Reprinted from JOURNAL OF MULTIVARIATE ANALYSIS

Vol. 30, No. 2, August 1989
Printed in Belgium

On the Weak Limit of the Largest Eigenvalue of a Large Dimensional Sample Covariance Matrix

JACK W. SILVERSTEIN*

North Carolina State University

Communicated by the Editors

Let $\{w_{ij}\}$, $i, j = 1, 2, \ldots$, be i.i.d. random variables and for each n let $M_n = (1/n) W_n W_n^{\mathrm{T}}$, where $W_n = (w_{ij})$, $i = 1, 2, \ldots, p$; $j = 1, 2, \ldots, n$; $p = p(n)$, and $p/n \to y > 0$ as $n \to \infty$. The weak behavior of the largest eigenvalue of M_n is studied. The primary aim of the paper is to show that the largest eigenvalue converges in probability to a nonrandom quantity if and only if $\mathbf{E}(w_{11}) = 0$ and $n^4 \mathbf{P}(|\omega_{11}| \geqslant n) = o(1)$, the limit being $(1 + \sqrt{y})^2 \mathbf{E}(w_{11}^2)$. © 1989 Academic Press, Inc.

1. INTRODUCTION

Let $\{w_{ij}\}$, $i, j = 1, 2, \ldots$, be i.i.d. random variables and for each n let $M_n = (1/n) W_n W_n^{\mathrm{T}}$, where $W_n = (w_{ij})$, $i = 1, 2, \ldots, p$; $j = 1, 2, \ldots, n$; $p = p(n)$, and $p/n \to y > 0$ as $n \to \infty$. The matrix M_n can be viewed as the sample covariance matrix of n samples of a p-dimensional vector containing i.i.d. components, where p and n are large but on the same order of magnitude. Denote the largest eigenvalue of M_n by $\lambda_{\max}(M_n)$. In [3] it is proven that if $\mathbf{E}(w_{11}^4) < \infty$ and $\mathbf{E}(w_{11}) = 0$ then $\lambda_{\max}(M_n) \to^{\text{a.s.}} (1 + \sqrt{y})^2 \mathbf{E}(w_{11}^2)$ as $n \to \infty$, while in [1] it is shown that $\limsup_n \lambda_{\max}(M_n) = \infty$ almost surely if $\mathbf{E}(w_{11}^4) = \infty$. Thus the almost sure behavior of $\lambda_{\max}(M_n)$ is essentially understood. The aim of this paper is to establish its weak limiting behavior. We will prove the following

THEOREM. *We have*

(a) $\lambda_{\max}(M_n)$ *converges in probability to a nonrandom quantity if and only if* $\mathbf{E}(w_{11}) = 0$ *and* $n^4 \mathbf{P}(|w_{11}| \geqslant n) = o(1)$, *the limit being* $(1 + \sqrt{y})^2 \mathbf{E}(w_{11}^2)$.

Received May 17, 1988; revised June 1988.

AMS 1980 subject classification: Primary 60F05; Secondary 62H99.

Key words and phrases: largest eigenvalue of sample covariance matrix, convergence in probability, bounded in probability.

* Supported by the National Science Foundation under Grant DMS-8603966.

0047-259X/89 $3.00

(b) *If* $\limsup_n n^4\mathbf{P}(|w_{11}| \geqslant n) > 0$, *then for any* $K > 0$ $\limsup_n \mathbf{P}(\lambda_{\max}(M_n) \geqslant K) > 0$.

(c) *If* $\limsup_n n^4\mathbf{P}(|w_{11}| \geqslant n) = \infty$, *then for any* $K > 0$ $\limsup_n \mathbf{P}(\lambda_{\max}(M_n) \geqslant K) = 1$.

(d) *If* $\mathbf{E}(w_{11}) = 0$ *and* $\limsup_n n^4\mathbf{P}(|w_{11}| \geqslant n) < \infty$, *then* $\{\lambda_{\max}(M_n)\}_{n=1}^{\infty}$ *is bounded in probability.*

Thus convergence in probability holds only under a slight weakening of the condition needed for almost sure convergence. For case (d) we remark here that, due to the known limiting behavior of the empirical distribution function of the eigenvalues of M_n [2], we have $\liminf_n \lambda_{\max}(M_n) \geqslant (1+\sqrt{y})^2\,\mathbf{E}(w_{11}^2)$ almost surely, implying any limiting distribution of a weakly converging subsequence of $\{\lambda_{\max}(M_n)\}_{n=1}^{\infty}$ must have mass in $[(1+\sqrt{y})^2\,\mathbf{E}(w_{11}^2), \infty)$. Other than that, no additional information is known, for example, whether $\lambda_{\max}(M_n)$ converges in distribution or how the distribution of any weakly convergent subsequence of $\{\lambda_{\max}(M_n)\}_{n=1}^{\infty}$ depends on w_{11}.

