

Probabilistic Estimates of the Largest Strictly Convex Singular Values of Pregaussian Random Matrices

Yang Liu

Department of Mathematics, Michigan State University, East Lansing, MI 48824, USA

Article history

Received: 28-02-2015

Revised: 27-03-2015

Accepted: 30-03-2015

Abstract: In this study, the p -singular values of random matrices with Gaussian entries defined in terms of the l_p -norm for $p > 1$, as is studied. Mainly, using analytical techniques, we show the probabilistic estimate, precisely, the decay, on the upper tail probability of the largest strictly convex singular values, when the number of rows of the matrices becomes very large and the lower tail probability of theirs as well. These results provide probabilistic description or picture on the behaviors of the largest p -singular values of random matrices in probability for $p > 1$. Also, we show some numerical experimental results, which verify the theoretical results.

Keywords: Probability, Random Matrices, Singular Value, Banach Norm

Introduction

The largest singular value and the smallest singular value of random matrices in l_2 -norm, including Gaussian random matrices, Bernoulli random matrices, subgaussian random matrices, etc, have attracted major research interest in recent years and have applications in compressed sensing, a technique for recovering sparse or compressible signals. For instance, (Soshnikov, 2002; Soshnikov and Fyodorov, 2004) studied the largest singular value of random matrices and (Rudelson and Vershynin, 2008a; 2008b; Tao and Vu, 2010) and some others, studied the smallest singular values.

In the study of the asymptotic behavior of eigenvalues of symmetric random matrices, Wigner symmetric matrix is a typical example, whose upper (or lower) diagonal entries are independent random variables with uniform bounded moments. Wigner proved in (Wigner, 1958) that the normalized eigenvalues are asymptotically distributed in the semicircular distribution. Precisely, let A be a symmetric gaussian random matrix of size $n \times n$ whose upper diagonal entries are independent and identically-distributed copies of the standard gaussian random variable, then the empirical distribution function of the eigenvalues of $\frac{1}{\sqrt{n}} A$ is asymptotically:

$$p(x) := \begin{cases} \frac{1}{2\pi} \sqrt{4-x^2} & \text{for } |x| \leq 2 \\ 0 & \text{for } |x| > 2 \end{cases} \quad (1.1)$$

As the matrix size n goes to infinity. This is the well-known Wigner's Semicircle law, which provides the precise description of the statistical behavior of eigenvalues of matrix of large size. In another case, for a random matrix whose entries are independent and identically-distributed (i.i.d.) copies of a complex random variable with mean 0 and variance 1, Tao and Vu, (2008; Tao *et al.*, 2010) that the eigenvalues of $\frac{1}{\sqrt{n}} a$ converges to the uniform distribution on the unit circle as n goes to ∞ and that holds not only for the random matrices with real entries but also for complex entries. Their result has also generalized (Girko, 1985) and solved the circular law conjecture open since the 1950's, that the smallest eigenvalue converges to the uniform distribution over the unit disk as n tends to infinity (Bai, 1997).

The largest singular values of matrices are actually their p -norm, which, from a geometric perspective, has connections with the Minkowski space, complex l^p space, in differential geometry, for which one can refer to (Liu, 2013; 2011), because one can view the p -norm of a matrix as a generalization of the p -norm of a vector.

For random matrices whose entries are i.i.d. random variable satisfying certain moment conditions, the largest singular value was studied in (Geman, 1980; Yin *et al.*, 1988). Tracy and Widom (1996) that the limiting law of largest eigenvalue distributions of Gaussian Orthogonal Ensemble (GOE) is given in terms of a particular Painlevé II function, which is the well-known Tracy-Widom law. Furthermore, the distribution of the

eigenvalue of Wishart matrices, $W_{N,n} = AA^*$, where $A = A_{N,n}$ is a Gaussian random matrix of size $N \times n$, was studied in (Johansson, 2000; Johnstone, 2001). They showed that the distribution of largest eigenvalue of Wishart matrices converges to the Tracy-Widom law as $\frac{n}{N}$ tend to some positive constant. More generally, the non-gaussian random matrices were studied in (Soshnikov, 2002). Seginer (2000) compared the Euclidean operator norm of a random matrix with i.i.d. mean zero entries to the Euclidean norm of its rows and columns. Later, (Latala, 2005) gave the upper bound on the expectation (or average value) of largest singular value namely the norm of any random matrix whose entries are independent mean zero random variables with uniformly bounded fourth moment.

The condition number, which is the ratio of the largest singular value over the smallest singular value of a matrix, is critical to the stability of linear systems. In (Edelman, 1988), the distribution of the condition number of Gaussian random matrices, was particularly investigated in numerical experiments. As a typical example of subgaussian random matrices, the invertibility of Bernoulli random matrices was also studied. Tao and Vu (2007) the probability of Bernoulli random matrices to be singular is shown to be at most $\left(\frac{3}{4} + o(1)\right)^n$, where n is the size of the matrices.

