Smallest singular value and limit eigenvalue distribution of a class of non-Hermitian random matrices with statistical application

Arup Bose
Indian Statistical Institute, Kolkata
Walid Hachem
Université Paris-Est Marne-la-Vallée, France

December 12, 2019

Empirical Spectral Distribution

R_N: an $N \times N$ (random) matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_N$.

Note that the eigenvalues can be complex random variables.

Empirical Spectral Distribution (ESD) of R_N is the (random) probability measure

$$\mu_N = \frac{1}{N} \sum_{i=1}^{N} \delta_{\lambda_i}.$$

When all eigenvalues are real, its cumulative form and its moments are respectively

$$ECDF(x) = \frac{1}{n} \#\text{eigenvalues} \leq x \quad \text{and,}$$

$$\beta_h(R_N) = \int x^h d\mu_N = N^{-1} \sum_{i=1}^{N} \lambda_i^h = N^{-1} \operatorname{Tr}(R_N^h).$$

Stieltjes transform

Stieltjes transform of any probablity distribution F on R is

$$m_F(z) = \int \frac{1}{x - z} dF(x), \quad z \in \mathbb{C}^+.$$

It is always defined.

Determines the distribution uniquely, convergence of Stieltjes transform if and only convergence in distribution...

Its moments are defined as

$$\beta_h = \int x^h dF(x).$$

LSD are often expressed through their Stietljes transform.

Two notions of convergence

Limiting spectral distribution (LSD): If this ESD converges weakly (for our purposes, almost surely or in probability) to a probability distribution, then the limit is called the LSD.

In this talk, all limit measures are non-random.

Tracial/algebraic/non-commutative convergence: For every polynomial π,

$$\lim N^{-1}\,\mathrm{Tr}(\pi(R_N, R_N^*)) \ \text{ exists } \ = \phi(\pi(r, r*)) \ \ (\mathit{say}).$$

(i) If R_N is symmetric, tracial convergence implies LSD provided the limit traces $\{\lim N^{-1}\,\mathrm{Tr}(R_N^h))\}$ identifies a unique probability distribution with these as the moments. This is indeed the moment method.

(ii) Tracial convergence notion can be extended to *joint tracial convergence* for multiple sequences. This would then define a *non-commutative algebra* $\mathcal{A}$ (say) (defined via dummy variables $r, r*$ etc..) along with a *linear functional* ϕ (say) (defined via the limit values as above). Such a pair $(\mathcal{A}, \phi)$ is an example of a *non-commutative* $*$*-probability space*.

Two methods to prove convergence

When the eigenvalues are all real (the matrix is real symmetric), there are two common methods to establish LSD:

Moment method.

Stieltjes transform method.

Moment method for real symmetric R_N

Recalll that the h-th order moment of the ESD of R_N equals

$$\beta_h(R_N) := \frac{1}{N}\operatorname{Tr}(R_N^h).$$

(M1) (Moment convergence) For every $h \geq 1, E(\beta_h(R_N)) \to \beta_h$,

(M2) $E(\beta_h(R_N) - E(\beta_h(R_N)))^2 \to 0, \quad \forall h$

(M4) (Borel-Cantelli) $\sum_{N=1}^{\infty} E(\beta_h(R_N) - E(\beta_h(R_N)))^4 < \infty, \quad \forall h \geq 1$, and

(C) (Unique limit) $\displaystyle\sum_{h=1}^{\infty} \beta_{2h}^{-\frac{1}{2h}} = \infty$ (Carleman's condition).

If (M1), (M2) and (C) hold, then ESD of R_N converges *in probability* to the distribution F which is determined uniquely by the moments $\{\beta_h\}$. The convergence is *almost sure* if $(M4)$ holds.

Usually (M1) is the hardest to establish. See Bose (book, 2018) for examples.

