Coverings of random ellipsoids, and invertibility of matrices with i.i.d. heavy-tailed entries

Liza Rebrova*
University of Michigan

Konstantin Tikhomirov
University of Alberta

March 6, 2016

Invertibility problem

Theorem (Rudelson, Vershynin)

Let $n \geq n_0$ and let $A = (a_{ij})$ be an $n \times n$ random matrix with i.i.d mean zero subgaussian entries. Then for any $\varepsilon > 0$ we have

$$\mathbb{P}\left\{s_n(A) \leq \varepsilon n^{-1/2}\right\} \leq L\varepsilon + u^n,$$

where $L > 0$ and $u \in (0,1)$ depend only on the distribution of a_{ij}.

Definition

A random variable ξ is called subgaussian, if for any $t > 0$

$$\mathbb{P}\{|\xi| > t\} \leq C\exp(-ct^2).$$

Invertibility problem

Theorem (R,Tikhomirov)

Let $n \geq n_0$ and let $A = (a_{ij})$ be an $n \times n$ random matrix with i.i.d mean zero $\mathbb{E}a_{ij}^2 = 1$ entries. Then for any $\varepsilon > 0$ we have

$$\mathbb{P}\{s_n(A) \leq \varepsilon n^{-1/2}\} \leq L\varepsilon + u^n,$$

where $L > 0$ and $u \in (0,1)$ depend only on the distribution of a_{ij}.

Corollary: i.i.d. matrices with heavy-tailed entries (as well as subgaussian) are well-invertible, as

$$\|A^{-1}\| = s_{max}(A^{-1}) = 1/s_n(A) \sim \sqrt{n}.$$

Subgaussian case, idea of the proof

$$s_n(A) := s_{min}(A) = \min_{x \in S^{n-1}} \|Ax\|$$

Approximation by the ε-net $\mathcal{N} \subset S^{n-1}$.
For any $x \in S^{n-1}$ find the closest $y \in \mathcal{N}$:

$$\|Ax\| \geq \|Ay\| - \|A(x-y)\| \geq \|Ay\| - \|A\|\|x-y\| \geq \inf_{y \in \mathcal{N}} \|Ay\| - \sqrt{n} \cdot \varepsilon$$

Lemma (norm of a subgaussian matrix $\sim \sqrt{n}$)

For A be an $n \times n$ random matrix with i.i.d. subgaussian entries.

$$\mathbb{P}\{\|A\| \geq t\sqrt{n}\} \leq \exp(-c_0 t^2 n) \qquad \text{for } t > C_0.$$

Challenge: find an ε-net with the sufficiently low cardinality

$$|\mathcal{N}| \sim \left(\frac{c}{\varepsilon\sqrt{n}}\right)^n.$$

Heavy-tailed case

$$s_n(A) := s_{min}(A) = \min_{x \in S^{n-1}} \|Ax\|$$

Approximation by the ε-net $\mathcal{N} \subset S^{n-1}$.
For any $x \in S^{n-1}$ find the closest $y \in \mathcal{N}$:

$$\|Ax\| \geq \|Ay\| - \|A(x-y)\| \geq \|Ay\| - \|A\|\|x-y\| \geq ???$$

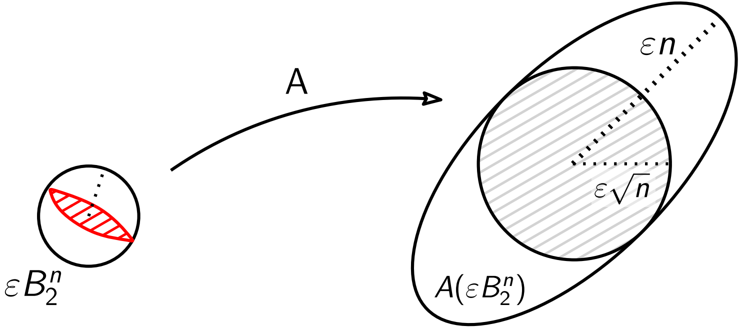
Norm is too large:

$$\|A\| \sim n >> \sqrt{n}$$

New challenge: obtain an estimate $\|A(x-y)\| \leq \sqrt{n}\varepsilon$, where x, y are in the same ε-net element.

