Convexité en grande dimension I

Problème 1 : Prékopa–Leindler, Brunn–Minkowski, Borell

Notations et conventions. On note $|A|$ la mesure de Lebesgue d'une partie Lebesgue mesurable $A \subseteq \mathbb{R}^n$ et, pour $A, B \subseteq \mathbb{R}^n$ et $s, t \in \mathbb{R}$, $sA + tB = \{sa + tb\,;\ a \in A,\ b \in B\}$. Tous les ensembles et toutes les fonctions considérés sont boréliens ; les fonctions sont à valeurs dans $[0, +\infty]$, avec la convention $0 \cdot \infty = 0$. On admettra que $A + B$ est toujours Lebesgue mesurable (il n'est pas forcément borélien, mais il l'est dans nos applications, par exemple quand les ensembles sont compacts).

L'objet du devoir est le théorème suivant et quelques-unes de ses conséquences.

Théorème (Prékopa–Leindler). *Soit $\lambda \in]0,1[$ et $f, g, h : \mathbb{R}^n \to [0, +\infty]$ telles que*

$$\forall x, y \in \mathbb{R}^n, \qquad h\big(\lambda x + (1-\lambda)y\big) \ \geq\ f(x)^\lambda g(y)^{1-\lambda}. \tag{H}$$

Alors

$$\int_{\mathbb{R}^n} h \ \geq\ \Big(\int_{\mathbb{R}^n} f\Big)^\lambda \Big(\int_{\mathbb{R}^n} g\Big)^{1-\lambda}. \tag{PL}$$

Remarque. Le théorème reste vrai si f, g, h sont seulement Lebesgue mesurables (c'est le cas de $h = \mathbf{1}_{\lambda A + (1-\lambda)B}$ dans la partie II) : en modifiant des représentants boréliens sur un borélien négligeable, on obtient des fonctions boréliennes $\tilde f \leq f$, $\tilde g \leq g$ et $\tilde h \geq h$ égales presque partout à f, g et h, et le triplet $(\tilde f, \tilde g, \tilde h)$ vérifie encore (H).

La partie I démontre ce théorème. Les parties II, III et IV n'utilisent que son énoncé et peuvent être traitées en l'admettant ; la partie III utilise aussi la question II.3.

Partie I — L'inégalité de Prékopa–Leindler

Dans cette partie, $\lambda \in]0,1[$ est fixé.

1. *Brunn–Minkowski en dimension 1.* Soit $A, B \subseteq \mathbb{R}$ non vides.
   (a) On suppose A et B compacts. En translatant A et B de sorte que $\max A = 0 = \min B$, montrer que $|A + B| \geq |A| + |B|$.
   (b) En déduire que $|A + B| \geq |A| + |B|$ dans le cas général.
2. *Une version renforcée en dimension 1.*
   (a) Montrer que pour toute fonction $u : \mathbb{R}^n \to [0, +\infty]$, $\int_{\mathbb{R}^n} u = \int_0^{+\infty} |\{u \geq s\}|\, ds$.
   (b) Soit $f, g, h : \mathbb{R} \to [0, +\infty]$ telles que $\sup f = \sup g$ et
   $$\forall x, y \in \mathbb{R}, \qquad h\big(\lambda x + (1-\lambda)y\big) \ \geq\ \min\big(f(x), g(y)\big).$$
   Montrer que $\int h \geq \lambda \int f + (1-\lambda) \int g$.
3. *Le théorème en dimension 1.* Soit $f, g, h : \mathbb{R} \to [0, +\infty]$ vérifiant (H).
   (a) Montrer qu'il suffit d'établir (PL) lorsque f et g sont bornées, nulles hors d'un compact et d'intégrales non nulles.
   (b) Dans ce cas, démontrer (PL) en appliquant la question I.2 à $f/\sup f$, $g/\sup g$ et à une fonction h convenablement renormalisée.
4. *Tensorisation.* On suppose le théorème établi dans $\mathbb{R}^n$, et l'on écrit les points de $\mathbb{R}^{n+1} = \mathbb{R} \times \mathbb{R}^n$ sous la forme $(s, \bar x)$. Soit f, g, h des fonctions sur $\mathbb{R}^{n+1}$ vérifiant (H). On pose $F(s) = \int_{\mathbb{R}^n} f(s, \bar x)\, d\bar x$ pour $s \in \mathbb{R}$, et l'on définit de même G et H à partir de g et h.
   (a) Montrer que $H(\lambda s + (1-\lambda)t) \geq F(s)^\lambda G(t)^{1-\lambda}$ pour tous $s, t \in \mathbb{R}$.

(b) Conclure que le théorème est vrai dans $\mathbb{R}^{n+1}$, puis dans $\mathbb{R}^n$ pour tout $n \geq 1$.