Their result shows that the probability of the smallest singular value of Bernoulli random matrices to be zero is exponentially small as n tends to infinity. Recently, the singularity probability $\left(\frac{3}{4} + o(1)\right)^n$ has been improved to $\left(\frac{1}{2} + o(1)\right)^n$ by (Bourgain *et al.*, 2010).

The recent studies of the smallest singular value have also been motivated, in a large sense, by some open questions or conjectures. Spielman and Teng (2002) the following conjecture was proposed in the International Congress of Mathematicians in 2002.

Conjecture 1.1

Let ξ be Bernoulli random variable, in other words, $P(\xi - 1) = P(\xi = -1) = \frac{1}{2}$. Then:

$$P\left(s_n\left(M(\xi) \leq \frac{t}{\sqrt{n}}\right) \leq t + c^n\right) \quad (1.2)$$

for all $t > 0$ and some $0 < c < 1$.

In the breakthrough work on the estimate on the smallest singular value, (Rudelson and Vershynin, 2008a), Rudelson and Vershynin obtained the upper tail

probabilistic estimate on the smallest value in l_2 -norm for square matrices of centered random variables, with unit variance and appropriate moment assumptions. In particular, they proved the Spielman-Teng conjecture up to a constant. The lower tail probabilistic estimate on the smallest value in l_2 -norm for square matrices was estimated in (Rudelson and Vershynin, 2008b). These results have shown that the smallest singular value of the

$n \times n$ subgaussian random matrices is of order $n^{-\frac{1}{2}}$ in high probability for large n . In a more explicit way, the distribution of the smallest singular value of random was given in (Tao and Vu, 2010) by using property testing from combinatorics and theoretical computer science. The pregaussian matrices were used to recover sparse image in (Rauhut, 2010) and matrix recovery, on which one can refer to (Oymak *et al.*, 2011; Lai *et al.*, 2012). Very recently, Rudelson and Vershynin (2010) gave a comprehensive survey on the extreme singular values of random matrices.

It is well-known that the classic singular value is defined in terms of l_2 -norm, then a natural question would be what if one defines the singular value by the l_q -quasinorm for $0 < q \leq 1$ and l_p -norm for $p > 1$. There were some remarkable results by other researchers on the largest singular values of random matrices in the l_2 -norm. Geman (1980; Yin *et al.*, 1988) showed that the largest singular value of random matrices of size $m \times N$ with independent entries of mean 0 and variance 1 tends to $\sqrt{m} + \sqrt{N}$ almost surely. The largest and smallest q -singular values of pregaussian random matrices for $0 < q \leq 1$ were studied in (Lai and Liu, 2014), which has applications in a technique of signal processing (Foucart and Lai, 2010; 2009; Lai and Liu, 2011) and other areas. Similar to the q -singular value when $0 < q \leq 1$, the strictly convex largest p -singular value, in which $p > 1$, can be defined and we will show the probabilistic estimate, precisely, the decay, on the upper tail probability of the largest strictly convex p -singular value, when the number of rows of the matrices becomes very large and the lower tail probability of theirs as well. These results provide probabilistic description or picture on the behaviors of the largest p -singular values of random matrices in probability.

The Largest p -Singular Value

The p -singular values of a matrix, in general, can be defined in the way of maximum of minimums or supremum of infimums. In largest p -singular values can be defined as follows:

Definition 2.1

For an $m \times N$ matrix A , the largest p -singular value of A denoted as $s_1^{(p)}(A)$ is defined as:

$$s_1^{(p)}(A) := \sup \{ \|Ax\|_p : x \in \mathbb{R}^N \text{ with } \|x\|_p = 1 \} \quad (2.1)$$

For given $p > 1$.

Lai and Liu (2014), the following lemma on a linear bound for partial binomial expansion was established.

Lemma 2.2

For every positive integer n :

$$\sum_{k=\lfloor \frac{n}{2} \rfloor + 1}^n \binom{n}{k} x^k (1-x)^{n-k} \leq 8x \quad (2.2)$$

For all $x \in [0, 1]$.

The above lemma can be applied to estimate probabilities.

Lemma 2.3

Suppose $\xi_1, \xi_2, \dots, \xi_n$ are i.i.d copies of a random variable ξ , then for any $\varepsilon > 0$:

$$P\left(\sum_{i=1}^n |\xi_i|^p \leq \frac{n\varepsilon}{2}\right) \leq 8P(|\xi| \leq \varepsilon) \quad (2.3)$$

For any given $p > 1$.