The proof of the theorem will be given in the next section. It relies on part of the proof of the main theorem in [3], as well as the fact that $\lambda_{\max}(M_n)$ is bounded below by each diagonal element of M_n.

2. Proof of the Theorem

Assume $\mathbf{E}(w_{11}) = 0$ and $n^4\mathbf{P}(|w_{11}| \geqslant n) = o(1)$. Without loss of generality we may assume $\mathbf{E}(w_{11}^2) = 1$. It then follows that $n^2\mathbf{P}(|w_{11}| \geqslant \varepsilon\sqrt{n}) = o(1)$ for every $\varepsilon > 0$. From this it is straightforward to construct a sequence $\{\delta_n\}_{n=1}^{\infty}$ of positive numbers such that $\delta_n \to 0$, $\delta_n \log n \to \infty$, and $n^2\mathbf{P}(|w_{11}| \geqslant \delta_n\sqrt{n}) \to 0$ as $n \to \infty$.

Let $\hat{w}_{ij} = \hat{w}_{ij}(n) = w_{ij} I_{(|w_{ij}| < \delta_n\sqrt{n})}$ (I_A being the indicator function on the set A), and define $\hat{M}_n = (1/n)\,\hat{W}_n \hat{W}_n^{\mathrm{T}}$, where $\hat{W}_n$ is $p \times n$ with (i, j)th element $\hat{w}_{ij}$. Then

$$\mathbf{P}(\lambda_{\max}(M_n) \neq \lambda_{\max}(\hat{M}_n)) \leqslant np\mathbf{P}(|w_{11}| \geqslant \delta_n\sqrt{n}) \to 0 \qquad \text{as} \quad n \to \infty. \tag{2.1}$$

Let $\tilde{w}_{ij} = \tilde{w}_{ij}(n) = \hat{w}_{ij} - \mathbf{E}(\hat{w}_{ij})$, and define $\tilde{M}_n = (1/n)\,\tilde{W}_n \tilde{W}_n^{\mathrm{T}}$, where $\tilde{W}_n$ is $p \times n$ with (i, j)th element $\tilde{w}_{ij}$. Let $\|B\|$ denote the spectral norm of any matrix B (that is, $\|B\| = \lambda_{\max}^{1/2}(BB^{\mathrm{T}})$), and let $\mathbf{1}_m$ denote the m-dimensional vector consisting of 1's. Then

$$|\lambda_{\max}^{1/2}(\tilde{M}_n) - \lambda_{\max}^{1/2}(\hat{M}_n)| \leqslant \frac{1}{\sqrt{n}}\,|\mathbf{E}(\hat{w}_{11})|\;\|\mathbf{1}_p \mathbf{1}_n^{\mathrm{T}}\| = \sqrt{p}\,|\mathbf{E}(\hat{w}_{11})|.$$

Since $\mathbf{E}(w_{11})=0$, $\mathbf{E}(w_{11}^2)=1$, we have

$$|\mathbf{E}(\hat{w}_{11})| = |\mathbf{E}(w_{11} I_{(|w_{11}| \geqslant \delta_n \sqrt{n})})| \leqslant \mathbf{P}^{1/2}(|w_{11}| \geqslant \delta_n \sqrt{n}).$$

Therefore,

$$|\lambda_{\max}^{1/2}(\tilde{M}_n) - \lambda_{\max}^{1/2}(\hat{M}_n)| \to 0 \qquad \text{as} \quad n \to \infty. \tag{2.2}$$

From (2.1) and (2.2) we see that

$$\lambda_{\max}(M_n) - \lambda_{\max}(\tilde{M}_n) \xrightarrow{\text{i.p.}} 0 \qquad \text{as} \quad n \to \infty.$$

At this point the same arguments in Section 4 of [3] can be applied to $\lambda_{\max}(\tilde{M}_n)$, since the assumptions needed on $\tilde{w}_{ij}$ are identical. Therefore, we have

$$\lambda_{\max}(\tilde{M}_n) \xrightarrow{\text{a.s.}} (1+\sqrt{y})^2 \qquad \text{as} \quad n \to \infty,$$

and the sufficiency part of (a) is proven.

