COMMUN. STATIST.-STOCHASTIC MODELS, 10(3), 525-532 (1994)

THE SPECTRAL RADII AND NORMS OF LARGE DIMENSIONAL NON-CENTRAL RANDOM MATRICES

Jack W. Silverstein

Department of Mathematics, Box 8205
North Carolina State University
Raleigh, North Carolina 27695-8205
e-mail: jack@jack.math.ncsu.edu

Key words and phrases: Random matrix, spectral norm, spectral radius.

ABSTRACT

Consider a matrix made up of i.i.d. random variables with positive mean and finite fourth moment. Results are given on its spectral norm and (if it is square) spectral radius as the dimension increases.

1. INTRODUCTION

In your favorite computer language, create a 100×100 matrix U full of i.i.d. random variables, uniformly distributed on $(0, 1)$. Compute λ, the Perron eigenvalue, the real eigenvalue of U equal to the spectral radius (the maximum, in absolute value, of its eigenvalues), guaranteed to exist for positive matrices. It turns out to be near 50. Why? The answer depends on the following result.

Theorem 1.1 ([1],[4]). For $n = 1, 2, \ldots$, and $s = s(n)$ for which $n/s \to y > 0$ as $n \to \infty$, let $V_n = (v_{ij})$, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, s$, where v_{ij}, $i\,j = 1, 2, \ldots, \infty$, are i.i.d. random variables with $\mathsf{E}(v_{1\,1}) = 0$ and $\mathsf{E}(v_{1\,1}^2) = \sigma^2 < \infty$. Then the spectral norm $\|\frac{1}{\sqrt{s}} V_n\|$ (where for any rectangular matrix A, $\|A\|$ equals the square root of the largest eigenvalue of AA^T) converges a.s. to $(1 + \sqrt{y})\sigma$ as $n \to \infty \iff \mathsf{E}(v_{1\,1}^4) < \infty$. If $\mathsf{E}(v_{1\,1}^4) = \infty$, then $\limsup_n \|\frac{1}{\sqrt{s}} V_n\| = \infty$ a.s.

The matrix U is of the form $V_n + \mu n e_n e_n^T$, where V_n is as above with $s = n$, $\mu > 0$, and $e_n = \frac{1}{\sqrt{n}}(1, 1, \ldots, 1)^T$. View U as a perturbation of $\mu n e_n e_n^T$, a rank-one matrix with positive eigenvalue μn, and exploit a perturbation theorem, such as Corollary 6.3.4 in [3]: $\hat{\lambda}$ eigenvalue of $A + E$, A normal $\Longrightarrow$ the existence of eigenvalue λ_i of A such that $|\hat{\lambda} - \lambda_i| \leq \|E\|$. Thus, when μn is eventually greater than $2\|V_n\|$ (which occurs a.s., since $\|V_n\| \sim 2\sigma\sqrt{n}$ a.s.), a simple continuity argument applied to the eigenvalues of $tV_n + \mu n e_n e_n^T$, $t \in [0, 1]$, will yield λ_n, the largest (in absolute value) eigenvalue of $V_n + \mu n e_n e_n^T$, to be real, positive, with multiplicity 1, and

$$|\lambda_n - \mu n| \leq K_n \sqrt{n} \text{ with } K_n \xrightarrow{a.s.} 2\sigma \text{ as } n \to \infty. \tag{1.1}$$

(Follow the n *continuously* changing eigenvalues of $tV_n + \mu n e_n e_n^T$ as t moves from 0 to 1. For fixed t they all must lie in the union of the two *disjoint* discs in the complex plane centered at the origin and μn, both having radius $t\|V_n\|$. Necessarily, for all $t \in [0, 1]$, one and only one eigenvalue of $tV_n + \mu n e_n e_n^T$ can lie in the disc centered at μn with radius $\|V_n\|$, and it must remain real, positive, and larger than all the other eigenvalues in absolute value.)

We see now where 50 comes into play, since μ for our U is simply $1/2$. But λ seems to be much closer to 50 than what is quaranteed by (1.1). Indeed, a simulation of 1000 generations of independent U's resulted in Perron eigenvalues ranging between 49.06 and 51.13. (1.1) would place λ merely between $50 - \frac{10}{\sqrt{3}}$ and $50 + \frac{10}{\sqrt{3}}$ (using the fact that the variance of a uniformly distributed r.v. on $(0, 1)$ is $1/12$).

