Random Matrices from the Classical Compact Groups: a Panorama

Part III: Concentration of Measure

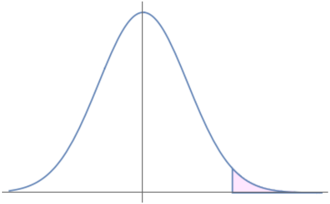
Mark Meckes

Case Western Reserve University

Oxford, 8 February 2021

Gaussian tails

The Gaussian distribution has extremely light tails:



$$\begin{aligned}\mathbb{P}[g \geq t] &= \mathbb{P}[e^{\lambda g} \geq e^{\lambda t}] \\ &\leq e^{-\lambda t}\mathbb{E}e^{\lambda g} \\ &= e^{-\lambda t + \lambda^2/2} \\ &= e^{-t^2/2} \quad \text{for } \lambda = t.\end{aligned}$$

Application: the norm of a Gaussian matrix

For $A \in M_n(\mathbb{R})$, $\|A\|_{op} = \sup_{x \in S^{n-1}} \|Ax\|_2 = \sup_{x,y \in S^{n-1}} \langle Ax, y \rangle$.

For an $n \times n$ Gaussian random matrix G,

$$\langle Gx, y \rangle = \sum_{ij} g_{ij} x_j y_i \sim N(0, 1),$$

so $\|G\|_{op}$ is the supremum of a Gaussian stochastic process.

Application: the norm of a Gaussian matrix

$\mathcal{N} \subseteq S^{n-1}$ is a (1/4)-net if: $\forall x \in S^{n-1}\ \exists y \in \mathcal{N}$ such that $\|x-y\|_2 \leq \frac{1}{4}$.

Lemma

1. $\|A\|_{op} \leq 2 \sup_{x,y \in \mathcal{N}} \langle Ax, y \rangle$.
2. *There is a* $\frac{1}{4}$*-net* $\mathcal{N} \subseteq S^{n-1}$ *with* $\#\mathcal{N} \leq 9^n$.

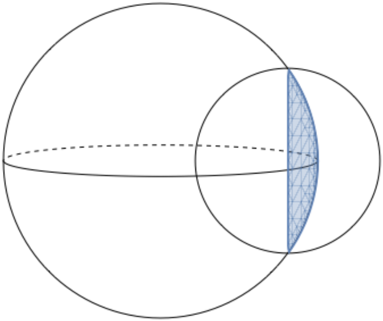
$$\begin{aligned}
\mathbb{P}[\|G\|_{op} \geq t] &\leq \mathbb{P}\left[\sup_{x,y\in\mathcal{N}} \langle Gx, y\rangle \geq t/2\right] \\
&\leq \sum_{x,y\in\mathcal{N}} \mathbb{P}[\langle Gx, y\rangle \geq t/2] \\
&\leq 81^n e^{-t^2/8}
\end{aligned}$$

$\rightsquigarrow \mathbb{E}\,\|G\|_{op} \leq C\sqrt{n}$ and $\|G\|_{op} \leq C'\sqrt{n}$ with high probability.

Spherical tails

Recall the Poincaré limit: if $\theta \sim \mathrm{unif}(S^{n-1})$ then $\sqrt{n}\theta_1 \approx N(0,1)$.

This phenomenon extends to the tails:



$$\mathbb{P}[\sqrt{n}\theta_1 \geq t] \leq e^{-ct^2}$$

$$\mathbb{P}[\theta_1 \geq t] \leq e^{-cnt^2}$$

Almost all the mass on S^{n-1} is within $\approx \frac{1}{\sqrt{n}}$ of an equator.

Isoperimetric inequalities

Classical: For $X \subseteq \mathbb{R}^n$, $\mathrm{vol}_{n-1}(\partial X) \geq \mathrm{vol}_{n-1}(\partial B)$, where B is a ball with $\mathrm{vol}_n(B) = \mathrm{vol}_n(X)$.

More refined version: Write $X_t = \{y \in \mathbb{R}^n | d(X, y) \leq t\}$. Then $\mathrm{vol}_n(X_t) \geq \mathrm{vol}_n(B_t)$.

Round balloons are the easiest to inflate!

Spherical version: For $X \subseteq S^{n-1}$, write $X_t = \{y \in S^{n-1} | d(X, y) \leq t\}$.

