Day 5 Talk 3

Konstantin Tikhomirov

("Invertibility via distance for random matrices with continuous distributions")

"Invertibility of square matrices with continuous distributions"

- A – $n \times n$ matrix (real entries)
- $s_{min}(A) = \inf_{x \in S^{n-1}} \|Ax\|$ $\quad s_{min}(A) = \frac{1}{s_{max}(A^{-1})} \geq \frac{1}{\|A^{-1}\|_{HS}}$.
- $\|A^{-1}\|_{HS}^2 = \sum_{i=1}^{n} \|row_i(A^{-1})\|_2^2$

Hilbert-Schmidt $= \sum_{i=1}^{n} dist(Col_i(A), H_i)^{-2}$

where $H_i = span\{Col_k(A),\ k \neq i\}$

$\ast$ distances $\rightsquigarrow$ bounds for $\|A^{-1}\|_{HS}$ $\rightsquigarrow$ bound for s_{min} $\rightsquigarrow$ invertibility (quantitative)

Q: What is the magnitude of s_{min} of a square random matrix (or the HS-norm of the inverse)?

Applications: • numerical analysis (solving system of equations)

Study singular values of:

- $A - zId$ — $s_{min}(A - zId)$ important for studying Limiting Spectral Distribution (LSD)

- statistics of $dist(X_i, span\{X_j, j \neq i\})$

<u>Some</u> known results on invertibility

(centered, non-Hermitian)

- Edelman: G - $n \times n$, i.i.d entries $N(0,1)$

$$\mathbb{P}\{s_{min}(G) \leq \epsilon n^{-1/2}\} \approx 1 - e^{(-\epsilon - \frac{\epsilon^2}{2})} \quad \forall \epsilon > 0$$

- Tao, Vu, Rudelson, Vershynin, Rudelson, Vershynin: $\mathbb{P}\{s_{min}(A) \leq \epsilon n^{-1/2}\} \lesssim C\epsilon + c^n$ — sub-gaussian

Sankar-Spielman-Teng:

$G + M$ — G - as before, M - fixed $n \times n$ matrix

$$\mathbb{P}\{s_{min}(G+M) \leq \epsilon n^{-1/2}\} \leq C\epsilon \quad \forall \epsilon > 0$$

independent of M!

- Conjectured that worst bound when $M = 0$

Tao–Vu : A $n\times n$ i.i.d. subgaussian, unit variance

$$\|M\| \geq \sqrt{n}$$

$$\mathbb{P}\{s_{min}(A+M) \leq \varepsilon n^{-1/2}\} \leq \left(\frac{\varepsilon \|M\|}{\sqrt{n}}\right)^{\frac{\log \sqrt{n}}{\log \|M\|}} + n^{-\omega(1)}$$

depends on M

For general distributions of A, can't hope for as good estimates as for Gaussian entries. Must be, in general, M-dependent. (Bernoulli matrices)

Question: when M-independent estimates are possible?

Thm [T'17] A – $n\times n$ independent columns.

each $Col_i(A)$ satisfies : $\forall\ F \subset \mathbb{R}^n$, $\dim F = 3$

column

$Proj_F(Col_i(A))$ has density $\rho(x)$ with

$$\rho(x) \leq \frac{C}{\max(1, \|x\|_2^{4000})}$$

Then $\forall$ fixed M we have $\boxed{\mathbb{P}\{\|(A+M)^{-1}\|_{HS} \geq t\sqrt{n}\} \leq \frac{C}{t} \quad \forall t > 0}$

M-independent.

Corollary

If A has independent isotropic log-concave columns, then

$$\mathbb{P}\{s_{min}(A+M) \le \epsilon n^{-1/2}\} \le C\epsilon \quad \forall \epsilon > 0$$

$$\forall M \in \mathbb{R}^{n \times n}$$

Same M-independent ~~est~~ estimates as for Gaussians.

Problems with existing methods.