Stieltjes transform method

The Stieltjes transform of the ESD of R_N equals

$$m_N(z) = \frac{1}{N} \sum \frac{1}{\lambda_i - z} = \frac{1}{N} \operatorname{Tr}(R_N - zI)^{-1}).$$

Express $m_{N+1}(z)$ in terms of $m_N(z)$ and use (martingale or any other) techniques to push the relation to a limiting functional equation.

The solution, (must show is unique) is the Stieltjes transform of the LSD.

Three basic matrices

Sample size: n.

Dimension: $N = N(n)$. Sometimes we write p instead of N.

Both $N, n \to \infty$,

$$N/n \to \gamma \in [0, \ \infty).$$

(A) The IID matrix: Suppose Z is the $N \times n$ matrix with iid random random variables (with mean 0, variance 1, plus usually finiteness of (some) moments..).

For the next few slides, assume finite fourth moment.

(B) The sample covariance matrix: $S_N = n^{-1}ZZ^*$.

(C) The Wigner matrix: $N^{-1/2}W_N$ where W_N is real symmetric whose elements are IID with mean zero, variance 1.

All limits in this talk are *universal*. That is, they do not depend on the underlying distribution of the random variables except through their second moments.

Basic LSD results for S_N and W_N

(1) The LSD of $N^{-1/2}W_N$ is the *semi-circle law*: Wigner (1955/1956...).

Suppose

$$S_N = n^{-1}ZZ^*.$$

(2) When $\gamma \neq 0$, LSD of S_N exists and is called the *Marchenko-Pastur Law* (1967).

(3) When $\gamma = 0$, LSD of S_N is degenerate (at 0). The LSD of

$$\sqrt{\frac{n}{N}}(S_N - I_N)$$

is the also the *semi-circle law*: Bai and Yin (1988).

All three results can be proved by either the moment method (Bose (2018, book)) or the Stieltjes transform method (Bai, book).

LSD of $A^{1/2}ZZ^*A^{1/2}$

Now suppose A is an $N \times N$ non-negative random matrix whose LSD exists.

(4) the LSD of $N^{-1/2}A^{1/2}WA^{1/2}$ exists: Bai and Zhang (2010).

(5) when $\gamma = 0$, the LSD of $\sqrt{\frac{n}{N}}\left(n^{-1}A^{1/2}ZZ^*A^{1/2} - A\right)$ exists: Pan and Gao (2009), Bao (2012) and is same as that in (4).

Now suppose B is an $n \times n$ *symmetric* non-random matrix (with tracial convergence and LSD...).

(6) when $\gamma = 0$, the LSD of $\sqrt{nN^{-1}}(n^{-1}A^{1/2}ZBZ^*A^{1/2} - n^{-1}\operatorname{Tr}(B)A)$ exists: Wang and Paul (2014).

Extensions?

Suppose we have several of the Z (independent), A and B type matrices.

What kind of LSD results should be valid? (different for $\gamma = 0$ and $\gamma \neq 0$).

How to establish them? (Moment method? Stieltjes transform?)

Any use for such results?

A general set up

- $\{Z_u = ((\varepsilon_{u,i,j}))_{N\times n}\},\ 1 \leq u \leq U.$

- $\{\varepsilon_{u,i,j}\}$ are independently distributed with mean 0, variance 1 and all moments uniformly bounded.

- $\{A_i\}$: class of $N \times N$ matrices, which converge **jointly**.

- $\{B_i\}$: class of $n \times n$ matrices, **each** of which converges tracially.

Theorem 1, joint convergence

Define

- $\mathbb{P} = \left(\prod_{i=1}^{k_l} \frac{1}{n} A_{t_i} \; Z_{j_i} B_{s_i} Z_{j_i}^* \right) A_{t_{k_l+1}}$,

- $\mathbb{G} = \left(\prod_{i=1}^{k_l} \frac{1}{n} \operatorname{Tr}(B_{s_i}) \right) \prod_{i=1}^{k_l+1} A_{t_i}$

- when $\gamma > 0$, the collection $\{\mathbb{P}\}$, converges jointly (in probability).

