Geometric properties of random matrices with independent log-concave rows/columns

Radosław Adamczak

University of Warsaw

Paris, May 2010

Based on joint work with
O. Guédon, A. Litvak, A. Pajor, N. Tomczak-Jaegermann

Isotropicity, the ψ_α condition

- A random vector X in $\mathbb{R}^n$ is **isotropic** if

$$\mathbb{E}X = 0$$

Isotropicity, the ψ_α condition

- A random vector X in $\mathbb{R}^n$ is **isotropic** if

$$\mathbb{E}X = 0$$

and

$$\mathbb{E}X \otimes X = \mathrm{Id}$$

Isotropicity, the ψ_α condition

- A random vector X in $\mathbb{R}^n$ is **isotropic** if

$$\mathbb{E}X = 0$$

and

$$\mathbb{E}X \otimes X = \mathrm{Id}$$

or equivalently for all $y \in \mathbb{R}^n$,

$$\mathbb{E}\langle X, y\rangle^2 = |y|^2.$$

Isotropicity, the ψ_α condition

- A random vector X in $\mathbb{R}^n$ is **isotropic** if

$$\mathbb{E}X = 0$$

and

$$\mathbb{E}X \otimes X = \mathrm{Id}$$

or equivalently for all $y \in \mathbb{R}^n$,

$$\mathbb{E}\langle X, y\rangle^2 = |y|^2.$$

- X is ψ_α ($\alpha \in [1, 2]$) with constant C if for all $y \in \mathbb{R}^n$,

$$\|\langle X, y\rangle\|_{\psi_\alpha} \le C|y|,$$

where

$$\|Y\|_{\psi_\alpha} = \inf\{a > 0 \colon \mathbb{E}\exp((Y/a)^\alpha) \le 2\}$$

Consequences

Consequences

- $\mathbb{E}|X_i|^2 = \sum_{j=1}^{n} \mathbb{E}\langle X_i, e_j\rangle^2 = n$

Consequences

- $\mathbb{E}|X_i|^2 = \sum_{j=1}^n \mathbb{E}\langle X_i, e_j\rangle^2 = n$
- For any $y \in S^{n-1}$ and $t \geq 0$,

$$\mathbb{P}(|\langle X_i, y\rangle| \geq t) \leq 2\exp(-(t/C)^\alpha).$$

Consequences

- $\mathbb{E}|X_i|^2 = \sum_{j=1}^{n} \mathbb{E}\langle X_i, e_j\rangle^2 = n$
- For any $y \in S^{n-1}$ and $t \geq 0$,

$$\mathbb{P}(|\langle X_i, y\rangle| \geq t) \leq 2\exp(-(t/C)^{\alpha}).$$

Fact

For every random vector X not supported on any $n-1$ dimensional hyperplane, there exists an affine map $T\colon \mathbb{R}^n \to \mathbb{R}^n$ such that TX is isotropic.

Consequences

- $\mathbb{E}|X_i|^2 = \sum_{j=1}^{n} \mathbb{E}\langle X_i, e_j\rangle^2 = n$
- For any $y \in S^{n-1}$ and $t \geq 0$,

$$\mathbb{P}(|\langle X_i, y\rangle| \geq t) \leq 2\exp(-(t/C)^{\alpha}).$$

Fact

For every random vector X not supported on any $n-1$ dimensional hyperplane, there exists an affine map $T: \mathbb{R}^n \to \mathbb{R}^n$ such that TX is isotropic.

If for a set $K \subseteq \mathbb{R}^n$ the random vector distributed uniformly on K is isotropic, we say that K is isotropic.

Log-concavity

A random vector X in $\mathbb{R}^n$ is log-concave if its law μ satisfies a Brunn-Minkowski type inequality

$$\mu(\theta A + (1-\theta)B) \geq \mu(A)^\theta \mu(B)^{1-\theta}.$$

Theorem (Borell)

A random vector not supported on any $(n-1)$ dimensional hyperplane is log-concave iff it has density of the form $\exp(-V(x))$, where $V\colon \mathbb{R}^n \to (-\infty, \infty]$ is convex.

Lemma (Borell)

An isotropic log-concave random vector is ψ_1 with a universal constant C.