Then $\mathrm{vol}_{n-1}(X_t) \geq \mathrm{vol}_{n-1}(B_t)$, where $B \subseteq S^{n-1}$ is a spherical cap with $\mathrm{vol}_{n-1}(B) = \mathrm{vol}_{n-1}(X)$.

Concentration of measure on the sphere

If $\mathrm{vol}_{n-1}(X) \geq \frac{1}{2}\,\mathrm{vol}_{n-1}(S^{n-1})$, then

$$\mathbb{P}[\theta \in X_t] \geq \mathbb{P}[\theta \in B_t] \geq 1 - e^{-cnt^2}.$$

Theorem (Lévy's lemma)

Suppose $F : S^{n-1} \to \mathbb{R}$ is 1-Lipschitz, and M is a median of $F(\theta)$. Then

$$\mathbb{P}[F(\theta) \geq M + t] \leq e^{-cnt^2}.$$

Fluctuations of $F(\theta)$ are of size $O\left(\frac{1}{\sqrt{n}}\right)$.

Gaussian concentration

This fact and the Poincaré limit lead to:

Theorem (Borell, Sudakov–Tsirelson)

Suppose $F : \mathbb{R}^n \to \mathbb{R}$ *is* 1*-Lipschitz, and* M *is a median of* $F(g)$*. Then*

$$\mathbb{P}[F(g) \geq M + t] \leq e^{-ct^2}.$$

Under a concentration result like this, all notions of the average value are basically equivalent:

- $|\mathbb{E}F(g) - M| \leq C$
- $\mathbb{E}F(g) \leq \sqrt{\mathbb{E}F(g)^2} \leq C\mathbb{E}F(g)$ if $F \geq 0$.

Quick application: concentration of norms

$$\mathbb{E}\,\|g\|_2^2 = n \rightsquigarrow \mathbb{P}\left[\left|\|g\|_2 - \sqrt{n}\right| \geq t\right] \leq 2e^{-ct^2}.$$

So for $x \in \mathbb{R}^n$ fixed, $\|Gx\|_2 \approx \sqrt{n}\,\|x\|_2$ with very high probability.

Similarly, $\|G\|_{op} \approx \sqrt{n}$ with $O(1)$ fluctuations.

From spheres to manifolds

Theorem (Bishop–Gromov comparison theorem)

Suppose M is an n-dimensional compact connected Riemannian manifold with Ricci curvature $\geq K > 0$.

Then the volume on M concentrates around balls at least as strongly as on an n-sphere of Ricci curvature K.

In particular, if $F : M \to \mathbb{R}$ is 1-Lipschitz and $X \sim \mathrm{unif}(M)$, then

$$\mathbb{P}[F(X) - \mathbb{E}F(X) \geq t] \leq 2e^{-cKt^2}.$$

Concentration on the classical compact groups (finally)

The Ricci curvature on $\mathbb{SO}(n)$, $\mathbb{SU}(n)$, and $\mathbb{S}\mathrm{p}(n)$ is $\geq cn$.

Theorem (Gromov, Gromov–Milman)

If $\mathbb{G} = \mathbb{SO}(n)$, $\mathbb{SO}^-(n)$, $\mathbb{SU}(n)$, *or* $\mathbb{S}\mathrm{p}(n)$ *and* $F : \mathbb{G} \to \mathbb{R}$ *is* 1*-Lipschitz, then*

$$\mathbb{P}[F(U) - \mathbb{E}F(U) \geq t] \leq 2e^{-cnt^2}.$$

But:

$\mathbb{O}(n)$ isn't connected, and $\mathbb{U}(n)$ doesn't have a positive lower bound on curvature.

Concentration on the classical compact groups

Dealing with $\mathbb{O}(n)$:

- Are you sure don't actually just want to work with $\mathbb{SO}(n)$?
- Condition on $\det U$: equal probability of being in $\mathbb{SO}(n)$ and $\mathbb{SO}^-(n)$.

Dealing with $\mathbb{U}(n)$:

- Let $V \in \mathbb{SU}(n)$ be Haar-distributed and $X \sim \text{unif}\left[0, \frac{\pi\sqrt{2}}{\sqrt{n}}\right]$ be independent.
- Then $U = e^{\sqrt{2}iX/\sqrt{n}}V \in \mathbb{U}(n)$ is Haar-distributed.
- The theorem on the last slide also applies when $\mathbb{G} = \mathbb{U}(n)$.