So, for any x, y taken from one net element we would like to have

$$\|A(x-y)\| \leq \sqrt{n}\varepsilon$$



The net should be refined without blowing up of cardinality $|\mathcal{N}|$. It is possible as A cannot have too many large directions!

(*) Our model: A is an $n \times n$ random matrix with i.i.d $\mathbb{E}a_{ij} = 0$ and $\mathbb{E}a_{ij}^2 = 1$ entries.

Definition

We call a diagonal matrix $D = (d_{ij})_{i,j=1}^n$ with $d_{ii} \in (0,1)$ a (positive diagonal) *contraction*.

Key proposition

Let $\delta \in (0,1]$ *and* A *satisfies (*). Then with probability* $1 - 4\exp(-\delta n/8)$ *there exists a contraction* D*, such that (1)* $\det D \geq \exp(-\delta n)$ *and (2)* $\|AD\|_{\infty\to 2} \leq \frac{C}{\sqrt{\delta}} n$.

Then the shifts of parallelepiped $D(n^{-\frac{1}{2}} B_\infty^n)$ can serve as "dilated net elements":

$$\|AD(n^{-\frac{1}{2}} B_\infty^n)\| \leq \frac{1}{\sqrt{n}} \|AD\|_{\infty\to 2} \leq \frac{C}{\sqrt{\delta}} \sqrt{n},$$

Consequences of the key proposition

1) This solves an invertibility problem for heavy-tailed matrices: there is a net refinement of cardinality $|\mathcal{N}'| \leq C^n \left(\frac{c}{\varepsilon\sqrt{n}}\right)^n$ such that $\|A(x-y)\| \sim \sqrt{n}\varepsilon$ for every A with high probability.

2) Also, it implies the following geometric theorem about the existence of a good covering for a random ellipsoid:

Theorem A

For any $\delta \in (0, 1/4]$ there is a subset $\mathcal{N} \subset B_2^n$ of cardinality at most $\exp(C_\delta n)$, such that for any random matrix A satisfying (*) with probability at least $1 - 4\exp(-n\delta/32)$ we have

$$A(B_2^n) \subset \bigcup_{y \in A(\mathcal{N})} \left(y + c_\delta \sqrt{n} B_2^n\right).$$

Proof of key proposition

Key proposition

Let $\delta \in (0,1]$ and A satisfies (*). Then with probability $1 - 4\exp(-\delta n/8)$ there exists a contraction D, such that (1) $\det D \geq \exp(-\delta n)$ and (2) $\|AD\|_{\infty\to 2} \leq \frac{C}{\sqrt{\delta}} n$.

Construction of D:

1. Step 1: construct D with the following properties:
   - Euclidean norms of all the rows of AD are bounded by $C_\delta\sqrt{n}$
   - $\det D \geq \exp(-\delta n)$
2. Step 2: show that the norms of the rows condition implies that $\|AD\|_{\infty\to 2} \leq C'_\delta n$ with high probability

Step 1: Bernoulli case, for one row

Suppose a vector $X := (a_{i1}^2, \dots, a_{in}^2)$ has Bernoulli coordinates.
We define D_i for X:

$$\begin{cases} d_{ii} := \frac{Cn}{\|X\|_1} \cdot \mathbb{I}_{\{x_i \neq 0\}} \text{ if } \|X\|_1 > Cn \\ d_{ii} := 1 \text{ otherwise} \end{cases}$$

Then $\|\langle X, D_i \rangle\|_1 \leq Cn$ and $\det D_i \geq \exp(-\delta)$.

