Random matrices and applications

Anna Lytova

University of Opole, Poland

Random matrices

$$M_n = \left(M_{jk}\right)_{j,k=1}^{n} = \begin{pmatrix} M_{11} & \dots & M_{1n} \\ \vdots & & \vdots \\ M_{n1} & \dots & M_{nn} \end{pmatrix}, \quad M_{ij} \text{ are random variables}$$

Random matrices

$$M_n = \left(M_{jk}\right)_{j,k=1}^{n} = \begin{pmatrix} M_{11} & \dots & M_{1n} \\ \vdots & & \vdots \\ M_{n1} & \dots & M_{nn} \end{pmatrix}, \quad M_{ij} \text{ are random variables}$$

RMT = PROBABILITY THEORY

+ ALGEBRA

+ ANALISYS (real, complex, functional)

+ Free probability theory

+ Combinatorics, graph theory

+ Convex geometry, Topology, Supersymmetry ...

$$M_n = \left(M_{jk}\right)_{j,k=1}^{n} = \begin{pmatrix} M_{11} & \dots & M_{1n} \\ \vdots & & \vdots \\ M_{n1} & \dots & M_{nn} \end{pmatrix}, \quad M_{ij} \text{ are random variables}$$

We are interested in

- eigenvalues $(\lambda_j)_j$ and eigenvectors $(\mathbf{v}_j)_j$, $M_n\mathbf{v}_j = \lambda_j\mathbf{v}_j$,
- singular values $(s_j)_j$ and singular vectors, $M_n^*M_n\mathbf{u}_j = s_j^2\mathbf{u}_j$,
- various spectral statistics $\varphi(\lambda_1, \dots, \lambda_n, \mathbf{v}_1, \dots, \mathbf{v}_n)$.

$$M_n = \left(M_{jk}\right)_{j,k=1}^{n} = \begin{pmatrix} M_{11} & \dots & M_{1n} \\ \vdots & & \vdots \\ M_{n1} & \dots & M_{nn} \end{pmatrix}, \quad M_{ij} \text{ are random variables}$$

We are interested in

- eigenvalues $(\lambda_j)_j$ and eigenvectors $(\mathbf{v}_j)_j$, $M_n \mathbf{v}_j = \lambda_j \mathbf{v}_j$,
- singular values $(s_j)_j$ and singular vectors, $M_n^* M_n \mathbf{u}_j = s_j^2 \mathbf{u}_j$,
- various spectral statistics $\varphi(\lambda_1, \dots, \lambda_n, \mathbf{v}_1, \dots, \mathbf{v}_n)$.

From now on let $M_n^* = M_n$, and $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$.

$$M_n = \left(M_{jk}\right)_{j,k=1}^{n} = \begin{pmatrix} M_{11} & \dots & M_{1n} \\ \vdots & & \vdots \\ M_{n1} & \dots & M_{nn} \end{pmatrix}, \quad M_{ij} \text{ are random variables}$$

We are interested in

- eigenvalues $(\lambda_j)_j$ and eigenvectors $(\mathbf{v}_j)_j$, $M_n \mathbf{v}_j = \lambda_j \mathbf{v}_j$,
- singular values $(s_j)_j$ and singular vectors, $M_n^* M_n \mathbf{u}_j = s_j^2 \mathbf{u}_j$,
- various spectral statistics $\varphi(\lambda_1, \dots, \lambda_n, \mathbf{v}_1, \dots, \mathbf{v}_n)$.

From now on let $M_n^* = M_n$, and $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$.

Main regimes:

- asymptotic, $n \to \infty$,

  global,
  local,

- non-asymptotic, $1 \ll n < \infty$.

Basic questions:

- Convergence of the *counting measures of eigenvalues*

$$N_n(\Delta) := |\{j : \lambda_j \in \Delta\}|/n = \frac{1}{n}\sum_{j=1}^{n} I_{\{\lambda_j \in \Delta\}}, \quad \forall \Delta \in \mathbb{R}.$$

Basic questions:

- Convergence of the *counting measures of eigenvalues*

$$N_n(\Delta) := |\{j : \lambda_j \in \Delta\}|/n = \frac{1}{n}\sum_{j=1}^{n} I_{\{\lambda_j \in \Delta\}}, \quad \forall \Delta \in \mathbb{R}.$$

$\exists N : \; N_n \xrightarrow[n\to\infty]{w} N$ a.s.?

Basic questions:

- Convergence of the *counting measures of eigenvalues*

$$N_n(\Delta) := |\{j : \lambda_j \in \Delta\}|/n = \tfrac{1}{n}\sum_{j=1}^{n} I_{\{\lambda_j \in \Delta\}}, \quad \forall \Delta \in \mathbb{R}.$$

$\exists N : \ N_n \xrightarrow[n\to\infty]{w} N$ a.s.?

i.e. for any bounded continuous function φ with probability 1

$$\int_{\mathbb{R}} \varphi(\lambda)dN_n(\lambda) = \frac{1}{n}\sum_{k=1}^{n}\varphi(\lambda_k) \xrightarrow[n\to\infty]{} \int_{\mathbb{R}} \varphi(\lambda)dN(\lambda).$$

Basic questions:

- Convergence of the *counting measures of eigenvalues*

$$N_n(\Delta) := |\{j : \lambda_j \in \Delta\}|/n = \tfrac{1}{n}\sum_{j=1}^{n} I_{\{\lambda_j\in\Delta\}}, \quad \forall\Delta \in \mathbb{R}.$$

  $\exists N : \ N_n \xrightarrow[n\to\infty]{w} N$ a.s.?

  i.e. for any bounded continuous function φ with probability 1

$$\int_{\mathbb{R}} \varphi(\lambda)dN_n(\lambda) = \frac{1}{n}\sum_{k=1}^{n}\varphi(\lambda_k) \xrightarrow[n\to\infty]{} \int_{\mathbb{R}} \varphi(\lambda)dN(\lambda).$$

- fluctuations and CLT for linear eigenvalue statistics:

  $\nu_n \sum_j(\varphi(\lambda_j) - \mathbf{E}\varphi(\lambda_j)) \xrightarrow[n\to\infty]{d} \xi \sim \mathcal{N}(0, \mathbf{Var}[\varphi])$?

Basic questions:

- Convergence of the *counting measures of eigenvalues*

$$N_n(\Delta) := |\{j : \lambda_j \in \Delta\}|/n = \tfrac{1}{n}\sum_{j=1}^n I_{\{\lambda_j \in \Delta\}}, \quad \forall \Delta \in \mathbb{R}.$$

  $\exists N : \; N_n \xrightarrow[n\to\infty]{w} N$ a.s.?

  i.e. for any bounded continuous function φ with probability 1

$$\int_{\mathbb{R}} \varphi(\lambda) dN_n(\lambda) = \frac{1}{n}\sum_{k=1}^{n} \varphi(\lambda_k) \xrightarrow[n\to\infty]{} \int_{\mathbb{R}} \varphi(\lambda) dN(\lambda).$$

- fluctuations and CLT for linear eigenvalue statistics:
  $\nu_n \sum_j (\varphi(\lambda_j) - \mathbf{E}\varphi(\lambda_j)) \xrightarrow[n\to\infty]{d} \xi \sim \mathcal{N}(0, \mathbf{Var}[\varphi])$?
- local laws, spacing distribution, bulk statistics, edge statistics, correlation functions,

Basic questions:

- Convergence of the *counting measures of eigenvalues*

$$N_n(\Delta) := |\{j : \lambda_j \in \Delta\}|/n = \tfrac{1}{n}\textstyle\sum_{j=1}^n I_{\{\lambda_j \in \Delta\}}, \quad \forall \Delta \in \mathbb{R}.$$

  $\exists N : \ N_n \xrightarrow[n\to\infty]{w} N$ a.s.?

  i.e. for any bounded continuous function φ with probability 1

$$\int_{\mathbb{R}} \varphi(\lambda) dN_n(\lambda) = \frac{1}{n}\sum_{k=1}^{n} \varphi(\lambda_k) \xrightarrow[n\to\infty]{} \int_{\mathbb{R}} \varphi(\lambda) dN(\lambda).$$

- fluctuations and CLT for linear eigenvalue statistics:
  $\nu_n \sum_j (\varphi(\lambda_j) - \mathbf{E}\varphi(\lambda_j)) \xrightarrow[n\to\infty]{d} \xi \sim \mathcal{N}(0, \mathbf{Var}[\varphi])$?
- local laws, spacing distribution, bulk statistics, edge statistics, correlation functions,
- invertibility of random matrices,

Basic questions:

- Convergence of the *counting measures of eigenvalues*

  $$N_n(\Delta) := |\{j : \lambda_j \in \Delta\}|/n = \tfrac{1}{n}\sum_{j=1}^n I_{\{\lambda_j \in \Delta\}}, \quad \forall \Delta \in \mathbb{R}.$$

  $\exists N : \; N_n \xrightarrow[n\to\infty]{w} N$ a.s.?

  i.e. for any bounded continuous function φ with probability 1

  $$\int_{\mathbb{R}} \varphi(\lambda) dN_n(\lambda) = \frac{1}{n}\sum_{k=1}^{n} \varphi(\lambda_k) \xrightarrow[n\to\infty]{} \int_{\mathbb{R}} \varphi(\lambda) dN(\lambda).$$

- fluctuations and CLT for linear eigenvalue statistics:

  $\nu_n \sum_j (\varphi(\lambda_j) - \mathbf{E}\varphi(\lambda_j)) \xrightarrow[n\to\infty]{d} \xi \sim \mathcal{N}(0, \mathbf{Var}[\varphi])$?
- local laws, spacing distribution, bulk statistics, edge statistics, correlation functions,
- invertibility of random matrices,
- quantitative estimates for the smallest and largest singular values,
- delocalization of eigenvectors...

