Nom :

Mathématiques approfondies
Cours ECG 2

Chapitres

Lycée Saint Louis 2022/2023

CHAPITRE 1

Rappels et compléments d'algèbre linéaire

Un mathématicien est une machine à transformer le café en théorèmes.

PAUL ERDÖS
Mathématicien hongrois (1913-1996).

1 Sommes de sous-espaces vectoriels

1.1 Rappels : sommes de deux s.e.v et supplémentaires

Définition 1 (somme de sous-espaces, somme directe)

Soient F et G deux sous-espaces vectoriels de E.

- **Le sous-espace somme** est défini par $\mathrm{F}+\mathrm{G}=\left\{u+v \,\middle|\, (u,v)\in \mathrm{F}\times\mathrm{G}\right\}$.
- On dit que F et G sont en **somme directe**, notée $\mathrm{F}\oplus\mathrm{G}$, si $\mathrm{F}\cap\mathrm{G}=\{0_{\mathrm{E}}\}$.

Remarque. $\mathrm{F}+\mathrm{G}$ est le plus petit sous-espace vectoriel de E (au sens de l'inclusion) contenant F et G.

Proposition 2 (unicité de la décomposition)

Soient F et G deux sous-espaces vectoriels de E. Les propriétés suivantes sont équivalentes.

i) Les sous-espaces F et G sont en somme directe.

ii) Tout vecteur $u\in\mathrm{F}+\mathrm{G}$ s'écrit de manière *unique* sous la forme :
$$u=u_{\mathrm{F}}+u_{\mathrm{G}} \quad \text{avec} \quad u_{\mathrm{F}}\in\mathrm{F},\quad u_{\mathrm{G}}\in\mathrm{G}.$$

Définition 3 (supplémentaire)

Soient E un espace vectoriel et F, G deux sous-espaces vectoriels de E.
On dit que F et G sont **supplémentaires** si tout vecteur de E se décompose de façon unique en une somme d'un vecteur de F et d'un vecteur de G. C'est-à-dire

$$\forall\, w\in\mathrm{E},\quad \exists!\,(u,v)\in\mathrm{F}\times\mathrm{G},\quad w=u+v.$$

⚠ **Attention.** Il n'y a pas unicité du supplémentaire. Il ne faut pas confondre supplémentaire et complémentaire.

Remarque. La méthode Analyse-Synthèse est particulièrement adaptée à cette définition.

Exercice 1

✦✦ Exemples

1. Dans $\mathbb{R}^3$, on pose $F = \{(x,y,z) \in \mathbb{R}^3 \mid x+y+z=0\}$, $u_1 = (1,0,0)$ et $u_2 = (1,1,1)$.
   Montrer que $\mathrm{Vect}(u_1)$ et $\mathrm{Vect}(u_2)$ sont deux supplémentaires de F dans $\mathbb{R}^3$.
2. Soit $n \in \mathbb{N}^*$. On se place dans $\mathcal{M}_n(\mathbb{R})$.
   Posons $\mathcal{S}_n$ et $\mathcal{A}_n$ les sous-espaces vectoriels des matrices symétriques et antisymétriques de taille n. Justifier que $\mathcal{S}_n$ et $\mathcal{A}_n$ sont supplémentaires.
   A *est symétrique si* ${}^t\mathrm{A} = \mathrm{A}$, *antisymétrique si* ${}^t\mathrm{A} = -\mathrm{A}$.

Proposition 4 (caractérisation des supplémentaires)

Soient F et G deux sous-espaces vectoriels de E. Les propriétés suivantes sont équivalentes.

i) Les sous-espaces F et G sont supplémentaires.

ii) $F \cap G = \{0_E\}$ et $F + G = E$.

On a donc : F et G sont supplémentaires si et seulement si $F \oplus G = E$.

Précisions en dimension finie

La preuve de l'énoncé suivant s'appuie sur le *théorème de la base incomplète* : en dimension finie, on peut compléter toute famille libre de vecteurs de E en une base de E.

Proposition 5 (existence d'un supplémentaire)

Soit E un espace vectoriel de dimension finie. Tout sous-espace vectoriel de E admet un supplémentaire.

Remarque. Concaténation de bases.
Soient $\mathcal{B}_F = (e_1, \ldots, e_p)$ et $\mathcal{B}_G = (f_1, \ldots, f_r)$ des bases respectivement de F et G. On montre que si F et G sont en somme directe, alors la famille $(e_1, \ldots, e_p, f_1, \ldots, f_r)$ est une base de $F \oplus G$.
On en déduit l'énoncé suivant.

Proposition 6 (cas de la somme directe)

Soient E un espace vectoriel de dimension finie et F, G deux sous-espaces vectoriels de E.
Si F et G sont en somme directe, alors

$$\dim(F \oplus G) = \dim(F) + \dim(G).$$

Théorème 7 (formule de Grassmann)

Soient E un espace vectoriel de dimension finie et F, G deux sous-espaces vectoriels de E. Alors,

$$\dim(F+G) = \dim(F) + \dim(G) - \dim(F \cap G).$$

Exercice 2

✦✦ Preuve

1. Soit F' un supplémentaire de $F \cap G$ dans F. Montrer que $F' \oplus G = F + G$.
2. Conclure en prouvant la formule de Grassmann.

Proposition 8 (caractérisation des supplémentaires)

Soient E un espace vectoriel de dimension finie et F, G deux sous-espaces vectoriels de E.
Les trois énoncés suivants sont équivalents.

i) F et G sont supplémentaires.

ii) $F+G=E$ et $\dim(F)+\dim(G)=\dim(E)$.

iii) $F\cap G=\{0_E\}$ et $\dim(F)+\dim(G)=\dim(E)$.

Exercice 3

✦✦ Soit $n\in\mathbb{N}\setminus\{0;1\}$. On considère $u=(a_1,a_2\dots,a_n)\in\mathbb{R}^n\setminus\{0_{\mathbb{R}^n}\}$ et

$$F=\mathrm{Vect}(u)\quad\text{et}\quad G=\left\{(x_1,x_2,\dots,x_n)\in\mathbb{R}^n\,\middle|\,\sum_{k=1}^{n}a_kx_k=0\right\}.$$

1. F et G sont des sous-espaces vectoriels de $\mathbb{R}^n$, préciser les dimensions de F et G.
2. Montrer que F et G sont supplémentaires dans $\mathbb{R}^n$.
3. Représenter dans le plan les s.e.v F et G lorsque $u=(1,1)$.

1.2 Compléments : sommes de n sous-espaces vectoriels

Sommes de sous-espaces vectoriels

Définition 9 (somme de s.e.v)

Soient $F_1,\dots,F_p$, des s.e.v d'un espace vectoriel E. On appelle **somme** de $F_1,\dots,F_p$, notée $\sum_{i=1}^{p}F_i$, l'ensemble

$$\sum_{i=1}^{p}F_i=\{\,u_1+\dots+u_p\mid\forall\,i\in[\![1,p]\!],\,u_i\in F_i\,\}.$$

Remarque. $\sum_{i=1}^{n}F_i$ est un s.e.v de E, c'est la plus petit sous-espace vectoriel contenant tous les F_i.

Exemples.

• Soit $\mathscr{F}=(e_1,e_2,\cdots,e_n)$ une famille de vecteurs de E.

$$\mathrm{vect}\left(e_1,\,e_2,\dots,\,e_p\right)=\mathrm{Vect}(e_1)+\mathrm{Vect}(e_2)+\ldots+\mathrm{Vect}(e_p).$$

• Soit $n\in\mathbb{N}\setminus\{0;1\}$. Notons $\mathscr{T}^+$, $\mathscr{T}^-$ et $\mathscr{D}$ respectivement l'ensemble des matrices triangulaires supérieures strictes, inférieures strictes et diagonales de taille (n,n). Ces trois ensembles sont des s.e.v de $\mathscr{M}_n(\mathbb{R})$ avec

$$\mathscr{M}_n(\mathbb{R})=\mathscr{T}^++\mathscr{T}^-+\mathscr{D}.$$

Exercice 4

✧ Soient $F_1,\dots,F_p$, des s.e.v d'un espace vectoriel E.

1. Montrer que $\sum_{i=1}^{p}F_i$ est bien un s.e.v de E.
2. Soit H, un sous-espace vectoriel de E tel que pour tout $i\in[\![1,n]\!]$, $F_i\subset H$.
   Montrer que $\sum_{i=1}^{p}F_i\subset H$.

Proposition 10 (somme et dimension)

Soient $F_1,\dots,F_p$, des s.e.v d'un espace vectoriel E.
Si les F_i sont tous de dimension finie, alors $\sum_{i=1}^{p}F_i$ est aussi de dimension finie avec

$$\dim\left(\sum_{i=1}^{p}F_i\right)\leqslant\sum_{i=1}^{p}\dim\left(F_i\right).$$

Sommes directes

Définition 11 (somme directe)

La somme de p sous-espaces vectoriels $F_1, F_2, \ldots, F_p$ d'un espace vectoriel E est dite **directe** si

$$\forall\,(u_1,\ldots,u_p) \in F_1 \times \cdots \times F_p, \qquad u_1 + \cdots + u_p = 0_E \quad \Rightarrow \quad u_1 = \cdots = u_p = 0_E.$$

La somme directe des s.e.v $F_1, F_2, \ldots, F_p$ est notée $F_1 \oplus F_2 \oplus \ldots \oplus F_p$.

Exemples.

- En reprenant l'exemple précédent, on a même $\mathscr{M}_n(\mathbb{R}) = \mathscr{T}^+ \oplus \mathscr{T}^- \oplus \mathscr{D}$.
- Si pour tout $i \in [\![1,p]\!]$, $F_i = \mathrm{Vect}(e_i)$ où e_i est un vecteur de E non nul alors la somme $\sum\limits_{i=1}^{p} F_i$ est directe si et seulement si la famille $(e_1, e_2, \ldots, e_p)$ est libre.

Proposition 12 (somme directe et unicité)

Soient $F_1, F_2, \ldots, F_p$ des s.e.v d'un e.v E. On a équivalence entre :

i) La somme $\sum\limits_{i=1}^{p} F_i$ est directe.

ii) Tout vecteur de $F_1 + F_2 + \ldots + F_p$ s'écrit d'une manière unique comme somme de vecteurs de $F_1, F_2, \ldots, F_p$. Autrement dit, pour tout $u \in F_1 + F_2 + \ldots + F_p$

$$\exists!\,(u_1, u_2, \ldots, u_p) \in F_1 \times F_2 \times \ldots \times F_p, \qquad u = u_1 + u_2 + \ldots + u_p.$$

Exercice 5

✦ Soient F_1, F_2 et F_3 trois s.e.v de E tels que

$$F_1 \cap F_2 = \{0_E\}, \quad F_2 \cap F_3 = \{0_E\} \quad \text{et} \quad F_1 \cap F_3 = \{0_E\}.$$

A-t-on nécessairement $F_1 \oplus F_2 \oplus F_3$?

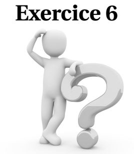
Théorème 13 (caractérisation par concaténation)

Soient $F_1, \ldots, F_p$ des s.e.v de dimension finie d'un espace vectoriel E. Soient $\mathscr{B}_1$, $\mathscr{B}_2, \ldots, \mathscr{B}_p$, des bases respectivement de F_1, F_2, …, F_p. On a équivalence entre les énoncés suivants.

i) La somme $\sum\limits_{i=1}^{p} F_i$ est directe.

ii) La famille $\mathscr{B}$, obtenue par concaténation des bases $\mathscr{B}_i$, est une base de $\sum\limits_{i=1}^{p} F_i$.

Exercice 6



✦✦ **Exemple**

Dans $\mathbb{R}^4$, montrer que $F \oplus G \oplus H = \mathbb{R}^4$ où

$$F = \{(x,y,z,t) \mid x+y=z+t=0\}, \quad G = \mathrm{Vect}(u), \quad H = \mathrm{Vect}(v) \quad \text{avec} \quad \begin{cases} u = (1,0,1,0) \\ v = (0,1,0,1). \end{cases}$$

Théorème 14 (caractérisation par les dimensions)

Soient $F_1, \ldots, F_p$ des s.e.v de dimension finie d'un espace E. On a l'équivalence entre les énoncés suivants.

i) La somme $\sum\limits_{i=1}^{p} F_i$ est directe.

ii) $\dim\left(\sum\limits_{i=1}^{p} F_i\right) = \sum\limits_{i=1}^{p} \dim(F_i)$.

Exemple. Reprenons l'exemple précédent.

$$\dim\left(\mathscr{T}^{+}\right)+\dim\left(\mathscr{T}^{-}\right)+\dim(\mathscr{D})=\frac{n(n-1)}{2}+\frac{n(n-1)}{2}+n=n^2=\dim\left(\mathscr{M}_n(\mathbb{R})\right).$$

Exercice 7

✦✦ Exemple

Dans $\mathbb{R}_3[x]$, on pose :

$$\mathrm{F}=\left\{\mathrm{P}\in\mathbb{R}_3[x]\mid \mathrm{P}(0)=\mathrm{P}(1)=\mathrm{P}(2)=0\right\},\qquad \mathrm{V}=\left\{\mathrm{P}\in\mathbb{R}_3[x]\mid \mathrm{P}(1)=\mathrm{P}(2)=0\right\},$$
$$\mathrm{G}=\left\{\mathrm{P}\in\mathbb{R}_3[x]\mid \mathrm{P}(1)=\mathrm{P}(2)=\mathrm{P}(3)=0\right\},\qquad \mathrm{H}=\left\{\mathrm{P}\in\mathbb{R}_3[x]\mid \mathrm{P}(x)=\mathrm{P}(-x)\right\}.$$

1. Préciser les dimensions de chacun de ces sous-espaces vectoriels.
2. Montrer que $\mathrm{V}\oplus\mathrm{H}=\mathbb{R}_3[x]$.
3. Justifier que $\mathrm{F}\oplus\mathrm{G}\oplus\mathrm{H}=\mathbb{R}_3[x]$.

2 Changement de bases

2.1 Rappels : liens entre matrices et applications linéaires

Si $\mathscr{B}=(e_1,\ldots,e_n)$ est une base d'un espace vectoriel E, alors pour tout vecteur u de E, il existe un unique n-uplet $(x_1,\ldots,x_n)\in\mathbb{R}^n$ tel que

$$u=\sum_{i=1}^{n}x_i\,e_i.$$

Dans ce cas, $(x_1,\ldots,x_n)$ sont les **coordonnées de u dans la base $\mathscr{B}$**. On définit la **matrice colonne des coordonnées** de u dans la base $\mathscr{B}$ par

$$\mathrm{Mat}_{\mathscr{B}}(u)=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}.$$

Généralisation. Soient $\mathscr{F}=(u_1,\ldots,u_p)$ une famille de vecteurs de E et $\mathscr{B}$ une base de E.
Pour tout $k\in[\![1,p]\!]$, on note $(x_{1,k},x_{2,k},\ldots,x_{n,k})$ les coordonnées de u_k dans la base $\mathscr{B}$.
On appelle **matrice de la famille $\mathscr{F}$ dans la base $\mathscr{B}$**, notée $\mathrm{Mat}_{\mathscr{B}}(\mathscr{F})$, la matrice dont les colonnes sont les matrices coordonnées des vecteurs de $\mathscr{F}$ dans la base $\mathscr{B}$:

$$\mathrm{Mat}_{\mathscr{B}}(\mathscr{F})=\begin{bmatrix}x_{1,1}&x_{1,2}&\cdots&x_{1,p}\\x_{2,1}&x_{2,2}&\cdots&x_{2,p}\\\vdots&\vdots&\ddots&\vdots\\x_{n,1}&x_{n,2}&\cdots&x_{n,p}\end{bmatrix}\in\mathscr{M}_{n,p}(\mathbb{R}).$$

Autrement dit, la matrice $\mathrm{Mat}_{\mathscr{B}}(\mathscr{F})$ s'obtient en concaténant les matrices colonnes des coordonnées des vecteurs de la famille.

Définition 15 (matrice d'une application linéaire)

Soient E, F deux espaces vectoriels de dimension finie et $\mathscr{B}_\mathrm{E}=(e_1,\ldots,e_p)$, $\mathscr{B}_\mathrm{F}$ deux bases respectives de E et F. **La matrice de** $\varphi\in\mathscr{L}(\mathrm{E},\mathrm{F})$ **relativement aux bases** $\mathscr{B}_\mathrm{E}$ **et** $\mathscr{B}_\mathrm{F}$ est la matrice de la famille $\left(\varphi(e_1),\ldots,\varphi(e_p)\right)$ dans la base $\mathscr{B}_\mathrm{F}$:

$$\mathrm{Mat}_{\mathscr{B}_\mathrm{F}}\left(\varphi(e_1),\ldots,\varphi(e_p)\right).$$

Elle est notée $\mathrm{Mat}_{\mathscr{B}_\mathrm{F},\mathscr{B}_\mathrm{E}}(\varphi)$.

Représentation

Notons $(e_1, \cdots, e_p)$ et $(\varepsilon_1, \cdots, \varepsilon_n)$ des bases de E et F. Une manière commode de se représenter la matrice est de placer en colonne les composantes des images de la base $(e_1, \cdots, e_p)$ par φ dans la base d'arrivée $(\varepsilon_1, \cdots, \varepsilon_n)$. Le coefficient en position (i, j) est la composante de $\varphi(e_i)$ suivant le vecteur ε_j.

$$\begin{array}{c c}
\begin{array}{ccccc} \varphi(e_1) & \varphi(e_2) & & & \varphi(e_p) \end{array} & \\
\begin{bmatrix} * & \cdots & \cdots & \cdots & * \\ \vdots & \ddots & \vdots & \iddots & \vdots \\ \vdots & \cdots & * & \cdots & \vdots \\ \vdots & \iddots & \vdots & \ddots & \vdots \\ * & \cdots & \cdots & \cdots & * \end{bmatrix} & \begin{array}{c} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \vdots \\ \varepsilon_n \end{array}
\end{array}$$

Exemples.
- La matrice de l'application identité $\mathrm{Mat}_{\mathscr{B}_E}(\mathrm{id}_E)$ est la matrice identité I_p où $p = \dim(E)$.
- La matrice de l'application nulle $u \in E \to 0_F \in F$ est $0_{n,p}$ avec $n = \dim(F)$, $p = \dim(E)$.

Théorème 16 (isomorphisme)

Soient E, F deux espaces vectoriels de dimension finie (notées respectivement p et n) et $\mathscr{B}_E$, $\mathscr{B}_F$ deux bases respectives de E et F.

L'application $\Phi : \left\{ \begin{array}{ccc} \mathscr{L}(E,F) & \to & \mathscr{M}_{n,p}(\mathbb{R}) \\ \varphi & \mapsto & \mathrm{Mat}_{\mathscr{B}_F,\mathscr{B}_E}(\varphi) \end{array} \right.$ est un isomorphisme d'espace vectoriel.

En particulier, pour tous φ, $\psi \in \mathscr{L}(E,F)$ et $\lambda \in \mathbb{R}$,

$$\mathrm{Mat}_{\mathscr{B}_F,\mathscr{B}_E}(\varphi+\psi) = \mathrm{Mat}_{\mathscr{B}_F,\mathscr{B}_E}(\varphi) + \mathrm{Mat}_{\mathscr{B}_F,\mathscr{B}_E}(\psi) \quad \text{et} \quad \mathrm{Mat}_{\mathscr{B}_F,\mathscr{B}_E}(\lambda\varphi) = \lambda \,\mathrm{Mat}_{\mathscr{B}_F,\mathscr{B}_E}(\varphi).$$

Rappel. Un isomorphisme est une application linéaire bijective.

Conséquences.
- À toute matrice $A \in \mathscr{M}_{n,p}(\mathbb{R})$ correspond une unique application linéaire de $\mathbb{R}^p$ dans $\mathbb{R}^n$ dont A soit la matrice dans les bases canoniques. Elle est appelée **l'application linéaire canoniquement associée à la matrice** A.

- En dimension finie, il ne peut avoir d'isomorphisme si la dimension de l'espace de départ ne coïncide pas avec la dimension de l'espace d'arrivée. Ainsi, l'espace vectoriel des applications linéaires de E dans F est aussi de dimension finie et

$$\dim\big(\mathscr{L}(E,F)\big) = \dim\big(\mathscr{M}_{n,p}(\mathbb{R})\big) = pn = \dim(E) \times \dim(F).$$

Théorème 17 (image d'un vecteur)

Soient E, F deux espaces vectoriels de dimension finie de bases respectives $\mathscr{B}_E$, $\mathscr{B}_F$.
Soient $\varphi \in \mathscr{L}(E,F)$ et $u \in E$.

Si on note
- U la matrice colonne des coordonnées de u dans la base $\mathscr{B}_E$.
- V la matrice colonne des coordonnées de $\varphi(u)$ dans la base $\mathscr{B}_F$.
- $A = \mathrm{Mat}_{\mathscr{B}_F,\mathscr{B}_E}(\varphi)$, la matrice de l'application φ dans les bases $\mathscr{B}_E$, $\mathscr{B}_F$,

alors $$AU = V.$$

Remarques.
- Cette relation s'écrit directement $\mathrm{Mat}_{\mathscr{B}_F}\big(\varphi(u)\big) = \mathrm{Mat}_{\mathscr{B}_F,\mathscr{B}_E}(\varphi)\,\mathrm{Mat}_{\mathscr{B}_E}(u)$.

- Cette formule s'étend par concaténation aux matrices d'une famille de vecteurs.

Soient $\varphi \in \mathscr{L}(E,F)$ et $(u_1, \cdots, u_q)$ une famille de vecteurs de E. Alors

$$\boxed{\mathrm{Mat}_{\mathscr{B}_F}\big(\varphi(u_1), \cdots, \varphi(u_q)\big) = \mathrm{Mat}_{\mathscr{B}_F,\mathscr{B}_E}(\varphi)\,\mathrm{Mat}_{\mathscr{B}_E}(u_1, \cdots, u_q)} \qquad (\bullet)$$

Théorème 18 (produit matriciel et composition d'applications)

Soient E, F, G trois espaces vectoriels de dimension finie de bases respectives $\mathscr{B}_E$, $\mathscr{B}_F$, $\mathscr{B}_G$.
Soient $\varphi \in \mathscr{L}(E,F)$ et $\psi \in \mathscr{L}(F,G)$. Alors

$$\mathrm{Mat}_{\mathscr{B}_G,\mathscr{B}_E}(\psi \circ \varphi) = \mathrm{Mat}_{\mathscr{B}_G,\mathscr{B}_F}(\psi) \cdot \mathrm{Mat}_{\mathscr{B}_F,\mathscr{B}_E}(\varphi).$$

Corollaire 19 (inversibilité et isomorphisme)

Soient E, F deux espaces vectoriels de dimension finie, $\mathscr{B}_E$, $\mathscr{B}_F$ deux bases respectives de E et F. Soit $\varphi \in \mathscr{L}(E,F)$. Les propriétés suivantes sont équivalentes.

i) L'application linéaire φ est bijective.

ii) La matrice $\mathrm{Mat}_{\mathscr{B}_F,\mathscr{B}_E}(\varphi)$ est inversible.

Dans ce cas, $$\mathrm{Mat}_{\mathscr{B}_E,\mathscr{B}_F}\left(\varphi^{-1}\right) = \mathrm{Mat}_{\mathscr{B}_F,\mathscr{B}_E}(\varphi)^{-1}.$$

Exercice 8

✦✦ *Les questions sont indépendantes*

1. Prouver cet énoncé grâce au théorème précédent.
   On pourra utiliser la caractérisation de la bijectivité.
2. Soient a, b, c, $d \in \mathbb{R}$ et $\varphi \in \mathscr{L}\left(\mathbb{R}^2\right)$ défini par $\varphi(x,y) \mapsto (ax+by, cx+dy)$.
   À quelle condition sur a, b, c et d, φ est bijective?
3. (a) Soit $n \in \mathbb{N}^*$. Déterminer A, la matrice de $\varphi : P \in \mathbb{R}_n[x] \mapsto P(x+1) \in \mathbb{R}_n[x]$ dans la base canonique.
   (b) Justifier que A est inversible et déterminer A^{-1}.

2.2 Compléments : matrices de passage

Définition 20 (cas particulier des matrices de passages)

Soient E, un espace vectoriel de dimension finie et $\mathscr{B}$, $\mathscr{C}$, deux bases.
La matrice $\mathrm{Mat}_{\mathscr{B}}(\mathscr{C})$ est appelée **matrice de passage de la base $\mathscr{B}$ à la base $\mathscr{C}$**. On la note $P_{\mathscr{B},\mathscr{C}}$.
Autrement dit, la j-ème colonne de $P_{\mathscr{B},\mathscr{C}}$ est la matrice coordonnée du j-ème vecteur de $\mathscr{C}$ dans la base $\mathscr{B}$.

Remarque. Pour $\mathscr{B}$ et $\mathscr{C}$, deux bases de E, on a

$$\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(\mathrm{id}_E) = \mathrm{Mat}_{\mathscr{B}}(\mathscr{C}) = P_{\mathscr{B},\mathscr{C}}.$$

Application. Soit u un vecteur de E dont $U_{\mathscr{B}}$ et $U_{\mathscr{C}}$ sont les matrices colonnes des coordonnées respectivement dans les bases $\mathscr{B}$ et $\mathscr{C}$ alors

$$U_{\mathscr{B}} = P_{\mathscr{B},\mathscr{C}} U_{\mathscr{C}}.$$

Remarque. Avec trois bases $\mathscr{B}$, $\mathscr{C}$ et $\mathscr{D}$, on a aussi $P_{\mathscr{B},\mathscr{D}} = P_{\mathscr{B},\mathscr{C}} P_{\mathscr{C},\mathscr{D}}$.

Exercice 9

1. ✦ **Exemple**
   On considère la base canonique $\mathscr{B} = \left(1, x, x^2\right)$ de $\mathbb{R}_2[x]$. Soient a, $b \in \mathbb{R}$ distincts. On considère les familles de polynômes

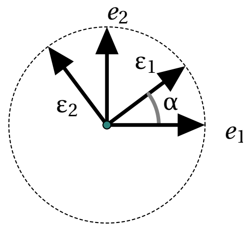
   $$\mathscr{C} = \left(1, x-a, (x-a)^2\right) \quad \text{et} \quad \mathscr{D} = \left((x-a)^2, (x-a)(x-b), (x-b)^2\right).$$

   (a) Justifier que $\mathscr{C}$ et $\mathscr{D}$ sont deux autres bases de $\mathbb{R}_2[x]$.
   (b) Calculer $P_{\mathscr{B},\mathscr{C}}$, $P_{\mathscr{B},\mathscr{D}}$ et $P_{\mathscr{C},\mathscr{D}}$. Vérifier par le calcul que $P_{\mathscr{B},\mathscr{D}} = P_{\mathscr{B},\mathscr{C}} \times P_{\mathscr{C},\mathscr{D}}$.
2. Prouver la remarque précédente dans le cas général.

Proposition 21 (inversibilité)

Soient $\mathscr{B}$, $\mathscr{C}$ deux bases d'un espace vectoriel de dimension finie E.
La matrice $P_{\mathscr{B},\mathscr{C}}$ est inversible avec $P_{\mathscr{B},\mathscr{C}} = \left(P_{\mathscr{C},\mathscr{B}}\right)^{-1}$.
Autrement dit, l'inverse de la matrice de passage de la base $\mathscr{B}$ à la base $\mathscr{C}$ est la matrice de passage de la base $\mathscr{C}$ à la base $\mathscr{B}$.

Exemple. Dans $\mathbb{R}^2$, notons $\mathscr{B}$ la base canonique et la famille $\mathscr{C}$ composée des deux vecteurs $\varepsilon_1 = \left(\cos(\alpha), \sin(\alpha)\right)$ et $\varepsilon_2 = \left(\sin(\alpha), \cos(\alpha)\right)$. La matrice inverse s'obtient en remplaçant α par $-\alpha$.



$$P_{\mathscr{B},\mathscr{C}}{}^{-1} = P_{\mathscr{C},\mathscr{B}} = \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{bmatrix}.$$

Théorème 22 (changement de bases)

Soient $\varphi \in \mathscr{L}(E)$ avec E de dimension finie, $\mathscr{B}$ et $\mathscr{C}$ deux bases de E. Alors

$$\mathrm{Mat}_{\mathscr{C}}(\varphi) = P_{\mathscr{B},\mathscr{C}}{}^{-1} \cdot \mathrm{Mat}_{\mathscr{B}}(\varphi) \cdot P_{\mathscr{B},\mathscr{C}}.$$

2.3 Compléments : matrices semblables

Définition et exemples

Définition 23 (matrices semblables)

Soient $A, B \in \mathscr{M}_n(\mathbb{R})$.
On dit que A est **semblable** à B s'il existe une matrice inversible P telle que $B = P^{-1} \cdot A \cdot P$.

Exercice 10

✦ **Vrai ou faux?** Pour tous $A, B, C \in \mathscr{M}_n(\mathbb{R})$.

1. A est semblable à A. ✓ ×
2. Si A est semblable à B, alors B est semblable à A. ✓ ×
3. Si A est semblable à B, et B est semblable à C, alors A est semblable à C. ✓ ×
4. Pour tout $p \in \mathbb{N}$, si A est semblable à B alors A^p est semblable à B^p. ✓ ×

Proposition 24 (endomorphisme et similitude)

Soient $A, B \in \mathscr{M}_n(\mathbb{R})$, deux matrices semblables et E un espace vectoriel de dimension finie.
Alors A et B représentent les matrices d'un même endomorphisme de E dans des bases différentes.
Autrement dit, il existe deux bases $\mathscr{B}$, $\mathscr{C}$ de E et $\varphi \in \mathscr{L}(E)$ tels que

$$A = \mathrm{Mat}_{\mathscr{B}}(\varphi) \quad \text{et} \quad B = \mathrm{Mat}_{\mathscr{C}}(\varphi).$$

Exemple. Les matrices suivantes sont semblables $M = \begin{bmatrix} 0 & 2 & -1 \\ 3 & -2 & 0 \\ -2 & 2 & 1 \end{bmatrix}$ et $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -4 \end{bmatrix}$

Exercice 11

✦ Notons φ l'endomorphisme de $\mathbb{R}^3$ canoniquement associé à M.

1. (a) Prouver l'existence de $u \in \mathbb{R}^3$ tel que $\mathrm{vect}(u) = \ker(\varphi - \mathrm{id}_E)$.
De même, on prouve que $v = (4,3,-2)$, $w = (-2,3,-2)$ *vérifient*
$$\mathrm{Vect}(v) = \ker(\varphi - 2\,\mathrm{id}_E) \quad et \quad \mathrm{Vect}(w) = \ker(\varphi + 4\,\mathrm{id}_E).$$
(b) Vérifier que (u, v, w) est une base de $\mathbb{R}^3$.
2. Démontrer que M et D sont semblables.

Lien avec le rang

Rappels.

• **Le rang** d'une matrice A, noté rg(A), est le rang de la famille de ses vecteurs colonnes dans $\mathcal{M}_{n,1}(\mathbb{R})$. Autrement dit, si $A = \begin{bmatrix} C_1 \, C_2 \, \dots \, C_p \end{bmatrix}$ alors
$$\mathrm{rg}(A) = \dim\big(\mathrm{Vect}(C_1, C_2, \dots, C_p)\big).$$

• **Propriétés de calculs.**

→ Le rang est invariant par transposition, autrement dit
$$\forall A \in \mathcal{M}_{n,p}(\mathbb{R}), \qquad \mathrm{rg}(A) = \mathrm{rg}({}^tA).$$

→ Le rang est invariant par les opérations élémentaires :

- L'échange de lignes ou de colonnes ($L_i \leftrightarrow L_j$ ou $C_i \leftrightarrow C_j$).
- La multiplication d'une ligne ou d'une colonne par $\lambda \in \mathbb{R}^*$ ($L_i \leftarrow \lambda L_i$ ou $C_i \leftarrow \lambda C_i$).
- L'addition d'une autre ligne ou colonne ($L_i \leftarrow L_i + L_j$ ou $C_i \leftarrow C_i + C_j$).

• Soient E, F deux espaces vectoriels de dimension finie et de bases respectives $\mathcal{B}_E$, $\mathcal{B}_F$. Soient $\varphi \in \mathcal{L}(E,F)$ et $A = \mathrm{Mat}_{\mathcal{B}_F,\mathcal{B}_E}(\varphi)$. On montre que
$$\mathrm{rg}(A) = \mathrm{rg}(\varphi).$$

Corollaire 25 (invariance du rang par similitude)

Soient $A, B \in \mathcal{M}_n(\mathbb{R})$.
Si A et B sont semblables, **alors** elles ont même rang.

Exercice 12

✧ La réciproque est fausse. Pouvez-vous donner un contre exemple ?

3 Trace d'une matrice

Définition 26 (trace)

Soit $A \in \mathcal{M}_n(\mathbb{R})$. On définit la **trace** de A, notée Tr(A), comme la somme des coefficients diagonaux de A.
Autrement dit, pour $A = (a_{i,j})_{i,j \in [\![1,n]\!]}$,
$$\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{i,i}.$$

Exercice 13

✧ Calculer $\mathrm{Tr}(\mathrm{I}_n)$, $\mathrm{Tr}(0_n)$, $\mathrm{Tr}(\mathrm{J})$ et $\mathrm{Tr}(\mathrm{K})$ où

$$\mathrm{J} = \begin{bmatrix} 1 & \dots & \dots & 1 \\ \vdots & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{bmatrix} \in \mathcal{M}_n(\mathbb{R}) \quad \text{et} \quad \mathrm{K} = \begin{bmatrix} 1 & 1 & 1 & \dots & \dots & 1 \\ 1 & 2 & 2 & \dots & \dots & 2 \\ 1 & 2 & 3 & \dots & \dots & 3 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & n-1 & n-1 \\ 1 & 2 & 3 & \dots & n-1 & n \end{bmatrix} \in \mathcal{M}_n(\mathbb{R}).$$

À partir des règles de calculs usuelles, on montre que :

Proposition 27 (forme linéaire)

L'application trace $\mathrm{Tr} : \mathcal{M}_n(\mathbb{R}) \to \mathbb{R}$ est une forme linéaire. C'est-à-dire, pour tous $\mathrm{A}, \mathrm{B} \in \mathcal{M}_n(\mathbb{R})$, $\lambda \in \mathbb{R}$,

$$\mathrm{Tr}(\mathrm{A}+\mathrm{B}) = \mathrm{Tr}(\mathrm{A}) + \mathrm{Tr}(\mathrm{B}) \quad \text{et} \quad \mathrm{Tr}(\lambda \mathrm{A}) = \lambda\,\mathrm{Tr}(\mathrm{A}).$$

Vocabulaire. Une forme linéaire est une application linéaire à valeurs dans $\mathbb{R}$.

Exercice 14

✦

1. Donner la dimension de $\ker(\mathrm{Tr})$. Préciser une base.
2. Expliciter un supplémentaire du noyau.

Proposition 28 (trace et produit)

Pour tous $\mathrm{A}, \mathrm{B} \in \mathcal{M}_n(\mathbb{R})$

$$\mathrm{Tr}(\mathrm{AB}) = \mathrm{Tr}(\mathrm{BA}).$$

Exercice 15

✦ Prouver cette proposition.
Si on note $[\mathrm{M}]_{i,j}$, *le coefficient en position* (i,j) *de la matrice* M, *on rappelle que pour* $\mathrm{A} \in \mathcal{M}_{n,p}(\mathbb{R})$ *et* $\mathrm{B} \in \mathcal{M}_{p,q}(\mathbb{R})$

$$[\mathrm{AB}]_{i,j} = \sum_{k=1}^{p} [\mathrm{A}]_{i,k} [\mathrm{B}]_{k,j}.$$

Corollaire 29 (invariance par similitude)

Soient $\mathrm{A} \in \mathcal{M}_n(\mathbb{R})$ et $\mathrm{P} \in \mathcal{M}_n(\mathbb{R})$, inversible. Alors

$$\mathrm{Tr}(\mathrm{A}) = \mathrm{Tr}\left(\mathrm{P}^{-1}\mathrm{AP}\right).$$

Remarque. Autrement dit, deux matrices semblables ont même trace. Cette propriété permet de définir la trace d'un endomorphisme (voir exercice 22).

Exercice 16

✦ Soient $\mathrm{A}, \mathrm{B} \in \mathcal{M}_n(\mathbb{R})$.

1. Que dire de A si $\mathrm{Tr}(\mathrm{A}^t\mathrm{A}) = 0$?
2. Que dire de A et B si pour tout $\mathrm{X} \in \mathcal{M}_n(\mathbb{R})$, $\mathrm{Tr}(\mathrm{AX}) = \mathrm{Tr}(\mathrm{BX})$?

4 Les espaces stables

Définition 30 (espace stable)

Soient $\varphi \in \mathscr{L}(\mathrm{E})$ et F une partie de E. On dit que F est une **partie stable** par φ si

$$\forall\, u \in \mathrm{F}, \qquad \varphi(u) \in \mathrm{F}.$$

Exemple. Soit $\varphi : f \in \mathscr{C}(\mathbb{R},\mathbb{R}) \mapsto f' \in \mathscr{C}(\mathbb{R},\mathbb{R})$. Les espaces $\mathbb{R}[x]$ et $\mathrm{Vect}(\cos,\sin)$ sont des parties stables de φ.

Exercice 17

✦ *Les questions sont indépendantes*

1. Montrer que toute somme de s.e.v stables par φ reste stable par φ.
2. Soient φ, $\psi \in \mathscr{L}(\mathrm{E})$ tels que $\varphi \circ \psi = \psi \circ \varphi$. Montrer que le noyau et l'image de φ sont stables par ψ.

Remarques.

- Soit F un s.e.v de E dont $(e_1, \ldots, e_p)$ est une famille génératrice. Alors F partie est stable par φ si et seulement si, pour tout $i \in [\![1;p]\!]$, $\varphi(e_i) \in \mathrm{F}$.
- Si F est un s.e.v stable par φ, on peut définir la restriction de φ à F par

$$\varphi|_{\mathrm{F}} : \left\{ \begin{array}{ccc} \mathrm{F} & \to & \mathrm{F} \\ u & \mapsto & \varphi(u). \end{array} \right.$$

L'application $\varphi|_{\mathrm{F}}$ définit alors un endomorphisme de F.

Exercice 18

✦ Soient $\varphi \in \mathscr{L}(\mathrm{E})$ injectif et F un sous-espace vectoriel de dimension finie de E stable par φ. Montrer que l'endomorphisme induit par φ sur F est un isomorphisme.

Exercices - TD

Exercice 19. ✦✦ On pose $A = \begin{bmatrix} -4 & 0 & 3 \\ 0 & 2 & 0 \\ -10 & 0 & 7 \end{bmatrix}$.

1. En examinant les instructions en Python suivantes, calculer les puissances de A.

Editeur

```
>>> import numpy as np
>>> A = np.array(([-4, 0, 3], [0, 2, 0], [-10, 0, 7]))
>>> P = np.array(([1, 0, 3], [0, 1, 0], [2, 0, 5]))
>>> P_inv = np.linalg.inv(P)  # calcule l'inverse de la matrice
>>> P_inv
array([[-5.,  0.,  3.],
       [ 0.,  1.,  0.],
       [ 2.,  0., -1.]])
>>> P_inv @ A @ P              # La commande @ fait un produit matriciel
array([[2., 0., 0.],
       [0., 2., 0.],
       [0., 0., 1.]])
```

2. Peut-on trouver $B \in \mathscr{M}_3(\mathbb{R})$ telle que $B^2 = A$?

Exercice 20. ✦
Peut-on trouver deux matrices $A, B \in \mathscr{M}_n(\mathbb{R})$ telles que $AB - BA = I_n$?

Exercice 21. ✦ **Somme directe dans $\mathbb{R}_n[x]$**
Soit $n \in \mathbb{N}^*$ fixé. Pour tout $i \in [\![0, n]\!]$, on définit $F_i = \left\{P \in \mathbb{R}_n[x] \mid \forall j \in [\![0, n]\!] \setminus \{i\},\ P(j) = 0\right\}$.

1. Justifier que pour tout $i \in [\![0, n]\!]$,
$$F_i = \mathrm{Vect}(L_i) \quad \text{où} \quad L_i = \prod_{\substack{j=0 \\ j \neq i}}^{n} (x - j).$$
2. Vérifier que la somme $F_0 + F_1 + \cdots + F_n$ est directe, puis l'égalité $\bigoplus_{i=0}^{n} F_i = \mathbb{R}_n[x]$.

Exercice 22. ✦ **Trace d'un endomorphisme**
Soit $\varphi \in \mathscr{L}(E)$ avec E de dimension finie. Soit $\mathscr{B}$ une base de E. On définit la trace de φ par
$$\mathrm{Tr}(\varphi) = \mathrm{Tr}\left(\mathrm{Mat}_{\mathscr{B}}(\varphi)\right).$$

1. Justifier que la trace de φ ne dépend pas du choix de la base.
2. *Exemples*
   (a) Soit $n \in \mathbb{N}^*$. Calculer la trace de φ où $\varphi : P \in \mathbb{R}_n[x] \mapsto P - P' \in \mathbb{R}_n[x]$.
   (b) On pose $\varphi : A \in \mathscr{M}_n(\mathbb{R}) \mapsto {}^tA \in \mathscr{M}_n(\mathbb{R})$.
      i. Justifier que le sous-espace des matrices symétriques de taille (n, n) (noté $\mathscr{S}_n$) et celui des matrices antisymétriques (noté $\mathscr{A}_n$) sont supplémentaires dans $\mathscr{M}_n(\mathbb{R})$. Préciser les dimensions.
      ii. En déduire la trace de φ.
   (c) Soit p, un projecteur de E de dimension n.
      i. Justifier que $\ker(p) \oplus \mathrm{Im}(p) = E$.
      ii. En déduire que la trace d'un projecteur d'un espace vectoriel de dimension finie est égale à son rang.

Exercice 23. ✦ *D'après HEC 2009 E*
Soit A un élément donné de $\mathscr{M}_2(\mathbb{R})$ non colinéaire à I_2. On note φ l'endomorphisme de $\mathbb{R}^2$ dont A est la matrice associée dans la base canonique (e_1, e_2) de $\mathbb{R}^2$ On pose : $w = e_1 + e_2$.

1. En considérant les trois vecteurs e_1, e_2 et w, montrer qu'il existe au moins un élément non nul x de $\mathbb{R}^2$ tel que la famille $(x, \varphi(x))$ soit une base de $\mathbb{R}^2$.

2. Montrer que la matrice M associée à φ dans la base $(x, \varphi(x))$ est de la forme

$$\begin{bmatrix} 0 & a \\ 1 & b \end{bmatrix}$$

où a et b sont deux réels, indépendants de la base $(x, \varphi(x))$, que l'on exprimera en fonction de $\det(\mathrm{A})$ et $\mathrm{Tr}(\mathrm{A})$.
On pourra admettre que le déterminant est un invariant de similitude.

3. En déduire que la matrice A est semblable à sa transposée ${}^t\mathrm{A}$.

Exercice 24. ✦✦✦ **Dimension du commutant** *Source : oraux HEC 2021*

Soit $\mathrm{A} \in \mathscr{M}_3(\mathbb{R})$ une matrice non nulle telle que $\mathrm{A}^2 = 0$.

1. Montrer que la matrice A est semblable à la matrice

$$\mathrm{B} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

2. Déterminer la dimension de l'espace vectoriel $\mathscr{C}_{\mathrm{A}} = \left\{ \mathrm{M} \in \mathscr{M}_3(\mathbb{R}) \;\middle|\; \mathrm{AM} = \mathrm{MA} \right\}$.

Exercice 25. ✦✦ **Base de Lagrange**

Soient a_0, a_1, …, a_n, $n+1$ réels deux à deux distincts. Notons $\mathscr{B} = \left(1, x, \ldots, x^n\right)$ la base canonique de $\mathbb{R}_n[x]$ et $(e_0, \ldots, e_n)$, la base canonique de $\mathbb{R}^{n+1}$.

1. Montrer que l'application suivante est un isomorphisme

$$\varphi : \left\{ \begin{array}{ccc} \mathbb{R}_n[x] & \longrightarrow & \mathbb{R}^{n+1} \\ \mathrm{P} & \longmapsto & \left(\mathrm{P}(a_0), \ldots, \mathrm{P}(a_n)\right). \end{array} \right.$$

2. Pour tout $i \in [\![0; n]\!]$, on note $\mathrm{L}_i = \varphi^{-1}\left(e_i\right)$. Montrer que la famille $\mathscr{C} = (\mathscr{L}_0, \ldots, \mathscr{L}_n)$ est une base de $\mathbb{R}_n[x]$ et que

$$\forall\, \mathrm{P} \in \mathbb{R}_n[x], \qquad \mathrm{P} = \sum_{i=0}^{n} \mathrm{P}\left(a_i\right) \mathrm{L}_i.$$

3. Préciser ${\mathrm{P}_{\mathscr{B},\mathscr{C}}}^{-1}$.
Pour rappel, $\mathrm{P}_{\mathscr{B},\mathscr{C}}$ *désigne la matrice de passage de la base* $\mathscr{B}$ *à* $\mathscr{C}$.

Exercice 26. ✦✦✦ **Vecteurs cycliques et espaces stables**

Soient E, un espace vectoriel de dimension finie, un entier $n \geqslant 2$ et $\varphi \in \mathscr{L}(\mathrm{E})$. On dit qu'un vecteur $u \in \mathrm{E}$ est cyclique pour φ s'il existe un entier m non nul tel que la famille $\mathscr{B}_{u,m} = \left(u, \varphi(u), \varphi^2(u), \ldots, \varphi^{m-1}(u)\right)$ soit une famille génératrice de E.
Pour rappel, $\varphi^i(u) = \varphi \circ \varphi \circ \ldots \circ \varphi(u)$ *avec* i *compositions.*

1. Quels sont les valeurs possibles pour m ?
2. Montrer que si u est cyclique pour φ, alors $\mathscr{B}_{u,n}$ est une base de E.
3. *Application.*
On suppose dans cette question que les seuls sous-espaces vectoriels stables par φ sont $\left\{0_{\mathrm{E}}\right\}$ et E.
   (a) Démontrer que φ admet un polynôme annulateur.
   (b) Justifier que tout vecteur $u \in \mathrm{E} \setminus \{0_{\mathrm{E}}\}$ est cyclique.
   On pourra considérer m*, le degré d'un polynôme annulateur de* φ.
   (c) On note $a_0, a_1, \ldots, a_{n-1}$ les coordonnées de $\varphi^n(u)$ dans la base $\mathscr{B}_{u,n}$. Justifier que

$$\varphi^n = a_0 \cdot \mathrm{id}_{\mathrm{E}} + a_1 \cdot \varphi + \ldots + a_{n-1} \cdot \varphi^{n-1}.$$

   (d) Donner la matrice de φ dans la base $\mathscr{B}_{u,n}$ à l'aide des réels a_i.

Exercice 27. ✦ Écrire un programme qui compte le nombre de matrices non-inversibles parmi les matrices de tailles $(25, 25)$:

$$\mathrm{M}(a, b) = \begin{bmatrix} a & b & \cdots & \cdots & b \\ b & a & \ddots & \cdots & b \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a & b \\ b & b & \cdots & b & a \end{bmatrix} \quad \text{avec} \quad -50 \leqslant a, b \leqslant 50.$$

Indication. On pourra utiliser les commandes `ones([n,p])` qui renvoie une matrice de $\mathscr{M}_{n,p}(\mathbb{R})$ ne contenant que des 1, `eye(n)` pour la matrice I_n et `np.linalg.matrix_rank(A)` pour le rang de la matrice A.

CHAPITRE 2

Valeurs propres et vecteurs propres

Une idée qui ne peut servir qu'une seule fois est une astuce. Sinon, elle devient une méthode.

GEORGE PÓLYA
Mathématicien américain d'origine hongroise et suisse (1887-1985).

1 Rappels : polynômes d'endomorphismes et de matrices

Rappelons que pour un endomorphisme φ de E, les applications $\varphi^2 = \varphi \circ \varphi$, $\varphi^3 = \varphi \circ \varphi \circ \varphi$ ... sont parfaitement définies et linéaires. Les **puissances** de φ sont les applications :

$$\varphi^0 = \mathrm{id}_E \quad \text{et} \quad \text{pour tout } n \in \mathbb{N}^*, \quad \varphi^n = \underbrace{\varphi \circ \cdots \circ \varphi}_{n \text{ compositions}} .$$

Remarque. Comme pour les matrices, il existe une version de la formule du binôme de Newton dans le cas des endomorphismes. Soient $\varphi, \psi \in \mathscr{L}(E)$ qui *commutent* ($\varphi \circ \psi = \psi \circ \varphi$). Alors pour tout entier naturel p,

$$(\varphi + \psi)^p = (\varphi + \psi) \circ (\varphi + \psi) \circ \cdots \circ (\varphi + \psi) = \sum_{i=0}^{p} \binom{p}{i} \varphi^i \circ \psi^{p-i}.$$

Définition 31 (polynôme de matrice, d'endomorphisme)

Soient $P(x) = \sum_{i=0}^{p} a_i x^i \in \mathbb{R}[x]$, $A \in \mathscr{M}_n(\mathbb{R})$ et $\varphi \in \mathscr{L}(E)$.

- Le **polynôme de matrice** P(A) est défini par $\quad P(A) = \sum_{i=0}^{p} a_i A^i \in \mathscr{M}_n(\mathbb{R})$.
Un polynôme P est **annulateur** de A si $P(A) = 0_n$.

- Le **polynôme d'endomorphisme** $P(\varphi)$ est défini par $\quad P(\varphi) = \sum_{i=0}^{p} a_i \varphi^i \in \mathscr{L}(E)$.
Un polynôme P est **annulateur** de φ si $P(\varphi) = 0_{\mathscr{L}(E)}$.

Règles de calculs

On démontre que pour tous $\lambda \in \mathbb{R}$, $P, Q \in \mathbb{R}[x]$,

$$(\lambda P)(A) = \lambda P(A) \quad \text{et} \quad (P + Q)(A) = P(A) + Q(A).$$

Retenons également la propriété de commutativité :

$$(PQ)(A) = P(A)Q(A) = Q(A)P(A) = (QP)(A).$$

On a de même avec des endomorphismes

$$(\lambda P)(\varphi) = \lambda P(\varphi), \qquad (P+Q)(\varphi) = P(\varphi) + Q(\varphi) \quad \text{et} \quad (PQ)(\varphi) = P(\varphi) \circ Q(\varphi) = Q(\varphi) \circ P(\varphi) = (QP)(\varphi).$$

Exercice 28

✦ Soit $A \in \mathcal{M}_n(\mathbb{R})$. *Les questions sont indépendantes.*

1. En étudiant la famille $(I_n, A, \ldots, A^p)$ pour un entier p bien choisi, montrer que A admet un polynôme annulateur.
2. Donner un polynôme annulateur à l'endomorphisme $\varphi : P \in \mathbb{R}_n[x] \mapsto P' \in \mathbb{R}_n[x]$.

Proposition 32 (polynôme d'endomorphisme, de matrice)

Soient E un espace vectoriel de dimension finie dont $\mathcal{B}$ est une base et φ un endomorphisme de E.
Pour tout $k \in \mathbb{N}$,

$$\mathrm{Mat}_{\mathcal{B}}(\varphi)^k = \mathrm{Mat}_{\mathcal{B}}(\varphi^k).$$

Plus généralement, pour tout $P \in \mathbb{R}[x]$,

$$P\Big(\mathrm{Mat}_{\mathcal{B}}(\varphi)\Big) = \mathrm{Mat}_{\mathcal{B}}\big(P(\varphi)\big).$$

Exercice 29

✦ *Les questions sont indépendantes*

1. Soient A, B, deux matrices carrées semblables. Montrer que tout polynôme annulateur de A est annulateur de B.
2. Soient $A \in \mathcal{M}_n(\mathbb{R})$ et $P \in \mathbb{R}[x]$ annulateur de A. Montrer que si $P(0) \neq 0$ alors A est inversible.

2 Valeurs propres, vecteurs propres, cas matriciel

2.1 Premières définitions

Définition 33 (valeurs propre, vecteur propre)

Soient $A \in \mathcal{M}_n(\mathbb{R})$, $\lambda \in \mathbb{R}$ et $X \in \mathcal{M}_{n,1}(\mathbb{R})$.
On dit que λ est une **valeur propre** de A et que X est un **vecteur propre** pour A associé à la valeur propre λ si

$$AX = \lambda X \quad \text{et} \quad X \neq 0_{n,1}.$$

Exemple. Posons

$$A = \begin{bmatrix} 1 & 3 & 0 \\ 0 & -2 & 0 \\ -1 & -2 & 0 \end{bmatrix} \quad \text{et} \quad X = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad Y = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad Z = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix},$$

de sorte que $$AX = 0 \cdot X, \quad AY = Y \quad \text{et} \quad AZ = -2Z.$$

Comme X, Y et Z sont des matrices colonnes non nuls, 0, 1 et -2 sont valeurs propres de A.

⚠ **Attention.** Un vecteur propre est toujours non nul.

Définition 34 (spectre)

Soit $A \in \mathcal{M}_n(\mathbb{R})$. Le **spectre** de A, noté Sp(A) est l'ensemble des valeurs propres de A.

Exemple. En reprenant l'exemple précédent, $\{0; 1; -2\} \subset \mathrm{Sp}(A)$.

Python. Voici le code pour obtenir une approximation des valeurs propres.

Editeur

```python
import numpy.linalg as al
 # On importe la sous-bibliothèque
     linalg
A=np.array([[1,3,0],[0,-2,0],[-1,-2,0]])
 # On définit la matrice A
print(al.eigvals(A))
```

Console

```
>>> # script executed

[ 0.   1.  -2.]
```

Selon ce calcul, 0, 1 et -2 sont toutes les valeurs propres de la matrice A. On conjecture donc que $\mathrm{Sp}(A) = \{0;1;-2\}$.

2.2 Caractérisation des valeurs propres

Théorème 35 (caractérisation avec l'inversibilité)

Soient $A \in \mathcal{M}_n(\mathbb{R})$ et $\lambda \in \mathbb{R}$. Les énoncés suivants sont équivalents.

i) Le réel λ est une valeur propre de A.

ii) La matrice $A - \lambda I_n$ n'est pas inversible.

Autrement dit, $\mathrm{Sp}(A)$ est l'ensemble des réels λ pour lesquels la matrice $A - \lambda I_n$ n'est pas inversible.

Remarque. En particulier, 0 est valeur propre si et seulement si, la matrice A n'est pas inversible.

Corollaire 36 (matrices triangulaires)

Soit $A \in \mathcal{M}_n(\mathbb{R})$.

Si A est une matrice triangulaire
Alors les valeurs propres de A sont les coefficients diagonaux de A.

2.3 Les sous-espaces propres $E_\lambda(A)$

Définition 37 ($E_\lambda(A)$)

Soit $A \in \mathcal{M}_n(\mathbb{R})$. Pour tout réel λ, on pose

$$E_\lambda(A) = \left\{X \in \mathcal{M}_{n,1}(\mathbb{R}) \mid AX = \lambda X\right\}.$$

En remarquant que $E_\lambda(A) = \ker\left(A - \lambda I_n\right)$, on en déduit que :

Proposition 38 (structure de $E_\lambda(A)$)

Soit $A \in \mathcal{M}_n(\mathbb{R})$. Pour tout $\lambda \in \mathbb{R}$,

- $E_\lambda(A)$ est un sous-espace vectoriel de $\mathcal{M}_{n,1}(\mathbb{R})$.
- $E_\lambda(A) \neq \{0_{n,1}\}$ si et seulement si λ est une valeur propre de A.

Remarque. La formule du rang donne alors $\dim\left(E_\lambda(A)\right) + \mathrm{rg}\left(A - \lambda I_n\right) = n$.

Vocabulaire. Si $E_\lambda(A) \neq \{0_{n,1}\}$ alors on parle de **l'espace propre associé à la valeur propre** λ. Dans ce cas,

$$1 \leqslant \dim\left(E_\lambda(A)\right) \leqslant n.$$

Exercice 30

✦ Soient $\mathrm{A}, \mathrm{B} \in \mathscr{M}_n(\mathbb{R})$. Montrer que si A et B sont semblables alors $\mathrm{Sp(A)} = \mathrm{Sp(B)}$ et les sous-espaces propres sont de même dimension.

Exemple. Le cas diagonal.
Si A est une matrice diagonale, alors les valeurs propres de A sont exactement les coefficients diagonaux de A et pour $\lambda \in \mathrm{Sp(A)}$, la dimension de $\mathrm{E}_\lambda(\mathrm{A})$ est égale au nombre de fois où λ apparait sur la diagonale de A.

Exercice 31

✦✦ **Vrai ou Faux ?**
Pour tous $\mathrm{A} \in \mathscr{M}_n(\mathbb{R})$ et $\lambda \in \mathbb{R}$,

1. Si $\lambda \in \mathrm{Sp(A)}$ alors $\lambda^2 \in \mathrm{Sp}(\mathrm{A}^2)$. ✓ ×
2. Si $\lambda^2 \in \mathrm{Sp}(\mathrm{A}^2)$ alors $\lambda \in \mathrm{Sp(A)}$. ✓ ×
3. $\mathrm{Sp(A)} = \mathrm{Sp}\left({}^t\mathrm{A}\right)$. ✓ ×
4. $\mathrm{E}_\lambda(\mathrm{A}) = \mathrm{E}_\lambda\left({}^t\mathrm{A}\right)$. ✓ ×
5. $\dim\left(\mathrm{E}_\lambda(\mathrm{A})\right) = \dim\left(\mathrm{E}_\lambda\left({}^t\mathrm{A}\right)\right)$. ✓ ×
6. Si la somme des coefficients de A sur chaque ligne vaut 1 alors 1 est valeur propre. ✓ ×
7. Si la somme des coefficients de A sur chaque colonne vaut 1 alors 1 est valeur propre. ✓ ×

3 Valeurs propres, vecteurs propres, cas des endomorphismes

3.1 Définitions et exemples

Définition 39 (valeurs propres, vecteurs propres, spectre)

Soit $\varphi \in \mathscr{L}(\mathrm{E})$.

- Soient $\lambda \in \mathbb{R}$ et $u \in \mathrm{E}$. On dit que λ est une **valeur propre** de φ et que u est un **vecteur propre** pour φ associé à la valeur propre λ si
$$\varphi(u) = \lambda u \quad \text{et} \quad u \neq 0_{\mathrm{E}}.$$
- L'ensemble des valeurs propres de φ est le **spectre de** φ, il est noté $\mathrm{Sp}(\varphi)$.

Exemple. Posons $u = (1, 0, -2,)$, $v = (0, 0, 1)$, $w = (1, 1, -2)$ et
$$\varphi : \left\{ \begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ (x, y, z) & \longmapsto & (x + 2y, 3y, 2x - 4y + 2z) \end{array} \right.$$
On a $\varphi(u) = u$, $\varphi(v) = 2v$ et $\varphi(w) = 3w$. Comme u, v et w sont non nuls, 1, 2 et 3 sont trois valeurs propres de φ. C'est-à-dire, $\{1; 2; 3\} \subset \mathrm{Sp}(\varphi)$.

Remarque. Le vecteur u (non nul) est un vecteur propre si et seulement si la droite vectorielle $\mathrm{Vect}(u)$ est stable par φ.

Exercice 32

✦ **Exemples** *Les questions sont indépendantes.*

On pose $\varphi : \left\{ \begin{array}{ccc} \mathbb{R}_n[x] & \longrightarrow & \mathbb{R}_n[x] \\ \mathrm{P} & \longmapsto & \mathrm{P}' + 2\mathrm{P} \end{array} \right.$ et $\psi : \left\{ \begin{array}{ccc} \mathscr{C}^\infty(\mathbb{R}, \mathbb{R}) & \longrightarrow & \mathscr{C}^\infty(\mathbb{R}, \mathbb{R}) \\ f & \longmapsto & f'. \end{array} \right.$

1. Étudier les valeurs propres de l'endomorphisme φ.
2. Montrer que tout réel est valeur propre de ψ.

Remarque. Soit $\mathrm{A} \in \mathscr{M}_n(\mathbb{R})$. Considérons l'endomorphisme
$$\varphi_{\mathrm{A}} : \left\{ \begin{array}{ccc} \mathscr{M}_{n,1}(\mathbb{R}) & \longrightarrow & \mathscr{M}_{n,1}(\mathbb{R}) \\ \mathrm{X} & \longmapsto & \mathrm{AX}. \end{array} \right.$$

Les valeurs propres (et vecteurs propres) de φ_A correspondent exactement aux valeurs propres (et vecteurs propres) de la matrice A.

Définition-proposition 40 ($E_\lambda(\varphi)$)

Soient $\varphi \in \mathscr{L}(E)$ et $\lambda \in \mathbb{R}$. On définit le sous-espace vectoriel $E_\lambda(\varphi)$ de E par

$$E_\lambda(\varphi) = \left\{ u \in E \,|\, \varphi(u) = \lambda u \right\} = \ker\left(\varphi - \lambda\, \mathrm{id}_E\right).$$

3.2 Précision en dimension finie

Théorème 41 (lien matrice et endomorphisme)

Soient $\varphi \in \mathscr{L}(E)$, $u \in E$ et $\mathscr{B}_E$, une base de E.

Si on note
- U la matrice colonne des coordonnées de u dans la base $\mathscr{B}$.
- $A = \mathrm{Mat}_{\mathscr{B}}(\varphi)$, la matrice de l'endomorphisme φ dans la base $\mathscr{B}$.

Alors, u est vecteur propre de l'endomorphisme φ pour la valeur propre λ si et seulement si U est vecteur propre de la matrice A pour la valeur propre λ.

Remarque. Soient A et B deux matrices semblables. Elle représente le même endomorphisme φ dans des bases différentes. On a alors

$$\mathrm{Sp}(A) = \mathrm{Sp}(\varphi) = \mathrm{Sp}(B).$$

Proposition 42 (caractérisations en dimension finie)

Soient $\varphi \in \mathscr{L}(E)$ avec E de dimension finie et $\lambda \in \mathbb{R}$. On a l'équivalence entre les énoncés suivants.

i) Le réel λ est valeur propre de φ.

ii) L'endomorphisme $\varphi - \lambda\, \mathrm{id}_E$ n'est pas injectif.

iii) L'endomorphisme $\varphi - \lambda\, \mathrm{id}_E$ n'est pas bijectif.

iv) $\mathrm{rg}\left(\varphi - \lambda\, \mathrm{id}_E\right) < \dim E$.

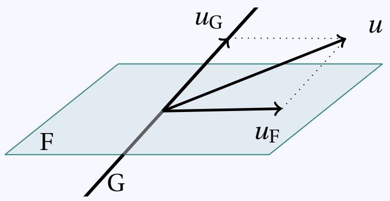
3.3 Précisions pour les endomorphismes remarquables

Les homothéties

Les projecteurs

Définition 43 (projecteur)

Soient E un espace vectoriel et F, G deux sous-espaces vectoriels supplémentaires.



Ainsi, pour tout $u \in E$, il existe une unique décomposition $u = u_F + u_G$ où $(u_F, u_G) \in F \times G$. On pose

$$p : \left\{ \begin{array}{ccc} E & \to & E \\ u & \mapsto & u_F. \end{array} \right.$$

Cette application est linéaire, elle est appelée le **projecteur** sur F parallèlement à G.

Remarque. On montre que p est un projecteur sur $\mathrm{F} = \mathrm{Im}(p)$ parallèlement à $\mathrm{G} = \ker(p)$. En particulier, le noyau et l'image d'un projecteur sont supplémentaires dans E.

$$\mathrm{Im}(p) \oplus \ker(p) = \mathrm{E}.$$

Exercice 33

✦ Savez-vous prouver cette remarque ?

Théorème 44 (caractérisation d'un projecteur)

Soit $p : \mathrm{E} \to \mathrm{E}$ une application. Les énoncés suivants sont équivalents.

i) p est un projecteur.

ii) p est linéaire et $p \circ p = p$.

Exercice 34

✦✦ Donner les valeurs propres et vecteurs propres pour un projecteur.

Les symétries

Reprenons E, un espace vectoriel et F, G deux sous-espaces vectoriels supplémentaires de E. Pour tout $u \in \mathrm{E}$, il existe un unique couple $(u_\mathrm{F}, u_\mathrm{G}) \in \mathrm{F} \times \mathrm{G}$ tel que $u = u_\mathrm{F} + u_\mathrm{G}$. On définit alors la symétrie par rapport à F parallèlement à G comme l'application :

$$s : u \in \mathrm{E} \mapsto u_\mathrm{F} - u_\mathrm{G} \in \mathrm{E}.$$

On montre que s est une application linéaire.

Exercice 35

✦✦ **Les symétries**

1. Faire un dessin similaire au cas des projecteurs pour illustrer la situation.
2. Préciser $s \circ s$. En déduire un polynôme annulateur.
3. On suppose que $s \neq \pm \mathrm{id}_\mathrm{E}$. Étudier les valeurs propres de s et préciser les espaces propres à l'aide de F et G.
4. Soit p le projecteur sur F parallèlement à G. Vérifier que $s = 2p - \mathrm{id}_\mathrm{E}$. Retrouver le résultat précédent.

4 Lien avec les polynômes et conséquences

4.1 Polynômes et valeurs propres

Théorème 45 (polynômes et valeurs propres)

Soient $\mathrm{Q} \in \mathbb{R}[x]$ et $\varphi \in \mathscr{L}(\mathrm{E})$.

Si u est un vecteur propre de φ associé à la valeur propre λ

Alors $$\mathrm{Q}(\varphi)(u) = \mathrm{Q}(\lambda) \cdot u.$$

Corollaire 46 (valeurs propres et racines)

Soient $\varphi \in \mathscr{L}(\mathrm{E})$, $\mathrm{P} \in \mathbb{R}[x]$ et $\lambda \in \mathbb{R}$.

Si
- λ est valeur propre de φ.
- P est un polynôme annulateur de φ.

Alors λ est une racine du polynôme P.

Remarque. On a des énoncés équivalents avec les matrices.
Si $\mathrm{A} \in \mathscr{M}_n(\mathbb{R})$, $\mathrm{Q} \in \mathbb{R}[x]$ et X est un vecteur propre de A associé à la valeur propre λ, alors

$$\mathrm{Q}(\mathrm{A})\mathrm{X} = \mathrm{Q}(\lambda)\mathrm{X}.$$

De plus, si Q est annulateur de A, λ est une racine de Q.

4.2 Conséquences

Rappels sur les polynômes de Lagrange.
Soient $p, n \in \mathbb{N}^*$. Soient $\lambda_1, \lambda_2, \cdots, \lambda_p$ des réels deux à deux distincts, il existe p polynômes, notés $(\mathrm{L}_i)_{i\in[\![1;p]\!]}$, tels que

$$\forall\, i,\, j \in [\![1, p]\!], \quad \mathrm{L}_i(\lambda_j) = \delta_{i,j} = \begin{cases} 0 & \text{si} \quad i \neq j, \\ 1 & \text{sinon.} \end{cases}$$

Exercice 36

✦✦

1. Prouver cet énoncé avec l'application linéaire
$$\varphi : \begin{cases} \mathbb{R}_{p-1}[x] & \longrightarrow & \mathbb{R}^p \\ \mathrm{P} & \longmapsto & \big(\mathrm{P}(\lambda_1), \mathrm{P}(\lambda_2), \ldots, \mathrm{P}(\lambda_p)\big). \end{cases}$$
2. Préciser les trois polynômes de Lagrange pour $a_1 = -1$, $a_2 = 0$ et $a_3 = 1$.
3. Donner l'expression de $\mathrm{L}_i(x)$ dans le cas général.

Théorème 47 (somme directe des espaces propres)

Soit φ un endomorphisme d'un espace vectoriel E.

Si $\lambda_1, \ldots, \lambda_p$ désignent p valeurs propres deux à deux distinctes de φ

Alors la somme $\sum\limits_{i=1}^{p} \mathrm{E}_{\lambda_i}(\varphi)$ est directe.

Corollaire 48 (en dimension finie)

Si on suppose de plus que E est de dimension finie, alors

$$\sum_{i=1}^{p} \dim\big(\mathrm{E}_{\lambda_i}(\varphi)\big) \leqslant \dim(\mathrm{E}).$$

Corollaire 49 (famille libre de vecteurs propres)

Soit φ un endomorphisme de E.

Si les vecteurs $u_1, u_2, \ldots, u_p$ sont des vecteurs propres de φ associés à des valeurs propres distinctes.

Alors la famille $\big(u_1, \ldots, u_p\big)$ est une famille libre de E.

Remarque. Dans le cas particulier où φ possède $n = \dim(E)$ valeurs propres distinctes. On peut construire une base de vecteurs propres de φ.

Corollaire 50 (majoration du nombre de valeurs propres)

Soit φ un endomorphisme de E de dimension finie. Alors

$$\operatorname{Card}\big(\operatorname{Sp}(\varphi)\big) \leqslant \dim(E).$$

Remarque. Matriciellement, si $A \in \mathcal{M}_n(\mathbb{R})$,

$$\operatorname{Card}\big(\operatorname{Sp}(A)\big) \leqslant n.$$

5 Recherche de valeurs propres et vecteurs propres

Cette section détaille les différentes méthodes pour calculer les valeurs propres et les vecteurs propres d'une matrice.

5.1 Recherche des valeurs propres

Calcul des valeurs propres de $A = \begin{bmatrix} 0 & 2 & -1 \\ 3 & -2 & 0 \\ -2 & 2 & 1 \end{bmatrix}$.

Rappelons que :

- Le réel λ est valeur propre de A si et seulement si $\operatorname{rg}(A - \lambda I_n) < n$.
- Le rang est invariant par transposition et opérations élémentaires sur les lignes et les colonnes.

Remarque. Si on dispose d'un polynôme annulateur, on peut se limiter à regarder les racines du polynôme pour déterminer les valeurs propres.

5.2 Recherche des vecteurs propres

Calcul des vecteurs propres associés à la matrice A de l'exemple précédent.

Exercice 37

✦ Soient $A = \begin{bmatrix} 7 & 3 & -9 \\ -2 & -1 & 2 \\ 2 & -1 & -4 \end{bmatrix}$ et $P(x) = x^3 - 2x^2 - 5x + 6$. On admet que $P(A) = 0_3$.

1. Préciser les racines de P.
2. Déterminer les valeurs propres de A et, pour chaque valeur propre, donner une base du sous-espace propre associé.

⋙ *Pour un autre exemple de calcul, voir exercice 41.*

5.3 Cas particulier de la dimension 2

Proposition 51 (lien avec le déterminant)

Soient $A \in \mathcal{M}_2(\mathbb{R})$ et $\lambda \in \mathbb{R}$. On a :

$$\lambda \in \operatorname{Sp}(A) \iff \det(A - \lambda I_2) = 0.$$

Exercice 38

✧ Donner en fonction de $\alpha \in \mathbb{R}$, les valeurs propres réelles de

$$R_\alpha = \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{bmatrix}.$$

Exercices - TD

Exercice 39. ✧ Expliquez l'intérêt de la fonction Python suivante.

Editeur

```
def bidule(A,x) : # A est une matrice carrée et x, un réel
    [n,p]=np.shape(A)
    if n!=p :
        print('Non, la matrice n est pas carrée ...')
    elif np.linalg.matrix_rank(A-x*np.eye(n))<n :
        print('Oui !')
    else :
        print('Non !')
```

Exercice 40. ✦ Soient φ, ψ deux endomorphismes nilpotents de $\mathbb{R}^n$.
Montrer que si ces deux endomorphismes commutent, la composée $\varphi \circ \psi$ et la somme $\varphi + \psi$ sont nilpotentes.
Un endomorphisme φ est dit nilpotent s'il existe un entier p tel que $\varphi^p = 0_{\mathscr{L}(\mathrm{E})}$ (composée p-ième).

Exercice 41. ✦ *D'après EDHEC 2019*
On note id l'endomorphisme identité de $\mathbb{R}^3$ et on considère l'endomorphisme f de $\mathbb{R}^3$ dont la matrice dans la base canonique est :

$$\mathrm{A} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 3 & -2 \\ -1 & 1 & 0 \end{bmatrix}.$$

1. (a) Déterminer un polynôme annulateur de A qui soit de degré 2 .
   (b) En déduire les deux valeurs propres possibles λ_1 et λ_2 de A (avec $\lambda_1 < \lambda_2$).
   (c) En Python, la commande `linalg.matrix_rank(M)` de la bibliothèque `numpy` renvoie le rang de la matrice `M`. On a saisi

Editeur

```
import numpy as np
A = np.array([[1, 0, 0],[-2, 3, -2],[-1,1,0]])
r1=np.linalg.matrix_rank(A-np.eye(3))
r2=np.linalg.matrix_rank(A-2*np.eye(3))
print('r1=',r1,'r2=',r2)
```

Python a répondu

Console

```
>>> # script executed

r1= 1 r2= 2
```

   Que peut-on conjecturer quant aux valeurs propres de f et à la dimension des sous-espaces propres associés ?
   (d) Donner une base de chacun des noyaux $\mathrm{Ker}\left(f - \lambda_1 \,\mathrm{id}\right)$ et $\mathrm{Ker}\left(f - \lambda_2 \,\mathrm{id}\right)$.
2. (a) Justifier qu'il existe une base (u_1, v_1, v_2) de $\mathbb{R}^3$, où (u_1, v_1) est une base de $\mathrm{Ker}\left(f - \lambda_1 \,\mathrm{id}\right)$ et (v_2) une base de $\mathrm{Ker}\left(f - \lambda_2 \,\mathrm{id}\right)$. On choisira ces vecteurs de façon que leurs composantes soient des entiers naturels les plus petits possible, la dernière composante de u_1 et la première de v_1 étant nulles.
   (b) On note $x = (a, b, c)$ un vecteur quelconque de $\mathbb{R}^3$. Déterminer, en fonction de a, b et c les coordonnées de x dans la base (u_1, v_1, v_2).

Exercice 42. ✦ **Un exemple détaillé : la matrice Attila**
Soient $n \in \mathbb{N} \setminus \{0; 1\}$, $\mathrm{J} \in \mathscr{M}_n(\mathbb{R})$ la matrice dont tous les coefficients valent 1 et $\mathrm{X} \in \mathscr{M}_{n,1}(\mathbb{R})$ la matrice colonne dont tous les coefficients sont égaux à 1.

1. (a) Préciser le rang de J. En déduire que 0 est valeur propre dont on précisera la dimension de l'espace propre associé. Préciser une base $\mathscr{C}$ de l'espace propre associée à la valeur propre 0.
   (b) Calculer JX. Que peut-on en déduire ?
2. En déduire que J ne possède que deux valeurs propres.
3. Montrer que J est semblable à une matrice diagonale dont on précisera les coefficients diagonaux.
   On pourra considérer la base $\mathscr{B}$ en ajoutant à $\mathscr{C}$, le vecteur X.
4. Conclure en donnant tous les polynômes annulateur de J.

Exercice 43. ✦ **Crochet de Lie et nilpotence**
Soient $n \in \mathbb{N}^*$ et $A, B \in \mathcal{M}_n(\mathbb{R})$ tels que $AB - BA = A$.

1. Montrer que pour tout $k \in \mathbb{N}$, on a $A^k B - BA^k = kA^k$.
2. On considère $\varphi_B : \mathcal{M}_n(\mathbb{R}) \to \mathcal{M}_n(\mathbb{R})$ défini par $\varphi_B(M) = MB - BM$. dVérifier que φ_B est un endomorphisme de $\mathcal{M}_n(\mathbb{R})$.
3. Traduire la question 1 en terme de vecteur propre et valeur propre.
4. En déduire l'existence d'un entier k non nul tel que $A^k = 0_n$.

Exercice 44. ✦ Soit $n \in \mathbb{N}^*$. On note φ l'application définie sur $\mathbb{R}_n[x]$ par $\varphi(P) = (x-1)P'(x)$.

1. Montrer que φ est un endomorphisme de $\mathbb{R}_n[x]$. Préciser sa matrice dans la base canonique de $\mathbb{R}_n[x]$.
2. En déduire le spectre de φ.
3. Soit P un vecteur propre de φ associé à une valeur propre λ non nul.
   (a) Justifier que 1 est une racine de P.
   (b) En étudiant la multiplicité de 1 en tant que racine de P, exprimer P en fonction de λ.
4. Conclure en donnant tous les espaces propres de φ.

Exercice 45. ✦✦ Soient $A, B \in \mathcal{M}(\mathbb{R})$ et $\lambda \in \mathbb{R}$.
Montrer que λ est valeur propre de AB si et seulement si λ est valeur propre de BA.

Exercice 46. ✦✦ **Matrices stochastiques**
Une matrice carrée de taille n stochastique est une matrice $P = (p_{ij})_{1 \leqslant i,j \leqslant n}$ vérifiant les deux conditions :

i) Pour tout $i \in [\![1, n]\!]$, $\sum\limits_{j=1}^{n} p_{ij} = 1$.

ii) Tous les coefficients de la matrice P sont positifs.

1. Traduire la condition **i)** en terme de vecteur propre. *On pourra poser* $X = {}^t(1, \dots, 1)$.
2. Montrer que l'ensemble $\mathscr{E}$ des matrices stochastiques est une partie convexe stable par produit.
   Une partie A *est convexe si pour tous* a, $b \in A$, *tout* $\lambda \in [0,1]$, $\lambda a + (1-\lambda) b \in A$.
   Une partie A *est stable par produit si pour tous* a, $b \in A$, $ab \in A$.
3. Démontrer que les valeurs propres réelles d'une matrice stochastique sont comprises entre -1 et 1.

≫ *Pour aller plus loin, on pourra consulter le sujet EM Lyon 2010 ECS, problème I.*

Exercice 47. ✦✦ **Polynôme annulateur minimal** *Questions sans préparation HEC 2015*
Soient E un espace vectoriel réel de dimension finie et f un endomorphisme de E.

1. Établir l'existence d'un polynôme P non nul tel que $P(f) = 0$.
2. Soit Q un polynôme tel que $Q(f) = 0$ et de degré minimal parmi les polynômes non nuls tels que $P(f) = 0$. Montrer que toute racine de Q est valeur propre de f.

Exercice 48. ✦✦✦ **Matrice à diagonale dominante, localisation des valeurs propres réelles**
Une matrice carrée $A = \left(a_{i,j}\right)_{1 \leqslant i,j \leqslant n}$ est dite à diagonale dominante si pour tout $i \in [\![1, n]\!]$,

$$|a_{i,i}| > |a_{i,1}| + \cdots + |a_{i,i-1}| + |a_{i,i+1}| + \cdots + |a_{i,n}|.$$

1. Montrer qu'une matrice à diagonale dominante est inversible.
   Indication. On pourra s'intéresser au noyau de la matrice.
2. Soit $B = \left(b_{i,j}\right)_{1 \leqslant i,j \leqslant n} \in \mathcal{M}_n(\mathbb{R})$. Montrer que $\mathrm{Sp}(B) \subset \bigcup\limits_{i=1}^{n} [b_{ii} - r_i, b_{ii} + r_i]$ où $r_i = \sum\limits_{\substack{j=1 \\ j \neq i}}^{n} |b_{i,j}|$.

Exercice 49. ✦✦ *D'après Oraux HEC 2018*
Soient $A \in \mathcal{M}_n(\mathbb{R})$ et $\varphi_A : M \in \mathcal{M}_n(\mathbb{R}) \mapsto AM \in \mathcal{M}_n(\mathbb{R})$.

1. Démontrer que A et φ_A ont le même spectre.
2. Comparer, pour chacune des valeurs propres, les dimensions des sous-espaces propres correspondants de A et φ_A.

Exercice 50. ✦✦ **Exemple dans un espace fonctionnel** *D'après EMLyon 2014 ECS*
On note E l'espace vectoriel des applications de $\mathbb{R}$ dans $\mathbb{R}$ continues, E_1 le sous-espace vectoriel des applications de $\mathbb{R}$ dans $\mathbb{R}$ de classe $\mathscr{C}^1$. On note, pour tout élément f de E, $T(f)$ l'application de $\mathbb{R}$ dans $\mathbb{R}$ définie, pour tout $x \in \mathbb{R}$, par

$$T(f)(x) = \frac{1}{2} \int_{x-1}^{x+1} f(t)\,dt.$$

1. Établir que, pour tout élément f de E, $T(f)$ appartient à E_1 et que, pour tout $x \in \mathbb{R}$: $\big(T(f)\big)'(x) = \frac{1}{2}\big(f(x+1) - f(x-1)\big)$.
   On note $T : E \to E$ l'application qui, à f, associe $T(f)$.
2. Montrer que T est un endomorphisme de E. Est-il surjectif?

3. Soit $s : \mathbb{R} \to \mathbb{R}$ définie par $s(t) = \sin(\pi t)$. Calculer $\mathrm{T}(s)$. Est-ce que T est injectif ?
4. Soit $\varphi : \mathbb{R} \to \mathbb{R}$, définie par

$$\varphi(t) = \begin{cases} \dfrac{e^t - e^{-t}}{2t} & \text{si } t \neq 0 \\ 1 & \text{si } t = 0. \end{cases}$$

   Pour tout $a \in \mathbb{R}$, on pose de plus $f_a : x \in \mathbb{R} \mapsto e^{ax} \in \mathbb{R}$. Vérifier que pour tout $a \in \mathbb{R}$, $\mathrm{T}(f_a) = \varphi(a) f_a$.
5. Justifier que $[1;+\infty[\subset \mathrm{Sp}(\mathrm{T})$. A-t'on égalité ?

CHAPITRE 3

Révisions et compléments sur les variables aléatoires

If there is a 50-50 chance that something can go wrong, then nine times out of 10 it will.

PAUL HARVEY

Animateur radio américain (1918-2009).

1 Rappels : espérance et variance (cas discret)

1.1 Espérance et la formule de transfert

Définition-Rappel 52 (espérance)

Soit X une variable aléatoire discrète infinie. On note $X(\Omega) = \{x_k \mid k \in \mathbb{N}^*\}$. On dit que X admet une **espérance** si la série de terme général $x_k \mathbf{P}(X = x_k)$ est *absolument convergente*. Alors on définit l'espérance de X par

$$\mathbf{E}(X) = \sum_{k=1}^{+\infty} x_k \mathbf{P}(X = x_k).$$

Attention. Il ne faut pas oublier de vérifier la convergence absolue. Elle est nécessaire pour justifier que l'espérance est bien définie et ne dépend pas du choix de l'indexation $X(\Omega)$.

Exercice 51

✦ On effectue une infinité de tirages successifs, mutuellement indépendants, dans une urne contenant initialement une boule rouge et une boule bleue. À chaque tirage, on note la couleur de la boule et on la remet dans l'urne en y ajoutant une boule bleue.
On note X (respectivement Y) la variable aléatoire égale au rang d'apparition de la première boule bleue (respectivement rouge). Pour que ces deux variables aléatoires soient bien définies, on admet que presque sûrement une boule bleue ou rouge finit par apparaître.

1. (a) Donner la loi de X.
   (b) X admet-elle une espérance ? Si oui, la calculer.
2. Reprendre la question 1 avec la variable aléatoire Y.

Proposition 53 (existence par domination)

Soient X et Y sont deux variables aléatoires discrètes sur un même espace probabilisé.

Si | $\rightharpoonup$ $0 \leqslant |X| \leqslant Y$.
$\rightharpoonup$ Y admet une espérance.

Alors X admet une espérance et $|\mathbf{E}(X)| \leqslant \mathbf{E}(Y)$.

⚠ **Attention.** Il ne faut pas confondre l'énoncé précédent avec la propriété de croissance de l'espérance dont on rappelle l'énoncé :

Si | $\rightharpoonup$ X et Y admettent une espérance,
$\rightharpoonup$ $X \leqslant Y$, **alors** $\mathbf{E}(X) \leqslant \mathbf{E}(Y)$.

Théorème 54 (formule de transfert)

Soient X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ et g une application définie sur $X(\Omega) = \{x_k \mid k \in I\}$. Les énoncés suivants sont équivalents.

i) La variable aléatoire $g(X)$ possède une espérance.

ii) La série de terme général $g(x_k)\,\mathbf{P}\big([X = x_k]\big)$ est absolument convergente.

Dans ce cas $$\mathbf{E}\big(g(X)\big) = \sum_{k \in I} g(x_k)\,\mathbf{P}\big([X = x_k]\big).$$

Exercice 52

✦ *Source oraux HEC-E 2009*

Soit X une variable aléatoire qui suit la loi de Poisson de paramètre $\lambda > 0$.

1. Établir l'existence de $\mathbf{E}\left(\frac{1}{1+X}\right)$.
2. Montrer que $\mathbf{E}\left(\frac{1}{1+X}\right) \leqslant \min\left(1, \frac{1}{\lambda}\right)$.

Remarque. La formule de transfert permet de prouver une version faible de la linéarité de l'espérance. Pour $a, b \in \mathbb{R}$ et X, une variable aléatoire admettant une espérance $\mathbf{E}(aX + b) = a\mathbf{E}(X) + b$.

Exercice 53

✧ Si X est une variable aléatoire admettant une espérance et X et $1 - X$ ont même loi.
Que dire de l'espérance de X ?

1.2 Moments et variance

Pour tout entier naturel s, le moment d'ordre s d'une variable aléatoire X est le nombre $m_s(X)$ défini par :

$$m_s(X) = \mathbf{E}(X^s).$$

En particulier pour $s = 1$, on retrouve l'espérance $m_1(X) = \mathbf{E}(X)$.
Dans le cas d'une variable aléatoire discrète, et sous réserve de convergence absolue, la formule de transfert donne :

$$m_s(X) = \sum_{k \in I} {x_k}^s\,\mathbf{P}\big(X = x_k\big).$$

Remarque. On montre que si X a un moment d'ordre r alors, X admet un moment d'ordre s pour tout entier $s \leqslant r$.

Exercice 54

✦

1. Prouver la remarque précédente.
2. Soit $r \in \mathbb{N}^*$, construire une variable aléatoire admettant un moment d'ordre r mais pas de moment d'ordre $r + 1$.

Définition 55 (variance et écart-type)

Soit X une variable aléatoire admettant un moment d'ordre 2.

- On appelle **variance** d'une variable aléatoire X, la quantité $\mathbf{V}(\mathrm{X}) = \mathbf{E}\left(\left(\mathrm{X} - \mathbf{E}(\mathrm{X})\right)^2\right)$.
- On appelle **écart-type** de X, la quantité $\sigma(\mathrm{X}) = \sqrt{\mathbf{V}(\mathrm{X})}$.

L'écart-type permet de quantifier les écarts par rapport à la moyenne. Un écart-type fort traduit un « grand éloignement » des valeurs de X par rapport à sa moyenne.

Proposition 56 (propriétés de la variance)

Soit X une variable aléatoire admettant un moment d'ordre 2.

- $\mathbf{V}(\mathrm{X}) = 0$ si et seulement si X est une application presque sûrement constante.
- Pour tous réels a, b,

$$\mathbf{V}(a\mathrm{X}+b) = a^2\mathbf{V}(\mathrm{X}) \quad ; \quad \sigma(a\mathrm{X}+b) = |a|\sigma(\mathrm{X}).$$

Vocabulaire. Une variable aléatoire X est dite **centrée** si $\mathbf{E}(\mathrm{X}) = 0$. Elle est dite **réduite** si $\sigma(\mathrm{X}) = 1$. Soit X une variable aléatoire. La variable aléatoire $\mathrm{Y} = \frac{\mathrm{X}-\mathbf{E}(\mathrm{X})}{\sigma(\mathrm{X})}$ est centrée, réduite.

Théorème 57 (formule de KOENIG-HUYGENS)

Soit X une variable aléatoire admettant un moment d'ordre 2, alors

$$\mathbf{V}(\mathrm{X}) = \mathbf{E}\left(\mathrm{X}^2\right) - \mathbf{E}(\mathrm{X})^2.$$

1.3 Cas les lois usuelles discrètes

On dit que X est une **variable aléatoire certaine**, ou **presque sûrement constante** s'il existe un réel c tel que $\mathbf{P}(\mathrm{X} = c) = 1$. Alors,

$$\mathbf{E}(\mathrm{X}) = c \quad \text{et} \quad \mathbf{V}(\mathrm{X}) = 0.$$

Proposition 58 (espérance et variance, lois usuelles)

Soit X, une variable aléatoire.

- Si $\mathrm{X} \hookrightarrow \mathscr{B}(p)$, alors $\quad \mathbf{E}(\mathrm{X}) = p \quad$ et $\quad \mathbf{V}(\mathrm{X}) = pq.$
- Si $\mathrm{X} \hookrightarrow \mathscr{B}(n,p)$, alors $\quad \mathbf{E}(\mathrm{X}) = np \quad$ et $\quad \mathbf{V}(\mathrm{X}) = npq.$
- Si $\mathrm{X} \hookrightarrow \mathscr{U}([\![1,n]\!])$, alors $\quad \mathbf{E}(\mathrm{X}) = \dfrac{n+1}{2} \quad$ et $\quad \mathbf{V}(\mathrm{X}) = \dfrac{n^2-1}{12}.$

Exercice 55

✦ Soient $n \in \mathbb{N}$ et X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ et à valeurs dans $[\![0;n]\!]$. On définit alors la fonction génératrice $\mathrm{G_X}$ de X par

$$\forall\, t \in \mathbb{R}, \qquad \mathrm{G_X}(t) = \sum_{k=0}^{n} \mathbf{P}([\mathrm{X}=k])\, t^k.$$

1. Justifier que $\mathbf{E}(\mathrm{X}) = \mathrm{G}'_{\mathrm{X}}(1)$.
2. Trouver une relation entre $\mathbf{V}(\mathrm{X}), \mathrm{G}''_{\mathrm{X}}(1)$ et $\mathrm{G}'_{\mathrm{X}}(1)$.
3. Retrouver les résultats énoncés à la proposition précédente.

Proposition 59 (espérance et variance, cas discret dénombrable)

- Si $X \hookrightarrow \mathcal{G}(p)$, alors X admet une espérance et une variance avec

$$\mathbf{E}(X) = \frac{1}{p} \quad \text{et} \quad \mathbf{V}(X) = \frac{q}{p^2}.$$

- Si $X \hookrightarrow \mathcal{P}(\lambda)$, alors X admet une espérance et une variance avec

$$\mathbf{E}(X) = \lambda \quad \text{et} \quad \mathbf{V}(X) = \lambda.$$

Exercice 56

✦ *Les questions sont indépendantes.*

1. Prouver les énoncés de la proposition précédente.
2. Soit X une v.a suivant une loi géométrique de paramètre $p \in]0,1[$. Calculer $\mathbf{E}(1/X)$.
   On pourra admettre que pour tout $x \in]-1,1[$, $\sum_{n=1}^{+\infty} \frac{x^n}{n} = -\ln(1-x)$.

2 Espérance conditionnelle

Définition 60 (espérance conditionnelle)

Soient X, une variable aléatoire sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ et A un événement de probabilité non nulle.
La variable X admet une **espérance sachant** A, notée $\mathbf{E}(X|A)$ si X admet une espérance pour la probabilité $\mathbf{P}_A$. Autrement dit, si la série $\sum x_k \mathbf{P}_A(X = x_k)$ converge absolument, alors

$$\mathbf{E}(X|A) = \sum_{x \in X(\Omega)} x \mathbf{P}_A\big([X = x]\big).$$

Remarque. Si X est une variable aléatoire finie, alors la convergence absolue est automatique. Pour tout événement A possible, X admet une espérance conditionnelle sachant A.

