Spectrum of Random Band Matrices

Indrajit Jana
(joint work with Alexander Soshnikov)

Temple University

March 30, 2018
Frontier Probability Days

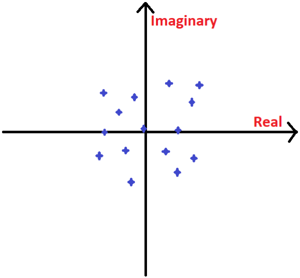
Definition (Random Matrix)

A random matrix is a matrix with random variables as the entries of the matrix. For example, $M_n = (m_{ij})_{k \times l}$ *with* $m_{ij} \overset{i.i.d}{\sim} \mathcal{N}(0,1)$ *is a rectangular random matrix.*

Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of an $n \times n$ random matrix M_n. Define the empirical spectral distribution

$$\mu_n := \sum_{i=1}^{n} \delta_{\lambda_i}.$$

Note that μ_n defines a random measure on the complex plane.



Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of an $n \times n$ random matrix M_n. Define the empirical spectral distribution

$$\mu_n := \sum_{i=1}^{n} \delta_{\lambda_i}.$$

Note that μ_n defines a random measure on the complex plane.

Questions

- What is the asymptotic behavior of μ_n, as $n \to \infty$?

Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of an $n \times n$ random matrix M_n. Define the empirical spectral distribution

$$\mu_n := \sum_{i=1}^{n} \delta_{\lambda_i}.$$

Note that μ_n defines a random measure on the complex plane.

Questions

- What is the asymptotic behavior of μ_n, as $n \to \infty$?
- Fluctuation of μ_n.

Three different matrix models

Definition (Wigner ensemble; symmetric)

Class of random matrices of the form $M = (m_{ij})_{n \times n}$ *such that* $m_{ij} = m_{ji}$ *for all* i, j *and* $\{m_{ij} : 1 \leq i \leq j \leq n\}$ *is a set of independent random variables.*

Three different matrix models

Definition (Wigner ensemble; symmetric)

Class of random matrices of the form $M = (m_{ij})_{n \times n}$ *such that* $m_{ij} = m_{ji}$ *for all* i, j *and* $\{m_{ij} : 1 \leq i \leq j \leq n\}$ *is a set of independent random variables.*

Definition (Sample covariance ensemble)

Class of random matrices of the form $M_{k \times k} = XX^*$*, where* $X_{k \times l}$ *is a random matrix with a set of independent random variables.*

Three different matrix models

Definition (Wigner ensemble; symmetric)

Class of random matrices of the form $M = (m_{ij})_{n \times n}$ *such that* $m_{ij} = m_{ji}$ *for all* i, j *and* $\{m_{ij} : 1 \leq i \leq j \leq n\}$ *is a set of independent random variables.*

Definition (Sample covariance ensemble)

Class of random matrices of the form $M_{k \times k} = XX^*$, *where* $X_{k \times l}$ *is a random matrix with a set of independent random variables.*

Definition (Ginibre ensemble; all iid entries)

Class of random matrices of the form $M = (m_{ij})_{n \times n}$, *where* m_{ij}*s are independently and identically distributed (iid) random variables.*

- Let $M_n = \frac{1}{\sqrt{n}} X_n = \frac{1}{\sqrt{n}} (x_{ij})_{n \times n}$ be a symmetric random matrix with independent entries such that $\mathbb{E}[x_{ij}] = 0$, $\mathbb{E}[x_{ij}^2] = 1$. Then the empirical spectral distribution of M_n converges almost surely to ρ_{sc}, where ρ_{sc} is the semicircle law whose pdf is given by

$$\rho_{sc}(x) = \frac{1}{2\pi} \sqrt{(4 - x^2)_+}.$$

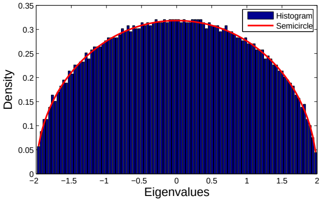


Figure: A MATLAB simulation done with a 4000×4000 Wigner matrix

- Let $M_n = \frac{1}{\sqrt{n}} X_n = \frac{1}{\sqrt{n}} (x_{ij})_{n \times n}$ be a symmetric random matrix with independent entries such that $\mathbb{E}[x_{ij}] = 0$, $\mathbb{E}[x_{ij}^2] = 1$. Then the empirical spectral distribution of M_n converges almost surely to ρ_{sc}, where ρ_{sc} is the semicircle law whose pdf is given by

$$\rho_{sc}(x) = \frac{1}{2\pi} \sqrt{(4 - x^2)_+}.$$

- Originally, this result was proved by using moment method (Wigner, 1955). It can be shown that the $2k$th moment of μ_n

$$\int x^{2k} \, d\mu_n(x) = \frac{1}{n} \sum_{j=1}^{n} \lambda_j^{2k} = \frac{1}{n} \text{tr} \left[\left(\frac{1}{\sqrt{n}} X_n \right)^{2k} \right] \to \frac{1}{k+1} \binom{2k}{k} = \int x^{2k} \rho_{sc}(x) \, dx.$$

And the odd moments vanishes.