Applications of random matrices

Physics: nuclear physics, quantum chaology, quantum field theory, condensed matter, statistical physics, wave propagation...

Statistics: multivariate statistics, principal component analysis, data compression, image processing...

Mathematics: number theory, combinatorics, integrable systems, graph theory...

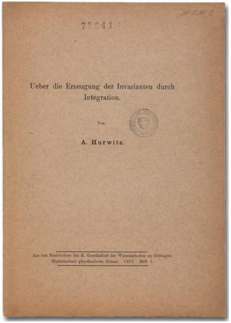
Information Theory: signal processing, wireless communications, quantum information theory, telecommunications, neural networks...

Biology: sequences matching, RNA folding, gene expressions network...

Economics and Finances: quantitative finances, time series analysis...

The origins of RMT

P. Diakonis and P. J. Forrester*: "**Hurwitz***'s paper "Über die Erzeugung der Invarianten durch Integration." [Gött. Nachrichten (***1897***), 71–90] should be regarded as the origin of random matrix theory in mathematics.*



Here Hurwitz introduced and developed the notion of an invariant measure for the matrix groups $SO(N)$ *and* $U(N)$*... Hurwitz's ideas and methods show themselves in the subsequent work of Weyl, Dyson and others on foundational studies in random matrix theory..."*

*Diaconis, Persi, and Peter J. Forrester. "Hurwitz and the origins of random matrix theory in mathematics." Random Matrices: Theory and Applications 6.01 (2017): 1730001.

The origins of RMT: multivariate statistics

Wishart, *The generalized product moment distribution in samples from a normal multivariate population*, Biometrika 20A (**1928**), 32–43.

John Wishart

THE GENERALISED PRODUCT MOMENT DISTRIBUTION IN SAMPLES FROM A NORMAL MULTIVARIATE POPULATION.

By JOHN WISHART, M.A., B.Sc. Statistical Department, Rothamsted Experimental Station.

1. *Introduction.*

For some years prior to 1915, various writers struggled with the problems that arise when samples are taken from uni-variate and bi-variate populations, assumed in most cases for simplicity to be normal. Thus "Student," in 1908*, by considering the first four moments, was led by K. Pearson's methods to infer the distribution of standard deviations, in samples from a normal population. His results, for comparison with others to be deduced later, will be stated in the form

$$dp = \frac{1}{\Gamma\left(\frac{N-1}{2}\right)} A^{\frac{N-1}{2}} \,.\, e^{-Aa} \,.\, a^{\frac{N-3}{2}} \, da \qquad \text{.....................(1)},$$

where N is the size of the sample, and

$$A = \frac{N}{2\sigma^2}, \quad a = s^2,$$

σ being the standard deviation of the sampled population, and s that estimated from the sample. Thus, if $x_1, x_2, \ldots x_N$ are the sample values,

$$N\bar{x} = \sum_1^N (x),$$

and

$$Ns^2 = \sum_1^N (x - \bar{x})^2.$$

When bi-variate populations were considered, other problems arose, such as the distribution of the correlation coefficient and of the regression coefficient in samples. These problems, taken by themselves, were found to be difficult, and only approximative results had been reached, when, in 1915, R. A. Fisher† gave a formula for the simultaneous distribution of the three quadratic statistical derivatives, namely the two variances (squared standard deviations) and the product moment coefficient. Thus, let $x_1, x_2, \ldots x_N$ represent the sample values of the x-variate, and $y_1, y_2, \ldots y$ the corresponding values for the y-variate, let σ_1 and σ_2 be the standard deviations of the sampled population and ρ the correlation between x and y. We then calculate the following statistical derivatives from the sample:

$$N\bar{x} = \sum_1^N (x) \qquad N\bar{y} = \sum_1^N (y)$$

$$N{s_1}^2 = \sum_1^N (x - \bar{x})^2 \qquad N{s_2}^2 = \sum_1^N (y - \bar{y})^2$$

$$Nrs_1s_2 = \sum_1^N (x - \bar{x})(y - \bar{y}).$$

* *Biometrika*, Vol. vi. 1908, pp. 4—6.
† *Biometrika*, Vol. x. 1915, p. 510.

One of the basic problems in the multivariate statistics is by sampling from a high-dimensional distribution to estimate its covariance matrix Σ.

Let $\mathbf{X} \in \mathbb{R}^n$, $\mathbf{E}\mathbf{X} = 0$, $\Sigma = \mathbf{E}\,\mathbf{X}\mathbf{X}^T$, and let

$$\mathbf{X}_1 = \begin{pmatrix} X_{11} \\ \vdots \\ X_{n1} \end{pmatrix}, \quad \ldots, \mathbf{X}_m = \begin{pmatrix} X_{1m} \\ \vdots \\ X_{nm} \end{pmatrix} \quad \text{be i.i.d. copies of } \mathbf{X}.$$

One of the basic problems in the multivariate statistics is by sampling from a high-dimensional distribution to estimate its covariance matrix Σ.

Let $\mathbf{X} \in \mathbb{R}^n$, $\mathbf{E}\mathbf{X} = 0$, $\Sigma = \mathbf{E}\,\mathbf{X}\mathbf{X}^T$, and let

$$\mathbf{X}_1 = \begin{pmatrix} X_{11} \\ \vdots \\ X_{n1} \end{pmatrix}, \quad \ldots, \mathbf{X}_m = \begin{pmatrix} X_{1m} \\ \vdots \\ X_{nm} \end{pmatrix} \quad \text{be i.i.d. copies of } \mathbf{X}.$$

Let $B_n = \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \ldots & \mathbf{X}_m \end{bmatrix}$. Then the Sample Covariance Matrix

$$S_n = m^{-1} B_n B_n^T = m^{-1} \sum_{\alpha=1}^{m} \mathbf{X}_\alpha \mathbf{X}_\alpha^T$$

is an unbiased estimator of Σ.

One of the basic problems in the multivariate statistics is by sampling from a high-dimensional distribution to estimate its covariance matrix Σ.

Let $\mathbf{X} \in \mathbb{R}^n$, $\mathbf{E}\mathbf{X} = 0$, $\Sigma = \mathbf{E}\,\mathbf{X}\mathbf{X}^T$, and let

$$\mathbf{X}_1 = \begin{pmatrix} X_{11} \\ \vdots \\ X_{n1} \end{pmatrix}, \quad \ldots, \mathbf{X}_m = \begin{pmatrix} X_{1m} \\ \vdots \\ X_{nm} \end{pmatrix} \quad \text{be i.i.d. copies of } \mathbf{X}.$$

Let $B_n = \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \ldots & \mathbf{X}_m \end{bmatrix}$. Then the Sample Covariance Matrix

$$S_n = m^{-1} B_n B_n^T = m^{-1} \sum_{\alpha=1}^{m} \mathbf{X}_\alpha \mathbf{X}_\alpha^T$$

is an unbiased estimator of Σ.

The Wishart matrix S_n^W corresponds to $\mathbf{X}_1, \ldots, \mathbf{X}_m \sim \mathcal{N}(0, I_n)$.

$$\text{jpdf}(\lambda_1^W, \ldots, \lambda_n^W) = c e^{-\sum_{j=1}^n \lambda_j/2} \prod_{j=1}^{n} \lambda_j^{m-n-1/2} \prod_{j<k} |\lambda_j - \lambda_k|$$

The origins of RMT: nuclear physics

Wigner: "*...it is tantalizing not to know what the probability of a certain spacing of the energy levels is.*"

Eugene Paul Wigner

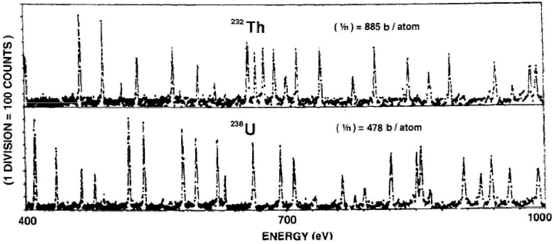


Figure 1.1. Slow neutron resonance cross-sections on thorium 232 and uranium 238 nuclei. Reprinted with permission from The American Physical Society, Rahn et al., Neutron resonance spectroscopy, X, *Phys. Rev. C* 6, 1854–1869 (1972).