Example 1 (making l_1 norm a bad row to be at most 2):

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & d & 0 & 0 & 0 \\ 0 & 0 & d & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & d \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & d & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & d & d & 0 & d \\ 1 & 0 & 0 & 0 & d \end{bmatrix}$$

Step 1: Bernoulli case, for a matrix

Define

$$D := \prod_{i=1}^{n} D_i$$

Then

- $\|\langle X, D\rangle\|_1 \leq Cn$ for every row $X = (a_{i1}^2, \ldots, a_{in}^2)$
- $\det D = \prod(\det D_i) \geq \exp(-\delta n)$.

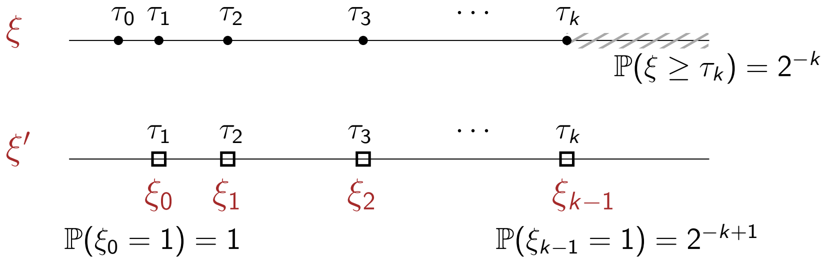
Example 2 (making l_1 norm of every row be at most 2):

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} d_1 \cdot 1 & 0 & 0 & 0 & 0 \\ 0 & d_1 d_2 & 0 & 0 & 0 \\ 0 & 0 & 1 \cdot d_2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & d_1 d_2 \end{bmatrix} = \begin{bmatrix} d_1 & 0 & 0 & 1 & 0 \\ 0 & d_1 d_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & d_1 d_2 & d_2 & 0 & d_1 d_2 \\ d_1 & d_1 d_2 & 0 & 0 & d_1 d_2 \end{bmatrix}$$

Step 1: General distribution case

Main idea: any random variable ξ (for us $\xi = a_{ij}^2$) can be almost surely approximated above by the sum of Bernoulli random variables ξ_i, such that $\mathbb{P}(\xi_i = 1) = 2^{-i}$,

$$\xi' := \sum_{i=1}^{\infty} \tau_i \xi_i \geq \xi, \qquad \text{and} \qquad \mathbb{E}\xi' \leq 2\mathbb{E}\xi.$$



Step 3: from bounded rows to bounded $\|.\|_{\infty\to 2}$ norm

Lemma

Let A has symmetrically distributed entries and satisfies (*). If there exists $D \in \mathcal{F}$ (any countable subset of the set of contractions), such that Euclidean norms of the rows AD are at most $K\sqrt{n}$, then

$$\|AD\|_{\infty\to 2} \le CKn$$

with exponentially high probability.

Main idea: as

$$\|AD\|^2_{\infty\to 2} = \sup_{v\in Vert(B^n_\infty)} \|ADv\|^2 = \sum \langle ADv, e_i\rangle^2,$$

after conditioning on $|A|$ and on D (only signs of the entries matter) we can use Bernstein's inequality to bound the norm.

References and related work

- E. Rebrova, K.Tikhomirov, Coverings of random ellipsoids, and invertibility of matrices with i.i.d. heavy-tailed entries Preprint. arXiv:1508.06690
- Tikhomirov K. E. The smallest singular value of random rectangular matrices with no moment assumptions on entries. Israel Journal of Mathematics, 2016. DOI: 10.1007/s11856-016-1287-8.
- Rudelson, M.; Vershynin, R. The Littlewood-Offord problem and invertibility of random matrices. Adv. Math. 218 (2008), no. 2, 600–633.
- M. Rudelson, Lecture notes on non-asymptotic random matrix theory, Notes from the AMS Short Course on Random Matrices, 2013

Thanks for your attention!

Please share your questions and comments with me.

My contact email is erebrova@umich.edu