Invertibility of random matrices

Mark Rudelson[1] Roman Vershynin[2]

[1]Department of Mathematics
University of Missouri

[2]Department of Mathematics
University of California, Davis

Phenomena in High Dimensions, Samos 2007

Invertibility problems

Let A be an $n \times n$ random matrix with independent entries.

Examples

1. Gaussian matrix: the entries are $N(0,1)$ normal random variables
2. random ± 1 matrix: the entries are Bernoulli random variables

Invertibility problems

Let A be an $n \times n$ random matrix with independent entries.

Examples

1. Gaussian matrix: the entries are $N(0,1)$ normal random variables
2. random ± 1 matrix: the entries are Bernoulli random variables

- Qualitative problem:
  - What is the probability that a random matrix A is non-singular?
- Quantitative problems:
  - What is the typical distance from a random matrix to the set of singular matrices?
  - What is the tail distribution of this distance?

Qualitative problem

Let A be an $n \times n$ random ± 1 matrix. Probability of non-singularity:

$$P_n = \mathbb{P}\left(\det(A) \neq 0\right).$$

Qualitative problem

Let A be an $n \times n$ random ± 1 matrix. Probability of non-singularity:

$$P_n = \mathbb{P}\left(\det(A) \neq 0\right).$$

- Komlos (1967): $P_n \geq 1 - c/\sqrt{n}$.

Qualitative problem

Let A be an $n \times n$ random ± 1 matrix. Probability of non-singularity:

$$P_n = \mathbb{P}\left(\det(A) \neq 0\right).$$

- Komlos (1967): $P_n \geq 1 - c/\sqrt{n}$.
- Kahn, Komlos, Szemeredy (1994): $P_n \geq 1 - 0.998^n$.
- Tao, Vu (2004): $P_n \geq 1 - 0.96^n$.
- Tao, Vu (2005): $P_n \geq 1 - (3/4)^n$.

Qualitative problem

Let A be an $n \times n$ random ± 1 matrix. Probability of non-singularity:

$$P_n = \mathbb{P}\left(\det(A) \neq 0\right).$$

- Komlos (1967): $P_n \geq 1 - c/\sqrt{n}$.
- Kahn, Komlos, Szemeredy (1994): $P_n \geq 1 - 0.998^n$.
- Tao, Vu (2004): $P_n \geq 1 - 0.96^n$.
- Tao, Vu (2005): $P_n \geq 1 - (3/4)^n$.

Conjecture: $P_n = 1 - \binom{n}{2} \cdot 2^{-(n-2)} \cdot (1 + o(1))$.

The main reason for singularity is that two rows or two columns of A are equal up to a sign.

Condition number of a matrix

Definition

Distortion (condition number):

$$D(A) = \sup_{x,y \in S^{n-1}} \frac{\|Ax\|}{\|Ay\|}$$

Condition number of a matrix

Definition

Distortion (condition number):

$$D(A) = \sup_{x,y \in S^{n-1}} \frac{\|Ax\|}{\|Ay\|} = \frac{s_1(A)}{s_n(A)} = \|A\| \cdot \left\|A^{-1}\right\|.$$

Here $s_1(A) \geq s_2(A) \geq \ldots \geq s_n(A) \geq 0$ are the singular values of A (the eigenvalues of $(A^*A)^{1/2}$):

$$s_1(A) = \left\|A : \mathbb{R}^n \to \mathbb{R}^N\right\|, \qquad s_n(A) = \left(\left\|A^{-1} : A\mathbb{R}^n \to \mathbb{R}^n\right\|\right)^{-1}.$$

We need an upper estimate for the first singular number and a lower estimate for the last one.

Applications of the condition number

Definition

Distortion (condition number):

$$D(A) = \sup_{x,y \in S^{n-1}} \frac{\|Ax\|}{\|Ay\|}.$$

- Error control.
  - Gaussian elimination for the system $Ax = b$.
- Rate of convergence.
  - Iteration methods for linear systems (conjugate gradients, Kaczmarz-Steinhaus algorithm)
  - Linear programming (smoothed analysis)

The first singular value

Assume that $\mathbb{E}a_{j,k} = 0$ and $\mathbb{E}|a_{j,k}|^2 \geq c$.

Theorem (Bai, Krishnaiah, Silverstein, Yin)

Let A_n be a family of $n \times n$ random matrices with i.i.d. entries

$$\lim_{n \to \infty} s_1(A_n)/\sqrt{n}$$

exists a.s. if and only if the fourth moment of the entries is finite.

The first singular value

Assume that $\mathbb{E}a_{j,k} = 0$ and $\mathbb{E}|a_{j,k}|^2 \geq c$.

Theorem (Bai, Krishnaiah, Silverstein, Yin)

Let A_n be a family of $n \times n$ random matrices with i.i.d. entries

$$\lim_{n \to \infty} s_1(A_n)/\sqrt{n}$$

exists a.s. if and only if the fourth moment of the entries is finite.

Theorem (Latała)

Let A be an $n \times n$ random matrix, whose entries have a uniformly bounded fourth moment. Then

$$\mathbb{E}s_1(A) \leq C\sqrt{n}.$$

The first singular value

Theorem (Large deviations)

Let A be an $n \times n$ random matrix with subgaussian entries. Then for any $t > t_0$

$$\mathbb{P}\left(s_1(A) \geq t\sqrt{n}\right) \leq e^{-cnt^2}.$$

Theorem (Concentration)

Let A be an $n \times n$ random matrix with bounded entries. Then for any $t > 0$

$$\mathbb{P}\left(\left|s_1(A) - \mathbb{M}(s_1(A))\right| \geq t\right) \leq 4e^{-t^2/4}.$$

$s_1(A)$ is a convex function of A.

