Spectral properties of large Non-Hermitian Random Matrices

David Renfrew

Department of Mathematics
University of Colorado, Boulder

May 9, 2016

Non-Hermitian random matrices

- C_N is an $N \times N$ real random matrix with i.i.d entries such that

$$\mathbb{E}[C_{ij}] = 0 \quad \mathbb{E}[C_{ij}^2] = \sigma^2/N$$

- We study in the large N limit of the empirical spectral measure:

$$\mu_N(z) = \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i}(z)$$

where $\{\lambda_i\}$ are the eigenvalues of C_N.

Circular law

- Girko (84), Bai (97), ..., Tao-Vu (12).
- As $N \to \infty$, $\mu_{N}(z)$ converges a.s. in distribution to μ_c, the uniform law on the disk of radius σ,

$$\frac{d\mu_c(z)}{dz} = \frac{1}{\sigma^2 \pi} \mathbf{1}_{|z| \leq \sigma},$$

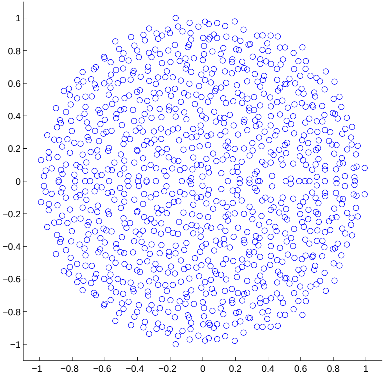


Figure : Eigenvalues of a 1000×1000 iid random matrix

Hermitian random matrices

- W_N is an $N \times N$ real random matrix with i.i.d entries on and above the diagonal such that

$$W_{ij} = W_{ji}$$

$$\mathbb{E}[W_{ij}] = 0 \quad \mathbb{E}[W_{ij}^2] = \sigma^2/N$$

- When W_{ij} are Gaussian, this is known as the Gaussian Orthogonal Ensemble (GOE).

Semi-circular law

- Wigner (55)
- As $N \to \infty$, $\mu_N(z)$ converges a.s. in distribution to μ_s, the semicircular law,

$$\frac{d\mu_s(z)}{dz} = \frac{1}{2\pi\sigma^2}\sqrt{4\sigma^2 - x^2},$$

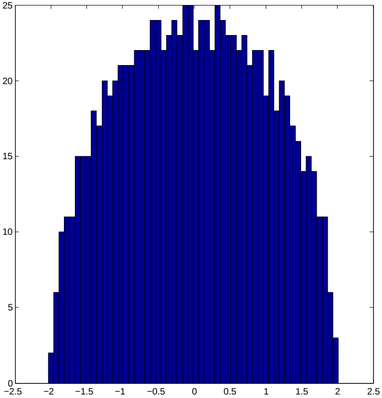


Figure : Eigenvalues of a 1000 $\times$ 1000 Wigner random matrix

Tools

- The key tool in studying the spectrum of Hermitian Matrices is the Stieltjes transform of μ

$$m(z) = \int \frac{d\mu(t)}{t - z}$$

- Inverse transform

$$\mu(x) = \frac{1}{\pi} \lim_{y \searrow 0} \mathfrak{Im}\, m(x + iy)$$

Tools

- The Stieltjes transform of an ESM is related to the is the resolvent

$$(X - z)^{-1}$$

by taking the trace

$$\int \frac{d\mu(x)}{x - z} = \frac{1}{N}\mathrm{Tr}(X - z)^{-1}$$

- In the non-Hermitian case the resolvent is not uniformly bounded inside the spectrum.

Semicircle Law

- The semicircle law is a central object in Free Probability.
- The Stieltjes transform of the semicircle law satisfies the relationship

$$m_\sigma(z) = -(z + \sigma^2 m_\sigma(z))^{-1}$$

Structured Random Matrices

- In a generalized model we divide the random matrix into D^2 blocks.
- Let g be a $D \times D$ matrix with real, positive entries.
- Let α be a D dimensional vector such that

$$\alpha_i > 0, \quad \sum_{i=1}^{D} \alpha_i = 1$$

- The cd^{th} block has size $\lfloor \alpha_c N \rfloor \times \lfloor \alpha_d N \rfloor$.
- Each entry of the cd^{th} block has variance $g_{cd}^2/N > 0$.

