DOI: 10.1214/ECP.v12-1243

Elect. Comm. in Probab. **12** (2007), 21–27

ELECTRONIC COMMUNICATIONS in PROBABILITY

SPECTRAL NORM OF RANDOM LARGE DIMENSIONAL NONCENTRAL TOEPLITZ AND HANKEL MATRICES

ARUP BOSE
Statistics and Mathematics Unit, Indian Statistical Institute, 203 B.T. Road, Kolkata 700108, India
email: abose@isical.ac.in

ARNAB SEN
Department of Statistics, University of California, Berkeley, CA 94720, USA
email: arnab@stat.berkeley.edu

Submitted October 11, 2006, *accepted in final form* January 16, 2007

AMS 2000 Subject classification: Primary 60F99, Secondary 60F05, 60F15
Keywords: Large dimensional random matrix, eigenvalues, Wigner matrix, Toeplitz matrix, Hankel matrix, spectral norm.

Abstract

Suppose s_n is the spectral norm of either the Toeplitz or the Hankel matrix whose entries come from an i.i.d. sequence of random variables with positive mean μ and finite fourth moment. We show that $n^{-1/2}(s_n - n\mu)$ converges to the normal distribution in either case. This behaviour is in contrast to the known result for the Wigner matrices where $s_n - n\mu$ is itself asymptotically normal.

1 Introduction

For an $n \times n$ real symmetric matrix A_n, let $\lambda_1(A_n) \leq \lambda_2(A_n) \leq \cdots \leq \lambda_n(A_n)$ be the eigenvalues of A_n. Let $\|A_n\|$ denote the spectral norm of A_n, i.e. the maximum of the eigenvalues in their modulus. In other words,

$$\|A_n\| = \max(-\lambda_1(A_n), \lambda_n(A_n)).$$

One of the most frequently studied large dimensional random matrix is the Wigner matrix. A (real) Wigner matrix (Wigner (1955, 1958)) of order n is a matrix whose entries above the diagonal are i.i.d. real random variables and whose diagonal elements are also i.i.d. real random variables, independent of the other elements. So this matrix is given by

$$W_n = \begin{bmatrix} w_{11} & w_{12} & w_{13} & \dots & w_{1(n-1)} & w_{1n} \\ w_{21} & w_{22} & w_{23} & \dots & w_{2(n-1)} & w_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ w_{n1} & w_{n2} & w_{n3} & \dots & w_{n(n-1)} & w_{nn} \end{bmatrix} \tag{1}$$

where $w_{kj} = w_{jk}\ j < k$, are i.i.d. (real) random variables and the diagonal elements w_{ii} are i.i.d. real random variables and are independent of the off diagonal variables.

There are a host of results known for the Wigner matrix and its variants. We quote below the results relevant to us on their spectral norm and extreme eigenvalues.

Theorem 1. *Suppose $\{W_n\}$ is a sequence of Wigner matrices of order n such that $\mathrm{E}(w_{11}^2) = 1$ and $E(w_{11}^4) < \infty$.*

(A) If $\mathrm{E}(w_{11}) = 0$, then the maximum eigenvalue of $n^{-\frac{1}{2}} W_n$ converges to 2 almost surely.

(B) Assume that the mean μ of the entries is positive. Let φ_n be the spectral norm of W_n. Then $\varphi_n - \mu n \xrightarrow{d} N(0, 1)$.

Part A is proved in Bai and Yin (1988). Part B is due to Silverstein (1994).

Observe that in Part B, the mean of the entries is assumed to be positive. We call this the noncentral case. It is interesting to note that for the distributional convergence, only centering suffices and no scaling is required.

Nonrandom Toeplitz and Hankel matrices are extremely well studied in mathematics, specially in operator theory. Let $\{x_0, x_1, \dots\}$ be a sequence of real numbers.