Examples

The following distributions are log-concave:

- Gaussian measures
- Uniform distributions on convex bodies
- Measures with density of the form $C\exp(-\|x\|)$, where $\|x\|$ is a norm.
- Products, affine images and convolutions of the above distributions.

The basic model

Definition

Let Γ be an $n \times N$ matrix with columns $X_1, \ldots, X_N$, where X_i's are independent isotropic log-concave random vectors in $\mathbb{R}^n$

Questions

The basic model

Definition

Let Γ be an $n \times N$ matrix with columns $X_1, \ldots, X_N$, where X_i's are independent isotropic log-concave random vectors in $\mathbb{R}^n$

Questions

- What is the operator norm of $\Gamma \colon \ell_2^N \to \ell_2^n$?

The basic model

Definition

Let Γ be an $n \times N$ matrix with columns $X_1, \ldots, X_N$, where X_i's are independent isotropic log-concave random vectors in $\mathbb{R}^n$

Questions

- What is the operator norm of $\Gamma \colon \ell_2^N \to \ell_2^n$?
- When is Γ^T close to a multiple of isometry?

The basic model

Definition

Let Γ be an $n \times N$ matrix with columns $X_1, \ldots, X_N$, where X_i's are independent isotropic log-concave random vectors in $\mathbb{R}^n$

Questions

- What is the operator norm of $\Gamma \colon \ell_2^N \to \ell_2^n$?
- When is Γ^T close to a multiple of isometry?
- How does Γ act on sparse vectors?

The basic model

Definition

Let Γ be an $n \times N$ matrix with columns $X_1, \ldots, X_N$, where X_i's are independent isotropic log-concave random vectors in $\mathbb{R}^n$

Questions

- What is the operator norm of $\Gamma\colon \ell_2^N \to \ell_2^n$?
- When is Γ^T close to a multiple of isometry?
- How does Γ act on sparse vectors?
- What is the smallest singular value of Γ?

Motivations: sampling convex bodies

Problem

Let $K \subseteq \mathbb{R}^n$ be a convex body, s.t. $B_2^n \subseteq K \subseteq RB_2^n$. Assume we have access to an oracle (a black box), which given $x \in \mathbb{R}^n$ tells us whether $x \in K$.

Motivations: sampling convex bodies

Problem

Let $K \subseteq \mathbb{R}^n$ be a convex body, s.t. $B_2^n \subseteq K \subseteq RB_2^n$. Assume we have access to an oracle (a black box), which given $x \in \mathbb{R}^n$ tells us whether $x \in K$.

How to generate random points uniformly distributed in K?
How to compute the volume of K?

Motivations: sampling convex bodies

Problem

Let $K \subseteq \mathbb{R}^n$ be a convex body, s.t. $B_2^n \subseteq K \subseteq RB_2^n$. Assume we have access to an oracle (a black box), which given $x \in \mathbb{R}^n$ tells us whether $x \in K$.

How to generate random points uniformly distributed in K?
How to compute the volume of K?

- This can be done by using Markov chains.
- Their speed of convergence depends on the position of the convex body.
- Preprocessing: First put K in the isotropic position (again by randomized algorithms).

- Centering the body is not comp. difficult – takes $\mathcal{O}(n)$ steps.
- The question boils down to:

  How to approximate the covariance matrix of X - uniformly distributed on K by the empirical covariance matrix

$$\frac{1}{N}\sum_{i=1}^{N} X_i \otimes X_i.$$

  or (after a linear transformation)

- Centering the body is not comp. difficult – takes $\mathcal{O}(n)$ steps.
- The question boils down to:

How to approximate the covariance matrix of X - uniformly distributed on K by the empirical covariance matrix

$$\frac{1}{N}\sum_{i=1}^{N} X_i \otimes X_i.$$

or (after a linear transformation)

Given an isotropic convex body in $\mathbb{R}^n$, how large N should we take so that

$$\Big\|\frac{1}{N}\sum_{i=1}^{N} X_i \otimes X_i - Id\Big\|_{\ell_2\to\ell_2} \le \varepsilon$$

with high probability?

Interpretation in terms of Γ.

