Statistics and Probability Letters 173 (2021) 109070

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Least singular value and condition number of a square random matrix with i.i.d. rows

M. Gregoratti, D. Maran *

Politecnico di Milano piazza Leonardo da Vinci 32, Milano, Italy

ARTICLE INFO

Article history:
Received 16 November 2020
Received in revised form 5 February 2021
Accepted 17 February 2021
Available online 24 February 2021

MSC:
60B20
15B52

Keywords:
Least singular value
Condition number
Random matrix

ABSTRACT

Introducing a new method for studying general probability distributions on $\mathbb{R}^n$, we generalize some results about the least singular value and the condition number of random matrices with i.i.d. Gaussian entries to the whole class of random matrices with i.i.d. rows.

© 2021 Elsevier B.V. All rights reserved.

1. Introduction

The first important results about the least singular value $\sigma_{\min}(\widetilde{X})$ and the condition number $\kappa(\widetilde{X})$ of a square $n \times n$ random matrix $\widetilde{X}$ were obtained in 1988. Edelman (1988) computed the exact distribution of $\sigma_{\min}(\widetilde{X})$ for a matrix of i.i.d. complex standard Gaussian entries and the limiting distribution in the i.i.d. real standard Gaussian case. Kostlan (1988) proved that $\mathbb{E}[\kappa(\widetilde{X})] = +\infty$ whenever the entries are i.i.d. real centred Gaussian, regardless of the matrix dimension. In recent years, further results were obtained for the behaviour of $\sigma_{\min}(\widetilde{X})$ and $\kappa(\widetilde{X})$: while Huang and Tikhomirov (2020) focused on the powers of Gaussian matrices, Rebrova and Tikhomirov (2018) and Rudelson and Vershynin (2008) weakened the Gaussianity assumption and Adamczak et al. (2012), Tikhomirov (2020), Livshyts et al. (2019), Tatarko (2018) even relaxed the assumption that the entries of the same row are i.i.d.

In particular, Adamczak et al. (2012) and Tikhomirov (2020) managed to prove, for $n \times n$ matrices of i.i.d. log-concave rows, bounds of the type

$$\mathbb{P}\left(\sigma_{\min}(\widetilde{X}) \leq \epsilon\right) < f(n, \epsilon)$$

which hold for fixed values of n and $\epsilon > 0$, and with $f(n, \epsilon) \to 0$ as $\epsilon \to 0^+$.

Then a natural aim could be to find how further these estimations can arrive and, moreover, to study what happens outside the log-concave case.

In this paper we can prove that, in a right neighbourhood of 0, there exists a linear lower bound for the cumulative distribution function of $\sigma_{\min}(\widetilde{X})$ of a square random matrix, of every fixed dimension n, in the general setting of i.i.d. rows without any additional assumption on the rows distribution.

* Corresponding author.
E-mail addresses: Matteo.Gregoratti@polimi.it (M. Gregoratti), Davide.maran@mail.polimi.it (D. Maran).

https://doi.org/10.1016/j.spl.2021.109070
0167-7152/© 2021 Elsevier B.V. All rights reserved.

Indeed, under these only assumptions, we can prove our main results:

- $\liminf_{\epsilon\to 0^+} \frac{\mathbb{P}\Big(\sigma_{\min}(\widetilde{X}) < \epsilon\Big)}{\epsilon} > 0,$
- $\mathbb{E}\left[\frac{1}{\sigma_{\min}(\widetilde{X})}\right] = \mathbb{E}\Big[\|\widetilde{X}^{-1}\|\Big] = +\infty,$
- $\mathbb{E}\Big[\kappa(\widetilde{X})\Big] = \mathbb{E}\Big[\|\widetilde{X}\|\|\widetilde{X}^{-1}\|\Big] = +\infty.$

The first item generalizes the behaviour of the least singular value of i.i.d. real Gaussian entries. The last item generalizes the result by Kostlan on the average condition number. Note that our results are trivial if $\mathbb{P}\Big(\sigma_{\min}(\widetilde{X}) = 0\Big) > 0$, which happens whenever the distribution of the rows is discrete, such as in the Littlewood–Offord problem (see Tao and Vu, 2009 and all the linked articles).