- Let $M_n = \frac{1}{\sqrt{n}} X_n = \frac{1}{\sqrt{n}} (x_{ij})_{n \times n}$ be a symmetric random matrix with independent entries such that $\mathbb{E}[x_{ij}] = 0$, $\mathbb{E}[x_{ij}^2] = 1$. Then the empirical spectral distribution of M_n converges almost surely to ρ_{sc}, where ρ_{sc} is the semicircle law whose pdf is given by

$$\rho_{sc}(x) = \frac{1}{2\pi} \sqrt{(4 - x^2)_+}.$$

- Originally, this result was proved by using moment method (Wigner, 1955). It can be shown that the $2k$th moment of μ_n

$$\int x^{2k} \, d\mu_n(x) = \frac{1}{n} \sum_{j=1}^{n} \lambda_j^{2k} = \frac{1}{n} \mathrm{tr} \left[\left(\frac{1}{\sqrt{n}} X_n \right)^{2k} \right] \to \frac{1}{k+1} \binom{2k}{k} = \int x^{2k} \rho_{sc}(x) \, dx.$$

  And the odd moments vanishes.

- It can be shown that the Stieltjes transform of μ_n

$$s_n(z) = \int_{\mathbb{R}} \frac{d\mu_n(\lambda)}{\lambda - z} = \frac{1}{n} \mathrm{tr} \left(\frac{1}{\sqrt{n}} X_n - z \right)^{-1} \to \frac{-z + \sqrt{z^2 - 4}}{2} = \int_{\mathbb{R}} \frac{\rho_{sc}(x) \, dx}{x - z},$$

  for any $z \in \{z \in \mathbb{C} : \Im(z) > 0\}$.

Let $M = \frac{1}{n}XX^*$, where X be an $m \times n$ random matrix with i.i.d. entries with mean 0 and variance σ^2. Suppose $m/n \to \gamma$ as $m, n \to \infty$, then the empirical spectral distribution of M converges to the Marchenko-Pastur law. The probability density function is given by

$$\mu_{MP}(x) = \begin{cases} f(x) & \text{if } 0 \leq \gamma \leq 1 \\ \left(1 - \frac{1}{\gamma}\right)\delta_0 + f(x) & \text{if } \gamma > 1, \end{cases}$$

where

$$f(x) = \frac{1}{2\pi\sigma^2}\frac{\sqrt{(\gamma_+ - x)(x - \gamma_-)}}{\gamma x}\mathbf{1}_{[\gamma_-, \gamma_+]}(x), \quad \gamma_\pm = \sigma^2(1 \pm \sqrt{\gamma})^2.$$

This was proved by Vladimir Marchenko, and Leonid Pastur in 1967.

Let $M = \frac{1}{n} X X^*$, where X be an $m \times n$ random matrix with i.i.d. entries with mean 0 and variance σ^2. Suppose $m/n \to \gamma$ as $m, n \to \infty$, then the empirical spectral distribution of M converges to the Marchenko-Pastur law. The probability density function is given by

$$\mu_{MP}(x) = \begin{cases} f(x) & \text{if } 0 \leq \gamma \leq 1 \\ \left(1 - \frac{1}{\gamma}\right) \delta_0 + f(x) & \text{if } \gamma > 1. \end{cases}$$

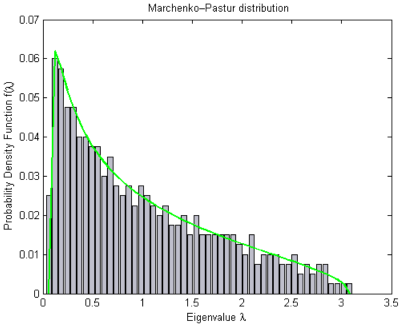


Figure: Done with a 400×700 random matrix with i.i.d. Gaussian entries
Downloaded from Mathworks.com

- Let M_n be an $n \times n$ matrix. If $M_n = \frac{1}{\sqrt{n}} X_n$, where x_{ij}, the entries of X_n, are iid complex normal variables with unit variance, then the joint density of $\lambda_1, \ldots, \lambda_n$ is given by

$$f(\lambda_1, \ldots, \lambda_n) = c_n \prod_{i<j} |\lambda_i - \lambda_j|^2 \prod_{i=1}^{n} e^{-n|\lambda_i|^2},$$

where c_n is the normalizing constant.

- Let M_n be an $n \times n$ matrix. If $M_n = \frac{1}{\sqrt{n}} X_n$, where x_{ij}, the entries of X_n, are iid complex normal variables with unit variance, then the joint density of $\lambda_1, \ldots, \lambda_n$ is given by

$$f(\lambda_1, \ldots, \lambda_n) = c_n \prod_{i<j} |\lambda_i - \lambda_j|^2 \prod_{i=1}^{n} e^{-n|\lambda_i|^2},$$

where c_n is the normalizing constant.

- Mehta (1967) Proved that, as $n \to \infty$, the eigenvalues of such matrices are uniformly distributed in the unit disk on the complex plane (Circular law).

- Let M_n be an $n \times n$ matrix. If $M_n = \frac{1}{\sqrt{n}} X_n$, where x_{ij}, the entries of X_n, are iid complex normal variables with unit variance, then the joint density of $\lambda_1, \ldots, \lambda_n$ is given by

$$f(\lambda_1, \ldots, \lambda_n) = c_n \prod_{i<j} |\lambda_i - \lambda_j|^2 \prod_{i=1}^{n} e^{-n|\lambda_i|^2},$$

where c_n is the normalizing constant.

