253 - Utilisation de la notion de convexité en analyse

Cadre : E un $\mathbb{K}$-ev ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$).

I - Parties convexes, fonctions convexes

1 - Définitions et exemples.

[OA] p71 **Def 1 :** Soient $A, B \in E$. On appelle segment d'extrémités A et B (noté $[A,B]$) l'ensemble $\{\lambda A + (1-\lambda)B,\ \lambda \in [0,1]\}$.

Def 2 : Soit $C \neq \emptyset$ partie de E. On dit que C est convexe si $\forall (A,B) \in C^2$, $[A,B] \subset C$.

Ex 3 : Un sev est convexe. Dans $\mathbb{R}$, les intervalles sont les parties convexes.

[OA] **Appli 4 :** E evn. Si C est un convexe de E, alors $\overline{C}$ et $\mathring{C}$ sont convexes.

Def 5 : I convexe de E, soit f une application définie sur I à valeurs réelles. On dit que f est convexe si : $\forall x, y \in I$, $\forall \lambda \in [0,1]$; $f((1-\lambda)x + \lambda y) \leq (1-\lambda) f(x) + \lambda f(y)$. [ROM] p233

Thm 6 : f est convexe $\Leftrightarrow$ son épigraphe $\mathcal{E}(f) = \{(x,y) \in I \times \mathbb{R} \text{ tq } y \geq f(x)\}$ est convexe dans $E \times \mathbb{R}$.

Def 7 : f est strictement convexe si $\forall x, y \in I$, $\forall \lambda \in]0,1[$, on a : $f((1-\lambda)x + \lambda y) < (1-\lambda) f(x) + \lambda f(y)$.

Def 8 : f est concave (resp. strictement concave) si la fonction $-f$ est convexe (resp. strictement convexe).

Prop 9 : $f : I \to \mathbb{R}$ convexe. Alors : $\forall x_1, \dots, x_n \in I$, $\forall \alpha_1, \dots, \alpha_n > 0$,
$$f\left(\frac{\alpha_1 x_1 + \dots + \alpha_n x_n}{\alpha_1 + \dots + \alpha_n}\right) \leq \frac{\alpha_1 f(x_1) + \dots + \alpha_n f(x_n)}{\alpha_1 + \dots + \alpha_n}.$$
[GOU] p95

2 - Enveloppe convexe

Def 10 : Soit A partie de E. Il existe une plus petite partie convexe de E contenant A. On l'appelle enveloppe convexe de A (notée $\mathrm{Conv}(A)$). C'est aussi l'ensemble des barycentres des points de A affectés de coefficients positifs, c'est l'ensemble des x tels que $\exists x_1, \dots, x_n \in A$, $\exists \lambda_1, \dots, \lambda_n \geq 0$, $\sum_{i=1}^n \lambda_i = 1$, $x = \sum_{i=1}^n \lambda_i x_i$. [GOU]

Ex 11 : (Thm de Lucas) Soit $P \in \mathbb{C}[X]$, $\deg(P) \geq 2$, alors les racines de P' appartiennent à l'enveloppe convexe des racines de P. [GOU] Algèbre p66

Thm 12 : (Thm de Carathéodory) Soit E un $\mathbb{R}$-ev de dim $n \in \mathbb{N}^*$, soit $x \in E$ le barycentre de p vecteurs $x_1, \dots, x_p \in E$ affectés de coefficients positifs $\lambda_1, \dots, \lambda_p$. Alors : $\exists I \subset \{1, \dots, p\}$ tel que $\mathrm{Card}(I) \leq n+1$ et tel que x soit barycentre des $(x_i)_{i \in I}$ affectés de coefficients positifs. [GOU] p54

Cor 13 : Soit E un $\mathbb{R}$-ev de dim finie et A partie compacte de E. Alors $\mathrm{Conv}(A)$ est compacte.

3 - Fonctions convexes et régularité

Thm 14 : f est convexe ssi pour tout couple (A,B) de points du graphe de f avec $A = (a, f(a))$, $B = (b, f(b))$, $a < b$, la courbe représentative de la restriction de f à l'intervalle $[a,b]$ est en dessous du segment $[A,B]$ (Figure 1). [ROM] p234

Prop 15 : $f : I \to \mathbb{R}$ est convexe ssi $\forall x_0 \in I$, $g_{x_0} : I \setminus \{x_0\} \to \mathbb{R}$, $x \mapsto \dfrac{f(x) - f(x_0)}{x - x_0}$ est croissante. [GOU] p94

Prop 16 : $f : I \to \mathbb{R}$ convexe. Alors f possède une dérivée à droite et à gauche en tout point de $\mathring{I}$. Elle est donc continue sur $\mathring{I}$. De plus, f'_d et f'_g sont croissantes sur $\mathring{I}$ et $f'_g(x) \leq f'_d(x)$ pour tout $x \in \mathring{I}$.

Rq 17 : $f : I \to \mathbb{R}$ convexe, alors $\forall a, b, c \in I$, $a < b < c$, on a :
$$\frac{f(b) - f(a)}{b - a} \leq \frac{f(c) - f(a)}{c - a} \leq \frac{f(c) - f(b)}{c - b}. \quad \text{(Figure 2)}.$$

Prop 18 : Soient C partie convexe de $\mathbb{R}^n$ et $f : C \to \mathbb{R}$ fonction convexe. Alors f est lipschitzienne sur tout compact contenu dans $\mathring{C}$. [OA] p28

Thm 19 : Soient U ouvert convexe de $\mathbb{R}^n$ et $f : U \to \mathbb{R}$. Si f est différentiable sur U, alors les propriétés suivantes sont équivalentes :
i) f est convexe
ii) le graphe de f est "au dessus de ses tangentes" : $\forall x, y \in U$, $f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle$
iii) $\nabla f : U \to \mathbb{R}^n$ est "monotone" : $\forall x, y \in U$, $\langle \nabla f(y) - \nabla f(x), y - x \rangle \geq 0$.
iv) Si f est 2 fois différentiable sur U, $Hf(x)$ est positive : $\forall x, h \in U$, $\langle Hf(x)h, h \rangle \geq 0$
($Hf(x)$ est la matrice hessienne de f en x).

