Convexité en grande dimension — Cours I

Problème 2 : mesure uniforme sur la sphère et sur la grassmannienne

Notations. On fixe $n \geq 2$. On note $|x|$ la norme euclidienne de $x \in \mathbb{R}^n$, $(e_1, \dots, e_n)$ la base canonique, $|A|$ la mesure de Lebesgue d'un borélien $A \subseteq \mathbb{R}^n$, $O(n)$ le groupe orthogonal et $\|\cdot\|$ la norme d'opérateur des matrices. Toutes les mesures sont boréliennes. La loi d'une variable aléatoire X à valeurs dans un espace métrique M est la probabilité $A \mapsto \mathbb{P}(X \in A)$ sur M. Un *vecteur gaussien standard* de $\mathbb{R}^n$ est un vecteur aléatoire de densité $(2\pi)^{-n/2} e^{-|x|^2/2}$.

Rappel : lemme de classe monotone. *Soit $\mathcal{A}$ une tribu sur un ensemble Ω, engendrée par une classe $\mathcal{C}$ stable par intersection finie, et μ, ν deux mesures sur $\mathcal{A}$ qui coïncident sur $\mathcal{C}$. Si $\mu(\Omega) = \nu(\Omega) < +\infty$ (par exemple si μ et ν sont des probabilités), ou s'il existe une suite croissante d'éléments de $\mathcal{C}$, de réunion Ω et de μ-mesure finie, alors $\mu = \nu$.*

On construit la probabilité uniforme σ sur S^{n-1} et la probabilité uniforme $\nu_{n,k}$ sur l'ensemble des sous-espaces de dimension k de $\mathbb{R}^n$, sans géométrie différentielle, et l'on établit les propriétés dont le cours se servira : invariance par $O(n)$, unicité, représentation gaussienne, projections, sous-espaces, repères et rotations aléatoires. La partie I est un lemme d'unicité, qui servira pour la sphère et pour la grassmannienne.

Partie I — Un lemme d'unicité

Soit (M, d) un espace métrique compact et Γ un groupe de bijections isométriques de M qui agit *transitivement* : pour tous $x, y \in M$, il existe $\gamma \in \Gamma$ tel que $\gamma x = y$. Une probabilité μ sur M est Γ-*invariante* si $\mu(\gamma A) = \mu(A)$ pour tout borélien A et tout $\gamma \in \Gamma$; autrement dit, si X est de loi μ, alors γX est de loi μ. On note $B(x, r) = \{y \in M : \ d(x, y) < r\}$.

Lemme. *Soit μ et ν deux probabilités Γ-invariantes sur M. Si $\nu(B(x, r)) > 0$ pour tous $x \in M$ et $r > 0$, alors $\mu = \nu$.*

1. Montrer que $\mu(B(x, r))$ ne dépend pas de x ; on le note μ_r. On définit de même $\nu_r > 0$.
2. Soit $f : M \to \mathbb{R}$ continue et $r > 0$. On pose $f_r(x) = \frac{1}{\nu_r} \int_{B(x,r)} f \, d\nu$. Montrer que $\int f_r \, d\mu = \frac{\mu_r}{\nu_r} \int f \, d\nu$, puis que $\mu_r = \nu_r$.
3. Montrer que f_r converge uniformément vers f quand $r \to 0$, et en déduire que $\int f \, d\mu = \int f \, d\nu$. Conclure.
   On peut approcher l'indicatrice d'un fermé F par les fonctions $x \mapsto \max(0, 1 - j\, d(x, F))$ puis utiliser le lemme de classe monotone (ou la régularité des mesures boréliennes).

Partie II — La sphère : mesure de cône

Pour un borélien $A \subseteq S^{n-1}$, on note $C(A) = \{t\theta : \ \theta \in A, \ 0 < t \leq 1\}$ le cône de base A et de sommet 0 (privé de 0), de sorte que $x \in C(A) \Leftrightarrow \frac{x}{|x|} \in A$ pour tout $x \in B_2^n \setminus \{0\}$; on pose alors

$$\sigma(A) := \frac{|C(A)|}{|B_2^n|}.$$

1. (a) Montrer que $C(A)$ est borélien, puis que σ est une probabilité sur S^{n-1}, invariante par $O(n)$.
   (b) Montrer que $\sigma(\mathcal{U}) > 0$ pour tout ouvert non vide $\mathcal{U}$ de S^{n-1}, et que $\sigma(S^{n-1} \cap F) = 0$ pour tout sous-espace $F \neq \mathbb{R}^n$.
2. *Coordonnées polaires.* Pour $x \neq 0$, on pose $\Phi(x) = x/|x|$ et $\Psi(x) = \big(|x|, \Phi(x)\big)$.
   (a) Soit $A \subseteq S^{n-1}$ borélien et $s > 0$. Montrer que $sC(A) = \{t\theta : \ \theta \in A, \ 0 < t \leq s\}$ et que $|sC(A)| = s^n |B_2^n| \, \sigma(A)$. En déduire que pour $0 \leq a < b$,

$$|\{t\theta : \ \theta \in A, \ a < t \leq b\}| = (b^n - a^n) |B_2^n| \, \sigma(A).$$

(b) Montrer que Ψ est un homéomorphisme de $\mathbb{R}^n \setminus \{0\}$ sur $]0,+\infty[\times S^{n-1}$, de réciproque $(r,\theta) \mapsto r\theta$.

