Quantitative version of a Silverstein's result

Alexander E. Litvak[1] Susanna Spektor

Abstract

We prove a quantitative version of a Silverstein's Theorem on the 4-th moment condition for convergence in probability of the norm of a random matrix. More precisely, we show that for a random matrix with i.i.d. entries, satisfying certain natural conditions, its norm cannot be small.

Let w be a real random variable with $\mathbb{E}w = 0$ and $\mathbb{E}w^2 = 1$, and let w_{ij}, $i, j \geq 1$ be its i.i.d. copies. For integers n and $p = p(n)$ consider the $p \times n$ matrix $W_n = \{w_{ij}\}_{i\leq p, j\leq n}$, and consider its sample covariance matrix $\Gamma_n := \frac{1}{n} W_n W_n^T$. We also denote by $X_j = (w_{j1}, \ldots, w_{jn})$, $j \leq p$, the rows of W_n.

The questions on behavior of eigenvalues are of great importance in random matrix theory. We refer to [4, 6, 14] for the relevant results, history and references.

In this note we study lower bounds on $\max_{i\leq p} |X_i|$, where $|\cdot|$ denotes the Euclidean norm of a vector, and on the operator (spectral) norms of matrices W_n and Γ_n. Note, as Γ_n is symmetric, its largest singular value $\lambda_{\max}$ is equal to the norm and that in general we have

$$\lambda_{\max}(\Gamma_n) = \|\Gamma_n\| = \frac{1}{n}\|W_n\|^2 \geq \frac{1}{n} \max_{i\leq p} |X_i|^2. \tag{1}$$

Assume that $p(n)/n \to \beta > 0$ as $n \to \infty$. In [19] it was proved that $\mathbb{E}w^4 < \infty$ then $\|\Gamma_n\| \to (1+\sqrt{\beta})^2$ a.s., while in [7] it was shown that $\limsup_{n\to\infty} \|\Gamma_n\| = \infty$ a.s. if $\mathbb{E}w^4 = \infty$.

In [16] Silverstein studied the weak behavior of $\|\Gamma_n\|$. In particular, he proved that assuming $p(n)/n \to \beta > 0$ as $n \to \infty$, $\|\Gamma_n\|$ converges to a non-random quantity (which must be $(1+\sqrt{\beta})^2$) in probability if and only if $n^4\mathbb{P}(|w| \geq n) = o(1)$.

The purpose of this note is to provide the quantitative counterpart of Silverstein's result. More precisely, we want to show an estimate of the type $\mathbb{P}\big(\|\Gamma_n\| \geq K\big) \geq \delta = \delta(K)$ for an arbitrary large K, provided that w has heavy tails (in particular, provided that w does not have 4-th moment). Our proof essentially follows ideas of [16]. It gives a lower bound on $\max_{i\leq p} |X_i|$ as well.

[1]Research partially supported by the E.W.R. Steacie Memorial Fellowship.

Theorem 1. *Let $\alpha \geq 2$, $c_0 > 0$. Let w be a random variable satisfying $\mathbb{E}w = 0$, $\mathbb{E}w^2 = 1$ and*

$$\forall t \geq 1 \qquad \mathbb{P}(|w| \geq t) \geq \frac{c_0}{t^\alpha}. \tag{2}$$

Let $W_n = \{w_{ij}\}_{i\leq p, j\leq n}$ be a $p\times n$ matrix whose entries are i.i.d. copies of w and let $X_i, i \leq p$, be the rows of W_n. Then, for every $K \geq 1$,

$$\mathbb{P}\left(\max_{i\leq p} |X_i| \geq \sqrt{Kn}\right) \geq \min\left\{\frac{c_0 p}{4n^{(\alpha-2)/2}K^{\alpha/2}},\ \frac{1}{2}\right\}. \tag{3}$$

In particular, $\Gamma_n = \frac{1}{n}W_nW_n^T$ satisfies for every $K \geq 1$,

$$\mathbb{P}\left(\|\Gamma_n\| \geq K\right) \geq \min\left\{\frac{c_0 p}{4n^{(\alpha-2)/2}K^{\alpha/2}},\ \frac{1}{2}\right\}.$$

Remark 2. Taking $K = (c_0pn/2)^{2/\alpha}/n$ we observe

$$\mathbb{P}\left(\|W_n\| \geq (c_0pn/2)^{1/\alpha}\right) \geq \mathbb{P}\left(\max_{i\leq p} |X_i| \geq (c_0pn/2)^{1/\alpha}\right) \geq \frac{1}{2}.$$