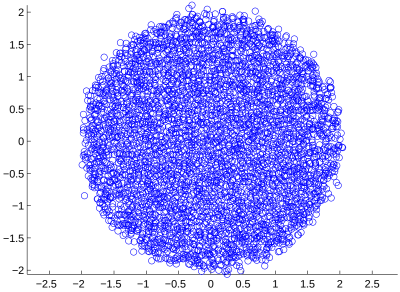


Figure: The circular law. Notice the inconsistency.

Random Band Matrices

Definition (Periodic band matrix)

An $n \times n$ *matrix* $M = (m_{ij})_{n \times n}$ *is called a periodic band matrix of bandwidth* b_n *if* $m_{ij} = 0$ *whenever* $b_n < |i - j| < n - b_n$*.*

Definition (Non-periodic band matrix)

M *is called a non-periodic band matrix of bandwidth* b_n *if* $m_{ij} = 0$ *whenever* $b_n < |i - j|$*.*

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & 0 & 0 & 0 & 0 & 0 & a_{19} & a_{1,10} \\ a_{21} & a_{22} & a_{23} & a_{24} & 0 & 0 & 0 & 0 & & a_{2,10} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & 0 & 0 & 0 & 0 & 0 \\ 0 & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} & 0 & 0 & 0 & 0 \\ & & & & & \vdots & & & & \\ 0 & 0 & 0 & 0 & 0 & a_{86} & a_{87} & a_{88} & a_{89} & a_{8,10} \\ a_{91} & 0 & 0 & 0 & 0 & 0 & a_{97} & a_{98} & a_{99} & a_{9,10} \\ a_{10,1} & a_{10,2} & 0 & 0 & 0 & 0 & 0 & a_{10,8} & a_{10,9} & a_{10,10} \end{bmatrix}.$$

The above is a 10×10 periodic band matrix of bandwidth 2

Definition (Periodic band matrix)

An $n \times n$ *matrix* $M = (m_{ij})_{n\times n}$ *is called a periodic band matrix of bandwidth* b_n *if* $m_{ij} = 0$ *whenever* $b_n < |i-j| < n - b_n$*.*

Definition (Non-periodic band matrix)

M *is called a non-periodic band matrix of bandwidth* b_n *if* $m_{ij} = 0$ *whenever* $b_n < |i-j|$*.*

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & a_{24} & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & 0 & 0 & 0 & 0 & 0 \\ 0 & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} & 0 & 0 & 0 & 0 \\ & & & & & \vdots & & & & \\ 0 & 0 & 0 & 0 & 0 & a_{86} & a_{87} & a_{88} & a_{89} & a_{8,10} \\ 0 & 0 & 0 & 0 & 0 & 0 & a_{97} & a_{98} & a_{99} & a_{9,10} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_{10,8} & a_{10,9} & a_{10,10} \end{bmatrix}.$$

The above is a 10×10 non-periodic band matrix of bandwidth 2

Definition (Periodic band matrix)

An $n \times n$ *matrix* $M = (m_{ij})_{n\times n}$ *is called a periodic band matrix of bandwidth* b_n *if* $m_{ij} = 0$ *whenever* $b_n < |i-j| < n - b_n$.

Definition (Non-periodic band matrix)

M *is called a non-periodic band matrix of bandwidth* b_n *if* $m_{ij} = 0$ *whenever* $b_n < |i-j|$.

- What is the joint density function of the eigen values of a random band matrix?

Definition (Periodic band matrix)

An $n \times n$ *matrix* $M = (m_{ij})_{n \times n}$ *is called a periodic band matrix of bandwidth* b_n *if* $m_{ij} = 0$ *whenever* $b_n < |i - j| < n - b_n$.

Definition (Non-periodic band matrix)

M *is called a non-periodic band matrix of bandwidth* b_n *if* $m_{ij} = 0$ *whenever* $b_n < |i - j|$.

- What is the joint density function of the eigen values of a random band matrix?
- What if we assume that the entries are standard Gaussian?

- In 1991, the Semicircle Law for random band matrices was conjectured by G. Casati et al.

- In 1991, the Semicircle Law for random band matrices was conjectured by G. Casati et al.
- In 1992, Molchanov et. al. proved it when $\frac{b_n}{n} \to 0$ and $b_n \to \infty$ as $n \to \infty$.

Let $X = (x_{ij})_{n\times n}$ be an $n \times n$ periodic band matrix of bandwidth b_n, where $b_n \to \infty$ as $n \to \infty$. Let R be a sequence of $n \times n$ deterministic periodic band matrices of bandwidth b_n. Let us denote $c_n = 2b_n + 1$ and μ_M be the ESD of M. Assume that

$$\begin{aligned}
&(a)\ \mu_{\frac{1}{c_n}RR^*} \to H, \text{for some non random probability distribution } H\\
&(b)\ \{x_{jk}:\ k \in I_j,\ 1 \le j \le n\} \text{ is an iid set of random variables,}\\
&(c)\ \mathbb{E}[x_{11}] = 0, \mathbb{E}[|x_{11}|^2] = 1,
\end{aligned} \tag{1}$$

and define $(d)\ Y = \dfrac{1}{\sqrt{c_n}}(R + \sigma X),$ where $\sigma > 0$ is fixed.