We have

$$\Big\|\frac{1}{N}\sum_{i=1}^{N} X_i \otimes X_i - Id\Big\|_{\ell_2\to\ell_2} = \sup_{y\in S^{n-1}}\Big|\frac{1}{N}\sum_{i=1}^{N}\langle X_i, y\rangle^2 - 1\Big|$$

$$= \sup_{y\in S^{n-1}}\Big|\frac{1}{N}|\Gamma^T y|^2 - 1\Big|$$

Interpretation in terms of Γ.

We have

$$\left\|\frac{1}{N}\sum_{i=1}^{N} X_i \otimes X_i - Id\right\|_{\ell_2\to\ell_2} = \sup_{y\in S^{n-1}} \left|\frac{1}{N}\sum_{i=1}^{N}\langle X_i, y\rangle^2 - 1\right|$$

$$= \sup_{y\in S^{n-1}} \left|\frac{1}{N}|\Gamma^T y|^2 - 1\right|$$

So the (geometric) question is

Interpretation in terms of Γ.

We have

$$\left\|\frac{1}{N}\sum_{i=1}^{N} X_i \otimes X_i - Id\right\|_{\ell_2\to\ell_2} = \sup_{y\in S^{n-1}}\left|\frac{1}{N}\sum_{i=1}^{N}\langle X_i, y\rangle^2 - 1\right|$$
$$= \sup_{y\in S^{n-1}}\left|\frac{1}{N}|\Gamma^T y|^2 - 1\right|$$

So the (geometric) question is

Let Γ be a matrix with independent columns $X_1, \ldots, X_N$ drawn from an isotropic convex body (log-concave measure) in $\mathbb{R}^n$.

Interpretation in terms of Γ.

We have

$$\left\|\frac{1}{N}\sum_{i=1}^{N} X_i \otimes X_i - Id\right\|_{\ell_2\to\ell_2} = \sup_{y\in S^{n-1}} \left|\frac{1}{N}\sum_{i=1}^{N}\langle X_i, y\rangle^2 - 1\right|$$
$$= \sup_{y\in S^{n-1}} \left|\frac{1}{N}|\Gamma^T y|^2 - 1\right|$$

So the (geometric) question is

Let Γ be a matrix with independent columns $X_1, \ldots, X_N$ drawn from an isotropic convex body (log-concave measure) in $\mathbb{R}^n$.

How large should N be so that $N^{-1/2}\Gamma^T\colon \mathbb{R}^n \to \mathbb{R}^N$ was an almost isometry?

History of the problem

- Kannan, Lovasz, Simonovits (1995) – $N = \mathcal{O}(n^2)$

History of the problem

- Kannan, Lovasz, Simonovits (1995) – $N = \mathcal{O}(n^2)$
- Bourgain (1996) – $N = \mathcal{O}(n \log^3 n)$

History of the problem

- Kannan, Lovasz, Simonovits (1995) – $N = \mathcal{O}(n^2)$
- Bourgain (1996) – $N = \mathcal{O}(n \log^3 n)$
- Rudelson (1999) – $N = \mathcal{O}(n \log^2 n)$

History of the problem

- Kannan, Lovasz, Simonovits (1995) – $N = \mathcal{O}(n^2)$
- Bourgain (1996) – $N = \mathcal{O}(n \log^3 n)$
- Rudelson (1999) – $N = \mathcal{O}(n \log^2 n)$
- Giannopoulos, Hartzoulaki, Tsolomitis (2005) – unconditional bodies: $N = \mathcal{O}(n \log n)$

History of the problem

- Kannan, Lovasz, Simonovits (1995) – $N = \mathcal{O}(n^2)$
- Bourgain (1996) – $N = \mathcal{O}(n \log^3 n)$
- Rudelson (1999) – $N = \mathcal{O}(n \log^2 n)$
- Giannopoulos, Hartzoulaki, Tsolomitis (2005) – unconditional bodies: $N = \mathcal{O}(n \log n)$
- Aubrun (2006) – unconditional bodies: $N = \mathcal{O}(n)$

