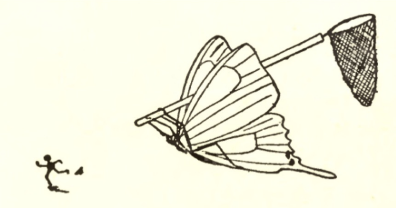
The smallest singular value of inhomogeneous random matrices and efficient net estimates

Galyna V. Livshyts

Georgia Institute of Technology

BIRS IMAG, Granada, Spain
June 2023



Question

a) Pick a large integer, say, $n = 1000000$. Flip a fair die n^2 times. 
Fill an $n \times n$ matrix with the outcomes. How likely is this matrix to be invertible?

$$\begin{bmatrix} 1 & 2 & 6 & 4 & 2 & 5 \\ 3 & 1 & 5 & 3 & 3 & 6 \\ 2 & 3 & 6 & 5 & 1 & 1 \\ 1 & 3 & 2 & 6 & 2 & 5 \\ 4 & 3 & 6 & 1 & 4 & 2 \\ 2 & 3 & 3 & 6 & 4 & 5 \end{bmatrix}$$

Question

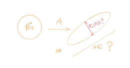
a) Pick a large integer, say, $n = 1000000$. Flip a fair die n^2 times. 
Fill an $n \times n$ matrix with the outcomes. How likely is this matrix to be invertible?

$$\begin{bmatrix} 1 & 2 & 6 & 4 & 2 & 5 \\ 3 & 1 & 5 & 3 & 3 & 6 \\ 2 & 3 & 6 & 5 & 1 & 1 \\ 1 & 3 & 2 & 6 & 2 & 5 \\ 4 & 3 & 6 & 1 & 4 & 2 \\ 2 & 3 & 3 & 6 & 4 & 5 \end{bmatrix}$$

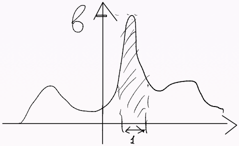
b) And what if now we do not roll the same die every time, but rather use *different* dice to determine different entries?

Notation and Preliminaries

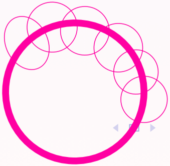
- The Hilbert-Schmidt norm of a matrix $A = (a_{ij})_{ij}$ is $||A||_{HS} = \sqrt{\sum_{i,j} a_{ij}^2}$;
- Singular values of A are the axi of the ellipsoid AB_2^n, denoted $\sigma_1(A) \geq ... \geq \sigma_n(A)$;
- The operator norm $||A|| = \sup_{x \in \mathbb{S}^{n-1}} |Ax| = \sigma_1(A)$;
- The smallest singular value $\sigma_n(A) = \inf_{x \in \mathbb{S}^{n-1}} |Ax|$;



- A random variable ξ is anti-concentrated if $\sup_{z \in \mathbb{R}} P(|\xi - z| < 1) < b \in [0, 1)$.

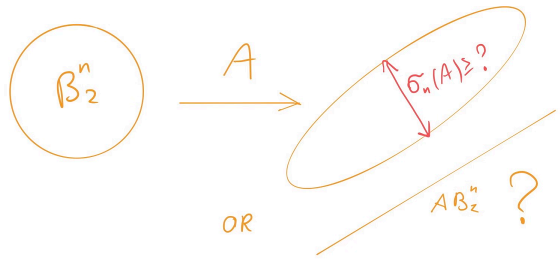


- Recall that for any $\epsilon > 0$ there exists a Euclidean ϵ-net covering the $n-$dimensional ball B_2^n of size $\left(\frac{3}{\epsilon}\right)^n$.



Main question

Question: how likely is a random $n \times n$ matrix A to be invertible?



A harder question: how likely is the smallest singular value $\sigma_n(A) = \inf_{x \in \mathbb{S}^{n-1}} |Ax|$ to be bigger than  ?

History

A is an $n \times n$ Gaussian, with i.i.d. entries $a_{ij} \sim N(0,1)$

$$\sigma_n(A) \approx \frac{1}{\sqrt{n}}.$$

Furthermore, for every $\epsilon \in (0,1)$,

$$P\left(\sigma_n(A) \leq \frac{\epsilon}{\sqrt{n}}\right) \leq \epsilon.$$

(Edelman, Szareck independently in 1990/1991)

History

A is $n \times n$ matrix with i.i.d. Bernoulli ± 1 entries

Conjecture (Erdos) 1950s: $P(\sigma_n(A) = 0) = Cn^2 \cdot 2^{-n}$
(when a pair of columns or rows coincide, and rarely elsewhere)

- Kolmos 60s: $P(\sigma_n(A) = 0) = o(1)$;
- Khan, Kolmos, Szemeredi 1995: $P(\sigma_n(A) = 0) \le 0.99^n$;
- Tao, Vu 2006, 2007: $P(\sigma_n(A) = 0) \le 0.75^n$;
- Bourgain, Vu, Wood, 2010: $P(\sigma_n(A) = 0) \le \sqrt{2}^{-n}$;
- Tikhomirov, 2019: $P(\sigma_n(A) = 0) \le (0.5 + o(1))^n$!