Ex 20 : La fonction $\ln$ est concave sur $\mathbb{R}_+^*$ car la dérivée seconde de $-\ln$ est positive.

Ex 21 : Soit $f : \mathbb{R} \to \mathbb{R}$ fonction convexe majorée sur $\mathbb{R}$. Alors f est constante. [GOU] p98

Ctre ex 22 : Ceci est faux si f est définie, convexe et majorée sur $\mathbb{R}^+$. Prendre $x \mapsto \dfrac{1}{x+1}$.

Prop 23 : Si f est continue sur I et vérifie $f\left(\frac{x+y}{2}\right) \leq \frac{f(x) + f(y)}{2}$ $\forall x, y \in I$, alors f est convexe. (admis) [GOU] p101

II - Inégalités de convexité

1) Quelques inégalités classiques

Prop 24 : La fonction exponentielle est strictement convexe sur $\mathbb{R}$. La fonction logarithme est strictement concave sur $\mathbb{R}_+^*$. [ROM] p246

Prop 25 : La fonction sinus est strictement convexe sur $[(2k+1)\pi, (2k+2)\pi]$ et strictement concave sur $[2k\pi, (2k+1)\pi]$ pour tout $k \in \mathbb{Z}$.

Ex 26: $e^x \geq x+1$, $\forall x \in \mathbb{R}$; $e^{\frac{a+b}{2}} \leq \frac{e^a + e^b}{2}$.

Ex 27: $\forall x \in \mathbb{R}_+^*$, $\ln(x) \leq x-1$.

Lem 28: Soit $(p,q) \in (\mathbb{R}_+^*)^2$ telsque $\frac{1}{p}+\frac{1}{q}=1$, alors: $\forall u, v \in \mathbb{R}_+^*$, $u^{\frac{1}{p}} v^{\frac{1}{q}} \leq \frac{1}{p}u + \frac{1}{q}v$

Ex 29: $\forall x \in]0, \frac{\pi}{2}[$, $\frac{2}{\pi}x < \sin(x) < x$.

Def 30: Soit $(x_i)_{1 \leq i \leq p}$ suite finie de points de E. On ditque $x \in E$ est combinaison linéaire convexe des x_i s'il existe des réels positifs ou nuls $\lambda_1, \dots, \lambda_p$ tels que: $\sum_{i=1}^{p} \lambda_i = 1$, $x = \sum_{i=1}^{p} \lambda_i x_i$.

Thm 31: Toute partie convexe d'un ev E est stable par combinaison linéaire convexe.

Thm 32: (Jensen) Si f est une fonction convexe définie sur une partie convexe d'un evn E, alors pour toute combinaison linéaire convexe $\sum_{i=1}^{p} \lambda_i x_i$ d'éléments de I, on a: $f\left(\sum_{i=1}^{p} \lambda_i x_i\right) \leq \sum_{i=1}^{p} \lambda_i f(x_i)$.

Thm 33: (Jensen) Si $f: \mathbb{R} \to \mathbb{R}$ fonction convexe alors $\forall u$ fonction continue sur un intervalle $[a,b]$, $a<b$, on a:
$$f\left(\frac{1}{b-a}\int_a^b u(t)\,dt\right) \leq \frac{1}{b-a}\int_a^b f \circ u(t)\,dt.$$

Thm 34: Soit $x = (x_i)_{1 \leq i \leq n}$ suite finie de réels strictement positifs. On note $H(x) = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$, $G(x) = \left(\prod_{i=1}^{n} x_i\right)^{1/n}$, $A(x) = \frac{\sum_{i=1}^{n} x_i}{n}$ les moyennes harmonique, géométrique et arithmétique de x.
On a: $H(x) \leq G(x) \leq A(x)$.

Rq 35: le lemme 28 peut aussi s'écrire: $\forall a \in \mathbb{R}_+^*$, $\forall b \in \mathbb{R}_+^*$, $ab < \frac{1}{p}a^p + \frac{1}{q}b^q$. Il s'agit de l'inégalité de Young et du cas particulier $n=2$ de l'inégalité entre $G(x)$ et $A(x)$ où $x=(a,b)$.

2) Inégalités dans les espaces L^p.

[BP] p159 **Thm 36:** (Inégalité de Hölder) Soient $f, g: X \to \mathbb{K}$ mesurables et $p,q>1$ vérifiant $\frac{1}{p}+\frac{1}{q}=1$.
a) si f et g sont réelles et positives alors $0 \leq \int fg\,d\mu \leq \|f\|_p \|g\|_q \leq +\infty$.
b) si $f \in L^p_{\mathbb{K}}(\mu)$ et $g \in L^q_{\mathbb{K}}(\mu)$, alors $fg \in L^1_{\mathbb{K}}(\mu)$ et $\|fg\|_1 \leq \|f\|_p \|g\|_q$

Prop 37: Si μ est une mesure vérifiant $\mu(X)=1$, alors l'application $k \mapsto \|f\|_k$ est croissante.