Exercice 57

✦ Soient X une variable aléatoire réelle définie sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ avec $X(\Omega) = \{x_k \mid k \in \mathbb{N}\}$ et A un événement de probabilité non nulle. Justifier que si $\mathbf{E}(X)$ existe, alors $\mathbf{E}(X \mid A)$ existe aussi et

$$\big|\mathbf{E}(X \mid A)\big| \leqslant \frac{1}{\mathbf{P}(A)} \mathbf{E}\big(|X|\big).$$

Théorème 61 (formule de l'espérance totale)

Soit X, une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$.

Si
- $(A_n)_{n \in I}$ est un système complet d'événements possibles.
- Pour tout $n \in I$, $\mathbf{E}(X \mid A_n)$ existe.
- La série $\sum_n \mathbf{P}(A_n) \mathbf{E}(|X| \, |A_n)$ converge absolument.

Alors X admet une espérance et $$\mathbf{E}(X) = \sum_{n \in I} \mathbf{P}(A_n) \mathbf{E}(X \mid A_n).$$

Exercice 58

✦ Lors d'une ponte, un esturgeon pond des œufs suivant une loi de Poisson de paramètre $\lambda = 500\,000$. Chaque œuf a une probabilité $p = 1/100\,000$ d'être un adulte. On fait l'hypothèse que les pontes sont indépendantes.
Combien, en moyenne, une ponte donne d'adulte esturgeon ?

⚠ **Attention.** Il faut bien noter les valeurs absolues dans la somme $\sum_n \mathbf{P}(A_n)\,\mathbf{E}(|X|\,|A_n)$ même s'il arrive souvent que la variable X soit positive. L'exercice **??** illustre cette condition.

Exercices - TD

Révisions et compléments

Exercice 59. ✦✦ *D'après Oraux HEC*

Soient $p \in]0;1[$ et la fonction Python X ci-contre.

1. En s'aidant d'un lancer d'une pièce dont la probabilité d'obtenir "face" est p, interpréter ce que simule la fonction X?
2. Soit X la variable aléatoire simulée par la fonction ci-dessus.
   (a) Quelle est la loi de X?
   (b) Quelle est l'espérance de X si elle existe?

Editeur

```
import random as rd

def X(p) :
    k=1
    y=0
    a=rd.random()
    while a>p :
        y=y+1
        k=k+1
        a=rd.random()
    while a<p :
        k=k+1
        a=rd.random()
    return k-1-y
```

Exercice 60. ✦

On tire (avec remise) une boule d'une urne contenant n boules numérotées.

1. On note T la variable aléatoire égale au numéro du tirage où pour la première fois deux boules différentes ont été tirées. Déterminer l'espérance de T.
2. Interpréter la variable aléatoire V_n dont la fonction Python suivante calcule une simulation.

Editeur

```
import random as rd
def V(n) :
    A=[]
    compteur=0
    while len(A)<n :
        u=rd.randint(1,n+1)
        if not(u in A) :
            A.append(u)
        compteur=compteur+1
    return compteur
```

Exercice 61. ✦ Soient $m \in \mathbb{N}^*$ et $p \in]0,1[$. On suppose que $X \hookrightarrow \mathscr{B}(m,p)$. On note $Y = \frac{1}{X+1}$. Calculer l'espérance de Y.

Exercice 62. ✦ Soit X une variable aléatoire de loi de Poisson de paramètre $\lambda \in \mathbb{R}_*^+$.

1. Soit $Y = (-1)^X X$. Est-ce que Y possède une espérance? Si oui, la calculer.
2. Soit Z définie par :
   - Si X prend une valeur paire, alors Z prend la valeur $\frac{X}{2}$
   - Si X prend une valeur impaire, alors Z prend la valeur 0.

   (a) Quelle relation a-t-on entre X, Y et Z?
   (b) Est-ce que Z possède une espérance? Si oui, la calculer.

Exercice 63. ✦ Soit X, une v.a à valeurs dans $\mathbb{N}^*$ et admettant une espérance.

1. Justifier que $1/X$ admet une espérance et exprimer $\mathbf{E}(1/X)$ en fonction des réels x_k et $p_k = \mathbf{P}([X = x_k])$.
2. En déduire que $\mathbf{E}(X)\mathbf{E}(1/X) \geqslant 1$.
   Indication. On pourra utiliser l'inégalité de Cauchy-Schwarz dans $\mathbb{R}^n$:

$$\forall (x_1,\ldots,x_n) \in \mathbb{R}^n, \quad \forall (y_1,\ldots,y_n) \in \mathbb{R}^n, \qquad \left(\sum_{k=1}^{n} x_k y_k\right)^2 \leqslant \left(\sum_{k=1}^{n} {x_k}^2\right) \cdot \left(\sum_{k=1}^{n} {y_k}^2\right).$$

Exercice 64. ✦✦ **Inégalité de Jensen**
Soit X une variable aléatoire finie et $\varphi : \mathbb{R} \to \mathbb{R}$ une fonction convexe. Démontrer que $\varphi(\mathbf{E}(X)) \leqslant \mathbf{E}(\varphi(X))$.

Exercice 65. ✦✦✦ **Nouvelle expression de l'espérance**
Soit X une variable aléatoire à valeurs dans $\mathbb{N}$.

1. On suppose dans cette question que X admet une espérance.
Montrer que la série $\sum_{k\geqslant 0} \mathbf{P}(X > k)$ converge, et que
$$\mathbf{E}(X) = \sum_{k=0}^{+\infty} \mathbf{P}(X > k) \qquad (E).$$
2. Réciproquement, on suppose dans cette question que la série $\sum_{k\geqslant 0} \mathbf{P}(X > k)$ converge. Montrer que X admet une espérance, et que la relation (E) est valable.
3. Adapter le raisonnement pour montrer que si X admet un moment d'ordre 2
$$\mathbf{E}(X^2) = \sum_{k=0}^{+\infty} (2k+1)\,\mathbf{P}(X > k).$$

Exercice 66. ✦✦ Soient X une variable aléatoire réelle et a, b deux réels. On note $m = \mathbf{E}(X)$.

1. Montrer que $\mathbf{E}\big((X-a)(b-X)\big) = (m-a)(b-m) - V(X)$.
2. On suppose $a < b$ et $X(\Omega) \subset [a, b]$.
   (a) Montrer que $(m-a)(b-m) \leqslant \frac{(b-a)^2}{4}$. En déduire $\mathbf{V}(X) \leqslant \frac{(b-a)^2}{4}$.
   (b) Ce majorant peut-il être amélioré ?

Exercice 67. ✦✦✦ *D'après oral HEC*
Soit $\alpha \in \mathbb{R}$. On considère une variable aléatoire réelle discrète X admettant des moments jusqu'à l'ordre 4 et vérifiant :
$$\begin{cases} \mathbf{E}(X) = \alpha \\ \mathbf{E}\left(X^2\right) = \mathbf{E}\left(X^4\right) = 1. \end{cases}$$

1. Montrer que α est nécessairement compris entre -1 et $+1$.
2. Trouver la loi de X.

Espérance conditionnelle

Exercice 68. ✧ **Vrai ou faux ?**
Pour tout A, un événements avec $\mathbf{P}(A), \mathbf{P}(B) \notin \{0;1\}$, et X, Y, deux variables aléatoires admettant une espérance.

1. $\mathbf{E}(X|\Omega) = \mathbf{E}(X)$.
2. $\mathbf{E}(X) = \mathbf{E}(X \mid A) + \mathbf{E}(X \mid \overline{A})$.
3. $\mathbf{E}(X) = \mathbf{P}(A)\cdot\mathbf{E}(X \mid A) + \mathbf{P}(\overline{A})\cdot\mathbf{E}(X \mid \overline{A})$.
4. Si, pour tout $x \in X(\Omega)$, les événements $[X = x]$ sont indépendants de A, $\mathbf{E}(X \mid A) = \mathbf{E}(X)$.
5. Si $X \leqslant Y$ alors $\mathbf{E}(X|A) \leqslant \mathbf{E}(Y|A)$.

Exercice 69. ✦ Alice et Bob font une partie de Tennis de Table. Alice commence et renvoie la balle avec la même probabilité p_A. Bob fait de même avec probabilité p_B. On suppose les coups mutuellement indépendants et on note X, la variable aléatoire correspondant au nombre de coups valides en fonctions de p_A et p_B.

1. Donner la loi de X. Vérifier que l'espérance existe et la calculer.
2. Préciser $\mathbf{E}\big(X \mid [X \geqslant 1]\big)$.

Exercice 70. *Source HEC 2014*
Soient $n_1 \in \mathbb{N}^*$, $n_2 \in \mathbb{N}^*$ et $p \in]0,1[$. On considère deux variables aléatoires indépendantes X_1 et X_2 définies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$: X_1 suit la loi binomiale de paramètres $\left(n_1, p\right)$ et X_2 suit la loi binomiale de paramètres $\left(n_2, p\right)$.

1. Rappeler la loi de $X_1 + X_2$.
2. Soit $n \in (X_1 + X_2)(\Omega)$. Préciser $\mathbf{P}_{[X_1+X_2=n]}\big([X_1 = k]\big)$.
3. Exprimer à l'aide d'une somme $\mathbf{E}\big(X_1 \mid [X_1 + X_2 = n]\big)$.

Exercice 71. ✦✦ **Formule de la variance totale**
Soient X une variable aléatoire discrète et A, un événement de probabilité non nulle. Sous réserve d'existence, on définit la variance de X sachant A, par la variance de X suivant la probabilité $\mathbf{P}_A$. On la note $\mathbf{V}(X \mid A)$.
Démontrer que si X admet une variance et $(A_i)_{i\in[[1,n]]}$ est un système complet d'événements alors
$$\begin{aligned}\mathbf{V}(X) = &\sum_{i=1}^{n} \mathbf{V}\left(X \mid A_i\right)\mathbf{P}\left(A_i\right) + \sum_{i=1}^{n} \mathbf{E}\left(X \mid A_i\right)^2\left(1 - \mathbf{P}\left(A_i\right)\right)\mathbf{P}\left(A_i\right) \\ &- 2\sum_{i<j} \mathbf{E}\left(X \mid A_i\right)\mathbf{P}\left(A_i\right)\mathbf{E}\left(X \mid A_j\right)\mathbf{P}\left(A_j\right).\end{aligned}$$

Exercice 72. ✦✦ Soit $(a_n)_{n\in\mathbb{N}^*}$, la suite définie par $a_n = \frac{1}{n(n+1)}$.

1. Déterminer la nature et la somme éventuelle des séries $\sum a_n$ et $\sum n a_n$.
2. Soit X, une variable aléatoire définie sur $(\Omega, \mathscr{A}, \mathbf{P})$ à valeurs dans $\mathbb{Z}^*$ telle que

$$\forall n \in \mathbb{Z}^*, \qquad \mathbf{P}\big([X = n]\big) = \alpha\, a_{|n|} \quad \text{où} \quad \alpha \in \mathbb{R}.$$

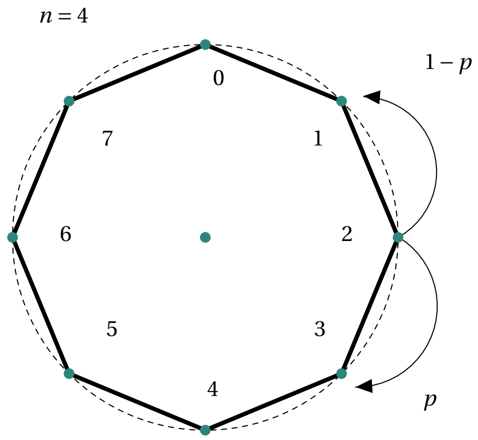
   Déterminer α.

3. (a) Que peut-on dire de la famille d'événements (A_n) avec $A_n = [X = n] \cap [X = -n]$ pour $n \in \mathbb{N}^*$?
   (b) Vérifier que $\mathbf{E}(X \mid A_n)$ est bien défini pour tout $n \in \mathbb{N}^*$. Préciser sa valeur.
   (c) Montrer que la série de terme général $\mathbf{E}(X \mid A_n)\, \mathbf{P}(A_n)$ est absolument convergente et calculer sa somme.
4. Est-ce que la variable X admet-elle une espérance ? Commenter ce résultat en comparant à la formule de l'espérance totale.

Exercice 73. ✦ **Mobile sur un polygone**

On place $2n$ points numérotés de 0 à $2n-1$ sur un cercle. Les points sont répartis uniformément.
Un mobile part initialement de 0 et à chaque étape, il avance d'une unité avec une probabilité p et recule avec probabilité $1-p$. Soit $N \in \mathbb{N}^*$, le nombre d'étapes
On note :

- Pour tout $k \in [\![0;N]\!]$, on note X_k la position du mobile à la k-ième étape.
- T_p la variable aléatoire égale au temps de retour à l'origine. Si le mobile ne retourne pas à l'origine au bout des N étapes, on pose $T_p(\omega) = 0$.



1. Écrire un programme qui prend en argument p, simule les 100 premières étapes du mobile et renvoie la position finale du mobile.
   La commande X%2n *renvoie le reste de la division euclidienne par* $2n$.
2. Modifier le programme pour afficher le temps T_p de premier retour à l'origine 0.
3. En déduire une approximation de l'espérance de T.
   Pour cela, on produira $m = 2000$ *réalisations de la variable* T. *La moyenne arithmétique de ces m évaluations donne une approximation de l'espérance. C'est la loi des grands nombres.*
4. Vérifier numériquement et expliquer l'égalité $\mathbf{E}(T_p) = \mathbf{E}(T_{1-p})$?

CHAPITRE 4

Espérance et espérance conditionnelle

Le hasard ne profite qu'aux esprits préparés.

LOUIS PASTEUR (1822-1895)

1 Rappels et compléments sur les séries

1.1 Définitions

Définition-Rappel 62 (série numérique)

Soit $(u_n)_{n\in\mathbb{N}}$ une suite réelle. La **série** associée à $(u_n)_{n\in\mathbb{N}}$ est la suite $(S_n)_{n\in\mathbb{N}}$ avec

$$\forall\, n \in \mathbb{N}, \qquad S_n = \sum_{k=0}^{n} u_k.$$

On dit que la série converge si la suite $(S_n)_{n\in\mathbb{N}}$ converge vers une limite finie.

Attention. Il ne faut pas confondre :

- $\sum u_n$: la série de terme général u_n ;
- $S_n = \sum_{k=0}^{n} u_k$: la somme partielle d'ordre n de la série ;
- $\sum_{k=0}^{+\infty} u_k$: la somme de la série, i.e, sous réserve de convergence, la limite des sommes partielles ;
- $\sum_{k=n+1}^{+\infty} u_k$: le reste d'ordre n de la série si cette dernière est convergente.

Remarques.

- Pour tout $n \in \mathbb{N}^*$, $\qquad u_n = S_n - S_{n-1}$.
- Si la série de terme général u_n converge, alors $u_n \underset{n\to\infty}{\longrightarrow} 0$.

La série harmonique $\sum 1/n$ montre que la réciproque est fausse.

- Soient u et v deux suites réelles telles que les séries de termes généraux u_n et v_n convergent.

Pour tous $\lambda, \mu \in \mathbb{R}$, la série de terme général $\lambda u_n + \mu v_n$ converge et

$$\forall\, n_0 \in \mathbb{N}, \qquad \sum_{k=n_0}^{+\infty} (\lambda u_k + \mu v_k) = \lambda \sum_{k=n_0}^{+\infty} u_k + \mu \sum_{k=n_0}^{+\infty} v_k.$$

Théorème 63 (convergence absolue)

- On dit que la série de terme général u_n **converge absolument** si la série de terme général $|u_n|$ converge.
- **Si** une série converge absolument, **alors** elle converge.

Remarque. La réciproque est fausse. La série $\sum(-1)^n/n$ donne un contre exemple.

Théorème 64 (famille sommable)

Soit I un ensemble dénombrable, indexé par $\mathbb{N}$ sous la forme $\mathrm{I} = \{\varphi(n) \mid n \in \mathbb{N}\}$ où φ est une bijection de $\mathbb{N}$ dans I.

Si la série $\sum u_{\varphi(n)}$ converge absolument,

alors sa somme est indépendante de l'indexation φ. On peut donc noter sans ambiguïté $\sum\limits_{i \in \mathrm{I}} u_i$.

1.2 Séries de références

Théorème 65 (séries géométriques, série exponentielle)

- Soit $x \in \mathbb{R}$. Les séries de terme généraux x^k, kx^{k-1} et $k(k-1)x^{k-2}$ sont convergentes si et seulement si $|x| < 1$. Dans ce cas,

$$\sum_{k=0}^{+\infty} x^k = \frac{1}{1-x}, \quad \sum_{k=1}^{+\infty} kx^{k-1} = \frac{1}{(1-x)^2} \quad \text{et} \quad \sum_{k=2}^{+\infty} k(k-1)x^{k-2} = \frac{2}{(1-x)^3}.$$

- Pour tout réel x, la série de terme général $\frac{x^k}{k!}$ est convergente et $\sum\limits_{k=0}^{+\infty} \frac{x^k}{k!} = \exp(x)$.
- On appelle **série de Riemann**, une série de terme général $\frac{1}{n^\alpha}$, avec $\alpha \in \mathbb{R}$.
Cette série est convergente si et seulement si $\alpha > 1$.

1.3 Critères de convergence pour les séries à termes positifs

Théorème 66 (critère de comparaison)

Soient u et v deux suites réelles telles qu'à partir d'un certain rang, $0 \leqslant u_n \leqslant v_n$.

- Si la série de terme général v_n converge, alors la série de terme général u_n aussi.
- Si la série de terme général u_n diverge, alors la série de terme général v_n aussi.

Exercice 74

✦ *Les questions 1 et 2 sont indépendantes.*

1. Soit $\sum u_k$ une série à termes positifs. On suppose que la série $\sum k^2 {u_k}^2$ est convergente.
   (a) Montrer que la série $\sum u_k$ est convergente. (i)
   (b) Étudier la réciproque. (i)
2. Soit $\sum u_n$ une série à termes positifs convergente. Justifier la convergence des séries :

$$\sum \frac{3u_n}{1+{u_n}^2}, \quad \sum \ln(1+u_n) \quad \text{et} \quad \sum \int_0^{u_n} \frac{\mathrm{d}t}{1+t^4}.$$

Théorème 67 (critères de négligeabilité et d'équivalence)

Soient u et v deux suites réelles.

- **Si**
  - $u_n = o(v_n)$
  - v est positive à partir d'un certain rang
  - la série de terme général v_n converge,

alors la série de terme général u_n converge.

- **Si**
  - $u_n \sim v_n$
  - v est positive à partir d'un certain rang,

alors les séries de termes généraux u_n et v_n sont de même nature.

Remarque. On a des critères analogues si v est négative à partir d'un certain rang.

Attention. Il ne faut pas oublier les conditions de signes lors de l'application de ces théorèmes.

Exercice 75

Exemples

✦ Déterminer la nature des séries suivantes :

- $\sum_{n\geqslant 1} \frac{1}{\sqrt{n}} \ln\left(1+\frac{3}{\sqrt{n}}\right)$;
- $\sum_{n\geqslant 1} \frac{n^3}{2^n}$;
- $\sum_{n\geqslant 0} \frac{n^n}{n!}$;
- $\sum_{n\geqslant 1} \frac{\ln(n)^4}{4^{\ln(n)}}$.

✦ Discuter, en fonction du paramètre $\alpha \in \mathbb{R}$, de la nature de la série :

- $\sum \arctan\left(\frac{n^\alpha}{1+n}\right)$,
- $\sum_{n\geqslant 1} \left(\frac{1}{1+2+\cdots+n}\right)^\alpha$;
- $\sum_{n\geqslant 1} \left(\sqrt{n+7}-\sqrt{n}\right)^\alpha$.

2 Révision sur les probabilités

2.1 L'application probabilité

On représente le résultat d'une expérience aléatoire comme un élément ω de l'ensemble Ω de tous les résultats possibles. Dans la suite, $\mathscr{A} \subset \mathscr{P}(\Omega)$ est un ensemble des événements. Il vérifie

- $\Omega \in \mathscr{A}$.
- *$\mathscr{A}$ est stable par passage au complémentaire :* $\quad \forall A \in \mathscr{A}, \quad \overline{A} = \Omega \setminus A \in \mathscr{A}$.
- *$\mathscr{A}$ est stable par union et intersection finie ou dénombrable* : si I est un ensemble fini ou dénombrable et si pour tout $i \in I$, $A_i \in \mathscr{A}$, alors
$$\bigcup_{i\in I} A_i \in \mathscr{A} \quad \text{et} \quad \bigcap_{i\in I} A_i \in \mathscr{A}.$$

Vocabulaire. On parle de tribu ou de σ-algèbre pour désigner $\mathscr{A}$.

Définition 68 (l'application probabilité)

Soient Ω, un univers des possibles et $\mathscr{A}$ un ensemble d'événements.
Une **probabilité** est une application **P** réelle définie sur $\mathscr{A}$ vérifiant les conditions suivantes.

- $\mathbf{P} : \mathscr{A} \to [0,1]$;
- $\mathbf{P}(\Omega) = 1$;
- **P** est σ-**additive** :
  Pour toute famille $(A_i)_{i\in I}$ finie ou dénombrable d'événements deux à deux disjoints,
$$\mathbf{P}\left(\bigcup_{i\in I} A_i\right) = \sum_{i\in I} \mathbf{P}(A_i).$$

Pour tout $A \in \mathscr{A}$, $\mathbf{P}(A)$ est appelée **la probabilité de l'événement** A.

Remarque. Dans le cas où l'ensemble des indices I est dénombrable (par exemple, $\mathbb{N}$), la définition suppose la convergence de la série $\sum_{i \in I} \mathbf{P}(A_i)$.

Vocabulaire.
- La donnée d'un triplet $(\Omega, \mathscr{A}, \mathbf{P})$ où Ω est un univers des possibles, $\mathscr{A}$ un ensemble d'événements sur Ω et $\mathbf{P}$ une probabilité sur $(\Omega, \mathscr{A})$, définit un **espace probabilisé**.
- Un événement de probabilité nulle est dit **négligeable**.
- Un événement de probabilité 1 est dit **presque-sûr**.

Théorème 69 (de la limite monotone)

Soient $(A_n)_{n\in\mathbb{N}}$ et $(B_n)_{n\in\mathbb{N}}$ deux suites d'événements sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$.

- **Si** la suite $(A_n)_{n\in\mathbb{N}}$ est **croissante** pour l'inclusion (c'est-à-dire, $\forall\, i \in \mathbb{N}$, $A_i \subset A_{i+1}$),

Alors $$\mathbf{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to+\infty} \mathbf{P}(A_n).$$

- **Si** la suite $(B_n)_{n\in\mathbb{N}}$ est **décroissante** pour l'inclusion ($\forall\, i \in \mathbb{N}$, $B_{i+1} \subset B_i$),

Alors $$\mathbf{P}\left(\bigcap_{n\in\mathbb{N}} B_n\right) = \lim_{n\to+\infty} \mathbf{P}(B_n).$$

2.2 Probabilité conditionnelle et indépendance

Définition 70 (probabilité conditionnelle)

Soient $\left(\Omega, \mathscr{A}, \mathbf{P}\right)$ un espace probabilisé et $A \in \mathscr{A}$ tel que $\mathbf{P}(A) \neq 0$. Soit $B \in \mathscr{A}$.
La **probabilité conditionnelle** de B sachant A est

$$\mathbf{P}_A(B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(A)}.$$

Remarques.
- À partir de la définition, on prouve une première **formule de Bayes**.

Si $\mathbf{P}(A) \neq 0$ et $\mathbf{P}(B) \neq 0$, alors $$\mathbf{P}_B(A) = \frac{\mathbf{P}_A(B) \cdot \mathbf{P}(A)}{\mathbf{P}(B)} \qquad (\bullet).$$

- Si A est un événement non négligeable, alors $(\Omega, \mathscr{A}, \mathbf{P}_A)$ est un espace probabilisé.

Définition 71 (indépendance deux à deux et mutuelle)

- Deux événements A et B sont dits **indépendants** pour la probabilité $\mathbf{P}$ si

$$\mathbf{P}(A \cap B) = \mathbf{P}(A) \times \mathbf{P}(B).$$

- Soit $(A_i)_{i\in I}$ une famille finie ou dénombrable d'événements. Les événements de cette famille sont dits **mutuellement indépendants** si pour toute partie finie non vide $\{i_1, i_2, \cdots, i_p\} \subset I$,

$$\mathbf{P}\left(\bigcap_{k=1}^{p} A_{i_k}\right) = \prod_{k=1}^{p} \mathbf{P}(A_{i_k}).$$

2.3 Formules des probabilités composées et probabilités totales

Proposition 72 (formule des probabilités composées)

Soient $A_1, A_2, \dots, A_{n-1}, A_n$, des événements d'un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ tels que

$$\mathbf{P}(A_1 \cap A_2 \cap \dots \cap A_{n-1}) \neq 0.$$

Alors $$\mathbf{P}(A_1 \cap A_2 \cap \dots \cap A_n) = \mathbf{P}(A_1) \cdot \mathbf{P}_{A_1}(A_2) \cdot \mathbf{P}_{A_1 \cap A_2}(A_3) \dots \mathbf{P}_{A_1 \cap \dots \cap A_{n-1}}(A_n).$$

Théorème 73 (formule des probabilités totales)

Soient $(\Omega, \mathscr{A}, \mathbf{P})$ un espace probabilisé et $(A_n)_{n \in I}$ un système complet d'événements non négligeables (I est un ensemble fini ou dénombrable). Pour tout événement B,

$$\mathbf{P}(B) = \sum_{n \in I} \mathbf{P}(A_n \cap B) = \sum_{n \in I} \mathbf{P}(A_n) \mathbf{P}_{A_n}(B).$$

Exercice 76

✦ Soient a, b, n, $m \in \mathbb{N}^*$. Considérons deux urnes contenant respectivement a et b boules blanches et n et m boules noires. On suppose de plus que le nombre de boules dans les deux urnes est identique.
On procède de la manière suivante :

- On choisit (de façon équiprobable) une des deux urnes.
- On effectue ensuite des tirages mutuellement indépendants avec remise dans cette **même** urne.

1. Calculer la probabilité d'obtenir une boule blanche au k-ième tirage, puis la probabilité d'obtenir k boules blanches consécutives.
2. Calculer la probabilité d'obtenir une boule blanche au $(k+1)$-ième tirage sachant que l'on a déjà obtenu k boules blanches aux tirages précédents.

3 Variable aléatoire

3.1 Définition

Définition 74 (variable aléatoire réelle)

Étant donné un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$, on appelle **variable aléatoire réelle** toute application X définie sur Ω à valeurs dans $\mathbb{R}$, telle que pour tout $t \in \mathbb{R}$,

$$\{\omega \in \Omega \mid X(\omega) \leqslant t\} \in \mathscr{A}.$$

Notation. Dans la suite, pour I un intervalle de $\mathbb{R}$ et $t \in \mathbb{R}$, on note :

$$[X \in I] = \{\omega \in \Omega \mid X(\omega) \in I\}, \quad [X \leqslant t] = \{\omega \in \Omega \mid X(\omega) \leqslant t\} \quad \text{et} \quad [X = t] = \{\omega \in \Omega \mid X(\omega) = t\}.$$

Ainsi, X est une variable aléatoire si pour tout réel t, l'ensemble $[X \leqslant t]$ est un événement.

Remarques.

- Soient a, $b \in \mathbb{R}$. Si X est une variable aléatoire réelle, alors $\big[X \in]a, b]\big]$ est un événement. Par conséquent, le réel $\mathbf{P}\big(\big[X \in]a, b]\big]\big)$ est bien défini.

- La donnée des probabilités $\mathbf{P}\big(\big[X \in]a, b]\big]\big)$ pour tous réels $a < b$ définit **la loi de probabilité** de la variable aléatoire réelle X.

Proposition 75 (Opérations sur les variables aléatoires)

Soient X, Y deux variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$. Alors

- la combinaison linéaire $\lambda X + \mu Y$ où $\lambda, \mu \in \mathbb{R}$,
- le produit $X \cdot Y$,
- le maximum $\max(X, Y)$ et le minimum $\min(X, Y)$

sont encore des variables aléatoires sur $(\Omega, \mathscr{A}, \mathbf{P})$.

Remarque. Cela s'étend à un nombre fini de variables aléatoires.

3.2 Fonction de répartition

Définition 76 (fonction de répartition)

Soit X, une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$. On définit la fonction F_X sur $\mathbb{R}$ par

$$\forall\, t \in \mathbb{R}, \qquad F_X(t) = \mathbf{P}\big([X \leqslant t]\big).$$

La fonction F_X est la **fonction de répartition** de la variable aléatoire X.

Exercice 77

✧ Soit X, la variable aléatoire donnant la valeur d'un dé équilibrée.
Donner le graphe sur $[-1;7]$ de F_X.

Proposition 77 (propriétés de la fonction de répartition)

Soient X une variable aléatoire réelle définie sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ et F_X sa fonction de répartition.

- Pour tout $t \in \mathbb{R}$, $F_X(t) \in [0;1]$.
- F_X est croissante : pour tous t_1, $t_2 \in \mathbb{R}$, $\quad t_1 \leqslant t_2 \quad \Rightarrow \quad F_X(t_1) \leqslant F_X(t_2)$.
- $\lim\limits_{t\to-\infty} F_X(t) = 0$ et $\lim\limits_{t\to+\infty} F_X(t) = 1$.
- F_X est continue à droite : pour tout $a \in \mathbb{R}$, $\lim\limits_{\substack{t\to a\\ t>a}} F_X(t) = F_X(a)$.
- Pour tous a, $b \in \mathbb{R}$ tels que $a < b$, $\qquad \mathbf{P}(a < X \leqslant b) = F_X(b) - F_X(a)$.

Astuce. À la fin de chaque calcul d'une fonction de répartition, il est toujours utile de vérifier rapidement les trois premiers points.

Proposition 78 (caractérisation de la loi)

Une fonction de répartition caractérise une loi de probabilité. Cela signifie que deux variables aléatoires ont la même fonction de répartition si et seulement si elles ont la même loi.

4 Révisions sur les variables aléatoires discrètes

4.1 Rappels : v.a finies et discrètes

Dans le cas où $X(\Omega)$ est fini ou dénombrable, la définition de la variable aléatoire se simplifie.

Définition 79 (variable aléatoire discrète)

Soit $(\Omega, \mathscr{A}, \mathbf{P})$ un espace probabilisé. Une **variable aléatoire discrète** est une application $X : \Omega \to \mathbb{R}$

telle que
- $X(\Omega) = \{x_i \mid i \in I\}$ où I est une partie finie ou infinie de $\mathbb{N}$,
- pour tout $i \in I$, $[X = x_i]$ est un événement.

Vocabulaire. Donner **la loi** d'une variable aléatoire discrète X signifie donner l'ensemble $X(\Omega)$ des valeurs prises par X et pour chaque $x \in X(\Omega)$, la probabilité $\mathbf{P}\big([X = x]\big)$.

Remarque. À une v.a discrète X est associée le système complet d'événements $\big([X = x]\big)_{x \in X(\Omega)}$. En particulier, il vient

$$\sum_{x \in X(\Omega)} \mathbf{P}\big([X = x]\big) = \mathbf{P}(\Omega) = 1.$$

Définition 80 (indépendance, cas discret)

- Soient X et Y sont deux variables aléatoires discrètes sur un même espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$.
On dit que X et Y sont **indépendantes** si

$$\forall (x, y) \in X(\Omega) \times Y(\Omega), \quad \mathbf{P}\big([X = x] \cap [Y = y]\big) = \mathbf{P}\big([X = x]\big) \cdot \mathbf{P}\big([Y = y]\big).$$

- Les variables aléatoires $X_1, \dots, X_n$ sont **mutuellement indépendantes** si

$$\forall (x_1, \cdots, x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega), \quad \mathbf{P}\left(\bigcap_{i=1}^{n} [X = x_i]\right) = \prod_{i=1}^{n} \mathbf{P}\big([X = x_i]\big).$$

4.2 Lois usuelles

On distinguera bien
- les cas finis : loi uniforme discrète, loi de Bernoulli, loi binomiale.
- les cas infinis dénombrables : loi géométrique, loi de Poisson.

Variable aléatoire certaine

Une variable X est dite **variable aléatoire certaine**, ou **presque sûrement constante** s'il existe un réel c tel que

$$\mathbf{P}\big([X = c]\big) = 1.$$

Loi de Bernoulli

Définition 81 (loi de Bernoulli)

Soit $p \in [0;1]$. La variable aléatoire X suit une **loi de Bernoulli**, noté $X \hookrightarrow \mathscr{B}(p)$, si $X(\Omega) = \{0;1\}$ et

$$\mathbf{P}\big([X = 1]\big) = p \quad \text{et} \quad \mathbf{P}\big([X = 0]\big) = 1 - p.$$

Exemple. Pour $A \in \mathscr{A}$ alors la fonction indicatrice $\mathbf{1}_A$ suit une loi de Bernoulli de paramètre $\mathbf{P}(A)$.

Exemples de modélisation.
⟶ Le résultat d'un lancer d'une pièce de monnaie équilibrée (1 pour pile, 0 face) suit une loi $\mathscr{B}(1/2)$.
⟶ Plus généralement, la variable aléatoire associée à une expérience aléatoire ayant seulement deux issues (0 pour échec, 1 pour succès) suit une loi $\mathscr{B}(p)$ où p est la probabilité de succès.

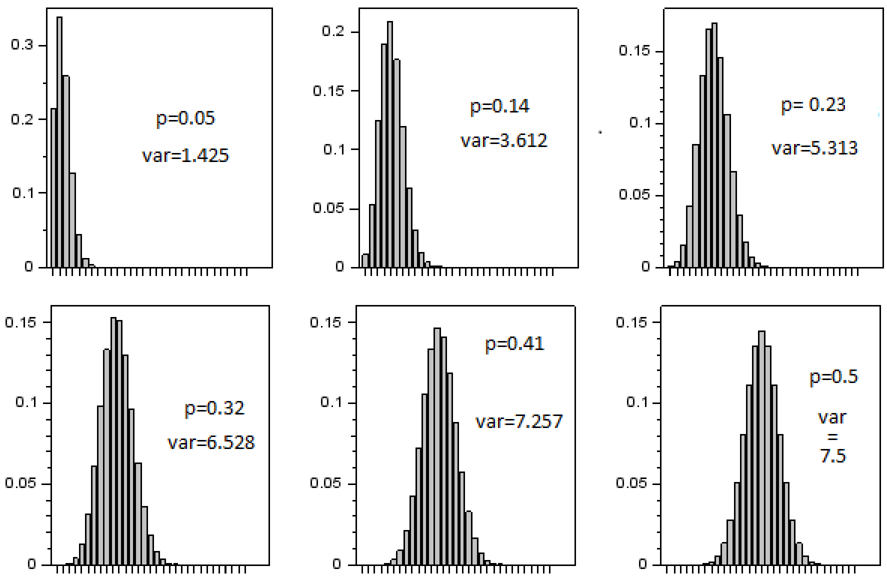
Loi binomiale

Définition 82 (loi binomiale)

Soient $n \in \mathbb{N}^*$, $p \in [0,1]$. On dit que X suit la **loi binomiale** de paramètres n et p, noté $X \hookrightarrow \mathscr{B}(n,p)$, si

$$X(\Omega) = [\![0,n]\!] \quad \text{et} \quad \forall k \in [\![0,n]\!], \quad \mathbf{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Représentation de $\mathscr{B}(n;p)$
Donnons quelques diagrammes en bâtons associés aux lois $\mathscr{B}(n;p)$ avec $n = 30$ et
$p \in \{0,05\,;0,14\,;0,23\,;0,32\,;0,41\,;0,5\}$.



Exemples de modélisation.
⟶ On lance n fois un dé et X compte le nombre de « 6 » obtenus. Alors $X \hookrightarrow \mathscr{B}(n, 1/6)$.
⟶ Plus généralement, lorsqu'on répète n *expériences* de Bernoulli (à deux issues : succès/échec) *identiques, mutuellement indépendantes,* dont la probabilité de succès est p, la variable X qui *compte le nombre de succès* suit alors une loi binomiale de paramètre (n,p).

⚠ **Attention.** Ne pas oublier la condition de mutuelle indépendance.

Loi uniforme

Définition 83 (loi discrète uniforme)

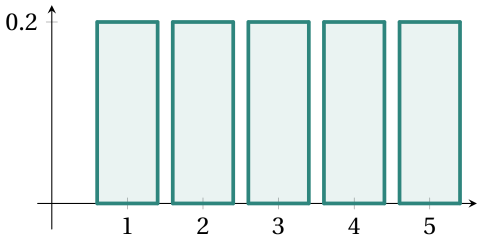
Soient $a, b \in \mathbb{Z}$ avec $a < b$.
La variable aléatoire X suit une **loi discrète uniforme sur** $[\![a,b]\!]$, noté $X \hookrightarrow \mathscr{U}([\![a,b]\!])$, si

$$X(\Omega) = [\![a,b]\!] \quad \text{et} \quad \forall k \in [\![a,b]\!], \quad \mathbf{P}\big([X = k]\big) = \frac{1}{b-a+1}.$$

Remarque. Si $X \hookrightarrow \mathscr{U}([\![1,n]\!])$ avec $n = b-a+1$, alors $Y = X + a - 1 \hookrightarrow \mathscr{U}([\![a,b]\!])$.

Représentation
Ci-contre le diagramme en bâtons de la loi :

$$\mathscr{U}([\![1;5]\!]).$$



Exemples de modélisation.
→ Le résultat d'un lancer d'un dé équilibré à 6 faces suit une loi uniforme sur $[\![1;6]\!]$.
→ Une urne contenant n boules indiscernables au toucher numérotées de 1 à n. On tire au hasard une boule. Le numéro obtenu suit une loi uniforme sur $[\![1;n]\!]$.

Loi géométrique

Définition 84 (loi géométrique $\mathscr{G}(p)$)

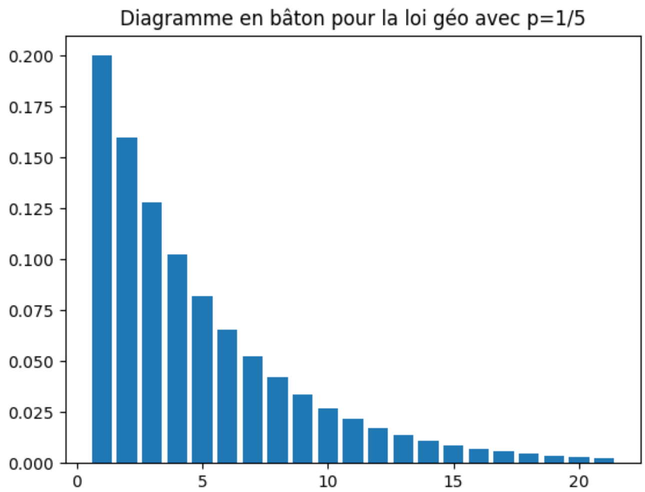
Soit $p \in]0;1[$ et $q = 1-p$.
On dit que la variable aléatoire X suit une **loi géométrique** de paramètre p, notée $\mathscr{G}(p)$, si

$$X(\Omega) = \mathbb{N}^* \quad \text{et} \quad \forall\, k \in \mathbb{N}^*, \quad \mathbf{P}\big([X = k]\big) = (1-p)^{k-1} p = q^{k-1} p.$$

Exercice 78

✦✦ Diagramme en bâton de la loi géométrique

1. Soient $p \in]0,1[$ et $X \hookrightarrow \mathscr{G}(p)$. Écrire un programme qui prend en argument p et renvoie la plus petite valeur entière n_p telle que $\mathbf{P}\big([X \geqslant n_p]\big) \leqslant 1\%$.
2. En déduire un second programme qui prend en argument p renvoie le diagramme en bâton sur $[\![0;n_p]\!]$.



Exemple de modélisation.
Une loi géométrique modélise un premier temps d'arrêt.
Si X renvoie *le rang du premier succès* dans une succession d'expériences de Bernoulli *identiques, mutuellement indépendantes*, alors X suit une loi géométrique où p est la probabilité de succès d'une expérience de Bernoulli.
Une autre caractéristique de la loi géométrique : c'est une loi sans mémoire :

Exercice 79

✦ Loi discrète sans mémoire
Soient $p \in]0;1[$ et X une variable aléatoire de loi $\mathscr{G}(p)$.

1. Justifier que : $\forall\, s, t \in \mathbb{N}^*, \qquad \mathbf{P}_{[X>s]}\big([X > s+t]\big) = \mathbf{P}\big([X > t]\big) \qquad (\bullet)$
2. *Réciproque.*
   Soit X une variable aléatoire sur $\mathbb{N}^*$ vérifiant $(\bullet)$. Posons $p = \mathbf{P}\big([X = 1]\big)$.
   (a) Soit $k \in \mathbb{N}$. Exprimer $\mathbf{P}\big([X > k]\big)$ en fonction de p.
   (b) En déduire X suit une loi géométrique de paramètre p.

Loi de Poisson

Définition 85 (loi de Poisson $\mathscr{P}(\lambda)$)

Soit λ un réel strictement positif.
On dit que la variable aléatoire X suit **une loi de Poisson** de paramètre λ, notée $\mathscr{P}(\lambda)$, si

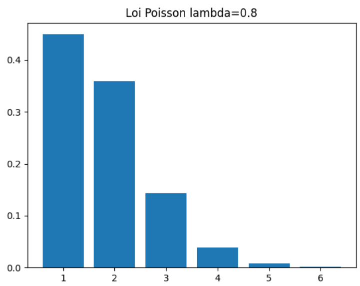
$$X(\Omega) = \mathbb{N} \quad \text{et} \quad \forall\, k \in \mathbb{N}, \quad \mathbf{P}\big([X = k]\big) = e^{-\lambda}\frac{\lambda^k}{k!}.$$

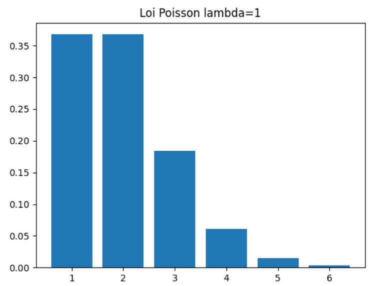
Exercice 80

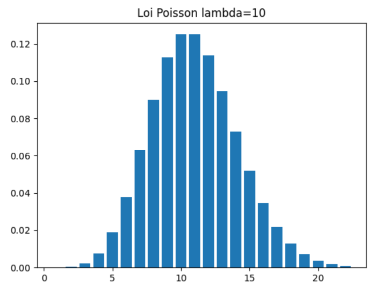
✦✦ **Diagramme en bâton de la loi de Poisson**

1. Soient $\lambda \in]0,+\infty[$ et $X \hookrightarrow \mathscr{P}(\lambda)$. Écrire un programme qui prend en argument n, λ et renvoie la liste
$$[p_0 \quad p_1 \quad \dots \quad p_n] \quad \text{où} \quad p_i = \frac{\lambda^i}{i!}e^{-\lambda}.$$
On pourra remarquer que $p_{i+1} = \lambda p_i/(i+1)$.
2. Écrire un programme qui prend en argument λ et renvoie la plus petite valeur entière n_p telle que $\mathbf{P}\big([X \geqslant n_\lambda]\big) \leqslant 1\%$.
3. En déduire un second programme qui prend en argument p renvoie le diagramme en bâton sur $[\![0; n_\lambda]\!]$.

Quelques exemples pour différentes valeurs du paramètre.







5 Simulation des variables aléatoires discrètes

5.1 Premiers exemples avec les lois usuelles

Pour simuler des variables aléatoires, on peut utiliser la bibliothèque `random` que l'on importe par :

Editeur

```
import numpy.random as rd
```

- **rd.random()**

La commande `rd.random()` renvoie un réel choisi au hasard dans l'intervalle $[0;1[$ suivant une loi de probabilité uniforme continue sur $[0;1[$.
Pour simuler à l'aide de Python un événement de probabilité p, on peut écrire `rd.random() < p` renverra `True` avec une probabilité p.

- **rd.randint(debut,fin)**

De la même façon, on peut tirer au hasard (et uniformément) un entier plutôt qu'un réel. Dans ce cas, la commande à utiliser est `rd.randint(début,fin)` qui tire au hasard avec une probabilité uniforme un entier dans l'intervalle `[début,fin[`.

Remarque. On peut aussi écrire `np.floor((b-a)*random()+a`.

- **rd.geometric(p)**

La fonction `rd.geometric(p)` permet de simuler une variable aléatoire qui suit une loi géométrique de paramètre p. Pour rappel, si X renvoie le rang du premier succès dans une infinité d'expériences de Bernoulli mutuellement indépendantes, $X \hookrightarrow \mathscr{G}(p)$ où p est la probabilité d'un succès.

- **rd.binomial(n,p)**

De la même façon, on peut simuler une variable aléatoire qui suit une loi binomiale de paramètres n, p avec la commande `rd.binomial(n,p)`. Pour rappel, le nombre de succès obtenus lors d'une répétition de n expériences de Bernoulli mutuellement indépendantes de probabilité de succès p suit une loi binomiale de paramètres n, p.

- **rd.poisson(lambda)**

Enfin, `rd.poisson(lambda)` simule une variable qui suit une loi de Poisson de paramètre `lambda`.

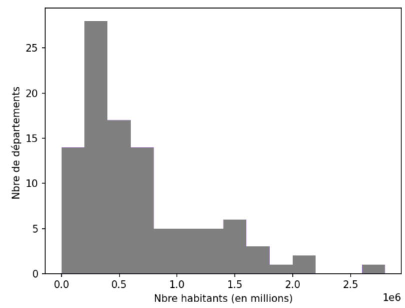
Remarque. Toutes les méthodes précédentes peuvent aussi renvoyer une liste de valeurs tirées selon les différentes lois de probabilité plutôt qu'une seule valeur. Pour faire cela, il suffit de donner comme argument supplémentaire à l'instruction la taille de la liste voulue. Par exemple, `rd.randint(1,100,200)` renvoie un tableau `numpy` contenant 200 entiers pris aléatoirement entre 1 et 100. De plus, la commande `rd.randint(1,100,[200,10])` renvoie une matrice de taille (200,10).

Histogrammes

Rappelons aussi le code pour le tracé d'un histogramme qui permet d'avoir un aperçu de la répartition des valeurs de l'échantillon. Ce graphique est obtenu en traçant, pour chaque $i \in [\![1;p]\!]$, le rectangle de base $[a_i; a_{i+1}]$ sur l'axe des abscisses et en ordonnées, l'effectif de la classe $[a_i; a_{i+1}]$.

Editeur

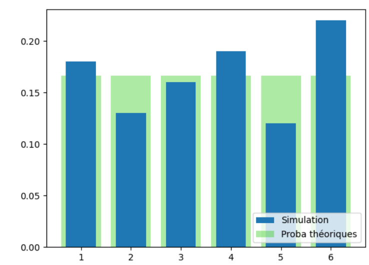
```
# Voici une exemple avec le nombre d'habitants
    par département français
# Création d'un tableau avec les intervalles de
    longueurs 200 000 (habitants)
inter =np.linspace(0,2.8*10**6,15)
plt.hist(L, bins=inter)
plt.xlabel('Nbre habitants (en millions)')
plt.ylabel('Nbre de départements ')
plt.show()
```



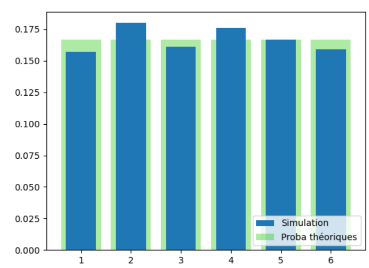
Editeur

```
# La loi de la variable
val=[1,2,3,4,5,6]
loi=[1/6,1/6,1/6,1/6,1/6,1/6]

# Simulation de la variable aléatoire
ech=np.floor(6*np.random.rand(m))+1

# Tracé du diagramme en bâtons
        # bar(abscisses,ordonnées)
plt.bar(val,loi,color=(0.2, 0.8, 0.1, 0.4),label=
    'Proba théoriques')

# Tracé de l'histogramme
        # hist(echantillon, bins=classe)
classe=[0.5,1.5,2.5,3.5,4.5,5.5,6.5]
plt.hist(ech,bins=classe,density='true',rwidth
    =0.6,label='Simulation')

plt.legend( loc = 'lower right')
plt.show()
```

Échantillon de taille $m = 100$:



Échantillon de taille $m = 1000$:



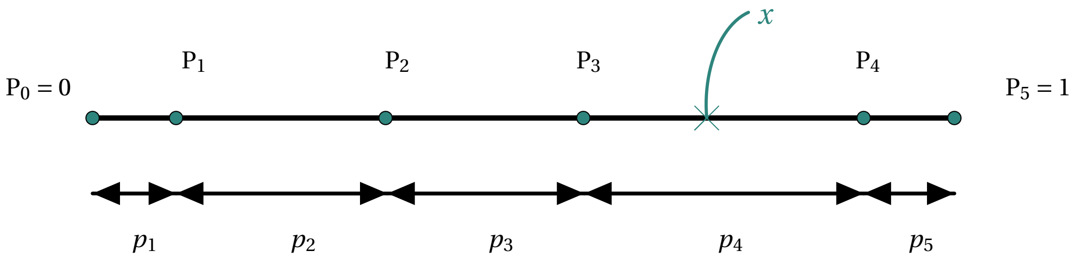
5.2 Simulation pour une loi discrète

Le cas fini

Soit X une variable aléatoire finie. Pour simplifier le programme, on suppose $X(\Omega) = [\![1;n]\!]$. On note

$$\forall\, i \in [\![1;n]\!], \qquad p_i = \mathbf{P}\big([X = i]\big) \quad \text{et} \quad P_i = \sum_{k=1}^{i} p_k.$$

On souhaite simuler la variable X uniquement avec la commande `rand()`. L'idée est la suivante : on simule la loi de X en tirant un réel, au hasard dans $[0;1]$ et en renvoyant l'entier k s'il appartient à l'intervalle $[P_{k-1};P_k[$. La probabilité que l'entier k soit choisi est alors $P_k - P_{k-1} = p_k = \mathbf{P}\big([X = k]\big)$ (avec la convention $P_0 = 0$).



Exercice 81

✦✦✦ Simulation des v.a finies

1. Que représente les réels P_i en termes de probabilités et de X ?
2. Écrire un programme qui prend en argument une liste $(p_i)_{i\in[\![1;n]\!]}$ correspond à la loi de X et simule X.
3. Tester votre programme sur la loi uniforme $\mathscr{U}([\![1;10]\!])$ et vérifier la cohérence du programme à l'aide d'un histogramme.

Le cas infini dénombrable avec la loi de Poisson

On peut toutefois étendre la méthode précédente au cas où X prend ses valeurs dans $\mathbb{N}$. Prenons le cas de $X \sim \mathscr{P}(\lambda)$. Notons pour $k \in \mathbb{N}$,

$$p_k = \mathbf{P}\big([X = k]\big) = e^{-\lambda}\frac{\lambda^k}{k!} \quad \text{et} \quad P_k = \sum_{j=0}^{k} p_j.$$

En particulier, on a

$$P_{k+1} = P_k + \frac{\lambda}{k+1} p_k.$$

Exercice 82

✦✦ Simulation d'une loi de Poisson

1. Adapter la méthode précédente pour simuler $X \hookrightarrow \mathscr{P}(\lambda)$.
2. Comparer l'histogramme obtenu aux valeurs théoriques (diagramme en bâtons construit à l'exercice **??**).

Exercices - TD

Révisions séries

Exercice 83. ✦

1. Montrer que pour tout n de $\mathbb{N}$, $(2+\sqrt{3})^n + (2-\sqrt{3})^n$ est un entier.
2. En déduire que la série de terme général $\sin\left(\pi(2+\sqrt{3})^n\right)$ est absolument convergente.

Exercice 84. Représenter dans le plan muni d'un repère $(O, \vec{\imath}, \vec{\jmath})$ l'ensemble des points de coordonnées (a, b) telles que la série de terme général $u_n = \frac{n^a}{n^2+n^b}$ soit convergente.

Exercice 85. ✦✦ **Formule du binôme négatif** *D'après EDHEC 2015*

1. Pour $p \in \mathbb{N}$ et $x \in]-1;1[$, vérifier l'équivalent $\displaystyle\binom{k}{p} \underset{k\to+\infty}{\sim} \frac{k^p}{p!}$. En déduire la convergence de la série $\sum\limits_{k=p}^{+\infty} \binom{k}{p} x^k$.
2. On définit, pour $x \in]-1;1[$, $\quad S_p(x) = \sum\limits_{k=p}^{+\infty} \binom{k}{p} x^k$.
   (a) Préciser $S_0(x)$.
   (b) À l'aide de la formule du triangle de Pascal, démontrer que $(1-x)S_{p+1}(x) = xS_p(x)$.
   (c) Conclure avec
$$\sum_{k=p}^{+\infty} \binom{k}{p} x^k = \frac{x^p}{(1-x)^{p+1}}.$$

Exercice 86. ✦✦ **Produit de Cauchy** *D'après EDHEC 2020*

On considère deux suites réelles $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in}$ à termes positifs et on suppose que les séries de termes généraux a_n et b_n sont convergentes, de sommes respectives $A = \sum\limits_{n=0}^{+\infty} a_n$ et $B = \sum\limits_{n=0}^{+\infty} b_n$. Pour tout entier naturel n, on pose : $c_n = \sum\limits_{k=0}^{n} a_k b_{n-k}$.

1. Montrer que : $\forall n \in \mathbb{N}$, $\sum\limits_{k=0}^{n} c_k \leqslant \left(\sum\limits_{k=0}^{n} a_k\right)\left(\sum\limits_{k=0}^{n} b_k\right) \leqslant \sum\limits_{k=0}^{2n} c_k$.
2. En déduire que la série de terme général c_n converge et que l'on a $\displaystyle\sum_{n=0}^{+\infty} c_n = \left(\sum_{n=0}^{+\infty} a_n\right)\left(\sum_{n=0}^{+\infty} b_n\right)$.
3. *Application*
   Soit $(a, b) \in \mathbb{R}^{+2}$. En utilisant le résultat précédent, démontrer que $\displaystyle\sum_{k=0}^{+\infty} \frac{(a+b)^k}{k!} = \left(\sum_{k=0}^{+\infty} \frac{a^k}{k!}\right) \cdot \left(\sum_{k=0}^{+\infty} \frac{b^k}{k!}\right)$.

Exercice 87. ✦✦ **Comparaison série-intégrale, une série de Bertrand**

Prouver que la série $\displaystyle\sum \frac{1}{n\left(\ln(n)\right)^2}$ est convergente.

Exercice 88. ✦✦ *D'après Oraux ESCP 2009*

Soient $(a_n)_{n\in\mathbb{N}}$, une suite réelle décroissante de limite nulle. Pour tout n de $\mathbb{N}^*$, on pose : $b_n = n(a_{n-1} - a_n)$.

1. Montrer que pour tout n de $\mathbb{N}^*$, on a : $\sum\limits_{k=1}^{n} b_k = \left(\sum\limits_{k=0}^{n-1} a_k\right) - na_n$.
2. On suppose dans cette question que la série de terme général a_n converge.
   (a) Montrer que $\lim\limits_{n\to+\infty} na_n = 0$.
   (b) En déduire que la série de terme général b_n converge et que $\sum\limits_{n=1}^{\infty} b_n = \sum\limits_{n=0}^{+\infty} a_n$.
3. On suppose dans cette question que la série de terme général b_n converge.
   (a) Montrer que pour tout n et k de $\mathbb{N}^*$, on a : $n\left(a_n - a_{n+k}\right) \leqslant \sum\limits_{j=n+1}^{n+k} b_j$.
   (b) En déduire que la série de terme général a_n converge et que $\sum\limits_{n=0}^{+\infty} a_n = \sum\limits_{n=1}^{+\infty} b_n$.

Probabilités

Exercice 89. ✦✦
Soit $n \in \mathbb{N}^*$. Considérons n urnes numérotées de 1 à n, contenant chacune r boules bleues et s boules rouges $(1 < s)$. On réalise l'expérience suivante. On choisit au hasard une première boule dans la première urne que l'on replace dans la seconde et on répète l'opération jusqu'à la dernière urne.
Quelle est la probabilité qu'une boule tirée au hasard dans la dernière urne soit bleue ?

Exercice 90. ✦✦✦ **Énigme : Monty-Hall**
Dans le film Las Vegas 21, Micky Rosa est professeur au M.I.T. Il soumet à ces étudiants, dont Ben, le fameux problème du Monty-Hall.

Micky Rosa : Ben, vous participez à un jeu télévisé et on vous donne la possibilité de choisir entre trois portes différentes, d'accord ? Derrière l'une des trois portes, il y a une voiture neuve, et derrière les deux autres, des chèvres. Quelle porte choisiriez-vous, Ben ?

Micky Rosa (Kevin Spacey)

Ben : Porte $n°1$?

Micky Rosa : Ben choisit la porte $n°1$ et là, l'animateur de jeu télé qui, soit dit en passant, sait ce qu'il y a derrière chacune des trois portes, décide d'ouvrir une autre des trois portes ; disons qu'il choisit la $n°3$, derrière laquelle se trouve une chèvre. Bien. Ben... L'animateur s'approche et dit "Ben, vous voulez garder la porte $n°1$ ou prendre la porte $n°2$?". Alors, est-il dans votre intérêt de changer de choix ?

Ben : Oui.

Micky Rosa : Attendez. Souvenez-vous : l'animateur sait où est la voiture. Alors, comment être sûr qu'il ne joue pas avec vous, qu'il n'essaie pas de vous déstabiliser pour vous faire choisir la chèvre ?

Est-ce que Ben a raison de changer de porte ?

Variables aléatoires discrètes

Exercice 91. ✧ Pour chaque énoncé, proposer une loi pour la variable aléatoire réelle X. On précisera les paramètres de la loi et les éventuelles hypothèses.

1. On lance un dé à 6 faces et on note X la variable égale au nombre obtenu.
2. On lance 10 fois un dé à 6 faces et on note X la variable égale au nombre de numéro pair obtenu.
3. Une urne contient 15 boules (5 noires, 3 blanches et 7 rouges). On tire successivement et avec remise 20 boules et on note X la variable égale au nombre de boules noires.
4. Un jeu de 52 cartes est aligné, faces cachées, sur une table de façon aléatoire. On découvre les cartes, de gauche à droite jusqu'à obtenir la dame de cœur. Soit X la variable égale au nombre de cartes découvertes.
5. Une urne contient n jetons numérotés de 1 à n. On tire au hasard un à un avec remise jusqu'à obtenir le jeton 1. On note X la variable égale au nombre de tirages effectués.
6. On pose m questions à un étudiant. Pour chaque question, n réponses sont possibles dont une seule uniquement est correcte. L'élève répond au hasard. Soit X la variable égale au nombre de bonnes réponses.

Exercice 92. ✦✦ **Matrices de rang 1 et indépendance de variables aléatoires** *D'après EMLyon 2013*

1. Montrer que $M \in \mathcal{M}_n(\mathbb{R})$ est de rang 1 si, et seulement si, il existe deux matrices colonnes U, V telles que $M = U\,{}^t V$.
2. On considère deux variables aléatoires X et Y définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$. On suppose de plus : $X(\Omega) = Y(\Omega) = [\![1; n]\!]$. On note, pour tout $(i, j) \in [\![1; n]\!]^2$,
$$m_{i,j} = \mathbf{P}\big([X = i] \cap [Y = j]\big),$$
puis $$M = \left(m_{i,j}\right)_{i,j} \in \mathcal{M}_n(\mathbb{R}), \quad U_X = \Big(\mathbf{P}\big([X = i]\big)\Big)_{1 \leqslant i \leqslant n} \in \mathcal{M}_{n,1}(\mathbb{R}) \quad \text{et} \quad U_Y = \Big(\mathbf{P}\big([Y = i]\big)\Big)_{1 \leqslant i \leqslant n} \in \mathcal{M}_{n,1}(\mathbb{R}).$$
   (a) On suppose, dans cette question, que les variables aléatoires X et Y sont indépendantes.
   Calculer $U_X\,{}^t U_Y$. En déduire que la matrice M est de rang 1 .
   (b) On suppose, dans cette question, que la matrice M est de rang 1. Notons $C_1, C_2, \ldots, C_n$, les colonnes de M.
      i. Vérifier que $C_1 + \cdots + C_n = U_X$.
      ii. En déduire que, pour tout $j \in [\![1; n]\!]$, il existe $\beta_j \in \mathbb{R}$ tel que $C_j = \beta_j U_X$.
      iii. Montrer que pour tout $j \in [\![1; n]\!]$, $\mathbf{P}\big([Y = j]\big) = \beta_j$.
      iv. En déduire que les variables aléatoires X et Y sont indépendantes.

Exercice 93. ✦✦ **Étude asymptotique de la queue d'une loi de Poisson.** *HEC 2012*

Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$, suivant une loi de Poisson de paramètre $\lambda \in \mathbb{R}_*^+$.

1. Montrer que pour tout $n \geqslant \lambda - 1$, $\mathbf{P}(X \geqslant n) \leqslant \mathbf{P}(X = n)\frac{n+1}{n+1-\lambda}$.
2. En déduire que $\mathbf{P}(X \geqslant n) \underset{n\to+\infty}{\sim} \mathbf{P}(X = n)$.
3. Montrer que $\mathbf{P}(X > n) = \underset{n\to+\infty}{o}\big(\mathbf{P}(X = n)\big)$.

– – –

Soit X *une variable aléatoire réelle discrète. Dans la suite, on appelle* ***mode*** *de* X*, toute valeur* $x_0 \in X(\Omega)$ *telle que*

$$\forall\, x \in X(\Omega), \qquad \mathbf{P}\Big([X = x]\Big) \leqslant \mathbf{P}\big([X = x_0]\big).$$

Exercice 94. ✦ Comment écrire un programme qui prend en argument la liste des probabilités $\Big(\mathbf{P}\big([X = x]\big)\Big)_{x\in X(\Omega)}$ et renvoie le ou les modes de X.

Exercice 95. ✦✦✦ **Mode(s) de la loi binomiale** $\mathscr{B}(n;p)$

1. Soit $n \in \mathbb{N}^*$. Soit $p \in]0,1[$. Montrer que, selon les valeurs de n et p, la loi binomiale $\mathscr{B}(n;p)$ admet un ou deux modes que l'on précisera.
2. Pour $n = 9$, illustrer par des représentations graphiques en Python les différents cas obtenus.

Exercice 96. ✦✦✦ **Mode(s) de la loi de Poisson**

1. Soit $\lambda \in \mathbb{R}_+^*$. Montrer que, selon la valeur de λ, la loi de Poisson $\mathscr{P}(\lambda)$ admet un ou deux modes que l'on précisera.
   Indication. On distinguera 3 cas : $\lambda \in]0,1]$, $\lambda \in \mathbb{N} \setminus \{0;1\}$ et $\lambda \in]1,+\infty[\setminus\mathbb{N}$.
2. Illustrer par des représentations graphiques en Python les différents cas obtenus.

CHAPITRE 5

Fonctions de plusieurs variables

La science consiste à passer d'un étonnement à un autre.

ARISTOTE
Philosophe grec de l'Antiquité (384-322 av. J.-C), disciple de Platon

1 Définitions et exemples

1.1 Norme euclidienne

Définition 86 (norme euclidienne)

Pour tout $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, on définit la **norme euclidienne** de x par

$$\|x\| = \sqrt{\sum_{i=1}^{n} {x_i}^2}.$$

Règles de calculs

- $\forall\, x \in \mathbb{R}^n, \qquad \|x\| = 0 \iff x = 0_{\mathbb{R}^n}.$
- Homogénéité : $\forall\, x \in \mathbb{R}^n, \quad \forall \lambda \in \mathbb{R}, \qquad \|\lambda x\| = |\lambda| \cdot \|x\|.$
- Inégalité triangulaire : $\forall\, x,\, y \in \mathbb{R}^n, \qquad \|x + y\| \leqslant \|x\| + \|y\|.$
- Inégalité de Cauchy-Schwarz : $\forall\, x,\, y \in \mathbb{R}^n, \qquad \left|\sum_{i=1}^{n} x_i y_i\right| \leqslant \|x\| \cdot \|y\|.$
  Il y a égalité si et seulement si les vecteurs x et y sont colinéaires.

1.2 Fonctions définies sur $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$

Exemples.
- Avec les fonctions usuelles :

$$f : (x, y) \in \mathbb{R}^2 \mapsto x^2 + xy - y^3, \quad g : (x, y, z) \in \mathbb{R}^3 \mapsto \ln\left(1 + x^2 e^{yz}\right), \quad h : x \in \mathbb{R}^n \mapsto \|x\|.$$

- **Les fonctions polynomiales**.

Une fonction polynomiale est une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$, combinaison linéaire de fonctions de la forme

$$(x_1, \ldots, x_n) \in \mathbb{R}^n \mapsto {x_1}^{\alpha_1} {x_2}^{\alpha_2} \ldots {x_n}^{\alpha_n},$$

où $\alpha_1, \alpha_2, \ldots, \alpha_n$ sont des entiers naturels. Par exemple • **Les fonctions affines**
Une fonction affine est une fonction polynomiale de $\mathbb{R}^n$ dans $\mathbb{R}$ de la forme

$$f : (x_1, \ldots, x_n) \in \mathbb{R}^n \mapsto a_1 x_1 + a_2 x_2 \ldots + a_n x_n + b$$

où $a_1, \ldots, a_n, b \in \mathbb{R}$. Par exemple

1.3 Graphes et lignes de niveau

Graphes

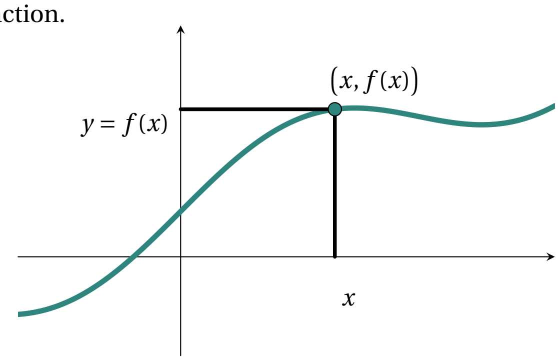
Définition 87 (graphe)

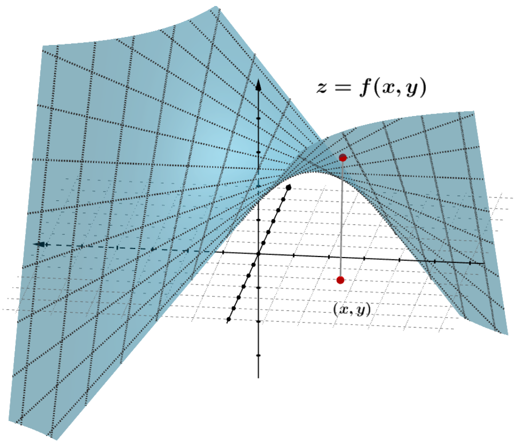
Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction. Le **graphe** de f est la partie de $\mathbb{R}^{n+1}$ définie par

$$\left\{ (x_1, \ldots, x_n, y) \in \mathbb{R}^{n+1} \,\middle|\, y = f(x_1, \ldots, x_n) \right\}.$$

Représentations

• Graphe des fonctions d'une variable réelle
Pour $n = 1$, on retrouve la courbe représentative d'une fonction.





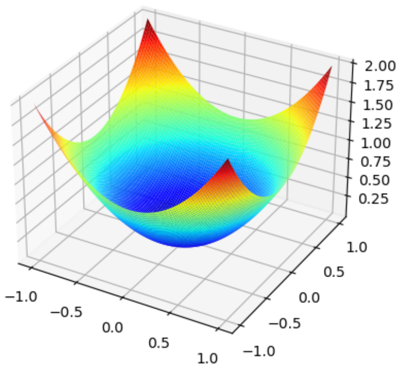
• Graphe des fonctions de deux variables
Ci-contre, la surface représentative de la fonction f : $(x, y) \in \mathbb{R}^2 \mapsto xy/10 + 5$.

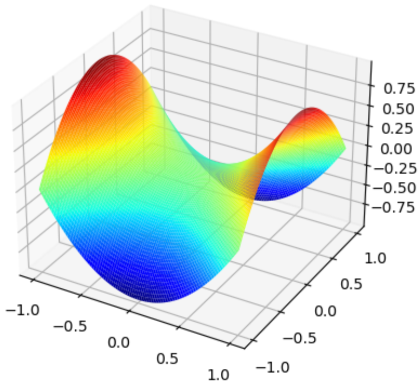
Suivant le même principe que pour le tracé de la courbe représentative d'une fonction d'une variable réelle, le code Python suivant permet de tracer la surface représentative d'une fonction de deux variables.

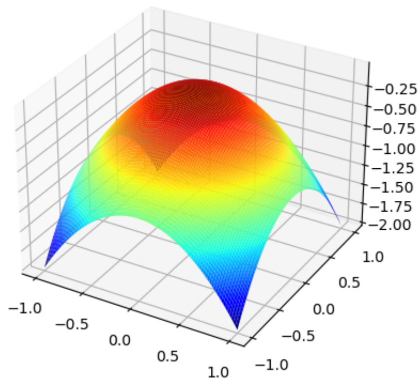
Editeur

```
import numpy as np
import matplotlib.pyplot as plt

def f(x,y):
    return x**2+y**2
 # Définition de la fonction f

x = np.linspace(-1, 1, 100)
 # 100 valeurs pour la variable x espacées régulièrement entre -1 et 1
y = np.linspace(-1, 1, 100) # De même pour la variable y

X, Y = np.meshgrid(x, y)
 # Tableau contenant les points (xi,yi) où xi et yi sont calculés précédemment
Z = f(X,Y)
 # Calcul des images pour tous les points (xi,yi)

fig = plt.figure()
ax = plt.axes(projection='3d')
ax.plot_surface(X, Y, Z, rstride=1, cstride=1,cmap='jet', edgecolor='none') # paramètres d
    'affichage
```

Exemples. Illustrons ce code à l'aide de trois exemples typiques de fonctions polynomiales de degré 2.







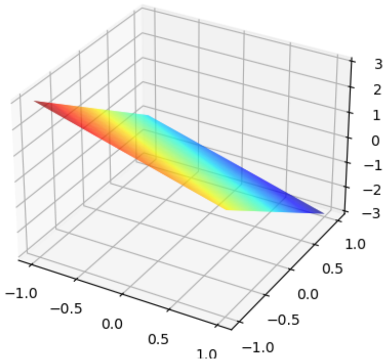
Editeur

```
def f1(x, y):
    return x**2 + y**2
```

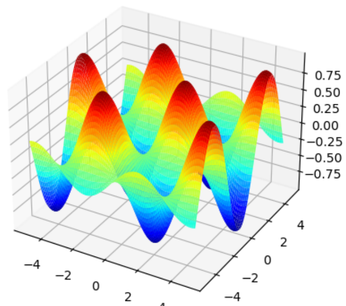
```
def f2(x, y):
    return x**2 - y**2
```

```
def f3(x, y):
    return -x**2 - y**2
```

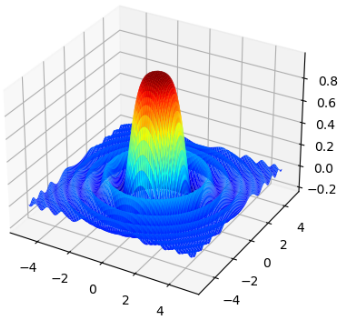
Exemple. Cas des fonctions affines.
Pour $n = 1$, le graphe d'une fonction affine sur $\mathbb{R}$ correspond à une droite.
Pour $n = 2$, le graphe d'une fonction affine sur $\mathbb{R}^2$ correspond à un plan d'équation $z = ax + by + c$.
Dans le cas général, on obtient "un hyperplan affine", c'est-à-dire un translaté d'un hyperplan vectoriel de $\mathbb{R}^{n+1}$.
Ci-contre, le plan d'équation $z + x + 2y = 0$ obtenu à parti de la fonction f définies sur $\mathbb{R}^2$ par $f(x, y) = -x - 2y$.



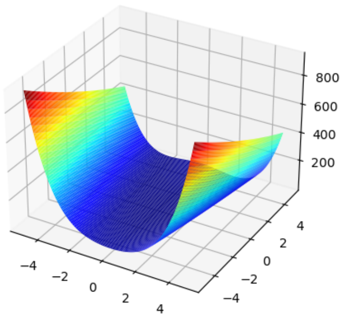
Exemples. Donnons de nouveaux graphes pour illustrer la diversité des cas possibles.



$(x, y) \in \mathbb{R}^2 \mapsto \sin(x)\cos(y)$



$(x, y) \in \mathbb{R}^2 \mapsto \dfrac{\sin(x^2 + y^2)}{x^2 + y^2 + 0.01}$



$(x, y) \in \mathbb{R}^2 \mapsto (1 - x)^2 + (y - x^2)^2$

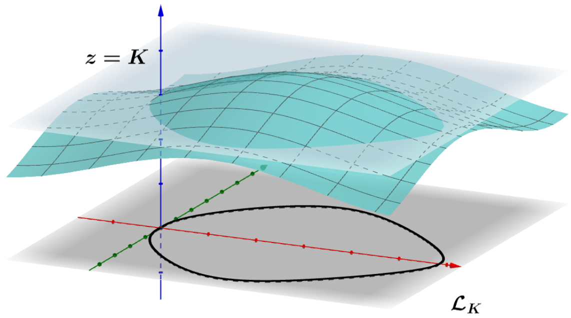
Lignes de niveau

Définition 88 (lignes de niveau)

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction.
La **ligne de niveau** de f associée au réel K est la partie de $\mathbb{R}^2$ définie par

$$\mathscr{L}_{\mathrm{K}} = \left\{ (x, y) \in \mathbb{R}^2 \mid f(x, y) = \mathrm{K} \right\}.$$

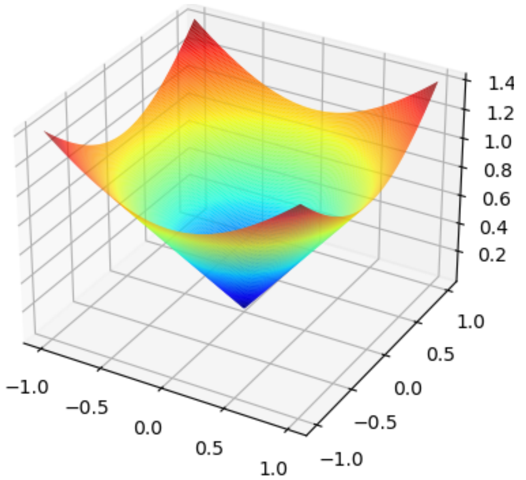
Pour obtenir la courbe de niveau $\mathscr{L}_K$, il suffit d'intersection le plan horizontal $z = K$ avec la surface définie par la fonction.

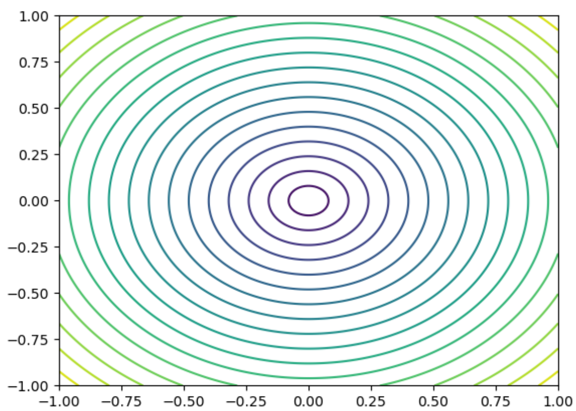


Ci-dessous, un code pour tracer la surface et les lignes de niveau.

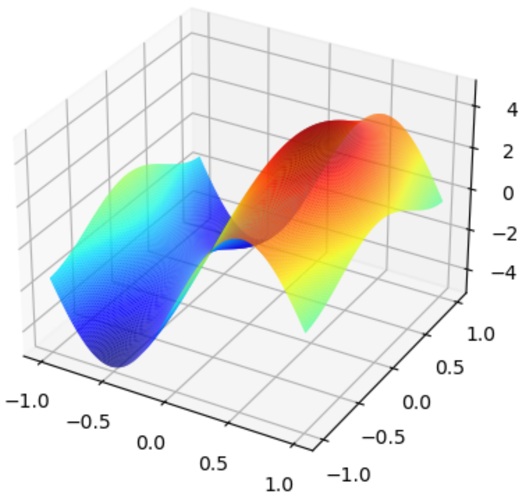
Editeur

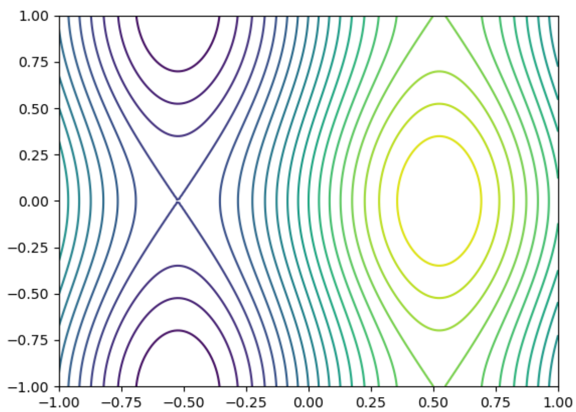
```
def f(x,y):
    return np.sqrt(x**2+y**2)

x=np.linspace(-1,1,200)
y=np.linspace(-1,1,200)
X , Y = np.meshgrid(x,y)
Z = f(X,Y)

graphe = plt.contour(X,Y,Z,20) # 20 pour obtenir 20 lignes de niveau
fig = plt.figure()
ax = plt.axes(projection='3d')
ax.plot_surface(x,y,z, rstride=1, cstride=1,cmap='jet', edgecolor='none')
plt.show()
```





- Donnons un deuxième exemple avec la fonction g définie sur $\mathbb{R}^2$ par $g(x, y) = 4\sin(x) + \cos(3y)$.





Remarque. Les symétries

I	$\forall (x,y) \in \mathbb{R}^2$	$f(x,y) = f(-x,-y)$	II	$\forall (x,y) \in \mathbb{R}^2$	$f(x,y) = -f(-x,-y)$
III	$\forall (x,y) \in \mathbb{R}^2$	$f(x,y) = f(y,x)$	IV	$\forall (x,y) \in \mathbb{R}^2$	$f(x,y) = f(-x,y)$
V	$\forall (x,y) \in \mathbb{R}^2$	$f(x,y) = -f(-x,y)$	VI	$\forall (x,y) \in \mathbb{R}^2$	$f(x,y) = f(x,-y)$
VII	$\forall\, a \in \mathbb{R}^2$	$f(a) = \varphi(\Vert a \Vert)$	avec	$\varphi : \mathbb{R} \to \mathbb{R}$.	

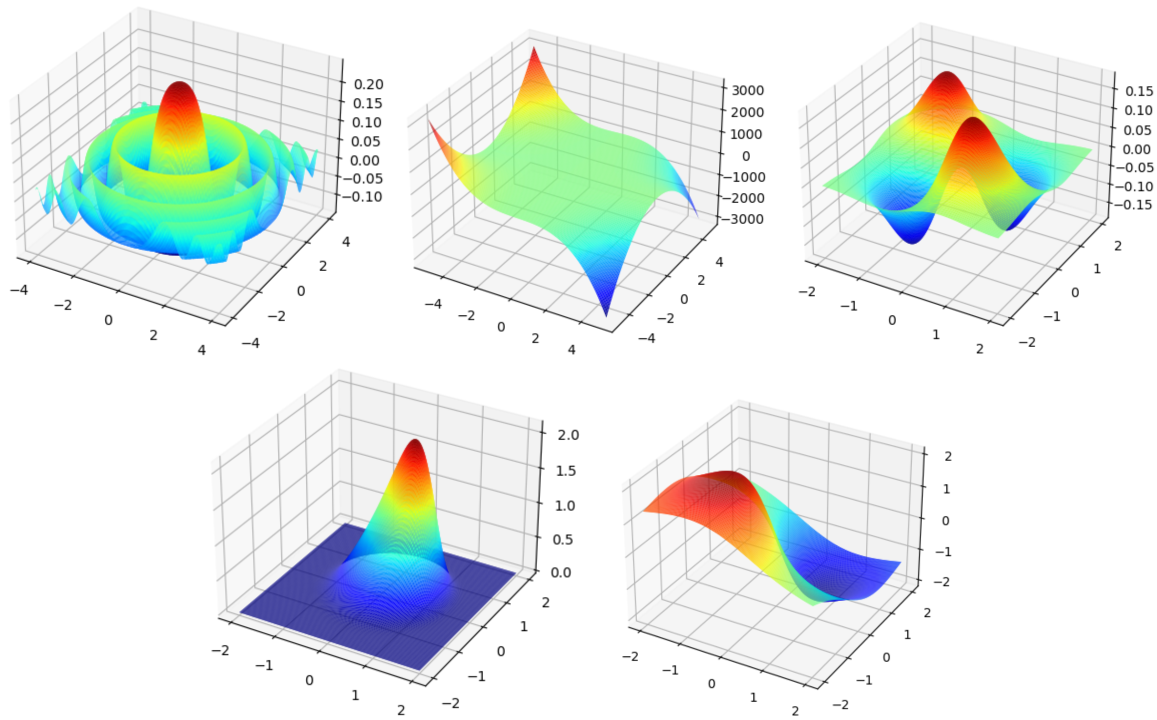
Exercice 97

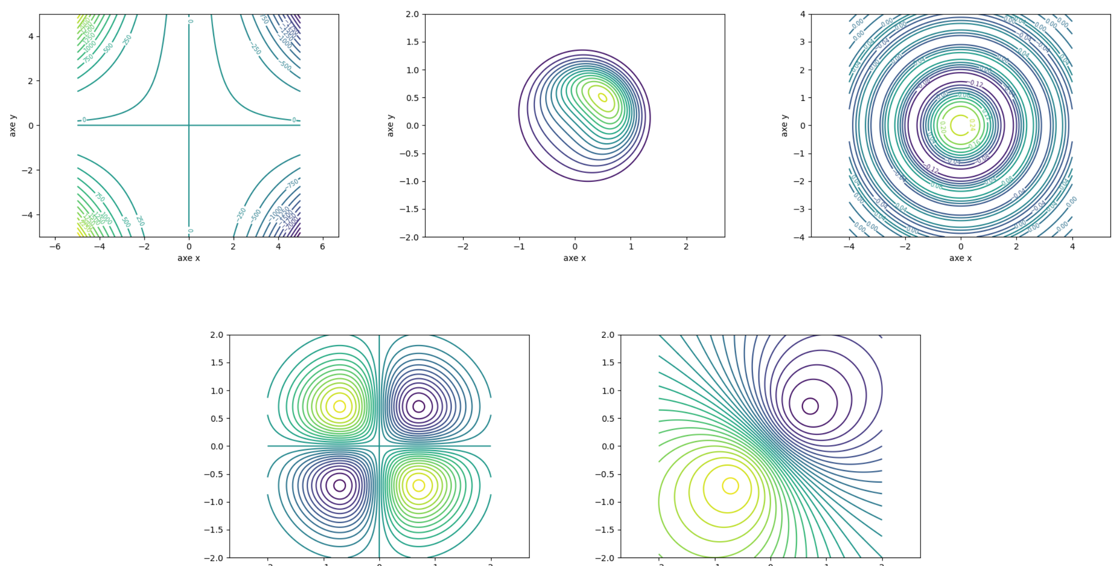
✧ Pour chacune des fonctions suivantes, préciser si la fonction vérifie une des symétries parmi I à VII.

$$f : (x,y) \mapsto \frac{\cos(x^2+y^2)}{4+x^2+y^2}, \qquad g : (x,y) \mapsto 5xy - x^3y^2, \quad h : (x,y) \in \mathbb{R}^2 \mapsto \mathrm{e}^{-(x^2+y^2)^2+x+y}$$

$$i(x,y) = -xy\exp(-x^2-y^2), \quad j(x,y) = -3*(x+y)/(1+x^2+y^2).$$

Associer à chaque fonction son graphe et ses lignes de niveau. Comment traduire géométriquement les symétries ?





1.4 Extrema

Définition 89 (extrema)

Soit $f : \mathbb{R}^n \to \mathbb{R}$, une fonction.
- On dit que f admet un **maximum** en $a \in \mathbb{R}^n$ si : $\forall\, x \in \mathbb{R}^n, \quad f(x) \leqslant f(a)$.
Dans ce cas, on dit que $f(a)$ est le maximum de f sur $\mathbb{R}^n$.
- On dit que f admet un **minimum** en $a \in \mathbb{R}^n$ si : $\forall\, x \in \mathbb{R}^n, \quad f(x) \geqslant f(a)$.
Dans ce cas, on dit que $f(a)$ est le maximum de f sur $\mathbb{R}^n$.

Attention. Il ne faut pas confondre le maximum $f(a)$ qui est unique, avec la valeur en laquelle le maximum est atteint (ici a) qui n'est pas unique.

Exemple. Pour $\alpha, \beta \in \mathbb{R}^+$, on définit la fonction de Rosenbrock par

$$f(x,y) = (\alpha - x)^2 + \beta\left(y - x^2\right)^2$$

On a un minimum atteint en $a = \left(\alpha, \alpha^2\right)$ avec $f(a) = 0$.

Exercice 98

✦ Déterminer les extrema (s'ils existent) des fonctions suivantes définies sur $\mathbb{R}^3$ ou $\mathbb{R}^2$.

$$\begin{array}{rclrcl} f_1(x,y,z) & = & 2x^2+4xy+2y^2+3z^2-1, & f_2(x,y,z) & = & 2xy+y^2+2xz-z^2, \\ f_3(x,y,z) & = & \mathrm{e}^{2xy+y^2+2xz-z^2}, & f_4(x,y) & = & \mathrm{e}^{1-x^2-y^2}, \\ f_5(x,y) & = & \sin(x)\cos(y), & f_6(x,y) & = & \sin(x+y)\cos(x-y), \\ f_7(x,y) & = & \sin(x+y)\cos(x+y). & & & \end{array}$$

Nous verrons aux chapitres suivants, une méthode plus systématique de recherche d'extrema.

2 Continuité des fonctions de plusieurs variables

2.1 Définitions et exemples

Définition 90 (continuité en un point, sur $\mathbb{R}^n$)

- Une fonction f, définie sur $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$, est continue au point $a \in \mathbb{R}^n$ si :

$$\forall\, \varepsilon \in \mathbb{R}_*^+, \quad \exists\, \alpha \in \mathbb{R}_*^+, \quad \forall\, x \in \mathbb{R}^n, \qquad \Big(\ \|x-a\| \leqslant \alpha \quad \Rightarrow \quad \left|f(x) - f(a)\right| \leqslant \varepsilon\ \Big).$$

- Une fonction f est continue sur $\mathbb{R}^n$ si et seulement si f est continue en tout point de $\mathbb{R}^n$.

Remarque. C'est exactement la même définition que pour les fonctions d'une variable à l'exception que la valeur absolue a été généralisée par la norme.

Exemples.
- Continuité de la norme.
- Continuité des applications coordonnées.

2.2 Opérations sur les fonctions continues

Proposition 91 (somme, produit, quotient)

Soient $f, g : \mathbb{R}^n \to \mathbb{R}$ et $\lambda \in \mathbb{R}$.

Si f et g sont continues sur $\mathbb{R}^n$ avec g ne s'annulant pas dans le dernier cas,

alors
- La somme $f+g$ est continue sur $\mathbb{R}^n$,
- La fonction $\lambda \cdot f$ est continue sur $\mathbb{R}^n$,
- Le produit $f \cdot g$ est continue sur $\mathbb{R}^n$,
- Le quotient f/g est continue sur $\mathbb{R}^n$.

Application. Comme les fonctions polynomiales s'obtiennent pas somme et produit des fonctions coordonnées (qui sont continues sur $\mathbb{R}^n$), on démontre ainsi la continuité des fonctions polynomiales.

Proposition 92 (composition)

Soient $\varphi : \mathrm{I} \to \mathbb{R}$ avec $\mathrm{I} \subset \mathbb{R}$ et $f : \mathbb{R}^n \to \mathbb{R}$.

Si
- f continue sur $\mathbb{R}^n$ et à valeurs dans I, c'est-à-dire, pour tout $x \in \mathbb{R}^n$, $f(x) \in \mathrm{I}$.
- φ est continue sur I.

Alors $\varphi \circ f : x \in \mathbb{R}^n \mapsto \varphi\big(f(x)\big) \in \mathbb{R}$ est continue sur $\mathbb{R}^n$.

Méthode

Rédaction d'une continuité

Posons $g : (x,y,z) \in \mathbb{R}^3 \mapsto z\sqrt{(xy)^2 + 3z^4}$.
La fonction $(x,y,z) \in \mathbb{R}^3 \mapsto (xy)^2 + 3z^4$ est continue sur $\mathbb{R}^3$ en tant que fonction polynomiale à valeurs dans $\mathbb{R}^+$.
La fonction racine carrée : $t \in \mathbb{R}^+ \mapsto \sqrt{t}$ est continue sur $\mathbb{R}^+$. Par composition, $(x,y,z) \in \mathbb{R}^3 \mapsto \sqrt{(xy)^2 + 3z^4}$ est continue sur $\mathbb{R}^3$.
De plus, $(x,y,z) \in \mathbb{R}^3 \mapsto z$ est continue et, par produit, g est continue sur $\mathbb{R}^3$.

Exercices - TD

Représentations graphiques

Exercice 99. ✧ Tracer les lignes de niveaux des fonctions définies sur $\mathbb{R}^2$.

$$f : (x, y) \in \mathbb{R}^2 \mapsto 2x - 3 + 1, \quad g(x, y) \in \mathbb{R}^2 \mapsto x^2 + y^2 - 5, \quad h(x, y) \in \mathbb{R}^2 \mapsto \mathrm{e}^{y - x^2}, \quad i(x, y) \in \mathbb{R} \mapsto xy.$$

Recherche d'extrema

Exercice 100. ✦ On définit la fonction f sur $\mathbb{R}^2$ par $f(x, y) = \mathrm{e}^x \left(x\mathrm{e}^x + (yx)^2\right)$.

1. Est-ce que f possède un maximum sur $\mathbb{R}^2$?
2. Étudier la fonction g définie sur $\mathbb{R}$ par $g(x) = f(x, 0)$.
3. En déduire un minimum pour f sur $\mathbb{R}^2$.

Exercice 101. ✦ Soit f la fonction définie sur $\mathbb{R}^2$ par $f(x, y) = x^4 + y^4 - 4xy$.

1. Comparer $f(-x, -y)$ et $f(x, y)$.
2. Soit $y \in \mathbb{R}^+$, montrer que la fonction $x \in \mathbb{R} \mapsto f(x, y)$ admet un minimum sur $\mathbb{R}$. On le note $g(y)$.
3. Étudier la fonction g et vérifier que f admet un minimum sur $\mathbb{R}^2$ atteint eu deux points.

Exercice 102. ✦✦✦

1. Étudier les extrema de $f : x = (x_1, \ldots, x_n) \in \mathbb{R}^n \mapsto \left(\sum_{i=1}^{n} {x_i}^2\right) \prod_{i=1}^{n} \mathrm{e}^{-{x_i}^2}$.
2. Soit $a = (a_i) \in \mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}$. Étudier les extrema de $f_a : x \in \mathbb{R}^n \mapsto \left(\sum_{i=1}^{n} a_i x_i\right)^2 \prod_{i=1}^{n} \mathrm{e}^{-{x_i}^2}$.

