right),$$

car, rappelons-le, nous avons montré que $\Phi : \mathbf{R} \rightarrow [0, 1]$ est bijective. Avec ce choix, nous donc montré que :

$$\begin{aligned} \lim_{n \rightarrow +\infty} \mathbf{P} \left(-\Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \leq \sqrt{n} \frac{\overline{X}_n - \mu}{\sigma_n} \leq \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \right) &= 1 - \alpha, \\ \lim_{n \rightarrow +\infty} \mathbf{P} \left(-\frac{\sigma_n}{\sqrt{n}} \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \leq \mu - \overline{X}_n \leq \frac{\sigma_n}{\sqrt{n}} \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \right) &= 1 - \alpha. \end{aligned} \quad \left. \begin{array}{l} \\ \end{array} \right\} \text{manipulation d'encadrement}$$

Il reste ensuite à ajouter \overline{X}_n de chaque côté de l'encadrement, on obtient alors la proposition suivante.

Proposition ALEA.17.21 | IC asymptotique pour la moyenne donné par le théorème central limite

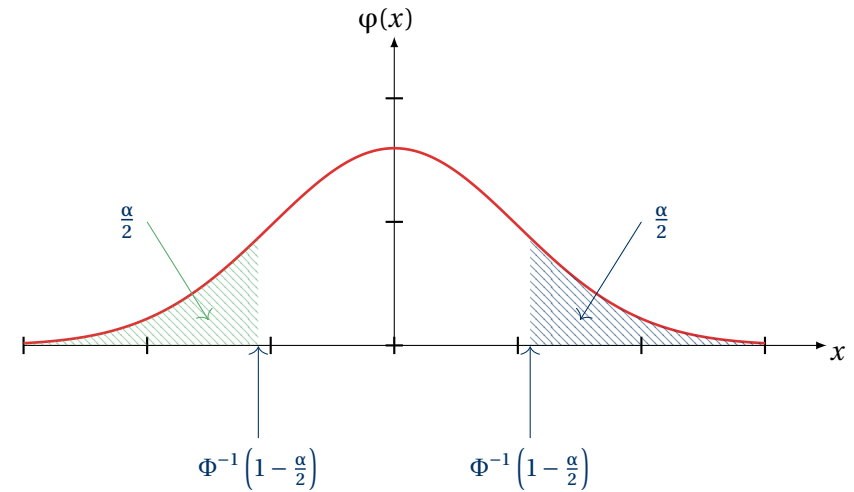
Soit (X_n) une suite de variables aléatoires indépendantes et de même loi, on note μ leur espérance commune, et admettant une variance et soient $\alpha \in]0, 1[$, $n \in \mathbf{N}^*$. Alors :

$$\mathbf{P}\left(\bar{X}_n - \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \frac{\sigma_n}{\sqrt{n}} \leq \mu \leq \bar{X}_n + \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \frac{\sigma_n}{\sqrt{n}}\right) \xrightarrow{n \rightarrow \infty} 1 - \alpha.$$

Ainsi,

$$\left[\bar{X}_n - \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \frac{\sigma_n}{\sqrt{n}} \leq \mu \leq \bar{X}_n + \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \frac{\sigma_n}{\sqrt{n}}\right]$$

est un intervalle de confiance asymptotique de niveau α pour l'espérance μ .





Attention Il faut savoir refaire la démarche qui précède l'énoncé de cette proposition.

Remarque 3.3 – Visualisation de $\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$ Représentons sur la densité Gaussienne la quantité $\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$.

Pour connaître les valeurs de $\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$ on utilisera la table de la loi normale (voir le tableau en fin de ce chapitre). On peut cependant garder à l'esprit les valeurs remarquables suivantes que l'on utilisera la plupart du temps :

- ▶ $\Phi^{-1}\left(1 - \frac{0.05}{2}\right) = 1.96$ donc pour un risque $\alpha = 0.05$,
- ▶ $\Phi^{-1}\left(1 - \frac{0.01}{2}\right) = 2.57$ donc pour un risque $\alpha = 0.01$.

Remarque 3.4 – Commentaires

1. Comment varie la taille de l'intervalle de confiance en fonction de n ? 
2. Comment simplifier cet intervalle de confiance lorsque l'écart-type σ est supposé connu? 

3. Pourquoi utiliser plutôt la seconde forme du théorème central limite?

Résumons les techniques pour obtenir un intervalle de confiance.

Méthode Résultats probabilistes pour établir un intervalle de confiance : le théorème central limite et l'inégalité de BIENAYMÉ-TCHEBYCHEV



1. (Si le « type de loi » (bernoulli, gaussienne etc.) du n -échantillon est connu)


On arrive parfois à calculer les probabilités $\mathbf{P}(\theta \in I_{X_1, \dots, X_n})$ explicitement pour n'importe quel intervalle I_{X_1, \dots, X_n} , les intervalles de confiance obtenus ne sont alors **pas asymptotiques**.

Pour des échantillons gaussiens, on a deux cas de figure :

- ▶ si σ est connue, on centre/réduit la moyenne empirique et on utilise la propriété de stabilité de la loi normale (cf. exemple de CRUELLA).
- ▶ [H.P] Si σ est inconnue, on peut avoir recours à la loi de STUDENT¹⁴ : voir **Remarque 6** ci-après pour une définition.


2. (Si la loi de départ n'est pas connue)¹⁵ On utilise soit :

- ▶ le théorème central limite en centrant réduisant la moyenne empirique (en approchant la variance par la version empirique si elle n'est pas connue), cela nous donne un intervalle de confiance seulement asymptotique,
- ▶ soit l'inégalité de BIENAYMÉ-TCHEBYCHEV.

Exemple 15 — Afin d'étudier la proportion p d'élèves satisfaits par le nouveau carambar bi-goût, on en interroge 100, et 56 d'entre eux déclarent être satisfaits par ce nouveau modèle. Donner un intervalle de confiance à 95 % pour p . 

¹⁴C'est la loi obtenue en remplaçant σ inconnue par σ_n^{cor} dans la centrée/réduite de la moyenne empirique d'un échantillon gaussien

¹⁵On sera dans ce contexte l'immense majorité du temps

Exemple 16 — Sur 250 ampoules, on observe une durée de vie moyenne de 600 heures et un écart-type de 50 heures. Donner un intervalle de confiance à 99 % de l'espérance de vie d'une ampoule. 

Remarque 3.5 — **Différence avec les intervalles de « fluctuation ».** On se donne par exemple une variable aléatoire $X \leftrightarrow \mathcal{N}(m, \sigma^2)$. Alors en utilisant la propriété de stabilité de la loi normale, on a :

$$\mathbf{P}(m - 1,96\sigma \leq X \leq m + 1,96\sigma) = 0,95.$$

L'intervalle $[m - 1,96\sigma, m + 1,96\sigma]$ est appelé « intervalle de *fluctuation* » pour X , car X

prend 95 % de ses valeurs dans cet intervalle, dont les bornes sont des nombres réels (dépendant des paramètres de la loi de X). Au contraire, un intervalle de *confiance* contient une valeur réelle mais inconnue, et ce sont ses bornes qui sont des variables aléatoires. Nous n'estimons pas la même chose dans les deux cas, mais on passe de l'un à l'autre par de simples manipulations sur les encadrements.

3.3. Test de conformité à la moyenne

Comme nous venons de le voir, l'une des fonctions des statistiques est de proposer, à partir d'observations d'un phénomène aléatoire (ou modélisé comme tel) une estimation de la loi de ce phénomène (ou plus précisément des paramètres associés). C'est ce que nous avons fait en construisant des intervalles de confiance. Les statistiques servent aussi à prendre des décisions. Peut-on considérer qu'un médicament est plus efficace qu'un placebo? Le nombre de consultations de Google par seconde suit-il une loi de Poisson? Les gènes pilotant la couleur des yeux et celle des cheveux sont-ils sur les mêmes chromosomes? Il y a deux points communs (au moins) à toutes ces questions : leurs réponses sont des oui-non et le phénomène sous-jacent est aléatoire. Les tests statistiques vont permettre d'apporter une réponse à des questions manichéennes en contrôlant l'aléa inhérent à la situation.

En statistique les deux éventualités sont appelées des hypothèses et sont notées \mathcal{H}_0 (hypothèse nulle) et \mathcal{H}_1 (hypothèse alternative : « $\mu \neq \mu_0$ » dans l'exemple qui suit).

Commençons directement par le test de conformité à la moyenne avant de présenter le vocabulaire général des tests statistiques.

PRINCIPE DU TEST D'ADÉQUATION À LA MOYENNE. Le seul test qui est au programme de BCPST est le *test d'adéquation à la moyenne*. On considère un n -échantillon (X_1, \dots, X_n) possédant une variance $\sigma^2 > 0$, et une espérance μ . Notons \mathcal{H}_0 « $\mu = \mu_0$ » : c'est l'hypothèse que la moyenne commune des X_i est $\mu = \mu_0$ avec $\mu_0 \in \mathbf{R}$. Si \mathcal{H}_0 est vraie, alors nous avons vu le résultat suivant.

Proposition ALEA.17.22 | Application du théorème central limite pour trouver une zone de rejet

$$\text{Si } \mathcal{H}_0 \text{ est vraie, alors pour tout } \alpha \in \mathbf{R}, \quad \mathbf{P} \left(\left| \frac{\bar{X}_n - \mu_0}{\frac{\sigma_n}{\sqrt{n}}} \right| > \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \right) \xrightarrow{n \rightarrow \infty} \alpha.$$

Preuve



Ainsi, en se fondant sur ce résultat probabiliste, une stratégie de décision pour accepter ou rejeter l'hypothèse \mathcal{H}_0 serait la suivante.

Définition/Proposition ALEA.17.3 | Test d'adéquation à la moyenne

1. si pour $n \geq 30$,¹⁶ $\mu_0 \in \left[\bar{X}_n - \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \frac{\sigma_n}{\sqrt{n}}, \bar{X}_n + \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \frac{\sigma_n}{\sqrt{n}} \right]$, alors on ne rejette pas l'hypothèse \mathcal{H}_0 avec risque d'erreur¹⁷ α ,
2. si pour $n \geq 30$, $\mu_0 \notin \left[\bar{X}_n - \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \frac{\sigma_n}{\sqrt{n}}, \bar{X}_n + \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \frac{\sigma_n}{\sqrt{n}} \right]$, alors on rejette l'hypothèse.

On appelle ce test le *test d'adéquation à la moyenne*.

¹⁶Un consensus pour que la convergence dans le théorème central limite soit suffisamment précise

Fixons un peu de vocabulaire :

▶ La variable aléatoire $\widehat{\theta}_n = \frac{\overline{X}_n - \mu_0}{\frac{\sigma_n}{\sqrt{n}}}$ est appelée *statistique de test*.

▶ L'intervalle

$$\left] -\infty, -\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \right[\cup \left] \Phi^{-1}\left(1 - \frac{\alpha}{2}\right), +\infty \right[$$

est appelé *la zone de rejet* ; car c'est lorsque la statistique de test est dans cette zone de rejet que l'on rejette \mathcal{H}_0 .

▶ Le risque d'erreur (de première espèce) est la probabilité de rejeter \mathcal{H}_0 alors qu'elle est vraie, c'est ce risque que l'on souhaite le plus petit possible. Dans notre test d'adéquation, il s'agit d' α qui est petit.

▶ L'hypothèse \mathcal{H}_1 qui est ici $\mu \neq \mu_0$ est appelée *hypothèse alternative*. Dans nos exemples, l'hypothèse alternative \mathcal{H}_1 sera toujours $^c \mathcal{H}_0$ — l'hypothèse contraire de \mathcal{H}_0 .

Résumé du test d'adéquation

1. **(But)** Tester une valeur possible de moyenne \mathcal{H}_0 « $\mu = \mu_0$ ».
2. **(Résultat probabiliste qui fonde le test)** le théorème central limite.
3. **(Décision)** On rejette \mathcal{H}_0 si $\mu_0 \notin \left[\overline{X}_n - \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \frac{\sigma_n}{\sqrt{n}}, \overline{X}_n + \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \frac{\sigma_n}{\sqrt{n}} \right]$.
Sinon on ne rejète pas \mathcal{H}_0 .

Remarque 3.6 — Que faire quand $n < 30$? Une solution : la loi de STUDENT lorsque l'échantillon est gaussien Dans le cas où $n < 30$ ¹⁸ et que l'échantillon est gaussien, on peut utiliser une version du test précédent faisant appel à la *loi de STUDENT*.

Plus précisément, on appelle loi de STUDENT à $k \in \mathbf{N}^*$ degrés de liberté la loi d'une variable aléatoire \mathcal{T}_k définie par :

$$\mathcal{T}_k = \frac{N}{\sqrt{(N_1^2 + \dots + N_k^2) / k}}$$

¹⁷ i.e. rejeter à tort l'hypothèse \mathcal{H}_0 alors qu'elle était vraie

¹⁸ et même pour tout n dans ce cas

où (N_1, \dots, N_k) est un k -échantillon de loi $\mathcal{N}(0, 1)$, et N également de loi $\mathcal{N}(0, 1)$ indépendante des $N_i, i \in \llbracket 1, k \rrbracket$.

Soit donc (X_1, \dots, X_n) notre échantillon de moyenne commune μ . Alors la centrée-réduite de \overline{X}_n où l'écart-type σ est remplacé par sa version empirique corrigée est :

$$\sqrt{n} \frac{\overline{X}_n - \mu}{\sigma_n^{\text{cor}}}$$

Alors, un théorème non trivial¹⁹ permet de montrer que cette variable aléatoire suit une loi \mathcal{T}_{n-1} . Cette loi étant tabulée, on peut alors en déduire facilement des intervalles de confiance comme nous l'avions fait avec le théorème central limite.

GÉNÉRALITÉS SUR LES TESTS STATISTIQUES.

Définition ALEA.17.27 | Test statistique, hypothèse nulle, zone de rejet

1. Une *hypothèse statistique* est un énoncé concernant un n -échantillon (valeur d'un paramètre, nature de la distribution, etc.).
2. Un *test statistique* est une démarche ayant pour but de fournir une règle de décision permettant, en se fondant sur l'observation d'un échantillon, de faire un choix entre deux hypothèses statistiques. L'*hypothèse nulle* est l'hypothèse fixant *a priori* une condition sur le paramètre, on la note \mathcal{H}_0 . Toute autre hypothèse est appelée hypothèse alternative, on la note \mathcal{H}_1 .
3. On appelle *région de rejet* (resp. *statistique de test*) une partie I de \mathbf{R} (resp. un estimateur $(\widehat{\theta}_n)$ associé au n -échantillon considéré) qui conduit à rejeter \mathcal{H}_0 pour \mathcal{H}_1 si $\widehat{\theta}_n \in I$. Dans le cas contraire, on dit que l'on *ne rejète pas* \mathcal{H}_0 .
4. On appelle *risque de première espèce* (ou *niveau du test*) la probabilité de rejeter \mathcal{H}_0 à tort (alors qu'elle est vraie).

¹⁹ théorème de COCHRAN

Définition ALEA.17.28 | Test statistique et niveau

Un test statistique est un algorithme qui conduit à accepter \mathcal{H}_0 ou à rejeter \mathcal{H}_0 à partir d'observations d'un phénomène aléatoire. On appelle *niveau du test* la probabilité de rejeter \mathcal{H}_0 alors qu'elle est vraie.

C'est le niveau du test que l'on souhaite le plus faible possible. En revanche, on se préoccupe en général peu de la probabilité d'accepter \mathcal{H}_0 alors que \mathcal{H}_1 est vraie. L'objectif d'un test est avant tout de valider l'hypothèse \mathcal{H}_0 . Ne pas rejeter \mathcal{H}_0 veut simplement dire que les observations ne sont pas incompatibles avec cette hypothèse.


Remarque 3.7 — Dissymétrie des hypothèses Retenez l'analogie avec la justice qui pose comme principe la présomption d'innocence. On souhaite contrôler en priorité la probabilité d'envoyer un innocent en prison en négligeant pour l'instant celle de relâcher un coupable. Dans cet exemple \mathcal{H}_0 est «la personne est innocente» et \mathcal{H}_1 est «la personne est coupable». Le risque de première espèce correspond donc au rejet de \mathcal{H}_0 (personne envoyée en prison) alors qu'elle est vraie (personne innocente).

Plutôt que de dire «on ne rejette pas l'hypothèse», on devrait dire «on ne rejette pas l'hypothèse avec un risque α de se tromper» (*i.e.* d'accepter l'hypothèse alors qu'elle est fautive). De même, plutôt que de dire «on rejette l'hypothèse», on devrait dire «on rejette l'hypothèse avec un risque $1 - \alpha$ de se tromper» (*i.e.* de rejeter l'hypothèse alors qu'elle est vraie).

**Méthode Démarche générale d'un test statistique**

1. Poser l'hypothèse (nulle) \mathcal{H}_0 que l'on souhaite tester.
2. Trouver un résultat de probabilité qui donne deux résultats différents selon que \mathcal{H}_0 est vraie ou non (dans le test d'adéquation *supra*, c'était le théorème central limite).
3. Donner la stratégie de décision, en fonction du résultat énoncé.

Remarque 3.8 — Il est illusoire, à cause de l'aléatoire sous-jacent au n -échantillon, de vouloir prendre à coup sûr la bonne décision. C'est pourquoi on se laisse une marge d'erreur. En général, on choisit $\alpha = 0,05$ ou $\alpha = 0,01$.

Exemple 17 — Chez le petit lapin, la durée moyenne de gestation est de 30 jours. On étudie un échantillon de 66 familles de gros lapins, pour lesquelles on observe une durée moyenne de gestation de 30,83 jours avec un écart-type de 4,07 jours. Peut-on conclure que la durée de gestation est significativement différente chez les petits et les gros lapins? 

Exemple 18 — Une étude commerciale, réalisée sur 100 personnes, montre que 49 % des internautes ont moins de 39 ans, alors que d'après le recensement cette tranche d'âge représente 41 % de la population française. Peut-on conclure que la population des internautes est plus jeune que la population française? La différence observée est-elle révélatrice d'un phénomène ou provient-elle des fluctuations d'échantillonnage? On pourra considérer le cas d'un niveau de confiance de 90 % puis de 95 %.



Soit (X_1, \dots, X_{100}) un 100-échantillon, de loi $\mathcal{B}(p)$ avec $p \in [0, 1]$, c'est pour un individu donné la probabilité qu'il ait moins de 39 ans. On sait par hypothèse que $\overline{X}_{100} = \frac{49}{100}$, puis par formule de KÖNIG-HUYGENS que $\sigma_{100}^2 = \frac{1}{100} \sum_{i=1}^{100} X_i^2 - \overline{X}_{100}^2 = \overline{X}_{100} - \overline{X}_{100}^2 = 0.2499$. Ainsi :

— un intervalle de confiance de niveau 95 % pour p est :

$$\left[0,49 - 1,96 \times \frac{\sqrt{0.2499}}{\sqrt{100}}; 0,49 + 1,96 \times \frac{\sqrt{0.2499}}{\sqrt{100}} \right] \approx [0.392, 0.588],$$

— un intervalle de confiance de niveau 90 % pour p est :

$$\left[0,49 - 1.64 \times \frac{\sqrt{0.2499}}{\sqrt{100}}; 0,49 + 1.64 \times \frac{\sqrt{0.2499}}{\sqrt{100}} \right] \approx [0.408, 0.572].$$

On constate que : 0,41 est dans les deux intervalles de confiance. On accepte donc l'hypothèse \mathcal{H}_0 « $p = \frac{41}{100}$ », i.e. que parmi les internautes et l'ensemble de la population la fréquence de personnes de moins de 39 ans est sensiblement la même. La petite différence observée provient de fluctuations d'échantillonnage, avec une étude commerciale menée sur plus de personnes nous aurions peut-être eu un résultat différent.