History

A random variable ξ is *sub-Gaussian* if for all $t > 0$,

$$P(|\xi| \geq t) \leq e^{-Kt^2}.$$

A is $n \times n$, has entries a_{ij} i.i.d. sub-Gaussian, $\mathbb{E}a_{ij} = 0$, $\mathbb{E}a_{ij}^2 = 1$

Rudelson, Vershynin 2008:

$$P\left(\sigma_n(A) \leq \frac{\epsilon}{\sqrt{n}}\right) \leq C\epsilon + e^{-cn}.$$

Note: this combines the behavior of Gaussian matrices and the Bernoulli ± 1 matrices.

A is $n \times n$, has entries a_{ij} uniformly anti-concentrated, i.i.d., $\mathbb{E}a_{ij} = 0$, $\mathbb{E}a_{ij}^2 = 1$

Rebrova, Tikhomirov 2016:

$$P\left(\sigma_n(A) \leq \frac{\epsilon}{\sqrt{n}}\right) \leq C\epsilon + e^{-cn}.$$

~~Sub-Gaussian~~

History

A is $n\times n$, has independent UAC entries, $\mathbb{E}||A||_{HS}^2\le Kn^2$, i.i.d. rows

L, 2018+

$$P\left(\sigma_n(A)\le\frac{\epsilon}{\sqrt{n}}\right)\le C\epsilon+e^{-cn}.$$

~~i.i.d. Columns, mean zero variance one~~

Remarks

• In fact, it is enough to assume for any $p>0$,

$$\sum_{i=1}^n\left(\mathbb{E}|Ae_i|^{2p}\right)^{\frac{1}{p}}\le Kn^2;\quad \sum_{i=1}^n\left(\mathbb{E}|A^Te_i|^{2p}\right)^{\frac{1}{p}}\le Kn^2.$$

Note: in principle, all entries may have infinite second moment, but distribution has to depend on n.

• It is much easier to prove this result, and to drop the i.i.d. rows assumption if e^{-cn} is replaced by $\frac{c}{\sqrt{n}}$.

Bai, Cook, Edelman, Gordon, Guedon, Huang, Koltchinckii, Latala, Litvak, Lytova, Meckes, Meckes, Mendelson, Pajor, Paouris, Rebrova, Rudelson, O'Rourke, Szarek, Tao, Tatarko, Tomczak-Jaegermann, Tikhomirov, Van Handel, Vershynin, Vu, Yaskov, Yin, Youssef,...

The smallest singular value: unstructured square case

Theorem (L, Tikhomirov, Vershynin 2019+)

Let A be an $n \times n$ random matrix with

- independent entries a_{ij}
- $\mathbb{E}||A||_{HS}^2 \leq Kn^2$
- a_{ij} are UAC, that is $\sup_{z\in\mathbb{R}} P(|a_{ij} - z| < 1) < b \in (0,1)$

Then for every $\epsilon \in (0,1)$,

$$P\left(\sigma_n(A) < \frac{\epsilon}{\sqrt{n}}\right) \leq C\epsilon + e^{-cn},$$

where C and c are absolute constants which depend (polynomially) only on K and b.

Arbitrary aspect ratio: history

Question: what if A is an $N \times n$ random matrix with $N \geq n$?

Litvak, Pajor, Rudelson, Tomczak-Jaegermann, 2005

$N \geq n + \frac{n}{C \log n}$, strong assumptions: $P(\sigma_n(A) \leq C_1 \sqrt{N}) \leq e^{-C_2 N}$.

Rudelson, Vershynin, 2009

$N \geq n$, a_{ij} i.i.d. sub-Gaussian, $\mathbb{E} a_{ij} = 0$, $\mathbb{E} a_{ij}^2 = 1$. Then for any $\epsilon \in (0,1)$,

$$P\left(\sigma_n(A) \leq \epsilon(\sqrt{N+1} - \sqrt{n})\right) \leq C_1 \epsilon^{N-n+1} + e^{-C_2 N};$$

Tao, Vu, 2010

Replaced sub-Gaussian with $\mathbb{E} a_{ij}^{C_1} \leq 1$, but $N \in [n, n+C_2]$