- when $\gamma = 0$, the collection $\{\sqrt{nN^{-1}}(\mathbb{P} - \mathbb{G})\}$, converges jointly (in probability).

Limits can be expressed in terms of **free variables**.

Proved by checking (M1) and (M2) conditions.

By-product:

- All traces are asymptotically normal.

Theorem 2, LSD for any symmetric polynomial

Then using the moment method, LSD exists for:

- any symmetric polynomial in $\{\mathbb{P}\}$ when $\gamma > 0$
- any symmetric polynomial in $\{\sqrt{nN^{-1}}(\mathbb{P} - \mathbb{G})\}$ when $\gamma = 0$.

Proved by checking condition (C).

For specific polynomials, moment condition can be reduced by truncation arguments.

See Bhattacharjee and Bose (IJPAM 2017/2018, RMTA 2016), Bose and Bhattacharjee (book 2018).

Stieltjes transform formulae (recursive equations) are also derived.

All existing results on covariance-type matrices, including the ones mentioned earlier follow.

Linear time series

General model, MA(q):

$$Y_t = \sum_{j=0}^{q} \psi_j X_{t-j} \;\; t \geq 1. \tag{0.1}$$

X_t's are *i.i.d.* N-dimensional vectors with mean 0 and variance-covariance matrix I_N.

ψ_j are $N \times N$ (non-random) *coefficient matrices*. $\psi_0 = I$. $N = N(n) \to \infty$ such that $\frac{N}{n} \to \gamma \in [0, \infty)$.

q can be infinite but that needs additional restrictions on $\{\psi_i\}$.

MA(0) is the i.i.d. process.

Sometimes we write p for N.

Autocovariance matrix sequence

The sample autocovariance matrix of order i of $\{Y_t\}$ equals

$$\hat{\Gamma}_i := \frac{1}{n} \sum_{t=i+1}^{n} Y_t Y_{t-i}^*.$$

They are all non-symmetric except $\hat{\Gamma}_0$.

What is the behaviour of (polynomials) of these matrices?

Consequences of Theorems 1 and 2

In both $\gamma = 0$ and $\gamma > 0$ cases,

• tracial convergence of polynomials in these matrices follow as consequence of Theorem 1.

• convergence of the ESD of any polynomial of the autocovariances which is a symmetric matrix, follows from Theorem 2 once we assume that the $\{\psi_j\}$ matrices converge jointly.

In particular *all* LSD results in the literature on symmetrized autocovariance matrices (such as $\Gamma_i + \Gamma_i^*$, $\Gamma_i\Gamma_i^*$, $\Gamma_i + \Gamma_i^* + \Gamma_j + \Gamma_j^*$) etc, follow.

One can even combine these matrices across several independent time series (should be useful for two or multi-sample problems..) and the above results are still automatically guaranteed.

We have only very limited success so far on the convergence of the spectral distribution in non-symmetric cases.

Some applications

- All traces are asymptotically normal (useful in testing).

- Applications: determination of the order q, testing for white noise.... (see book).

Simulation 1: ECDF of $\hat{\Gamma}_0$ for MA(0)

$$\hat{\Gamma}_0 = \frac{1}{n}\sum_{t=1}^{n} X_t X_t^* .$$

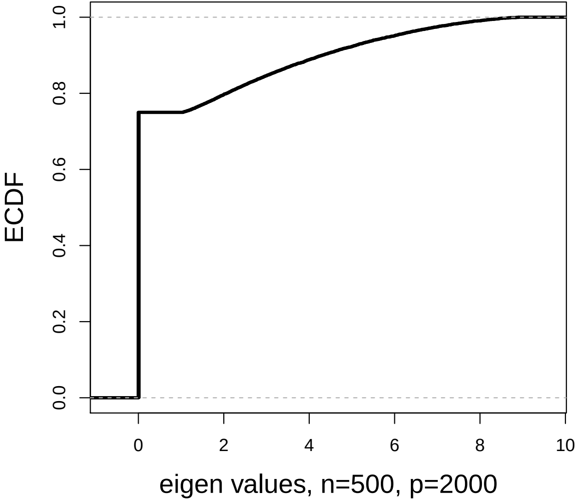


Figure: LSD of $\hat{\Gamma}_0$ is the Marchenko-Pastur law (well known in RMT).