Continuité

Exercice 103. ✧ Justifier la continuité des fonctions suivantes :

$$f : x \in \mathbb{R}^n \mapsto \|x\|^{\|x\|} \quad \text{et} \quad g : (x, y) \in \mathbb{R}^2 \mapsto \int_{x+y^2}^{y^2} \exp(t^2)\,\mathrm{d}t$$

On choisit dans le premier cas, la convention $0^0 = 1$.

Exercice 104. ✦✦ Soit $h : \mathbb{R} \to \mathbb{R}$ continue.
Montrer que l'application $f : x \in \mathbb{R}^n \mapsto h\big(\sin(\|x\|)\big)$ admet un minimum et un maximum.

Exercice 105. ✦ Soit $\mathrm{C} = \left\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\right\}$ et f une application continue de $\mathbb{R}^2$ dans $\mathbb{R}$.

1. En considérant l'application $g : [0, 2\pi] \to \mathbb{R}$ définie par

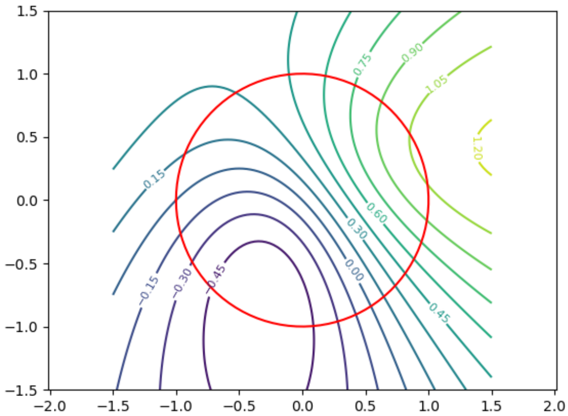
$$g(t) = f\big(\cos(t), \sin(t)\big) - f\big(-\cos(t), -\sin(t)\big),$$

justifier qu'il existe $(x_0, y_0) \in \mathrm{C}$ tel que $f(x_0, y_0) = f(-x_0, -y_0)$.

2. Voici quelques lignes de niveau de $f : (x, y) \in \mathbb{R}^2 \mapsto (x^2 + y + x)/(x^2 + y^2 + 1)$. Dans quelle zone trouver les couples (x_0, y_0) solutions ? Vérifier par le calcul.

Editeur

```
x = np.linspace(-1.5,1.5,100)
y = np.linspace(-1.5,1.5,100)
X, Y = np.meshgrid(x,y)
Z =(X**2+Y+X)/(X**2+Y**2+1)

graphe = plt.contour(X,Y,Z,15)
plt.axis('equal')
 # Pour un repère orthonormé
plt.clabel(graphe,inline=1,fontsize=8)
 # Tracé des lignes de niveau
t = np.linspace(-np.pi,np.pi,100)
plt.plot(np.cos(t),np.sin(t),'r')

plt.show()
```



Exercice 106. ✦✦✦

1. Nous savons que si une fonction $h : \mathbb{R} \to \mathbb{R}$ est strictement croissante alors elle est injective. La réciproque est fausse. Pouvez-vous donner le graphe d'un contre-exemple ?
2. L'objectif des questions suivantes est de prouver que la réciproque devient vraie si on suppose en plus que h est continue.

   On dit qu'une partie A de $\mathbb{R}^2$ est convexe si

$$\forall\, a, a' \in \mathrm{A}, \qquad \forall t \in [0;1] \quad ta + (1-t)a' \in \mathrm{A}.$$

   Soit $f : \mathbb{R}^2 \to \mathbb{R}$, continue. On définit $f(\mathrm{A})$, la partie de $\mathbb{R}$ par

$$f(\mathrm{A}) = \left\{ y \in \mathbb{R} \mid \exists x \in \mathrm{A},\ y = f(x) \right\}.$$

   (a) Montrer que pour toute partie convexe A, $f(\mathrm{A})$ est un intervalle de $\mathbb{R}$.
   Pour rappel, I *est un intervalle de* $\mathbb{R}$ *si pour tous* $\alpha, \beta \in \mathrm{I}$ *avec* $\alpha < \beta$, *on a* $[\alpha, \beta] \subset \mathrm{I}$.

   (b) Soit $h : \mathbb{R} \to \mathbb{R}$ injective et continue.

      i. Vérifier que $\mathrm{A} = \left\{ (x_1, x_2) \in \mathbb{R}^2 \mid x_1 < x_2 \right\}$ est une partie convexe de $\mathbb{R}^2$.

      ii. En considérant l'application f définie sur $\mathbb{R}^2$ par $f(x_1, x_2) = h(x_1) - h(x_2)$, conclure que h est strictement monotone sur $\mathbb{R}$.

CHAPITRE 6

Variables aléatoires à densité

If people do not believe that mathematics is simple, it is only because they do not realize how complicated life is.

JOHN VON NEUMANN
Mathématicien et physicien américano-hongrois (1903-1957)

1 Rappels : intégrales impropres

1.1 Convergence et convergence absolue

Définition 93 (intégrale généralisée en b)

Soit f une fonction continue sur un intervalle $[a;b[$ avec $-\infty < a < b \leqslant +\infty$.
L'intégrale généralisée (ou impropre) de f sur $[a;b[$ est notée $\int_a^b f(t)\,\mathrm{d}t$. Elle est dite **convergente** si $\int_a^x f(t)\,\mathrm{d}t$ admet une limite *finie* lorsque x tend vers b avec $x < b$. Dans ce cas, on pose

$$\int_a^b f(t)\,\mathrm{d}t = \lim_{\substack{x\to b \\ x<b}} \int_a^x f(t)\,\mathrm{d}t.$$

Si la limite n'existe pas ou qu'elle est infinie, l'intégrale est dite **divergente**.

Remarques.

- La définition s'étend aux fonctions continues sur $]a,b]$ avec $-\infty \leqslant a < b < +\infty$.
- Si f est une fonction continue sur un intervalle $]a,b[$ avec $-\infty \leqslant a < b \leqslant +\infty$.

L'intégrale généralisée de f sur $]a,b[$ est notée $\int_a^b f(t)\,\mathrm{d}t$. Elle est dite **convergente** si pour un réel $c \in]a,b[$, les intégrales généralisées $\int_a^c f(t)\,\mathrm{d}t$ et $\int_c^b f(t)\,\mathrm{d}t$ sont convergentes. Dans ce cas, on pose

$$\int_a^b f(t)\,\mathrm{d}t = \int_a^c f(t)\,\mathrm{d}t + \int_c^b f(t)\,\mathrm{d}t.$$

Théorème 94 (absolue convergence)

$\displaystyle\int_a^b f(t)\,\mathrm{d}t$ est dite **absolument convergente** si $\displaystyle\int_a^b |f(t)|\,\mathrm{d}t$ est convergente.

De plus, **si** $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ est absolument convergente, **alors** elle est convergente.

⚠ **Attention.** La réciproque est fausse. Il existe des intégrales généralisées convergentes sans être absolument convergente.

1.2 Critères de convergence

À l'aide du théorème de convergence monotone, on montre que, pour $f : [a;b[\to \mathbb{R}$ une fonction continue positive, l'intégrale $\int_a^b f(t)\,\mathrm{d}t$ converge si et seulement si l'application $x \in [a;b[\mapsto \int_a^x f(t)\,\mathrm{d}t$ est majorée.

Proposition 95 (critère de comparaison et de négligeabilité)

Soient f, g deux fonctions continues définies sur $[a,b[$.

- **Si**
  - f et g sont positives,
  - pour tout $x \in [a,b[$, $\quad f(x) \leqslant g(x)$,
  - $\displaystyle\int_a^b g(t)\,\mathrm{d}t$ est convergente, **alors** $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ est aussi convergente.
- **Si**
  - g est positive sur un voisinage de b ;
  - $f \underset{b^-}{=} o(g)$;
  - $\displaystyle\int_a^b g(t)\,\mathrm{d}t$ est convergente, **alors** $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ est aussi convergente.

Proposition 96 (critère d'équivalence)

Soient f, g deux applications continues définies sur $[a,b[$.

Si
- f et g sont de signe constant sur un voisinage de b,
- $f \underset{b^-}{\sim} g$,

alors les intégrales généralisées $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ et $\displaystyle\int_a^b g(t)\,\mathrm{d}t$ sont de même nature.

Remarque. En pratique, on compare souvent aux **intégrales de Riemann.**

- L'intégrale généralisée $\displaystyle\int_0^1 \frac{\mathrm{d}t}{t^\alpha}$ est convergente en 0 si et seulement si $\alpha < 1$;
- L'intégrale généralisée $\displaystyle\int_1^{+\infty} \frac{\mathrm{d}t}{t^\alpha}$ est convergente en $+\infty$ si et seulement si $\alpha > 1$.

1.3 Règles de calculs

Sous réserve de convergence, les propriétés de linéarité, de croissance de l'intégrale, l'inégalité triangulaire et la relation de Chasles sont encore valables. Par contre, pour effectuer une intégration par parties, on se ramène à un segment.

Théorème 97 (changement de variable)

Soit f continue sur un intervalle $]a,b[$.

Si $\quad \varphi :]\alpha,\beta[\to]a,b[$ est une bijection croissante de classe $\mathscr{C}^1$,

alors $\quad$ les intégrales généralisées $\int_a^b f(u)\,\mathrm{d}u$ et $\int_\alpha^\beta \varphi'(t) f\big(\varphi(t)\big)\,\mathrm{d}t$ sont de même nature.

Dans le cas de convergence,
$$\int_a^b f(u)\,\mathrm{d}u = \int_\alpha^\beta \varphi'(t) f\big(\varphi(t)\big)\,\mathrm{d}t.$$

Remarque. L'énoncé s'étend au cas décroissant.

2 Densité d'une variable aléatoire

Définition 98 (variable aléatoire à densité)

Soit X une variable aléatoire sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$. Notons F_X la fonction de répartition de X. On dit que X est une **variable aléatoire à densité** si :

- F_X est continue sur $\mathbb{R}$;
- F_X est de classe $\mathscr{C}^1$ sur $\mathbb{R}$ sauf en un nombre fini de points.

Exemples.
- Si F_X vérifie :

$$\forall\, x \in]-\infty; 0[,\ F_X(x) = 0, \quad \forall\, x \in [0;1],\ F_X(x) = x, \quad \forall\, x \in]1; +\infty[,\ F_X(x) = 1,$$

alors X est une variable à densité.
On dira dans la suite que X suit une loi uniforme sur $[0;1]$.
- Les variables aléatoires discrètes ne sont pas des variables à densité.

Remarque. Si X est une variable aléatoire, on montre que pour $b \in \mathbb{R}$: $\boxed{\mathbf{P}\Big([X = b]\Big) = 0.}$
Plus généralement, pour tous $a, b \in \mathbb{R}$

$$\mathbf{P}\Big([X \in]a;b[]\Big) = \mathbf{P}\Big([X \in [a;b[]\Big) = \mathbf{P}\Big([X \in]a;b]]\Big) = \mathbf{P}\Big([X \in [a;b]]\Big).$$

Définition-proposition 99 (densité de probabilité)

- On dit qu'une application $f : \mathbb{R} \to \mathbb{R}$ est une **densité de probabilité** si elle vérifie :
  - f est positive sur $\mathbb{R}$.
  - f est continue sur $\mathbb{R}$ sauf en un nombre fini de points.
  - L'intégrale suivante est convergente et $\displaystyle\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = 1.$

Soit X une variable aléatoire à densité.

- Toute application $f_X : \mathbb{R} \to \mathbb{R}$ à valeurs positives qui coïncide avec $F_X{}'$ sur $\mathbb{R}$ privé d'un ensemble fini de points est une **densité de probabilité**. On dit que f_X est *une* densité de probabilité de la variable aléatoire X.

Méthode

Comment montrer qu'une variable aléatoire est à densité ?

Il suffit de vérifier les trois points de la définition.

Par exemple, justifions que la fonction $f : t \in \mathbb{R} \mapsto \dfrac{1}{\pi(1+t^2)}$ définit une densité de probabilité.
- f est clairement une fonction positive.
- f est continue sur $\mathbb{R}$ en tant que fraction rationnelle dont le dénominateur ne s'annule pas.
- Pour $A \in \mathbb{R}^+$, on a

$$\int_0^A f(t)\,\mathrm{d}t = \int_0^A \frac{\mathrm{d}t}{\pi(1+t^2)} = \left[\frac{\arctan(t)}{\pi}\right]_0^A = \frac{\arctan(A) - \arctan(0)}{\pi} \underset{A\to+\infty}{\longrightarrow} \frac{\pi/2}{\pi} = \frac{1}{2}.$$

De même, pour $B \in \mathbb{R}^-$, par parité de f, $\displaystyle\int_B^0 f(t)\,\mathrm{d}t \underset{B\to-\infty}{\longrightarrow} \frac{1}{2}$. On a bien la convergence et l'égalité $\displaystyle\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = 1.$
En conclusion, f est une densité de probabilité.

Proposition 100 (densité et fonction de répartition)

Soient $X : \Omega \to \mathbb{R}$ une variable aléatoire à densité sur $(\Omega, \mathscr{A}, \mathbf{P})$ et F_X, f_X respectivement la fonction de répartition et une densité de X.

- Pour tout $x \in \mathbb{R}$, l'intégrale $\displaystyle\int_{-\infty}^{x} f_X(t)\,\mathrm{d}t$ est convergente, et $\quad F_X(x) = \displaystyle\int_{-\infty}^{x} f_X(t)\,\mathrm{d}t$.
- Soient $a, b \in \mathbb{R} \cup \{-\infty; +\infty\}$ tels que $a < b$, $\quad \mathbf{P}\Big([X \in [a;b]]\Big) = \displaystyle\int_a^b f_X(t)\,\mathrm{d}t$.

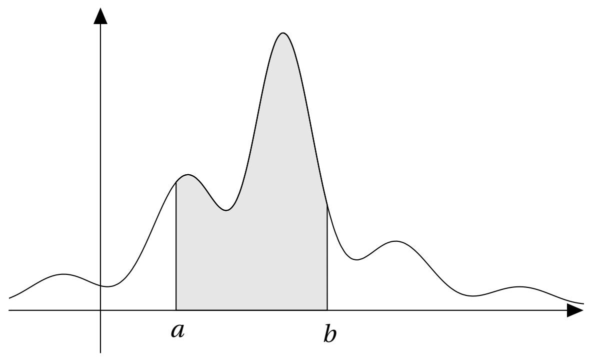
Remarque. En particulier, on a $\mathbf{P}\big([X \geqslant x]\big) = \mathbf{P}\big([X > x]\big) = 1 - F(x) = \displaystyle\int_x^{+\infty} f(t)\,\mathrm{d}t$.

⚠ **Attention.** Il ne faut pas confondre fonction de répartition et densité.
Rappelons qu'une fonction de répartition est une fonction croissante de limite 0 et 1 en $-\infty$ et $+\infty$. Le lien est, pour tout $t \in \mathbb{R}$, sauf un nombre fini,

$$\boxed{f_X(t) = F_X{}'(t).}$$

La densité d'une variable X n'est pas unique alors que la fonction de répartion l'est.

Interprétation graphique



Traçons la courbe représentative d'une densité.
L'aire comprise entre la courbe, l'axe des abscisses et les droites d'équation $x = a$, $x = b$ correspond exactement à la probabilité que la variable aléatoire prenne les valeurs comprises entre a et b.

$$\boxed{\text{Aire} = \int_a^b f(t)\,\mathrm{d}t = \mathbf{P}\Big(\big[a \leqslant X \leqslant b\big]\Big).}$$

Lorsque $a \to -\infty$ et $b \to +\infty$, on retrouve bien

$$\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = 1 = \mathbf{P}(\Omega).$$

De nouveau, on constate que pour une variable aléatoire continue, pour tout $a \in \mathbb{R}$, $\mathbf{P}\Big([X = a]\Big) = 0$.

Proposition 101 (la densité caractérise la loi)

Soient X et Y deux variables aléatoires à densité sur $(\Omega, \mathscr{A}, \mathbf{P})$ et $f : \mathbb{R} \to \mathbb{R}$ une application.
Si f est à la fois une densité de X et Y, alors X et Y ont même loi.

Remarque. En pratique, donner la loi d'une variable aléatoire à densité est équivalent à donner

$$\mathbf{P}\Big(\big[X \in [a,b]\big]\Big) \text{ pour tous } a,\ b \in \mathbb{R},$$

C'est aussi équivalent à donner la fonction de répartition ou une densité de probabilité.

Méthode

Comment montrer qu'une variable aléatoire est à densité et obtenir une densité ?

Soit X, une première variable aléatoire à densité et $\varphi : \mathbb{R} \to \mathbb{R}$, une application. On pose $Y = \varphi(X)$. Pour justifier que Y est une nouvelle variable aléatoire, on peut procéder comme suit :

- → On précise $Y(\Omega)$, l'ensemble des valeurs prises par Y.
- → On exprime la fonction de répartition de Y en fonction de celle de X. En particulier, on cherche pour tout $t \in Y(\Omega)$, la probabilité $\mathbf{P}\big([Y \leqslant t]\big)$.
- → Ensuite, on s'assure que F_Y est bien continue sur $\mathbb{R}$ et de classe $\mathscr{C}^1$ sauf en un nombre fini de points (noté dans la suite D).
- → On calcule la dérivée $F_Y{}'$ sur $\mathbb{R} \setminus D$ afin d'avoir une densité de Y.

Exercice 107

✧ **Un exemple avec la loi de Laplace**
Soit $\lambda \in \mathbb{R}$. On suppose que $X : \Omega \to \mathbb{R}$ est une variable à densité et que l'application $f : \mathbb{R} \to \mathbb{R}$ définie pour tout réel t par $f(t) = \lambda \exp(-|t|)$ est une densité de X.

1. Quelle est la valeur de λ ? Préciser le graphe de f.
2. Déterminer la fonction de répartition de X. *Indication. On distinguera $x \leqslant 0$ et $x \geqslant 0$.*
3. On pose $Y = 3X - 2$.
   (a) Établir une relation entre F_X et F_Y, les fonctions de répartition de X et Y.
   (b) Vérifier que Y est une variable aléatoire à densité et préciser une densité.

Exemple. Calcul de la loi de $Y = aX + b$ en fonction de la loi de X
Soient X une variable aléatoire à densité et $a, b \in \mathbb{R}$ avec $a \neq 0$.
On montre que $Y = aX + b$ est une variable aléatoire à densité avec une densité :

$$\forall\, t \in \mathbb{R}, \qquad f_Y(t) = \frac{1}{|a|} f_X\left(\frac{t-b}{a}\right).$$

Exercice 108

✦✦ Soit X une variable aléatoire à densité suivant une loi uniforme sur $[-1;1])$. C'est-à-dire, une densité de X est donnée par $f = 1/2 \cdot \mathbf{1}_{[-1;1]}$.

1. Justifier que $U = |X|$ est une variable à densité et donner une densité.
2. Faire de même avec $V = X^2$.
3. *Cas général*
   Soit Z une variable aléatoire réelle à densité f. Justifier que $Y = Z^2$ est une variable aléatoire réelle admettant une densité nulle sur $\mathbb{R}^-$ et donnée sur $\mathbb{R}_*^+$ par
   $$g(t) = \frac{1}{2\sqrt{t}}\big(f(\sqrt{t}) + f(-\sqrt{t})\big).$$

Proposition 102 (conditions pour une v.a.d)

Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction.

Si
- La fonction f est positive.
- La fonction f est continue sur $\mathbb{R}$, éventuellement privé d'un ensemble fini de points.
- $\displaystyle\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = 1.$

Alors il existe une variable aléatoire X dont f est une densité de probabilité.

3 Espérance

3.1 Définition

Définition 103 (espérance)

Soit X une variable aléatoire à densité f_X sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$.
On définit **l'espérance de** X, *sous réserve de convergence absolue*, comme le réel

$$\mathbf{E}(X) = \int_{-\infty}^{+\infty} t\, f_X(t)\,\mathrm{d}t.$$

⚠ **Attention.** Comme pour les séries, il ne faut pas oublier la convergence absolue.

Exercice 109

✧ **Loi de Pareto**
Soient $\alpha \in \mathbb{R}_*^+$ et X une variable aléatoire de densité f définie par :

$$f(t) = \frac{1}{\alpha} t^{-\frac{1+\alpha}{\alpha}} \quad \text{si} \quad t \in [1; +\infty[\quad \text{et } f(t) = 0 \text{ sinon.}$$

1. Vérifier que f est bien une densité de probabilité.
2. Calculer $\mathbf{P}\big([0 \leqslant X \leqslant 1+\alpha]\big)$ en fonction de α.
3. Pour quelles valeurs de α, X admet-elle une espérance ? La calculer quand elle existe.

3.2 Règles de calculs sur l'espérance

Les résultats suivants sont identiques au cas des v.a discrètes dénombrables.

Proposition 104 (linéarité de l'espérance, existence par domination)

Soient X et Y des variables aléatoires à densité admettant une espérance.

- Alors pour tout $\lambda, \mu \in \mathbb{R}$, la variable aléatoire $\lambda X + \mu Y$ admet une espérance et on a :

$$\mathbf{E}(\lambda X + \mu Y) = \lambda \mathbf{E}(X) + \mu \mathbf{E}(Y).$$

- **Si**
  - → $|X| \leqslant Y$.
  - → Y admet une espérance.

  Alors, X admet une espérance et $|\mathbf{E}(X)| \leqslant \mathbf{E}(Y)$.

Exemple. Si X est une variable aléatoire bornée, alors X admet une espérance.

Proposition 105 (positivité et croissance de l'espérance)

Soient X, Y des variables aléatoires à densité admettant une espérance.

- Positivité de l'espérance : si $X \geqslant 0$ presque sûrement, alors $\mathbf{E}(X) \geqslant 0$.
- Croissance de l'espérance : si $X \leqslant Y$ presque sûrement, alors $\mathbf{E}(X) \leqslant \mathbf{E}(Y)$.

Vocabulaire. Soit $X : \Omega \to \mathbb{R}$ une variable aléatoire à densité sur $(\Omega, \mathscr{A}, \mathbf{P})$. On dit que X est **centrée** si X admet une espérance et $\mathbf{E}(X) = 0$.
D'après la proposition précédente, la variable aléatoire $X - \mathbf{E}(X)$ est toujours une variable aléatoire centrée.

3.3 Formule de transfert

Théorème 106 (de transfert)

Soient X une variable aléatoire admettant une densité f_X nulle en dehors de $]a,b[$ avec $(a, b \in \mathbb{R} \cup \{\pm\infty\})$ et $\varphi :]a,b[\to \mathbb{R}$ une fonction continue sauf éventuellement en un nombre fini de points. On a l'équivalence entre

i) La variable $\varphi(X)$ admet une espérance.

ii) L'intégrale $\displaystyle\int_a^b \varphi(t) f_X(t)\,\mathrm{d}t$ converge absolument.

Et en cas de convergence absolue, $$\mathbf{E}\big(\varphi(X)\big) = \int_a^b \varphi(t) f_X(t)\,\mathrm{d}t.$$

Exercice 110

✧ **Applications**

1. Soit X, un v.a.d admettant une espérance et a, $b \in \mathbb{R}$. Montrer que $a\mathrm{X}+b$ admet une espérance et
$$\mathbf{E}(a\mathrm{X}+b) = a\mathbf{E}(\mathrm{X})+b.$$
2. Soit X, suivant une loi uniforme sur $]0;1]$. Justifier l'existence de l'espérance $\mathbf{E}\big(\ln(\mathrm{X})\big)$ et la calculer.

4 Moments et variance

4.1 Moments

Pour tout entier naturel s, le moment d'ordre s d'une variable aléatoire X est le nombre $m_s(\mathrm{X})$ défini par :

$$m_s(\mathrm{X}) = \mathbf{E}(\mathrm{X}^s).$$

En particulier pour $s=1$, on retrouve l'espérance $m_1(\mathrm{X}) = \mathbf{E}(\mathrm{X})$.
Dans le cas d'une variable aléatoire à densité, et sous réserve d'absolue convergence, la formule de transfert donne :

$$m_s(\mathrm{X}) = \int_{-\infty}^{+\infty} x^s f(x)\,\mathrm{d}x.$$

Exercice 111

✦ On définit la fonction f sur $\mathbb{R}$ par

$$f(x) = \begin{cases} -\ln(x) & \text{si } 0 < x \leqslant 1 \\ 0 & \text{sinon.} \end{cases}$$

1. Vérifier que f est une densité de probabilité.
2. Soit X une variable aléatoire de densité f.
Montrer que X admet des moments à tout ordre et les calculer.

Remarque. On montre que si X a un moment d'ordre r alors, X admet un moment d'ordre s pour tout entier $s \leqslant r$.

Exercice 112

✦

1. Prouver la remarque précédente.
2. Soit $r \in \mathbb{N}^*$, construire une variable aléatoire à densité admettant un moment d'ordre r mais pas de moment d'ordre $r+1$.

4.2 Variances

De nouveau, les résultats qui suivent sont calqués sur le cas des v.a discrètes.

Définition 107 (variance et écart-type)

Soit X une variable aléatoire admettant un moment d'ordre 2.

- On appelle **variance** d'une variable aléatoire X, la quantité $\mathbf{V}(\mathrm{X}) = \mathbf{E}\Big(\big(\mathrm{X}-\mathbf{E}(\mathrm{X})\big)^2\Big)$.
- On appelle **écart-type** de X, la quantité $\sigma(\mathrm{X}) = \sqrt{\mathbf{V}(\mathrm{X})}$.

Remarque. La quantité $\sigma(\mathrm{X})$ est bien définie car $\big(\mathrm{X}-\mathbf{E}(\mathrm{X})\big)^2 \geqslant 0$. Donc, par croissance de l'espérance, $\mathbf{V}(\mathrm{X}) \geqslant 0$.

L'écart-type permet de quantifier les écarts par rapport à la moyenne. Un écart-type fort traduit un « grand éloignement » des valeurs de X par rapport à sa moyenne.
En reprenant les preuves dans le cas discret, on prouve les deux énoncés suivants.

Proposition 108 (propriétés de la variance)

Soit X une variable aléatoire admettant un moment d'ordre 2.

- $\mathbf{V}(\mathrm{X}) = 0$ si et seulement si X est une application presque sûrement constante ;
- Pour tous réels a, b,
$$\mathbf{V}(a\mathrm{X}+b) = a^2\mathbf{V}(\mathrm{X}) \qquad ; \qquad \sigma(a\mathrm{X}+b) = |a|\sigma(\mathrm{X}).$$

Vocabulaire. Une variable aléatoire X est dite **centrée** si $\mathbf{E}(\mathrm{X}) = 0$. Elle est dite **réduite** si $\sigma(\mathrm{X}) = 1$. Soit X une variable aléatoire. La variable aléatoire $\mathrm{Y} = \frac{\mathrm{X}-\mathbf{E}(\mathrm{X})}{\sigma(\mathrm{X})}$ est centrée, réduite.

En pratique, on calcule la variance avec :

Théorème 109 (formule de Koenig-Huygens)

Soit X une variable aléatoire admettant un moment d'ordre 2.
$$\mathbf{V}(\mathrm{X}) = \mathbf{E}\left(\mathrm{X}^2\right) - \mathbf{E}(\mathrm{X})^2.$$

5 Simulation avec Python

5.1 Rappels : les histogrammes

Un histogramme donne une idée graphique de la répartition des valeurs d'un **échantillon** $\mathrm{E} = \{x_1, \dots, x_n\}$. Pour cela, on découpe l'ensemble des valeurs possibles en un certain nombre d'intervalles. Notons $p \in \mathbb{N}^*$ le nombre d'intervalles choisis. Ces intervalles doivent être deux à deux disjoints, et leur réunion est égale à (ou contient) l'ensemble des valeurs possibles. Plus précisément, en notant $a_1, a_2, \dots, a_{p+1} \in \mathbb{R}$, les bornes de tous ces intervalles avec $a_1 < a_2 < \dots < a_{p+1}$, on a
$$\mathrm{E} \subset \bigcup_{i=1}^{p} [a_i ; a_{i+1}[.$$
L'intervalle $[a_i ; a_{i+1}[$ est une **classe**. On peut ainsi définir l'**effectif** d'une classe comme le nombre d'éléments de E appartenant à la classe puis sa fréquence f_i comme le rapport de l'effectif sur le nombre total d'éléments de E.

Ensuite, le graphique est obtenu en traçant, pour chaque $i \in [\![1, p]\!]$, le rectangle de base $[a_i, a_{i+1}]$ sur l'axe des abscisses, et

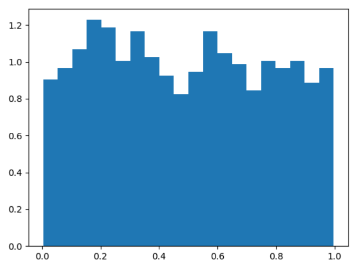
→ d'aire égale à m_i (on parle d'histogramme en effectif) ;

d'aire égale à f_i (on parle d'histogramme en effectif). Dans ce cas, la somme des aires des rectangles est égale à 1.

Exemple. Simulation d'une loi uniforme

La fonction Python ci-dessous prend en argument un entier naturel m non nul et tire m réel uniformément dans $[0; 1]$, puis affiche l'histogramme.

Editeur
```
L= np.random.rand(1000)
classes =np.linspace(0,1,10)

plt.hist(L, bins=20, density=True)  # Création de
    l'histogramme
plt.show()
```



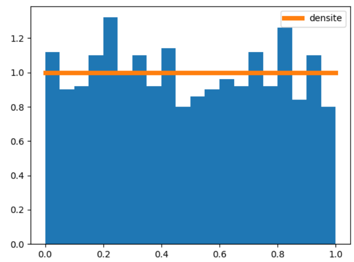
Expliquons les commandes :

- `density=True` : Trace l'histogramme en fréquence.
- `bins=...` : Si on met une valeur entière en argument, on définit le nombre de classes. Sinon, c'est une liste (ou un tableau) qui définit les différents intervalles des classes.
- `rwidth=...` : Largeur relative de la barre par rapport à l'intervalle.

Dans ce dernier exemple, il est utile de rajouter la courbe représentative d'une densité. Plus l'échantillon est important, plus on a de chance que l'histogramme « se rapproche » de la courbe.

Editeur

```
L= np.random.rand(1000)
classes =np.linspace(0,1,10)
plt.hist(L, bins=20, density=True)
 # Création de l'histogramme
x=[0,1]; y=[1,1]
plt.plot(x,y,linewidth=5,label="densite")
plt.legend()
plt.show()
```



5.2 Méthode d'inversion, illustration avec la loi de Cauchy

L'idée repose sur l'énoncé suivant :

Si
- U suit la loi uniforme sur $]0;1[$.
- X est une variable aléatoire à densité dont la fonction de répartition F est une bijection de $]a;b[$ avec $-\infty \leqslant a < b \leqslant +\infty$ sur $]0;1[$.

Alors La variable $Y = F^{-1}(U)$ suit la même loi que X.

Exercice 113

✦ Prouver l'énoncé précédent.

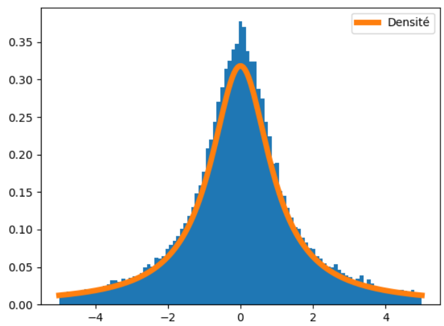
Exemple. On dit qu'une variable aléatoire X à densité suit une loi de Cauchy si une densité est donnée par $f : t \in \mathbb{R} \mapsto \frac{1}{\pi(1+t^2)}$. On a vu que la fonction de répartition est définie sur $\mathbb{R}$ par $F(x) = \frac{\arctan(x)}{\pi} + \frac{1}{2}$. La fonction F est strictement croissante, continue de $\mathbb{R}$ dans $]0;1[$. Le théorème de la bijection s'applique

$$\forall\, t \in]0;1[, \qquad F^{-1}(t) = \tan\big(\pi(t-1/2)\big).$$

Ci-dessous, un histogramme d'un échantillon obtenu par simulation d'une loi de Cauchy par la méthode d'inversion.

Editeur

```
plt.clf()
U= np.random.rand(50000)
L=np.tan(np.pi*(U-1/2))
inter=np.linspace(-5,5,100)
plt.hist(L, bins=inter, density=True)  # Création
     de l'histogramme
x=np.linspace(-5,5,200)
y=1/(np.pi*(1+x**2))
plt.plot(x,y,linewidth=5,label="Densité")
plt.legend()
plt.show()
```

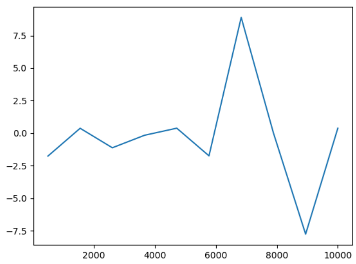


Exercice 114

✦✦ Expliquer l'intérêt du code suivant ? Qu'illustre-t on sur la loi de Cauchy ?

Editeur

```
Liste=100*np.linspace(5,100,10)
E=[]
for m in Liste :

    U= np.random.rand(int(m))
    L=np.tan(np.pi*(U-1/2))
    e=sum(L)/m
    E.append(e)
plt.plot(Liste,E)
plt.show()
```



5.3 Méthode par rejet

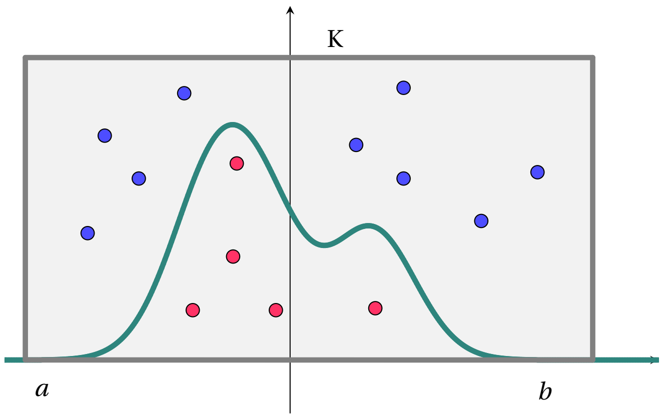
Soit X, une variable à densité et f, une densité. On suppose que f est bornée et à support bornée. C'est-à-dire, qu'il existe $\mathrm{K} \in \mathbb{R}^+$, a, $b \in \mathbb{R}$ avec $a < b$ tels que

$$\forall x \in \mathbb{R}, \qquad 0 \leqslant f(x) \leqslant \mathrm{K} \quad \text{et} \quad f(x) = 0 \text{ si } x \notin [a;b].$$

Rappelons que l'aire comprise entre la courbe, l'axe des abscisses et les droites d'équation $x = a$, $x = b$ correspond exactement à la probabilité que la variable aléatoire prenne les valeurs comprises entre a et b.

$$\text{Aire} = \int_a^b f(t)\,\mathrm{d}t = \mathbf{P}\Big(\big[a \leqslant \mathrm{X} \leqslant b\big]\Big).$$

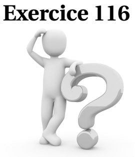
Ainsi, pour simuler la variable X, on peut de tirer un point au hasard sous la courbe représentative de f de manière uniforme. On prend alors pour réalisation de la loi f l'abscisse de ce point.
Précisons que pour tirer un point sous la courbe on simule des variables uniformes sur $[a;b] \times [0;\mathrm{K}]$ jusqu'à ce que l'un des points tirés soient situés sous la courbe de f.



Exercice 115

✦ Quelle est la loi du nombre d'essais avant de tirer un point qui est bien situé sous la courbe de f ?

Exercice 116



✦ **Simulation par la méthode de rejet**
Soit X une variable aléatoire dont une densité f est définie sur $\mathbb{R}$ par

$$f(x) = \begin{cases} 6x(1-x) & \text{si } x \in [0;1] \\ 0 & \text{sinon.} \end{cases}$$

1. Tracer le graphe de f.
2. Écrire un programme Python qui simule la variable X à l'aide de la commande **random**.

Exercices - TD

Révisions : intégrales

Exercice 117. ✦ Sommes de Riemann

Justifier que pour $\alpha \in \mathbb{R}_*^+$
$$\sum_{k=1}^{n} k^\alpha \underset{n\to+\infty}{\sim} \frac{n^{\alpha+1}}{\alpha+1}.$$

Exercice 118. ✦ Lemme de Riemann-Lebesgue

Soit $f : [0;\pi] \to \mathbb{R}$ de classe $\mathscr{C}^1$. Justifier que
$$\int_0^\pi f(t)\sin(nt)\,\mathrm{d}t \underset{n\to+\infty}{\longrightarrow} 0.$$

Exercice 119. ✦ Soit f une fonction continue bornée sur $[0, +\infty[$.

1. Démontrer que les intégrales I et J sont convergentes où
$$\mathrm{I} = \int_0^{+\infty} \frac{f(x)}{1+x^2}\mathrm{d}x \quad \text{et} \quad \mathrm{J} = \int_0^{+\infty} \frac{f(1/x)}{1+x^2}\mathrm{d}x.$$
2. Vérifier que $\mathrm{I} = \mathrm{J}$.
3. *Application.* Pour $n \in \mathbb{N}$, calculer $\displaystyle\int_0^{+\infty} \frac{\mathrm{d}x}{(1+x^2)(1+x^n)}$ et $\displaystyle\int_0^{+\infty} \frac{x^n}{(1+x^2)(1+x^n)}\mathrm{d}x$.

Exercice 120. ✦✦ Soit $f : \mathbb{R} \to \mathbb{R}$ continue telle que $\displaystyle\int_0^{+\infty} f(t)\,\mathrm{d}t$ converge.

1. Donner un exemple où $f(t) \underset{t\to+\infty}{\not\longrightarrow} 0$. *Un graphe suffit...*
2. Vérifier que $f(t) \underset{t\to+\infty}{\longrightarrow} 0$ est vrai si on suppose que f de classe $\mathscr{C}^1$ avec $\displaystyle\int_0^{+\infty} f'(t)\,\mathrm{d}t$ convergente.

Variable aléatoire à densité

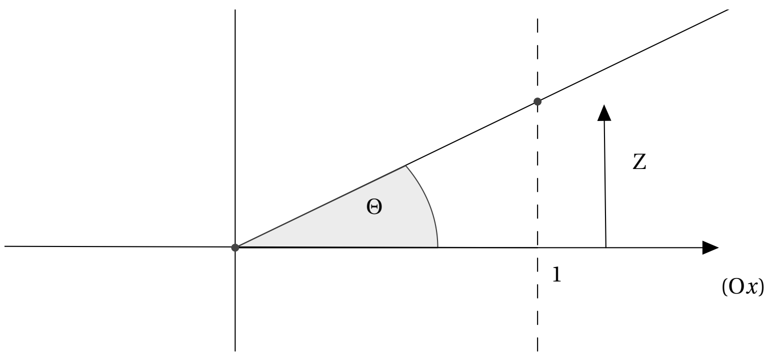
Exercice 121. ✧ Vrai ou faux?

Toute variance de variable aléatoire réelle à densité est strictement positive.

Exercice 122. ✦ Autour de la loi de Cauchy

Une variable aléatoire à densité suit une loi de Cauchy si une densité est $f : t \in \mathbb{R} \mapsto \dfrac{1}{\pi(1+t^2)}$.

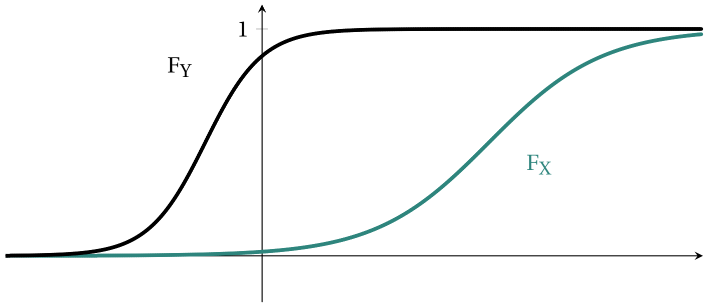
1. *Préliminaires.*
   (a) Simplifier $\arctan(t) + \arctan(1/t)$ lorsque $t \in \mathbb{R}^*$.
   (b) Donner la fonction de répartition d'une loi de Cauchy.
2. Soit X une variable aléatoire à densité suivant une loi de Cauchy. On pose $\mathrm{Y} = 1/\mathrm{X}$.
   (a) Calculer $\mathbf{P}([\mathrm{Y} < 0])$, puis, pour tout $x \in \mathbb{R}_*^+$, $\mathbf{P}([0 < \mathrm{Y} \leqslant x])$.
   (b) Montrer que Y suit encore une loi de Cauchy.
3. Un tireur envoie sa balle sur un demi-plan droit. On suppose que l'angle Θ entre le trajectoire et l'axe (Ox) est une variable aléatoire suivant une loi uniforme sur $]-\frac{\pi}{2}; \frac{\pi}{2}[$.
   Justifier que la position Z de la balle sur la cible (d'axe $x = 1$) suit une loi de Cauchy.



Exercice 123. ✦ Soit X une variable aléatoire à densité avec une densité de probabilité f paire.

1. Préciser $\mathbf{P}\big([X \geqslant 0]\big)$.
2. On note F la fonction de répartition de X. Justifier que $x \in \mathbb{R} \mapsto F(x) - \frac{1}{2}$ est impaire.
3. On suppose que $\displaystyle\int_0^{+\infty} t f(t)\,\mathrm{d}t$ converge. Démontrer que X admet une espérance et la calculer.

Exercice 124. ✦✦ *D'après Oral HEC-E*



La figure ci-dessous est la représentation graphique de la fonction de répartition de deux variables aléatoires X et Y admettant des espérances.
Quelle est la variable dont l'espérance est la plus grande ? Justifier.

Exercice 125. ✦✦ **Définition de la médiane.**
Soit X une variable aléatoire à densité sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$.

1. Justifier qu'il existe un réel m tel que $\quad \mathbf{P}\big([X \leqslant m]\big) = \dfrac{1}{2}$.
   Un tel réel est une médiane de X.
2. Démontrer que l'ensemble des médianes d'une variable à densité X est un segment.
   Une partie I *de* $\mathbb{R}$ *est un intervalle si pour tous* $x, y \in I$, *le segment* $[x; y] \subset I$.

Exercice 126. ✦ **Nouvelle expression de l'espérance**
Soit X une v.a à densité. On note f une densité de X, et F sa fonction de répartition. On suppose que f est continue sur $\mathbb{R}$ et que $f = 0$ sur $\mathbb{R}^-$.

1. Montrer que pour tout $x \in \mathbb{R}^+$,
$$\int_0^x \big(1 - F(t)\big)\,\mathrm{d}t = x\big(1 - F(x)\big) + \int_0^x t f(t)\,\mathrm{d}t. \qquad (\bullet)$$
2. On suppose dans cette question que X possède une espérance.
   (a) Montrer que pour tout $x \in \mathbb{R}_+^*$, $0 \leqslant x \displaystyle\int_x^{+\infty} f(t)\,\mathrm{d}t \leqslant \int_x^{+\infty} t f(t)\,\mathrm{d}t$.
   (b) En déduire que $x\big(1 - F(x)\big) \underset{x \to +\infty}{\longrightarrow} 0$.
   (c) À l'aide de la relation $(\bullet)$, montrer que l'intégrale $\displaystyle\int_0^{+\infty} \big(1 - F(t)\big)\,\mathrm{d}t$ converge, et que $\displaystyle\int_0^{+\infty} \big(1 - F(t)\big)\,\mathrm{d}t = \mathbf{E}(X)$.
3. On étudie la réciproque.
   On suppose dans cette question que l'intégrale $\displaystyle\int_0^{+\infty} \big(1 - F(t)\big)\,\mathrm{d}t$ converge.
   (a) Montrer que l'application $\varphi : x \to \displaystyle\int_0^x t f(t)\,\mathrm{d}t$ est croissante sur $\mathbb{R}_+$.
   (b) Montrer en utilisant $(\bullet)$ que φ est majorée.
   (c) En déduire que $\mathbf{E}(X)$ existe et que
$$\mathbf{E}(X) = \int_0^{+\infty} \big(1 - F(t)\big)\,\mathrm{d}t.$$

Exercice 127. ✦✦✦ *Oraux, sans préparation, HEC 2008*
Soit F la fonction de répartition d'une variable aléatoire à densité. Montrer que la fonction g définie par :
$$\forall x \in \mathbb{R}, \qquad g(x) = F(x+1) - F(x)$$
est une densité de probabilité.

Exercice 128. ✦ Soit X une variable aléatoire admettant une densité f continue sur $\mathbb{R}$ et un moment d'ordre 2.
Caractériser la valeur du réel a qui minimise l'expression
$$\varphi(a) = \mathbf{E}\Big((X - a)^2\Big) = \int_{-\infty}^{+\infty} (t - a)^2 f(t)\,\mathrm{d}t.$$

CHAPITRE 7

Lois à densité usuelles

J'étais alors en proie à la mathématique.
Temps sombre ! enfant ému du frisson poétique,
Pauvre oiseau qui heurtais du crâne mes barreaux,
On me livrait tout vif aux chiffres, noirs bourreaux ; [...]
On me tordait depuis les ailes jusqu'au bec,
Sur l'affreux chevalet des X et des Y ; [...]
VICTOR HUGO, *Les Contemplations*, 1856

1 Lois uniformes continues

Définition, espérance et variance

Définition 110 (lois uniformes continues)

Soient $X : \Omega \to \mathbb{R}$ une variable aléatoire à densité sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ et $a, b \in \mathbb{R}$ avec $a < b$. On dit que X suit la **loi uniforme continue** sur l'intervalle $[a;b]$, noté $X \hookrightarrow \mathscr{U}([a;b])$, si X a pour densité

$$\forall\, t \in \mathbb{R}, \quad f_X(t) = \begin{cases} \dfrac{1}{b-a} & \text{si} \quad t \in [a;b] \\ 0 & \text{sinon.} \end{cases}$$

Remarque. Le facteur $\frac{1}{b-a}$ est imposé par le fait qu'une densité est toujours d'intégrale 1 sur $\mathbb{R}$.

Proposition 111 (fonction de répartition et espérance d'une loi uniforme)

Soit X une variable aléatoire suivant une loi uniforme $\mathscr{U}([a;b])$. La fonction de répartition de X est

$$\forall\, x \in \mathbb{R}, \qquad F_X(x) = \begin{cases} 0 & \text{si} \quad x < a \\ \dfrac{x-a}{b-a} & \text{si} \quad a \leqslant x \leqslant b \\ 1 & \text{si} \quad x > b. \end{cases}$$

De plus, X admet une espérance et une variance avec

$$\mathbf{E}(X) = \frac{a+b}{2} \quad \text{et} \quad V(X) = \frac{(b-a)^2}{12}.$$

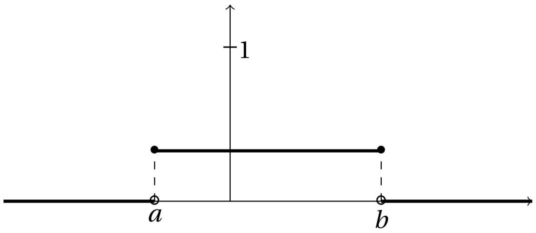
Remarque. Comme X est une variable aléatoire bornée, X admet des moments à tout ordre.

Remarque. Autrement dit, si $X \hookrightarrow \mathscr{U}([a;b])$, alors pour tous $c, d \in \mathbb{R}$ tels que $a \leqslant c \leqslant d \leqslant b$, on a

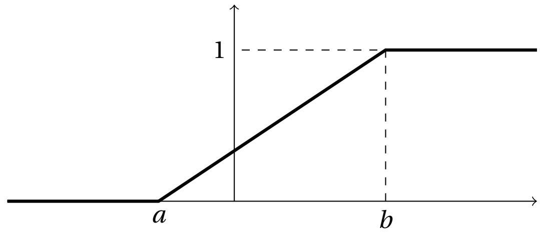
$$\mathbf{P}\Big([X \in [c;d]]\Big) = \frac{d-c}{b-a}.$$

En particulier, X est presque sûrement à valeurs dans $[a;b]$.

Les graphes



Une densité de $X \hookrightarrow \mathscr{U}([a;b])$



La fonction de répartition de $X \hookrightarrow \mathscr{U}([a;b])$

Exemple de modélisation

Souvent lorsque qu'on ne précise pas, un tirage *au hasard* correspond à une loi uniforme.
Par exemple, choisir un nombre réel *au hasard* dans $[0;1]$ signifie choisir un nombre suivant une loi uniforme $\mathscr{U}([0;1])$.

Exercice 129

✦ À partir de 8 heures, les bus passent toutes les quinze minutes à un arrêt donné. Ils passent donc à 8h00, 8h15, 8h30, etc... Un usager se présente à cet arrêt entre 8h00 et 8h30. On suppose que l'heure exacte de son arrivée est une variable aléatoire X de loi uniforme sur cette période.

1. Quelle est la probabilité que l'usager doive attendre moins de cinq minutes ? Plus de dix minutes ?
2. On note T le temps d'attente du bus (en minutes).
   (a) Déterminer F_T, la fonction de répartition de T.
   (b) En déduire la loi de T. En moyenne, combien de temps attend l'usager ?

Transformation affine et lois uniformes

X suit une loi uniforme sur $[0;1]$
si et seulement si
$Y = (b-a)X + a$ suit une loi uniforme sur $[a;b]$.

Exercice 130

✦ **Vrai ou faux ?** Si $U \hookrightarrow \mathscr{U}([-1;1])$, alors U et $-U$ ont même loi.

2 Lois exponentielles

Définition, espérance et variance

Définition 112 (lois exponentielles)

Soient $X : \Omega \to \mathbb{R}$ une variable aléatoire à densité sur $(\Omega, \mathscr{A}, \mathbf{P})$ et $\lambda \in \mathbb{R}_*^+$.
On dit que X suit la **loi exponentielle** de paramètre λ, notée $\mathscr{E}(\lambda)$, si X a pour densité

$$\forall t \in \mathbb{R}, \qquad f_X(t) = \begin{cases} \lambda e^{-\lambda t} & \text{si} \quad t \in \mathbb{R}^+ \\ 0 & \text{sinon.} \end{cases}$$

Proposition 113 (fonction de répartition et espérance d'une loi exponentielle)

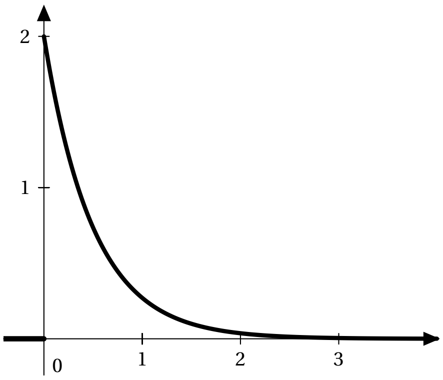
Soit X une variable aléatoire suivant une loi exponentielle $\mathscr{E}(\lambda)$. La fonction de répartition de X est

$$\forall\, x \in \mathbb{R}, \qquad F_X(x) \;=\; \begin{cases} 0 & \text{si} \quad x < 0 \\ 1 - e^{-\lambda x} & \text{si} \quad x \geqslant 0. \end{cases}$$

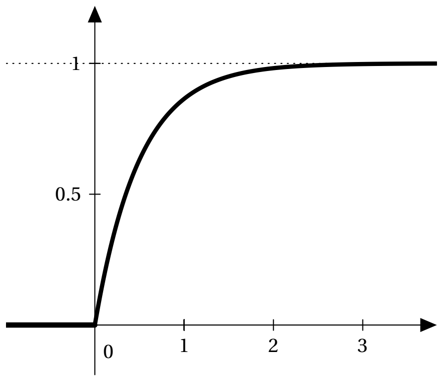
De plus, X admet une espérance et une variance avec

$$\mathbf{E}(X) = \frac{1}{\lambda} \quad \text{et} \quad V(X) = \frac{1}{\lambda^2}.$$

Les graphes



Une densité de $X \hookrightarrow \mathscr{E}(2)$



La fonction de répartition de $X \hookrightarrow \mathscr{E}(2)$

Exemples de modélisation

Proposition 114 (loi sans mémoire)

Soit $X \hookrightarrow \mathscr{E}(\lambda)$.
Alors X est une **variable aléatoire sans mémoire**. C'est-à-dire :

$$\forall\, s, t \in \mathbb{R}^+, \qquad \mathbf{P}\big([X > t]\big) = \mathbf{P}_{[X>s]}\big([X > s + t]\big).$$

Exercice 131

✦ Prouver cette proposition.
L'exercice 143 p. 82 étudie la réciproque.

Remarque. On parle aussi de propriété de *non vieillissement.*
Plus généralement, la loi exponentielle intervient dans les processus continus sans mémoire comme le temps d'émission d'un électron, le temps de vie d'un composant électronique, ...

Transformation linéaire et lois exponentielles

Soient $\lambda \in \mathbb{R}_*^+$, $X \hookrightarrow \mathscr{E}(1)$ et $Y = \frac{1}{\lambda}X$. On a vu la relation entre les fonctions de répartition :

$$\forall\, t \in \mathbb{R}, \qquad F_Y(t) = F_X(\lambda t) \quad \Rightarrow \quad f_Y(t) = \lambda\, f_X(\lambda t) = \begin{cases} \lambda e^{-\lambda t} & \text{si} \quad t \geqslant 0 \\ 0 & \text{sinon.} \end{cases}$$

Retenons :

X suit une loi exponentielle de paramètre 1
si et seulement si
$Y = \frac{1}{\lambda}X$ suit une loi exponentielle de paramètre λ.

Exercice 132

✦ **Moments de la loi exponentielle**

Pour tout $n \in \mathbb{N}$, on pose $\mathrm{I}_n = \displaystyle\int_0^{+\infty} t^n \mathrm{e}^{-t}\,\mathrm{d}t$.

1. (a) Justifier que I_n est convergente et donner une relation simple entre I_{n+1} et I_n.
   (b) En déduire I_n pour tout $n \in \mathbb{N}$.
2. Conclure en justifiant que la loi $\mathscr{E}(\lambda)$ admet un moment à tout ordre et retrouver les expressions de l'espérance et la variance.

Exercice 133

✦ **Simulation de la loi exponentielle par la méthode d'inversion**

1. Soient $\mathrm{U} \hookrightarrow \mathscr{U}([0;1])$ et $\mathrm{X} = -\ln(\mathrm{U})$. Donner la loi de X.
2. En déduire un programme qui prend en argument λ et simule une variable aléatoire de loi $\mathscr{E}(\lambda)$.
3. Tester la simulation en traçant un histogramme et la densité.

3 Lois γ

Avant de donner la définition, reprenons un exercice classique sur la définition et les propriétés de la fonction Γ définie sur $\mathbb{R}_*^+$ par

$$\Gamma(x) = \int_0^{+\infty} t^{x-1}\mathrm{e}^{-t}\,\mathrm{d}t.$$

Exercice 134

✦ **Propriétés de la loi Γ**

1. Justifier que Γ est une application bien posée sur $\mathbb{R}_*^+$.
2. (a) Préciser $\Gamma(1)$.
   (b) En utilisant l'intégrale de Gauss, démontrer que $\Gamma(1/2) = \sqrt{\pi}$.
3. (a) Justifier que pour tout $x \in \mathbb{R}_*^+$, $\Gamma(x+1) = x\Gamma(x)$.
   (b) En déduire que que pour tout $n \in \mathbb{N}^*$, $\Gamma(n) = (n-1)!$.

Définition, espérance et variance

Définition 115 (lois γ)

Soient $\mathrm{X} : \Omega \to \mathbb{R}$, une variable à densité sur $(\Omega, \mathscr{A}, \mathbf{P})$ et $\nu \in \mathbb{R}_*^+$.
On dit que X suit une **loi γ de paramètre** ν, noté $\mathrm{X} \hookrightarrow \gamma(\nu)$, si X a pour densité

$$f_\nu : t \mapsto \begin{cases} 0 & \text{si } t \leqslant 0 \\ \frac{1}{\Gamma(v)} t^{v-1}\mathrm{e}^{-t} & \text{si } t > 0 \end{cases}.$$

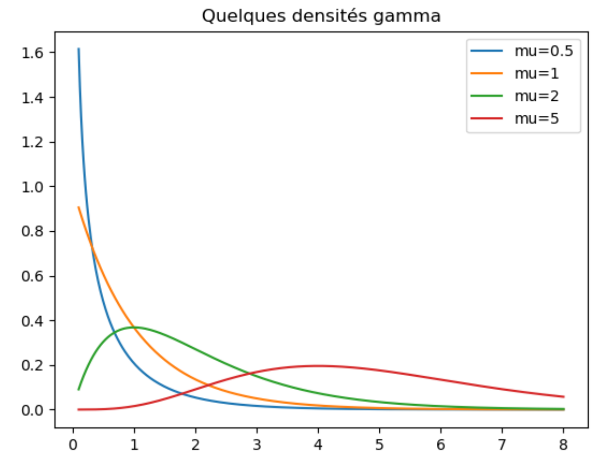
Remarques.
• Le facteur $1/\Gamma(v)$ permet justement de s'assurer que l'intégrale $\int_{-\infty}^{+\infty} f_\nu(t)\,\mathrm{d}t$ vaut bien 1 pour avoir une densité.
• Remarquons que la loi $\gamma(1)$ n'est autre que la loi exponentielle $\mathscr{E}(\lambda)$ qui s'utilise pour modéliser des durées de vie sans vieillissement. Nous verrons au chapitre suivant l'intérêt de ces lois lors de sommes de v.a.

Les graphes

Ci-dessous, le tracé avec Python de quelques densités pour différentes valeurs du paramètre ν. Noter l'utilisation de la bibliothèque **spicy** qui permet, entre autre, d'importer des fonctions de référence (dans notre cas, la fonction Γ).

Editeur

```
import scipy.special as sp
import numpy as np
import matplotlib.pyplot as plt
plt.clf()

X=np.linspace(0.1,8,400)
for mu in [0.5, 1, 2,5] :
    Gamma=sp.gamma(mu)
    Y=X**(mu-1)*np.exp(-X)/Gamma

    plt.plot(X,Y)
plt.legend(['mu=0.5', 'mu=1', 'mu=2','mu=5'])
plt.title('Quelques densités gamma')
plt.show()
```



Proposition 116 (espérance et variance d'une loi γ)

Si X est une variable aléatoire suivant une loi $\gamma(\nu)$ alors X admet une espérance et une variance avec

$$\mathbf{E}(X) = \nu \quad \text{et} \quad \mathbf{V}(X) = \nu.$$

4 Lois normales

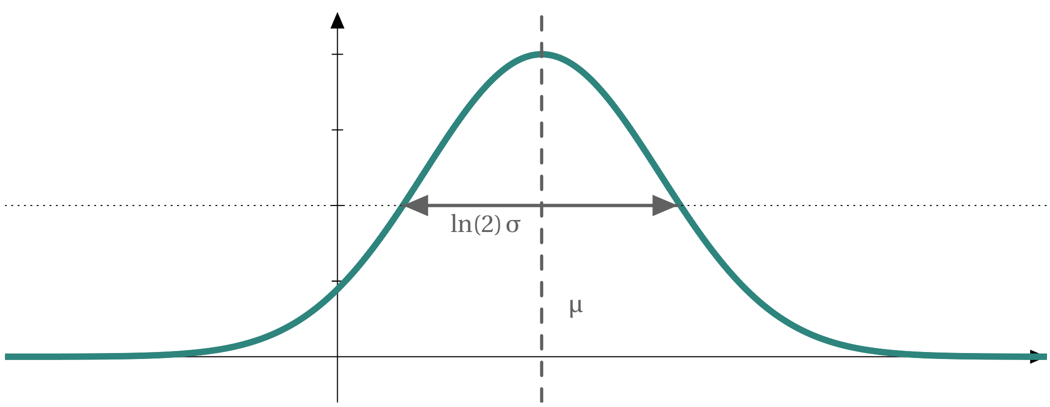
Définition, espérance et variance

Définition 117 (lois normales)

Soient $X : \Omega \to \mathbb{R}$ une variable aléatoire à densité sur $(\Omega, \mathscr{A}, \mathbf{P})$ et $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}_*^+$.
On dit que X suit la **loi normale** de paramètres μ, σ, noté $X \hookrightarrow \mathcal{N}(\mu; \sigma^2)$, si X a pour densité

$$\forall\, t \in \mathbb{R}, \qquad f_{\mu,\sigma}(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right).$$

Ci-dessous, l'allure en cloche de la courbe de la densité $f_{\mu,\sigma}$ de la loi normale. Plus σ est proche de 0^+, plus la courbe est étroite.



Transformation affine et lois normales

Proposition 118 (loi normale et transformation affine)

Soient $\sigma \in \mathbb{R}_*^+$, $\mu \in \mathbb{R}$.

Si X suit une loi normale $\mathcal{N}(0;1)$.

Alors la variable aléatoire $Y = \sigma X + \mu$ suit une loi normale $\mathcal{N}(\mu;\sigma^2)$.

Plus généralement, toute transformation affine d'une loi normale est encore une loi normale.

Remarque. Dans la suite, $\mathcal{N}(0;1)$ désigne la **loi normale centrée réduite**. La densité est

$$f : t \in \mathbb{R} \mapsto \frac{1}{\sqrt{2\pi}} e^{-t^2/2}.$$

Remarque. Il est d'usage de noter Φ la fonction de répartition de la loi normale centrée réduite. On démontre qu'il n'existe pas d'expression avec les fonctions usuelles de Φ.

Retenons les relations suivantes :

$$\boxed{\forall x \in \mathbb{R}, \quad \Phi(x) + \Phi(-x) = 1 \quad \text{et} \quad \Phi(0) = \frac{1}{2}.}$$

Espérance et variance

Proposition 119 (espérance et variance d'une loi normale)

Si X est une variable aléatoire suivant une loi normale $\mathcal{N}(\mu;\sigma^2)$, alors X admet une espérance et une variance avec

$$\mathbf{E}(X) = \mu \quad \text{et} \quad \mathbf{V}(X) = \sigma^2.$$

Remarque. Si $X \hookrightarrow \mathcal{N}(\mu;\sigma^2)$, alors on obtient une variable aléatoire centrée réduite via

$$\boxed{X^* = \frac{X-\mu}{\sigma} \hookrightarrow \mathcal{N}(0;1).}$$

Modélisation

Nous verrons l'importance de la loi normale grâce au théorème limite central au chapitre CONVERGENCES DES VARIABLES ALÉATOIRES.

Les lois normales sont utilisées dans de nombreux domaines des mathématiques. Les mathématiques financières en font une large utilisation (parfois abusive). Citons, par exemple, le modèle de Black-Scholes pour l'évolution d'une action en bourse.

Exercices - TD

Exercice 135. ✦ Soit X une variable aléatoire suivant la loi normale $\mathcal{N}(\mu;\sigma^2)$. Calculer l'espérance de X^2.

Exercice 136. ✦ **Un exemple détaillé**
Soit $X \hookrightarrow \mathcal{U}([-1;1])$.

1. On pose pour tout $t \in]-1;1[$, $g(t) = \dfrac{1}{2}\ln\left(\dfrac{1+t}{1-t}\right)$.
   (a) Justifier que g est une bijection de $]-1;1[$ dans $\mathbb{R}$.
   (b) Donner une expression simple de $g^{-1}(x)$ pour tout $x \in \mathbb{R}$. Justifier que g^{-1} est de classe $\mathscr{C}^1$ et donner la dérivée.
   (c) On définit la variable $T = g(X)$.
      i. Prouver que pour tout $x \in \mathbb{R}$, $F_T(x) = F_X\left(g^{-1}(x)\right)$.
      ii. En déduire que T est une variable aléatoire à densité. *On précisera une densité.*
2. (a) Justifier que $U = |X|$ est une variable à densité et donner une densité.
   (b) Faire de même avec $V = X^2$.

Exercice 137. ✦ Soit X est une variable aléatoire qui suit une loi normale centrée réduite.

1. Donner la loi de $Y = |X|$. Préciser son espérance et sa variance
2. Trouver la fonction de répartition de $Z = \frac{X+|X|}{2}$. La variable Z est-elle à densité ?

Exercice 138. ✦✦ **Moments de la loi normale**

1. Justifier l'existence pour tout $n \in \mathbb{N}$ de $I_n = \displaystyle\int_0^{+\infty} x^n e^{-x^2/2}\,dx$.
2. Montrer que pour tout $n \in \mathbb{N}$, $I_{n+2} = (n+1)I_n$.
3. Vérifier que $I_0 = \sqrt{\frac{\pi}{2}}$.
4. Démontrer que pour tout $p \in \mathbb{N}$:
$$I_{2p} = \frac{(2p)!}{2^p p!}\sqrt{\frac{\pi}{2}} \quad \text{et} \quad I_{2p+1} = 2^p p!.$$
5. Conclure en donnant les moments de la loi normale centrée réduite à tout ordre.

Exercice 139. ✦✦ **Tracé de la fonction de répartition de la loi normale centrée réduite**

1. ✧ *Premières propriétés*
   (a) Préciser les variations de Φ et donner les limites en $\pm\infty$ de Φ.
   (b) Vérifier pour tout réel x, $\quad \Phi(x) - \frac{1}{2} = -\left(\Phi(-x) - \frac{1}{2}\right)$ et $\Phi(0) = \dfrac{1}{2}$.
   (c) Justifier le développement limité en 0 : $\quad \Phi(x) = \dfrac{1}{2} + \dfrac{1}{\sqrt{2\pi}}x + o_0(x)$.
   (d) En utilisant les résultats précédents, tracer l'allure de la courbe représentative de Φ.
2. ✦✦✦ *Tracé avec Python* *D'après ESCP ORAUX 1999*
   (a) Soit $x \in \mathbb{R}^+$. Montrer que, pour tout entier naturel $n \geqslant 1$,
$$\frac{x}{n}\sum_{k=1}^{n} e^{-\frac{(kx)^2}{2n^2}} \leqslant \int_0^x e^{-\frac{t^2}{2}}\,dt \leqslant \frac{x}{n}\sum_{k=0}^{n-1} e^{-\frac{(kx)^2}{2n^2}}.$$
   En déduire que
$$\left|\int_0^x e^{-\frac{t^2}{2}}\,dt - \frac{x}{n}\sum_{k=0}^{n-1} e^{-\frac{(kx)^2}{2n^2}}\right| \leqslant \frac{x}{n}\left(1 - e^{-\frac{x^2}{2}}\right).$$
   (b) Quelle inégalité doit vérifier n pour trouver une valeur approchée de $\Phi(x)$ à 10^{-6} près ?
   (c) Proposer, en langage Python, une fonction `I(f, a, b, n)` qui prend en entrée une fonction f à valeurs réelles, deux réels a et b et un entier naturel n et qui renvoie une valeur approchée avec la méthode des rectangles de $\int_a^b f(t)\,dt$ calculée avec n rectangles.
   (d) En déduire une nouvelle fonction qui renvoie une approximation de $\displaystyle\int_0^x e^{-\frac{t^2}{2}}\,dt$, puis de $\Phi(x)$.

Exercice 140. ✦✦ On note Φ la fonction de répartition d'une variable aléatoire de loi $\mathcal{N}(0,1)$.

1. Montrer grâce à une intégration par partie, puis un encadrement d'intégrale, que

$$\forall x \in \mathbb{R}_+^*, \qquad \frac{1}{\sqrt{2\pi}}\left(\frac{1}{x}-\frac{1}{x^3}\right)e^{-x^2/2} \leqslant 1-\Phi(x) \leqslant \frac{1}{\sqrt{2\pi}}\frac{1}{x}e^{-x^2/2}.$$

2. En déduire un équivalent de $1-\Phi(x)$ quand $x \to +\infty$.

Exercice 141. ✦✦ **Simulation d'une géométrique à partir de la loi exponentielle**
Soit $\lambda \in \mathbb{R}_+^*$. Soit X une variable aléatoire sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ de loi $\mathscr{E}(\lambda)$. On pose $Y = \lfloor X \rfloor + 1$.

1. Montrer $Y \hookrightarrow \mathscr{G}\left(1-e^{-\lambda}\right)$.
2. En déduire un programme qui simule une loi géométrique de paramètre p à partir de la loi exponentielle.

Exercice 142. ✦✦ Soit X, une variable aléatoire dont f est une densité paire et continue sur $\mathbb{R}$. On suppose que $X^2 \hookrightarrow \mathscr{E}(\lambda)$. Déterminer la densité f.

Exercice 143. ✦✦✦ **Loi sans mémoire**

1. *Préliminaires*
   Soit f une fonction continue sur $\mathbb{R}$, non nulle et vérifiant

   $$\forall x, y \in \mathbb{R}, \qquad f(x+y) = f(x)\cdot f(y).$$

   (a) Montrer que $\forall x \in \mathbb{R}$, $f(x) > 0$.

   (b) i. Déterminer une expression de $f(n)$, où $n \in \mathbb{N}$. En déduire $f(n)$, où $n \in \mathbb{Z}$.

   ii. Déterminer ensuite une expression de $f(p/q)$, où $p \in \mathbb{Z}$, $q \in \mathbb{N}^*$.

   (c) Soit $x \in \mathbb{R}$. Étudier la limite de $\lfloor nx \rfloor / n$ lorsque n tend vers $+\infty$.

   (d) En déduire $f(x)$, où $x \in \mathbb{R}$.

2. (a) Montrer que toute v.a. X de loi exponentielle vérifie la relation :

   $$\forall x, y \in \mathbb{R}^+, \qquad \mathbf{P}_{[X>y]}(X > x+y) = \mathbf{P}(X > x).$$

   (b) Réciproquement, soit X une v.a. à densité définie sur $\mathbb{R}$ vérifiant la relation précédente et telle que sa fonction de répartition F soit nulle sur $]-\infty, 0]$.

   i. Montrer que s'il existe $x_0 > 0$ tel que $\mathbf{P}(X > x_0) = 0$, alors pour tout $x \geqslant 0$, $\mathbf{P}(X > x) = 0$.

   ii. On suppose que pour tout $x \in \mathbb{R}^+$, $\mathbf{P}(X > x) \neq 0$. Montrer que X suit une loi exponentielle.

Exercice 144. ✦ **Loi du χ^2**
Soit X une variable de loi normale centrée réduite. On note Φ la fonction de répartition de X. On pose $Y = \frac{1}{2}X^2$.

1. Exprimer la fonction de répartition F de Y à l'aide de Φ.
2. En déduire que Y est une variable à densité, et déterminer une densité de Y.

CHAPITRE 8

Diagonalisation

Die Mathematiker sind eine Art Franzosen : redet man zu ihnen, so übersetzen sie es in ihre Sprache, und dann ist es alsobald ganz etwas anderes. [a]

GOETHE

Écrivain allemand (1749-1832)

a. Les mathématiciens sont comme les Français : quoi que vous leur disiez, ils le traduisent dans leur propre langue et le transforment en quelque chose de totalement différent.

1 Définitions

Définition 120 (endomorphisme diagonalisable)

Soit $\varphi \in \mathscr{L}(E)$.
On dit que φ est **diagonalisable** s'il existe une base de E qui soit composée de vecteurs propres de φ.

Exercice 145

✦ Soit $\varphi \in \mathscr{L}(E)$ avec E de dimension finie. *Les questions sont indépendantes.*

1. Que dire de φ si ce dernier est diagonalisable et n'admet qu'une seule valeur propre ?
2. Que dire de φ si ce dernier est diagonalisable et $\mathrm{rg}\left(\varphi^2\right) = 0$?

Définition 121 (matrice diagonalisable)

Soit $A \in \mathscr{M}_n(\mathbb{R})$.
On dit que A est **diagonalisable** s'il existe une matrice inversible P et une matrice diagonale $D \in \mathscr{M}_n(\mathbb{R})$ telles que

$$A = P \cdot D \cdot P^{-1}.$$

Autrement dit, une matrice est diagonalisable si et seulement si elle est semblable à une matrice diagonale.

Remarque. La matrice A est diagonalisable si les colonnes de P forment une base de $\mathscr{M}_{n,1}(\mathbb{R})$ constituée de vecteurs propres de A. De plus, le spectre de A s'identifie au spectre de D qui correspond donc aux coefficients diagonaux de D.

Exemples. • Soit $A = \begin{bmatrix} 0 & 2 & -1 \\ 3 & -2 & 0 \\ -2 & 2 & 1 \end{bmatrix}$. On a montré que

$$\mathrm{Sp}(A) = \{1, 2, -4\} \quad \text{avec} \quad E_1(A) = \mathrm{Vect}\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right), \quad E_2 = \mathrm{Vect}\begin{bmatrix} 4 \\ 3 \\ -2 \end{bmatrix} \quad \text{et} \quad E_{-4} = \mathrm{Vect}\begin{bmatrix} 2 \\ -3 \\ 2 \end{bmatrix}.$$

On pose
$$P = \begin{bmatrix} 1 & 4 & 2 \\ 1 & 3 & -3 \\ 1 & -2 & 2 \end{bmatrix}.$$

Noter que P est la matrice de passage de la base canonique de $\mathcal{M}_{3,1}(\mathbb{R})$ à la base de vecteurs propres de A. On vérifie numériquement :

Editeur

```
D = np.array(([1, 0, 0], [0, 2,
      0], [0, 0, -4]))
 #attention à l'orde des
      valeurs propres
P = np.array(([1, 4, 2], [1, 3,
      -3], [1, -2, 2]))
P_inv = np.linalg.inv(P)

print(P @ D @ P_inv)
```

Editeur

```
>>> print(P @ D @ P_inv)
[[-2.22044605e-16  2.00000000e+00 -1.00000000e+00]
 [ 3.00000000e+00 -2.00000000e+00  0.00000000e+00]
 [-2.00000000e+00  2.00000000e+00  1.00000000e+00]]
```

On retrouve bien A (attention aux arrondis près).

- La matrice $T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ n'est pas diagonalisable.

Exercice 146

✦ **Vrai ou faux ?**

1. La somme de deux matrices diagonalisables est diagonalisable. ✓ ×
2. Si A est diagonalisable alors A^2 est diagonalisable. ✓ ×
3. Si A^2 est diagonalisable alors A est diagonalisable. ✓ ×
4. Si A est inversible, A est diagonalisable si et seulement si A^{-1} est diagonalisable. ✓ ×

Exercice 147

✦ Montrer que si $\mathrm{rg}\left(A^2\right) < \mathrm{rg}(A)$, alors A ne peut-être diagonalisable.

Proposition 122 (lien en dimension finie)

Soit E, un espace vectoriel de dimension finie.
Soient $\varphi \in \mathcal{L}(E)$ et $A \in \mathcal{M}_n(\mathbb{R})$, la matrice de φ dans une base $\mathcal{B}$ de E. On a l'équivalence entre les énoncés.

i) L'endomorphisme φ est diagonalisable.

ii) La matrice A est diagonalisable.

2 Caractérisations

2.1 Version « endomorphisme »

Proposition 123 (caractérisation avec les s.e.p)

Soit $\varphi \in \mathcal{L}(E)$. On a l'équivalence entre les énoncés suivants.

i) L'espace vectoriel E est somme directe des sous-espaces propres de φ.

ii) L'endomorphisme φ est diagonalisable.

Corollaire 124 (caractérisation avec les dimensions)

Soit $\varphi \in \mathscr{L}(E)$ avec E de dimension finie. On a l'équivalence entre les énoncés suivants.

i) $\sum_{\lambda \in \mathrm{Sp}(\varphi)} \dim\left(E_\lambda(\varphi)\right) = \dim(E)$.

ii) L'endomorphisme φ est diagonalisable.

Remarque. Comme $\dim\left(E_\lambda(\varphi)\right) \geqslant 1$, on retrouve le fait qu'un endomorphisme de dimension finie a au plus $\dim(E)$ valeurs propres.

Corollaire 125 (cas particulier)

Soit $\varphi \in \mathscr{L}(E)$ avec E de dimension finie.

Si φ possède $\dim(E)$ valeurs propres distinctes,

alors φ est diagonalisable et les sous-espaces propres sont tous de dimension 1.

⚠ **Attention.** La réciproque est fausse. Par exemple, pour E de dimension $n \geqslant 2$, l'endomorphisme id_E est diagonalisable avec seulement une valeur propre (1).

Exercice 148

✦ **Exemple.**
Soit $n \in \mathbb{N} \setminus \{0; 1\}$. Posons pour tout polynôme $P \in \mathbb{R}[x]$, le polynôme $\varphi(P)$ défini par

$$\varphi(P)(x) = \frac{1}{n}x(1-x)P'(x) + xP(x).$$

1. Vérifier φ induit un endomorphisme de $\mathbb{R}_n[x]$.
2. Pour tout $k \in [\![0; n]\!]$, on pose $P_k(x) = x^k(1-x)^{n-k}$. Calculer $\varphi(P_k)$.
3. Justifier que φ est diagonalisable.

2.2 Version « matricielle »

Regroupons et traduisons les résultats précédents.

Théorème 126 (caractérisations)

Soit $A \in \mathscr{M}_n(\mathbb{R})$. Les énoncés suivants sont équivalents.

i) La matrice A est diagonalisable.

ii) Il existe une base de $\mathscr{M}_{n,1}(\mathbb{R})$ formée de vecteurs propres de A.

iii) $\mathscr{M}_{n,1}(\mathbb{R})$ est somme directe des sous-espaces propres de A.

iv) $\sum_{\lambda \in \mathrm{Sp}(A)} \dim\left(E_\lambda(A)\right) = n$.

Exemple. La matrice Attila.

Proposition 127 (n valeurs propres)

Soit $A \in \mathscr{M}_n(\mathbb{R})$.

Si A admet n valeurs propres distinctes,

alors A est diagonalisable et les sous-espaces propres sont tous de dimension 1.

Python. La commande `eigvals` permet le calcul de valeurs propres. Par exemple :

Editeur
```
import numpy.linalg as al
 # On importe la sous-bibliothèque
     linalg
A=np.array([[1,3,0],[0,-2,0],[-1,-2,0]])
 # On définit la matrice A
print(al.eigvals(A))
```

Console
```
>>> # script executed
[ 0.   1.  -2.]
```

Selon ce calcul, 0, 1 et -2 sont toutes les valeurs propres de A. La matrice A est diagonalisable.

3 Compléments

3.1 Cas particuliers

Cas des matrices de taille 2

Rappelons que pour $A \in \mathcal{M}_2(\mathbb{R})$ et $\lambda \in \mathbb{R}$.

$$\lambda \in \mathrm{Sp}(A) \iff \det(A - \lambda I_2) = 0.$$

Exercice 149

✦

Soit $A \in \mathcal{M}_2(\mathbb{R})$ diagonalisable. Notons λ_1 et λ_2, les deux valeurs propres éventuellement confondues de la matrice A.

1. Montrer que $\lambda_1 + \lambda_2 = \mathrm{tr}(A)$ et $\lambda_1\lambda_2 = \det(A)$.
2. En minimisant le nombre de calcul, montrer que la matrice $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ n'est pas diagonalisable dans $\mathbb{R}$.

p. ??

Cas des matrices triangulaires

Exercice 150

✦ *Les questions 1 et 2 sont indépendantes.*

1. Est-ce que les matrices suivantes sont diagonalisables ?

$$T_1 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad T_2 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad \text{et} \quad T_3 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

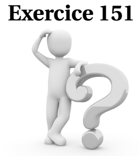
2. À quelle condition sur $\alpha \in \mathbb{R}$, la matrice $M_\alpha = \begin{bmatrix} \alpha^2 & 1 \\ 0 & \alpha \end{bmatrix}$ est diagonalisable ?

Cas des matrices symétriques réelles

Théorème 128 (cas symétrique, première version)

Toute matrice symétrique réelle est diagonalisable.

Exercice 151



✦ *Les questions 1 et 2 sont indépendantes*

1. Montrer que l'endomorphisme suivant est diagonalisable.

$$\varphi : \left\{ \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R}^3 \\ (x, y, z) & \mapsto & (2x + y + z, x + 3z, x + 3y - z). \end{array} \right.$$

2. Soient $n \in \mathbb{N}^*$ et A une matrice symétrique appartenant à $\mathcal{M}_n(\mathbb{R})$ vérifiant $A^n = I_n$. Calculer A^2.

Cas des projecteurs

- Soit p, un projecteur de E (avec $p \neq 0_{\mathscr{L}(\mathrm{E})}$ et $p \neq \mathrm{id}_{\mathrm{E}}$). En reprenant l'étude effectué au chapitre précédent, on a
$$\mathrm{E} = \mathrm{E}_0(p) \oplus \mathrm{E}_1(p).$$

Les projecteurs sont des endomorphismes diagonalisables. Si $\mathscr{B}$ est une base adaptée à la décomposition en sous-espaces propres, la matrice de p dans la base $\mathscr{B}$ est diagonale.
En particulier, on constate que $\mathrm{Tr}\big(\mathrm{Mat}_{\mathscr{B}}(p)\big) = \mathrm{rg}\big(\mathrm{Mat}_{\mathscr{B}}(p)\big) = \mathrm{rg}(p)$.

$$\mathrm{Mat}_{\mathscr{B}}(p) = \left[\begin{array}{cccccc} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ \underbrace{0 \quad \cdots \quad 0}_{\dim \mathrm{E}_1(p)} & & & \underbrace{0 \quad \cdots \quad 0}_{\dim \mathrm{E}_0(p)} & & \end{array}\right].$$

3.2 Pratique de la diagonalisation et applications

En reprenant les méthodes étudiées au chapitre VALEURS PROPRES, VECTEURS PROPRES, traiter les exercices suivants.

Exercice 152

✦ Si possible, diagonaliser les matrices suivantes :
$$\mathrm{A} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix} \quad \text{et} \quad \mathrm{B} = \begin{bmatrix} -4 & 6 & -3 \\ -1 & 3 & -1 \\ 4 & -4 & 3 \end{bmatrix}.$$

Exercice 153

✦ Considérons l'application $\varphi : \mathrm{P} \in \mathbb{R}_2[x] \mapsto x(1-x)\mathrm{P}'(x) + 2x\mathrm{P}(x) \in \mathbb{R}_2[x]$.

1. Montrer que φ est un endomorphisme de $\mathbb{R}_2[x]$.
2. Exprimer la matrice de φ dans la base canonique. La diagonaliser.
3. Conclure en donnant une base de vecteurs propres de φ.

Astuce. Dans la recherche des valeurs, il ne faut pas oublier que pour une matrice diagonalisable
$$\mathrm{Tr}(\mathrm{A}) = \sum_{\lambda \in \mathrm{Sp}(\mathrm{A})} \lambda \times \dim\big(\mathrm{E}_\lambda(\mathrm{A})\big).$$

Exercice 154

✦

1. Prouver la remarque précédente.
2. Soit A définie par $\begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix}$.
   Sachant que $\mathrm{rg}(\mathrm{A} + 2\mathrm{I}_n) = 1$, que peut-on en déduire sur la diagonalisation de A ?

Exercice 155

✧ **Calcul des puissances**
Calculer pour tout $p \in \mathbb{N}$, A^p où la matrice A est étudiée à l'exercice 152.

Exercice 156

✦ **Polynôme de matrices et racine carrée d'une matrice**

On pose
$$\mathrm{A} = \begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix}.$$

1. Vérifier que A est diagonalisable et la diagonaliser.
2. En déduire l'inversibilité de A et A^{-1}.
3. Expliquer comment calculer Q(A) où $\mathrm{Q} \in \mathbb{R}[x]$. Préciser un polynôme annulateur non nul de A.
4. Déterminer une matrice B telle que $\mathrm{B}^2 = \mathrm{A}$.

Les applications sont nombreuses. Citons par exemple : la recherche du commutant (voir exercice 165, p.89) ; la résolution des suites récurrentes linéaires d'ordre 2 (voir exercice 166, p.89) ; la résolution de systèmes différentiels linéaires (voir exercice 167, p.90).

Exercices - TD

Exercice 157. ✧ Montrer que les matrices suivantes sont semblables

$$A=\begin{bmatrix}1&2&3\\0&2&3\\0&0&3\end{bmatrix}\quad\text{et}\quad B=\begin{bmatrix}3&0&0\\3&2&0\\3&2&1\end{bmatrix}.$$

Exercice 158. ✦✦ **Diagonalisation avec un paramètre**
Pour tout réel a, on pose

$$M_a=\begin{bmatrix}a+2&-(2a+1)&a\\1&0&0\\0&1&0\end{bmatrix}.$$

On vérifie par le calcul que $Q(x)=x^3-(a+2)x^2+(2a+1)x-a$ est annulateur de M_a.

1. Justifier que pour $a=1$, M_a ne peut être diagonalisable.
2. Étudier le cas général.

Exercice 159. ✦
On considère l'application φ, qui à tout polynôme P de $\mathbb{R}_n[x]$ associe $\varphi(P)=\sum_{k=0}^{n}P^{(k)}$, où $P^{(k)}$ désigne la dérivée $k^{\text{ème}}$ de P avec la convention $P^{(0)}=P$.

1. Montrer que φ est un endomorphisme de $\mathbb{R}_n[x]$.
2. Est-ce que φ est diagonalisable ?

Exercice 160. ✦✦ *D'après EDHEC 2014*
Soient $n\in\mathbb{N}\setminus\{0;1\}$ et A une matrice non nulle donnée de $\mathcal{M}_n(\mathbb{R})$. On considère l'application f qui à toute matrice M de $\mathcal{M}(\mathbb{R})$ associe :

$$f(M)=\mathrm{Tr}(A)\,M-\mathrm{Tr}(M)\,A.$$

1. Montrer que f est un endomorphisme de $\mathcal{M}_n(\mathbb{R})$.
2. (a) Pour toute matrice M de $\mathcal{M}_n(\mathbb{R})$, exprimer $(f\circ f)(M)$ à l'aide de Tr(A) et $f(M)$.
   (b) En déduire un polynôme annulateur de f. Que peut-on en déduire sur les valeurs propres de f ?
3. (a) Montrer que 0 est valeur propre de f.
   (b) Montrer que, si $\mathrm{Tr}(A)=0$, alors f n'est pas diagonalisable.
4. On suppose dans cette question que la trace de A est non nulle.
   (a) Préciser la dimension de ker(Tr).
   (b) En déduire que f est diagonalisable.

Exercice 161. ✦✦ **Diagonalisation des matrices de rang 1** *D'après EMLyon 2013*

1. Montrer que $M\in\mathcal{M}_n(\mathbb{R})$ est de rang 1 si et seulement si il existe deux matrices colonnes non nulles U, V telles que $M=U\,{}^t V$.
2. Soit $A\in\mathcal{M}_n(\mathbb{R})$ une matrice de rang 1. On note U et V deux matrices colonnes non nulles de $\mathcal{M}_{n,1}(\mathbb{R})$ telles que $A=U\,{}^tV$ et on note $a=\mathrm{Tr}(A)$.
   (a) Montrer que 0 est valeur propre de A et déterminer la dimension du sous-espace propre associé.
   (b) Montrer : ${}^tVU=(a)$, puis : $A^2=aA$.
   (c) Montrer que si $a=0$, alors A n'est pas diagonalisable dans $\mathcal{M}_n(\mathbb{R})$.
   (d) On suppose $a\neq 0$. Calculer AU. Déduire des questions précédentes que A est diagonalisable.
   (e) Énoncer une condition nécessaire et suffisante pour qu'une matrice de $\mathcal{M}_n(\mathbb{R})$ de rang 1 soit diagonalisable.

Exercice 162. ✦✦✦ *D'après Oraux HEC 2014*
Soit φ l'endomorphisme de $\mathbb{R}^3$ dont la matrice dans la base canonique $\mathscr{B}=(e_1,e_2,e_3)$ est

$$A=\begin{bmatrix}2&0&0\\1&3&-2\\1&1&0\end{bmatrix}.$$

1. Montrer que $\varphi-\mathrm{id}_{\mathbb{R}^3}$ est un projecteur.
2. En déduire les valeurs propres de φ ?
3. Combien existe-t-il de droites vectorielles de $\mathbb{R}^3$ stables par φ ?

Exercice 163. ✦✦ *D'après Oraux HEC 2013*

Soient $n \in \mathbb{N}^*$ et a et b deux réels tels que $ab \neq 0$. On note $\mathrm{M}(a,b)$ la matrice de $\mathscr{M}_{n+1}(\mathbb{R})$ donnée par :

$$\mathrm{M}(a,b) = \begin{bmatrix} 0 & a & a & \cdots & a \\ b & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ b & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

1. (a) Calculer $\mathrm{M}(a,b)^2$.

   (b) Montrer que $\mathrm{M}(a,b)^2$ est diagonalisable et trouver ses valeurs propres.

2. Montrer que $\mathrm{M}(c,d) = \begin{bmatrix} 0 & c & c & \cdots & c \\ d & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ d & 0 & 0 & \cdots & 0 \end{bmatrix}$ est semblable à $\mathrm{M}(a,b)$ si et seulement si $ab = cd$.

3. (a) Est-ce que la matrice $\mathrm{M}(a,b)$ est semblable à sa transposée ?

   (b) À l'aide de la trace, montrer que si la matrice $\mathrm{M}(a,b)$ est diagonalisable alors $ab > 0$.

Exercice 164. ✦✦✦ Soit E un espace vectoriel de dimension finie.

1. Soit f, un endomorphisme de E diagonalisable. Montrer que f admet un polynôme annulateur scindé à racines simples.
   Pour rappel, un polynôme P *est scindé à racines simples s'il existe r réels* $a_1, \ldots, a_r$ *tels que* $\mathrm{P}(x) = \prod_{i=1}^{r}(x - a_i)$.

2. La question suivante a pour objectif de prouver la réciproque. Si f admet un polynôme annulateur scindé à racines simples, alors f est diagonalisable.

   (a) Soient $f, g \in \mathscr{L}(\mathrm{E})$, montrer que $\dim \mathrm{Ker}(f \circ g) \leqslant \dim\big(\mathrm{Ker}(f)\big) + \dim\big(\mathrm{Ker}(g)\big)$.

   (b) Montrer plus généralement que pour $f_1, f_2, \ldots, f_r \in \mathscr{L}(\mathrm{E})$,

   $$\dim\big(\mathrm{Ker}\big(f_1 \circ \cdots \circ f_r\big)\big) \leqslant \sum_{j=1}^{r} \dim\big(\mathrm{Ker}(f_j)\big).$$

   (c) En déduire la réciproque.

3. *Applications*

   (a) Retrouver le fait qu'un projecteur et qu'une symétrie de E sont diagonalisables.

   (b) En déduire que si f est diagonalisable et stabilise un sous-espace vectoriel F, alors il induit sur celui-ci un endomorphisme diagonalisable.

Quelques applications de la diagonalisation

Exercice 165. ✦✦ **Recherche du commutant** *D'après ESCP 2012*

Soit $\mathrm{A} \in \mathscr{M}_n(\mathbb{R})$ admettant n valeurs propres distinctes. On définit le commutant de A par

$$\mathscr{C} = \big\{\mathrm{M} \in \mathscr{M}_n(\mathbb{R}) \mid \mathrm{AM} = \mathrm{MA}\big\}.$$

1. Justifier que la famille $\big(\mathrm{I}_n, \mathrm{A}, \mathrm{A}^2, \ldots, \mathrm{A}^{n-1}\big)$ est libre.
2. Vérifier que $\mathscr{C}$ est un sous-espace vectoriel de $\mathscr{M}_n(\mathbb{R})$ de dimension supérieure ou égale à n.
3. Montrer l'existence d'une matrice P de $\mathscr{M}_n(\mathbb{R})$, inversible et d'une matrice Δ de $\mathscr{M}_n(\mathbb{R})$ diagonale, telles que

   $$\mathrm{A} = \mathrm{P}\Delta\mathrm{P}^{-1}.$$

4. Soit $\mathrm{M} \in \mathscr{C}$. Montrer que tout vecteur colonne propre de A est un vecteur colonne propre de M. En déduire que la matrice $\mathrm{P}^{-1}\mathrm{MP}$ est diagonale. En déduire que $\mathscr{C}$ est de dimension inférieure ou égale à n.
5. Montrer que $\big(\mathrm{I}_n, \mathrm{A}, \ldots, \mathrm{A}^{n-1}\big)$ est une base de $\mathscr{C}$.
6. On suppose A inversible. Montrer qu'il existe $\mathrm{Q} \in \mathbb{R}[x]$ tel que $\mathrm{A}^{-1} = \mathrm{Q}(\mathrm{A})$.

