Concentration de la mesure et convexité en grande dimension

Joseph Lehec

Séminaire Bourbaki
1er avril 2022

Isopérimétrie sur la sphère

Pour $A \subset \mathbb{S}^{n-1}$ et $r > 0$ on définit son élargissement :

$$A_r = \{x \in \mathbb{S}^{n-1} \colon d(x, A) \leq r\},$$

où d est la distance géodésique.

Théorème [P. Lévy]

Pour tout $A \subset \mathbb{S}^{n-1}$ si C est une calotte sphérique de même mesure que A alors

$$\sigma_{n-1}(A_r) \geq \sigma_{n-1}(C_r).$$

Autrement dit à mesure fixée les calottes sphériques minimisent la mesure de l'élargissement.

Concentration de la mesure

Corollaire [V. Milman]

Si $\sigma_{n-1}(A) = \frac{1}{2}$ alors

$$\sigma_{n-1}(A_r^c) \leq \exp\left(-cn \cdot r^2\right)$$

où $c > 0$ est une constante universelle.

Démonstration : D'après l'isopérimétrie

$$\sigma_{n-1}(A_r^c) \leq \sigma_{n-1}(C_r^c)$$

où C est une demi-sphère. Il est facile de voir que

$$\sigma_n(C_r^c) \leq \exp(-cn \cdot r^2).$$

Concentration pour les fonctions lipschitziennes

Théorème

Si $f : \mathbb{S}^{n-1} \to \mathbb{R}$ est lipschitizienne avec constante L et m_f est une médiane de f :

$$\sigma_{n-1}(|f - m_f| \geq r) \leq 2 \exp\left(-c \cdot \frac{nr^2}{L^2}\right).$$

Démonstration : à cause du caractère lipschitzien

$$\{f \leq m_f\}_{r/L} \subset \{f \leq m_f + r\}$$

Ensuite on applique le corollaire à $\{f \leq m_f\}$ qui est de mesure $\frac{1}{2}$ par définition.

Remarque : on peut remplacer médiane par moyenne quitte à perdre un peu sur la constante c.

Le théorème de Dvoretzky

Théorème [Dvoretzky '61]

Si X est un Banach de dimension infinie, pour tout $\epsilon > 0$ et tout entier k il existe un sous-espace E de dimension k à distance $1 + \epsilon$ de l'espace euclidien ℓ_2^k.

Ici la distance est la distance de Banach-Mazur.
Concrètement cela signifie qu'il existe une norme euclidienne (i.e. provenant d'un produit scalaire) sur E telle que

$$\|x\|_2 \leq \|x\|_X \leq (1+\epsilon)\|x\|_2, \quad \forall x \in E.$$

Le théorème de Dvoretzky-Milman

Théorème [V. Milman '71]

Si X est un Banach de dimension n la plupart des sous-espaces de dimension $k = c(\epsilon)\log n$ sont à distance $1+\epsilon$ de ℓ_2^k.

Remarques :

- La plupart au sens de la mesure de Haar sur les sous-espaces de dimension k (i.e. unique mesure invariante par rotation)
- $\log n$ est optimal (pour l'espace ℓ_∞^n)
- ℓ_1^n admet des sections euclidiennes de dimension proportionnelle à n.

Schéma de la preuve de Dvoretzky par Milman

- Supposons que la boule euclidienne est incluse dans B_X.
- Alors $x \in \mathbb{S}^{n-1} \mapsto \|x\|$ est 1-Lipschtzienne.
- Grâce à la concentration on sait que si θ un vecteur aléatoire uniforme sur la sphère alors $\|\theta\|$ est proche de
$$M := \int_{\mathbb{S}^{n-1}} \|x\| \, d\sigma_{n-1}$$
avec (très) grande probabilité.
- On peut alors avoir cette propriété pour suffisamment de points $\theta_1, \ldots, \theta_N$ au sein d'un même espace de dimension k pour que celui-ci soit à distance $1 + \epsilon$ de l'espace euclidien. Ça marche si
$$k = c(\epsilon) n \cdot M^2$$
- De plus, quitte à faire une transformation linéaire de X on a
$$M \geq c \cdot \frac{\sqrt{\log n}}{\sqrt{n}}.$$