Proof. Given $p > 1$, we have the relation on the probability events that:

$$\left\{ (\xi_1, \dots, \xi_n) : \sum_{i=1}^n |\xi_i|^p \leq \frac{m\varepsilon}{2} \right\} \quad (2.4)$$

Is contained in:

$$\bigcup_{k=\lfloor \frac{n}{2} \rfloor + 1}^n \{ (\xi_1, \dots, \xi_n) : |\xi_{i_1}|^p \leq \varepsilon, \dots, |\xi_{i_k}|^p > \varepsilon, \dots, |\xi_{i_m}|^p > \varepsilon \} := \varepsilon \quad (2.5)$$

where, $\{i_1, i_2, \dots, i_k\}$ is a subset of $\{1, 2, \dots, n\}$ and $\{i_{k+1}, \dots, i_n\}$ is its complement.

Let $x = P(|\xi_1|^p \leq \varepsilon)$, then by the union probability:

$$P(\varepsilon) = \sum_{k=\lfloor \frac{n}{2} \rfloor + 1}^n \binom{n}{k} x^k (1-x)^{n-k} \quad (2.6)$$

And applying Lemma 2.2, we have:

$$P(\varepsilon) \leq 8x = 8P(|\xi_1| \leq \varepsilon) \quad (2.7)$$

Since the event (2.4) is contained in the event (2.5):

$$P\left(\sum_{i=1}^n |\xi_i|^p \leq \frac{n\varepsilon}{2}\right) \leq P(\varepsilon) \leq 8P(|\xi_1|^p \leq \varepsilon) \quad (2.8)$$

To estimate the lower tail probability of the largest p -singular value, we have the following theorem on

the lower tail probability of the largest p -singular value for $p > 1$.

Theorem 2.4

Let ξ be a pregaussian variable normalized to have variance 1 and A is an $m \times N$ matrix with i.i.d. copies of ξ in its entries, then for every $p > 1$ and any $\varepsilon > 0$, there exists $\gamma > 0$ such that:

$$P\left(s_1^{(p)}(A) \leq \gamma m^{\frac{1}{p}}\right) \leq \varepsilon \quad (2.9)$$

Which γ only depends on p , ε and the pregaussian variable ξ .

Proof. Since a_{ij} is pregaussian with variance 1, then any $\varepsilon > 0$, there is some $\delta > 0$, such that:

$$P(|a_{ij}|^p \leq \delta) \leq \frac{\varepsilon}{8} \quad (2.10)$$

But we know:

$$s_{\frac{1}{p}}^{(p)}(A) \geq \left(\sum_{i=1}^m |a_{ij}|^p\right)^{\frac{1}{p}} \quad (2.11)$$

For all j , because by the definition of the largest p -singular value 2.1, choosing x to be the standard basis vectors of \mathbb{R}^N gives us $\max_j \left(\sum_{i=1}^m |a_{ij}|^p\right)^{\frac{1}{p}} \leq s_1^{(p)}(A)$.

Therefore, by Lemma 2.3:

$$P\left(s_1^{(p)}(A) \leq \left(\frac{\delta}{2}\right)^{\frac{1}{p}} m^{\frac{1}{p}}\right) \leq P\left(\sum_{i=1}^m |a_{ij_0}|^p \leq \frac{m\delta}{2}\right) \leq 8P(|a_{ij}|^p \leq \delta) \leq \varepsilon \quad (2.12)$$

Thus let $\gamma = \left(\frac{\delta}{2}\right)^{\frac{1}{p}}$, then (2.9) follows.

For the upper tail probability of the largest p -singular value, $p > 1$, we can derive the following lemma first by using the Minkowski inequality and discrete Hölder inequality.

Lemma 2.5

For $p \geq 1$, (2.1) defines a norm on the space of $m \times N$ matrices and:

$$\max_j \|a_j\|_p \leq s_1^{(p)}(A) \leq N^{\frac{p-1}{p}} \max_j \|a_j\|_p \quad (2.13)$$

In which $a_j, j = 1, 2, \dots, N$, are the column vectors of A .

Applying the above lemma, an estimate we can derive easily for Bernoulli random matrices, whose every entry equals to 1 or -1 with equal probability (Tao and Vu, 2009), is the following theorem on the upper tail probability of the largest p -singular value of Bernoulli matrices for $p > 1$.

Theorem 2.6

Let ξ be a Bernoulli random variable normalized to have variance 1 and A be an $m \times N$ matrix with i.i.d. copies of ξ in its entries, then:

$$\frac{1}{m^p} \leq s_1^{(p)}(A) \leq m^{\frac{1}{p}} N^{\frac{p-1}{p}} \tag{2.14}$$