Quantitative problems: the median

The last singular value $s_n(A) = \min_{x \in S^{n-1}} \|Ax\|$
is the distance of A to the set of singular matrices.

Conjecture (von Neumann, Smale):

$s_n(A) \sim n^{-1/2}$ with probability close to 1.

Quantitative problems: the median

The last singular value $s_n(A) = \min_{x \in S^{n-1}} \|Ax\|$
is the distance of A to the set of singular matrices.

Conjecture (von Neumann, Smale):

$s_n(A) \sim n^{-1/2}$ with probability close to 1.

Theorem (Edelman, 1988)

Let A be an $n \times n$ Gaussian matrix. Then for any $\varepsilon > 0$

$$\mathbb{P}\left(s_n(A) \le \varepsilon n^{-1/2}\right) \sim \varepsilon.$$

Quantitative problems: the median

The last singular value $s_n(A) = \min_{x \in S^{n-1}} \|Ax\|$
is the distance of A to the set of singular matrices.

Conjecture (von Neumann, Smale):

$s_n(A) \sim n^{-1/2}$ with probability close to 1.

Theorem (Edelman, 1988)

Let A be an $n \times n$ Gaussian matrix. Then for any $\varepsilon > 0$

$$\mathbb{P}\left(s_n(A) \leq \varepsilon n^{-1/2}\right) \sim \varepsilon.$$

Theorem (R, 2005)

Let A be an $n \times n$ matrix with i.i.d. subgaussian entries. Then for any $\varepsilon \geq cn^{-1/2}$

$$\mathbb{P}\left(s_n(A) \leq C\varepsilon \cdot n^{-3/2}\right) \leq \varepsilon.$$

Quantitative problems: the tail distribution

Theorem (Edelman, 1988)

Let A be an $n \times n$ Gaussian matrix. Then for any $\varepsilon > 0$

$$\mathbb{P}\left(s_n(A) \le \varepsilon \cdot n^{-1/2}\right) \sim \varepsilon.$$

Conjecture (Spielman, Teng):

Let A be an $n \times n$ random ± 1 matrix. Then for any $\varepsilon > 0$

$$\mathbb{P}\left(s_n(A) \le \varepsilon \cdot n^{-1/2}\right) \le \varepsilon + e^{-cn}.$$

Quantitative problems: the tail distribution

Theorem (Edelman, 1988)

Let A be an $n \times n$ Gaussian matrix. Then for any $\varepsilon > 0$

$$\mathbb{P}\left(s_n(A) \le \varepsilon \cdot n^{-1/2}\right) \sim \varepsilon.$$

Conjecture (Spielman, Teng):

Let A be an $n \times n$ random ± 1 matrix. Then for any $\varepsilon > 0$

$$\mathbb{P}\left(s_n(A) \le \varepsilon \cdot n^{-1/2}\right) \le \varepsilon + e^{-cn}.$$

Theorem (Tao, Vu, 2005)

Let A be an $n \times n$ random ± 1 matrix. Then for any $\alpha > 0$ there exists $\beta > 0$ such that

$$\mathbb{P}\left(s_n(A) \le n^{-\beta}\right) \le n^{-\alpha}.$$

Quantitative problems: the results

Conjecture (von Neumann, Smale):

$s_n(A) \sim n^{-1/2}$ with probability close to 1.

Quantitative problems: the results

Conjecture (von Neumann, Smale):

$s_n(A) \sim n^{-1/2}$ with probability close to 1.

Theorem (Median)

Let A be an $n \times n$ random matrix, whose entries have a uniformly bounded fourth moment. Then for any $\varepsilon > 0$ there exists $\delta > 0$, such that

$$\mathbb{P}\left(s_n(A) \leq \delta \cdot n^{-1/2}\right) \leq \varepsilon$$

for all $n \geq n_0(\varepsilon)$.

Quantitative problems: the results

Definition

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

$$\mathbb{P}\left(|\xi| > t\right) \leq Ce^{-ct^2}.$$

Examples

- Gaussian random variable
- Any bounded random variable (including random ± 1)

Therefore, $\mathbb{P}\left(\det(A) = 0\right) \leq e^{-cn}$.

Quantitative problems: the results

Definition

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

$$\mathbb{P}\left(|\xi| > t\right) \le Ce^{-ct^2}.$$

Examples

- Gaussian random variable
- Any bounded random variable (including random ± 1)

Theorem (Tail distribution)

Let A be an $n \times n$ random matrix with i.i.d. subgaussian entries. Then for any $\varepsilon > 0$

$$\mathbb{P}\left(s_n(A) \le c'\varepsilon \cdot n^{-1/2}\right) \le \varepsilon + e^{-cn}.$$

Therefore, $\mathbb{P}\left(\det(A) = 0\right) \le e^{-cn}$.

Compressible and incompressible vectors

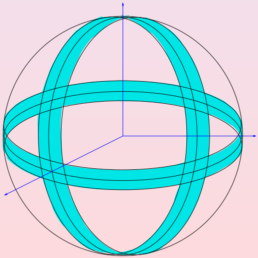
Assume that A is a random matrix with i.i.d. subgaussian entries.

Definition

A vector $x \in S^{n-1}$ is called compressible if it is close to a sparse vector in the ℓ_2-norm.

We decompose the sphere in two parts: $S^{n-1} = Comp \cup Incomp$.