Moreover, in the cases of a random matrix described by Adamczak et al. (2012) and Tikhomirov (2020), we get additional results by combining our lower bound with their upper bounds. We prove that the probability of $\sigma_{\min}(\widetilde{X}) \in [0, \epsilon)$ grows linearly with ϵ in a neighbourhood of 0, as well as we prove an interesting property of the moments of $\kappa(\widetilde{X})$ showing that the isotropic log-concave distribution has the best behaviour in terms of the well conditioning.

2. Notations

Given $p \in \mathbb{N} \cup \{+\infty\}$, a vector $x \in \mathbb{R}^n$ and a square matrix $A \in \mathbb{R}^{n\times n}$, we introduce operator p-norms, based on the usual vector $\|\cdot\|_p$ norm as

$$\|A\|_p = \max_{\|x\|_p=1} \|Ax\|_p.$$

Moreover, if we denote by $\sigma_{\min}(A)$ and $\sigma_{\max}(A)$ the smallest and the largest *singular value* of A respectively, then we have

$$\sigma_{\max}(A) = \|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2, \qquad \sigma_{\min}(A) = \min_{\|x\|_2=1} \|Ax\|_2 = \frac{1}{\|A^{-1}\|_2}.$$

Finally, the *condition number* of A in matrix norm $\|\cdot\|$ on $\mathbb{R}^{n\times n}$ is

$$\kappa(A) = \begin{cases} \|A\|\,\|A^{-1}\|, & \text{if A is invertible,} \\ +\infty, & \text{otherwise.} \end{cases}$$

The condition number depends on the choice of the matrix norm, but different condition numbers are always pairwise equivalent thanks to the pairwise equivalence of the norms. Lastly, in a given space $\mathbb{R}^n$ we are indicating with $B_d(x)$, where $x \in \mathbb{R}^n$ and $d > 0$, the 2-norm ball of radius d centred at x.

3. The moulds

Our results are based on the introduction of moulds, whose definition is motivated by the following well known result: given a positive random variable W, if $\liminf_{t\to+\infty}\Big(1 - \mathbb{P}(W \le t)\Big)t = q > 0$, then $\mathbb{E}[W] = +\infty$.

Definition 1. Let X be a random vector in $\mathbb{R}^n$. For every integer number $m \ge 0$, the *m-dimensional mould* of X, denoted by $\mathbf{C}_m(X)$, is the set of all $x \in \mathbb{R}^n$ such that

$$\liminf_{\epsilon\to 0^+} \frac{\mathbb{P}\Big(\|X - x\|_2 < \epsilon\Big)}{\epsilon^m} > 0.$$

Of course, every mould $\mathbf{C}_m(X)$ only depends on the distribution of X and, moreover, it does not change if we change the vector norm in the definition. Then we can immediately prove the following important feature of the moulds.

Theorem 1. *Let X be a random vector in $\mathbb{R}^n$ and let $x \in \mathbf{C}_m(X)$, $m \ge 1$. Then*

$$\mathbb{E}\left[\frac{1}{\|X - x\|^m}\right] = +\infty.$$

Proof. By definition of mould, we know that $\liminf_{\epsilon\to 0^+} \frac{\mathbb{P}(\|X-x\|<\epsilon)}{\epsilon^m} > 0$. Then the thesis comes from the above mentioned property after the change of variable $\epsilon = \sqrt[m]{1/t}$.

In order to usefully apply such a theorem, we need to explore some other features of the moulds. First of all, moulds are a sequence of Borel sets that obviously grows with the index:

$$\mathbf{C}_\ell(X) \subseteq \mathbf{C}_m(X), \qquad \forall \ell \leq m. \tag{1}$$

The measurability of each $\mathbf{C}_m(X)$ can be deduced from the following lemma, which allows to substitute the limit over the continuous index $\epsilon \to 0^+$ with a limit over a sequence $t_n \to 0$. The proof is straightforward.