Structured random matrices

Let C_N be an $N \times N$ matrix with iid random entries with zero mean, variance $1/N$, and finite fourth moment.

let

$$c_i = \left\{ c \,\middle|\, \frac{i}{N} \in \left(\sum_{d=1}^{c-1} \alpha_d, \sum_{d=1}^{c} \alpha_d \right] \right\}$$

Let X_N be an $N \times N$ random matrix whose i, j entry is

$$X_{ij} := g_{c_i c_j} C_{ij}.$$

Main Theorem

The support of the limiting is determined by the matrix $D \times D$, G:

$$G_{cd} = \alpha_c g_{cd}^2.$$

This matrix has positive entries.
It's largest eigenvalue is real and equal to the spectral radius of G.
The associated eigenvector has strictly positive entries.

Main Theorem

Theorem (Aljadeff, R., Stern)

The empirical spectral measure converges almost surely to a deterministic measure μ.

The density of μ is radially symmetric and its support is a disk with radius $\sqrt{\rho(G)}$.

The density of μ is determined by a fixed point equation.

Fixed point equation

$$a_c = \frac{[Ga]_c + \eta}{|z|^2 - ([\widehat{G}\widehat{a}]_c + \eta)([Ga]_c + \eta)}$$

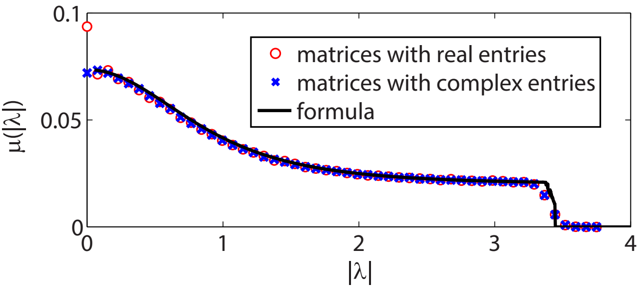
$$\widehat{a}_c = \frac{([\widehat{G}\widehat{a}]_c + \eta) + \eta}{|z|^2 - ([\widehat{G}\widehat{a}]_c + \eta)([Ga]_c + \eta)}$$

$$\mu(D(0,|z|)) = 1 - \sum_c \widehat{a}_c(z)[Ga(z)]_c$$

for $0 \leq |z| \leq \sqrt{\rho(G)}$

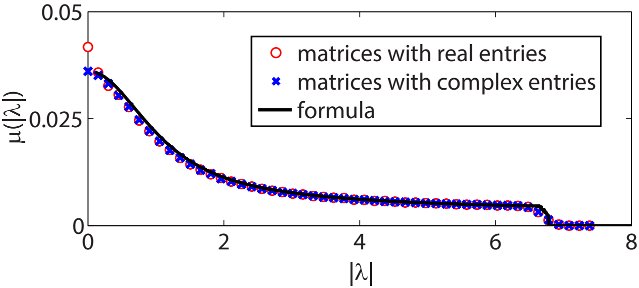
Example

$$N = 1000,\ D = 2,\ \alpha = (0.3, 0.7) \text{ and } g = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$



Example

$$N = 1000,\ D = 3,\ \alpha = (0.25, 0.30, 0.45) \text{ and } g = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$



Motivation

The linearized mean-field dynamics of a neural network model satisfy the following system of differential equations

$$x_i'(t) = -x_i(t) + \sum_j J_{ij} x_j(t)$$

If the real part of the eigenvalues of J are less than 1 then $x_i \to 0$.

If J has eigenvalues greater than one the system exhibits chaotic behavior.

Main Theorem

Our theorem should be compared with the case that each entry of the random matrix has variance g_{cd}^2 with probability $\alpha_c\alpha_d$.

Then the limiting density is the uniform law on the disk of radius

$$\sqrt{\sum_{c,d} g_{cd}^2 \alpha_c \alpha_d}$$

Which can be bigger than or less than $\sqrt{\rho(G)}$ depending on the choice of parameters.

Hermitization

The spectral measure can be recovered from the log potential.

$$2\pi\mu_N(z) = \Delta \int \log|z - s| d\mu_N(s)$$

This can be written in terms of the resolvent (when justified)

$$\partial_{\bar{z}} \int \frac{1}{z - s} d\mu_N(s)$$

Hermitization

The log potential allows one to connect eigenvalues of a non-Hermitian matrix to those of a family of Hermitian matrices.

$$\int \log|z-s|d\mu_N(s) = \frac{1}{N}\log(|\det(X_N - z)|) = \int_0^\infty \log(x)\nu_{N,z}(x)$$

Where $\nu_{N,z}(x)$ is the empirical spectral measure of

$$\begin{pmatrix} 0 & X_N - z \\ X_N^* - \overline{z} & 0 \end{pmatrix}$$

Hermitization

The quantity

$$\frac{1}{2\pi}\Delta\int_0^\infty \log(x)\nu_{N,z}(x)$$

is known as the Brown measure.