- Compressible vectors: the norm is concentrated on a few coordinates
- Incompressible vectors: many coordinates of the order $n^{-1/2}$.



Invertibility for compressible vectors

Lemma (Litvak, Pajor, R', Tomczak-Jaegermann)

$$\mathbb{P}\left(\inf_{x \in Comp} \|Ax\|_2 \leq cn^{1/2} \right) \leq e^{-c'n}.$$

This bound is much stronger than we need.

Invertibility for compressible vectors

Lemma (Litvak, Pajor, R', Tomczak-Jaegermann)

$$\mathbb{P}\left(\inf_{x \in Comp} \|Ax\|_2 \leq cn^{1/2} \right) \leq e^{-c'n}.$$

This bound is much stronger than we need.

Proof.

1. Individual estimate: let $x \in S^{n-1}$ be any vector. Then
$$\mathbb{P}\left(\|Ax\|_2 \leq cn^{1/2} \right) \leq e^{-c'n}.$$
2. The set of sparse vectors admits a small ε-net.
3. This ε-net is a 2ε-net for the set of compressible vectors
4. Approximation. □

Invertibility via distance

Lemma

Let $X_1, \dots, X_n$ denote the column vectors of A, and let H_k denote the span of all column vectors except the k-th. Then for every $\varepsilon > 0$, one has

$$\mathbb{P}\left(\inf_{x \in Incomp} \|Ax\|_2 < c\varepsilon n^{-1/2} \right) \le C\, \mathbb{P}\left(\mathrm{dist}(X_n, H_n) < \varepsilon \right).$$

Invertibility via distance

Lemma

Let $X_1, \ldots, X_n$ denote the column vectors of A, and let H_k denote the span of all column vectors except the k-th. Then for every $\varepsilon > 0$, one has

$$\mathbb{P}\left(\inf_{x \in Incomp} \|Ax\|_2 < c\varepsilon n^{-1/2}\right) \leq C\,\mathbb{P}\left(\mathrm{dist}(X_n, H_n) < \varepsilon\right).$$

Proof.

Writing $Ax = \sum_{k=1}^n x_k X_k$, we have

$$\|Ax\|_2 \geq \max_{k=1,\ldots,n} \mathrm{dist}(Ax, H_k) = \max_{k=1,\ldots,n} |x_k|\,\mathrm{dist}(X_k, H_k).$$

Invertibility via distance

Lemma

Let $X_1, \dots, X_n$ denote the column vectors of A, and let H_k denote the span of all column vectors except the k-th. Then for every $\varepsilon > 0$, one has

$$\mathbb{P}\left(\inf_{x \in Incomp} \|Ax\|_2 < c\varepsilon n^{-1/2} \right) \le C\, \mathbb{P}\left(\operatorname{dist}(X_n, H_n) < \varepsilon \right).$$

Proof.

Writing $Ax = \sum_{k=1}^n x_k X_k$, we have

$$\|Ax\|_2 \ge \max_{k=1,\dots,n} \operatorname{dist}(Ax, H_k) = \max_{k=1,\dots,n} |x_k| \operatorname{dist}(X_k, H_k).$$

Let $x \in Incomp$. Then $|x_k| \sim n^{-1/2}$ for at least cn indices k. Hence,

$$\|Ax\|_2 < c\varepsilon n^{-1/2} \Rightarrow \operatorname{dist}(X_k, H_k) < \varepsilon \text{ for at least } cn \text{ indices } k$$

Invertibility via distance

Lemma

Let $X_1, \dots, X_n$ denote the column vectors of A, and let H_k denote the span of all column vectors except the k-th. Then for every $\varepsilon > 0$, one has

$$\mathbb{P}\left(\inf_{x \in Incomp} \|Ax\|_2 < c\varepsilon n^{-1/2} \right) \leq C\,\mathbb{P}\left(\mathrm{dist}(X_n, H_n) < \varepsilon\right).$$

Proof.

Writing $Ax = \sum_{k=1}^{n} x_k X_k$, we have

$$\|Ax\|_2 \geq \max_{k=1,\dots,n} \mathrm{dist}(Ax, H_k) = \max_{k=1,\dots,n} |x_k|\,\mathrm{dist}(X_k, H_k).$$

Let $x \in Incomp$. Then $|x_k| \sim n^{-1/2}$ for at least cn indices k. Hence,

$\exists x \in Incomp$ $\|Ax\|_2 < c\varepsilon n^{-1/2} \Rightarrow \mathrm{dist}(X_k, H_k) < \varepsilon$ for at least cn indices k

Invertibility via distance

Lemma

Let $X_1, \dots, X_n$ denote the column vectors of A, and let H_k denote the span of all column vectors except the k-th. Then for every $\varepsilon > 0$, one has

$$\mathbb{P}\left(\inf_{x \in Incomp} \|Ax\|_2 < c\varepsilon n^{-1/2} \right) \le C\, \mathbb{P}\left(\mathrm{dist}(X_n, H_n) < \varepsilon \right).$$

Proof.

$\exists x \in Incomp \ \|Ax\|_2 < c\varepsilon n^{-1/2} \Rightarrow \ \mathrm{dist}(X_k, H_k) < \varepsilon$ for at least cn indices k

Invertibility via distance

Lemma

Let $X_1, \ldots, X_n$ denote the column vectors of A, and let H_k denote the span of all column vectors except the k-th. Then for every $\varepsilon > 0$, one has

$$\mathbb{P}\left(\inf_{x \in Incomp} \|Ax\|_2 < c\varepsilon n^{-1/2}\right) \leq C\,\mathbb{P}\left(\operatorname{dist}(X_n, H_n) < \varepsilon\right).$$

Proof.