Let $X = (x_{ij})_{n \times n}$ be an $n \times n$ periodic band matrix of bandwidth b_n, where $b_n \to \infty$ as $n \to \infty$. Let R be a sequence of $n \times n$ deterministic periodic band matrices of bandwidth b_n. Let us denote $c_n = 2b_n + 1$ and μ_M be the ESD of M. Assume that

$$\begin{aligned} &(a)\ \mu_{\frac{1}{c_n}RR^*} \to H, \text{for some non random probability distribution } H \\ &(b)\ \{x_{jk} :\ k \in I_j,\ 1 \le j \le n\} \text{ is an iid set of random variables,} \\ &(c)\ \mathbb{E}[x_{11}] = 0, \mathbb{E}[|x_{11}|^2] = 1, \\ \text{and define } &(d)\ Y = \frac{1}{\sqrt{c_n}}(R + \sigma X), \text{ where } \sigma > 0 \text{ is fixed.} \end{aligned} \tag{1}$$

Definition (Poincaré inequality)

Let X be a $\mathbb{R}^k$ valued random variable with probability measure μ. The probability measure μ is said to satisfy the Poincaré inequality with constant $m > 0$, if for all continuously differentiable functions $f : \mathbb{R}^k \to \mathbb{R}$,

$$Var(f(X)) \le \frac{1}{m}\mathbb{E}(|\nabla f(X)|^2).$$

Let $X = (x_{ij})_{n\times n}$ be an $n \times n$ periodic band matrix of bandwidth b_n, where $b_n \to \infty$ as $n \to \infty$. Let R be a sequence of $n \times n$ deterministic periodic band matrices of bandwidth b_n. Let us denote $c_n = 2b_n + 1$ and μ_M be the ESD of M. Assume that

$$\begin{aligned}
&(a)\ \mu_{\frac{1}{c_n}RR^*} \to H, \text{for some non random probability distribution } H \\
&(b)\ \{x_{jk} :\ k \in I_j,\ 1 \leq j \leq n\} \text{ is an iid set of random variables,} \\
&(c)\ \mathbb{E}[x_{11}] = 0, \mathbb{E}[|x_{11}|^2] = 1,
\end{aligned} \tag{1}$$

and define $(d)\ Y = \frac{1}{\sqrt{c_n}}(R + \sigma X),$ where $\sigma > 0$ is fixed.

Definition (Poincaré inequality)

Let X be a $\mathbb{R}^k$ valued random variable with probability measure μ. The probability measure μ is said to satisfy the Poincaré inequality with constant $m > 0$, if for all continuously differentiable functions $f : \mathbb{R}^k \to \mathbb{R}$,

$$\mathit{Var}(f(X)) \leq \frac{1}{m}\mathbb{E}(|\nabla f(X)|^2).$$

- For example, the uniform distribution on $[0, 1]$, the standard Gaussian distribution satisfy the Poincaré inequality.

Theorem (without Poincaré)

Let Y be the band matrix as defined in (1). In addition to the existing assumption, assume that

$$(i)\ \frac{n}{c_n^2} \to 0,$$
$$(ii)\ H \text{ is compactly supported}$$
$$(iii)\ \mathbb{E}[|x_{11}|^{2p}] < \infty, \text{ for some } p \in \mathbb{N}.$$

Then $\mathbb{E}|m_n(z) - m(z)|^p \to 0$ uniformly for all $z \in \{z : \Im(z) > \eta\}$ for any fixed $\eta > 0$, where $m_n(z) = \frac{1}{n}\sum_{i=1}^{n}(\lambda_i(YY^*) - z)^{-1}$ is the empirical Stieltjes transform of YY^*, and $m(z) = \int_{\mathbb{R}} \frac{d\mu(x)}{x-z}$. In particular, the expected ESD of YY^* converges. In addition, the Stieltjes transform of μ satisfies

$$m(z) = \int_{\mathbb{R}} \frac{dH(t)}{\frac{t}{1+\sigma^2 m(z)} - (1 + \sigma^2 m(z))z} \quad \text{for any } z \in \mathbb{C}^+.$$

Theorem (without Poincaré)

Let Y be the band matrix as defined in (1). In addition to the existing assumption, assume that

$$(i)\ \frac{n}{c_n^2} \to 0,$$
$$(ii)\ H \text{ is compactly supported}$$
$$(iii)\ \mathbb{E}[|x_{11}|^{2p}] < \infty, \text{ for some } p \in \mathbb{N}.$$

Then $\mathbb{E}|m_n(z) - m(z)|^p \to 0$ uniformly for all $z \in \{z : \Im(z) > \eta\}$ for any fixed $\eta > 0$, where $m_n(z) = \frac{1}{n}\sum_{i=1}^{n}(\lambda_i(YY^*) - z)^{-1}$ is the empirical Stieltjes transform of YY^*, and $m(z) = \int_{\mathbb{R}} \frac{d\mu(x)}{x-z}$. In particular, the expected ESD of YY^* converges. In addition, the Stieltjes transform of μ satisfies

$$m(z) = \int_{\mathbb{R}} \frac{dH(t)}{\frac{t}{1+\sigma^2 m(z)} - (1+\sigma^2 m(z))z} \quad \text{for any } z \in \mathbb{C}^+.$$

Remark

If $R = 0$ and $\sigma = 1$, then the above integral equation becomes $m(z)(1 + m(z))z + 1 = 0$ which yields the Stieltjes transform of the Marchenko-Pastur law.