Exercice 166. ✦✦ **Suite récurrente linéaire d'ordre 2**

Soit $a \in \mathbb{R} \setminus \{1\}$. On note E l'espace vectoriel des suites réelles $(u_n)_{n\in\mathbb{N}}$ telles que

$$\forall\, n \in \mathbb{N}, \qquad u_{n+2} = (1+a)u_{n+1} - a u_n.$$

Soit u, une suite de E. On pose $\mathrm{U}_n = \begin{bmatrix} u_n \\ u_{n+1} \end{bmatrix}$.

1. Déterminer une matrice A telle que $\mathrm{U}_{n+1} = \mathrm{AU}_n$.

2. (a) Montrer que la matrice A est diagonalisable sur $\mathbb{R}$. Puis, préciser une matrice inversible et une matrice D diagonale telles que $A = PDP^{-1}$.

   (b) En déduire A^n, pour tout $n \in \mathbb{N}$.

3. À partir des questions précédentes, donner l'expression de u_n en fonction de n, a, u_0 et u_1.
4. Donner une base de E. Comparer les résultats obtenus avec la méthode classique des suites récurrentes linéaires d'ordre 2.

Exercice 167. ✦✦ **Système différentiel linéaire**

1. *Préliminaires*
   Soient I intervalle de $\mathbb{R}$ et a, une fonction continue sur I. On considère l'équation différentielle

$$\forall x \in I, \qquad y'(x) = a(x)y(x).$$

   Soit A, une primitive de a sur I. Montrer qu'il existe $C \in \mathbb{R}$ tel que pour tout $x \in I$, $y(x) = C e^{A(x)}$.

2. On considère le système différentiel suivant :

$$(\mathscr{S}) : \quad \begin{cases} x' = 8x - 18y + 27z \\ y' = -3x + \frac{7}{2}y - 6z \\ z' = -4x + 7y - 11z \end{cases}$$

   avec les conditions initiales : $x(0) = 1, \quad y(0) = 0, \quad z(0) = 0.$

   (a) Écrire le système $(\mathscr{S})$ ci-dessus sous la forme $X' = AX$, pour une certaine matrice A de taille 3×3 à coefficients réels qu'on déterminera où on a posé :

$$X(t) = \begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix} \quad \text{et} \quad X'(t) = \begin{bmatrix} x'(t) \\ y'(t) \\ z'(t) \end{bmatrix}.$$

   (b) Vérifier que la matrice A est diagonalisable et déterminer une matrice inversible Q et une matrice diagonale D telles que $A = Q^{-1}DQ$.
   Pour commencer, on pourra calculer AX_1, AX_2 *où :*

$$X_1 = \begin{bmatrix} -\frac{3}{2} \\ 1 \\ 1 \end{bmatrix} \quad \text{et} \quad X_2 = \begin{bmatrix} 0 \\ \frac{3}{2} \\ 1 \end{bmatrix}.$$

   On admet dans la suite que pour toute matrice Q *à coefficients constants, si* $Y = Q \cdot X$ *alors* $Y' = Q \cdot X'$.

   (c) Pour tout $t \in \mathbb{R}$, on pose $Y(t) = QX(t)$. Montrer que X est solution du système (E) si et seulement si les coordonnées u, v et w de Y sont solutions d'un système différentiel diagonal.

   (d) Donner l'expression de $Y(t)$ puis les expressions de x, y et z.

CHAPITRE 9

Algèbre bilinéaire

> Ἀγεωμέτρητος μηδεὶς εἰσίτω
>
> *Que nul n'entre ici s'il n'est géomètre.*
>
> Inscription que Platon aurait fait graver à l'entrée de l'Académie, son école d'Athènes.

1 Produits scalaires

1.1 Définitions et exemples

Définition 129 (forme bilinéaire)

Soient E un espace vectoriel et $\varphi : E \times E \to \mathbb{R}$ une application.
On dit que φ est une **forme bilinéaire** si elle est

- linéaire à droite
$$\forall\, u,\ v,\ w \in E, \quad \forall \lambda,\ \mu \in \mathbb{R}, \qquad \varphi(u\,,\,\lambda v + \mu w) = \lambda\varphi(u,v) + \mu\varphi(u,w).$$
- linéaire à gauche
$$\forall\, u,\ v,\ w \in E, \quad \forall \lambda,\ \mu \in \mathbb{R}, \qquad \varphi(\lambda v + \mu w\,,\ u) = \lambda\varphi(v,u) + \mu\varphi(w,u).$$

Définition 130 (produit scalaire)

Soit $\varphi : E \times E \to \mathbb{R}$ une application. On dit que φ est un **produit scalaire** si :

- φ est une forme bilinéaire.
- φ est symétrique : $\forall\, u,\ v \in E, \quad \varphi(u,v) = \varphi(v,u).$
- φ est positive : $\forall\, u \in E, \quad \varphi(u,u) \geqslant 0.$
- φ est définie : $\forall\, u \in E, \quad \varphi(u,u) = 0 \quad \Rightarrow \quad u = 0_E.$

Remarque. Si φ est linéaire à droite et symétrique alors φ est linéaire à gauche et donc bilinéaire. En pratique, on démontre donc d'abord la symétrie et une linéarité (à gauche ou droite) pour obtenir la condition de bilinéarité.

Notation. On écrit souvent $\langle u, v\rangle$ au lieu de $\varphi(u,v)$.

Exemples.

- *Dans* $\mathbb{R}^n$.

Pour tous $u = (x_1, \ldots, x_n)$ et $v = \big(y_1, \ldots, y_n\big)$ dans $\mathbb{R}^n$, posons $\langle u, v\rangle = \sum_{i=1}^{n} x_i y_i$.

Alors $\langle . , . \rangle$ est un produit scalaire sur $\mathbb{R}^n$, appelé **produit scalaire canonique sur** $\mathbb{R}^n$.

- *Dans* $\mathbb{R}_n[x]$. ⋙ *Voir exercice 173.*

Donnons deux exemples dans le cas polynomial.

→ Soient $a_0, a_1, \dots, a_n$, $n+1$ réels deux à deux distincts. L'application suivante est un produit scalaire

$$(\mathrm{P}, \mathrm{Q}) \in \mathbb{R}_n[x] \mapsto \langle \mathrm{P}, \mathrm{Q} \rangle = \sum_{i=0}^{n} \mathrm{P}(a_k)\mathrm{Q}(a_k).$$

→ On peut aussi poser pour $a, b \in \mathbb{R}$ avec $a < b$

$$(\mathrm{P}, \mathrm{Q}) \in \mathbb{R}_n[x] \mapsto \langle \mathrm{P}, \mathrm{Q} \rangle = \int_a^b \mathrm{P}(t)\mathrm{Q}(t)\, \mathrm{d}t.$$

- *Dans* $\mathcal{M}_{n,1}(\mathbb{R})$.

L'application $\varphi : (\mathrm{X}, \mathrm{Y}) \mapsto {}^t\mathrm{X}\mathrm{Y}$ est un produit scalaire sur $\mathcal{M}_{n,1}(\mathbb{R})$. On l'appelle produit scalaire canonique de $\mathcal{M}_{n,1}(\mathbb{R})$. Noter, comme souvent, qu'on identifie $\mathcal{M}_{1,1}(\mathbb{R})$ et $\mathbb{R}$.

- *Dans* $\mathcal{M}_n[\mathbb{R}]$. ⋙ *Voir exercice 176.*

Pour $\mathrm{A}, \mathrm{B} \in \mathcal{M}_n(\mathbb{R})$, on pose $\langle \mathrm{A}, \mathrm{B} \rangle = \mathrm{Tr}\left(\mathrm{A}^t\mathrm{B}\right).$

- *Dans* $\mathscr{C}^0([a;b])$ avec $a < b$. ⋙ *Voir exercice 171.*

Pour $f, g \in \mathscr{C}^0([a;b])$ $\langle f, g \rangle = \displaystyle\int_a^b f(t)g(t)\, \mathrm{d}t.$

Définition 131 (norme)

Soit E, un espace vectoriel muni d'un produit scalaire φ. L'application

$$\mathrm{N} : \begin{cases} \mathrm{E} & \to & \mathbb{R} \\ u & \mapsto & \mathrm{N}(u) = \sqrt{\varphi(u, u)} \end{cases}$$

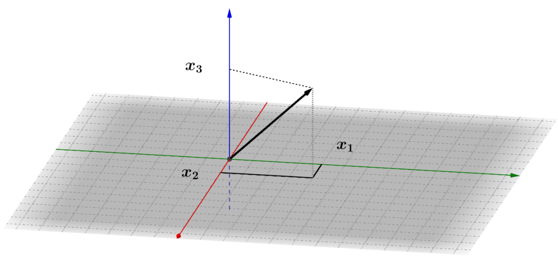
est appelée **norme associée au produit scalaire** φ.

Notation. On note souvent $\|\cdot\|$ au lieu de N.

Exemple. Dans $\mathbb{R}^n$, la norme euclidienne associée au produit scalaire canonique est définie par :

$$\forall\, u = (x_1, \dots, x_n) \in \mathbb{R}^n, \qquad \|u\| = \sqrt{\sum_{i=1}^{n} {x_i}^2}.$$

Dans $\mathbb{R}^2$ ou $\mathbb{R}^3$, on constate que la norme représente la distance à l'origine, ou encore la longueur du vecteur.



Remarque. Par extension, pour $u, v \in \mathrm{E}$, $\|u\|$ représente la « longueur » du vecteur u alors que $\|u - v\|$ correspond à la « distance » entre les deux vecteurs.

1.2 Propriétés du produit scalaire, de la norme

Proposition 132 (règles de calcul)

Soit E un espace vectoriel muni d'un produit scalaire $\langle \cdot, \cdot \rangle$ dont $\|\cdot\|$ est la norme associée.

- $\forall\, u \in \mathrm{E}, \qquad \|u\| = 0 \iff u = 0_{\mathrm{E}}.$
- $\forall\, \lambda \in \mathbb{R}, \forall\, u \in \mathrm{E}, \qquad \|\lambda u\| = |\lambda| \cdot \|u\|.$
- $\forall\, u, v \in \mathrm{E}, \qquad \|u + v\|^2 = \|u\|^2 + 2\langle u, v \rangle + \|v\|^2.$

Exercice 168

✦ **Identité du parallélogramme et formule de polarisation**
Montrer que pour tous $u, v \in \mathrm{E}$,

$$\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2 \quad \text{et} \quad \langle u, v\rangle = \frac{1}{4}\left(\|u+v\|^2 - \|u-v\|^2\right).$$

Théorème 133 (inégalité de Cauchy-Schwarz)

Soit E un espace vectoriel muni d'un produit scalaire $\langle\cdot,\cdot\rangle$ dont $\|\cdot\|$ est la norme associée.

$$\forall\, u, v \in \mathrm{E}, \qquad |\langle u, v\rangle| \leqslant \|u\| \cdot \|v\|.$$

De plus, on a égalité si et seulement si la famille (u, v) est liée.

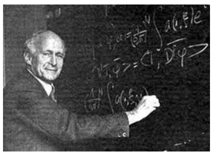
Exercice 169

✦✦ **Preuve**
On suppose $u \neq 0_{\mathrm{E}}$ et on définit l'application P sur $\mathbb{R}$ par $\mathrm{P}(\lambda) = \|\lambda u + t v\|^2$.

1. Vérifier que P est un polynôme de degré 2. Préciser son discriminant
2. Conclure (sans oublier le cas d'égalité).

Exercice 170

✧ **Qui est qui ?**
Parmi les photos ci-dessous, reconnaître Hermann Amandus Schwarz (mathématicien allemand) et Laurent Schwar**t**z (mathématicien français, médaille Fields pour ses travaux sur la théorie des distributions).



Exemple. En reprenant le produit scalaire canonique sur $\mathbb{R}^n$, l'inégalité de Cauchy-Schwarz devient

$$\forall\,(x_i)_{i\in[\![1,n]\!]} \in \mathbb{R}^n, \quad \forall\,(y_i)_{i\in[\![1,n]\!]} \in \mathbb{R}^n, \qquad \left|\sum_{i=1}^{n} x_i y_i\right| \leqslant \sqrt{\sum_{i=1}^{n} {x_i}^2}\,\sqrt{\sum_{i=1}^{n} {y_i}^2}.$$

"Le cas d'égalité de l'inégalité de Cauchy-Schwarz est méconnu et beaucoup des tentatives pour prouver l'inégalité et son cas d'égalité ne sont que paraphrases et esbroufe."

Rapport de Jury : HEC 2019

Exercice 171

✦ 1. Justifier que l'application suivante est un produit scalaire sur $\mathscr{C}^0([a;b])$.

$$\forall\, f, g \in \mathscr{C}^0([a;b]), \qquad \varphi(f,g) = \int_a^b f(t)g(t)\,\mathrm{d}t.$$

2. Expliciter l'inégalité de Cauchy-Schwarz pour ce produit scalaire.
3. Montrer que si $f \in \mathscr{C}^0([0;1])$ avec f ne s'annulant pas sur $[0;1]$

$$\int_0^1 f(t)\,\mathrm{d}t \cdot \int_0^1 \frac{1}{f(t)}\,\mathrm{d}t \geqslant 1.$$

Préciser le cas d'égalité.

Proposition 134 (inégalité triangulaire)

Pour tous $u, v \in \mathrm{E}$,

$$\|u+v\| \leqslant \|u\| + \|v\|.$$

Exercice 172

✦ Soient E un espace euclidien et $B = \{x \in E \mid \|x\| \leqslant 1\}$.

1. Soient $u, v \in E$.
   Que peut-on dire de u et v si on a le cas d'égalité $\|u+v\| = \|u\| + \|v\|$?
2. Démontrer que B est une partie strictement convexe de E, c'est-à-dire que, pour tous $x, y \in B$ avec $x \neq y$, tout $t \in]0,1[$, on a $\|tx + (1-t)y\| < 1$.
3. Illustrer ce résultat dans $\mathbb{R}^2$ et $\mathbb{R}^3$ pour le produit scalaire canonique.

Remarque. Par récurrence, on montre que pour tout famille finie $(u_1, u_2, \ldots, u_p)$ de vecteurs de E

$$\left\| \sum_{i=1}^{p} u_i \right\| \leqslant \sum_{i=1}^{p} \|u_i\|.$$

1.3 Orthogonalité

Orthogonalité et vecteurs

Définition 135 (vecteurs orthogonaux)

Soient E, un espace vectoriel muni d'un produit scalaire $\langle \cdot, \cdot \rangle$.
Deux vecteurs u et v de E sont dits **orthogonaux**, noté $u \perp v$, si $\langle u, v \rangle = 0$.

Remarque. Le seul vecteur de E qui soit orthogonal à tous les autres vecteurs de E est le vecteur nul. Autrement dit, pour tout $u \in E$, on a l'équivalence :

$$\big(\forall\, v \in E, \qquad \langle u, v\rangle = 0\big) \quad \Longleftrightarrow \quad u = 0_E.$$

⚠ **Attention.** La notion d'orthogonalité est relative au produit scalaire.

Exercice 173

✦✦ **Deux produits scalaires sur** $\mathbb{R}_2[x]$
Pour tous $P, Q \in \mathbb{R}_2[x]$, on pose

$$\varphi_1(P,Q) = P(0)Q(0) + P(1)Q(1) + P(2)Q(2) \quad \text{et} \quad \varphi_2(P,Q) = \int_{-1}^{1} P(t)Q(t)\,dt.$$

1. Vérifier que φ_1 et φ_2 définissent deux produits scalaires sur $\mathbb{R}_2[x]$.
2. Notons P_1 le polynôme d'expression $P_1(x) = x$.
   Donner un vecteur orthogonal à P_1 relativement à φ_1 mais non orthogonal pour φ_2 et inversement.

Théorème 136 (de Pythagore)

Deux vecteurs u et v sont **orthogonaux** si et seulement si

$$\|u+v\|^2 = \|u\|^2 + \|v\|^2.$$

Définition 137 (famille normée, orthogonale, orthonormée)

Soit $\mathscr{F} = (u_1, \ldots, u_p)$ une famille de vecteurs de E.

- La famille est dite **normée** si pour tout $i \in [\![1;p]\!]$, $\|u_i\| = 1$.
- La famille est dite **orthogonale** si, pour tout $(i,j) \in [\![1;p]\!]^2$ avec $i \neq j$, $u_i \perp u_j$.
- La famille est dite une famille **orthonormée** (ou orthonormale) si c'est une famille orthogonale et normée.

Remarques. Autrement dit, la famille de vecteurs $(u_1, \dots, u_p)$ est orthonormée si et seulement si

$$\forall (i,j) \in [\![1,p]\!]^2, \qquad \langle u_i, u_j \rangle = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j. \end{cases}$$

Le théorème de Pythagore se généralise, si $(u_1, \dots, u_p)$ est une famille orthogonale de vecteurs de E, alors

$$\left\| \sum_{i=1}^{p} u_i \right\|^2 = \sum_{i=1}^{p} \|u_i\|^2.$$

Proposition 138 (orthogonalité implique liberté)

Soit $\mathscr{F}$ une famille de vecteurs de E.

- **Si** $\mathscr{F}$ est orthogonale, et si aucun des vecteurs de $\mathscr{F}$ n'est le vecteur nul, **alors** $\mathscr{F}$ est libre.
- **Si** $\mathscr{F}$ est orthonormée, **alors** $\mathscr{F}$ est libre.

Remarque. Les réciproques sont fausses.

Exercice 174

✦ Soit $E = \mathscr{C}^0([-\pi, \pi])$ muni du produit scalaire défini par

$$\forall f, g \in E, \qquad \langle f, g \rangle = \int_{-\pi}^{\pi} f(t) g(t)\, dt.$$

1. Justifier que la famille $(x \mapsto \cos(kx))_{1 \leqslant k \leqslant n}$ est orthogonale. Est-elle orthonormée ?
2. En déduire que la famille $(x \mapsto \cos(kx),\ x \mapsto \sin(kx))_{1 \leqslant k \leqslant n}$ est une famille libre de E.

Orthogonalité et sous-espaces vectoriels

Définition 139 (sous-espaces orthogonaux)

Soient F et G deux sous-espaces vectoriels de E. On dit que F et G sont **orthogonaux** si

$$\forall u \in F, \quad \forall v \in G, \qquad u \perp v.$$

Remarque. Soient $(e_i)_{i \in [\![1;n]\!]}$ et $(\varepsilon_j)_{j \in [\![1;p]\!]}$ deux familles respectivement de F et G.
Si les familles sont génératrices, alors le sous espace vectoriel F est orthogonal à G si

$$\forall i \in [\![1,n]\!], \quad \forall j \in [\![1,p]\!], \qquad e_i \perp \varepsilon_j.$$

Exercice 175

✧ Dans $\mathbb{R}^3$, on pose $u = (1,2,3)$ et

$$F = \left\{ (x,y,z) \in \mathbb{R}^3 \mid x + 2y + 3z = 0 \right\} \quad \text{et} \quad G = \mathrm{Vect}(u).$$

1. Donner une famille génératrice de F.
2. Vérifier que F et G sont orthogonaux.

Exercice 176

✦ *Exemple dans* $\mathscr{M}_n(\mathbb{R})$.

1. Montrer que l'application φ définie sur $\mathscr{M}_n(\mathbb{R})^2$ par $\varphi(A,B) = \mathrm{Tr}(A^t B)$ est un produit scalaire.
2. Notons $\mathscr{S}_n(\mathbb{R})$ (resp. $\mathscr{A}_n(\mathbb{R})$), l'ensemble des matrices symétriques (resp. antisymétriques). Montrer que $\mathscr{S}_n(\mathbb{R})$ et $\mathscr{A}_n(\mathbb{R})$ sont orthogonaux pour ce produit scalaire.

Définition-proposition 140 (le s.e.v orthogonal)

Soit F une partie de E muni d'un produit scalaire $\langle\cdot,\cdot\rangle$. Posons

$$\mathrm{F}^{\perp} = \{x \in \mathrm{E} \mid \forall\, y \in \mathrm{F}, \quad x \perp y\}.$$

Alors :

- L'ensemble $\mathrm{F}^{\perp}$ est un sous-espace vectoriel de E.
- Les sous-espaces F et $\mathrm{F}^{\perp}$ sont orthogonaux.

Le sous-espace vectoriel $\mathrm{F}^{\perp}$ s'appelle **l'orthogonal** de F dans E.

Exemples. On a $\mathrm{E}^{\perp} = \{0_{\mathrm{E}}\}$ et $\{0_{\mathrm{E}}\}^{\perp} = \mathrm{E}$.

Remarques.
• Si G est un s.e.v de E tel que G est orthogonal à F, alors $\mathrm{G} \subset \mathrm{F}^{\perp}$. Autrement dit, $\mathrm{F}^{\perp}$ est le plus grand (au sens de l'inclusion) sous-espace vectoriel de E qui soit orthogonal à F.
• Si $\mathrm{F} \subset \mathrm{G}$ alors $\mathrm{G}^{\perp} \subset \mathrm{F}^{\perp}$.

Exercice 177

✦ Soient F, G deux sous-espaces vectoriels de E. Prouver les énoncés suivants.

1. $\mathrm{F} \subset (\mathrm{F}^{\perp})^{\perp}$.
2. $\mathrm{F}^{\perp} + \mathrm{G}^{\perp} \subset (\mathrm{F} \cap \mathrm{G})^{\perp}$.
3. $(\mathrm{F} \cup \mathrm{G})^{\perp} = \mathrm{F}^{\perp} \cap \mathrm{G}^{\perp}$.
4. $\mathrm{F}^{\perp} \cap \mathrm{G}^{\perp} \subset (\mathrm{F} + \mathrm{G})^{\perp}$.

2 Espaces euclidiens

2.1 Définitions et exemples

Définition 141 (espaces euclidiens)

Un **espace euclidien** est la donnée :

→ d'un espace vectoriel E de dimension finie,
→ d'un produit scalaire sur E.

On note $(\mathrm{E}, \langle\cdot,\cdot\rangle)$.

2.2 Bases orthonormées

Définitions, exemples et coordonnées

Comme son nom l'indique une famille $\mathscr{B} = (e_i)_{i\in[\![1;n]\!]}$ est une **base orthonormée** (abrégé en b.o.n) si c'est à la fois :

- une base : $$\forall\, u \in \mathrm{E}, \quad \exists!\, (\lambda_i)_{i\in[\![1;n]\!]}, \quad u = \sum_{i=1}^{n} \lambda_i\, e_i.$$
- une famille orthonormée : $$\forall\, i,\, j \in [\![1;n]\!], \qquad \langle u_i, u_j\rangle = \delta_{i,j} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j. \end{cases}$$

Exemples. Les bases canoniques de $\mathbb{R}^n$ et $\mathscr{M}_{n,1}(\mathbb{R})$ sont des b.o.n pour les produits scalaires canoniques.

Exercice 178

✧ Soient $\mathscr{B} = (e_1, e_2, e_3)$ la base canonique de $\mathbb{R}^3$ et $\theta \in \mathbb{R}$. Justifier que la base $\mathscr{C} = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$ définie par

$$\varepsilon_1 = \cos(\theta) e_1 + \sin(\theta) e_2 \quad \varepsilon_2 = \sin(\theta) e_1 - \cos(\theta) e_2 \quad \text{et} \quad \varepsilon_3 = e_3$$

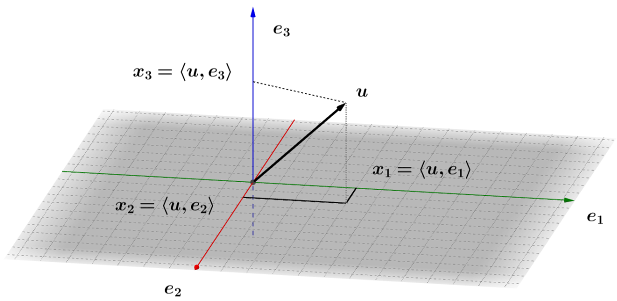
reste une base orthonormée. Donner une interprétation graphique.

Théorème 142 (coordonnées dans une b.o.n)

Soient $\left(\mathrm{E},\langle\cdot,\cdot\rangle\right)$, un espace euclidien et $\mathscr{B}=(e_1,\ldots,e_n)$, une base orthonormée.
Pour tout vecteur $u\in\mathrm{E}$,

$$u=\sum_{i=1}^{n}\langle u,e_i\rangle\, e_i.$$

Autrement dit, les coordonnées de u dans la base $\mathscr{B}$ sont les réels $\langle u,e_1\rangle,\ldots,\langle u,e_n\rangle$.



⚠ **Attention.** Il ne faut pas oublier l'hypothèse sur une base orthonormée.

Proposition 143 (norme et b.o.n)

Soit $(e_1,e_2,\ldots,e_n)$, une base orthonormée d'un espace euclidien $\left(\mathrm{E},\langle\cdot,\cdot\rangle\right)$. Pour tous $u,v\in\mathrm{E}$,

$$\langle u,v\rangle=\sum_{i=1}^{n}\langle u,e_i\rangle\langle v,e_i\rangle \quad\text{et}\quad \|u\|^2=\sum_{i=1}^{n}\langle u,e_i\rangle^2.$$

Théorèmes d'existence d'une b.o.n et procédé d'orthonormalisation de Schmidt

Soit $\mathscr{C}=\left(f_1,\ldots,f_n\right)$, une base quelconque d'un espace euclidien $\left(\mathrm{E},\langle\cdot,\cdot\rangle\right)$. Prouvons qu'il existe une famille $\mathscr{B}=(e_1,\ldots,e_n)$ telle que :

i) $\mathscr{B}$ est une base orthonormée.

ii) Pour tout $k\in[\![1,n]\!]$, $\quad \mathrm{vect}(e_1,e_2,\ldots,e_k)=\mathrm{vect}\left(f_1,f_2,\ldots,f_k\right)$.

La preuve s'effectue par récurrence et donne un procédé de construction de la base $\mathscr{B}$ à partir de la base $\mathscr{C}$.

- *Initialisation.* On pose $\quad v_1=f_1 \quad\text{et}\quad e_1=\dfrac{v_1}{\|v_1\|}.$
- Pour tout entier $k\geqslant 2$, on pose $\quad e_k=\dfrac{v_k}{\|v_k\|} \quad\text{où}\quad v_k=f_k-\displaystyle\sum_{i=1}^{k-1}\langle f_k,e_i\rangle e_i.$

Exercice 179

✦✦ Exemples

1. On considère $\mathbb{R}^3$ muni du produit scalaire usuel.
   Orthonormaliser les bases $\left((-1,1,1),(1,-1,1),(1,1,-1)\right)$ et $\left((1,1,0),(1,0,1),(0,1,1)\right)$.
2. On considère l'espace vectoriel $\mathbb{R}_2[x]$ muni du produit scalaire $\langle\mathrm{P},\mathrm{Q}\rangle=\displaystyle\int_0^1\mathrm{P}(t)\mathrm{Q}(t)\,\mathrm{d}t$.
   Orthonormaliser les bases $\left(1,x,x^2\right)$ et $\left((x-1)^2,(x-1)(x+1),(x+1)^2\right)$.

Corollaire 144 (existence d'une base orthonormée)

Tout espace euclidien admet une base orthonormée.

Corollaire 145 (base orthonormée incomplète)

Toute famille orthonormée d'une espace euclidien $\left(\mathrm{E}, \langle\cdot,\cdot\rangle\right)$ peut être complétée en une base orthonormée de E.
Autrement dit, si $\left(e_1, e_2, \dots, e_p\right)$ est une famille orthonormée de E de dimension n, il existe des vecteurs $e_{p+1}, e_{p+2}, \dots, e_n$ tels que la famille $\left(e_1, e_2, \dots, e_p, e_{p+1}, \dots, e_n\right)$ soit une base orthonormée de E.

Bases orthonormées et matrices

Proposition 146 (expression du produit scalaire avec les matrices colonnes)

Soient $\left(\mathrm{E}, \langle\cdot,\cdot\rangle\right)$, un espace euclidien, $\mathscr{B} = (e_1, \dots, e_n)$ une base orthonormée de E et $u, v \in \mathrm{E}$.

Si on note
- $\mathrm{U} = \mathrm{Mat}_{\mathscr{B}}(u)$.
- $\mathrm{V} = \mathrm{Mat}\mathscr{B}(v)$.

Alors $\langle u, v\rangle = {}^t\mathrm{UV}$ et $\|u\|^2 = {}^t\mathrm{U}\,\mathrm{U}$.

⚠ **Attention.** Rappelons que ces énoncés ne sont valables que dans le cadre d'une base orthonormée.

Définition 147 (matrice orthogonale)

Soit $\mathrm{M} \in \mathscr{M}_n(\mathbb{R})$.
On dit que M est une **matrice orthogonale** si M est inversible et $\mathrm{M}^{-1} = {}^t\mathrm{M}$.

Remarque. D'après les résultats sur les matrices, M est orthogonale si et seulement si ${}^t\mathrm{MM} = \mathrm{I}_n$ ou $\mathrm{M}{}^t\mathrm{M} = \mathrm{I}_n$.

Exemple. Pour tout $\theta \in \mathbb{R}$, la matrice $\mathrm{R}_\theta = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$ est orthogonale.

Remarques.
- Les colonnes d'une matrice M orthogonale sont non nulles et forment une base orthonormée de $\mathscr{M}_{n,1}(\mathbb{R})$ pour le produit scalaire canonique.
- **Structure de l'ensemble $\mathscr{O}_n$ des matrices orthogonales de taille** (n, n)
  - Présence d'un élément neutre : $\mathrm{I}_n \in \mathscr{O}_n$.
  - Stabilité par passage à l'inverse : $\forall\, \mathrm{P} \in \mathscr{O}_n, \quad \mathrm{P}^{-1} \in \mathscr{O}_n$.
  - Stabilité par produit : $\forall\, \mathrm{P}, \mathrm{Q} \in \mathscr{O}_n, \quad \mathrm{PQ} \in \mathscr{O}_n$.

Exercice 180

Les questions 1, 2 et 3 sont indépendantes

✧ 1. Justifier les trois points de la remarque précédente.
✧ 2. Que dire d'une matrice diagonale et orthogonale ?
✦✦ 3. (a) Montrer que si $\mathrm{P} \in \mathscr{M}_2(\mathbb{R})$ est orthogonale avec $\det(\mathrm{P}) > 0$ alors il existe $\theta \in \mathbb{R}$ tel que $\mathrm{P} = \mathrm{R}_\theta$.
(b) Comment interpréter géométriquement l'action de la matrice R_θ ?

Proposition 148 (matrice de passage orthogonale)

Soient $\left(\mathrm{E}, \langle\cdot,\cdot\rangle\right)$ un espace euclidien et $\mathscr{B} = (e_1, \dots, e_n)$ une base orthonormée de E.
Soit $\mathscr{C}$ une autre base de E. On a équivalence entre les énoncés :

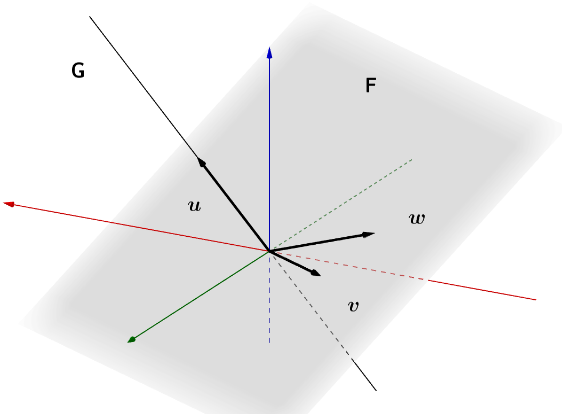
i) La base $\mathscr{C}$ est orthonormée.

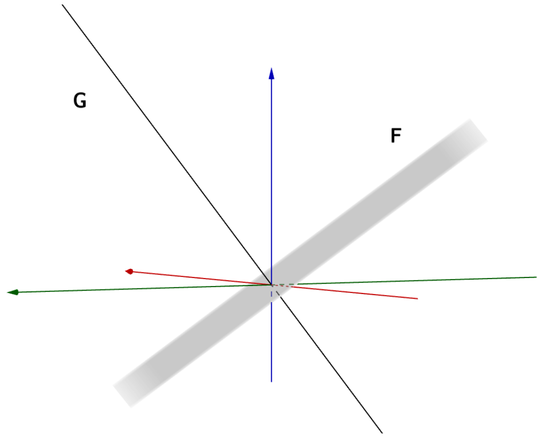
ii) La matrice de passage $\mathrm{P}_{\mathscr{B},\mathscr{C}}$ est une matrice orthogonale.

2.3 Le supplémentaire orthogonal

Tout s.e.v d'un espace euclidien $\left(E, \langle\cdot,\cdot\rangle\right)$ admet un supplémentaire orthogonal.

Exemple. Reprenons F et G de l'exercice 175, p. 95. Les vecteurs $v = (-2, 1, 0)$ et $w = (-3, 0, 1)$ forment une base du plan F et $u = (1, 2, 3)$, une base de la droite vectorielle G. On constate que F et G sont supplémentaires et orthogonaux.





Proposition 149 (propriétés)

Soit F un sous-espace vectoriel d'un espace euclidien $\left(E, \langle\cdot,\cdot\rangle\right)$.
- F et $F^{\perp}$ sont supplémentaires dans E.
- $\dim\left(F^{\perp}\right) = \dim(E) - \dim(F)$.

Remarque. Il n'y a pas en général unicité du supplémentaire, mais il y a bien unicité du supplémentaire orthogonal.

Exemples.
- Soit $E = \mathbb{R}^3$ muni du produit scalaire canonique. Posons $u = (1, 1, 1)$ et $F = \mathrm{Vect}(u)$. Déterminons une base de $F^{\perp}$.
- Soit $E = \mathbb{R}_2[x]$ muni du produit scalaire $\langle P, Q\rangle = P(-1)Q(-1) + P(0)Q(0) + P(1)Q(1)$. Donnons une base de l'orthogonal de $F = \mathbb{R}_1[x]$.
- Dans le cas de $\mathscr{M}_n(\mathbb{R})$ muni du produit scalaire canonique, on montre que $\mathscr{S}_n(\mathbb{R})^{\perp} = \mathscr{A}_n$.

Exercice 181

✦ *Les questions sont indépendantes*

Soient F et G deux sous-espaces vectoriels d'un espace euclidien $\left(E, \langle\cdot,\cdot\rangle\right)$.

1. Montrer que $\left(F^{\perp}\right)^{\perp} = F$.
2. Montrer que $(F+G)^{\perp} = F^{\perp} \cap G^{\perp}$ puis $F^{\perp} + G^{\perp} = (F \cap G)^{\perp}$.

Exercice 182

✦ Soient $n \in \mathbb{N}^*$ et $E = \mathbb{R}_{2n}[x]$. Considérons F et G les s.e.v de E constitué des polynômes pairs (resp. impairs). Montrer que F et G sont des sous-espaces supplémentaires orthogonaux pour le produit scalaire

$$\forall P, Q \in E, \qquad \langle P, Q\rangle = \int_{-1}^{1} P(t)Q(t)\,dt.$$

Exercices - TD

Exercice 183. ✧ Montrer que pour tous réels $x_1, x_2, \dots, x_n$, on a $\sum\limits_{i=1}^{n} x_i \leqslant \sqrt{n}\sqrt{\sum\limits_{i=1}^{n} x_i^2}$. Préciser le cas d'égalité.

Exercice 184. ✦

1. Rappeler l'inégalité de Cauchy-Schwarz puis vérifier que

$$\forall \mathrm{A} \in \mathscr{M}_n(\mathbb{R}), \quad \mathrm{Tr}\left(\mathrm{A}^2\right) \leqslant \mathrm{Tr}\left({}^t\mathrm{A}\mathrm{A}\right).$$

2. Montrer également que $\mathrm{Tr}\left(\mathrm{A}^2\right) = \mathrm{Tr}\left({}^t\mathrm{A}\mathrm{A}\right)$ si et seulement si A est une matrice symétrique.

Exercice 185. ✦✦ **Condition nécessaire et suffisante d'orthogonalité**
Soient u et v deux vecteurs non nuls d'un espace euclidien $\left(\mathrm{E}, \langle\cdot,\cdot\rangle\right)$. Établir l'équivalence entre les énoncés suivants :

i) Les vecteurs u et v sont orthogonaux.

ii) Pour tout $\lambda \in \mathbb{R}$, $\|\lambda u + v\| \geqslant \|v\|$.

Exercice 186. ✦✦✦ **Frames**
Soient $\left(\mathrm{E}, \langle\cdot,\cdot\rangle\right)$ un espace euclidien et $\mathscr{B} = (e_1, e_2, \dots, e_n)$ une famille de vecteurs de norme 1 de E.

1. On suppose que

$$\forall\, u \in \mathrm{E}, \qquad \|u\|^2 = \sum_{i=1}^{n} \langle u, e_i \rangle^2.$$

Justifier que $\mathscr{B}$ est une base de E.

2. On suppose maintenant qu'il existe deux réels strictement positifs A et B tels que

$$\forall\, u \in \mathrm{E}, \qquad \mathrm{A}\,\|u\|^2 \leqslant \sum_{i=1}^{n} \langle u, e_i \rangle^2 \leqslant \mathrm{B}\,\|u\|^2 \qquad (\star)$$

   (a) Justifier que $\mathscr{B}$ reste une famille génératrice.

   (b) En considérant sur $\mathrm{E} = \mathbb{R}^2$ la famille (e_1, e_2, e_3) définie par

$$e_1 = (1,0), \quad e_2 = \frac{1}{2}(-1, \sqrt{3}) \quad \text{et} \quad e_3 = \frac{1}{2}(-1, -\sqrt{3}),$$

   montrer qu'une famille peut vérifier $(\star)$ sans être libre.

La théorie des frames permet d'étudier la stabilité et la redondance des représentations linéaires discrètes d'un signal. On la retrouve notamment dans la théorie des ondelettes, particulière utile en analyse d'images.

Exercice 187. ✦✦✦ On considère un espace euclidien E ainsi qu'une famille $\mathscr{F} = \left(e_1, \cdots, e_p\right)$ de vecteurs de E. Montrer que si la famille $\mathscr{F}$ est génératrice de E, alors l'endomorphisme

$$\varphi : \left\{ \begin{array}{ccc} \mathrm{E} & \longrightarrow & \mathrm{E} \\ u & \longmapsto & \sum\limits_{i=1}^{p} \langle u, e_i \rangle e_i \end{array} \right. \qquad \text{est bijectif.}$$

Indication. On pourra, pour un vecteur u bien choisi, considérer $\langle u, \varphi(u) \rangle$.

Exercice 188. ✦ **Probabilité de collision**
Soient X et Y deux variables aléatoires finies indépendantes et de même loi. Notons

$$\mathrm{X}(\Omega) = \mathrm{Y}(\Omega) = \{x_1, x_2, \cdots, x_n\} \quad \text{et} \quad \forall\, i \in [\![1,n]\!] \quad \mathbf{P}(\mathrm{X} = x_i) = p_i.$$

1. Démontrer que $\mathbf{P}(\mathrm{X} = \mathrm{Y}) = \sum\limits_{k=1}^{n} p_k^2$.

2. En déduire que $\mathbf{P}(\mathrm{X} = \mathrm{Y}) \geqslant \frac{1}{n}$. Préciser le cas d'égalité.

Endomorphisme conservant la norme, les angles ...

Exercice 189. ✦✦
Soient E un espace euclidien et f un endomorphisme de E.

1. On dit que f est une isométrie si pour tout $x \in \mathrm{E}$, $\|f(x)\| = \|x\|$.
   (a) Montrer que toute isométrie est un isomorphisme de E.
   (b) Prouver l'équivalence entre les énoncés :
      i) f est une isométrie ;
      ii) Pour tous $x, y \in \mathrm{E}$, $\langle f(x), f(y)\rangle = \langle x, y\rangle$.
2. On suppose que f est une isométrie vérifiant $f^2 = -\mathrm{id}_\mathrm{E}$. Montrer que pour tout $x \in \mathrm{E}$, $f(x)$ est orthogonal à x.
3. On suppose que f est une isométrie et pour tout $x \in \mathrm{E}$, $f(x)$ est orthogonal à x.
   En calculant $\langle f(x)+x, f^2(x)+f(x)\rangle$, déduire $\langle x, f^2(x)\rangle = -\|x\|^2$.
   En développant $\|f^2(x)+x\|^2$, montrer $f^2 = -\mathrm{id}_\mathrm{E}$.
4. Maintenant soit f un endomorphisme vérifiant $f^2 = -\mathrm{id}_\mathrm{E}$ et pour tout $x \in \mathrm{E}$, $f(x)$ est orthogonal à x. Montrer que f est une isométrie.

Exercice 190. ✦✦ **Endomorphisme qui conserve l'orthogonalité** *Source : HEC 2007.*
Soit E un espace euclidien de dimension n. On note $\langle\cdot,\cdot\rangle$ le produit scalaire et $\|\cdot\|$ la norme associée. Soit f un endomorphisme de E qui vérifie la propriété suivante :
$$\forall (u,v) \in \mathrm{E}^2, \qquad \langle u, v\rangle = 0 \quad \Rightarrow \quad \langle \varphi(u), \varphi(v)\rangle = 0.$$

1. Vérifier que si u et v sont deux vecteurs de même norme, alors $(u-v)$ et $(u+v)$ sont orthogonaux.
2. Démontrer qu'il existe $k \in \mathbb{R}^+$ tel que pour tout $u \in \mathrm{E}$, $\|\varphi(u)\| = k\|u\|$.

Produits scalaires sur $\mathbb{R}_n[x]$ et familles de polynômes orthogonaux

Exercice 191. ✦ À quelles conditions sur les réels α, β et γ l'application définie sur $\mathbb{R}_2[x]^2$ par
$$\varphi(\mathrm{A},\mathrm{B}) = \alpha\mathrm{P}(-1)\mathrm{Q}(-1) + \beta\mathrm{P}(0)\mathrm{Q}(0) + \gamma\mathrm{P}(1)\mathrm{Q}(-1)$$
est un produit scalaire sur $\mathbb{R}_2[x]$?

Exercice 192. ✦✦ Soient $\alpha_0, \alpha_1, \ldots, \alpha_n$ des réels 2 à 2 distincts. Pour tous les polynômes P et Q appartenant à $\mathrm{E} = \mathbb{R}_n[x]$, on pose
$$\langle \mathrm{P}, \mathrm{Q}\rangle = \sum_{i=0}^{n} \mathrm{P}^{(i)}(\alpha_i)\cdot \mathrm{Q}^{(i)}(\alpha_i).$$
Montrer que $(\mathrm{E}, \langle\cdot,\cdot\rangle)$ est un espace euclidien.

Exercice 193. ✦✦ **Un classique : les polynômes de Tchebychev**
On définit $(\mathrm{T}_n)_{n\in\mathbb{N}}$, une suite de polynômes par la récurrence
$$\mathrm{T}_0 = 1, \quad \mathrm{T}_1(x) = x \quad \text{et} \quad \forall n \in \mathbb{N}, \quad \mathrm{T}_{n+2}(x) = 2x\mathrm{T}_{n+1}(x) - \mathrm{T}_n(x).$$

1. (a) Soit $n \in \mathbb{N}$, justifier que pour tout $\theta \in \mathbb{R}$,
   $$\cos\big((n+2)\theta\big) + \cos(n\theta) = 2\cos\big((n+1)\theta\big)\cos(\theta).$$
   (b) Montrer que pour réel θ, tout entier naturel n,
   $$\cos(n\theta) = \mathrm{T}_n\big(\cos(\theta)\big) \qquad (\bullet)$$
   (c) Vérifier que T_n est l'unique polynôme vérifiant les relations $(\bullet)$. Préciser le degré de T_n.
2. Soit $n \in \mathbb{N}^*$. Soit $\langle\cdot,\cdot\rangle$ l'application définie sur $\mathbb{R}_n[x]^2$ par :
   $$\langle \mathrm{P}, \mathrm{Q}\rangle = \int_{-1}^{1} \frac{\mathrm{P}(t)\mathrm{Q}(t)}{\sqrt{1-t^2}}\,\mathrm{d}t.$$
   (a) Montrer que $\langle\cdot,\cdot\rangle$ est un produit scalaire bien défini sur $\mathbb{R}_n[x]$.
   (b) Vérifier que la famille $(\mathrm{T}_k)_{0\leqslant k\leqslant n}$ est une base orthogonale pour ce produit scalaire.
   (c) Déterminer $\|\mathrm{T}_k\|$ pour tout entier $k \in [\![0,n]\!]$.

Exercice 194. ✦✦✦ **Polynômes de Legendre**

On munit $\mathbb{R}[x]$ du produit scalaire : $$\langle \mathrm{P}, \mathrm{Q}\rangle = \int_{-1}^{1} \mathrm{P}(t)\mathrm{Q}(t)\,\mathrm{d}t$$

et, pour tout $n \in \mathbb{N}$, on définit les polynômes $\mathrm{P}_n(x) = (x^2-1)^n$ et $\mathrm{L}_n = {\mathrm{P}_n}^{(n)}$.

1. Préciser L_n pour $n \in \{0,1,2,3\}$.
2. Déterminer le degré, le coefficient dominant et le coefficient constant de L_n.
3. Calculer $L_n(1)$.
4. Prouver que, pour tout $n \geqslant 1, L_n$ possède n racines simples dans $]-1,1[$.
5. (a) Justifier que si f et $g : [a,b] \to \mathbb{R}$ sont de classe $\mathscr{C}^n$ alors
$$\int_a^b f^{(n)}(t)g(t)\,\mathrm{d}t = \left[\sum_{k=0}^{n-1}(-1)^k f^{(n-1-k)}(t)g^{(k)}(t)\right]_a^b + (-1)^n\int_a^b f(t)g^{(n)}(t)\,\mathrm{d}t.$$
   (b) Posons pour tout $p \in \mathbb{N}$, $E_p(x) = x^p$. Calculer $\langle L_n, E_p\rangle$ pour $p \in [\![0; n-1]\!]$.
   (c) En déduire que la famille $(L_n)_{n\geqslant 0}$ est orthogonale pour le produit scalaire $\langle\cdot,\cdot\rangle$.
6. Pour tout $p \in \mathbb{N}$, on pose $W_p = \displaystyle\int_0^{\pi/2}\cos^{2p+1}(u)\,\mathrm{d}u$.
   (a) Donner une relation simple entre W_{p+1} et W_p. En déduire W_p pour tout $p \in \mathbb{N}$.
   (b) Vérifier que
$$\|L_n\|^2 = 2(2n)!W_n.$$
   (c) En déduire $\|L_n\|$ pour tout $n \in \mathbb{N}$.
7. En décomposant le polynôme $xL_n(x)$ dans la base $(L_k)_{0\leqslant k\leqslant n+1}$ de $\mathbb{R}_{n+1}[x]$, justifier que
$$\forall n \in \mathbb{N}^*, \qquad L_{n+1}(x) - 2(2n+1)xL_n(x) + 4n^2L_{n-1}(x) = 0.$$

Algèbre bilinéaire et réduction

Exercice 195. ✦✦ Soit $(E, \langle\cdot,\cdot\rangle)$ un espace euclidien avec $\dim E \geqslant 2$. Pour tout vecteur u non nul et pour tous les réels $(\lambda,\mu) \neq (0,0)$, on définit l'endomorphisme f de E par
$$f(x) = \lambda x + \mu\langle x, u\rangle u.$$

1. Déterminer les valeurs propres et les vecteurs propres de f. Est-ce que f est diagonalisable ?
2. À quelles conditions sur λ et μ, l'application f est une isométrie ?
   f est une isométrie si pour tout $x \in E$, $\|f(x)\| = \|x\|$.

Exercice 196. ✦✦✦ **Matrices de Gram et valeurs propres**
Soient E un espace euclidien de dimension n et $(u_1, u_2, \cdots, u_n)$ une famille de vecteurs de E. On pose
$$G = \left(\left\langle u_i, u_j\right\rangle\right)_{1\leqslant i,j\leqslant n} \quad \text{et} \quad M = \mathrm{Mat}_{\mathscr{B}}(u_1, u_2, \cdots, u_n),$$
où $\mathscr{B}$ est une base orthonormée de E.

1. Vérifier que $G = {}^t MM$.
2. En déduire que $\ker(G) = \ker(M)$, puis $\mathrm{rg}(G) = \mathrm{rg}(u_1, u_2, \cdots, u_n)$.
   Indication. Considérer ${}^t XGX$ pour $X \in \mathscr{M}_{n,1}(\mathbb{R})$.
3. Vérifier que G est inversible si et seulement si la famille $(u_1, u_2, \cdots, u_n)$ est une base de E.
4. Justifier que les valeurs propres de G sont positives ou nulles.
5. Justifier que les valeurs propres sont majorées par $\sum\limits_{i=1}^n \|u_i\|^2$.

Compléments en dimension infinie

Exercice 197. ✦✦✦ **Orthogonal d'un hyperplan**
On munit $E = \mathscr{C}^0([0,1],\mathbb{R})$ du produit scalaire défini par : $\forall (f,g) \in E^2, \quad \langle f, g\rangle = \displaystyle\int_0^1 f(t)g(t)\,\mathrm{d}t$
Considérons le sous-ensemble A de E, constitué des applications qui s'annulent en 0 et le sous-ensemble B de E, constitué des applications dont l'intégrale sur $[0,1]$ est nulle. A et B sont les noyaux de formes linéaires non nulle
$$f \in E \mapsto f(0) \in \mathbb{R} \quad \text{et} \quad f \in E \mapsto \int_0^1 f(t)\,\mathrm{d}t \in \mathbb{R}.$$
Ce sont donc des hyperplans de E. Déterminer $A^\perp$ et $B^\perp$.

Exercice 198. ✦✦ **Un sujet de concours** *D'après Emlyon S 2014*
On note E l'espace vectoriel des fonctions définies et continues sur $\mathbb{R}^+$ et à valeurs dans $\mathbb{R}$ et E_2 l'ensemble des fonctions f de E telles que l'intégrale $\int_0^{+\infty}(f(x))^2\,\mathrm{d}x$ converge. Pour toute fonction f de E, on note toujours $\Phi(f)$ la fonction définie dans cette partie sur $\mathbb{R}^+$ par :
$$\forall x \in \mathbb{R}^+, \quad \Phi(f)(x) = \begin{cases} \dfrac{1}{x^2}\displaystyle\int_0^x tf(t)\mathrm{d}t & \text{si } x > 0 \\ \dfrac{f(0)}{2} & \text{si } x = 0. \end{cases}$$

1. (a) Justifier : $\forall (x,y)\in\mathbb{R}^2, |xy| \leqslant \frac{1}{2}\left(x^2+y^2\right)$.

   (b) En déduire que, pour toutes fonctions f et g de E_2, l'intégrale $\displaystyle\int_0^{+\infty} f(x)g(x)\,\mathrm{d}x$ est absolument convergente.

   (c) Montrer alors que E_2 est un sous-espace vectoriel de E.

2. On considère l'application $\langle\cdot,\cdot\rangle$ de $E_2\times E_2$ dans $\mathbb{R}$ définie par :

$$\forall (f,g)\in E_2\times E_2,\quad \langle f,g\rangle = \int_0^{+\infty} f(x)g(x)\mathrm{d}x.$$

   Montrer que $\langle\cdot,\cdot\rangle$ est un produit scalaire sur E_2. On munit E_2 de ce produit scalaire et de la norme associée notée $\|\cdot\|$.

3. Soit f une fonction de E_2.
   On note pour tout x de $\mathbb{R}^+$: $\quad h(x) = \displaystyle\int_0^x t f(t)\mathrm{d}t$.

   (a) Calculer les limites de $x\mapsto \frac{(h(x))^2}{x^4}$ et de $x\mapsto \frac{(h(x))^2}{x^3}$ en 0.

   (b) Montrer, à l'aide d'une intégration par parties :

$$\forall X\in\mathbb{R}_*^+,\quad \int_0^X \frac{\left(h(x)\right)^2}{x^4}\,\mathrm{d}x = -\frac{1}{3}\frac{\left(h(X)\right)^2}{X^3}+\frac{2}{3}\int_0^X f(x)\Phi(f)(x)\mathrm{d}x \qquad (\bullet)$$

   (c) Soit $X\in\mathbb{R}_*^+$. En étudiant le signe de la fonction polynomiale $\lambda\in\mathbb{R}\mapsto \displaystyle\int_0^\lambda \left(\lambda f(x)+\Phi(f)(x)\right)^2\mathrm{d}x$, montrer l'inégalité de Cauchy-Schwarz suivante :

$$\int_0^X f(x)\Phi(f)(x)\mathrm{d}x \leqslant \left(\int_0^X \left(f(x)\right)^2\mathrm{d}x\right)^{1/2}\left(\int_0^X \left(\Phi(f)(x)\right)^2\mathrm{d}x\right)^{1/2}.$$

   (d) En déduire :

$$\forall X\in\mathbb{R}_*^+,\qquad \left(\int_0^X (\Phi(f)(x))^2\,\mathrm{d}x\right)^{1/2} \leqslant \frac{2}{3}\left(\int_0^X \left(f(x)\right)^2\mathrm{d}x\right)^{1/2}.$$

   (e) Montrer alors que la fonction $\Phi(f)$ appartient à E_2 et que l'on a : $\|\Phi(f)\| \leqslant \dfrac{2}{3}\|f\|$.

   (f) En utilisant la relation $(\bullet)$, justifier que la limite de $X\mapsto X\left(\Phi(f)(X)\right)^2$ en $+\infty$ est finie, puis en raisonnant par l'absurde, montrer que cette limite est nulle.

   (g) En déduire :

$$\|\Phi(f)\|^2 = \frac{2}{3}\left\langle \Phi(f), f\right\rangle.$$

CHAPITRE 10

Introduction au calcul différentiel

It's like asking why Beethoven's Ninth Symphony is beautiful. If you don't see why, someone can't tell you. I know numbers are beautiful. If they aren't beautiful, nothing is.

PAUL ERDÖS

1 Rappels : dérivation des fonctions d'une variable réelle

1.1 Définition du nombre dérivé et interprétation géométrique

Définition 150 (nombre dérivé, fonction dérivée)

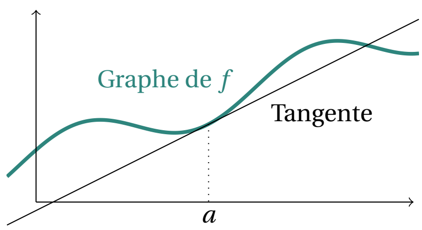
Soient I, un intervalle, $f : \mathrm{I} \to \mathbb{R}$ et $a \in \mathrm{I}$.
- f est **dérivable en** a si le quotient $\frac{f(x)-f(a)}{x-a}$ admet une limite *finie* en a. Si cette dernière existe, elle est unique et notée $f'(a)$.
- f est **dérivable sur** I si elle est dérivable pour tout réel de I. Ainsi, on définit la fonction dérivée par

$$f' : \left\{ \begin{array}{ccc} \mathrm{I} & \to & \mathbb{R} \\ a & \mapsto & f'(a). \end{array} \right.$$

- Graphiquement, f est dérivable en a s'il existe une tangente à la courbe. L'équation de la tangente est alors

$$y = f(a) + f'(a)(x-a).$$

- Le terme $\frac{f(x)-f(a)}{x-a}$ est le taux d'accroissement de f entre a et x.



1.2 Les théorèmes

Théorème 151 (développement limité à l'ordre 1)

Soient $f : \mathrm{I} \to \mathbb{R}$ et $a \in \mathrm{I}$. Les énoncés suivants sont équivalents.

i) La fonction f est dérivable en a et $\lambda = f'(a)$.

ii) f admet un développement limité à l'ordre 1 en a.
C'est-à-dire, il existe $\lambda \in \mathbb{R}$ tel que pour tout $x \in \mathrm{I}$, $\quad f(x) = f(a) + \lambda(x-a) + o_a(x-a)$.

Remarque. On peut réécrire le développement limité sous la forme : il existe une fonction ε définie sur un voisinage $\mathcal{V}$ de 0 telle que pour tout $h \in \mathcal{V}$,

$$f(a+h) = f(a) + f'(a)h + h\varepsilon(h) \quad \text{avec} \quad \varepsilon(h) \underset{h\to 0}{\longrightarrow} 0.$$

Théorème 152 (extremum - condition nécessaire)

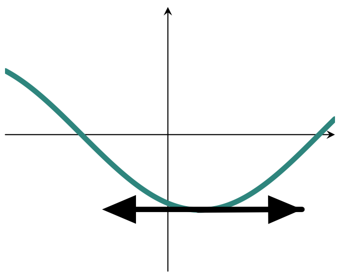
Soit $f : \mathrm{I} \to \mathbb{R}$ où I est un intervalle.

Si
- f admet un extremum local en a.
- f est dérivable en a.
- a n'est pas un des bords de I,

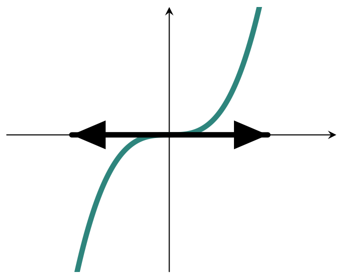
alors $f'(a) = 0$.

Vocabulaire. On dit que $\mathrm{A} = \big(a, f(a)\big)$ est un point critique de f lorsque $f'(a) = 0$.

Remarque. La réciproque est fausse : tout point critique ne donne pas un extremum. La fonction $x \in \mathbb{R} \mapsto x^3$ en 0 fournit un contre-exemple.



Point critique et minimum



Point critique sans extremum

2 Dérivées partielles et gradient

Soient $f : \mathbb{R}^n \to \mathbb{R}$ et $a = (a_1, a_2, \dots, a_n) \in \mathbb{R}^n$. Pour tout $k \in [\![1;n]\!]$, on considère $f_{k,a}$ définie par

$$f_{k,a} : \left\{ \begin{array}{ccl} \mathbb{R} & \to & \mathbb{R} \\ t & \mapsto & f(a_1, a_2, \dots, a_{k-1}, t, a_{k+1}, \dots, a_n). \end{array} \right.$$

L'application $f_{k,a}$ est la k-ième application partielle de f en a. C'est une fonction d'une variable réelle, on peut donc utiliser les définitions et résultats de première année.

Définition 153 (dérivée partielle)

Soient $f : \mathbb{R}^n \to \mathbb{R}$ et $a \in \mathbb{R}$.
On dit que la fonction f admet une **dérivée partielle d'ordre** k **en** a si l'application partielle $f_{k,a}$ est dérivable en a_k. On note alors $\partial_k f(a)$ le nombre dérivée $f'_{k,a}(a)$. Autrement dit, la limite suivante existe et

$$\partial_k f(a) = \lim_{t \to a_k} \frac{f(a_1, a_2, \dots, a_{k-1}, t, a_{k+1}, \dots, a_n) - f(a)}{t - a_k}.$$

Remarque. Si f admet pour tout $a \in \mathbb{R}^n$, une dérivée partielle en a, on peut définir la i-ème dérivée partielle par

$$\partial_i f : a \in \mathbb{R}^n \mapsto \partial_i f(a).$$

C'est encore une fonction de n variables réelles.

Exercice 199

✧ Préciser les dérivées partielles des fonctions définies sur $\mathbb{R}^2$ ou $\mathbb{R}^3$ suivantes.

$$f(x,y) = x^2 \exp(xy), \quad g(x,y) = \ln\left(1 + \sqrt{x^2+y^2}\right), \quad h(x,y) = \sin(x)^2 + \cos(y)^2,$$

$$i(x,y,z) = x^2y^2z^2, \quad j(x,y,z) = \arctan(xyz).$$

On pourra utiliser les symétries pour simplifier certains calculs.

Définition 154 (gradient)

Soient $a \in \mathbb{R}^n$ et $f : \mathbb{R}^n \to \mathbb{R}$ tels que pour tout $i \in [\![1;n]\!]$, f admet une i-ème dérivée partielle en a.
On définit le **gradient** de f en a, noté $\nabla f(a)$, par le vecteur de $\mathbb{R}^n$

$$\nabla f(a) = \big(\partial_1 f(a), \dots, \partial_n f(a)\big).$$

3 Fonctions de classe $\mathscr{C}^1$

3.1 Définitions, exemples et règles de calculs

Définition 155 (fonctions de classe $\mathscr{C}^1$)

Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction.
On dit que f est de **classe** $\mathscr{C}^1$ sur $\mathbb{R}^n$ si pour toutes les dérivées partielles existent et sont continues sur $\mathbb{R}^n$.

Exemple. Les dérivées partielles d'une fonction polynomiale sont encore des fonctions polynomiales, elles sont donc continues sur $\mathbb{R}^n$. Ainsi, les fonctions polynomiales sont de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$.

⚠ **Attention.** La condition $\mathscr{C}^1$ est une condition plus restrictive que la simple existence des dérivées partielles.

Proposition 156 (linéarité, produit et quotient de fonctions $\mathscr{C}^1$)

Soient f, g deux fonctions de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$. Alors

- Pour tout réel λ, λf est de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$.
- La somme $f + g$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$.
- Le produit $f \cdot g$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$.
- Si de plus, la fonction g ne s'annule pas sur $\mathbb{R}^n$, alors f/g est bien définie et de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$.

Exemple. Toute fonction rationnelle définie sur $\mathbb{R}^n$ est de classe $\mathscr{C}^1$. Pour rappel, une fonction est dite rationnelle si elle peut s'écrire comme le quotient de deux fonctions polynomiales.

Proposition 157 (composition de fonctions $\mathscr{C}^1$)

Soient $f : \mathbb{R}^n \to \mathbb{R}$, une partie I de $\mathbb{R}$ et $\varphi : \mathrm{I} \to \mathbb{R}$.

Si
→ Pour tout $x \in \mathbb{R}^n$, $f(x) \in \mathrm{I}$.
→ La fonction de plusieurs variables f est de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$.
→ La fonction d'une variable φ est de classe $\mathscr{C}^1$ sur I.

Alors la composée $\varphi \circ f$ est bien définie et de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$.

Exemple. Justifions que $g : (x,y) \in \mathbb{R}^2 \mapsto x \ln\big(1 + x^2 + y^2\big)$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}$.
→ Pour tout $(x,y) \in \mathbb{R}^2$, $1 + x^2 + y^2 \in \mathbb{R}_*^+$.
→ $(x,y) \in \mathbb{R}^2 \mapsto 1 + x^2 + y^2$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$ en tant que fonction polynomiale.
→ La fonction logarithme $\ln : \mathbb{R}_*^+ \mapsto \mathbb{R}$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}_*^+$.
Par composition, $(x,y) \in \mathbb{R}^2 \mapsto \ln\big(1 + x^2 + y^2\big)$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$.
De plus, $(x,y) \in \mathbb{R}^2 \mapsto x$ est aussi de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$. Par produit, la fonction g est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$.

3.2 Développement limité d'ordre 1

Condition d'existence et exemples

Rappelons la définition de la norme euclidienne et du produit scalaire canonique sur $\mathbb{R}^n$.

$$\forall\, x = (x_i) \in \mathbb{R}^n, \quad \forall\, y = (y_i) \in \mathbb{R}^n, \qquad \|x\| = \sqrt{\sum_{i=1}^{n} x_i{}^2} \quad \text{et} \quad \langle x, y\rangle = \sum_{i=1}^{n} x_i y_i.$$

Dans la suite, pour une fonction $\varepsilon : \mathbb{R}^n \to \mathbb{R}$, la notation $\varepsilon(h) \underset{h\to 0}{\longrightarrow} 0$ signifie que ε est continue en $0_{\mathbb{R}^n}$ avec $\varepsilon(0_{\mathbb{R}^n}) = 0$. C'est-à-dire

$$\forall \eta \in \mathbb{R}_*^+, \quad \exists \alpha \in \mathbb{R}_*^+, \quad \forall\, h \in \mathbb{R}^n \qquad \Big(\|h\| \leqslant \alpha \quad \Rightarrow \quad \|\varepsilon(h)\| \leqslant \eta \Big).$$

Théorème 158 (développement limité d'ordre 1)

Soit $f : \mathbb{R}^n \to \mathbb{R}$ de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$.
Alors pour tout $a \in \mathbb{R}^n$, f admet en a un **unique développement limité à l'ordre 1**.
C'est-à-dire, il existe une fonction $\varepsilon : \mathbb{R}^n \to \mathbb{R}$ telle que, pour $h \in \mathbb{R}^n$,

$$f(a+h) = f(a) + \big\langle \nabla f(a), h \big\rangle + \|h\|\varepsilon(h) \quad \text{avec} \quad \varepsilon(h) \underset{h\to 0}{\longrightarrow} 0.$$

Exercice 200

✦
Donner le développement limité de $f : (x, y) \in \mathbb{R}^2 \mapsto \ln(1 + x^2 + y^2)$ en $(2, 1)$ et $g : x \in \mathbb{R}^n \mapsto \|x\|^2$ en tout point $a \in \mathbb{R}^n$.

Remarques.

- On peut reformuler l'équation précédente :

→ Avec les dérivées partielles :

$$f(a+h) = f(a) + \sum_{k=1}^{n} \partial_k f(a) \times h_k + \|h\|\varepsilon(h) \quad \text{où} \quad \left\{ \begin{array}{rcl} h & = & (h_1, h_2, \dots, h_n) \\ \nabla f(a) & = & \big(\partial_1 f(a), \dots, \partial_n f(a)\big). \end{array} \right.$$

→ Avec le changement de variable $x = a + h$ $\qquad f(x) = f(a) + \sum_{k=1}^{n} (x_k - a_k)\,\partial_k f(a) + \|x - a\|\varepsilon(x - a).$

- L'existence de dérivées partielles ne suffit pas à assurer l'existence d'un développement limité. On peut montrer qu'elle n'assure même pas la continuité de l'application.

Exercice 201

✦✦ **Unicité du développement limité**
Soit $f : \mathbb{R}^n \to \mathbb{R}$ de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$ et $a \in \mathbb{R}^n$. On suppose qu'il existe une forme linéaire $\mathrm{L} : \mathbb{R}^n \to \mathbb{R}$ telle que pour tout $h \in \mathbb{R}^n$

$$f(a+h) = f(a) + \mathrm{L}(h) + \|h\|\varepsilon(h) \quad \text{avec} \quad \varepsilon(h) \underset{h\to 0}{\longrightarrow} 0.$$

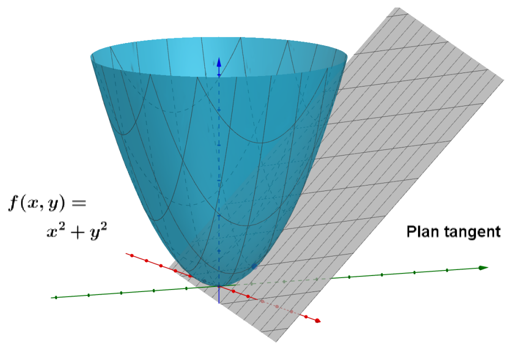
Montrer que pour tout $h \in \mathbb{R}^n$, $\mathrm{L}(h) = \big\langle \nabla f(a), h \big\rangle$.

Interprétation géométrique : le plan tangent

Exemple. Illustrons la situation avec la fonction f définie sur $\mathbb{R}^2$ par $f(x, y) = x^2 + y^2$ et $a = (0.1, 0.5)$. L'équation du plan affine est

$$z = f(a) + \partial_1 f(a)(x - 0.1) + \partial_2 f(a)(y - 0.5).$$

À l'instar de la droite tangente pour une fonction d'une variable réelle, on obtient ici un plan tangent à la surface.

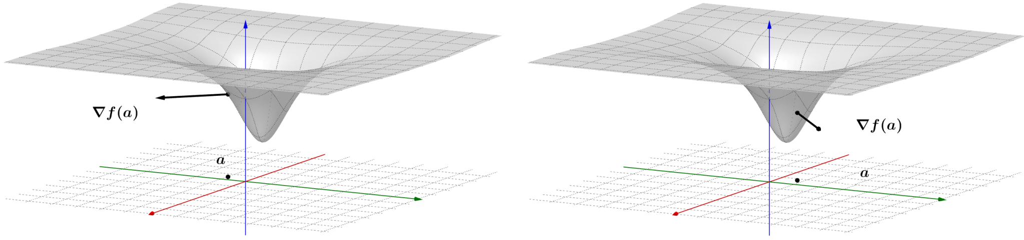


Le gradient donne la direction de plus grande pente

Soit $f : \mathbb{R}^n \to \mathbb{R}$ de classe $\mathscr{C}^1$. Supposons de plus que $\nabla f(a) \neq 0_{\mathbb{R}^n}$. D'après le développement limité d'ordre 1, $f(a+h) - f(a) \simeq \langle \nabla f(a), h \rangle$ lorsque h est « proche » de 0. Or, l'inégalité de Cauchy-Schwarz donne

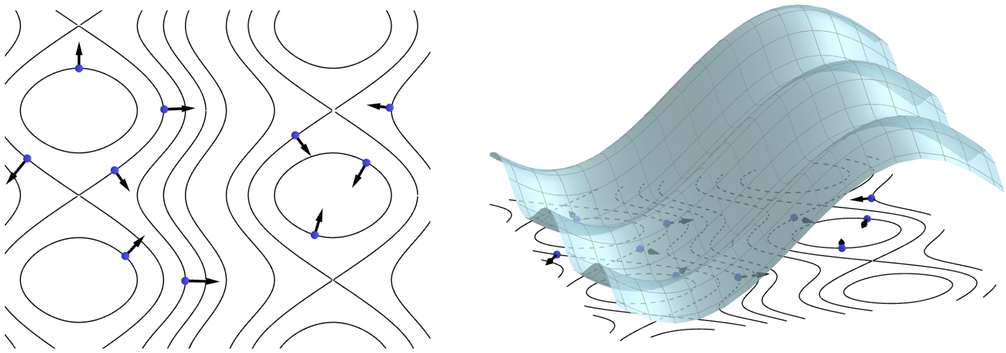
$$\left|\langle \nabla f(a), h\rangle\right| \leqslant \|\nabla f(a)\| \cdot \|h\|$$

avec égalité si et seulement si les vecteurs sont colinéaires. Autrement dit, la variation $|f(a+h)-f(a)|$ est « localement » maximale lorsque le vecteur h est colinéaire au gradient $\nabla f(a)$. On dit que $\nabla f(a)$ donne la direction de plus grande pente (et dirigé dans le sens des pentes croissantes).



Le gradient est orthogonal aux lignes de niveau

Commençons par un exemple. On a tracé ci-dessous quelques lignes de niveaux de la fonction $f : (x, y) \in \mathbb{R}^2 \mapsto 4\sin(x/3) + \cos(y) + 5$ ainsi que quelques gradients $\nabla f(a)$ pour différentes valeurs de a.



On constate que le gradient $\nabla f(a)$ est systématiquement orthogonal à la tangente à la ligne de niveau en a. Prouvons le cas général. Soient $f : \mathbb{R}^2 \to \mathbb{R}$ de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$ et γ_1, $\gamma_2 : \mathbb{R} \to \mathbb{R}$ de classe $\mathscr{C}^1$ sur $\mathbb{R}$. On pose

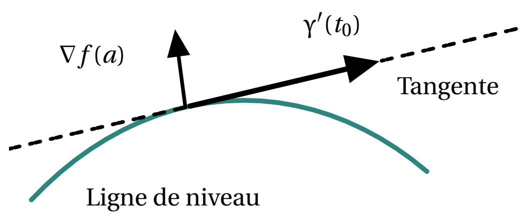
$$\gamma : \left\{ \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R}^2 \\ t & \mapsto & \big(\gamma_1(t), \gamma_2(t)\big). \end{array} \right.$$

Comme γ_1 et γ_2 sont de classe $\mathscr{C}^1$, on dit que γ est de classe $\mathscr{C}^1$, et on pose $\gamma' = \left(\gamma_1', \gamma_2'\right)$. Soient $t_0 \in \mathbb{R}$ et $a = \gamma(t_0) \in \mathbb{R}^2$.

Exercice 202

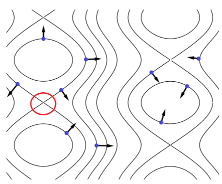
1. Montrer que $\gamma(t) - \gamma(t_0) = \gamma'(t_0)(t - t_0) + o_{t_0}(t - t_0)$.
2. Démontrer que pour $t \in \mathbb{R}$, $f(\gamma(t)) = f(a) + \left\langle \nabla(f)(a), \gamma'(t_0)\right\rangle (t - t_0) + o_{t_0}(t - t_0)$ et en déduire que $(f \circ \gamma)'(t_0) = \left\langle \nabla f(a), \gamma'(t_0)\right\rangle$.
3. On pose $\mathrm{K} = f(a)$ de sorte que a appartienne à la ligne de niveau $\mathscr{L}_{\mathrm{K}} = \left\{x \in \mathbb{R}^2 \mid f(x) = \mathrm{K}\right\}$. On suppose de plus que γ est à valeurs dans $\mathscr{L}_{\mathrm{K}}$, c'est à dire, pour tout $t \in \mathbb{R}$, $\gamma(t) \in \mathscr{L}_{\mathrm{K}}$. Montrer que $\nabla f(a) \perp \gamma'(t_0)$.

Comme γ est à valeurs dans $\mathscr{L}_{\mathrm{K}}$, le vecteur $\gamma'(t_0)$ est un vecteur tangent à la ligne de niveau $\mathscr{L}_{\mathrm{K}}$. On montre donc ici que le gradient de f en a est orthogonal à tout vecteur tangent (en a) à $\mathscr{L}_{\mathrm{K}}$.



Exercice 203

✦ Que dire du gradient au niveau d'un croisement ?

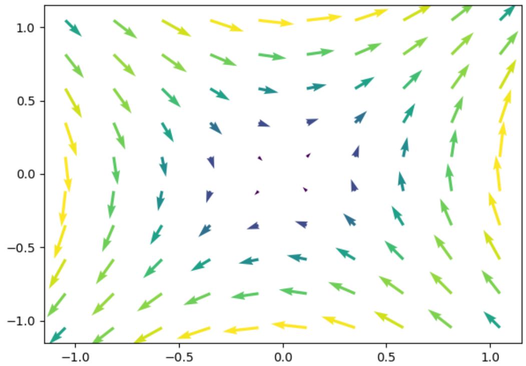


Tracé des vecteurs gradients avec Python

Pour le tracé de champs de vecteurs, on peut utiliser la commande `quiver`.

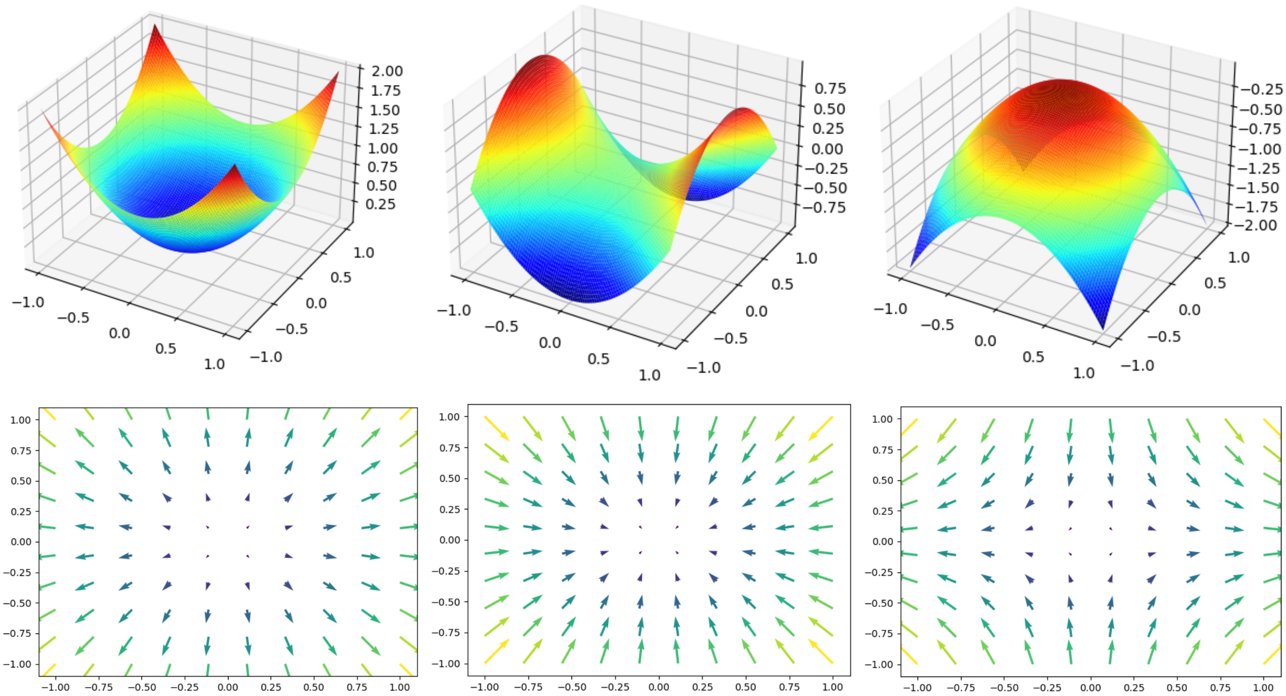
Editeur

```
x = np.linspace(-np.pi/3,np.pi/3,10)
y = np.linspace(-np.pi/3,np.pi/3,10)
X, Y = np.meshgrid(x, y)
Z = np.sin(X*Y)
dx=Y*np.cos(X*Y)
 # expression de la première dérivée partielle
dy=X*np.cos(X*Y)
 # expression de la seconde
color_array = np.sqrt((dx)**2+(dy)**2)
 # Pour que la couleur dépende
 # de la norme du gradient
fig, ax = plt.subplots(figsize=(7,7))
ax.quiver(X,Y,dx,dy,color_array)
plt.show()
```



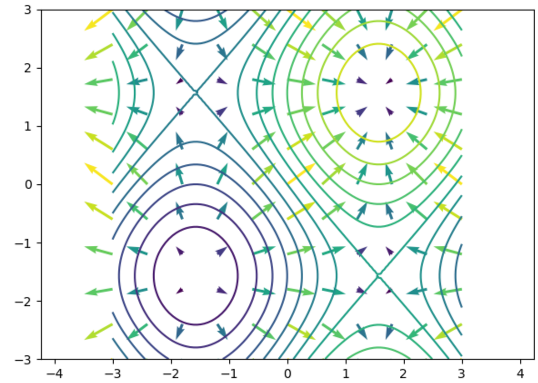
Exercice 204

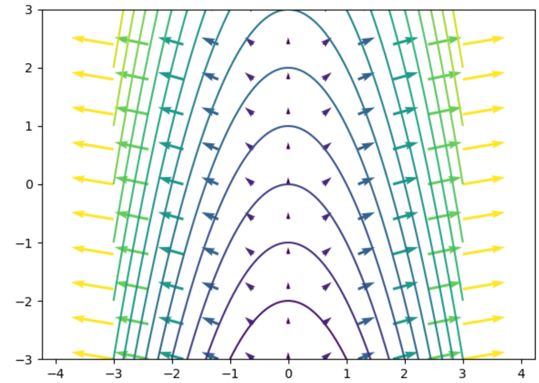
✦ Associer à chaque surface ci-dessous, la représentation des gradients.

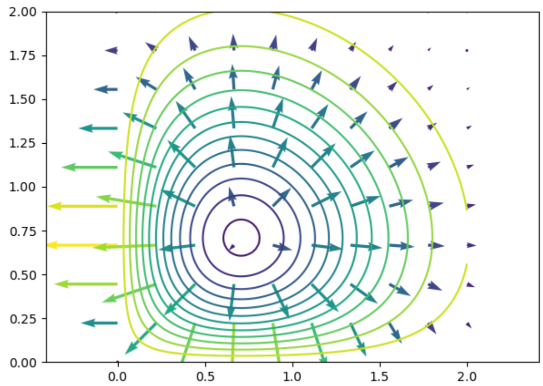


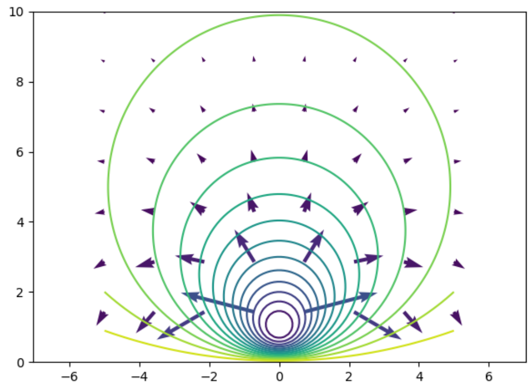
Remarque. En rajoutant, les lignes de niveaux, on peut de nouveau vérifier l'orthogonalité du gradient avec les lignes de niveau. Voici une succession d'exemples avec les fonctions. (Attention, il faut bien rendre les axes normés).

- $f(x, y) = 4\sin(x) + 3\sin(y)$ sur $[-3;3]^2$.
- $f(x, y) = x^2 + y$ sur $[-3;3]^2$.
- $f(x, y) = -xye^{-x^2-y^2}$ sur $[0;2]^2$.
- $f(x, y) = -\frac{3\cdot y}{1+x^2+y^2}$ sur $[-5,5] \times [0;10]$.



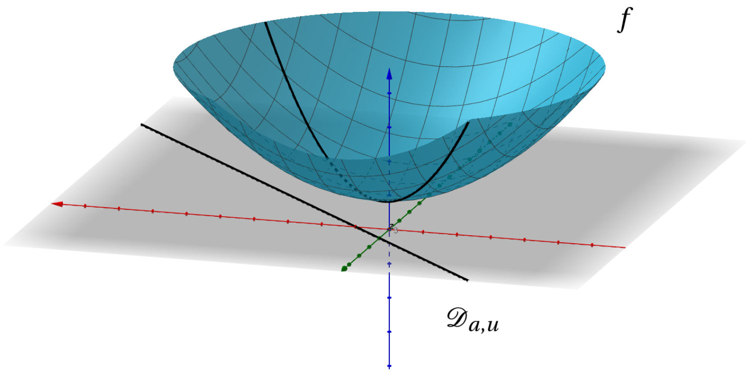


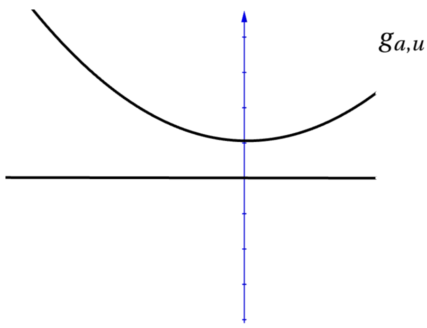




3.3 Dérivées directionnelles

Soient $a,\ u \in \mathbb{R}^n \setminus \{0\}$. La droite affine passant par a et de direction u est l'ensemble $\mathscr{D}_{a,u} = \left\{a + t \cdot u \,|\, t \in \mathbb{R}\right\}$. Afin de ramener le problème à une fonction d'une seule variable, on peut restreindre l'étude de la fonction f à la droite $\mathscr{D}_{a,u}$. On pose donc $g_{a,u} :\ t \in \mathbb{R} \mapsto f(a + t \cdot u) \in \mathbb{R}$.





Théorème 159 (gradient et dérivée directionnelle)

Soient $f : \mathbb{R}^n \to \mathbb{R}$ de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$ et $a,\ u \in \mathbb{R}^n$ avec u, non nul. Alors la fonction

$$g_{a,u} :\ t \in \mathbb{R} \mapsto f(a + t \cdot u) \in \mathbb{R}$$

est de classe $\mathscr{C}^1$ sur $\mathbb{R}$ avec
$$\forall t \in \mathbb{R}, \qquad g'_{a,u}(t) = \langle \nabla f(a + tu), u \rangle.$$

Remarques.

- Si on précise les dérivées partielles, on obtient $g'_{a,u}(t) = \sum\limits_{i=1}^{n} u_i \partial_i f(a + tu)$ où $u = (u_1, \ldots, u_n)$.

- Lorsqu'on considère le vecteur e_i de la base canonique de $\mathbb{R}^n$, on retrouve l'expression de la dérivée partielle $\partial_k f(a)$.
- Pour $t = 0$, on a directement

$$\boxed{g'_{a,u}(0) = \langle u, \nabla f(a) \rangle.}$$

Lorsque u est de norme 1, cette quantité représente **la dérivée directionnelle de f au point a** dans la direction (et le sens) de u.

4 Optimisation : condition d'ordre 1

4.1 Extrema locaux

Définition 160 (extrema locaux)

Soient $f : \mathbb{R}^n \to \mathbb{R}$ et $a \in \mathbb{R}^n$.

→ On dit que f a un **maximum local** en a s'il existe $r \in \mathbb{R}_*^+$ tel que

$$\forall x \in \mathbb{R}^n, \qquad \Big(\|x - a\| < r \quad \Rightarrow \quad f(a) \geqslant f(x) \Big).$$

→ On dit que f a un **minimum local** en a s'il existe $r \in \mathbb{R}_*^+$ tel que

$$\forall x \in \mathbb{R}^n, \qquad \Big(\|x - a\| < r \quad \Rightarrow \quad f(a) \leqslant f(x) \Big).$$

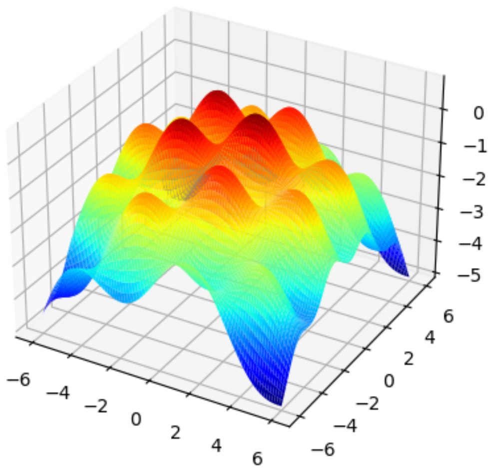
→ On dit que f a un **extremum local** si f a un maximum local ou un minimum local.

Remarque. En pratique, pour étudier un extremum, on regardera le signe de $f(x) - f(a)$.

Exemple. Ci-dessous, un exemple de fonction avec une multitude d'extrema locaux qui ne sont tous pas globaux.

Editeur

```
def f(x, y):
    z=np.sin(x)*np.cos(2*y) -(x**2+y**2)/15
    return z

x = np.linspace(-6, 6, 100)
y = np.linspace(-6, 6, 100)
X, Y = np.meshgrid(x, y)
Z = f(X,Y)

fig = plt.figure()
ax = plt.axes(projection='3d')
ax.plot_surface(X, Y, Z, rstride=1,
cstride=1,cmap='jet', edgecolor='none')
```



4.2 Point critique et condition nécessaire d'extremum

Définition 161 (point critique)

Soient $f : \mathbb{R}^n \to \mathbb{R}$ de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$ et $a \in \mathbb{R}^n$.
On dit que a est un **point critique** de f si $\nabla f(a) = 0_{\mathbb{R}^n}$. Autrement dit,

$$\forall i \in [\![1; n]\!], \qquad \partial_i f(a) = 0.$$

Exercice 205

1. ✧ Calculer les points critiques des fonctions suivantes.
$$f(x,y)=2x^2+2y^2+2xy-x-y \quad \text{et} \quad g(x,y)=2\mathrm{e}^{2x}+2\mathrm{e}^{2y}+2\mathrm{e}^{x+y}-\mathrm{e}^x-\mathrm{e}^y.$$
2. Justifier que la fonction h définie sur $\mathbb{R}^2$ par $h(x,y)=\mathrm{e}^x\left(x+y^2+\mathrm{e}^x\right)$ admet un unique point critique.

Théorème 162 (condition nécessaire d'extremum)

Soient $f:\mathbb{R}^n\to\mathbb{R}$ de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$ et $a\in\mathbb{R}^n$.

Si f a un extremum en $a\in\mathbb{R}^n$,

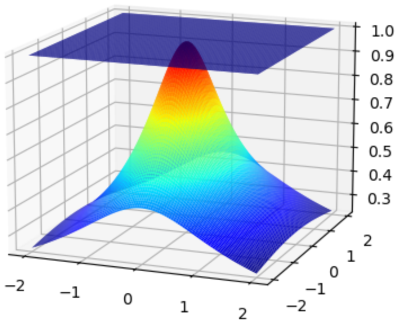
alors a est un point critique de f.

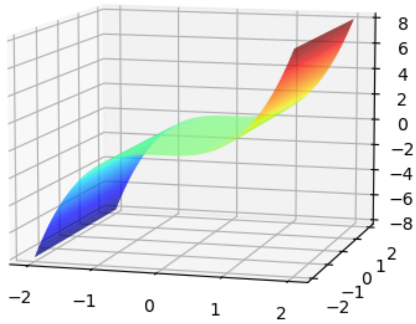
Exercice 206

✦ **Preuve**
Soit f une fonction de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$ admettant un maximum en $a\in\mathbb{R}^n$.

1. Soit u un vecteur de $\mathbb{R}^n$ non nul. En utilisant la fonction $g_{a,u}: t\in\mathbb{R}\mapsto f(a+tu)$, montrer que $\langle\nabla f(a),u\rangle=0$.
2. Conclure que a est un point critique de f.

Attention. La réciproque est fausse.
Ci-dessous, deux exemples de fonctions admettant un point critique en $(0,0)$. Mais dans le second cas, il n'existe pas d'extremum en $(0,0)$.





Vocabulaire. Un point critique qui ne correspond pas à un extremum est un point selle.

Exemple.
→ *La fonction f est de classe $\mathscr{C}^1$ et calcul des dérivées partielles.*
La fonction f est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$. Pour tout $(x,y)\in\mathbb{R}^2$, on a
$$\partial_1 f(x,y)=2x-3+y \quad \text{et} \quad \partial_2 f(x,y)=x+2y.$$
→ *Recherche des points critiques.*
Si f possède un extremum local en (x,y), alors $\partial_1 f(x,y)=\partial_2(x,y)=0$, donc $2x+y=3$ et $x+2y=0$. On obtient un unique point critique avec $\mathrm{A}=(2,-1)$.
→ *Vérifications du ou des points candidats.*
On étudie le signe de la différence $f(x,y)-f(\mathrm{A})$. Pour tout $(h,k)\in\mathbb{R}^2$:
$$f(2+h,-1+k)-f(2,-1)=(2+h)^2-3(2+h)+(2+h)(-1+k)+(-1+k)^2+3=h^2+hk+k^2=\left(h+\frac{k}{2}\right)^2+\frac{3k^2}{4}\geqslant 0,$$
donc f possède un minimum local en $(2,-1)$.

Exercice 207

✦ En reprenant la méthode précédente, déterminer les extrema de f définie sur $\mathbb{R}^2$ par
$$f(x,y)=x^2y+y^2+x.$$

Exercices - TD

Dérivées partielles premières

Exercice 208. ✦ Calculer les dérivées partielles de f définie sur $\mathbb{R}^2$ par $f(x,y) = \displaystyle\int_{y^2}^{y^2+x^2} \cos(u^2)\,\mathrm{d}u$.