This figure was copied from Mehta's book [1]

In the 50s **Wigner** proposed to construct a statistical theory of energy levels.

Dyson: "*We picture a complex nucleus as a "black box" in which a large number of particles are interacting according to unknown laws... The statistical theory will not predict the detailed level sequence of any one nucleus, but it will describe the general appearance of the level structure...* "

Example* (The spacing distribution and Wigner surmise)

Let $n = 2,\ M = \begin{pmatrix} x_1 & x_3 \\ x_3 & x_2 \end{pmatrix}$, where $x_1, x_2 \sim \mathcal{N}(0, 1),\ x_3 \sim \mathcal{N}(0, 1/2)$.

Example* (The spacing distribution and Wigner surmise)

Let $n = 2$, $M = \begin{pmatrix} x_1 & x_3 \\ x_3 & x_2 \end{pmatrix}$, where $x_1, x_2 \sim \mathcal{N}(0,1)$, $x_3 \sim \mathcal{N}(0,1/2)$.

Eigenvalues $\lambda_{1,2} = \left(x_1 + x_2 \pm \sqrt{(x_1 - x_2)^2 + 4x_3^2}\right)/2$, $\lambda_1 \geq \lambda_2$.

Example* (The spacing distribution and Wigner surmise)

Let $n = 2$, $M = \begin{pmatrix} x_1 & x_3 \\ x_3 & x_2 \end{pmatrix}$, where $x_1, x_2 \sim \mathcal{N}(0, 1)$, $x_3 \sim \mathcal{N}(0, 1/2)$.

Eigenvalues $\lambda_{1,2} = \left(x_1 + x_2 \pm \sqrt{(x_1 - x_2)^2 + 4x_3^2}\right)/2$, $\lambda_1 \geq \lambda_2$.

Spacing $S := \lambda_1 - \lambda_2 = \sqrt{(x_1 - x_2)^2 + 4x_3^2}$,

Example* (The spacing distribution and Wigner surmise)

Let $n = 2,\ M = \begin{pmatrix} x_1 & x_3 \\ x_3 & x_2 \end{pmatrix}$, where $x_1,\, x_2 \sim \mathcal{N}(0,1),\ x_3 \sim \mathcal{N}(0,1/2)$.

Eigenvalues $\lambda_{1,2} = \left(x_1 + x_2 \pm \sqrt{(x_1 - x_2)^2 + 4x_3^2}\right)/2,\ \lambda_1 \geq \lambda_2$.

Spacing $S := \lambda_1 - \lambda_2 = \sqrt{(x_1 - x_2)^2 + 4x_3^2},\ p(s) := \frac{d}{ds}\mathbb{P}(S \leq s)$ – ?

Example* (The spacing distribution and Wigner surmise)

Let $n = 2$, $M = \begin{pmatrix} x_1 & x_3 \\ x_3 & x_2 \end{pmatrix}$, where $x_1, x_2 \sim \mathcal{N}(0,1)$, $x_3 \sim \mathcal{N}(0,1/2)$.

Eigenvalues $\lambda_{1,2} = \left(x_1 + x_2 \pm \sqrt{(x_1 - x_2)^2 + 4x_3^2}\right)/2$, $\lambda_1 \geq \lambda_2$.

Spacing $S := \lambda_1 - \lambda_2 = \sqrt{(x_1 - x_2)^2 + 4x_3^2}$, $p(s) := \frac{d}{ds}\mathbb{P}(S \leq s)$ – ?

$$p(s) = \int\int\int dx_1 dx_2 dx_3 \frac{e^{-x_1^2/2}}{\sqrt{2\pi}} \frac{e^{-x_2^2/2}}{\sqrt{2\pi}} \frac{e^{-x_3^2}}{\sqrt{\pi}} \delta\left(s - \sqrt{(x_1 - x_2)^2 + 4x_3^2}\right), \quad s > 0$$

Example* (The spacing distribution and Wigner surmise)

Let $n = 2,\ M = \begin{pmatrix} x_1 & x_3 \\ x_3 & x_2 \end{pmatrix}$, where $x_1, x_2 \sim \mathcal{N}(0,1),\ x_3 \sim \mathcal{N}(0,1/2)$.

Eigenvalues $\lambda_{1,2} = \left(x_1 + x_2 \pm \sqrt{(x_1 - x_2)^2 + 4x_3^2}\right)/2,\ \lambda_1 \geq \lambda_2$.

Spacing $S := \lambda_1 - \lambda_2 = \sqrt{(x_1 - x_2)^2 + 4x_3^2}$, $p(s) := \frac{d}{ds}\mathbb{P}(S \leq s)$ – ?

$$p(s) = \int\int\int dx_1 dx_2 dx_3 \frac{e^{-x_1^2/2}}{\sqrt{2\pi}} \frac{e^{-x_2^2/2}}{\sqrt{2\pi}} \frac{e^{-x_3^2}}{\sqrt{\pi}} \delta\left(s - \sqrt{(x_1 - x_2)^2 + 4x_3^2}\right),\ \ s > 0$$

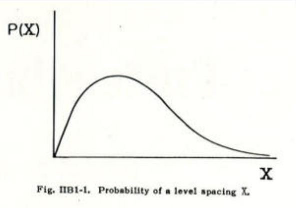
$$x_1 - x_2 = r\cos\varphi,\ 2x_3 = r\sin\varphi,\ x_1 + x_2 = y$$

$$p(s) = \frac{1}{8\pi^{3/2}} \int_0^\infty dr\, r\delta(s - r) \int_0^{2\pi} d\varphi$$

$$\int_{-\infty}^{\infty} dy e^{-\frac{1}{2}\left[\left(\frac{r\cos\varphi + y}{2}\right)^2 + \left(\frac{-r\cos\varphi + y}{2}\right)^2 + \frac{r^2\sin^2\varphi}{2}\right]} = \frac{s}{2}e^{-s^2/4}$$

*Livan, G., Novaes, M. and Vivo, P., 2018. *Introduction to random matrices theory and practice*. Monograph Award, p.63.

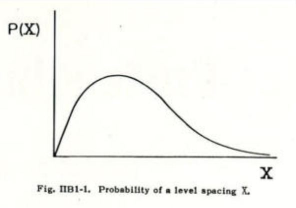
Wigner surmise: $p(s) = \frac{s}{2} e^{-s^2/4}$



Original picture in Wigner's proceedings

Eugene Wigner and Edward Teller
https://djalil.chafai.net/

Wigner surmise: $p(s) = \frac{s}{2} e^{-s^2/4}$

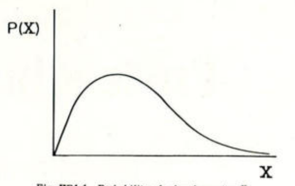


Original picture in Wigner's proceedings

Eugene Wigner and Edward Teller
https://djalil.chafai.net/

1. $\lambda_1, \ldots, \lambda_n$ are eigenvalues of a symmetric random matrix with independent Gaussian entries: $p(s)$ was computed by Gaudin and Mehta (60's)

Wigner surmise: $p(s) = \frac{s}{2}e^{-s^2/4}$



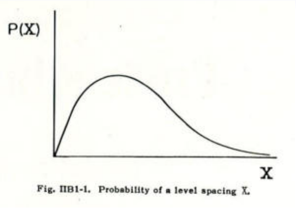
Original picture in Wigner's proceedings

Eugene Wigner and Edward Teller
https://djalil.chafai.net/

1. $\lambda_1, \ldots, \lambda_n$ are eigenvalues of a symmetric random matrix with independent Gaussian entries: $p(s)$ was computed by Gaudin and Mehta (60's)

2. $x_1, \ldots, x_n$ are i.i.d. $U[0,1]$: $p(s) = e^{-s}$

Wigner surmise: $p(s) = \frac{s}{2} e^{-s^2/4}$



Original picture in Wigner's proceedings

Eugene Wigner and Edward Teller
https://djalil.chafai.net/

1. $\lambda_1, \ldots, \lambda_n$ are eigenvalues of a symmetric random matrix with independent Gaussian entries: $p(s)$ was computed by Gaudin and Mehta (60's)

2. $x_1, \ldots, x_n$ are i.i.d. $U[0,1]$: $p(s) = e^{-s}$

3. The Wigner-Dyson-Gaudin-Mehta universality conjecture asserts that the local eigenvalue statistics of large random matrices depends only on the symmetry class of the matrix ensemble.

