Reprinted from JOURNAL OF MULTIVARIATE ANALYSIS
All Rights Reserved by Academic Press, New York and London

Vol. 26, No. 2, August 1988
Printed in Belgium

A Note on the Largest Eigenvalue of a Large Dimensional Sample Covariance Matrix

Z. D. BAI

University of Pittsburgh

JACK W. SILVERSTEIN*

North Carolina State University

AND

Y. Q. YIN

University of Arizona

Communicated by the Editors

Let $\{v_{ij};\ i, j=1, 2, ...\}$ be a family of i.i.d. random variables with $E(v_{11}^4)=\infty$. For positive integers p, n with $p=p(n)$ and $p/n \to y>0$ as $n \to \infty$, let $M_n=(1/n)\ V_n V_n^{\mathrm{T}}$, where $V_n=(v_{ij})_{1 \leqslant i \leqslant p, 1 \leqslant j \leqslant n}$, and let $\lambda_{\max}(n)$ denote the largest eigenvalue of M_n. It is shown that $\overline{\lim}_n\ \lambda_{\max}(n)=\infty$ a.s. This result verifies the boundedness of $E(v_{11}^4)$ to be the weakest condition known to assure the almost sure convergence of $\lambda_{\max}(n)$ for a class of sample covariance matrices. © 1988 Academic Press, Inc.

For each $n=1, 2, 3, ...$ let $V_n=(v_{ij}(n))$ be a $p \times n$ matrix, where the $v_{ij}(n)$'s for i, j are independent and are identically distributed for all i, j, n with $E(v_{11})=0$, $E(v_{11}^2)=1$, and $p=p(n)$, with $p/n \to y>0$ and $n \to \infty$. Let $M_n=(1/n)\ V_n V_n^{\mathrm{T}}$. The matrix M_n is then the sample covariance matrix of n samples of a p-dimensional random vector.

Several papers have studied the asymptotic behavior of $\lambda_{\max}(n)$, the largest eigenvalue of M_n [1–3]. In [2] it is shown that $\lambda_{\max}(n)$ converges a.s. (i.p.) to $(1+\sqrt{y})^2$ if $E(|v_{11}|^{6+\delta})<\infty$ $(E(|v_{11}|^{4+\delta}<\infty)$ for any positive δ. In [3] it is shown that if V_n is the first p rows and n columns of a doubly infinite array of v_{ij}'s, then $E(v_{11}^4)<\infty$ implies $\lambda_{\max}(n) \to^{\text{a.s.}}$

Received June 5, 1986; revised February 1, 1988.
AMS 1980 subject classifications: Primary 60F15; Secondary 62H99.
Key words and phrases: sample covariance matrix, largest eigenvalue.
* Supported by National Science Foundation Grant DMS-8603966.

0047-259X/88 $3.00

$(1+\sqrt{y})^2$ as $n \to \infty$. In this paper we will show the boundedness of $E(v_{11}^4)$ to be the weakest condition on the limiting behavior of $\lambda_{\max}(n)$ for doubly infinite arrays. In fact, we will show:

If $\{v_{ij};\, i, j=1, 2, ...\}$ is a family of i.i.d. random variables with $E(v_{11}^4)=\infty$ and M_n, $\lambda_{\max}(n)$ are defined as above with $V_n=(v_{ij})_{1 \leqslant i \leqslant p, 1 \leqslant j \leqslant n}$, $p=p(n)$, $p/n \to y>0$ as $n \to \infty$, then

$$\varlimsup_n \lambda_{\max}(n)=\infty \qquad \text{a.s.}$$

The result will be established by showing for any $M>0$

$$P(\lambda_{\max}(n) \geqslant M^2 \text{ i.o.})=1. \tag{1}$$

Using the fact that the largest eigenvalue of any symmetric matrix is bounded below by the largest diagonal element of the matrix, we have

$$\lambda_{\max}(n) \geqslant \frac{1}{n} v_{ij}^2, \qquad \text{for any} \quad i \in \{1, 2, ..., p(n)\}, j \in \{1, 2, ..., n\}.$$

Defining $A_k=\{v_{ij};\, 2^{k-1}<j \leqslant 2^k,\, 1 \leqslant i \leqslant p(2^k)\}$ for $k=1, 2, ...,$ we therefore have $P(\lambda_{\max}(n) \geqslant M^2 \text{ i.o.}) \geqslant P(\exists v_{ij} \in A_k \text{ s.t. } |v_{ij}| \geqslant M2^{k/2}$ for infinitely many k's).

By the Borel–Cantelli theorem, (1) will follow if we can show

$$\sum_k P\,(\exists v_{ij} \in A_k \text{ s.t. } |v_{ij}| \geqslant M2^{k/2})=\infty. \tag{2}$$

The left side of (2) is

$$\sum_k 1-P(|v_{11}|<M2^{k/2})^{p(2^k)2^{k-1}}.$$

Since $p(n)/n \to y>0$ as $n \to \infty$ we have for any positive $\underline{y}<y$,

$$p(2^k)\,2^{k-1}>\tfrac{1}{2}\underline{y}4^k \text{ for all } k \text{ sufficiently large.}$$

Therefore, (1) follows if

$$\sum_k 1-P(|v_{11}|<M2^{k/2})^{1/2\,\underline{y}4^k}=\infty \qquad (0<\underline{y}<y). \tag{3}$$

Using the relation $E(v_{11}^4) \leqslant \sum_k E(v_{11}^4 I_{(M2^{k/2} \leqslant |v_{11}|<M2^{(k+1)/2})})+M^4$, I_A being the indicator function on the set A, we have

$$\sum_k \tfrac{1}{2}\underline{y}4^k P(|v_{11}| \geqslant M2^{k/2})=\infty, \tag{4}$$

since $E(v_{11}^4)=\infty$.

Using the relation $(1-a)^b \leqslant e^{-ab}$, valid for $a \in [0, 1]$, $b \geqslant 0$ and properties of finite products, it is straightforward to show

$$\sum_k 1-(1-a_k)^{b_k} < \infty \Rightarrow \sum_k a_k b_k < \infty \qquad (a_k \in [0, 1], b_k \geqslant 0).$$

Therefore, (4) implies (3) and we are done.

References

[1] Geman, S. (1980). A limit theorem for the norm of random matrices. *Ann. Probab.* **8** 252–261.

[2] Silverstein, J. W. (1984). On the largest eigenvalue of a large dimensional sample covariance matrix. Technical Report.

[3] Yin, Y. Q., Bai, Z. D., and Krishnaiah, P. R. (1984). *On the Limit of the Largest Eigenvalue of the Large Dimensional Sample Covariance Matrix.* Technical Report 84-44, Center for Multivariate Analysis, University of Pittsburgh.

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