This estimate seems to be sharp in view of the following result (see Corollary 2 in [5]). Let $0 < \alpha < 4$ and let w be defined by

$$\mathbb{P}(|w| > t) = \min\{1, t^{-\alpha}\} \quad \text{for} \quad t > 0.$$

Let W_n and X_i's be as in Theorem 1. Assume that $p/n \to \beta > 0$ as $n \to \infty$. Then

$$\lim_{n\to\infty} \mathbb{P}\left(\|W_n\| \leq (pn)^{1/\alpha}\, t\right) = \exp(-t^{-\alpha}).$$

Remark 3. If p is proportional to n, say $p = \beta n$, the theorem gives

$$\mathbb{P}\left(\|\Gamma_n\| \geq K\right) \geq \mathbb{P}\left(\max_{i\leq p} |X_i| \geq \sqrt{Kn}\right) \geq \min\left\{\frac{c_0\, \beta}{4n^{(\alpha-4)/2}K^{\alpha/2}},\ \frac{1}{2}\right\},$$

in particular, taking $K = (c_0\beta/2)^{2/\alpha}\, n^{4/\alpha-1}$, we observe

$$\mathbb{P}\left(\|W_n\| \geq (c_0\beta/2)^{1/\alpha}\, n^{2/\alpha}\right) \geq \mathbb{P}\left(\max_{i\leq p} |X_i| \geq (c_0\beta/2)^{1/\alpha}\, n^{2/\alpha}\right) \geq \frac{1}{2}.$$

Remark 4. Note that by Chebychev's inequality one has $\mathbb{P}(|w| \geq t) \leq t^{-2}$. Note also that we use condition (2) in the proof only once, with $t = \sqrt{Kn}$.

Remark 5. If $p \geq (2/c_0)K^{\alpha/2}n^{(\alpha-2)/2}$, then, by condition (2), we have

$$\frac{n}{2}\mathbb{P}(w^2 \geq Kn) \geq \frac{nc_0}{2(Kn)^{\alpha/2}} = \frac{c_0}{2K^{\alpha/2}n^{(\alpha-2)/2}} \geq \frac{1}{p}.$$

Therefore in this case the proof below gives

$$\mathbb{P}(\|\Gamma_n\| \geq K) \geq \mathbb{P}\left(\max_{i\leq p} |X_i| \geq \sqrt{Kn}\right) \geq \frac{1}{2}.$$

In particular, if $\alpha = 4$ and $p \geq (2K^2/c_0)n$ then $\|\Gamma_n\| \geq K$ with probability at least $1/2$.

Before we prove the theorem we would like to mention that last decade many works appeared on non-limit behavior of the norms of random matrices with random entries. In most of them $\max_{i\le p}|X_i|$ appears naturally (or $\sqrt{n}$, when X_i is with high probability bounded by $\sqrt{n}$). For earlier works on Gaussian matrices we refer to [9, 10, 18] and references therein. For the general case of centered i.i.d. $w_{i,j}$ (as in our setting) Seginer [15] proved that

$$\mathbb{E}\|W_n\| \le C\left(\mathbb{E}\max_{i\le p}|X_i| + \mathbb{E}\max_{j\le n}|Y_j|\right),$$

where Y_j, $j \le n$, are the columns of W_n. Later Latała [11] was able to remove the condition that $w_{i,j}$ are identically distributed (his formula involves 4-th moments). Moreover, Mendelson and Paouris [12] have recently proved that for centered i.i.d. $w_{i,j}$ of variance one satisfying $\mathbb{E}|w_{1,1}|^q \le L$ for some $q > 4$ and $L > 0$ with high probability one has

$$\mathbb{E}\|W_n\| \le \max\{\sqrt{p}, \sqrt{n}\} + C(q, L)\min\{\sqrt{p}, \sqrt{n}\}.$$

In [1, 3, 12, 13, 17] matrices with independent columns (which can have dependent coordinates) were investigated. In particular, in [1] (see Theorem 3.13 there) it was shown that if columns of $p \times n$ matrix A satisfy

$$\sup_{q\ge 1}\sup_{i\le p}\sup_{y\in S^{n-1}} \tfrac{1}{q}\left(\mathbb{E}|\langle X_i, y\rangle|^q\right)^{1/q} \le \psi$$

then with probability at least $1 - \exp\left(-c\sqrt{p}\right)$ one has

$$\|A\| \le 6\max_{i\le p}|X_i| + C\psi\sqrt{p} \tag{4}$$

(using Theorem 5.1 in [2] the factor 6 can be substituted by $(1+\varepsilon)$ in which case constants C and c will be substituted with $C\ln(2/\varepsilon)$ and $c\ln(2/\varepsilon)$ correspondingly). Moreover, very recently (4) was extended to the case of matrices whose (independent) columns satisfy

$$\sup_{i\le p}\sup_{y\in S^{n-1}} \left(\mathbb{E}|\langle X_i, y\rangle|^q\right)^{1/q} \le \psi$$

for some $q > 4$ with the constant C depending on q ([8]).