ANNEXE : TABLES DE VALEURS POUR LA FONCTION DE RÉPARTITION DE LA LOI NORMALE CENTRÉE RÉDUITE $\mathcal{N}(0, 1)$.



Méthode Obtenir $\Phi(x)$ pour un certain $x \in \mathbf{R}$ à l'aide d'une table

Si l'on souhaite avoir, par exemple, $\Phi(0,96)$, on :

1. se place sur la ligne «0.9»,
2. se place ensuite sur la colonne «0.06».
3. On obtient alors la valeur désirée. Dans cet exemple, $\Phi(0,96) = 0,8315$.



Méthode Chercher $x \in \mathbf{R}$ tel que $\Phi(x) = \alpha$ à l'aide d'une table, $\alpha \in [0, 1]$

Si l'on souhaite avoir, par exemple, $x \in \mathbf{R}$ tel que $\Phi(x) = 0.975$.

1. On cherche dans la grille l'endroit où se trouve une valeur suffisamment proche de $\alpha = 0.975$.
2. Dans cet exemple, on constate que $\Phi(1.96) = 0.975$.

Si l'on souhaite avoir, par exemple, $x \in \mathbf{R}$ tel que $\Phi(x) = 0.160$.

1. En parcourant la table, on constate que 0.160 n'y apparaît.
2. On reformule alors la condition en passant au complémentaire :

$$1 - \Phi(x) = 0,84 = \Phi(-x).$$

3. On cherche donc dans la table 0.84, on trouve alors

$$0.84 = \Phi(1.00) \quad \text{donc} \quad -x = 1.00, \quad x = -1.00.$$

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.090
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224

0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

$$\Phi(x) = \mathbf{P}(X \leq x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \quad \text{et} \quad \Phi(-x) = 1 - \Phi(x).$$

On retiendra en particulier la valeur typique $\Phi(1.96) = 0.975$, de sorte que :

$$\mathbf{P}(X \in [-1.96; 1.96]) = 2\Phi(1.96) - 1 = 0.95$$

*** **Fin du chapitre** ***

4.

EXERCICES

4.1. Descriptives

Exercice ALEA.17.1 | Un médecin effectue des recherches sur l'efficacité d'un nouveau bêta-bloquant. Cette famille de médicaments est destinée à diminuer le rythme cardiaque des malades atteints de tachycardie (pouls supérieur à 100 battements par minute au repos). Il a donc séparé les malades en 2 groupes : le groupe A reçoit le traitement d'un nouveau médicament, le groupe B reçoit un placebo. Voici les résultats.

- ▶ A : 74 - 91 - 91 - 84 - 95 - 93 - 95 - 102 - 81 - 116 - 88 - 95,
- ▶ B : 94 - 95 - 113 - 95 - 104 - 113 - 94 - 144 - 105 - 153.

1. Calculer l'étendue et la médiane pour chacune de ces deux séries.
2. Construire le diagramme de TUCKER de ces deux séries.
3. L'effet du médicament semble-t-il satisfaisant ?

Exercice ALEA.17.2 | L'indice moyen d'un salaire a évolué de la façon suivante :

Année	1	2	3	4	5	6	7
Indice	165	176	193	202	222	245	253

1. Représenter cette série statistique par un nuage de points.
2. Déterminer la droite de régression linéaire de l'indice en fonction de l'année.
3. Prévoir l'indice à l'année 9.

Exercice ALEA.17.3 | **Ajustement d'ordre deux – Extrait Agro—Véto 2019** Dans le cas de la régression linéaire on cherche à approcher un nuage de points à l'aide d'une droite, donc un polynôme de degré 1. Dans cet exercice, on essaie de généraliser à un

polynôme de degré 2.

Soient $n \in \mathbf{N}^*$, et $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$ deux éléments de \mathbf{R}^n , on suppose que les x_i sont deux à deux distincts. Soit de plus

$$F \begin{cases} \mathbf{R}^3 & \longrightarrow & \mathbf{R}, \\ (a, b, c) & \longmapsto & \sum_{i=1}^n (y_i - a - bx_i - cx_i^2)^2. \end{cases}$$

1. Expliquer ce que représente géométriquement $F(a, b, c)$ pour tout $(a, b, c) \in \mathbf{R}^3$. On pourra placer sur le graphique un nuage de points $(x_i, y_i)_{1 \leq i \leq n}$ pour n petit.
2. Justifier que F admet des dérivées partielles dans toutes les directions, et calculer $\text{grad} F(a, b, c)$ pour tout $(a, b, c) \in \mathbf{R}^3$.

3. On note ici $Y = {}^T y$ et $\beta = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ avec $(a, b, c) \in \mathbf{R}^3$.

3.1) Chercher une matrice $T \in \mathfrak{M}_{n,3}(\mathbf{R})$ telle que :

$$\begin{cases} \frac{\partial F}{\partial a}(a, b, c) = -2 \left\langle Y - T\beta \left| \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \right. \right\rangle, \\ \frac{\partial F}{\partial b}(a, b, c) = -2 \left\langle Y - T\beta \left| \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right. \right\rangle, \\ \frac{\partial F}{\partial c}(a, b, c) = -2 \left\langle Y - T\beta \left| \begin{pmatrix} x_1^2 \\ \vdots \\ x_n^2 \end{pmatrix} \right. \right\rangle, \end{cases}$$

où $\langle \cdot | \cdot \rangle$ désigne le produit scalaire euclidien sur $\mathfrak{M}_{n,1}(\mathbf{R})$.

- 3.2)** En écrivant matriciellement les produits scalaires précédents, déduire que tout point critique $(\hat{a}, \hat{b}, \hat{c}) \in \mathbf{R}^3$ de F vérifie :

$$({}^T T T) \hat{\beta} = {}^T T Y, \quad \text{avec : } \hat{\beta} = \begin{pmatrix} \hat{a} \\ \hat{b} \\ \hat{c} \end{pmatrix}.$$

4. On souhaite ensuite trouver une condition pour que la matrice ${}^T T T$ soit inversible.

- 4.1) Préciser le format de ${}^T T T$.
 - 4.2) Montrer que $\ker(T) = \ker({}^T T T)$.
 - 4.3) Que vaut $\dim \ker({}^T T T) + \dim \text{Rg}({}^T T T)$? Justifier.
 - 4.4) Que vaut $\dim \ker(T) + \dim \text{Rg}(T)$? Justifier.
 - 4.5) En déduire que ${}^T T T$ est inversible si et seulement si T est de rang 3.
5. Dans ce cas-là, donner une expression de $\hat{\beta}$.

4.2. Estimateurs

Exercice ALEA.17.4 | Barycentre de deux estimateurs

1. Dans une population de porcs, on veut estimer le gain moyen quotidien (GMQ) noté μ , on suppose que les gains sont d'écart-type $\sigma \in \mathbf{R}$. À cet effet, on choisit deux échantillons indépendants dans cette population. On observe deux échantillons : l'un (X_1, \dots, X_{10}) de 10 individus, et (Y_1, \dots, Y_{30}) de 30 individus. On propose deux estimateurs de μ :

$$T_1 = \frac{\overline{X}_{10} + \overline{Y}_{30}}{2} \text{ et } T_2 = \frac{10\overline{X}_{10} + 30\overline{Y}_{30}}{40}.$$

On cherche à déterminer le meilleur des deux estimateurs.

- 1.1) Calculer le biais de chaque estimateur pour le paramètre μ . Cela permet-il de les départager?
 - 1.2) Calculer les variances de T_1 et T_2 en fonction de la variance σ^2 du gain quotidien de la population. Conclure.
 - 1.3) **(Généralisation)** Soient T_1 et T_2 deux estimateurs de $\mu \in \mathbf{R}$, sans biais et indépendants. Pour tout $a \in \mathbf{R}$, on pose $\Theta_a = aT_1 + (1-a)T_2$.
 - 1.4) Soit $a \in \mathbf{R}$. Calculer le biais de Θ_a pour le paramètre μ .
2. Parmi tous les Θ_a , $a \in \mathbf{R}$, lequel a le plus petit risque quadratique? Est-ce cohérent avec la première question?

Exercice ALEA.17.5 | Soit (X_1, \dots, X_n) un n -échantillon de loi de BERNOULLI de para-

mètre $p \in]0, 1[$, $n \geq 1$. On pose $S_n = \sum_{k=1}^n X_k$ ainsi que

$$\overline{X}_n = \frac{S_n}{n}, \quad T_n = \frac{S_n + 1}{n + 2}.$$

Comparer les biais, ainsi que les risques quadratiques de \overline{X}_n et T_n en tant qu'estimateurs de p . Peut-on privilégier l'un de ces estimateurs par rapport à l'autre?

Exercice ALEA.17.6 | Soit (X_1, \dots, X_n) un n -échantillon, $n \geq 2$, d'une loi $\mathcal{P}(\lambda)$ de paramètre $\lambda > 0$ inconnu. On souhaite estimer le paramètre $\theta = e^{-\lambda}$. Pour $k \in [1, n]$, on pose $Y_k = \mathbb{1}_{\{X_k=0\}}$, et on introduit $\overline{Y}_n = \frac{1}{n} \sum_{k=1}^n Y_k$, $S_n = \sum_{k=1}^n X_k$.

- 1. 1.1) Montrer que \overline{Y}_n est un estimateur sans biais de θ .
- 1.2) Calculer $\text{Var}(\overline{Y}_n)$. Que dire de $\mathbf{P}(|\overline{Y}_n - \theta| \geq \varepsilon)$ pour tout $\varepsilon > 0$? On dit que \overline{Y}_n est un estimateur convergent vers θ .
- 2. Pour $j \in \mathbf{N}$, calculer $\varphi(j) = \mathbf{P}(X_1 = 0 | S_n = j)$.
- 3. 3.1) Montrer que $T_n = \varphi(S_n)$ est un estimateur sans biais de θ .
- 3.2) Calculer $\text{Var}(T_n)$ et en déduire que T_n converge vers θ .
- 4. Lequel des deux vous semble plus efficace?

Exercice ALEA.17.7 | Estimer l'amplitude d'une uniforme Soit (X_1, \dots, X_n) un $n \in \mathbf{N}^*$ -échantillon de loi $\mathcal{U}[a, b]$ avec $(a, b) \in \mathbf{R}^2$. On souhaite estimer $b - a$.

- 1. Qu'estime de manière non biaisée $2\overline{X}_n$?
- 2. Déterminer une densité de $\max(X_1, \dots, X_n)$ et $\min(X_1, \dots, X_n)$, puis calculer leur espérance.
- 3. Déterminer quel paramètre estime $\widehat{\theta}_n = \max(X_1, \dots, X_n) - \min(X_1, \dots, X_n)$ de manière non biaisée. Commentez.

Exercice ALEA.17.8 | Loi exponentielle généralisée Soient $a \in \mathbf{R}$, $b \in]0, +\infty[$, et

$$f \left| \begin{array}{l} \mathbf{R} \longrightarrow \mathbf{R} \\ x \longrightarrow \begin{cases} 0 & \text{si } x \leq a, \\ \frac{1}{b} e^{-\frac{x-a}{b}} & \text{sinon.} \end{cases} \end{array} \right.$$

1. Vérifier que f est une densité de probabilité.
2. Soit X une variable aléatoire de densité f , on note $X \hookrightarrow \mathcal{E}(a, b)$.
 - 2.1) Déterminer la loi de $\frac{X-a}{b}$.
 - 2.2) En déduire l'espérance et la variance de X .
3. Soit (X_1, \dots, X_n) un n -échantillon de X . On pose $Y_n = \min(X_1, \dots, X_n)$.
 - 3.1) Montrer que $Y_n \hookrightarrow \mathcal{E}\left(a, \frac{b}{n}\right)$.
 - 3.2) Montrer que Y_n est un estimateur convergeant vers a .
4. On pose $Z_n = \frac{1}{n} \sum_{i=1}^n (X_i - Y_n)$, $U_n = \sum_{i=1}^n X_i$.
 - 4.1) Calculer $\mathbf{E}(Z_n)$, $\mathbf{Var}(Z_n)$, en fonction de $\mathbf{Cov}(U_n, Y_n)$, b et $n \in \mathbf{N}$ pour tout entier n .
 - 4.2) En utilisant l'inégalité $|\rho(U_n, Y_n)| \leq 1$, montrer que : $\lim_{n \rightarrow \infty} \mathbf{Var}(Z_n) = 0$.
 - 4.3) En déduire que Z_n est un estimateur convergeant de b .

Exercice ALEA.17.9 | Améliorons les résultats du baccalauréat Lors de l'examen national du baccalauréat, $N \in \mathbf{N}^*$ candidats obtiennent des moyennes générales entre 0 et 20. Des tests statistiques d'adéquation ont montré que ces moyennes se répartissent selon une $\mathcal{N}(12, 2^2)$. On considère donc dans la suite X_1, \dots, X_N une suite de variables aléatoires réelles indépendantes de même loi $\mathcal{N}(12, 2^2)$. Le candidat $i \in \llbracket 1, N \rrbracket$ est déclaré admis (on note A_i l'évènement associé) si :



1. il a obtenu une moyenne supérieure ou égale à 10,
2. il a obtenu une moyenne entre 8 et 10, et les oraux de rattrapage lui ont permis d'atteindre 10. On suppose que ceci se produit avec une probabilité $\alpha \in]0, 1[$.

On note dans tout l'exercice Φ la fonction de répartition de la loi $\mathcal{N}(0, 1)$, $Y_N = \sum_{i=1}^N \mathbb{1}_{A_i}$ et P_N la variable aléatoire donnant la proportion de candidats admis. L'objectif du ministre de l'Éducation Nationale est d'obtenir un taux de réussite de 90 % — légèrement supérieur à l'année précédente — et souhaiterait ainsi connaître le paramètre α lui permettant d'arriver à ses fins.

1. Soit $i \in \llbracket 1, N \rrbracket$. Montrer que :

$$\mathbf{P}(A_i) = 1 - \Phi(-1) + \alpha(\Phi(-1) - \Phi(-2)).$$

On notera $\beta_\alpha = 1 - \Phi(-1) + \alpha(\Phi(-1) - \Phi(-2))$ dans la suite.

2. Donner le lien entre P_N et Y_N . Que représente Y_N et quelle est sa loi?
3. On suppose N très grand. Par quelle loi normale peut-on approcher la loi de Y_N ? Justifier.
4. Montrer alors que : $\mathbf{P}(P_N \geq 0.9) \approx 1 - \Phi\left(\frac{\sqrt{N}(0.9 - \beta_\alpha)}{\sqrt{\beta_\alpha(1 - \beta_\alpha)}}\right)$.
5.  Tracer la fonction $\alpha \in [0, 001, 0, 999] \mapsto 1 - \Phi\left(\frac{\sqrt{N}(0.9 - \beta_\alpha)}{\sqrt{\beta_\alpha(1 - \beta_\alpha)}}\right)$ à l'aide de Python par exemple pour $N = 1000$. *Indication* : On pourra faire appel à la commande norm.cdf pour obtenir Φ , après avoir réalisé l'import from scipy.stats import norm. Commenter.
6.  Quel algorithme vous permettrait d'obtenir une valeur approchée de α , à 10^{-3} près, qui permettrait d'atteindre une probabilité de 95 % pour l'évènement $\{P_N \geq 0.9\}$? Le mettre en place, et donner une valeur approchée de α pour $N = 1000$ candidats.
7. À la session qui suit, le nouveau ministre de l'Éducation Nationale Jean-Michel B. (qui souhaite garder son anonymat) décide de faire comme son prédécesseur et d'obtenir un taux de réussite supérieur à 90%. Il a à sa disposition un N -échantillon (X_1, \dots, X_N) des notes obtenues à la session précédente (où le taux de réussite était supérieur à 90 %) mais ne connaît pas *a priori* la valeur de α ayant permis une telle prouesse. Proposer une démarche statistique, à l'aide d'un estimateur classique, permettant d'obtenir une estimation de α . On supposera à nouveau N grand.

Solution (exercice ALEA.17.9)

1. Le fonctionnement de l'examen invite ici à conditionner. Nous avons, pour $i \in \llbracket 1, N \rrbracket$,

$$\begin{aligned} \mathbf{P}(A_i) &= \mathbf{P}(A_i | X_i > 10) \mathbf{P}(X_i > 10) + \mathbf{P}(A_i | 8 \leq X_i \leq 10) \mathbf{P}(8 \leq X_i \leq 10), \\ &= \mathbf{P}(A_i | X_i > 10) \mathbf{P}(X_i > 10) + \mathbf{P}(A_i | 8 \leq X_i \leq 10) \mathbf{P}(8 \leq X_i \leq 10), \\ &= 1 \cdot \mathbf{P}(X_i > 10) + \alpha \mathbf{P}(8 \leq X_i \leq 10). \end{aligned}$$

$$\begin{aligned} \text{Or, } X_i^* &\stackrel{\text{(déf.)}}{=} \frac{X_i - 12}{2} \hookrightarrow \mathcal{N}(0, 1) \text{ par propriété de stabilité de la loi normale, donc} \\ &= \mathbf{P}(X_i^* > -1) + \alpha \mathbf{P}(-2 \leq X_i^* \leq -1) \\ &= \boxed{1 - \Phi(-1) + \alpha(\Phi(-1) - \Phi(-2))}. \end{aligned}$$

2. L'interprétation de Y_N est la suivante : cette variable aléatoire correspond au nombre d'admis au baccalauréat sur les N candidat(e)s. Or, P_N est par définition la proportion, donc $P_N = \frac{Y_N}{N}$. De l'interprétation qui précède, nous avons également qu' Y_N est le nombre de succès dans une succession d'expériences de BERNOULLI (l'admission ou non d'un candidat) indépendantes (on suppose que la réussite d'un candidat n'influe pas sur les autres). Donc :

$$Y_N \leftrightarrow \mathcal{B}(N, 1 - \Phi(-1) + \alpha(\Phi(-1) - \Phi(-2))).$$

3. Puisque N est assez grand, le théorème de Moivre-Laplace d'approximation de la loi normale nous livre l'approximation suivante : la loi de Y_N est proche d'un $\mathcal{N}(N\beta_\alpha, N\beta_\alpha(1 - \beta_\alpha))$. Rappelons que ce théorème est bien applicable dans ce contexte puisque les variables aléatoires $\mathbb{1}_{A_i}$ sont indépendantes de même loi et possèdent une variance.