Thm 38: (Inégalité de Minkowski) Si $p \in [1, +\infty[$, alors $(L^p_{\mathbb{K}}(\mu), \|.\|_p)$ est un $\mathbb{K}$-ev semi normé. En particulier, $\forall f, g \in L^p_{\mathbb{K}}(\mu)$, $\|f+g\|_p \leq \|f\|_p + \|g\|_p$ [B

3- Inégalités en probabilité

Thm 39: (Inégalité de Jensen) Si ϕ est convexe sur $\mathbb{R}$ et si X est un va réelle telle que X et $\phi(X)$ sont intégrables, alors $\phi(\mathbb{E}[X]) \leq \mathbb{E}[\phi(X)]$

Ex 40: Pour une va X intégrable, $|\mathbb{E}[X]| \leq \mathbb{E}[|X|]$
Pour une va X de carré intégrable, $(\mathbb{E}[X])^2 \leq \mathbb{E}[X^2]$
Pour une va X à valeurs strictement positives, $\mathbb{E}[\frac{1}{X}] \geq \frac{1}{\mathbb{E}[X]}$

Prop 41: Soit X va réelle bornée par 1 ($|X| \leq 1$). On suppose que X est centrée. Alors on a $\mathbb{E}[e^{tX}] \leq e^{t^2/2}$

App 42: (Inégalité de Hoeffding) On considère une suite $(X_n)_{n \in \mathbb{N}^*}$ de va réelles indépendantes, bornées et centrées. On suppose que $|X_n| \leq c_n$ ($c_n > 0$).
On note, $\forall n \in \mathbb{N}^*$, $S_n = \sum_{j=1}^{n} X_j$
Alors $\forall \varepsilon > 0$, $\mathbb{P}(|S_n| \geq \varepsilon) \leq 2\exp\left(\frac{-\varepsilon^2}{2\sum_{j=1}^{n} c_j^2}\right)$ [OUV] p

App 43: Soit $\alpha \in]0,1[$, un intervalle de confiance par excès pour $g(\theta)$ de niveau $1-\alpha$ est une statistique I à valeurs dans les intervalles de $\mathbb{R}$ telle que pour chaque θ: $\mathbb{P}_\theta(g(\theta) \in I) \geq 1-\alpha$ [C p3

Rq 44: Il est possible de déterminer des intervalles de confiance par excès de la moyenne d'une variable aléatoire grâce à l'app 42.

III - Résultats d'optimisation

1) Fonctions convexes et extrémums

Prop 45: Soient C partie convexe et $f: C \to \mathbb{R}$ fonction convexe. Alors les ensembles de niveau de f, càd les parties de E de la forme $\{x \in C, f(x) \leq \alpha\}$ sont convexes. [O p2

Rq 46: Une fonction dont tous les ensembles de niveau sont convexes peut ne pas être convexe.

Lemme 47: (unicité du minimum). Soient C un convexe non vide et $f: C \to \mathbb{R}$ application strictement convexe sur C. Alors il existe au plus un point $x' \in C$ minimisant f sur C. [O p3

Prop 48: Soit A et B deux matrices symétriques réelles définies positives, α et β deux réels positifs tels que $\alpha + \beta = 1$. Alors
$$\det(\alpha A + \beta B) \geq (\det A)^\alpha (\det B)^\beta.$$

App 49: (Ellipsoïde de John Loewner) Soit K un compact d'intérieur non vide de $\mathbb{R}^n$. Alors il existe un unique ellipsoïde centré en 0 de volume minimal contenant K.

Prop 50: Soit f fonction convexe sur un ouvert U de $\mathbb{R}^n$. Si f est différentiable en $a \in U$ et $Df(a)=0$ alors f admet en a un minimum global sur U.

2 - Points fixes

Thm 51: (Méthode de Newton) Soit $f : [c,d] \to \mathbb{R}$ fonction de classe C^2, on suppose $c<d$, $f(c)<0$, $f(d)>0$ et $f'(x)>0$ $\forall x \in [c,d]$. On considère la suite récurrente : $x_{n+1} = F(x_n)$, $n \geq 0$ avec $F(x) = x - \dfrac{f(x)}{f'(x)}$. Alors f admet un zéro unique a et $\forall x \in [c,d]$ $\exists z \in [a,x]$ tel que $F(x) - a = \dfrac{1}{2}\dfrac{f''(z)}{f'(x)}(x-a)^2$, càd $\exists \Pi > 0$ tel que $|F(x)-a| \leq \Pi |x-a|^2$ $\forall x \in [c,d]$.
De plus, $\exists \alpha > 0$ tel que $I = [a-\alpha, a+\alpha]$ soit stable par F et tel que $\forall x_0 \in I$, (x_n) converge quadratiquement vers a dans I.
On peut voir l'interprétation géométrique avec les figures 3 et 4.

Ex 52: On fixe $y>0$ et on prend $f(x) = x^2 - y$. On a donc $F(x) = \frac{1}{2}\left(x + \frac{y}{x}\right)$, $F(x) - a = \dfrac{(x-a)^2}{2x}$ et $x_n = F^n(x_0)$ ce qui nous donne la relation : $\dfrac{x_n - a}{x_n + a} = \left(\dfrac{x_0 - a}{x_0 + a}\right)^{2^n}$

Thm 53: Soit X une partie convexe fermée non vide d'un espace de Banach E et soit $F : U \to E$ application différentiable sur un ouvert U de E contenant X. On suppose qu'il existe $K<1$ telle que $\|DF(x)\| \leq K$, $\forall x \in X$. Si de plus, $F(X) \subset X$, alors il existe un unique point $a \in X$ tel que $F(a)=a$. De plus, ce point peut s'obtenir comme limite de la suite (x_n) des itérés définie par récurrence à partir d'un point quelconque $x_0 \in X$ selon $x_{n+1} = F(x_n)$.