Compute the spectral norm of U. It cannot be smaller than λ. A relation corresponding to (1.1) can be derived using similar continuity

arguments from a perturbation theorem on the singular values of rectangular matrices, such as Corollary 7.3.8 in [3]:
Let A and B be $n \times s$ rectangular matrices, with respective singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_q$, and $\tau_1 \geq \tau_2 \geq \cdots \geq \tau_q$, where $q = \min(n, s)$. Then for all $i = 1, 2, \ldots, q$,

$$|\sigma_i - \tau_i| \leq \|B - A\|.$$

It follows that

$$|\lambda_n - \mu\sqrt{ns}| \leq K_n\sqrt{n} \text{ with } K_n \xrightarrow{a.s.} (1 + 1/\sqrt{y})\sigma \text{ as } n \to \infty, \tag{1.2}$$

where $\lambda_n = \|V_n + \mu\sqrt{ns}e_n e_s^T\|$, and V_n, $n \times s$, y are defined as in Theorem 1.1. But, again, simulations show $\|U\|$ to be much closer than $10/\sqrt{3}$ away from 50.

The purpose of this paper is to provide more detailed information on the limiting behavior of the spectral radii and spectral norms of random matrices as the dimension increases, with entries having positive means. The following theorems will be proved.

Theorem 1.2. Let λ_n be the largest (in absolute value) eigenvalue of $V_n + \mu n e_n e_n^T$, where $\mu > 0$, and V_n is defined in Theorem 1.1 with $s = n$ and $\mathsf{E}(v_{1\,1}^4) < \infty$. Then, with probability one, λ_n is real and positive (that is, it is the spectral radius of $V_n + \mu n e_n e_n^T$) for all n sufficiently large. Moreover,

$$\lambda_n = \mu n + X_n + \frac{1}{\sqrt{n}} Z_n$$

where $\{Z_n\}$ is a tight sequence, and $X_n = \dfrac{1}{n} \displaystyle\sum_{1 \leq i,j \leq n} v_{ij}$. Thus, by the central limit theorem, $\lambda_n - \mu n \xrightarrow{D} N(0, \sigma^2)$.

Theorem 1.3. With V_n $n \times s$, y defined as in Theorem 1.1, $\mathsf{E}(v_{1\,1}^4) < \infty$ and $\mu > 0$

$$\|V_n + \mu\sqrt{ns}e_n e_s^T\| = \mu\sqrt{ns} + \frac{1}{2}\frac{\sigma^2}{\mu}\left(\sqrt{\frac{n}{s}} + \sqrt{\frac{s}{n}}\right) + X_n + \frac{1}{\sqrt{n}} Z_n$$

where $\{Z_n\}$ is tight, and

$$X_n = \frac{1}{\sqrt{ns}} \sum_{\substack{1 \le i \le n \\ 1 \le j \le s}} v_{ij} \xrightarrow{D} N(0, \sigma^2).$$

These theorems obviously account for the small variation of the Perron eigenvalue and spectral norm of U about 50, the latter being approximately $\frac{\sigma^2}{\mu} = 1/6$ larger than the former.

Notice results for negative μ can be trivially derived from these theorems.

The proofs, given in the next section, rely mainly on Theorem 1.1, (1.1), and (1.2), and require little additional probabilistic arguments. They can easily be extended to allow for complex entries in V_n (the proof of Theorem 1.1 can be modified to the complex case).

As will be seen, Z_n in either case can be expressed in a form for which further analysis is possible. A more detailed study of Z_n will undoubtably yield it to be asymptotically normal.

It is remarked here that a similar result is obtained for the largest eigenvalue of non-central random matrices of Wigner type, that is, symmetric matrices with independent entries on and above the diagonal ([2]), although with a proof more probabilistic in nature. The techniques used in the next section can easily be applied to the Wigner case.

2. PROOFS OF THEOREMS 1.2,1.3

For the following, Z_n will denote a generic random variable, not necessarily the same quantity from one appearance to the next, for which $\{Z_n\}$ is tight.