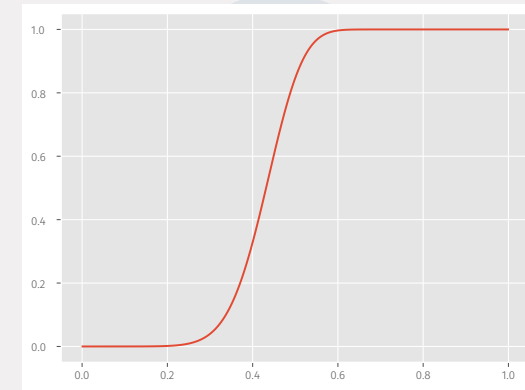
4.

$$\begin{aligned} \mathbf{P}(P_N \geq 0.9) &= \mathbf{P}(Y_N \geq N0.9) \\ &= \mathbf{P}\left(\frac{Y_N - N\beta_\alpha}{\sqrt{N\beta_\alpha(1 - \beta_\alpha)}} \geq \frac{0.9N - N\beta_\alpha}{\sqrt{N\beta_\alpha(1 - \beta_\alpha)}}\right) \\ &\approx 1 - \Phi\left(\frac{0.9N - N\beta_\alpha}{\sqrt{N\beta_\alpha(1 - \beta_\alpha)}}\right) \\ &= 1 - \Phi\left(\frac{\sqrt{N}(0.9 - \beta_\alpha)}{\sqrt{\beta(1 - \beta_\alpha)}}\right). \end{aligned}$$

5. Passons au tracé de la fonction. On commence par définir en Python la fonction $\alpha \mapsto \beta_\alpha$.

```
from scipy.stats import norm
import numpy as np
import matplotlib.pyplot as plt
def beta(alpha):
    return 1 + alpha*(norm.cdf(-1) - norm.cdf(-2)) -
        norm.cdf(-1)
N = 1000
X = np.linspace(0.001, 0.999, 10**3)
```

```
Y = 1 - norm.cdf(np.sqrt(N)*(0.9-beta(X))/np.sqrt(beta(X)*(1 -
    -beta(X))))
plt.plot(X, Y)
```



Nous constatons que la fonction est croissante. Ceci s'explique aisément : plus on laisse passer de candidats après l'oral de rattrapage, plus la proportion d'admis sera grande! À l'aide de ce graphique nous pourrions trouver une valeur approchée de β_α permettant de réaliser la condition $\mathbf{P}(P_N \geq 0.9) \geq 0,95$. On en déduit alors une valeur approchée de α . Nous allons faire autrement dans la question suivante.

6. Il s'agit de résoudre l'équation suivante en β_α :

$$1 - \Phi\left(\frac{\sqrt{N}(0.9 - \beta_\alpha)}{\sqrt{\beta(1 - \beta_\alpha)}}\right) = 0.95,$$

ou de manière équivalente

$$g(\alpha) = 0.05 - \Phi\left(\frac{\sqrt{N}(0.9 - \beta_\alpha)}{\sqrt{\beta(1 - \beta_\alpha)}}\right) = 0$$

Nous allons donc :

- ▶ créer dans Python la fonction g , apparaissant dans le membre de gauche,
- ▶ puis mettre en place un algorithme de Dichotomie pour obtenir la valeur désirée de β_α .
- ▶ Enfin on en déduira facilement une valeur approchée pour α .

```

from scipy.stats import norm
import numpy as np
N = 1000
def beta(alpha):
    return 1 + alpha*(norm.cdf(-1) - norm.cdf(-2)) -
        norm.cdf(-1)
def g(alpha):
    return 0.05 - norm.cdf(np.sqrt(N)*(0.9,
        -beta(alpha))/np.sqrt(beta(alpha)*(1-beta(alpha))))
def dichot(a, b, f, prec):
    """
    Retourne une valeur approchée d'un zéro de f entre a et b
    avec précision prec
    """
    while b-a > prec:
        c = (a+b)/2
        if f(a)*f(c) <= 0:
            b = c
        else:
            a = c
    return (a+b)/2
alpha_suffisant = dichot(0.1, 0.9, g, 10**(-3))

```

Voici ce que l'exécution nous donne : 0.7.

7. Il suffit de regarder les notes entre 8 et 10 des candidats, puis de calculer la fréquence notée f dans la suite de ceux ayant été admis. En remarque (non demandée), donnons une explication Mathématique. Constatons que :
- ▶ $\sum_{j=1}^N \mathbb{1}_{\{8 \leq X_j \leq 10\}}$ est le nombre de candidats au rattrapage,
 - ▶ $\sum_{i=1}^N \mathbb{1}_{A_i \cap \{8 \leq X_i \leq 10\}}$ est le nombre de candidats admis au rattrapage.

La fréquence mentionnée précédemment est donc

$$\begin{aligned}
 f &= \frac{1}{\sum_{j=1}^N \mathbb{1}_{\{8 \leq X_j \leq 10\}}} \sum_{i=1}^N \mathbb{1}_{A_i} \mathbb{1}_{\{8 \leq X_i \leq 10\}} \\
 &= \frac{N}{\sum_{j=1}^N \mathbb{1}_{\{8 \leq X_j \leq 10\}}} \times \frac{\sum_{i=1}^N \mathbb{1}_{A_i \cap \{8 \leq X_i \leq 10\}}}{N}.
 \end{aligned}$$

Rappelons que l'espérance d'une indicatrice est égale à la probabilité de l'évènement. Ainsi, d'après la loi faible des grands nombres, le premier terme est « proche de » $\frac{1}{\mathbf{P}(8 \leq X_i \leq 10)}$. De même, le second est « proche de » $\mathbf{P}(A_i \cap 8 \leq X_i \leq 10) = \mathbf{P}(A_i | 8 \leq X_i \leq 10) \mathbf{P}(8 \leq X_i \leq 10) = \alpha \mathbf{P}(8 \leq X_i \leq 10)$. Donc, *in fine*,

$$f \approx \frac{\alpha}{\mathbf{P}(8 \leq X_i \leq 10)} \mathbf{P}(8 \leq X_i \leq 10) = \boxed{\alpha}.$$

Seconde remarque : les manipulations du « proche de » nécessitent des connaissances sur la convergence en probabilités, qui dépassent très largement le cadre du programme.

4.3. Intervalles de confiance

Exercice ALEA.17.10 | Contrairement à ce que pense Popeye, l'épinard n'est pas l'aliment le plus riche en fer. La lentille, par exemple, en apporte davantage. On a procédé à des analyses sur 10 échantillons de lentilles (de loi parente X) et d'épinards (de loi parente Y), et on a relevé la teneur en fer (en mg pour 100 g de produit frais).

Echantillon	1	2	3	4	5	6	7	8	9	10
Epinard ($\sim Y$)	2.64	2.75	2.82	2.72	2.56	2.59	2.83	2.70	2.67	2.62

Lentille ($\sim X$)	9.02	9.08	8.82	8.94	8.95	9.11	9.14	9.02	9.04	8.85
-----------------------	------	------	------	------	------	------	------	------	------	------

- Calculer la teneur moyenne en fer, l'écart-type associé, la médiane et les quartiles pour les épinards et les lentilles, sur les relevés statistiques donnés dans le tableau.
- Déterminer un intervalle de confiance à 95 % pour la teneur moyenne en fer des épinards et des lentilles, *i.e.* pour l'espérance de X et Y.
- Proposer une représentation graphique illustrant le propos initial.

Exercice ALEA.17.11 | Une usine fabrique des câbles. On suppose que la charge maximale supportée par un câble, exprimée en tonnes, est une variable aléatoire qui suit une loi normale $\mathcal{N}(\mu, 0, 5^2)$. Une étude portant sur 50 câbles a donné une moyenne des charges maximales supportées égales à 12,2 tonnes.

- Déterminer l'intervalle de confiance à 99% de la charge maximale moyenne de tous les câbles fabriqués par l'usine.
- Déterminer la taille minimale de l'échantillon étudié pour que la longueur de l'intervalle de confiance à 99% soit inférieure ou égale à 0,2?

Exercice ALEA.17.12 | **Principe CMR : capture/marquage/recapture** Dans cet exercice, il va être question d'évaluer, de diverses manières, le nombre L (inconnu) de souris dans la cantine. Pour cela, on a prélevé dans la cantine 200 souris que l'on a marquées avant de les relâcher. Il y a donc maintenant dans la cantine L souris dont 200 sont marquées. En capturant de nouveau un certain nombre (on prendra 100) de souris on observe $\overline{X}_{100} = 0,6$, on va devoir estimer au mieux le nombre L.

- On capture une souris, quelle est la probabilité p qu'elle soit marquée?
- Quel estimateur sans biais et convergent de p connaissez-vous?
- En utilisant l'inégalité de BIENAYMÉ-TCHEBYCHEV, proposer un intervalle de confiance non-asymptotique dans lequel le nombre L a une probabilité supérieure à 0.95 % de se trouver. On utilisera la majoration classique $p(1-p) \leq \frac{1}{4}$.
- En utilisant le théorème central limite, trouver un intervalle de confiance asymptotique dans lequel le nombre L a une probabilité supérieure à 0.95 % de se trouver.
- Comparer les deux intervalles de confiance.



Exercice ALEA.17.13 | Agro—Véto, Sujet 3, 2018

On pourra utiliser pour les programmes Python la fonction `linalg.matrix_rank()` du module `numpy`, qui permet de déterminer le rang d'une matrice, comme le montre l'exemple suivant :

```
>>> import numpy as np
>>> A = np.array([[1, 2, 1], [2, 3, 2], [3, 5, 3]])
>>> np.linalg.matrix_rank(A)
2
```

La dernière ligne affiche le rang de la matrice $\begin{pmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 3 & 5 & 3 \end{pmatrix}$, c'est-à-dire 2. On pourra aussi utiliser la fonction `randint()` du module `random`. Pour a et b deux entiers `randint(a, b)` retourne un entier équiprobablement entre a et b (a et b étant inclus). On considère la matrice :

$$A = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 0 & 1 \\ -3 & 0 & -1 \end{pmatrix}.$$

- 1.1)  Écrire une fonction Python prenant en arguments deux vecteurs de taille 3 et renvoyant un booléen (True ou False) indiquant s'ils sont colinéaires. (On pourra représenter les vecteurs par des listes).
 - 1.2)  Écrire une fonction Python `vecteurs_propres(u)` prenant en argument un vecteur de taille 3 et renvoyant un booléen (True ou False) indiquant s'il est un vecteur propre de A.
- 2.1) Vérifier que $-1, 1, 2$ sont valeurs propres de A et préciser pour chacune un vecteur propre associé.
 - 2.2) La matrice A est-elle diagonalisable?
- Soient X_1, \dots, X_n , n variables aléatoires indépendantes suivant la loi de BERNOULLI

de paramètre $p \in]0; 1[$. On note :

$$M_n = \frac{1}{n} \sum_{k=1}^n X_k, \quad M_n^* = \frac{M_n - p}{\sqrt{\frac{p(1-p)}{n}}}$$

3.1) Donner, pour $\alpha \in \mathbf{R}_+^*$, l'approximation de la probabilité $P(-\alpha < M_n^* < \alpha)$ donnée par le théorème central limite.

3.2) En déduire que $\left[M_n - \frac{1}{\sqrt{n}}, M_n + \frac{1}{\sqrt{n}} \right]$ est un intervalle de confiance de p au seuil de 95%. On pourra admettre que, $\forall x \in [0; 1]$, $x(1-x) \leq \frac{1}{4}$ et si Φ désigne la fonction de répartition d'une variable suivant une loi normale centrée réduite, alors $\Phi(1; 96) \approx 0,975$.

4. On note N_V le nombre de vecteurs propres de A dont les coefficients sont des entiers de $\llbracket -5, 5 \rrbracket$.

4.1) Expliquer comment le programme suivant permet d'estimer la valeur de N_V :

```
def simul ():
    u = [ randint (-5 ,5) for k in range (3) ]
    return vecteurs_propres(u)
n = 10000 #Valeur de n a definir.
nb = 0
for k in range (n):
    if simul ():
        nb += 1
print(round (nb/n *11**3)) # round (x) = l'entier le plus
    - proche de x.
```

4.2) Comment choisir n pour que l'on soit sûr à 95% de la valeur affichée?

4.3) Commenter le résultat obtenu.

Solution (exercice ALEA.17.13)

On considère la matrice :

$$A = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 0 & 1 \\ -3 & 0 & -1 \end{pmatrix}.$$

1. 1.1) On rappelle que deux vecteur u et v sont colinéaires si et seulement si $\text{Rg}(u, v) < 2$.

```
def colineaires(u, v):
    a = np.array([u, v])
    return np.linalg.matrix_rank(a) < 2 #retourne un
    - booléen
```

Nous pouvons alors tester si $[1, 1], [2, 2]$ sont colinéaires : True.

1.2) Le vecteur u est vecteur propre de A si et seulement si u est non nul et Au est colinéaire à u . Il suffit alors de tester la condition $u \neq 0$ et la colinéarité entre le produit Au et u :

```
def vecteurs_propres(u):
    return u != [0,0,0] and colineaires(np.dot(A,u), u)
```

2. 2.1) On pose $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$; $AX = -X \iff (A + I)X = 0 \iff \begin{cases} 4x + y + z = 0 \\ x + y + z = 0 \\ -3x = 0 \end{cases} \iff \begin{cases} x = 0 \\ z = -y \end{cases}$ Le système admet d'autres solutions que $(0, 0, 0)$ donc -1 est valeur propre de A et

$$E_{-1}(A) = \text{Vect} \left(\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right).$$

$$AX = X \iff \begin{cases} 2x + y + z = 0 \\ x - y + z = 0 \\ -3x - 2z = 0 \end{cases} \iff \begin{cases} y + 2x + z = 0 \\ 3x + 2z = 0 \\ -3x - 2z = 0 \end{cases} \iff$$

$\begin{cases} y = -1/2x \\ z = -3/2x \end{cases}$ De même, on en déduit que 1 est valeur propre de A et

$$E_1(A) = \text{Vect} \left(\begin{pmatrix} -2 \\ 1 \\ 3 \end{pmatrix} \right).$$

$$AX = 2X \iff \begin{cases} x + y + z = 0 \\ x - 2y + z = 0 \\ -3x - 3z = 0 \end{cases} \iff \begin{cases} y + x + z = 0 \\ 3x + 3z = 0 \\ -3x - 3z = 0 \end{cases} \iff \begin{cases} y = 0 \\ z = -x \end{cases}$$

Ainsi 2 est bien valeur propre de A et

$$E_2(A) = \text{Vect} \left(\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right).$$

Comme A ne peut avoir plus de trois valeurs propres, on en déduit que

$$\text{Spec}(A) = \{-1, 1, 2\}.$$

2.2) A est carrée d'ordre 3 et possède 3 valeurs propres distinctes donc

A est diagonalisable.

3. Soient X_1, \dots, X_n , n variables aléatoires indépendantes suivant la loi de BERNOULLI de paramètre $p \in]0; 1[$. On note : $M_n = \frac{1}{n} \sum_{k=1}^n X_k$ et $M_n^* = \frac{M_n - p}{\sqrt{\frac{p(1-p)}{n}}}$.

3.1) X_k suit la loi de BERNOULLI $\mathbf{B}(p)$ donc $\mathbf{E}(X_k) = p$ et $\mathbf{Var}(X_k) = p(1-p)$ pour tout $k \in \llbracket 1, n \rrbracket$. On en déduit, par linéarité de l'espérance que $\mathbf{E}(M_n) = \frac{1}{n} \sum_{k=1}^n \mathbf{E}(X_k) = \frac{1}{n} \sum_{k=1}^n p = p$. De plus, comme les variables aléatoires X_1, \dots, X_n sont indépendantes, on a :

$$\mathbf{Var}(M_n) = \frac{1}{n^2} \sum_{k=1}^n \mathbf{Var}(X_k) = \frac{p(1-p)}{n} \text{ et par conséquent } \sigma_{M_n} = \sqrt{\frac{p(1-p)}{n}}.$$

Ainsi si $M_n^* = \frac{M_n - \mathbf{E}(M_n)}{\sigma(M_n)}$: M_n^* correspond à la variable centrée réduite associée à M_n . Comme les variables aléatoires X_1, \dots, X_n sont indépendantes et de même loi, on peut appliquer le théorème central limite : on en déduit que

M_n^* suit approximativement la loi normale $\mathcal{N}(0, 1)$. Ainsi, pour $\alpha \in \mathbf{R}_+^*$,

$\mathbf{P}(-\alpha < M_n^* < \alpha) \approx \Phi(\alpha) - \Phi(-\alpha)$ où Φ désigne la fonction de répartition de la loi normale centrée réduite. Sachant qu'on a $\Phi(-\alpha) = 1 - \Phi(\alpha)$ et $\mathbf{P}(-\alpha \leq X \leq \alpha) = \mathbf{P}(-\alpha < X < \alpha)$ pour toute variable X à densité, on peut aussi écrire que

$$\mathbf{P}([-\alpha \leq M_n^* \leq \alpha]) \approx 2\Phi(\alpha) - 1.$$

3.2)

$$\begin{aligned} \mathbf{P} \left(p \in \left[M_n - \frac{1}{\sqrt{n}}; M_n + \frac{1}{\sqrt{n}} \right] \right) &= \mathbf{P} \left(-\frac{1}{\sqrt{n}} \leq M_n - p \leq \frac{1}{\sqrt{n}} \right) \\ &= \mathbf{P} \left(-\frac{1}{\sqrt{p(1-p)}} \leq M_n^* \leq \frac{1}{\sqrt{p(1-p)}} \right) \end{aligned}$$

Comme $\sqrt{p(1-p)} \leq \sqrt{1/4} = 1/2$ alors $-\frac{1}{\sqrt{p(1-p)}} \leq -2$ et $2 \leq \frac{1}{\sqrt{p(1-p)}}$.²⁰ Donc :

$$\mathbf{P} \left(-\frac{1}{\sqrt{p(1-p)}} \leq M_n^* \leq \frac{1}{\sqrt{p(1-p)}} \right) \geq \mathbf{P}(-2 \leq M_n^* \leq 2). \text{ Or } \mathbf{P}(-2 \leq M_n^* \leq 2) = 2\Phi(2) - 1 \geq 2\Phi(1,96) - 1 = 0,95 \text{ car la fonction } \Phi \text{ est croissante sur } \mathbf{R}. \text{ Dès lors, on en déduit que } \mathbf{P} \left(-\frac{1}{\sqrt{p(1-p)}} \leq M_n^* \leq \frac{1}{\sqrt{p(1-p)}} \right) \geq 0,95.$$

$\mathbf{P} \left(p \in \left[M_n - \frac{1}{\sqrt{n}}; M_n + \frac{1}{\sqrt{n}} \right] \right) \geq 95\%$ ce qui signifie que

$$\left[M_n - \frac{1}{\sqrt{n}}, M_n + \frac{1}{\sqrt{n}} \right] \text{ est un intervalle de confiance de } p \text{ au seuil de } 95\%.$$

4. 4.1)

On note N_V (respectivement p) le nombre (respectivement la proportion) de vecteurs propres de A qui appartient à $\llbracket -5, 5 \rrbracket^3$. Comme $\# \llbracket -5, 5 \rrbracket^3 = 11^3$, alors $p = \frac{N_V}{11^3}$ soit $N_V = p \times 11^3$. On considère l'épreuve de BERNOULLI qui consiste à choisir au hasard un vecteur de $\llbracket -5, 5 \rrbracket^3$, puis à renvoyer 1 si le vecteur en question est un vecteur propre de la matrice A (la probabilité de succès est notre paramètre de BERNOULLI). On réalise $n = 10000$ fois dans des conditions indépendantes cette expérience (ce qui est réalisé dans la boucle for). On note $X_k = 1$ si le k -ième vecteur tiré est vecteur propre de A, $X_k = 0$ sinon. Les variables aléatoires X_k sont indépendantes et de même loi $\mathcal{B}(p)$. On sait alors que la variable aléatoire $M_n = \frac{1}{n} \sum_{k=1}^n X_k$ est un estimateur sans biais de p . Le nombre nb/n en sortie de boucle correspond à une réalisation de la variable M_n et donne une estimation de p . En multipliant par 11^3 et en arrondissant à l'entier le plus proche (car $N_V \in \mathbf{N}$)²¹, on obtient donc une estimation de N_V .

²⁰On « remontre » ici, dans ce cas particulier, que l'on peut épaissir tout intervalle de confiance de seuil $1 - \alpha$, la version épaissie reste un intervalle de confiance de seuil $1 - \alpha$.