Lemma 2.

$$\liminf_{\epsilon \to 0^+} \frac{\mathbb{P}\Big(\|X-x\|_2 < \epsilon\Big)}{\epsilon^m} > 0 \iff \liminf_{k\to\infty} \frac{\mathbb{P}\Big(\|X-x\|_2 < 2^{-k}\Big)}{2^{-mk}} > 0.$$

Even if some $\mathbf{C}_m(X)$ can be empty, an important result holds for $m = n$.

Theorem 3. *Let X be a random vector in $\mathbb{R}^n$. Then $\mathbb{P}\Big(X \in \mathbf{C}_n(X)\Big) = 1$.*

Proof. By contradiction, suppose that $\mathbb{P}\Big(X \in \mathbf{C}_n(X)\Big) < 1$. Then $\mathbb{P}\Big(X \in \mathbf{C}_n(X)^c\Big) > 0$ and so there is a closed $C \subseteq \mathbf{C}_n(X)^c$ with $\mathbb{P}(X \in C) > 0$.

Thus, $\mu(E) := \mathbb{P}(X \in C \cap E)$ defines a non-zero measure on the Borel sets. Anyway,

$$\liminf_{\epsilon\to 0^+} \frac{\mu(B_\epsilon(x))}{\epsilon^n} = 0, \qquad \forall x \in \mathbb{R}^n.$$

Indeed, the liminf holds both in C, by definition of $\mathbf{C}_n(X)$, and in C^c, as this set is open and μ vanishes on the subsets of C^c. Therefore, by Lemma 2,

$$\liminf_{m\to\infty} \frac{\mu(B_{2^{-m}}(x))}{2^{-mn}} = 0, \qquad \forall x \in \mathbb{R}^n. \tag{2}$$

Since μ is not the zero measure, there exists a hypercube Q with $\mu(Q) > 0$, which means that $\frac{\mu(Q)}{\mathrm{diam}(Q)^n} = k > 0$.

Now, if we cut Q into 2^n identical hypercubes $\{Q_i\}_{i\in 1,\dots,2^n}$ with half its diameter, for at least one of them we will have $\frac{\mu(Q_\star)}{\mathrm{diam}(Q_\star)^n} \geq k$.

Iterating this process with $Q = Q_\star$, the centres of the hypercubes form a Cauchy sequence $\{x_m\}_m$ converging to $x \in Q$ such that

$$\liminf_{m\to\infty} \frac{\mu\Big(\{y : \ \|x-y\|_\infty < 2^{-m}\,\mathrm{diam}(Q)\}\Big)}{(2^{-m}\,\mathrm{diam}(Q))^n} \geq k > 0,$$

and so, by the equivalence of the infinity norm and the 2-norm, we obtain $\liminf_{m\to\infty} \frac{\mu(B_{2^{-m}}(x))}{2^{-mn}} > 0$ which contradicts (2) and completes the proof.

Thus, every random vector X in $\mathbb{R}^n$ takes values almost surely in its n-dimensional mould $\mathbf{C}_n(X)$. In particular $\mathbf{C}_n(X)$ cannot be empty. Depending on the distribution of X, such a property can be extended also to lower m-dimensional moulds $\mathbf{C}_m(X)$. The proof is not difficult but a bit technical.

Theorem 4. *Let X be random vector in $\mathbb{R}^n$ such that $X \in B$ a.s., B being a Borelian subset of $\mathbb{R}^n$. Suppose that there exists a measurable function $d : B \to \mathbb{R}^m$ and a number $c > 0$ such that*

$$\|d(x) - d(y)\| \geq c\,\|x-y\|, \qquad \forall x, y \in B.$$

Then $\mathbb{P}\Big(X \in \mathbf{C}_m(X)\Big) = 1$.

For example, Theorem 4 immediately implies that $\mathbb{P}\Big(X \in \mathbf{C}_m(X)\Big) = 1$ if X takes values almost surely in some m-dimensional linear subspace of $\mathbb{R}^n$.

4. *n* i.i.d. *n*-dimensional random vectors

Assumption 2. $X_1, \dots, X_n$ satisfy Assumption 2 if they are i.i.d. random vectors in $\mathbb{R}^n$ such that $X_1, \dots, X_{n-1}$ are linearly independent a.s. ($n \geq 2$).