Vershynin, 2011

Replaced sub-Gaussian with $\mathbb{E} a_{ij}^4 < \infty$ but

$$P\left(\sigma_n(A) \leq \epsilon(\sqrt{N+1} - \sqrt{n})\right) \leq \delta(\epsilon) \to_{\epsilon \to 0} 0.$$

Arbitrary aspect ratio

Theorem (L. 2018+)

Let $N \geq n \geq 1$ be integers. Let A be an $N \times n$ random matrix with

- independent UAC entries a_{ij}
- i.i.d. rows
- $\mathbb{E}a_{ij} = 0$
- $\mathbb{E}a_{ij}^2 = 1$

Then for every $\epsilon > 0$,

$$P\left(\sigma_n(A) < \epsilon(\sqrt{N+1} - \sqrt{n})\right) \leq (C\epsilon \log 1/\epsilon)^{N-n+1} + e^{-cN},$$

where C and c are absolute constants which depend (polynomially) only on the concentration function bounds.

Remark: a more general result in fact follows...

Very tall case

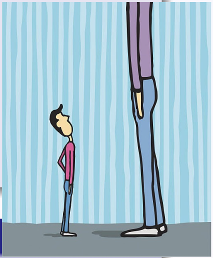
Proposition 1 (L. 2018+) tall case with dependent columns

Suppose A is an $N \times n$ random matrix with independent rows, $\mathbb{E}||A||_{HS}^2 \le KNn$, $N \ge C_0 n$, and assume for every $x \in \mathbb{S}^{n-1}$,

$$\sup_{y \in \mathbb{R}} P(|\langle A^T e_i, x\rangle - y| \le 1) \le b \in (0,1).$$

Then

$$\mathbb{E}\sigma_n(A) \ge c\sqrt{N}.$$



Proposition 2 (L. 2018+) tall case with low moments

Fix $p > 0$. Suppose $N \ge C_0' n$, A is an $N \times n$ random matrix with independent UAC entries. Suppose

$$\sum_{i=1}^{n} \left(\mathbb{E}|Ae_i|^{2p}\right)^{\frac{1}{p}} \le KnNe^{\frac{c_0 N}{n}}.$$

Then

$$P(\sigma_n \le C_1\sqrt{N}) \le e^{-C_2 \min(p,1) N}.$$

A naive attempt

Goal: $P(\sigma_n(A) \leq 2\heartsuit) \leq \diamondsuit$.

Discretize $\mathbb{S}^{n-1}$:

Suppose we find a small finite set $\mathcal{N} \subset \mathbb{R}^n$ with

- $\#\mathcal{N} \leq \spadesuit$;
- $\forall x \in \mathbb{S}^{n-1} \, \exists y \in \mathcal{N} : |A(x-y)| \leq \heartsuit$ with probability $\geq 1 - \clubsuit$.

Then we write:

$$P(\sigma_n(A) \leq \heartsuit) = P\left(\inf_{x \in \mathbb{S}^{n-1}} |Ax| \leq \heartsuit\right) \leq$$

$$P\left(\inf_{y \in \mathcal{N}} |Ay| \leq 2\heartsuit\right) + \clubsuit = P(\exists y \in \mathcal{N} : |Ax| \leq 2\heartsuit) + \clubsuit \leq$$

$$\spadesuit \cdot \sup_{y \in \mathcal{N}} P(|Ay| \leq 2\heartsuit) + \clubsuit.$$

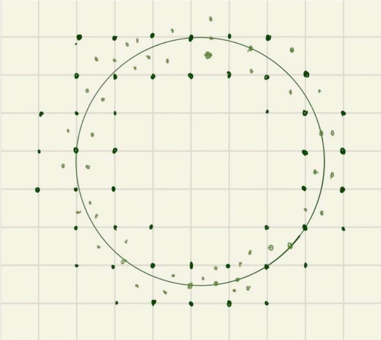
So **IF** we know that for each y, $P(|Ay| \leq 2\heartsuit) \leq \frac{\diamondsuit - \clubsuit}{\spadesuit}$, we are done!

The net result

Theorem (L. 2018+) – Lite version

There exists a deterministic net $\mathcal{N} \subset \frac{3}{2}B_2^n \setminus \frac{1}{2}B_2^n$ of cardinality 1000^n such that for any integer N and any $N \times n$ random matrix A with independent columns, with probability at least $1 - e^{-5n}$, for every $x \in \mathbb{S}^{n-1}$ there exists $y \in \mathcal{N}$ such that

$$|A(x-y)| \leq \frac{100}{\sqrt{n}} \sqrt{\mathbb{E}||A||_{HS}^2}.$$



Previously known cases

Folklore: A has **sub-gaussian** independent entries a_{ij}, $\mathbb{E}a_{ij} = 0$, $\mathbb{E}a_{ij}^2 = const$.