²¹Attention, cette fonction n'est pas la partie entière, qui est `int()` dans Python

4.2) On a vu précédemment que $\mathbf{P}\left(M_n - \frac{1}{\sqrt{n}} \leq p \leq M_n + \frac{1}{\sqrt{n}}\right) \geq 0,95$. Cela équivaut à : $\mathbf{P}\left(11^3 M_n - \frac{11^3}{\sqrt{n}} \leq N_V \leq 11^3 M_n + \frac{11^3}{\sqrt{n}}\right) \geq 0,95$. Si on choisit n tel que $\frac{11^3}{\sqrt{n}} \leq 0,5$, on aura donc

$$\mathbf{P}(11^3 M_n - 0,5 \leq N_V \leq 11^3 M_n + 0,5) \geq 0,95.$$

Par ailleurs, l'entier N_n le plus proche de $11^3 M_n$ vérifie

$$11^3 M_n - 0,5 \leq N_n \leq 11^3 M_n + 0,5.$$

On en déduit que l'écart entre N_n et N_V est inférieur ou égal à 1 (avec une probabilité d'au moins 95%). Or $\frac{11^3}{\sqrt{n}} \leq 0,5 \iff \sqrt{n} \geq 2 * 11^3 \iff n \geq 4 * 11^6$. Donc en choisissant $n \geq 4 * 11^6$ (soit $n \geq 7086244$), la valeur affichée $\text{round}(nb/n * 11^{**3})$ donne une estimation de N_V à 95%.

4.3) On reprend le programme du début de la question, en remplaçant n par 7086244. On obtient $\boxed{22}$ après exécution. Calculons la valeur exacte de N_V afin de la comparer à 22 : D'après l'étude réalisée en seconde question, les vecteurs propres de A à coefficients entiers sont de la forme $(0, k, -k)$ ou $(-2k, k, 3k)$ ou $(k, 0, -k)$ avec $k \in \mathbb{Z}^*$. Comme il y a 10 entiers non nuls compris entre -5 et 5 , on dénombre :

- ▶ 10 vecteurs propres $(0, k, -k)$ éléments de $\llbracket -5, 5 \rrbracket^3$
- ▶ 10 vecteurs propres $(k, 0, -k)$ éléments de $\llbracket -5, 5 \rrbracket^3$

Reste à dénombrer ceux qui sont de la forme $(-2k, k, 3k)$ avec $k \neq 0$. Il faut

$$\text{que l'on ait : } \begin{cases} 0 < |2k| \leq 5 \\ 0 < |k| \leq 5 \\ 0 < |3k| \leq 5 \end{cases}$$

Les seuls entiers k qui conviennent sont -1 et 1 .

Il y a donc 2 vecteurs propres de la forme $(-2k, k, 3k)$ qui appartiennent à $\llbracket -5, 5 \rrbracket^3$.

On en déduit que $\boxed{N_V = 10 + 10 + 2 = 22}$ ce qui correspond à la valeur obtenue par estimation dans la question précédente.

4.4. Tests

Exercice ALEA.17.14 | On considère un test de dépistage d'une maladie qui a donné les résultats suivants lors d'une phase de test sur 1600 individus sains et 1600 individus malades :

	Malades	Sains
Tests positifs	1522	92
Tests négatifs	78	1508

On considère p_1 la probabilité d'avoir un test positif pour une personne saine et p_2 la probabilité d'avoir un test négatif pour une personne malade.

1. Déterminer un intervalle de confiance pour p_1 au risque 5 %. On arrondira au millième.
2. Combien doit on avoir au plus de personnes malades testées négatives pour valider l'hypothèse $p_2 = 0.001$? On arrondira les calculs à 10^{-4} .

Exercice ALEA.17.15 |

1. Afin d'étudier le pourcentage p de consommateurs satisfait par le produit A, on interroge 100 consommateurs et 56 déclarent être satisfaits. Est-ce suffisant pour continuer l'exploitation du produit A ? *Indication : on cherchera un intervalle de confiance à 95%.*
2. En supposant qu'on garde la même moyenne empirique de 0,56, et le même risque $\alpha = 0,05$, combien de personnes doit-on interroger pour prendre une décision ?

IV

Quatrième partie

Annexes

Chapitre ANN.18.

L'alphabet grec

Cet alphabet, très souvent utilisé pour écrire des symboles mathématiques, compte vingt-quatre lettres minuscules et majuscules. Certaines majuscules se retrouvent dans l'alphabet latin.

Minuscules	Majuscules	Prononciation
α	Α	alpha
β	Β	beta
γ	Γ	gamma
δ	Δ	delta
ε	Ε	epsilon
ζ	Ζ	zêta
η	Η	êta
θ	Θ	thêta
ι	Ι	iota
κ	Κ	kappa
λ	Λ	lambda
μ	Μ	mu
ν	Ν	nu
ξ	Ξ	xi
ο	Ο	omicron
π	Π	pi
ρ	Ρ	rho

σ	Σ	sigma
τ	Τ	tau
υ	Υ	upsilon
φ	Φ	phi
χ	Χ	chi
ψ	Ψ	psi
ω	Ω	omega

TAB. ANN.18.1. : L'alphabet grec

En minuscules, on utilise en général vingt-un de ces vingt-quatre lettres, à l'exception des lettres ι, ο et υ. En effet ces lettres ressemblent trop aux lettres latines i, o et v. On veillera d'ailleurs à ne pas confondre la lettre grecque ω « *omega* » avec la lettre latine w.

En majuscules, on utilise les dix lettres qui ne font pas partie de l'alphabet latin. En l'occurrence, il s'agit de : Γ Δ Θ Λ Ξ Π Σ Φ Ψ Ω.

Chapitre ANN.19.

Résumé des méthodes

1. EN ALGÈBRE

Méthode Unicité de l'écriture algébrique et identification

Une reformulation de l'unicité de l'écriture $z = x + iy$ est la suivante :

$$x + iy = x' + iy' \iff x = x', \quad y = y', \quad (x, y, x', y') \in \mathbf{R}^4.$$

On peut donc *identifier* partie réelle et partie imaginaire.

Méthode Expression conjuguée

Dans la preuve précédente on a utilisé une technique classique pour obtenir l'inverse d'un nombre complexe écrit sous forme algébrique. On peut la résumer comme suit :

$$\frac{1}{x + iy} = \frac{x - iy}{(x + iy)(x - iy)} = \frac{x}{x^2 + y^2} + i \left(\frac{-y}{x^2 + y^2} \right).$$

Méthode Développement d'une norme de somme au carré

Soit $|z + z'|^2$ avec $z, z' \in \mathbf{C}$.

1. Écrire la quantité en fonction du conjugué : $|z + z'|^2 = (z + z')\overline{(z + z')}$.

2. Développer.

Méthode Mettre sous forme exponentielle un nombre complexe

Soit $z \neq 0$.

- Calculer $|z|$, puis $\frac{z}{|z|}$.
- Chercher $\theta \in [0, 2\pi[$ tel que : $\frac{z}{|z|} = e^{i\theta}$, *i.e.* tel que


$$\cos(\theta) = \frac{\operatorname{Re}(z)}{|z|}, \quad \sin(\theta) = \frac{\operatorname{Im}(z)}{|z|}.$$

La forme exponentielle est alors : $z = |z|e^{i\theta}$. Il arrive parfois que l'angle θ

Méthode Technique de l'angle moitié (forme trigonométrique d'une somme d'exponentielles imaginaires)

Soient deux nombres complexes z, z' de module un donné sous forme trigonométrique : $z = e^{i\theta}, z' = e^{i\theta'}$ avec $(\theta, \theta') \in [0, 2\pi[$. Alors la forme trigonométrique de $z + z'$ s'obtient par le calcul suivant :

$$\begin{aligned} z + z' &= e^{i\theta} + e^{i\theta'} = e^{i\frac{\theta+\theta'}{2}} \left(e^{i\frac{\theta-\theta'}{2}} + e^{-i\frac{\theta-\theta'}{2}} \right) \\ &= 2e^{i\frac{\theta+\theta'}{2}} \cos\left(\frac{\theta-\theta'}{2}\right). \end{aligned}$$

 La méthode s'adapte à $z - z'$ en faisant apparaître un sinus. On obtient alors facilement module et argument :

$$|z + z'| = 2 \left| \cos \left(\frac{\theta - \theta'}{2} \right) \right|, \quad \text{Arg}(z + z') \equiv \frac{\theta + \theta'}{2} \quad [2\pi].$$

Méthode Calculs de racines n -ième de complexes

On cherche donc les solutions de $z^n = \alpha$ avec $\alpha \neq 0$ (si $\alpha = 0$ il n'y a que zéro comme solution).

1. Calculer la forme trigonométrique de $\alpha = \rho e^{i\theta}$.
2. Chercher z sous la forme $z = \rho' e^{i\theta'}$.
3. En remplaçant, on obtient comme conditions $(\rho')^n = \rho$ et $n\theta' = \theta + 2k\pi$ avec $k \in \mathbf{Z}$. Résoudre ces deux équations puis conclure.

Dans la pratique, la méthode sous-entend que l'on est capable de trouver une racine n -ième de α .

Méthode Écriture d'une combinaison linéaire de fonctions trigonométriques sous « forme déphasée »

Soient $a, b, x \in \mathbf{R}$. On souhaite transformer l'expression $E(x) = a \cos x + b \sin x$ en $\rho \cos(x + \varphi)$, avec $\rho \in \mathbf{R}^+, \varphi \in \mathbf{R}$.

1. Mettre $\rho = \sqrt{a^2 + b^2}$ en facteur, de sorte que

$$E(x) = \rho \left(\frac{a}{\rho} \cos x + \frac{b}{\rho} \sin x \right).$$

2. Comme $\left(\frac{a}{\rho}, -\frac{b}{\rho} \right)$ est sur le cercle unité, puisque $\left(\frac{a}{\rho} \right)^2 + \left(-\frac{b}{\rho} \right)^2 = 1$, il existe $\varphi \in [0, 2\pi[$ tel que :

$$\frac{a}{\sqrt{a^2 + b^2}} = \cos \varphi, \quad -\frac{b}{\sqrt{a^2 + b^2}} = \sin \varphi.$$

3. Alors $E(x) = \cos x \cos \varphi - \sin x \sin \varphi = \cos(x + \varphi)$.

Une méthode analogue existe si l'on souhaite une forme déphasée de la forme $\rho \sin(x + \varphi)$, il suffit de choisir l'angle différemment.

Méthode Linéarisation & Antilinéarisation avec des complexes

1. (Pour linéariser $\cos^k \theta, \sin^k \theta$) écrire

$$\cos^k \theta = \left(\frac{e^{i\theta} + e^{-i\theta}}{2} \right)^k, \quad \sin^k \theta = \left(\frac{e^{i\theta} - e^{-i\theta}}{2i} \right)^k,$$

puis développer avec le binôme, regrouper les termes avec leur conjugué, utiliser les formules d'EULER.

2. (Pour antilinéariser $\cos(k\theta), \sin(k\theta)$) écrire

$$\begin{aligned} \cos(k\theta) &= \text{Re} \left(e^{i k \theta} \right) \underset{\text{MOIVRE}}{=} \text{Re} \left(\left(e^{i\theta} \right)^k \right) = \text{Re} \left((\cos \theta + i \sin \theta)^k \right), \\ \sin(k\theta) &= \text{Im} \left(e^{i k \theta} \right) \underset{\text{MOIVRE}}{=} \text{Im} \left(\left(e^{i\theta} \right)^k \right) = \text{Im} \left((\cos \theta + i \sin \theta)^k \right), \end{aligned}$$

puis développer avec le binôme et calculer les parties réelles et imaginaires.

Méthode Calculs de sommes trigonométriques

1. Écrire \cos, \sin comme des parties réelles/imaginaires d'exponentielles complexes.
2. Utiliser la linéarité de $\text{Re}(\dots), \text{Im}(\dots)$, i.e. : $\text{Re}(\sum \dots) = \sum \text{Re}(\dots), \text{Im}(\sum \dots) = \sum \text{Im}(\dots)$.
3. Utiliser la formule donnant la somme de termes géométriques. Conclure.

Méthode Montrer qu'un polynôme est nul

Pour montrer qu'un polynôme est nul, on peut au choix :

1. montrer que tous ses coefficients sont nuls,
2. montrer qu'il admet plus de racines que son degré (en particulier s'il en admet une infinité).

Le plus souvent, on utilise 2 pour en déduire la nullité de tous les coefficients.

 **Méthode Montrer que deux polynômes sont égaux**

Pour montrer que deux polynômes sont égaux, on peut au choix :

1. montrer que leurs coefficients sont identiques,
2. montrer que la différence admet plus de racines que son degré (en particulier si elle en admet une infinité).

 **Méthode Factorisation d'un polynôme**

Soit $P \in \mathbf{K}[X]$. Pour transformer P en un produit de polynômes de degré 1 ou 2, on :

1. cherche une racine $\lambda \in \mathbf{K}$.
2. On écrit P sous la forme $(X - \lambda) \times Q = P$ avec $Q \in \mathbf{K}[X]$.
3. On recommence le processus avec Q .

En résumé : cela revient à chercher les racines de P .

 **Méthode Lien entre la factorisation sur \mathbf{C} et \mathbf{R}**

Pour décomposer un polynôme en produit d'irréductibles dans $\mathbf{R}[X]$, on peut déjà le décomposer dans $\mathbf{C}[X]$ (en cherchant ses racines), puis on regroupe les racines complexes conjuguées entre elles.

 **Méthode Système à somme et produit fixés**

Ainsi, pour résoudre le système

$$\begin{cases} x_1 + x_2 = s, \\ x_1 x_2 = p \end{cases}$$

en $(x_1, x_2) \in \mathbf{K}^2$, il suffit de chercher les racines de $X^2 - sX + p$ à l'aide du discriminant.

 **Méthode Montrer l'appartenance « à un Vect »**

Pour montrer que $x \in \text{Vect}(x_i)_{i \in I}$, on cherche $\lambda_1, \dots, \lambda_n$ des scalaires, tels que :

$$x = \sum_{i=1}^n \lambda_i x_i. \text{ En particulier, si } E = \mathbf{K}^n, \text{ on tombe sur la résolution d'un système}$$

 linéaire.

 **Méthode Montrer qu'un ensemble n'est pas un espace vectoriel**

Pour montrer qu'un ensemble n'est pas un espace vectoriel, on peut montrer qu'il n'est pas un sous-espace vectoriel d'un espace vectoriel de référence. En

- ▶ vérifiant qu'il ne contient pas le neutre dudit ensemble de référence.
- ▶ Ou, on montre qu'il n'est pas stable par combinaison linéaire.

 **Méthode Montrer qu'un ensemble est un espace vectoriel**

Deux options sont possibles.

1. Justifier qu'il est un sous-espace vectoriel d'un espace vectoriel de référence (\mathbf{K}^n , de polynômes, de fonctions, de suites ...) **ou**
2. Montrer que l'ensemble s'écrit comme Vect d'une famille.

 **Méthode Lien entre paramétrisation & équations cartésiennes**

- ▶ **Paramétrisation** → **Équations implicites / cartésiennes** : Résolution d'un système en les paramètres de la forme paramétrique (généralement notés λ, μ, ν, \dots). Le système présentera alors une condition de compatibilité qui sera l'équation cartésienne cherchée.
- ▶ **Équations implicites / cartésiennes** → **Paramétrisation** : Résolution du système linéaire en les inconnues (généralement notées x, y, z, \dots), l'ensemble des solutions correspond alors à la forme paramétrique cherchée.

 **Méthode Montrer la liberté/liaison d'une famille**

1. Pour montrer qu'une famille $(x_k)_{1 \leq k \leq n}$ est libre, on écrit :

« Soit $(\lambda_k)_{1 \leq k \leq n} \in \mathbf{K}^n$ tel que $\sum_{k=1}^n \lambda_k x_k = 0_E$. Alors [...] donc les λ_k sont tous nuls. » En général, l'étape [...] consiste en les arguments suivants :

- ▶ Dans $\mathbf{R}^n, \mathbf{C}^n$ ou $\mathbf{K}_n[X]$: on arrive à la résolution d'un système linéaire.
 - ▶ Dans $\mathbf{R}^1, \mathbf{R}^N$ on fait de l'analyse (limites, dérivation, etc.).
2. Pour montrer qu'une famille $(x_k)_{1 \leq k \leq n}$ est liée, on écrit :

« [...] Posons alors $\lambda_1 = \dots, \dots, \lambda_n = \dots$: on a alors $\sum_{k=1}^n \lambda_k x_k = 0_E$, mais les λ_k ne sont pas tous nuls. »

Méthode Montrer qu'une famille est génératrice

Pour montrer qu'une famille $(x_k)_{1 \leq k \leq n}$ ¹ de vecteurs de F est une famille génératrice de F, on écrit :

« Soit $x \in F$. Alors cherchons $\lambda_1, \dots, \lambda_n$ tels que $x = \sum_{k=1}^n \lambda_k x_k$. [...] On a donc déterminé des λ_i qui conviennent, la famille est génératrice. »

En général, l'étape [...] consiste en les arguments suivants :

- ▶ Dans \mathbf{R}^n ou \mathbf{C}^n : on résout un système linéaire.
- ▶ Dans \mathbf{R}^1 , \mathbf{R}^n on fait de l'analyse (limites, dérivation, etc.).

Méthode Montrer qu'une application n'est pas linéaire

Pour montrer qu'une application n'est pas linéaire, on peut :

- ▶ vérifier si l'égalité $u(0_E) = 0_F$ est vérifiée.
- ▶ Si c'est le cas, on cherche s'il existe $\lambda, \mu \in \mathbf{K}$ et $x, y \in E$ tels que : $u(\lambda.x + \mu.y) \neq \lambda.u(x) + \mu.u(y)$.

Méthode Image d'une application si une famille génératrice de l'espace de départ est connue

1. On commence par chercher une famille génératrice \mathcal{G} de l'ensemble de départ E.
2. On calcule les images de chacun des vecteurs de \mathcal{G} .
3. Si l'on souhaite une base, on cherche à extraire une sous-famille libre.

Méthode Construction d'applications linéaires à l'aide d'une base

À la question « construisez une application linéaire entre E et F », si vous connais-

sez une base $\mathcal{B} = (e_1, \dots, e_n)$ de E, vous pouvez répondre :

je pose $u(e_1) = \text{Truc}_1 \in F$, ..., je pose $u(e_n) = \text{Truc}_n$,

en découlera alors automatiquement $u(x)$ pour tout $x \in E$ par linéarité.

Méthode Binôme et calculs des puissances

Si on arrive à écrire une matrice comme somme d'une matrice D diagonale et d'une matrice nilpotente N (i.e. telle que $N^p = 0$ pour un certain $p \in \mathbf{N}$), qui **commutent**, on utilise la formule du binôme matricielle :

$$(D + N)^p = \sum_{k=0}^p \binom{p}{k} D^{p-k} N^k.$$

Méthode Inverse matriciel à l'aide d'un polynôme annulateur

Supposons qu'il existe $a_0, \dots, a_p \in \mathbf{K}$ tel que $a_0 \neq 0$, et soit $A \in \mathcal{M}_{n,n}(\mathbf{K})$ une matrice carrée vérifiant

$$a_0 I_n + a_1 A + \dots + a_p A^p = 0_n. \quad (1)$$

On dit que $P = a_0 + a_1 X + \dots + a_p X^p$ est *annulateur de A*. Alors (1) est équivalente à

$$a_1 A + \dots + a_p A^p = -a_0 I_n,$$

puis étant donné que a_0 est non nul,

$$A \left(-\frac{a_1}{a_0} I_n + \dots - \frac{a_p}{a_0} A^{p-1} \right) = I_n.$$

La matrice A est alors inversible d'inverse $-\frac{a_1}{a_0} I_n + \dots - \frac{a_p}{a_0} A^{p-1}$.