For example, Assumption 2 is satisfied by n i.i.d. random vectors with an absolutely continuous distribution in $\mathbb{R}^n$. Note that, in particular, this assumption ensures that there is a random vector Y with unit ∞-norm which is almost surely orthogonal to $X_1, \dots, X_{n-1}$. So we can state the main result of this section.

Theorem 5. *Let $X_1, \dots, X_n$ be random vectors satisfying Assumption 2. Let Y be a random vector with unit ∞-norm which is almost surely orthogonal to $X_1, \dots, X_{n-1}$. Then*

$$0 \in \mathbf{C}_1(X_n \cdot Y), \qquad \mathbb{E}\left[\frac{1}{|X_n \cdot Y|}\right] = +\infty.$$

The proof of Theorem 5 takes the whole section and, of course, it relies on the introduction of moulds and their basic properties. We begin with the following property of $\mathbf{C}_{n-1}(Y)$.

Theorem 6. *Let $X_1, \dots, X_n$ be random vectors satisfying Assumption 2. Let Y be a random vector with unit ∞-norm which is almost surely orthogonal to $X_1, \dots, X_{n-1}$. Then*

$$Y \in \mathbf{C}_{n-1}(Y) \text{ a.s.}$$

Proof. By construction, the random vector Y belongs to $S_\infty^{n-1} = \left\{ v \in \mathbb{R}^n \ : \ \|v\|_\infty = 1 \right\}$ a.s. Since there exists a measurable dilation $d : S_\infty^{n-1} \to \mathbb{R}^{n-1}$, the thesis follows immediately by Theorem 4.

The next step is to study the special case of *bounded* $X_1, \dots, X_n$, where we can prove the desired results by showing a link between $\mathbf{C}_{n-1}(Y)$ and the properties of X_n.

Theorem 7. *Let $X_1, \dots, X_n$ be random vectors satisfying Assumption 2 and, moreover, let them be bounded. Let Y be a random vector with unit ∞-norm which is almost surely orthogonal to $X_1, \dots, X_{n-1}$. Then*

1. $y \in \mathbf{C}_{n-1}(Y) \implies 0 \in \mathbf{C}_1(X_n \cdot y)$,
2. $0 \in \mathbf{C}_1(X_n \cdot Y)$,
3. $\mathbb{E}\left[\frac{1}{|X_n \cdot Y|}\right] = +\infty$.

Proof. We prove the theorem thesis by thesis.

1. Since $X_1, \dots, X_n$ are i.i.d., for every $y \in \mathbb{R}^n$ and for every $\epsilon > 0$ we have

$$\mathbb{P}\Big(|X_n \cdot y| < \epsilon\Big) = \sqrt[n-1]{\mathbb{P}\left(\bigcap_{j=1}^{n-1}\Big(|X_j \cdot y| < \epsilon\Big)\right)}.$$

Now, let us take $r > 0$ such that $\|X_j\|_1 < r$ a.s., and let us denote by $\widehat{X}$ the $\mathbb{R}^{(n-1)\times n}$ random matrix with rows $X_j : 1 \le j \le n-1$.
Then we have the following relationships among events

$$\bigcap_{j=1}^{n-1}\Big(|X_j \cdot y| < \epsilon\Big) = \Big(\|\widehat{X}y\|_\infty < \epsilon\Big) = \Big(\|\widehat{X}(y - Y)\|_\infty < \epsilon\Big)$$
$$\supseteq \Big(\|\widehat{X}\|_\infty \|y - Y\|_\infty < \epsilon\Big) \supseteq \Big(r\|y - Y\|_\infty < \epsilon\Big),$$

so that

$$\liminf_{\epsilon \to 0^+} \frac{\mathbb{P}\Big(|X_n \cdot y| < \epsilon\Big)}{\epsilon} \ge \liminf_{\epsilon \to 0^+} \sqrt[n-1]{\frac{\mathbb{P}\Big(\|Y - y\|_\infty < \epsilon/r\Big)}{\epsilon^{n-1}}}.$$

Therefore $0 \in \mathbf{C}_1(X_n \cdot y)$ for every $y \in \mathbf{C}_{n-1}(Y)$.