The limit of the Brown measures might not be limit of the empirical spectral measures because the logarithm is unbounded.

This issue is resolved by showing $\nu_{N,z}(x)$ is uniformly integrable with respect to the logarithm.

Hermitization

The measure $\nu_{N,z}(x)$ can be recovered from the resolvent

$$R(z) = \begin{pmatrix} \eta & X_N - z \\ X_N^* - \overline{z} & \eta \end{pmatrix}^{-1}$$

Then the limiting brown measure μ can be computed from $\Delta \int \log |z - s| d\nu_{N,z}(s)$.

Then if $\log(x)$ is uniformly integrable, then μ_N converges weakly to μ.

Except in special cases (the circular law, for example) it is hard to compute the μ by this strategy.

Hermitization

Instead treating the resolvent as a 2×2 block matrix

$$\begin{pmatrix} \eta & X_N - z \\ X_N^* - \overline{z} & \eta \end{pmatrix}^{-1}_{21} = (X_N - z)^*(\eta^2 - (X_N - z)(X_N - z)^*)^{-1}$$

Then

$$\lim_{\eta = it \to 0} (X_N - z)^*(\eta^2 - (X_N - z)(X_N - z)^*)^{-1} = (X_N - z)^{-1}.$$

Hermitization

- This gives another way to compute the density of the limiting measure

$$\mu_N = -\frac{1}{\pi}\partial_{\overline{z}} \lim_{\eta = it \to 0} \mathrm{tr}_N R^{21}$$

Bordenave and Chafaï (12) and in physics literature

Structured random matrices

- We instead consider the $2D \times 2D$ matrix formed by taking the trace over each block.
- The diagonal entries within each block are exchangeable, so suffices to consider the $(1,1)$ entry of each block.
- Schur's complement

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1}_{11} = (A - BD^{-1}C)^{-1}$$

Resolvent

By Schur's complement

$$R_{N;11} = \left(H_{11} - q \otimes I_d - H_{1\cdot}^{(1)} R_N^{(1)} H_{\cdot 1}^{(1)}\right)^{-1}$$

Where $q = \begin{pmatrix} \eta & z \\ \bar{z} & \eta \end{pmatrix}$,

$$
\begin{pmatrix}
a_1 & & & c_1 & & \\
& \ddots & & & \ddots & \\
& & a_D & & & c_D \\
b_1 & & & \widehat{a}_1 & & \\
& \ddots & & & \ddots & \\
& & b_D & & & \widehat{a}_D
\end{pmatrix}
$$

$$
\approx \begin{pmatrix}
(\widehat{G}\widehat{a})_1 + \eta & & & z & & \\
& \ddots & & & \ddots & \\
& & (\widehat{G}\widehat{a})_D + \eta & & & z \\
\overline{z} & & & (Ga)_1 + \eta & & \\
& \ddots & & & \ddots & \\
& & \overline{z} & & & (Ga)_D + \eta
\end{pmatrix}^{-1}
$$

Matrix valued semi-circular elements

- The matrix-valued Stieltjes transform of A satisfies

$$M_N(Z) = -(\Sigma(M_N(Z)) + Z)^{-1}$$

where Σ is a a linear operator on $d \times d$ matrices such that

$$\Sigma(B) = (I_d \otimes \phi)(A(B \otimes I)A)$$

- This equation has a unique solution that is a matrix valued Stietljes transform (J. Helton, R. Far, R. Speicher ('07))

Analysis of the fixed point equations

$$a_c = \frac{[Ga]_c + \eta}{|z|^2 - ([\widehat{G}\widehat{a}]_c + \eta)([Ga]_c + \eta)}$$

$$b_c = \frac{\overline{z}}{|z|^2 - ([\widehat{G}\widehat{a}]_c + \eta)([Ga]_c + \eta)} \to 1 - \widehat{a}_c[Ga]_c$$

as $\eta \to 0$.

Smallest singular value

- Theorem (Tao, Vu) Let C_N be an iid random matrix and F_N a deterministic matrix, for any $B > 0$, there exists $A > 0$

$$\mathbb{P}\left(\sigma_N(C_N + F_N) \leq N^{-A}\right) = O(N^{-B}).$$

- Where $\sigma_N(C_N + F_N)$ is the least singular value of $C_N + F_N$.

Smallest singular value

The least singular value of the block matrix

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

is controlled by the least singular values

$$(D + CA^{-1}B) \text{ and } (A + BD^{-1}C)$$

but we assume that A, B, C, D are independent so the least singular values are only polynomially small by the Tao-Vu theorem.

Thank you

Thank you