Simulation 2

Model 1 (MA(0)): $Y_t = X_t$.

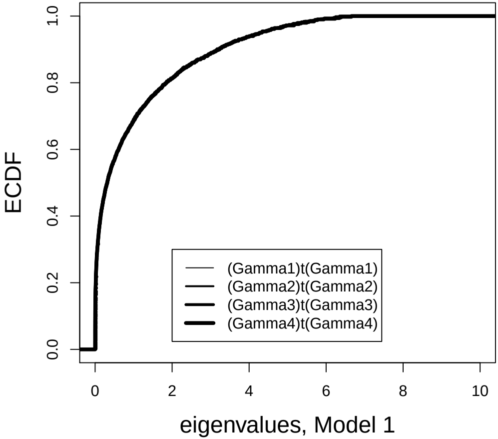


Figure: Identical ECDF of $\hat{\Gamma}_u \hat{\Gamma}_u^*$, $1 \leq u \leq 4$ for $N = n = 300$. LSD known from BB.

Simulation 3

Model 2 $\quad Y_t = X_t + A_N X_{t-1}, \quad A_N = 0.5 I_N.$

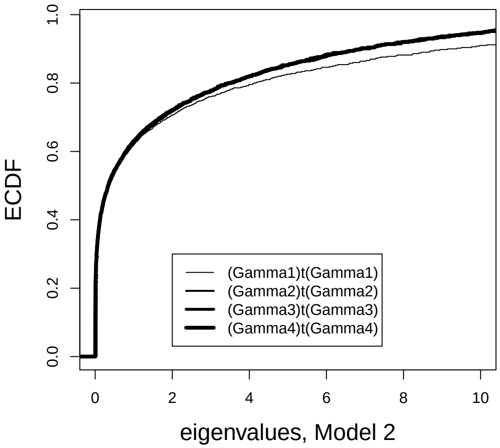


Figure: ECDF of $\hat{\Gamma}_1 \hat{\Gamma}_1^*$ different from ECDF of $\hat{\Gamma}_u \hat{\Gamma}_u^*$, $2, 3, 4$. $N = n = 300$. Spectral distribution of A_N is degenerate at 0.5. LSD known from BB.

ESD of $\hat{\Gamma}_1$ for MA(0), $n = 500$

$$\hat{\Gamma}_1 = \frac{1}{n} \sum_{t=1}^{n} X_t X_{t-1}^*.$$

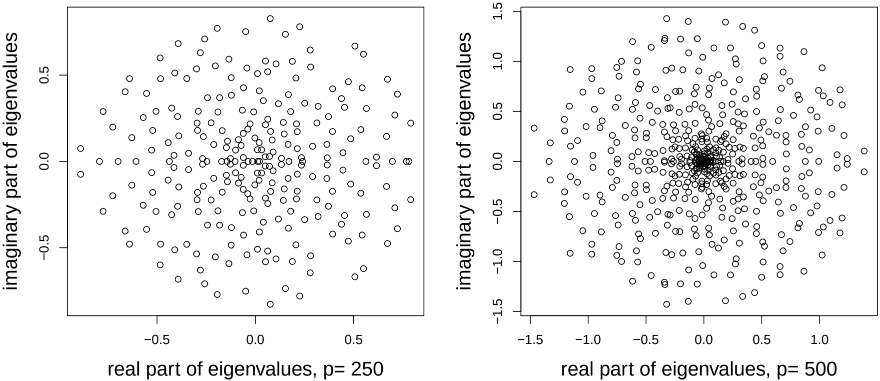


Figure: That this LSD exists shall be argued soon.