Exercice 209. ✦ Déterminer toutes les fonctions $f : \mathbb{R}^2 \to \mathbb{R}$ de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$ telles que

$$\forall (x,y) \in \mathbb{R}^2, \qquad \partial_1 f(x,y) = 4x^3y + 3y + 1 \quad \text{et} \quad \partial_2 f(x,y) = x^4 + 3x.$$

Exercice 210. ✦✦ **Dérivation composée**

1. *Préliminaires*
   Soient f une fonction définie et de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$ et $\varphi_1, \varphi_2, \ldots, \varphi_n$, n fonctions définies et de classe $\mathscr{C}^1$ sur un intervalle ouvert I de $\mathbb{R}$. On définit la fonction g par :
   $$\forall t \in \mathrm{I}, \qquad g(t) = f\left(\varphi_1(t), \varphi_2(t), \ldots, \varphi_n(t)\right).$$
   Montrer que la fonction g est également de classe $\mathscr{C}^1$ sur I et
   $$\forall t \in \mathrm{I}, \qquad g'(t) = \sum_{i=1}^{n} \partial_i f\left(\varphi_1(t), \varphi_2(t), \ldots, \varphi_n(t)\right)\varphi_i'(t).$$
2. *Application 1*
   Soit f une fonction de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$ telle que :
   $$\forall (x,y) \in \mathbb{R}^2, \qquad \nabla f(x,y) = (x,y).$$
   Démontrer que tous les points $\left(\cos(t), \sin(t)\right) \in \mathbb{R}^2$ avec $t \in \mathbb{R}$ appartiennent tous à la même ligne de niveau de f.
3. *Application 2 : les fonctions homogènes*
   On dit qu'une fonction $f : \mathbb{R}^2 \mapsto \mathbb{R}$ est homogène de degré $\alpha \in \mathbb{R}$ si pour tous $(x,y) \in \mathbb{R}^2$ et $\lambda \in \mathbb{R}_*^+$:
   $$f(\lambda x, \lambda y) = \lambda^\alpha f(x,y).$$
   (a) Soit f une fonction homogène.
      i. Montrer que les dérivées partielles d'une fonction homogène de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$ sont aussi homogènes.
      ii. Justifier que
      $$\forall (x,y) \in \mathbb{R}^2, \quad x\partial_1 f(x,y) + y\partial_2 f(x,y) = \alpha f(x,y) \qquad (\bullet)$$
   (b) On suppose réciproquement que f vérifie $(\bullet)$.
      i. Montrer que, pour tout $(x,y) \in \mathbb{R}^2$, l'application $\varphi : t \mapsto f\left(tx, ty\right)$ vérifie l'équation différentielle
      $$\forall t \in \mathbb{R}_*^+, \quad \varphi'(t) = \frac{\alpha}{t}\varphi(t).$$
      ii. En étudiant la fonction $h : t \in \mathbb{R}_*^+ \mapsto \varphi(t)t^{-\alpha}$, en déduire que f est homogène de degré α.

Exercice 211. ✦✦ **Inégalité des accroissements finis**
On considère une application $f : \mathbb{R}^2 \to \mathbb{R}$ de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$. On suppose qu'il existe $k \in [0,1[$ tel que, pour tout $(x,y) \in \mathbb{R}^2$,
$$\left|\partial_1 f(x,y)\right| + \left|\partial_2 f(x,y)\right| \leqslant k.$$

1. Soient $(x,y), (x',y') \in \mathbb{R}^2$. Justifier que
   $$\left|f(x,y) - f(x',y')\right| \leqslant k\max\left(\left|x-x'\right|, \left|y-y'\right|\right).$$
2. Soit $(u_n)_{n\in\mathbb{N}}$ la suite définie par
   $$u_0 = 1, \quad u_1 = 2 \quad \text{et} \quad \forall n \in \mathbb{N}, \quad u_{n+2} = f\left(u_{n+1}, u_n\right).$$
   Pour $n \in \mathbb{N}$, on pose $a_n = \max\left(|u_{n+2} - u_{n+1}|, |u_{n+1} - u_n|\right)$.
   (a) Montrer que, pour tout $n \in \mathbb{N}$, $a_{n+1} \leqslant k a_n$. en déduire la convergence de la série $\sum a_n$.
   (b) Conclure sur la convergence de la suite (u_n).
3. Justifier que la convergence ne dépend pas des conditions initiales.
4. *Python.* On définit la suite u par $u_0 = 1$, $u_1 = 2$ et
   $$\forall n \in \mathbb{N}, \quad u_{n+2} = 1 + \frac{1}{4}\cdot\sin(u_n + u_{n+1}).$$
   (a) Écrire un programme qui prend en argument n et renvoie u_n.
   (b) Vérifier que la suite u est convergente. Donner une approximation de la limite à 10^{-3} près.

Points critiques et optimisation

Exercice 212. ✦ Recherche d'un minimum

Soit $f : \mathbb{R}^3 \to \mathbb{R}$ définie par $$f(x,y,z) = \left(x+z^2\right)\mathrm{e}^{x\left(y^2+z^2+1\right)}.$$

1. (a) Montrer que f est de classe $\mathscr{C}^1$ sur $\mathbb{R}^3$.
   (b) Vérifier que f possède un unique point critique, noté $\mathrm{A} \in \mathbb{R}^3$. Calculer $f(\mathrm{A})$.
2. (a) Montrer que si $x \geqslant 0$, alors $f(x,y,z) \geqslant 0$, et que si $x \leqslant 0$, alors $f(x,y,z) \geqslant x\mathrm{e}^x$.
   (b) Déterminer le minimum de la fonction $x \in \mathbb{R} \mapsto x\mathrm{e}^x$, et en déduire que f atteint son minimum en A.

Exercice 213. ✦✦ Une infinité de points critiques! *D'après oral ESCP*

Soit f l'application de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par

$$\forall (x,y) \in \mathbb{R}^2, \qquad f(x,y) = (2x-y)^2\mathrm{e}^{2x-y}.$$

1. Justifier que f est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$ et vérifier que

$$\forall \mathrm{A} \in \mathbb{R}^2, \qquad \partial_1 f(\mathrm{A}) + 2\partial_2 f(\mathrm{A}) = 0.$$

2. Montrer que f possède une infinité de points critiques. Trouver ceux en lesquels f admet un extremum local ou global.

Exercice 214. ✦✦✦ Mélange avec des probabilités *D'après oral ESCP*

Toutes les variables aléatoires de cet exercice sont à densité et définies sur un même espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$. Sous réserve d'existence, on note $\mathbf{E}(\mathrm{X})$ l'espérance d'une variable aléatoire X.

Dans tout l'exercice, X, Y et Z sont trois variables aléatoires ayant des moments d'ordre 2 . On admet que chacune des variables aléatoires XY, XZ et YZ admet une espérance et on suppose que la condition suivante est vérifiée : $\mathbf{E}\left(\mathrm{X}^2\right)\mathbf{E}\left(\mathrm{Y}^2\right) - (\mathbf{E}(\mathrm{XY}))^2 \neq 0$ Pour tout $(x,y) \in \mathbb{R}^2$, on pose : $f(x,y) = \mathbf{E}\left((\mathrm{Z}-x\mathrm{X}-y\mathrm{Y})^2\right)$.

1. Quelle est, selon les valeurs des réels a, b, c et d, le rang de la matrice $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$?
2. (a) Établir les inégalités strictes : $\mathbf{E}\left(\mathrm{X}^2\right) > 0$ et $\mathbf{E}\left(\mathrm{Y}^2\right) > 0$.
   (b) Montrer que pour tout couple $(x,y) \in \left(\mathbb{R}^*\right)^2$, on a : $\mathbf{E}\left((x\mathrm{X}+y\mathrm{Y})^2\right) > 0$.
3. (a) Montrer que la fonction f est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$ et qu'elle admet un unique point critique $\left(x_0, y_0\right)$.
   (b) Montrer que pour tout couple $(x,y) \in \mathbb{R}^2$, on a :

   $$\mathbf{E}\left(\left(\mathrm{Z}-x_0\mathrm{X}-y_0\mathrm{Y}\right)(x\mathrm{X}+y\mathrm{Y})\right) = 0.$$

   (c) En déduire l'égalité :

   $$f(x,y) = \mathbf{E}\left(\left(\mathrm{Z}-x_0\mathrm{X}-y_0\mathrm{Y}\right)^2\right) + \mathbf{E}\left(\left[\left(x-x_0\right)\mathrm{X}-\left(y_0-y\right)\mathrm{Y}\right]^2\right)$$

   (d) Étudier les extrema de f.
4. Dans cette question, on suppose que X et Y sont indépendantes et suivent toutes les deux la loi uniforme sur l'intervalle $[0,1]$ et on pose : $\mathrm{Z} = \mathrm{X}^2$. Déterminer l'ensemble des couples $\left(x_0, y_0\right)$ pour lesquels $\mathbf{E}\left((\mathrm{Z}-x\mathrm{X}-y\mathrm{Y})^2\right)$ est minimale.

Exercice 215. ✦✦✦ Fonctions convexes sur $\mathbb{R}^n$ *D'après oral ESCP 2001*

On munit $\mathbb{R}^2$ du produit scalaire canonique $\langle x, y\rangle = x_1y_1 + x_2y_2$ où $x = (x_1, x_2)$ et $y = \left(y_1, y_2\right)$.

Soit f une application de $\mathbb{R}^2$ dans $\mathbb{R}$ et de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$.

1. Soient $x = (x_1, x_2)$, $y = \left(y_1, y_2\right) \in \mathbb{R}^2$ et $\varphi_{x,y}$ définie sur $\mathbb{R}$ par $\varphi_{x,y}(t) = f\left(x + t(y-x)\right)$.
   Justifier que $\varphi_{x,y}$ est de classe $\mathscr{C}^1$ et rappeler une expression de sa dérivée.
2. On suppose que f est convexe sur $\mathbb{R}^2$. C'est-à-dire

   $$\forall\, x,\, y \in \mathbb{R}^2, \quad \forall \lambda \in [0,1], \quad f\left(\lambda x + (1-\lambda)y\right) \leqslant \lambda f(x) + (1-\lambda) f(y).$$

   (a) Montrer que pour tous $x,\ y \in \mathbb{R}^2$, $\varphi_{x,y}$ est une fonction convexe (d'une seule variable). Que dire sur les variations de la fonction $\varphi'_{x,y}$?
   (b) En déduire que

   $$\forall\, x,\, y \in \mathbb{R}^2, \quad f(y) - f(x) \geqslant \left\langle \nabla f(x), y - x\right\rangle \qquad (\bullet)$$

3. Réciproquement, on suppose que f vérifie la relation $(\bullet)$ ci-dessus. Montrer que f est convexe.
4. On suppose que f est convexe sur $\mathbb{R}^2$.
   (a) Montrer que si f présente en $x_0 \in \mathbb{R}^2$ un minimum relatif, alors f présente en x_0 un minimum global.
   (b) Montrer que si l'ensemble des points où f admet un minimum, noté $\mathscr{A}$, est non vide, alors cet ensemble est une partie convexe de $\mathbb{R}^2$.

   $$\forall\, (x,y) \in \mathscr{A}^2, \quad \forall \lambda \in [0,1], \quad \lambda x + (1-\lambda) y \in \mathscr{A}.$$

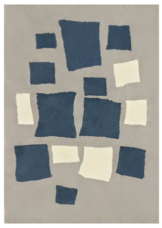
CHAPITRE 11

Vecteurs aléatoires

Le hasard est ma matière première.

JEAN ARP (1886-1966)
Co-fondateur du mouvement Dada.

Ci-contre : "Collage avec des carrés disposés selon les lois du hasard".



1 Rappels : couples de variables aléatoires

1.1 Lois, lois marginales et indépendance

Définition 163 (loi d'un couple)

La loi (conjointe) d'un couple (X, Y) de variables aléatoires discrètes, c'est la donnée de la valeur de $\mathbf{P}\big([X = x] \cap [Y = y]\big)$ pour tout couple $(x, y) \in X(\Omega) \times Y(\Omega)$.

Définition 164 (indépendance)

Soient X et Y deux variables aléatoires discrètes. On dit que X et Y sont **indépendantes** si

$$\forall (x, y) \in X(\Omega) \times Y(\Omega), \qquad \mathbf{P}\big([X = x] \cap [Y = y]\big) = \mathbf{P}(X = x) \cdot \mathbf{P}(Y = y).$$

Remarque. Les lois de X et Y sont appelées **lois marginales**. Elles s'obtiennent à partir de la loi du couple en utilisant la formule des probabilités totales :

$$\forall x \in X(\Omega), \qquad \mathbf{P}(X = x) = \sum_{y \in Y(\Omega)} \mathbf{P}\big([X = x] \cap [Y = y]\big)$$

$$\forall y \in Y(\Omega), \qquad \mathbf{P}(Y = y) = \sum_{x \in X(\Omega)} \mathbf{P}\big([X = x] \cap [Y = y]\big)$$

⚠ **Attention.** Les lois marginales de X et Y ne permettent pas de retrouver la loi du couple.
Par exemple, soient X et Y deux variables de Bernoulli de paramètre 1/2. Dans les trois cas suivants, les lois marginales sont identiques alors que les lois de couples différent.

- X et Y indépendantes.
- $Y = X$.
- $Y = 1 - X$.

Par contre si X et Y sont indépendantes, la loi du couple (X, Y) est connue.

Proposition 165 (loi d'une fonction, cas général)

Soient X et Y deux variables aléatoires réelles discrètes, et soit g une fonction à valeurs réelles définie sur le sous-ensemble $X(\Omega) \times Y(\Omega)$ de $\mathbb{R}^2$. Notons Z la variable aléatoire $g(X,Y)$. Alors, pour tout élément k de $Z(\Omega)$,

$$\mathbf{P}(Z = k) = \sum_{\substack{i \in X(\Omega),\ j \in Y(\Omega) \\ g(i,j)=k}} \mathbf{P}([X = i] \cap [Y = j]),$$

où la somme porte sur le sous-ensemble de $X(\Omega) \times Y(\Omega)$ constitué par les couples (i, j) tels que $g(i, j) = k$.

1.2 Calculs d'espérance

Sans hypothèses particulières, l'espérance d'une fonction de deux variables aléatoires s'obtient à l'aide du théorème de transfert. La loi conjointe des deux variables permet le calcul, lorsqu'elle existe de l'espérance de $g(X,Y)$ sans nécessité de calculer la loi de la fonction associée.

Théorème 166 (de transfert pour un couple de variables)

Soient X et Y deux variables aléatoires réelles discrètes, et soit g une fonction à valeurs réelles définie sur le sous-ensemble $X(\Omega) \times Y(\Omega)$ de $\mathbb{R}^2$. Sous réserve de convergence absolue,

$$\mathbf{E}\big(g(X,Y)\big) = \sum_{\substack{x \in X(\Omega) \\ y \in Y(\Omega)}} g(x,y) \cdot \mathbf{P}\big([X = x] \cap [Y = y]\big).$$

Remarque. C'est l'hypothèse de convergence absolue qui assure que la valeur de la somme ne dépend pas de l'ordre de ses termes.

Voici deux applications de ce théorème, dans le cas d'une combinaison linéaire d'une part, et dans le cas d'un produit de deux variables indépendantes d'autre part.

Corollaire 167 (linéarité de l'espérance)

Soient X et Y deux variables aléatoires définies sur $(\Omega, \mathscr{A}, \mathbf{P})$ et λ et μ deux réels.

Si X et Y admettent une espérance,

alors alors $\lambda X + \mu Y$ admet une espérance et $\mathbf{E}(\lambda X + \mu Y) = \lambda \mathbf{E}(X) + \mu \mathbf{E}(Y)$.

Corollaire 168 (espérance d'un produit de variables indépendantes)

Soient X et Y deux variables aléatoires.

Si X et Y
- admettent une espérance et
- sont indépendantes,

alors $X \cdot Y$ admet une espérance et $\mathbf{E}(X \cdot Y) = \mathbf{E}(X) \cdot \mathbf{E}(Y)$.

Exercice 216

✧ Soit X une variable aléatoire suivant la loi uniforme sur $X(\Omega) = \{-1; 0; 1\}$. Notons $Y = X^2$.

1. Déterminer la loi du couple (X,Y). En déduire la loi de Y.
2. Les variables X et Y sont-elles indépendantes ?
3. Comparer $\mathbf{E}(X \cdot Y)$ et $\mathbf{E}(X) \cdot \mathbf{E}(Y)$. Commenter.

1.3 Loi d'une somme, exemples

Soient (X,Y) un couple de variables aléatoires. Donnons la loi de la somme $Z = X+Y$. On a

$$Z(\Omega) = \{x+y \mid (x,y) \in X(\Omega) \times Y(\Omega)\}.$$

Pour tout $z \in Z(\Omega)$.

$$\mathbf{P}(Z=z) = \sum_{\substack{(x,y)\in X(\Omega)\times Y(\Omega)\\ x+y=z}} \mathbf{P}\big([X=x] \cap [Y=y]\big) = \sum_{\substack{x\in X(\Omega)\\ z-x\in Y(\Omega)}} \mathbf{P}\big([X=x]\cap[Y=z-x]\big).$$

Si les variables X et Y sont indépendantes, alors la loi de Z est donnée par la formule du **produit de convolution discret**. Pour tout $z \in Z(\Omega)$:

$$\mathbf{P}(Z=z) = \sum_{\substack{x\in X(\Omega)\\ z-x\in Y(\Omega)}} \mathbf{P}(X=x)\mathbf{P}(Y=z-x).$$

Dans le cas particulier où les variables X et Y sont supposées indépendantes, de loi binomiale ou de Poisson, et où la fonction g est simplement l'addition, la loi de $g(X,Y)$ est en fait totalement connue :

Proposition 169 (somme de binomiale)

Soient $m, n \in \mathbb{N}^*$, $p \in [0;1]$ et $X_1, X_2 : \Omega \to \mathbb{R}$ deux variables aléatoires.

Si
- X_1 et X_2 sont indépendantes.
- $X_1 \hookrightarrow \mathscr{B}(n,p)$ et $X_2 \hookrightarrow \mathscr{B}(m,p)$.

Alors $$X_1 + X_2 \hookrightarrow \mathscr{B}(n+m,p).$$

Proposition 170 (somme de loi de Poisson)

Soient $\lambda, \mu \in \mathbb{R}_+^*$ et $X_1, X_2 : \Omega \to \mathbb{R}$ deux variables aléatoires.

Si
- X_1 et X_2 sont indépendantes.
- $X_1 \hookrightarrow \mathscr{P}(\lambda)$ et $X_2 \hookrightarrow \mathscr{P}(\mu)$.

Alors $$X_1 + X_2 \hookrightarrow \mathscr{P}(\lambda+\mu).$$

Remarque. Par récurrence, ces résultats s'étendent pour des sommes de n variables aléatoires.

1.4 Loi du maximum, du minimum

Pour déterminer la loi de $\max(X,Y)$, on passe par la fonction de répartition.
Pour la loi du $T = \min(X,Y)$, on passe par la fonction d'anti-répartition ($x \in \mathbb{R} \mapsto \mathbf{P}(T > x)$).

Exercice 217

✦ **Un exemple classique**
Soient p et r deux réels de $]0;1[$ et X et Y deux v.a indépendantes. On suppose de plus que $X \hookrightarrow \mathscr{G}(p)$, $Y \hookrightarrow \mathscr{G}(r)$.

1. Donner la loi de la variable aléatoire $\min(X,Y)$.
2. Calculer $\mathbf{P}\big([X \geqslant Y]\big)$.
3. Comment interpréter les résultats précédents en termes de lancer de pièces ?

2 Généralisation aux vecteurs aléatoires, indépendances

Dans la suite, un vecteur aléatoire sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ est la donnée d'un n-uplet $(X_1, \ldots, X_n)$ où chaque X_i désigne une variable aléatoire réelle sur $(\Omega, \mathscr{A}, \mathbf{P})$.

2.1 Lois, lois marginales

Définition 171 (loi d'un vecteur aléatoire, loi marginale)

- La **loi d'un vecteur aléatoire** $(X_1,\dots,X_n)$ de variables aléatoires réelles est donnée par la fonction $F_{(X_1,\dots,X_n)}$ définie sur $\mathbb{R}^n$ par

$$F_{(X_1,\dots,X_n)}(t_1,\dots,t_n) = \mathbf{P}\left(\bigcap_{i=1}^{n}[X_i \leqslant t_i]\right).$$

- On appelle i**-ème loi marginale** de $(X_1,X_2,\dots,X_n)$ la loi de X_i.

Proposition 172 (égalité en loi)

Soient $(X_i)_{i\in[\![1;n]\!]}$, $(Y_i)_{i\in[\![1;n]\!]}$ deux vecteurs aléatoires définis sur un même espace probabilisé $(\Omega,\mathscr{A},\mathbf{P})$.

Si
- → Les vecteurs $(X_1,\dots,X_n)$ et $(Y_1,\dots,Y_n)$ ont la même loi.
- → La fonction g est continue sur $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$.

Alors les variables aléatoires $g(X_1,\dots,X_n)$ et $g(Y_1,\dots,Y_n)$ ont la même loi.

Exercice 218

✦ Soient X et Y deux variables aléatoires indépendantes définies sur un même espace probabilisé $(\Omega,\mathscr{A},\mathbf{P})$, de même loi, à valeurs dans $[1;+\infty[$.
Montrer que $\mathbf{E}\left(\frac{X}{X+Y}\right)$ existe et vaut $\frac{1}{2}$.
On pourra remarquer que $\frac{X}{X+Y}$ *et* $\frac{Y}{X+Y}$ *ont même loi.*

2.2 Indépendance

Définition 173 (indépendance)

Soit $(X_1,X_2,\dots,X_n)$, un vecteur aléatoire défini sur un espace probabilisé $(\Omega,\mathscr{A},\mathbf{P})$.
Les variables X_1, X_2, ..., X_n sont dites **mutuellement indépendantes** si pour tout $(t_1,\dots,t_n)\in\mathbb{R}^n$, les événements $[X_1\leqslant t_1],\dots,[X_n\leqslant t_n]$ sont mutuellement indépendants. Autrement dit,

$$\forall\,(t_1,\dots,t_n)\in\mathbb{R}^n,\qquad \mathbf{P}\left(\bigcap_{i=1}^{n}[X_i\leqslant t_i]\right)=\prod_{i=1}^{n}\mathbf{P}(X_i\leqslant t_i).$$

Remarques.

- Autrement dit, les variables aléatoires $X_1,\dots,X_n$ sont mutuellement indépendantes si et seulement si

$$\forall\,(t_1,t_2,\dots,t_n)\in\mathbb{R}^n,\qquad F_{(X_1,\dots,X_n)}(t_1,\dots,t_n)=\prod_{k=1}^{n}F_{X_k}(t_k).$$

- Il est clair que, si les variables aléatoires $X_1,X_2,\dots,X_n$ sont mutuellement indépendantes, alors, pour tout $k\in[\![1;n]\!]$, les variables aléatoires $X_1,X_2,\dots,X_k$ sont mutuellement indépendantes.

- On dit qu'une suite de variables aléatoires $(X_n)_{n\in\mathbb{N}}$ est composée de variables mutuellement indépendantes si pour toute partie finie $I\subset\mathbb{N}$, les variables $(X_i)_{i\in I}$ sont mutuellement indépendantes.

Proposition 174 (des coalitions)

Soit $(X_1, X_2, \dots, X_n)$, un vecteur aléatoire défini sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$.

Si $X_1, X_2, \dots, X_n$, sont indépendantes,

alors toute variable aléatoire fonction de $X_1, X_2, \dots, X_p$ est indépendante de toute variable aléatoire fonction de $X_{p+1}, X_{p+2}, \dots, X_n$.

Vocabulaire. La fonction $\varphi(X_1, \dots, X_n)$ de $X_1, X_2, \dots, X_p$ est une variable aléatoire

$$\varphi : \left\{ \begin{array}{lcl} \Omega & \to & \mathbb{R} \\ \omega & \mapsto & \varphi\big(X_1(\omega), \dots, X_n(\omega)\big) \end{array} \right. \qquad \text{où} \quad \varphi : \mathbb{R}^n \to \mathbb{R}.$$

Proposition 175 (caractérisation par des intervalles)

Soit $(X_1, X_2, \dots, X_n)$, un vecteur aléatoire défini sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$. Les deux énoncés suivants sont équivalents.

i) $X_1, \dots, X_n$ sont mutuellement indépendantes.

ii) Pour tous intervalles $I_1, \dots, I_n$ de $\mathbb{R}$, $\quad \mathbf{P}\left(\left[\bigcap_{i=1}^{n} X_i \in I_i\right]\right) = \prod_{i=1}^{n} \mathbf{P}\big([X_i \in I_i]\big).$

Proposition 176 (caractérisation de l'indépendance, cas discret)

Soit $(X_1, X_2, \dots, X_n)$, un vecteur aléatoire défini sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$. Les variables $X_1, \dots, X_n$ sont mutuellement indépendantes si et seulement si pour tout $(t_1, \dots, t_n) \in X_1(\Omega) \times \dots \times X_n(\Omega)$,

$$\mathbf{P}\left(\bigcap_{i=1}^{n} [X_i = t_i]\right) = \prod_{i=1}^{n} \mathbf{P}(X_i = t_i).$$

2.3 Calculs d'espérance et de la variance

Le programme limite la définition de l'espérance et de la variance aux cas de variables aléatoires finies, discrètes dénombrables et à densité. Mais les propositions s'étendent au cas général.

L'espérance

Proposition 177 (linéarité de l'espérance)

Soient $(X_1, X_2, \dots, X_n)$ un vecteur aléatoire défini sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ et $(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$.

Si pour tout $i \in [\![1;n]\!]$, X_i admet une espérance,

Alors $\lambda_1 X_1 + \dots + \lambda_n X_n$ admet une espérance avec

$$\mathbf{E}(\lambda_1 X_1 + \dots + \lambda_n X_n) = \lambda_1 \mathbf{E}(X_1) + \dots + \lambda_n \mathbf{E}(X_n).$$

Proposition 178 (espérance d'un produit)

Soit $(X_1, X_2, \dots, X_n)$, un vecteur aléatoire défini sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$.

Si
- Pour tout $i \in [\![1;n]\!]$, X_i admet une espérance.
- Les variables $(X_i)_{i\in[\![1;n]\!]}$ sont mutuellement indépendantes.

Alors $X_1 \times \cdots \times X_n$ admet une espérance avec

$$\mathbf{E}(X_1 \times \ldots \times X_n) = \mathbf{E}(X_1) \times \ldots \times \mathbf{E}(X_n).$$

Exercice 219

✦ *D'après oral ESCP 2022*

1. Soient S et T deux variables aléatoires réelles indépendantes, telles que T et $-T$ aient la même loi. Montrer que :
$$\mathbf{E}\big(\cos(S+T)\big) = \mathbf{E}\big(\cos(S)\big)\mathbf{E}\big(\cos(T)\big).$$
2. On considère une suite $\big(R_k\big)_{k\in\mathbb{N}^*}$ de variables aléatoires mutuellement indépendantes, à valeurs dans $\{-1, 1\}$ et telles que :
$$\forall k \in \mathbb{N}^*, \quad \mathbf{P}\big(R_k = -1\big) = \mathbf{P}\big(R_k = 1\big) = \frac{1}{2}.$$
et on pose, pour tout $n \in \mathbb{N}^*$, $S_n = \sum\limits_{k=1}^{n} R_k$. Montrer que pour tous $n \in \mathbb{N}^*$ et $t \in \mathbb{R}$, on a :
$$\mathbf{E}\big(\cos(S_n t)\big) = (\cos t)^n.$$

La variance

Proposition 179 (variance et indépendance)

Soient X et Y deux variables aléatoires sur $(\Omega, \mathscr{A}, \mathbf{P})$.

Si
- X et Y admettent une variance.
- X et Y sont indépendantes.

Alors $X+Y$ admet une espérance avec $\mathbf{V}(X+Y) = \mathbf{V}(X) + \mathbf{V}(Y)$.

Le résultat suivant s'en déduit par récurrence en rappelant que, par le lemme des coalitions, X_{k+1} est indépendants de $(X_1 + \ldots + X_k)$.

Corollaire 180 (variance et indépendance)

Soit $(X_1, X_2, \dots, X_n)$, un vecteur aléatoire défini sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$.

Si
- Pour tout $i \in [\![1;n]\!]$, X_i admet une variance.
- Les variables $(X_i)_{i\in[\![1;n]\!]}$ sont mutuellement indépendantes.

Alors $X_1 + \cdots + X_n$ admet une espérance avec

$$\mathbf{V}(X_1 + \cdots + X_n) = \mathbf{V}(X_1) + \cdots + \mathbf{V}(X_n).$$

Exemple. Nous avons vu que si $X \hookrightarrow \mathscr{B}(n,p)$ alors il existe $X_1, X_2, \dots, X_n$, des variables aléatoires mutuellement indépendantes suivant des lois de Bernoulli de paramètre p telles que $X = \sum\limits_{i=1}^{n} X_i$. On retrouve alors l'espérance et la variance

$$\mathbf{E}(\mathrm{X}) = \mathbf{E}\left(\sum_{i=1}^{n} \mathrm{X}_i\right) = \sum_{i=1}^{n} \mathbf{E}(\mathrm{X}_i) = \sum_{i=1}^{n} p = np \quad \text{et} \quad \mathbf{V}(\mathrm{X}) = \mathbf{V}\left(\sum_{i=1}^{n} \mathrm{X}_i\right) = \sum_{i=1}^{n} \mathbf{V}(\mathrm{X}_i) = \sum_{i=1}^{n} pq = npq.$$

3 Compléments sur les couples : la covariance

3.1 Définitions et premières propriétés

Définition-proposition 181 (covariance)

Soient X, Y deux variables aléatoires sur $(\Omega, \mathscr{A}, \mathbf{P})$.
Si X et Y admettent un moment d'ordre 2, alors la variable aléatoire $\left(\mathrm{X}-\mathbf{E}(\mathrm{X})\right)\left(\mathrm{Y}-\mathbf{E}(\mathrm{Y})\right)$ admet une espérance. On définit la covariance de X et Y par

$$\mathrm{Cov}(\mathrm{X},\mathrm{Y}) = \mathbf{E}\Big(\left(\mathrm{X}-\mathbf{E}(\mathrm{X})\right)\cdot\left(\mathrm{Y}-\mathbf{E}(\mathrm{Y})\right)\Big).$$

Remarque. Lorsque $\mathrm{X}=\mathrm{Y}$, nous retrouvons la définition de la variance.

$$\mathbf{V}(\mathrm{X}) = \mathrm{Cov}(\mathrm{X},\mathrm{X}).$$

Soit $(\Omega, \mathscr{A}, \mathbf{P})$, un espace probabilisé. Dans la suite, $\mathscr{M}_2$ désigne l'ensemble des variables aléatoires de $(\Omega, \mathscr{A}, \mathbf{P})$ admettant un moment d'ordre 2.

Proposition 182 (propriété de la covariance)

La covariance est :

- **Symétrique.** $\forall \mathrm{X}, \mathrm{Y} \in \mathscr{M}_2, \quad \mathrm{Cov}(\mathrm{X},\mathrm{Y}) = \mathrm{Cov}(\mathrm{Y},\mathrm{X}).$
- **Positive.** $\forall \mathrm{X} \in \mathscr{M}_2, \quad \mathrm{Cov}(\mathrm{X},\mathrm{X}) \geqslant 0.$
  De plus, il y a égalité $\mathrm{Cov}(\mathrm{X},\mathrm{X}) = 0$ si et seulement si X est presque sûrement constante.
- **Bilinéaire.**
  - Linéaire à droite. Pour tous $\mathrm{X}_1, \mathrm{X}_2, \mathrm{Y} \in \mathscr{M}_2$, tous $\lambda, \mu \in \mathbb{R}$,
  $$\mathrm{Cov}\left(\lambda_1 \mathrm{X}_1 + \mu \mathrm{X}_2, \mathrm{Y}\right) = \lambda_1 \mathrm{Cov}(\mathrm{X}_1, \mathrm{Y}) + \mu \mathrm{Cov}(\mathrm{X}_2, \mathrm{Y}).$$
  - Linéaire à gauche. Pour tous $\mathrm{X}_1, \mathrm{X}_2, \mathrm{Y} \in \mathscr{M}_2$, tous $\lambda, \mu \in \mathbb{R}$,
  $$\mathrm{Cov}\left(\mathrm{Y}, \lambda \mathrm{X}_1 + \mu \mathrm{X}_2\right) = \lambda \mathrm{Cov}(\mathrm{Y}, \mathrm{X}_1) + \mu \mathrm{Cov}(\mathrm{Y}, \mathrm{X}_2).$$

Attention. Ces propriétés sont à rapprocher de la définition d'un produit scalaire. Il y a toutefois, une grande différence : la covariance n'est pas définie.

$$\mathrm{Cov}(\mathrm{X},\mathrm{X}) = 0 \quad \not\Rightarrow \quad \mathrm{X} = 0.$$

Vocabulaire. Deux variables sont dites décorrélées si la covariance est nulle.

Exercice 220

✧ *Les questions sont indépendantes.*

1. Montrer que $\mathrm{Cov}(a\mathrm{X}+b, c\mathrm{Y}+d) = ac\,\mathrm{Cov}(\mathrm{X},\mathrm{Y})$.
2. Soit (X,Y) un couple de variables aléatoires admettant la même variance. Montrer qu'alors les variables aléatoires $\mathrm{X}+\mathrm{Y}$ et $\mathrm{X}-\mathrm{Y}$ sont décorrélées.

Proposition 183 (dépendance et corrélation)

Soient X, Y deux variables aléatoires sur $(\Omega, \mathscr{A}, \mathbf{P})$ admettant un moment d'ordre 2.

Si X et Y sont indépendants

Alors X et Y sont des variables décorrélées : $\mathrm{Cov}(X, Y) = 0$.

Exercice 221

✧ **La réciproque est fausse!**
Soit X une variable aléatoire suivant la loi uniforme sur $\{-1; 0; 1\}$. Vérifier que X et $Y = X^2$ sont décorrélées mais non indépendantes.

Théorème 184 (formule de Huygens)

Soient X, Y deux variables aléatoires sur $(\Omega, \mathscr{A}, \mathbf{P})$ admettant un moment d'ordre 2. On a

$$\mathrm{Cov}(X, Y) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y).$$

Exercice 222

✦ Soient X et Y deux variables aléatoires sur $(\Omega, \mathscr{A}, \mathbf{P})$, indépendantes, admettant une espérance $\mu \neq 0$ et une variance $\nu \neq 0$.

1. Exprimer l'espérance et la variance de XY en fonction de μ et ν.
2. Est-ce que les variables $X + Y$ et XY sont indépendantes ?

Proposition 185 (variance d'une somme)

Soient X, Y deux variables aléatoires sur $(\Omega, \mathscr{A}, \mathbf{P})$ admettant un moment d'ordre 2.
Alors pour tout $(\lambda, \mu) \in \mathbb{R}^2$, la variable $\lambda X + \mu Y$ admet un moment d'ordre 2, et

$$\mathbf{V}(\lambda X + \mu Y) = \lambda^2 \mathbf{V}(X) + 2\lambda\mu\, \mathrm{Cov}(X, Y) + \mu^2 \mathbf{V}(Y).$$

Remarques.

- En particulier, $X + Y$ admet un moment d'ordre 2 et $\mathbf{V}(X + Y) = \mathbf{V}(X) + 2\,\mathrm{Cov}(X, Y) + \mathbf{V}(Y)$.
- Cette formule est à comparer avec la formule avec le produit scalaire. Si $\left(E, \langle\cdot,\cdot\rangle\right)$ est un espace euclidien

$$\forall\, x, y \in E, \qquad \|x + y\|^2 = \|x\|^2 + 2\left\langle x, y\right\rangle + \|y\|^2.$$

- Soient $X_1, \ldots, X_n$ de variables aléatoires définies sur un même espace probabilisé et admettant des variances. On montre que

$$\mathbf{V}\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} \mathbf{V}(X_k) + 2 \sum_{1 \leqslant i < j \leqslant n} \mathrm{cov}\left(X_i, X_j\right).$$

Théorème 186 (inégalité de Cauchy-Schwarz)

Soient X, Y. On a
$$\left|\mathrm{Cov}(X, Y)\right| \leqslant \sigma(X)\,\sigma(Y)$$
où $\sigma(X)$ et $\sigma(Y)$ désignent respectivement l'écart-type de X et Y.

Exercice 223

✦ **Preuve**

1. Prouver cette inégalité en introduisant la fonction polynomiale $t \in \mathbb{R} \mapsto \mathbf{V}(tX + Y)$.
2. Préciser le cas d'égalité.

3.2 Coefficient de corrélation

Définition 187 (coefficients de corrélation)

Soient X, Y deux variables aléatoires sur $(\Omega, \mathscr{A}, \mathbf{P})$ admettant un moment d'ordre 2. On suppose de plus que $\mathbf{V}(X) \neq 0$ et $\mathbf{V}(Y) \neq 0$. On appelle **coefficient de corrélation linéaire**, noté $\rho(X, Y)$, le nombre réel défini par

$$\rho(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sigma(X)\sigma(Y)}$$

où $\sigma(X)$ et $\sigma(Y)$ désignent respectivement l'écart-type de X et Y.

Remarques.

- Si X et Y sont indépendantes, leur coefficient de corrélation linéaire est nul. Mais la réciproque est fausse.
- À l'aide de l'inégalité de Cauchy-Schwarz,

$$\left|\rho(X, Y)\right| \leqslant 1.$$

Précisons que le cas d'égalité. Si $\left|\rho_{X,Y}\right| = 1$, alors il existe des réels a et b tels que $\mathbf{P}(Y = aX + b) = 1$.

Exercice 224

✧ Soient X et Y des variables aléatoires et a et b deux réels non nuls, b et d des réels.

1. Prouver que $\rho_{aX+b,cY+d} = \rho_{X,Y}$. *Propriété d'invariance par changement d'échelle.*
2. Soient X et Y des variables aléatoires admettant des variances non nulles. Soit

$$Z = \left(\frac{1}{\sigma_Y}\right)Y - \left(\frac{\rho_{X,Y}}{\sigma_X}\right)X.$$

   (a) Calculer $\mathbf{V}(Z)$.

   (b) Que peut-on en déduire si $\left|\rho_{X,Y}\right| = 1$?

Exercices - TD

Révisions : couples de v.a discrètes

Exercice 225. ✧ On considère n boîtes numérotées de 1 à n. La boîte numéro k contient k boules numérotées de 1 à k. On choisit au hasard une boîte, puis une boule dans cette boîte. Soient X et Y les numéros de la boîte et de la boule obtenues.

1. Déterminer la loi du couple (X, Y).
2. Calculer $\mathbf{P}(X = Y)$.
3. Déterminer la loi de Y ainsi que son espérance.

Exercice 226. ✦
Une urne contient $N-2$ boules vertes, 1 boule blanche et 1 boule rouge. On tire les boules de l'urne, une à une et sans remise.

1. Soient X_1 le rang d'apparition de la boule blanche et X_2 le rang d'apparition de la boule rouge. Déterminer la loi de X_1, la loi de X_2 et la loi du couple (X_1, X_2).
2. Soit X le rang où on obtient pour la première fois soit la boule blanche, soit la boule rouge. Soit Y le rang où on a obtenu pour la première fois les deux boules blanche et rouge. Déterminer la loi de X et la loi de Y.

Exercice 227. ✦✦ *D'après Oral HEC 2014*
Soient X et Y deux variables aléatoires définies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$, à valeurs dans $\mathbb{N}$ telles que :

$$\forall (i,j) \in \mathbb{N}^2, \qquad \mathbf{P}\big([X = i] \cap [Y = j]\big) = \frac{\lambda}{(i+j+1)!}.$$

1. Déterminer le réel λ.
2. Les variables aléatoires X et Y sont-elles indépendantes ?

Compléments

Exercice 228. ✧ **Vrai ou faux ?**
Dans la suite, X et Y désignent deux variables aléatoires définies sur un même espace probabilisé admettant un moment d'ordre 2.

1. Si $\mathbf{E}(X) = 0$ alors $\mathbf{E}(X^2) = 0$. ✓ ×
2. Si X et Y sont indépendantes, $\mathrm{Cov}(X, X+Y) = \mathbf{V}(X)$. ✓ ×

Exercice 229. ✦ **Un cas très particulier !**
Soient X et Y sont deux variables aléatoires définies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ telles que $X \hookrightarrow \mathscr{B}(p)$ et $Y \hookrightarrow \mathscr{B}(q)$.
Montrer que si X et Y sont décorrélées alors elles sont indépendantes.

Exercice 230. ✦✦ Soient X_n des variables aléatoires indépendantes et suivant une loi de Bernoulli de paramètre p. On pose

$$Y_n = X_n X_{n+1} \quad \text{et} \quad U_n = Y_1 + \ldots + Y_n.$$

1. Quelle est la loi de Y_n ? Les Y_i sont-elles deux à deux indépendantes ?
2. Calculer l'espérance et la variance de U_n.

Exercice 231. ✦✦ **Loi de Pascal**
On dispose d'une urne contenant une proportion $p \in]0,1[$ de boules blanches. On effectue une suite infinie de tirages, indépendants. Pour chaque $r \in \mathbb{N}^*$, on note R_r la variable aléatoire qui renvoie le rang de la r-ième «boule blanche», (avec 0 si le r-ième boule blanche n'apparaît pas).
Pour $i \in \mathbb{N}^*$, on note X_i la v.a qui renvoie 1 si l'on a obtenu une boule blanche au i-ème tirage, et 0 sinon.
Pour $n \in \mathbb{N}^*$, on note $Y_n = \sum\limits_{i=1}^{n} X_i$. On note aussi $Y_0 = 0$.

1. On note A l'évènement « obtenir un nombre fini de boules blanches » et, pour tout $n \in \mathbb{N}$, A_n l'évènement « ne plus obtenir de boules blanches à partir du n-ième lancer ».
   (a) En utilisant le théorème de la limite monotone, montrer que pour tout $n \in \mathbb{N}$, $\mathbf{P}(A_n) = 0$.
   (b) Exprimer A en fonction des évènements $(A_n)_{n \in \mathbb{N}^*}$. En déduire $\mathbf{P}(A)$.
   (c) En déduire que, pour tout $r \in \mathbb{N}^*$, $\mathbf{P}(R_r = 0) = 0$.
2. Identifier la loi de R_1, en déduire l'existence et la valeur de $\mathbf{E}(R_1)$ et $\mathbf{V}(R_1)$.
3. Soit $n \in \mathbb{N}^*$. Quelle est la loi de Y_n ?

4. Soient $r \in \mathbb{N}^*$ et $n \in \mathbb{N}^*$ tel que $n \geqslant r$. Exprimer l'évènement $[R_r = n]$ en fonction d'événements formés à partir des variables Y_{n-1} et X_n.
5. En déduire la loi de R_r.
6. Écrire une fonction Python qui prend en argument r et simule la variable aléatoire R_r.

Exercice 232. ✦✦✦ Somme aléatoire de variables aléatoires discrètes - identité de Wald

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires variables aléatoires mutuellement indépendantes, à valeurs dans $\mathbb{N}$, de même loi et admettant une espérance. Soit N, une nouvelle variable aléatoire à valeurs dans $\mathbb{N}$, indépendante de X_n pour tout $n \in \mathbb{N}^*$. On pose pour tout $\omega \in \Omega$

$$S(\omega) = \begin{cases} X_1(\omega) + \cdots + X_{N(\omega)}(\omega) & \text{si } N(\omega) \geqslant 1 \\ 0 & \text{si } N(\omega) = 0 \end{cases}$$

On admet que S définit une variable aléatoire.

1. Soit $n \in \mathbb{N}$ tel que $\mathbf{P}(N = n) \neq 0$. Justifier l'existence et calculer l'espérance $\mathbf{E}(S \mid [N = n])$.
2. En déduire l'existence de l'espérance de S et l'égalité $\mathbf{E}(S) = \mathbf{E}(X_1) \cdot \mathbf{E}(N)$.

Exercice 233. ✦✦ Un peu de Python

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires mutuellement indépendantes telles que pour tout $n \in \mathbb{N}$,

$$\mathbf{P}(X_n = 1) = p \quad \text{et} \quad \mathbf{P}(X_n = -1) = 1 - p.$$

1. Soit $p \in [0;1]$. En utilisant **rd.random()<p**, écrire une fonction prenant en argument p et simule la variable aléatoire X_1.
2. On pose $S_0 = 0$ et pour tout $n \in \mathbb{N}^*$, $S_n = X_1 + X_2 + \cdots + X_n$.
   Écrire une fonction d'arguments p et n qui renvoie la liste $[S_0, S_1, \cdots, S_n]$.
3. Conjecturer la limite de la suite (S_n) en fonction de p.
4. Soit $n \in \mathbb{N}^*$. On note $T_n = \min\{k \in \mathbb{N}, |S_k| = n\}$. Écrire une fonction d'arguments p, n qui renvoie une simulation de la variable aléatoire T_n.

Exercice 234. ✦✦✦ *d'après HEC 2014*

On lance indéfiniment un dé équilibré et, pour tout $n \in \mathbb{N}^*$, on note X_n le numéro sorti au n-ième tirage. Les variables aléatoires X_n, définies sur un espace probabilisé $(\Omega, \mathscr{A}, P)$, sont donc supposées indépendantes et de même loi uniforme sur $[\![1,6]\!]$. Pour tout $i \in [\![1,6]\!]$, on note T_i le temps d'attente de la sortie du numéro i.

1. (a) Donner la loi de T_1 ainsi que son espérance et sa variance.
   (b) Trouver l'espérance des variables aléatoires $\mathrm{Inf}(T_1, T_2)$ et $\mathrm{Sup}(T_1, T_2)$.
2. Justifier l'existence de la covariance de T_1 et de T_2, que l'on notera $\mathrm{Cov}(T_1, T_2)$.
3. (a) Établir, pour tout $i \in [\![2,6]\!]$, la relation : $\mathbf{E}\big(T_1 \mid [X_1 = i]\big) = 7$.
   (b) Montrer que pour tout $i \in [\![3,6]\!]$, on a : $\mathbf{E}\big(T_1T_2 \mid [X_1 = i]\big) = \mathbf{E}\big((1+T_1)(1+T_2)\big)$.
   (c) Calculer $\mathbf{E}(T_1T_2)$.
   (d) En déduire $\mathrm{Cov}(T_1, T_2)$ ainsi que le coefficient de corrélation linéaire de T_1 et T_2.
4. (a) Trouver un réel α tel que les variables aléatoires T_1 et $T_2 + \alpha T_1$ soient non corrélées.
   (b) Calculer l'espérance conditionnelle $\mathbf{E}\big(T_2 + \alpha T_1 \mid [T_1 = 1]\big)$.
   (c) Les variables aléatoires T_1 et $T_2 + \alpha T_1$ sont-elles indépendantes ?

Avec un peu d'algèbre linéaire...

Exercice 235. ✦✦

Soit $(X_1, X_2, \ldots, X_n)$, un vecteur aléatoire. On définit la matrice colonne aléatoire $\mathscr{X}$ et, sous réserve d'existence, la matrice $\mathscr{E}$ des espérances et $\mathscr{V}$ des variances-covariances par

$$\mathscr{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}, \quad \mathscr{E} = \begin{bmatrix} \mathbf{E}(X_1) \\ \mathbf{E}(X_2) \\ \vdots \\ \mathbf{E}(X_n) \end{bmatrix} \quad \text{et} \quad \mathscr{V} = \begin{bmatrix} \mathbf{V}(X_1) & \mathrm{Cov}(X_1,X_2) & \cdots & \mathrm{Cov}(X_1,X_n) \\ \mathrm{Cov}(X_2,X_1) & \mathbf{V}(X_2) & \cdots & \mathrm{Cov}(X_1,X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}(X_n,X_1) & \mathrm{Cov}(X_n,X_2) & \cdots & \mathbf{V}(X_n) \end{bmatrix}.$$

1. Soient $a_1, a_2, \ldots, a_n$ des réels et A la matrice ligne $A = \begin{bmatrix} a_1 & a_2 & \ldots & a_n \end{bmatrix}$. On note Z la variable aléatoire réelle définie par
   $$Z = A\mathscr{X} = \sum_{k=1}^{n} a_i X_i.$$
   Montrer que $\mathbf{E}(Z) = A\mathscr{E}$ et $\mathbf{V}(Z) = A\mathscr{V}\,{}^t\!A$.
2. Justifier que la matrice $\mathscr{V}$ est diagonalisable et que les valeurs propres sont positives.
3. Que dire des variables aléatoires X_i si la matrice $\mathscr{V}$ n'est pas inversible ?

Exercice 236. ✦✦ **Matrice des lois conditionnées**

On dit qu'une matrice carrée à coefficients positifs M est stochastique si la somme des coefficients de chaque colonne est égale à 1. Soient X, Y deux variables aléatoires à valeurs dans $[\![1,n]\!]$. On appelle matrice des lois conditionnées de X sachant Y, la matrice carrée d'ordre n,

$$\mathrm{M} = \left(a_{ij}\right)_{i,j\in[\![1;n]\!]}, \quad \text{où} \quad a_{ij} = \mathbf{P}_{[\mathrm{Y}=j]}(\mathrm{X}=i).$$

1. Montrer que M est stochastique et que 1 est une valeur propre de M.
   Indication. 1 *est valeur propre si et seulement si* ${}^t\mathrm{M}-\mathrm{I}_n$ *n'est pas inversible.*
2. Montrer que X et Y sont indépendantes si et seulement si les colonnes de M sont proportionnelles.

Exercice 237. ✦✦✦ *d'après oral ESCP 2022*

Soient X et Y deux variables aléatoires définies sur un même espace probabilisé $(\Omega,\mathscr{A},\mathrm{P})$. On suppose que X et Y admettent des moments d'ordre 2 et on suppose que X n'a pas une variance nulle. Soit la matrice :

$$\mathrm{A} = \begin{bmatrix} \mathbf{E}\left(\mathrm{X}^2\right) & \mathbf{E}(\mathrm{X}) \\ \mathbf{E}(\mathrm{X}) & 1 \end{bmatrix}.$$

1. Pour tout $(a,b)\in\mathbb{R}^2$, montrer l'existence de l'espérance $\mathbf{E}\big((\mathrm{Y}-a\mathrm{X}-b)^2\big)$ et trouver une matrice colonne B et un nombre $\mathrm{C}\in\mathbb{R}$ (qui ne dépendent ni de a ni de b) tels que :

$$\mathbf{E}\Big((\mathrm{Y}-a\mathrm{X}-b)^2\Big) = {}^t\mathrm{UAU} - 2\,{}^t\mathrm{BU} + \mathrm{C} \quad \text{avec} \quad \mathrm{U} = \begin{bmatrix} a \\ b \end{bmatrix}.$$

   On souhaite montrer l'existence et trouver $(a,b)\in\mathbb{R}^2$ tel que $\mathbf{E}\big((\mathrm{Y}-a\mathrm{X}-b)^2\big)$ soit minimal. On pose $f(a,b) = \mathbf{E}\big((\mathrm{Y}-a\mathrm{X}-b)^2\big)$.
2. Montrer que f admet une borne inférieure sur $\mathbb{R}^2$.
3. Montrer que A est diagonalisable et que ses valeurs propres sont strictement positives.
4. En déduire l'existence d'un minimum pour f sur $\mathbb{R}^2$.
5. Trouver explicitement tous les couples (a,b) pour lesquels ce minimum est atteint.

Exercice 238. ✦✦✦ **Application de l'inversibilité de la matrice de Vandermonde** *d'après Oraux HEC 2014*

1. Soient $a_0, a_1, \ldots, a_n$, $n+1$ réels deux à deux distincts. Notons $\mathscr{B} = \left(1, x, \ldots, x^n\right)$ la base canonique de $\mathbb{R}_n[x]$ et $\mathscr{C} = (e_0,\ldots,e_n)$, la base canonique de $\mathbb{R}^{n+1}$.
   (a) Montrer que l'application suivante est un isomorphisme

$$\varphi : \left\{ \begin{array}{ccc} \mathbb{R}_n[x] & \to & \mathbb{R}^{n+1} \\ \mathrm{P} & \mapsto & \big(\mathrm{P}(a_0),\ldots,\mathrm{P}(a_n)\big). \end{array} \right.$$

   (b) Expliciter A, la matrice de φ de la base $\mathscr{B}$ à $\mathscr{C}$. Est-t-elle inversible ?
2. Soit (X, Y) un couple de variables aléatoires réelles discrètes finies définies sur $(\Omega,\mathscr{A},\mathbf{P})$, et soit n et m deux entiers supérieurs ou égaux à 2 . On suppose que $\mathrm{X}(\Omega) = \{x_1, x_2, \ldots, x_n\}$ et $\mathrm{Y}(\Omega) = \left\{y_1, y_2, \ldots, y_m\right\}$. On pose pour tout $i\in[\![1,n]\!]$ et tout $j\in[\![1,m]\!]$:

$$p_i = \mathbf{P}\left(\mathrm{X}=x_i\right), \quad q_j = \mathbf{P}\left(\mathrm{Y}=y_j\right), \quad \pi_{i,j} = \mathbf{P}\left([\mathrm{X}=x_i]\cap[\mathrm{Y}=y_j]\right) \quad \text{et} \quad \delta_{i,j} = \pi_{i,j} - p_i q_j.$$

   On suppose que pour tout $h\in[\![1;n-1]\!]$ et tout $k\in[\![1;m-1]\!]$, la covariance de X^h et Y^k est nulle.
   (a) Soit $k\in[\![1;m-1]\!]$. Montrer que pour tout $i\in[\![1;n]\!]$, on a : $\sum\limits_{j=1}^{m} \delta_{i,j}\, {y_j}^k = 0$.
   (b) En déduire que X et Y sont indépendantes.

CHAPITRE 12

Compléments sur les variables à densité

Il n'y a rien de plus triste qu'une vie sans hasard.

HONORÉ DE BALZAC (1799-1850)

1 Loi du maximum, loi du minimum

Soient X, Y des variables aléatoires *indépendantes* à densité définies sur un même espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$. Notons f_X et f_Y des densités respectivement de X et Y. On définit les applications

$$Z : \left\{ \begin{array}{ccc} \Omega & \to & \mathbb{R} \\ \omega & \mapsto & \max\big(X(\omega); Y(\omega)\big) \end{array} \right. \quad \text{et} \quad T : \left\{ \begin{array}{ccc} \Omega & \to & \mathbb{R} \\ \omega & \mapsto & \min\big(X(\omega); Y(\omega)\big). \end{array} \right.$$

Noter simplement $Z = \max(X, Y)$ et $T = \min(X, Y)$. Les applications Z et T sont des variables aléatoires.

- Pour tout $t \in \mathbb{R}$, on montre que

$$F_Z(t) = F_X(t) \cdot F_Y(t).$$

Par produit, F_Z est continue sur $\mathbb{R}$ et de classe $\mathscr{C}^1$ sur $\mathbb{R}$ sauf en un nombre fini de points (noté D). Z est une variable aléatoire à densité. Une densité est donnée par dérivation, pour $t \in \mathbb{R} \setminus D$

$$f_Z(t) = {F_Z}'(t) = {F_X}'(t) \cdot F_Y(t) + F_X(t) \cdot {F_Y}'(t) = f_X(t) F_Y(t) + f_Y(t) F_X(t).$$

La dernière égalité s'étend à tout réel t.

- Le calcul est similaire pour le minimum en rajoutant le passage au complémentaire. Pour tout $t \in \mathbb{R}$,

$$1 - F_T(t) = (1 - F_X(t)) \cdot (1 - F_Y(t)).$$

La variable T est à densité. Une densité est donnée par dérivation, pour $t \in \mathbb{R}$

$$f_T(t) = f_X(t)\big(1 - F_Y(t)\big) + f_Y(t)\big(1 - F_X(t)\big).$$

Exercice 239

✧ Soient X, Y des variables aléatoires suivant respectivement des lois exponentielles de paramètres λ et μ. Donner la loi de $T = \min(X, Y)$.

Le calcul précédent se généralise au cas de n variables aléatoires indépendantes à densité $(X_1, X_2, \ldots, X_n)$. Avec

$$Z = \max(X_1, \ldots, X_n) \quad \text{et} \quad T = \min(X_1, \ldots, X_n),$$

on a

$$F_Z = \prod_{i=1}^{n} F_{X_i} \quad \text{et} \quad 1 - F_T = \prod_{i=1}^{n} \big(1 - F_{X_i}\big).$$

Les variables Z et T restent des variables à densité.

2 Loi d'une somme de variables aléatoires à densité

2.1 Calcul d'un produit de convolution

Soient f, g deux fonctions continues sur $\mathbb{R}$ (sauf éventuellement en un nombre fini de points). Pour tout réel x, on définit, sous réserve d'existence, $f * g(x)$ par

$$f * g(x) = \int_{-\infty}^{+\infty} f(x-t)g(t)\,\mathrm{d}t.$$

Cela définit alors une nouvelle fonction $x \in \mathbb{R} \mapsto f * g(x)$. On parle alors de produit de convolution, noté $f * g$.

Exercice 240

✧ **Conditions suffisantes d'existence et propriétés**

Soient f et g, deux fonctions continues sur $\mathbb{R}$ telles que f soit bornée sur $\mathbb{R}$ et $\int_{-\infty}^{+\infty} g(t)\,\mathrm{d}t$ soit absolument convergente.

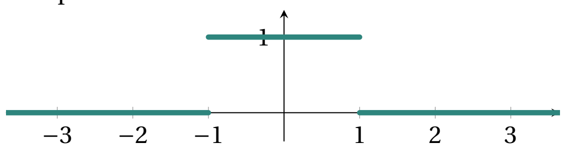
1. Montrer que $f * g$ est bien définie sur $\mathbb{R}$.
2. Vérifier que $f * g = g * f$ et pour tout $\lambda \in \mathbb{R}$, $f * (g + \lambda h) = (f * g) + \lambda(f * h)$.

Exemples.

- *Exemple 1 avec la gaussienne.*

- *Exemple 2 avec une fonction à support borné.* Soit c définie sur $\mathbb{R}$ par

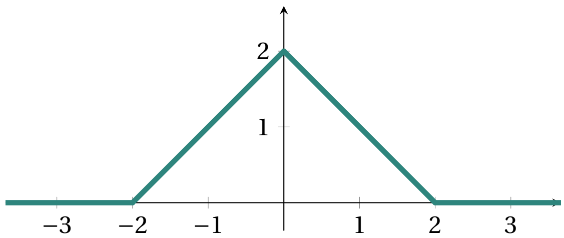
$$c(t) = \begin{cases} 1 & \text{si } t \in [-1;1] \\ 0 & \text{sinon.} \end{cases}$$



On montre que

$$c * c(x) = \begin{cases} 0 & \text{si} & x \notin [-2;2] \\ 2+x & \text{si} & x \in [-2;0] \\ 2-x & \text{si} & x \in [0;2]. \end{cases}$$

On obtient un graphe en "triangle".



Exercice 241

✦✦✦ On définit les fonctions f, g et h sur $\mathbb{R}$ par

$$f(t) = \frac{1}{1+t^2}, \qquad g(t) = \left\{ \begin{array}{ll} \mathrm{e}^{-t} & \text{si} \quad t \geqslant 0 \\ 0 & \text{sinon} \end{array} \right. \qquad \text{et} \qquad h(t) = \left\{ \begin{array}{ll} 1 & \text{si} \quad t \in [0;1] \\ 0 & \text{sinon.} \end{array} \right.$$

Calculer $f * h$, $g * g$, $g * h$.

2.2 Le théorème de sommation

Théorème 188 (loi d'une somme)

Soient X et Y deux variables aléatoires sur $(\Omega, \mathscr{A}, \mathbf{P})$, de densités respectives f_X et f_Y.

Si
- → Les variables X et Y sont indépendantes,
- → La fonction h définie sur $\mathbb{R}$ par
$$h(x) = \int_{-\infty}^{+\infty} f_X(t) f_Y(x-t)\,\mathrm{d}t$$
est bien définie et continue sauf en un nombre fini de points.

Alors la variable $X + Y$ est à densité et h est une densité.

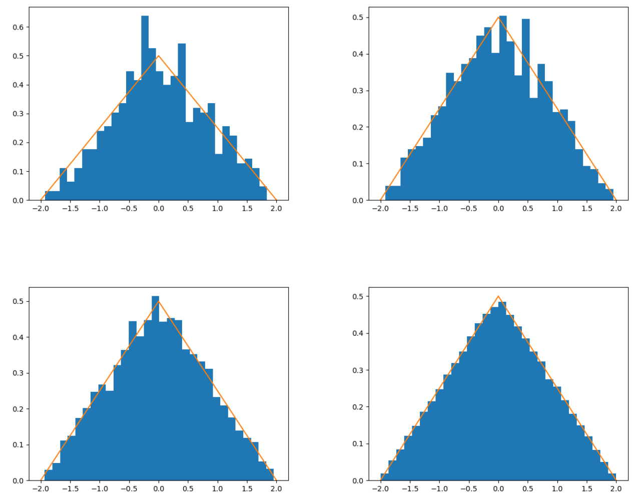
Exemple. Le deuxième calcul de l'*exemple 2* traduit le fait que la somme de deux lois uniformes sur $[-1;1]$ est encore une variable aléatoire à densité dont une densité est donnée par

$$h(x) = \begin{cases} 0 & \text{si} \quad x \notin [-2;2] \\ (2+x)/4 & \text{si} \quad x \in [-2;0] \\ (2-x)/4 & \text{si} \quad x \in [0;2]. \end{cases}$$

Vérification expérimentale avec Python

Editeur

```
m=500
 # puis, 1000, 5000 et 10^4
 # taille de l'échantillon
x=-1+2*np.random.rand(m)
y=-1+2*np.random.rand(m)
 # simulation d'une loi uniforme
 # continue sur [-1;1]

plt.hist(x+y,30,density=True)
 # tracé de l'histogramme
plt.plot([-2,0,2],[0,0.5,0])
 # tracé de la densité
plt.show()
 # affichage
```



⚠ **Attention.** La somme de deux variables aléatoires à densité n'est pas toujours une variable aléatoire à densité. Il suffit de considérer $X + (-X)$ pour s'en convaincre.

Exercice 242

✦✦ Probabilité de collision
Soient X et Y deux variables aléatoires indépendantes à densité bornée.
Que dire de la probabilité $\mathbf{P}([X = Y])$?

3 Application aux lois usuelles

3.1 Stabilité par somme des lois γ

On rappelle les définitions des fonctions Γ et β

$$\Gamma : x \in \mathbb{R}_*^+ \mapsto \int_0^{+\infty} t^{x-1} e^{-t} \, dt \quad \text{et} \quad \beta : (x,y) \in \mathbb{R}^2 \mapsto \int_0^1 t^{x-1}(1-t)^{y-1} \, dt.$$

Théorème 189 (somme de deux lois γ)

Soient $\nu_1, \nu_2 \in \mathbb{R}_+^*$ et X_1, X_2 deux variables aléatoires sur $(\Omega, \mathscr{A}, \mathbf{P})$.

Si
- X_1 et X_2 sont indépendantes.
- $X_1 \hookrightarrow \gamma(\nu_1)$ et $X_2 \hookrightarrow \gamma(\nu_2)$.

Alors $X_1 + X_2 \hookrightarrow \gamma(\nu_1 + \nu_2)$.

Remarque. Ce calcul a permis de retrouver l'égalité

$$\forall \nu_1, \nu_2 \in \mathbb{R}_*^+, \quad \beta(\nu_1, \nu_2) = \int_0^1 (1-u)^{\nu_1 - 1} u^{\nu_2 - 1} du = \frac{\Gamma(\nu_1)\Gamma(\nu_2)}{\Gamma(\nu_1 + \nu_2)}.$$

Corollaire 190 (cas particulier des lois exponentielles)

Soient $X_1, \cdots, X_n$, n variables aléatoires définies sur un même espace probabilisé.

Si
- Les variables $X_1, \cdots, X_n$ sont mutuellement indépendantes.
- Pour tout $i \in [\![1;n]\!]$, $X_i \hookrightarrow \mathscr{E}(1)$.

Alors $$X_1 + X_2 + \cdots + X_n \hookrightarrow \gamma(n).$$

Exercice 243

✦ **Cas général**
Soient $X_1, \cdots, X_n$ des variables mutuellement indépendantes suivant une loi $\mathscr{E}(\lambda)$. Donner une densité de la somme $X_1 + X_2 + \cdots + X_n$.

3.2 Stabilité par somme des lois normales

Proposition 191 (somme de deux lois normales)

Soient $m_1, m_2 \in \mathbb{R}$, $\sigma_1, \sigma_2 \in \mathbb{R}_+^*$ et $X_1, X_2 : \Omega \to \mathbb{R}$ deux variables aléatoires.

Si
- X_1 et X_2 sont indépendantes.
- $X_1 \hookrightarrow \mathcal{N}\left(m_1, {\sigma_1}^2\right)$ et $X_2 \hookrightarrow \mathcal{N}\left(m_2, {\sigma_2}^2\right)$.

Alors $$X_1 + X_2 \hookrightarrow \mathcal{N}\left(m_1 + m_2, {\sigma_1}^2 + {\sigma_2}^2\right).$$

Exercice 244

✦✦ **Preuve**
Prouver l'énoncé en adaptant le premier exemple du chapitre et en admettant l'égalité suivante :

$$-\frac{(x-t)^2}{2{\sigma_1}^2} - \frac{t^2}{2{\sigma_2}^2} = -\frac{\sigma^2}{2{\sigma_1}^2{\sigma_2}^2}\left(t - \frac{{\sigma_2}^2}{\sigma^2}x\right)^2 - \frac{x^2}{2\sigma^2} \quad \text{où} \quad \sigma = \sqrt{{\sigma_1}^2 + {\sigma_2}^2}.$$

Exercice 245

✦ On considère X, Y deux variables aléatoires indépendantes de loi respectives $\mathcal{N}\left(m_1, {\sigma_1}^2\right)$ et $\mathcal{N}\left(m_2, {\sigma_2}^2\right)$. Soient $S = X + Y$ et $T = X - Y$.

1. Déterminer la loi de S puis celle de T.
2. On suppose que S et T sont indépendantes. Donner la loi de S+T. En déduire $\sigma_1 = \sigma_2$.

Corollaire 192 (somme de n lois normales)

Soient $X_1, \cdots, X_n$, n variables aléatoires indépendantes telles que pour tout $i \in [\![1;n]\!]$, $X_i \hookrightarrow \mathcal{N}(m_i, {\sigma_i}^2)$. Alors
$$X_1 + \cdots + X_n \hookrightarrow \mathcal{N}\left(\sum_{i=1}^n m_i, \sum_{i=1}^n {\sigma_i}^2\right).$$

Exemple. En particulier, retenons que si $(X_i)_{i \in [\![1;n]\!]}$ est un vecteur aléatoire composé de variables mutuellement indépendantes et de loi identique $\mathcal{N}(\mu, \sigma^2)$, alors
$$\overline{X_n}^* = \frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}} \hookrightarrow \mathcal{N}(0;1).$$

Remarque. Nous avons vu que les lois normales sont stables par transformations affines, ainsi toute combinaison linéaire ed variables aléatoires suivant des lois normales restent une loi normale.

4 Un exemple détaillé

Soient X, Y deux variables aléatoires indépendantes suivant une loi uniforme sur $]0;1[$. On pose

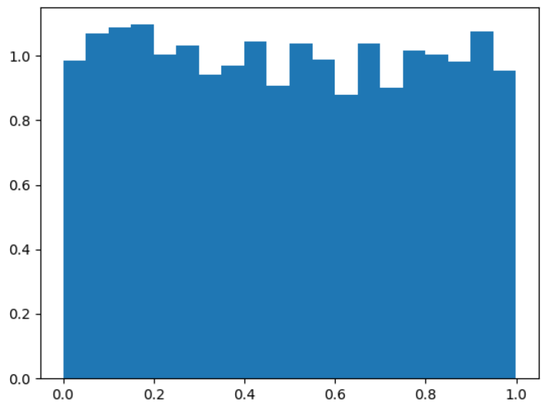
$$Z = \frac{\min(X,Y)}{\max(X,Y)}.$$

Déterminons la loi de Z.

Simulation Python et conjecture

Editeur

```
import random as rd
import matplotlib.pyplot as plt

def simulation():
    X=rd.random()
    Y=rd.random()
    Z=min(X,Y)/max(X,Y)
    return Z

Ech=[]
for i in range(5000):
    Z=simulation()
    Ech.append(Z)
plt.clf()
plt.hist(Ech, bins=20, density=
    True)
plt.show()
```



De manière assez surprenante, la loi semble uniforme sur l'intervalle $[0;1]$. Une preuve s'impose !

Exercice 246

✦✦ *Les questions sont indépendantes.*

1. Soient X, Y deux variables aléatoires indépendantes de loi $\mathscr{U}(]0;1)$. Donner la loi de
$$T = -\ln\big(\max\{X,Y\}\big).$$
2. Reprendre le problème avec trois variables aléatoires indépendantes X, Y, T suivant des lois uniformes sur $]0;1]$ et
$$Z = \frac{\min(X,Y,T)}{\max(X,Y,T)}.$$

Exercices - TD

Compléments sur les variables à densité

Exercice 247. ✦
Soient $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes de même loi normale centrée réduite et $\alpha \in \mathbb{R}\setminus\{\pm 1;0\}$. On définit

$$Y_0 = X_0 \quad \text{et} \quad \forall\, n\in\mathbb{N}, \quad Y_{n+1} = \alpha Y_n + X_{n+1}.$$

1. Calculer pour tout $n\in\mathbb{N}$, $\mathbf{E}(Y_n)$ et $\mathbf{V}(Y_n)$.
2. Justifier que pour tout $n\in\mathbb{N}^*$, la variable Y_n suit une loi normale dont on précisera les paramètres.
3. Calculer pour tout $p\in\mathbb{N}^*$, $\mathrm{Cov}\left(Y_n, Y_{n+p}\right)$.

Exercice 248. ✦✦
Soient X et Y deux variables aléatoires indépendantes sur $(\Omega,\mathscr{A},\mathbf{P})$ telles que $X\hookrightarrow\mathscr{N}\left(a,\sigma^2\right)$ et $Y\hookrightarrow\mathscr{N}\left(b,\sigma^2\right)$. Prouver l'équivalence

$$\mathbf{P}(X\leqslant Y)\geqslant\frac{1}{2} \quad\Longleftrightarrow\quad a\leqslant b.$$

Exercice 249. ✦✦
Soient f, g deux fonctions positives non nulles bornées, continues et telles que les intégrales $\int_{-\infty}^{+\infty} f(t)\mathrm{d}t$ et $\int_{-\infty}^{+\infty} g(t)\mathrm{d}t$ soient convergentes. Montrer que

$$\int_{-\infty}^{+\infty}(f*g)(t)\mathrm{d}t = \left(\int_{-\infty}^{+\infty} f(t)\mathrm{d}t\right)\left(\int_{-\infty}^{+\infty} g(t)\mathrm{d}t\right).$$

Exercice 250. ✦✦ *D'après oraux HEC 2013*
Soient X et Y deux variables aléatoires indépendantes, définies sur un espace probabilisé $(\Omega,\mathscr{A},\mathbf{P})$ et de même loi $\mathscr{N}(0,1)$. On note α, la probabilité que la matrice ci-dessous soit diagonalisable.

$$M = \begin{bmatrix} 0 & X & 0 \\ Y & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

1. Calculer $\mathbf{P}(X=Y)$ et $\mathbf{P}(XY>0)$.
2. Calculer α.

Loi d'une somme, d'une différence, d'un produit, d'un maximum, etc

Exercice 251. ✦ **Loi d'un produit d'une variable discrète et d'une continue**
Soient X et Y deux variables aléatoires indépendantes définies sur un même espace probabilisé. On suppose que X est une variable discrète qui prend les valeurs -1 et $+1$ équiprobablement, c'est-à-dire $\mathbf{P}(X=-1)=\mathbf{P}(X=1)=1/2$ et que Y suit la loi normale centrée réduite $\mathscr{N}(0,1)$. Donner la loi de $Z=XY$? Commenter.

Exercice 252. ✦ **Exemples avec les lois uniformes continues** *D'après EDHEC*
On considère deux variables aléatoires X et Y, définies sur le même espace probabilisé $(\Omega,\mathscr{A},\mathbf{P})$, indépendantes et suivant la loi uniforme sur $[0;1]$. On pose : $Z=X+Y$.

1. (a) Déterminer une densité de Z.
   (b) Montrer que, pour tout $x\in]0,1[$, les événements $[Z>1]$ et $[1-x<Z\leqslant 1+x]$ sont indépendants.
2. On pose : $T=\max(X,Y)$. On admet que T est une variable aléatoire définie sur $(\Omega,\mathscr{A},\mathbf{P})$.
   (a) Montrer que T est une variable aléatoire à densité, puis donner une densité de T.
   (b) En déduire que T possède une espérance et la déterminer.
3. On pose : $U=|X-Y|$ et on admet que U est une variable aléatoire définie sur $(\Omega,\mathscr{A},\mathbf{P})$. Montrer que U est combinaison linéaire de Z et T, puis en déduire l'espérance de U.

Exercice 253. ✦ **Loi d'une différence de lois exponentielles** *D'après Ecricome*
Soient a et b deux réels strictement positifs, X et Y deux variables aléatoires définies sur le même espace probabilisé $(\Omega,\mathscr{A},\mathbf{P})$, indépendantes et suivant chacune une loi exponentielle de paramètres respectifs a et b.

1. Déterminer la fonction de répartition, puis donner une densité de la variable $T=-X$.

2. Montrer que $\mathrm{Y}-\mathrm{X}$ admet une densité, notée h, définie par :

$$h(t)=\begin{cases}\dfrac{ab}{a+b}\mathrm{e}^{-bt}, & \text{pour } t>0\\ \dfrac{ab}{a+b}\mathrm{e}^{at}, & \text{pour } t\leqslant 0.\end{cases}$$

3. On considère la variable $\mathrm{Z}=|\mathrm{Y}-\mathrm{X}|$.
   (a) Montrer que pour tout $s\in[0,+\infty[$, $\mathbf{P}(\mathrm{Z}\leqslant s)=1-\dfrac{b\mathrm{e}^{-as}+a\mathrm{e}^{-bs}}{a+b}$.
   (b) Vérifier que Z est une variable aléatoire à densité, puis donner une densité de Z.
   (c) Montrer que Z admet une espérance et la calculer.

Exercice 254. ✦✦✦ Minimum sur un nombre aléatoire de variables *D'après Oral ESCP*

1. Soit $q\in]0,1[$ et $\mathrm{I}=\displaystyle\int_0^1-\frac{\ln(t)}{(1-qt)^2}\,\mathrm{d}t$.
   (a) Déterminer deux réels a et b tels que :
   $$\forall t\in\mathbb{R}\setminus\left\{0;\frac{1}{q}\right\},\qquad \frac{1}{t(1-qt)}=\frac{a}{t}+\frac{b}{1-qt}.$$
   (b) En déduire l'existence et la valeur de I.

   Dans la suite, on considère une suite $(\mathrm{X}_n)_{n\geqslant1}$ de variables aléatoires définies sur un espace probabilisé $(\Omega,\mathscr{A},\mathbf{P})$, indépendantes, de même loi exponentielle de paramètre 1.
2. *Cas 1.*
   Soit $n\in\mathbb{N}^*$, fixé. On pose : $\mathrm{U}_n=\min(\mathrm{X}_1,\cdots,\mathrm{X}_n)$.
   Déterminer la loi de U_n, ainsi que son espérance et sa variance.
3. *Cas 2.*
   Soit N une variable aléatoire qui suit une loi géométrique de paramètre $p\in]0,1[$. On suppose que N est indépendant des variables X_k pour tout $k\in\mathbb{N}^*$, et on pose :
   $$\mathrm{U}_\mathrm{N}=\min\left(\mathrm{X}_1,\cdots,\mathrm{X}_\mathrm{N}\right),$$
   c'est-à-dire $\mathrm{U}_\mathrm{N}(\omega)=\min\left(\mathrm{X}_1(\omega),\cdots,\mathrm{X}_{\mathrm{N}(\omega)}(\omega)\right)$ pour tout $\omega\in\Omega$. On admet que U_N est une variable aléatoire.
   Déterminer la loi de U_N, ainsi que son espérance sous réserve d'existence.
   Indication : on pourra utiliser le changement de variable $\varphi(t)=\mathrm{e}^{-t}$.

Exercice 255. ✦✦ Maximum et loi normale *D'après EDHEC 2006*

On considère deux variables aléatoires X et Y, définies sur un espace probabilisé $(\Omega,\mathscr{A},\mathbf{P})$, indépendantes et suivant toutes deux la loi normale centrée réduite (de densité notée φ et de fonction de répartition notée Φ).
On pose $\mathrm{Z}=\max(\mathrm{X},\mathrm{Y})$ et l'on se propose de déterminer la loi de Z, ainsi que son espérance et sa variance.

1. (a) Montrer que Z est une variable aléatoire à densité.
   (b) Vérifier que Z admet pour densité la fonction f définie pour tout réel x par :
   $$f(x)=2\varphi(x)\Phi(x).$$
2. (a) Rappeler la valeur de l'intégrale $\displaystyle\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t$.
   (b) En déduire la convergence et la valeur de $\displaystyle\int_{-\infty}^{+\infty}\mathrm{e}^{-t^2}\,\mathrm{d}t$.
   (c) En remarquant que, pour tout réel x, $\varphi'(x)=-x\varphi(x)$, montrer, grâce à une intégration par parties, que :
   $$\int_0^{+\infty}xf(x)\,\mathrm{d}x=\frac{1}{\sqrt{2\pi}}+\frac{1}{\pi}\int_0^{+\infty}\mathrm{e}^{-t^2}\,\mathrm{d}t.$$
   (d) Montrer de même que : $\displaystyle\int_{-\infty}^{0}xf(x)\,\mathrm{d}x=-\frac{1}{\sqrt{2\pi}}+\frac{1}{\pi}\int_{-\infty}^{0}\mathrm{e}^{-t^2}\,\mathrm{d}t$. En déduire que Z a une espérance et donner sa valeur.
3. (a) Montrer que X^2 et Z^2 suivent la même loi.
   (b) Déterminer $\mathbf{E}\left(\mathrm{Z}^2\right)$, puis donner la valeur de la variance de Z.

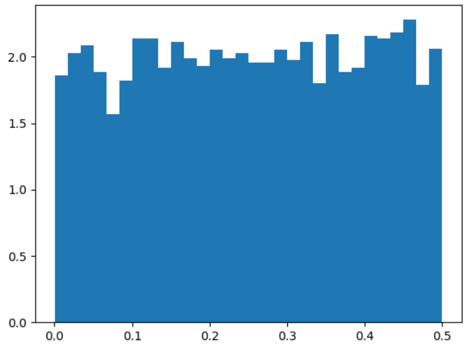
Exercice 256. ✦✦ Un exemple sans l'hypothèse d'indépendance

Soit X une variable aléatoire suivant une loi uniforme sur $[0;1]$.

1. Déterminer la loi et l'espérance de $\mathrm{Y}=\min(\mathrm{X},1-\mathrm{X})$, ainsi que de $\mathrm{Z}=\max(\mathrm{X},1-\mathrm{X})$.
2. Vérifier que vos résultats sont en accord avec les simulations numériques suivantes :

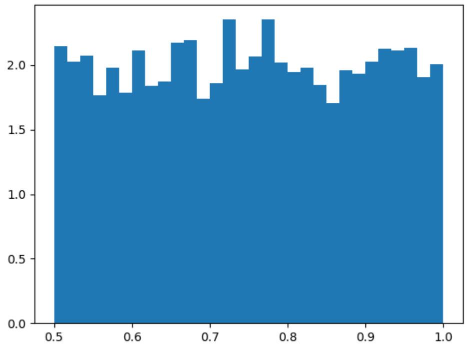
Editeur

```
ech=[]

for i in range(5000):
    x=np.random.rand()
    y=min(x,1-x)
    ech.append(y)

plt.hist(ech,30,density=True)
plt.show()
```



Editeur

```
ech=[]

for i in range(5000):
    x=np.random.rand()
    z=max(x,1-x)
    ech.append(z)

plt.hist(ech,30,density=True)
plt.show()
```



3. Calculer, si elles existent, les espérances de Y/Z et de Z/Y.

Exercice 257. ✦✦✦ **Somme de lois de Cauchy** *Oraux HEC 2008*

Les variables aléatoires considérées sont définies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$. Soit X une variable aléatoire réelle. On dit que X vérifie la propriété $(\mathscr{D})$ si, pour tout entier $n \in \mathbb{N}^*$, il existe n variables aléatoires réelles $X_{1,n}, X_{2,n}, \ldots, X_{n,n}$ mutuellement indépendantes, de même loi et dont la somme a même loi que X.

1. (a) Montrer que si X suit une loi de Poisson, alors X vérifie $(\mathscr{D})$.
   (b) Montrer qu'il en est de même si X suit une loi normale.
2. Une variable aléatoire prenant un nombre fini de valeurs vérifie-t-elle la propriété $(\mathscr{D})$?
3. Soit $(a, b) \in \mathbb{R}^2$, avec $a < b$. On considère dans cette question une variable aléatoire X à valeurs dans $[a, b]$ et vérifiant la propriété $(\mathscr{D})$.
   (a) Montrer que $\mathbf{V}\left(X_{1,n}\right) \leqslant \frac{(b-a)^2}{n^2}$.
   (b) Que peut-on en déduire sur X ?
4. (a) Déterminer le réel c tel que, pour tout $\lambda \in \mathbb{R}^*$, la fonction $x \in \mathbb{R} \longmapsto f_\lambda(x) = \dfrac{c\lambda}{\lambda^2 + x^2}$ soit une densité de probabilité sur $\mathbb{R}$. Dans toute la suite, c aura cette valeur.
   (b) Soit X_1 et X_2 deux variables aléatoires indépendantes de densité f_1. Montrer qu'une densité g de $X_1 + X_2$ vérifie :
   $$\forall x \in \mathbb{R}, \quad g(2x) = \int_{-\infty}^{+\infty} f_1(x+t) f_1(x-t)\,\mathrm{d}t$$
   (c) Soit $x \in \mathbb{R}^*$. Déterminer quatre réels $\left(a, a', b, b'\right)$ (dépendant de x) tels que
   $$\forall t \in \mathbb{R}, \quad \frac{1}{\left[1+(x-t)^2\right]\left[\left(1+(x+t)^2\right]\right.} = \frac{a(x-t)+b}{1+(x-t)^2} + \frac{a'(x+t)+b'}{1+(x+t)^2}.$$
   (d) En déduire une expression simple de g sur $\mathbb{R}^*$. On admet que g est continue sur $\mathbb{R}$. Calculer $g(0)$.
   (e) Une variable aléatoire X de densité f_λ vérifie-t-elle la propriété $(\mathscr{D})$?

Applications à la simulation des variables aléatoires

Pensez à utiliser basthon *pour tester vos codes!*

Exercice 258. ✦ **Simulation de la loi du χ^2**
Soient $n \in \mathbb{N}^*$ et $(X_i)_{i\in[\![1;n]\!]}$ un vecteur aléatoire de variables aléatoires mutuellement indépendantes et de même loi $\mathcal{N}(0;1)$. On dit que la somme $S_n = X_1 + \ldots + X_n$ suit une loi du χ^2 à n degrés de liberté.

1. Écrire un programme qui prend en argument n et simule la variable S_n.
   On pourra utiliser dans la bibliothèque ***numpy****, la commande* ***np.random.normal(mu,nu,m)*** *pour simuler m réalisation d'une variable aléatoire suivant une loi normale $\mathcal{N}(\mu, v)$.*
2. En déduire un second programme qui prend en argument n et $m \in \mathbb{N}^*$ et renvoie un échantillon de taille m de S_n.
   Tracer les histogrammes.
3. Parmi ces trois densités, une seule correspond à la loi du χ^2 de degré 4, laquelle?

$$f : x \in \mathbb{R} \mapsto \begin{cases} 0 & \text{si } x < 0 \\ \frac{1}{4}xe^{-\frac{x}{2}} & \text{si } x \geqslant 0, \end{cases} \qquad g : x \in \mathbb{R} \mapsto \frac{e^x}{(e^x+1)^2} \quad \text{et} \quad h : x \in \mathbb{R} \mapsto \begin{cases} 0 & \text{si } x < 0 \\ 4xe^{-2x} & \text{si } x \geqslant 0. \end{cases}$$

Exercice 259. ✦ **Simulation de la loi $\gamma(n)$**

1. Soit X une variable aléatoire suivant une loi uniforme sur $]0;1[$. Donner la loi de $Y = -\ln(X)$.
2. Soit $n \in \mathbb{N}^*$. En déduire un programme Python pour simuler une variable aléatoire de loi $\gamma(n)$.
3. Tester votre programme en affichant les histogrammes et en comparant avec une densité de $\gamma(n)$.

Exercice 260. ✦✦ **Simulation des lois normales par la méthode de Box-Müller**

1. *La théorie*
   Soient X et Y deux variables aléatoires définies par
   $$\begin{cases} X = \sqrt{W}\cos(\Theta) \\ Y = \sqrt{W}\sin(\Theta) \end{cases} \quad \text{où} \quad W \sim \mathcal{E}(1/2) \quad \text{et} \quad \Theta \sim \mathcal{U}_{[0;2\pi]}$$
   avec W et Θ indépendantes. On admet que les variables X et Y sont indépendantes et de même loi $\mathcal{N}(0;1)$.
   En déduire que si U et V sont deux variables aléatoires indépendantes suivant la loi uniforme sur $]0,1]$ alors
   $$\sqrt{-2\ln U}\cos(2\pi V) \quad \text{et} \quad \sqrt{-2\ln U}\sin(2\pi V)$$
   sont indépendantes et suivent une loi normale centrée réduite.
2. *La pratique*
   (a) Écrire une fonction Python **Normale** qui simule une loi normale centrée réduite.
   (b) En déduire une seconde fonction Python qui prend en argument $m \in \mathbb{N}^*$ et renvoie un échantillon de taille $2m$ de la loi normale centrée réduite avec l'histogramme de l'échantillon.
   (c) En déduire une fonction en Python qui, à partir de $(\mu,\sigma) \in \mathbb{R} \times \mathbb{R}_+^*$, simule la loi $\mathcal{N}\left(\mu,\sigma^2\right)$.

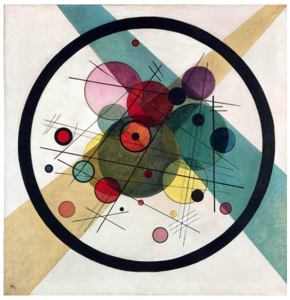
Exercice 261. ✦✦✦ **Simulation d'une loi de Poisson par des lois exponentielles**
Soit $\lambda \in \mathbb{R}_*^+$. Soient $X_1, X_2, \ldots$ des variables aléatoires indépendantes, de même loi exponentielle de paramètre 1. On note N le plus grand entier n tel que $X_1 + \cdots + X_n < \lambda$, en convenant que $N = 0$ si $X_1 > \lambda$. Autrement dit, pour $n \in \mathbb{N}$,

$$[N = n] = [X_1 + \cdots + X_n < \lambda] \cap [X_1 + \cdots + X_n + X_{n+1} \geqslant \lambda].$$

1. (a) Vérifier que pour tout $n \in \mathbb{N}$, $\mathbf{P}\big([N \geqslant n]\big) = \mathbf{P}\big([N \geqslant n+1]\big) + \frac{\lambda^n}{n!}e^{-\lambda}$.
   (b) En déduire que N suit une loi de Poisson de paramètre λ.
2. En déduire un programme qui simule une loi de Poisson de paramètre λ en utilisant uniquement la commande **random()**.
   On rappelle que si $U \hookrightarrow \mathcal{U}([0;1])$, alors $-\ln(U)/\lambda \hookrightarrow \mathcal{E}(\lambda)$.

CHAPITRE 13

Endomorphismes symétriques



Cercles dans un cercle, 1923, VASSILY KANDINSKY

1 Matrices et endomorphismes symétriques

1.1 Les définitions et exemples

Définition 193 (matrice symétrique)

On dit qu'une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est **symétrique** si ${}^tA = A$.
Autrement dit, si $(a_{i,j})_{i,j}$ sont les coefficients de la matrice A : $\forall (i,j) \in [\![1;n]\!]^2, \qquad a_{i,j} = a_{j,i}$.

Exercice 262

✦ Donner la dimension de $\mathcal{S}_n(\mathbb{R})$ défini comme le sous-espace vectoriel des matrices symétriques de $\mathcal{M}_n(\mathbb{R})$.

Définition 194 (endomorphisme symétrique)

Soient, E un espace vectoriel muni d'un produit scalaire $\langle \cdot, \cdot \rangle$ et $\varphi \in \mathcal{L}(E)$. On dit que φ est un **endomorphisme symétrique** si

$$\forall u, v \in E, \qquad \langle \varphi(u), v \rangle = \langle u, \varphi(v) \rangle.$$

Exemples.

- L'application $\varphi : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R}^2 \\ (x,y) & \mapsto & (2x-6y, -6x-7y) \end{array} \right.$ est symétrique sur $\mathbb{R}^2$ muni du produit scalaire canonique.
- Soient E, un espace euclidien de dimension $n \geqslant 2$ et $u_0 \in E \setminus \{0_E\}$. Pour tout réel $a \in \mathbb{R}^+$, on définit l'endomorphisme $\varphi_a : E \to E$ par $\varphi_a(u) = u + a\langle u, u_0 \rangle u_0$. On vérifie que φ_a est symétrique.

- On pourra consulter l'exercice 276, p. 145, pour un exemple en dimension infinie.

Exercice 263

✦✦ Soient $\left(E,\langle\cdot,\cdot\rangle\right)$ un espace euclidien, et f, g deux endomorphismes symétriques de E.

1. Justifier que si f et g commutent alors $f\circ g$ est symétrique.
2. On souhaite prouver la réciproque. On suppose donc $f\circ g$ symétrique.
   (a) Simplifier pour tous $x,\ y\in E$, $\langle x, f\circ g(y)-g\circ f(y)\rangle$.
   (b) En déduire que f et g commutent.

1.2 Premières propriétés

Proposition 195 (caractérisation via une base)

Soient $\mathscr{B}=(e_1,\dots,e_n)$ une base de E et $\varphi\in\mathscr{L}(E)$. Les deux énoncés suivants sont équivalents.

i) L'endomorphisme φ est symétrique.

ii) $\forall (i,j)\in[\![1,n]\!]^2,\qquad \left\langle \varphi\left(e_i\right),e_j\right\rangle=\left\langle e_i,\varphi\left(e_j\right)\right\rangle.$

Théorème 196 (lien avec les matrices)

Soit $\varphi\in\mathscr{L}(E)$ où $\left(E,\langle\cdot,\cdot\rangle\right)$ est un espace euclidien. Les trois énoncés suivants sont équivalents.

i) L'endomorphisme φ est un endomorphisme symétrique de E.

ii) Il existe une base orthonormée $\mathscr{B}$ de E telle que la matrice $\mathrm{Mat}_{\mathscr{B}}(\varphi)$ soit une matrice symétrique.

iii) Pour toutes les bases orthonormées $\mathscr{B}$ de E, la matrice $\mathrm{Mat}_{\mathscr{B}}(\varphi)$ est une matrice symétrique.

⚠ **Attention.** Il ne faut pas oublier que $\mathscr{B}$ doit être une base orthonormée!