$\exists x \in Incomp\ \|Ax\|_2 < c\varepsilon n^{-1/2} \Rightarrow \operatorname{dist}(X_k, H_k) < \varepsilon$ for at least cn indices k

Denote $p := \mathbb{P}\left(\operatorname{dist}(X_k, H_k) < \varepsilon\right)$. Then $\mathbb{E}\left|\{k : \operatorname{dist}(X_k, H_k) < \varepsilon\}\right| = np$.

Invertibility via distance

Lemma

Let $X_1, \ldots, X_n$ denote the column vectors of A, and let H_k denote the span of all column vectors except the k-th. Then for every $\varepsilon > 0$, one has

$$\mathbb{P}\left(\inf_{x \in Incomp} \|Ax\|_2 < c\varepsilon n^{-1/2}\right) \le C\,\mathbb{P}\left(\mathrm{dist}(X_n, H_n) < \varepsilon\right).$$

Proof.

$\exists\, x \in Incomp\ \|Ax\|_2 < c\varepsilon n^{-1/2} \Rightarrow\ \mathrm{dist}(X_k, H_k) < \varepsilon$ for at least cn indices k

Denote $p := \mathbb{P}\left(\mathrm{dist}(X_k, H_k) < \varepsilon\right)$. Then $\mathbb{E}\left|\{k :\ \mathrm{dist}(X_k, H_k) < \varepsilon\}\right| = np$.
Therefore, by Chebychev's inequality,

$$\mathbb{P}\left(\mathrm{dist}(X_k, H_k) < \varepsilon \text{ for at least } cn \text{ indices } k\right) \le \frac{np}{nc}.$$

Invertibility via distance

Lemma

Let $X_1, \dots, X_n$ denote the column vectors of A, and let H_k denote the span of all column vectors except the k-th. Then for every $\varepsilon > 0$, one has

$$\mathbb{P}\left(\inf_{x \in Incomp} \|Ax\|_2 < c\varepsilon n^{-1/2}\right) \le C\,\mathbb{P}\left(\operatorname{dist}(X_n, H_n) < \varepsilon\right).$$

Proof.

$\exists x \in Incomp \ \|Ax\|_2 < c\varepsilon n^{-1/2} \Rightarrow \operatorname{dist}(X_k, H_k) < \varepsilon$ for at least cn indices k

Denote $p := \mathbb{P}\left(\operatorname{dist}(X_k, H_k) < \varepsilon\right)$. Then $\mathbb{E}\left|\{k : \operatorname{dist}(X_k, H_k) < \varepsilon\}\right| = np$.
Therefore, by Chebychev's inequality,

$$\mathbb{P}\left(\operatorname{dist}(X_k, H_k) < \varepsilon \text{ for at least } cn \text{ indices } k\right) \le (1/c)\,\mathbb{P}\left(\operatorname{dist}(X_k, H_k) < \varepsilon\right)$$

□

Random normal

$$\mathbb{P}\left(\mathrm{dist}(X_n, H_n) < \varepsilon\right) \leq ?$$

Random normal

$$\mathbb{P}\left(\operatorname{dist}(X_n, H_n) < \varepsilon\right) \le ?$$

Let $X^* =: (a_1, \ldots, a_n)$ be any unit vector orthogonal to $X_1, \ldots, X_{n-1}$. We can choose X^* so that it depends only on $X_1, \ldots, X_{n-1}$ and is independent of $X_n = (\xi_1, \ldots, \xi_n)$.

Random normal

$$\mathbb{P}\left(\mathrm{dist}(X_n, H_n) < \varepsilon\right) \leq ?$$

Let $X^* =: (a_1, \dots, a_n)$ be any unit vector orthogonal to $X_1, \dots, X_{n-1}$. We can choose X^* so that it depends only on $X_1, \dots, X_{n-1}$ and is independent of $X_n = (\xi_1, \dots, \xi_n)$.

$$\mathrm{dist}(X_n, H_n) \geq |\langle X^*, X_n \rangle|.$$

We use the small ball probability estimates to bound

$$\mathbb{P}\left(|\langle X^*, X_n \rangle| < \varepsilon\right) = \mathbb{P}\left(|\sum_{k=1}^{n} a_k \xi_k| < \varepsilon\right).$$

Small ball probability

$$P_a(\varepsilon) := \mathbb{P}\left(\left|\sum_{k=1}^{n} a_k \xi_k\right| < \varepsilon \right) \leq ?$$

Small ball probability

$$P_a(\varepsilon) := \mathbb{P}\left(\left| \sum_{k=1}^{n} a_k \xi_k \right| < \varepsilon \right) \leq ?$$

Examples (Let $\xi_1, \ldots, \xi_n$ be Bernoulli random variables.)

- Compressible vector: if $a = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, \ldots, 0)$, then $P_a(\varepsilon) = 1/2$.
- Incompressible vector: if $a = (\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, , \ldots, \frac{1}{\sqrt{n}})$, then $P_a(\varepsilon) \sim t + 1/\sqrt{n}$.

We have to treat compressible and incompressible normals separately.

Random normal is incompressible

Lemma

$$\mathbb{P}\left(X^* \in Comp\right) \leq e^{-cn}.$$

Proof.

Let A' be the $(n-1) \times n$ random matrix with rows $X_1, \ldots, X_n$.
By the definition of the random normal,

$$A'X^* = 0.$$

Therefore, if $X^* \in Comp$ then $\inf_{x \in Comp} \|A'x\|_2 = 0.$

Random normal is incompressible

Lemma

$$\mathbb{P}\left(X^* \in Comp\right) \leq e^{-cn}.$$

Proof.