- Let $\mathcal{N}$ be the standard ε-net, i.e. such that

$$\mathbb{S}^{n-1} \subset \cup_{x \in \mathcal{N}} \left(x + \varepsilon B_2^n\right),$$

  and $\#\mathcal{N} \leq \left(\frac{3}{\varepsilon}\right)^n$.
- Then we can estimate $|A(x-y)| \leq ||A||\varepsilon \leq C\varepsilon \cdot \frac{||A||_{HS}}{\sqrt{n}}$?
- Recall, for any matrix A: $\frac{1}{\sqrt{n}}||A||_{HS} \leq ||A|| \leq ||A||_{HS}$.
- But specifically for sub-gaussian mean zero variance 1 case,

$$P\left(||A|| \geq \frac{100}{\sqrt{n}}\sqrt{\mathbb{E}||A||_{HS}^2}\right) \leq e^{-5n}. \tag{1}$$

- Without strong assumptions, (1) is not true.

Previously known cases

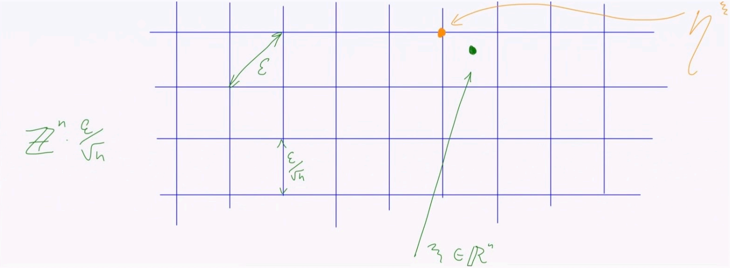
Rebrova, Tikhomirov (2016) proved this Theorem assuming i.i.d. UAC entries a_{ij}, with $\mathbb{E}a_{ij} = 0$, $\mathbb{E}a_{ij}^2 = const$, and $N = n$.

Guedon, Litvak, Tatarko (2019) extended the result of Rebrova and Tikhomirov in the case of arbitrary n, N, and replaced i.i.d. entries with i.i.d. columns.

- Advantage: the Theorem only assumes independence of columns, and no other structural assumptions!
- In particular, allowing dependent columns is crucial for the proof of the arbitrary aspect ratio result.
- Not requiring mean zero is another cool feature.

Step 1: randomized rounding and comparison via Hilbert-Schmidt

Randomized rounding (Raghavan-Tompson 1987, Beck 1987, Kannan-Vempala 1997, Srinivasan 1999, Alon-Klartag 2017, Klartag-L 2018+, L 2018+, Tikhomirov 2019+,...)



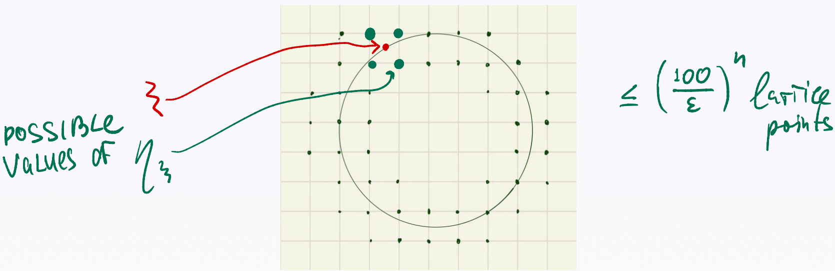
Definition

For $\xi \in \mathbb{S}^{n-1}$, write each $\xi_i = \frac{\epsilon}{\sqrt{n}}(k_i + p_i)$ for $k_i \in \mathbb{Z}$ and $p_i \in [0,1)$. Consider a random vector $\eta^\xi \in (\epsilon/\sqrt{n})\mathbb{Z}^n$:

$$\eta_i^\xi = \begin{cases} \frac{\epsilon}{\sqrt{n}} k_i, & \text{with probability } 1 - p_i \\ \frac{\epsilon}{\sqrt{n}}(k_i + 1), & \text{with probability } p_i. \end{cases}$$

Step 1: randomized rounding and comparison via Hilbert-Schmidt

- $\mathbb{S}^{n-1} \subset \bigcup_{j=1}^{(100/\epsilon)^n} \left(x_j + \frac{\epsilon}{\sqrt{n}} B_\infty^n \right)$.