Exercice 264

✦

1. Justifier que l'ensemble des endomorphismes symétriques de E est un sous-espace vectoriel de $\mathscr{L}(E)$.
2. Si E est de dimension finie, pouvez-vous préciser sa dimension?

2 Réduction

2.1 Diagonalisation des endomorphismes symétriques

Premières propriétés

Proposition 197 (espace stable)

Soient φ un endomorphisme symétrique d'un espace euclidien $\left(E,\langle\cdot,\cdot\rangle\right)$ et F un sous-espace vectoriel de E.

Si F est stable par φ,

alors $F^{\perp}$ est également stable par φ.

Exercice 265

✦ Prouver cette proposition.

Proposition 198 (vecteurs propres orthogonaux)

Soit φ un endomorphisme symétrique d'un espace euclidien $\left(\mathrm{E}, \langle\cdot,\cdot\rangle\right)$.

Si u et v sont deux vecteurs propres de φ associés à des valeurs propres distinctes,

alors les vecteurs u et v sont orthogonaux.

Remarque. On a la généralisation suivante. Si $e_1, \ldots, e_p$ sont des vecteurs propres de f associés à des valeurs propres deux à deux distinctes, alors la famille $\left(e_1, \ldots, e_p\right)$ est orthogonale.

Corollaire 199 (espaces propres orthogonaux)

Soit φ un endomorphisme symétrique d'un espace euclidien $\left(\mathrm{E}, \langle\cdot,\cdot\rangle\right)$.
Alors les sous-espaces propres de φ sont deux à deux orthogonaux.

Exemple. On considère $\mathscr{M}_n(\mathbb{R})$ muni du produit scalaire $\langle \mathrm{A}, \mathrm{B}\rangle = \mathrm{Tr}({}^t\mathrm{AB})$. On vérifie que $\varphi : \mathrm{M} \in \mathscr{M}_n(\mathbb{R}) \mapsto {}^t\mathrm{M} \in \mathscr{M}_n(\mathbb{R})$ est un endomorphisme symétrique. φ possède deux valeurs propres : -1 et 1 où $\mathrm{E}_1(\varphi)$, $\mathrm{E}_{-1}(\varphi)$ désigne respectivement l'ensemble des matrices symétriques et antisymétriques. Ces sous-espaces sont donc orthogonaux pour ce produit scalaire.

Le théorème spectral

Théorème 200 (spectral)

Soit φ un endomorphisme d'un espace euclidien $\left(\mathrm{E}, \langle\cdot,\cdot\rangle\right)$.

Si φ est symétrique,

alors
- L'endomorphisme φ est diagonalisable, les valeurs propres sont réelles.
- Il existe une base orthonormée de E formée de vecteurs propres de φ.

⚠ **Attention.** Il ne faut pas oublier que la base des vecteurs propres peut être choisie orthonormée.

Remarque. La réciproque est vraie mais elle est beaucoup moins utile.

Exercice 266

✦✦ *Les questions sont indépendantes.*

Soit φ un endomorphisme symétrique d'un espace euclidien $\left(\mathrm{E}, \langle\cdot,\cdot\rangle\right)$.

1. Que dire de φ si pour tout $u \in \mathrm{E}$, $\langle u, \varphi(u)\rangle = 0$?
2. Justifier que $\mathrm{Sp}(\varphi) \subset \mathbb{R}^+$ si et seulement si φ vérifie

$$\forall u \in \mathrm{E}, \quad \langle u, \varphi(u)\rangle \geqslant 0 \quad (\bullet)$$

2.2 Diagonalisation des matrices symétriques réelles

Théorème spectral dans le cas matriciel

Théorème 201 (spectral, version matricielle)

Soit $A \in \mathcal{M}_n(\mathbb{R})$.

Si A est symétrique,

alors
- A est diagonalisable, les valeurs propres sont réelles.
- Il existe une matrice orthogonale P et une matrice diagonale réelle D telles que

$$A = PDP^{-1} = PD\,{}^tP.$$

Remarque. Les colonnes de la matrice P forment une b.o.n de vecteurs propres de A. Pour rappel, une matrice est orthogonale si et seulement si les matrices colonnes forment une base orthonormée pour le produit scalaire canonique de $\mathcal{M}_{n,1}(\mathbb{R})$.

Exercice 267

Les questions sont indépendantes.

1. ✦ On considère la matrice
$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$
Justifier que A est diagonalisable. Calculer A^2. En déduire que $\mathrm{Sp}(A) = \{-1; 1\}$.

2. ✦✦ Soit $M \in \mathcal{M}_n(\mathbb{R})$ telle que $M + {}^tM$ soit nilpotente.
Montrer que la matrice M est antisymétrique.

Proposition 202 (décomposition d'une matrice symétrique)

Soit A une matrice symétrique de $\mathcal{M}_n(\mathbb{R})$.

Notons
- $(\lambda_1, \dots, \lambda_n)$ les valeurs propres de A.
- $(X_1, \dots, X_n)$ une b.o.n de vecteurs propres de A telle que $AX_i = \lambda_i X_i$ pour tout $i \in [\![1; n]\!]$.

Alors
$$A = \sum_{i=1}^{n} \lambda_i X_i {}^tX_i = \lambda_1 X_1 {}^tX_1 + \cdots + \lambda_n X_n {}^tX_n.$$

Remarque. En particulier, A est combinaison linéaire de n matrices de projecteurs de rang 1.

Exercice 268

✦ Soit $A = (a_{i,j})_{i,j}$ une matrice symétrique réelle, et soient $\lambda_1, \dots, \lambda_n$ ses valeurs propres.
Grâce au calcul de $\mathrm{Tr}\left({}^tAA\right)$, démontrer que

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j}{}^2 = \sum_{i=1}^{n} \lambda_i{}^2.$$

Pratique de la réduction des matrices symétriques

Méthode

Comment obtenir une b.o.n de vecteurs propres d'une matrice/endomorphisme symétrique ?

- Déterminer les valeurs propres.
  (Par un calcul du rang, un polynôme annulateur, le déterminant ...)
- Pour chaque valeur propre, déterminer une base de vecteurs propres.
- À l'aide du procédé d'orthonormalisation de Schmidt, déterminer une base orthonormée pour chacun des sous-espaces propres.
- On obtient une base de E par concaténation des bases orthonormées des sous-espaces propres.

Exercice 269

✦ Donner une b.o.n de vecteurs propres pour la matrice

$$A = \begin{bmatrix} 2 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & 2 \end{bmatrix}.$$

Exercice 270

✦ On considère la matrice carrée d'ordre 3 : *d'après EMLyon 2007 E*

$$A = \frac{1}{2}\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

1. Montrer, sans calcul, que A est diagonalisable.
2. Déterminer une matrice diagonale D et une matrice inversible et symétrique P, de première ligne $\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$ et de deuxième ligne $\begin{bmatrix} 1 & -1 & 0 \end{bmatrix}$ telles que

$$A = PDP^{-1}.$$

3 Formes quadratiques associées à une matrice

Définitions

Définition 203 (forme quadratique d'une matrice symétrique)

Soit $A \in \mathscr{M}_n(\mathbb{R})$, symétrique. La **forme quadratique associée à** A est l'application définie sur $\mathbb{R}^n$ par

$$q(h) = {}^t HAH$$

où H est la matrice des coordonnées de h dans la base canonique de $\mathbb{R}^n$.

Remarque. On constate que pour $A = \left(a_{ij}\right)_{(i,j)\in[\![1;n]\!]^2}$ et $h = (h_i)_{i\in[\![1;n]\!]}$

$$q(h) = \sum_{i,j\in[\![1;n]\!]} a_{ij} h_i h_j.$$

Par symétrie de A, on peut réécrire cette expression

$$q(h) = \sum_{i=1}^{n} a_{ii} {h_i}^2 + 2\sum_{i<j} a_{ij} h_i h_j.$$

En particulier, si A est diagonale avec $A = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, on a simplement

$$q(h) = \sum_{i=1}^{n} \lambda_i {h_i}^2.$$

Exercice 271

✦ **Forme quadratique associé à un endomorphisme symétrique**
Soient φ un endomorphisme symétrique de $\mathbb{R}^n$ et A la matrice de φ dans la base canonique.
Justifier que si q est la forme quadratique associée à A alors

$$\forall\, h \in \mathbb{R}^n, \quad q(h) = \left\langle h, \varphi(h) \right\rangle$$

où $\langle \cdot, \cdot \rangle$ est le produit scalaire canonique sur $\mathbb{R}^n$.

Expression dans une b.o.n

Théorème 204 (expression dans une b.o.n)

Soit q, une forme quadratique associée à une matrice symétrique A. Alors il existe une base orthonormée $\mathscr{B}$ de $\mathbb{R}^n$ telle que si h a pour coordonnées $\tilde{h_1}, \dots, \tilde{h_n}$ dans $\mathscr{B}$, on a

$$q(h) = \sum_{i=1}^{n} \lambda_i \tilde{h_i}^2,$$

où $\lambda_1, \dots, \lambda_n$ sont les valeurs propres de A.

Exercice 272

✦ **Un classique : l'encadrement de Rayleigh**
Si on pose $\alpha = \min \mathrm{Sp}(A)$ et $\beta = \max \mathrm{Sp}(A)$, montrer que

$$\forall\, h \in \mathbb{R}^n \setminus \{0\}, \qquad \alpha \leqslant \frac{q(h)}{\|h\|^2} \leqslant \beta.$$

Signe d'une forme quadratique

Exercice 273

✧ À quelles conditions nécessaires et suffisantes sur le spectre de A, a-t-on

1. $\forall\, u \in E, \qquad q(u) \geqslant 0$?
2. $\forall\, u \in E, \qquad q(u) \leqslant 0$?
3. $\forall\, u \in E \setminus \{0_E\}, \qquad q(u) > 0$?
4. $\forall\, u \in E \setminus \{0_E\}, \qquad q(u) < 0$?

Exercices - TD

Endomorphismes symétriques

Exercice 274. ✦
Soient $\left(E,\langle\cdot,\cdot\rangle\right)$ un espace euclidien et φ un endomorphisme symétrique de E. Démontrer que $\ker(\varphi)$ et $\mathrm{Im}(\varphi)$ sont supplémentaires orthogonaux.

Exercice 275. ✦ **La symétrie implique la linéarité**
Soit $\varphi : E \to E$ tel que, pour tous $u, v \in E$, on a $\langle\varphi(u), v\rangle = \langle u, \varphi(v)\rangle$. Justifier que φ est une application linéaire.

Exercice 276. ✦ **Exemple d'endomorphisme symétrique en dimension infinie** *d'après EMLyon 2011*
On note $E = \mathscr{C}^\infty([0;1];\mathbb{R})$, muni du produit scalaire $\langle\cdot,\cdot\rangle$ défini par :

$$\forall f, g \in E, \qquad \langle f, g\rangle = \int_0^1 f(x)g(x)\mathrm{d}x$$

et, pour toute fonction $f \in E$, on pose

$$T(f) : \left\{ \begin{array}{ccc} [0;1] & \to & \mathbb{R} \\ x & \mapsto & \left(x^2 - x\right) f''(x) + (2x-1)f'(x). \end{array} \right.$$

Montrer que T est un endomorphisme symétrique de E.

Exercice 277. ✦ **Endomorphisme symétrique et produit scalaire** *d'après EDHEC 2015*
On considère l'espace euclidien $\mathbb{R}^n$ muni du produit scalaire canonique. Pour tout couple (x, y) de $\mathbb{R}^n \times \mathbb{R}^n$, on note $\langle x, y\rangle$ le produit scalaire canonique de x et y. On note $\mathscr{B} = (e_1, e_2, \ldots, e_n)$ la base canonique de $\mathbb{R}^n$ et on rappelle que $\mathscr{B}$ est orthonormale pour le produit scalaire $\langle\cdot,\cdot\rangle$. On considère un endomorphisme f de $\mathbb{R}^n$, symétrique, dont les valeurs propres sont toutes strictement positives.

1. Justifier l'existence d'une base orthonormale de $\mathbb{R}^n$, $\mathscr{B}' = (u_1, u_2, \ldots, u_n)$, formée de vecteurs propres de f.
2. (a) Montrer que, pour tout x de $\mathbb{R}^n$, on a : $\langle x, f(x)\rangle \geqslant 0$.
   (b) Vérifier que l'égalité $\langle x, f(x)\rangle = 0$ a lieu si et seulement si $x = 0$.
   (c) En déduire que l'application φ, de $\mathbb{R}^n \times \mathbb{R}^n$ dans $\mathbb{R}$, définie par $\varphi(x, y) = \langle x, f(y)\rangle$, est un produit scalaire sur $\mathbb{R}^n$.
3. (a) En utilisant $\mathscr{B}'$, montrer qu'il existe un endomorphisme g de $\mathbb{R}^n$, symétrique pour le produit scalaire canonique, dont les valeurs propres sont strictement positives, et tel que $g^2 = f$.
   (b) Établir que g est bijectif.
   (c) Montrer que la famille $\left(g^{-1}(e_1), g^{-1}(e_2), \ldots, g^{-1}(e_n)\right)$ est une base orthonormée de $\mathbb{R}^n$ pour le produit scalaire φ.

Exercice 278. ✦✦
Soient $\left(E,\langle\cdot,\cdot\rangle\right)$ un espace euclidien de dimension n, et f un endomorphisme symétrique de E. En notant $\lambda_1, \ldots, \lambda_p$ ses valeurs propres telles que $\lambda_1 \leqslant \cdots \leqslant \lambda_p$, montrer que

$$\forall x \in E, \qquad \lambda_1 \|x\|^2 \leqslant \langle f(x), x\rangle \leqslant \lambda_p \|x\|^2.$$

Exercice 279. ✦✦✦ *Oraux HEC 2009*
Soient $\left(E,\langle\cdot,\cdot\rangle\right)$ un espace euclidien de dimension n et f, g deux endomorphismes de E symétriques et ayant des valeurs propres strictement positives.

1. Prouver qu'il existe un endomorphisme φ de E ayant des valeurs propres positives tel que $f = \varphi^2 = (\varphi \circ \varphi)$.
2. Montrer que : $\mathrm{Ker}(f + g) = \mathrm{Ker}\, f \cap \mathrm{Ker}\, g$.

Endomorphismes particuliers d'un espace euclidien

Exercice 280. ✦ **Et le cas antisymétrique ?**
Soient $n \in \mathbb{N}^*$ et $\mathbb{R}^n$ est muni de son produit scalaire usuel. Soit A une matrice de $\mathscr{M}_n(\mathbb{R})$ antisymétrique.

1. Montrer que la seule valeur propre réelle possible de A est 0.
2. Montrer que $B = A^2$ est une matrice symétrique réelle.
3. Quel est le signe des valeurs propres non nulles de B ?

Exercice 281. ✦ **Un exemple d'endomorphisme antisymétrique**
Soit n un entier au moins égal à 3. On travaille dans l'espace $\mathrm{E} = \mathbb{R}^n$ muni de son produit scalaire canonique. On considère deux vecteurs a et b de $\mathbb{R}^n$ de norme 1 et orthogonaux. On définit sur E l'application f par :

$$\forall x \in \mathrm{E}, \quad f(x) = \langle a, x\rangle b - \langle b, x\rangle a.$$

1. Vérifier que f est un endomorphisme de E.
2. (a) Déterminer $\ker f$ et une base de $\operatorname{Im} f$.
   (b) Vérifier que $\operatorname{Ker} f$ et $\operatorname{Im} f$ sont supplémentaires.
3. Montrer que :
$$\forall (x, y) \in \mathrm{E}^2, \quad \langle f(x), y\rangle = -\langle x, f(y)\rangle$$
4. En déduire que $f \circ f$ est un endomorphisme symétrique.

Exercice 282. ✦✦ Soit E un espace euclidien de dimension n. Soient f et g deux endomorphismes de E tels que $f \circ g = g \circ f$. On note S (resp. T) la matrice de f (resp. g) dans une base orthonormée $\mathscr{B}$ de E. On suppose que S est symétrique et T antisymétrique. Montrer que :

$$\forall\, x \in \mathrm{E}, \qquad \left\|(f-g)(x)\right\| = \left\|(f+g)(x)\right\|.$$

Exercice 283. ✦ **Adjoint u^* d'un endomorphisme u et endomorphismes normaux** *d'après EDHEC 2019*
Soient $n \in \mathbb{N}^*$ et $\left(\mathrm{E}, \langle\cdot,\cdot\rangle\right)$ un espace euclidien de dimension n. On note $\mathscr{B} = (e_1, e_2, \ldots, e_n)$ une base orthonormale de E.

Partie A. *Définition de l'adjoint d'un endomorphisme de* E
Dans toute cette partie, u désigne un endomorphisme de E. On se propose de montrer qu'il existe un unique endomorphisme de E, noté u^*, qui à tout vecteur y de E associe le vecteur $u^*(y)$ vérifiant :

$$\forall x \in \mathrm{E}, \quad \langle u(x), y\rangle = \left\langle x, u^*(y)\right\rangle.$$

1. Montrer que si u^* existe, alors on a, pour tout y de E :
$$u^*(y) = \sum_{i=1}^{n} \left\langle u\left(e_i\right), y\right\rangle e_i.$$
   En déduire que si u^* existe, alors u^* est unique.
2. Vérifier que l'application u^* définie par l'égalité établie à la question 1 est effectivement un endomorphisme de E.
3. Conclure que cette application est solution du problème posé, c'est-à-dire que c'est l'unique endomorphisme de E, appelé adjoint de u, vérifiant :
$$\forall (x, y) \in \mathrm{E}^2, \quad \langle u(x), y\rangle = \left\langle x, u^*(y)\right\rangle.$$

Partie B. *Étude des endomorphismes normaux*
On dit que u est un endomorphisme normal quand on a l'égalité :

$$u \circ u^* = u^* \circ u.$$

3. Soit f un endomorphisme symétrique de E. Donner son adjoint et vérifier que f est normal.

   Dans la suite, u désigne un endomorphisme normal.
4. (a) Montrer que : $\forall x \in \mathrm{E}$, $\|u(x)\| = \left\|u^*(x)\right\|$.
   (b) En déduire que $\operatorname{Ker}(u) = \operatorname{Ker}\left(u^*\right)$.
5. Montrer que si F est un sous-espace vectoriel de E stable par u, alors $\mathrm{F}^\perp$ est stable par u^*.
6. On suppose que u possède une valeur propre λ et on note E_λ le sous espace propre associé.
   (a) Montrer que E_λ est stable par u^*.
   (b) Établir que $\left(u^*\right)^* = u$ puis en déduire que $\mathrm{E}_\lambda^\perp$ est stable par u.

≫ *Pour aller plus loin, HEC 2019 Maths I, Essec 2014*

Compléments

Exercice 284. ✦✦✦ **Une descente de gradient** *D'après oral ESCP 2012*
Soit $n \in \mathbb{N} \setminus \{0; 1\}$. On considère $\mathbb{R}^n$ muni de son produit scalaire canonique noté $\langle\cdot,\cdot\rangle$, $\|\cdot\|$, la norme associée, et $\mathrm{A} \in \mathscr{M}_n(\mathbb{R})$, symétrique réelle dont les valeurs propres sont toutes strictement positives.
On confond vecteur de $\mathbb{R}^n$ et matrice colonne canoniquement associée et on pose, pour tout $\mathrm{X} \in \mathbb{R}^n$,

$$\Phi(\mathrm{X}) = {}^t\mathrm{X}\mathrm{A}\mathrm{X}.$$

1. Soit B un élément de $\mathbb{R}^n$. Montrer que l'équation $AX = B$ d'inconnue $X \in \mathbb{R}^n$ admet une unique solution qu'on notera R.
2. Montrer qu'il existe deux réels α et β strictement positifs tels que pour tout X de $\mathbb{R}^n$

$$\alpha\|X\|^2 \leqslant \Phi(X) \leqslant \beta\|X\|^2.$$

3. Dans la suite de l'exercice, on pose pour $X \in \mathbb{R}^n$: $F(X) = \Phi(X) - 2^tBX$.
   (a) Déterminer le gradient ∇F_X de F en X.
   (b) Soient X et H deux éléments de $\mathbb{R}^n$. Montrer que

$$F(X+H) = F(X) + \langle \nabla F_X, H\rangle + \Phi(H).$$

   (c) En déduire que F possède un minimum sur $\mathbb{R}^n$. En quel point est-il atteint ?
4. Soit $X \in \mathbb{R}^n$ fixé, $X \neq 0$. Déterminer $\alpha \in \mathbb{R}$ de façon à ce que $F\left(X - \alpha\nabla F_X\right)$ soit minimal. Calculer ce minimum.
5. Soit $X_0 \in \mathbb{R}^n$. On définit une suite $\left(X_k\right)_{k\in\mathbb{N}}$ de vecteurs de $\mathbb{R}^n$ par,

$$\text{pour tout } k \in \mathbb{N} : X_{k+1} = X_k - \alpha_k \nabla F_{X_k}, \text{ où } \alpha_k = \frac{\left\|\nabla F_{X_k}\right\|^2}{2\Phi\left(X_k\right)} \text{ si } X_k \neq R \text{ et 0 sinon.}$$

   (a) Montrer que la suite $\left(F\left(X_k\right)\right)_{k\in\mathbb{N}}$ converge.
   (b) Exprimer $F\left(X_{k+1}\right) - F\left(X_k\right)$ en fonction de α_k et de ∇F_{X_k}.
6. Une suite $\left(Y_k\right)_{k\in\mathbb{N}}$ de vecteurs de $\mathbb{R}^n$ sera dite convergente vers un vecteur $Z \in \mathbb{R}^n$ si $\lim\limits_{k\to+\infty} \left\|Y_k - Z\right\| = 0$, ce qui revient à dire que les coordonnées de Y_k convergent vers les coordonnées correspondantes de Z.
   (a) Montrer que la suite $\left(\nabla F_{X_k}\right)_{k\in\mathbb{N}}$ converge vers 0.
   (b) En déduire la limite de la suite $\left(X_k\right)_{k\in\mathbb{N}}$.

Exercice 285. ✦✦✦ **Matrices symétriques positives et définies positives** *d'après l'écrit HEC 2006*

Soient $n \in \mathbb{N}^*$, $S = \left(a_{ij}\right)_{1\leqslant i,j\leqslant n} \in \mathscr{M}_n(\mathbb{R})$ symétrique telle que :

$$\forall (i,j) \in [\![1;n]\!]^2, \quad a_{ij} > 0.$$

On note β la plus grande valeur propre de S et V le sous-espace propre de S associé à β. On munit $\mathscr{M}_{n,1}(\mathbb{R})$ de son produit scalaire canonique et de la norme associée $\|\cdot\|$.

1. Soit $X_0 = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in V\setminus\{0\}$. On note $|X_0| = \begin{bmatrix} |x_1| \\ \vdots \\ |x_n| \end{bmatrix}$.
   (a) Montrer ${}^tX_0SX_0 \leqslant {}^t|X_0|S|X_0|$ et en déduire : $|X_0| \in V$.
   (b) Montrer que les coordonnées de $S|X_0|$ sont toutes strictement positives et en déduire que X_0 n'a aucune coordonnée nulle.
   (c) Montrer : ${}^tX_0SX_0 = {}^t|X_0|S|X_0|$ et en déduire que les coordonnées de X_0 sont toutes de même signe.
2. (a) En déduire qu'il n'existe pas deux vecteurs de $V\setminus\{0\}$ orthogonaux entre eux.
   (b) Conclure : $\dim(V) = 1$.

CHAPITRE 14

Projections orthogonales

L'art des mathématiques consiste à trouver le cas particulier qui contient tous les germes de la généralité.

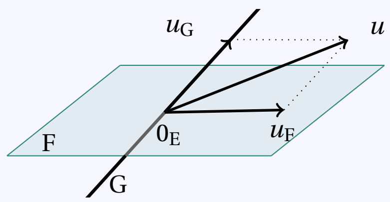
DAVID HILBERT
Mathématicien allemand (1862-1943)

1 Rappels

1.1 Les projecteurs

Définition 205 (projecteur)

Soient E un espace vectoriel et F, G deux sous-espaces vectoriels supplémentaires.



Ainsi, pour tout $u \in E$, il existe une unique décomposition $u = u_F + u_G$ où $(u_F, u_G) \in F \times G$. On pose

$$p : \left\{ \begin{array}{ccc} E & \longrightarrow & E \\ u & \longmapsto & u_F. \end{array} \right.$$

Cette application est linéaire, elle est appelée le **projecteur** sur F parallèlement à G.

Remarque. Rappelons que $F = \mathrm{Im}(p) = \mathrm{Ker}\left(p - \mathrm{Id}_E\right)$ et $G = \mathrm{Ker}(p)$. En particulier, on a

$$E = \mathrm{Im}(p) \oplus \mathrm{Ker}(p).$$

Si $\mathscr{B}$ est une base adaptée à cette décomposition en supplémentaire, on obtient une base de diagonalisation avec

$$\begin{bmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}.$$
$$\underbrace{\qquad\qquad\quad}_{\mathrm{rg}\,A}\underbrace{\qquad\qquad\quad}_{\dim\ker A}$$

De plus, $\mathrm{id}_E - p$ est le projecteur sur G parallèlement à F.

Théorème 206 (caractérisation d'un projecteur)

Soit $p : \mathrm{E} \to \mathrm{E}$ une application. Les propriétés suivantes sont équivalentes.

i) L'application p est un projecteur.

ii) L'application p est linéaire et $p \circ p = p$.

1.2 Rappels sur les sous-espaces orthogonaux

Soit F un sous-espace vectoriel de E. On appelle orthogonal de F, et on note $\mathrm{F}^{\perp}$, l'ensemble des vecteurs orthogonaux à F, c'est-à-dire :

$$\mathrm{F}^{\perp} = \left\{u \in \mathrm{E},\ \forall\, v \in \mathrm{F},\ \langle u, v\rangle = 0\right\}.$$

Proposition 207 (espaces supplémentaires orthogonaux)

Soit F, un sous-espace vectoriel d'un espace euclidien $\left(\mathrm{E}, \langle\cdot,\cdot\rangle\right)$. Alors

$$\mathrm{E} = \mathrm{F} \oplus \mathrm{F}^{\perp}.$$

Exercice 286

✦ Soient $\left(\mathrm{E}, \langle\cdot,\cdot\rangle\right)$ un espace euclidien et F, G deux sous-espaces vectoriels.
Justifier que F et G sont supplémentaires si et seulement si $\mathrm{F}^{\perp}$ et $\mathrm{G}^{\perp}$ sont supplémentaires.

Exercice 287

✦ *Les questions sont indépendantes.*

1. Dans $\mathbb{R}^3$ muni du produit scalaire canonique, on considère les plans F et G d'équations respectives :
$$x + 2y + 3z = 0 \quad \text{et} \quad x - y - z = 0.$$
Déterminer $\mathrm{F}^{\perp}$ puis $(\mathrm{F} \cap \mathrm{G})^{\perp}$.
2. Soit $\mathbb{R}_3[x]$ muni du produit scalaire $\langle \mathrm{P}, \mathrm{Q}\rangle = \displaystyle\int_0^1 \mathrm{P}(t)\mathrm{Q}(t)\,\mathrm{d}t$. On pose $\mathrm{F} = \mathbb{R}_1[x]$. Déterminer une base de $\mathrm{F}^{\perp}$.

Proposition 208 (condition suffisante d'appartenance à l'orthogonal)

Soit $\mathrm{F} = \mathrm{Vect}\,(e_1, \ldots, e_k)$ un sous-espace vectoriel de E. On a

$$u \in \mathrm{F}^{\perp} \iff \forall\, i \in [\![1, k]\!], \qquad \langle u, e_i\rangle = 0.$$

Exercice 288

✦✦ **Vecteur normal à un hyperplan**
Soit F un hyperplan d'un espace euclidien $\left(\mathrm{E}, \langle\cdot,\cdot\rangle\right)$.

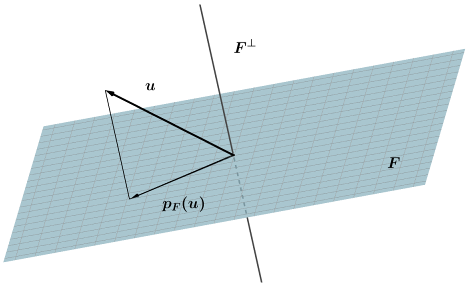
1. Montrer qu'il existe $u_0 \in \mathrm{E}$ tel que pour tout $v \in \mathrm{E}$: $\quad v \in \mathrm{F} \iff \langle u_0, v\rangle = 0$.
2. *Exemples*
   (a) On considère $\mathrm{E} = \mathbb{R}^3$ muni du produit scalaire canonique et le plan $\mathrm{F} = \left\{(x, y, z) \in \mathbb{R}^3, 2x + 3y - z = 0\right\}$. Déterminer un vecteur normal à F.
   (b) On considère maintenant $\mathrm{E} = \mathbb{R}_3[x]$ et le produit scalaire défini par $\langle \mathrm{P}, \mathrm{Q}\rangle = \displaystyle\int_{-1}^1 \mathrm{P}(t)\mathrm{Q}(t)\,\mathrm{d}t$. Déterminer un vecteur normal à $\mathbb{R}_2[x]$.

2 Projecteurs orthogonaux

2.1 Définitions et exemples

Définition 209 (projecteur orthogonal)

Soit $\left(E, \langle \cdot, \cdot \rangle\right)$, un espace euclidien.
On appelle **projection orthogonale** sur F, notée p_F, la projection sur F parallèlement à $F^{\perp}$.
Pour tout $u \in E$, $p_F(u)$ est appelé le **projeté orthogonal de** u **sur** F.



Exercice 289

✦ Soit p un projecteur orthogonal. Justifier que pour tout $u \in E$

$$\left\langle p(u), u \right\rangle = \| p(u) \|^2.$$

Remarque. D'après le rappel du début de chapitre, un projecteur p est orthogonal si et seulement si $\ker(p)$ et $\operatorname{Im}(p)$ sont supplémentaires orthogonaux, si et seulement si $E_0(p)$ et $E_1(p)$ sont des supplémentaires orthogonaux.

Exercice 290

✦ **Exemple**
Soit $\mathscr{M}_n(\mathbb{R})$ muni du produit scalaire $\langle A, B\rangle = \operatorname{Tr}\left({}^tAB\right)$. On définit l'endomorphisme p de $\mathscr{M}_n(\mathbb{R})$ par

$$p(M) = \frac{M + {}^tM}{2}.$$

Préciser le noyau et l'image de p et vérifier que p est un projecteur orthogonal.

Proposition 210 (caractérisation)

Soit p, un projecteur d'un espace euclidien $\left(E, \langle \cdot, \cdot \rangle\right)$. On a l'équivalence entre les énoncés suivants.

i) Le projecteur p est orthogonal.

ii) L'endomorphisme p est symétrique.

⚠ **Attention.** Il ne faut pas oublier la condition de base orthonormée.

Proposition 211 (caractérisation matricielle)

Soient E un espace euclidien, $\mathscr{B}$ une base orthonormée de E, p un endomorphisme et $A = \operatorname{Mat}_{\mathscr{B}}(p)$.
On a l'équivalence entre :

i) L'endomorphisme p est un projecteur orthogonal.

ii) La matrice A est symétrique et $A^2 = A$.

2.2 Expression et calcul explicite du projeté

Théorème 212 (expression du projeté)

Soit F, un sous-espace vectoriel d'un espace euclidien $\left(\mathrm{E}, \langle\cdot,\cdot\rangle\right)$. Si $\mathscr{B}_{\mathrm{F}} = \left(e_1, \ldots, e_p\right)$ est une base orthonormée de F, alors

$$\forall\, u \in \mathrm{E}, \qquad p_{\mathrm{F}}(u) = \sum_{i=1}^{p} \langle u, e_i\rangle\, e_i.$$

Exemple. Considérons le cas où F est une droite vectorielle. Il existe donc $e \in \mathrm{E} \setminus \{0_{\mathrm{E}}\}$ tel que

$$\mathrm{F} = \mathrm{Vect}(e).$$

$(e_1) = (e/\|e\|)$ est une base orthonormée de F et d'après ce qui précède

$$\forall\, u \in \mathrm{E}, \quad p_{\mathrm{F}}(u) = \langle u, e_1\rangle\, e_1 = \frac{\langle u, e\rangle}{\|e\|^2}\, e.$$

Remarque. Retour sur le procédé d'orthonormalisation de Schmidt en terme de projecteurs orthogonaux.

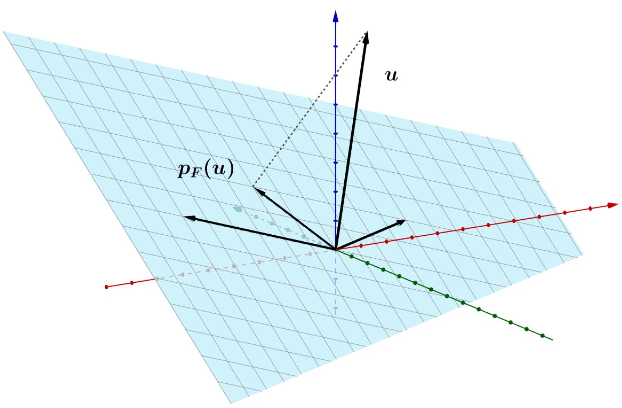
En pratique, on préféra toutefois la méthode suivante qui ne nécessite pas le calcul en amont d'une base orthonormée.

Méthode

Comment calculer un projeté ?

Soient $u \in \mathrm{E}$ et F, un sous-espace vectoriel de E. Calculons $p_{\mathrm{F}}(u)$, le projeté orthogonal de u sur F.

- *Étape 1*
  On trouve une base $\left(u_1, \ldots, u_p\right)$ de F (non nécessairement orthonormée).
- *Étape 2*
  Comme $p_{\mathrm{F}}(u) \in \mathrm{F}$, il existe un unique p-uplet $(\lambda_1, \ldots, \lambda_p) \in \mathbb{R}^p$ tels que $p_{\mathrm{F}}(u) = \sum_{i=1}^{p} \lambda_i u_i$. On détermine les réels $(\lambda_i)_{i \in [\![1;p]\!]}$ en remarquant que
  $$\begin{cases} \left\langle u - p_{\mathrm{F}}(u), u_1 \right\rangle = 0 \\ \quad\vdots \\ \left\langle u - p_{\mathrm{F}}(u), u_p \right\rangle = 0. \end{cases}$$
- *Étape 3*
  On résout le système linéaire précédent à p équations pour trouver les p inconnues $(\lambda_i)_{i \in [\![1;p]\!]}$. On conclut par le calcul de $p_{\mathrm{F}}(u)$.



Exercice 291

✦ **Exemple**

Soient $\mathbb{R}_3[x]$ muni du produit scalaire $\langle P, Q\rangle = \displaystyle\int_0^1 P(t)Q(t)\,dt$ et $F = \mathbb{R}_1[x]$.

Donner l'expression du projeté orthogonal de $P = 1 + x + x^2 + x^3$ sur F.

3 Applications à l'optimisation

3.1 Distance à un sous-espace vectoriel

Soient F un sous-espace vectoriel de E et $u \in E$. On définit (sous réserve d'existence), la distance du vecteur u à F par

$$d(u, F) = \min_{u \in F} \|x - u\|.$$

Théorème 213 (distance à un s.e.v)

Soient F un sous-espace vectoriel de E, p_F la projection orthogonale sur F et $u \in E$. Alors la distance $d(u, F)$ est bien définie et

$$d(u, F) = \min_{v \in F} \|u - v\| = \left\| u - p_F(u) \right\|.$$

De plus, le minimum est atteint uniquement pour $p_F(u)$.

Remarque. Le projeté orthogonal de u sur F est caractérisé par

$$\forall\, v \in F, \quad \|u - v\| \geqslant \|u - p(u)\|.$$

Autrement dit :

$$v = p_F(u) \quad \Longleftrightarrow \quad \left(v \in F \quad \text{et} \quad \|u - v\| = \min_{w \in F} \|u - w\| \right).$$

Ainsi, $d(u, F) = 0$ si et seulement si $u \in F$.

Exercice 292

✦✦ **Exemples**

Soient $u_0 \in E \neq 0_E$, et un hyperplan H.

1. Exprimer la distance d'un vecteur x à la droite $\mathrm{Vect}(u_0)$.
2. Faire de même avec la distance à H (on pourra exprimer le résultat à l'aide de n, un vecteur normal à H, voir exercice 288).

Exercice 293

✦✦ Justifier que la quantité

$$\inf_{(a,b) \in \mathbb{R}^2} \int_{-1}^{1} \left(t^2 - at - b\right)^2 \, dt$$

est bien définie et la calculer.

3.2 Problème des moindre carrés, droite de régression

Théorème 214 (problème des moindre carrés, pseudo-solution)

Soient $n, p \in \mathbb{N}^*$ avec $n \geqslant p$, $A \in \mathcal{M}_{n,p}(\mathbb{R})$ de rang p et $B \in \mathcal{M}_{n,1}(\mathbb{R})$.
Alors il existe un unique vecteur $X_0 \in \mathcal{M}_{p,1}(\mathbb{R})$ minimisant la quantité $\|AX - B\|$.

Remarques.

- $\|\cdot\|$ désigne ici la norme associée au produit scalaire canonique sur $\mathcal{M}_{n,1}(\mathbb{R})$.
- Le vecteur X_0 est l'unique solution du système de Cramer ${}^t\!AAX = {}^t\!AB$. On parle alors de **pseudo-solution**.

Exercices - TD

Exercice 294. ✦✦ À bonne distance d'Attila
On considère $\mathscr{M}_n(\mathbb{R})$ muni du produit scalaire

$$\langle A, B\rangle = \operatorname{Tr}\left({}^t AB\right).$$

Soit H, le sous-espace vectoriel des matrices de trace nulle.

1. Donner la dimension de $H^{\perp}$. Préciser une base.
2. Soit J la matrice de $\mathscr{M}_n(\mathbb{R})$ dont tous les coefficients sont égaux à 1. Calculer

$$\min_{A \in H} \|A - J\|.$$

Exercice 295. ✦✦✦ CNS pour un projecteur orthogonal
Soient $\left(E, \langle\cdot,\cdot\rangle\right)$ un espace euclidien et p, un projecteur de E. Montrer l'équivalence entre les énoncés suivants :

i) Le projecteur p est orthogonal.

ii) $\forall x \in E, \quad \left\langle x, p(x)\right\rangle \geqslant 0.$

iii) $\forall x \in E, \quad \|p(x)\| \leqslant \|x\|.$

Exercice 296. ✦✦ Matrice d'un projecteur orthogonal
Soient $\left(E, \langle\cdot,\cdot\rangle\right)$, un espace euclidien et F, un sous-espace vectoriel.
Soient $\mathscr{B}$ une base orthonormée de E et $\left(u_1, \dots, u_p\right)$ une base orthonormée de F. On note $U_1, \dots, U_p$ les vecteurs colonnes des coordonnées de $u_1, \dots, u_p$ dans la base $\mathscr{B}$. Montrer que la matrice du projeteur orthogonal sur F, noté p_F, dans la base $\mathscr{B}$ est donnée par

$$\operatorname{Mat}_{\mathscr{B}}\left(p_F\right) = \sum_{i=1}^{p} U_i \, {}^t U_i.$$

Exercice 297. ✦✦✦ Un problème d'optimisation, deux méthodes *Oraux ESCP 1999*

1. Soit F la fonction définie sur $\mathbb{R}^3$ par : $F(x, y, z) = 24x^2 + 2y^2 + z^2 + 12xy + 2yz + 4zx - 240x - 48y - 12z$.
   (a) Déterminer les points critiques de F.
   (b) On admet (par le calcul) que
   $$F(x, y, z) = (2x + y + z - 6)^2 + (4x + y - 18)^2 + 4(x - 9)^2 - 684$$
   Montrer que F atteint son minimum sur $\mathbb{R}^3$ en un unique point. Préciser ses coordonnées, ainsi que la valeur du minimum.
2. (a) Rappeler la valeur de $I_n = \displaystyle\int_0^{+\infty} e^{-t} t^n \, dt$.
   (b) Justifier la convergence et exprimer en fonction de F, l'intégrale :
   $$I(a, b, c) = \int_0^{+\infty} e^{-t}\left(t^3 - at^2 - bt - c\right)^2 dt.$$
   (c) Justifier l'existence et calculer
   $$I = \inf_{(a,b,c) \in \mathbb{R}^3} I(a, b, c).$$
3. (a) Justifier (brièvement) que
   $$\langle P, Q\rangle = \int_0^{+\infty} e^{-t} P(t) Q(t) \, dt$$
   définit un produit scalaire sur l'espace vectoriel $\mathbb{R}_3[x]$ des polynômes à coefficients réels de degré inférieur ou égal à 3.
   (b) Calculer la distance du polynôme $P_0(x) = x^3$ au sous-espace $H = \mathbb{R}_2[x]$.
   (c) Soit $T(x) = ax^2 + bx + c$ un polynôme appartenant à H. Montrer que T est le projeté orthogonal du polynôme x^3 sur le sous-espace H si et seulement si
   $$\partial_1 F(a, b, c) = 0, \quad \partial_2 F(a, b, c) = 0 \quad \text{et} \quad \partial_3 F(a, b, c) = 0$$
   et retrouver le résultat précédent.

Exercice 298. ✦✦✦ **Sommes de projecteurs orthogonaux** *d'après ESCP 2001*

Soit $s \in \mathbb{N}$, $s \geqslant 2$. On rappelle l'expression du produit scalaire canonique sur $\mathcal{M}_{s,1}(\mathbb{R})$,

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_s \end{bmatrix},\ y = \begin{bmatrix} y_1 \\ \vdots \\ y_s \end{bmatrix} \in \mathcal{M}_{s,1}(\mathbb{R}), \quad \langle x, y\rangle = \sum_{k=1}^{s} x_k y_k.$$

On note $\|\cdot\|$, la norme associée à ce produit scalaire.

1. On suppose dans cette question que $s = 4$, et on considère la matrice :

$$\mathrm{P} = \frac{1}{4}\begin{bmatrix} 1 & -1 & 1 & 1 \\ -1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & 1 \end{bmatrix}.$$

   (a) Calculer ${}^t\mathrm{P}$ et P^2.

   (b) Déterminer les valeurs propres de P et les sous-espaces propres associés. Montrer que P est la matrice d'une projection orthogonale sur un sous-espace de $\mathcal{M}_{4,1}(\mathbb{R})$ que l'on déterminera.

2. On revient maintenant au cas général (s quelconque). Montrer que la matrice $\mathrm{P} \in \mathcal{M}_s(\mathbb{R})$ est la matrice d'une projection orthogonale si et seulement si on a $\mathrm{P}^2 = \mathrm{P}$ et ${}^t\mathrm{P} = \mathrm{P}$.

3. Soient P et Q deux matrices de $\mathcal{M}_s(\mathbb{R})$ représentant chacune une projection orthogonale. On suppose de plus que pour tout $x \in \mathcal{M}_{s,1}(\mathbb{R})$:

$$\|\mathrm{P}x\|^2 + \|\mathrm{Q}x\|^2 \leqslant \|x\|^2 \quad (\bullet)$$

   (a) Montrer que $\mathrm{PQ} = \mathrm{QP} = 0$.

   (b) En déduire que $\mathrm{P} + \mathrm{Q}$ est la matrice d'une projection orthogonale.

4. Soient $\mathrm{P}_1, \mathrm{P}_2, \ldots, \mathrm{P}_n$, n matrices représentant chacune une projection orthogonale de $\mathcal{M}_{s,1}(\mathbb{R})$ et telles que $\mathrm{P}_1 + \mathrm{P}_2 + \cdots + \mathrm{P}_n = \mathrm{I}$, où I est la matrice identité de $\mathcal{M}_s(\mathbb{R})$.
   Montrer que pour toute partie non vide E de $\{1, 2, \ldots, n\}$, $\sum_{k \in \mathrm{E}} \mathrm{P}_k$ est la matrice d'une projection orthogonale de $\mathcal{M}_{s,1}(\mathbb{R})$.

5. On se place à nouveau dans le cas $s = 4$ et on considère la matrice :

$$\mathrm{Q} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}.$$

   (a) Montrer que Q est la matrice d'une projection orthogonale de $\mathcal{M}_{4,1}(\mathbb{R})$.

   (b) Montrer que P et Q vérifient la relation $(\bullet)$, où P est la matrice de la première question. Que peut-on dire de $\mathrm{P} + \mathrm{Q}$?

Exercice 299. ✦✦✦ **Sujet de révision** *extrait de ESSEC 2012*

Dans tout le problème, les lettres m et n désignent des entiers naturels supérieurs ou égaux à 1. Par ailleurs, on munit $\mathcal{M}_{m,1}(\mathbb{R})$ de sa structure euclidienne canonique. Ainsi si

$$\mathrm{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \quad \text{et} \quad \mathrm{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \in \mathcal{M}_{m,1}(\mathbb{R}),$$

le produit scalaire de X et Y s'obtient par la relation ${}^t\mathrm{XY} = \sum_{i=1}^{m} x_i y_i$ et la norme euclidienne de Y par : $\|\mathrm{Y}\|_m^2 = {}^t\mathrm{YY} = \sum_{i=1}^{m} {y_i}^2$.

1. *Question préliminaire.*
   Soit F un sous-espace vectoriel de $\mathcal{M}_{n,1}(\mathbb{R})$ de dimension k non nulle et $(\mathrm{U}_1, \mathrm{U}_2, \ldots, \mathrm{U}_k)$ une base orthonormée de vecteurs colonnes de F. On envisage la projection orthogonale sur F représentée par sa matrice P dans la base canonique de $\mathcal{M}_{n,1}(\mathbb{R})$.
   Montrer que $\mathrm{P} = \sum_{i=1}^{k} \mathrm{U}_i^t \mathrm{U}_i$ et vérifier que P est une matrice symétrique.

2. **Partie I.** *Décomposition spectrale de la matrice* ${}^t\mathrm{AA}$ *associée à une matrice* A *de* $\mathcal{M}_{m,n}(\mathbb{R})$.
   On envisage dans toute cette partie une matrice A appartenant à $\mathcal{M}_{m,n}(\mathbb{R})$.

   (a) Préciser la taille de la matrice ${}^t\mathrm{AA}$ et vérifier que $\mathrm{Ker\,A} \subset \mathrm{Ker}\,{}^t\mathrm{AA}$.

   (b) Montrer que si $\mathrm{X} \in \mathrm{Ker}\,{}^t\mathrm{AA}$ alors $\|\mathrm{AX}\|_m = 0$ et établir que $\mathrm{Ker\,A} = \mathrm{Ker}\,{}^t\mathrm{AA}$. Montrer que A et ${}^t\mathrm{AA}$ sont nulles simultanément.

   (c) Justifier l'égalité : $\mathrm{Im}\,{}^t\mathrm{A} = \mathrm{Im}\,{}^t\mathrm{AA}$.

3. (a) Établir que la matrice ${}^t\mathrm{AA}$ est diagonalisable et en calculant $\|\mathrm{AX}\|_m^2$ pour X vecteur propre de la matrice ${}^t\mathrm{AA}$, montrer que ses valeurs propres sont des réels positifs.

   (b) On désigne par $\left(\lambda_1, \lambda_2, \ldots, \lambda_p\right)$ la liste des valeurs propres distinctes de la matrice ${}^t\mathrm{AA}$, classée dans l'ordre croissant. On rappelle que

$$\mathscr{M}_{n,1}(\mathbb{R}) = \bigoplus_{i=1}^{p} \mathrm{E}_{\lambda_i}\left({}^t\mathrm{AA}\right) \quad \text{où} \quad \mathrm{E}_{\lambda_i}\left({}^t\mathrm{AA}\right) = \mathrm{Ker}\left({}^t\mathrm{AA} - \lambda_i \mathrm{I}_n\right).$$

   Pour i entier naturel compris entre 1 et p, on note P_i la matrice de la projection orthogonale sur $\mathrm{E}_{\lambda_i}\left({}^t\mathrm{AA}\right)$ dans la base canonique de $\mathscr{M}_{n,1}(\mathbb{R})$.
   Vérifier que pour i et j distincts compris entre 1 et p, $\mathrm{P}_i\mathrm{P}_j$ est la matrice nulle.

   (c) Justifier les relations : $\mathrm{I}_n = \sum\limits_{i=1}^{p} \mathrm{P}_i$ et ${}^t\mathrm{AA} = \sum\limits_{i=1}^{p} \lambda_i \mathrm{P}_i$.
   Cette dernière écriture s'appelle *la décomposition spectrale* de ${}^t\mathrm{AA}$.

4. *Exemples*

   (a) Déterminer la décomposition spectrale de ${}^t\mathrm{AA}$ lorsque A est la matrice 3,3 égale à

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ -1 & 1 & 2 \end{bmatrix}.$$

   (b) On envisage la matrice ligne $\mathrm{A} = [a_1\ a_2\ \cdots\ a_n]$ où les réels $a_1, a_2, \ldots, a_n$ sont fixés, non tous nuls simultanément. Ainsi, $\mathrm{A}{}^t\mathrm{A}$ est un réel. Montrer que le polynôme $\mathrm{S}(x) = x^2 - \left(\mathrm{A}{}^t\mathrm{A}\right)x$ est annulateur pour la matrice ${}^t\mathrm{AA}$.
   Préciser la liste des valeurs propres et la décomposition spectrale de la matrice ${}^t\mathrm{AA}$.

Partie II. *Pseudo solution d'une équation linéaire.*
On s'intéresse dans cette partie à l'équation AX = B où $\mathrm{A} \in \mathscr{M}_{m,n}(\mathbb{R})$ et $\mathrm{B} \in \mathscr{M}_{m,1}(\mathbb{R})$. Une matrice X appartenant à $\mathscr{M}_{n,1}(\mathbb{R})$ est dite solution de cette équation si elle vérifie la relation AX = B. Elle est dite pseudo solution de cette équation si elle vérifie :

$$\forall \mathrm{Z} \in \mathscr{M}_{n,1}(\mathbb{R}), \quad \|\mathrm{AX} - \mathrm{B}\|_m \leqslant \|\mathrm{AZ} - \mathrm{B}\|_m.$$

5. On suppose que l'équation AX = B admet au moins une solution. Montrer que X est une pseudo-solution si et seulement si elle est solution de l'équation.

6. On suppose que X est une pseudo-solution de l'équation. Montrer que, pour tout réel λ et toute matrice Y de $\mathscr{M}_{n,1}(\mathbb{R})$, on a :

$$\lambda^2 \|\mathrm{AY}\|_m^2 + 2\lambda\, {}^t\mathrm{Y}\, {}^t\mathrm{A}(\mathrm{AX} - \mathrm{B}) \geqslant 0.$$

   En déduire que ${}^t\mathrm{AAX} = {}^t\mathrm{AB}$.

7. Montrer que tout X de $\mathscr{M}_{n,1}(\mathbb{R})$ vérifiant la relation ${}^t\mathrm{AAX} = {}^t\mathrm{AB}$ est pseudo-solution et en déduire qu'il existe toujours au moins une pseudo-solution de l'équation.

8. *Exemple*
   Déterminer toutes les pseudo-solutions de l'équation AX = B lorsque :

$$\mathrm{A} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ -1 & 1 & 2 \end{bmatrix} \quad \text{et} \quad \mathrm{B} = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}.$$

   Parmi celles-ci, préciser celle dont la norme euclidienne est minimale.

9. Donner une condition sur le rang de A pour que l'équation admette une unique pseudo-solution.

CHAPITRE 15

Convergences et approximations

> *Ce calcul délicat s'étend aux questions les plus importantes de la vie, qui ne sont en effet, pour la plupart, que des problèmes de probabilité.*
>
> PIERRE-SIMON, MARQUIS DE LAPLACE
> Mathématicien, physicien français (1749-1827)

1 Inégalités de concentration

Proposition 215 (inégalités de Markov et de Bienaymé-Tchebychev)

- Soit Z une variable aléatoire positive admettant une espérance, alors

$$\forall \lambda \in \mathbb{R}_*^+, \qquad \mathbf{P}\big([Z \geqslant \lambda]\big) \leqslant \frac{\mathbf{E}(Z)}{\lambda}.$$

- Soit X une variable aléatoire réelle admettant une variance, alors

$$\forall \varepsilon \in \mathbb{R}_*^+, \qquad \mathbf{P}\big(|X - \mathbf{E}(X)| \geqslant \varepsilon\big) \leqslant \frac{\mathbf{V}(X)}{\varepsilon^2}.$$

Exercice 300

✦ En moyenne, une personne sur 100 sait placer dans le bon ordre les pays baltes sur une carte. On choisit au hasard n personnes et de manière indépendante, notons Y_n le pourcentage des personnes capables de donner le bon ordre.

1. Donner l'espérance et la variance de Y_n.
2. Par application de l'inégalité de Bienaymé-Tchebychev, donner une valeur de n à partir de laquelle Y_n se trouve dans l'intervalle

$$I =]0,009; 0,011[$$

avec une probabilité supérieure à 0,9.

Remarque. Ces inégalités ont peu d'applications pratiques, car la majoration qu'elles fournissent est la plupart du temps excessive, mais elles sont valables quelle que soit la loi de X, pourvu que l'on puisse définir une espérance ou une variance. L'inégalité de Bienaymé-Tchebychev nous permettra toutefois de démontrer la loi faible des grands nombres (voir théorème page 159).

Exercice 301

✦ Soit X une variable aléatoire possédant une espérance de 6 et une variance de 2. Appliquer, lorsque cela est possible, l'inégalité de Bienaymé-Tchebychev pour majorer ou minorer les probabilités des événements suivants. Préciser si le résultat obtenu est intéressant.

1. $[2 \leqslant X \leqslant 10]$;
2. $[5 < X < 7]$;
3. $[X \leqslant 7]$;
4. $[|X-6| \geqslant 1]$;
5. $[X \geqslant 11]$;
6. $[X \geqslant 4]$.

2 Convergence en probabilité

2.1 Définition et exemples

Définition 216 (Convergence en probabilité)

Soient $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ et X une variable aléatoire définie aussi sur cet espace.
On dit que la suite $(X_n)_n$ **converge en probabilité** vers la variable aléatoire X si pour tout nombre réel $\varepsilon > 0$, on a

$$\lim_{n\to+\infty} \mathbf{P}\left(|X_n - X| \geqslant \varepsilon\right) = 0.$$

On note
$$X_n \xrightarrow[n\to+\infty]{\mathbf{P}} X.$$

Exemple. Soit (X_n) une suite de variables aléatoires mutuellement indépendantes. On suppose de plus que

$$\forall\, n \in \mathbb{N}, \quad X_n \hookrightarrow \mathscr{U}([0;1]).$$

On montre que $(Y_n)_{n\in\mathbb{N}^*}$ converge en probabilité vers la variable aléatoire presque surement constante à 1.

Exercice 302

✧ Pour tout $n \in \mathbb{N}^*$, on pose $Z_n = \min(X_1, X_2, \ldots, X_n)$.
Reprendre le calcul précédent pour justifier la convergence en probabilité de $(Z_n)_n$ vers une variable aléatoire que l'on précisera.

2.2 Les théorèmes de convergence en probabilité

Règles de calcul

⚠ **Attention.** Contrairement au cas des suites réelles ou des fonctions numériques, il n'y a pas unicité de la limite (si elle existe). Plus précisément, si $(X_n)_{n\in\mathbb{N}}$ est une suite de variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ et X, X′ des variables aléatoires définies sur le même espace et telles que

$$X_n \xrightarrow[n\to+\infty]{\mathbf{P}} X \quad \text{et} \quad X_n \xrightarrow[n\to+\infty]{\mathbf{P}} X' \quad \text{alors} \quad \mathbf{P}\left(X \neq X'\right) = 0.$$

Proposition 217 (convergence en probabilité et somme)

Soient $(X_n)_{n\in\mathbb{N}}$, $(Y_n)_{n\in\mathbb{N}}$ deux suites de variables aléatoires sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$.

Si $\quad X_n \xrightarrow[n\to+\infty]{\mathbf{P}} X \quad \text{et} \quad Y_n \xrightarrow[n\to+\infty]{\mathbf{P}} Y$

Alors pour tout $\lambda \in \mathbb{R}$, $\quad \lambda X_n \xrightarrow[n\to+\infty]{\mathbf{P}} \lambda X \quad \text{et} \quad X_n + Y_n \xrightarrow[n\to+\infty]{\mathbf{P}} X + Y.$

Proposition 218 (composition par une fonction continue)

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$.

Si
- La suite $(X_n)_{n\in\mathbb{N}}$ converge en probabilité vers X.
- La fonction f est continue sur $\mathbb{R}$ à valeurs réelles.

Alors $$f(X_n) \xrightarrow[n\to+\infty]{\mathbf{P}} f(X).$$

Remarque. On peut affiner le théorème en ne supposant que f est continue sur l'intervalle I si pour tout $n \in \mathbb{N}$, $\mathbf{P}(X_n \in I) = 1$.

Loi faible des grands nombres

Théorème 219 (loi faible des grands nombres)

Soient $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ et X une variable aléatoire définie aussi sur cet espace.

Si
- La variable X admet un moment d'ordre 2.
- Les variables $(X_n)_{n\in\mathbb{N}}$ sont mutuellement indépendantes et de même loi que X.

Alors la suite des variables aléatoires $\overline{X_n}$, moyenne arithmétique des n variables X_1, X_2, $\cdots$, X_n, converge en probabilité vers son espérance mathématique $\mathbf{E}(X)$. Autrement dit,

$$\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_n \xrightarrow[n\to+\infty]{\mathbf{P}} \mathbf{E}(X).$$

Remarque. Loi faible des grands nombres dans le cas de la loi binomiale
Soit $(Y_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ avec $Y_n \hookrightarrow \mathscr{B}(n;p)$. On a

$$\frac{Y_n}{n} \xrightarrow[n\to+\infty]{\mathbf{P}} Z, \qquad \text{où Z est une variable aléatoire certaine égale à } p.$$

Application. Considérons une expérience aléatoire, et un événement A de probabilité théorique p associé à cette expérience. Répétons n fois l'expérience de manière indépendante et désignons par Y_n le nombre de succès (c'est-à-dire le nombre de fois où A est réalisé). Y_n suit donc une loi binomiale de paramètres n, p. Posons de plus,

$$F_n(A) = \frac{Y_n}{n}, \text{ la fréquence empirique d'apparition de l'événement A.}$$

Corollaire 220 (interprétation d'une probabilité)

Lorsque le nombre d'expériences aléatoires augmente indéfiniment, la fréquence d'apparition $F_n(A)$ d'un événement A converge en probabilité vers sa probabilité théorique p. Autrement dit

$$\forall\,\varepsilon \in \mathbb{R}_*^+, \quad \mathbf{P}\Big(\big|F_n(A) - p\big| > \varepsilon\Big) \xrightarrow[n\to+\infty]{} 0.$$

On a une première formulation mathématique de l'interprétation intuitive d'une probabilité d'un événement.

> La probabilité d'un événement est la fréquence que l'on observerait si on effectuait une infinité de fois l'expérience dans des conditions parfaitement identiques.

Remarque. Il existe de nombreuses versions de la loi des grands nombres. Par exemple, la loi *forte* des grands nombres établit la convergence *presque sûr* de la suite $(X_n)_{n\in\mathbb{N}}$ (voir exercice 328, p.172).

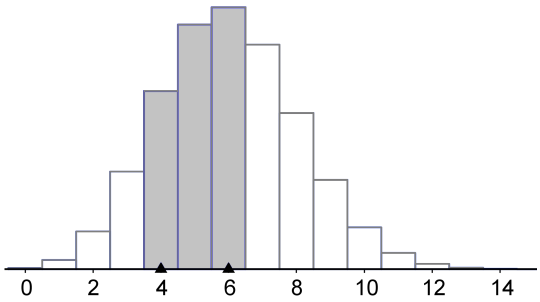
3 Convergence en loi

3.1 Rappels : représentations graphiques des lois

Cas des variables aléatoires discrètes

- Soit X une variable finie avec $X(\Omega) = \{x_1, x_2, \cdots, x_m\}$.

Nous avons vu que l'on peut résumer une loi d'une variable finie par un tableau. Pour chaque indice i, on indique la probabilité $\mathbf{P}(X = x_i)$.
On peut aussi utiliser les histogrammes, en abscisse, on place les valeurs x_i. L'aire du rectangle partant de x_i s'identifie à $\mathbf{P}(X = x_i)$.

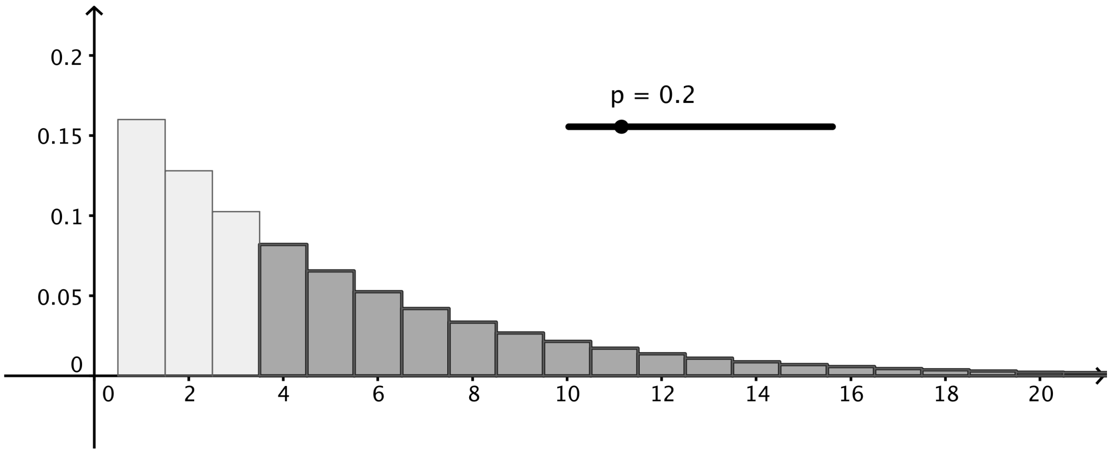


Exemple. Ci contre, le cas de la loi binomiale de paramètres $n = 20$, $p = 0,3$.
Notons que pour avoir la probabilité $\mathbf{P}\Big([X \in [a;b]]\Big)$, il suffit de sommer les aires des rectangles compris entre les abscisses a et b. En particulier, la somme des aires totales des rectangles est $\mathbf{P}(\Omega) = 1$.
Dans notre exemple, la partie grisée a pour aire

$$\mathbf{P}\Big([X \in [4;6]]\Big).$$

- Lorsque X est une variable aléatoire discrète dénombrable ($X(\Omega) = \{x_1, x_2, \cdots, x_n, \cdots\}$), on ne considère qu'un nombre fini de valeurs $\{x_1, x_2, \cdots, x_n\}$. En géneral, les valeurs où la probabilité n'est pas négligeable.

Donnons l'exemple de la loi géométrique de paramètre $p = 0,2$ où on s'est limité à $[\![0;20]\!]$.



L'aire de la partie la plus grisée correspond à une approximation de $\mathbf{P}([X \geqslant 4])$.

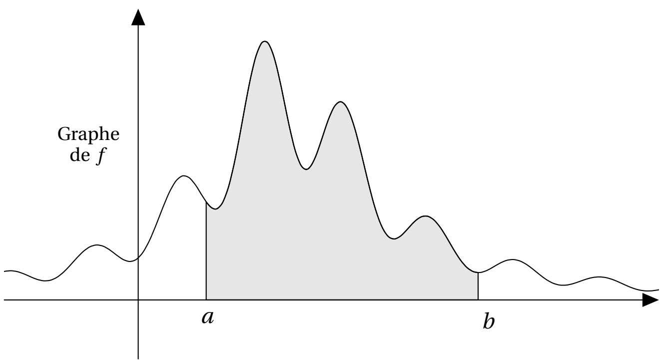
Graphe des densités des variables aléatoires à densité

Soit X une variable aléatoire à densité dont f, est une densité.
L'aire de la partie grisée comprise entre la courbe, l'axe des abscisses et les droites d'équation $x = a$, $x = b$ correspond exactement à la probabilité que X prenne les valeurs entre a et b.

$$\text{Aire} = \mathbf{P}([a \leqslant X \leqslant b]) = \int_a^b f(t)\,\mathrm{d}t.$$

Lorsque $a \to -\infty$ et $b \to +\infty$, on retrouve bien

$$1 = \mathbf{P}(\Omega) = \int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t.$$



3.2 Définition et exemples

Définition 221 (Convergence en loi)

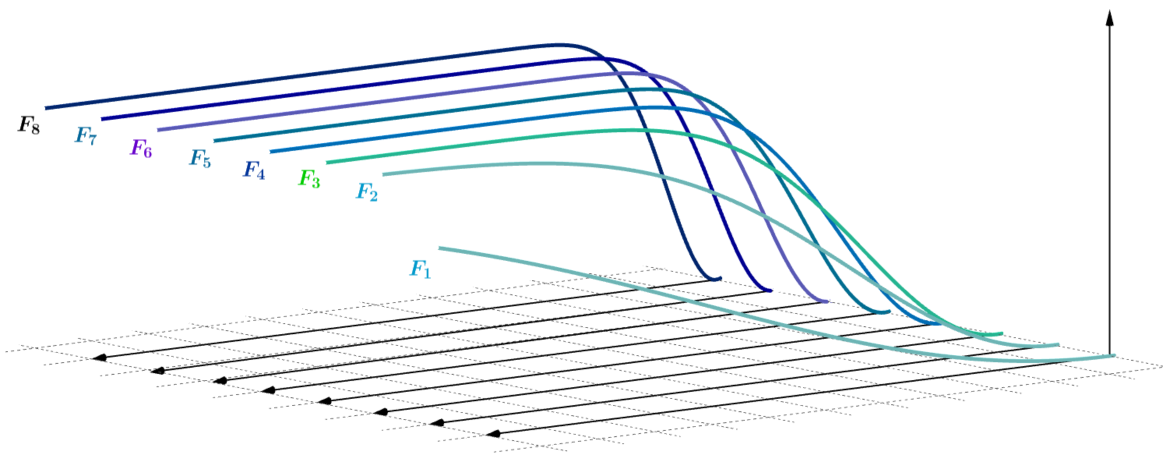
Soient $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires définies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ et X une variable aléatoire définie aussi sur cet espace. Pour tout $n \in \mathbb{N}$, on note F_n, la fonction de répartition de la variable X_n et F celle de X.
On dit que la suite $(X_n)_{n\in\mathbb{N}}$ **converge en loi** vers X si en tout point x de continuité de F, on a

$$F_n(x) \xrightarrow[n\to\infty]{} F(x).$$

On note

$$X_n \xrightarrow[n\to+\infty]{\mathscr{L}} X.$$

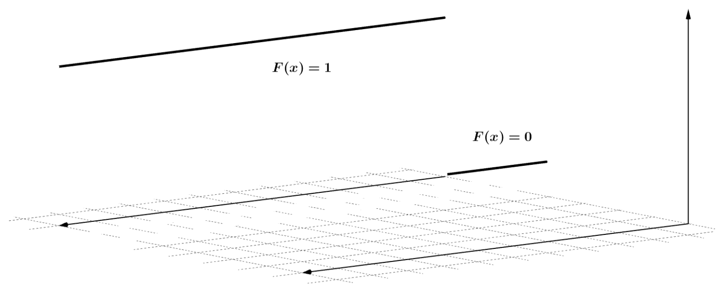
Exemple. Pour tout $n \in \mathbb{N}^*$, on vérifie que la fonction f_n définie sur $\mathbb{R}$ par $f_n(x) = n^2 x \exp\left(-n^2x^2/2\right) \mathbf{1}_{\mathbb{R}_+}(x)$ est une densité de probabilité. Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires à densité où f_n est une densité de X_n. Notons F_n la fonction de répartition de X_n.



On constate que « les courbes convergent vers une courbe limite » qui n'est autre que la courbe de la fonction de répartition d'une loi presque surement constante :

$$\forall x \in \mathbb{R}, \quad F(x) = \begin{cases} 1 & \text{si } x \geqslant 0 \\ 0 & \text{sinon.} \end{cases}$$

On vérifie, par le calcul, que $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers X où X est une variable presque surement constante à 0.



Exercice 303

✦✦ Reprenons les notations de l'exemple précédent et posons, pour tout $n \in \mathbb{N}^*$,

$$M_n = n(1 - Y_n).$$

Étudier la convergence en loi de la suite $(M_n)_{n\in\mathbb{N}^*}$.

⚠ **Attention.** Il n'y a pas unicité de la limite lors d'une convergence en loi. Si X et Y sont deux variables aléatoires de même loi, alors

$$X_n \xrightarrow[n\to+\infty]{\mathscr{L}} X \iff X_n \xrightarrow[n\to+\infty]{\mathscr{L}} Y.$$

Proposition 222 (Convergence en loi)

La convergence en loi de la suite $(X_n)_{n\in\mathbb{N}}$ vers une variable X impose pour tous points a, b de continuité de F
$$\mathbf{P}\Big([a < X_n \leqslant b]\Big) \underset{n\to\infty}{\longrightarrow} \mathbf{P}\Big([a < X \leqslant b]\Big).$$

Proposition 223 (convergence en loi dans le cas discret)

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires telles que pour tout $n\in\mathbb{N}$,
$$X_n(\Omega)\subset\mathbb{N} \quad \text{et} \quad \forall\, k\in\mathbb{N},\quad \mathbf{P}\big([X_n = k]\big) \underset{n\to\infty}{\longrightarrow} p_k \in [0;1].$$
Alors, $(X_n)_{n\in\mathbb{N}}$ converge en loi vers une variable aléatoire X avec
$$X(\Omega)\subset\mathbb{N} \quad \text{et} \quad \forall\, k\in X(\Omega), \qquad \mathbf{P}\big([X = k]\big) = p_k.$$

Exemple. Si $X_n \hookrightarrow \mathscr{B}(p_n)$ avec $p_n \underset{n\to\infty}{\longrightarrow} p \in]0;1[$. On montre que $(X_n)_{n\in\mathbb{N}}$ converge en loi vers une variable aléatoire suivant une loi de Bernoulli $\mathscr{B}(p)$.

Exercice 304

✦ **Exemples** *Les questions 1 et 2 sont indépendantes*

1. Soient $n\in\mathbb{N}^*$ et X_n une variable aléatoire de loi
$$X_n(\Omega) = \{0;1;2\} \quad \text{et} \quad \mathbf{P}\big([X_n = 0]\big) = \frac{n-\alpha}{3n}, \quad \mathbf{P}\big([X_n = 1]\big) = \frac{n+\cos(n)^2}{3n} \quad \mathbf{P}\big([X_n = 2]\big) = \frac{n+\sin(n)^2}{3n}.$$
   (a) Déterminer la valeur de α.
   (b) Vérifier que la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers une loi usuelle.
2. Soit $X_n \hookrightarrow \mathscr{P}(\lambda_n)$ avec $\lambda_n \underset{n\to\infty}{\longrightarrow} \lambda \in \mathbb{R}_*^+$.
   Montrer que $(X_n)_{n\in\mathbb{N}}$ converge en loi vers une variable aléatoire dont on précisera la loi.

3.3 Les théorèmes de convergence en loi

Règles de calculs sur les limites

⚠ **Attention.** Contrairement à la convergence en probabilité, la convergence en loi de $(X_n)_{n\in\mathbb{N}}$ et $(Y_n)_{n\in\mathbb{N}}$ vers X et Y n'implique pas nécessairement la convergence de la suite $(X_n + Y_n)_{n\in\mathbb{N}}$ vers $X+Y$.

Remarques. *Un peu de hors-programme*

- On montre que la convergence en probabilité implique la convergence en loi :
$$X_n \xrightarrow[n\to+\infty]{\mathbf{P}} X \quad\Rightarrow\quad X_n \xrightarrow[n\to+\infty]{\mathscr{L}} X.$$

Il est à noter que la réciproque est fausse.

Exercice 305

✧ Soient c un réel et $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles convergeant en loi vers une variable aléatoire X. Alors la suite de variables aléatoires $(X_n + c)$ converge en loi vers $X + c$.

- L'exercice s'étend avec le théorème de Slutsky :

Si $X_n \xrightarrow[n\to+\infty]{\mathscr{L}} X$ et si (Y_n) converge en probabilité vers une constante c, alors :

- $(X_n + Y_n)$ converge en loi vers $X + c$
- (X_nY_n) converge en loi vers cX.

Proposition 224 (composition et convergence en loi)

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$.

Si
- La suite $(X_n)_{n\in\mathbb{N}}$ converge en loi vers X.
- La fonction f est continue sur $\mathbb{R}$ à valeurs réelles.

Alors
$$f(X_n) \underset{n\to+\infty}{\xrightarrow{\mathscr{L}}} f(X).$$

Convergence de la loi binomiale vers la loi de Poisson

Théorème 225 (convergence de la loi binomiale vers la loi de Poisson)

Soient $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires binomiales $\mathscr{B}(n;p_n)$ telles que

$$np_n \underset{n\to+\infty}{\longrightarrow} \lambda \in \mathbb{R}_*^+.$$

Alors la suite $(X_n)_{n\in\mathbb{N}}$ converge en loi vers une variable aléatoire suivant une loi de Poisson $\mathscr{P}(\lambda)$.

$$X_n \underset{n\to+\infty}{\xrightarrow{\mathscr{L}}} Z \quad \text{avec} \quad Z \hookrightarrow \mathscr{P}(\lambda).$$

Exercice 306

✦✦ Voici la preuve du théorème. Complétez-la.

Soit $k\in\mathbb{N}$. Pour $n \geqslant k$, on a $\quad \mathbf{P}([X_n = k]) = \binom{n}{k} {p_n}^k (1-p_n)^{n-k} = \binom{n}{k}(1-p_n)^n \left(\frac{p_n}{1-p_n}\right)^k$.

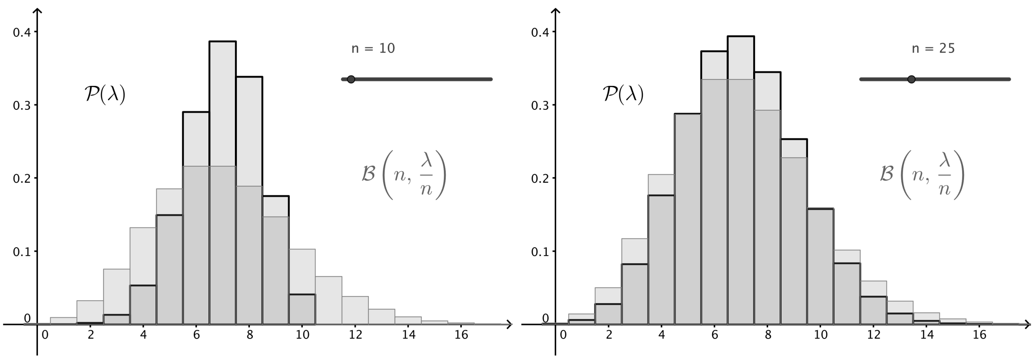
1. Justifier les équivalents de chacun des facteurs lorsque n tend vers $+\infty$ (k est fixé).

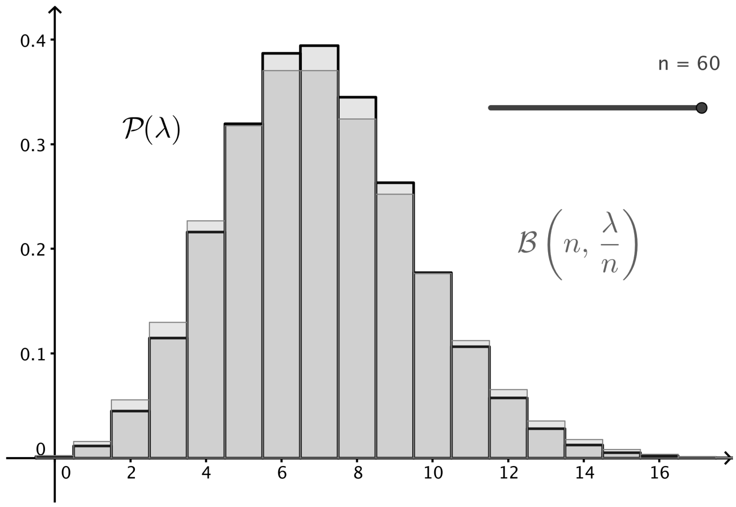
(a) $\left(\frac{p_n}{1-p_n}\right)^k \underset{n+\infty}{\sim} \left(\frac{\lambda}{n}\right)^k$; $\quad (b)$ $\ln(1-p_n) \underset{n+\infty}{\sim} -\frac{\lambda}{n}$, $(1-p_n)^n \underset{n+\infty}{\sim} e^{-\lambda}$; $\quad (c)$ $\binom{n}{k} \underset{n+\infty}{\sim} \frac{n^k}{k!}$.

2. Conclure.

Interprétation graphique

On trace les histogrammes représentant les lois binomiales $\mathscr{B}(n;\lambda/n)$ et de Poisson $\mathscr{P}(\lambda)$. On constate que plus n est grand, plus les histogrammes associés aux lois $\mathscr{P}(\lambda)$ et $\mathscr{B}(n;\lambda/n)$ se confondent.





Application à l'approximation

Exercice 307. Soit $Z \hookrightarrow \mathscr{P}(\lambda)$.

1. Justifier que pour tout réel x, $\dfrac{e^x + e^{-x}}{2} = \displaystyle\sum_{k=0}^{+\infty} \frac{x^{2k}}{(2k)!}$. En déduire $\mathbf{P}\big([Z \text{ est pair}]\big)$.
2. *Application*. Une ligne de transmission entre émetteur et récepteur transporte des données représentées par 10 000 bits (un bit est un élément de $\{0;1\}$). La probabilité que la transmission d'un bit soit erronée est estimée à 10^{-5} et on admet que les erreurs sont mutuellement indépendantes les unes des autres. On contrôle la qualité de la transmission avec un calcul de parité sur le nombre de « 1 » envoyés :
   - S'il y a un nombre impair d'erreurs, un message d'erreur apparaît.
   - Sinon, c'est-à-dire s'il y a un nombre pair d'erreurs, la transmission est acceptée.

   (a) Considérons X la variable aléatoire associant à chaque envoi de données, le nombre d'erreurs lors de la transmission, c'est-à-dire le nombre de bits parmi les 8 192 dont la transmission est erronée. Quelle est la loi de X ?

   (b) Calculer la probabilité qu'il n'y ait aucune erreur sachant que la transmission est acceptée.
   On admettra que l'on peut approximer le problème par une loi de Poisson.

4 Théorème limite central

4.1 Le théorème

Théorème 226 (limite central)

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$.

Si
- Les variables $(X_n)_{n\in\mathbb{N}^*}$ sont mutuellement indépendantes.
- Les variables $(X_n)_{n\in\mathbb{N}^*}$ ont même loi et admettent une espérance m et une variance $\sigma^2 \neq 0$.
- On note $\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i$ et $\overline{X_n}^* = \frac{\sqrt{n}}{\sigma}(\overline{X_n} - m)$.

Alors $$\left(\overline{X_n}^*\right)_{n\in\mathbb{N}^*} \xrightarrow[n\to+\infty]{\mathscr{L}} Z \quad \text{avec} \quad Z \hookrightarrow \mathscr{N}(0;1).$$

Autrement dit, pour tous $a < b$, $$\mathbf{P}\Big(\big[a \leqslant \overline{X_n}^* \leqslant b\big]\Big) \underset{n\to\infty}{\longrightarrow} \Phi(b) - \Phi(a) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2}\,dt,$$

où Φ est la fonction de répartition de la loi normale centrée réduite.

L'énoncé du théorème central limite est parfois surprenant et n'a souvent rien à voir avec celui du programme.

Rapport de Jury : Oral HEC 2021

4.2 Cas particuliers

Rappelons que si X est une variable aléatoire admettant une variance σ^2 (et donc une espérance), on définit la variable aléatoire centrée réduite associée à X, notée X^* par

$$\boxed{X^* = \frac{X - \mathbf{E}(X)}{\sigma} \quad \text{avec} \quad \mathbf{E}(X^*) = 0 \quad \text{et} \quad \mathbf{V}(X^*) = 1.}$$

Par exemple, si $X \hookrightarrow \mathscr{B}(n;p)$ alors $X^* = \dfrac{X - np}{\sqrt{np(1-p)}}$.

Théorème 227 (de Moivre-Laplace)

Si $(X_n)_{n\in\mathbb{N}}$ est une suite de variables aléatoires suivant des lois binomiales $\mathscr{B}(n;p)$ avec $p \in]0;1[$, alors

$$X_n{}^* = \frac{X_n - np}{\sqrt{np(1-p)}} \xrightarrow[n\to+\infty]{\mathscr{L}} Z \quad \text{avec} \quad Z \hookrightarrow \mathcal{N}(0;1).$$

Autrement dit, pour tous $a < b$, on a

$$\mathbf{P}\Big(\big[a \leqslant X_n{}^* \leqslant b\big]\Big) \xrightarrow[n\to\infty]{} \Phi(b) - \Phi(a) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2}\, dt,$$

où Φ est la fonction de répartition de la loi normale centrée réduite.

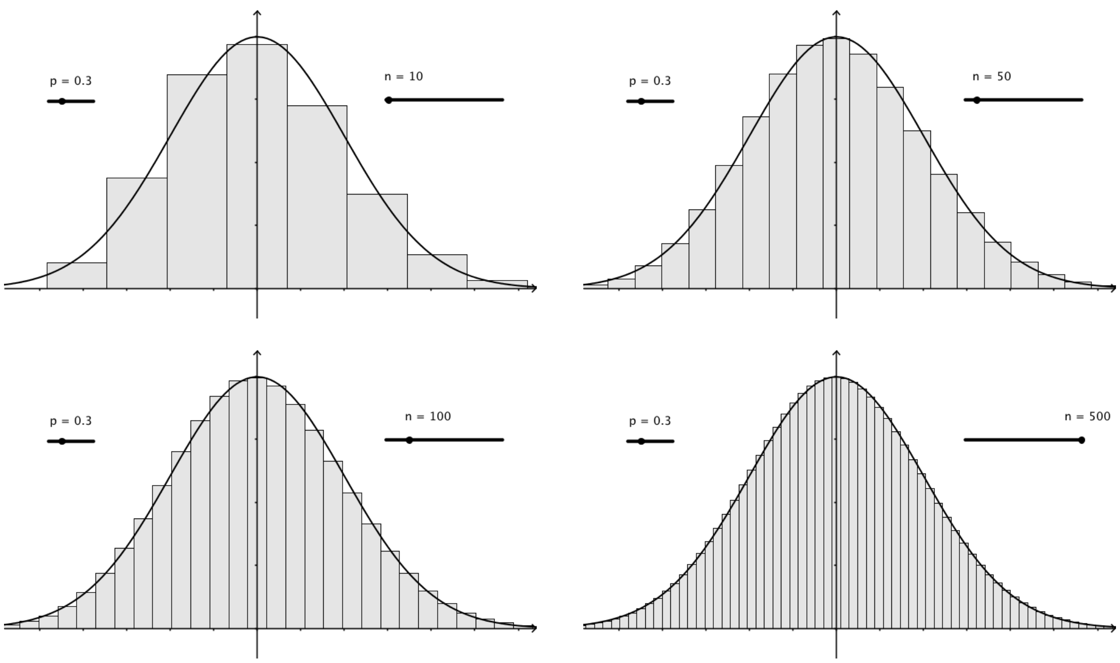
Remarque. Dans la pratique, dès que $n \geqslant 30$, $np \geqslant 5$ et $nq \geqslant 5$, on approche $\mathscr{B}(n,p)$ par $\mathcal{N}(np, npq)$.

⚠ **Attention.** Il y a deux théorèmes de convergence impliquant des lois binomiales. Précisons la différence :

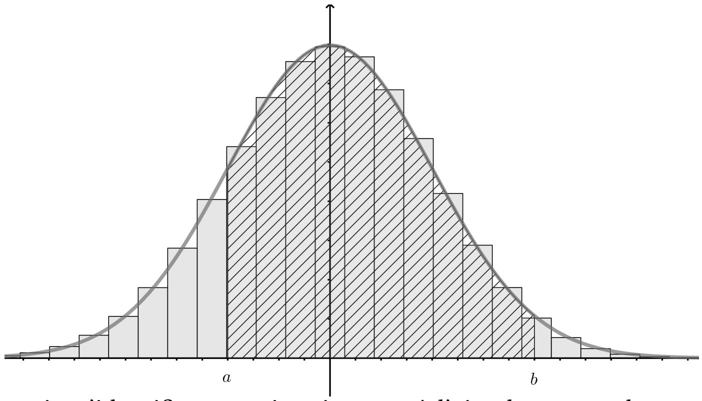
- Dans le premier, $n \to +\infty$ mais $np_n \xrightarrow[n\to\infty]{} \lambda > 0$. Sous-entendu, $p_n \xrightarrow[n\to\infty]{} 0$.
- Dans le second, p correspond à une probabilité fixée strictement positive.

Interprétation graphique

On trace les histogrammes associés aux lois de $X_n{}^*$ pour différentes valeurs de n (10, 50, 100 et 500). On superpose la courbe représentative de la densité de loi normale centrée réduite.



On constate que plus n est grand, plus les histogrammes épousent la forme de la courbe.



Interprétons.
Soit $X_n \hookrightarrow \mathscr{B}(n;p)$. De nouveau, plaçons l'histogramme associé à la loi ${X_n}^*$. L'aire hachurée est l'aire sous la courbe représentative de la densité. Elle vaut donc

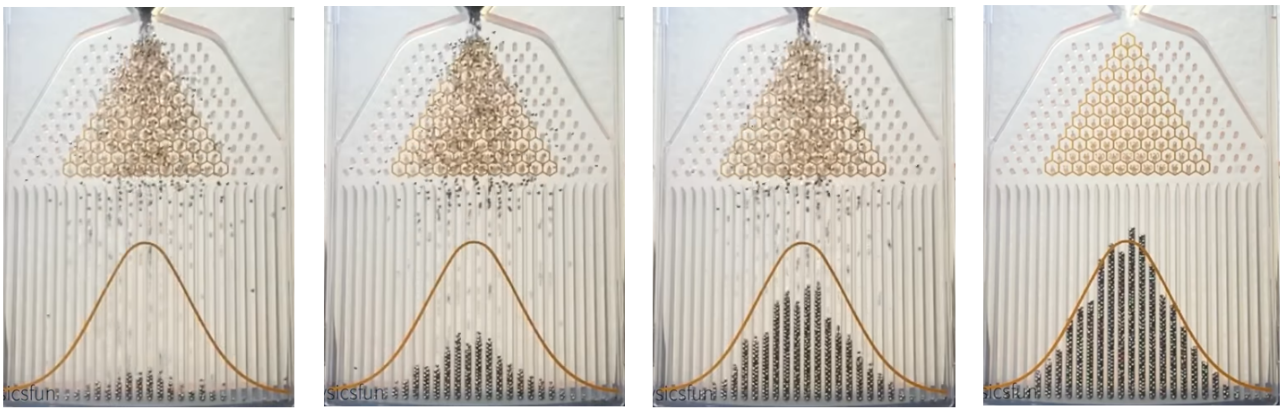
$$\int_a^b f(t)\,\mathrm{d}t = \mathbf{P}\big([a \leqslant Z \leqslant b]\big),$$

où $f : t \mapsto \dfrac{1}{\sqrt{2\pi}}\exp(-t^2)$ et $Z \hookrightarrow \mathcal{N}(0;1)$.

Cette aire s'identifie approximativement à l'aire des rectangles compris entre les droites $x = a$ et $x = b$. Or, cette dernière est par construction $\mathbf{P}\big([a \leqslant {X_n}^* \leqslant b]\big)$. On en déduit l'approximation :

$$\mathbf{P}\Big(\big[a \leqslant X_n^* \leqslant b\big]\Big) \simeq \mathbf{P}\Big(\big[a \leqslant Z \leqslant b\big]\Big).$$

Exemple. *La planche de Galton.*



Théorème 228 (convergence des lois de Poisson)

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires telle que $X_n \hookrightarrow \mathscr{P}(n\lambda)$ pour tout $n \in \mathbb{N}^*$.
Alors la suite des variables aléatoires centrées réduites $({X_n}^*)_{n\in\mathbb{N}}$ converge en loi vers une variable aléatoire suivant une loi normale centrée réduite.

$${X_n}^* = \frac{X_n - n\lambda}{\sqrt{n\lambda}} \xrightarrow[n\to+\infty]{\mathscr{L}} Z \quad \text{avec} \quad Z \hookrightarrow \mathcal{N}(0;1).$$

Remarque. Dans la pratique, dès que $\lambda \geqslant 18$, on approche la loi $\mathscr{P}(\lambda)$ par la loi normale $\mathcal{N}(\lambda,\lambda)$.

4.3 Applications à l'approximation

Méthode. Comment approximer une probabilité à l'aide d'une loi normale ?

Exemple. *Surbooking.*

5 Python

5.1 Simulation d'une loi normale par la méthode des 12 uniformes

D'après le théorème limite central, si (X_n) est une suite de variables aléatoires mutuellement indépendantes suivant une loi uniforme sur $[0;1]$, la variable aléatoire $\overline{X_n}^*$ centrée réduite associée à la moyenne arithmétique des variables $X_1, \dots, X_n$ converge en loi vers la loi normale $\mathcal{N}(0,1)$.

Exercice 308

✦✦

1. En se limitant à douze variables X_1, ..., X_{12}, écrire une fonction Python qui renvoie une simulation de la loi normale centrée réduite.
2. En déduire une seconde fonction qui prend en arguments μ, σ, m et renvoient m simulations d'une loi $\mathcal{N}(\mu, \sigma^2)$.
3. Tester votre programme en superposant sur une même figure l'histogramme de 2000 simulations de loi $\mathcal{N}(1,4)$ et la densité de cette loi.

5.2 La méthode de Monte-Carlo

• **Principe**

– Soit $f : [0,1] \to [a;b]$ une fonction continue dont on souhaite calculer l'intégrale $\displaystyle\int_0^1 f(t)\,\mathrm{d}t$.
– $(U_i)_{i\in\mathbb{N}}$ une suite de variables aléatoires indépendantes suivant la loi uniforme sur $[0,1]$.
– Pour tout $i \in \mathbb{N}$, on pose $f(U_i)$. D'après les lemme de l'égalité en loi et le lemme des coalitions, les variables sont indépendantes et de même loi.

Par le théorème de transfert, les variables $X_i = f(U_i)$ admettent toutes une même espérance donnée par

$$\mathbf{E}(X_i) = \int_0^1 f(t)\,\mathrm{d}t.$$

De même, ces variables admettent une variance et la loi faible des grands nombres donne

$$\frac{X_1 + \cdots + X_n}{n} \xrightarrow[n\to+\infty]{\mathbf{P}} \mathbf{E}(X_1) = \int_0^1 f(t)\,\mathrm{d}t.$$

Dès lors, pour calculer une valeur approchée de l'intégrale, on peut simuler un grand nombre de fois les variables U_i, calculer les images $f(U_i)$ et en faire leur moyenne arithmétique.

Exercice 309

✦ **La théorie : estimation de la probabilité de l'erreur**

Soit X, une variable aléatoire à valeurs dans $[a;b]$.