5. *Complément : une preuve par transport en dimension* 1. On suppose ici que f et g sont deux fonctions continues et strictement positives sur $\mathbb{R}$, avec $\int f = \int g = 1$. On pose $\Phi(x) = \int_{-\infty}^{x} f$ et $\Psi(x) = \int_{-\infty}^{x} g$.

   (a) Montrer que $T = \Psi^{-1} \circ \Phi$ est une bijection croissante de classe C^1 de $\mathbb{R}$ sur $\mathbb{R}$, et que $g(T(x))\,T'(x) = f(x)$ pour tout x. (On dit que T transporte la mesure $f(x)\,dx$ sur la mesure $g(y)\,dy$.)

   (b) On pose $T_\lambda(x) = \lambda x + (1-\lambda)T(x)$. Justifier l'égalité $\int h = \int h(T_\lambda(x))\,T_\lambda'(x)\,dx$, puis démontrer (PL).

   Cet argument s'étend à la dimension n, avec une application de transport bien choisie (de Knothe ou de Brenier) dont la différentielle a des valeurs propres positives : l'inégalité arithmético-géométrique s'applique alors à ces valeurs propres.

Partie II — Brunn–Minkowski et premières applications

Dans cette partie, λ désigne un réel quelconque de $]0,1[$.

1. Soit $A, B \subseteq \mathbb{R}^n$. Montrer que $|\lambda A + (1-\lambda)B| \geq |A|^\lambda |B|^{1-\lambda}$.

2. Soit $A, B \subseteq \mathbb{R}^n$ non vides. Montrer que

$$|A+B|^{1/n} \;\geq\; |A|^{1/n} + |B|^{1/n}.$$

   On pourra appliquer la question II.1 à $A/|A|^{1/n}$ et $B/|B|^{1/n}$, pour un λ bien choisi. Montrer réciproquement que cette inégalité entraîne celle de la question II.1.

3. *Mesures log-concaves.* Une fonction $\rho : \mathbb{R}^n \to [0,+\infty[$ est *log-concave* si $\rho = e^{-V}$ avec $V : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ convexe. Par exemple : l'indicatrice $\mathbf{1}_K$ d'un convexe K, la densité gaussienne $(2\pi)^{-n/2}e^{-|x|^2/2}$, ou le produit de deux fonctions log-concaves. Soit μ une mesure de densité log-concave. Montrer que pour tous $A, B \subseteq \mathbb{R}^n$,

$$\mu(\lambda A + (1-\lambda)B) \;\geq\; \mu(A)^\lambda \mu(B)^{1-\lambda}.$$

4. *Principe de Brunn.* Soit $K \subseteq \mathbb{R}^n$ ($n \geq 2$) un convexe compact d'intérieur non vide et $\theta \in S^{n-1}$. Pour $t \in \mathbb{R}$, on note $K_t = \{y \in \theta^\perp ;\ y + t\theta \in K\}$ (la section de K par l'hyperplan $t\theta + \theta^\perp$, ramenée dans $\theta^\perp$) et $S(t) = |K_t|$, où $|\cdot|$ désigne ici la mesure de Lebesgue de l'hyperplan $\theta^\perp \simeq \mathbb{R}^{n-1}$.

   (a) Montrer que $\lambda K_s + (1-\lambda)K_t \subseteq K_{\lambda s + (1-\lambda)t}$, et en déduire que $S^{1/(n-1)}$ est concave sur l'intervalle $I = \{t ;\ K_t \neq \emptyset\}$. (L'exposant est optimal : penser à un cône.)

   (b) On suppose K symétrique : $K = -K$. Montrer que S atteint son maximum en 0 : parmi les sections de K par des hyperplans orthogonaux à θ, la section centrale est de volume maximal.

5. *Lemme d'Anderson.* Soit μ une mesure de densité log-concave et paire, $K \subseteq \mathbb{R}^n$ un convexe symétrique, et $\varphi(x) = \mu(K + x)$ pour $x \in \mathbb{R}^n$.

   (a) Montrer que φ est paire et log-concave, au sens où $\varphi(\lambda x + (1-\lambda)y) \geq \varphi(x)^\lambda \varphi(y)^{1-\lambda}$ pour tous $x, y \in \mathbb{R}^n$.

   (b) En déduire que pour tout $x \in \mathbb{R}^n$, la fonction $t \mapsto \varphi(tx)$ est décroissante sur $[0,+\infty[$. En particulier $\mu(K+x) \leq \mu(K)$: parmi les translatés d'un convexe symétrique, c'est le convexe centré qui a la plus grande mesure.

6. *Applications.* Soit $K \subseteq \mathbb{R}^n$ un convexe symétrique.

   (a) Soit X un vecteur aléatoire de densité log-concave et paire, et Y un vecteur aléatoire indépendant de X, de loi quelconque. Montrer que $\mathbb{P}(X+Y \in K) \leq \mathbb{P}(X \in K)$.