One may conjecture that the bound might be $m^{\frac{1}{p}}$. However, considering the Bernoulli matrices, whose entries are in Bernoulli distribution, as special subgaussian matrices, the expectation of the largest p -singular value may not be $m^{\frac{1}{p}}$. Indeed, let A be an $m \times m$ Bernoulli matrix and x be a non-zero vector in \mathbb{R}^m . The expectation of the largest p -singular value:

$$E(s_1^{(p)}(A)) \leq E \frac{\|Ax\|_p}{\|x\|_p} \tag{2.15}$$

For all $x \in \mathbb{R}^m$ and particularly for $x = (1, \dots, 1) \in \mathbb{R}^m$, we have:

$$E \frac{\|Ax\|_p}{\|x\|_p} = n^{-\frac{1}{p}} E \left(\sum_{i=1}^n |\epsilon_{i1} + \dots + \epsilon_{im}|^p \right)^{\frac{1}{p}} \tag{2.16}$$

Now let $X_i = \epsilon_{i1} + \dots + \epsilon_{im}$, then $E \left(\sum_{i=1}^n |\epsilon_{i1} + \dots + \epsilon_{im}|^p \right)^{\frac{1}{p}}$ is the expectation of the l_p -norm of the vector (X_1, X_2, \dots, X_n) .

We also have the following result on the upper tail probability of the largest p -singular value of Bernoulli matrices for $p > 1$.

Theorem 2.7

Let A be an $m \times m$ Bernoulli matrix with every entry equal to 1 or -1 with equal probability, then one has:

$$P(s_1^{(p)}(A) \geq K_m) \leq \exp(-cm) \tag{2.17}$$

For some $K > 0$ and some absolute constant $c > 0$.

Proof. Let $A = (\epsilon_{ij})_{m \times m}$ and S_p^{m-1} be the unit sphere with respect to l_p -norm in \mathbb{R}^m , then for any $x \in S_p^{m-1}$, by the convexity of the function $f(t) := t^p$ for $p > 1$:

$$\|Ax\|_p = \left(\sum_{i=1}^m \left| \sum_{j=1}^m \epsilon_{ij} x_j \right|^p \right)^{\frac{1}{p}} \leq m^{\frac{p-1}{p}} \tag{2.18}$$

$$\left(\sum_{i=1}^m \sum_{j=1}^m |\epsilon_{ij} x_j|^p \right)^{\frac{1}{p}} = m \|x\|_p = m$$

Therefore we have:

$$E \|Ax\|_p \leq m \tag{2.19}$$

For all $x \in S_p^{m-1}$. By Chernoff bound, we get:

$$P(\|Ax\|_p \geq Km) \leq P(\|Ax\|_p \geq KE \|Ax\|_p) \leq \exp(-cKm) \tag{2.20}$$

For any $K > 2$ and some absolute constant $c > 0$.

By Lemma 4.10 in (Pisier, 1999), there is a subset N which is a δ -net of S_p^{m-1} with cardinality:

$$\text{card}(N) \leq \left(1 + \frac{2}{\delta} \right)^m \tag{2.21}$$

Finally, using the union bound of probability and an approximation of any point on the sphere by points of the δ -net, we obtain (2.17).

For the rectangular matrices, we have the following theorem on the upper tail probability of the largest p -singular value of rectangular matrices for $1 < p \leq 2$.

Theorem 2.8

Let ξ be a pregaussian variable normalized to have variance 1 and A is an $m \times N$ matrix with i.i.d. copies of ξ in its entries, then for every $1 < p \leq 2$ and any $\epsilon > 0$, there exists $K > 0$ such that:

$$P \left(s_1^{(p)}(A) \geq K \left(m^{\frac{1}{p}} + m^{\frac{2-p}{2p}} N^{\frac{1}{2}} \right) \right) \leq \epsilon \tag{2.22}$$

where, K only depends on p , ϵ and the pregaussian variable ξ .

Proof. By the discrete Hölder inequality and the definition of the largest p -singular value:

$$s_1^{(p)}(A) = \sup_{x \in \mathbb{R}^N, x \neq 0} \frac{\|Ax\|_p}{\|x\|_p} \leq \sup_{x \in \mathbb{R}^N, x \neq 0} \frac{1}{m^{\frac{p-1}{2}}} \|Ax\|_2 = m^{\frac{1-p}{2}} s_1^{(2)}(A) \tag{2.23}$$

We also know that there exists $K > 0$ such that:

$$P\left(s_1^{(2)}(A) \geq K\left(m^{\frac{1}{2}} + N^{\frac{1}{2}}\right)\right) \leq \varepsilon \quad (2.24)$$

Therefore, we have:

$$\begin{aligned} &P\left(s_1^{(2)}(A) \geq K\left(m^{\frac{1}{2}} + m^{\frac{1}{p}} N^{\frac{1}{2}}\right)\right) \\ &\leq P\left(s_1^{(2)}(A) \geq K\left(m^{\frac{1}{2}} + N^{\frac{1}{2}}\right)\right) \leq \varepsilon \end{aligned} \quad (2.25)$$

To have a full generalization, let us derive the following useful lemma.

In general, for the relation between $s_1^{(q)}$ and $s_1^{(q)}$, $\frac{1}{p} + \frac{1}{q} = 1, q > 1$, one can deduce the following duality lemma on general rectangular matrices.