Suppose $n^4 \mathbf{P}(|w_{11}| \geqslant n) = o(1)$ but $\mathbf{E}(w_{11}) = a \neq 0$. Then the matrix $M_n' \equiv (1/n)(W_n - \mathbf{1}_p \mathbf{1}_n^{\mathrm{T}})(W_n - \mathbf{1}_p \mathbf{1}_n^{\mathrm{T}})^{\mathrm{T}}$ satisfies the conditions of (a). Moreover, we have

$$\left| \lambda_{\max}^{1/2}(M_n) - \left\| \frac{1}{\sqrt{n}}\, a \mathbf{1}_p \mathbf{1}_n^{\mathrm{T}} \right\| \right| \leqslant \lambda_{\max}^{1/2}(M_n').$$

However, $\|(1/\sqrt{n})\, a \mathbf{1}_p \mathbf{1}_n^{\mathrm{T}}\| = \sqrt{p}\,|a|$. Therefore, $\lambda_{\max}(M_n)/pa^2 \to^{\text{i.p.}} 1$ as $n \to \infty$. This verifies the necessity of $\mathbf{E}(w_{11}) = 0$ in (a).

Assume now $\limsup_n n^4 \mathbf{P}(|w_{11}| \geqslant n) > 0$. For any $K > 0$ we have

$$\begin{aligned} \mathbf{P}(\lambda_{\max}(M_n) \geqslant K) &\geqslant \mathbf{P}\left(\max_i \frac{1}{n} \sum_{j=1}^{n} w_{ij}^2 \geqslant K \right) \\ &= 1 - \left(1 - \mathbf{P}\left(\frac{1}{n} \sum_{j=1}^{n} w_j^2 \geqslant K \right) \right)^p, \end{aligned} \tag{2.3}$$

where $w_1, w_2, \ldots$ are i.i.d. having the same distribution as w_{11}. We need then to investigate the limiting behavior of $n\mathbf{P}((1/n)\sum_{j=1}^{n} w_j^2 \geqslant K)$. Since $\mathbf{E}(w_{11}^2) = 1$ we have $n\mathbf{P}(w_{11}^2 \geqslant nK) = o(1)$ (we remark here the need for the second moment of w_{11} to be finite. If $\mathbf{E}(w_{11}^2)$ were infinite, then from the strong law of large numbers applied to $(1/n)\sum_{j=1}^{n} w_{1j}^2$ we would have $\lambda_{\max}(M_n) \to^{\text{a.s.}} \infty$ as $n \to \infty$). We have

$$\mathbf{P}\left(\frac{1}{n}\sum_{j=1}^{n} w_j^2 \geqslant K\right)$$
$$\geqslant \mathbf{P}\left(\bigcup_j (w_j^2 \geqslant nK)\right)$$
$$\geqslant n\mathbf{P}(w_{11}^2 \geqslant nK) - \frac{n(n-1)}{2}\mathbf{P}^2(w_{11}^2 \geqslant nK) \geqslant \frac{n}{2}\mathbf{P}(w_{11}^2 \geqslant nK) \tag{2.4}$$

for n sufficiently large.

It is a simple matter to show $\limsup_n n^2\mathbf{P}(w_{11}^2 \geqslant nK) > 0$, and if $\limsup_n n^4\mathbf{P}(|w_{11}| \geqslant n) = \infty$, then $\limsup_n n^2\mathbf{P}(w_{11}^2 \geqslant nK) = \infty$. Therefore, from (2.3) and (2.4) we get $\limsup_n \mathbf{P}(\lambda_{\max}(M_n) \geqslant K) > 0$ for any positive K which verifies (b) and the necessity of $n^4\mathbf{P}(|w_{11}| \geqslant n) = o(1)$ in (a). Moreover, we get $\limsup_n \mathbf{P}(\lambda_{\max}(M_n) \geqslant K) = 1$ if $\limsup_n n^4\mathbf{P}(|w_{11}| \geqslant n) = \infty$, proving (c).

