Université Claude Bernard Mathématiques
M2 Phenomena of High Dimension

Final exam

Warm-up: Gaussian concentration around mean or median

Let $G \sim N(0, \mathrm{Id}_n)$. Recall the following result from the course: if $X = f(G)$ is a random variable with median M, with $f : (\mathbf{R}^n, |\cdot|) \to \mathbf{R}$ a 1-Lipschitz function, then for every $t \geqslant 0$,

$$\mathbf{P}(X \geqslant M + t) \leqslant \frac{1}{2} \exp(-t^2/2).$$

Deduce the following inequality for some absolute constants C, c

$$\mathbf{P}(X \geqslant \mathbf{E}[X] + t) \leqslant C \exp(-ct^2).$$

In this exam (this is relevant in Exercices 1 and 3) you are allowed to use the values $C = 1$ and $c = 1/2$.

Exercise 1 Convex hull of a Gaussian cloud

Let $G_1, \dots, G_N$ be i.i.d. $N(0, \mathrm{Id}_n)$ random vectors in $\mathbf{R}^n$. Show the existence of constants C, $c > 0$ such that the following holds: if $N \geqslant C^n$, then with high probability

$$c\sqrt{\log(N)} B_2^n \subset \mathrm{conv}\{G_1, \dots, G_N\} \subset C\sqrt{\log(N)} B_2^n.$$

Hint. *For the first inclusion, show that* $\sup_i \langle G_i, x\rangle \geqslant c\sqrt{\log N}$ *for every* $x \in S^{n-1}$ *by a union bound argument over a* ε*-net.*

Exercise 2 Covering and packing in the discrete cube

In this exercise you can use the following estimate: for integers $0 \leqslant k \leqslant n$ we have

$$\frac{1}{n+1} 2^{nH(k/n)} \leqslant \sum_{j=0}^{k} \binom{n}{j} \leqslant 2^{nH(k/n)}$$

where $H(t) = -t \log_2 t - (1-t) \log_2 (1-t)$ is the binary entropy function.

Let $Q_n = \{0,1\}^n$. For x, $y \in Q_n$, define $d(x,y) = \frac{1}{n} \mathrm{card}\{i \ \ x_i \neq y_i\}$. For $\varepsilon \in (0,1)$, denote by $N(Q_n, \varepsilon)$ and $P(Q_n, \varepsilon)$ the covering and packing numbers for the metric space (Q_n, d).

1. For $0 < \varepsilon < 1/2$, show that
$$1 - H(\varepsilon) \leqslant \limsup_{n\to\infty} n^{-1} \log_2 P(Q_n, \varepsilon) \leqslant 1 - H(\varepsilon/2)$$
2. For $0 < \varepsilon < 1/2$, show by a random covering argument that
$$\lim_{n\to\infty} n^{-1} \log_2 N(Q_n, \varepsilon) = 1 - H(\varepsilon).$$

Exercise 3 Diameter of random sections

1. Let $1 \leqslant m < n$ and M be a $m \times n$ random matrix with i.i.d. $N(0,1)$ entries. Show that $\ker M$ is distributed according to the measure $\mu_{n,n-m}$ on the Grassmann manifold $\mathsf{G}_{n,n-m}$.
2. Let $L \subset S^{n-1}$ a closed subset, and define $w_G(L) := \mathbf{E} \sup_{x \in L} \langle G, x\rangle$ for $G \sim N(0, \mathrm{Id}_n)$. Prove using Gordon's lemma that
$$\mathbf{E} \min_{x \in L} |Mx| \geqslant \kappa_m - w_G(L).$$
If $\kappa_m \geqslant w_G(L)$, deduce that for a random $(n-m)$-dimensional subspace E with distibution $\mu_{n,n-m}$,
$$\mathbf{P}(E \cap L \neq \emptyset) \leqslant \exp(-(\kappa_m - w_G(L))^2/2).$$
3. Deduce the following theorem: if $K \subset \mathbf{R}^n$ is a symmetric convex body and $E \subset G_{n,k}$ is a random k-dimensional subspace with distribution $\mu_{n,k}$ for $k = \lfloor n/2 \rfloor$, then with high probability,
$$K \cap E \subset 2w(K) B_2^n.$$

Exercise 4 A "cheap" form of the reverse Santaló inequality up to $\log n$ factor

You can use freely the formula

$$\int_0^1 t^n (1-t)^m \,\mathrm{d}t = \frac{n!\; m!}{(n+m+1)!}$$

for m, $n \geqslant 0$. If m and n are not integers, it still holds true with the convention $n! = \Gamma(n+1) = n\Gamma(n)$. We denote by vol_n the Lebesgue measure in any n-dimensional Euclidean (sub)space. For K a convex body in $\mathbf{R}^n$, we denote $s(K) = \mathrm{vol}_n(K)\,\mathrm{vol}_n(K^\circ)$.