(c) On compare deux mesures sur $]0,+\infty[\times S^{n-1}$: l'image m par Ψ de la mesure de Lebesgue sur $\mathbb{R}^n \setminus \{0\}$, définie par $m(B) = |\Psi^{-1}(B)|$ pour B borélien, et $m' = \rho \otimes \sigma$, où ρ est la mesure sur $]0,+\infty[$ de densité $n|B_2^n|\, r^{n-1}$. Montrer que m et m' coïncident sur les *pavés* $]a,b] \times A$, où $0 \le a < b$ et A est un borélien de S^{n-1}, et en déduire que $m = m'$.

(d) En déduire que pour toute fonction borélienne $f : \mathbb{R}^n \to [0,+\infty]$,

$$\int_{\mathbb{R}^n} f(x)\,dx = n|B_2^n| \int_0^{+\infty} r^{n-1}\Big(\int_{S^{n-1}} f(r\theta)\,d\sigma(\theta)\Big)dr.$$

On pourra écrire $f = h \circ \Psi$ sur $\mathbb{R}^n \setminus \{0\}$, où $h(r,\theta) = f(r\theta)$.

3. (a) Soit X de loi uniforme sur B_2^n. Noter que, par définition, $X/|X|$ est de loi σ. Montrer que $|X|$ a pour densité nr^{n-1} sur $[0,1]$, et que ces deux variables sont indépendantes.

(b) Soit g un vecteur gaussien standard de $\mathbb{R}^n$. Montrer que $g/|g|$ est de loi σ et indépendant de $|g|$.

(c) Montrer que $\int_{S^{n-1}} \langle x,\theta\rangle^2\, d\sigma(\theta) = |x|^2/n$ pour tout $x \in \mathbb{R}^n$.

4. Montrer que σ est l'unique probabilité sur S^{n-1} invariante par $O(n)$. En déduire que si Y est un vecteur aléatoire de $\mathbb{R}^n$ dont la loi est invariante par $O(n)$ et si $\mathbb{P}(Y=0)=0$, alors $Y/|Y|$ est de loi σ.

Partie III — La grassmannienne

Soit $1 \le k \le n-1$ et $G_{n,k}$ l'ensemble des sous-espaces vectoriels de dimension k de $\mathbb{R}^n$. Pour $E \in G_{n,k}$, on note P_E la projection orthogonale sur E, et l'on pose $d(E,F) = \|P_E - P_F\|$. Pour $U \in O(n)$, on note $UE = \{Ux :\ x \in E\}$. Enfin $E_0 = \mathrm{Vect}(e_1,\dots,e_k)$.

1. (a) Montrer que $E \mapsto P_E$ est une bijection isométrique de $G_{n,k}$ sur

$$\mathcal{P}_k = \{P \in M_n(\mathbb{R}) :\ P^T = P = P^2,\ \mathrm{tr}\, P = k\},$$

et que $\mathcal{P}_k$ est compact. En déduire que $(G_{n,k}, d)$ est un espace métrique compact.

(b) Montrer que $P_{UE} = UP_EU^T$, et que $O(n)$ agit transitivement sur $G_{n,k}$ par isométries.

(c) Montrer que $E \mapsto E^\perp$ est une isométrie de $G_{n,k}$ sur $G_{n,n-k}$, et que $(UE)^\perp = U(E^\perp)$.

2. *Construction.* Soit $\xi_1,\dots,\xi_k$ des vecteurs aléatoires indépendants de loi σ.

(a) Montrer que $(\xi_1,\dots,\xi_k)$ est presque sûrement libre.

(b) Soit $\Omega_k \subseteq (\mathbb{R}^n)^k$ l'ensemble des familles libres et, pour $x = (x_1,\dots,x_k) \in \Omega_k$, X la matrice $n \times k$ de colonnes $x_1,\dots,x_k$. Montrer que Ω_k est ouvert, que la projection orthogonale sur $\mathrm{Vect}(x_1,\dots,x_k)$ est $X(X^TX)^{-1}X^T$, et en déduire que $x \mapsto \mathrm{Vect}(x_1,\dots,x_k)$ est continue de Ω_k dans $G_{n,k}$.

(c) On note $\nu_{n,k}$ la loi de $\mathrm{Vect}(\xi_1,\dots,\xi_k)$. Montrer que $\nu_{n,k}$ est invariante par $O(n)$.

3. (a) Montrer que $\nu_{n,k}(B(E,r)) > 0$ pour tous $E \in G_{n,k}$ et $r > 0$.