Let A' be the $(n-1) \times n$ random matrix with rows $X_1, \ldots, X_n$.
By the definition of the random normal,

$$A'X^* = 0.$$

Therefore, if $X^* \in Comp$ then $\inf_{x \in Comp} \|A'x\|_2 = 0.$ □

Lemma

$$\mathbb{P}\left(\inf_{x \in Comp} \|A'x\|_2 \leq cn^{1/2}\right) \leq e^{-c'n}.$$

Small ball probability for incompressible vectors

$$P_a(\varepsilon) := \mathbb{P}\left(\left|\sum_{k=1}^{n} a_k \xi_k\right| < \varepsilon \right) \leq ?$$

Lemma (CLT bound)

Let $a \in S^{n-1}$ be an incompressible vector. Then for every $\varepsilon > 0$

$$P_a(\varepsilon) \leq C(\varepsilon + n^{-1/2}),$$

Small ball probability for incompressible vectors

$$P_a(\varepsilon) := \mathbb{P}\left(\left|\sum_{k=1}^{n} a_k \xi_k\right| < \varepsilon \right) \leq ?$$

Lemma (CLT bound)

Let $a \in S^{n-1}$ be an incompressible vector. Then for every $\varepsilon > 0$

$$P_a(\varepsilon) \leq C(\varepsilon + n^{-1/2}),$$

Proof.

1. An incompressible vector has at least cn coordinates of the order $n^{-1/2}$.
2. Condition on the other coordinates and apply Berry–Esseen Theorem: $\mathbb{P}\left(\left|\sum_{k=1}^{n} a_k \xi_k\right| < \varepsilon\right) \leq \mathbb{P}\left(|\gamma| < \varepsilon\right) + Cn^{-1/2} \leq C(\varepsilon + n^{-1/2})$. □

Polynomial bound

Theorem (Polynomial bound)

Let A be an $n \times n$ random matrix, with i.i.d. subgaussian entries. Then for any $\varepsilon > 0$, such that

$$\mathbb{P}\left(s_n(A) \le c\varepsilon \cdot n^{-1/2}\right) \le \varepsilon + cn^{-1/2}.$$

Polynomial bound

Theorem (Polynomial bound)

Let A be an $n \times n$ random matrix, with i.i.d. subgaussian entries. Then for any $\varepsilon > 0$, such that

$$\mathbb{P}\left(s_n(A) \leq c\varepsilon \cdot n^{-1/2}\right) \leq \varepsilon + cn^{-1/2}.$$

Proof.

1. $$\mathbb{P}\left(s_n(A) \leq c\varepsilon \cdot n^{-1/2}\right) \leq \mathbb{P}\left(\inf_{x \in Comp} \|Ax\|_2 < \varepsilon \cdot n^{-1/2}\right) + \mathbb{P}\left(\inf_{x \in Incomp} \|Ax\|_2 < \varepsilon \cdot n^{-1/2}\right)$$

Polynomial bound

Theorem (Polynomial bound)

Let A be an $n \times n$ random matrix, with i.i.d. subgaussian entries. Then for any $\varepsilon > 0$, such that

$$\mathbb{P}\left(s_n(A) \le c\varepsilon \cdot n^{-1/2}\right) \le \varepsilon + cn^{-1/2}.$$

Proof.

1. $\mathbb{P}\left(s_n(A) \le c\varepsilon \cdot n^{-1/2}\right) \le e^{-cn}$

$$+\mathbb{P}\left(\inf_{x \in Incomp} \|Ax\|_2 < \varepsilon \cdot n^{-1/2}\right)$$

Polynomial bound

Theorem (Polynomial bound)

Let A be an $n \times n$ random matrix, with i.i.d. subgaussian entries. Then for any $\varepsilon > 0$, such that

$$\mathbb{P}\left(s_n(A) \le c\varepsilon \cdot n^{-1/2}\right) \le \varepsilon + cn^{-1/2}.$$

Proof.

1. $\mathbb{P}\left(s_n(A) \le c\varepsilon \cdot n^{-1/2}\right) \le e^{-cn}$
$$+\mathbb{P}\left(\inf_{x \in Incomp} \|Ax\|_2 < \varepsilon \cdot n^{-1/2}\right)$$
2. $\mathbb{P}\left(\inf_{x \in Incomp} \|Ax\|_2 < \varepsilon \cdot n^{-1/2}\right) \le \mathbb{P}\left(|\langle X_n, X^* \rangle| < \varepsilon\right)$

Polynomial bound

Theorem (Polynomial bound)

Let A be an $n \times n$ random matrix, with i.i.d. subgaussian entries. Then for any $\varepsilon > 0$, such that

$$\mathbb{P}\left(s_n(A) \le c\varepsilon \cdot n^{-1/2}\right) \le \varepsilon + cn^{-1/2}.$$

Proof.