Theorem (under Poincaré assumption)

Let Y be the same as (1). In addition to the existing assumption, assume that

$(i)\ \log n = O(c_n)$

(ii) H is compactly supported

(iii) Both $\Re(x_{ij})$ and $\Im(x_{ij})$ satisfy Poincaré inequality with constant m.

Then $\mathbb{E}|m_n(z) - m(z)| \to 0$ uniformly for all $z \in \{z : \Im(z) > \eta\}$ for any fixed $\eta > 0$, and $m(z)$ satisfies

$$m(z) = \int_{\mathbb{R}} \frac{dH(t)}{\frac{t}{1+\sigma^2 m(z)} - (1+\sigma^2 m(z))z} \quad \text{for any } z \in \mathbb{C}^+. \qquad (2)$$

- Eigenvalues are sensitive to small changes of the matrix entries

$$M = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad M_\epsilon = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \epsilon & 0 & 0 \end{pmatrix}$$

Eigenvalues of M are all zero and those of M_ϵ is $\lambda_k = \epsilon^{1/3} e^{2k\pi i/3},\ k = 0, 1, 2.$

- Eigenvalues are sensitive to small changes of the matrix entries

$$M = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad M_\epsilon = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \epsilon & 0 & 0 \end{pmatrix}$$

  Eigenvalues of M are all zero and those of M_ϵ is $\lambda_k = \epsilon^{1/3} e^{2k\pi i/3},\ k = 0, 1, 2.$

- If we take $\epsilon_n = 1/n$, then $\lambda_k^{(n)} = n^{-1/n} e^{2k\pi i/n},\ k = 0, \ldots, n-1.$ Observation; $|\lambda_k^{(n)}| \to 1.$

- By Cauchy integral formula, $\mathbb{E}[Z^k] = 0$ for any random variable Z which is uniformly distributed over any bounded simply connected region. So proving $\frac{1}{n}\text{tr}(X_n/\sqrt{n})^k \to 0$ does not prove the circular law.

- By Cauchy integral formula, $\mathbb{E}[Z^k] = 0$ for any random variable Z which is uniformly distributed over any bounded simply connected region. So proving $\frac{1}{n}\text{tr}(X_n/\sqrt{n})^k \to 0$ does not prove the circular law.

- The Stieltjes transform $s_n(z) = \frac{1}{n}\text{tr}(X_n/\sqrt{n} - zI)^{-1}$ should satisfy $s_n(z) \to -1/z$ as $n \to \infty$. But again this does not uniquely identify the uniform distribution over unit disk.

- Let $z = s + it$. The real part of the Stieltjes transform can be written as

$$
\begin{aligned}
m_{nr}(z) &:= \Re(m_n(z)) \\
&= \frac{1}{n}\sum_{i=1}^{n} \frac{\Re(\lambda_i - z)}{|\lambda_i - z|^2} \\
&= -\frac{1}{2}\frac{\partial}{\partial s}\int_0^{\infty} \log x \nu_n(dx, z),
\end{aligned}
$$

where $\nu_n(\cdot, z)$ is the ESD of $(\frac{1}{\sqrt{n}}X_n - zI)(\frac{1}{\sqrt{n}}X_n - zI)^*$.

- Let $z = s + it$. The real part of the Stieltjes transform can be written as

$$\begin{aligned} m_{nr}(z) &:= \Re(m_n(z)) \\ &= \frac{1}{n}\sum_{i=1}^{n} \frac{\Re(\lambda_i - z)}{|\lambda_i - z|^2} \\ &= -\frac{1}{2}\frac{\partial}{\partial s}\int_0^\infty \log x \nu_n(dx, z), \end{aligned}$$

  where $\nu_n(\cdot, z)$ is the ESD of $(\frac{1}{\sqrt{n}}X_n - zI)(\frac{1}{\sqrt{n}}X_n - zI)^*$.

- Lemma (Girko): Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of an $n \times n$ random matrix M_n and

$$\mu_n(x, y) := \frac{1}{n}\#\{\lambda_i, 1 \le i \le n : \Re(\lambda_i) \le x, \Im(\lambda_i) \le y\}$$

  be the empirical spectral distribution (ESD) of M_n.

$$\int\int e^{i(ux+vy)}\mu_n(dx, dy) = \frac{u^2 + v^2}{i4\pi u}\int\int \frac{\partial}{\partial s}\left[\int_0^\infty \log x \nu_n(dx, z)\right] e^{i(us+vt)}\, dtds,$$

Possible application

- Girko's lemma leads a way to prove the Circular law.