1. Pour quelle valeur de m, $\mathbf{E}\left((X-m)^2\right)$ atteint son minimum ? Avec $m = \frac{a+b}{2}$, déduire :
$$\mathbf{V}(X) \leqslant \frac{(b-a)^2}{4}.$$
2. En reprenant les notations du début, justifier que
$$\forall \varepsilon \in \mathbb{R}_*^+, \quad \mathbf{P}\left(\left|\int_0^1 f(t)\mathrm{d}t - \frac{1}{n}\sum_{k=1}^n f(U_k)\right| > \varepsilon\right) \leqslant \frac{(b-a)^2}{4n\varepsilon^2} \qquad (\bullet)$$

Exercice 310

✦ **La pratique**

On pose
$$I = \int_0^1 \frac{4}{1+t^2}\,\mathrm{d}t.$$

1. Donner la valeur explicite de I. En déduire, une fonction Python qui prend en argument n qui permet d'approcher I.
2. Déterminer n afin d'obtenir une valeur à 10^{-3} près de I avec une probabilité d'au moins 95%. Commenter.

Exercices - TD

Révisions

Exercice 311. ✦ Estimation
Une urne contient une proportion p de boules blanches. On souhaite obtenir expérimentalement une approximation de p. Pour cela, on effectue $n \in \mathbb{N}^*$ tirages avec remise et on note X_n le nombres de boules blanches obtenues au cours de ces n tirages. On suppose les tirages mutuellement indépendants.

1. Donner la loi de X_n. Préciser l'espérance et la variance.
2. Justifier que pour tout réel $\varepsilon > 0$, $\quad \mathbf{P}\left(\left|\frac{X_n}{n} - p\right| \geqslant \varepsilon\right) \leqslant \frac{1}{4n\varepsilon^2}$.
3. Combien de tirages faut-il effectuer pour pouvoir affirmer, avec un risque inférieur à 5%, que la fréquence d'obtention de boules blanches au cours des n tirages diffère de p d'au plus 1% ?

Exercice 312. ✦✦ Les souris mutantes
Un laboratoire élève des souris dont 1/4 sont mutantes. La durée de vie d'une souris mutante est une variable aléatoire dont la moyenne est de 3 ans avec un écart-type de 9 mois, mais elle ne vit jamais plus de 4 ans. La durée de vie d'une souris normale a une moyenne d'un an, avec un écart-type de 6 mois. On ne prend en compte que les souris dont la durée de vie est strictement positive.

Une souris est vivante au bout de deux ans. On note α la probabilité qu'elle soit mutante.

On considère l'événement M : « La souris est une souris mutante »et on note X la variable aléatoire égale à la durée de vie de la souris.

1. Exprimer $\dfrac{\mathbf{P}(M \cap [X \geqslant 2])}{\mathbf{P}(\overline{M} \cap [X \geqslant 2])}$ en fonction de α.
2. En utilisant l'inégalité de Bienaymé-Tchebychev, donner une minoration de α.

Exercice 313. ✦✦ Inégalité de Chernov

1. Soit $t \in \mathbb{R}_+^*$. Soit X une variable aléatoire discrète telle que e^{tX} admette une espérance. Montrer, à l'aide de l'inégalité de Markov, que pour tout $a \in \mathbb{R}$,
$$\mathbf{P}(X \geqslant a) \leqslant \frac{\mathbf{E}\left(e^{tX}\right)}{e^{ta}}.$$
2. Soient $n \in \mathbb{N}^*$ et $p \in]0,1[$. On suppose que $X \hookrightarrow \mathscr{B}(n,p)$.
   (a) Montrer que pour tout $t \in \mathbb{R}_+^*$, e^{tX} admet une espérance et que :
   $$\mathbf{E}\left(e^{t\mathbf{X}}\right) = \left(1 - p + pe^t\right)^n.$$
   (b) Étudier les variations de la fonction $f : t \longmapsto (1-p)e^{-\frac{t}{2}} + pe^{\frac{t}{2}}$ sur $\mathbb{R}_+^*$. En déduire que f admet un minimum sur $\mathbb{R}_+^*$, égal à $2\sqrt{p(1-p)}$.
   (c) À l'aide de la question 1, montrer que $\mathbf{P}\left(X \geqslant \frac{n}{2}\right) \leqslant 2^n\left(p(1-p)\right)^{\frac{n}{2}}$.

Exercice 314. ✦✦ Comparaison entre la médiane et l'espérance
Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$. On suppose que X admet une variance.

1. Soient $\alpha \in \mathbb{R}_*^+, \beta \in \mathbb{R}^+$. Démontrer que $\quad \mathbf{P}\left([X \geqslant \mathbf{E}(X) + \alpha]\right) \leqslant \dfrac{\mathbf{E}\left(\left(X - \mathbf{E}(X) + \beta\right)^2\right)}{(\alpha+\beta)^2}$.
2. Avec $\beta = \mathbf{V}(X)/\alpha$, en déduire que $\quad \mathbf{P}\left([X \geqslant \mathbf{E}(X) + \alpha]\right) \leqslant \dfrac{\mathbf{V}(X)}{\mathbf{V}(X) + \alpha^2}$.
3. On suppose dans cette question que X est une variable aléatoire à densité.
   (a) Justifier qu'il existe un réel m tel que
   $$\mathbf{P}\left([X \leqslant m]\right) = \frac{1}{2}.$$
   Un tel réel est une médiane de X.
   (b) À l'aide de la question 1, justifier que : $\quad \left|m - \mathbf{E}(X)\right| \leqslant \sigma(X)$.

Convergences en probabilité et en loi

Exercice 315. ✦✦
Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes suivant la même loi de Bernoulli de paramètre $p\in]0,1[$. Pour tout $n\in\mathbb{N}^*$, on pose :

$$Y_n = X_n + X_{n+1} \quad \text{et} \quad T_n = \frac{1}{n}\sum_{i=1}^{n} Y_i.$$

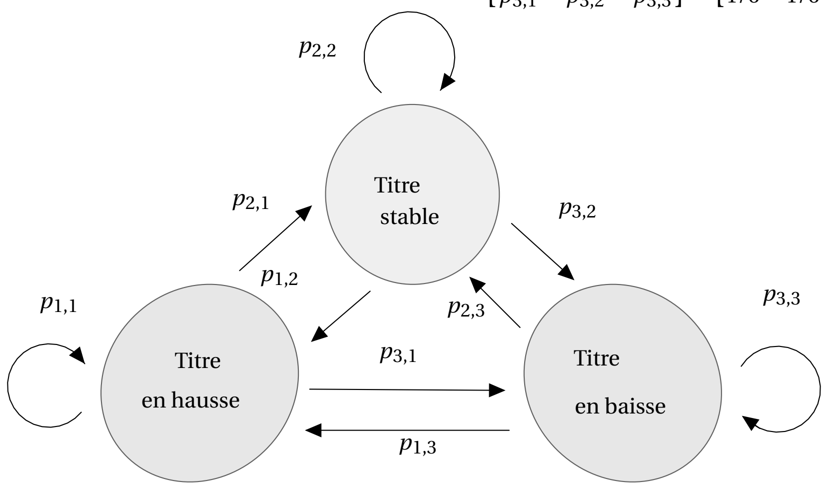
1. Calculer l'espérance et la variance de T_n.
2. Peut-on appliquer la loi des grands nombres pour étudier la convergence en probabilité de la suite $(T_n)_{n\in\mathbb{N}^*}$?
3. Justifier que $(T_n)_{n\in\mathbb{N}^*}$ converge en probabilité vers la variable aléatoire presque sûrement constante.

Exercice 316. ✦ **Chaîne de markov : évolution d'un titre boursier**
Dans une bourse de valeurs, un titre peut monter, descendre ou rester stable. On modélise l'évolution du titre.

- Si un jour n, le titre monte, le jour suivant, il montera avec la probabilité $2/3$, restera stable avec la probabilité $1/6$, et baissera avec la probabilité $1/6$.
- Si un jour n, le titre est stable, le jour $n+1$, il montera avec la probabilité $1/6$, restera stable avec la probabilité $2/3$, et baissera avec la probabilité $1/6$.
- Si un jour n, le titre baisse, le jour $n+1$, il montera avec la probabilité $1/6$, restera stable la probabilité $1/6$, et baissera avec la probabilité $2/3$.

Le premier jour, le titre est stable.

Les probabilités sont spécifiées par une matrice dite de transition : $M = \begin{bmatrix} p_{1,1} & p_{1,2} & p_{1,3} \\ p_{2,1} & p_{2,2} & p_{2,3} \\ p_{3,1} & p_{3,2} & p_{3,3} \end{bmatrix} = \begin{bmatrix} 2/3 & 1/6 & 1/6 \\ 1/6 & 2/3 & 1/6 \\ 1/6 & 1/6 & 2/3 \end{bmatrix}$.



On souhaite connaître l'évolution de ce titre. Pour cela, on introduit pour tout $n\in\mathbb{N}^*$, la variable aléatoire aléatoire X_n définie par

$$X_n = \begin{cases} 1 & \text{si le titre donné monte le jour } n \\ 0 & \text{si le titre est stable le jour } n \\ -1 & \text{si le titre donné baisse le jour } n. \end{cases} \quad \text{et} \quad U_n = \begin{bmatrix} \mathbf{P}(X_n=1) \\ \mathbf{P}(X_n=0) \\ \mathbf{P}(X_n=-1) \end{bmatrix}.$$

1. (a) Vérifier que $U_{n+1} = MU_n$.
   (b) En déduire U_n en fonction de M et U_1.
2. Donner la loi de X_n.
3. Justifier que $(X_n)_n$ converge en loi vers une variable aléatoire X.
4. Comparer MU et U où $U = \begin{bmatrix} \mathbf{P}(X=1) \\ \mathbf{P}(X=0) \\ \mathbf{P}(X=-1) \end{bmatrix}$. Commenter.

Exercice 317. ✧ Pour $n\in\mathbb{N}^*$, on considère une variable X_n dont la loi est donnée par :

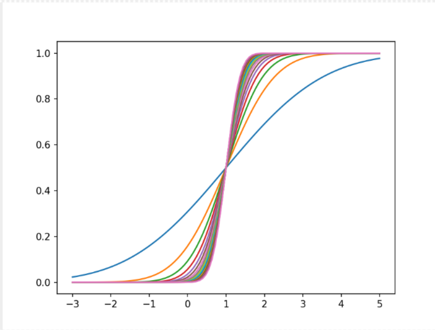
$$X_n(\Omega) = \{0,n\}, \quad \mathbf{P}(X_n=0) = 1-\frac{1}{n} \quad \text{et} \quad \mathbf{P}(X_n=n) = \frac{1}{n}.$$

Étudier la convergence en loi de la suite $(X_n)_{n\in\mathbb{N}^*}$ et la convergence de la suite numérique $\left(\mathbf{E}(X_n)\right)_{n\in\mathbb{N}^*}$. Commenter.

Exercice 318. ✦ Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires telles que pour tout $n\in\mathbb{N}^*$, $X_n \hookrightarrow \mathcal{N}(1,1/n)$.

1. Étudier la convergence en loi de la suite $(X_n)_{n\in\mathbb{N}^*}$.
2. Expliquer et commenter le programme Python suivant.

```python
import scipy.special as sp
import matplotlib.pyplot as plt
import numpy as np

plt.clf()

def Fnormal(x,k):
    return sp.ndtr(k**(1/2)*(x-1)/2)

x =np.linspace(-3,5,100)
for k in range(1,50,3):
    y=Fnormal(x,k)
    plt.plot(x,y)

plt.show()
```



Exercice 319. ✦✦ **Convergence en loi avec des lois de Cauchy**
Pour tout $n \in \mathbb{N}^*$, on pose

$$\forall t \in \mathbb{R}, \quad f_n(t) = \frac{n}{\pi\left(1+n^2 t^2\right)}.$$

1. Justifier que f_n est une densité de probabilité. Soit X_n une variable aléatoire dont f_n est une densité.
2. Peut-on appliquer l'inégalité de Markov à X_n ?
3. Donner la fonction de répartition de X_n. En déduire la convergence en loi de la suite de variable aléatoire $(X_n)_n$.

Exercice 320. ✦✦ **Variante de la loi faible des grands nombres**
Soit $(p_n)_{n\in\mathbb{N}^*}$ une suite de réels appartenant à $[0,1]$ et $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires indépendantes. On suppose que pour tout $k \in \mathrm{N}^*$, X_k suit une loi de Bernoulli de paramètre p_k. On pose pour tout $n \in \mathbb{N}^*$,

$$Y_n = \frac{1}{n}\sum_{k=1}^{n} X_k \quad \text{et} \quad m_n = \frac{1}{n}\sum_{k=1}^{n} p_k.$$

1. (a) Montrer que pour tout $k \in \mathbb{N}^*$: $\mathbf{V}\left(X_k\right) \leqslant \frac{1}{4}$. En déduire, pour tout $n \in \mathbb{N}^*$, une majoration de $\mathbf{V}(Y_n)$. On admet que la variance d'une somme de variables de Bernoulli indépendantes est la somme des variances.
   (b) En déduire, à l'aide de l'inégalité de Bienaymé-Tchebychev, que, pour tout $\varepsilon > 0$,

$$\mathbf{P}\left(|Y_n - m_n| < \varepsilon\right) \underset{n\to\infty}{\longrightarrow} 1.$$

2. On suppose que la suite $(m_n)_{n\in\mathbb{N}^*}$ converge vers m.
   (a) Soit $\varepsilon > 0$. On suppose $|m_n - m| < \frac{\varepsilon}{2}$. Comparer les événements $\left[|Y_n - m_n| < \frac{\varepsilon}{2}\right]$ et $[|Y_n - m| < \varepsilon]$. En déduire que

$$\mathbf{P}\left(|Y_n - m_n| < \frac{\varepsilon}{2}\right) \leqslant \mathbf{P}\left(|Y_n - m| < \varepsilon\right).$$

   (b) En déduire la convergence en probabilité de la suite $(Y_n)_{n\in\mathbb{N}^*}$.

Exercice 321. ✦ **Convergence de loi discrète vers une loi continue**
Pour tout $n \in \mathbb{N}^*$, on considère la variable aléatoire discrète $X_n \hookrightarrow \mathscr{U}([\![1;n]\!])$. On pose $Y_n = X_n/n$.
Justifier que la suite $(Y_n)_{n\in\mathbb{N}}$ converge en loi vers une variable aléatoire que l'on précisera.
Indication. On pourra utiliser l'expression de la fonction de répartition de X_n,

$$\forall x \in [0;n], \qquad F_n(x) = \frac{\lfloor x \rfloor}{n}.$$

Exercice 322. ✦✦ **Un classique** *D'après EMLyon 2017*
On considère la fonction f définie sur $\mathbb{R}$ par :

$$\forall t \in \mathbb{R}, \quad f(t) = \begin{cases} 0 & \text{si } t < 0 \\ \frac{2}{\pi(1+t^2)} & \text{si } t \geqslant 0 \end{cases}$$

1. Montrer que f est une densité de probabilité.
2. On considère une variable aléatoire réelle X à densité, de densité f.
   (a) Déterminer la fonction de répartition F_X de X.
   (b) La variable aléatoire X admet-elle une espérance ? une variance ?
3. On considère une suite de variables aléatoires réelles $(X_n)_{n\in\mathbb{N}^*}$ à densité, à valeurs strictement positives, mutuellement indépendantes, dont chacune a pour densité f.
   On définit, pour tout n de $\mathbb{N}^*$, les variables aléatoires $M_n = \max(X_1,\ldots,X_n)$ et $Z_n = \frac{n}{M_n}$.
   (a) Déterminer, pour tout n de $\mathbb{N}^*$, la fonction de répartition F_{M_n} de M_n.

(b) Justifier :
$$\forall u \in]0;+\infty[,\quad \mathrm{Arctan}(u)+\mathrm{Arctan}\left(\frac{1}{u}\right)=\frac{\pi}{2}\quad \text{et}\quad \mathrm{Arctan}(u)\underset{u\to 0}{\sim} u.$$

(c) Montrer alors, pour tout n de $\mathbb{N}^*$:
$$\forall x\in]0;+\infty[,\quad \mathbf{P}(Z_n\leqslant x)=1-\left(1-\frac{2}{\pi}\mathrm{Arctan}\left(\frac{x}{n}\right)\right)^n.$$

(d) En déduire que la suite de variables aléatoires $(Z_n)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire à densité dont on reconnaîtra la loi.

⋙ *Pour une version similaire, voir EDHEC 2019 Exercice 2.*

Approximations

Exercice 323. ✦✦ Aiguilles de Buffon

On dispose d'un sol recouvert de parquet, dont les lattes ont une largeur de 20 cm. On dispose aussi de grandes aiguilles, dont la longueur est égale à 20 cm. On admet que si on laisse tomber une aiguille par terre, la probabilité que l'aiguille soit à cheval sur deux lattes est égale à $\frac{2}{\pi}$.

1. Proposer une expérience, à base de lancers d'aiguille, permettant de donner une approximation de $\frac{2}{\pi}$ (dont on déduira une approximation de π).
2. De façon théorique, et en utilisant la valeur de π, combien de lancers faudrait-il réaliser, approximativement, pour obtenir une approximation de $\frac{2}{\pi}$ à moins de 0,01 valable avec une probabilité de 95% ?
   Données : $\Phi(1,645)\simeq 0,95$, $\Phi(1,960)\simeq 0,975$.

Exercice 324. ✦✦✦ Approximation de π via la méthode de Monte Carlo *d'après oraux ESCP 2014*

Soit $(U_n)_{n\in\mathbb{N}^*}$ et $(V_n)_{n\in\mathbb{N}^*}$ deux suites de variables aléatoires définies sur un espace probabilisé $(\Omega,\mathscr{A},P)$ et toutes de même loi uniforme sur $[0,1]$. On suppose que toutes les variables U_n et V_n (pour $n\in\mathbb{N}^*$) sont indépendantes.

1. Pour tout réel $x\in]0,1]$, calculer l'intégrale :
$$J(x)=\int_0^x \frac{1}{\sqrt{t(x-t)}}\,\mathrm{d}t.$$
   On pourra justifier et utiliser le changement de variable (à x fixé) :
$$\varphi :]-\pi/2,\pi/2[\to\mathbb{R},\qquad \theta\mapsto t=\frac{x}{2}\sin\theta+\frac{x}{2}.$$
2. (a) Déterminer la loi de U_n^2.
   (b) Justifier que la variable $U_n^2+V_n^2$ possède une densité h, que l'on exprimera sous forme d'une intégrale.
   (c) Déterminer $h(x)$ pour $x\in[0,1]$.
3. On pose :
$$\forall n\in\mathbb{N}^*,\quad X_n=\begin{cases}1 & \text{si } U_n^2+V_n^2\leqslant 1\\ 0 & \text{sinon.}\end{cases}$$
   Déterminer la loi de X_n.
4. (a) Prouver que la suite $(Z_n)_{n\in\mathbb{N}^*}$ définie par $\forall n\in\mathbb{N}^*, Z_n=\frac{4}{n}\sum_{k=1}^n X_k$, converge en probabilité vers la constante π. C'est-à-dire, pour tout $\varepsilon\in\mathbb{R}_*^+$,
$$\mathbf{P}(|Z_n-\pi|\geqslant\varepsilon)\underset{n\to\infty}{\longrightarrow}0.$$
   (b) Soit $\alpha\in]0,1[$ et $\delta>0$.
   Montrer qu'il existe un entier n_0, qu'on exprimera en fonction de α et δ, tel que
$$\forall n\geqslant n_0,\quad \mathbf{P}(|Z_n-\pi|>\delta)\leqslant\alpha$$
   (c) En déduire un programme Python qui donne une approximation de π.

Les inclassables

Exercice 325. ✦✦ Application de la formule de Stirling *D'après EDHEC 2007*

On considère une suite $(X_n)_{n\in\mathbb{N}^*}$ de variables aléatoires définies sur le même espace probabilisé $(\Omega,\mathscr{A},P)$, mutuellement indépendantes, et qui suivent toutes la loi exponentielle de paramètre 1. On pose $S_n=\sum\limits_{k=1}^n X_k$.

1. Rappeler quelle est la loi suivie par S_n. Donner l'espérance et la variance de S_n.
2. À l'aide du théorème central limite, établir que $\mathbf{P}(S_n\leqslant n)\underset{n\to\infty}{\longrightarrow}\frac{1}{2}$.
3. En déduire la valeur de
$$\lim_{n\to+\infty}\int_0^n \frac{t^{n-1}}{(n-1)!}\mathrm{e}^{-t}\,\mathrm{d}t.$$

4. (a) Utiliser le résultat précédent pour montrer que $\int_0^1 z^{n-1}\mathrm{e}^{-nz}\,\mathrm{d}z \underset{n\to+\infty}{\sim} \frac{n!}{2n^{n+1}}$.

   (b) On admet que $n! \underset{+\infty}{\sim} \sqrt{2\pi n}n^n\mathrm{e}^{-n}$. En déduire un nouvel équivalent de $\int_0^1 z^{n-1}\mathrm{e}^{-nz}\,\mathrm{d}z$.

Exercice 326. ✦✦✦
Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes telles que : $X_n \hookrightarrow \mathscr{B}\left(n, \frac{1}{4}\right)$.

1. Calculer $\mathbf{P}\left(X_n \leqslant \frac{n}{4}\right)$.
2. En utilisant le théorème de la limite centrée déterminer un équivalent de $u_n = \sum_{k=0}^{\lfloor n/4\rfloor} \binom{n}{k} 3^{n-k}$.

Exercice 327. ✦✦✦

1. En appliquant le théorème limite central à une suite de variables aléatoires indépendantes (X_n) suivant toutes une loi de Poisson $\mathscr{P}(1)$, démontrer que
$$\mathrm{e}^{-n}\sum_{k=0}^{n}\frac{n^k}{k!} \underset{n\to\infty}{\longrightarrow} \frac{1}{2}.$$
2. Écrire la formule de Taylor à la fonction exponentielle entre 0 et n, en déduire
$$\int_0^n \mathrm{e}^{-t}t^n\,\mathrm{d}t \underset{n+\infty}{\sim} \frac{n!}{2}.$$

Compléments théoriques sur les différentes convergences

Exercice 328. ✦✦✦ **Convergence presque sûr**
Soient $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires et X une variable aléatoire. Toutes les variables sont définies sur un même espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$. On dit que la suite (X_n) converge presque sûrement vers X si :
$$\mathbf{P}\left(\left\{\omega\in\Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\}\right) = 1.$$
Justifier que si la suite $(X_n)_n$ converge presque sûrement vers X, alors elle converge aussi en probabilité vers X.

Exercice 329. ✦✦✦ **La convergence en probabilité implique la convergence en loi** *d'après oraux HEC*
Soient $(X_n)_{n\in\mathbb{N}^*}$ et X des variables aléatoires définies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ telles que $(X_n)_{n\in\mathbb{N}^*}$ converge en probabilité vers X. On note F_n la fonction de répartition de X_n et F la fonction de répartition de X.
Soit x un point de continuité de F, et soit $\delta > 0$ fixé.

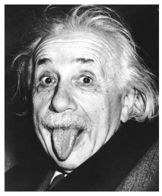
1. Montrer qu'on peut choisir ε tel que $F(x-\varepsilon) > F(x) - \delta$ et $F(x+\varepsilon) < F(x) + \delta$.
2. Montrer que
$$[X_n \leqslant x] \subset [X \leqslant x+\varepsilon] \cup [|X_n - X| > \varepsilon].$$
Montrer de même que
$$[X \leqslant x-\varepsilon] \subset [X_n \leqslant x] \cup [|X_n - X| > \varepsilon].$$
3. En déduire que
$$F_n(x) \leqslant F(x) + \delta + P(|X_n - X| > \varepsilon) \text{ et } F(x) - \delta \leqslant F_n(x) + P(|X_n - X| > \varepsilon).$$
4. Conclure en prouvant que la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers X.

CHAPITRE 16

Compléments sur les fonctions de plusieurs variables, optimisation

Do not worry too much about your difficulties in mathematics, I can assure you that mine are still greater.

ALBERT EINSTEIN



1 Rappels sur les fonctions d'une variable réelle

Avec la continuité

Dans la suite, I est un intervalle de $\mathbb{R}$. Rappelons la définition de l'image d'une fonction $f : \mathrm{I} \to \mathbb{R}$, noté $f(\mathrm{I})$, comme l'ensemble des réels qui admettent un antécédent par f. Autrement dit,

$$f(\mathrm{I}) = \{y \in \mathbb{R} \mid \exists\, x \in \mathrm{I},\ f(x) = y\} = \{f(x) \mid x \in \mathrm{I}\}.$$

Théorème 229 (image d'un segment)

Pour toute application continue $f : \mathrm{I} \to \mathbb{R}$ où I est un segment de $\mathbb{R}$, l'image $f(\mathrm{I})$ est aussi un segment. Alors, f admet un minimum et un maximum et

$$f(\mathrm{I}) = \left[\min_{\mathrm{I}} f, \max_{\mathrm{I}} f\right].$$

Remarque. On dit alors que f est bornée et atteint ses bornes. Il existe α, β tels que

$$\forall\, x \in \mathrm{I}, \quad f(\alpha) \leqslant f(x) \leqslant f(\beta).$$

Avec la dérivée

Théorème 230 (condition nécessaire pour un extrema)

Soit $f : \mathrm{I} \to \mathbb{R}$.

Si
- f admet un extremum local en a,
- f est dérivable en a, et
- a appartient à un *intervalle ouvert* inclus dans I, **alors** $f'(a) = 0$.

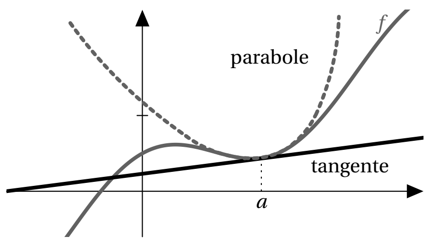
Remarque. La réciproque est fausse : tout point critique ne donne pas un extremum. La fonction $x \in \mathbb{R} \mapsto x^3$ en 0 fournit un contre-exemple.

Avec les dérivées successives

Théorème 231 (formule de Taylor-Young)

Soient $f \in \mathscr{C}^n(\mathrm{I})$, $a \in \mathrm{I}$ et $n \in \mathbb{N}$. Au voisinage du réel a,

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + o\big((x-a)^n\big).$$



Exemples.

- $n = 1$

$$f(x) = \underbrace{f(a) + f'(a)(x-a)}_{\text{équation de la tangente}} + o(x-a).$$

- $n = 2$

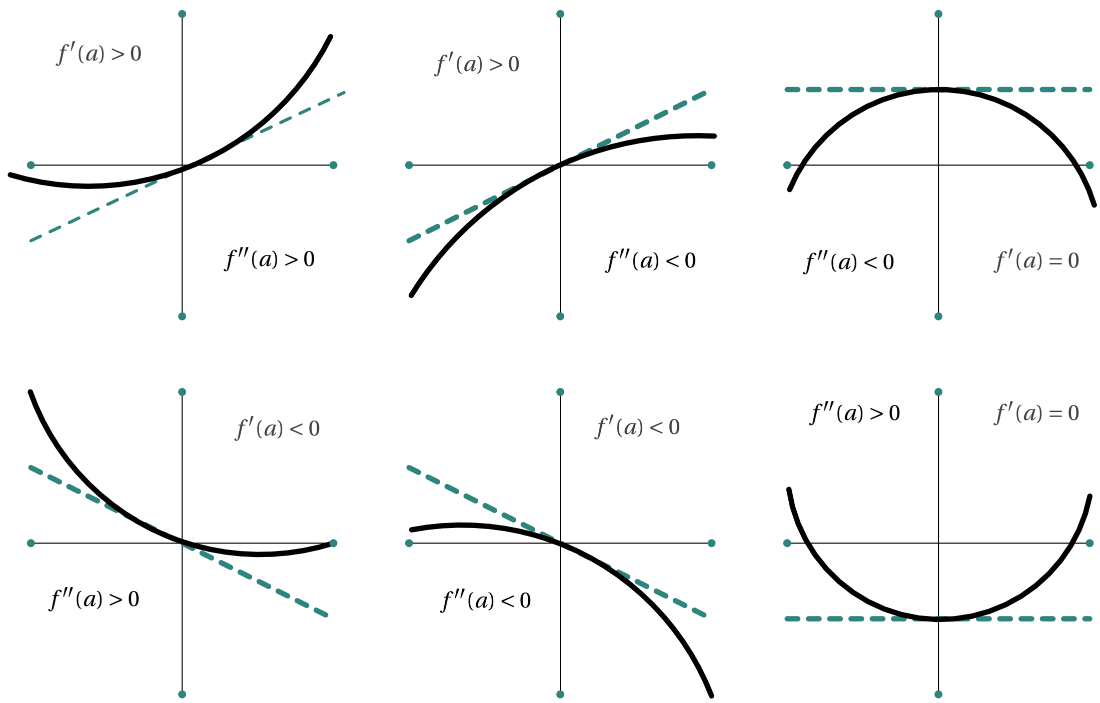
$$f(x) = \underbrace{f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2}_{\text{équation d'une parabole si } f''(a) \neq 0} + o_a\big((x-a)^2\big).$$

Allure du graphe d'une fonction au voisinage d'un point

En reprenant le cas $n = 2$, au voisinage de a, la courbe représentative de f est « proche » de la courbe d'équation

$$y = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2.$$

Lorsque $f''(a) \neq 0$, la courbe est une parabole.
Distinguons suivant le signe de $f'(a)$ et $f''(a)$. En gras, la parabole, en pointillés, la tangente.



Exercice 330

✦✦ **Égalité de Taylor-Lagrange**
Soient f une fonction de classe $\mathscr{C}^2$ sur un intervalle I et $a \in$ I. Justifier que pour tout $x \in$ I, il existe c compris entre a et x tel que

$$f(x) = f(a) + f'(a)(x-a) + \frac{1}{2}(x-a)^2 f''(c).$$

On pourra introduire la fonction définie par $\varphi(t) = f(x) - f(t) - f'(t)(x-t) + (x-t)^2\lambda/2$ *pour un réel* λ *bien choisi.*

Avec la convexité

Théorème 232 (condition suffisante pour un minimum global)

Soient $f : \mathrm{I} \to \mathbb{R}$ et a appartenant à un intervalle ouvert inclus dans I.

Si
- f est convexe.
- f est dérivable en a.
- a est un point critique ($f'(a) = 0$).

Alors, f admet un minimum global en a.

Remarque. On a un énoncé similaire dans le cas d'une fonction concave où on obtient un maximum global.

2 Fonctions définies sur une partie de $\mathbb{R}^n$

2.1 Fermés, ouverts et bornés

Dans la suite, on considère le produit scalaire canonique sur $\mathbb{R}^n$ et on note $\|\cdot\|$ la norme associée.

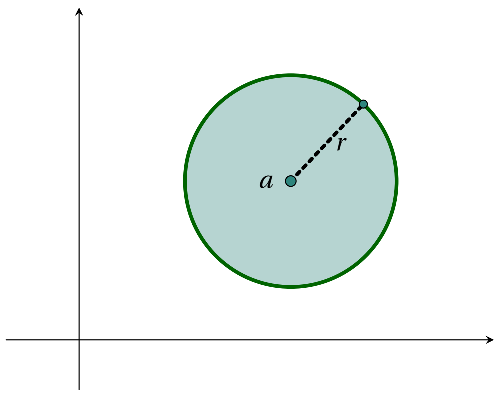
Définition 233 (boule ouverte)

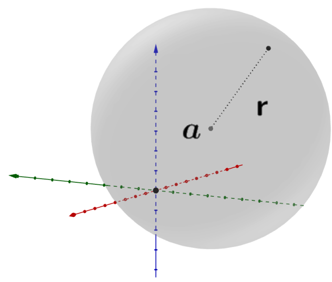
Soient $a \in \mathbb{R}^n$ et $r \in \mathbb{R}_*^+$, on définit la **boule ouverte** de centre a et de rayon r par :

$$\mathscr{B}(a, r) = \left\{x \in \mathbb{R}^n \;\middle|\; \|x - a\| < r\right\}.$$

C'est l'ensemble des points de $\mathbb{R}^n$ situés à une distance strictement inférieure à r du point a.

À gauche, une boule ouverte dans $\mathbb{R}^2$ et à droite dans $\mathbb{R}^3$. Il faut exclure dans les deux cas la « frontière ».





Définition 234 (partie ouverte)

Une partie $\mathcal{O}$ de $\mathbb{R}^n$ est dite **ouverte** si, en chaque point de $\mathcal{O}$, il existe une boule ouverte centrée en ce point et contenue dans $\mathcal{O}$. Autrement dit,

$$\forall\, a \in \mathcal{O}, \quad \exists\, r \in \mathbb{R}_*^+, \qquad \mathcal{B}(a,r) \subset \mathcal{O}.$$

Exemples.
→ $\mathbb{R}^n$ est une partie ouverte puisqu'une boule ouverte en un point quelconque de $\mathbb{R}^n$ est toujours contenue dans $\mathbb{R}^n$.
→ Toute boule ouverte est un ouvert.
→ $\mathbb{R}^n$ privé d'un ensemble fini de points est ouvert.
→ Soient $\mathrm{I}_1, \mathrm{I}_2, \dots, \mathrm{I}_n$, n intervalles ouverts de $\mathbb{R}$. La partie $\mathrm{I}_1 \times \mathrm{I}_2 \times \dots \times \mathrm{I}_n$ est une partie ouverte de $\mathbb{R}^n$ (on parle de pavé ouvert).

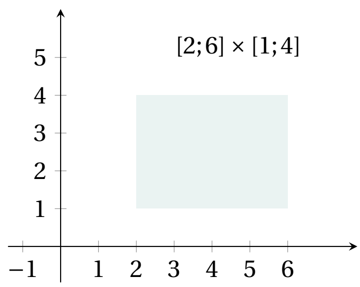
Remarque. Une union quelconque d'ouverts est un ouvert. Une intersection *finie* d'ouverts reste un ouvert.

Définition 235 (partie fermée)

Une partie F de $\mathbb{R}^n$ est dite **fermée** de $\mathbb{R}^n$ si son complémentaire $\overline{\mathrm{F}}$ est un ouvert de $\mathbb{R}^n$.

Remarque. En reprenant la remarque précédente, et par passage au complémentaire, une union finie de parties fermées reste une partie fermée et une intersection quelconque de parties fermées est une partie fermée.

Exemple. On montre que si I_1, I_2, …, I_n sont des intervalles fermés de $\mathbb{R}$. La partie $\mathrm{I}_1 \times \mathrm{I}_2 \times \dots \times \mathrm{I}_n$ est une partie fermée de $\mathbb{R}^n$.



Proposition 236 (conditions suffisantes pour une partie ouverte/fermée)

Soient $\varphi : \mathbb{R}^n \to \mathbb{R}$, une fonction continue sur $\mathbb{R}^n$ et $r \in \mathbb{R}$.
→ La partie $\{x \in \mathbb{R}^n \mid \varphi(x) < r\}$ est un ouvert de $\mathbb{R}^n$.
→ La partie $\{x \in \mathbb{R}^n \mid \varphi(x) \leqslant r\}$ est un fermé de $\mathbb{R}^n$.
→ La partie $\{x \in \mathbb{R}^n \mid \varphi(x) = r\}$ est un fermé de $\mathbb{R}^n$.

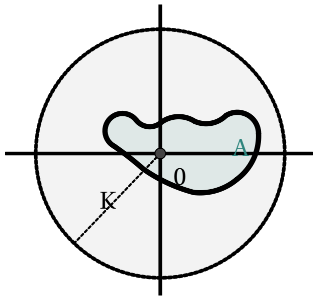
Exemples.
- Pour $a \in \mathbb{R}^n$ et $r \in \mathbb{R}^+$, on définit la partie de $\mathbb{R}^n$: $\overline{\mathcal{B}}(a,r) = \{x \in \mathbb{R}^n \mid \|x-a\| \leqslant r\}$.
- De même, le **disque** $\mathcal{D}(a,r) = \{x \in \mathbb{R}^n \mid \varphi(x) = r\} = \{x \in \mathbb{R}^n \mid \|x-a\| = r\}$ est une partie fermée.

Définition 237 (partie bornée)

Soit A une partie de $\mathbb{R}^n$.
La partie A est dite **bornée** s'il existe $\mathrm{K} \in \mathbb{R}^+$ tel que pour tout $x \in \mathrm{A}$, $\|x\| \leqslant \mathrm{K}$.

Remarque. Autrement dit, une partie A de $\mathbb{R}^n$ est bornée si elle est incluse dans une boule (fermée) de centre l'origine :

$$\exists K \in \mathbb{R}_*^+, \qquad A \subset \overline{\mathscr{B}}(0, K).$$



Exercice 331

Exemples

✦ 1. Parmi les parties suivantes, préciser les parties ouvertes, fermées, bornées.

$$A_1 = \left\{(x, y) \in \mathbb{R}^2 \mid x + 2y = 5\right\}, \quad A_2 = \left\{(x, y) \in \mathbb{R}^2 \mid x > 0, y > 0 \text{ et } x + y < 1\right\}$$
$$A_3 = \left\{x \in \mathbb{R}^n \mid \langle u, x\rangle \geqslant 0\right\}, \quad A_4 = \left\{x \in \mathbb{R}^n \mid \langle u, x\rangle = 0\right\} \quad \text{où} \quad u \in \mathbb{R}^n \setminus \{0\}.$$

✦✦ 2. Justifier que tout sous-espace vectoriel d'un espace vectoriel de $\mathbb{R}^n$ est un fermé.
On pourra exprimer le sous-espace comme intersection de parties du type A_4.

2.2 Rappels et compléments sur le cas $\mathscr{C}^1$

Les définitions vues au premier semestre s'étendent à une partie ouverte de $\mathbb{R}^n$. Ainsi, une fonction $f : \mathcal{O} \to \mathbb{R}$ est de classe $\mathscr{C}^1$ sur $\mathcal{O}$ si les dérivées partielles sont continues en chaque point de $\mathcal{O}$.
On rappelle que pour tout $a \in \mathcal{O}$, la fonction f admet en a un unique développement limité à l'ordre 1.
C'est-à-dire, il existe $\varepsilon : \mathcal{O} \to \mathbb{R}$ telle que, pour $h \in \mathbb{R}^n$ avec $a + h \in \mathcal{O}$,

$$f(a+h) = f(a) + \left\langle \nabla(f)(a), h\right\rangle + \|h\|\varepsilon(h) \quad \text{et} \quad \varepsilon(h) \underset{h\to 0}{\longrightarrow} 0.$$

3 Fonctions de classe $\mathscr{C}^2$

3.1 Définitions et exemples

Définition 238 (dérivées partielles d'ordre 2)

Soient un ouvert $\mathcal{O}$ de $\mathbb{R}^n$, une fonction $f : \mathcal{O} \to \mathbb{R}$, $a \in \mathcal{O}$ et $i, j \in [[1, n]]$.
On suppose que la dérivée partielle première $\partial_j f$ est définie sur $\mathcal{O}$. Si la dérivée partielle $\partial_i\left(\partial_j f\right)$ est définie en a, on dit que f admet une **dérivée partielle d'ordre 2 d'ordre** (i, j) **en** a et on note $\partial^2_{i,j} f(a)$ pour $\partial_i\left(\partial_j f\right)(a)$.

Remarque. Si la dérivée partielle $a \in \mathcal{O} \mapsto \partial_i\left(\partial_j f\right)$ est définie sur $\mathcal{O}$, on dit que f admet, sur $\mathcal{O}$, une dérivée partielle d'ordre 2 d'ordre (i, j) et on note

$$\partial^2_{i,j} f = \partial_i\left(\partial_j f\right).$$

On obtient alors une nouvelle fonction de plusieurs variables $\partial^2_{i,j} f : \mathcal{O} \mapsto \mathbb{R}$.

Exercice 332

✧ Calculer $\partial^2_{1,2} f$ et $\partial^2_{1,1} f + \partial^2_{2,2} f + \partial^2_{3,3} f$ dans les deux cas suivants :

I $\quad f : (x_1, x_2, x_3) \in \mathbb{R}^3 \mapsto {x_1}^2 + {x_2}^2 - 2{x_3}^2 - x_2 + x_1 + x_1 x_2$ $\qquad$ II $\quad f : x \in \mathbb{R}^3 \setminus \{0_{\mathbb{R}^3}\} \mapsto 1/\|x\|$.

Définition 239 (fonction de classe $\mathscr{C}^2$)

Soient un ouvert $\mathcal{O}$ de $\mathbb{R}^n$ et une fonction $f : \mathcal{O} \to \mathbb{R}$.
On dit que f est une fonction de **classe** $\mathscr{C}^2$ sur $\mathcal{O}$

Si
- → La fonction f admet des dérivées partielles d'ordre 2 d'indice (i,j) pour tout $(i,j) \in [\![1,n]\!]^2$.
- → Les dérivées partielles sont des fonctions continues de $\mathcal{O}$ dans $\mathbb{R}$.

Exemple. Les fonctions polynomiales sont de classe $\mathscr{C}^2$ sur $\mathbb{R}^n$ car les dérivées partielles d'ordre 2 existent et restent polynomiales (donc continues sur $\mathbb{R}^n$).

Remarque. Comme pour le cas $\mathscr{C}^1$, les combinaisons linéaires, les produits et quotients de fonctions de classe $\mathscr{C}^2$ sur un ouvert $\mathcal{O}$ ou encore la composition par une fonction d'une variable de classe $\mathscr{C}^2$ restent de classe $\mathscr{C}^2$ sur $\mathcal{O}$.

Définition 240 (la matrice hessienne)

Soit $f : \mathcal{O} \subset \mathbb{R}^n \to \mathbb{R}$ admettant des dérivées partielles d'ordre 2 pour tout couple $(i,j) \in [\![1;n]\!]^2$.
La **matrice hessienne** de f au point $a \in \mathcal{O}$, notée $\nabla^2 f(a) \in \mathscr{M}_n(\mathbb{R})$, est définie par

$$\nabla^2 f(a) = \left[\partial^2_{i,j} f(a)\right]_{(i,j)\in[\![1,n]\!]^2} = \begin{bmatrix} \partial^2_{1,1} f(a) & \partial^2_{1,2} f(a) & \cdots & \partial^2_{1,n} f(a) \\ \partial^2_{2,1} f(a) & \partial^2_{2,2} f(a) & \cdots & \partial^2_{2,n} f(a) \\ \vdots & \vdots & \vdots & \vdots \\ \partial^2_{n,1} f(a) & \partial^2_{2,2} f(a) & \cdots & \partial^2_{n,n} f(a) \end{bmatrix}.$$

Exercice 333

✦ Soit f la fonction définie par *d'après EDHEC 2020*

$$\forall (x,y,z) \in \mathbb{R}^3, \qquad f(x,y,z) = x e^{x(y^2+z^2+1)}.$$

1. Montrer que f est de classe $\mathscr{C}^2$ sur $\mathbb{R}^3$.
2. Déterminer le seul point critique a de f.
3. Former la hessienne de f au point a et vérifier qu'elle est diagonale.

3.2 Les énoncés

Théorème 241 (de Schwarz)

Soit f définie sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$.

Si f est de classe $\mathscr{C}^2$ sur $\mathcal{O}$,

alors pour tout couple $(i,j) \in [\![1,n]\!]^2$ et pour tout point $a \in \mathcal{O}$, on a $\partial^2_{j,i} f(a) = \partial^2_{i,j} f(a)$.

Remarque. D'après le théorème de Schwarz, la matrice hessienne d'une fonction de classe $\mathscr{C}^2$ est symétrique réelle. En particulier, la matrice est diagonalisable dans une base orthonormée (théorème spectral).

3.3 Forme quadratique et développement limité d'ordre 2

Préliminaires

- La forme quadratique associée à une matrice symétrique $A \in \mathscr{M}_n(\mathbb{R})$ est l'application définie sur $\mathbb{R}^n$ par

$$q(h) = {}^t\!HAH$$

où H est la matrice des coordonnées de h dans la base canonique de $\mathbb{R}^n$.

- Nous avons vu aussi que le signe de la forme quadratique est précisé par le signe des valeurs propres de la matrice associée. Plus précisément

  i) $\forall\, u \in \mathrm{E}, \qquad q(u) \geqslant 0 \quad \Longleftrightarrow \quad \mathrm{Sp(A)} \subset \mathbb{R}^+$;

  ii) $\forall\, u \in \mathrm{E}, \qquad q(u) \leqslant 0 \quad \Longleftrightarrow \quad \mathrm{Sp(A)} \subset \mathbb{R}^-$.

- Soient $\mathscr{O}$ un ouvert de $\mathbb{R}^n$, $f : \mathscr{O} \to \mathbb{R}$ de classe $\mathscr{C}^2$.

On peut considérer la forme quadratique associée à la matrice hessienne de f en a, notée $q_a : \mathbb{R}^n \to \mathbb{R}$ et définie par

$$\forall\, h = (h_1, \dots, h_n) \in \mathbb{R}^n, \qquad q_a(h) = {}^t\mathrm{H}\left(\nabla^2 f(a)\right)\mathrm{H}, \quad \text{où } \mathrm{H} = \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix}.$$

Cela s'écrit aussi :
$$q_a(h) = \sum_{(i,j)\in[\![1,n]\!]^2} \partial^2_{i,j} f(a)\, h_i h_j = \sum_{i=1}^{n} \partial^2_{ii} f(a)\, {h_i}^2 + 2\sum_{i<j} \partial^2_{i,j} f(a)\, h_i h_j.$$

- Soit $h \in \mathbb{R}^n$ tel que le segment $[a, a+h]$ soit inclus dans $\mathscr{O}$. On pose pour tout $t \in [0;1]$,

$$g_{a,h}(t) = f(a + th).$$

En reprenant le résultat du premier semestre sur les dérivées directionnelles, on montre que $g_{a,h}$ est une fonction de classe $\mathscr{C}^2$ sur $[0;1]$ avec pour tout $t \in [0;1]$,

$$\boxed{g'_{a,h}(t) = \left\langle \nabla f(a+th), h \right\rangle = \sum_{i=1}^{n} h_i \partial_i f(a+th) \quad \text{où} \quad h = (h_1, \dots, h_n) \quad \text{et} \quad g''_{a,h}(t) = q_{a+th}(h).}$$

Ensuite, en reprenant le résultat de l'exercice 330, p.175, il vient

$$f(a+h) = f(a) + \left\langle \nabla f(a), h \right\rangle + \frac{1}{2} q_{a+\theta h}(h).$$

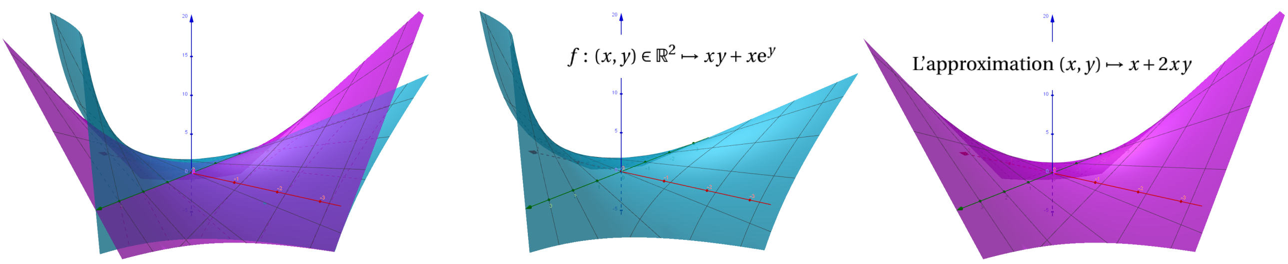
Développement limité d'ordre 2

Théorème 242 (développement limité d'ordre 2)

Soient $\mathscr{O}$ un ouvert de $\mathbb{R}^n$, $f : \mathscr{O} \to \mathbb{R}$ de classe $\mathscr{C}^2$.
Alors pour tout $a \in \mathscr{O}$, il existe un voisinage $\mathscr{V}$ de $0_{\mathbb{R}^n}$ dans $\mathbb{R}^n$, une fonction $\varepsilon : \mathscr{V} \to \mathbb{R}$ telle que pour tout $h \in \mathscr{V}$,
$$f(a+h) = f(a) + \langle \nabla f(a), h \rangle + \frac{1}{2} q_a(h) + \|h\|^2 \varepsilon(h) \quad \text{et} \quad \varepsilon(h) \underset{h\to 0}{\longrightarrow} 0$$
où q_a est la forme quadratique associée à $\nabla^2 f(a)$.

Exemple. Posons $f : (x, y) \in \mathbb{R}^2 \mapsto xy + xe^y$. La fonction f est de classe $\mathscr{C}^2$ sur $\mathbb{R}^2$. On vérifie que le développement limité d'ordre 2 en $(0,0)$ est

$$f(h_1, h_2) = f(0,0) + \langle \nabla f(0,0), h \rangle + \frac{1}{2} q_{(0,0)}(h) + \|h\|^2 \varepsilon(h) = 0 + \left\langle (1,0), (h_1, h_2) \right\rangle + \frac{1}{2} \cdot 4h_1h_2 + \|h\|^2 \varepsilon(h) = h_1 + 2h_1h_2 + \|h\|^2 \varepsilon(h).$$



4 Applications à l'optimisation

4.1 Rappels : extrema locaux/globaux

Définition 243 (extrema locaux)

Soient $f : \mathbb{R}^n \to \mathbb{R}$ et $a \in \mathbb{R}^n$.

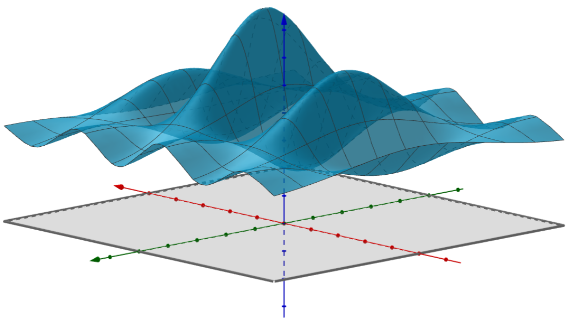
- On dit que f a un **maximum local** en a s'il existe $r > 0$ tel que
$$\forall x \in \mathbb{R}^n, \qquad \Big(\|x-a\| < r \quad \Rightarrow \quad f(a) \geqslant f(x) \Big).$$
- On dit que f a un **minimum local** en a s'il existe $r > 0$ tel que
$$\forall x \in \mathbb{R}^n, \qquad \Big(\|x-a\| < r \quad \Rightarrow \quad f(a) \leqslant f(x) \Big).$$
- On dit que f a un **extremum local** si f a un maximum local ou un minimum local.

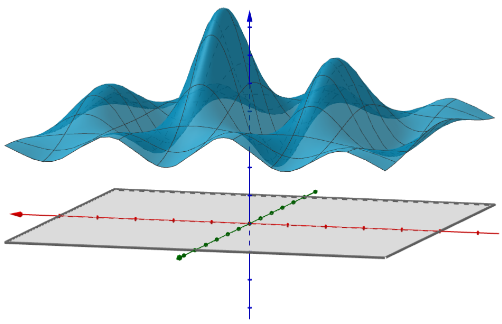
4.2 Condition d'existence d'extremum

Théorème 244 (sur un fermé borné)

Une fonction continue sur une partie fermée bornée admet un maximum global et un minimum global.

Remarque. On peut traduire mathématiquement. Soit $f : \mathcal{O} \subset \mathbb{R}^n \to \mathbb{R}$ continue. Il existe $\alpha, \beta \in \mathcal{O}$ tels que
$$\forall x \in \mathcal{O}, \quad f(\alpha) \leqslant f(x) \leqslant f(\beta).$$





4.3 Condition nécessaire d'ordre 1

L'énoncé du premier semestre s'étend au cas d'un ouvert en reprenant la preuve.

Théorème 245 (condition d'ordre 1)

Soient $f : \mathcal{O} \to \mathbb{R}$ de classe $\mathscr{C}^1$ sur $\mathcal{O}$ et $a \in \mathcal{O}$.

Si
- $\mathcal{O}$ est un ouvert de $\mathbb{R}^n$.
- f a un extremum en $a \in \mathbb{R}^n$.

Alors a est un point critique de f, c'est-à-dire $\nabla f(a) = 0$.

⚠ **Attention.** Comme pour le cas d'une variable :
- L'énoncé précédent n'est valable que sur un ouvert.
- La réciproque est fausse. On parle alors de **point col**.

4.4 Condition suffisante d'ordre 2

Théorème 246 (condition d'ordre 2)

Si a est un point critique de f :

- **Si** $\text{Sp}\left(\nabla^2 f(a)\right) \subset \mathbb{R}_+^*$, **alors** f admet un minimum local en a.
- **Si** $\text{Sp}\left(\nabla^2 f(a)\right) \subset \mathbb{R}_-^*$, **alors** f admet un maximum local en a.
- **Si** $\text{Sp}\left(\nabla^2 f(a)\right)$ contient deux réels non nuls de signes distincts, **alors** f n'admet pas d'extremum en a.

Remarque. Les réciproques sont fausses. Par exemple, on peut considérer la fonction définie sur $\mathbb{R}^2$ par $f(x,y) = x^4 + y^4$.

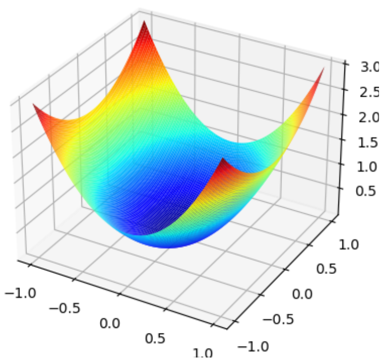
Illustration dans $\mathbb{R}^2$

Posons pour $\alpha, \beta \in \mathbb{R}$, $f_{\alpha,\beta}(x,y) = \alpha x^2 + \beta y^2$. La matrice Hessienne en point critique $(0,0)$ est

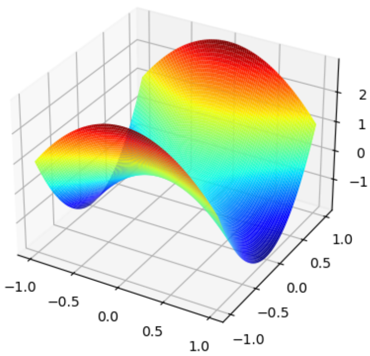
$$\mathrm{H}_{\alpha,\beta} = \begin{bmatrix} 2\alpha & 0 \\ 0 & 2\beta \end{bmatrix}.$$

Trois cas sont bien à distinguer :

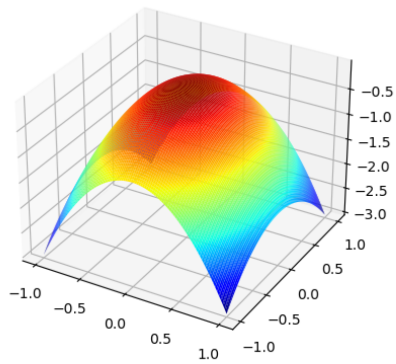
I. $\alpha, \beta \in \mathbb{R}_*^+$,



II. α, β de signes opposés,



III. $\alpha, \beta \in \mathbb{R}_*^-$.



Exemples

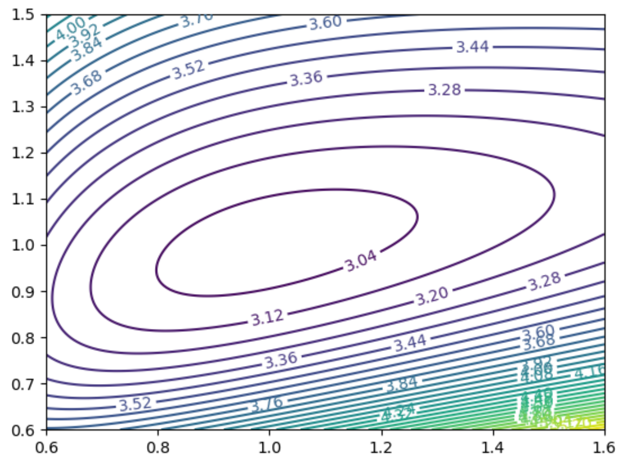
Exemple. On considère la fonction f définie sur l'ouvert $\mathbb{R}_+^* \times \mathbb{R}_+^*$ par :

$$\forall (x,y) \in \mathbb{R}_+^* \times \mathbb{R}_+^*, \quad f(x,y) = \frac{x}{y^2} + y^2 + \frac{1}{x}.$$

Justifions que la fonction f admet un minimum local en $(1,1)$. Commençons par tracer les courbes de niveaux :

Editeur

```
import matplotlib.pyplot as plt
import numpy as np

def f(x,y):
    return x/y**2+y**2+1/x

x=np.linspace(0.4,1.6,200)
y=np.linspace(0.5,1.5,200)
X , Y = np.meshgrid(x,y)
Z = f(X,Y)


graphe = plt.contour(X,Y,Z,20)
# Le 20 pour 20 lignes de niveau
plt.show()
```



On vérifie que f est de classe $\mathscr{C}^2$ sur $\mathbb{R}_+^* \times \mathbb{R}_+^*$. La matrice hessienne de f au point $(1,1)$ a deux valeurs propres strictement positives. On a un minimum local en $(1,1)$.

Exercice 334

✦ *D'après Oraux HEC E*

Soit f la fonction définie sur $\mathcal{O} = \mathbb{R}_+^* \times \mathbb{R}$ par :

$$\forall (x, y) \in \mathbb{R}_+^* \times \mathbb{R}, \quad f(x, y) = x\left((\ln x)^2 + 2y^2\right).$$

1. Démontrer qu'il existe dans $\mathcal{O}$ un unique point-col pour f.
2. La fonction f admet-elle sur $\mathcal{O}$ un maximum global ? un minimum global ?

• Complétons l'étude de la dimension 2 avec *les notations de Monge* (hors programme).
Soit f une fonction de classe $\mathscr{C}^2$ sur un ouvert $\mathcal{O}$ de $\mathbb{R}^2$. Pour $(a, b) \in \mathbb{R}^2$, on notera $\mathrm{H}_{(a,b)}$ la matrice hessienne de l'application au point (a, b).

$$\mathrm{H}_{(a,b)} = \begin{bmatrix} r & s \\ s & t \end{bmatrix}.$$

Exercice 335

✦✦ Supposons que $(a, b) \in \mathcal{O}$ est un point critique de f. Justifier que

- si $rt - s^2 < 0$, alors $f(a, b)$ est un point col.
- si $rt - s^2 > 0$, alors $f(a, b)$ est un extremum local. Dans ce cas, si $r < 0$, alors $f(a, b)$ est un maximum local et, si $r > 0$, alors $f(a, b)$ est un minimum local.

Remarque. Nous verrons au chapitre suivant le cas d'une fonction définie sur un fermé borné.

4.5 Convexité, condition suffisante dans le cas d'extremum global

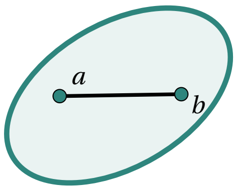
Définition 247 (partie convexe)

Une partie C de $\mathbb{R}^n$ est dite **convexe** si, pour tout couple (a, b) d'éléments de C, le segment $[a, b]$ est tout entier inclus dans C. Autrement dit, C est convexe lorsque

$$\forall\, a, b \in \mathrm{C}, \quad \forall \lambda \in [0, 1], \quad \lambda a + (1 - \lambda) b \in \mathrm{C}.$$

Exemple. $\mathbb{R}^n$ est une partie convexe.

Graphiquement, tout segment dont les extrémités sont dans C est inclus dans C.



Théorème 248 (convexité, extremum global)

Soient $\mathcal{O}$ est un ouvert convexe de $\mathbb{R}^n$, $f : \mathcal{O} \to \mathbb{R}$ et a,un point critique de f.

- **Si** pour tout $x \in \mathcal{O}$, $\mathrm{Sp}\left(\nabla^2 f(x)\right) \subset \mathbb{R}^+$, **alors** f admet un minimum global en a.
- **Si** pour tout $x \in \mathcal{O}$, $\mathrm{Sp}\left(\nabla^2 f(x)\right) \subset \mathbb{R}^-$, **alors** f admet un maximum global en a.

Exercice 336

✦✦✦ **Preuve**
On conserve les mêmes notations et hypothèses. Soit $t \in \mathrm{I} \setminus \{0\}$. À l'aide des résultats préliminaires sur la fonction $g_{a,h}$, justifier que

$$f(a + th) = f(a) + \langle \nabla f(a), th \rangle + \int_0^t (t - u) q_{a+uh}(h)\, \mathrm{d}u.$$

Conclure sur le théorème.

Exemple. Soit $f(x,y,z)=11x^2+3y^2+3z^2-10xy+10xz-6yz+1$.
→ La fonction f est de classe $\mathscr{C}^2$ sur $\mathbb{R}^3$ et on vérifie que la matrice hessienne en (x,y,z) est

$$\nabla^2 f(x,y,z)=2\begin{bmatrix} 11 & -5 & 5 \\ -5 & 3 & -3 \\ 5 & -3 & 3 \end{bmatrix}.$$

→ Utilisons Python pour obtenir rapidement le spectre :

Editeur

```
import numpy as np

H= 2*np.array
    ([[11,-5,5],[-5,3,-3],[5,-3,3]])

Sp=np.linalg.eigvals(H)
```

Console

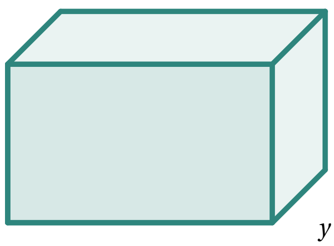
```
>>> # script executed
>>> Sp
array([32.,  2.,  0.])
```

→ On vérifie que f admet un unique point critique donné par $\mathrm{A}=(1,1,1)$.
→ Concluons : en tout point $(x,y,z)\in\mathbb{R}^3$, la matrice hessienne a un spectre inclus dans $\mathbb{R}^+$. On peut donc en déduire que f admet en A un minimum global.

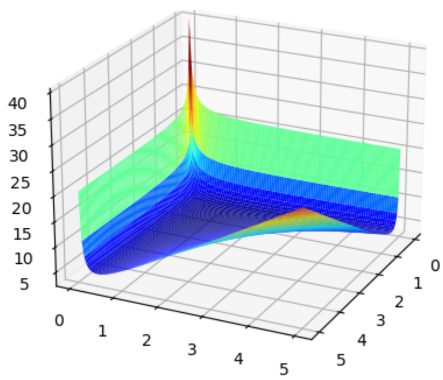
4.6 Exemple détaillé sur un ouvert

On souhaite minimiser la surface d'un pavé de volume 1. Notons x, y, z les longueurs des cotés. L'aire de la surface a pour expression $2(xy+xz+yz)$. Le volume vaut $xyz=1$ et $z=1/(xy)$. Finalement, on cherche à minimiser sur $\mathscr{O}=\mathbb{R}_*^{+2}$, la fonction f définie par

$$f(x,y)=xy+xz+yz=xy+\frac{1}{x}+\frac{1}{y}.$$



- Les dérivées successives.
- Étude du ou des points critiques.
- Étude locale au niveau du point critique.
- Étude globale. Deux méthodes : par continuité sur un fermé borné / par un argument de convexité.
- Astuce par une inégalité de convexité.



Exercices - TD

Exercice 337. ✦ *Oraux HEC 2008*

Soit f la fonction définie par :

$$\forall (x;y) \in \mathbb{R}^2, \qquad f(x,y) = x^3 + y^3 - 3xy + 1.$$

1. Calculer les dérivées partielles d'ordre 1 et d'ordre 2 de f.
2. Déterminer les points critiques de f.
3. La fonction f a-t-elle des extrema locaux? globaux?

Exercice 338. ✦ *EML Lyon2020 E*

On considère la fonction F de classe $\mathscr{C}^2$ sur l'ouvert $]0,+\infty[^2$. définie par :

$$F(x,y) = x^2y + x^2 - \frac{y^2}{2} - 2x.$$

1. Calculer les dérivées partielles d'ordre 1 de F en tout point (x,y) de $]0,+\infty[^2$.
2. Montrer que la fonction F admet un unique point critique du type (α,α^2) où $\alpha \in \mathbb{R}$.
3. Écrire la matrice hessienne, notée H, de la fonction F au point (α,α^2).
4. Montrer que la matrice H admet deux valeurs propres distinctes, notées λ_1 et λ_2, vérifiant $\lambda_1\lambda_2 = -6\alpha^2 - 2$.
5. La fonction F présente-t-elle des extrema locaux sur $]0,+\infty[^2$?

Exercice 339. ✦✦ *EML Lyon 2007*

On considère l'application $f : [0;+\infty[\to \mathbb{R}$ définie par

$$f(x) = \begin{cases} \frac{\ln(1+x)}{x} & \text{si } x > 0 \\ 1 & \text{si } x = 0. \end{cases}$$

1. Vérifier que f est de classe $\mathscr{C}^1$ sur $\mathbb{R}^+$.
2. On définit les fonctions $F : [0;+\infty[\to \mathbb{R}$ et $G :]0;+\infty[^2 \to \mathbb{R}$ par

$$F(x) = \int_0^x f(t)\,dt \quad \text{et} \quad G(x,y) = F(xy) - F(x) - F(y).$$

   (a) Montrer que G est de classe $\mathscr{C}^2$ sur $]0;+\infty[^2$, puis exprimer, pour tout $(x,y) \in]0;+\infty[^2$, les dérivées partielles premières et secondes de G en (x,y) en fonction de x, y, $f(x)$, $f(y)$, $f(xy)$, $f'(x)$, $f'(y)$, $f'(xy)$.
   (b) Établir que G admet $(1,1)$ comme unique point critique.
   (c) Est-ce que G admet un extremum local?

Exercice 340. ✦✦ *d'après EDHEC*

Soient $n \in \mathbb{N} \setminus \{0;1\}$ et la fonction f définie sur $\mathbb{R}^n$ par

$$\forall (x_1,x_2,\ldots,x_n) \in \mathbb{R}^n, f(x_1,x_2,\ldots,x_n) = \sum_{k=1}^n {x_k}^2 + \left(\sum_{k=1}^n x_k\right)^2 - \sum_{k=1}^n x_k.$$

1. (a) Montrer que f est de classe $\mathscr{C}^2$ sur $\mathbb{R}^n$.
   (b) Calculer les dérivées premières et secondes de f sur $\mathbb{R}^n$.
2. (a) Déterminer le seul point critique $(a_1,a_2,\ldots,a_n)$ de f sur $\mathbb{R}^n$.
   (b) Vérifier que la hessienne de f en ce point est la matrice $A_n = 2(I_n + J_n)$, où I_n désigne la matrice unité de $\mathscr{M}_n(\mathbb{R})$ et J_n la matrice de $\mathscr{M}_n(\mathbb{R})$ dont tous les éléments sont égaux à 1.
3. (a) Déterminer le rang de J_n. En déduire que 0 est valeur propre de J_n et déterminer la dimension du sous-espace propre associé.
   On pose :

$$U = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \in \mathscr{M}_{n,1}(\mathbb{R}) \quad \text{et} \quad H = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix} \in \mathscr{M}_{n,1}(\mathbb{R}).$$

   (b) Calculer le produit J_nU.
   (c) À l'aide des questions précédentes, donner les valeurs propres de J_n puis celles de A_n.

4. (a) Montrer que, pour tout $\mathrm{H} \in \mathbb{R}^n$ non nul, on a : ${}^t\mathrm{H}\mathrm{A}_n\mathrm{H} > 0$.
   (b) En déduire que f admet un minimum local en $(a_1, a_2, \dots, a_n)$ et vérifier que ce minimum est égal à $-\dfrac{n}{4(n+1)}$.

Exercice 341. ✦✦ **Dérivation d'une intégrale à paramètre** *d'après Ecricome*

On considère la fonction définie sur $\mathscr{F} = [0, +\infty[\times [0, +\infty[$ par

$$f(x, t) = \mathrm{e}^{-t^2}\sqrt{1+xt}.$$

1. (a) Vérifier que f est de classe $\mathscr{C}^2$ sur $\mathscr{F}$ et donner $\partial_1 f(x, t)$ et $\partial^2_{1,1} f(x, t)$.
   (b) Justifier que pour tout $(x, t) \in \mathscr{F}$,
   $$\left|\partial^2_{1,1} f(x, t)\right| \leqslant \frac{t^2}{4}\mathrm{e}^{-t^2}.$$
2. Montrer que pour tout réel α strictement positif, l'intégrale $\displaystyle\int_0^{+\infty} t^\alpha \mathrm{e}^{-t^2}\,\mathrm{d}t$ est convergente.
3. En déduire, pour tout réel x positif, la convergence des intégrales
$$\int_0^{+\infty} \mathrm{e}^{-t^2}\sqrt{1+xt}\,\mathrm{d}t \quad \text{et} \quad \int_0^{+\infty} \frac{t\mathrm{e}^{-t^2}}{\sqrt{1+xt}}\,\mathrm{d}t.$$
4. On considère maintenant la fonction g définie sur $[0, +\infty[$ par
$$g(x) = \int_0^{+\infty} f(x, t)\,\mathrm{d}t = \int_0^{+\infty} \mathrm{e}^{-t^2}\sqrt{1+xt}\,\mathrm{d}t.$$
   (a) Sans chercher à calculer la dérivée de g, montrer que g est croissante sur $[0, +\infty[$.
   (b) Soit $x_0 \in [0, +\infty[$, montrer que, pour $(x, t) \in \mathscr{F}$
   $$\left|f(x, t) - f(x_0, t) - (x - x_0)\partial_1 f(x_0, t)\right| \leqslant \frac{t^2}{8}\mathrm{e}^{-t^2}|x - x_0|^2.$$
   (c) En déduire que, pour $x_0 \in [0, +\infty[$,
   $$\left|g(x) - g(x_0) - (x - x_0)\int_0^{+\infty} \partial_1 f(x_0, t)\,\mathrm{d}t\right| \leqslant \frac{|x - x_0|^2}{8}\int_0^{+\infty} t^2\mathrm{e}^{-t^2}\,\mathrm{d}t.$$
   (d) Montrer que g est dérivable sur $[0, +\infty[$ et que g' est définie par
   $$g'(x) = \int_0^{+\infty} \partial_1 f(x, t)\,\mathrm{d}t.$$
   Retrouver le sens de variation de g.

Exercice 342. ✦✦✦

Soit $f : \mathbb{R}^n \to \mathbb{R}$ de classe $\mathscr{C}^1$. On note S la sphère unité de $\mathbb{R}^n$ et B la boule unité ouverte. On suppose que f est constante sur S. Démontrer l'existence de $a \in \mathrm{B}$ tel que $\nabla f(a) = 0_{\mathbb{R}^n}$.

Exercice 343. ✦✦✦ **Existence d'une valeur propre pour une matrice symétrique et quotient de Rayleight**

Considérons sur $\mathscr{M}_{n,1}(\mathbb{R})$ le produit scalaire usuel $\langle \mathrm{X}, \mathrm{Y}\rangle = {}^t\mathrm{X}\mathrm{Y}$. Soit S une matrice symétrique réelle de $\mathscr{M}_n(\mathbb{R})$, on pose

$$\mathscr{B} = \left\{x \in \mathbb{R}^n \mid \|x\| = 1\right\} \quad \text{et} \quad \mathrm{R_S} : \left\{\begin{array}{ccc} \mathbb{R}^n \setminus \{0_n\} & \to & \mathbb{R} \\ x & \mapsto & \mathrm{R_S}(\mathrm{X}) = \dfrac{\langle \mathrm{X}, \mathrm{SX}\rangle}{\langle \mathrm{X}, \mathrm{X}\rangle} \end{array}\right.$$

où X est le vecteur-colonne des coordonnées de x dans la base canonique de $\mathbb{R}^n$.

1. (a) Justifier que la restriction de $\mathrm{R_S}$ à $\mathscr{B}$ admet un minimum et un maximum.
   (b) En remarquant que, pour tous $\alpha \in \mathbb{R}^*$, $x \in \mathbb{R}^n \setminus \{0_n\}$, $\mathrm{R_S}(x) = \mathrm{R_S}(\alpha x)$, montrer que l'application $\mathrm{R_S}$ admet un minimum et un maximum global sur $\mathbb{R}^n \setminus \{0_n\}$.
2. *Application.*
   (a) Montrer que $\mathrm{R_S}$ est de classe $\mathscr{C}^1$.
   (b) Soit x_0 un point où la maximum est atteint. En écrivant que x_0 est un point critique de $\mathrm{R_S}$, montrer que x_0 est vecteur propre de S.

Exercice 344. ✦✦✦ *D'après ESCP 01*

Dans cet exercice, n désigne un entier naturel supérieur ou égal à 2 . On note I_n la matrice identité de $\mathscr{M}_n(\mathbb{R})$ et J_n la matrice de $\mathscr{M}_n(\mathbb{R})$ dont tous les éléments valent 1.

1. Déterminer le rang de J_n et en déduire ses valeurs propres. La matrice J_n est-elle diagonalisable ?

   Dans toute la suite, on considère la fonction f_n définie sur $\mathbb{R}^n$ par :

$$f_n(x_1,\dots,x_n) = \left(\sum_{i=1}^{n} x_i\right)\exp\left(-\sum_{i=1}^{n} {x_i}^2\right).$$

2. Montrer que f_n est de classe $\mathscr{C}^2$ sur $\mathbb{R}^n$.
3. Montrer que f_n possède deux points critiques a et $b = -a$, avec a dont les coordonnées sont positives.
4. Justifier que la hessienne de f_n en a est $\mathrm{H}_n(a) = \dfrac{-2}{\sqrt{2n\mathrm{e}}}(n\mathrm{I}_n + \mathrm{J}_n)$.
5. Établir que f_n possède un extremum local en a. Quelle est sa nature ? Donner sa valeur. On admet que f_n possède un extremum local de nature et de valeur opposées en b.
6. (a) Étudier la fonction h définie sur $\mathbb{R}^+$, par : $\forall t \geqslant 0,\ h(t) = t\mathrm{e}^{-t^2}$.

   (b) Montrer que, pour tout $(x_1,\dots,x_n)$ de $\mathbb{R}^n$,

$$\left(\sum_{k=1}^{n} x_k\right)^2 \leqslant n\sum_{k=1}^{n} {x_k}^2.$$

   (c) Déduire des deux questions précédentes que f_n admet en a et b des extremums globaux.

CHAPITRE 17

Estimations

> *If there is a 50-50 chance that something can go wrong, then nine times out of 10 it will.*
>
> PAUL HARVEY
> Animateur radio américain (1918-2009)

1 Estimation ponctuelle

1.1 Principe : modéliser, estimer, tester/prédire

On considère un phénomène aléatoire et on s'intéresse à une variable aléatoire réelle X qui pourrait le décrire. On suppose que la loi de probabilité de X n'est pas complètement spécifiée et appartient à une famille de lois dépendant d'un paramètre θ décrivant un sous-ensemble Θ. Le paramètre θ est une quantité inconnue, fixée dans toute l'étude, que l'on cherche à déterminer ou pour laquelle on cherche une information partielle.

Le problème de l'*estimation* consiste alors à préciser la vraie valeur du paramètre θ (ou plus généralement d'une fonction $g(\theta)$), à partir d'un échantillon de données $x_1, \ldots, x_n$ obtenues en observant n fois le phénomène. Cette fonction du paramètre représentera en général une valeur caractéristique de la loi inconnue comme son espérance, sa variance, son étendue, etc.

Exemples.
- *Exemple 1. Nombre de buts en Ligue 1.*

→ Voici le nombre de buts par journée lors de la saison 2021/2022 :

$$26, 34, 29, 31, 28, 25, 32, 30, 25, 26, 29, 29, 29, 30, 21, 29, 29, 27,$$

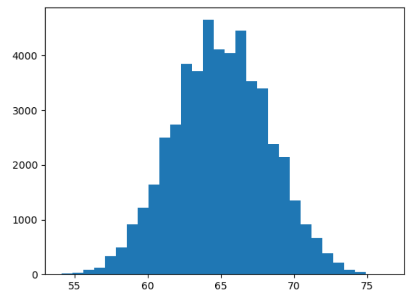
$$27, 15, 26, 35, 30, 23, 23, 22, 21, 19, 23, 32, 38, 30, 26, 35, 30, 36, 30, 37.$$

Ces nombres constituent *l'échantillon de données.*
→ On pose un *modèle* en supposant que le nombre de buts lors d'une journée est une variable aléatoire X qui suit une loi de Poisson. Notons de paramètre λ.
→ Pour donner une valeur à λ, (on dira *estimer*), on peut utiliser le fait que X admet une espérance $\mathbf{E}(X) = \lambda$. Ainsi, on peut préciser λ en prenant la valeur moyenne du nombre de buts (à savoir ici $\lambda = 2.81$).
→ Une fois le modèle posé, on peut le *tester* s'il donne des résultats cohérents et ainsi l'utiliser pour faire de la *prédiction.*

- *Exemple 2. Durée de vie d'un composant électrique.*

- *Exemple 3.*

Vincent, éleveur à Armentières sur Ourcq (02) produit des œufs de poules élevées en plein air. Il souhaite anticiper sa production de gros œufs pour l'année 2023 en analysant sa production actuelle. Un très gros œuf a un poids supérieur ou égal à 73 grammes. Voici la répartition de la production des 51 243 œufs de l'année.



Console

```
>>> len(L)
51243

>>> np.mean(L) # poids moyen d'un oeuf
65.115928

>>> np.std(L)  # écart-type
3.2336541402592824
```

Vincent suppose que le poids moyen d'un œuf correspond à une variable aléatoire X qui une loi normale de paramètre (μ, σ^2) avec $\mu = 65.1$ et $\sigma = 3.23$.

La suite de ce chapitre propose de donner un cadre théorique rigoureux à ces trois exemples et de juger aussi de la pertinence de cette approche. Nous discuterons essentiellement de l'étape d'estimation. Notamment nous verrons deux types d'estimations : l'estimation ponctuelle et l'estimation par intervalle de confiance.

1.2 Définitions et exemples

Dans la suite, on se fixe :

- un espace probabilisable $(\Omega, \mathscr{A})$.
- un espace des paramètres Θ qui est une partie de $\mathbb{R}^n$.
  On suppose de plus que pour chaque paramètre $\theta \in \Theta$, il existe une probabilité $\mathbf{P}_\theta$ définie sur $(\Omega, \mathscr{A})$.
- une application X qui est bien une variable aléatoire sur les espaces probabilisés $(\Omega, \mathscr{A}, \mathbf{P}_\theta)$ (pour tout $\theta \in \Theta$).

Notation. Sous réserve d'existence :

- $\mathbf{E}_\theta(X)$ désigne l'espérance de X pour la probabilité $\mathbf{P}_\theta$.
- $\mathbf{V}_\theta(X)$ est la variance de X pour la probabilité $\mathbf{P}_\theta$.

Exemples.

- *Exemple 1.*

Dans ce cas, les variables aléatoires X_i suivent une loi de Poisson de paramètre θ. On a $\Theta = \mathbb{R}_*^+$, et

$$\forall\, \theta \in \mathbb{R}_*^+, \quad \forall\, n \in \mathbb{N}, \quad \mathbf{P}_\theta(X_i = n) = e^{-\theta}\frac{\theta^n}{n!}.$$

On a de plus, $\mathbf{E}_\theta(X_i) = \theta$ et $\mathbf{V}_\theta(X_i) = \theta$.

- *Exemple 2.* Dans le cas où les X_i suivent une loi exponentielle de paramètre $\lambda = \theta$, alors on a

$$\forall\, \theta \in \mathbb{R}_+^*, \quad \forall\, t \in \mathbb{R}_+^*, \quad \mathbf{P}_\theta\,(X \leqslant t) = 1 - e^{-\theta\, t}.$$

- *Exemple 3.* Dans ce cas, les X_i suivent une loi normale $\mathcal{N}\left(\mu, \sigma^2\right)$ dont on ne connait ni l'espérance ni la variance, alors $\theta = (\mu, \sigma)$, et $\Theta = \mathbb{R} \times \mathbb{R}_+^*$.

Définition 249 (échantillon)

Soient $X : \Omega \to \mathbb{R}$ une variable aléatoire sur les espaces $(\Omega, \mathscr{A}, \mathbf{P}_\theta)$ et $n \in \mathbb{N}^*$.
On appelle **n-échantillon** de la loi de X toute famille $(X_1, \ldots, X_n)$ telle que :

- Les applications $X_1, \ldots, X_n$ sont des variables aléatoires pour la probabilité $\mathbf{P}_\theta$, et ce pour tout $\theta \in \Theta$.
- Les variables $X_1, \ldots, X_n$ sont $\mathbf{P}_\theta$-indépendantes et de même loi que X pour tout $\theta \in \Theta$.

Vocabulaire.

- On dit aussi que la loi de X est la loi parente (ou encore loi mère) de l'échantillon.
- On note souvent $(X_1, \ldots, X_n)$ *i.i.d* pour signaler que les variables sont indépendantes, et identiquement distribuées (c'est-à-dire de même loi).

En pratique, un échantillon de données $x_1, \dots, x_n$ est la réalisation de n variables aléatoires $X_1, \dots, X_n$. L'objectif de l'estimation ponctuelle est de déterminer le paramètre θ qui « explique » au mieux les valeurs de l'échantillon.

Définition 250 (estimateur)

- On appelle **estimateur** de θ toute variable aléatoire de la forme $\varphi(X_1, \dots, X_n)$, où $(X_1, \dots, X_n)$ est un n-échantillon et φ, une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$.

- Plus généralement, pour $g : \Theta \to \mathbb{R}$, une fonction, un estimateur de $g(\theta)$ est une variable aléatoire de la forme $\varphi(X_1, \dots, X_n)$ où $(X_1, \dots, X_n)$ est un n-échantillon.

Remarques.
- Un estimateur ne peut pas dépendre de θ puisque c'est la valeur que l'on souhaite déterminer. Par contre, il peut très bien dépendre de n.
- Estimer ponctuellement $g(\theta)$ par $\varphi(x_1, \dots, x_n)$ où $\varphi(X_1, X_2, \dots, X_n)$ est un estimateur de $g(\theta)$ et $(x_1, \dots, x_n)$ est une réalisation de l'échantillon $(X_1, \dots, X_n)$, c'est décider d'accorder à $g(\theta)$ la valeur $\varphi(x_1, \dots, x_n)$.

La réalisation $\varphi(x_1, \dots, x_n)$ de l'estimateur T_n est l'estimation de $g(\theta)$. Cette estimation ne dépend que de l'échantillon $(x_1, x_2, \dots, x_n)$ observé.

Exemples. Avec les notations de la définition, soit $(X_1, \dots, X_n)$ un n-échantillon de la loi de X. Il est souvent utile de considérer **l'estimateur de la moyenne empirique** donné par :

$$\overline{X_n} = \frac{X_1 + X_2 + \cdots + X_n}{n}.$$

Mais on peut inventer toute une gamme d'estimateurs :

$$T_n = \max_{1 \leqslant k \leqslant n} X_k, \quad U_n = \min_{1 \leqslant k \leqslant n} X_k, \quad V_n = \ln(1 + |U_n|), \quad W_n = \max\left\{k \in [\![1, n]\!] \mid X_k = T_n\right\}, \quad Y_n = 1, \quad \text{etc.}$$

1.3 Biais, convergence et comparaison des estimateurs

Les définitions qui suivent permettent de quantifier la qualité d'un estimateur et de les comparer entre-eux.

Définition 251 (biais)

Soit T_n un estimateur de $g(\theta)$ tel que tout $\theta \in \Theta$, T_n admet une espérance pour la probabilité $\mathbf{P}_\theta$.

- On définit le **biais** de T_n en $g(\theta)$ par

$$b_\theta(T_n) = \mathbf{E}_\theta(T_n) - g(\theta).$$

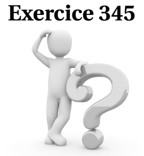
- Si pour tout $\theta \in \Theta$, $b_\theta(T_n) = 0$, on dit l'estimateur est **sans biais**. Sinon, on dit que l'estimateur est biaisé.

Exemple. Avec les notations de la définition et si X admet une espérance μ, l'estimateur $\overline{X_n}$ est un estimateur sans biais de μ.

Remarque. On peut donner une définition moins contraignante : un estimateur est **asymptotiquement sans biais** si

$$b_\theta(T_n) \xrightarrow[n \to \infty]{} 0.$$

Exercice 345



✦✦ **Estimateur de la variance**

1. On suppose, dans cette question, que la variable X admet un espérance μ connue et une variance σ^2, inconnue.
   Montrer que la variable $T_n = \frac{1}{n} \sum_{k=1}^{n} (X_k - \mu)^2$ est un estimateur sans biais de σ^2.
2. On suppose, dans cette question, que μ est inconnu. On note $V_n = \frac{1}{n} \sum_{k=1}^{n} \left(X_k - \overline{X_n}\right)^2$.
   (a) Montrer que V_n est un estimateur asymptotiquement sans biais de σ^2 et calculer le biais de cet estimateur.
   (b) Construire, à partir de V_n, un estimateur sans biais de σ^2.

Définition 252 (estimateur convergent)

Soit $(\mathrm{T}_n)_{n\in\mathbb{N}^*}$, une suite d'estimateurs de $g(\theta)$.
On dit que la suite $(\mathrm{T}_n)_{n\in\mathbb{N}}$ est **convergente** si pour tout $\theta\in\Theta$, la suite (T_n) converge en probabilité vers la variable aléatoire presque sûrement constante $g(\theta)$. Autrement dit

$$\forall\,\theta\in\Theta,\quad \forall\varepsilon>0,\quad \mathbf{P}_\theta\left(\left|\mathrm{T}_n-g(\theta)\right|\geqslant\varepsilon\right)\xrightarrow[n\to\infty]{}0.$$

Vocabulaire. Par abus de langage, on dit aussi que T_n est un estimateur convergent de $g(\theta)$.

Exercice 346

✦ **Exemples**

1. Considérons X une variable aléatoire dont la loi est uniforme sur $[0,\theta]$, où le paramètre θ est inconnu. Vérifier que $\mathrm{T}_n=2\bar{\mathrm{X}}_n$ est un estimateur convergent de θ.
2. Soit X une variable à densité donnée pour $\theta\in\mathbb{R}_*^+$ par

$$\forall\,x\in\mathbb{R},\quad f_\theta(x)=\frac{2}{\theta}\left(1-\frac{x}{\theta}\right)\mathbf{1}_{[0,\theta]}(x).$$

   (a) Pour un échantillon $(\mathrm{X}_1,\ldots,\mathrm{X}_n)$, montrer que $3\overline{\mathrm{X}_n}$ est un estimateur sans biais de θ.
   (b) Vérifier que cet estimateur est convergent.

Remarque. La notion de convergence des estimateurs ne donne aucune assurance pratique que la valeur prise par un estimateur à partir de l'échantillon de données sera assez proche de la vraie valeur du paramètre. On quantifie la qualité des estimateurs par la notion de **risque quadratique**. Cette notion est maintenant hors-programme mais on pourra consulter l'exercice 355, page 196, pour des précisions.

Proposition 253 (composition et estimateurs)

Soit $(\mathrm{T}_n)_{n\in\mathbb{N}}$, une suite d'estimateurs.

Si
- La suite $(\mathrm{T}_n)_{n\in\mathbb{N}}$ est une suite d'estimateurs convergente de $g(\theta)$.
- La fonction f est continue sur $\mathbb{R}$.

Alors $\big(f(\mathrm{T}_n)\big)_{n\in\mathbb{N}}$ est une suite d'estimateurs convergente de $f\big(g(\theta)\big)$.

Exercice 347

✦ Soit $(\mathrm{X}_1,\mathrm{X}_2,\ldots,\mathrm{X}_n)$ un n-échantillon issu d'une loi normale $\mathcal{N}(0,\sigma^2)$.

1. Pour tout i, calculer l'espérance de la variable aléatoire $|\mathrm{X}_i|$.
2. En déduire un estimateur sans biais et convergent de σ.

Exercice 348

✦✦ **Biais et composition**

Soit $(\mathrm{X}_1,\mathrm{X}_2,\ldots,\mathrm{X}_n)$ un n-échantillon d'une variable aléatoire X suivant une loi de Poisson de paramètre λ inconnu. On sait que la moyenne empirique $\overline{\mathrm{X}}_n=\frac{1}{n}\sum_{k=1}^{n}\mathrm{X}_k$ est un estimateur sans biais et convergent de λ. On cherche à estimer $\mathrm{e}^{-\lambda}$.
Est-ce que l'estimateur $\mathrm{T}_n=\mathrm{e}^{-\overline{\mathrm{X}}_n}$ est un estimateur sans biais de $\mathrm{e}^{-\lambda}$? asymptotiquement sans biais ? convergent ?

Proposition 254 (condition suffisante de convergence)

Soit $(\mathrm{T}_n)_{n\in\mathbb{N}}$, une suite d'estimateurs de $g(\theta)$.

Si $\mathbf{E}(\mathrm{T}_n)\xrightarrow[n\to\infty]{}g(\theta)$ et $\mathbf{V}(\mathrm{T}_n)\xrightarrow[n\to\infty]{}0.$

Alors $(\mathrm{T}_n)_{n\in\mathbb{N}}$ est une suite d'estimateurs convergente de $g(\theta)$.

Remarque. Un estimateur sans biais Y_n est meilleur qu'un autre estimateur sans biais Z_n si $\mathbf{V}(Y_n) \leqslant \mathbf{V}(Z_n)$ pour tout entier n (le fait d'être meilleur se matérialise dans l'inégalité de Bienaymé-Tchebychev).

Exemple. L'estimateur de la moyenne empirique est convergent si X admet un variance.

Exercice 349

✧ Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires mutuellement indépendantes et suivant toutes la loi $\mathscr{B}(p)$, où p est un paramètre inconnu que l'on chercher à estimer. On pose

$$\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{et} \quad T_n = \frac{2}{n(n+1)}\sum_{i=1}^{n} iX_i.$$

1. Montrer que $\overline{X_n}$ et T_n sont deux estimateurs sans biais de p.
2. Calculer et comparer les variances de $\overline{X_n}$ et de T_n.
3. Montrer que $\overline{X_n}$ et T_n sont deux estimateurs convergents de p.

Exemple. Questions sensibles lors d'un sondage d'opinion.

Comparaison des estimateurs sans biais

Simulation Python.
Reprenons le cas d'une variable X suivant une loi uniforme sur $[0;\theta]$ et des deux estimateurs sans biais :

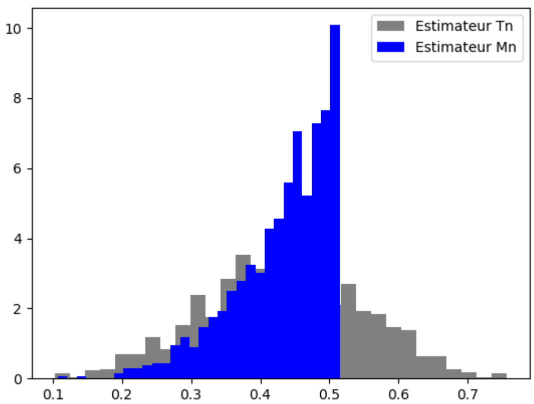
$$\overline{X_n} = \frac{2}{n}\sum_{i=1}^{n} X_i \quad \text{et} \quad \tilde{M_n} = \frac{n+1}{n}\max(X_1,\ldots,X_n).$$

Simulons un grand nombre d'échantillons de données et affichons les réalisations de ces deux estimateurs à l'aide des histogrammes.