Concentration de la mesure

« Dvoretzky's theorem, especially as proved by Milman, is a milestone in the local (that is, finite-dimensional) theory of Banach spaces. While I feel sorry for a mathematician who cannot see its intrinsic appeal, this appeal on its own does not explain the enormous influence that the proof has had, well beyond Banach space theory, as a result of planting the idea of measure concentration in the minds of many mathematicians. Huge numbers of papers have now been published exploiting this idea or giving new techniques for showing that it holds. »

T. Gowers

Concentration de la mesure

(X,d) un espace métrique, μ une mesure de probabilité sur X

Définition

Fonction de concentration de μ :

$$\alpha_\mu(r) = \sup\left\{\mu(A_r^c)\colon \mu(A) = \frac{1}{2}\right\}$$

De manière équivalente, c'est la plus petite fonction telle qu'on ait

$$\mu(f \geq m_f + r) \leq \alpha_\mu(r)$$

pour toute fonction 1-lipschitzienne.
Remarque : ici m_f est la médiane de f mais ça ne change pas grand chose de prendre la moyenne.

Isopérimétrie gaussienne

$$\gamma_n(dx) = \frac{1}{(2\pi)^{n/2}} \mathrm{e}^{-|x|^2/2}\, dx$$

Théorème [Borell '75] [Sudakov-Tsirel'son '74]

Soit $A \subset \mathbb{R}^n$ et $r > 0$:

$$\gamma_n(A_r) \geq \gamma_n(H_r)$$

où H est un demi-espace de même mesure que A.

Corollaire

$$\alpha_{\gamma_n}(r) \leq \exp\left(-\frac{1}{2} r^2\right).$$

Limites de cette approche de la concentration

La mesure uniforme sur la sphère et la mesure gaussienne sont essentiellement les deux seuls cas dans lesquels on sache résoudre complètement le problème isopérimétrique. Dans les deux cas on procède par symétrisation. Pour obtenir de la concentration dans d'autres contextes il faut trouver des méthodes qui soient moins dépendantes des symétries.

Mesures uniformément log-concaves

Théorème [Bakry-Émery]

Soit μ une probabilité sur $\mathbb{R}^n$ de la forme

$$\mu(dx) = \mathrm{e}^{-V(x)}\,dx$$

avec $x \mapsto V(x) - \rho|x|^2/2$ convexe pour un certain $\rho > 0$. Alors

$$\alpha_\mu(r) \leq 2\exp\left(-c\rho \cdot r^2\right).$$

Moralité : une mesure plus log-concave que la mesure gaussienne est aussi bien concentrée que la gaussienne (à des constantes universelles près)

Inégalité de Prékopa-Leindler

Théorème

Soit $f, g, h \colon \mathbb{R}^n \to \mathbb{R}_+$ et $\lambda \in [0,1]$ vérifiant

$$f(x)^{1-\lambda} g(y)^{\lambda} \le h\left((1-\lambda)x + \lambda y\right), \quad \forall x, y \in \mathbb{R}^n$$

alors

$$\left(\int_{\mathbb{R}^n} f(x)\,dx\right)^{1-\lambda} \left(\int_{\mathbb{R}^n} g(x)\,dx\right)^{\lambda} \le \int_{\mathbb{R}^n} h(x)\,dx.$$

C'est à la fois une forme inverse de Hölder et une forme fonctionnelle de Brunn-Minkowski :

$$|(1-\lambda)A + \lambda B| \ge |A|^{1-\lambda} |B|^{\lambda}.$$

Concentration pour les mesures uniformément log-concaves

- Hypothèse :

$$V\left(\frac{x+y}{2}\right) \leq \frac{1}{2}(V(x)+V(y)) - \frac{\rho}{8}|x-y|^2.$$

- On fixe $\theta > 0$. Les trois fonctions

$$f(x) = \mathbf{1}_A(x)\exp(-V(x)), \quad g(y) = \exp\left(\theta d(y,A) - V(y)\right)$$

$$h(z) = \exp\left(\frac{2\theta^2}{\rho} - V(z)\right)$$

  vérifient l'hypothèse de Prékopa-Leindler (avec $\lambda = 1/2$).