Lemma 2.9

For any $q \geq 1$ and $m \times N$ matrix A :

$$s_1^{(q)}(A) = s_1^{(p)}(A^T) \quad (2.26)$$

where, $\frac{1}{p} + \frac{1}{q} = 1$.

Proof. By the discrete Hölder inequality, we know that if $\frac{1}{p} + \frac{1}{q} = 1$ then:

$$\langle Ax, y \rangle \leq \|Ax\|_q \quad (2.27)$$

For all $x \in \mathbb{R}^N$ and $y \in \mathbb{R}^m$ with $\|y\|_p = 1$ and furthermore the equality holds for some y_0 with $\|y_0\|_p = 1$. Thus:

$$\|Ax\|_p = \sup_{y \in \mathbb{R}^m, \|y\|_p = 1} \langle Ax, y \rangle \quad (2.28)$$

By the definition of the largest q -singular value:

$$\begin{aligned} s_1^{(q)}(A) &= \sup_{x \in \mathbb{R}^N, \|x\|_q = 1} \|Ax\|_q \\ &= \sup_{x \in \mathbb{R}^N, \|x\|_q = 1} \sup_{y \in \mathbb{R}^m, \|y\|_p = 1} \langle Ax, y \rangle \end{aligned} \quad (2.29)$$

In the same way, we also have:

$$s_1^{(p)}(A^T) = \sup_{y \in \mathbb{R}^m, \|y\|_p = 1} \sup_{x \in \mathbb{R}^N, \|x\|_q = 1} \langle Ax, y \rangle \quad (2.30)$$

Finally, using $\langle Ax, y \rangle = \langle A^T y, x \rangle$ and switching the supremums, we get $s_1^{(q)}(A) = s_1^{(p)}(A^T)$.

We have the following remarks on the above lemma.

Remark 2.10

One can also obtain the above lemma the operator duality on the dual spaces.

Remark 2.11

The above lemma allows us to obtain the probabilistic estimates on $s_1^{(p)}(A)$ for $p > 2$ by taking the transpose of A and using the estimates on $s_1^{(q)}(A^T)$.

Thus using the duality lemma, we obtain.

Theorem 2.12

(Lower tail probability of the largest p -singular value of rectangular matrices, $p > 2$). Let ξ be a pregaussian random variable normalized to have variance 1 and A be an $m \times N$ matrix with i.i.d. copies of ξ in its entries, then for every $p > 2$ and any $\varepsilon > 0$, there exists $\gamma > 0$ such that:

$$P\left(s_1^{(p)}(A) \leq \gamma m^{\frac{p-1}{p}}\right) \leq \varepsilon \quad (2.31)$$

which, γ only depends on p , ε and the pregaussian random variable ξ .

Moreover, we have the upper tail probability of the largest p -singular value of rectangular matrices for $p > 2$.

Theorem 2.13

(Upper tail probability of the largest p -singular value of rectangular matrices, $p > 2$). Let ξ be a pregaussian variable normalized to have variance 1 and A is an $m \times N$ matrix with i.i.d. copies of ξ in its entries, then for every $p > 2$ and any $\varepsilon > 0$, there exists $K > 0$ such that:

$$P\left(s_1^{(p)}(A) \geq K\left(N^{\frac{p-1}{p}} + m^{\frac{1}{2}} N^{\frac{p-2}{2p}}\right)\right) \leq \varepsilon \quad (2.32)$$

where, K only depends on p , ε and the pregaussian variable ξ .

Numerical Experiments

In general, matrix p -norms are, in fact, NP-hard to approximate if $p \neq 1, 2, \infty$, on which one can refer to (Hendrickx and Olshevsky, 2010; Liu, 2014; Higham, 1992). In this section, however, we would like to show the results from some numerically computable experiments on the p -singular value for $p > 1$ and q -singular value for $0 < q \leq 1$ of random matrices.

For $p = 2$, we plot the largest 2-singular value of Gaussian random matrices of size $n \times n$, where n runs

from 1 through 100. Figure 1 this graph shows that the 2-singular value is $O(\sqrt{n})$.

For $p = 1$, in the first numerical experiment we plot the largest 1-singular value of Gaussian random of size $n \times n$, where n runs from 1 through 100. Figure 2 the graph shows that the largest 1-singular value is $O(n)$.

In the second numerical experiment for $p = 1$, we plot the largest 1-singular value of Gaussian random matrices of size $n \times n$, where n runs from 1 through 200. Figure 3 the graph shows that the largest 1-singular value is $O(n)$.

In the third experiment for $p = 1$, we plot the largest 1-singular value of Gaussian random matrices of size

$n \times n$, where n runs from 1 through 400. Figure 4 the graph shows that the largest 1-singular value is $O(n)$.

