2.8

Singular value estimates for approximate isometries

Recall that an $N \times n$ matrix B is an isometry if $\|Bx\|_2^2 = \|x\|_2^2$ for all $x \in \mathbb{R}^n$.

$$\|Bx\|_2^2 = \langle x, B^*Bx \rangle$$

$\Leftrightarrow B^*B = I$

In fact, we have good control of the singular values iff B is an approximate isometry.

Lemma 2.33 (Approximate Isometries)

Consider a matrix B satisfying $\|B^*B - I\| \leq \max(\delta, \delta^2)$ for some $\delta > 0$. Then we have

$$1-\delta \leq s_{min}(B) \leq s_{max}(B) \leq 1+\delta \quad (*)$$

Conversely, if B satisfies $(*)$ for some $\delta > 0$, then $\|B^*B - I\| \leq 3\max(\delta, \delta^2)$.

Proof: One has $\|B^*B - I\| \leq \max(\delta, \delta^2)$ iff for all $x \in S^{n-1}$

$\left| \|Bx\|_2^2 - 1 \right| \leq \max(\delta, \delta^2)$ as above.

The geometric interpretation of the singular values yields that $(*) \Leftrightarrow \forall x \in S^{n-1}, \left| \|Bx\|_2 - 1 \right| \leq \delta$.

So we need to show

$$\max(|z-1|, |z-1|^2) \leq |z^2-1| \leq 3\max(|z-1|, |z-1|^2).$$

3. Random matrices with independent entries

Our main goal: require only independence of columns / rows.

For comparison purposes:

results with the stronger assu of inde entries

(without proof)

Thm 3.1 (Bai-Yin's law)

Let $A = A_{N,n}$ be an $N \times n$ random matrix whose entries are independent copies of a r.v. with zero mean, unit var, and finite fourth moment. Suppose that the dimension N and n grows to ∞, while the aspece ratio $\frac{n}{N}$ converges to a constant in $[0,1]$.

Then

$$\left.\begin{aligned} s_{min}(A) &= \sqrt{N} - \sqrt{n} + o(\sqrt{n}) \\ s_{max}(A) &= \sqrt{N} + \sqrt{n} + o(\sqrt{n}) \end{aligned}\right\} \text{almost surely}$$

Gaussian matrices satisfying the bounds in expectation even in the non-asymptotic case:

Thm 3.2 (Gordon's theorem for Gaussian matrices)

Let A be an $N \times n$ matrix whose entries are inde standard normal var. Then

$$\sqrt{N} - \sqrt{n} \leq E\, s_{min}(A) \leq E\, s_{max}(A) \leq \sqrt{N} + \sqrt{n}.$$

There is also an associated deviation equality:

Prop 3.3

Let A be as in Thm 3.2. Then for every $t \geq 0$, with probability at least $1 - \exp(-\frac{t^2}{2})$ one has

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

Minimal Requirement for being able to estimate the maximal singular values are quite weak:

Thm 3.4 Latala's Theorem

Let A be a random matrix whose entries a_{ij} are independent random variables with finite fourth moments. Then

$$E s_{max}(A) \leq C\left(\max_i \left(\sum_j E a_{ij}^2\right)^{\frac{1}{2}} + \max_j \left(\sum_i E a_{ij}^2\right)^{\frac{1}{2}} + \left(\sum_{i,j} E a_{ij}^4\right)^{\frac{1}{4}} \right)$$

If the moment are uni bounded, then this yields

$$s_{max}(A) = O(\sqrt{N} + \sqrt{n})$$

weaker than $\sqrt{N} + O(\sqrt{n})$, which is our goal, but with weak ass. $(N \gg n)$

For small + singular value (hard edge)

Thm 3.5:

Let A be an $N \times n$ random matrix whose entries are i.i.d. subgaussian random variables with zero mean and unit variance. Then for $\varepsilon \geq 0$

$$P(s_{min}(A) \leq \varepsilon(\sqrt{N} - \sqrt{n-1})) \leq (C\varepsilon)^{N-n+1} + c^N.$$

we need this term to include the case $\varepsilon = 0$

$\downarrow$

where $c > 0$ and $c \in (0,1)$ depend ones on the subgaussian norm of the entries.

very strong result, as it gives non-trivial bounds even for quadratic matrices

4. Random matrices with independent rows

Model: Given an arbitrary prob dist in $\mathbb{R}^n$, one takes a sample of N independent points, arrange them to the rows of an $N \times n$ matrix A.

(Recall: tall mat, $N > n$)

Two important cases with _____ different analysis:

- subgaussian row
- heavy-tailed row

4.1 Subgaussian rows

Thm 4.1 Subgaussian rows

Let A be an $N \times n$ matrix whose rows are inde iso subgaussian r. vec in $\mathbb{R}^n$.