Wigner surmise: $p(s) = \frac{s}{2} e^{-s^2/4}$

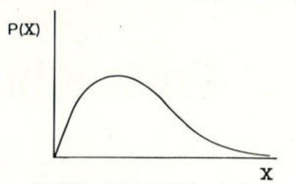


Fig. IIB1-1. Probability of a level spacing X.

Original picture in Wigner's proceedings

Eugene Wigner and Edward Teller
https://djalil.chafai.net/

1. $\lambda_1, \ldots, \lambda_n$ are eigenvalues of a symmetric random matrix with independent Gaussian entries: $p(s)$ was computed by Gaudin and Mehta (60's)

2. $x_1, \ldots, x_n$ are i.i.d. $U[0,1]$: $p(s) = e^{-s}$

3. The Wigner-Dyson-Gaudin-Mehta universality conjecture asserts that the local eigenvalue statistics of large random matrices depends only on the symmetry class of the matrix ensemble.

4. The same spacing distribution have: eigenvalues of large Hermitian random matrices, resonances of various heavy nuclei, zeroes of the Riemann zeta function, bus arrival times, birds perching on an electric wire...

Wigner real symmetric matrices

$$M_n = \frac{1}{\sqrt{n}} W_n = \frac{1}{\sqrt{n}} \begin{pmatrix} W_{11} & \dots & W_{1n} \\ \vdots & & \vdots \\ W_{n1} & \dots & W_{nn} \end{pmatrix}$$

- $W_{jk} = W_{kj} \in \mathbb{R}$,
- $(W_{jk})_{1 \le j < k \le n}$ and $(W_{jj})_{1 \le j \le n}$ are two independent families of i.i.d. zero mean random variables
- $\mathbf{E}\, W_{jk}^2 = 1$, $1 \le j < k \le n$.

Wigner real symmetric matrices

$$M_n = \frac{1}{\sqrt{n}} W_n = \frac{1}{\sqrt{n}} \begin{pmatrix} W_{11} & \dots & W_{1n} \\ \vdots & & \vdots \\ W_{n1} & \dots & W_{nn} \end{pmatrix}$$

- $W_{jk} = W_{kj} \in \mathbb{R}$,
- $(W_{jk})_{1 \le j < k \le n}$ and $(W_{jj})_{1 \le j \le n}$ are two independent families of i.i.d. zero mean random variables
- $\mathbf{E}\, W_{jk}^2 = 1,\ 1 \le j < k \le n.$

In particular, if all entries of M_n are independent Gaussian random variables,

$$W_{jk} \sim N(0, 1 + \delta_{jk}), \quad 1 \le j \le k \le n,$$

then we say that M_n belongs to the Gaussian Orthogonal Ensemble (GOE). In this case

$$\text{jpdf}\big((M_{jk})_{j \le k}\big) = c_n \exp\{-n \operatorname{Tr} M_n^2/4\}$$

Wigner's Semicircle Law (1955)

Theorem. *Let M_n belongs to GOE, and N_n be the counting measure of its eigenvalues,*

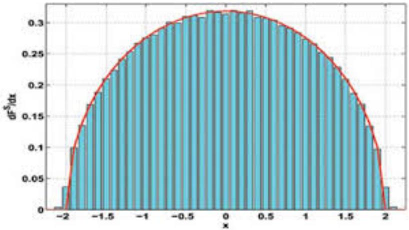
$$N_n(\Delta) = \frac{1}{n}|\{j : \lambda_j \in \Delta\}| = \frac{1}{n}\sum_j I_{\{\lambda_j \in \Delta\}}, \quad \forall \Delta \subset \mathbb{R}.$$

Then almost surely

$$N_n \xrightarrow[n\to\infty]{w} N_{sc},$$

where $N_{sc}(d\lambda) = \rho_{sc}(\lambda)d\lambda$,

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



Normalised empirical eigenvalue distribution for a 100 × 100 GUE matrix
(Image by Alan Edelman)

In other words, for any bounded continuous function φ, with probability 1,

$$\lim_{n\to\infty}\int_{\mathbb{R}} \varphi(\lambda)dN_n(\lambda) = \int_{-2}^{2} \varphi(\lambda)\rho_{sc}(\lambda)d\lambda$$

Note that

$$L_{M_n}[\varphi] := \int \varphi(\lambda) N_n(d\lambda) = \frac{1}{n}\sum_j \varphi(\lambda_j) = \frac{1}{n}\mathrm{Tr}\,\varphi(M_n).$$

$L_{M_n}[\varphi]$ is a *linear eigenvalue statistic* corresponding to a test function φ.

The Wigner's theorem on convergence to the semicircle law,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \varphi(\lambda_j) = \int_{-2}^{2} \varphi(\lambda)\rho_{sc}(\lambda) d\lambda \qquad (1)$$

can be considered as an analog of the *Law of Large Numbers*.

Note that

$$L_{M_n}[\varphi] := \int \varphi(\lambda) N_n(d\lambda) = \frac{1}{n} \sum_j \varphi(\lambda_j) = \frac{1}{n} \mathrm{Tr}\, \varphi(M_n).$$

$L_{M_n}[\varphi]$ is a *linear eigenvalue statistic* corresponding to a test function φ.

The Wigner's theorem on convergence to the semicircle law,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \varphi(\lambda_j) = \int_{-2}^{2} \varphi(\lambda) \rho_{sc}(\lambda) d\lambda \tag{1}$$

can be considered as an analog of the *Law of Large Numbers*.

Moment method: To prove (1) it is enough to consider $\varphi(\lambda) = \lambda^k$, $k \in \mathbb{N}$.

Note that

$$L_{M_n}[\varphi] := \int \varphi(\lambda) N_n(d\lambda) = \frac{1}{n}\sum_j \varphi(\lambda_j) = \frac{1}{n}\mathrm{Tr}\,\varphi(M_n).$$

$L_{M_n}[\varphi]$ is a *linear eigenvalue statistic* corresponding to a test function φ.

The Wigner's theorem on convergence to the semicircle law,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \varphi(\lambda_j) = \int_{-2}^{2} \varphi(\lambda)\rho_{sc}(\lambda)d\lambda \qquad (1)$$

can be considered as an analog of the *Law of Large Numbers*.

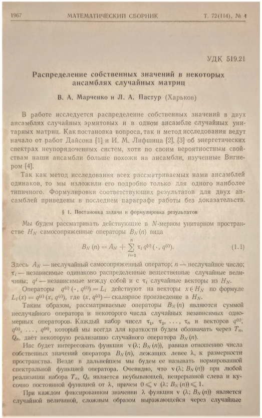
Moment method: To prove (1) it is enough to consider $\varphi(\lambda) = \lambda^k$, $k \in \mathbb{N}$.

Method of Stielties transform: To prove (1) it is enough to consider

$$\varphi(\lambda) = (\lambda - z)^{-1}, \quad z \in \mathbb{C} \setminus \mathbb{R}.$$

Stielties transform and convergence to the Marchenko-Pastur law

Marchenko, V., Pastur, L. (1967). *The eigenvalue distribution in some ensembles of random matrices*. Math. USSR Sbornik, **1**, 457–483.


1967 МАТЕМАТИЧЕСКИЙ СБОРНИК Т. 72(114), № 4

УДК 519.21

Распределение собственных значений в некоторых ансамблях случайных матриц

В. А. Марченко и Л. А. Пастур (Харьков)

В работе исследуется распределение собственных значений в двух ансамблях случайных эрмитовых и в одном ансамбле случайных унитарных матриц. Как постановка вопроса, так и метод исследования ведут начало от работ Дайсона [1] и И. М. Лифшица [2], [3] об энергетических спектрах неупорядоченных систем, хотя по своим вероятностным свойствам наши ансамбли больше похожи на ансамбли, изученные Вигнером [4].

Так как метод исследования всех рассматриваемых нами ансамблей одинаков, то мы изложили его подробно только для одного наиболее типичного. Формулировки соответствующих результатов для двух ансамблей приведены в последнем параграфе работы без доказательств.

§ 1. Постановка задачи и формулировка результатов

Мы будем рассматривать действующие в N-мерном унитарном пространстве H_N самосопряженные операторы $B_N(n)$ вида

$$B_N(n) = A_N + \sum_{i=1}^{n} \tau_i q^{(i)}(\cdot, q^{(i)}). \tag{1.1}$$

Здесь A_N — неслучайный самосопряженный оператор; n — неслучайное число; τ_i — независимые одинаково распределенные вещественные случайные величины; q^i — независимые между собой и с τ_i случайные векторы из H_N.