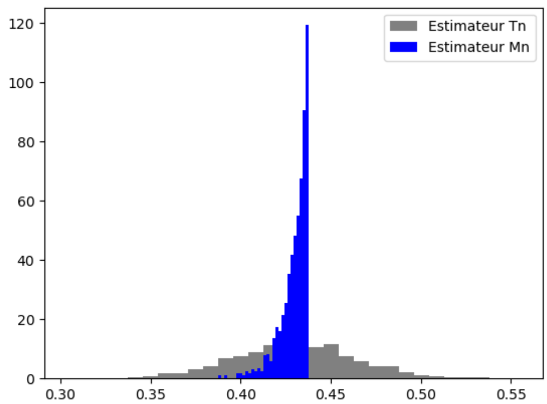
Editeur

```
# Importation des bibliothèques
# et définition des deux estimateurs

import numpy as np
import random as rd
import matplotlib.pyplot as plt

theta=0.4292081919392057
# choix du paramètre "inconnu"

def estimateurT(n):
    return 2*np.sum(theta*np.random.rand(n))/n

def estimateurM(n):
    return ((n+1)/n)*max(theta*np.random.rand(n))

# affichage des histogrammes

def histogrammes(n):
    LM=[]
    LT=[]
    for i in range(1000):
        LM.append(estimateurM(n))
        LT.append(estimateurT(n))
    plt.clf()
    plt.hist(LM,30)
    plt.hist(LT,30)
    plt.show()
```

```
>>> histogrammes(5)
```




```
>>> histogrammes(50)
```



Le meilleur estimateur sans biais est celui qui a la variance la plus faible.

2 Estimation par intervalle de confiance

S'il existe des critères pour juger des qualités d'un estimateur ponctuel T_n de $g(\theta)$, aucune certitude ne peut jamais être apportée quant au fait que l'estimation donnée par l'échantillon de donnée soit une « bonne » valeur du paramètre $g(\theta)$. L'estimation par intervalle de confiance permet de trouver un intervalle aléatoire qui contienne $g(\theta)$ avec une probabilité minimale donnée. Dans tout ce paragraphe, $(U_n)_{n\in\mathbb{N}^*}$ et $(V_n)_{n\in\mathbb{N}^*}$ désigneront deux suites d'estimateurs de $g(\theta)$ telles que pour tous $\theta\in\Theta$ et $n\in\mathbb{N}^*$,

$$\mathbf{P}_\theta\left([U_n \leqslant V_n]\right) = 1.$$

2.1 Intervalle de confiance

Définition

Définition 255 (intervalle de confiance)

Soient $\alpha \in]0,1[$, U_n et V_n deux estimateurs de $g(\theta)$ tels que pour tout $\theta\in\Theta$

$$\mathbf{P}_\theta\left(U_n \leqslant g(\theta) \leqslant V_n\right) \geqslant 1-\alpha.$$

On dit que l'intervalle $[U_n, V_n]$ est un **intervalle de confiance** de $g(\theta)$ avec un risque d'au plus α ou au niveau de confiance au moins égal à $1-\alpha$.

Remarque. En pratique, on part d'un échantillon de données $x_1, x_2, \dots, x_n$. On en déduit les valeurs

$$u_n = U_n(x_1, x_2, \dots, x_n) \quad \text{et} \quad v_n = V_n(x_1, x_2, \dots, x_n).$$

Ainsi on construit un intervalle aléatoire $[u_n, v_n]$ dans lequel $g(\theta)$ à une probabilité supérieure à $1-\alpha$ de s'y trouver.

Remarque. Dans les conditions usuelles, on considère des niveaux de confiance de 95% ou 99% (soit $\alpha = 0,05$ ou $\alpha = 0,01$).

Estimation par intervalle de confiance en utilisant l'inégalité de Bienaymé-Tchebychev

Méthode

Comment obtenir un intervalle de confiance à partir de l'inégalité de Bienaymé-Tchebychev ?

Soit T_n, une variable aléatoire admettant une variance, l'inégalité se réécrit sous la forme

$$\mathbf{P}_\theta\left(|T_n - \mathbf{E}_\theta\left(T_n\right)| \leqslant \varepsilon\right) \geqslant 1 - \frac{\mathbf{V}_\theta(T_n)}{\varepsilon^2}.$$

Si T_n est un estimateur sans biais de $g(\theta)$ (c'est-à-dire $\mathbf{E}_\theta\left(T_n\right) = g(\theta)$) et si on peut trouver un entier n tel que $\frac{\mathbf{V}_\theta(T_n)}{\varepsilon^2} \leqslant \alpha$, on peut récrire l'inégalité

$$\mathbf{P}_\theta\left(T_n - \varepsilon \leqslant g(\theta) \leqslant T_n + \varepsilon\right) \geqslant 1-\alpha.$$

L'intervalle de confiance est alors $[T_n - \varepsilon ; T_n + \varepsilon]$, il est de longueur 2ε.

Exemple. *Estimation du paramètre p d'une loi de Bernoulli.*
On réalise un sondage sur n personnes avec une unique question. On suppose que les réponses des personnes sont indépendantes et on veut déterminer un intervalle de confiance d'au moins 0.95 de la probabilité p de répondre positivement à l'unique question posée. D'après la loi faible des grands nombres, cet intervalle sera d'autant plus petit (au sens de l'inclusion) que le nombre de personnes interrogés sera grand. Désignons, pour un entier naturel non nul quelconque n, par X_n la variable aléatoire égale à 1 si la $n^{\text{ième}}$ personne répond positivement et 0 sinon. La

suite $(X_n)_{n\in\mathbb{N}^*}$ est une suite de variables de loi de Bernoulli indépendantes de paramètre p.

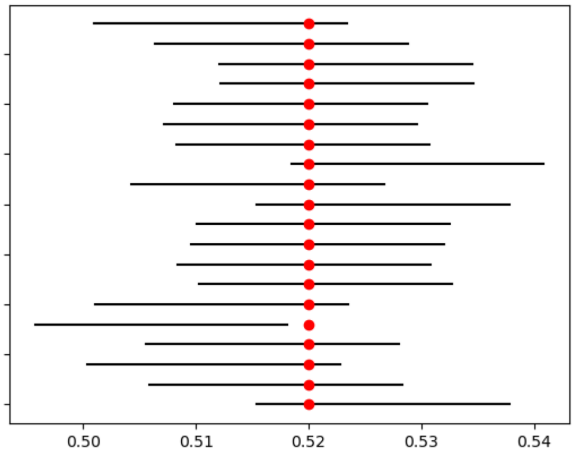
Exercice 350

✦ Soit $\left(\overline{X}_n\right)_{n\in\mathbb{N}^*}$ la suite des moyennes empiriques. Justifier qu'on a un intervalle de confiance de p à un niveau de confiance $1-\alpha$ avec

$$\mathbf{P}\left(p \in \left[\overline{X}_n - \frac{1}{2\sqrt{n\alpha}}, \overline{X}_n + \frac{1}{2\sqrt{n\alpha}}\right]\right) \geqslant 1-\alpha$$

Editeur

```
import numpy as np
import matplotlib.pyplot as plt

p=0.52 # le paramètre "inconnu" à estimer

n=10000 # Nombre de sondés
alpha=0.20 # niveau de risque

NbreTest=20
for i in range(NbreTest):
    Ech=np.random.rand(n)<p
    u=np.mean(Ech)-1/(2*(n*alpha)**(1/2))
    v=np.mean(Ech)+1/(2*(n*alpha)**(1/2))
    plt.plot([u,v],[i,i],'k')
    plt.plot([p],[i],'ro')
plt.show()
```



Estimation par intervalle de confiance de la moyenne d'une loi normale dont l'écart type est connu

Soit $(X_1, X_2, \dots, X_n)$ un n-échantillon issu d'une loi normale $\mathcal{N}(\mu, \sigma^2)$. On suppose σ connu mais l'espérance μ est inconnu et on cherche à l'estimer. On considère $\overline{X_n}$ la moyenne empirique de l'échantillon. Nous avons vu que c'est un estimateur sans biais et convergent de μ.

Exercice 351

✦ Posons t_α positif tel que $t_\alpha = \Phi^{-1}\left(1-\frac{\alpha}{2}\right)$. Avec ce choix, montrer que

$$\left[\overline{X}_n - t_\alpha\frac{\sigma}{\sqrt{n}}, \overline{X}_n + t_\alpha\frac{\sigma}{\sqrt{n}}\right]$$

est un intervalle de confiance de μ avec un niveau de confiance égal à $1-\alpha$.

Applications numériques

- Pour un risque $\alpha = 5\%$, $1-\frac{\alpha}{2} = 0{,}975$ et $t_{0.05} = \Phi^{-1}(0{,}975) = 1{,}96$.
- Pour un risque de $\alpha = 1\%$, $1-\frac{\alpha}{2} = 0{,}995$ et $t_{0.01} = \Phi^{-1}(0{,}995) = 2{,}58$.

Exercice 352

✦✦ **Comparaison des méthodes précédentes**
Dans une population donnée, une étude statistique faite sur un groupe de 100 personnes donne lieu à la série statistique suivante.

Poids	48	49	50	51	52	53	54	55
Effectif	3	5	2	6	6	10	12	10
Poids	56	57	58	59	60	61	62	63
Effectif	9	8	8	6	5	4	3	3

On suppose que le poids d'un individu du groupe est une variable aléatoire X qui suit une loi normale d'écart-type $\sigma = 3{,}5$. Dans chaque groupe de 100 personnes étudié, on désigne par X_i la variable aléatoire égale au poids du i ème individu, pour tout $i \in [\![1, 100]\!]$.

1. En utilisant les deux méthodes précédentes, déterminer une valeur approchée d'un intervalle de confiance à 95%, de la moyenne des poids des individus.
2. Comparer les deux méthodes.

2.2 Intervalle de confiance asymptotique

Définition 256 (intervalle de confiance asymptotique)

Soient $\alpha \in]0,1[$, U_n et V_n deux estimateurs de $g(\theta)$ tels que pour tout $\theta \in \Theta$

$$\mathbf{P}_\theta\left(U_n \leqslant g(\theta) \leqslant V_n\right) \geqslant 1-\alpha_n \quad \text{et} \quad \alpha_n \underset{n\to\infty}{\longrightarrow} \alpha.$$

On dit que l'intervalle $[U_n, V_n]$ est un **intervalle de confiance asymptotique** de $g(\theta)$ avec un risque d'au plus α ou au niveau de confiance au moins égal à $1-\alpha$.

Intervalle de confiance asymptotique du paramètre d'une loi de Bernoulli

Soit $(X_1, X_2, \ldots, X_n)$ un n-échantillon issu d'une loi de Bernoulli de paramètre p.

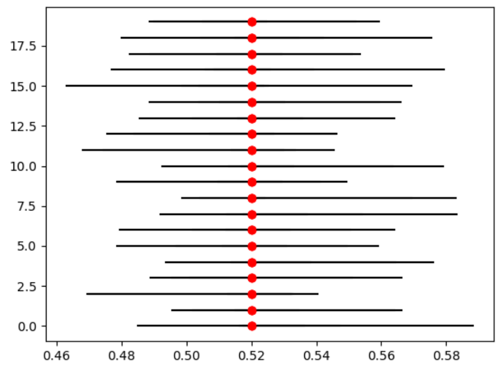
Exercice 353

✦✦ Justifier qu'on a un intervalle de confiance asymptotique de p au niveau de confiance $1-\alpha$ avec

$$\left[\overline{X_n} - \frac{t_\alpha}{2\sqrt{n}}, \overline{X_n} + \frac{t_\alpha}{2\sqrt{n}}\right].$$

Editeur

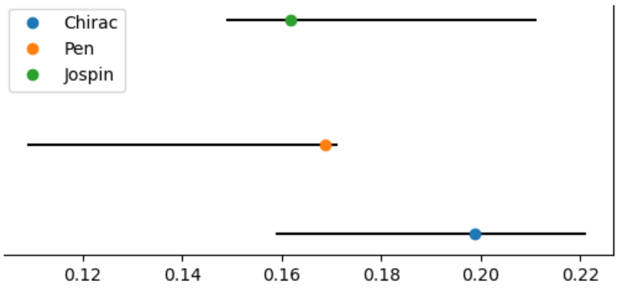
```
p=0.52 # le paramètre "inconnu" à estimer

n=10000 # Nombre de sondés
alpha=0.05
talpha=1.96

NbreTest=20

for i in range(NbreTest):
    Ech=np.random.rand(n)<p
    u=np.mean(Ech)-talpha/(2*(n)**(1/2))
    v=np.mean(Ech)+talpha/(2*(n)**(1/2))
    plt.plot([u,v],[i,i],'k')
    plt.plot([p],[i],'ro')
plt.show()
```



Remarque. Les intervalles obtenus sont plus précis que ceux obtenus par l'inégalité de Bienaymé-Tchebychev. En effet, les longueurs des intervalles obtenus par le théorème central limite est plus petite que celles obtenues par l'inégalité de Bienaymé-Tchebychev.

Exemple. *Premier tour de l'élection présidentielle de 2002 (21 avril)*
Voici quelques résultats des sondages d'opinion quelques jours avant l'élection.

	Jacques CHIRAC	Jean Marie LE PEN	Lionel JOSPIN
CSA - 11 avril	21%	12%	19%
IFOP - 12 avril	19%	11,5%	17%
SOFRES - 13 avril	20%	13%	18%
CSA - 18 avril	19,15%	14%	18%
IPSOS - 18 avril	20%	14%	18%
BVA - 19 avril	**19%**	**14%**	**18%**
Sondage confidentiel - 21 avril	18%	14,5%	17%
Résultat final	**19,88%**	**16,88%**	**16,18%**



Extrait du site du conseil constitutionnel : "Dans le cas précis de l'élection présidentielle, la publication d'un sondage d'intention de vote prévoyant un faible écart entre les candidats (51-49 par exemple) rend nécessaire une telle précaution si l'on considère que, dans le cas des sondages portant sur un échantillon de 1 000 personnes (qui est l'échantillon standard en France pour les élections nationales et notamment l'élection présidentielle), la marge d'erreur est estimée à 3%.
L'obligation pour les instituts de sondage de rendre publique leur marge d'erreur est prévue par la proposition de loi sénatoriale."

Exercices - TD

Exercice 354. ✦ **Estimation d'un paramètre d'une loi uniforme**
Soit $\theta \in \mathbb{R}_*^+$. Soit X une variable aléatoire suivant une loi uniforme sur $]0;\theta[$ et $(X_1,\dots,X_n)$ un n-échantillon de X.

1. Montrer que pour tout $n \in \mathbb{N}^*, T_n = \frac{2}{n}(X_1+\cdots+X_n)$ est un estimateur sans biais de θ.
2. Considérons $Y_{\min}$ et $Y_{\max}$ définies par :
$$Y_{\min} = \min_{i\in[\![1;n]\!]} X_i \quad \text{et} \quad Y_{\max} = \max_{i\in[\![1;n]\!]} X_i.$$
Pour chacune de ces deux variables, préciser :
   (a) La fonction de répartition,
   (b) La densité de probabilité,
   (c) L'espérance et la variance. *On pourra remarquer que* $Y_{\min}$ *et* $\theta - Y_{\max}$ *ont même loi.*
3. Posons $T'_n = \frac{n+1}{n} Y_{\max}$. Justifier que T'_n est un estimateur sans biais de θ.
4. Quel est le meilleur estimateur de θ entre T'_n et T_n ?
5. Posons $T''_n = Y_{\min} + Y_{\max}$. Montrer sans calculs superflus que $\mathbf{V}(T''_n) \leqslant 4\mathbf{V}(Y_{\max})$.
6. En déduire que T''_n est meilleur estimateur de θ que T_n. *Un estimateur sans biais est d'autant meilleur que sa variance est faible.*
7. Est-ce que les estimateurs T_n, T'_n et T''_n sont convergents ?

≫ Solution p. ??

Exercice 355. ✦ **Risque quadratique**
Soit T_n un estimateur de $g(\theta)$. On suppose que pour tout $\theta \in \Theta, T_n$ admet un moment d'ordre 2 pour la probabilité $\mathbf{P}_\theta$.
Pour tout $\theta \in \Theta$, on appelle risque quadratique de T_n en $g(\theta)$ et on note $r_\theta(T_n)$ le réel :
$$r_\theta(T_n) = \mathbf{E}_\theta\left(\left(T_n - g(\theta)\right)^2\right).$$

1. Justifier que $r_\theta(T_n) = \left[b_\theta(T_n)\right]^2 + \mathbf{V}_\theta(T_n)$.
2. (a) Justifier que si le risque quadratique de T_n tend vers 0 quand n tend vers l'infini, alors T_n est un estimateur convergent de $g(\theta)$.
   (b) En déduire que, si T_n est un estimateur asymptotiquement sans biais de $g(\theta)$ et si la variance de T_n tend vers 0 lorsque n tend vers l'infini, alors T_n est un estimateur convergent de $g(\theta)$.
3. Pour comparer deux estimateurs T_n et U_n de $g(\theta)$, on peut calculer leur risque quadratique. Si on a :
$$\forall\, \theta \in \Theta, \quad r_\theta(T_n) \leqslant r_\theta(U_n),$$
alors on dira que T_n est un meilleur estimateur de $g(\theta)$ que U_n.
   (a) *Exemple 1.*
   Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires indépendantes suivant la loi exponentielle $\mathscr{E}(\lambda)$. On pose
$$\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{et} \quad T_n = \frac{1}{n+1}\sum_{i=1}^{n} X_i.$$
   Comparer ces deux estimateurs de $1/\lambda$.
   (b) *Exemple 2.*
   Soient T_0 et T_1 deux estimateurs de θ, sans biais et indépendants. Pour tout a réel, on pose $T_a = aT_1 + (1-a)T_0$.
      i. T_a est-il un estimateur sans biais de θ ?
      ii. Parmi tous les estimateurs T_a, lequel est le meilleur ?

Exercice 356. ✦✦ **Estimateur sans biais de l'écart-type σ d'une loi normale centrée** *extrait HEC 2005*
Soit X une variable aléatoire qui suit une loi normale centrée et d'écart-type σ, le paramètre réel inconnu σ, est strictement positif.

1. Montrer que la variable aléatoire $T = \frac{X^2}{2\sigma^2}$ suit une loi γ de paramètre $1/2$. En déduire la valeur de $\Gamma(1/2)$.
2. Pour n entier naturel non nul, on considère un n-échantillon $(X_1, X_2, \dots, X_n)$ constitué de variables indépendantes et de même loi que X.
   (a) On désigne par S_n la variable aléatoire $S_n = \sum_{i=1}^{n} \frac{{X_i}^2}{2\sigma^2}$. Quelle est la loi de probabilité de S_n ?
   (b) En déduire que la variable aléatoire Y_n définie par $Y_n = \frac{1}{n}\sum_{i=1}^{n} {X_i}^2$ est un estimateur sans biais de σ^2.

Exercice 357. ✦✦ *d'après EDHEC 2000*

Un sondage consiste à proposer l'affirmation « A » à certaines personnes d'une population donnée. Le sujet abordé étant délicat, le stratagème suivant est mis en place afin de mettre en confiance les personnes sondées pour qu'elles ne mentent pas...
L'enquêteur dispose d'un paquet de 20 cartes, numérotées de 1 à 20 , qu'il remet à la personne sondée. Celle-ci tire une carte au hasard et ne la montre pas à l'enquêteur. La règle est alors la suivante :

- si la carte porte le numéro 1, la personne sondée répond "vrai" si elle est d'accord avec l'affirmation « A » et "faux" sinon.
- si la carte porte un autre numéro, la personne sondée répond "vrai" si elle n'est pas d'accord avec l'affirmation « A » et "faux" sinon.

Le but de l'enquête est d'évaluer la proportion p de gens de cette population qui sont réellement d'accord avec l'affirmation « A ».

1. On interroge une personne selon ce procédé et on considère l'événement suivant, noté V : «la personne répond "vrai"». On note $\theta = \mathbf{P}(V)$. En utilisant la formule des probabilités totales, exprimer θ en fonction de p, puis en déduire p en fonction de θ.
2. Certaines considérations théoriques laissent penser que $p = \frac{17}{18}$.
   (a) Vérifier que $\theta = \frac{1}{10}$.
   (b) Calculer la probabilité pour qu'une personne ayant répondu "vrai" soit d'accord avec l'affirmation « A».

   On revient au cas général où l'on ne connaît ni p, ni θ.
3. On considère un échantillon aléatoire, de taille n, extrait de la population considérée et on note S_n le nombre de réponses "vrai" obtenues. On suppose n assez grand pour pouvoir considérer que cet échantillonnage est assimilable à un tirage avec remise.
   (a) Donner la loi de S_n ainsi que son espérance et sa variance.
   (b) Montrer que $\frac{S_n}{n}$ est un estimateur sans biais et convergent de θ.
4. Dans cette question, on suppose que l'on a réalisé un échantillon de 100 personnes et on constate que 23 personnes ont répondu "vrai".
   (a) Donner une estimation ponctuelle de θ et de p.
   (b) Donner un intervalle de confiance à 95% de θ puis de p.
   On rappelle que, si Φ désigne la fonction de répartition d'une variable X suivant la loi normale $\mathcal{N}(0,1)$, alors $\Phi(1,96) = 0,975$

Exercice 358. ✦✦ **Estimation des paramètres d'une loi de Pareto**

Soient $(b,\theta) \in \mathbb{R}^{+*} \times \mathbb{R}$. On suppose que X est une va à densité qui prend ses valeurs dans $[\theta;+\infty[$, dont une densité f est définie par :

$$\forall x \in \mathbb{R}, \quad f(x) = \begin{cases} 0 & \text{si } x < \theta \\ \frac{1}{b}\exp\left(-\frac{x-\theta}{b}\right) & \text{si } x \geqslant \theta. \end{cases}$$

Soient $n \in \mathbb{N} \setminus \{0;1\}$ et $(X_1,\ldots,X_n)$ un n-échantillon de même loi que X. On pose :

$$\overline{X_n} = \frac{X_1+\cdots+X_n}{n} \quad \text{et} \quad T_n = \min(X_1,\ldots,X_n).$$

1. Soit $k \in [\![1;n]\!]$. Reconnaitre la loi de $Y_k = (X_k - \theta)/b$.
2. On pose $\overline{Y_n} = \frac{Y_1+\cdots+Y_n}{n}$ et $U_n = \min(Y_1,\ldots,Y_n)$. Calculer $\mathbf{E}\left(\overline{Y_n}\right)$ et $\mathbf{E}(U_n)$.
3. Exprimer les variables aléatoires $\overline{X_n}$ et T_n en fonction des variables $\overline{Y_n}$ et U_n. En déduire $\mathbf{E}\left(\overline{X_n}\right)$ et $\mathbf{E}(T_n)$.
4. Déterminer un estimateur sans biais $\widehat{\theta}_n$ de θ et un estimateur sans biais $\widehat{b}_n$ de b sous la forme de combinaisons linéaires de $\overline{X_n}$ et T_n.

Exercice 359. ✦✦ **Le boulanger** *d'après oral ESCP 2022, sujet 23*

Un boulanger vend du pain chaque jour.

- La quantité de pain produite chaque jour est une quantité fixée Q choisie par le boulanger, Q étant exprimée en kilogramme.
- La demande de pain de la part des clients est une variable aléatoire X strictement positive, toujours exprimée en kilogramme.
- On suppose que la variable X admet une densité f strictement positive sur $\mathbb{R}_+^*$, nulle sur $\mathbb{R}_-$, continue sur $\mathbb{R}$ et on note F la fonction de répartition de X.
- Le coût de fabrication par kilogramme est c euros et le prix de vente est v euros par kilogramme.
- On note B la variable aléatoire égale au bénéfice quotidien.
- La variable indicatrice d'un événement A est notée $\mathbb{1}_A$.
- On suppose que $0 < c < v$.

Si la demande de pain X est inférieure à l'offre Q, le boulanger ne vend que la quantité X (le pain invendu un jour donné n'est pas remis en vente le lendemain!); si la demande est supérieure à l'offre, il ne vend que la quantité Q. Dans ces conditions, on cherche la quantité optimale à produire, c'est-à-dire la quantité Q_0 qui maximise l'espérance de B.

1. Établir la relation suivante :
$$B = v\left[Q + (X-Q)\mathbf{1}_{[X<Q]}\right] - cQ.$$
2. Montrer que la variable $X\mathbf{1}_{[X<Q]}$ admet une espérance et donner son expression sous forme d'intégrale.
3. En déduire l'égalité suivante :
$$\mathbf{E}(B) = (v-c)Q + v\left[\int_0^Q tf(t)\mathrm{d}t - QF(Q)\right].$$
4. Exprimer Q_0 à l'aide de F, de v et de c. Le boulanger cherche à prévoir sa demande journalière. La demande aléatoire X_n qui va s'exprimer le jour n n'est pas connu à l'avance mais le boulanger fait l'hypothèse que la demande ne variera pas beaucoup d'un jour à l'autre et que :
$$X_{n+1} = X_n + U_{n+1}$$
où :
   - X_0 est une constante strictement positive fixée.
   - Les U_k sont des variables aléatoires indépendantes, de même loi, d'espérance nulle et de variance σ^2 non nulle.
5. (a) Exprimer X_n en fonction des U_i et de X_0.
   (b) Montrer que la suite $\left(\frac{X_0}{n}\right)$ converge en probabilité vers 0.
   (c) Démontrer que si deux suites de variables aléatoires (A_n) et (B_n) convergent en probabilité respectivement vers des variables aléatoires a et b, alors la suite $(A_n + B_n)$ converge en probabilité vers $a+b$.
   (d) En déduire que $\left(\frac{X_n}{n}\right)$ converge en probabilité vers une variable que l'on précisera.
   (e) Montrer que la suite $\left(\frac{X_0}{n}\right)$ converge en loi vers 0.
   (f) Montrer que la suite $\left(\frac{X_n}{n}\right)$ converge en loi et préciser la loi limite.
   On pourra utiliser le Théorème de Slutsky : Si X_n *converge en loi vers* X, *et si* Y_n *converge en probabilité vers une constante* c, *alors* X_nY_n *converge en loi vers* cX.

Exercice 360. ✦✦ **Estimation et loi de Poisson** *d'après oral ESCP 2022, sujet 35*

Soit $X_1,\dots,X_n$ un n-échantillon de variables aléatoires indépendantes qui suivent la loi de Poisson de paramètre $\lambda>0$ inconnu. On cherche dans cet exercice à estimer $e^{-\lambda}$.
Pour tout $k\in[\![1,n]\!]$, on note Y_k la fonction indicatrice de l'événement $\left[X_k=0\right]$. Pour tout $n\in\mathbb{N}^*$, on pose :
$$\overline{Y_n} = \frac{1}{n}\sum_{k=1}^{n} Y_k \quad \text{et} \quad S_n = \sum_{k=1}^{n} X_k$$

1. (a) Déterminer la loi de Y_k.
   (b) Montrer que $\overline{Y_n}$ est un estimateur sans biais de $e^{-\lambda}$.
   (c) $\overline{Y_n}$ est-il un estimateur convergent de $e^{-\lambda}$?
2. Pour $j\in\mathbb{N}$, on pose $\varphi(j) = \mathbf{P}_{[S_n=j]}(X_1=0)$. Calculer $\varphi(j)$.
3. On pose à présent $T_n = \varphi(S_n)$.
   (a) Montrer que T_n est un estimateur sans biais de $e^{-\lambda}$.
   (b) T_n est-il un estimateur convergent de $e^{-\lambda}$?
4. Comparer les variances des deux estimateurs T_n et $\overline{Y}_n$.

Lien entre estimation et optimisation

Exercice 361. ✦

Soit $(X_1, X_2, \dots, X_n, Y_1, Y_2, \dots, Y_m)$ un échantillon de variables aléatoires indépendantes, de même loi de Bernoulli de paramètre inconnu p. On pose :
$$\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad \overline{Y_m} = \frac{1}{m}\sum_{i=1}^{m} Y_i \quad \text{et} \quad Z = a\overline{X_n} + b\overline{Y_m}, \quad \text{où} \quad (a,b)\in\mathbb{R}^2.$$
Pour quelle valeur de (a,b), Z est-il le meilleur estimateur sans biais de p ?

Exercice 362. ✦✦ **Maximum de vraisemblance** *D'après EDHEC 2014, voie E*

Dans cet exercice, θ désigne un réel strictement positif et n un entier naturel supérieur ou égal à 2. Pour tout k de $\mathbb{N}$, on pose :
$$u_k = \frac{1}{1+\theta}\left(\frac{\theta}{1+\theta}\right)^k.$$

1. Montrer que la suite $\left(u_k\right)_{k\in\mathbb{N}}$ définit bien une loi de probabilité. On considère maintenant une variable aléatoire X prenant ses valeurs dans $\mathbb{N}$ et dont la loi est donnée par :
$$\forall k\in\mathbb{N}, \quad \mathbf{P}(X=k) = u_k.$$

2. (a) On pose $Y = X + 1$. Reconnaître la loi de Y, puis en déduire l'espérance et la variance de X.
   (b) Écrire une fonction Python qui prend en argument θ et simule la loi d'une variable aléatoire X.
3. Dans cette question, on souhaite estimer le paramètre θ par la méthode du maximum de vraisemblance. Pour ce faire, on considère un échantillon $(X_1, X_2, \dots, X_n)$ composé de variables aléatoires indépendantes ayant toutes la même loi que X et on introduit la fonction L, de $\mathbb{R}_+^*$ dans $\mathbb{R}$, définie par :

$$\forall \theta \in \mathbb{R}_+^*, \quad L(\theta) = \prod_{k=1}^{n} \mathbf{P}\left(X_k = x_k\right)$$

   où $x_1, x_2, \dots, x_n$ désignent des entiers naturels éléments de $X(\Omega)$.
   L'objectif est de choisir la valeur de θ qui rend $L(\theta)$ maximale.
4. (a) Écrire $\ln(L(\theta))$ en fonction de θ et de $S_n = \sum_{k=1}^{n} x_k$.
   (b) On considère la fonction φ, définie par :

$$\forall \theta \in]0; +\infty[, \quad \varphi(\theta) = S_n \ln \theta - (S_n + n) \ln(1 + \theta).$$

   Montrer que la fonction φ admet un maximum, atteint en un seul réel que l'on notera $\widehat{\theta}_n$ et que l'on exprimera en fonction de S_n. Que représente $\widehat{\theta}_n$ pour la fonction L?

   On pose dorénavant : $T_n = \frac{1}{n} \sum_{i=1}^{n} X_i$. La variable T_n est appelée estimateur du maximum de vraisemblance pour θ.
   (c) Vérifier que T_n est un estimateur sans biais de θ.
   (d) On définit le risque quadratique par

$$r_\theta(T_n) = \mathbf{E}_\theta\left(\left(T_n - g(\theta)\right)^2\right).$$

   Calculer $r_\theta(T_n)$ de T_n et vérifier que $\lim_{n \to +\infty} r_\theta(T_n) = 0$.

Exercice 363. ✦✦✦ **Maximum de vraisemblance** *D'après HEC, 2020*

Lorsque l'on cherche à estimer un paramètre inconnu à partir d'un échantillon de données, on appelle statistique exhaustive toute fonction de ces données qui résume à elle seule l'information que ces données fournissent sur le paramètre. On donne ici une définition précise de cette notion d'exhaustives dans le cas des échantillons de variables aléatoires discrètes, illustrée de plusieurs exemples qui en montrent l'intérêt. On s'intéressera dans ce problème à l'estimation d'un paramètre réel inconnu θ appartenant à un intervalle Θ.

On dispose pour cela de plusieurs observations $x_1, \dots, x_n$ considérées comme les réalisations de variables aléatoire discrètes $X_1, \dots, X_n$ définie sur le même espace probabilisable $(\Omega, \mathscr{A})$, à valeurs dans une partie B de $\mathbb{N}$.

L'espace probabilisable $(\Omega, \mathscr{A})$ est muni d'une famille $\left(\mathbf{P}_\theta\right)_{\theta \in \Theta}$ de probabilités indexées par le paramètre θ.

On fait, pour toutes les valeurs du paramètre θ, les trois hypothèses suivantes.

- Les variables aléatoires $X_1, \dots, X_n$ sont mutuellement indépendantes, c'est-à-dire :

$$\forall (x_1, \dots, x_n) \in B^n, \quad \mathbf{P}_\theta\left(\bigcap_{i=1}^{n} [X_i = x_i]\right) = \prod_{i=1}^{n} \mathbf{P}_\theta\left([X_i = x_i]\right) \tag{1}$$

- Les variables aléatoires $X_1, \dots, X_n$ suivent toutes la même loi qu'une variable aléatoire de référence, dotée X, à valeurs dans B, c'est-à-dire :

$$\forall i \in [1, n], \quad \forall x \in B, \quad \mathbf{P}_\theta\left([X_i = x]\right) = \mathbf{P}_\theta([X = x]) \tag{2}$$

- Tous les éléments de B sont des valeurs effectivement possibles de X, c'est-à-dire :

$$\forall x \in B, \quad \mathbf{P}_\theta([X = x]) > 0 \tag{3}$$

On appelle statistique toute variable aléatoire S de la forme $\omega \mapsto s(X_1(\omega), \dots, X_n(\omega))$, où s désigne une application définie sur B^n et à valeurs réelles. On note alors $S = s(X_1, \dots, X_n)$.

Pour tout $\theta \in \Theta$, on note $\mathbf{E}_\theta(S)$ l'espérance de S lorsque $(\Omega, \mathscr{A})$ est muni de la probabilité $\mathbf{P}_\theta$ (si cette espérance existe). On note de même $\mathbf{V}_\theta(S)$ la variance de S (si elle existe).

Partie 1 : développements en série

1. Dans cette question, x désigne un nombre réel strictement compris entre 0 et 1.
   (a) Justifier la convergence de la série $\sum_{k \geqslant 1} x^k / k$.
   (b) Vérifier, pour tout $m \in \mathbb{N}^*$ et tout $t \in]0, 1[$, l'égalité : $\dfrac{1}{1-t} = \dfrac{t^m}{1-t} + \sum_{k=0}^{m-1} t^k$.
   (c) Démontrer que l'intégrale $\displaystyle\int_0^x \frac{t^m}{1-t}\,dt$ tend vers 0 quand l'entier m tend vers l'infini.
   (d) En déduire la somme de la série $\sum_{k \geqslant 1} x^k / k$.

2. Dans cette question, indépendante de la précédente, $(a_k)_{k\in\mathbb{N}}$ désigne une suite de nombres réels telle que la série $\sum_{k\geqslant 0} a_k c^k$ est absolument convergente pour un réel strictement positif c.

   (a) Justifier que la fonction $f : x \mapsto a_0 + \sum_{k=1}^{+\infty} a_k x^k$ est bien définie sur le segment $[-c, c]$.

   (b) Pour un entier naturel m, on pose : $M_m = \sum_{k=m+1}^{+\infty} |a_k|\, c^{k-m-1}$ Justifier, pour tout $x \in [-c, c]$, l'inégalité :

$$\left|\sum_{k=m+1}^{+\infty} a_k x^k\right| \leqslant M_m |x|^{m+1}.$$

   (c) Justifier, pour tout $m \in \mathbb{N}^*$, le développement limite au voisinage de 0 :

$$f(x) = a_0 + \sum_{k=1}^{m} a_k x^k + o\left(x^m\right)$$

   (d) Démontrer que si la fonction f est nulle sur l'intervalle $]0, c]$, alors $(a_k)_{k\in\mathbb{N}}$ est la suite nulle.

- Dans toute la suite du problème, pour tout $\theta \in \Theta$ et tout $(x_1, \ldots, x_n) \in B^n$, on note :

$$L(x_1, \ldots, x_n, \theta) = \prod_{i=1}^{n} \mathbf{P}_\theta\left([X_i = x_i]\right) \tag{4}$$

Cette quantité, qui s'écrit aussi $\prod_{i=1}^{n} \mathbf{P}_\theta\left([X = x_i]\right)$ d'après (2), est appelée *la vraisemblance* de la valeur θ du paramètre au vu des observations $x_1, \ldots, x_n$.

Partie II : estimateur du maximum de vraisemblance, un exemple
Dans cette partie, Θ est l'intervalle ouvert $]0, 1[$, B est égal à $\mathbb{N}^*$ et on a :

$$\forall x \in B, \quad \mathbf{P}_\theta([X = x]) = (1-\theta)^{x-1}\theta.$$

On note $\overline{X}$ la variable aléatoire $\dfrac{1}{n}\sum_{i=1}^{n} X_i$.

3. Soit $\theta \in \Theta$

   (a) Reconnaître la loi de X lorsque $(\Omega, \mathscr{A})$ est muni de la probabilité $\mathbf{P}_\theta$.

   (b) En déduire que $\overline{X}$ est un estimateur sans biais du paramètre $1/\theta$

   (c) Quel est le risque quadratique de cet estimateur ? Rappelons que le risque quadratique est défini par

$$r_\theta\left(\overline{X}\right) = \mathbf{E}_\theta\left(\left(\overline{X} - 1/\theta\right)^2\right).$$

4. On note T la variable aléatoire $\dfrac{1}{n}\sum_{i=1}^{n}\dfrac{1}{X_i}$.

   (a) En utilisant le résultat de la question 1.d, justifier que :

$$\forall \theta \in \Theta, \quad \mathbf{E}_\theta(T) = \frac{\theta \ln(\theta)}{\theta - 1}.$$

   (b) En déduire que T est un estimateur de θ dont le biais $b_\theta(T)$ est strictement positif.

5. Soit $(x_1, \ldots, x_n) \in B^n$.

   (a) Justifier, pour tout $\theta \in \Theta$, l'égalité :

$$\ln\left(L(x_1, \ldots, x_n, \theta)\right) = n\ln(\theta) - \left(n - \sum_{i=1}^{n} x_i\right)\ln(1-\theta).$$

   (b) En déduire que, lorsque les x_i ne sont pas tous égaux à 1, le nombre $n/\left(\sum_{i=1}^{n} x_i\right)$ est l'unique valeur de θ qui maximise la vraisemblance $L(x_1, \ldots, x_n, \theta)$

6. On note U la variable aléatoire $n/\left(\sum_{i=1}^{n} X_i\right)$.

   (a) établir, pour tout $\theta \in \Theta$ et tout entier $k \geqslant n$, l'égalité :

$$\frac{n}{k} = \theta - \theta^2\left(\frac{k}{n} - \frac{1}{\theta}\right) + \int_{1/\theta}^{k/n}\left(\frac{k}{n} - t\right)\frac{2}{t^3}\,dt.$$

(b) En déduire que U est un estimateur de θ dont le biais $b_\theta(\mathrm{U})$ est donné par :

$$\forall \theta \in \Theta, \quad b_\theta(\mathrm{U}) = \sum_{k=n}^{+\infty} \mathbf{P}\left(\left[\sum_{i=1}^{n} \mathrm{X}_i = k\right]\right) \int_{1/\theta}^{k/n} \left(\frac{k}{n} - t\right) \frac{2}{t^3} \mathrm{d}t.$$

(c) Justifier que $b_\theta(\mathrm{U})$ est strictement positif, quelle que soit la valeur du paramètre θ

7. Dans cette question, on suppose que le nombre des observations est illimité. On dispose donc, pour estimer le paramètre θ, d'une suite $(\mathrm{X}_n)_{n\in\mathbb{N}^*}$ de variables aléatoires mutuellement indépendantes et de même loi. Pour tout entier $n \in \mathbb{N}^*$, on note

$$\mathrm{T}_n = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\mathrm{X}_i} \quad \text{et} \quad \mathrm{U}_n = \frac{n}{\sum\limits_{i=1}^{n} \mathrm{X}_i}.$$

Étudier la convergence des deux suites d'estimateurs $(\mathrm{T}_n)_{n\in\mathbb{N}^*}$ et $(\mathrm{U}_n)_{n\in\mathbb{N}^*}$ du paramètre θ.

≫ *to be continued..*

Python

Exercice 364. ✦ **Retour sur la méthode de Monte-Carlo**
Soient $\mathrm{U}_1, \mathrm{U}_2, \dots, \mathrm{U}_\mathrm{N}$ des variables aléatoires réelles indépendantes de loi uniforme sur $[0,1]$ et N une variable aléatoire de loi géométrique de paramètre p indépendante de la suite $(\mathrm{U}_i)_{i\in\mathbb{N}}$. On pose

$$\mathrm{X} = \max_{1\leqslant i\leqslant \mathrm{N}} \mathrm{U}_i.$$

1. Déterminer la fonction de répartition de la variable aléatoire X.
2. Calculer l'espérance de X.
3. Simuler la variable et vérifier votre résultat.

Exercice 365. ✦ Soit $(\mathrm{X}_1, \dots, \mathrm{X}_n)$ un échantillon suivant la loi $\mathscr{E}(\lambda)$. Nous disposons des deux estimateurs de $\theta = \dfrac{1}{\lambda}$:

- La moyenne empirique $\overline{\mathrm{X}_n} = \frac{1}{n}\sum\limits_{i=1}^{n} \mathrm{X}_i$;
- L'écart-type empirique $\sigma_n = \sqrt{\frac{1}{n-1}\sum\limits_{k=1}^{n}\left(\mathrm{X}_k - \overline{\mathrm{X}_n}\right)^2}$.

On souhaite comparer ces deux estimateurs.

1. Écrire un programme qui prend en argument n et λ et simule $\overline{\mathrm{X}_n}$, puis σ_n.
2. Afficher des histogrammes et comparer les estimateurs.

Exercice 366. ✦
Soient $n \in \mathbb{N}^*$ et $(\mathrm{X}_1, \mathrm{X}_2, \dots, \mathrm{X}_n)$ un échantillon d'une loi de Poisson $\mathscr{P}(\theta)$. On cherche à estimer $\exp(-\theta)$. Pour cela, on pose :

$$\mathrm{A}_n = \left(1 - \frac{1}{n}\right)^{\sum\limits_{i=1}^{n} \mathrm{X}_i} \quad \text{et} \quad \mathrm{B}_n = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{[\mathrm{X}_i = 0]}.$$

1. Écrire deux programmes qui prend en arguments n, θ et simulent respectivement A_n et B_n.
2. On suppose que $\theta = 1$.
   En déduire un programme qui prend en argument n et renvoie une approximation de l'espérance de A_n et B_n.
   Que peut-on conjecturer sur le biais de chacun de ces estimateurs ?
3. Afficher l'histogramme de 10000 réalisations de T_{100} et U_{100}.
   Que peut-en déduire sur la qualité de ces estimateurs ?

CHAPITRE 18

Optimisation sous contraintes linéaires

La science naît du jour où des erreurs, des échecs, des surprises désagréables, nous poussent à regarder le réel de plus près.

RENÉ THOM
Mathématicien français, fondateur de la théorie des catastrophes (1923-2002)

1 Position du problème

Soient $n \in \mathbb{N} \setminus \{0;1\}$ et $p \in \mathbb{N}^*$ avec $p < n$. Considérons de plus $g_1, g_2, \dots, g_p$, p formes linéaires sur $\mathbb{R}^n$. Autrement dit, pour chaque $i \in [\![1;p]\!]$, il existe $\lambda_1, \lambda_2, \dots, \lambda_n$ tels que

$$\forall\, x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n, \quad g_i(x) = \sum_{k=1}^{n} \lambda_k x_k.$$

Notons $\mathscr{C}$, l'ensemble des solutions du système linéaire

$$\begin{cases} g_1(x) & = & b_1 \\ \vdots & \vdots & \vdots \\ g_p(x) & = & b_p, \end{cases}$$

où l'inconnue est un élément x de $\mathbb{R}^n$, et $\mathscr{H}$ l'ensemble des solutions du système homogène associé. En notant que $\mathscr{H} = \bigcap_{i=1}^{p} \operatorname{Ker}(g_i)$, on constate que $\mathscr{H}$ est un sous-espace vectoriel de $\mathbb{R}^n$.
Si $\mathscr{C}$ n'est pas vide, on parle d'espace affine. Dans ce cas, si $x_0 \in \mathscr{C}$ alors on peut écrire

$$\mathscr{C} = \left\{x_0 + h \in \mathbb{R}^n \mid h \in \mathscr{H}\right\}.$$

Définition 257 (extremum sous contrainte)

Soit $f : \mathscr{O} \to \mathbb{R}$ où $\mathscr{O}$ est une partie de $\mathbb{R}^n$.
- On dit que f admet un **extremum local sous la contrainte** $\mathscr{C}$ en un point a si la restriction de f à $\mathscr{C}$ admet un extremum local en a.
- On dit que f admet un **extremum global sous la contrainte** $\mathscr{C}$ en un point a si la restriction de f à $\mathscr{C}$ admet un extremum global en a.

Exemple. Une société de production d'épingles possède deux usines avec des fonctionnements différents. La première usine produit x tonne(s) d'épingles pour un coût $C_1(x)$ et la seconde produit la même quantité avec un coût $C_2(x)$ (en euros). Une analyse des coûts donne :

$$C_1(x) = 3 + 4x \quad \text{et} \quad C_2(x) = 4x^2 + 2.$$

Le coût total d'une production d'une quantité x par la première usine et y, pour la seconde est

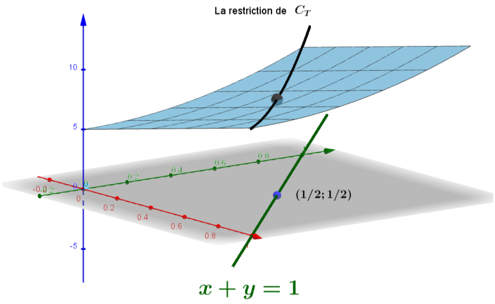
$$C_T(x, y) = C_1(x) + C_2(y) = 5 + 4x + 4y^2.$$

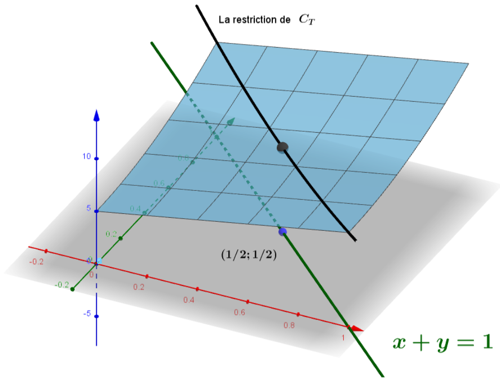
Pour une production d'une tonne, il faut rajouter la contrainte linéaire $g(x,y) = x + y = 1$.
Dans cet exemple, si on souhaite minimiser le coût d'une tonne d'épingles en répartissant la production entre les deux usines, on cherche le minimum global de la fonction C_T sous la contrainte $\mathscr{C} = \{(x,y) \in \mathbb{R}^2 \mid x + y = 1\}$.

2 Recherche des extrema sous contraintes d'égalités linéaires

2.1 Par substitution

Reprenons l'exemple des épingles. Répartissons la production de la tonne entre les deux usines : soit $x \in [0;1]$, la production de l'usine 1. La production de la seconde usine est $y = 1 - x$.
Le coût total est alors $C(x) = C_1(x) + C_2(1-x) = 3 + 4x + 4(1-x)^2 + 2 = 9 - 4x + 4x^2$. La fonction C est polynomiale du second degré. On vérifie l'existence d'un minimum en $x = 1/2$, le coût minimal est alors $C_{min} = C(1/2) = 8$.
En conclusion, on minimise le coût de production en répartissant équitablement la production de la tonne entre les deux usines.





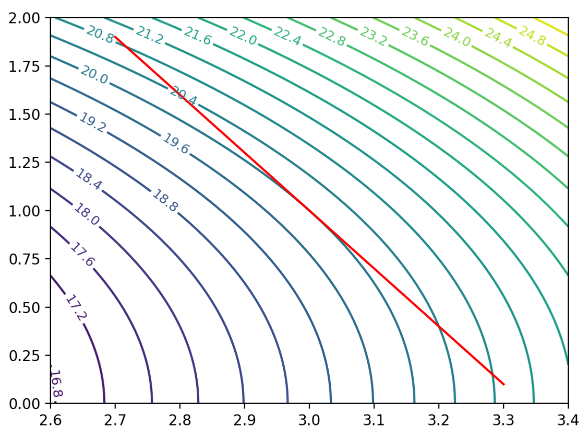
Exercice 367

✧

1. Déterminer le minimum de $f : (x,y) \in \mathbb{R}^2 \mapsto x^2 + y^2 + 10$ sous la contrainte $3x + y = 10$.
2. Tester la pertinence du résultat à l'aide du code Python ci-dessous.

Editeur

```
import numpy as np
import matplotlib.pyplot as plt

x=np.linspace(2.6,3.4,200)
y=np.linspace(0,2,200)
X , Y = np.meshgrid(x,y)
Z = X**2+Y**2+10

graphe = plt.contour(X,Y,Z,25)
plt.clabel(graphe,inline=1,fontsize=9)

x=np.array([2.7,3.3])
plt.plot(x,-3*x+10,'r')
plt.show()
```



Exercice 368

✧ Soit f la fonction définie par : *D'après EDHEC 2020*

$$\forall (x,y,z) \in \mathbb{R}^3, \qquad f(x,y,z) = xe^{x(y^2+z^2+1)}$$

On souhaite étudier les extrema de f sous la contrainte linéaire : $\begin{cases} x = 1 \\ y + z = 0 \end{cases}$.
Montrer que, sous la contrainte, f présente un minimum global au point $(1,0,0)$. Quelle est sa valeur ?

2.2 Par du calcul différentiel

Les énoncés

Théorème 258 (condition nécessaire d'ordre 1)

Soient $f : \mathcal{O} \to \mathbb{R}$ où $\mathcal{O}$ est une partie ouverte de $\mathbb{R}^n$ et $a \in \mathcal{O}$.

Si
- f de classe $\mathscr{C}^1$ sur $\mathcal{O}$.
- f admet un extremum sous la contrainte $\mathscr{C}$ en un point A.

Alors pour tout $h \in \mathscr{H}$, la dérivée de f en a dans la direction h est nulle, c'est-à-dire

$$\forall\, h \in \mathscr{H}, \quad \langle \nabla f(a), h \rangle = 0.$$

Autrement dit, le gradient $\nabla f(a)$ appartient $\mathscr{H}^\perp$.

Vocabulaire. Soient f une fonction de classe $\mathscr{C}^1$ sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$, a un point de $\mathcal{O}$. Si ∇f_a appartient à $\mathscr{H}^\perp$, on dit que a est un point critique de f sous la contrainte $\mathscr{C}$.

Remarque. Avec les notations du début de chapitre, pour tout $i \in \{1; p\}$, g_i est de classe $\mathscr{C}^1$ avec pour tout $x \in \mathcal{O}$

$$\nabla g_i(x) = (\lambda_1, \lambda_2, \ldots, \lambda_n) = \text{ cste.}$$

On notera donc simplement ∇g_i (on confond ici la constante avec la fonction). En particulier, on a

$$g_i(x) = \langle \nabla g_i, x \rangle.$$

Proposition 259 (famille génératrice de l'orthogonal)

Avec les notations introduites précédemment,

$$\mathscr{H}^\perp = \operatorname{Vect}\left(\nabla g_1, \ldots, \nabla g_p\right)$$

Théorème 260 (condition d'ordre 1)

Soient f une fonction de classe $\mathscr{C}^1$ sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$ et $a \in \mathcal{O}$.

Si
- f de classe $\mathscr{C}^1$ sur $\mathcal{O}$.
- f admet un extremum sous la contrainte $\mathscr{C}$ en un point a.

Alors il existe $\left(\lambda_1, \ldots, \lambda_p\right) \in \mathbb{R}^p$ tels que $\nabla f(a) = \sum_{i=1}^{p} \lambda_i \nabla g_i$.

Vocabulaire. Si la famille $\left(\nabla g_1, \ldots, \nabla g_p\right)$ est libre, les réels $\lambda_1, \ldots, \lambda_p$ sont uniques. Ils sont appelés multiplicateurs de Lagrange.

Exercice 369

✦ Déterminer les extrema possibles de la fonction f définie sur $\mathbb{R}^3$ par $f(x, y, z) = xyz$ sous la contrainte

$$\mathscr{C} : \quad \begin{cases} x + y + z &= 0 \\ x - y + z &= 2. \end{cases}$$

⚠ **Attention.** La condition précédente n'est qu'une condition nécessaire et non suffisante. Elle ne fournit que des candidats pour les extrema et il convient ensuite de poursuivre l'analyse.

- Si l'on recherche un extremum *global*, On peut revenir à la définition et étudier le signe de
$$f(x)-f(a) \quad \text{pour } x\in\mathscr{C}$$
ou, à l'aide du changement de variable $h=x-a$, on regarde le signe de
$$f(a+h)-f(a) \quad \text{pour } h\in\mathscr{H}.$$

- Si on cherche un extremum *local*, on peut revenir ou développement limité d'ordre 2. Notons q_a, la forme quadratique de f en a.
$$f(a+h)=f(a)+\langle\nabla f(a),h\rangle+\frac{1}{2}q_a(h)+\|h\|^2\varepsilon(h)$$
avec $\varepsilon(h)\xrightarrow[h\to 0]{}0$ et $h\in\mathscr{H}$. Dans le cas d'un point critique sous la contrainte $\mathscr{H}$, on obtient la simplification :
$$\forall\, h\in\mathscr{H},\quad f(a+h)-f(a)=\frac{1}{2}q_a(h)+\|h\|^2\varepsilon(h).$$
On regarde ensuite le signe de la forme quadratique (restreinte à $\mathscr{H}$).
  - Si $q_a(h)>0$ pour tout $h\in\mathscr{H}$, on a un minimum local.
  - Si $q_a(h)\leqslant 0$ pour tout $h\in\mathscr{H}$, on a un maximum local.
  - Sinon, on ne peut pas conclure.

- Dernier argument : se souvenir qu'il existe toujours un maximum et un minimum pour une fonction continue sur un fermé borné.

Exemples d'études globales

Exemples.
- Soient $f:(x,y,z)\in\mathbb{R}^3\mapsto x^2+y^2+z^2\in\mathbb{R}$ et $g:(x,y,z)\in\mathbb{R}^3\mapsto x+y+z\in\mathbb{R}$.

Déterminer les extremums globaux de f sous la contrainte $\mathscr{C}: g(x,y,z)=3$.
- On considère
$$f: x=(x_1,\dots,x_n)\in(\mathbb{R}_+)^n\mapsto\prod_{k=1}^{n}x_k \quad\text{et}\quad g: x\in\mathbb{R}^n\mapsto\sum_{k=1}^{n}x_k.$$

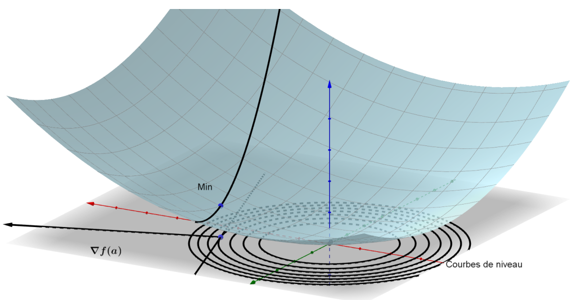
Étudier les extrema de f sous la contrainte $g(x)=c$ où $c\in\mathbb{R}_*^+$.

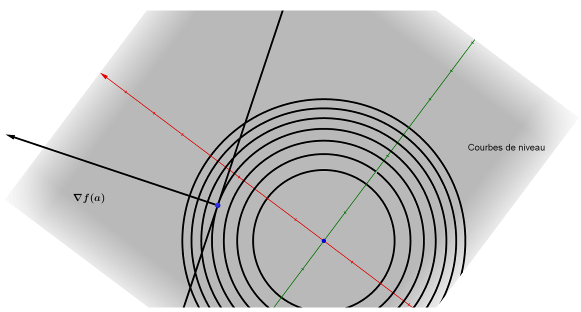
Exercice 370

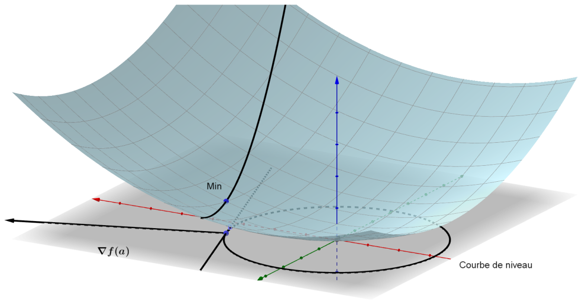
✦ Soit f définie que $\mathbb{R}^3$ par
$$f(x,y,z)=x^2-2xy+yz+y-z.$$
Vérifier que $a=\left(\frac{4}{5},\frac{3}{5},\frac{1}{5}\right)$ est l'unique point critique sous la contrainte
$$\begin{cases}2x-y &=1\\ x+z &=1.\end{cases}$$
Est-ce un maximum ou minimum global? ni l'un ni l'autre?

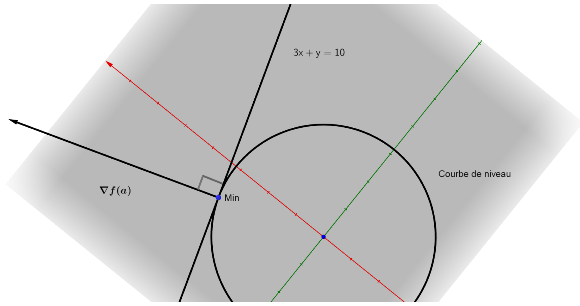
2.3 Interprétation géométrique

Reprenons l'exemple de la fonction $f:(x,y)\in\mathbb{R}^2\mapsto x^2+y^2+10$ sous la contrainte $3x+y=10$.









3 Compléments sur les fonctions définies sur des fermés bornés

Méthode

Existence et calcul des extrema d'une fonction définie sur un fermé borné

Dans la suite, on considère une fonction f de classe $\mathscr{C}^1$ sur un ensemble fermé borné, noté $\mathscr{K}$.

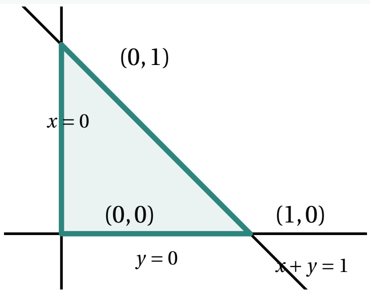
- *I. Existence*
  On vérifie que l'ensemble de départ de f est bien une partie :
  - fermé, par exemple, en montrant qu'il existe φ, une fonction de $\mathbb{R}^n$ et $\mathscr{C}^0$, telle que
  $$\mathscr{K} = \left\{x \in \mathbb{R}^n, \varphi(x) \leqslant c\right\}.$$
  - bornée, en montrant qu'il existe un réel M tel que pour tout $x \in \mathscr{K}$, $\|x\| \leqslant \mathrm{M}$.
  On sait alors que f admet un maximum global et un minimum global sur $\mathscr{K}$, mais on ne sait pas en quels points ils sont atteints.
- *II. Étude sur l'ouvert*
  Soit $\mathscr{O}$, le plus grand ouvert inclus dans $\mathscr{K}$. Autrement dit, $\mathscr{O}$ correspond à $\mathscr{K}$ privé de ses bords. On étudie alors les points critiques pour avoir des points candidats possibles où les extrema sont atteints.
- *III. Étude sur le bord*
  On étudie maintenant la fonction sur le bord $\mathscr{K} \setminus \mathscr{O}$. On retrouve le cas d'une recherche d'extremum sous contrainte.
- *IV. Conclusion*
  On compare les différents valeurs de f aux points obtenus sur l'ouvert et sur le bord pour déterminer le maximum et le minimum global.

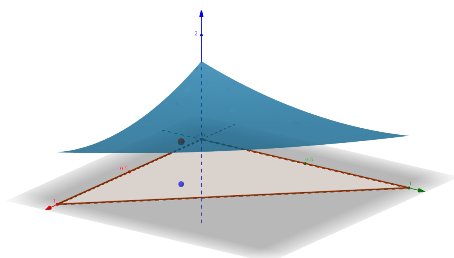
Exemple. Considérons la fonction f définie sur $\mathbb{R}^2$ par

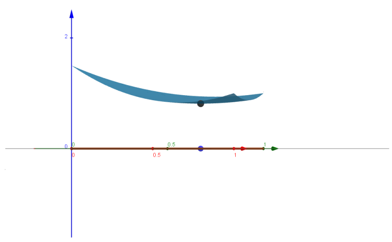
$$f : (x, y) \mapsto 3x^2 + 2y^2 - 4x - 3y + 4xy + 3.$$

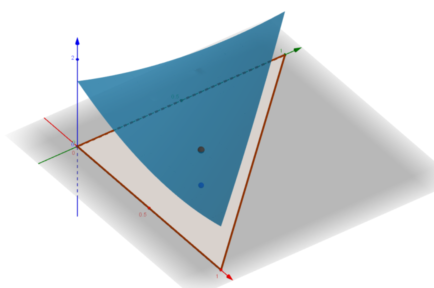
Étudier les extrema de la restriction de f à l'ensemble

$$\mathscr{K} = \{(x, y) \in \mathbb{R}^2 \mid x \geqslant 0,\ y \geqslant 0, x + y \leqslant 1\}.$$









Exercices - TD

Exercice 371. ✦ Soit f définie sur $\mathcal{O} = \left(\mathbb{R}_+^*\right)^3$ par $f(x,y,z) = x\ln x + y\ln y + z\ln z$..

1. Justifier soigneusement que f est de classe $\mathcal{C}^1$ sur $\mathcal{O}$.
2. Soit $c \in \mathbb{R}_*^+$ et $\mathscr{C} = \left\{f(x,y,z) \in \mathcal{O} \mid x+y+z = 3c\right\}$. Montrer que f admet un unique point critique sous la contrainte $\mathscr{C}$.
3. Justifier que ce point critique correspond à un minimum local sous contrainte pour f.

Exercice 372. ✦✦ *D'après oraux ESCP*

Soit $n \in \mathbb{N} \setminus \{0;1\}$ et f la fonction définie sur $\left(\mathbb{R}_*^+\right)^n$ par $f(x_1,\dots,x_n) = \sum\limits_{k=1}^{n} {x_k}^4$.

Étudier les extrema locaux et globaux de f sous la contrainte $\sum\limits_{i=1}^{n} x_i = n$.

Exercice 373. ✦✦

Soient $n \in \mathbb{N}^*$, $\alpha_1,\dots,\alpha_n$ des réels strictement positifs tels que $\sum\limits_{i=1}^{n} \alpha_i = 1$. Soient f et g définies sur $\left(\mathbb{R}^+\right)^n$ par

$$f(x_1,\dots,x_n) = \prod_{i=1}^{n} {x_i}^{\alpha_i} \quad \text{et} \quad g(x_1,\dots,x_n) = \sum_{i=1}^{n} \alpha_i x_i.$$

Posons de plus $\mathscr{K} = \left\{(x_1,\dots,x_n) \in \left(\mathbb{R}^+\right)^n \mid g(x_1,\dots,x_n) = 1\right\}$.

1. Déterminer les points critiques de f sur $\left(\mathbb{R}_+^*\right)^n$ sous la contrainte $g(x_1,\dots,x_n) = 1$.
2. Montrer que f possède un maximum M sur $\mathscr{K}$ et que celui-ci est atteint sur $\mathscr{K} \cap \left(\mathbb{R}_+^*\right)^n$.
3. En déduire que
$$\forall\ (x_1,\dots,x_n) \in \left(\mathbb{R}^+\right)^n, \quad \prod_{i=1}^{n} x_i^{\alpha_i} \leqslant \sum_{i=1}^{n} \alpha_i x_i.$$

Exercice 374. ✦✦ *d'après EDHEC 2001*

On désigne par n et r deux entiers naturels vérifiant : $n \geqslant 2$ et $r \geqslant 3$. On considère une épreuve aléatoire pouvant aboutir à r résultats différents $R_1, R_2, \dots, R_r$ de probabilités respectives $x_1, x_2, \dots, x_r$. On admet que, pour tout i de $[\![1,r]\!]$, $0 < x_i < 1$. On effectue n épreuves indépendantes du type de celle décrite ci-dessus.
Pour tout i de $[\![1,r]\!]$, on note X_i la variable aléatoire qui vaut 1 si le résultat numéro i n'est pas obtenu à l'issue de ces n épreuves et qui vaut 0 sinon.
On désigne par X la variable aléatoire égale au nombre de résultats qui n'ont pas été obtenus à l'issue des n épreuves.

1. Exprimer la variable X en fonction des variables $X_1, X_2, \dots, X_r$. Donner la loi de X_i pour tout i de $\{1,2,\dots,r\}$.
2. En déduire que l'espérance de X est
$$\mathbf{E}(X) = \sum_{i=1}^{r} \left(1 - x_i\right)^n.$$
3. En optimisant sous la contrainte $\sum\limits_{i=1}^{n} x_i = 1$, justifier l'existence et donner la valeur du minimum local de $\mathbf{E}(X)$.

≫ Voir sujet DS 6 pour une démonstration par substitution.

Exercice 375. ✦✦✦ **Optimisation et estimateur**

Soit $n \in \mathbb{N}^*$. On désigne par $(X_1,\dots,X_n)$ un n-échantillon d'une loi de Poisson de paramètre λ inconnu.
Montrer que parmi toutes les combinaisons linéaires des $X_1,\dots,X_n$, il existe un unique estimateur T_n sans biais de λ et qui est de variance minimale. Reconnaître cet estimateur.

Exercice 376. ✦✦✦ **Entropie dans le cas discret** *d'après ESCP 2021, num 2.15*

Soit un entier $n \in \mathbb{N}^*$. Pour toute variable aléatoire Z définie sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ qui prend n valeurs réelles distinctes $z_1,\dots,z_n$ (i.e. $Z(\Omega) = \{z_1,\dots,z_n\}$ et $p_k = \mathbf{P}\left(Z = z_k\right) \neq 0$ pour tout $k \in [\![1,n]\!]$), on appelle entropie de Z, le nombre réel défini par :

$$H(Z) = -\sum_{k=1}^{n} p_k \ln\left(p_k\right).$$

Pour tout $x = (x_1,\dots,x_n) \in]0,1[^n$, on note $h_n(x) = h_n(x_1,\dots,x_n) = -\sum\limits_{k=1}^{n} x_k \ln\left(x_k\right)$.

1. Calculer l'entropie d'une variable aléatoire U qui suit une loi uniforme sur un ensemble $\{u_1,\dots,u_n\}$.
2. Justifier que la fonction h_n est de classe $\mathscr{C}^2$ sur l'ouvert $]0,1[^n$ et calculer en tout point $x \in]0,1[^n$., le gradient $\nabla h_n(x)$ et la matrice hessienne $\nabla^2\left(h_n\right)(x)$ de h_n.
3. Montrer que la condition nécessaire d'ordre 1 pour un extremum de la fonction h_n est vérifiée en un unique point critique x^* que l'on précisera.

4. Fixons $x \in]0,1[^n$ tel que $x_1 + \cdots + x_n = 1$ et notons $u = x - x^*$.
   (a) Vérifier que, pour tout $t \in [0,1]$, on a : $x^* + tu \in]0,1[^n$.
   On note alors ψ la fonction définie sur $[0,1]$ par : $\psi(t) = h_n\left(x^* + tu\right)$.
   (b) En utilisant la formule de Taylor avec reste intégral à l'ordre 1 pour ψ entre les points 0 et 1, montrer que h_n admet en x^* un maximum global sous la contrainte $x_1 + \cdots + x_n = 1$. Ce maximum est-il atteint en d'autres points que x^* ?
5. Parmi les variables aléatoires prenant n valeurs (chacune avec une probabilité non nulle), quelles sont les lois de celles qui ont la plus grande entropie ?

Exercice 377. ✦✦ **Exemple sur un fermé borné** *d'après oral ESCP 2002*

Soit $\mathrm{D} = \left\{(x, y) \in \mathbb{R}^2, -1 \leqslant x \leqslant y \leqslant 1\right\}$ et f la fonction définie sur D par

$$f(x, y) = (y - x)^2 + 6xy.$$

1. La fonction f admet-elle des extremums ?
2. Déterminer les points critiques de f. Déterminer les extremums de f.

Toute bonne chose a une fin, sauf le saucisson qui en a deux.

PROVERBE DANOIS

CHAPITRE 19

Statistiques

> *La statistique est aujourd'hui un fait social total : elle règne sur la société, régente les institutions et domine la politique. Un vêtement de courbes, d'indices, de graphiques, de taux recouvre l'ensemble de la vie. L'éducation disparaît derrière les enquêtes PISA, l'université derrière le classement de Shanghai, les chômeurs derrière la courbe du chômage... La statistique devait refléter l'état du monde, le monde est devenu un reflet de la statistique.*
>
> OLIVIER REY
> Mathématicien et philosophe contemporain

Que ce soit à l'issue d'une enquête, par observation directe ou à la suite d'une expérimentation, nous disposons à présent d'un ensemble de données qu'il convient maintenant de traiter. C'est-à-dire, les organiser, les représenter et en déterminer les éléments caractéristiques comme la moyenne, la variance ou l'écart-type. Tel est l'enjeu de la statistique descriptive.

1 Rappels en statistiques univariées

Soit Ω un ensemble.

- Ω désigne **la population** qui fera l'objet de l'étude.
- Les éléments $\omega_i \in \Omega$ sont les **individus**.
- Une **variable statistique** (ou **caractère**) X sur la population Ω est une application $X : \Omega \to F$. Dans le cas où F est une partie de $\mathbb{R}$, on dit que le caractère est **quantitatif**. Sinon, on dit que le caractère est qualitatif.

  Remarque. Dans la suite, on ne considère que les caractères quantitatifs.

- Si Ω est un ensemble fini, le nombre d'éléments de Ω, noté $\mathrm{Card}(\Omega)$ est **l'effectif de la population**.
- Quand il n'est pas possible d'étudier chaque individu de la population, on étudie seulement les individus d'une partie finie E de la population Ω. Dans ce cas, la partie E est appelée **échantillon**. Le nombre d'individus de l'échantillon, $\mathrm{Card}(E)$, est **la taille de l'échantillon**. Ainsi, si N désigne la taille, on peut écrire $E = \{e_1, e_2, \cdots, e_N\}$ où les e_i sont distincts deux à deux.

On distinguera deux types de variables statistiques.

- Si l'ensemble des valeurs prises par la variable, noté $X(\Omega)$, est un ensemble fini, on dit que la variable quantitative est **discrète**.
- Dans le cas contraire, on dit que la variable quantitative est **continue.**

1.1 Cas discret

Soit X une variable statistique discrète et $E = \{e_1, e_2, \cdots, e_N\}$ un échantillon.

- La donnée du N-uplet des observations $x = \big(X(e_1), X(e_2), \cdots, X(e_N)\big)$ définit une **série statistique**.
- L'échantillon étant un ensemble fini, l'ensemble des valeurs prises par la variable X sur E est aussi fini. On peut l'écrire $X(E) = \{m_1, m_2, \cdots, m_p\}$ où $m_1 < m_2 < \cdots < m_p$ où chaque m_i est une **valeur** ou **modalité.**

- **L'effectif d'une modalité** m_i est le nombre d'individu de E pour lequel le caractère prend la modalité m_i. C'est-à-dire $n_i = \text{Card}\left\{e \in \text{E} \mid \text{X}(e) = m_i\right\}$. Si N est la taille de l'échantillon, $\text{N} = \sum_{i=1}^{p} n_i$.
- **La fréquence d'une modalité** m_i est la quantité $f_i = \dfrac{\text{Effectif de la modalité}}{\text{Taille de l'échantillon}} = \dfrac{n_i}{\text{N}}$. On a alors $\sum_{i=1}^{p} f_i = 1$.
- **La fréquence cumulée d'une modalité** m_i est la somme de toutes les fréquences des modalités qui lui sont inférieures. Autrement dit, $\text{F}_i = \sum_{m_j \leqslant m_i} f_j$.

Remarque. Donner une série statistique est équivalent à la donnée du couple (m, n) où

- $m = (m_1, m_2, \cdots, m_p)$ est le p-uplet constitué des modalités,
- $n = (n_1, n_2, \cdots, n_p)$ est le p-uplet constitué des effectifs.

Il est alors commode de représenter une série statistique à l'aide d'un tableau :

Modalités	m_1	m_2	...	m_p	Total
Effectifs	n_1	n_2	...	n_p	N

1.2 Cas continu

On découpe l'ensemble des valeurs possibles $\text{X}(\Omega)$ en un certain nombre d'intervalles. Notons $p \in \mathbb{N}^*$ le nombre d'intervalles choisis. Ces intervalles doivent être deux à deux disjoints, et leur réunion est égale à (ou contient) l'ensemble des valeurs possibles. Plus précisément, en notant $a_1, a_2, \ldots, a_{p+1} \in \mathbb{R}$, les bornes de tous ces intervalles avec $a_1 < a_2 < \cdots < a_{p+1}$, on a

$$\text{X}(\Omega) \subset \bigcup_{i=1}^{p} [a_i ; a_{i+1}[.$$

L'intervalle $[a_i ; a_{i+1}[$ est une **classe**.

Remarques. On peut ainsi définir l'effectif d'une classe et sa fréquence. On peut aussi utiliser ces définitions pour une variable discrète lorsque l'ensemble des valeurs prises est trop important.

2 Paramètres

Statistics may be regarded as
(i) the study of populations,
(ii) as the study of variation,
(iii) as the study of methods of the reduction of data.

Ronald Aylmer Fisher
biologiste et statisticien britannique (1890-1962).

2.1 Paramètres de position

Définition 261 (La moyenne empirique)

Soit $x = (x_i)_{1 \leqslant i \leqslant \text{N}}$, une série statistique. On définit la **moyenne** de la série par

$$\overline{x} = \frac{1}{\text{N}} \sum_{i=1}^{\text{N}} x_i.$$

Si $(m_i)_{1 \leqslant i \leqslant p}$, $(n_i)_{1 \leqslant i \leqslant p}$ et $(f_i)_{1 \leqslant i \leqslant p}$ désignent respectivement les modalités, les effectifs et les fréquences, on a aussi

$$\overline{x} = \frac{1}{\text{N}} \sum_{i=1}^{p} n_i\, m_i = \sum_{i=1}^{p} f_i\, m_i$$

La moyenne empirique s'obtient avec Python par la commande `np.mean(x)`.

Définition 262 (La médiane d'une série statistique)

Soit $x = (x_i)_{1 \leqslant i \leqslant N}$ une série statistique ordonnées suivant l'ordre croissant. On définit la **médiane** de la série statistique par

- x_p si N est impair avec $p = (N+1)/2$,
- $\frac{x_p + x_{p+1}}{2}$ si N est pair avec $p = N/2$.

Autrement dit, la médiane un nombre réel m_e tel que le nombre d'individus pour lesquels X prend une valeur inférieure ou égale à m_e soit égal au nombre d'individus pour lesquels X prend une valeur supérieure ou égale à m_e.

2.2 Paramètres de dispersion

Définition 263 (La variance empirique et l'écart-type)

Soit x, une série statistique.
- **La variance** de la série est le réel, noté s_x^2, défini par

$$s_x^2 = \frac{1}{N}\left((x_1 - \overline{x})^2 + (x_2 - \overline{x})^2 + \ldots + (x_N - \overline{x})^2\right) = \frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})^2.$$

- **L'écart-type**, s_x d'une série statistique est défini comme la racine carrée de la variance.

Remarques.
- Si les modes $m_1, m_2, \cdots, m_p$ de la série sont donnés avec leurs effectifs $n_1, n_2, \cdots, n_p$ ou leurs fréquences $f_1, f_2, \ldots, f_p$, alors

$$s_x^2 = \frac{1}{N}\sum_{i=1}^{p} n_i (m_i - \overline{x})^2 = \sum_{i=1}^{p} f_i (m_i - \overline{x})^2.$$

- Une variance est toujours un nombre positif. L'écart-type est donc bien défini.
- Une série a une variance nulle si et seulement si toutes les valeurs de la série sont identiques.

Les commandes Python sont `np.var(x)`, pour la variance et `np.std(x)`, pour l'écart-type.

En pratique, on utilise la formule suivante pour calculer la variance.

Théorème 264 (Formule de Koenig-Huygens)

Soit x une série statistique de modalités $(m_i)_{1 \leqslant i \leqslant p}$, d'effectifs $(n_i)_{1 \leqslant i \leqslant p}$ et de variance s_x^2, alors

$$s_x^2 = \left(\frac{1}{N}\sum_{i=1}^{p} n_i {m_i}^2\right) - \overline{x}^2.$$

Remarque. Si x^2 désigne la série dont chaque donnée est ${x_i}^2$. On démontre que $\frac{1}{N}\sum_{i=1}^{p} n_i {m_i}^2$ est la moyenne $\overline{(x^2)}$. La formule de formule de Koenig-Huygens devient alors

$$s_x^2 = \overline{(x^2)} - (\overline{x})^2.$$

3 Statistiques bivariées

3.1 Premières définitions

Définition 265 (La covariance empirique)

Soient $x = (x_i)_{1\leqslant i\leqslant N}$ et $y = (y_i)_{1\leqslant i\leqslant N}$, deux séries statistiques. On définit la **covariance** des deux séries par

$$\mathrm{Cov}(x, y) = \frac{1}{N}\sum_{i=1}^{N}\left(x_i - \overline{x}\right)\left(y_i - \overline{y}\right).$$

On définit aussi le **coefficient de corrélation empirique** par $\rho_{x,y} = \dfrac{\mathrm{Cov}(x,y)}{\sigma_x\sigma_y} \in [-1;1]$.

Proposition 266 (Calcul de la covariance)

Soient $x = (x_i)_{1\leqslant i\leqslant N}$ et $y = (y_i)_{1\leqslant i\leqslant N}$, deux séries statistiques. On a

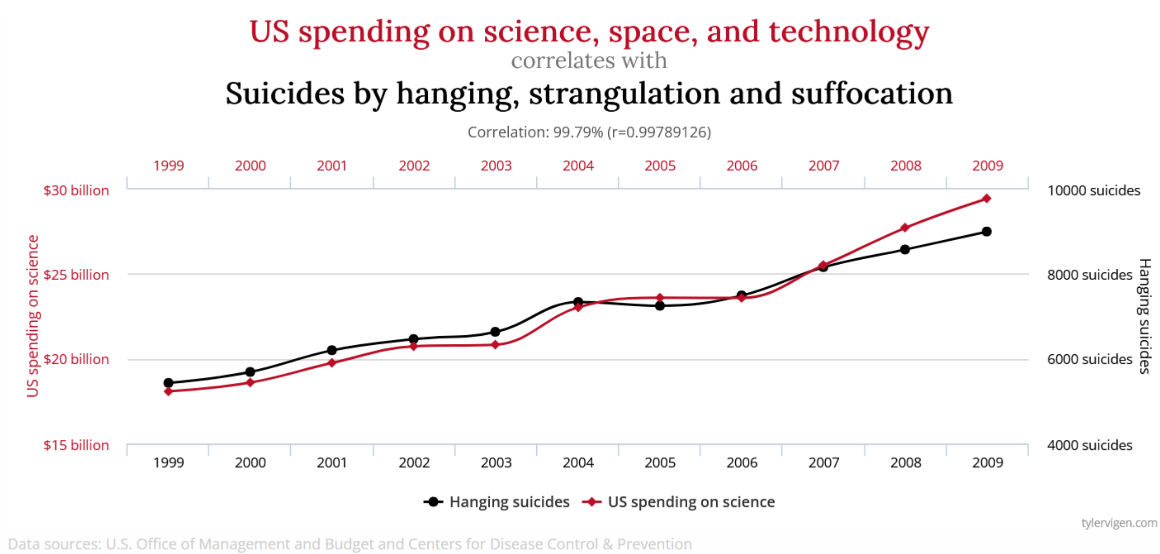
$$\mathrm{Cov}(x, y) = \overline{x * y} - \overline{x}\cdot\overline{y}$$

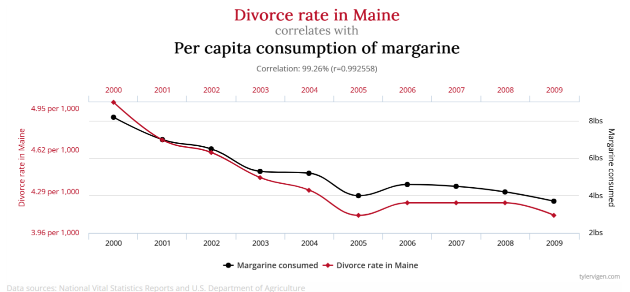
où $\overline{x * y}$ désigne la moyenne de la série $x * y = (x_i \cdot y_i)_{1\leqslant i\leqslant N}$.
Dès lors, la commande Python pour obtenir la covariance empirique est

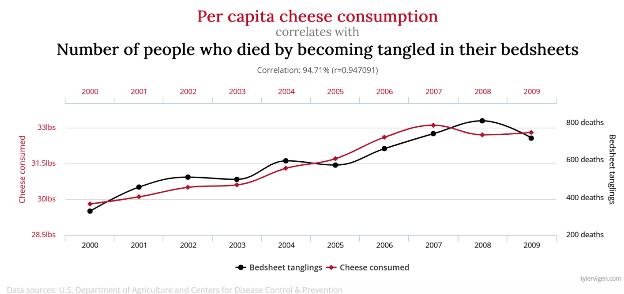
```
np.mean(x*y)-np.mean(x)*np.mean(y)
```

Corrélation n'est pas causalité

Quelques exemples issus du site : *http ://tylervigen.com/spurious-correlations.*







Les faits sont têtus, il est plus facile de s'arranger avec les statistiques.

Mark Twain

3.2 Nuage de points, droite de régression

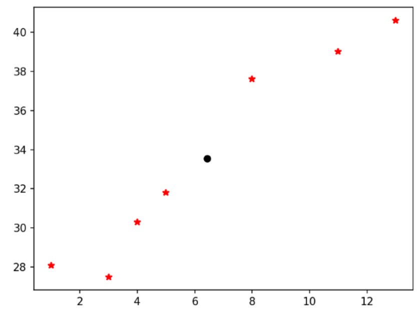
Dans la suite, x, y désigne deux séries statistiques. La question est de savoir dans quelle mesure l'une des deux, dite **expliquée**, dépend de l'autre, dite **explicative**.

Définition 267 (Nuage de points)

Soient $x = (x_i)_{1\leqslant i\leqslant \mathrm{N}}$ et $y = (y_i)_{1\leqslant i\leqslant \mathrm{N}}$, deux séries statistiques.
Le **nuage de points** associé à ces deux séries est l'ensemble des points M_i du plan de coordonnées (x_i, y_i) où $i \in [\![1, \mathrm{N}]\!]$.

Remarque. Le point $(\overline{x}, \overline{y})$ est le **point moyen** du nuage.

• Voici le code pour tracer un nuage de points associé à deux séries statistiques de même effectif, ainsi que sont point moyen.

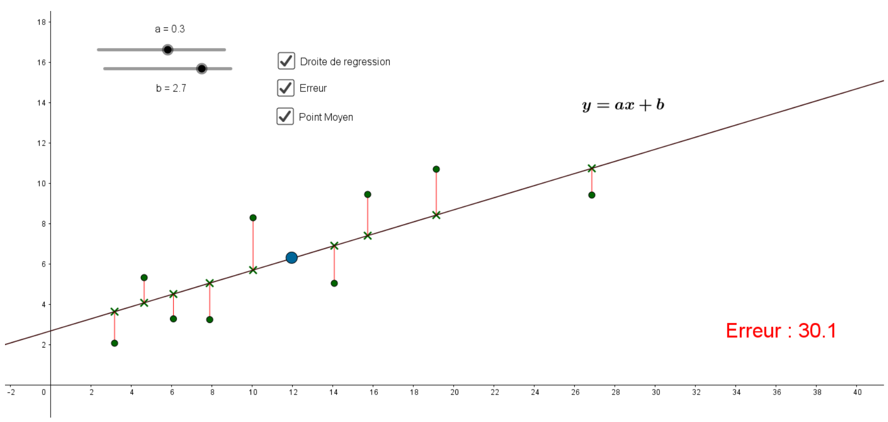


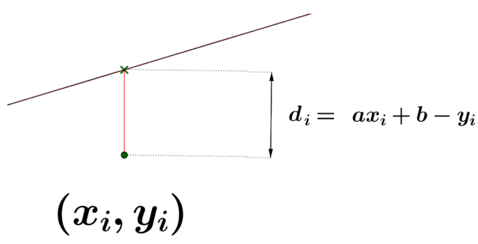
Editeur

```
import matplotlib.pyplot as plt
import numpy as np

# Exemples avec deux séries
x=np.array([1,3,4,5,8,11,13])
y=np.array([28.1,27.5,30.3,31.8,37.6,39,40.6])


# Tracé du nuage de points et du point moyen

plt.plot(x,y,'r*')
plt.plot(np.mean(x),np.mean(y),'ko')
```

Soit n un entier supérieur à 2.
Considérons n points de $\mathbb{R}^2$, $(x_1, y_1), \ldots, (x_n, y_n)$ non alignés verticalement. On cherche la droite qui « approxime » au mieux ces n points. Si on note $y = ax + b$, l'équation d'une droite, on cherche à minimiser l'erreur

$$\mathrm{E}_r = \sum_{i=1}^{n} {d_i}^2 = \sum_{i=1}^{n} \left(ax_i + b - y_i\right)^2.$$





Point de vue algèbre linéaire

Traduisons matriciellement le problème. Posons.

$$\mathrm{X} = \begin{bmatrix} a \\ b \end{bmatrix}, \quad \mathrm{A} = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \quad \text{et} \quad \mathrm{B} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \text{de sorte que} \quad \mathrm{AX} - \mathrm{B} = \begin{bmatrix} ax_1 + b - y_1 \\ ax_2 + b - y_2 \\ \vdots \\ ax_n + b - y_n \end{bmatrix}.$$

Si on considère le produit scalaire canonique sur $\mathcal{M}_{n,1}(\mathbb{R})$ et la norme associée

$$\mathrm{E}_r = \|\mathrm{AX} - \mathrm{B}\|^2.$$

D'après les résultats sur les projecteurs et le théorème des moindres carrés, on a montré qu'il existe une unique solution X_0 donnée par l'équation ${}^t\mathrm{AAX}_0 = {}^t\mathrm{AB}$. On obtient alors ses composantes a et b :

$$a = \rho(x, y)\sqrt{\frac{\mathbf{V}(y)}{\mathbf{V}(x)}} \quad \text{et} \quad b = \overline{y} - a\overline{x}.$$

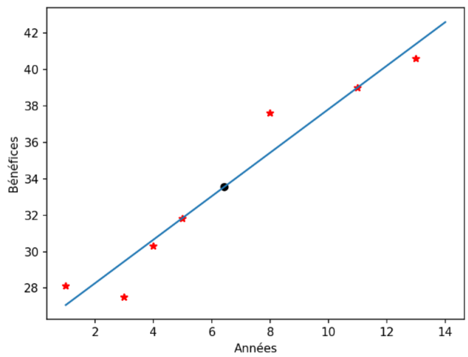
L'équation de la droite est alors :

$$y - \overline{y} = \rho(x, y)\frac{\sigma(y)}{\sigma(x)}(x - \overline{x}).$$

Exemple. On peut compléter l'exemple Python précédent en affichant la droite de régression. On comprend aussi sur cet exemple que la droite de régression peut être utiliser pour faire de la prédiction en prolongeant la droite. Précisons aussi qu'il existe des commandes Python plus directes pour afficher la droite de régression.

Editeur

```
# Tracé de la droite de régression

plt.xlabel('Années')
plt.ylabel('Bénéfices')
COVxy=np.mean(x*y)-np.mean(x)*np.mean(y)
a=COVxy/np.std(x)**2
b=np.mean(y)-a*np.mean(x)
t=np.linspace(1,14,2)
plt.plot(t,a*t+b)
```



Exercice 378

✦ Vérifier que le point moyen appartient à la droite de régression.

Point de vue fonctions de plusieurs variables

Exercice 379. ✦✦ *D'après HEC 2008 E*

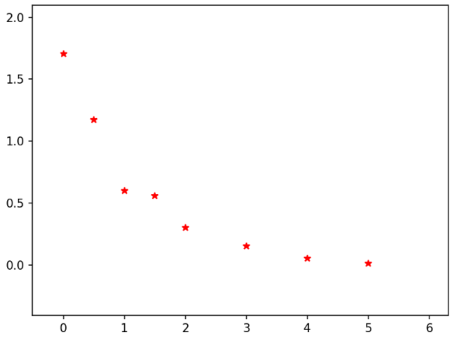
Soit f la fonction définie sur $\mathbb{R}^2$ à valeurs réelles qui, à tout couple (a, b) de $\mathbb{R}^2$, associe le réel $f(a, b)$ tel que :

$$f(a, b) = \sum_{k=1}^{n} \left(ax_k + b - y_k\right)^2$$

1. Justifier que f est de classe $\mathscr{C}^2$ sur $\mathbb{R}^2$.
2. (a) Écrire le système d'équations (S) permettant de déterminer les points critiques de f.
   (b) Résoudre le système (S).
   En déduire que f admet un unique point critique $(\hat{a}, \hat{b})$ que l'on exprimera en fonction de $\overline{x}, \overline{y}, \sigma_x^2$ et $\mathrm{cov}(x, y)$.
   (c) Montrer que ce point critique correspond à un minimum local de f.
   (d) Établir la formule suivante : $f(\hat{a}, \hat{b}) = n\sigma_y^2\left(1 - r^2(x, y)\right)$.
3. (a) Montrer que l'on a : $|r(x, y)| \leqslant 1$.
   (b) Que peut-on dire du nuage de points lorsque $|r(x, y)| = 1$?

3.3 Un exemple d'ajustement linéaire

Il arrive parfois que le nuage ne s'accorde pas à une droite mais à une courbe d'une fonction classique.

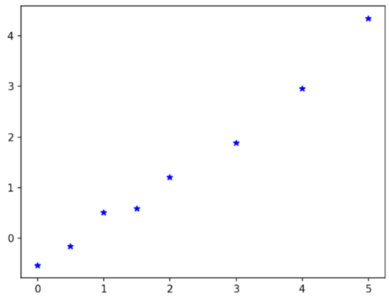


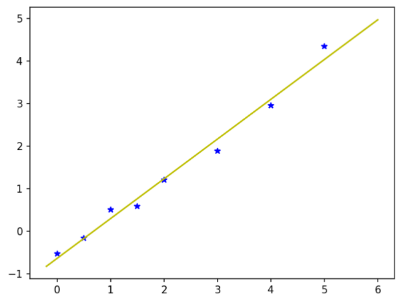
On peut alors supposer que le nuage s'organise autour de la courbe d'équation $y = f(ax+b)$ pour une certaine fonction f. Si f est bijective, on introduit la série $z = f^{-1}(x)$. Prenons l'exemple des séries :

Editeur

```
x=np.array([0,0.5,1,1.5,2,3,4,5])
y=np.array([1.707, 1.173, 0.601, 0.558,
    0.299,0.152, 0.052, 0.013])

plt.plot(x,y,'r*')
```

Dans l'exemple, on va tester avec la fonction f définie sur $\mathbb{R}$ par $f(t) = \exp(-t)$.
On procède alors à une régression linéaire sur les séries x, z pour estimer les réels a et b.





Console

```
>>> a
0.9349559793558717
>>> b
-0.6381866319796354
```

Testons le résultat, en comparant la courbe obtenue avec le nuage de points.

Editeur

```
z=-np.log(y)

COVxz=np.mean(x*z)-np.mean(x)*np.mean(z)
a=COVxz/np.std(x)**2
b=np.mean(z)-a*np.mean(x)

plt.plot(x,y,'r*')
t=np.linspace(-0.2,6,100)
plt.plot(t,np.exp(-(a*t+b)))
plt.show()
```

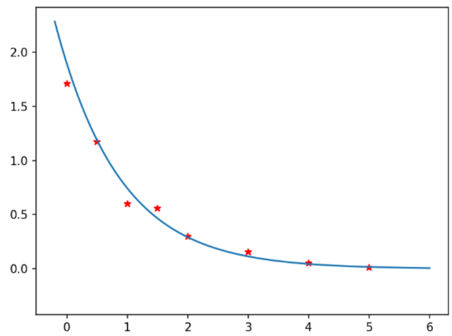


Table des matières