Non-symmetric matrices

Tracial moments *cannot* give the existence of the LSD since they do not capture the moments of the ESD which is now on the complex plane.

One of the major steps in the proof for any non-symmetric matrix is to establish suitable bounds for the *smallest singular value* of a related matrix.

We shall clarify at the end how this step helps via *log-potential*.

$(X^{(n)} = [x_{ij}^{(n)}]_{i,j=0}^{N^{(n)}-1,n-1})_{n\geq 1}$ is a sequence of **complex** random matrices such that

Assumption 1. For each $n \geq 1$, $\{x_{ij}^{(n)}\}_{i,j=0}^{N^{(n)}-1,n-1}$ are i.i.d. with

$\mathbb{E}x_{00}^{(n)} = 0$,

$\mathbb{E}|x_{00}^{(n)}|^2 = 1/n$, and

$\sup_n n^2\mathbb{E}|x_{00}^{(n)}|^4 = \boldsymbol{m}_4 < \infty$.

Note the different scaling of the entries.

$(N^{(n)})_{n\geq 1}$, a sequence of positive integers, diverges to ∞ as $n \to \infty$.

$N/n \to \gamma$, $0 < \gamma < \infty$ as $n \to \infty$.

Think of N as the dimension of the vectors and n as the sample size (number of observations).

For any matrix $M \in \mathbb{C}^{n \times n}$, $s_0(M) \geq \cdots \geq s_{n-1}(M)$: the singular values of M.

Assumption 2. $A^{(n)} \in \mathbb{C}^{n \times n}$: sequence of deterministic matrices such that,

$$0 < \inf_n s_{n-1}(A^{(n)}) \leq \sup_n s_0(A^{(n)}) < \infty.$$

- The *Identity matrix* trivially satisfies Assumption 2.
- The *circulant matrix*

$$J^{(n)} = \begin{bmatrix} 0 & & & 1 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix} \in \mathbb{R}^{n \times n} \tag{0.2}$$

also satisfies Assumption 2. Its ESD converges to the uniform distribution on the circle of unit radius.

Assumption 3 The random variables $x_{00}^{(n)}$ satisfy $\sup_n |n\mathbb{E}(x_{00}^{(n)})^2| < 1$.

Assumption 3 essentially says that the x_{ij} are *not* real.

This is fine for applications in wireless communications and signal processing.

Not acceptable for time series applications.

This assumption is due to the application of a Berry-Esseen bound as a tool in exactly one of the steps in the proof. We are trying to remove that.

A general smallest singular value result

Let $\|\cdot\|$ denote the spectral norm of a matrix.

Consider the random matrix $X^{(n)}A^{(n)}X^{(n)^*} - zI_N$, where z is an arbitrary non-zero complex number and I_N is the identity matrix of dimension N.

Theorem 3 Suppose Assumptions 1, 2 and 3 hold. Let C be a positive constant. Then, there exist $\alpha, \beta > 0$ such that for each $z \in \mathbb{C} \setminus \{0\}$,

$$\mathbb{P}\left[s_{N-1}(X^{(n)}A^{(n)}X^{(n)^*} - z) \leq t, \ \|X\| \leq C\right] \leq c\left(n^{\alpha}t^{1/2} + n^{-\beta}\right) + \exp(-c'n),$$

where the constants $c, c' > 0$ depend on C, z, and $\boldsymbol{m}_4$ only. In particular the result is true for XJX^*.

Note that we DO NOT (yet) have the result for matrices with real entries.

We next state our main LSD result. Then we shall explain the connection how Theorem 3 helps in proving Theorem 4.

LSD result for $\hat{\Gamma}_1 = \sum X_t X_{t-1}^*$

To state our result on LSD, let for any $0 < \gamma < \infty$,

$$g(y) = \frac{y}{y+1}(1-\gamma+2y)^2, \ (0 \vee (\gamma-1)) \leq y \leq \gamma. \tag{0.3}$$

Then g^{-1} exists on the interval $[0 \vee ((\gamma-1)^3/\gamma), \gamma(\gamma+1)]$ and maps it to $[0 \vee (\gamma - \gamma^{-1}), \gamma]$. It is an analytic increasing function on the interior of the interval.