History of the problem

- Kannan, Lovasz, Simonovits (1995) – $N = \mathcal{O}(n^2)$
- Bourgain (1996) – $N = \mathcal{O}(n \log^3 n)$
- Rudelson (1999) – $N = \mathcal{O}(n \log^2 n)$
- Giannopoulos, Hartzoulaki, Tsolomitis (2005) – unconditional bodies: $N = \mathcal{O}(n \log n)$
- Aubrun (2006) – unconditional bodies: $N = \mathcal{O}(n)$
- Paouris (2006) – $N = \mathcal{O}(n \log n)$

History of the problem

- Kannan, Lovasz, Simonovits (1995) – $N = \mathcal{O}(n^2)$
- Bourgain (1996) – $N = \mathcal{O}(n \log^3 n)$
- Rudelson (1999) – $N = \mathcal{O}(n \log^2 n)$
- Giannopoulos, Hartzoulaki, Tsolomitis (2005) – unconditional bodies: $N = \mathcal{O}(n \log n)$
- Aubrun (2006) – unconditional bodies: $N = \mathcal{O}(n)$
- Paouris (2006) – $N = \mathcal{O}(n \log n)$
- Litvak, Pajor, Tomczak-Jaegermann, R.A. (2008) – $N = \mathcal{O}(n)$

History of the problem

- Kannan, Lovasz, Simonovits (1995) – $N = \mathcal{O}(n^2)$
- Bourgain (1996) – $N = \mathcal{O}(n \log^3 n)$
- Rudelson (1999) – $N = \mathcal{O}(n \log^2 n)$
- Giannopoulos, Hartzoulaki, Tsolomitis (2005) – unconditional bodies: $N = \mathcal{O}(n \log n)$
- Aubrun (2006) – unconditional bodies: $N = \mathcal{O}(n)$
- Paouris (2006) – $N = \mathcal{O}(n \log n)$
- Litvak, Pajor, Tomczak-Jaegermann, R.A. (2008) – $N = \mathcal{O}(n)$

For arbitrary isotropic random vectors, if you do not assume any uniform bound on $\langle X_i, y\rangle$, $y \in S^{n-1}$, you cannot remove the logarithm (the optimal bound $N = \mathcal{O}(n \log^\beta n)$ is due to M. Rudelson). Recently $N = O(n \log\log n)$ was proven under a uniform bound on $(4+\varepsilon)$-th moments of $\langle X_i, y\rangle$ (R. Vershynin).

Remark

If $\frac{1}{\sqrt{N}}\Gamma^T$ is an almost isometry then obviously $\|\Gamma\| \leq C\sqrt{N}$, so the KLS question and the question about $\|\Gamma\|$ are related.

Remark

If $\frac{1}{\sqrt{N}}\Gamma^T$ is an almost isometry then obviously $\|\Gamma\| \le C\sqrt{N}$, so the KLS question and the question about $\|\Gamma\|$ are related.

It turns out that to answer KLS it is enough to have good bounds on

$$A_m := \sup_{\substack{z \in S^{N-1} \\ |\mathrm{supp}\ z| \le m}} |\Gamma z|$$

Remark

If $\frac{1}{\sqrt{N}}\Gamma^T$ is an almost isometry then obviously $\|\Gamma\| \le C\sqrt{N}$, so the KLS question and the question about $\|\Gamma\|$ are related.

It turns out that to answer KLS it is enough to have good bounds on

$$A_m := \sup_{\substack{z\in S^{N-1} \\ |\mathrm{supp}\ z|\le m}} |\Gamma z|$$

Theorem (Litvak, Pajor, Tomczak-Jaegermann, R.A.)

If $N \le \exp(c\sqrt{n})$ and the vectors X_i are log-concave then for $t > 1$, with probability at least $1 - \exp(-ct\sqrt{n})$,

$$\forall_{m\le N}\ A_m \le Ct\Big(\sqrt{n} + \sqrt{m}\log\Big(\frac{2N}{m}\Big)\Big).$$

In particular, with high probability $\|\Gamma\| \le C(\sqrt{n} + \sqrt{N})$.

Sketch of the proof

A modification of Bourgain's approach. One approximates an arbitrary vector z with $|\mathrm{supp}\, z| \le m$ by $x_0 + x_1 + \ldots + x_l$ ($l < \log_2 m$), where

$$|\mathrm{supp}\, x_i| \simeq m/2^i,\ \|x_i\|_\infty \simeq \sqrt{2^i/m},\ i \ge 1$$
$$|\mathrm{supp}\, x_0| \simeq m/2^l,\ \|x_0\|_\infty \le 1$$

and x_i comes from a 2^{-i}–net in the set of sparse vectors of support at most $m/2^i$.