Операторы $q^{(i)}(\cdot, q^{(i)}) = L_i$ действуют на векторы $x \in H_N$ по формуле $L_i(x) = q^{(i)}(x, q^{(i)})$, где $(x, q^{(i)})$ — скалярное произведение в H_N.

Таким образом, рассматриваемые операторы $B_N(n)$ являются суммой неслучайного оператора и некоторого числа случайных независимых одномерных операторов. Каждый набор чисел $\tau_1, \tau_2, \ldots, \tau_n$ и векторов $q^{(1)}, q^{(2)}, \ldots, q^{(n)}$, который мы всегда для краткости будем обозначать через T_n, Q_n, дает некоторую реализацию случайного оператора $B_N(n)$.

Нас будет интересовать функция $\nu(\lambda; B_N(n))$, равная отношению числа собственных значений оператора $B_N(n)$, лежащих левее λ, к размерности пространства. Везде в дальнейшем мы будем ее называть нормированной спектральной функцией оператора. Очевидно, что $\nu(\lambda; B_N(n))$ при любой реализации набора T_n, Q_n является неубывающей, непрерывной слева и кусочно постоянной функцией от λ, причем $0 \leqslant \nu(\lambda; B_N(n)) \leqslant 1$.

При каждом фиксированном значении λ функция $\nu(\lambda; B_N(n))$ является случайной величиной, сложным образом выражающейся через случайные

Vladimir Marchenko and Leonid Pastur

Stieltjes transform of a non-negative finite measure m:

$$s(z) = \int_{\mathbb{R}} \frac{m(d\lambda)}{\lambda - z}, \quad \Im z \neq 0$$

Stieltjes transform of a non-negative finite measure m:

$$s(z) = \int_{\mathbb{R}} \frac{m(d\lambda)}{\lambda - z}, \quad \Im z \neq 0$$

- the *Stieltjes - Perron inversion formula*:

$$m(\Delta) = \lim_{\varepsilon \to 0^+} \frac{1}{\pi} \int_{\Delta} \Im s(\lambda + i\varepsilon) d\lambda;$$

Stieltjes transform of a non-negative finite measure m:

$$s(z) = \int_{\mathbb{R}} \frac{m(d\lambda)}{\lambda - z}, \quad \Im z \neq 0$$

- the *Stieltjes - Perron inversion formula*:

$$m(\Delta) = \lim_{\varepsilon \to 0^+} \frac{1}{\pi} \int_{\Delta} \Im s(\lambda + i\varepsilon) d\lambda;$$

- There is a one-to-one correspondence between finite non-negative measures and their Stieltjes transforms. This correspondence is continuous if we use the uniform convergence of analytic functions on compact subsets of $\mathbb{C} \setminus \mathbb{R}$ for Stieltjes transforms and the vague convergence of measures.

Stieltjes transform of a non-negative finite measure m:

$$s(z) = \int_{\mathbb{R}} \frac{m(d\lambda)}{\lambda - z}, \quad \Im z \neq 0$$

- the *Stieltjes - Perron inversion formula*:

$$m(\Delta) = \lim_{\varepsilon \to 0^+} \frac{1}{\pi} \int_{\Delta} \Im s(\lambda + i\varepsilon) d\lambda;$$

- There is a one-to-one correspondence between finite non-negative measures and their Stieltjes transforms. This correspondence is continuous if we use the uniform convergence of analytic functions on compact subsets of $\mathbb{C} \setminus \mathbb{R}$ for Stieltjes transforms and the vague convergence of measures.
- For the counting measure of eigenvalues we have

$$s_n(z) := \int_{\mathbb{R}} \frac{N_n(d\lambda)}{\lambda - z} = n^{-1} \mathrm{Tr}(M_n - z)^{-1},$$

$$s_n(z) \to s(z) \Leftrightarrow N_n \to N = \int_{\mathbb{R}} \frac{N(d\lambda)}{\lambda - z}$$

Convergence of the empirical eigenvalue distributions of the sample covariance matrices

Let

$$\mathbf{X}_1 = \begin{pmatrix} X_{11} \\ \vdots \\ X_{n1} \end{pmatrix}, \quad \ldots, \mathbf{X}_m = \begin{pmatrix} X_{1m} \\ \vdots \\ X_{nm} \end{pmatrix} \quad \text{be i.i.d. random vectors.}$$

The sample covariance matrix:

$$S_n = B_n B_n^T = \sum_{\alpha=1}^{m} \mathbf{X}_\alpha \mathbf{X}_\alpha^T, \quad B_n = \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \ldots & \mathbf{X}_m \end{bmatrix}.$$

We suppose that $m = m(n)$ and that

$$m \to \infty, \, m/n \to c \geq 1 \quad \text{as} \quad n \to \infty$$

Marchenko-Pastur distribution

Theorem (MP'67). *Let* $M_n = \sum_{\alpha=1}^{m} \mathbf{X}_\alpha \mathbf{X}_\alpha^T$, *where* $\{\mathbf{X}_\alpha\}_\alpha$ *are i.i.d. copies of* $\mathbf{X} \in \mathbb{R}^n$ *s.t.*

$$\mathbf{E}\mathbf{X} = 0, \quad \mathbf{E}\,\mathbf{X}\mathbf{X}^T = \frac{1}{n} I_n,$$

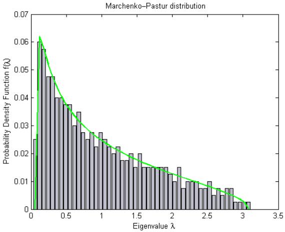
and the components of $\mathbf{X}$ *are i.i.d.*

Then as $m, n \to \infty$, $m/n \to c \geq 1$, *we have a.s.:*

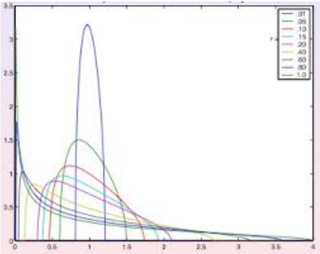
$$N_n \xrightarrow{w} N_{MP}, \quad N_{MP}(d\lambda) = \rho_{MP}(\lambda) d\lambda,$$

$$\rho_{MP}(\lambda) = \frac{\sqrt{((\lambda - a^-)(a^+ - \lambda))_+}}{2\pi\lambda},$$

$$a^\pm = (\sqrt{c} \pm 1)^2.$$



Pic. by Youssef Khmou



djalil.chafai.net

Following Pajor, A. and Pastur, L. (2009). *On the limiting empirical measure of eigenvalues of the sum of rank one matrices with log-concave distribution*, Studia Math., **195**(1), 11–29., we prove theorem not supposing that the coordinates of $\mathbf{X}$ are independent. Instead we suppose that for any deterministic $n \times n$ matrix A_n

$$\mathbf{Var}\{(A_n\mathbf{X},\, \mathbf{X})\} \leq ||A_n||^2_{op}\delta_n, \quad \delta_n = o(1),\ n \to \infty.$$

Following Pajor, A. and Pastur, L. (2009). *On the limiting empirical measure of eigenvalues of the sum of rank one matrices with log-concave distribution*, Studia Math., **195**(1), 11–29., we prove theorem not supposing that the coordinates of $\mathbf{X}$ are independent. Instead we suppose that for any deterministic $n \times n$ matrix A_n

$$\mathbf{Var}\{(A_n\mathbf{X},\ \mathbf{X})\} \leq ||A_n||^2_{op}\delta_n, \quad \delta_n = o(1),\ n \to \infty.$$

Main steps of the proof ($N_n \to N_{MP}$ in probability)

Notations:

$$M_n = \sum_{\beta=1}^m \mathbf{X}_\beta \mathbf{X}_\beta^T, \qquad G(z) = (M_n - z)^{-1}, \qquad s_n = \frac{1}{n}\mathrm{Tr}\, G,$$

$$M_n^\alpha = \sum_{\beta\neq\alpha}^m \mathbf{X}_\beta \mathbf{X}_\beta^T = M_n - \mathbf{X}_\alpha \mathbf{X}_\alpha^T, \qquad G^\alpha(z) = (M_n^\alpha - z)^{-1} \qquad s_n^\alpha = \frac{1}{n}\mathrm{Tr}\, G^\alpha$$

Following Pajor, A. and Pastur, L. (2009). *On the limiting empirical measure of eigenvalues of the sum of rank one matrices with log-concave distribution*, Studia Math., **195**(1), 11–29., we prove theorem not supposing that the coordinates of $\mathbf{X}$ are independent. Instead we suppose that for any deterministic $n \times n$ matrix A_n

$$\mathbf{Var}\{(A_n\mathbf{X},\ \mathbf{X})\} \le ||A_n||^2_{op}\delta_n, \quad \delta_n = o(1),\ n \to \infty.$$

Main steps of the proof ($N_n \to N_{MP}$ in probability)

Notations:

$$M_n = \sum_{\beta=1}^m \mathbf{X}_\beta \mathbf{X}_\beta^T, \qquad G(z) = (M_n - z)^{-1}, \qquad s_n = \frac{1}{n}\mathrm{Tr}\, G,$$

$$M_n^\alpha = \sum_{\beta \neq \alpha}^m \mathbf{X}_\beta \mathbf{X}_\beta^T = M_n - \mathbf{X}_\alpha \mathbf{X}_\alpha^T, \qquad G^\alpha(z) = (M_n^\alpha - z)^{-1} \qquad s_n^\alpha = \frac{1}{n}\mathrm{Tr}\, G^\alpha$$

Note that $\|G(z)\|_{op} \le |\Im z|^{-1},\ |s_n(z)| \le |\Im z|^{-1},\ s_n^\alpha = s_n + O(n^{-1})$.

