12

Last time:

$A \in \mathbb{R}^{N \times n}$ with rows A_i drawn independently from isotropic distributions.

Let m be such that $\|A_i\|_2 \leq \sqrt{m}$ a.s. $\forall i$.

Then for every $t \geq 0$

$$\sqrt{N} - t\sqrt{m} \leq s_{min}(A) \leq s_{max}(A) \leq \sqrt{N} + t\sqrt{m}$$

with prob $\geq 1 - 2n\exp(-ct^2)$ for $c > 0$ absolute constant.

① necessity of boundness assumption

② dimensional factor in probability bound

i.e.

$$\sqrt{N} - C\sqrt{n\log n} \le s_{min}(A) \le s_{max}(A) \le \sqrt{N} + C\sqrt{n\log n}$$

with prob ≥ 0.

This factor is also necessary.

For that, consider the coordinate distribution

Again we try to find a condition when no trivial lower bound of s_{min} is possible.

In fact, $s_{min} = 0$.

If $\exists\, i$ s.t. e_i is never chosen.

$$Pr(\exists i : A_j \neq \sqrt{n} e_i \ \forall j)$$

$$\geq \sum_i Pr(A_j \neq \sqrt{n} e_i \ \forall j)$$

we need A to be isotropic so use factor n

$$- \sum_{\substack{i_1, i_2 \\ i_1 \neq i_2}} P(A_j \neq \sqrt{n}\, e_{i_1} \text{ and } A_j \neq \sqrt{n}\, e_{i_2}, \forall j)$$

$$= n\left(1-\frac{1}{n}\right)^N - \binom{n}{2}\left(1-\frac{2}{n}\right)^N$$

$$\sim n e^{-\frac{N}{n}} - \frac{1}{2}\left(n e^{-\frac{N}{n}}\right)^2$$

For $N = D n \log n$

Lower bound $\sim n^{1-D} - \frac{1}{2}(n^{1-D})^2$

For $D=1$ lower bound $\sim \frac{1}{2}$

So we need $N > n \log n$

otherwise we get trivial bound

$\rightarrow$ coupon collector's problem

Similar argument shows the optimality of the prob bound in Thm 4.2

Thm 4.3 (Heavy-tailed rows, non-isotropic)

Let A be an $N \times n$ matrix whose rows A_i are independent random vectors in $\mathbb{R}^n$ with the common second moment matrix $\Sigma = E A_i \otimes A_i$

Let m be a number such that $\|A_i\|_2 \le \sqrt{m}$ almost surely for all i

Then for every $t \ge 0$, the following inequality holds with prob at least $1 - 2n\exp(-ct^2)$

$$\left\|\frac{1}{N}A^*A - \Sigma\right\| \le \max\left(\|\Sigma\|^{\frac{1}{2}}\delta, \delta^2\right) \text{ where } \delta = t\sqrt{\frac{m}{N}}.$$

Here $c > 0$ is an absolute constant.

In particular, this inequality yields $\|A\| \le \|\Sigma\|^{\frac{1}{2}}\sqrt{N} + t\sqrt{m}$.

who also holds with prob ...

Pf: $\|\Sigma\| = \|E A_i \otimes A_i\| \overset{\text{Jesen}}{\leq} E\|A_i \otimes A_i\|$

$$= E \sup_{x \in S^n} \langle x, (A_i \otimes A_i) x \rangle$$

what kind of norm? ?

$$= E \sup_{x \in S^n} |\langle A_i, x \rangle|^2$$

$$= E \|A_i\|_2^2$$

$$\leq m$$

we obtain for $X_i = \frac{1}{N}(A_i \otimes A_i - \Sigma)$

$$E X_i^2 = \frac{1}{N^2}\left(E (A_i \otimes A_i)^2 - \Sigma^2\right)$$

$$\Rightarrow \|E X_i^2\| \leq \frac{1}{N^2} \|E \underbrace{\|A_i\|_2^2}_{\leq m} A_i \otimes A_i - \Sigma^2\|$$

$$\leq \frac{2m\|\Sigma\|}{N^2}$$

$\|EX_i\| \le K := \frac{2m}{N}$ as before

Berstein ineq

$$P\left(\left\|\frac{1}{N}A^*A - \Sigma\right\| \ge \varepsilon\right)$$

$$\le 2n \exp\left(-\tilde{c} \min\left(\frac{\varepsilon^2}{\|\Sigma\|}, \varepsilon\right) \frac{N}{2m}\right)$$

$$\varepsilon := \max\left(\|\Sigma\|^{\frac{1}{2}} \delta, \delta^2\right)$$

$$\le 2n \exp\left(-\tilde{c}\, \delta \frac{N}{2m}\right)$$

$$= 2n \exp\left(-\tilde{c} \frac{t'}{2}\right)$$

set $c = \frac{\tilde{c}}{2}$

Then $\frac{1}{N}\|A\|^2 = \|\frac{1}{N}A^*A\|$

$$\leq \|\Sigma\| + \|\frac{1}{N}A^*A - \Sigma\|$$

with prob ...

$$\leq \|\Sigma\| + \|\Sigma\|^{\frac{1}{2}} + \delta + \delta^2$$

$$\leq (\|\Sigma\|^{\frac{1}{2}} + \delta)^{\frac{1}{2}}$$

$\max\{a, b\} \leq a+b$

✓

Almost sure boundness can be relaxed for a bound in expectation.

Theorem 4.4

Heavy-tailed rows, expected singular value

Let A be an $N \times n$ matrix whose rows A_i are independent isotropic vector in $\mathbb{R}^n$.

Let $E \max_{i \leq N} \|A_i\|_2^2 = m < \infty$.

Then

$$E \max_{j \leq n} \|S_j(A) - \sqrt{N}\| \leq C \sqrt{m \log \min(N, n)}$$

For proof, want to apply Rudelson's inequality (coro 2.29)

Need Rademacher random variables.

Lemma 4.5 (Symmetrization)

Let (X_i) be a finite sequence of independent random vectors in some Banach space and (ε_i) be independent symmetric Bernoulli random var

(= Rademacher var)

Then $E\|\sum_i (X_i - EX_i)\| \le 2E\|\sum_i \varepsilon_i X_i\|$.

Pf: Define $\tilde{X}_i = X_i - X_i'$, where (X_i') is an independent copy of the sequence (X_i). Then $\tilde{X}_i$ are independent symmetric random variables, i.e. the sequence $(\tilde{X}_i)$ is dis id with $(-\tilde{X}_i)$ and thus also with $\varepsilon_i \tilde{X}_i$.

Now

$$E\left\|\sum_i (X_i - EX_i)\right\|$$ ← $\|\cdot\|$ is norm in Banach

$$= E\left\|\sum_i (X_i - EX_i')\right\|$$

$$\leq E\left\|\sum_i \tilde{X}_i\right\|$$

↑

Jensen ☆

condition on X_i

convexity, tower property

$$= E\left\|\sum_i \varepsilon_i \tilde{X}_i\right\| + E\left\|\sum_i \varepsilon_i X_i'\right\|$$

$$\leq 2E\left\|\sum_i \varepsilon_i X_i\right\|$$

Lemma 4.6

must be real-valued

Let Z be a non-negativ r.v.

Then $E|Z^2-1| \geq \max(E|Z-1|, E|Z-1|^2)$

Pf: Since

$$|z-1| \leq |z-1| \cdot |z+1|$$

$$\leq |z^2-1|$$

$$|z-1|^2 \leq |z-1| \cdot |z+1|$$

Jensen, take square root.