Then the $n \times n$ Toeplitz Matrix is the matrix whose (i, j)-th entry is $x_{|i-j|}$. So it is given by

$$T_n = \begin{bmatrix} x_0 & x_1 & x_2 & \dots & x_{n-2} & x_{n-1} \\ x_1 & x_0 & x_1 & \dots & x_{n-3} & x_{n-2} \\ x_2 & x_1 & x_0 & \dots & x_{n-4} & x_{n-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{n-1} & x_{n-2} & x_{n-3} & \dots & x_1 & x_0 \end{bmatrix}.$$

Hankel matrices have very close connections with the Toeplitz matrices. The $n \times n$ Hankel Matrix is the matrix whose (i, j)-th entry is x_{i+j-2}. So it is given by

$$H_n = \begin{bmatrix} x_0 & x_1 & x_2 & \dots & x_{n-2} & x_{n-1} \\ x_1 & x_2 & x_3 & \dots & x_{n-1} & x_n \\ x_2 & x_3 & x_4 & \dots & x_n & x_{n+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{n-1} & x_n & x_{n+1} & \dots & x_{2n-3} & x_{2n-2} \end{bmatrix}.$$

The question of existence of limiting spectral distribution for the eigenvalues of random Toeplitz and Hankel matrices has been settled recently. See Bryc et. al (2006).

Theorem 2 (Bryc, Dembo and Jiang (2006)). *Let the $\{x_i\}$ in the Toeplitz (Hankel) matrix T_n (H_n) be i.i.d. with mean zero and variance one. Then with probability one, the empirical spectral distribution of $\frac{1}{\sqrt{n}}T_n$ $(\frac{1}{\sqrt{n}}H_n)$ converges weakly as $n \to \infty$ to a nonrandom symmetric probability measure which does not depend on the distribution of the entries $\{x_i\}$ and has unbounded support.*

Also see Hammond and Miller (2005) for some detailed information on the behavior of empirical spectral moments of random Toeplitz matrices. Unlike the Wigner case, apparently, there are no results known for the behavior at the edge of the spectrum of the random Toeplitz and the Hankel matrices. In the next section we consider Toeplitz and Hankel matrices where $\{x_i\}$ are i.i.d. with mean $\mu > 0$. We show that the spectral norm of both the Toeplitz and the Hankel matrices obey a strong law and also converges to a normal distribution under appropriate centering and scaling.

2 Main results and proofs

Suppose $\{x_i\}$ are i.i.d. and have mean μ. Let T_n be the Toeplitz matrix formed by these $\{x_i\}$. Let $u_n = n^{-1/2}(\underbrace{1, 1, \cdots, 1}_{n\text{times}})^T$. Then $T_n^\circ = T_n - \mu n u_n u_n^T$ is the corresponding centered Toeplitz matrix whose entries have mean zero. We now state our main theorem.

Theorem 3. *Suppose T_n is a Toeplitz matrix where $\mathrm{E}(x_0) = \mu > 0$ and $\mathrm{Var}(x_0) = 1$. Let $T_n^\circ = T_n - \mu n u_n u_n^T$. Then,*

(A)

$$\frac{\|T_n\|}{n} \to \mu \quad \textit{almost surely and} \quad \|\frac{T_n^\circ}{\|T_n\|}\| \to 0 \quad \textit{almost surely.}$$

(B) Further assume $\mathrm{E}(x_0^4) < \infty$. Then for $M_n = \|T_n\|$ or $M_n = \lambda_n(T_n)$,

$$\frac{M_n - \mu n}{\sqrt{n}} \to N(0, 4/3) \quad \textit{in distribution.}$$

(C) If T_n and T_n° are replaced by the corresponding Hankel matrices, then (A) holds. Further, (B) holds with the limiting variance being changed from $4/3$ to $2/3$.

Before going to the proof of the theorem let us state the following Lemma.

Lemma 1. Let T_n and T_n° be as above. Then

(i) If $\mathrm{E}(x_0^2) < \infty$, then $n^{-1}\|T_n^\circ\| \to 0$ a.s.

(ii) If $\mathrm{E}(x_0^4) < \infty$, then $n^{-3/4}\|T_n^\circ\|$ is tight.