- Therefore, there is a set $\mathcal{N}$ such that for all $\xi \in \mathbb{S}^{n-1}$, we have $\eta^\xi \in \mathcal{N}$, and $\#\mathcal{N} \leq \left(\frac{100}{\epsilon}\right)^n$;
- We have $\|\xi - \eta^\xi\|_\infty \leq \frac{\epsilon}{\sqrt{n}}$ and $\mathbb{E}\eta^\xi = \xi$;
- Hence, using the fact that $\mathbb{E}(\eta^\xi - \xi) = 0$, we get:

$$\mathbb{E}|\langle \eta^\xi - \xi, \theta \rangle|^2 \leq \frac{\epsilon^2 |\theta|^2}{n}. \ (\heartsuit)$$

Step 1: randomized rounding and comparison via Hilbert-Schmidt

Lemma 1 (comparison via Hilbert-Schmidt)

There exists a collection of points $\mathcal{F}$ with $\#\mathcal{F} \leq (\frac{C}{\epsilon})^{n-1}$ such that for any (deterministic) matrix $A : \mathbb{R}^n \to \mathbb{R}^N$, for every $\xi \in \mathbb{S}^{n-1}$ there exists an $\eta \in \mathcal{F}$ satisfying

$$|A(\eta - \xi)| \leq \frac{\epsilon}{\sqrt{n}} ||A||_{HS}.$$

Proof.

- Recall: $|Ax|^2 = \sum_{i=1}^N \langle A^T e_i, x\rangle^2$, where $A^T e_i$ are the rows of A;
- By ($\heartsuit$), $\mathbb{E}_\eta |\langle \eta^\xi - \xi, A^T e_i\rangle|^2 \leq C\frac{\epsilon^2 |A^T e_i|^2}{n}$;
- Summing up, we get

$$\mathbb{E}_\eta |A(\eta^\xi - \xi)|^2 = \mathbb{E}_\eta \sum_{i=1}^N \langle A^T e_i, \eta^\xi - \xi\rangle^2 \leq \left(C' \frac{\epsilon}{\sqrt{n}} ||A||_{HS} \right)^2;$$

- If $P(\textit{find a red ball in a box}) \geq 0.1$ then there exists a red ball in a box.

□

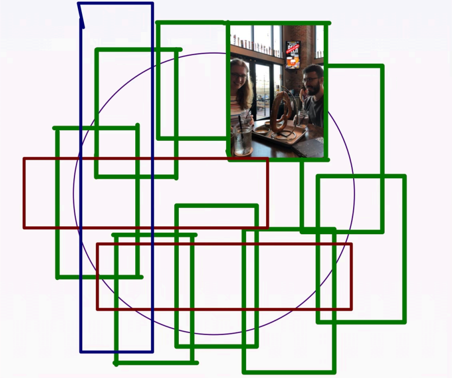
Step 2: parallelepipeds

Remark

$$P(||A||_{HS}^2 \geq 10\mathbb{E}||A||_{HS}^2) \leq 0.1.$$

Thus Lemma 1 implies the Theorem with probability 0.9 rather than $1 - e^{-5n}$. Not good:(

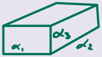
Idea of Rebrova and Tikhomirov, 2016: cover with parallelepipeds and not just cubes!



Step 2: parallelepipeds

Admissible set of parallelepipeds

- For $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$ with $\alpha_i > 0$, we fix the parallelepiped



$$P_\alpha = \{x \in \mathbb{R}^n : |x_i| \leq \alpha_i\}.$$

- For $\kappa > 1$, denote $\Omega_\kappa = \left\{\alpha \in \mathbb{R}^n : \alpha_i \in [0,1], \prod_{i=1}^n \alpha_i > \kappa^{-n}\right\}$.
- Note: if $\alpha \in \Omega_\kappa$ then $P_\alpha \geq (0.5\kappa)^{-n}$ – hence the covering is not too big.

Lemma 2 (comparison via parallelepipeds)

Pick any $\alpha \in \Omega_\kappa$. Let A be any $N \times n$ matrix. There exists a net $\mathcal{F}_\alpha$ with $\#\mathcal{F}_\alpha \leq \left(\frac{100\kappa}{\epsilon}\right)^n$ such that for every $\xi \in \mathbb{S}^{n-1}$ there exists an $\eta \in \mathcal{F}_\alpha$ satisfying

$$|A(\eta - \xi)| \leq \frac{\epsilon}{\sqrt{n}} \sqrt{\sum_{i=1}^n \alpha_i^2 |Ae_i|^2}.$$

Step 3: $\mathcal{B}_\kappa$ and nets on nets

Key definition: for any matrix A

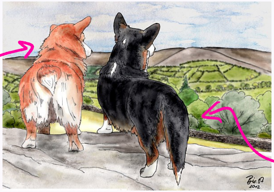
$$\mathcal{B}_\kappa(A) := \min_{\alpha_i \in [0,1],\, \prod_{i=1}^n \alpha_i \geq \kappa^{-n}} \sum_{i=1}^n \alpha_i^2 |Ae_i|^2.$$

Step 3: $\mathcal{B}_\kappa$ and nets on nets

Key definition: for any matrix A

$$\mathcal{B}_\kappa(A) := \min_{\alpha_i \in [0,1],\, \prod_{i=1}^n \alpha_i \geq \kappa^{-n}} \sum_{i=1}^{n} \alpha_i^2 |Ae_i|^2.$$

Minimum cuts off heavy tails!