For $p = \infty$, we plot the largest ∞ -singular value of Gaussian random matrices of size $n \times n$, where n runs from 1 through 500. Figure 5 this graph shows that the ∞ -singular value is $O(n)$.

Higham (1992), the p -norm of a matrix of size m by n was estimated reliably in $O(mn)$ operations and an algorithm that can estimate the p -norm in a specific accuracy, within a factor of $n^{1-\frac{1}{p}}$, was provided. Using this algorithm, we plot the largest 4-singular value of Gaussian random matrices and Bernoulli random matrices of size $m \times n$, where m and n run from 1 through 81 Fig. 6 and 7.

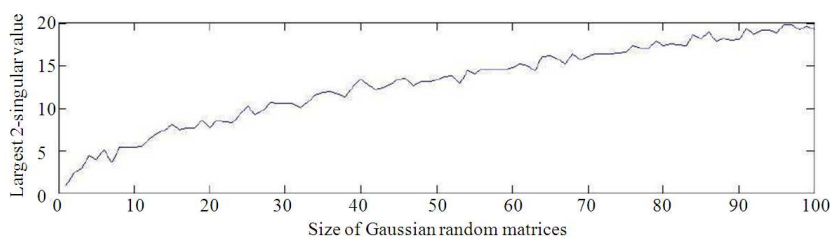


Fig.1. Largest 2-singular value of Gaussian random matrices

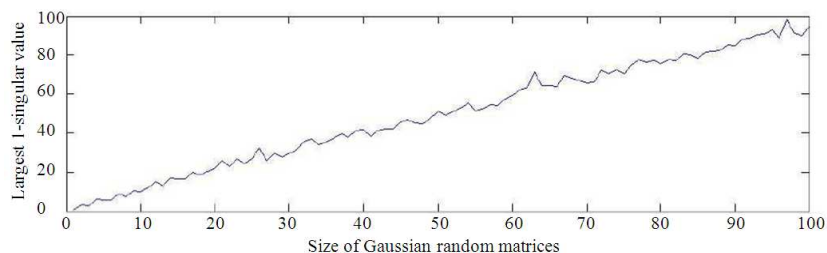


Fig. 2. Largest 1-singular value of Gaussian random matrices: Experiment 1

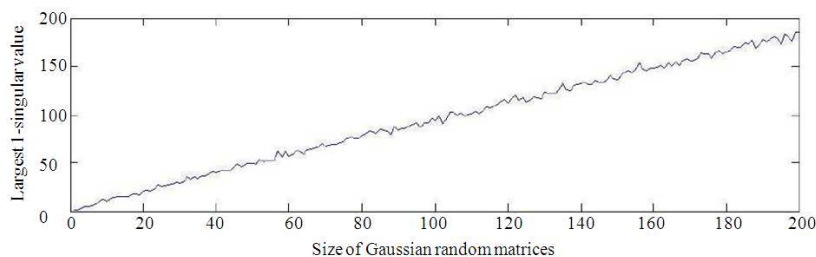


Fig. 3. Largest 1-singular value of Gaussian random matrices: Experiment 2

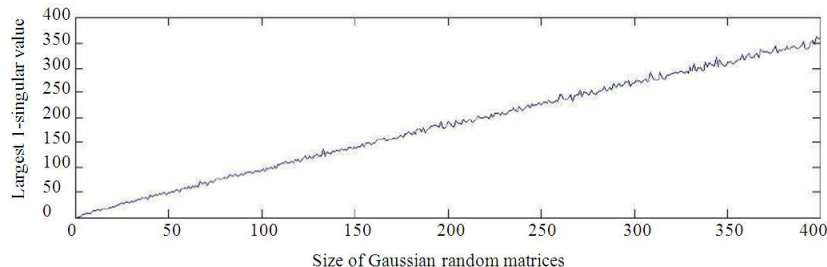