¹adapter la méthode pour une famille quelconque $\text{Vect}(x_i)_{i \in I}$

Méthode Passer d'une forme échelonnée en ligne à une forme échelonnée réduite en ligne

On élimine de droite à gauche, colonne par colonne, tous les coefficients au-dessus de la diagonale à l'aide de transvections, donc d'opérations de la forme $L_i \leftarrow L_i + \mu L_j$ avec $j > i$ et $\mu \in \mathbf{R}$.

Méthode du miroir

Supposons que l'on arrive, après avoir effectué certaines opérations élémentaires sur les lignes de A (i.e. après avoir multiplié par $E = \dots E_3 E_2 E_1$, un produit de matrices d'opérations élémentaires, à gauche la matrice A), à transformer A en I_n . Alors on obtient l'égalité matricielle :

$$\begin{matrix} EA \\ I_n \end{matrix} = \begin{matrix} EI_n A \\ A^{-1} \end{matrix}, \quad \text{qui fournit ainsi l'inverse de A,}$$

la matrice inverse sera alors $EI_n = E$.

Méthode Opérations pour la recherche d'éléments propres

- ▶ Si le coefficient (3, 1) n'est pas nul :
 1. l'opération optimale à effectuer en premier pour des matrices de taille 3×3 est la permutation $L_1 \longleftrightarrow L_3$ qui permettra d'obtenir un coefficient indépendant de λ en position (1, 1). On élimine alors avec celui-ci les coefficients (2, 1), (3, 1).
 2. Ensuite, en position (2, 2), nous avons un coefficient affine en λ et que l'on souhaite utiliser en nouveau pivot afin d'éliminer le coefficient (3, 2). Pour éliminer λ en (2, 2), on peut faire une opération simple en fonction de L_3 .
 3. Un pivot indépendant de λ est alors obtenu en (2, 2), on peut alors éliminer le coefficient (3, 2).
- ▶ Si le coefficient (3, 1) est nul :
 1. on fait la permutation $L_1 \longleftrightarrow L_2$ qui permettra d'obtenir un coefficient indépendant de λ en position (1, 1). On élimine alors avec celui-ci le co-

²Afin de ne pas ajouter des termes non nuls dans les colonnes à gauche du pivot de la ligne i

efficient (3, 1).

2. En positions (2, 2), nous avons un coefficient indépendant de λ qui sert à éliminer le coefficient (3, 2).

Méthode Montrer qu'une famille est une base à l'aide de matrice

Soit \mathcal{B} une base de E un espace vectoriel, et \mathcal{F} une famille telle que $\#\mathcal{B} = \#\mathcal{F}$.

- ▶ Calculer $M = \underset{\mathcal{B}}{\mathcal{M}at}(\mathcal{F})$.
- ▶ Montrer que M est inversible.

Méthode Calculer une matrice

Pour déterminer $\underset{\mathcal{B}, \mathcal{C}}{\mathcal{M}at}(u)$, il faut donc :

1. calculer les vecteurs de $u(\mathcal{B})$, i.e. les $u(e_j)$ pour tout $j \in \llbracket 1, p \rrbracket$.
2. Pour tout $i \in \llbracket 1, p \rrbracket$, calculer les coordonnées de $u(e_j)$ dans la base image \mathcal{C} , i.e. chercher les $\lambda_{i,j}$ tels que :

$$u(e_j) = \sum_{i=1}^n \lambda_{i,j} f_i,$$

où (f_1, \dots, f_n) est une base de \mathcal{F} . Ces coordonnées existent bien, et sont uniques, puisque \mathcal{F} est une base de F.

3. Conclure : la j -ème colonne de $\underset{\mathcal{B}, \mathcal{C}}{\mathcal{M}at}(u)$ sera donc : $\begin{pmatrix} \lambda_{1,j} \\ \vdots \\ \lambda_{n,j} \end{pmatrix}$.

Méthode Montrer que deux matrices sont semblables

Soient $A, B \in \mathfrak{M}_{n,n}(\mathbf{K})$. Alors pour montrer $A \sim B$:

- ▶ (Si $n = 2$) on cherche P sous la forme $P = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathfrak{M}_{2,2}(\mathbf{R})$ de sorte que $A \cdot P = P \cdot B$. On peut utiliser aussi la méthode ci-après.
- ▶ (Si $n = 3$) une seule méthode est à privilégier : celle du changement de

base. Notons

$$B = \begin{pmatrix} b_{1,1} & \cdots & b_{1,n} \\ \vdots & & \vdots \\ b_{n,1} & \cdots & b_{n,n} \end{pmatrix}.$$

L'idée est de chercher $P = \left(e_1 \mid \cdots \mid e_n \right)$ inversible, de sorte que

$$A = P \cdot B \cdot P^{-1} = P \cdot \begin{pmatrix} \overset{A \cdot e_1}{b_{1,1}} & \cdots & \overset{A \cdot e_n}{b_{1,n}} \\ \vdots & & \vdots \\ b_{n,1} & \cdots & b_{n,n} \end{pmatrix} \cdot P^{-1}.$$

D'après la formule de changement de base³, les $e_i, i \in \llbracket 1, n \rrbracket$ doivent vérifier :

$$Ae_i = b_{1,i}e_1 + \cdots + b_{n,i}e_n, \quad \forall i \in \llbracket 1, n \rrbracket.$$

Cette condition est simplement un ensemble de systèmes linéaires à résoudre. Dans la pratique, ces calculs seront rapides à mener puisque la matrice B aura très souvent que quelques coefficients non nuls.

Méthode Opérations pour la recherche d'éléments propres

- ▶ Si le coefficient (3, 1) n'est pas nul :
 1. l'opération optimale à effectuer en premier pour des matrices de taille 3×3 est la permutation $L_1 \longleftrightarrow L_3$ qui permettra d'obtenir un coefficient indépendant de λ en position (1, 1). On élimine alors avec celui-ci les coefficients (2, 1), (3, 1).
 2. Ensuite, en position (2, 2), nous avons un coefficient affine en λ et que l'on souhaite utiliser en nouveau pivot afin d'éliminer le coefficient (3, 2). Pour éliminer λ en (2, 2), on peut faire une opération simple en fonction

³appliquée à $f : X \in \mathfrak{M}_{n,1}(\mathbf{K}) \longrightarrow AX$ l'endomorphisme canoniquement associé à A , $\mathcal{B} = (e_1, \dots, e_n)$, P correspondant alors à $P^{\mathcal{B} \xrightarrow{\text{can}} \mathcal{B}}$

de L_3 .

3. Un pivot indépendant de λ est alors obtenu en (2, 2), on peut alors éliminer le coefficient (3, 2).
- ▶ Si le coefficient (3, 1) est nul :
 1. on fait la permutation $L_1 \longleftrightarrow L_2$ qui permettra d'obtenir un coefficient indépendant de λ en position (1, 1). On élimine alors avec celui-ci le coefficient (3, 1).
 2. En positions (2, 2), nous avons un coefficient indépendant de λ qui sert à éliminer le coefficient (3, 2).

Méthode Éléments propres en dimension 2 l'aide du déterminant

Soit $A \in \mathfrak{M}_{2,2}(\mathbf{K})$. Alors :

$$\lambda \in \text{Spec}(A) \iff \det(A - \lambda I_2) = 0.$$

Ainsi,

1. il suffit de résoudre en λ l'équation du second degré $\det(A - \lambda I_2) = 0$.
2. On calcule ensuite $E_\lambda(A) = \ker(A - \lambda I_2)$ pour chaque solution trouvée précédemment.

Méthode Éléments propres en dimension 2 à l'aide du pivot de Gauß

- ▶ Si le coefficient (2, 1) est nul : la matrice $A - \lambda I_2$ est triangulaire, donc c'est terminé.
- ▶ Sinon, on fait la permutation $L_1 \longleftrightarrow L_2$ qui permettra d'obtenir un coefficient indépendant de λ en position (1, 1). On élimine alors avec celui-ci le coefficient (2, 1). La matrice $A - \lambda I_2$ est alors elle aussi échelonnée.

Méthode Calcul d'éléments propres d'une avec la définition

Soit $A \in \mathfrak{M}_{n,n}(\mathbf{K})$. Alors pour trouver les éléments propres de A , on peut, si elle possède beaucoup de zéros :

- ▶ considérer le système $AX = \lambda X$,
- ▶ le résoudre en les $(\lambda, X) \in \mathbf{K} \times \left(\mathfrak{M}_{n,1}(\mathbf{K}) \setminus \{0_{\mathfrak{M}_{n,1}(\mathbf{K})}\} \right)$.



► Conclure.



Méthode Valeur propre et somme sur chaque ligne constante

Lorsque la somme des coefficients sur les lignes d'une matrice est constante égale à $\lambda \in \mathbf{K}$,

le vecteur $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ est un vecteur propre de valeur propre associée λ .



Méthode Diagonalisabilité des matrices possédant une unique valeur propre – Important

- Si une matrice⁴ A possède une unique valeur propre λ , alors $A = PDP^{-1}$, où $D = \lambda I$. On en déduit que $A = \lambda I$.
- **(Conséquence)** Si A est une matrice possédant une unique valeur propre $\lambda \in \mathbf{K}$ et n'est pas égale à λI_n , alors elle **n'est pas** diagonalisable.



Méthode Comment trouver les puissances d'une matrice diagonalisable ?

1. Diagonaliser la matrice A . On obtient $A = PDP^{-1}$ avec D diagonale et P inversible.
2. Chercher D^n pour tout $n \in \mathbf{N}$.
3. $A^n = PD^nP^{-1}$, que l'on montre généralement par récurrence, on en déduit A^n .



Méthode Développement d'une norme de somme au carré

Soit $\|x + y\|^2$ avec $x, y \in E$.

1. Écrire la quantité en fonction du produit scalaire : $\|x + y\|^2 = \langle x + y | x + y \rangle$.
2. Développer en utilisant la bilinéarité du produit scalaire.

⁴Le principe est naturellement le même pour un endomorphisme, en passant par une matrice associée qui aura elle aussi une unique valeur propre.



Méthode Majorer des sommes avec CAUCHY-SCHWARZ

Interpréter une somme comme un produit scalaire. On retiendra l'exemple très classique suivant : si $(x_1, \dots, x_n) \in \mathbf{R}^n$,

$$\left| \sum_{i=1}^n x_i \right| = |\langle (1, \dots, 1) | (x_1, \dots, x_n) \rangle| \leq \sqrt{n} \sum_{i=1}^n x_i^2.$$



Méthode Orthonormalisation d'une famille de deux vecteurs

Lors (e_1, e_2) est une base quelconque (non forcément orthonormée) de F , une version orthonormée est obtenue en posant :

$$\begin{aligned} f_1 &= \frac{e_1}{\|e_1\|}, \\ f_2 &= \frac{e_2 - p_{\text{Vect}(e_1)}(e_2)}{\|e_2 - p_{\text{Vect}(e_1)}(e_2)\|} \\ &= \frac{e_2 - \frac{1}{\|e_1\|^2} \langle x | e_1 \rangle e_1}{\left\| e_2 - \frac{1}{\|e_1\|^2} \langle x | e_1 \rangle e_1 \right\|}. \end{aligned}$$

Le vecteur $e_2 - p_{\text{Vect}(e_1)}(e_2)$ est une version « redressée » du vecteur e_2 , orthogonale à e_1 (ou f_1).



Méthode Calcul d'une projection orthogonale


Deux méthodes pour calculer un projeté orthogonal sur F :

1. **(En utilisant la définition)** Si on ne connaît pas une base orthonormale de F , mais une famille génératrice (e_1, \dots, e_q) de F : soit $x \in E$, alors on cherche l'unique vecteur $p_F(x)$ vérifiant

$$p_F(x) \in F, \quad x - p_F(x) \in F^\perp.$$

On caractérise $x' = p_F(x)$ de la manière suivante :

$$x' = p_F(x) \iff x' \in F, \quad \text{et} \quad \langle x - x' | e_i \rangle = 0, \quad \forall i \in \llbracket 1, q \rrbracket.$$

 **2. (En utilisant la formule dans une base orthonormale)** Si une base **orthonormale** (e_1, \dots, e_q) de F est connue, alors :

$$\forall u \in E, \quad p_F(u) = \sum_{i=1}^q \langle u | e_i \rangle e_i.$$

Lorsque F est de dimension 1 ou 2, on peut se ramener facilement à une base orthonormée et donc utiliser **2**, en dimension 3 ou plus : si aucune base orthonormée n'est donnée, on utilisera **1**.

2. EN ANALYSE

 **Méthode Nier l'existence d'une limite : fonctions d'une variable**

Exhiber deux suites (u_n) et (v_n) d'éléments de I tendant vers a , mais telles que :

$$((f(u_n))_n \text{ et } (f(v_n))_n \text{ ne convergent pas vers la même limite.}$$

 **Méthode Continuité et permutation de limites**


Il faut surtout retenir la caractérisation séquentielle de la manière suivante : si f est continue, et avec les mêmes notations que *supra*,

$$f\left(\lim_{n \rightarrow \infty} (\dots)\right) = \lim_{n \rightarrow \infty} f(\dots).$$

 **Méthode Utilisation du théorème de la bijection ou du théorème des valeurs intermédiaires**

On souhaite justifier l'existence et l'unicité éventuelle d'une solution $x \in \mathbf{R}$ à l'équation $f(x) = \alpha$ avec $\alpha \in \mathbf{R}$.

1. Si l'unicité n'est pas souhaitée : on applique le théorème des valeurs intermédiaires.
2. Si l'unicité est souhaitée : on applique le théorème de la bijection sur un in-

 tervalle I tel que $\alpha \in f(I)$.

Notez également que l'équation $f(x) = \alpha$ est équivalente à $g(x) = \alpha$ avec $g = f - \text{Id}$ et $\alpha = 0$.

 **Méthode Translation x_0 dans un développement limité**

Si $x_0 \neq 0$, alors la recherche d'un $\text{DL}_n(x_0)$ pour une fonction f se fera en se ramenant au voisinage de 0 par le changement de variable « $h = x - x_0$ ». Plus précisément,

1. considérer $g : h \mapsto f(x_0 + h)$,
2. faire un $\text{DL}_n(0)$ de g : on obtient une expression du type $g(h) = R_n(h) + o(h^n)$ (avec R_n fonction polynomiale de degré n définie au voisinage de zéro),
3. un $\text{DL}_n(x_0)$ de f est alors : $f(x) =_{x \rightarrow x_0} R_n(x - x_0) + o((x - x_0)^n)$.

 **Méthode Prolongement**

Pour montrer qu'une fonction est prolongeable par continuité et/ou dérivable en un point x_0 , il suffit de faire un développement limité d'ordre 0/1 au voisinage de ce point.

 **Méthode Recherche d'équivalent/Signe local d'une fonction**


On lit :

1. un équivalent de fonction au voisinage de x_0 en regardant le **premier terme non nul** du $\text{DL}_n(x_0)$ pour un certain entier n assez grand (de sorte que le développement limité ne soit pas nul).
2. le signe de f au voisinage de x_0 grâce à celui de $a_0(x - x_0)^p$. Cela dépend donc de la parité de p notamment.

 **Méthode Nier l'existence d'une limite pour les fonctions de plusieurs variables**

Si $f : P \subset \mathbf{R}^n \rightarrow \mathbf{R}$ et $A \in P$. Alors si f admet une limite $\ell \in \overline{\mathbf{R}}$ en A , pour toute fonction $\varphi : Q \subset \mathbf{R}^n \rightarrow P$ (avec $A \in \overline{Q}$) telle que $\lim_B \varphi = A$ pour $B \in \overline{Q}$, on a :

$$\lim_{X \rightarrow B} f \circ \varphi(X) = \ell.$$

 Ainsi, si l'on trouve deux telles fonctions $\varphi_1 : Q_1 \subset \mathbf{R}^n \longrightarrow P$ (associée à un point $B_1 \in \overline{Q_1}$) et $\varphi_2 : Q_2 \subset \mathbf{R}^n \longrightarrow P$ (associée à un point $B_2 \in \overline{Q_2}$), vérifiant :


$$\lim_{X \rightarrow B_1} f \circ \varphi_1(X) \neq \lim_{X \rightarrow B_2} f \circ \varphi_2(X),$$

la fonction n'admet donc pas de limite en A. On dit que l'on a trouvé des « chemins le long desquels la fonction ne converge pas vers la même limite ».

 **Méthode Trouver la monotonie d'une suite**

1. On étudie le signe de la différence $u_{n+1} - u_n$ pour tout entier n .
2. Pour les suites qui ne s'annulent pas, on peut également comparer le quotient $\frac{u_{n+1}}{u_n}$ à 1 pour tout entier $n \in \mathbf{N}$. Cette seconde version est bien adaptée aux produits.

Méthode Exprimer les solutions d'un système de suites récurrentes linéaires – Cas diagonalisable

 On considère trois suites $(u_n), (v_n), (w_n)$ telles que

$$\forall n \in \mathbf{N}, \begin{pmatrix} u_{n+1} \\ v_{n+1} \\ w_{n+1} \end{pmatrix} = A \begin{pmatrix} u_n \\ v_n \\ w_n \end{pmatrix}, \quad \text{avec } A \in \mathfrak{M}_{3,3}(\mathbf{R}) \text{ supposée diagonalisable.}$$

Alors pour exprimer les suites $(u_n), (v_n), (w_n)$ en fonction de n , on :

1. diagonalise A, i.e. on cherche P inversible, D diagonale telle que $A = PDP^{-1}$,
2. on pose $X_n = \begin{pmatrix} u_n \\ v_n \\ w_n \end{pmatrix}$, et $Y_n = P^{-1}X_n$ pour tout $n \in \mathbf{N}$. On montre que $Y_{n+1} = DY_n$ pour tout $n \in \mathbf{N}$.
3. On peut alors exprimer Y_n en fonction de $n \in \mathbf{N}$.
4. Puis on conclut à l'aide de P : $X_n = PY_n$.

 **Méthode Plan d'étude d'une suite implicite**

1. Établir l'existence et l'unicité de la suite grâce au théorème de la bijection.
2. Chercher la monotonie en comparant $f_n(x_{n+1})$ à $f_n(x_n) = 0$ (ou $f_{n+1}(x_n)$ à $f_n(x_n) = 0$).
3. Trouver la valeur de la limite : en général on raisonne par l'absurde dans l'identité $f_n(x_n) = 0$.

 **Méthode Principe de superposition**

Lorsque le second membre b de (E_n) s'écrit comme une somme fonctions, par exemple $b = \lambda b_1 + \mu b_2$ avec b_i continue pour tout $i \in \llbracket 1, 2 \rrbracket$, $\lambda, \mu \in \mathbf{R}$, on peut appliquer le *principe de superposition*. Il s'agit de considérer les deux équations différentielles linéaires :

$$y^{(n)} + a_{n-1}^1(t)y^{(n-1)} + \dots + a_1^1(t)y' + a_0^1(t)y = b_1(t) \quad (E_n^1)$$

$$y^{(n)} + a_{n-1}^2(t)y^{(n-1)} + \dots + a_1^2(t)y' + a_0^2(t)y = b_2(t) \quad (E_n^2)$$

Si l'on a :

- ▶ déterminé une solution particulière y_1^{par} de (E_n^1) ,
- ▶ et une solution particulière y_2^{par} de (E_n^2) .