- Girko's lemma leads a way to prove the Circular law.
- Bai, Silverstein (1995) proved the limiting ESD of $(\frac{1}{\sqrt{n}}X_n - zI)(\frac{1}{\sqrt{n}}X_n - zI)^*$ type of matrices, and proved the Circular law for random matrices which have bounded $(2+\delta)$ moments of the entries. More recently, Tao, Vu (2007) proved it only under the second moment assumption.

- Girko's lemma leads a way to prove the Circular law.
- Bai, Silverstein (1995) proved the limiting ESD of $(\frac{1}{\sqrt{n}}X_n - zI)(\frac{1}{\sqrt{n}}X_n - zI)^*$ type of matrices, and proved the Circular law for random matrices which have bounded $(2+\delta)$ moments of the entries. More recently, Tao, Vu (2007) proved it only under the second moment assumption.
- Controlling the smallest eigenvalue of $(\frac{1}{\sqrt{n}}X_n - zI)(\frac{1}{\sqrt{n}}X_n - zI)^*$ is nontrivial.

- Girko's lemma leads a way to prove the Circular law.
- Bai, Silverstein (1995) proved the limiting ESD of $(\frac{1}{\sqrt{n}}X_n - zI)(\frac{1}{\sqrt{n}}X_n - zI)^*$ type of matrices, and proved the Circular law for random matrices which have bounded $(2+\delta)$ moments of the entries. More recently, Tao, Vu (2007) proved it only under the second moment assumption.
- Controlling the smallest eigenvalue of $(\frac{1}{\sqrt{n}}X_n - zI)(\frac{1}{\sqrt{n}}X_n - zI)^*$ is nontrivial.
- It is believed that the circular law should hold for random band matrices of bandwidth $b_n >> \sqrt{n}$.

- Girko's lemma leads a way to prove the Circular law.
- Bai, Silverstein (1995) proved the limiting ESD of $(\frac{1}{\sqrt{n}}X_n - zI)(\frac{1}{\sqrt{n}}X_n - zI)^*$ type of matrices, and proved the Circular law for random matrices which have bounded $(2+\delta)$ moments of the entries. More recently, Tao, Vu (2007) proved it only under the second moment assumption.
- Controlling the smallest eigenvalue of $(\frac{1}{\sqrt{n}}X_n - zI)(\frac{1}{\sqrt{n}}X_n - zI)^*$ is nontrivial.
- It is believed that the circular law should hold for random band matrices of bandwidth $b_n >> \sqrt{n}$.

¡Thank you for your attention!

The following notations are introduced for convenience of writing the proof.

$$\begin{aligned} A &= \frac{RR^*}{c_n(1+\sigma^2 m_n)} - \sigma^2 z m_n I \\ B &= A - zI \\ C &= YY^* - zI \\ C_j &= C - y_j y_j^*, \end{aligned}$$

where p_j is the jth column of the matrix P.

The following notations are introduced for convenience of writing the proof.

$$
\begin{aligned}
A &= \frac{RR^*}{c_n(1+\sigma^2 m_n)} - \sigma^2 z m_n I \\
B &= A - zI \\
C &= YY^* - zI \\
C_j &= C - y_j y_j^*,
\end{aligned}
$$

where p_j is the jth column of the matrix P.

Remark

The eigenvalues of $A - zI$ are given by $\frac{\lambda_i}{1+\sigma^2 m_n} - (1+\sigma^2 m_n)z$, where λ_is are eigenvalue of $\frac{1}{c_n} RR^*$. Therefore $\int_{\mathbb{R}} \frac{dH(t)}{\frac{t}{1+\sigma^2 m} - (1+\sigma^2 m)z}$ can be thought of as $\frac{1}{n}\text{tr}(A - zI)^{-1}$ for large n. So heuristically, proving the theorem is same as showing that $\frac{1}{n}\text{tr}(A - zI)^{-1} - m_n \to 0$ as $n \to \infty$.

Lemma (Sherman-Morrison formula)

Let $P_{n\times n}$ and $(P+vv^*)$ be invertible matrices, where $v\in\mathbb{C}^n$. Then we have

$$(P+vv^*)^{-1} = P^{-1} - \frac{P^{-1}vv^*P^{-1}}{1+v^*P^{-1}v}.$$

In particular,

$$v^*(P+vv^*)^{-1} = \frac{v^*P^{-1}}{1+v^*P^{-1}v}.$$

Lemma (Sherman-Morrison formula)

Let $P_{n\times n}$ and $(P+vv^*)$ be invertible matrices, where $v \in \mathbb{C}^n$. Then we have

$$(P+vv^*)^{-1} = P^{-1} - \frac{P^{-1}vv^*P^{-1}}{1+v^*P^{-1}v}.$$

In particular,

$$v^*(P+vv^*)^{-1} = \frac{v^*P^{-1}}{1+v^*P^{-1}v}.$$

Using the Sherman-Morrison formula we have

$$I + zC^{-1} = YY^*C^{-1} = \sum_{j=1}^{n} y_j \frac{y_j^* C_j^{-1}}{1+y_j C_j^{-1} y_j^*}.$$

Lemma (Sherman-Morrison formula)