Fig. 4. Largest 1-singular value of Gaussian random matrices: Experiment 3

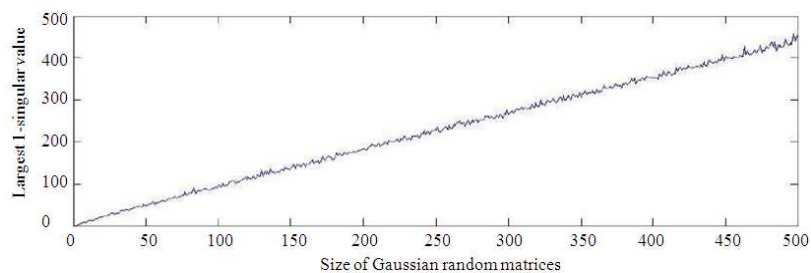


Fig. 5. Largest ∞ -singular value of Gaussian random matrices

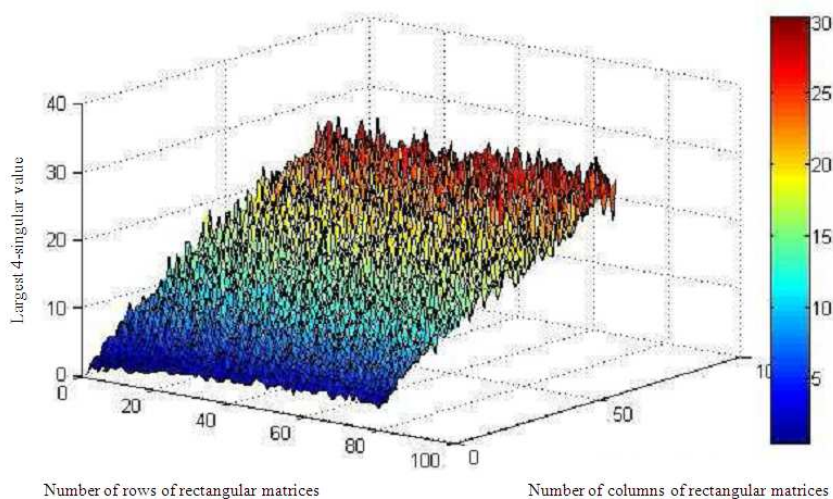


Fig. 6. Largest 4-singular value of Gaussian random matrices

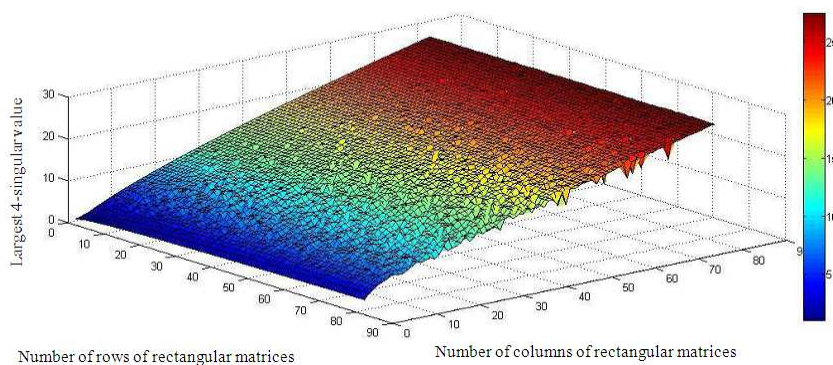


Fig. 7. Largest 4-singular value of Bernoulli random matrices

Acknowledgment

The author would like to thank Prof. M.J. Lai for suggesting the research problem. The author is partially supported by the Air Force Office of Scientific Research under grant AFOSR 9550-12-1-0455.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of

the other authors have read and approved the manuscript and no ethical issues involved.

References

- Bai, Z.D., 1997. Circular law. *Ann. Probab.*, 25: 494-529.
- Bourgain, J., V.H. Vu and P.M. Wood, 2010. On the singularity probability of discrete random matrices. *J. Functional Analysis*, 258: 559-603.
DOI: 10.1016/j.jfa.2009.04.016