(b) En déduire que $\nu_{n,k}$ est l'unique probabilité sur $G_{n,k}$ invariante par $O(n)$.

(c) Montrer que si E est de loi $\nu_{n,k}$, alors $E^\perp$ est de loi $\nu_{n,n-k}$.

Un *sous-espace aléatoire de dimension k* est une variable aléatoire de loi $\nu_{n,k}$.

4. *Projection aléatoire d'un vecteur fixe.* Soit E de loi $\nu_{n,k}$ et $x \in S^{n-1}$.

(a) Montrer que la loi de $|P_Ex|$ ne dépend pas de $x \in S^{n-1}$.

(b) Soit θ de loi σ, indépendant de E. Montrer que $|P_E x|$, $|P_E \theta|$ et $|P_{E_0}\theta| = (\theta_1^2 + \cdots + \theta_k^2)^{1/2}$ ont même loi. En déduire que $\mathbb{E}|P_E x|^2 = k/n$, et que $|P_E x|^2$ a même loi que $(g_1^2 + \cdots + g_k^2)/(g_1^2 + \cdots + g_n^2)$, où g est un vecteur gaussien standard.

C'est le point de départ du lemme de Johnson–Lindenstrauss : ce quotient est très concentré autour de k/n.

Partie IV — Approche gaussienne : sous-espaces, repères et rotations aléatoires

Soit $1 \le k \le n$ et Z une matrice aléatoire $k \times n$ dont les coefficients sont des variables $\mathcal{N}(0,1)$ indépendantes. Ses lignes $g_1, \ldots, g_k$ sont donc des vecteurs gaussiens standard de $\mathbb{R}^n$ indépendants.

1. (a) Montrer que ZU^T a même loi que Z pour tout $U \in O(n)$.

   (b) On suppose $k \le n-1$. Montrer que Z est presque sûrement de rang k, que $F = \mathrm{Vect}(g_1, \ldots, g_k)$ est de loi $\nu_{n,k}$ et que $\ker Z$ est de loi $\nu_{n,n-k}$.

   Un sous-espace aléatoire de codimension k est donc le noyau d'une matrice gaussienne $k \times n$: c'est sous cette forme qu'il apparaît dans le low M^* et en acquisition comprimée.

2. *Repère aléatoire.* Sur l'événement presque sûr où Z est de rang k, on pose $V = Z^T(ZZ^T)^{-1/2}$, matrice $n \times k$.

   (a) Montrer que $V^T V = I_k$, que $V(\mathbb{R}^k) = \mathrm{Vect}(g_1, \ldots, g_k)$, et que UV a même loi que V pour tout $U \in O(n)$.

   (b) Montrer que Va est de loi σ pour tout $a \in S^{k-1}$. En déduire que pour toute partie finie $N \subseteq S^{k-1}$ et tout borélien $A \subseteq S^{n-1}$,

   $$\mathbb{P}(\exists a \in N : \ Va \notin A) \ \le \ |N|\, \sigma(S^{n-1} \setminus A).$$

   (c) Soit σ_k la probabilité uniforme sur S^{k-1}. Montrer que pour toute fonction borélienne $f : S^{n-1} \to [0, +\infty]$, $\mathbb{E} \int_{S^{k-1}} f(Va)\, d\sigma_k(a) = \int_{S^{n-1}} f \, d\sigma$.

   *L'intégrale $\int f(Va)\, d\sigma_k(a)$ ne dépend que de $V(\mathbb{R}^k)$, car σ_k est invariante par $O(k)$: c'est la moyenne de f sur la sphère unité du sous-espace aléatoire $V(\mathbb{R}^k)$. L'inégalité de (b) est l'*union bound *sur un réseau transporté, dans la preuve du théorème de Dvoretzky.*

3. *Rotation aléatoire.* On prend ici $k = n$, de sorte que $R := V$ est dans $O(n)$, et l'on note η sa loi.

   (a) Montrer que WR et RW ont même loi que R pour tout $W \in O(n)$.

   (b) Montrer que $R\theta_0$ est de loi σ pour tout $\theta_0 \in S^{n-1}$, et que RE est de loi $\nu_{n,k}$ pour tout $E \in G_{n,k}$ $(1 \le k \le n-1)$.

   (c) ($\star$) Montrer que η est l'unique probabilité sur $O(n)$ invariante par multiplication à gauche : c'est la *mesure de Haar* de $O(n)$.

Références. K. Ball, *An elementary introduction to modern convex geometry*, leçon 1 (la boule faite de cônes) ; A. Naor, *Local theory of Banach spaces* (mesure de cône, mesure de Haar) ; P. Mattila, *Geometry of sets and measures in Euclidean spaces*, ch. 3 (mesures uniformément distribuées : le lemme de la partie I en est une version simplifiée) ; E. Meckes, *The random matrix theory of the classical compact groups*, ch. 1 (constructions de la mesure de Haar de $O(n)$, dont l'orthonormalisation d'une matrice gaussienne).