Note that g is a cubic polynomial and so a formula can be given for its inverse.

Theorem 4. Suppose Assumptions 1 and 3 hold. Then, the LSD of $\sum X_t X_{t-1}^*$ exists in probability. The limit measure μ is rotationally invariant on $\mathbb{C}$. Let $F(r) = \mu(\{z \in \mathbb{C} : |z| \le r\}), 0 \le r < \infty$ be the distribution function of the radial component.

If $\gamma \le 1$, then

$$F(r) = \begin{cases} \gamma^{-1} g^{-1}(r^2) & \text{if } 0 \le r \le \sqrt{\gamma(\gamma+1)}, \\ 1 & \text{if } r > \sqrt{\gamma(\gamma+1)}. \end{cases}$$

If $\gamma > 1$, then

$$F(r) = \begin{cases} 1 - \gamma^{-1} & \text{if } 0 \le r \le (\gamma-1)^{3/2}/\sqrt{\gamma}, \\ \gamma^{-1} g^{-1}(r^2) & \text{if } (\gamma-1)^{3/2}/\sqrt{\gamma} < r \le \sqrt{\gamma(\gamma+1)}, \\ 1 & \text{if } r > \sqrt{\gamma(\gamma+1)}. \end{cases}$$

The support of $\boldsymbol{\mu}$ is the disc $\{z : |z| \leq \sqrt{\gamma(\gamma+1)}\}$ when $\gamma \leq 1$.

When $\gamma > 1$, the support is the ring $\{z : (\gamma-1)^{3/2}/\sqrt{\gamma} \leq |z| \leq \sqrt{\gamma(\gamma+1)}\}$ together with the point $\{0\}$ where there is a mass $1-\gamma^{-1}$.

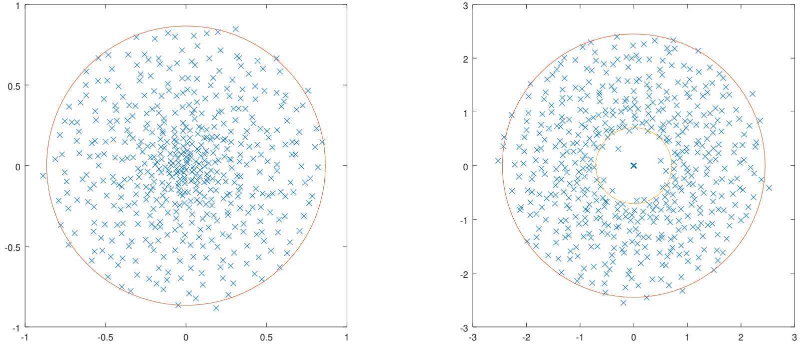
Moreover, $F(r)$ has a positive and analytical density on the open interval $(0 \vee \mathrm{sign}(\gamma-1)|\gamma-1|^{3/2}/\sqrt{\gamma}, \sqrt{\gamma(\gamma+1)})$.

This density is bounded if $\gamma \neq 1$. If $\gamma = 1$, then the density is bounded everywhere except when $r \downarrow 0$.

A cumbersome closed form expression for g^{-1} (and hence for $F(\cdot)$) can be obtained by calculating the root of a third degree polynomial. For the special case $\gamma = 1$, g^{-1} is given by

$$g^{-1}(t) = \frac{t^{1/3}}{2}\left(\left[1+\sqrt{1-\frac{t}{27}}\right]^{1/3} + \left[1-\sqrt{1-\frac{t}{27}}\right]^{1/3}\right), \quad 0 \le t \le 2.$$