Following Pajor, A. and Pastur, L. (2009). *On the limiting empirical measure of eigenvalues of the sum of rank one matrices with log-concave distribution*, Studia Math., **195**(1), 11–29., we prove theorem not supposing that the coordinates of $\mathbf{X}$ are independent. Instead we suppose that for any deterministic $n \times n$ matrix A_n

$$\mathbf{Var}\{(A_n\mathbf{X},\ \mathbf{X})\} \le ||A_n||_{op}^2 \delta_n, \quad \delta_n = o(1),\ n \to \infty.$$

Main steps of the proof ($N_n \to N_{MP}$ in probability)

Notations:

$$M_n = \sum_{\beta=1}^m \mathbf{X}_\beta \mathbf{X}_\beta^T, \qquad G(z) = (M_n - z)^{-1}, \qquad s_n = \frac{1}{n}\mathrm{Tr}\, G,$$

$$M_n^\alpha = \sum_{\beta \neq \alpha}^m \mathbf{X}_\beta \mathbf{X}_\beta^T = M_n - \mathbf{X}_\alpha \mathbf{X}_\alpha^T, \qquad G^\alpha(z) = (M_n^\alpha - z)^{-1} \qquad s_n^\alpha = \frac{1}{n}\mathrm{Tr}\, G^\alpha$$

Note that $\|G(z)\|_{op} \le |\Im z|^{-1},\ |s_n(z)| \le |\Im z|^{-1},\ s_n^\alpha = s_n + O(n^{-1})$.

Step 1: $\mathbf{Var}\, s_n = o(1), n \to \infty$.

Step 2: $s-?:\ \mathbf{E}s_n = \frac{1}{n}\mathbf{E}\operatorname{Tr} G \to s.$

Step 2: $s-?$: $\mathbf{E}s_n = \frac{1}{n}\mathbf{E}\,\mathrm{Tr}\, G \to s.$ $\Rightarrow$ $\rho_{MP}(\lambda) = \frac{1}{\pi}\lim_{\varepsilon\to 0}\Im s(\lambda + i\varepsilon).$

Step 2: $s-?$: $\mathbf{E}s_n = \frac{1}{n}\mathbf{E}\,\mathrm{Tr}\,G \to s$. $\Rightarrow$ $\rho_{MP}(\lambda) = \frac{1}{\pi}\lim_{\varepsilon\to 0}\Im s(\lambda + i\varepsilon)$.

- $\mathbf{E}\,\mathbf{X}\mathbf{X}^T = \frac{1}{n}I_n \Rightarrow \mathbf{E}(G^\alpha \mathbf{X}_\alpha, \mathbf{X}_\alpha) = \frac{1}{n}\mathbf{E}\,\mathrm{Tr}G^\alpha = \mathbf{E}s_n + O(n^{-1})$

Step 2: $s-?$: $\mathbf{E}s_n = \frac{1}{n}\mathbf{E}\,\mathrm{Tr}\,G \to s$. $\Rightarrow$ $\rho_{MP}(\lambda) = \frac{1}{\pi}\lim_{\varepsilon\to 0}\Im s(\lambda + i\varepsilon)$.

- $\mathbf{E}\,\mathbf{X}\mathbf{X}^T = \frac{1}{n}I_n \Rightarrow \mathbf{E}(G^\alpha \mathbf{X}_\alpha, \mathbf{X}_\alpha) = \frac{1}{n}\mathbf{E}\,\mathrm{Tr}G^\alpha = \mathbf{E}s_n + O(n^{-1})$
- $\mathbf{Var}\{(A_n\mathbf{X},\ \mathbf{X})\} = o(1) \Rightarrow \mathbf{Var}(G^\alpha \mathbf{X}_\alpha, \mathbf{X}_\alpha) = o(1)$

Step 2: $s-?$: $\mathbf{E}s_n = \frac{1}{n}\mathbf{E}\,\mathrm{Tr}\,G \to s.$ $\Rightarrow$ $\rho_{MP}(\lambda) = \frac{1}{\pi}\lim_{\varepsilon\to 0}\Im s(\lambda + i\varepsilon)$.

- $\mathbf{E}\,\mathbf{X}\mathbf{X}^T = \frac{1}{n}I_n \Rightarrow \mathbf{E}(G^\alpha \mathbf{X}_\alpha, \mathbf{X}_\alpha) = \frac{1}{n}\mathbf{E}\,\mathrm{Tr}G^\alpha = \mathbf{E}s_n + O(n^{-1})$
- $\mathbf{Var}\{(A_n\mathbf{X},\ \mathbf{X})\} = o(1) \Rightarrow \mathbf{Var}(G^\alpha \mathbf{X}_\alpha, \mathbf{X}_\alpha) = o(1)$
- $\mathbf{E}\frac{1}{\xi} = \frac{1}{\mathbf{E}\xi} + \frac{1}{\mathbf{E}\xi}\mathbf{E}\frac{\xi - \mathbf{E}\xi}{\xi} \Rightarrow \mathbf{E}\dfrac{1}{1 + (G^\alpha \mathbf{X}_\alpha, \mathbf{X}_\alpha)} = \dfrac{1}{1 + \mathbf{E}s_n} + o(1)$

Step 2: $s-?:\ \mathbf{E}s_n=\frac{1}{n}\mathbf{E}\,\mathrm{Tr}\,G\to s.\ \Rightarrow\ \rho_{MP}(\lambda)=\frac{1}{\pi}\lim_{\varepsilon\to 0}\Im s(\lambda+i\varepsilon).$

- $\mathbf{E}\,\mathbf{X}\mathbf{X}^T=\frac{1}{n}I_n\ \Rightarrow\ \mathbf{E}(G^\alpha\mathbf{X}_\alpha,\mathbf{X}_\alpha)=\frac{1}{n}\mathbf{E}\,\mathrm{Tr}G^\alpha=\mathbf{E}s_n+O(n^{-1})$
- $\mathbf{Var}\{(A_n\mathbf{X},\ \mathbf{X})\}=o(1)\Rightarrow\mathbf{Var}(G^\alpha\mathbf{X}_\alpha,\mathbf{X}_\alpha)=o(1)$
- $\mathbf{E}\frac{1}{\xi}=\frac{1}{\mathbf{E}\xi}+\frac{1}{\mathbf{E}\xi}\mathbf{E}\frac{\xi-\mathbf{E}\xi}{\xi}\ \Rightarrow\mathbf{E}\dfrac{1}{1+(G^\alpha\mathbf{X}_\alpha,\mathbf{X}_\alpha)}=\dfrac{1}{1+\mathbf{E}s_n}+o(1)$
- $G-G^\alpha=-\frac{G^\alpha\mathbf{X}_\alpha\mathbf{X}_\alpha^T G^\alpha}{1+(G^\alpha\mathbf{X}_\alpha,\mathbf{X}_\alpha)}\ \Rightarrow\ (G\mathbf{X}_\alpha,\mathbf{X}_\alpha)=1-\frac{1}{1+(G^\alpha\mathbf{X}_\alpha,\mathbf{X}_\alpha)}$

Step 2: $s-?$: $\mathbf{E}s_n = \frac{1}{n}\mathbf{E}\,\mathrm{Tr}\,G \to s$. $\Rightarrow$ $\rho_{MP}(\lambda) = \frac{1}{\pi}\lim_{\varepsilon\to 0}\Im s(\lambda + i\varepsilon)$.