$\|A\|_{HS}^2$

Corollary of Lemma 2

Let A be any $N \times n$ matrix. There exists a small enough net $\mathcal{F}$ such that for every $\xi \in \mathbb{S}^{n-1}$ there exists an $\eta \in \mathcal{F}$ satisfying

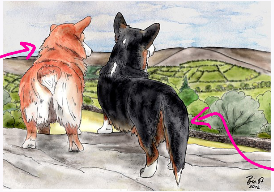
$$|A(\eta - \xi)| \leq \frac{\epsilon}{\sqrt{n}} \sqrt{\mathcal{B}_\kappa(A)}.$$

Step 3: $\mathcal{B}_\kappa$ and nets on nets

Key definition: for any matrix A

$$\mathcal{B}_\kappa(A) := \min_{\alpha_i \in [0,1],\, \prod_{i=1}^n \alpha_i \geq \kappa^{-n}} \sum_{i=1}^n \alpha_i^2 |Ae_i|^2.$$

Minimum cuts off heavy tails!



$\|A\|_{HS}^2$

Corollary of Lemma 2

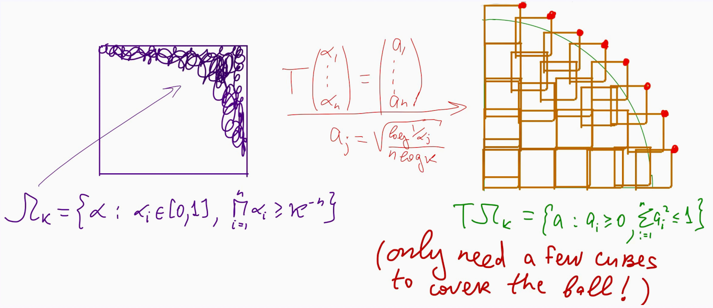
Let A be any $N \times n$ matrix. There exists a small enough net $\mathcal{F}$ such that for every $\xi \in \mathbb{S}^{n-1}$ there exists an $\eta \in \mathcal{F}$ satisfying

$$|A(\eta - \xi)| \leq \frac{\epsilon}{\sqrt{n}} \sqrt{\mathcal{B}_\kappa(A)}.$$

But the net depends on the matrix! Not good:(

Step 3: $\mathcal{B}_\kappa$ and nets on nets

Way out: discretize the admissible set Ω_κ.



The "nets on nets" Lemma

There exists a collection $\mathcal{F} \subset \Omega_{\kappa^2}$ of cardinality 30^n such that for any $\alpha \in \Omega_\kappa$ there exists a $\beta \in \mathcal{F}$ so that for all $i = 1, \ldots, n$ we have $\alpha_i^2 \geq \beta_i^2$.
In particular, for any $N \times n$ matrix A, we have

$$\mathcal{B}_\kappa(A) \geq \min_{\beta \in \mathcal{F}} \sum_{i=1}^{n} \beta_i^2 |Ae_i|^2.$$

A net for deterministic matrices: combining steps 1-3.

Theorem about deterministic matrices

There exists a deterministic net $\mathcal{N}$ of cardinality 1000^n such that for any integer N and any $N \times n$ deterministic matrix A, for every $x \in \mathbb{S}^{n-1}$ there exists $y \in \mathcal{N}$ such that

$$|A(x-y)| \leq \frac{100}{\sqrt{n}} \sqrt{\mathcal{B}_{10}(A)}.$$

This reduces the proof of the Theorem to estimating the large deviation of $\mathcal{B}_\kappa(A)$ when A is a random matrix coming from an appropriate model.

Step 4: Large deviation of $\mathcal{B}_\kappa$.

Recall: $\mathcal{B}_\kappa(A) := \min_{\alpha_i\in[0,1],\, \prod_{i=1}^n \alpha_i \geq \kappa^{-n}} \sum_{i=1}^n \alpha_i^2 |Ae_i|^2$.

Lemma

Let A be a random matrix with independent columns. Pick any $\kappa > 1$. Then

$$P\left(B_\kappa(A) \geq 10\mathbb{E}||A||_{HS}^2\right) \leq (C\kappa)^{-2n}.$$

Step 4: Large deviation of $\mathcal{B}_\kappa$.

Recall: $\mathcal{B}_\kappa(A) := \min_{\alpha_i \in [0,1], \prod_{i=1}^n \alpha_i \geq \kappa^{-n}} \sum_{i=1}^n \alpha_i^2 |Ae_i|^2$.