Sankar-Spielman-Teng

Uses rational invariance

~~A~~ Doesn't seem to work here

Covering arguments

Problem: no control of the matrix norm

Observation: if all 1-dim proj of $Col_i(A)$ have uniformly bounded densities, then

$$\mathbb{P}\{dist(Col_i(A+M), span\{Col_j(A+M)\}_{j \ne i}) \le t\} \le Ct.$$

From this: $\mathbb{P}[\|A^{-1}\|_{HS} \ge tn] \le \frac{C}{t} \quad \forall t > 0$

Worse by $\sqrt{n}$.

Must use dependence/correlations of the dist( ) terms in $\|A^{-1}\|_{HS}$

Consider Correlations of pairs of distances

Given $A+M$, $Col_i := Col_i(A+M)$,

Fix $i \neq j$, denote $H^{i,j} = span\{Col_k, k \neq i,j\}$

$X := Proj_{(H^{ij})^\perp}(Col_i) = X_A + X_M$ ← $Proj\ Col_i(M)$

$\uparrow$ $proj\ Col_i(A)$

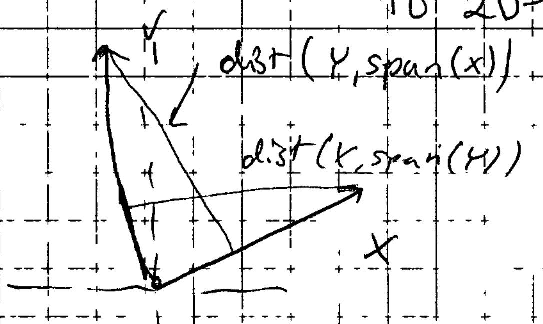
$Y := Proj_{(H^{ij})^\perp}(Col_j) = Y_A + Y_M$ similarly

Then, $dist(Col_i, span\{Col_k, k \neq i\})$

$= dist(X, span\{Y\})$

, $dist(Col_j, span\{Col_k, k \neq j\})$ ⇝ reduces to 2D-setting

$= dist(Y, span\ X)$

Consider $(H^{ij})^\perp$ ⟶ 

To understand ratios of distances

⇝ need to understand $\dfrac{dist(X, span(Y))}{dist(Y, span(X))}$

Next, more precisely.

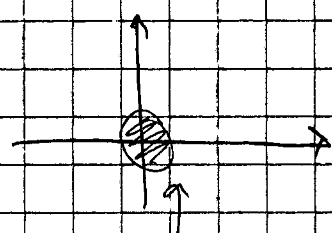
Fix realization of H^{ij} (so, X_M, Y_M are fixed)

$$\mathbb{P}\left\{ \frac{\operatorname{dist}(X, \operatorname{span}(Y))}{\operatorname{dist}(Y, \operatorname{span}(X))} \geq t \, \frac{\|X_M\|+1}{\|Y_M\|+1} \right\} \overset{*}{\leq}$$

If num & denom independent, would, roughly only get $\frac{1}{t}$. But, we get much better estimate.

$$\overset{*}{\leq} \underbrace{\mathbb{P}\left\{ \|X\| \geq t^{\delta}(\|X_M\|+1) \right\}}_{\text{very small for } t \text{ big}} + \underbrace{\mathbb{P}\left\{ \|Y\| \leq t^{\delta-1}(\|Y_M\|+1) \right\}}$$

$\delta \approx \frac{1}{100}$ (say)



probability ~~of~~ $\approx$ area of disc

by density bound proportional to (disregard Y_M)

$\approx (t^{\delta-1})^2$

$\lesssim \frac{1}{t^{2-4\delta}}$ (taking into account Y_M)

Note: Main idea, actual proof involves three dimensions, instead of two.

(~~Major problem~~)

in Actual proof:

* Conditioning on $\{Col_k, \; k \neq i, j, q\} = H_{i,j,q}$

Study projections of Col_i, Col_j, Col_q

onto orthogonal complement of $H_{i,j,q}$.

[(i,j,q) - triples of indices]

Remark: Q1: Interesting to understand if there is a simpler proof?

Q2: If this can be proved with only bounded density?

(Condition for density slightly odd, only log-concave automatically satisfies).

Main theorem sharp up to universal constants.