Sketch of the proof

A modification of Bourgain's approach. One approximates an arbitrary vector z with $|\mathrm{supp}\, z| \le m$ by $x_0 + x_1 + \ldots + x_l$ ($l < \log_2 m$), where

$$|\mathrm{supp}\, x_i| \simeq m/2^i,\ \|x_i\|_\infty \simeq \sqrt{2^i/m},\ i \ge 1$$
$$|\mathrm{supp}\, x_0| \simeq m/2^l,\ \|x_0\|_\infty \le 1$$

and x_i comes from a 2^{-i}–net in the set of sparse vectors of support at most $m/2^i$.

Then using the ψ_1 condition one shows that with high probability

$$A_m^2 \lesssim \max_i |X_i|^2 + A_m(\sqrt{n} + \sqrt{m}\log(2N/m)).$$

Sketch of the proof

A modification of Bourgain's approach. One approximates an arbitrary vector z with $|\mathrm{supp}\, z| \leq m$ by $x_0 + x_1 + \ldots + x_l$ ($l < \log_2 m$), where

$$|\mathrm{supp}\, x_i| \simeq m/2^i, \ \|x_i\|_\infty \simeq \sqrt{2^i/m}, \ i \geq 1$$
$$|\mathrm{supp}\, x_0| \simeq m/2^l, \ \|x_0\|_\infty \leq 1$$

and x_i comes from a 2^{-i}–net in the set of sparse vectors of support at most $m/2^i$.

Then using the ψ_1 condition one shows that with high probability

$$A_m^2 \lesssim \max_i |X_i|^2 + A_m(\sqrt{n} + \sqrt{m}\log(2N/m)).$$

Theorem (G. Paouris)

$$\mathbb{P}(|X_i| \geq Ct\sqrt{n}) \leq \exp(-ct\sqrt{n})$$

Sketch of the proof

A modification of Bourgain's approach. One approximates an arbitrary vector z with $|\mathrm{supp}\ z| \le m$ by $x_0 + x_1 + \ldots + x_l$ ($l < \log_2 m$), where

$$|\mathrm{supp}\ x_i| \simeq m/2^i,\ \|x_i\|_\infty \simeq \sqrt{2^i/m},\ i \ge 1$$
$$|\mathrm{supp}\ x_0| \simeq m/2^l,\ \|x_0\|_\infty \le 1$$

and x_i comes from a 2^{-i}–net in the set of sparse vectors of support at most $m/2^i$.

Then using the ψ_1 condition one shows that with high probability

$$A_m^2 \lesssim \max_i |X_i|^2 + A_m(\sqrt{n} + \sqrt{m}\log(2N/m)).$$

Theorem (G. Paouris)

$$\mathbb{P}(|X_i| \ge Ct\sqrt{n}) \le \exp(-ct\sqrt{n})$$

Thus $\max_i |X_i| \le C\sqrt{n}$ with high probability and we can solve the inequality for A_m.

Compressed sensing and neighbourly polytopes

Imagine we have a vector $x \in \mathbb{R}^N$ (N large), which is supported on a small number of coordinates (say $|\mathrm{supp}\ \mathrm{x}| = m << N$).

Compressed sensing and neighbourly polytopes

Imagine we have a vector $x \in \mathbb{R}^N$ (N large), which is supported on a small number of coordinates (say $|\mathrm{supp}\ \mathrm{x}| = m << N$).

If we knew the support of x, to determine x it would be enough to take m measurements along basis vectors.

Compressed sensing and neighbourly polytopes

Imagine we have a vector $x \in \mathbb{R}^N$ (N large), which is supported on a small number of coordinates (say $|\mathrm{supp\ x}| = m << N$).

If we knew the support of x, to determine x it would be enough to take m measurements along basis vectors.

What if we don't know the support?

Compressed sensing and neighbourly polytopes

Imagine we have a vector $x \in \mathbb{R}^N$ (N large), which is supported on a small number of coordinates (say $|\mathrm{supp}\ \mathrm{x}| = m << N$).