- Edelman, A., 1988. Eigenvalues and condition numbers of random matrices. *SIAM J. Matrix Anal. Appl.*, 9: 543-560. DOI: 10.1137/0609045
- Foucart, S. and M.J. Lai, 2009. Sparsest solutions of underdetermined linear systems via ℓ_q -minimization for $0 < q \leq 1$. *Applied Comput. Harmonic Analysis*, 26: 395-407, 2009. DOI: 10.1016/j.acha.2008.09.001
- Foucart, S. and M.J. Lai, 2010. Sparse recovery with pre-Gaussian random matrices. *Studia Math.*, 200: 91-102. DOI: 10.4064/sm200-1-6
- Geman, S., 1980. A limit theorem for the norm of random matrices. *Ann. Probab.*, 8: 252-261. DOI: 10.1214/aop/1176994775
- Girko, V.L., 1985. Circular law. *Theory Probab. Appl.*, 29: 694-706. DOI: 10.1137/1129095
- Hendrickx, J.M. and A. Olshevsky, 2010. Matrix p -norms are NP-hard to approximate if $p \neq 1, 2, \infty$. *SIAM. J. Matrix Anal. Appl.*, 31: 2802-2812. DOI: 10.1137/09076773X
- Higham, N.J., 1992. Estimating the matrix p -norm. *Numerische Mathematik*, 62: 539-555. DOI: 10.1007/BF01396242
- Johansson, K., 2000. Shape fluctuations and random matrices. *Commun. Math. Phys.*, 209: 437-476. DOI: 10.1007/s002200050027
- Johnstone, I.M., 2001. On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.*, 29: 295-327.
- Lai, M.J., S. Li, L.Y. Liu and H. Wang, 2012. Two results on the Schatten p -quasi-norm minimization for low-rank matrix recovery.
- Lai, M.J. and Y. Liu, 2011. The null space property for sparse recovery from multiple measurement vectors. *Applied Comput. Harmonic Anal.*, 30: 402-406. DOI: 10.1016/j.acha.2010.11.002
- Lai, M.J. and Y. Liu, 2014. The probabilistic estimates on the largest and smallest q -singular values of random matrices. *Math. Comp.* DOI: 10.1090/S0025-5718-2014-02895-0
- Latala, R., 2005. Some estimates of norms of random matrices. *Proc. Am. Math. Soc.*, 133: 1273-1282. DOI: 10.1090/S0002-9939-04-07800-1
- Liu, Y., 2011. On the Lagrangian subspaces of complex Minkowski space. *J. Math. Sci. Adv. Appl.*, 7: 87-93.
- Liu, Y., 2013. On the Kähler form of complex l_p space and its Lagrangian subspaces.
- Liu, Y., 2014. Low rank approximations of linear operators in p -norms and their algorithms.
- Oymak, S., K. Mohan, M. Fazel and B. Hassibi, 2011. A simplified approach to recovery conditions for low rank matrices. *Proceedings of the IEEE International Symposium on Information Theory*, Jul. 31-Aug. 5, IEEE Xplore press, St. Petersburg, pp: 2318-2322. DOI: 10.1109/ISIT.2011.6033976
- Pisier, G., 1999. *The Volume of Convex Bodies and Banach Space Geometry*. 1st Edn., Cambridge University Press, Cambridge, ISBN-10: 052166635X, pp: 250.
- Rauhut, H., 2010. *Compressive Sensing and Structured Random Matrices*. In: *Theoretical Foundations and Numerical Methods for Sparse Recovery*, Fornasier, M. (Ed.), Walter de Gruyter, New York, ISBN-10: 3110226154, pp: 1-92.
- Rudelson, M. and R. Vershynin, 2008a. The least singular value of a random square matrix is $O(n^{-1/2})$. *Comptes Rendus Mathematique*, 346: 893-896. DOI: 10.1016/j.crma.2008.07.009
- Rudelson, M. and R. Vershynin, 2008b. The Littlewood-Offord problem and invertibility of random matrices. *Adv. Math.*, 218: 600-633. DOI: 10.1016/j.aim.2008.01.010
- Rudelson, M. and R. Vershynin, 2010. Non-asymptotic theory of random matrices: Extreme singular values. *Proceedings of the International Congress of Mathematicians, (ICM' 10)*, Hindustan Book Agency, New Delhi, pp: 1576-1602.
- Seginer, Y., 2000. The expected norm of random matrices. *Combinatorics Probability Comput.*, 9: 149-166. DOI: 10.1017/S096354830000420X
- Soshnikov, A., 2002. A note on universality of the distribution of the largest eigenvalues in certain sample covariance matrices. *J. Stat. Phys.*, 108:1033-1056. DOI: 10.1023/A:1019739414239
- Soshnikov, A. and Y.V. Fyodorov, 2004. On the largest singular values of random matrices with independent Cauchy entries. *J. Math. Phys.*, 46: 033302-033316. DOI: 10.1063/1.1855932
- Spielman, D. and S.H. Teng, 2002. Smoothed analysis of algorithms. *Proceedings of the International Congress of Mathematicians, (ICM' 02)*, Beijing, pp: 597-606.
- Tao, T. and V. Vu, 2007. On the singularity probability of random Bernoulli matrices. *J. Am. Math. Soc.*, 20: 603-628. DOI: 10.1090/S0894-0347-07-00555-3
- Tao, T. and V. Vu, 2008. Random matrices: The circular law. *Commun. Contemp. Math.*, 10: 261-307. DOI: 10.1142/S0219199708002788
- Tao, T. and V. Vu, 2009. On the permanent of random Bernoulli matrices. *Adv. Math.*, 220: 657-669. DOI: 10.1016/j.aim.2008.09.006
- Tao, T. and V. Vu, 2010. Random matrices: The distribution of the smallest singular values. *Geometric Functional Analysis*, 20: 260-297. DOI: 10.1007/s00039-010-0057-8
- Tao, T., V. Vu and M. Krishnapur, 2010. Random matrices: Universality of ESDs and the circular law. *Ann. Probab.*, 38: 2023-2065. DOI: 10.1214/10-AOP534

Tracy, C.A. and H. Widom 1996. On orthogonal and symplectic matrix ensembles. *Commun. Math. Phys.*, 177: 727-754. DOI: 10.1007/BF02099545

Wigner, E.P., 1958. On the distribution of the roots of certain symmetric matrices. *Ann. Math.*, 67: 325-327.

Yin, Y.Q., Z.D. Bai and P.R. Krishnaiah, 1988. On the limit of the largest eigenvalue of the large dimensional sample covariance matrix. *Probability Theory Related Fields*, 78: 509-521. DOI: 10.1007/BF00353874