Lemma

Let A be a random matrix with independent columns. Pick any $\kappa > 1$. Then

$$P\left(B_\kappa(A) \geq 10\mathbb{E}||A||_{HS}^2\right) \leq (C\kappa)^{-2n}.$$

Theorem follows now!

Step 4: Large deviation of $\mathcal{B}_\kappa$.

Recall: $\mathcal{B}_\kappa(A) := \min_{\alpha_i \in [0,1], \prod_{i=1}^n \alpha_i \geq \kappa^{-n}} \sum_{i=1}^n \alpha_i^2 |Ae_i|^2$.

Lemma

Let A be a random matrix with independent columns. Pick any $\kappa > 1$. Then

$$P\left(B_\kappa(A) \geq 10\mathbb{E}||A||_{HS}^2\right) \leq (C\kappa)^{-2n}.$$

Theorem follows now!

Small Question:

Could the above estimate be improved if we consider

$$P\left(B_\kappa(A) \geq t \cdot \mathbb{E}||A||_{HS}^2\right) \leq ?,\ t \to \infty$$

If yes, then one could potentially remove the extra $\left(\log \frac{1}{\epsilon}\right)^n$ factor in the "tall" theorem.

Step 4: Large deviation of $\mathcal{B}_\kappa$.

Lemma

Let A be a random matrix with independent columns. Pick any $\kappa > 1$. Then

$$P\left(B_\kappa(A) \geq 10\mathbb{E}||A||_{HS}^2\right) \leq (C\kappa)^{-2n}.$$

Proof.

- Denote $Y_i = |Ae_i|$. If $B_\kappa(A) \geq 10\sum_{i=1}^n \mathbb{E}Y_i^2$, then for any collection $\alpha_1, ..., \alpha_n \in [0,1]$, either

$$\sum_{i=1}^n \alpha_i^2 Y_i^2 \geq 10\sum_{i=1}^n \mathbb{E}Y_i^2,$$

or

$$\prod_{i=1}^n \alpha_i < \kappa^{-n}.$$

- Consider a collection of random variables $\alpha_i^2 = \min\left(1, \frac{\mathbb{E}Y_i^2}{Y_i^2}\right)$.

Step 4: Large deviation of $\mathcal{B}_\kappa$.

Proof.

- We estimate

$$P\left(B_\kappa(A) \geq 10\mathbb{E}||A||_{HS}^2\right) \leq$$

$$P\left(\sum_{i=1}^{n} \min\left(1, \frac{\mathbb{E}Y_i^2}{Y_i^2}\right) Y_i^2 \geq 10\mathbb{E}||A||_{HS}^2\right) +$$

$$P\left(\prod_{i=1}^{n} \min\left(1, \frac{\mathbb{E}Y_i^2}{Y_i^2}\right) < \kappa^{-2n}\right) =: P_1 + P_2.$$

- $P_1 = 0$.
- By Markov's inequality, $P_2 \leq (C\kappa)^{-2n}$.

□

Summary: the non-lite version of the net theorem

Theorem (NON-lite)

Fix $n \in \mathbb{N}$. Consider any $S \subset \mathbb{R}^n$. Pick any $\gamma \in (1, \sqrt{n})$, $\epsilon \in (0, \frac{1}{20\gamma})$, $\kappa > 1$, $p > 0$ and $s > 0$.
There exists a (deterministic) net $\mathcal{N} \subset S + 4\epsilon\gamma B_2^n$, with

$$\#\mathcal{N} \leq \begin{cases} N(S, \epsilon B_2^n) \cdot (C_1\gamma)^{\frac{C_2 n}{\gamma^{0.08}}}, & \text{if } \log\kappa \leq \frac{\log 2}{\gamma^{0.09}}, \\ N(S, \epsilon B_2^n) \cdot (C\kappa\log\kappa)^n, & \text{if } \log\kappa \geq \frac{\log 2}{\gamma^{0.09}}, \end{cases}$$

such that for every $N \in \mathbb{N}$ and every random $N \times n$ matrix A with independent columns, with probability at least

$$1 - \kappa^{-2pn}\left(1 + \frac{1}{s^p}\right)^n,$$

for every $x \in S$ there exists $y \in \mathcal{N}$ such that

$$|A(x-y)| \leq C_3 \frac{\epsilon\gamma\sqrt{s}}{\sqrt{n}} \sqrt{\sum_{i=1}^{n} (\mathbb{E}|Ae_i|^{2p})^{\frac{1}{p}}}.$$

Here C, C_1, C_2, C_3 are absolute constants. γ is the "sparsity" parameter

The distance theorem

Theorem about distances (L, Tikhomirov, Vershynin, 2019+)

Let an $n \times n$ A have independent UAC entries and $\mathbb{E}||A||_{HS}^2 \leq Kn^2$. Denote

$$H_j = span\{Ae_i : i \neq j, i = 1, \ldots, n\};$$

Take any $j \leq n$ such that $\mathbb{E}|Ae_j|^2 \leq rn^2$. Then

$$P\left(dist(A_j, H_j) \leq \varepsilon\right) \leq C\varepsilon + 2e^{-cn}, \quad \varepsilon \geq 0.$$

Theorem (Fernandez, L, Tatarko, TBD) – ongoing

The analogous result is also true for inhomogeneous matrices with arbitrary aspect ratio.