La somme $\lambda y_1^{\text{par}} + \mu y_2^{\text{par}}$ est alors une solution particulière de (E_n) . Le principe reste naturellement vrai si l'on considère une somme de n fonctions dans le second membre.

 **Méthode Variation de la constante**

Chercher y^{par} sous la forme $t \in I \longrightarrow C(t)e^{-A(t)}$, où la fonction $C : I \longrightarrow \mathbf{R}$ est dérivable et est à déterminer.

Méthode Résolution par changement de fonction inconnue

Soit (E) une équation différentielle en une fonction y que l'on ne sait pas résoudre *a priori*.

1. Soit une fonction z dépendant de y donnée par l'exercice (généralement « de la forme $z(t) = y \circ \varphi(t)$ »).
2. Calculer les dérivées successives z, z', z'', \dots (en fonction de l'ordre de l'équation différentielle en y).
3. Évaluer (E) en $\varphi(t)$ pour tout $t \in \mathbf{R}$.
4. Combiner 2) et 3) pour trouver une équation différentielle en z .

Méthode d'EULER

1. On commence par subdiviser l'intervalle $[0, \tau]$ de manière uniforme à l'aide de $N + 1$ points espacés d'un pas de $h = \frac{\tau}{N}$. Plus précisément, on pose :

$$t_0 = 0, \quad t_1 = h, \quad t_2 = 2h, \quad \dots, \quad t_k = kh, \quad \dots, \quad t_N = \tau.$$

2. **(Heuristique)** On considère que pour h petit :

$$\text{Si } y \text{ est une solution, alors } y'(t) = f(t, y(t)) \approx \frac{y(t+h) - y(t)}{h}.$$

Donc que pour toute solution y :

$$y(t+h) \simeq y(t) + hf(t, y(t)) \quad (2.1)$$

3. On définit donc une suite de points y_i satisfaisant le plus possible l'approximation précédente, en posant :

$$\forall k \in \llbracket 0, N-1 \rrbracket, \quad y_{k+1} = y_k + h \times f(t_k, y_k).$$

Méthode Résolution d'un système différentiel

Notons $Y'(t) = AY(t)$ le système avec $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix} \in \mathfrak{M}_{p,1}(\mathbf{K})$ un vecteur de fonctions

dérivables, et $A \in \mathfrak{M}_{p,p}(\mathbf{K})$ une matrice diagonalisable à coefficients constants. Pour résoudre le système, on commence par

1. diagonaliser la matrice A . On obtient $A = PDP^{-1}$ avec D diagonale et P inversible.
2. Considérer la nouvelle fonction inconnue $Z = P^{-1}Y$.
3. Établir un système en Z : puisque P est une matrice constante, on a $Z' = P^{-1}Y' = P^{-1}PDP^{-1}Y \implies Z' = DZ$.
4. Résoudre $Z' = DZ$, c'est un système d'équations différentielles indépendantes.
5. On déduit $Y = PZ$.

Méthode Série quasi-télescopique

Lorsque le terme général d'une série est de la forme $u_{n+p} - u_n$, avec $p \in \mathbf{N}$ fixé, alors on se ramène à une série télescopique en écrivant

$$\forall n \in \mathbf{N}, \quad u_{n+p} - u_n = (u_{n+p} - u_{n+p-1}) + \dots + (u_{n+1} - u_n).$$

Chacune des séries intervenant est alors télescopique.

Méthode Calcul d'une somme de série polynôme fois géométrique

Une série du type $\left(\sum_{n \geq 0} n^k q^n\right)$, avec $|q| < 1, k \geq 0$, peut s'exprimer en fonction de séries géométriques et géométriques dérivées. En effet,

- ▶ pour $k = 1$, on écrit

$$n^1 q^n = q(nq^{n-1}).$$

- ▶ Pour $k = 2$, on écrit

$$n^2 q^n = q^2(n(n-1)q^{n-2}) + q(nq^{n-1}).$$

Méthode Calcul d'une somme de série polynôme fois exponentielle

Une série du type $\left(\sum_{n \geq 0} \frac{n^k}{n!}\right)$, avec $k \in \mathbf{N}$, peut s'exprimer en fonction de séries exponentielles après changement d'indice. En effet,

- ▶ pour $k = 1$, on écrit pour $n \geq 1$

$$\frac{n^1}{n!} = \frac{1}{(n-1)!}.$$

- ▶ Pour $k = 2$, on écrit

$$\begin{aligned} \frac{n^2}{n!} &= \frac{n}{(n-1)!} \quad \text{si } n \geq 1, \\ &= \frac{n-1+1}{(n-1)!} = \frac{1}{(n-2)!} + \frac{1}{(n-1)!}. \end{aligned} \quad \left. \vphantom{\frac{n^2}{n!}} \right\} \text{ si } n \geq 2$$

Méthode Convergence des séries exponentielles convergeant vers zéro

Soit (u_n) une suite telle que $u_n \xrightarrow{n \rightarrow \infty} \infty$. Alors :

1. $n^2 e^{-u_n} \xrightarrow{n \rightarrow \infty} 0$ par croissances comparées,
2. donc pour n assez grand, $n^2 e^{-u_n} \leq 1 \implies 0 \leq e^{-u_n} \leq \frac{1}{n^2}$.
3. Comme $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converge, le théorème de comparaison donne la convergence de $\sum_{n=0}^{\infty} e^{-u_n}$.

Méthode Justifier l'existence d'une primitive

Il suffit de montrer la continuité de la fonction.

Méthode Justifier que l'intégrale d'une fonction définie sur un segment existe

Montrer que la fonction est :

- ▶ continue sur le segment,
- ▶ ou que l'on peut la prolonger en une fonction continue,
- ▶ ou encore qu'elle est continue par morceaux.

Méthode Primitiver une fonction en utilisant une intégrale

Lorsque vous avez besoin d'une technique d'intégration (intégration par parties ou changement de variable par exemple) pour primitiver une fonction $f : I \rightarrow \mathbf{R}$, choisir $a \in I$, puis calculer $\int_a^x f$ pour tout $x \in I$.

Méthode Calcul d'une intégrale à bornes variables

Soient I un intervalle, a un point de I et $f : I \rightarrow \mathbf{R}$ une fonction continue. Soient par ailleurs $u, v : J \rightarrow I$ deux fonctions définies sur J où J est un intervalle réel.

Pour étudier la dérivabilité de $x \in J \mapsto \int_{u(x)}^{v(x)} f$, on :

1. introduit une primitive de f , notée F .
2. Alors : $\int_{u(x)}^{v(x)} f = F \circ v(x) - F \circ u(x)$.
3. Justifier la dérivabilité et dériver à l'aide la formule de dérivation d'une composée.
4. On obtient *in fine* $\frac{d}{dx} \left(\int_{u(x)}^{v(x)} f \right) = f \circ v(x) v'(x) - f \circ u(x) u'(x)$ — cette formule ne doit pas être apprise par coeur, il faut savoir la retrouver en dérivant une composée.

Un point important est que le résultat ne dépend pas de F ; inutile donc de chercher à calculer F explicitement.

Méthode Quand utiliser l'intégration par parties ?

Pour intégrer un *produit* de deux fonctions, dont l'une est facile à *primitiver* et l'autre est facile à *dériver*. Exemple : une exponentielle multipliée par un polynôme.

Méthode Mise en place d'un changement de variable

En pratique, on écrit les calculs formels ci-après au brouillon⁵ :

$$\begin{cases} u = \varphi(t) \\ du = \varphi'(t) dt \\ x = \varphi(\alpha) \iff t = \alpha \\ x = \varphi(\beta) \iff t = \beta \end{cases}$$

en **justifiant que φ est de classe \mathcal{C}^1** ⁶. Ce n'est qu'après que l'on pourra écrire l'égalité :

$$\int_{\varphi(\alpha)}^{\varphi(\beta)} f(u) du = \int_{\alpha}^{\beta} f(\varphi(t))\varphi'(t) dt.$$

Méthode Essayer de se ramener à une notation sous forme de fonction puissance

Beaucoup d'expression peuvent se mettre sous la forme « $(x - a)^\alpha$ » : la formule de primitivation de cette expression est donc centrale.

Méthode Primitives de fractions rationnelles

On sait déterminer une primitive des fonctions de la forme $x \mapsto \frac{1}{ax^2 + bx + c}$ où a, b et c sont des constantes réelles et $a \neq 0$. Il suffit de discuter selon la valeur du discriminant Δ :

- si $\Delta > 0$, alors on factorise le dénominateur pour se ramener à $x \mapsto \frac{1}{(x - \alpha)(x - \beta)}$, puis on écrit la fraction comme somme de deux autres qui se primitivent avec un logarithme.
- Si $\Delta = 0$, alors on factorise le dénominateur pour se ramener à $x \mapsto \frac{1}{(x - \alpha)^2}$,
- si $\Delta < 0$, alors on met le dénominateur **sous forme canonique** et on effectue un changement de variable pour se ramener à $u \mapsto \frac{1}{u^2 + 1}$.

⁵Ils n'ont pas vocation à servir de preuve de la formule de changement de variable

⁶Le changement de variable vous sera donné dans la pratique

Méthode Calcul d'une primitive avec des fonctions trigonométriques

- Commencer par linéariser l'expression, à l'aide de nombres complexes si besoin.
- Primitiver avec les formules usuelles.

Méthode intégration par parties pour les intégrales généralisées

- Revenir à une intégrale partielle.
- Utiliser la formule déjà connue sur le segment.
- Chercher à passer à la limite.

Méthode Convergence d'intégrale à intégrande exponentielle décroissante

Soit $f : \mathbf{R} \rightarrow \mathbf{R}$ une fonction telle que $f(t) \xrightarrow{t \rightarrow \infty} \infty$. Alors :

- $t^2 e^{-f(t)} \xrightarrow{t \rightarrow \infty} 0$ par croissances comparées,
- donc pour t assez grand, $t^2 e^{-f(t)} \leq 1 \implies 0 \leq e^{-f(t)} \leq \frac{1}{t^2}$.
- Comme $\int_1^\infty \frac{1}{t^2} dt$ converge, le théorème de comparaison donne la convergence de $\int_1^\infty e^{-f(t)}$.⁷

Méthode Plan d'étude d'une intégrale

Soit f continue ou continue sauf en un nombre fini de points sur $]a, b[$ ou $[a, b[$ ou $]a, b]$ ($-\infty \leq a < b \leq +\infty$). Il s'agit de se poser les questions dans l'ordre suivant afin d'analyser l'existence de l'intégrale.

- Suis-je capable de calculer l'intégrale $\int_a^A f$ (ou $\int_A^a f$ en fonction du cas), ou les deux dans le cas de $]a, b[$ explicitement pour A dans l'intervalle d'intégration ? Si oui, on la calcule et on analyse l'existence d'une limite en A .
- Sinon, et ce sera l'immense majorité des cas, on se demande si :
 - elle est positive, dans ce cas on essaie de la majorer ou minorer par une fonction simple dont on connaît la nature de l'intégrale. On utilise éventuellement des développements limités et relations de comparai-

⁷Attention à bien démarrer l'intégrale à une borne strictement différente de zéro, 1 par exemple.



sons pour cela.

- ▶ elle n'est pas positive, on étudie la convergence absolue.



$X = k$	k_1	k_2	...
$\mathbf{P}(X = k)$	$\mathbf{P}(X = k_1)$	$\mathbf{P}(X = k_2)$...

3. EN ALÉATOIRE



Méthode Quand utiliser la formule des probabilités totales

Pour calculer la probabilité d'un évènement pour lequel on a besoin de faire une disjonction de cas. Exemple typique : deux urnes dont le tirage se fait dans l'une ou l'autre en fonction du résultat du tirage précédent, on utilise alors le résultat du tirage précédent comme système complet d'évènements puis on applique la formule des probabilités totales.



Méthode Quand utiliser la formule des probabilités composées

Pour calculer la probabilité d'un évènement qui est une intersection d'évènements non indépendants.⁸ Exemple typique : une urne dont on change les proportions de boules de chaque type étape par étape, piocher pour la première fois une boule d'un type au rang n revient à piocher que des boules des autres types jusqu'au rang $n - 1$ puis une boule du bon type au rang n .



Méthode Répondre à la question « déterminer la loi d'une variable aléatoire réelle discrète X »

1. Commencer par déterminer son support $X(\Omega)$ s'il n'est pas déjà donné *i.e.* l'ensemble de départ de \mathbf{P}_X .
2. Calculer les $\mathbf{P}(X = x)$ pour tout $x \in X(\Omega)$. Si $X(\Omega)$ est fini, il n'y a donc qu'un nombre fini de probabilités à déterminer, on peut alors les représenter sous forme d'un tableau comme ceci :



Méthode Trouver la loi d'un min ou max de variables aléatoires discrètes indépendantes

Pour le max $X = \max(X_1, \dots, X_n)$, si X_1, \dots, X_n sont par exemple à valeurs dans \mathbf{N} .

1. On calcule la fonction de répartition : $\mathbf{P}(X \leq k) = \mathbf{P}(X_1 \leq k) \dots \mathbf{P}(X_n \leq k)$ pour tout $k \in \mathbf{N}$. On invoque l'indépendance au moment adéquat.
2. On calcule ensuite $\mathbf{P}(X = k)$ en fonction de $\mathbf{P}(X \leq k)$ et $\mathbf{P}(X \leq k - 1)$, *i.e.* $\mathbf{P}(X = k) = \mathbf{P}(X \leq k) - \mathbf{P}(X \leq k - 1)$ pour tout entier k .

Pour $X = \min(X_1, \dots, X_n)$, remplacer dans **1**) \leq par \geq , puis en déduire la fonction de répartition. Étape **2**) inchangée, mais utiliser la relation $\mathbf{P}(X = k) = \mathbf{P}(X \geq k) - \mathbf{P}(X \geq k + 1)$.



Méthode Étudier l'existence d'une variance dans le cas discret

Cela revient à étudier la convergence de :

$$\sum_{x \in X(\Omega)} |x|^2 \mathbf{P}(X = x).$$



Méthode Pour montrer qu'une variable aléatoire est à densité

Deux méthodes sont possibles :

1. on revient à la définition en devinant une densité,
2. ou⁹ on montre que la fonction de répartition F_X est continue et de classe \mathcal{C}^1 sauf éventuellement en un nombre fini points. Une densité est alors obtenue en dérivant F_X là où elle est dérivable. On met en général la valeur zéro¹⁰ pour une densité là où F_X n'est pas dérivable.

⁸ sinon c'est plus facile, on utilise l'indépendance.

⁹ C'est cette méthode que nous utiliserons le plus souvent

¹⁰ valeur arbitraire

Méthode Obtenir des renseignements sur $X(\Omega)$ lorsque X est à densité

Au choix :

- ▶ on peut déduire des informations si X est une fonction de variable aléatoire (par exemple un carré, une valeur absolue, une racine carrée, *etc.*)
- ▶ on regarde une densité puis on analyse les points où elle est non nulle.
- ▶ On regarde là où la fonction de répartition est différente de zéro et un.

Méthode Montrer que l'image d'une variable aléatoire réelle à densité est à densitéSoit X une variable à densité et $Y = g(X)$ avec g une fonction au moins définie sur $X(\Omega)$ (valeur absolue, logarithme, *etc.*).

1. Deviner un ensemble contenant $Y(\Omega)$ au moyen de l'image de la fonction g .
2. Calculer la fonction de répartition de la variable aléatoire Y .
3. Calculer la fonction de répartition de Y *i.e.* $\mathbf{P}(Y \leq y)$ pour tout $y \in \mathbf{R}$ et vérifier le **Théorème ALEA.14.1**. Pour effectuer ledit calcul, faire des disjonctions de cas en y en utilisant l'ensemble trouvé en 1.

Méthode Trouver la loi d'un max de variables aléatoires à densité indépendantesPour le max $X = \max(X_1, \dots, X_n)$.

1. on calcule la fonction de répartition : $\mathbf{P}(X \leq x) = \mathbf{P}(X_1 \leq x) \dots \mathbf{P}(X_n \leq x)$ pour tout $x \in \mathbf{R}$. On invoque l'indépendance au moment adéquat.
2. On utilise ensuite l'expression de la fonction de répartition à l'aide de la densité donnée.
3. On vérifie le **Théorème ALEA.14.1** *i.e.* que $x \mapsto \mathbf{P}(X \leq x)$ est bien continue et de classe \mathcal{C}^1 sauf en un nombre fini de points. On déduit alors également une densité.

Méthode Trouver la loi d'un min de variables aléatoires à densité indépendantesPour le min $X = \min(X_1, \dots, X_n)$.

1. on calcule la fonction d'antirépartition : $\mathbf{P}(X > x) = \mathbf{P}(X_1 > x) \dots \mathbf{P}(X_n > x)$ pour tout $x \in \mathbf{R}$. On invoque l'indépendance au moment adéquat.
2. On utilise ensuite l'expression de la fonction de répartition à l'aide de la densité donnée.

3. On vérifie le **Théorème ALEA.14.1** *i.e.* que $x \mapsto 1 - \mathbf{P}(X > x) = \mathbf{P}(X \leq x)$ est bien continue et de classe \mathcal{C}^1 sauf en un nombre fini de points. On déduit alors également une densité.

Méthode Étudier l'existence d'une variance dans le cas à densitéOn étudie l'existence d'un moment d'ordre deux, *i.e.* la convergence de

$$\int_{-\infty}^{\infty} |t|^2 f_X(t) dt.$$

Méthode Lecture d'une tableSi l'on souhaite avoir, par exemple, $\Phi(0,96)$, on :


1. se place sur la ligne «0.9»,
2. se place ensuite sur la colonne «0.06».
3. On obtient alors la valeur désirée.

Méthode Répondre à la question «déterminer la loi conjointe du vecteur aléatoire X, Y »

1. Commencer par déterminer son support $(X, Y)(\Omega)$ s'il n'est pas déjà donné.
2. Calculer les $\mathbf{P}(X = x, Y = y)$ pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$. Si $X(\Omega), Y(\Omega)$ sont finis, il n'y a donc qu'un nombre fini de probabilités à déterminer, on peut les présenter sous forme d'un tableau comme ceci :

$Y = \ell \mid X = k$	k_1	k_2	...
ℓ_1	$\mathbf{P}(Y = \ell_1, X = k_1)$	$\mathbf{P}(Y = \ell_1, X = k_2)$...
\vdots			\vdots

Méthode Déterminer des probabilités d'évènements du type $\{X = Y\}, \dots, \{X_1 = \dots = X_n\}$ Par exemple, si X, Y sont deux variables aléatoires discrètes, on souhaite calculer $\mathbf{P}(X = Y)$.

-  1. Introduire le système complet associé à X (ou Y) dans l'évènement $\{X = Y\}$:

$$\{X = Y\} = \bigcup_{k \in X(\Omega)} \{X = k, Y = k\}.$$

2. On passe ensuite aux probabilités :

$$\mathbf{P}(X = Y) = \sum_{k \in X(\Omega)} \mathbf{P}(X = k, Y = k).$$

En cas d'indépendance, on écrit alors

$$\mathbf{P}(X = Y) = \sum_{k \in X(\Omega)} \mathbf{P}(X = k) \mathbf{P}(Y = k).$$

De manière générale, pour $\mathbf{P}(X_1 = \dots = X_n)$, on introduit le système complet d'évènements associé $n - 1$ variables aléatoires parmi les n .

 **Méthode Développement d'une variance de somme**

Soit la quantité **Var** $(X + Y)$ avec X, Y deux variables aléatoires discrètes ayant une variance.