Quick application: concentration of norms again

Let $P_k \in M_{n,k}(\mathbb{R})$ be the first k columns of a random $U \in \mathbb{O}(n)$ (equivalently, $U \in \mathbb{SO}(n)$).

P_k is essentially the projection onto a random $E \in G_{n,k}$.

Easy computations:

- For fixed $x \in \mathbb{R}^n$, $F(U) = \|P_k x\|_2$ is $\|x\|_2$-Lipschitz.
- $\mathbb{E}\,\|P_k x\|_2^2 = \|x\|_2^2\,\mathbb{E}\,\|P_k e_1\|_2^2 = \frac{k}{n}\,\|x\|_2^2$.

Therefore,

$$\mathbb{P}\left[\left|\|P_k x\|_2 - \sqrt{\frac{k}{n}}\,\|x\|_2\right| \geq \left(\sqrt{\frac{k}{n}}\,\|x\|_2\right) t\right] \leq 2e^{-ckt^2}.$$

Convergence of the spectral measure: a no-work proof

Let $f : \mathbb{C} \to \mathbb{R}$ be 1-Lipschitz, and define $F(U) = \frac{1}{n}\sum_{i=1}^{n} f(\lambda_i(U))$.

By invariance, if $U \in \mathbb{U}(n)$ is random, then

$$\mathbb{E}F(U) = \frac{1}{2\pi}\int_0^{2\pi} f(e^{i\theta})\, d\theta.$$

By the Hoffman–Wielandt inequality, F is $\frac{1}{\sqrt{n}}$-Lipschitz.

Convergence of the spectral measure: a no-work proof

So for each fixed 1-Lipschitz $f : \mathbb{C} \to \mathbb{R}$,

$$\mathbb{P}\left[\left|\frac{1}{n}\sum_{i=1}^{n} f(\lambda_i(U)) - \frac{1}{2\pi}\int_0^{2\pi} f(e^{i\theta})\, d\theta\right| \geq C\frac{\sqrt{\log n}}{n}\right] \leq \frac{1}{n^2}.$$

By the Borel–Cantelli lemma, if $U_n \in \mathbb{U}(n)$ is random for each n, then with probability 1

$$\left|\frac{1}{n}\sum_{i=1}^{n} f(\lambda_i(U_n)) - \frac{1}{2\pi}\int_0^{2\pi} f(e^{i\theta})\, d\theta\right| < C\frac{\sqrt{\log n}}{n}$$

for all sufficiently large n.

Tensorizable concentration

Using logarithmic Sobolev inequalities (Bakry–Émery, Herbst) the Gromov–Milman result generalizes to:

Theorem

Suppose $U_1, \ldots, U_m \in \mathbb{G}$ are Haar-distributed in any of the connected groups and independent and $F : M_n(\mathbb{C})^m \to \mathbb{R}$ is 1-Lipschitz. Then

$$\mathbb{P}[F(U_1, \ldots, U_m) - \mathbb{E}F(U_1, \ldots, U_m) \geq t] \leq e^{-cnt^2}.$$

The upper bound here is independent of m.

Another tool for next time: Dudley's inequality

A subgaussian random process is a collection of random variables $\{X_u | u \in T\}$ indexed by metric space T such that

$$\mathbb{P}[|X_u - X_v| \geq t] \leq 2e^{-t^2/d(u,v)^2}.$$

Theorem (Dudley's entropy bound)

If $\{X_u | u \in T\}$ is a centered subgaussian random process then

$$\mathbb{E} \sup_{u \in T} X_u \leq C \int_0^\infty \sqrt{\log N(T, \varepsilon)}\, d\varepsilon,$$

where $N(T, \varepsilon)$ is the smallest number of ε-balls needed to cover T.

$\log N(T, \varepsilon)$ is called the metric entropy.

Additional references

- Roman Vershynin, *High-Dimensional Probability: An Introduction with Applications in Data Science*, Cambridge, 2018.

- Keith Ball, "An elementary introduction to modern convex geometry", in *Flavors of Geometry*, Cambridge, 1997.

- Mikhael Gromov, "Isoperimetic inequalities in Riemannian manifolds", appendix to *Asymptotic Theory of Finite Dimensional Normed Spaces* by Vitali Milman and Gideon Schechtman, Springer, 1986.