Prop 54: Soit K compact convexe d'un evn et $f : K \to K$ une application continue telle que $\forall (x,y) \in K^2$, $\|f(x) - f(y)\| \leq \|x-y\|$. Alors f admet au moins un point fixe. [GOU] p52

3 - Théorèmes de projection et de séparation

Thm 55: (Projection sur un convexe fermé). [OA] p95
Soient $(H, \langle .,. \rangle)$ espace de Hilbert et C un convexe fermé non vide de H. Alors $\forall x \in H$, il existe un unique élément de C, qui réalise la distance de x à C. Ce point est appelé la projection de x sur C et il est noté $p_C(x)$.
On a ainsi : $\forall y \in C$, $\|x - p_C(x)\| \leq \|x - y\|$
L'élément $p_C(x)$ est de plus caractérisé par :
i) $p_C(x) \in C$
ii) $\mathrm{Re}(\langle x - p_C(x), y - p_C(x) \rangle) \leq 0$, $\forall y \in C$.

Thm 56: (Hahn Banach large) Soit $A \subset E$, $B \subset E$ deux ensembles convexes non vides et disjoints. On suppose A ouvert. Alors il existe un hyperplan fermé qui sépare A et B au sens large. [BRE] p5

Lemme 57: (Jauge d'un convexe). Soit $C \subset E$ un convexe ouvert ($0 \in C$). $\forall x \in E$, on pose : $p(x) = \inf\{\alpha > 0, \alpha^{-1} x \in C\}$ (on dit que p est la jauge de C)
Alors p vérifie : 1) $p(\lambda x) = \lambda p(x)$, $\forall x \in E$, $\forall \lambda > 0$.
2) $p(x+y) \leq p(x) + p(y)$, $\forall x, y \in E$.
3) $\exists \Pi > 0$ tel que $0 \leq p(x) \leq \Pi \|x\|$, $\forall x \in E$.
4) $C = \{x \in E, p(x) < 1\}$.

Lemme 58: Soit $C \subset E$ un convexe ouvert non vide et soit $x_0 \in E$ avec $x_0 \notin C$. Alors $\exists f \in E'$ tel que $f(x) < f(x_0)$, $\forall x \in C$. En particulier, l'hyperplan d'équation $[f = f(x_0)]$ sépare $\{x_0\}$ et C au sens large.

Thm 59: (Hahn-Banach strict) Soient $A \subset E$ et $B \subset E$ deux convexes, non vides, disjoints. On suppose que A est fermé et que B est compact. Alors il existe un hyperplan fermé qui sépare A et B au sens strict.

Biblio : [ROM], Rombaldi
[BP], Briane Pagès
[CV], Cadre Vial
[GOU], Gourdon
[BL], Barbe Ledoux
[OUV], Ouvrard 2
[ROU], Rouvière
[OA], Objectif Agrégation.