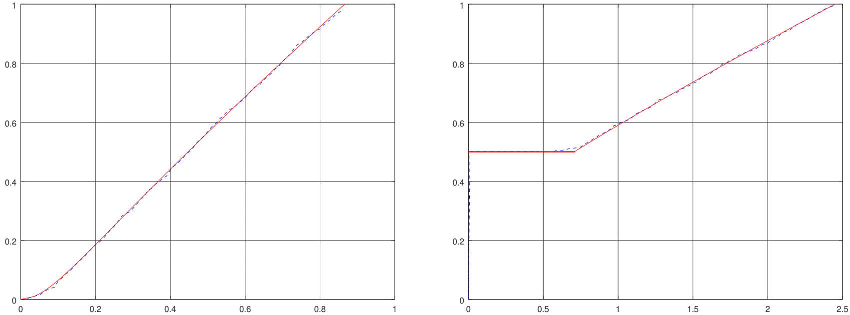
Eigenvalue realizations corresponding to the cases where $\gamma = 0.5$ and $\gamma = 2$ are shown in the next Figure.



: $(N, n) = (500, 1000)$: $(N, n) = (1000, 500)$

Figure: Eigenvalue realizations and LSD support.

Plot of F



: $(N, n) = (500, 1000)$: $(N, n) = (1000, 500)$

Figure: Plots of $F(r)$ (plain curves) and their empirical realizations (dashed curves).

LSD results and also normality of trace results are useful for graphical and significance tests of hypothesis.

The book has examples using the symmetrized autocovariances. This amounts to using the singular values. There is loss of information in dealing with singular values rather than eigenvalues.

Statistical applications with eigenvalues

Theorem 4 gives the so-called null distribution of the eigenvalues under white noise (IID) hypothesis. To test other hypothesis, we need LSD results for general MA(q) models. This problem is non-trivial.

log-potential

The *logarithmic potential* of a probability measure μ on $\mathbb{C}$ is the $\mathbb{C} \to (-\infty, \infty]$ superharmonic function defined as

$$U_\mu(z) = -\int_{\mathbb{C}} \log|\lambda - z| \, \mu(d\lambda) \ \text{(whenever the integral is finite).}$$

μ can be recovered from $U_\mu(\cdot)$: let $\Delta = \partial_x^2 + \partial_y^2 = 4\partial_z \partial_{\bar{z}}$ for $z = x + \imath y \in \mathbb{C}$ be the Laplace operator defined on

$$C_c^\infty(\mathbb{C}) = \{\varphi : \varphi \text{ is a compactly supported real valued smooth function on } \mathbb{C}\}.$$

Then

$$\mu = -(2\pi)^{-1} \Delta U_\mu \quad \text{in the sense that} \tag{0.4}$$

$$\int_{\mathbb{C}} \varphi(z) \, \mu(dz) = -\frac{1}{2\pi} \int_{\mathbb{C}} \Delta\varphi(z) \, U_\mu(z) \, dz, \ \forall \varphi \in C_c^\infty(\mathbb{C}).$$

Convergence of the logarithmic potentials for Lebesgue almost all $z \in \mathbb{C}$ implies the weak convergence of the underlying measures under a tightness criterion.

Observe that

$\sum X_t X_{t-1}^*$ is essentially the same as the matrix as $XJX^* = Y$ (say).

The logarithmic potential of the spectral measure of Y equals

$$
\begin{aligned}
U_{\mu_n}(z) &= -\frac{1}{N}\sum \log|\lambda_i - z| = -\frac{1}{N}\log|\det(Y-z)| \\
&= -\frac{1}{2N}\log\det(Y-z)(Y-z)^* = -\int \log\lambda \, \nu_{n,z}(d\lambda),
\end{aligned}
$$

where the probability measure $\nu_{n,z}$ is the *singular value* distribution of $Y - z$, given as

$$
\nu_{n,z} = \frac{1}{N}\sum_{i=0}^{N-1} \delta_{s_i(Y-z)}.
$$

Thus we need to study the asymptotic behavior of $U_{\mu_n}(z)$ for Lebesgue almost all $z \in \mathbb{C}$.

