2.7 sums of independent random matrices

In this section: matrix analogs of Khintzeine's ? and Berstein's inequality.

Thm 2.28 (Non-commutative Khintchine ineq)

Let $A_1, \cdots, A_N$ be self-adjoint $n \times n$ matrices, and $\varepsilon_1, \cdots, \varepsilon_N$ be independent symmetric Bernoulli r.v., then for every $2 \le p < \infty$, we have

$$\left\| \left(\sum_{i=1}^{N} A_i^2 \right)^{\frac{1}{2}} \right\|_{S_p^n} \le \left(E \left\| \sum_{i=1}^{N} \varepsilon_i A_i \right\|_{S_p^n}^p \right)^{\frac{1}{p}}$$

$$\le C \sqrt{p} \left\| \left(\sum_{i=1}^{N} A_i^2 \right)^{\frac{1}{2}} \right\|_{S_p^n}$$

absolute constant

Pf: heavy on combinatorics, skipped.

Remark: For $p = \log n$

$$\|A\| \le \|A\|_{S_n^p} \quad (p \nearrow \text{ norm } \searrow)$$

$$\le \left(\sum_{i=1}^{n} s_i(A)^p\right)^{\frac{1}{p}}$$

$$\le \left(n \sup_i s_i(A)^{\log n}\right)^{\frac{1}{\log n}}$$

$$= n^{\frac{1}{\log n}} \|A\|$$

$$= e\|A\|$$

$$\Rightarrow \frac{1}{e}\left\|\left(\sum_{i=1}^{N} A_i^2\right)^{\frac{1}{2}}\right\| \le E\left\|\sum_{i=1}^{N} \varepsilon_i A_i\right\|$$

$$\le Ce\sqrt{\log n}\left\|\left(\sum_{i=1}^{N} A_i^2\right)^{\frac{1}{2}}\right\|$$

(log factor is estimated more later)

Corollary 2.29 (Rudelon's inequality)

Let $X_1, \cdots, X_N$ be vectors in $\mathbb{R}^n$ and $\varepsilon_1, \cdots, \varepsilon_N$ be independent Rademacher r.v.,

then $$E\left\|\sum_{i=1}^{N} \varepsilon_i X_i \otimes X_i\right\| \le C\sqrt{\log \min\{N, n\}} \cdot \max_{1\le i\le N} \|x_i\|_2 \cdot \left\|\sum_{i=1}^{N} X_i \otimes X_i\right\|^{\frac{1}{2}}$$

where C is an ab-constant.

Pf: If $N<n$, we can replace $\mathbb{R}^n$ with $\operatorname{span}\{X_1, \cdots, X_N\}$, i.e. w.l.o.g $n \le N$.

($\sum \varepsilon_i X_i \otimes X_i$ is rank-1)

?(听下回放, 10:43)

Then $$\left\|\left(\sum_{i=1}^{N} (X_i \otimes X_i)^2\right)^{\frac{1}{2}}\right\| = \left\|\left(\sum_{i=1}^{N} x_i \underbrace{x_i^T x_i}_{=\|x_i\|_2^2} x_i^T\right)^{\frac{1}{2}}\right\|$$

all mat are posi-semidef

$$\leq \max_{1\leq i\leq N} \|X_i\| \left\| \sum_{i=1}^{N} X_i \otimes X_i \right\|^{\frac{1}{2}}$$

Then $E\left\| \sum_{i=1}^{N} \varepsilon_i X_i \otimes X_i \right\| \leq C_i \sqrt{\log n} \left\| \left(\sum_{i=1}^{N} (X_i \otimes X_i)^2 \right)^{\frac{1}{2}} \right\|$

Non-commutative Khintchine

$$\leq C_1 \sqrt{\log \min\{N, n\}} \cdot \max_{1\leq i\leq N} \|X_i\|_2 \left\| \sum_{i=1}^{N} X_i \otimes X_i \right\|^{\frac{1}{2}}$$

Thm 2.30 (Non-commutative Bernstein type Ineq)

Consider a finite sequence of independent centered self-adjoint random matrices. Assume we have for some numbers K and σ that

$$\|X_i\| \leq K \quad a.s., \quad \|\sum_i EX_i^2\| \leq \sigma^2 .$$

Then for every $t \geq 0$, we have

$$P(\|\sum X_i\| \geq t) \leq 2n\exp\left(\frac{-\frac{t^2}{2}}{\sigma^2 + \frac{Kt}{3}}\right) .$$

Proof we need two lemmas:

Lemma 2.31 Consider a finite sequence $\{X_k\}$ of independent, random, self-adjoint random mat of dimension d.