2. Using Fubini–Tonelli theorem and Fatou's lemma, we have

$$\liminf_{\epsilon \to 0^+} \frac{\mathbb{P}\Big(|X_n \cdot Y| < \epsilon\Big)}{\epsilon} = \liminf_{\epsilon \to 0^+} \frac{\int_{\mathbb{R}^n} \mathbb{P}\Big(|X_n \cdot y| < \epsilon\Big)\, \mathrm{d}P^Y(y)}{\epsilon}$$
$$\ge \int_{\mathbb{R}^n} \liminf_{\epsilon \to 0^+} \frac{\mathbb{P}\Big(|X_n \cdot y| < \epsilon\Big)}{\epsilon}\, \mathrm{d}P^Y(y)$$

Since $\liminf_{\epsilon\to 0^+} \frac{\mathbb{P}(|X_n\cdot y|<\epsilon)}{\epsilon} > 0$ for every y in the support of Y (see previous point), by monotonicity we have

$$\int_{\mathbb{R}^n} \liminf_{\epsilon\to 0^+} \frac{\mathbb{P}\Big(|X_n\cdot y| < \epsilon\Big)}{\epsilon} dP^Y(y) > 0$$

which entails the thesis.

3. Thesis 3 follows immediately from thesis 2 thanks to Theorem 1.

Finally we can prove Theorem 5.

Proof of Theorem 5. The result is already proved for bounded random vectors thanks to Theorem 7. Then, taken a $\rho > 0$ such that the event

$$E_\rho = \bigcap_{i=1}^{n} \Big(\|X_i\| < \rho\Big)$$

has positive probability, it is enough to consider the conditional probability $\mathbb{P}_\rho(\cdot) = \mathbb{P}(\cdot|E_\rho)$. Indeed for every $\epsilon > 0$

$$\frac{\mathbb{P}\Big(|X_n\cdot Y| < \epsilon\Big)}{\epsilon} \geq \frac{\mathbb{P}_\rho\Big(|X_n\cdot Y| < \epsilon\Big)}{\epsilon}\,\mathbb{P}(E_\rho),$$

where the right hand side has a strictly positive liminf as $\epsilon \to 0^+$ by Theorem 7, as the random vectors $X_1,\ldots,X_n$ are bounded under $\mathbb{P}_\rho$ and it is a straightforward verification that they are also $\mathbb{P}_\rho$-i.i.d. and still satisfy Assumption 2.

Therefore $0 \in \mathbf{C}_1(X_n\cdot Y)$ and the full thesis follows by Theorem 1.

5. Least singular value $\sigma_{\min}(\widetilde{X})$

Thanks to the introduction of the definition of moulds for a random vector (Section 3) and thanks to the properties deduced for an n-uple of i.i.d. random vectors in $\mathbb{R}^n$ (Section 4), we can finally come to our main results. Let us start with the least singular value.

5.1. The main result for $\sigma_{min}(\widetilde{X})$

Theorem 8. *Let $\widetilde{X}$ be a square random matrix with i.i.d. rows. Then*

$$0 \in \mathbf{C}_1\Big(\sigma_{min}(\widetilde{X})\Big) \qquad \text{i.e.} \qquad \liminf_{\epsilon\to 0^+} \frac{\mathbb{P}\Big(\sigma_{min}(\widetilde{X}) < \epsilon\Big)}{\epsilon} > 0,$$

and, if $\widetilde{X}$ is invertible almost surely,

$$\mathbb{E}\left[\frac{1}{|\sigma_{min}(\widetilde{X})|}\right] = \mathbb{E}\Big[\|\widetilde{X}^{-1}\|\Big] = +\infty.$$