Proof Let $y_i = x_i - \mu$. Define $y_i^c = y_i \mathrm{I}(|y_i| \le c) - \mathrm{E}\big[y_i \mathrm{I}(|y_i| \le c)\big]$, the truncated and centered version of y_i. Let $T_n^{(c)}$ be the Toeplitz matrix formed by the sequence $\{y_i^c\}$.

Then $n^{-1}\big|\|T_n^\circ\| - \|T_n^{(c)}\|\big| \le n^{-1}\|T_n^\circ - T_n^{(c)}\| \le \big[n^{-2}\,\mathrm{Tr}[(T_n^\circ - T_n^{(c)})^2]\big]^{1/2}$.

Define $r_i^c = y_i - y_i^c, \quad i = 0, 1, 2, \cdots$. Clearly, $\{r_i^c\}$'s are i.i.d. and $\mathrm{E}\,(r_0^c)^2 \to 0$ as $c \to \infty$.

Now $n^{-2}\,\mathrm{Tr}[(T_n^\circ - T_n^{(c)})^2] = n^{-2}[2\sum_{i=1}^{n-1}(n-i)(r_i^c)^2 + n(r_0^c)^2] \leq 2n^{-1}\sum_{i=0}^{n-1}(r_i^c)^2 \to \mathrm{E}(r_0^c)^2$ as $n \to \infty$ by SLLN. So, we just need to show that $n^{-1}\|T_n^{(c)}\| \to 0$.

To complete the proof of part (i), we extract a crucial fact from the proof of Theorem **??** given in Bryc, Dembo and Jiang (2006). If we carefully follow their argument, it is clear that for bounded mean zero random variables, $n^{-3}\,\mathrm{Tr}[(T_n^{(c)})^4]$ converges to some positive constant almost surely. Hence,

$$(n^{-1}\|T_n^{(c)}\|)^4 \leq n^{-4}\,\mathrm{Tr}[(T_n^{(c)})^4] \to 0 \quad \text{almost surely.}$$

To prove part (ii), it is enough to show the fourth moment is uniformly bounded. But this is true since,

$\mathrm{E}\big[(n^{-3/4}\|T_n^\circ\|)^4\big] \leq n^{-3}\,\mathrm{E}[\mathrm{Tr}\big((T_n^\circ)^4\big)] = n^{-3}\,\mathrm{E}(\sum_{i_1,i_2,i_3,i_4} y_{|i_1-i_2|}y_{|i_2-i_3|}y_{|i_3-i_4|}y_{|i_4-i_1|}) = O(1)$ from Bryc, Dembo and Jiang (2006).

Thus the proof of the Lemma is complete.

Proof of Theorem ?? We will prove only Parts A and B. The proof of Part C is similar and will be omitted.

Using the triangle inequality for norms,

$$\|\mu n u_n u_n^T\| - \|T_n^\circ\| \leq \|T_n\| \leq \|\mu n u_n u_n^T\| + \|T_n^\circ\|$$

$$\text{or,}\ \ \mu - \|\frac{T_n^\circ}{n}\| \leq \frac{\|T_n\|}{n} \leq \mu + \|\frac{T_n^\circ}{n}\|.$$

Thus using Lemma **??**(i) , we easily conclude $\frac{\|T_n\|}{n} \to \mu$ almost surely. This proves the first part of (A). The second part now follows again from Lemma **??**(i)

We now prove part (B). Define the three sets,

$$\Omega_1^{(n)} = \{\|\frac{T_n^\circ}{\|T_n\|}\| \leq 1/2\}, \quad \Omega_2^{(n)} = \{n^{-1}\|T_n^\circ\| < \mu/2\}, \quad \Omega^{(n)} = \Omega_1^{(n)} \cap \Omega_2^{(n)}.$$

For simplicity we will drop the superscript and write Ω_1, Ω_2 and Ω for the above three sets respectively. Note that from Lemma **??**(i) and first part of the Theorem, given $\epsilon > 0$, for all large n,

$$P(\Omega) > 1 - \epsilon.$$

Suppose A is any matrix such that $\|A\| < \alpha < 1$. Then $(I-A)^{-1} = \sum_{j=0}^\infty A^j$ and $\|(I-A)^{-1}\| < (1-\alpha)^{-1}$.