1. Écrire la quantité en fonction de la covariance : $\mathbf{Var}(X + Y) = \mathbf{Cov}(X + Y, X + Y)$.
2. Développer en utilisant la bilinéarité de la covariance.


 **Méthode Déterminer la loi de la somme de deux variables aléatoires discrètes indépendantes**

1. Introduire le système complet associé à X ou Y dans l'évènement $\{X + Y = k\}$ pour tout $k \in (X + Y)(\Omega)$.
2. Utiliser la formule des probabilités totales pour avoir


$$\mathbf{P}(X + Y = k) = \sum_{\ell \in X(\Omega)} \mathbf{P}(X = \ell) \cdot \mathbf{P}(Y = k - \ell).$$

3. Pour poursuivre le calcul, regarder pour quels $\ell \in X(\Omega)$, on a :

$$\mathbf{P}(X = \ell) \cdot \mathbf{P}(Y = k - \ell) \neq 0.$$

 Puis conclure en finissant le calcul de la somme.

Méthode Déterminer une densité de la somme de deux variables aléatoires indépendantes à densité

-  1. Rappeler l'hypothèse d'indépendance puis écrire l'expression de la convolée des densités :

$$\forall z \in \mathbf{R}, \quad (f_X \star f_Y)(z) = \begin{cases} \int_{-\infty}^{+\infty} f_X(x) f_Y(z - x) dx & \text{en cas de convergence,} \\ 0 & \text{sinon.} \end{cases}$$

2. Pour poursuivre le calcul, regarder pour quels $x \in \mathbf{R}$, on a :

$$f_X(x) f_Y(z - x) \neq 0.$$

Puis conclure en réduisant l'intervalle d'intégration.

 **Méthode >_☞ Approcher une espérance (ou une probabilité) par simulation**

Soient X, X_1, \dots, X_n une suite i.i.d. d'espérance μ , et admettant une variance. Alors :


1. $\overline{X}_n \approx \mu$ pour n assez grand, et \overline{X}_n s'obtient en simulant un grand nombre de fois la loi commune aux $X_i, i \in \llbracket 1, n \rrbracket$.
2. **(Pour l'espérance)**

$$\mathbf{E}(X) \approx \overline{X}_n = \frac{\sum_{i=1}^n X_i}{n}, \quad n \text{ assez grand.}$$

On forme la moyenne de simulations X_i . Si cette moyenne ne semble pas converger, alors X n'a probablement pas d'espérance.

3. **(Pour une probabilité/fonction de répartition)** Soit I un intervalle de \mathbf{R} . Alors comme $\mathbf{P}(X \in I) = \mathbf{E}(\mathbb{1}_{\{X \in I\}})$, on peut utiliser 1) pour approcher la probabilité :

$$\mathbf{P}(X \in I) \approx \frac{\sum_{i=1}^n \mathbb{1}_{\{X_i \in I\}}}{n}, \quad n \text{ assez grand.}$$

 On compte le nombre de simulations X_i dans I et on renvoie la moyenne. Pour obtenir une approximation de $F_X(x)$, on compte le nombre de simulations « $\leq x$ ».

Méthode Pour retenir l'approximation de \overline{X}_n par une loi normale

Penser aux paramètres : $E(\overline{X}_n) = \mu$, $\text{Var}(\overline{X}_n) = \frac{\sigma^2}{n}$, « à la limite » on obtient une loi $\mathcal{N}(\mu, \frac{\sigma^2}{n})$ de mêmes paramètres.

Méthode Pour retenir l'approximation de la loi binomiale par la normale

Penser aux paramètres : $E(S_n) = np$, $\text{Var}(S_n) = np(1-p)$, « à la limite » on obtient une loi $\mathcal{N}(np, np(1-p))$ qui a même espérance/variance que S_n .

Méthode Pour retenir l'approximation de la loi de Poisson par la normale

Penser aux paramètres : $E(S_n) = n\lambda$, $\text{Var}(S_n) = n\lambda$, « à la limite » on obtient une loi $\mathcal{N}(n\lambda, n\lambda)$ de même espérance/variance que S_n .

Méthode Lecture d'une table

Si l'on souhaite avoir, par exemple, $\Phi(0,96)$, on :

1. se place sur la ligne « 0.9 »,
2. se place ensuite sur la colonne « 0.06 ».
3. On obtient alors la valeur désirée.

Méthode Résultats probabilistes pour établir un intervalle de confiance : le théorème central limite et l'inégalité de BIENAYMÉ-TCHEBYCHEV

1. (Si le « type de loi » (bernoulli, gaussienne etc.) du n -échantillon est connu)
On arrive parfois à calculer les probabilités $\mathbf{P}(\theta \in I_{X_1, \dots, X_n})$ explicitement pour n'importe quel intervalle I_{X_1, \dots, X_n} , les intervalles de confiance obtenus ne sont alors **pas asymptotiques**.

Pour des échantillons gaussiens, on a deux cas de figure :

- ▶ si σ est connue, on centre/réduit la moyenne empirique et on utilise la

propriété de stabilité de la loi normale (cf. exemple de CRUELLA).

- ▶ **[H.P]** Si σ est inconnue, on peut avoir recours à la loi de STUDENT¹¹ : voir Remarque 6 ci-après pour une définition.

2. (Si la loi de départ n'est pas connue)¹² On utilise soit :

- ▶ le théorème central limite en centrant réduisant la moyenne empirique (en approchant la variance par la version empirique si elle n'est pas connue), cela nous donne un intervalle de confiance seulement asymptotique,
- ▶ soit l'inégalité de BIENAYMÉ-TCHEBYCHEV.

Méthode Démarche générale d'un test statistique

1. Poser l'hypothèse (nulle) \mathcal{H}_0 que l'on souhaite tester.
2. Trouver un résultat de probabilité qui donne deux résultats différents selon que \mathcal{H}_0 est vraie ou non (dans le test d'adéquation *supra*, c'était le théorème central limite).
3. Donner la stratégie de décision, en fonction du résultat énoncé.

Méthode Obtenir $\Phi(x)$ pour un certain $x \in \mathbf{R}$ à l'aide d'une table

Si l'on souhaite avoir, par exemple, $\Phi(0,96)$, on :

1. se place sur la ligne « 0.9 »,
2. se place ensuite sur la colonne « 0.06 ».
3. On obtient alors la valeur désirée. Dans cet exemple, $\Phi(0,96) = 0,8315$.

Méthode Chercher $x \in \mathbf{R}$ tel que $\Phi(x) = \alpha$ à l'aide d'une table, $\alpha \in [0, 1]$

Si l'on souhaite avoir, par exemple, $x \in \mathbf{R}$ tel que $\Phi(x) = 0.975$.

1. On cherche dans la grille l'endroit où se trouve une valeur suffisamment proche de $\alpha = 0.975$.
2. Dans cet exemple, on constate que $\Phi(1.96) = 0.975$.

Si l'on souhaite avoir, par exemple, $x \in \mathbf{R}$ tel que $\Phi(x) = 0.160$.

¹¹ C'est la loi obtenue en remplaçant σ inconnue par σ_n^{cor} dans la centrée/réduite de la moyenne empirique d'un échantillon gaussien

¹² On sera dans ce contexte l'immense majorité du temps



1. En parcourant la table, on constate que 0.160 n'y apparaît.
2. On reformule alors la condition en passant au complémentaire :

$$1 - \Phi(x) = 0,84 = \Phi(-x).$$

3. On cherche donc dans la table 0.84, on trouve alors

$$0.84 = \Phi(1.00) \quad \text{donc} \quad -x = 1.00, \quad x = -1.00.$$

Chapitre ANN.20.

Annexe – Questions de cours posées au concours Agro–Véto

Cette liste n'est pas exhaustive et est vouée à être modifiée à chaque session ; elle sera ainsi complétée après chaque rapport publié par le SCAV.

Question	Réponse ou Réponse du chapitre	Commentaire
----------	--------------------------------	-------------

1. ALGÈBRE

APPLICATIONS, NOMBRES COMPLEXES, TRIGONOMÉTRIE, POLYNÔMES

SUP Définition du module d'un nombre complexe	Chapter ALG.1 $ z = \sqrt{x^2 + y^2}$ si $z = x + iy$ avec $(x, y) \in \mathbf{R}^2$	Connaitre également l'interprétation géométrique
SUP Pour $\theta \in \mathbf{R}$, exprimer $\cos(2\theta)$, $\sin(2\theta)$ en fonction de $\cos(\theta)$ et $\sin(\theta)$		

SUP Pour $\theta \in \mathbf{R}$, exprimer $\cos\left(\theta - \frac{\pi}{2}\right)$, $\sin\left(\theta - \frac{\pi}{2}\right)$ en fonction de $\cos(\theta)$ et $\sin(\theta)$

SUP Somme et produit des racines d'une équation du second degré $ax^2 + bx + c = 0$ avec $(a, b, c) \in \mathbf{R}^3$, $a \neq 0$

SUP Qu'appelle-t-on racine d'un polynôme ?

SUP Qu'appelle-t-on ordre de multiplicité d'une racine d'un polynôme ?

SUP Pour $n \in \mathbf{N}$, et $k \in \mathbf{N}$, que vaut la dérivée k -ième du polynôme X^n ?

Chapter ALG.2 Notant x_1, x_2 les deux racines, $x_1 + x_2 = \frac{-b}{a}$, $x_1 x_2 = \frac{c}{a}$

Chapter ALG.2 un $\lambda \in \mathbf{K}$ (donc \mathbf{R} ou \mathbf{C}) tel que $P(\lambda) = 0$

Chapter ALG.2 Le $m \in \mathbf{N}$ maximal tel que $(X - \lambda)^m \mid P$ mais $(X - \lambda)^{m+1} \nmid P$

Chapter ALG.2 $(X^n)^{(k)} = n(n-1) \dots (n-k+1)X^{n-k}$ si $k \leq n$ et 0 si $k > n$

à retrouver rapidement si besoin à l'aide des formules $\frac{-b \pm \sqrt{\Delta}}{2a}$

Ici on attend la **définition**, ne pas confondre avec la **caractérisation** à l'aide du polynôme dérivé

Ne pas oublier le cas particulier

SUP Définition d'une application $f : E \rightarrow F$ injective	Pour tout $x, x' \in E$, $f(x) = f(x') \implies x = x'$	Ne surtout pas parler de noyau, on ne vous dit pas que f est linéaire
SUP Définition d'une application $f : E \rightarrow F$ surjective	Pour tout $y \in F$, il existe $x \in E$ tel que $f(x) = y$	Ne surtout pas parler de rang, on ne vous dit pas que f est linéaire
SUP Pour une application $f : E \rightarrow F$ bijective, définition de f^{-1}	L'unique application $f^{-1} : F \rightarrow E$ vérifiant $f \circ f^{-1} = \text{Id}_F, f^{-1} \circ f = \text{Id}_E$	

ALGÈBRE LINÉAIRE		
Donner la définition d'une base d'un espace vectoriel	Chapter ALG.3 , <i>i.e.</i> donner la définition d'une famille libre et d'une famille génératrice finie	Attention aux quantificateurs!
Définition d'une famille libre (u_1, \dots, u_n) de vecteurs dans un espace vectoriel E	Chapter ALG.3	Attention aux quantificateurs!
Définition d'une famille génératrice dans un espace vectoriel E	Chapter ALG.3	Attention aux quantificateurs!
Donner la définition d'une base d'un espace vectoriel	Chapter ALG.3 (<i>i.e.</i> donner la définition d'une famille libre et d'une famille génératrice finie)	Attention aux quantificateurs!

Donner la définition d'un endomorphisme d'un espace vectoriel E	Chapter ALG.3 Pour tous $\lambda, \mu \in \mathbf{K}$ et $x, y \in E$, $f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$.	
Définition du noyau d'une application linéaire $f : E \rightarrow F$	Chapter ALG.3 , $\ker f = \{x \in E, f(x) = 0_F\}$. C'est un sous-espace vectoriel de E (l'ensemble de départ)	Savoir que $\ker f = \{0_E\}$ caractérise l'injectivité mais uniquement pour les applications linéaires
Énoncer le théorème du rang pour une application linéaire $f : E \rightarrow F$	Chapter ALG.3	Ne pas oublier l'hypothèse dim E finie , et ne pas mélanger ensemble de départ et d'arrivée
Condition nécessaire et suffisante pour qu'une application linéaire soit injective	Chapter ALG.3 Noyau réduit à zéro	Ici f est linéaire, mais ne surtout pas dire que c'est équivalent à la surjectivité (pas d'hypothèse de dimension finie)
Formule de changement de base pour les vecteurs	Chapter ALG.3 (de la forme « $Y = PX$ »)	Ne pas confondre avec celle pour les matrices/applications linéaires de la forme « $Y = P^{-1}XP$ »
SUP Inversibilité d'une matrice carrée 2×2	Chapter ALG.4 (condition de déterminant non nul (par exemple) et expression de l'inverse)	Connaitre aussi la définition : il existe N de même format que M telle que $MN = NM = I_2$

SUP Définition d'une matrice carrée inversible	Chapter ALG.5	Ici on attend clairement la définition générale : il existe N de même format que M telle que $MN = NM = I_n$
Matrices semblables : définition	Chapter ALG.4	Si A, B sont deux matrices, il existe P inversible telle que $A = PBP^{-1}$
Donner la définition d'une valeur propre et d'un vecteur propre pour un endomorphisme f d'un espace vectoriel E de dimension finie	Chapter ALG.5	Ne pas oublier la condition $x \neq 0_E$ dans la définition de vecteur propre
Donner la définition d'une valeur propre ainsi que d'un sous-espace propre pour une matrice $\mathfrak{M}_{3,2}(\mathbf{R})$	Chapter ALG.5	Ne pas oublier la condition $X \neq 0_{\mathfrak{M}_{n,1}(\mathbf{K})}$ dans la définition de vecteur propre
Soit u , un endomorphisme de \mathbf{R}^n . Donner une condition nécessaire et suffisante pour que u soit diagonalisable	Chapter ALG.5 n valeurs propres distinctes (condition suffisante), la somme des dimensions des espaces propres est égale à n (condition nécessaire et suffisante)	

Énoncer une condition suffisante et non nécessaire pour qu'une matrice soit diagonalisable, puis une condition nécessaire et suffisante pour qu'une matrice soit diagonalisable	Chapter ALG.5 n valeurs propres distinctes (condition suffisante), la somme des dimensions des espaces propres est égale à n (condition nécessaire et suffisante)	
Donner deux conditions suffisantes et non nécessaires de diagonalisabilité d'une matrice carrée réelle.	Chapter ALG.5 n valeurs propres distinctes (condition suffisante 1), Diagonale (condition suffisante 2, mais il y a beaucoup d'autres possibilités)	

ALGÈBRE BILINÉAIRE		
Donner la définition du produit scalaire de deux vecteurs (x_1, \dots, x_n) et (y_1, \dots, y_n) de \mathbf{R}^n	Chapter ALG.6	
Énoncer l'inégalité de CAUCHY-SCHWARZ	Chapter ALG.6 pour tout $x, y \in \mathbf{R}^n$, $ \langle x y \rangle \leq \ x\ \ y\ $	Ne pas oublier les valeurs absolues
Définition d'une base orthonormale de \mathbf{R}^n	Chapter ALG.6	Bien penser aux deux conditions : orthogonalité ET norme un pour chaque vecteur

Énoncer le théorème de PYTHAGORE dans \mathbf{R}^n	Chapter ALG.6 pour tout $x, y \in \mathbf{R}^n$, $\ x + y\ ^2 = \ x\ ^2 + \ y\ ^2$ si $x \perp y$	Ne pas oublier l'hypothèse d'orthogonalité et les carrés sur les normes
Définition de la distance d'un vecteur x de \mathbf{R}^n à un sous-espace vectoriel F de \mathbf{R}^n	Chapter ALG.6 $d(x, F) = \min_{y \in F} \ x - y\ $ ou $\left(= \ x - p_F(x)\ \right)$	La question est un peu vague, mais sans renseignement supplémentaire on attend donc plutôt la définition (que la partie entre parenthèses)

2. GÉOMÉTRIE

SUP Donner une représentation paramétrique de la droite de l'espace passant par le point A de coordonnées (x_A, y_A, z_A) et dirigée par le vecteur \vec{u} de coordonnées (a, b, c)	Un point $M(x, y, z)$ est sur ladite droite si et seulement si il existe $\lambda \in \mathbf{R}$ tel que $\overrightarrow{AM} = \lambda \vec{u}$, on déduit la représentation paramétrique $x(\lambda) = x_A + \lambda a, y(\lambda) = y_A + \lambda b, z(\lambda) = z_A + \lambda c.$	
SUP Équation cartésienne d'un plan de \mathbf{R}^3 normal au vecteur $u = (1, 2, -1)$	Le plan est d'après le cours de la forme $x + 2y - z + d = 0$ avec $d \in \mathbf{R}$	Si une condition de passage en un point avait été donnée, nous aurions eu une unique constante d , à trouver en injectant les coordonnées du point

3. ANALYSE

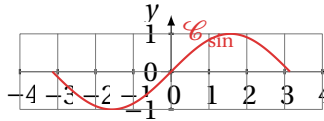
FONCTIONNELLE		
SUP Si $\alpha \in \mathbf{R}$, rappeler les solutions de l'équation $\cos(x) = \cos(\alpha)$ en $x \in \mathbf{R}$	$x = \pm\alpha + 2k\pi$ pour $k \in \mathbf{Z}$	Faire un dessin de cercle trigonométrique pour le retrouver (et l'expliquer se besoin)
SUP Si $\alpha \in \mathbf{R}$, rappeler les solutions de l'équation $\sin(x) = \sin(\alpha)$ en $x \in \mathbf{R}$	$x = \alpha + 2k\pi$ ou $x = \pi - \alpha + 2k\pi$ pour $k \in \mathbf{Z}$	Faire un dessin de cercle trigonométrique pour le retrouver (et l'expliquer se besoin)
SUP Si $\alpha \in \mathbf{R}$, rappeler les solutions de l'équation $\tan(x) = \tan(\alpha)$ en $x \in \mathbf{R} \setminus \left\{ \frac{\pi}{2} + k\pi, k \in \mathbf{Z} \right\}$	$x = \alpha + 2k\pi$ ou $x = \pi + \alpha + 2k\pi$ pour $k \in \mathbf{Z}$ pour $k \in \mathbf{Z}$	
SUP Donner la définition de la partie entière d'un réel x	Chapter ANA.7 Soit $x \in \mathbf{R}$, c'est l'unique entier $n \in \mathbf{Z}$ vérifiant $n \leq x < n + 1$	Attention aux inégalités strictes et larges : on ne les met pas au hasard. Il est important aussi de connaître le graphe (et de mentionner les points d'ouverture/fermeture de la courbe).
SUP Minorant et minimum d'une partie non vide de \mathbf{R} .	Si $A \subset \mathbf{R}$, un minorant m de A vérifie $\forall a \in A, a \geq m$, un minimum est un minorant appartenant à la partie	Savoir aussi que le plus grand minorant est ce que l'on appelle la borne inférieure de A notée $\inf A$

SUP Si f est la fonction définie sur $]0, 1[$ par : $f(x) = \sqrt{1-x}$ sur l'intervalle $]0, 1[$, déterminer l'expression de sa dérivée sur $]0, 1[$

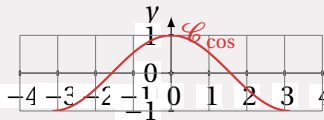
Chapter ANA.7 Soit $x \in]0, 1[$, $f'(x) = \frac{-1}{2\sqrt{1-x}}$

Eh oui, certaines questions sont très simples

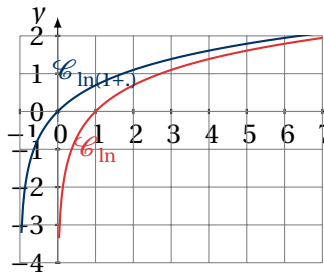
SUP Allure de la représentation graphique de la fonction sin sur l'intervalle $[-\pi, \pi]$



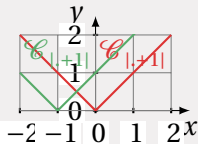
SUP Allure de la représentation graphique de la fonction cos sur l'intervalle $[-\pi, \pi]$



SUP Allure des représentations graphiques des fonctions $x \mapsto \ln(x)$ et $x \mapsto \ln(1+x)$



SUP Allure des représentations graphiques des fonctions $x \mapsto |x|$ et $x \mapsto |x+1|$

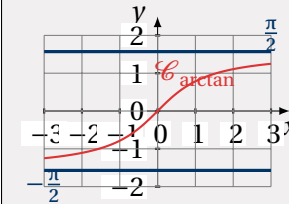


SUP Rappeler les deux expressions de la dérivée de la fonction tan

$$\tan' = 1 + \tan^2 = \frac{1}{\cos^2}$$

Savoir aussi qu'elle est dérivable sur son ensemble de définition

SUP Allure de la représentation graphique de la fonction arctan en précisant ses asymptotes en $\pm\infty$



Savoir aussi justifier l'existence à l'aide du théorème de la bijection, bien dire qu'il est nécessaire de restreindre tan à $]-\frac{\pi}{2}, \frac{\pi}{2}[$.