We start with V_n $n \times n$, satisfying the conditions of Theorem 1.2. Concentrating for the moment on realizations for which $\|V_n\| \sim 2\sigma\sqrt{n}$, for any fixed realization we assume n is large enough so that λ_n, defined in Theorem 1.2, satisfies $\|\frac{1}{\lambda_n}V_n\| \le 1/2$. Notice, then, $\lambda_n I - V_n$ is invertible, and $\|(I - \frac{1}{\lambda_n}V_n)^{-1}\| \le 2$ (use the fact that, for square A, $(I-A)^{-1} = \sum_{j=0}^{\infty} A^j$ whenever $\|A\| < 1$). Let f be an eigenvector of $V_n + \mu n e_n e_n^T$ corresponding to λ_n. Then $V_n f + e_n^T f \mu n e_n = \lambda_n f$, which implies $\mu n e_n^T f(\lambda_n I - V_n)^{-1} e_n = f$. Multiplying on both sides by e_n^T, we find (noting that $e_n^T f$ cannot be 0) $\lambda_n = \mu n e_n^T (I - \frac{1}{\lambda_n}V_n)^{-1} e_n$.

Write

$$\lambda_n = \mu n(1 + \tfrac{1}{\lambda_n} e_n^T V_n e_n + \tfrac{1}{\lambda_n^2} e_n^T V_n^2 e_n + e_n^T (\tfrac{1}{\lambda_n} V_n)^3 (I - \tfrac{1}{\lambda_n} V_n)^{-1} e_n).$$

We have $|\mu n e_n^T (\frac{1}{\lambda_n} V_n)^3 (I - \frac{1}{\lambda_n} V_n)^{-1} e_n| \le 2\mu n \|(\frac{1}{\lambda_n} V_n)^3\| = K_n \cdot \frac{1}{\sqrt{n}}$, where $K_n \to 16(\frac{\sigma}{\mu})^3$.

For all $n \ge 1$ and all realizations, let

$$Y_n = \begin{cases} \mu n^{3/2} e_n^T (\frac{1}{\lambda_n} V_n)^3 (I - \frac{1}{\lambda_n} V_n)^{-1} e_n & \text{if } \|\frac{1}{\lambda_n} V_n\| \le 1/2 \\ \sqrt{n}(\lambda_n - \mu n(1 + \frac{1}{\lambda_n} e_n^T V_n e_n + \frac{1}{\lambda_n^2} e_n^T V_n^2 e_n)) & \text{o.w.} \end{cases}$$

Then, since $|Y_n| \le \sup_{k \ge n} |Y_k| \to \limsup_n |Y_n| \le 16(\dfrac{\sigma}{\mu})^3$ a.s., we have the tightness of $\{Y_n\}$.

We have

$$\mathsf{E}((\frac{1}{\sqrt{n}} e_n^T V_n^2 e_n)^2) = \frac{1}{n^3} \mathsf{E}((\sum_{i\,j\,k} v_{ij} v_{jk})^2) =$$

$$\frac{1}{n^3}(n\mathsf{E}(v_{1\,1}^4) + (3n(n-1) + 2n^2(n-1))\sigma^4).$$

Therefore, $\{\frac{1}{\sqrt{n}} e_n^T V_n^2 e_n\}$ is tight, which implies $\{\frac{\mu n^{3/2}}{\lambda_n^2} e_n^T V_n^2 e_n + Y_n\}$ is tight.

At this stage we have $\lambda_n = \mu n + \frac{\mu n}{\lambda_n} X_n + \frac{1}{\sqrt{n}} Z_n$ (X_n defined in Theorem 1.2). Therefore, $\frac{\mu n}{\lambda_n} - 1 = \frac{1}{n} Z_n$, and the proof of Theorem 1.2 is now complete.

We proceed to the proof of Theorem 1.3, where V_n is $n \times s$. To facilitate the exposition, we will write $V = V_n$. Let $\lambda = \|V + \mu\sqrt{ns} e_n e_s^T\|$. Write $V + \mu\sqrt{ns} e_n e_s^T$ in its singular value decomposition $\underline{U}\Lambda\underline{V}$ ([3], p. 415), where $\underline{U}$ is $n \times n$, $\underline{V}$ is $s \times s$, both orthogonal, and Λ is non-negative diagonal, its diagonal elements arranged in non-increasing order. Then $\lambda = \Lambda_{1\,1}$. Let $\underline{u}$, $\underline{v}$ be the first columns of $\underline{U}$,$\underline{V}^T$, respectively. Then

$$V\underline{v} + e_s^T \underline{v} \mu \sqrt{ns} e_n = \lambda \underline{u} \text{ and } V^T \underline{u} + e_n^T \underline{u} \mu \sqrt{ns} e_s = \lambda \underline{v}.$$

For the moment we concentrate on a realization for which $\|V\| \sim (1 + \sqrt{y})\sigma\sqrt{s}$, and n sufficiently large so that $\|\frac{1}{\lambda} V\| \le \frac{1}{\sqrt{2}}$. Then neither $e_n^T \underline{u}$ nor $e_s^T \underline{v}$ will be 0, $\lambda^2 I - VV^T$, $\lambda^2 I - V^T V$ are invertible (I denoting the generic square identity matrix), and $\max(\|(I - \frac{1}{\lambda^2}VV^T)^{-1}\|, \|(I - \frac{1}{\lambda^2}V^TV)^{-1}\|) \le 2$.