If we knew the support of x, to determine x it would be enough to take m measurements along basis vectors.

What if we don't know the support?

Answer (Donoho, Candes, Tao, Romberg) Take measurements in random directions $Y_1, \ldots, Y_n$ and set

$$\hat{x} = \mathrm{argmin}\ \{\|y\|_1 : \langle Y_i, y\rangle = \langle Y_i, x\rangle\}$$

Compressed sensing and neighbourly polytopes

Definition

*A polytope $K \subseteq \mathbb{R}^n$ is called m-**neighbourly** if any set of vertices of K of cardinality at most $m+1$ is the vertex set of a face.*

Compressed sensing and neighbourly polytopes

Definition

A (centraly symetric) polytope $K \subseteq \mathbb{R}^n$ is called m**-(symmetric)-neighbourly** *if any set of vertices of K of cardinality at most $m+1$ (containing no opposite pairs) is the vertex set of a face.*

Compressed sensing and neighbourly polytopes

Definition

A (centraly symetric) polytope $K \subseteq \mathbb{R}^n$ is called m**-(symmetric)-neighbourly** *if any set of vertices of K of cardinality at most $m+1$ (containing no opposite pairs) is the vertex set of a face.*

Theorem (Donoho)

Let Γ be an $n \times N$ matrix with columns $X_1, \ldots, X_N$. The following conditions are equivalent

(i) *For any $x \in \mathbb{R}^N$ with $|\mathrm{supp}\, x| \le m$, x is the unique solution of the minimization problem*

$$\min \|t\|_1, \quad \Gamma t = \Gamma x.$$

(ii) *The polytope $K(\Gamma) = \mathrm{conv}(\pm X_1, \ldots, \pm X_N)$ has $2N$ vertices and is m-symmetric-neighbourly.*

Compressed sensing and neighbourly polytopes

Definition (Restricted Isometry Property (Candes, Tao))

For an $n \times N$ matrix Γ define the **isometry constant** $\delta_m = \delta_m(\Gamma)$ as the smallest number such that

$$(1 - \delta_m)|x|^2 \leq |\Gamma x|^2 \leq (1 + \delta_m)|x|^2$$

for all m-sparse vectors $x \in \mathbb{R}^N$.

Compressed sensing and neighbourly polytopes

Definition (Restricted Isometry Property (Candes, Tao))

For an $n \times N$ matrix Γ define the **isometry constant** $\delta_m = \delta_m(\Gamma)$ as the smallest number such that

$$(1-\delta_m)|x|^2 \le |\Gamma x|^2 \le (1+\delta_m)|x|^2$$

for all m-sparse vectors $x \in \mathbb{R}^N$.

Theorem (Candes)

If $\delta_{2m}(\Gamma) < \sqrt{2} - 1$ then for every m-sparse $x \in \mathbb{R}^n$, x is the unique solution to

$$\min \|t\|_1, \quad \Gamma t = \Gamma x.$$

In consequence, the polytope $K(\Gamma)$ (resp. $K_+(\Gamma) = \mathrm{conv}(X_1, \ldots, X_N)$) is m-symmetric-neighbourly (resp. m-neighbourly)

The following matrices satisfy RIP

- Gaussian matrices (Candes, Tao), $m \simeq n/\log(2N/n)$
- Matrices with rows selected randomly from the Fourier matrix (Candes & Tao, Rudelson & Vershynin), $m \simeq n/\log^4(N)$
- Matrices with independent subgaussian isotropic rows (Mendelson, Pajor, Tomczak-Jaegermann), $m \simeq n/\log(2N/n)$
- Matrices with independent log-concave isotropic columns (LPTA), $m \simeq n/\log^2(2N/n)$

Neighbourly polytopes

Theorem (LPTA)

Assume that X_i's are ψ_r. Let $\theta \in (0, 1/4)$ and assume that $N \leq \exp(c\theta^C n^c)$ and $m \log^{2/r}\left(\frac{2N}{\theta m}\right) \leq \theta^2 n$. Then, with probability at least $1 - \exp(-c\theta^C n^c)$

$$\delta_m\Big(\frac{1}{\sqrt{n}}\Gamma\Big) \leq \theta.$$

Corollary (LPTA)