- Donc

$$\mu(A) \cdot \int_{\mathbb{R}^n} \exp(\theta d(x,A))\,\mu(dx) \leq \exp\left(4\frac{\theta^2}{\rho}\right).$$

- Ce qui implique facilement

$$\alpha_\mu(r) \leq 2\exp\left(-c\rho \cdot r^2\right).$$

Mesures log-concaves

On dit que μ est log-concave sur $\mathbb{R}^n$ si elle est supportée par un convexe K d'intérieur non vide et elle possède une densité de la forme

$$\mathrm{e}^{-V(x)}$$

avec $V : K \to \mathbb{R}$ convexe.
Cette classe contient les mesures uniformes sur les convexes, mais aussi les gaussiennes.

Question

Que peut-on dire de la concentration des mesures log-concaves ?

Quelques propriétés

Quelques propriétés des mesures log-concaves

1. Si μ est log-concave alors

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

   pour tous $A, B \subset \mathbb{R}^n$ et $\lambda \in [0,1]$ (Prékopa-Leindler)
2. Le produit de deux mesures log-concaves est log-concave (évident)
3. Les marginales des mesures log-concaves sont log-concaves (encore Prékopa)
4. Idem pour les convolées de mesures log-concaves
5. Les mesures log-concaves ont des moments exponentiels

Moments

Lemme de Borell

Si X est un vecteur aléatoire log-concave et θ un vecteur de norme 1, les moments de $\langle X, \theta \rangle$ vérifient

$$\|\langle X, \theta \rangle\|_p \leq Cp \cdot \|\langle X, \theta \rangle\|_1. \quad \forall p \geq 1.$$

Supposons pour fixer les idées que X est *isotrope* :

$$\mathbb{E}X = 0 \quad \text{et} \quad \mathrm{cov}(X) = \mathrm{Id}$$

Alors on peut reformuler ce lemme ainsi

$$\mathbb{P}(|\langle X, \theta \rangle| \geq t) \leq C \exp(-ct), \quad \forall t > 0.$$

Conjecture KLS, première version

Conjecture [Kannan, Lovasz, Simonovits '95]

Si μ est log-concave et isotrope sur $\mathbb{R}^n$

$$\alpha_\mu(r) \leq C \exp(-c \cdot r),$$

avec $c, C > 0$ universelles (en particulier ne dépendant pas de n).

Dit autrement : les demi-espaces sont les ensembles extrémaux pour la fonction de concentration, à des constantes universelles près. (Rapprocher ceci de l'isopérimétrie gaussienne).

Inégalité de Poincaré et concentration

$C_P(\mu)$: constante de Poincaré de μ. I.e. meilleure constante dans

$$\mathrm{var}_\mu(f) \leq C \int |\nabla f|^2 \, d\mu$$

Théorème [E. Milman '06]

Pour les mesures log-concaves la constante de Poincaré et la meilleure constante dans la concentration exponentielle sont les mêmes, à un facteur universel près.

Remarque : l'implication Poincaré $\Rightarrow$ Concentration exponentielle est facile (et ne nécessite pas la log-concavité) c'est la réciproque qui constitue le théorème de Milman.

Conjecture KLS, deuxième version

Définition (Constante KLS)

$$\psi_n^2 = \sup\{\ C_P(\mu) \colon \mu \text{ log-concave isotrope sur } \mathbb{R}^n\}$$

Conjecture [KLS '95]

$$\psi_n \leq C.$$

Théorème [KLS '95]

$$\psi_n \leq C\sqrt{n}.$$

Leur motivation était de nature algorithmique (vitesse de convergence de certaines marches aléatoires sur les convexes) mais l'essor de cette conjecture s'explique plus par ses implications pour la géométrie des convexes de grande dimension.

La conjecture de l'hyperplan

Conjecture [Bourgain '80]

Si K est un convexe de $\mathbb{R}^n$ de volume 1 il existe un hyperplan affine H tel que $|K \cap H| \geq c$ où c est une constante universelle.

La question peut sembler anecdotique au premier abord mais elle admet tout un tas de formulations équivalentes et c'est une question qu'on est forcé de rencontrer quand on s'intéresse à la distribution de la masse dans des convexes de grande dimension.

Définition

Constante de l'hyperplan :

$$\frac{1}{L_n} = \inf_K \sup_H |K \cap H|$$

l'inf étant pris sur les convexes de $\mathbb{R}^n$ de volume 1 et le sup sur les hyperplans.