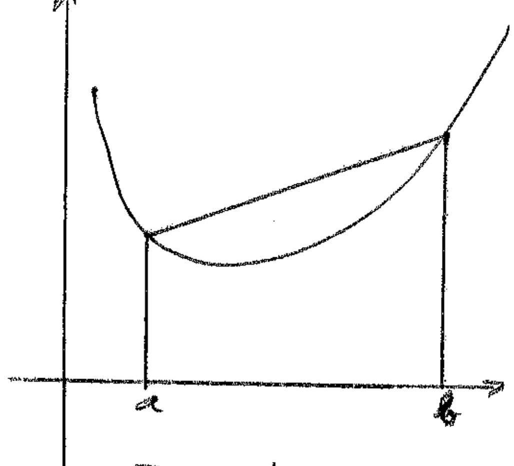


Figure 1

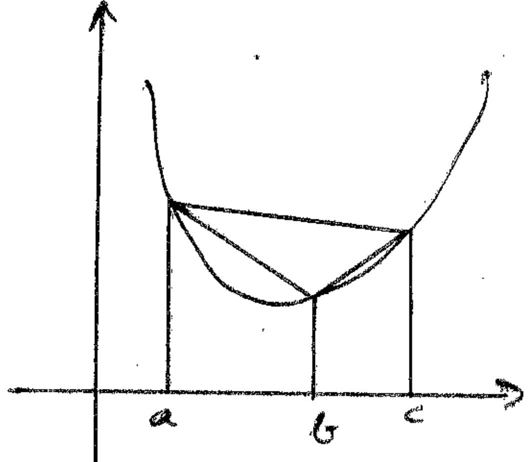


Figure 2

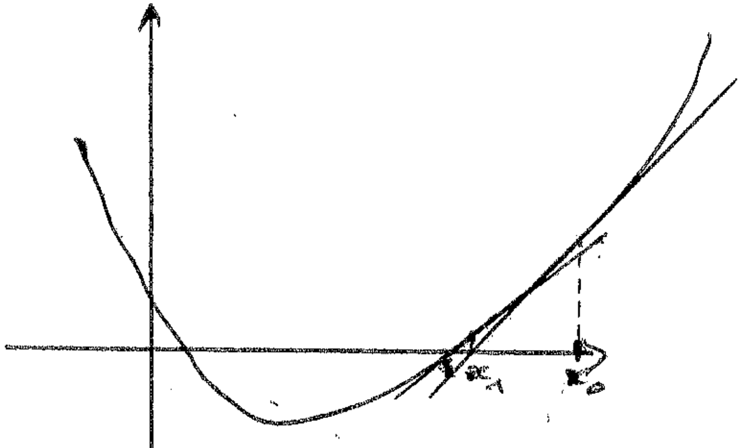


Figure 3

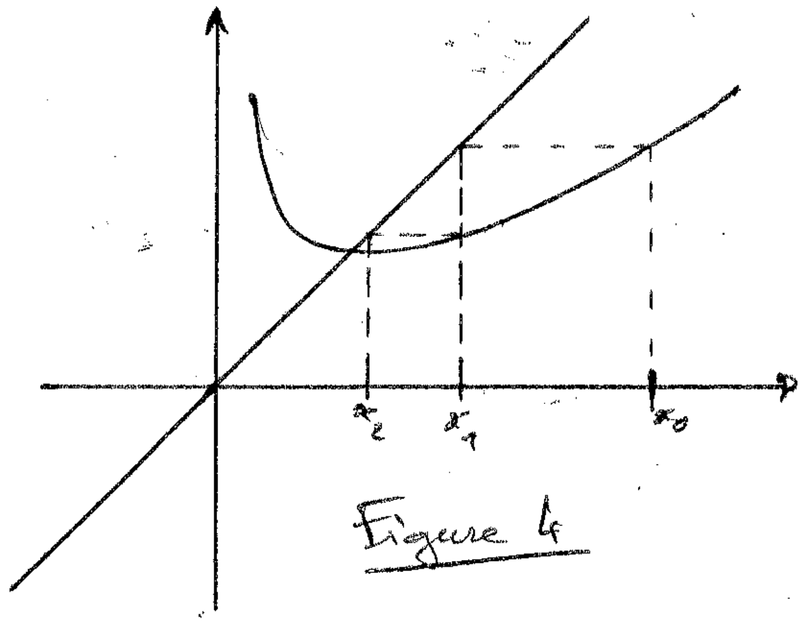


Figure 4

Chapitre 18

Inégalité de Hoeffding

Références : Ouvrard, *Probabilités 2*, 10.11

Théorème.

Soit $(X_n)_n$ une suite de variables aléatoires réelles indépendantes et centrées. On suppose de plus $|X_n| \leqslant c_n$ presque partout, où $c_n > 0$. Alors, en notant $S_n = \sum_{j=1}^{n} X_j$, on a $\forall \varepsilon > 0$,

$$\mathbb{P}(|S_n| > \varepsilon) \leqslant 2 \exp\left(-\frac{\varepsilon^2}{2\sum_{j=1}^{n} c_j^2}\right).$$

Lemme.

Soit X une variable aléatoire réelle centrée et bornée par 1 ps, alors $\forall t \in \mathbb{R}$, $\mathbb{E}[e^{tX}] \leqslant e^{\frac{t^2}{2}}$.

Démonstration. Soit $x \in [-1, 1]$, alors par convexité de l'exponentielle, on a

$$e^{tx} = \exp\left(\frac{1-x}{2} \times (-t) + \frac{1+x}{2} \times t\right) \leqslant \frac{1-x}{2} e^{-t} + \frac{1+x}{2} e^{t}.$$

On rappelle que l'exponentielle est convexe car de dérivée seconde positive, et on peut utiliser l'inégalité car $\frac{1-x}{2} + \frac{1+x}{2} = 1$ et $\frac{1-x}{2}$, $\frac{1+x}{2}$ sont dans $[0,1]$.

X étant bornée presque surement par 1, on peut utiliser l'inégalité précédente : $e^{tX} \leqslant \frac{1-X}{2} e^{-t} + \frac{1+X}{2} e^{t}$. En intégrant, on obtient $\mathbb{E}[e^{tX}] \leqslant \frac{1}{2}(e^{-t} + e^{t}) = \cosh(t)$ car X est centrée.

Puis $\cosh(t) = \sum \frac{t^{2n}}{(2n)!} \leqslant \sum \frac{t^{2n}}{2^n n!} = e^{\frac{t^2}{2}}$. En effet, $\frac{(2n)!}{n!} = (2n)(2n-1)...(n+1) \geqslant 2 \times 2 \times ... \times 2 = 2^n$.
D'où il vient $\mathbb{E}[e^{tX}] \leqslant e^{\frac{t^2}{2}}$. □

On va maintenant prouver l'inégalité de Hoeffding.

Démonstration. Commençons par étudier $\mathbb{P}(S_n > \varepsilon)$.
On a $\forall t > 0$, $\mathbb{P}(S_n > \varepsilon) = \mathbb{P}(e^{tS_n} > e^{t\varepsilon})$.
L'inégalité de Markov donne alors $\mathbb{P}(e^{tS_n} > e^{t\varepsilon}) \leqslant \frac{\mathbb{E}[e^{tS_n}]}{e^{t\varepsilon}}$ car e^{tS_n} est positive.
De plus, $\mathbb{E}[e^{tS_n}] = \prod_{j=1}^{n} \mathbb{E}[e^{tX_j}]$ par indépendance et $\mathbb{E}[e^{tX_j}] = \mathbb{E}[e^{(c_j t)\frac{X_j}{c_j}}] \leqslant e^{\frac{c_j^2 t^2}{2}}$ par le lemme.
On en déduit $\mathbb{E}[e^{tS_n}] \leqslant \prod_{j=1}^{n} e^{\frac{c_j^2 t^2}{2}} = \exp\left(\frac{t^2 \sum_{j=1}^{n} c_j^2}{2}\right)$.

Finalement $\mathbb{P}(S_n > \varepsilon) \leqslant \exp\left(\frac{t^2 \sum_{j=1}^n c_j^2}{2} - t\varepsilon\right)$.

On définit $a := \sum_{j=1}^n c_j^2$, alors on va tenter de minimiser sur $\mathbb{R}^{+*}$ la fonction de t, $\frac{at^2}{2} - t\varepsilon$. Elle atteint son minimum en $t_0 = \frac{\varepsilon}{a}$ et elle vaut alors $-\frac{\varepsilon^2}{2a}$. On a donc $\mathbb{P}(S_n > \varepsilon) \leqslant \exp\left(-\frac{\varepsilon^2}{2a}\right)$.

Pour finir, si on refait le même raisonnement avec la suite $(-X_i)_i$, on trouve $\mathbb{P}(S_n < -\varepsilon) = \mathbb{P}(-S_n > \varepsilon) \leqslant \exp\left(-\frac{\varepsilon^2}{2a}\right)$.