Sketch of the proof of the distance theorem

Esseen's Lemma

Given a variable ξ with the characteristic function $\varphi(\cdot) = \mathbb{E}\exp(2\pi \mathbf{i} \xi \cdot)$,

$$P(|\xi| < t) \le C \int_{-1}^{1} \left| \varphi\left(\frac{s}{t}\right) \right| ds, \quad t > 0,$$

where $C > 0$ is an absolute constant.

RLCD – definition

For a random vector X in $\mathbb{R}^n$, a (deterministic) vector v in $\mathbb{R}^n$, and parameters $L > 0$, $u \in (0,1)$, define

$$RLCD^X_{L,u}(v) := \inf\left\{\theta > 0 : \mathbb{E}dist^2(\theta v \star \overline{X}, \mathbb{Z}^n) < \min(u|\theta v|^2, L^2)\right\}.$$

Here by $\star$ we denote the Schur product

$$v \star X := (v_1 X_1, \ldots, v_n X_n).$$

Note: Rudelson-Vershynin previously defined LCD, a parameter which worked well to study the i.i.d. case.

Sketch of the proof of the distance theorem

Geometric meaning of RLCD

$RLCD^X(v)$ is roughly how much the X-associated lattice has to be scaled down to get close to v.

Anticoncentration via RLCD

Let $X = (X_1, \ldots, X_n)$ be a random vector with independent coordinates satisfying $\max_i P(\sup_{z\in\mathbb{R}} |X_i - z| < 1) \le b$ for some $b \in (0,1)$. Let $c_0 > 0$, $L > 0$ and $u \in (0,1)$. Then for any vector $v \in \mathbb{R}^n$ with $|v| \ge c_0$ and any $\varepsilon \ge 0$, we have

$$P(\langle X, v\rangle < \varepsilon) \le C\varepsilon + C\exp(-\widetilde{c}L^2) + \frac{C}{RLCD^X_{L,u}(v)}.$$

Here $C > 0, \widetilde{c} > 0$ may only depend on b, c_0, u.

In words

If RLCD of a vector v is large, then the scalar product $\langle X, v\rangle$ has great anti-concentration properties!

Sketch of the proof of the distance theorem using the double counting idea

- Let ν be the random normal, orthogonal to columns $Ae_2, \ldots, Ae_n$.
- Goal: $RLCD(\nu) = LARGE$.
- Let $M = [Ae_2, \ldots, Ae_n]^T$, then $M\nu = 0$ (since it is orthogonal to all the rows of M).
- Consider a net $\mathcal{N}$ (from the net theorem) on $\mathbb{S}^{n-1}$ with respect to M.
- Let $S_{bad} = \{y \in \mathbb{S}^{n-1} : RLCD(y) = small\ but\ not\ too\ small\}$ (a level set)
- $$P(\nu \in S_{bad}) = P(\inf_{x \in S_{bad}} |Mx| = 0) \leq \#\mathcal{F} \cdot P(|Mx| < \epsilon\sqrt{n}),$$

  where $\mathcal{F} \subset \mathcal{N}$ which forms a net on S_{bad}.
- Since on S_{bad} RLCD is not too bad, $P(|Mx| < \epsilon\sqrt{n})$ is small.
- **Most of the points on a lattice have large RLCD! – the double counting method.**
- $\#\mathcal{F} \leq e^{-Cn} \#\mathcal{N}$, since $RLCD$ is stable.
- Combining these bounds allows to iterate on the level sets and to obtain the distance theorem.

Open Questions

- How to remove the $\left(\log \frac{1}{\epsilon}\right)^n$ from the "tall" Theorem?
- What can one say about the invertibility of inhomogeneous sparse matrices?
- Given a random matrix A with an inhomogeneous profile, determine the expectation of $\sigma_n(A)$ explicitly depending on the profile.
- Estimate $\sigma_n(B+M)$ where B is the Bernoulli matrix and M has Hilbert-Schmidt norm larger than Cn.
- What can be said about the invertibility of the inhomogeneous symmetric Wigner matrices?

Thanks for your attention!