Lemma 4.3 of Bordenave and Chafai Let (M_n) be a sequence of random matrices with complex entries. Let ζ_n be its spectral measure and let $\sigma_{n,z}$ be the empirical singular value distribution of $M_n - z$. Assume that

(i) for almost every $z \in \mathbb{C}$, there exists a probability measure $\boldsymbol{\sigma}_z$ such that $\sigma_{n,z} \Rightarrow \boldsymbol{\sigma}_z$ in probability,

(ii) log is uniformly integrable in probability with respect to the sequence $(\sigma_{n,z})$.

Then, there exists a probability measure $\boldsymbol{\zeta}$ such that $\zeta_n \Rightarrow \boldsymbol{\zeta}$ in probability, and furthermore,

$$U_{\boldsymbol{\zeta}}(z) = -\int \log \lambda \, \boldsymbol{\sigma}_z(d\lambda) \quad \mathbb{C} - \text{a.e.}$$

Thus we need to establish that:

Step 1: for almost all $z \in \mathbb{C}$, $\nu_{n,z} \Rightarrow \nu_z$ (a deterministic probability measure) in probability.

Step 2: the function log is uniformly integrable with respect to the measure $\nu_{n,z}$ for almost all $z \in \mathbb{C}$ in probability. That is,

$$\forall \varepsilon > 0, \quad \lim_{T\to\infty} \limsup_{n\geq 1} \mathbb{P}\left[\int_0^\infty |\log \lambda| \, 1_{|\log \lambda| \geq T} \; \nu_{n,z} d(\lambda) > \varepsilon\right] = 0. \tag{0.5}$$

Then there exists a probability measure $\boldsymbol{\mu}$ such that $\mu_n \Rightarrow \boldsymbol{\mu}$ in probability, and $U_{\boldsymbol{\mu}}(z) = -\int \log|\lambda| \, \check{\boldsymbol{\nu}}_z(d\lambda)$ $\mathbb{C}$-almost everywhere. It would then remain to identify the measure $\boldsymbol{\mu}$ to complete the proof of Theorem 4.

Step 1 Proved by convergence of the trace of resolvent. There one first replaces trace by its expectation, then uses Gaussian approximation, PN inequality.

Step 2 requires control of both, the small eigenvalues and the large eigenvalues. The latter is achieved trivially (since spectral norm of X is almost surely bounded). The former is achieved by Theorem 3.

For detailed proof, see Bose and Hachem (2018).

Why is the general case difficult?

Consider

$$Y_t = X_t + \psi_1 X_{t-1}.$$

Then

$$\hat{\Gamma}_1 = \sum Y_t Y_{t-1}.$$

Substituting the first equation in the second, we get several terms which involve the autocovariance matrices of order 0 and 1 of $\{X_t\}$ together with the matrix ψ_1. This is now non-symmetric and a complicated expression to deal with. It gets even tougher with the increase in the order of the process.

Very selective references

Bhattacharjee and Bose (2019). Joint convergence of sample autocovariance matrices when $p/n \to 0$ with application. AoS.

Bhattacharjee and Bose (2018) (Book). *Large Covariance and Autocovariance Matrices*. Chapman and Hall.

Bhattacharjee and Bose (2016). Large sample behavior of high dimensional autocovariance matrices. AoS.

Bhattacharjee, Monika and Bose, Arup (2017). Matrix polynomial generalizations of the sample variance-covariance matrix when $pn^{-1} \to y \in (0, \infty)$. IJPAM, Errata, 49, 4, 783–788 (2018).

Bhattacharjee, Monika and Bose, Arup (2016). Matrix polynomial generalizations of the sample variance-covariance matrix when $pn^{-1} \to 0$. RMTA.

Bordenave and Chafai (2012). Around the circlar law. Prob Surveys.

Bose (2018). *Patterned Random Matrices*. Chapman & Hall.

Bose and Hachem (2018). Arxiv.

References to other relevant works are available in the above works.

THANK YOU !