Let $X_1, \ldots, X_N$ be random vectors drawn from an isotropic ψ_r ($r \in [1, 2]$) convex body in $\mathbb{R}^n$. Then, for $N \leq \exp(cn^c)$, with probability at least $1 - \exp(-cn^c)$, the polytope $K(\Gamma)$ (resp. $K_+(\Gamma)$) is m-symmetric-neighbourly (resp. m-neighbourly) with

$$m = \lfloor c \frac{n}{\log^{2/r}(CN/n)} \rfloor.$$

Method of proof

We use the same approximation techniques as for the KLS problem to bound

$$B_m = \sup_{|\mathrm{supp}\ z|\leq m,\ |z|=1} \left| \Big| \sum_{i\leq N} z_i X_i \Big|^2 - \sum_{i\leq N} z_i^2 |X_i|^2 \right|^{1/2}$$

Theorem (B. Klartag)

$$\mathbb{P}\left(\max_{i\leq N} \left| \frac{|X_i|^2}{n} - 1 \right| \geq \varepsilon \right) \leq C \exp(-c\varepsilon^C n^c).$$

Thus

$$\delta_n(n^{-1/2}\Gamma) \leq n^{-1} B_m^2 + \varepsilon$$

with overwhelming probability.

Smallest singular value

Definition

For an $n \times n$ matrix Γ let $s_1(\Gamma) \geq s_2(\Gamma) \geq \ldots \geq s_n(\Gamma)$ be the singular values of Γ, i.e. eigenvalues of $\sqrt{\Gamma\Gamma^T}$. In particular

$$s_1(\Gamma) = \|A\|, \; s_n(\Gamma) = \inf_{x \in S^{n-1}} |\Gamma x| = \frac{1}{\|A^{-1}\|}$$

Theorem (Edelman, Szarek)

Let Γ be an $n \times n$ random matrix with independent $\mathcal{N}(0,1)$ entries. Let s_n denote the smallest singular values of Γ. Then, for every $\varepsilon > 0$,

$$\mathbb{P}(s_n(\Gamma) \leq \varepsilon n^{-1/2}) \leq C\varepsilon,$$

where C is a universal constant.

Theorem (Rudelson, Vershynin)

Let Γ be a random matrix with independent entries X_{ij}, satisfying $\mathbb{E}X_{ij} = 0$, $\mathbb{E}X_{ij}^2 = 1$, $\|X_{ij}\|_{\psi_2} \leq B$. Then for any $\varepsilon \in (0,1)$,

$$\mathbb{P}(s_n(\Gamma) \leq \varepsilon n^{-1/2}) \leq C\varepsilon + c^n,$$

where $C > 0, c \in (0,1)$ depend only on B.

Theorem (Guédon, Litvak, Pajor, Tomczak-Jaegermann, R.A.)

Let Γ be an $n \times n$ random matrix with independent isotropic log-concave rows. Then, for any $\varepsilon \in (0,1)$,

$$\mathbb{P}(s_n(\Gamma) \leq \varepsilon n^{-1/2}) \leq C\varepsilon + C\exp(-cn^c)$$

and

$$\mathbb{P}(s_n(\Gamma) \leq \varepsilon n^{-1/2}) \leq C\varepsilon^{n/(n+2)} \log^C(2/\varepsilon).$$

Corollary

For any $\delta \in (0,1)$ *there exists* C_δ *such that for any* n *and* $\varepsilon \in (0,1)$*,*

$$\mathbb{P}(s_n(\Gamma) \le \varepsilon n^{-1/2}) \le C_\delta \varepsilon^{1-\delta}.$$

Definition

For an $n \times n$ matrix Γ define the **condition number** $\kappa(\Gamma)$ as

$$\kappa(\Gamma) = \|\Gamma\| \cdot \|\Gamma^{-1}\| = \frac{s_1(\Gamma)}{s_n(\Gamma)}.$$

Corollary

If Γ *has independent isotropic log-concave columns, then for any* $\delta > 0$, $t > 0$*,*

$$\mathbb{P}(\kappa(\Gamma) \ge nt) \le \frac{C_\delta}{t^{1-\delta}}.$$

Thank you