D'où $\mathbb{P}(|S_n| > \varepsilon) = \mathbb{P}(S_n > \varepsilon) + \mathbb{P}(S_n < -\varepsilon) \leqslant 2\exp\left(-\frac{\varepsilon^2}{2a}\right) = 2\exp\left(-\frac{\varepsilon^2}{2\sum_{j=1}^n c_j^2}\right)$. □

Corollaire.

Si il existe $\alpha, \beta > 0$ tels que $\sum_{j=1}^n c_j^2 \leqslant n^{2\alpha-\beta}$, alors $\frac{S_n}{n^\alpha}$ converge presque surement vers 0.

Démonstration. Soit $\varepsilon > 0$, on utilise l'inégalité de Hoeffding : $\mathbb{P}(|S_n| > n^\alpha \varepsilon) \leqslant 2\exp\left(-\frac{n^{2\alpha}\varepsilon^2}{2\sum_{j=1}^n c_j^2}\right)$.

On a alors sous nos hypothèses : $\mathbb{P}\left(\frac{|S_n|}{n^\alpha} > \varepsilon\right) \leqslant 2\exp\left(-\frac{n^\beta \varepsilon^2}{2}\right)$.

La série de terme général $\exp\left(-\frac{n^\beta\varepsilon^2}{2}\right)$ converge car à termes positifs et $\exp\left(-\frac{n^\beta\varepsilon^2}{2}\right) = o\left(\frac{1}{n^2}\right)$.

On en déduit par Borel-Cantelli que $\mathbb{P}\left(\liminf_n \left\{\frac{|S_n|}{n^\alpha} < \varepsilon\right\}\right) = 1$, ce qui est une manière de dire que $\frac{|S_n|}{n^\alpha} \to 0$ presque surement. [1] □

Remarques : • Si on compare les inégalités de Bienaymé-Tchebychev et Hoeffding sur des lois de Bernoulli ($c_n = 1$), on a

$\mathbb{P}(|Sn - np| > \sqrt{n}\varepsilon) \leqslant \frac{p(1-p)}{\varepsilon^2}$ pour Bienaymé-Tchebychev,

et $\mathbb{P}(|Sn - np| > \sqrt{n}\varepsilon) \leqslant 2\exp\left(-\frac{\varepsilon^2}{2}\right)$ pour Hoeffding.

C'est quand même carrément mieux avec Hoeffding !

• Cette inégalité permet d'obtenir des intervalles de confiance... mais on préfère souvent utiliser le TCL.

• Une généralisation de cette inégalité est l'inégalité d'Azuma :

Théorème.

Soit $(X_n)_n$ une martingale issue de 0 dont les accroissements sont contrôlés par une suite déterministe (c_n), c'est à dire $|X_n - X_{n-1}| \leqslant c_n$ presque surement pour tout $n \geqslant 1$. Alors pour tout $\varepsilon > 0$, on a

$$\mathbb{P}(|X_n| \geqslant \varepsilon) \leqslant 2e^{-\frac{\lambda^2}{2\sigma_n^2}}, \text{ où } \sigma_n^2 = \sum_{i=1}^n c_i^2 .$$

1.

$$\begin{aligned} X_n \to X ps &\Leftrightarrow \mathbb{P}\left(\bigcap_{\varepsilon>0}\bigcup_{n\in\mathbb{N}}\bigcap_{m\geqslant n}\{|X_m - X| < \varepsilon\}\right) = 1 \\ &\Leftrightarrow \mathbb{P}\left(\bigcap_{k>0}\bigcup_{n\in\mathbb{N}}\bigcap_{m\geqslant n}\left\{|X_m - X| < \frac{1}{k}\right\}\right) = 1 \\ &\Leftrightarrow \lim_{k\to\infty} \mathbb{P}\left(\liminf_n \left\{|X_n - X| < \frac{1}{k}\right\}\right) = 1 \end{aligned}$$

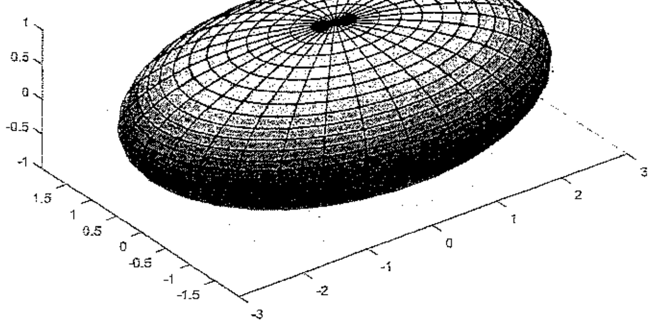
On peut appliquer ce résultat au calcul du nombre de couleurs nécessaires pour colorier un graphe à n sommets avec la condition que deux sommets reliés par une arête ne peuvent être de la même couleur.
Il y a au maximum $\binom{n}{k}$ arêtes possibles. On note Y_k la variable aléatoire de loi de Bernoulli $b(p)$ décidant de si on rajoute l'arête k ou non. On note enfin χ le nombre minimal de couleurs nécessaires selon les arêtes présentes.
On montre alors $\mathbb{P}(|\chi - \mathbb{E}[\chi]| \geqslant \varepsilon\sqrt{n-1}) \leqslant 2e^{-\frac{\lambda^2}{2}}$.

Chapitre 5

Ellipsoïde de John-Loewner

Références : : Francinou, Gianella, Nicolas, *Oraux X-ENS - Algèbre 3*, 3.37

On se place sur $\mathbb{R}^n$ muni de sa structure euclidienne usuelle.
Un ellipsoïde centré en l'origine est une quadrique définie par une équation du type $\sum_{i=1}^{n} \frac{x_i^2}{\alpha_i^2} = 1$, quitte à appliquer la méthode de Gauss. Un ellipsoïde plein centré en l'origine est donc défini intuitivement par l'équation $q(x) \leqslant 1$ où q est une forme quadratique définie positive (id est de signature $(n, 0)$). On notera ε_q l'ellipsoïde plein associé à q et Q (resp. Q^+, Q^{++}) l'ensemble des formes quadratiques (resp. positives, définies positives).