1. $\mathbb{P}\left(s_n(A) \le c\varepsilon \cdot n^{-1/2}\right) \le e^{-cn} + \mathbb{P}\left(\inf_{x \in Incomp} \|Ax\|_2 < \varepsilon \cdot n^{-1/2}\right)$
2. $\mathbb{P}\left(\inf_{x \in Incomp} \|Ax\|_2 < \varepsilon \cdot n^{-1/2}\right) \le \mathbb{P}\left(|\langle X_n, X^* \rangle| < \varepsilon \text{ and } X^* \in Comp\right) + \mathbb{P}\left(|\langle X_n, X^* \rangle| < \varepsilon \text{ and } X^* \in Incomp\right)$

Polynomial bound

Theorem (Polynomial bound)

Let A be an $n \times n$ random matrix, with i.i.d. subgaussian entries. Then for any $\varepsilon > 0$, such that

$$\mathbb{P}\left(s_n(A) \le c\varepsilon \cdot n^{-1/2}\right) \le \varepsilon + cn^{-1/2}.$$

Proof.

1. $\mathbb{P}\left(s_n(A) \le c\varepsilon \cdot n^{-1/2}\right) \le e^{-cn}$
$$+\mathbb{P}\left(\inf_{x \in Incomp} \|Ax\|_2 < \varepsilon \cdot n^{-1/2}\right)$$
2. $\mathbb{P}\left(\inf_{x \in Incomp} \|Ax\|_2 < \varepsilon \cdot n^{-1/2}\right) \le e^{-cn}$
$$+\mathbb{P}\left(|\langle X_n, X^* \rangle| < \varepsilon \text{ and } X^* \in Incomp\right)$$

Polynomial bound

Theorem (Polynomial bound)

Let A be an $n \times n$ random matrix, with i.i.d. subgaussian entries. Then for any $\varepsilon > 0$, such that

$$\mathbb{P}\left(s_n(A) \le c\varepsilon \cdot n^{-1/2}\right) \le \varepsilon + cn^{-1/2}.$$

Proof.

1. $\mathbb{P}\left(s_n(A) \le c\varepsilon \cdot n^{-1/2}\right) \le e^{-cn}$ $+\mathbb{P}\left(\inf_{x \in Incomp} \|Ax\|_2 < \varepsilon \cdot n^{-1/2}\right)$
2. $\mathbb{P}\left(\inf_{x \in Incomp} \|Ax\|_2 < \varepsilon \cdot n^{-1/2}\right) \le e^{-cn}$ $+\mathbb{P}\left(|\langle X_n, X^* \rangle| < \varepsilon \text{ and } X^* \in Incomp\right)$
3. $\mathbb{P}\left(|\langle X_n, X^* \rangle| < \varepsilon \text{ and } X^* \in Incomp\right) \le C(\varepsilon + n^{-1/2})$. □

Stratification of the sphere

Theorem (Exponential bound)

Let A be an $n \times n$ random matrix, with i.i.d. subgaussian entries. Then for any $\varepsilon > 0$, such that

$$\mathbb{P}\left(s_n(A) \le c\varepsilon \cdot n^{-1/2}\right) \le \varepsilon + e^{-cn}.$$

Stratification of the sphere

Theorem (Exponential bound)

Let A be an $n \times n$ random matrix, with i.i.d. subgaussian entries. Then for any $\varepsilon > 0$, such that

$$\mathbb{P}\left(s_n(A) \leq c\varepsilon \cdot n^{-1/2}\right) \leq \varepsilon + e^{-cn}.$$

Recall that $S^{n-1} = Comp \cup Incomp$. We proved that

$$\mathbb{P}\left(s_n(A) \leq c\varepsilon \cdot n^{-1/2}\right) \leq e^{-cn} + \mathbb{P}\left(|\langle X_n, X^* \rangle| < \varepsilon \text{ and } X^* \in Incomp\right)$$

Stratification of the sphere

Theorem (Exponential bound)

Let A be an $n \times n$ random matrix, with i.i.d. subgaussian entries. Then for any $\varepsilon > 0$, such that

$$\mathbb{P}\left(s_n(A) \le c\varepsilon \cdot n^{-1/2}\right) \le \varepsilon + e^{-cn}.$$

Recall that $S^{n-1} = Comp \cup Incomp$. We proved that

$$\mathbb{P}\left(s_n(A) \le c\varepsilon \cdot n^{-1/2}\right) \le e^{-cn} + \mathbb{P}\left(|\langle X_n, X^* \rangle| < \varepsilon \text{ and } X^* \in Incomp\right)$$

We decompose *Incomp* further into the set of typical and atypical vectors: $Incomp = Typ \cup Atyp$.

$$\mathbb{P}\left(|\langle X_n, X^* \rangle| < \varepsilon \text{ and } X^* \in Incomp\right)$$
$$\le \mathbb{P}\left(|\langle X_n, X^* \rangle| < \varepsilon \text{ and } X^* \in Atyp\right) + \mathbb{P}\left(|\langle X_n, X^* \rangle| < \varepsilon \text{ and } X^* \in Typ\right)$$

Stratification of the sphere

Theorem (Exponential bound)

Let A be an $n \times n$ random matrix, with i.i.d. subgaussian entries. Then for any $\varepsilon > 0$, such that

$$\mathbb{P}\left(s_n(A) \le c\varepsilon \cdot n^{-1/2}\right) \le \varepsilon + e^{-cn}.$$

Recall that $S^{n-1} = Comp \cup Incomp$. We proved that

$$\mathbb{P}\left(s_n(A) \le c\varepsilon \cdot n^{-1/2}\right) \le e^{-cn} + \mathbb{P}\left(|\langle X_n, X^* \rangle| < \varepsilon \text{ and } X^* \in Incomp\right)$$

We decompose *Incomp* further into the set of typical and atypical vectors: $Incomp = Typ \cup Atyp$.

