On the concentration of Haar measures

Sourav Chatterjee

Motivational example

- Let M be an $n \times n$ hermitian matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$.
- Let $F_M(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}_{\{\lambda_i \le x\}}$ denote the empirical distribution function of M.
- Let ρ_M denote the uniform distribution on the set of all hermitian matrices having the same spectrum as M.

Example contd.

- Let M and N be two hermitian matrices of order n.
- Let $A \sim \rho_M$ and $B \sim \rho_N$ be independent random matrices. Let $H = A + B$.
- Results of Gromov, Milman, Szarek give concentration bounds for $\frac{1}{n}\sum_i f(\lambda_i^H)$ when f is smooth.
- What about the concentration of $F_H(x)$?

Example contd.

Theorem

Let M and N be two $n \times n$ hermitian matrices. Suppose $A \sim \rho_M$ and $B \sim \rho_N$ are independent, and $H = A + B$. Then, for every $x \in \mathbb{R}$, $\mathrm{Var}(F_H(x)) \leq \kappa n^{-1} \log n$, and

$$\mathbb{P}\{|F_H(x) - \mathbb{E}(F_H(x))| \geq t\} \\ \leq 2 \exp\left(-\frac{nt^2}{2\kappa \log n}\right).$$

Here κ is a numerical constant independent of everything else.

(Recall: ρ_M is the uniform distribution on all $n \times n$ hermitian matrices having the same spectrum as M.)

Ideas

- Gromov-Milman-Szarek give concentration under Hilbert-Schmidt distance.
- $H \mapsto F_H(x)$ is not Lipschitz under that metric.
- Idea # 1: Change the metric to make the map Lipschitz. New metric: $d(M, N) = \mathrm{rank}(M - N)$.
- Fact:
$$\|F_M - F_N\|_\infty \leq n^{-1}\mathrm{rank}(M - N).$$
- Need: A measure concentration technique that can be adapted to arbitrary metrics.

Ideas contd.

- Idea # 2: Concentration inequalities from rates of convergence of Markov chains that take "small steps" w.r.t. a given metric.
- For rank distance, we have the "random reflection walk" on hermitian matrices.
  - At step i, let
    $$U = I - (1 - e^{i\varphi})uu^*,$$
    where
    - u is chosen from the uniform distribution on the unit sphere in $\mathbb{C}^n$, and
    - φ has density proportional to $(\sin(\varphi/2))^{n-1}$ on $[0, 2\pi)$.
  - Let $M_{i+1} = UM_iU^*$.
- Easy: $\mathrm{rank}(M_{i+1} - M_i) \leq 3$.
- Difficult: Convergence to stationary measure ρ_{M_1} in $Cn \log n$ steps (Porod '96).

General theorem about Haar measures

- Let G be a compact topological group.
- Let X be a Haar-distributed random variable on G.
- Let Y be another G-valued r.v. having the following properties:
    1. Y^{-1} has the same distribution as Y.
    2. For any $x \in G$, xYx^{-1} has the same distribution as Y.
- Let $Y_1, Y_2, \ldots,$ be i.i.d. copies of Y. Suppose a and b are two positive constants such that

$$d_{TV}(Y_1 Y_2 \cdots Y_k, X) \leq ae^{-bk}$$

  for every k, where d_{TV} is the total variation distance.

(Think of X as a Haar-distributed unitary matrix and Y as a random reflection. Porod '96 gives $a \sim n$ and $b \sim 1/n$.)

General Theorem (contd.)

Theorem

Suppose $f : G \to \mathbb{R}$ *is bounded and measurable. Let* $\|f\|_\infty = \sup_{x\in G} |f(x)|$ *and*

$$\|f\|_Y := \sup_{x\in G} \left[\mathbb{E}(f(x) - f(Yx))^2\right]^{1/2}.$$

Let A *and* B *be two numbers such that* $\|f\|_\infty \le A$ *and* $\|f\|_Y \le B$*, and* a, b *as above. Let*

$$C = \frac{B^2}{b}\left[\left(\log\frac{4aA}{B}\right)^+ + \frac{b(2-e^{-b})}{1-e^{-b}}\right].$$

Then $\mathrm{Var}(f(X)) \le C/2$*, and for any* $t \ge 0$*,*

$$\mathbb{P}\{|f(X) - \mathbb{E}f(X)| \ge t\} \le 2e^{-t^2/C}.$$

Application

- In our example, X is a Haar-distributed unitary matrix, and $f(X) = F_H(x)$, where $x \in \mathbb{R}$ is fixed, and
$$H = XMX^* + N,$$
M and N being fixed hermitian matrices.
- We have $A = 1$, $B \sim 1/n$, $a \sim n$, and $b \sim 1/n$. Plugging in, we get our result.
- Also applicable to products of random matrices.
- Applies to other groups, e.g. symmetric groups.
- Can go beyond Haar measures and groups. Will not discuss today.

Idea of proof

- Let $X_0 = X$, $X_0' = XY_0$. Construct

$$X_0 \to X_1 \to X_2 \to \cdots \text{ and } X_0' \to X_1' \to X_2' \to \cdots$$

  as $X_{i+1} = X_i Y_{i+1}$ and $X_{i+1}' = X_i' Y_{i+1}'$, where $Y_{i+1}' = Y_0^{-1} Y_{i+1} Y_0$.
- Key step:

$$\mathrm{Cov}(g(X)f(X)) = \frac{1}{2}\sum_{i=0}^{\infty} \mathbb{E}\big[(g(X_0) - g(X_0'))(f(X_i) - f(X_i'))\big]$$

  for any g and f.
- Note that $X_i' = X_i Y_0$ for all i. Thus, $|\mathbb{E}(g(X)f(X))| \leq O(B^2\tau)$, where $B \geq$ the "size" of $g(X) - g(XY)$ and $f(X) - f(XY)$, and $\tau =$ mixing time of the chain.
- For variance bound, take $g = f$. For concentration, take $g = e^{\theta f}$.

Proof of the key step

- By the symmetry between the X and X' chains we have

$$\mathbb{E}[g(X_0)(f(X_i) - f(X_i'))] = \mathbb{E}[g(X_0')(f(X_i') - f(X_i))]$$

- Thus,

$$\mathbb{E}[(g(X_0) - g(X_0')(f(X_i) - f(X_i'))] = 2\mathbb{E}[g(X_0)(f(X_i) - f(X_i'))].$$

- Again, $\mathbb{E}(f(X_i')|X_0) = \mathbb{E}(f(X_{i+1})|X_0)$, since $X_i' = X_i Y_0$ and $X_{i+1} = X_i Y_{i+1}$.
- Thus,

$$\begin{aligned}\frac{1}{2}\sum_{i=0}^{\infty}&\mathbb{E}\big[(g(X_0) - g(X_0'))(f(X_i) - f(X_i'))\big]\\ &= \sum_{i=0}^{\infty}\mathbb{E}[g(X_0)(f(X_i) - f(X_{i+1}))] = \mathrm{Cov}(g(X), f(X)).\end{aligned}$$

Summary

- Given a random variable X on some space, and a function f, we want to get tail bounds for $f(X)$.
  - Step 1: Choose a metric d so that f is Lipschitz w.r.t. d.
  - Step 2: Find a Markov chain whose stationary distribution is the law of X, and which takes small steps w.r.t. d.
  - Step 3: Under some further conditions, compute concentration inequalities for $f(X)$ using the rate of convergence of the Markov chain.
- For Haar measures, take random walks that are "constant on conjugacy classes", like the random reflections walk.
- Applications to random matrices, random permutations, etc.