Let $P_{n\times n}$ and $(P+vv^*)$ be invertible matrices, where $v\in\mathbb{C}^n$. Then we have

$$(P+vv^*)^{-1}=P^{-1}-\frac{P^{-1}vv^*P^{-1}}{1+v^*P^{-1}v}.$$

In particular,

$$v^*(P+vv^*)^{-1}=\frac{v^*P^{-1}}{1+v^*P^{-1}v}.$$

Using the Sherman-Morrison formula we have

$$I+zC^{-1}=YY^*C^{-1}=\sum_{j=1}^{n}y_j\frac{y_j^*C_j^{-1}}{1+y_jC_j^{-1}y_j^*}.$$

Taking trace and dividing by n on the both sides we obtain

$$zm_n=\frac{1}{n}\sum_{j=1}^{n}\frac{y_j^*C_j^{-1}y_j}{1+y_jC_j^{-1}y_j^*}-1=-\frac{1}{n}\sum_{j=1}^{n}\frac{1}{1+y_j^*C_j^{-1}y_j}. \qquad (3)$$

Using the resolvent identity,

$$
\begin{aligned}
& B^{-1} - C^{-1} = B^{-1}(YY^* - A)C^{-1} \\
& = \frac{1}{c_n} B^{-1} \left[RR^* + \sigma R X^* + \sigma X R^* + \sigma^2 X X^* - \frac{1}{1+\sigma^2 m_n} RR^* + c_n \sigma^2 z m_n \right] C^{-1} \\
& = \frac{1}{c_n} \sum_{j=1}^{n} B^{-1} \left[\frac{\sigma^2 m_n}{1+\sigma^2 m_n} r_j r_j^* + \sigma r_j x_j^* + \sigma x_j r_j^* + \sigma^2 x_j x_j^* - \frac{c_n}{n} \frac{1}{1 + y_j^* C_j^{-1} y_j} \sigma^2 \right] C^{-1}.
\end{aligned}
$$

Using the resolvent identity,

$$B^{-1} - C^{-1} = B^{-1}(YY^* - A)C^{-1}$$
$$= \frac{1}{c_n} B^{-1} \left[RR^* + \sigma RX^* + \sigma XR^* + \sigma^2 XX^* - \frac{1}{1+\sigma^2 m_n} RR^* + c_n \sigma^2 z m_n \right] C^{-1}$$
$$= \frac{1}{c_n} \sum_{j=1}^{n} B^{-1} \left[\frac{\sigma^2 m_n}{1+\sigma^2 m_n} r_j r_j^* + \sigma r_j x_j^* + \sigma x_j r_j^* + \sigma^2 x_j x_j^* - \frac{c_n}{n} \frac{1}{1+y_j^* C_j^{-1} y_j} \sigma^2 \right] C^{-1}.$$

Taking the trace, dividing by n, and using (3), we have

$$\frac{1}{n}\mathrm{tr} B^{-1} - m_n$$
$$= \frac{1}{n} \sum_{j=1}^{n} \left[\frac{\sigma^2 m_n}{1+\sigma^2 m_n} \frac{1}{c_n} r_j^* C^{-1} B^{-1} r_j + \frac{1}{c_n} \sigma x_j^* C^{-1} B^{-1} r_j + \frac{1}{c_n} \sigma r_j^* C^{-1} B^{-1} x_j \right.$$
$$\left. + \frac{1}{c_n} \sigma^2 x_j^* C^{-1} B^{-1} x_j - \frac{1}{1+y_j^* C_j^{-1} y_j} \frac{1}{n} \sigma^2 \mathrm{tr} C^{-1} B^{-1} \right]$$
$$\equiv \frac{1}{n} \sum_{j=1}^{n} \left[T_{1,j} + T_{2,j} + T_{3,j} + T_{4,j} + T_{5,j} \right]. \tag{4}$$

We introduce the following notations for convenience

$$
\begin{aligned}
\rho_j &= \frac{1}{c_n} r_j^* C_j^{-1} r_j, \quad \omega_j = \frac{1}{c_n} \sigma^2 x_j^* C_j^{-1} x_j, \\
\beta_j &= \frac{1}{c_n} \sigma r_j^* C_j^{-1} x_j, \quad \gamma_j = \frac{1}{c_n} \sigma x_j^* C_j^{-1} r_j, \\
\hat{\rho}_j &= \frac{1}{c_n} r_j^* C_j^{-1} B^{-1} r_j, \quad \hat{\omega}_j = \frac{1}{c_n} \sigma^2 x_j^* C_j^{-1} B^{-1} x_j, \\
\hat{\beta}_j &= \frac{1}{c_n} \sigma r_j^* C_j^{-1} B^{-1} x_j, \quad \hat{\gamma}_j = \frac{1}{c_n} \sigma x_j^* C_j^{-1} B^{-1} r_j, \\
\alpha_j &= 1 + \frac{1}{c_n} (r_j + \sigma x_j)^* C_j^{-1} (r_j + \sigma x_j) = 1 + \rho_j + \beta_j + \gamma_j + \omega_j.
\end{aligned} \tag{5}
$$