(b) Soit Σ_1, Σ_2 des matrices symétriques, avec Σ_1 définie positive et $\Sigma_2 - \Sigma_1$ positive, et soit G_i un vecteur gaussien centré de covariance Σ_i. Montrer que $\mathbb{P}(G_2 \in K) \leq \mathbb{P}(G_1 \in K)$.

(c) En déduire que $\mathbb{E}\|G_1\| \leq \mathbb{E}\|G_2\|$ pour toute norme $\|\cdot\|$ sur $\mathbb{R}^n$. Par exemple, si $g_1, \dots, g_n$ sont des gaussiennes standard indépendantes et $0 < |a_i| \leq |b_i|$ pour tout i, alors $\mathbb{E}\|\sum_i a_i g_i e_i\| \leq \mathbb{E}\|\sum_i b_i g_i e_i\|$.

Partie III — Lemme de Borell

Dans cette partie, μ est une probabilité sur $\mathbb{R}^n$ de densité log-concave, X un vecteur aléatoire de loi μ, et $K \subseteq \mathbb{R}^n$ un convexe symétrique tel que $a = \mu(K) > 0$.

1. Soit $r > 1$. Montrer que
$$\frac{2}{r+1}\,(\mathbb{R}^n \setminus rK) + \frac{r-1}{r+1}\,K \;\subseteq\; \mathbb{R}^n \setminus K.$$

2. En déduire le *lemme de Borell* :
$$\forall r \geq 1, \qquad \mu(\mathbb{R}^n \setminus rK) \;\leq\; a\Big(\frac{1-a}{a}\Big)^{\frac{r+1}{2}}.$$
En particulier, si $\mu(K) \geq 3/4$, alors
$$\mu(\mathbb{R}^n \setminus rK) \leq e^{-r/2}, \qquad \forall r \geq 1.$$

3. Soit $N : \mathbb{R}^n \to [0, +\infty[$ une semi-norme.

(a) En appliquant la question III.2 à $\{N \leq s\}$ pour s assez grand, montrer que $\mathbb{E}N(X) < \infty$.

(b) Soit $m = \mathbb{E}N(X)$, que l'on suppose non nul. En considérant $\{N < 4m\}$, montrer que
$$\forall r \geq 1, \qquad \mathbb{P}\big(N(X) \geq 4mr\big) \leq e^{-r/2}.$$

(c) En déduire qu'il existe une constante numérique C telle que pour tout $p \geq 1$,
$$\big(\mathbb{E}N(X)^p\big)^{1/p} \;\leq\; C\,p\,\mathbb{E}N(X).$$

C'est une inégalité de Hölder inverse, et le point essentiel est que C ne dépend ni de n, ni de μ, ni de N. Avec $N(x) = |\langle x, \theta\rangle|$, elle dit que les formes linéaires d'un vecteur log-concave ont des queues au moins exponentielles.

Partie IV — Une autre application : la concentration gaussienne

On note γ_n la mesure gaussienne standard sur $\mathbb{R}^n$, de densité $(2\pi)^{-n/2}e^{-|x|^2/2}$, et $d(x, A) = \inf_{a \in A} |x - a|$ pour $A \subseteq \mathbb{R}^n$ non vide.

1. Soit $A \subseteq \mathbb{R}^n$ tel que $\gamma_n(A) > 0$. En appliquant le théorème avec $\lambda = 1/2$ à $f = \mathbf{1}_A e^{-|\cdot|^2/2}$, $g = e^{d(\cdot,A)^2/4}e^{-|\cdot|^2/2}$ et $h = e^{-|\cdot|^2/2}$, montrer que
$$\int_{\mathbb{R}^n} e^{d(x,A)^2/4}\,d\gamma_n(x) \;\leq\; \frac{1}{\gamma_n(A)}.$$
On pourra utiliser l'identité du parallélogramme.

2. En déduire que $\gamma_n\big(\{x\,;\; d(x, A) \geq r\}\big) \leq e^{-r^2/4}/\gamma_n(A)$ pour tout $r > 0$.

3. Soit $F : \mathbb{R}^n \to \mathbb{R}$ une fonction 1-lipschitzienne et M une médiane de F sous γ_n, c'est-à-dire telle que $\gamma_n(\{F \leq M\}) \geq 1/2$ et $\gamma_n(\{F \geq M\}) \geq 1/2$. Montrer que
$$\forall r > 0, \qquad \gamma_n\big(\{|F - M| \geq r\}\big) \;\leq\; 4e^{-r^2/4}.$$
Là encore, la borne ne dépend pas de la dimension. Par exemple, pour $F(x) = |x|$, la médiane est de l'ordre de $\sqrt{n}$, mais $|x|$ s'en écarte de plus de r avec probabilité au plus $4e^{-r^2/4}$.