1. Let $K \subset \mathbf{R}^n$ be a symmetric convex body. Show the formulas, for $m \in \mathbf{N}$

$$\int_K (1-\|x\|_K)^m \,\mathrm{dvol_n}(x) = \frac{m!\; n!}{(m+n)!}\,\mathrm{vol}_n(K),$$

$$\int_K (1-\|x\|_K^2)^{m/2} \,\mathrm{dvol_n}(x) = \frac{(m/2)!(n/2)!}{(m/2+n/2)!}\,\mathrm{vol}_n(K).$$

2. (a) Let $C \subset \mathbf{R}^N$ be a centrally symmetric convex body, $E \subset \mathbf{R}^N$ a k-dimensional linear subspace, and P_E the orthogonal projection onto E (so that $E^\perp = \ker P_E$). Show that for every $x \in P_E C$, the set $C \cap (x + E^\perp)$ contains a translate of $(1-\|x\|_{P_E C})(C \cap E^\perp)$.

   (b) Deduce the estimate

$$\mathrm{vol}_n(C) \geqslant \frac{k!\;(n-k)!}{n!}\,\mathrm{vol}_k(P_E C)\,\mathrm{vol}_{n-k}(C \cap E^\perp).$$

3. Given symmetric convex bodies K_1, K_2 in $\mathbf{R}^n$, define $K_1 +_2 K_2$ and $K_2 \cap_2 K_2$ by the formulas

$$K_1 +_2 K_2 = \{t_1 x_1 + t_2 x_2 \,:\, x_i \in K_i,\ t_1^2 + t_2^2 \leqslant 1\},$$

$$\|\cdot\|_{K_1 \cap_2 K_2} = \sqrt{\|\cdot\|_{K_1}^2 + \|\cdot\|_{K_2}^2}.$$

   (a) Show that $(K_1 +_2 K_2)^\circ = K_1^\circ \cap_2 K_2^\circ$.

   (b) Fix a convex body $K \subset \mathbf{R}^n$, and define $C \subset \mathbf{R}^{2n} \simeq \mathbf{R}^n \oplus \mathbf{R}^n$ by

$$C = \{(sx, ty) \,:\, x \in K,\ y \in K^\circ,\ s^2 + t^2 \leqslant 1\}.$$

   Consider the subspace $E = \{(x,x) \,:\, x \in \mathbf{R}^n\}$. Show that

$$\mathrm{vol}_{2n}(C) = \frac{(n/2)!(n/2)!}{n!} s(K),$$

$$\mathrm{vol}_n(P_E C) = 2^{-n/2}\,\mathrm{vol}_n(K +_2 K^\circ),$$

$$\mathrm{vol}_n(C \cap E^\perp) = 2^{n/2}\,\mathrm{vol}_n(K \cap_2 K^\circ).$$

   (c) Using (2b), conclude that $s(K) \geqslant 2^{-n} s(K \cap_2 K^\circ)$
   You can use the inequality $\binom{2n}{n} \leqslant 2^n \binom{n}{n/2}$.

4. The goal of this question is to prove the following.

   Theorem. If K is a convex body and $\mathcal{E}$, $\mathcal{F}$ are ellispoids such that $\mathcal{E} \subset K \subset \mathcal{F}$ and $\frac{\mathrm{vol}_n(\mathcal{F})}{\mathrm{vol}_n(\mathcal{E})} = r^n$ with $r \in [2, +\infty)$, then $s(K) \geqslant (2\log_2 r)^{-n} s(B_2^n)$.

   (a) Show that we can reduce to the case when $\mathcal{E}^\circ = \mathcal{F}$.

   (b) Assuming $\mathcal{E}^\circ = \mathcal{F}$, show that $\frac{1}{\sqrt{2}}\mathcal{E} \subset K \cap_2 K^\circ \subset \frac{1}{\sqrt{2}} B_2^n$.

   (c) Using (3c), prove the theorem by induction on r.