Hence, on the set Ω_1, $(I - \frac{1}{\|T_n\|}T_n^\circ)^{-1}$ exists and

$$||(I - \frac{1}{\|T_n\|}T_n^\circ)|| \leq 2. \tag{2}$$

The following fact is well known in the theory of matrices (See Horn and Johnson (1985) Corollary 6.3.4).

Fact. Suppose $\hat{\lambda}$ is an eigenvalue of $A+P$, where A is normal (that is $AA^T = A^TA$). Then there exists an eigenvalue λ of A such that $|\hat{\lambda} - \lambda| \leq \|P\|$.

Using this fact and noting that the distinct eigenvalues of $\mu n u_n u_n^T$ are 0 and μn, we get

$$\lambda_1(T_n) \geq -\|T_n^\circ\|.$$

and

$$\lambda_n(T_n) \geq \mu n - \|T_n^\circ\|.$$

Hence on Ω_2 (i.e. if $n^{-1}\|T_n^\circ\| < \mu/2$), we have

$$\lambda_n(T_n) > -\lambda_1(T_n).$$

So, on Ω_2

$$\|T_n\| = \lambda_n(T_n). \tag{3}$$

Then there exists an eigenvector $g \neq 0$ such that $T_n g = \lambda_n(T_n)g$ or equivalently, $(T_n^\circ + \mu n u_n u_n^T)g = \lambda_n(T_n)g$. Noting that $u_n^T g$ is just a real number, this implies that $\mu n u_n^T g u_n = \big(\lambda_n(T_n)I - T_n^\circ\big)g$.

Now we will work on the set Ω, on which

$$g = \mu n u_n^T g(\lambda_n(T_n)I - T_n^\circ)^{-1}u_n.$$

Note that by the last relation, $u_n^T g \neq 0$ and hence premultiplying both sides of this relation by u_n^T,

$$\lambda_n(T_n) = \mu n u_n^T (I - \frac{1}{\lambda_n(T_n)}T_n^\circ)^{-1}u_n.$$

Motivated by the above relation, define

$$\tilde{\lambda}_n = \begin{cases} \mu n\Big(1 + \frac{1}{\lambda_n(T_n)}u_n^T T_n^\circ u_n + (\frac{1}{\lambda_n(T_n)})^2 u_n^T (T_n^\circ)^2 u_n + u_n^T(\frac{T_n^\circ}{\lambda_n(T_n)})^3(I - \frac{1}{\lambda_n(T_n)}T_n^\circ)^{-1}u_n\Big) & \text{on the set } \Omega \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

We have $\tilde{\lambda}_n = \lambda_n(T_n)$ on Ω_1. Also recall that on the set Ω_2, $\lambda_n(T_n) = \|T_n\|$ and $P(\Omega) \to 1$. So, $\tilde{\lambda}_n - \lambda_n(T_n) = o_p(1)$ and $\tilde{\lambda}_n - \|T_n\| = o_p(1)$. Hence, it is enough to find the limiting distribution of $\tilde{\lambda}_n$.

Consider the last three terms of $\tilde{\lambda}_n$ on Ω. Call them B_2, B_3 and B_4.

$\mathbf{B_4}$. Using inequality (**??**),

$$\begin{aligned} \frac{|B_4|}{\sqrt{n}} &= \Big|\mu n^{1/2} u_n^T (\frac{T_n^\circ}{\lambda_n(T_n)})^3 (I - \frac{1}{\lambda_n(T_n)}T_n^\circ)^{-1}u_n\Big| \\ &\leq 2\mu n^{1/2}\Big(\frac{\|T_n^\circ\|}{\lambda_n(T_n)}\Big)^3 \\ &= \frac{2\mu}{n^{1/4}}\Big(\frac{n}{\lambda_n(T_n)}\Big)^3\Big(\frac{\|T_n^\circ\|}{n^{3/4}}\Big)^3 \to 0 \quad \text{in probability by Lemma } \textbf{??}(ii) \end{aligned}$$

Hence $\frac{B_4}{\sqrt{n}} = o_p(1)$.