Finally, assume $\limsup_n n^4\mathbf{P}(|w_{11}| \geqslant n) < \infty$. Define $\hat{w}_{ij}$, $\hat{W}_n$, $\hat{M}_n$, $\tilde{w}_{ij}$, $\tilde{W}_n$, $\tilde{M}_n$ as above but with $\delta_n = 1$ for each n. We immediately get (2.2), since $n\mathbf{P}(|w_{11}| \geqslant \sqrt{n}) = o(1)$. We have

$$|\lambda_{\max}^{1/2}(M_n) - \lambda_{\max}^{1/2}(\hat{M}_n)|$$
$$\leqslant \frac{1}{\sqrt{n}}\|W_n - \hat{W}_n\| = \lambda_{\max}^{1/2}\left(\frac{1}{n}(W_n - \hat{W}_n)(W_n - \hat{W}_n)^{\mathrm{T}}\right)$$
$$\leqslant \left(\frac{1}{n}\operatorname{tr}(W_n - \hat{W}_n)(W_n - \hat{W}_n)^{\mathrm{T}}\right)^{1/2}.$$

Therefore, for any $K > 0$ we get from Chebyshev's inequality

$$\mathbf{P}(|\lambda_{\max}^{1/2}(M_n) - \lambda_{\max}^{1/2}(\hat{M}_n)| \geqslant K)$$
$$\leqslant \frac{\mathbf{E}(\sum_{i=1}^{p}\sum_{j=1}^{n} w_{ij}^2 I_{(w_{ij}^2 \geqslant n)})}{nK^2} = \frac{p\mathbf{E}(w_{11}^2 I_{(w_{11}^2 \geqslant n)})}{K^2}. \tag{2.5}$$

Using integration by parts, it is straightforward to show for any random variable X and $m > 1$

$$n^{m-1}\mathbf{E}(|X| I_{(|X| \geqslant n)}) \leqslant \frac{m}{m-1}\sup_{x \in [n, \infty)} x^m \mathbf{P}(|X| \geqslant x).$$

Therefore, with $m = 2$ we conclude from (2.2) and (2.5) that

$$\{\lambda_{\max}^{1/2}(M_n) - \lambda_{\max}^{1/2}(\tilde{M}_n)\}_{n=1}^{\infty} \tag{2.6}$$

is bounded in probability.

Now, the main part of the proof in [3] of showing $\lambda_{\max}(\tilde{M}_n) \to^{\text{a.s.}} (1+\sqrt{y})^2$ for an appropriately chosen sequence $\{\delta_n\}_{n=1}^{\infty}$ involves proving for any $z > (1+\sqrt{y})^2$,

$$\sum_{n=1}^{\infty} \mathbf{E}\left(\left(\frac{\lambda_{\max}(\tilde{M}_n)}{z}\right)^{k_n}\right) < \infty \tag{2.7}$$

for appropriately chosen $\{k_n\}_{n=1}^{\infty}$. Of course, (2.7) implies

$$\mathbf{P}(\lambda_{\max}(\tilde{M}_n) > z \text{ infinitely often}) = 0.$$

However, if $\delta_n = 1$ and $k_n = [\log n]$ (where $[x]$ is the integer part of x), then the same arguments in [3] can be used to show (2.7) for all z sufficiently large. Therefore, for some $z > 0$, $\limsup_n \lambda_{\max}(\tilde{M}_n) \leqslant z$ almost surely. This together with (2.6) being bounded in probability give us (d).

Acknowledgment

The author thanks the referee for some helpful suggestions.

References

[1] Bai, Z. D., Silverstein, J. W., and Yin, Y. Q. (1988). A note on the largest eigenvalue of a large dimensional sample covariance matrix. *J. Multivariate Anal.* **26** 166–168.

[2] Yin, Y. Q. (1986). Limiting spectral distribution for a class of random matrices, *J. Multivariate Anal.* **20** 50–68.

[3] Yin, Y. Q., Bai, Z. D., and Krishnaiah, P. R. (1988). On limit of the largest eigenvalue of the large dimensional sample covariance matrix. *Probab. Theory Related Fields* **78** 509–521.

Printed by Catherine Press, Ltd., Tempelhof 41, B-8000 Brugge, Belgium