We introduce the following notations for convenience

$$\begin{aligned}
\rho_j &= \frac{1}{c_n} r_j^* C_j^{-1} r_j, \quad \omega_j = \frac{1}{c_n} \sigma^2 x_j^* C_j^{-1} x_j, \\
\beta_j &= \frac{1}{c_n} \sigma r_j^* C_j^{-1} x_j, \quad \gamma_j = \frac{1}{c_n} \sigma x_j^* C_j^{-1} r_j, \\
\hat{\rho}_j &= \frac{1}{c_n} r_j^* C_j^{-1} B^{-1} r_j, \quad \hat{\omega}_j = \frac{1}{c_n} \sigma^2 x_j^* C_j^{-1} B^{-1} x_j, \\
\hat{\beta}_j &= \frac{1}{c_n} \sigma r_j^* C_j^{-1} B^{-1} x_j, \quad \hat{\gamma}_j = \frac{1}{c_n} \sigma x_j^* C_j^{-1} B^{-1} r_j, \\
\alpha_j &= 1 + \frac{1}{c_n} (r_j + \sigma x_j)^* C_j^{-1} (r_j + \sigma x_j) = 1 + \rho_j + \beta_j + \gamma_j + \omega_j.
\end{aligned} \tag{5}$$

Using the Sherman-Morrison formula, (4) can be written as

$$\begin{aligned}
\frac{1}{n} \mathrm{tr} B^{-1} - m_n = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\alpha_j} \Bigg[&\frac{1}{1 + \sigma^2 m_n} (\sigma^2 m_n - \gamma_j - \omega_j) \hat{\rho}_j \\
&+ \frac{1}{1 + \sigma^2 m_n} (1 + \rho_j + \beta_j + \sigma^2 m_n) \hat{\gamma}_j + \hat{\beta}_j + \hat{\omega}_j - \frac{1}{n} \sigma^2 \mathrm{tr} C^{-1} B^{-1} \Bigg].
\end{aligned} \tag{6}$$

Lemma (Effect of rank one perturbation on the partial trace of resolvent)

Let P and Q be $n \times n$ Hermition matrices, and $I \subset \{1, 2, \ldots, n\}$, then

$$\left| \sum_{k \in I} (P - zI)^{-1}_{kk} - \sum_{k \in I} (Q - zI)^{-1}_{kk} \right| \leq \frac{2}{\Im(z)} \text{rank}(P - Q).$$

Lemma (Effect of rank one perturbation on the partial trace of resolvent)

Let P and Q be $n \times n$ Hermition matrices, and $I \subset \{1, 2, \ldots, n\}$, then

$$\left| \sum_{k \in I} (P - zI)^{-1}_{kk} - \sum_{k \in I} (Q - zI)^{-1}_{kk} \right| \leq \frac{2}{\Im(z)} \text{rank}(P - Q).$$

The above and Hoeffding's inequality together yield the following tail bounds

$$\mathbb{P}\left(\left| \sum_{k \in I_j} M_{kk} - \mathbb{E} \sum_{k \in I_j} M_{kk} \right| > t \right) \leq 2 \exp\left\{ -\frac{\Im(z)^2 t^2}{32n} \right\},$$

for $M = C_j^{-1}$, $C_j^{-1} B_j^{-1}$, $C_j^{-1} r_j r_j^* C_j^{-1*}$, and $C_j^{-1} B_j^{-1} r_j r_j^* B_j^{-1*} C_j^{-1*}$.

Lemma (Effect of rank one perturbation on the partial trace of resolvent)

Let P and Q be $n \times n$ Hermition matrices, and $I \subset \{1, 2, \ldots, n\}$, then

$$\left| \sum_{k \in I} (P - zI)^{-1}_{kk} - \sum_{k \in I} (Q - zI)^{-1}_{kk} \right| \leq \frac{2}{\Im(z)} \text{rank}(P - Q).$$

The above and Hoeffding's inequality together yield the following tail bounds

$$\mathbb{P}\left(\left| \sum_{k \in I_j} M_{kk} - \mathbb{E} \sum_{k \in I_j} M_{kk} \right| > t \right) \leq 2 \exp\left\{ -\frac{\Im(z)^2 t^2}{32n} \right\},$$

for $M = C_j^{-1}$, $C_j^{-1} B_j^{-1}$, $C_j^{-1} r_j r_j^* C_j^{-1*}$, and $C_j^{-1} B_j^{-1} r_j r_j^* B_j^{-1*} C_j^{-1*}$.

- Applying these tail estimates on (6), we have the result.

Lemma (Effect of rank one perturbation on the partial trace of resolvent)

Let P and Q be $n \times n$ Hermition matrices, and $I \subset \{1, 2, \ldots, n\}$, then

$$\left| \sum_{k \in I} (P - zI)^{-1}_{kk} - \sum_{k \in I} (Q - zI)^{-1}_{kk} \right| \leq \frac{2}{\Im(z)} \text{rank}(P - Q).$$

The above and Hoeffding's inequality together yield the following tail bounds

$$\mathbb{P}\left(\left| \sum_{k \in I_j} M_{kk} - \mathbb{E} \sum_{k \in I_j} M_{kk} \right| > t \right) \leq 2 \exp\left\{ -\frac{\Im(z)^2 t^2}{32n} \right\},$$

for $M = C_j^{-1}$, $C_j^{-1} B_j^{-1}$, $C_j^{-1} r_j r_j^* C_j^{-1*}$, and $C_j^{-1} B_j^{-1} r_j r_j^* B_j^{-1*} C_j^{-1*}$.

- Applying these tail estimates on (6), we have the result.

¡Thanks Again!