L'ellipsoïde d'équation $\frac{x^2}{3^2} + \frac{y^2}{2^2} + \frac{z^2}{1^2} = 1$

Théorème.

Soit K un compact de $\mathbb{R}^n$ d'intérieur non vide, alors il existe un unique ellipsoïde centré en 0 de volume minimal contenant K.

Démonstration. • Commençons par nous donner une forme quadratique q définie positive et calculons le volume V_q de ε_q.
On utilise le théorème spectral pour dire qu'il existe une base orthonormale pour le produit scalaire euclidien et orthogonale pour q. On la note $\mathcal{B} = (e_1, ..., e_n)$.
q s'écrit dans cette base $q(x) = \sum_{i=1}^{n} a_i x_i^2$ avec $a_i > 0$, donc $\det(q) = a_1 ... a_n$.
Le volume de l'ellipsoïde associé est $V_q = \int_{q(y) \leqslant 1} d\lambda(y)$ avec λ la mesure de Lebesgue (sur la base canonique

donc). On fait alors le changement de base orthonormée $x = Py$ avec P la matrice de passage de la base canonique à la base $\mathcal{B}$. Le jacobien vaut donc 1 car P est orthogonale.
On a donc $V_q = \int_{a_1 x_1^2 + \ldots + a_n x_n^2 \leqslant 1} \mathrm{d}x_1 \ldots \mathrm{d}x_n$.

On fait le changement de variable avec $\phi : \begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R}^n \\ (x_1, \ldots, x_n) & \mapsto & (\frac{x_1}{\sqrt{a_1}}, \ldots, \frac{x_n}{\sqrt{a_n}}) \end{array}$. On a bien un $\mathcal{C}^1$ difféomorphisme et le jacobien est $\frac{1}{\sqrt{\prod a_i}}$.

On note $D(q)$ le déterminant de q dans n'importe quelle base orthonormée, alors on a $V_q = \frac{1}{\sqrt{D(q)}} \int_{x_1^2 + \ldots + x_n^2 \leqslant 1} \mathrm{d}x_1 \ldots \mathrm{d}x_n = \frac{V_0}{\sqrt{D(q)}}$ où V_0 est le volume de la boule unité pour la norme euclidienne.[1]

$\rightarrow$ Le problème se ramène donc à celui-ci : il existe une unique forme quadratique $q \in Q^{++}$ telle que $D(q)$ soit maximal et pour tout $x \in K$, $q(x) \leqslant 1$.
On introduit maintenant la norme $N(q) = \sup_{\|x\| \leqslant 1} |q(x)|$ sur l'espace Q[2] et on définit l'ensemble $\mathcal{A} = \{q \in Q^+, \forall x \in K, q(x) \leqslant 1\} \subset Q$. On va chercher à maximiser D sur ce domaine.

- $\mathcal{A}$ est convexe :

$\lambda q + (1 - \lambda) q'$ est une forme quadratique positive.
De plus, $\lambda q(x) + (1 - \lambda) q'(x) \leqslant \lambda + 1 - \lambda = 1$ pour $x \in K$. Donc $\lambda q + (1 - \lambda) q' \in \mathcal{A}$.

- $\mathcal{A}$ est fermé :

Soit q_n une suite de $\mathcal{A}$ convergeant vers q dans Q. On remarque que $\forall x \in \mathbb{R}^n$, $|q(x) - q_n(x)| \leqslant N(q - q_n) \|x\|^2$, donc $q_n(x) \to q(x)$.
Comme pour tout n, pour tout x, $q_n(x) \geqslant 0$, on a $1 \geqslant q(x) \geqslant 0$, et pour $x \in K$, $q_n(x) \leqslant 1$, donc $q(x) \leqslant 1$. Il vient $q \in \mathcal{A}$.

- $\mathcal{A}$ est borné :

K est d'intérieur non vide[3] donc il existe $a \in K$, $r > 0$ tels que $B(a, r) \subset K$.
Soit $q \in \mathcal{A}$ et $x \in B(0, r)$, alors $q(a + x) \leqslant 1$.
On applique l'inégalité de Minkowski[4] et on a $\sqrt{q(x)} \leqslant \sqrt{q(x + a)} + \sqrt{q(-a)} = \sqrt{q(x + a)} + \sqrt{q(a)} \leqslant 2$ d'où $q(B(0, r)) \leqslant 4$.
On en déduit $\forall x \in B(0, 1)$, $q(x) = q(rx) \frac{1}{r^2} \leqslant \frac{4}{r^2}$ et donc $N(q) \leqslant \frac{4}{r^2}$.

- $\mathcal{A}$ est non vide :

K est borné donc inclus dans une boule $B(0, M)$. On pose $q(x) = \frac{\|x\|^2}{M^2}$. On a bien $q \in \mathcal{A}$ car $\forall x \in K$, $q(x) \leqslant 1$.

- Conclusion sur l'existence :

L'application D est continue (car le déterminant est continu) sur $\mathcal{A}$ compact, donc D atteint son maximum en q_0.
On a $D(x \mapsto \frac{\|x\|^2}{M^2}) > 0$ et c'est une forme quadratique dans $\mathcal{A}$, donc $D(q_0) > 0$, d'où $q_0 \in Q^{++}$.
On a donc existence d'un ellipsoïde de volume minimal contenant K.