We have $V^TV\underline{v} + e_s^T\underline{v}\mu\sqrt{ns}V^Te_n = \lambda\underline{v} - \lambda e_n^T\underline{u}\mu\sqrt{ns}e_s$. Therefore

$$\mu\sqrt{ns}(e_s^T\underline{v}V^Te_n + \lambda e_n^T\underline{u}e_s) = (\lambda^2 I - V^TV)\underline{v},$$

which implies

$$\mu\sqrt{ns}(e_s^T\underline{v}e_s^T(\lambda^2 I - V^TV)^{-1}V^Te_n + \lambda e_n^T\underline{u}e_s^T(\lambda^2 I - V^TV)^{-1}e_s) = e_s^T\underline{v}.$$

Similarly, we find

$$\mu\sqrt{ns}(e_n^T\underline{u}e_n^T(\lambda^2 I - VV^T)^{-1}Ve_s + \lambda e_s^T\underline{v}e_n^T(\lambda^2 I - VV^T)^{-1}e_n) = e_n^T\underline{u}.$$

Noting that $e_s^T(\lambda^2 I - V^TV)^{-1}V^Te_n = e_n^T(\lambda^2 I - VV^T)^{-1}Ve_s$, we arrive at the 2×2 system

$$\mu\sqrt{ns}\begin{pmatrix} e_n^T(\lambda^2 I - VV^T)^{-1}Ve_s & \lambda e_n^T(\lambda^2 - VV^T)^{-1}e_n \\ \lambda e_s^T(\lambda^2 - V^TV)^{-1}e_s & e_n^T(\lambda^2 I - VV^T)^{-1}Ve_s \end{pmatrix}\begin{pmatrix} e_n^T\underline{u} \\ e_s^T\underline{v} \end{pmatrix} = \begin{pmatrix} e_n^T\underline{u} \\ e_s^T\underline{v} \end{pmatrix}.$$

Since $\begin{pmatrix} e_n^T\underline{u} \\ e_s^T\underline{v} \end{pmatrix} \ne 0$, it is an eigenvector of the above matrix. Thus

$$(\mu\sqrt{ns}e_n^T(\lambda^2 I - VV^T)^{-1}Ve_s - 1)^2$$
$$= \lambda^2 e_n^T(\lambda^2 I - VV^T)^{-1}e_n e_s^T(\lambda^2 I - V^TV)^{-1}e_s\mu^2 ns,$$

which implies

$$\lambda^2 = \mu^2 ns\frac{e_n^T(I - \frac{1}{\lambda^2}VV^T)^{-1}e_n e_s^T(I - \frac{1}{\lambda^2}V^TV)^{-1}e_s}{(\frac{\mu\sqrt{ns}}{\lambda^2}e_n^T(I - \frac{1}{\lambda^2}VV^T)^{-1}Ve_s - 1)^2}.$$

For the following K_n will denote a generic positive constant converging to a constant not depending on the realization.

Notice $|\frac{\mu\sqrt{ns}}{\lambda^2}e_n^T(I - \frac{1}{\lambda^2}VV^T)^{-1}Ve_s| \le K_n\frac{1}{\sqrt{n}}$. Write

$$\frac{\mu\sqrt{ns}}{\lambda^2}e_n^T(I - \frac{1}{\lambda^2}VV^T)^{-1}Ve_s = \frac{\mu\sqrt{ns}}{\lambda^2}X_n + Y_n,$$

where X_n is defined in Theorem 1.3, and

$$Y_n = \frac{\mu\sqrt{ns}}{\lambda^2} e_n^T VV^T (I - \frac{1}{\lambda^2} VV^T)^{-1} V e_s.$$

Notice $|Y_n| \leq K_n \frac{1}{n^{3/2}}$.