Then for every $t \geq 0$, with probability at least $1 - 2\exp(-ct^2)$ one has $\sqrt{N} - C\sqrt{n} - t \leq s_{min}(A) \leq s_{max}(A) \leq \sqrt{N} + C\sqrt{n} + t$

Here $C = C_k$, $c_k = c_k > 0$ depend only on the subgaussian norm $k = \max_i \|A_i\|_{\psi_2}$ of the rows.

This is a generalized version of Prop 3.3 relaxing the ass of inde Gaussian entries at the cost of a constant C.

Proof: Basic version of covering argument.

$\rightarrow$ Need to find $\sup_{x \in S^{n-1}} \|Bx\|_2$

Three steps: 1. Approximation step

$\rightsquigarrow$ discrete S^{n-1} using a net $\mathcal{N}$

2. Concentration Step: control $\|Ax\|_2$ for a fixed $x \in \mathcal{N}$

3. union bound

Proof:

Applying Lemma 2.33 (Approximate isometries) to the matrix $B = \frac{A}{\sqrt{N}}$

we see that we need to show that

$$\left\| \frac{1}{N} A^* A - I \right\| \leq \underbrace{\max(\delta, \delta^2)}_{= \varepsilon}, \text{ where } \delta = C\sqrt{\frac{n}{N}} + \frac{t}{\sqrt{N}}.$$

Recall Lemma 1.7: It suffices to calculate

$\|Cx\|_2$ on a $\frac{1}{4}$-net $\mathcal{N}$

More precisely, $\|\frac{1}{N}A^*A - I\| \le \underset{(1-\frac{1}{2})^{-1}}{2} \max_{x\in\mathcal{N}} |\langle (\frac{1}{N}A^*A - I)x, x\rangle|$

$$= 2\max_{x\in\mathcal{N}} \left|\frac{1}{N}\|Ax\|_2^2 - 1\right|$$

To complete the proof, it suffices to show that with the required proba

$$\max_{x\in\mathcal{N}} \left|\frac{1}{N}\|Ax\|_2^2 - 1\right| \le \frac{\varepsilon}{2}$$

Size of net: $|\mathcal{N}| \le 9^n$ (lemma 2.5)

Fix $x \in S^{n-1}$, then $\|Ax\|_2^2 = \sum_{i=1}^{N} \langle A_i, x\rangle^2$

$$:= \sum_{i=1}^{N} Z_i^2$$

$\{Z_i\}$ are inde subgau r.v.

One has

$EZ_i^2 = \|x\|_2^2 = 1$ as the A_i's are isotropic $\|Z_i\|_{\psi_2} \le K$ by the definition of K and the definition of vector ψ_2-norm.

Now $\|Z_i^2 - 1\|_{\psi_2} \underset{\substack{\uparrow \\ \text{centering}}}{\le} 2\|Z_i^2\|_{\psi_1} \underset{\substack{\uparrow \\ \text{Lemma 2.19}}}{\le} 4\|Z_i\|_{\psi_2}^2 \le 4K^2$

So by Corollary 2.22 (renormalized Berstein)

with $K \ge \|Z_i\|_{\psi_2} \ge \frac{1}{\sqrt{2}} (E|Z_i|^2)^{\frac{1}{2}} = \frac{1}{\sqrt{2}}$

$$P\left\{\left|\frac{1}{N}\|Ax\|_2^2 - 1\right| \ge \frac{\varepsilon}{2}\right\} = P\left(\left|\frac{1}{N}\sum_{i=1}^{N} Z_i^2 - 1\right| \ge \frac{\varepsilon}{2}\right)$$

$$\underset{\substack{\uparrow \\ \text{Coro 2.22}}}{\le} 2\exp\left(-\frac{c_1}{K^4}\min(\varepsilon^2, \varepsilon) N\right)$$

$$= 2\exp\left(-\frac{c_1}{K^4}\delta^2 N\right)$$

$$\underset{\uparrow}{\leq} 2\exp\left(-\frac{c_1}{K^4}(C^2 n + t^2)\right)$$

$$\delta^2 = \cdots \geq C^2\frac{n}{N} + \frac{t^2}{N}$$

$$\min\Big(\underset{\delta > 1 \Uparrow}{\max(\delta, \delta^2)}, \underset{\Uparrow \delta \leq 1}{\max(\delta^2, \delta^4)}\Big)$$

Step 3:

$$P\{\max | \quad | \geq \frac{\varepsilon}{2}) \leq 9^n \cdot 2\exp\left(-\frac{c_1}{K^4}(C^2 n + t^2)\right)$$

$$\underset{\uparrow}{\leq} 2\exp\left(-\frac{c_1 t^2}{K^4}\right)$$

for C large enough $C \geq K^2\sqrt{\frac{\log 9}{c_1}}$