- Unicité :

Soit $q \in \mathcal{A}$ tel que $D(q) = D(q_0)$ et $q \neq q_0$. Comme $\mathcal{A}$ est convexe, $\frac{q + q_0}{2} \in \mathcal{A}$, or comme det est strictement log-concave sur les matrices symétriques définies positives, on a $D(\frac{q + q_0}{2}) > \sqrt{D(q)}\sqrt{D(q_0)} \geqslant D(q_0)$, ce qui contredit la maximalité de $D(q_0)$. $\square$

1. On a l'impression que le calcul du volume V_q fait ici dépend du choix de la base $\mathcal{B}$, mais ce n'est pas le cas. Si on prend une autre base orthonormée, la matrice de passage de l'une à l'autre est orthogonale, donc le jacobien du changement de variable correspondant est 1.
2. Les sup ou max de normes sont toujours de bons candidats de normes...
3. C'est ici qu'on l'utilise !!!
4. On a le droit de l'appliquer car on est sur Q^+.

Lemme (Log-concavité du déterminant sur S_n^{++}).

Soient A, $B \in S_n^{++}(\mathbb{R})$, α, $\beta \in [0,1]$ tels que $\alpha+\beta=1$, alors $\det(\alpha A+\beta B) \geqslant (\det A)^\alpha(\det B)^\beta$.
Si $A \neq B$ et α, $\beta \neq 0$, l'inégalité est stricte.

Démonstration. On utilise le théorème de pseudo-réduction simultanée[5] pour écrire $A = {}^tPP$ et $B = {}^tPDP$ avec $P \in \mathrm{GL}_n(\mathbb{R})$ et $D = \mathrm{Diag}(\lambda_1, ..., \lambda_n)$ avec $\lambda_i > 0$.[6]
On a donc $(\det A)^\alpha(\det B)^\beta = \det P^2 (\det D)^\beta$ et $\det(\alpha A+\beta B) = \det P^2(\det(\alpha I_n+\beta D)$.
On veut donc montrer $\det(\alpha I_n+\beta D) \geqslant (\det D)^\beta$, soit $\prod(\alpha+\beta\lambda_i) \geqslant \left(\prod \lambda_i\right)^\beta$, soit $\sum \ln(\alpha+\beta\lambda_i) \geqslant \beta\sum\ln(\lambda_i)$.
Or $\ln(\alpha+\beta\lambda_i) \geqslant \alpha\ln(1)+\beta\ln(\lambda_i) = \beta\ln(\lambda_i)$ par concavité du logarithme. On a ainsi le résultat en sommant ces inégalités et en remontant le raisonnement.
Si $A \neq B$, un des λ_i est différent de 1 et on utilise la stricte concavité du log pour conclure. □

Remarques : → On peut déplacer le problème en n'importe quel point de l'espace. En effet, si on prend $a \in \mathbb{R}^n$ comme nouveau centre du repère de l'espace, on opère une translation de notre compact et il reste d'intérieur non vide. On applique notre théorème et on obtient un unique ellipsoïde centré en 0 de volume minimal contenant K. On opère la translation inverse pour résoudre le problème. En effet, un ellipsoïde translaté reste un ellipsoïde.
→ On peut appliquer ce théorème sur n'importe quel espace vectoriel réel euclidien.
→ Application : FGN 3.38 : les sous-groupes compacts de $\mathrm{GL}(E)$ maximaux pour l'inclusion sont les $\mathrm{O}(q)$ où E est un espace vectoriel réel de dimension finie et $q \in Q^{++}$.
→ Si K est un triangle équilatéral dont un des sommets est l'origine, il faut savoir déterminer l'ellipsoïde correspondant. Celui-ci passe par deux points du triangle et cela le détermine entièrement. Si on appelle a le côté du triangle, on a l'équation $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ avec b tel que $\cos^2(\frac{\pi}{3})+\frac{a^2}{b^2}\sin^2(\frac{\pi}{3})=1$. C'est donc un cercle.
Pour montrer qu'il passe par les deux sommets. Il suffit de dire que via une affinité orthogonale, on peut faire passer l'ellipse par un sommet. Puis si il était minimal alors il serait unique or si on tourne notre ellipsoïde, il passe par l'autre côté et contient toujours le triangle. On en déduit qu'on passe nécessairement par les deux sommets.
→ **Un prolongement du théorème :** on peut poser l'application φ qui à $x \in \mathbb{R}^n$ associe le volume de $\mathcal{E}_{q_x}$ avec q_x l'unique forme quadratique définissant l'ellipsoïde $\mathcal{E}_{q_x}$ centré en x de volume minimal contenant K.
Si on prouve que φ est continue, et que hors d'un certain compact le volume devient grand (l'enveloppe convexe ?), alors on aura montré qu'il existe un ellipsoïde de volume minimal contenant K (et ceci sans fixer son centre).
→ On peut se poser la question de savoir s'il existe un unique ellipsoïde de volume maximal à l'intérieur d'un compact. (Baptiste)
C'est faux ! On prend deux compacts connexes identiques disjoints. Leur union forme un compact. Mais si il existe un unique ellipsoïde dans ce compact, il est dans l'un deux sous compacts connexe. Comme ils sont identiques, on perd l'unicité.
→ Dans le Alessandri, il y a une application de l'ellipsoïde de John Loewner à la recherche des sous-groupes compacts de $\mathrm{GL}_n(\mathbb{R})$.

5.

Théorème.

Si A et B sont dans $S_n(\mathbb{R})$ et si A est définie positive, alors il existe $P \in \mathrm{GL}_n(\mathbb{R})$ telle que ${}^tPAP = I_n$ et ${}^tPBP$ soit diagonale.

6. On rappelle qu'on n'a pas diagonalisé A et B ! Les λ_i ne sont pas les valeurs propres de B. On a juste trouvé une base orthogonale commune pour deux produits scalaires.