Proof of the Theorem. By (1) the "In particular" part of the Theorem follows immediately from (3). Thus, it is enough to prove (3).

Since $X_1, \ldots, X_p$ are i.i.d. random vectors and since $|X_1|^2$ is distributed as $\sum_{j=1}^{n} w_{1,j}^2$, we observe for every $K \ge 1$,

$$\begin{aligned}\mathbb{P}\left(\max_{i\le p}|X_i| \ge \sqrt{Kn}\right) &= 1 - \mathbb{P}\left(\max_{i\le p}|X_i| < \sqrt{Kn}\right) = 1 - \mathbb{P}\left(\left\{\forall i \,:\, |X_i| < \sqrt{Kn}\right\}\right)\\ &= 1 - \left(\mathbb{P}(|X_1| < \sqrt{Kn})\right)^p = 1 - \left(\mathbb{P}\left(\sum_{j=1}^{n} w_{1,j}^2 < Kn\right)\right)^p.\end{aligned} \tag{5}$$

For $j \le n$ consider the events $A_j := \{w_{1,j}^2 \ge nK\}$. Clearly,

$$A := \left\{ \sum_{j=1}^n w_{1,j}^2 \ge nK \right\} \supset \bigcup_{j=1}^n A_j.$$

By the inclusion-exclusion principle, we have

$$\begin{aligned}
\mathbb{P}(A) \ge \mathbb{P}\left\{ \bigcup_{j=1}^n A_j \right\} &\ge \sum_{j=1}^n \mathbb{P}(A_j) - \sum_{j \ne k} \mathbb{P}\left(A_j \cap A_k\right) = \sum_{j=1}^n \mathbb{P}\left(w^2 \ge nK\right) - \sum_{j \ne k} \Big(\mathbb{P}\left(w^2 \ge nK\right)\Big)^2 \\
&= n\mathbb{P}\left(w^2 \ge nK\right) - \frac{n^2 - n}{2} \Big(\mathbb{P}\left(w^2 \ge nK\right)\Big)^2 \\
&= \frac{n}{2} \mathbb{P}(w^2 \ge nK)(2 - (n-1)\mathbb{P}(w^2 \ge nK)).
\end{aligned}$$

By Chebychev's inequality we have $\mathbb{P}(w^2 \ge nK) \le \dfrac{1}{nK}$, hence, $2 - (n-1)\mathbb{P}(w^2 \ge nK) \ge 1$. Thus, by (5),

$$\mathbb{P}\left(\max_{i \le p} |X_i| \ge \sqrt{Kn} \right) \ge 1 - \left(1 - \mathbb{P}\left(\frac{1}{n} \sum_{j=1}^n w_{1,j}^2 \ge K \right) \right)^p \ge 1 - \Big(1 - \frac{n}{2}\mathbb{P}\left(w^2 \ge nK\right)\Big)^p.$$

If $\dfrac{n}{2}\mathbb{P}(w^2 \ge Kn) \ge \dfrac{1}{p}$, then

$$\mathbb{P}\left(\max_{i \le p} |X_i| \ge \sqrt{Kn} \right) \ge 1 - \left(1 - \frac{1}{p}\right)^p \ge 1 - \frac{1}{e} \ge \frac{1}{2}.$$

Finally assume that

$$\frac{n}{2}\mathbb{P}(w^2 \ge Kn) \le \frac{1}{p}. \tag{6}$$

Using that $(1-x)^p \le (1+px)^{-1}$ on $[0,1]$, we get

$$\mathbb{P}\left(\max_{i \le p} |X_i| \ge \sqrt{Kn} \right) \ge 1 - \frac{1}{(np/2)\mathbb{P}(w^2 \ge Kn) + 1}.$$

Applying condition (2) with $t = \sqrt{Kn}$ and using (6) again, we observe

$$1 \ge \frac{np}{2}\mathbb{P}(w^2 \ge Kn) \ge \frac{np}{2} \frac{c_0}{(Kn)^{\alpha/2}}.$$

Thus,

$$\mathbb{P}\left(\max_{i \le p} |X_i| \ge \sqrt{Kn} \right) \ge \frac{c_0 p}{4n^{(\alpha-2)/2} K^{\alpha/2}},$$

which completes the proof. □

Acknowledgment. We are grateful to A. Pajor for useful comments and to S. Sodin for bringing reference [5] to our attention.