- $\mathbf{E}\,\mathbf{X}\mathbf{X}^T = \frac{1}{n}I_n \Rightarrow \mathbf{E}(G^\alpha \mathbf{X}_\alpha, \mathbf{X}_\alpha) = \frac{1}{n}\mathbf{E}\,\mathrm{Tr}G^\alpha = \mathbf{E}s_n + O(n^{-1})$
- $\mathbf{Var}\{(A_n\mathbf{X},\ \mathbf{X})\} = o(1) \Rightarrow \mathbf{Var}(G^\alpha \mathbf{X}_\alpha, \mathbf{X}_\alpha) = o(1)$
- $\mathbf{E}\frac{1}{\xi} = \frac{1}{\mathbf{E}\xi} + \frac{1}{\mathbf{E}\xi}\mathbf{E}\frac{\xi - \mathbf{E}\xi}{\xi} \Rightarrow \mathbf{E}\dfrac{1}{1 + (G^\alpha \mathbf{X}_\alpha, \mathbf{X}_\alpha)} = \dfrac{1}{1 + \mathbf{E}s_n} + o(1)$
- $G - G^\alpha = -\frac{G^\alpha \mathbf{X}_\alpha \mathbf{X}_\alpha^T G^\alpha}{1 + (G^\alpha \mathbf{X}_\alpha, \mathbf{X}_\alpha)} \Rightarrow (G\mathbf{X}_\alpha, \mathbf{X}_\alpha) = 1 - \frac{1}{1 + (G^\alpha \mathbf{X}_\alpha, \mathbf{X}_\alpha)}$
- $zG(z) = -I + G(z)M = -I + \sum_{\alpha=1}^m G\mathbf{X}_\alpha \mathbf{X}_\alpha^T$

Step 2: $s-?$: $\mathbf{E}s_n = \frac{1}{n}\mathbf{E}\,\mathrm{Tr}\,G \to s.$ $\Rightarrow$ $\rho_{MP}(\lambda) = \frac{1}{\pi}\lim_{\varepsilon\to 0}\Im s(\lambda + i\varepsilon)$.

- $\mathbf{E}\,\mathbf{X}\mathbf{X}^T = \frac{1}{n}I_n \Rightarrow \mathbf{E}(G^\alpha \mathbf{X}_\alpha, \mathbf{X}_\alpha) = \frac{1}{n}\mathbf{E}\,\mathrm{Tr}G^\alpha = \mathbf{E}s_n + O(n^{-1})$
- $\mathbf{Var}\{(A_n\mathbf{X},\ \mathbf{X})\} = o(1) \Rightarrow \mathbf{Var}(G^\alpha \mathbf{X}_\alpha, \mathbf{X}_\alpha) = o(1)$
- $\mathbf{E}\frac{1}{\xi} = \frac{1}{\mathbf{E}\xi} + \frac{1}{\mathbf{E}\xi}\mathbf{E}\frac{\xi - \mathbf{E}\xi}{\xi} \Rightarrow \mathbf{E}\dfrac{1}{1 + (G^\alpha \mathbf{X}_\alpha, \mathbf{X}_\alpha)} = \dfrac{1}{1 + \mathbf{E}s_n} + o(1)$
- $G - G^\alpha = -\frac{G^\alpha \mathbf{X}_\alpha \mathbf{X}_\alpha^T G^\alpha}{1 + (G^\alpha \mathbf{X}_\alpha, \mathbf{X}_\alpha)} \Rightarrow (G\mathbf{X}_\alpha, \mathbf{X}_\alpha) = 1 - \frac{1}{1 + (G^\alpha \mathbf{X}_\alpha, \mathbf{X}_\alpha)}$
- $zG(z) = -I + G(z)M = -I + \sum_{\alpha=1}^m G\mathbf{X}_\alpha \mathbf{X}_\alpha^T \Rightarrow$

$$z\mathbf{E}\,s_n = \frac{1}{n}\mathbf{E}\,\mathrm{Tr}(zG) = -1 + \frac{1}{n}\sum_{\alpha=1}^m \mathbf{E}\,\mathrm{Tr}\,G\mathbf{X}_\alpha \mathbf{X}_\alpha^T$$

$$= 1 - \frac{m}{n} - \frac{1}{n}\sum_{\alpha=1}^m \mathbf{E}\frac{1}{1 + (G^\alpha \mathbf{y}_\alpha, \mathbf{y}_\alpha)} = -1 + \frac{m}{n} - \frac{1}{n}\sum_{\alpha=1}^m \frac{1}{1 + \mathbf{E}s_n} + o(1).$$

Step 2: $s-?$: $\mathbf{E}s_n = \frac{1}{n}\mathbf{E}\,\mathrm{Tr}\,G \to s.$ $\Rightarrow$ $\rho_{MP}(\lambda) = \frac{1}{\pi}\lim_{\varepsilon\to 0}\Im s(\lambda+i\varepsilon)$.

- $\mathbf{E}\,\mathbf{X}\mathbf{X}^T = \frac{1}{n}I_n \Rightarrow \mathbf{E}(G^\alpha\mathbf{X}_\alpha, \mathbf{X}_\alpha) = \frac{1}{n}\mathbf{E}\,\mathrm{Tr}G^\alpha = \mathbf{E}s_n + O(n^{-1})$
- $\mathbf{Var}\{(A_n\mathbf{X},\ \mathbf{X})\} = o(1) \Rightarrow \mathbf{Var}(G^\alpha\mathbf{X}_\alpha, \mathbf{X}_\alpha) = o(1)$
- $\mathbf{E}\frac{1}{\xi} = \frac{1}{\mathbf{E}\xi} + \frac{1}{\mathbf{E}\xi}\mathbf{E}\frac{\xi-\mathbf{E}\xi}{\xi} \Rightarrow \mathbf{E}\dfrac{1}{1+(G^\alpha\mathbf{X}_\alpha, \mathbf{X}_\alpha)} = \dfrac{1}{1+\mathbf{E}s_n} + o(1)$
- $G - G^\alpha = -\frac{G^\alpha\mathbf{X}_\alpha\mathbf{X}_\alpha^T G^\alpha}{1+(G^\alpha\mathbf{X}_\alpha,\mathbf{X}_\alpha)} \Rightarrow (G\mathbf{X}_\alpha, \mathbf{X}_\alpha) = 1 - \frac{1}{1+(G^\alpha\mathbf{X}_\alpha,\mathbf{X}_\alpha)}$
- $zG(z) = -I + G(z)M = -I + \sum_{\alpha=1}^m G\mathbf{X}_\alpha\mathbf{X}_\alpha^T \Rightarrow$

$$z\mathbf{E}\,s_n = \frac{1}{n}\mathbf{E}\,\mathrm{Tr}(zG) = -1 + \frac{1}{n}\sum_{\alpha=1}^m \mathbf{E}\,\mathrm{Tr}\,G\mathbf{X}_\alpha\mathbf{X}_\alpha^T$$

$$= 1 - \frac{m}{n} - \frac{1}{n}\sum_{\alpha=1}^m \mathbf{E}\frac{1}{1+(G^\alpha\mathbf{y}_\alpha, \mathbf{y}_\alpha)} = -1 + \frac{m}{n} - \frac{1}{n}\sum_{\alpha=1}^m \frac{1}{1+\mathbf{E}s_n} + o(1).$$

Hence, $zs(z) = -1 + c - c(1+s(z))^{-1}$.

Literature:

- M.L. Mehta, *Random Matrices*, Elsevier, Amsterdam, 2004.
- G. Akemann, J. Baik, and P. Di Francesco, *The Oxford handbook of random matrix theory*, Oxford University Press, Oxford, 2011.
- G. W. Anderson, A. Guionnet and O. Zeitouni, *An introduction to random matrices*, Cambridge Studies in Advanced Mathematics, 118, Cambridge University Press, Cambridge, 2010.
- Z. Bai and J. W. Silverstein, *Spectral analysis of large dimensional random matrices*, second edition, Springer Series in Statistics, Springer, New York, 2010.
- L. Pastur and M. Shcherbina, *Eigenvalue distribution of large random matrices*, Mathematical Surveys and Monographs, 171, American Mathematical Society, Providence, RI, 2011.
- T. Tao, *Topics in random matrix theory*, **132**, American Mathematical Soc., 2012.
- R. Vershynin, *High-dimensional probability: An introduction with applications in data science*, **4**, Cambridge university press, 2018.

Convergence of empirical spectral distributions: as $n \to \infty$, for any bounded continuous φ we have with probability 1:

$$\int \varphi(\lambda) dN_n(\lambda) = n^{-1} \sum_{j=1}^{n} \varphi(\lambda_j) \to \int \varphi(\lambda)\rho(\lambda) d\lambda,$$

This is an analog of the Law of Large Numbers.