Proof. If the random matrix $\widetilde{X}$ is singular with positive probability the thesis is trivial. Otherwise its rows $X_1,\ldots,X_n$ satisfy Assumption 2 and we can consider the random vector Y of Theorem 5. Then it is enough to observe that, since $\|Y\|_\infty = 1$ and so $\|Y\|_2 \geq 1$,

$$\begin{aligned}\Big(\sigma_{\min}(\widetilde{X}) < \epsilon\Big) = \Big(\min_{\|y\|_2=1} \|\widetilde{X}y\|_2 < \epsilon\Big) &\supseteq \left(\frac{\|\widetilde{X}Y\|_2}{\|Y\|_2} < \epsilon\right) \\ &\supseteq \Big(\|\widetilde{X}Y\|_2 < \epsilon\Big) = \Big(|X_n\cdot Y| < \epsilon\Big),\end{aligned}$$

to deduce

$$\liminf_{\epsilon\to 0^+} \frac{\mathbb{P}\Big(\sigma_{\min}(\widetilde{X}) < \epsilon\Big)}{\epsilon} \geq \liminf_{\epsilon\to 0^+} \frac{\mathbb{P}\Big(|X_n\cdot Y| < \epsilon\Big)}{\epsilon} > 0.$$

The full thesis then follows thanks to Theorem 1.

Since the least singular value is invariant under transposition, the theorem holds for matrices with i.i.d. columns, too.

Theorem 8 implies that, in a right neighbourhood of 0, there exists a linear lower bound for the cumulative distribution function of $\sigma_{\min}(\widetilde{X})$. Of course, such a bound is not always optimal, as for example in the discrete case. However, for the large class of log-concave distributions, some results from the literature will complete our results with proper upper bounds.

5.2. Additional results for $\sigma_{min}(\widetilde{X})$ in the isotropic log-concave case

A distribution is log-concave if it is absolutely continuous with a density f such that

$$f(\lambda x+(1-\lambda)y)\le f(x)^{\lambda}f(y)^{1-\lambda},\qquad \forall\lambda\in(0,1).$$

The study of random matrices with log-concave rows or columns is related to some geometric applications which arise in the study of sampling on convex bodies, as shown by Adamczak et al. (2012).

Summing up our result and the ones of Adamczak et al. (2012, corollary 2.14) and Tikhomirov (2020, corollary 2.14) we get the following corollary.

Theorem 9. *Let $\widetilde{X}$ be a random matrix with i.i.d. rows drawn from a zero-mean log-concave distribution. Then, for every $\delta>0$ there exist $0<k_1<k_2$ such that*

$$k_1\epsilon<\mathbb{P}\Big(\sigma_{min}(\widetilde{X})<\epsilon\Big)<k_2\epsilon^{1-\delta}$$

(where $k_2=C_\delta\sqrt{n}$ and C_δ only depends on δ) holds for sufficiently small $\epsilon>0$. Moreover, there exists a universal constant n_0 such that, if the size of $\widetilde{X}$ is greater than n_0, then

$$k_1\epsilon<\mathbb{P}\Big(\sigma_{min}(\widetilde{X})<\epsilon\Big)<k_2\epsilon$$

(where $k_2=C\sqrt{n}$ and C is a universal constant) holds for sufficiently small $\epsilon>0$.

6. Condition number $\kappa(\widetilde{X})$

From the results proved for $\sigma_{\min}(\widetilde{X})$, we can derive a theorem on $\mathbb{E}\Big[\kappa(\widetilde{X})\Big]$.

6.1. The main result for $\kappa(\widetilde{X})$

Theorem 10. *Let $\widetilde{X}$ be a square random matrix with i.i.d. rows. Then, for every choice of the matrix norm,*

$$\mathbb{E}\Big[\kappa(\widetilde{X})\Big]=+\infty.$$

Proof. If the random matrix $\widetilde{X}$ is singular with positive probability the thesis is trivial. Otherwise, when $\widetilde{X}$ is invertible a.s., it is enough to prove the theorem for the operator norm induced by the norm infinity of $\mathbb{R}^n$, as condition numbers are pairwise equivalent for a change of the matrix norm.

We prove the theorem in two steps, first for rows $X_1,\dots,X_n$ bounded from below, then for the general case of $\widetilde{X}$ invertible a.s.