SUP Si f est une fonction définie sur un intervalle I , et si $a \in I$, définition de la continuité de f en a

Chapter ANA.7

SUP Dérivée d'une composée $f \circ g$

Chapter ANA.7

SUP Donner la formule de TAYLOR-YOUNG

Chapter ANA.7

Ne pas oublier l'hypothèse de classe \mathcal{C}^n

SUP Définition de la négligeabilité d'une fonction par rapport à une autre au voisinage de ∞

Chapter ANA.7 Soient deux fonctions f, g définies au voisinage de a , et telle que g ne s'annule pas au voisinage de a , $f(x) = o(g(x))$ signifie $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 0$

SUP Développement limité à l'ordre 5 au voisinage de 0 de la fonction sinus

Chapter ANA.7

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} o(x^5)$$

Vérifier à l'aide de la parité

SUP Développement limité d'ordre 5 au voisinage de 0 de la fonction cosinus	Chapter ANA.7 $\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + o(x^5)$	Vérifier à l'aide de la parité	SUP Quelles sont les solutions de l'équation différentielle $y' + a(t)y = 0$?	Chapter ANA.9 $\{x \mapsto Ce^{-A(x)}, C \in \mathbf{R}\}$ où A est une primitive de a	Dire également que A existe dès que a est continue, bien mentionner un ensemble de solutions (donc avec des accolades).
SUP Donner le développement limité au voisinage de 0, à l'ordre 4 de la fonction $x \mapsto \frac{x}{1+x}$	Chapter ANA.7 $\frac{x}{1+x} = x(1 - x + x^2 - x^3 + o(x^3)) = x - x^2 + x^3 - x^4 + o(x^4)$	Vérifier à l'aide de la parité	SUP Soit $(a, b) \in \mathbf{R}^2$. Donner l'ensemble des solutions de l'équation différentielle : $y'' + ay' + by = 0$	Chapter ANA.9 (Considérer l'équation caractéristique $x^2 + ax + b = 0$, distinguer les cas $\Delta = a^2 - 4b$ positif, nul ou négatif)	Montrer simplement que vous connaissez le résultat (donner des noms génériques pour les racines réels ou complexes)
SUP Développement limité à l'ordre 4 de $x \mapsto \ln(1+x)$ lorsque x est au voisinage de 0	Chapter ANA.7 $\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + o(x^4)$		SUP Si f est la fonction définie sur]0, 1[par : $f(x) = \sqrt{1-x}$ sur l'intervalle]0, 1[, déterminer l'expression d'une de ses primitives sur]0, 1[$\sqrt{1-x} = (1-x)^{1/2}$ se primitive en $x \mapsto -\frac{(1-x)^{3/2}}{3/2} = -\frac{2}{3}(1-x)^{3/2}$	Se ramener à des fonctions puissances permet de ne retenir qu'une seule formule de primitivation/dérivation
SUP Définition de la dérivée d'une fonction f en un point a	Chapter ANA.7 $\lim_{x \rightarrow a} \frac{f(x)-f(a)}{x-a}$ existe et est finie	Savoir expliquer géométriquement ce que cela signifie	SUP Donner la dérivée et une primitive de $t \mapsto \frac{1}{t^3}$ sur $]0, +\infty[$		Se ramener à des fonctions puissances $\frac{1}{t^3} = t^{-3}$ permet de ne retenir qu'une seule formule de primitivation/dérivation
SUP Énoncer le théorème de ROLLE	Chapter ANA.7	Ne pas oublier les hypothèses précises : continue sur le segment , dérivable sur l'intervalle ouvert			
SUP Rappeler la formule des accroissements finis	Chapter ANA.7	Ne pas oublier les hypothèses précises : continue sur le segment , dérivable sur l'intervalle ouvert			

<p>SUP Si $\alpha \in \mathbf{R}$, déterminer l'expression d'une primitive de $x \mapsto \frac{1}{x^\alpha}$ sur \mathbf{R}^{+*}</p>	<p>Chapter ANA.7 $\frac{1}{x^\alpha} = x^{-\alpha}$ se primitive en $\frac{x^{-\alpha+1}}{-\alpha+1} = \frac{1}{1-\alpha}x^{1-\alpha}$ si $\alpha \neq 1$. Si $\alpha = 1$, alors $x \mapsto \ln x$ est une primitive</p>	<p>Ne pas oublier de cas particulier sur α, et la valeur absolue dans le cas particulier</p>
<p>SUP Énoncer le théorème d'intégration par parties sur une intégrale</p>	<p>Chapter ANA.11</p>	<p>Ne pas oublier les hypothèses \mathcal{C}^1, aussi importantes que la formule</p>
<p>Donner la définition de la convergence d'une intégrale généralisée sur l'intervalle $[a, \infty[$, i.e. définition de $\int_a^\infty d(t) dt$ avec f une fonction continue</p>	<p>Chapter ANA.11 Convergence vers une limite finie de $x \mapsto \int_a^x f(t) dt$ lorsque $x \mapsto \int_a^\infty f(t) dt$</p>	<p>On note alors $\int_a^\infty f(t) dt$ la valeur de la limite</p>
<p>SUP Définition et convergence de la somme de Riemann d'une fonction f continue sur $[0, 1]$</p>	<p>$S_n(f) = \frac{1}{n} \sum_{k=0}^{n-1} f\left(0 + \frac{k(1-0)}{n}\right) \xrightarrow{n \rightarrow \infty} \int_0^1 f(t) dt$</p>	<p>Bien mentionner que f est continue si ce n'est pas précisé. S'aider d'un dessin en cas de besoin. La somme converge encore avec pour bornes 1 et n.</p>
<p>Dérivation d'une fonction de la forme $f(x(t), y(t))$, la fonction f étant de classe \mathcal{C}^1 et les fonctions x et y étant dérivables</p>	<p>Chapter ANA.8 $\frac{d(f(x(t), y(t)))}{dt} = \frac{\partial f}{\partial x}(x(t), y(t))x'(t) + \frac{\partial f}{\partial y}(x(t), y(t))y'(t)$</p>	<p>Attention au fait que x, y n'ont qu'une seule variable, donc on attend la « première » formule de la chaîne du cours</p>

Définition d'un point critique pour une fonction de deux variables

Chapter ANA.8 On suppose que la fonction admet des dérivées partielles, un point critique $c \in \mathbf{R}^n$ est un point tel que $\text{grad} f(c) = 0$, équivalent à l'annulation de toutes les dérivées partielles en c

Supposer que la fonction admet des dérivées partielles

SUITES & SÉRIES

SUP Donner la définition de $u_n \xrightarrow{n \rightarrow \infty} \ell$

Chapter ANA.9
 $\forall \varepsilon > 0, \exists n_0 \in \mathbf{N}, \forall n \geq n_0, |u_n - \ell| < \varepsilon$

Attention aux quantificateurs

SUP Énoncer la définition et le théorème des suites adjacentes

Définition : deux suites de monotonie différente dont la différence tend vers zéro. Théorème : les deux suites convergent vers la même limite

Attention au mélange entre les deux!

Pour $|q| < 1$, donner l'expression des sommes suivantes : $\sum_{n=0}^{\infty} nq^{n-1}$ et $\sum_{n=0}^{\infty} n(n-1)q^{n-2}$

Chapter ANA.10
 $\frac{1}{(1-q)^2}, \frac{2}{(1-q)^3}$

Pour rappel il s'agit de $\frac{d}{dq} \left(\frac{1}{1-q} \right), \frac{d^2}{dq^2} \left(\frac{1}{1-q} \right)$, les formules se retiennent donc facilement

DÉNOMBREMENT

SUP Pour n et k entiers naturels, donner l'expression du coefficient binomial $\binom{n}{k}$		Ne pas oublier les conventions.
SUP Formule de PASCAL sur les coefficients binomiaux		

PROBABILITÉS

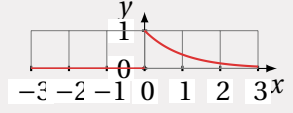
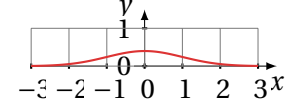
SUP Énoncer la formule de BAYES	Chapter ALEA.12	Ne pas oublier l'hypothèse de système complet d'évènements
Définition d'un système complet d'évènements	Chapter ALEA.12	On attend une version générale : en 2A les système complet d'évènements ont un nombre infini de parties
Énoncer la formule des probabilités totales	Chapter ALEA.12	Ne pas oublier l'hypothèse système complet d'évènements. Préciser aussi que les termes $\mathbf{P}(A A_i)$ $\mathbf{P}(A_i)$ sont par convention égaux à zéro lorsque l'un des A_i est de probabilité nulle

SUP Énoncer la formule des probabilités composées	Chapter ALEA.12	Ne pas oublier l'hypothèse de non-négligeabilité de l'intersection de plus grande longueur
SUP Définition de la notion d'indépendance mutuelle d'une famille finie d'évènements	Chapter ALEA.12 Si (A_1, \dots, A_n) est une famille de n évènements, l'indépendance mutuelle signifie que toute sous-famille A_{i_1}, \dots, A_{i_p} on a $\mathbf{P}(A_{i_1} \cap \dots \cap A_{i_p}) = \mathbf{P}(A_{i_1}) \times \dots \times \mathbf{P}(A_{i_p})$	Attention ce n'est pas uniquement $\mathbf{P}(A_1 \cap \dots \cap A_p) = \mathbf{P}(A_1) \times \dots \times \mathbf{P}(A_p)$

VARIABLES ET VECTEURS ALÉATOIRES

Donner la définition de la fonction de répartition d'une variable aléatoire réelle. Donner son tableau de variations	Chapter ALEA.13 $F_X : t \in \mathbf{R} \longrightarrow \mathbf{P}(X \leq t)$, c'est une fonction croissante, $\lim_{t \rightarrow -\infty} F_X(t) = 0, \lim_{t \rightarrow \infty} F_X(t) = 1$	On demande ici une allure générale du tableau, mentionnez également les limites en $\pm\infty$
Si X est une variable aléatoire admettant une espérance et une variance, et si a et b sont deux réels, rappeler l'espérance et la variance de la variable aléatoire $aX + b$	Chapters ALEA.13 et ALEA.14 et	peu importe si elles sont discrètes ou à densité, on a $\mathbf{E}(aX + b) = a\mathbf{E}(X) + b, \mathbf{Var}(aX + b) = a^2\mathbf{Var}(X)$ sous hypothèse d'existence d'espérance et variance

Énoncer le théorème de transfert dans le cadre d'une variable aléatoire réelle discrète	Chapter ALEA.13	L'hypothèse clef est la convergence absolue, qui garantit que l'ordre des termes du support n'a aucune importance dans la somme obtenue	SUP Rappeler la valeur de $\mathbf{P}(X = k)$ si X suit une loi hypergéométrique	Chapter ALEA.13	Pour retrouver rapidement l'expression, et le support, bien mémoriser l'interprétation de la loi hypergéométrique en terme de tirages sans remise dans une urne
Définition de la variance d'une variable aléatoire discrète admettant un moment d'ordre 2	$\mathbf{Var}(X) = \mathbf{E}((X - \mathbf{E}(X))^2)$ si X admet un moment d'ordre deux	Il faut être prêt à donner la version sous forme de série (application du théorème de transfert si l'examineur le demande). Possibilité de donner aussi la version KÖNIG-HUYGENS $\mathbf{Var}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2$	Quand dit-on qu'une variable aléatoire suit la loi géométrique de paramètre p ? Quelles sont alors son espérance et sa variance?	Chapter ALEA.16	Rappeler la définition de la loi (y compris le support $X(\Omega)$), l'interprétation en terme d'expérience aléatoire (temps d'attente de succès dans une répétition indépendante de schémas de BERNOULLI), puis espérance/variance
Variance de la différence de deux variables aléatoires. Cas particulier où les deux variables sont indépendantes	ALEA.13 Soient X, Y deux telles variables, $\mathbf{Var}(X - Y) = \mathbf{Var}(X) + \mathbf{Var}(-Y) = \mathbf{Var}(X) + (-1)^2 \mathbf{Var}(Y) = \mathbf{Var}(X) + \mathbf{Var}(Y)$, à la première égalité nous avons utilisé l'indépendance de X et $-Y$	La variance n'est pas linéaire	Soit X une variable suivant la loi géométrique de paramètre $\frac{1}{3}$. Donner espérance et variance	Chapter ALEA.13	
Définition de l'espérance d'une variable aléatoire X discrète à valeurs dans \mathbf{N}	Chapter ALEA.16	Ne pas oublier l'hypothèse de convergence absolue (qui garantit que la valeur de la somme ne dépend pas de l'énumération des éléments du support)	Soit X une variable suivant la loi de POISSON de paramètre $\lambda > 0$. Expliciter la loi de X , et donner son espérance et sa variance	Chapter ALEA.13	Ne pas oublier de préciser $X(\Omega)$

<p>Comment déterminer les lois marginales d'un couple (X, Y) de variables aléatoires réelles discrètes si on connaît la loi conjointe?</p>	<p>Chapter ALEA.15 On attend $\mathbf{P}(X = x) = \sum_{z \in Y(\Omega)} \mathbf{P}(X = x, Y = z)$ et $\mathbf{P}(Y = y) = \sum_{x \in X(\Omega)} \mathbf{P}(X = x, Y = y)$ d'après la formule des probabilités totales, pour tous $(x, y) \in X(\Omega) \times Y(\Omega)$</p>	<p>Savoir le rejustifier dans la pratique.</p>	<p>À quelle(s) condition(s) sur sa fonction de répartition une variable aléatoire X admet-elle une densité de probabilité? Comment détermine-t-on alors une densité de X?</p>	<p>Chapter ALEA.14 On calcule la fonction de répartition F_X, on vérifie qu'elle est continue et de classe \mathcal{C}^1 sauf en un nombre fini de points</p>	<p>Attention à ne pas confondre avec la définition de densité (qui requiert de la continuité sauf en un nombre fini de points)</p>
<p>Formule de l'espérance de $Z = u(X, Y)$ où X et Y sont des variables aléatoires discrètes à valeurs dans $\llbracket 1, n \rrbracket$ et u une fonction de \mathbf{R}^2 dans \mathbf{R}</p>	<p>Chapter ALEA.15 $\sum_{k=1}^n \sum_{\ell=1}^n u(k, \ell) \mathbf{P}(X = k, Y = \ell)$ — ou d'abord une somme en ℓ puis en k</p>	<p>Ici on suppose que le support est fini, la somme double associée est donc finie (nécessairement convergente)</p>	<p>Définition de l'espérance d'une variable aléatoire X admettant une densité f continue sur \mathbf{R}</p> <p>Fonction de répartition de l'uniforme sur $[-\pi, \pi]$</p>	<p>Chapter ALEA.14</p> <p>Chapter ALEA.14</p>	<p>Ne pas oublier l'hypothèse de convergence absolue de l'intégrale</p> <p>Attention à ne pas oublier les indicatrices $\mathbb{1}_{[-\pi, \pi]}$</p>
<p>Énoncer la définition de l'indépendance de deux variables aléatoires réelles discrètes</p>	<p>Chapter ALEA.13</p>		<p>Densité d'une loi exponentielle de paramètre λ</p>	<p>Chapter ALEA.14</p>	<p>Attention à ne pas oublier les indicatrices $\mathbb{1}_{\mathbf{R}^+}$</p>
<p>Définition de la covariance de deux variables aléatoires réelles discrètes</p>	<p>Chapter ALEA.15 Soient X, Y deux variables aléatoires discrètes possédant un moment d'ordre deux, on pose $\mathbf{Cov}(X, Y) = \mathbf{E}((X - \mathbf{E}(X))(Y - \mathbf{E}(Y)))$</p>	<p>La covariance de X et Y existe si X et Y ont chacune un moment d'ordre deux, vous pouvez aussi donner la version KÖNIG-HUYGENS : $\mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y)$</p>	<p>Allure de la représentation graphique d'une densité de la loi exponentielle de paramètre 1</p>		<p>Se tenir prêt à rappeler la densité si l'examinateur vous le demande (et l'espérance/variance)</p>
<p>Lien(s) entre l'indépendance de deux variables aléatoires discrètes et leur covariance</p>	<p>Chapter ALEA.15 L'indépendance implique la non-corrélation (covariance nulle)</p>	<p>Attention à ne pas s'emmêler les pincesaux, la réciproque est fautive (voir cours).</p>	<p>Allure de la représentation graphique d'une densité de la loi normale d'espérance 1 et de variance 1</p>		<p>Se tenir prêt à rappeler la densité si l'examinateur vous le demande (et l'espérance/variance)</p>

Densité d'une loi normale centrée réduite	Chapter ALEA.14 Une variable aléatoire de densité $x \rightarrow \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$	Avoir en tête également l'espérance/variance, ainsi que le graphe.
Densité d'une loi normale d'espérance nulle et variance 1/2	Chapter ALEA.14 Une variable aléatoire de densité $x \rightarrow \frac{\sqrt{2}}{\sqrt{2\pi}} e^{-\frac{\sqrt{2}x^2}{2}}$	Avoir en tête également l'espérance/variance, ainsi que le graphe.
Énoncer le théorème de transfert dans le cas d'une variable aléatoire admettant une densité	Chapter ALEA.14	Ne pas oublier l'hypothèse de convergence absolue de l'intégrale
Énoncer l'inégalité de MARKOV	Chapter ALEA.12	Ne pas oublier l'hypothèse de moment d'ordre 1 et de positivité . Ici on demande la version générale (avec symbole « E (X) »), donc vraie dans le cas discret ou continue
Énoncer l'inégalité de BIENAYMÉ-TCHEBYCHEV	Chapter ALEA.12	Ne pas oublier l'existence d'un moment d'ordre deux
Énoncer la loi faible des grands nombres	Chapter ALEA.16	Ne pas oublier les hypothèses de moment d'ordre deux et sur le caractère i.i.d. de la suite
Énoncer le théorème central limite	Chapter ALEA.16	Ne pas oublier les hypothèses de moment d'ordre deux et sur le caractère i.i.d. de la suite

STATISTIQUES