References

[1] R. Adamczak, A.E. Litvak, A. Pajor and N. Tomczak-Jaegermann, *Quantitative estimates of the convergence of the empirical covariance matrix in log-concave ensembles*, J. Amer. Math. Soc. **23** (2010), 535–561.

[2] R. Adamczak, A.E. Litvak, A. Pajor and N. Tomczak-Jaegermann, *Restricted isometry property of matrices with independent columns and neighborly polytopes by random sampling*, Constructive Approximation, **34** (2011), 61–88.

[3] R. Adamczak, A.E. Litvak, A. Pajor and N. Tomczak-Jaegermann, *Sharp bounds on the rate of convergence of empirical covariance matrix*, C.R. Math. Acad. Sci. Paris, **349** (2011), 195–200.

[4] G.W. Anderson, A. Guionnet, O. Zeitouni, *An introduction to random matrices*, Cambridge Studies in Advanced Mathematics, 118. Cambridge University Press, Cambridge, 2010.

[5] A. Auffinger, G. Ben Arous, S. Péché, *Sandrine Poisson convergence for the largest eigenvalues of heavy tailed random matrices*, Ann. Inst. Henri Poincar Probab. Stat. **45** (2009), 589–610.

[6] Z.D. Bai, J.W. Silverstein, *Spectral analysis of large dimensional random matrices.* Second edition. Springer Series in Statistics. Springer, New York, 2010.

[7] Z. D. Bai, J. Silverstein, Y. Q. Yin, A note on the largest eigenvalue of a large dimensional sample covariance matrix, *Journal of Multivariate Analysis*, Vol 26, **2** (1988), 166–168.

[8] O. Guédon, A.E. Litvak, A. Pajor and N. Tomczak-Jaegermann, *Restricted isometry property for random matrices with heavy tailed columns*, C.R. Math. Acad. Sci. Paris, to appear.

[9] Y. Gordon, *On Dvoretzky's theorem and extensions of Slepian's lemma,* Israel seminar on geometrical aspects of functional analysis (1983/84), II, Tel Aviv Univ., Tel Aviv, 1984.

[10] Y. Gordon, *Some inequalities for Gaussian processes and applications,* Israel J. Math. 50 (1985), 265–289.

[11] R. Latala, *Some estimates of norms of random matrices*, Proc. Amer. Math. Soc. 133 (2005), 1273–1282

[12] S. Mendelson, G. Paouris, *On generic chaining and the smallest singular values of random matrices with heavy tails*, Journal of Functional Analysis, 262 (2012), 3775-3811.

[13] S. Mendelson, G. Paouris, *On the singular values of random matrices*, Journal of the European Mathematical Society, 16 (2014), 823–834.

[14] L. Pastur, M. Shcherbina, *Eigenvalue distribution of large random matrices*. Mathematical Surveys and Monographs, 171. American Mathematical Society, Providence, RI, 2011.

[15] Y. Seginer, *The expected norm of random matrices*, Combin. Probab. Comput. 9 (2000), 149–166.

[16] J. Silverstein, On the weak limit of the largest eigenvalue of a large dimensional sample covariance matrix, *J. of Multivariate Anal.*, **30** (1989), 2, 307–311.

[17] N. Srivastava, R. Vershynin, *Covariance estimation for distributions with 2+epsilon moments*, Annals of Probability 41 (2013), 3081–3111.

[18] S. J. Szarek, *Condition numbers of random matrices*, J. Complexity 7 (1991), no. 2, 131–149.

[19] Y. Q. Yin, Z. D. Bai, P. R. Krishnaiah, On the limit of the largest eigenvalue of the large dimensional sample covariance matrix, *Probab. Th. Rel. Fields.*, **78** (1988), 509–527.

Alexander Litvak and Susanna Spektor,
Dept. of Math. and Stat. Sciences,
University of Alberta,
Edmonton, Alberta, Canada, T6G 2G1.
E-mails: `aelitvak@gmail.com` and `sanaspek@gmail.com`