1. If $\|X_1\|_1>\rho$ a.s. for some $\rho>0$, then the thesis immediately follows. Indeed, such a condition gives $\|\widetilde{X}\|_\infty=\max_i\|X_i\|_1>\rho$ a.s. and so, by Theorem 8,
$$\mathbb{E}\Big[\kappa_\infty(\widetilde{X})\Big]=\mathbb{E}\Big[\|\widetilde{X}\|_\infty\,\|\widetilde{X}^{-1}\|_\infty\Big]>\rho\,\mathbb{E}\Big[\|\widetilde{X}^{-1}\|\Big]=+\infty.$$
2. If $\widetilde{X}$ is invertible a.s., then $\mathbb{P}\Big(\|X_i\|>0\Big)=1$ and, by monotonicity, there exists $\rho>0$ such that $\mathbb{P}\Big(\|X_1\|_1>\rho\Big)>0$. Thus, the event $E_\rho=\bigcap_{i=1}^n\Big(\|X_i\|_1>\rho\Big)$ has positive probability and we can consider the conditional probability $\mathbb{P}_\rho(\cdot)=\mathbb{P}(\cdot|E_\rho)$. As $\mathbb{P}(A)\ge\mathbb{P}_\rho(A)\,\mathbb{P}(E_\rho)$ for every event A, we also have $\mathbb{E}[W]\ge\mathbb{E}_\rho[W]\,\mathbb{P}(E_\rho)$ for every random variable $W\ge 0$. Thus $\mathbb{E}\Big[\kappa(\widetilde{X})\Big]\ge\mathbb{E}_\rho\Big[\kappa(\widetilde{X})\Big]\,\mathbb{P}(E_\rho)=+\infty$ by step 1, as the random vectors $X_1,\dots,X_n$ are bounded from below under $\mathbb{P}_\rho$ and it is a straightforward verification that they are also $\mathbb{P}_\rho$-i.i.d. and satisfy Assumption 2.

This theorem is a generalization of Kostlan (1988, theorem 5.2), which is proved for a random matrix with i.i.d. Gaussian entries.

6.2. Additional results for $\kappa(\widetilde{X})$ in the isotropic log-concave case

Again, Adamczak et al. (2012) proved an upper bound for $\kappa(\widetilde{X})$ (Corollary 2.15) in the isotropic log-concave case. Their result bounds the probability that $\kappa(\widetilde{X})$ is high, and it is equivalent to say that $\mathbb{E}\Big[\kappa(\widetilde{X})^{1-\delta}\Big]<+\infty$ for any $\delta>0$; thus, it can be merged with Theorem 10 to obtain the following corollary.

Theorem 11. *Let $\widetilde{X}$ be a square random matrix with i.i.d. isotropic log-concave rows. Then*

$$\mathbb{E}\Big[\kappa(\widetilde{X})^{\alpha}\Big]<+\infty\iff\alpha<1.$$

Acknowledgement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Adamczak, R., Guedon, O., Litvak, A.E., Pajor, A., Tomczak-Jaegermann, N., 2012. Condition number of a square matrix with I.I.D. columns drawn from a convex body. Proc. Amer. Math. Soc. 140 (3), 987–998.
Edelman, A., 1988. Eigenvalues and condition numbers of random matrices. SIAM J. Matrix Anal. Appl. 9, 543–560.
Huang, H., Tikhomirov, K., 2020. Remark on the smallest singular value of powers of Gaussian matrices. Electron. Commun. Probab. 25, paper no. 10.
Kostlan, E., 1988. Complexity theory of numerical linear algebra. J. Comput. Appl. Math. 22, 219–230.
Livshyts, G., Tikhomirov, K., Vershynin, R., 2019. The smallest singular value of inhomogeneous square random matrices. ArXiv. arXiv:1909.04219v4. 24:118–173.
Rebrova, E., Tikhomirov, K., 2018. Coverings of random ellipsoids, and invertibility of matrices with I.I.D. heavy-tailed entries. Israel J. Math. 227, 507–544.
Rudelson, M., Vershynin, R., 2008. Invertibility of random matrices: Norm of the inverse. Ann. of Math. 168, 575–600.
Tao, T., Vu, V., 2009. Inverse Littlewood-Offord theorems and the condition number of random discrete matrices. Ann. of Math. 169, 595–632.
Tatarko, K., 2018. An upper bound on the smallest singular value of a square random matrix. J. Complexity 48, 119–128.
Tikhomirov, K., 2020. Invertibility via distance for non-centered random matrices with continuous distributions. ArXiv. arXiv:1707.09656v3. 24:118–173.