What can be said about fluctuations (CLT)?

$$\nu_n \cdot \sum_{j=1}^{n} (\varphi(\lambda_j) - \mathbf{E}\varphi(\lambda_j)) \to Z \sim \mathcal{N}(0, V) \ \textit{in distribution?}$$

The CLT for Linear Eigenvalue Statistics for Wigner Random Matrices

Theorem (AL, Pastur'09). Let $M_n = n^{-1/2}\left(W_{jk}\right)_{j,k=1}^{n}$ be a Wigner matrix,

- $W_{jk} = W_{kj} \in \mathbb{R}, j \le k$, are independent,
- $\mathbf{E}\{W_{jk}\} = 0, \quad \mathbf{E}\{W_{jk}^2\} = (1 + \delta_{jk})$,
- $\mu_{3,4} := \mathbf{E}\{W_{jk}^{3,4}\}$, $\kappa_4 := \mu_4 - 3$,
- $\varphi \in \mathcal{H}_{5/2}$.

Then as $n \to \infty$, $\sum_{j=1}^{n}(\varphi(\lambda_j) - \mathbf{E}\varphi(\lambda_j)) \to Z \sim \mathcal{N}(0, V[\varphi])$ in distribution,

$$V[\varphi] = \frac{1}{2\pi^2}\int_{-2}^{2}\int_{-2}^{2}\left(\frac{\varphi(\lambda_1)-\varphi(\lambda_2)}{\lambda_1-\lambda_2}\right)^2 \frac{(4-\lambda_1\lambda_2)d\lambda_1 d\lambda_2}{\sqrt{4-\lambda_1^2}\sqrt{4-\lambda_2^2}}$$

$$+ \frac{\kappa_4}{2\pi^2}\left(\int_{-2}^{2}\varphi(\mu)\frac{2-\mu^2}{\sqrt{4-\mu^2}}d\mu\right)^2.$$

Circular law

Let now $M_n = n^{-1/2}\{M_{ij}\}_{i,j=1}^n$, where $(M_{ij})_{i,j}$ are i.i.d. copies of ξ,

$$\mathbf{E}\,\xi = 0, \quad \mathbf{Var}\,\xi = 1,$$

and let $\lambda_1,\ldots,\lambda_n \in \mathbb{C}$ be the eigenvalues M_n.

The *density of the empirical spectral distribution* of M_n:

$$\mu_{M_n}(\lambda) := \frac{1}{n}\sum_{j=1}^{n}\delta(\lambda - \lambda_j).$$

It was conjectured in the 1950s, that μ_{M_n} converges to the density of the uniform probability measure on the unit disk:

$$\mu_{M_n} \to \pi^{-1}\mathbf{1}_D\,dxdy, \qquad \text{where} \quad D = \{|z| \le 1\}.$$

Problem: Neither the method moment nor the Stieltjes transform method work in this setting!

The hermitization trick (Girko, 1984)

$$
\begin{aligned}
\int_{\mathbb{C}} \ln|\lambda - z| \mu_{M_n}(\lambda) d\lambda &= \frac{1}{n} \sum_j \ln|\lambda_j(M_n) - z| \\
&= \frac{1}{n} \ln\Big|\prod_j (\lambda_j(M_n) - z)\Big| \\
&= \frac{1}{n} \ln\big|\det(M_n - z)\big| \\
&= \frac{1}{2n} \ln\det(M_n - z)(M_n^* - \bar{z}) \\
&= \frac{1}{2n} \sum_j \ln\lambda_j((M_n - z)(M_n^* - \bar{z})) \\
&= \frac{1}{2} \int_0^\infty \ln(t)\, \mu_{(M_n - z)(M_n^* - \bar{z})}(t) dt.
\end{aligned}
$$

Need to estimate the smallest singular value $s_{min}(M_n - zI)$ of $M_n - zI$!

History and references:

$M_n = n^{-1/2}(M_{ij})_{i,j=1}^n$, $(M_{ij})_{i,j}$ are i.i.d. copies of ξ: $\mathbf{E}\,\xi = 0$, $\mathbf{Var}\,\xi = 1$.

Mehta **(1967)**: ξ is a standard complex Gaussian variable (using the joint density function of the eigenvalues, discovered by Ginibre **(1965)**)

Girko **(1984)**: $\mathbf{E}\,|\xi|^{2+\varepsilon} < \infty$ (but the proof has gaps)

Edelman **(1997)**: ξ is a standard real Gaussian variable

Bai **(1997)**: ξ has bounded density and bounded 6th moment (later improved to $(2+\varepsilon)$-moment in his book with Silverstein **(2010)**)

Girko **(2004)**: $\mathbf{E}\,|\xi|^{4+\varepsilon} < \infty$ (no density conditions!)

Pan, Zhou **(2010)**: $\mathbf{E}\,|\xi|^4 < \infty$

Tao, Vu **(2008)**: $\mathbf{E}\,|\xi|^{2+\varepsilon} < \infty$

Götze, Tikhomirov **(2010)**: $\mathbf{E}\,|\xi|^2(\ln|\xi|)^{20} < \infty$

Tao, Vu **(2010)**: Universality: No additional conditions!

Many recent works on matrices with non i.i.d. entries. In particular, for sparse matrices: Götze–Tikhomirov, Tao–Vu, Basak–Rudelson.

Circular law for regular graphs

(Based on a joint work with: A. Litvak, K. Tikhomirov, N. Tomczak-Jaegermann, P. Youssef)

$G \in \mathcal{D}_{n,d} \Leftrightarrow$ every vertex of G has exactly d in-neighbors and d out-neighbors

$$\mathrm{P}\{G \in \Gamma\} = \frac{|\Gamma|}{|\mathcal{D}_{n,d}|}, \quad \Gamma \subset \mathcal{D}_{n,d}.$$

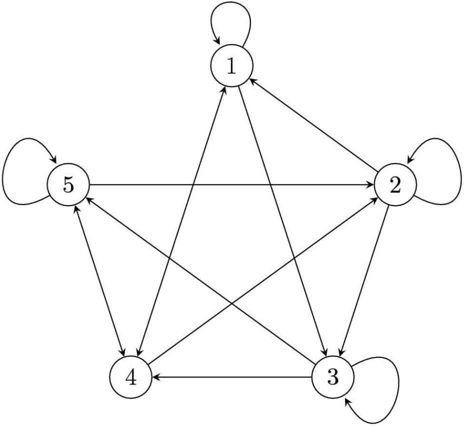
$M \in \mathcal{M}_{n,d} \Leftrightarrow$

$$M_{ij} = \begin{cases} 1, & \text{if there is an edge from } i \text{ to } j; \\ 0, & \text{otherwise.} \end{cases}$$

$$\sum_{i=1}^{n} M_{ij} = \sum_{j=1}^{n} M_{ij} = d$$

A closely related model: Erdös-Renyi graphs.

$n = 5, d = 3$



$G \in \mathcal{D}_{n,d}$

1	0	1	1	0
1	1	1	0	0
0	0	1	1	1
1	1	0	0	1
0	1	0	1	1

$M \in \mathcal{M}_{n,d}$

Quantitative estimates for the smallest singular value

N.Cook, 2017: Let $d > C \ln^{11} n$. Then

$$\mathbb{P}\Big(s_{\min} > 1/n^{C(\ln n)/\ln d}\Big) > 1 - C \ln^{5.5} n/\sqrt{d}.$$

Theorem (LLTTP, 2017). Let $C < d < n/\ln^2 n$. Then

$$\mathbb{P}\Big(s_{\min} > 1/n^6\Big) > 1 - C \ln^2 d/\sqrt{d}.$$

Conjecture: $s_{\min} \approx \sqrt{d}/n.$

Proof: Need to estimate

$$\mathrm{P}(s_{\min}(A) < \delta) = \mathrm{P}\Big(\inf_{x \in S^{n-1}} \|Ax\|_2 < \delta\Big) = \mathrm{P}\Big(\exists x \in S^{n-1} : \|Ax\|_2 < \delta\Big)$$

- ε-net argument,
- anti-concentration inequalities,
- study of normal vectors of hyperplanes spanned by the rows of A_n.

Circular law for adjacency matrices

Let M be uniformly distributed in the set of $n \times n$ matrices with $0/1$ entries, such that sums in rows and in columns are equal to d.

N.Cook, 2017: The circular law holds for $d^{-1/2}M$ provided that $d > \ln^{96} n$.

Theorem (LLTTP, 2018). The circular law holds for $d^{-1/2}M$ provided that $d = d(n) \to \infty$ as $n \to \infty$.

Conjecture (complex Kesten–McKay distribution): For every fixed d, as $n \to \infty$ the normalized counting measures of eigenvalues of $M \in \mathcal{M}_{n,d}$ converge to the probability measure (called the Kesten–McKay distribution) with the density

$$\frac{1}{\pi}\frac{d^2(d-1)}{(d^2-|z|^2)^2}\chi_{\{|z|<\sqrt{d}\}}dxdy.$$