Ass there is a function $g: (0, +\infty) \to [0, +\infty]$

and a sequence $\{A_k\}$ of fixed self-adjoint matrices that satisfy the relation:

$$Ee^{\theta X_k} \preccurlyeq e^{g(\theta)\cdot A_k} \quad \text{for } \theta > 0$$

$A \preccurlyeq B \Leftrightarrow B-A$ is posi-semi-defi

matrix exponential, self-adjoint, then $e^B = I + B + \cdots$ so both sides are matrices! ($B = U\Sigma U^*$, $e^B = U(I+\Sigma+\cdots)U^*$)

Define the scale para $\rho := \lambda_{max}\left|\sum_k A_k\right|$

Then for all $t \in \mathbb{R}$,

$$P\left\{\lambda_{max}\left(\sum_k X_k\right) \geq t\right\} \leq d \inf_{\theta>0} e^{-\theta t + g(\theta)\cdot\rho}$$

(proof: homework)

Lemma 2.32 (bounded Berstein mgf)

moment generating function

Suppose X is a random self-adjoint matrix that satisfies $EX = 0$ and $\lambda_{max} \leq 1$.

Then for all $\theta > 0$,

$$Ee^{\theta X} \preccurlyeq \exp\left((e^\theta - \theta - 1)E(X^2)\right).$$

Pf: Fix $\theta > 0$ and define a smooth func f on $\mathbb{R}$

$$f(x) = \begin{cases} \dfrac{e^{\theta x} - \theta x - 1}{\theta^2 x^2} & x \neq 0 \\ \dfrac{1}{2} & x = 0 \end{cases}$$

$$= \sum_{j=0}^{\infty} \frac{(\theta x)^j}{(j+2)!} \quad \text{for all } x$$

Thm.

$$f'(x) = \sum_{j=1}^{\infty} \frac{\theta j (\theta x)^{j-1}}{(j+2)!}$$

$$= \sum_{e=0}^{\infty} \left(\frac{\theta (2e+1)(\theta x)^{2e}}{(2e+3)!} + \frac{\theta (2e+2)(\theta x)^{2e+1}}{(2e+4)!} \right)$$

$$= \theta \sum_{e=0}^{\infty} \frac{(\theta x)^{2e}}{(2e+3)!} \underbrace{\left(2e+1+\frac{2e+2}{\underbrace{2e+4}_{<1}}\theta x\right)}_{>0 \Rightarrow \text{we require } \theta x \geq -1}$$

$$\geq 0 \longleftarrow \theta x \geq -1$$

Similarly (with opposite pairing)

for $\theta x \leq -1$

$\Rightarrow f$ is strictly increasing.

Therefore $f(x) \leq f(1)$ when $x \leq 1$.

Now $\forall i : \lambda_i(X) \leq \lambda_{max}(X) \leq 1$

so if $X = U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^*$

$$f(X) = \sum_{j=0}^{\infty} \frac{\theta^j}{(j+2)!} \left(U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^* \right)^j$$

$$= U \left(\sum_{j=0}^{\infty} \frac{\theta^j}{(j+2)!} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}^j \right) U^*$$

$$= U \begin{pmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{pmatrix} U^*$$

$$\preceq U \begin{pmatrix} f(1) & & \\ & \ddots & \\ & & f(1) \end{pmatrix} U^*$$

$$= f(1) I$$

$$e^{\theta X} = I + \theta X + X\theta^2 f(X)\cdot X$$

power series!

$$\lesssim I + \theta X + \theta^2 f(1) X^2$$

and

$$Ee^{\theta X} \leq I + f(1)\theta^2 EX^2$$

$$\leq I + f(1)\cdot\left(\theta^2 EX^2 + \sum_{k=2}^{\infty} \frac{(\theta^2 f(1)')^k (E(X^2))^k}{k!}\right)$$

$$= \exp(\theta^2 + f(1) EX^2)$$

$$= \exp((e^\theta - \theta - 1) EX^2)$$

Pf of Thm 2.30

Homogeneity ? w.l.o.g $K=1$

By lemma 2.3.2

$$E e^{\theta X_n} \preccurlyeq \exp(g(\theta) E X_k^2)$$

where $g(\theta) = e^{\theta} - \theta - 1$ for $\theta > 0$

Set $A_k := E(X_k^2)$, then this is exactly the assumption of Lemma 2.3.1

We conclude with the lemma

$$P\{\lambda_{max}(\sum_k X_k) \geq t\} \leq n \inf_{\theta > 0} e^{-\theta t + g(\theta)\sigma^2}$$

where $\sigma^2 = \| \sum_i E X_i^2 \|$ in line with the Thm.

choosing $\theta = \log(1 + \frac{t}{\sigma^2})$ (opt para!)

$$P\{\lambda_{max}(\sum_k X_k) \geq t\} \leq n \exp(-\sigma^2 \ln(\frac{t}{\sigma}))$$

for $h(u) = (1+u)\log(1+u) - u$

compare power series of the derivatives.

one obtains

$$h(u) \geq \frac{\frac{u^2}{2}}{1+\frac{u}{3}} \quad \text{for } u \geq 0$$

which completes the proof of the theorem.