$\mathbf{B_3}$.

$$\frac{\mu n}{(\lambda_n(T_n))^2} u_n^T (T_n^\circ)^2 u_n = \mu \Big(\frac{n}{\lambda_n(T_n)} \Big)^2 \underbrace{n^{-1} u_n^T (T_n^\circ)^2 u_n}_{A_n \text{ say}}.$$

We have

$$\begin{aligned} \mathrm{E}(A_n^2) &= n^{-4}\,\mathrm{E}\Big[\big(\sum_{i,j,k} y_{|i-j|} y_{|j-k|} \big)^2 \Big] \\ &= n^{-4}\,\mathrm{E}\big(\sum y_{|i1-j1|} y_{|j1-k1|} y_{|i2-j2|} y_{|j2-k2|} \big) = O(1). \end{aligned}$$

where $y_i = x_i - \mu$ is the centered version of $\{x_i\}$. The last step above follows from argument similar to those given in Bryc, Dembo and Jiang (2006). So, $\{A_n\}$ is tight. Hence $\frac{B_3}{\sqrt{n}} \to 0$.

$\mathbf{B_2}$.

$\frac{n\mu}{\lambda_n(T_n)} \to 1$ and $u_n^T T_n^\circ u_n = \frac{1}{n} \sum_{i,j} y_{|i-j|}$

$= \frac{2}{n} \sum_{i=0}^{n-1} (n-i) y_i + \frac{n y_0}{n} = \frac{1}{n} [2 \sum_{i=0}^{n-1} (n-i) y_i - n y_0]$.

By central limit theorem for sums of independent random variables, it easily follows that $\frac{B_2}{\sqrt{n}} \to N(0, 4/3)$ in distribution.

Combining the above steps, we have $\frac{\tilde{\lambda}_n - \mu n}{\sqrt{n}} \xrightarrow{d} N(0, 4/3)$. This completes the proof of part (B) since as we have already observed that on Ω, $\tilde{\lambda}_n = \lambda_n(T_n) = \|T_n\|$ and $P(\Omega) \to 1$. $\square$

Acknowledgement The authors are grateful to the Referee for his extremely prompt and meticulous report. This manuscript has greatly benefited from his detailed comments.

References

[1] Bai, Zhidong and Yin, J. F. (1988). Necessary and sufficient condition for the almost sure convergence of the largest eigenvalue of Wigner Matrices *Ann. Probab.*, 16, No. 4, 1729–1741. MR0958213

[2] Bryc, Włodzimierz, Dembo, Amir and Jiang, Tiefeng (2006). Spectral measure of large random Hankel, Markov and Toeplitz matrices. *Ann. Probab.*, 34, no. 1, 1–38. Also available at *http://arxiv.org/abs/math.PR/0307330* MR2206341

[3] Hammnond, Christopher and Miller, Steven J. (2005). Distribution of eigenvalues for the ensemble of real symmetric Toeplitz matrices. *J. Theoret. Probab.*, 18, no. 3, 537–566. MR2167641

[4] Horn, R. A. and Johnson C. R.(1985) Matrix Analysis *Cambridge Univ. Press.* MR0832183

[5] Silverstein, J. W. (1994). The Spectral Radii and norms of large dimensional non-central random matrices *Commun. Statist.-Stochastic Models*, 10(3), 525–532. MR1284550

[6] Wigner, E. P. (1955). Characteristic vectors of bordered matrices with infinite dimensions. *Ann. of Math.*, (2), 62, 548–564 MR0077805

[7] Wigner, E. P. (1958). On the distribution of the roots of certain symmetric matrices. *Ann. of Math.*, (2), 67, 325–327 MR0095527