Write $e_n^T(I - \frac{1}{\lambda^2}VV^T)^{-1}e_n = 1 + \frac{1}{\lambda}X_n^{(1)}$, where $X_n^{(1)} = \frac{1}{\lambda}e_n^T VV^T e_n + Y_n^{(1)}$, $Y_n^{(1)} = \frac{1}{\lambda^3}e_n^T(VV^T)^2(I - \frac{1}{\lambda^2}VV^T)^{-1}e_n$, and $e_s^T(I - \frac{1}{\lambda^2}V^TV)^{-1}e_s = 1 + \frac{1}{\lambda}X_n^{(2)}$, where $X_n^{(2)} = \frac{1}{\lambda}e_s^T V^T V e_s + Y_n^{(2)}$, $Y_n^{(2)} = \frac{1}{\lambda^3}e_s^T(V^TV)^2(I - \frac{1}{\lambda^2}V^TV)^{-1}e_s$. Notice $X_n^{(1)}$,$X_n^{(2)}$,$Y_n^{(1)}$,$Y_n^{(2)}$ are non-negative, with $\max(X_n^{(1)}, X_n^{(1)}) \leq K_n$, and $\max(Y_n^{(1)}, Y_n^{(1)}) \leq K_n \frac{1}{n}$.

For all n sufficiently large, we have

$$\lambda = \frac{\mu\sqrt{ns}\sqrt{(1 + \frac{1}{\lambda}X_n^{(1)})(1 + \frac{1}{\lambda}X_n^{(2)})}}{1 - (\frac{\mu\sqrt{ns}}{\lambda^2}X_n + Y_n)}.$$

Using $|\sqrt{1+x} - (1 + \frac{1}{2}x)| \leq \frac{1}{8}x^2$ for $x \geq 0$, $|\frac{1}{1-x} - (1+x)| \leq 2x^2$ for $|x| \leq \frac{1}{2}$, and arguing in the same manner in the proof of Theorem 1.2, we find for all $n \geq 1$ and all realizations

$$\lambda = \mu\sqrt{ns} + \frac{1}{2}\frac{\mu\sqrt{ns}}{\lambda^2}(e_n^T VV^T e_n + e_s V^T V e_s) + \left(\frac{\mu\sqrt{ns}}{\lambda}\right)^2 X_n + \frac{1}{\sqrt{n}} Z_n.$$

Write

$$e_n^T VV^T e_n + e_s V^T V e_s = (\frac{1}{n} + \frac{1}{s})\sum_{ij} v_{ij}^2 + A_n,$$

where

$$A_n = \frac{1}{n}\sum_{\substack{i\neq \underline{i} \\ j}} v_{ij}v_{\underline{i}j} + \frac{1}{s}\sum_{\substack{j\neq \underline{j} \\ i}} v_{ij}v_{i\underline{j}}.$$

From the central limit theorem $\{\frac{1}{ns}\sum_{ij}(v_{ij}^2 - \sigma^2)\}$ is tight, and since $\frac{1}{n}\mathsf{E}(A_n^2) = 2(\frac{(n-1)s}{n^2} + \frac{(s-1)n}{ns})$, we find $\{\frac{1}{\sqrt{n}}A_n\}$ to be tight. Therefore

$$\lambda = \mu\sqrt{ns} + \frac{1}{2}\sigma^2\frac{\mu\sqrt{ns}(n+s)}{\lambda^2} + \left(\frac{\mu\sqrt{ns}}{\lambda}\right)^2 X_n + \frac{1}{\sqrt{n}} Z_n.$$

Since $\lambda = \mu\sqrt{ns} + Z_n$, we find

$$\left(\frac{\mu\sqrt{ns}}{\lambda}\right)^2 - 1 = \frac{1}{n}Z_n \text{ and } \frac{\mu\sqrt{ns}(n+s)}{\lambda^2} - \frac{1}{\mu}\left(\sqrt{\frac{n}{s}} + \sqrt{\frac{s}{n}}\right) = \frac{1}{n}Z_n.$$

This completes the proof of Theorem 1.3.

ACKNOWLEDGMENTS

The author would like to thank the referees for their helpful suggestions. This paper was written while visiting the Center for Stochastic Processes, University of North Carolina at Chapel Hill, and supported in part by NSF and AFOSR Grant F49620-92-J-0154.

REFERENCES

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

[2] Füredi, Z. and Komlós, J. The eigenvalues of random symmetric matrices. *Combinatorica* **3** (1981), pp.233-241.

[3] Horn, R.A. and Johnson, C.R. *Matrix Analysis* Cambridge Univ. Press, 1985.

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