Estimée de Bourgain

Théorème [Bourgain '90]

$$L_n = Cn^{1/4} \log n.$$

Bourgain lui-même n'était pas très satisfait de ce résultat mais c'est resté la meilleure borne pendant des décennies (Klartag a supprimé le $\log n$ en 2006).

Théorème de la Limite Centrale pour les convexes

La philosophie de la théorie asymptotique des convexes est que la grande dimension s'accompagne de phénomènes d'universalité.

Conjecture de la Limite Centrale pour les convexes

Quand n est grand la plupart des marginales d'une mesure log-concave isotrope de $\mathbb{R}^n$ sont à peu près gaussiennes.

Rapprocher ceci de Dvoretzky : la plupart des sections de dimension $\log n$ d'un convexe de dimension n sont à peu près euclidiennes.

La conjecture de la variance

Constante *Thin-Shell*

$$\sigma_n^2 = \frac{1}{n} \cdot \sup \operatorname{var}_\mu(|x|^2).$$

le sup étant pris sur les mesures log-concaves isotropes de $\mathbb{R}^n$.
[Antilla, Ball, Perissinaki] [Bobkov, Koldobsky]

A priori σ_n est de l'ordre de $\sqrt{n}$, la question est de savoir si σ_n est en fait négligeable devant $\sqrt{n}$. Cela signifierait que pour une mesure log-concave isotrope la masse est concentrée dans une mince couronne de rayon $\sqrt{n}$ et d'épaisseur négligeable devant $\sqrt{n}$.

Conjecture de la variance

$$\sigma_n \leq C$$

Lien avec le TCL

Théorème [Antilla, Ball, Perissinaki '03]

1. Si $\sigma_n = o(\sqrt{n})$ alors le TCL est vrai.
2. La conjecture $\sigma_n \leq C$ donnerait une vitesse de convergence optimale (type Berry-Esseen).
3. Cette conjecture est vraie si on se restreint aux boules ℓ_p.

Moralité : le TCL se ramène à trouver une borne non triviale sur σ_n.

Liens entre les différentes conjectures

Il est évident que

$$\sigma_n \leq 2\psi_n$$

(prendre $f(x) = |x|^2$ dans Poincaré).

Théorème [Eldan-Klartag '06]

$$L_n \leq C\sigma_n$$

On a donc la chaine d'inégalités

$$L_n \leq C\sigma_n \leq C'\psi_n.$$

En particulier

$$\text{KLS} \;\Rightarrow\; \text{variance} \;\Rightarrow\; \text{hyperplan}$$

Quelques résultats sur la conjecture de la variance

1. Klartag '08 : $\sigma_n = o(\sqrt{n})$ ($\Rightarrow$ TCL pour les convexes)
2. Fleury, Guédon, Paouris '08 : idem
3. ...
4. Guédon, E. Milman '11 : $\sigma_n = O(n^{1/3})$

Ces résultats utilisent la concentration de la mesure sur la sphère et le théorème de Dvoretzky (sous diverses formes).

Le théorème d'Eldan

Theorem [Eldan '13]

$$\psi_n \leq C\sigma_n \cdot \log n.$$

Autrement dit l'implication triviale entre la conjecture KLS et la conjecture de la variance peut-être renversée, à un logarithme près. C'est un théorème absolument remarquable, mais plus encore que le résultat, c'est la méthode de preuve développée par Eldan, totalement novatrice, qui va complètement révolutionner le problème.

Les résultats « post-Eldan »

Les résultats qui suivent reprennent tous la méthode d'Eldan.

1. [Lee-Vempala '17]

$$\psi_n \leq C n^{1/4}$$

   Noter que ceci améliore aussi le $\sigma_n \leq C n^{1/3}$ de Guédon-Milman et retrouve le $L_n \leq C n^{1/4}$ de Bourgain. Pour cette raison, plusieurs personnes ont pensé à ce moment là que $n^{1/4}$ était vraisemblablement le bon ordre de grandeur.

2. [Chen '21]

$$\psi_n \leq \exp\left(C \sqrt{\log n} \sqrt{\log \log n} \right)$$

   C'est une borne meilleure que toute puissance de n.

3. [Klartag-Lehec '22]

$$\psi_n \leq C (\log n)^{\alpha}$$

   (La preuve donne $\alpha = 5 + o(1)$).