$$\begin{aligned}
&\mathbb{P}\left(|\langle X_n, X^* \rangle| < \varepsilon \text{ and } X^* \in Incomp\right) \\
&\le \mathbb{P}\left(X^* \in Atyp\right) + \mathbb{P}\left(|\langle X_n, X^* \rangle| < \varepsilon \text{ and } X^* \in Typ\right) \\
&\le e^{-cn} + C(\varepsilon + e^{-cn}).
\end{aligned}$$

Stratification by the LCD

Theorem (Small ball probability via the LCD)

Let $\xi_1, \ldots, \xi_n$ be i.i.d. subgaussian random variables. Then for any $a \in Incomp$ and for any $\varepsilon > 0$

$$P_a(\varepsilon) := \mathbb{P}\left(\left|\sum_{j=1}^{n} a_j \xi_j\right| < \varepsilon\right) \leq C\left(\varepsilon + \frac{1}{LCD(a)}\right) + Ce^{-cn}.$$

Stratification by the LCD

Theorem (Small ball probability via the LCD)

Let $\xi_1, \dots, \xi_n$ be i.i.d. subgaussian random variables. Then for any $a \in Incomp$ and for any $\varepsilon > 0$

$$P_a(\varepsilon) := \mathbb{P}\left(\Big| \sum_{j=1}^{n} a_j \xi_j \Big| < \varepsilon \right) \le C\left(\varepsilon + \frac{1}{LCD(a)} \right) + Ce^{-cn}.$$

- Typical vectors: LCD $> e^{cn} \;\Rightarrow\; P_a(\varepsilon) \le C\varepsilon + Ce^{-cn}$
- Atypical vectors: LCD $\le e^{cn}$

Stratification by the LCD

Theorem (Small ball probability via the LCD)

Let $\xi_1, \dots, \xi_n$ be i.i.d. subgaussian random variables. Then for any $a \in Incomp$ and for any $\varepsilon > 0$

$$P_a(\varepsilon) := \mathbb{P}\left(\Big|\sum_{j=1}^{n} a_j \xi_j\Big| < \varepsilon\right) \le C\left(\varepsilon + \frac{1}{LCD(a)}\right) + Ce^{-cn}.$$

- Typical vectors: LCD $> e^{cn} \;\Rightarrow\; P_a(\varepsilon) \le C\varepsilon + Ce^{-cn}$
- Atypical vectors: LCD $\le e^{cn}$

Examples (Atypical vectors)

- $a = (\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}})$ LCD$(a) = \sqrt{n}$
- $a = (\frac{1+1/n}{\sqrt{n}}, \frac{1+2/n}{\sqrt{n}}, \dots, \frac{1+n/n}{\sqrt{n}})$ LCD$(a) = n^{3/2}$

Stratification by the LCD

Theorem (Small ball probability via the LCD)

Let $\xi_1, \ldots, \xi_n$ be i.i.d. subgaussian random variables. Then for any $a \in$ Incomp and for any $\varepsilon > 0$

$$P_a(\varepsilon) := \mathbb{P}\left(\left| \sum_{j=1}^{n} a_j \xi_j \right| < \varepsilon \right) \le C\left(\varepsilon + \frac{1}{LCD(a)} \right) + Ce^{-cn}.$$

- Typical vectors: LCD $> e^{cn} \;\Rightarrow\; P_a(\varepsilon) \le C\varepsilon + Ce^{-cn}$
- Atypical vectors: LCD $\le e^{cn}$

Theorem (Random normal is typical)

Let $X_1, \ldots, X_{n-1}$ be vectors with i.i.d. subgaussian coordinates and let X^* be any unit vector orthogonal to $X_1, \ldots, X_{n-1}$. Then

$$\mathbb{P}\left(LCD(X^*) < e^{cn} \right) \le e^{-c'n}.$$

Random normal is typical

Theorem (Random normal is typical)

Let $X_1, \dots, X_{n-1}$ be vectors with i.i.d. subgaussian coordinates and let X^* be any unit vector orthogonal to $X_1, \dots, X_{n-1}$. Then

$$\mathbb{P}\left(LCD(X^*) < e^{cn}\right) \le e^{-c'n}.$$

Random normal is typical

Theorem (Random normal is typical)

Let $X_1, \ldots, X_{n-1}$ be vectors with i.i.d. subgaussian coordinates and let X^* be any unit vector orthogonal to $X_1, \ldots, X_{n-1}$. Then

$$\mathbb{P}\left(LCD(X^*) < e^{cn}\right) \leq e^{-c'n}.$$

Proof.

We further partition the set *Atyp* into the level sets according to the values of the LCD:

$$S_D := \{x \in Atyp : \; D \leq LCD(x) < 2D\}.$$

Here $D = 2^j$, where $j = 1, \ldots, cn$.

Random normal is typical

Theorem (Random normal is typical)

Let $X_1, \dots, X_{n-1}$ be vectors with i.i.d. subgaussian coordinates and let X^* be any unit vector orthogonal to $X_1, \dots, X_{n-1}$. Then

$$\mathbb{P}\left(LCD(X^*) < e^{cn}\right) \le e^{-c'n}.$$

Proof.

We further partition the set $Atyp$ into the level sets according to the values of the LCD:

$$S_D := \{x \in Atyp : \; D \le LCD(x) < 2D\}.$$

Here $D = 2^j$, where $j = 1, \dots, cn$.
On each set S_D it is enough to prove that

$$\mathbb{P}\left(X^* \in S_D\right) \le e^{-n}.$$

Then taking the union bound over the level sets S_D completes the proof. □

Norm minimization via the LCD

Lemma

Let $S_D \subset S^{n-1}$ be the set of all incompressible points such that $D \le LCD(x) < 2D$. Let X^* be a random normal. Then

$$\mathbb{P}\left(X^* \in S_D\right) \le e^{-n}.$$

- Individual probability estimate via LCD.

$$\mathbb{P}\left(\|X^* - y\| \text{ is small }\right) \text{ is exponentially small}$$

  for any fixed $y \in S_D$.
- Estimate of the cardinality of the ε-net.
- Approximation.

The volumetric estimate of the cardinality of the ε-net is not good enough!

The size of an ε-net

Lemma

Let $W_D \subset S^{n-1}$ be the set of vectors for which $LCD(x) \leq D$. Then there exists a (α/D)-net in W_D in the Euclidean metric, of cardinality at most

$$(CD/\alpha^{c'})^n \quad \text{for some } c' < 1.$$

The size of an ε-net

Lemma

Let $W_D \subset S^{n-1}$ be the set of vectors for which $LCD(x) \leq D$. Then there exists a (α/D)-net in W_D in the Euclidean metric, of cardinality at most

$$(CD/\alpha^{c'})^n \quad \text{for some } c' < 1.$$

Volumetric estimate:

$$|t\text{-net}| \leq (3/t)^n.$$

We gain α^c instead of α.

The size of an ε-net

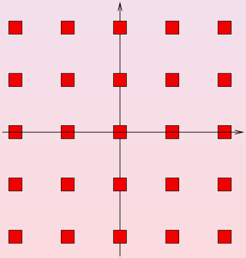
Lemma

Let $W_D \subset S^{n-1}$ be the set of vectors for which $LCD(x) \le D$. Then there exists a (α/D)-net in W_D in the Euclidean metric, of cardinality at most

$$(CD/\alpha^{c'})^n \quad \text{for some } c' < 1.$$

Proof.

- For any $x \in W_D$, the vector $n^{1/2} \cdot LCD(x) \cdot x$ has cn coordinates α-close to the integers.



The size of an ε-net

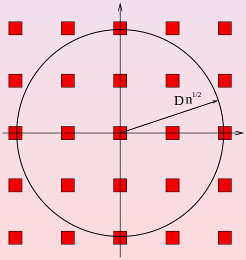
Lemma

Let $W_D \subset S^{n-1}$ be the set of vectors for which $LCD(x) \le D$. Then there exists a (α/D)-net in W_D in the Euclidean metric, of cardinality at most

$$(CD/\alpha^{c'})^n \quad \text{for some } c' < 1.$$

Proof.

- For any $x \in W_D$, the vector $n^{1/2} \cdot LCD(x) \cdot x$ has cn coordinates α-close to the integers.
- The restrictions of these vectors to such coordinates admit a (α/D)-net of cardinality $(CD)^{cn}$.



The size of an ε-net

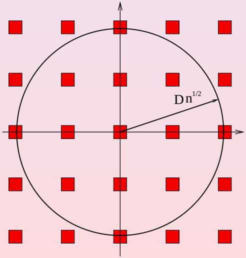
Lemma

Let $W_D \subset S^{n-1}$ be the set of vectors for which $LCD(x) \le D$. Then there exists a (α/D)-net in W_D in the Euclidean metric, of cardinality at most

$$(CD/\alpha^{c'})^n \quad \text{for some } c' < 1.$$

Proof.

- For any $x \in W_D$, the vector $n^{1/2} \cdot LCD(x) \cdot x$ has cn coordinates α-close to the integers.
- The restrictions of these vectors to such coordinates admit a (α/D)-net of cardinality $(CD)^{cn}$.
- On the rest of the coordinates we use the volumetric estimate: $(CD/\alpha)^{(1-c)n}$. □



Norm minimization via the LCD

ε-net argument

Lemma

Let $S_D \subset S^{n-1}$ be the set of all incompressible points such that $D \leq LCD(x) < 2D$. Then

$$\mathbb{P}\left(X^* \in S_D\right) \leq e^{-n}.$$

Norm minimization via the LCD

ε-net argument

Lemma

Let $S_D \subset S^{n-1}$ be the set of all incompressible points such that $D \le LCD(x) < 2D$. Then

$$\mathbb{P}\left(X^* \in S_D\right) \le e^{-n}.$$

Proof.

- Let $x \in S^{n-1}$ be an incompressible vector such that $LCD(x) > D$. Then for any $y \in S_D$
$$\mathbb{P}\left(\|X^* - y\| < \alpha/D\right) \le (C\alpha/D)^{n-1}.$$
- There exists a (α/D)-net $\mathcal{N}$ in S_D in the Euclidean metric, of cardinality at most $(CD/\alpha^c)^n$ for some $c < 1$.

Norm minimization via the LCD

ε-net argument

Lemma

Let $S_D \subset S^{n-1}$ be the set of all incompressible points such that $D \le LCD(x) < 2D$. Then

$$\mathbb{P}\left(X^* \in S_D\right) \le e^{-n}.$$

Proof.

- Let $x \in S^{n-1}$ be an incompressible vector such that $LCD(x) > D$. Then for any $y \in S_D$
$$\mathbb{P}\left(\|X^* - y\| < \alpha/D\right) \le (C\alpha/D)^{n-1}.$$
- There exists a (α/D)-net $\mathcal{N}$ in S_D in the Euclidean metric, of cardinality at most $(CD/\alpha^c)^n$ for some $c < 1$.
- $\mathbb{P}\left(\exists x \in \mathcal{N}\ \|A'x\|_2 < \alpha/D\right) \le (C\alpha^{1-c})^n \le e^{-n}$. □