An introduction to random matrices and free probability in portfolio theory: a data science approach.

Andrés García Medina

CONACyT Research Fellow at Center for Research in Mathematics, Monterrey (Mexico).
https://sites.google.com/view/andresgm/home

Observatory of Complex Systems at Dipartimento di Fisica e Chimica – Emilio Segrè, Palermo University. July 11-12, 2022.

FOSEC SEP-INVESTIGACION BASICA-CONACY [A1-S-43514]

1. Introduction
2. Elements of Random Matrices
3. Classical Ensembles
4. Simulation: semicircle and Marchenko-Pastur laws
5. Wigner's Surmise
6. Simulation: Wigner's Surmise
7. Portfolio Theory
8. Simulation: optimal frontier and classical RMT technique

Motivation

- Universality[1]

Werner Heisenberg and Eugene Wigner (1928)

[1] The statistical properties of the city transport in Cuernavaca (Mexico) and random matrix ensembles. J. Phys. A: Math. Gen. 33 (2000) L229–L234.

Background

Matrices appear in many areas of the sciences, from mathematics to physics, computer science, biology, economics, and quantitative finance.

In many circumstances the matrices that we encounter when modeling a phenomenon are large and without a particular structure.

Eugene Wigner discovered that it is possible (in many circumstances) to replace a large and complex (but deterministic) matrix by a typical element of a certain ensemble of random matrices.

Wigner's idea supposes the concept of *universality: the statistical distribution of eigenvalues does not depend on the specific matrix that represents the system, but on its symmetry*.

This idea has been incredibly fruitful and has led to the development of a subfield of mathematical physics called *Random Matrices Theory*.

Next we will investigate the simplest cases of ensembles of random matrices.

Background

Random matrix theory is a new type of statistical mechanics, where instead of having an ensemble of states governed by the same Hamiltonian, one has an ensemble of Hamiltonians governed by the same symmetry.

- Introduced into mathematical statistics by Wishart in 1928.
- Many mathematicians later worked out of purely theoretical interest.
- In 1935 Elie Cartan classifies ensembles according to the symmetry that is preserved (in French).
- The first book containing the most relevant mathematical results was published by L. K. Hua in 1958 (in Chinese).
- In the 1950s Wigner used it to deal with the eigenvalues and eigenvectors of quantum many-body complex systems.
- F.J. Dyson and M.L. Mehta made analytical calculations (1960).
- Mehta published in 1967 a book with the main techniques developed up to that time.
- Quantum Chaos Conjecture, was published in a paper by Bohigas, Giannoni, and Schmidt in 1984.

Background

- For systems with time invariance and rotational symmetry, the Hamiltonian matrix H can be chosen real and symmetric

$$H = H'. \tag{1}$$

- For systems without time reversal invariance, the matrix H is Hermitian

$$H = H^{\dagger}. \tag{2}$$

- For time-reversal invariant systems with spin $1/2$ and no rotational symmetry, the Hamiltonian is written in terms of the Pauli matrices σ_γ

$$H^0_{nm} I_2 - i \sum_{\gamma=1}^{3} H^{(\gamma)}_{nm} \sigma_\gamma, \tag{3}$$

where H^0 is symmetric and H^γ are antisymmetric.

Background

- The probability density of finding a particular matrix within one of these ensembles is given by

$$P_{N\beta}(H) \propto e^{-\frac{\beta N}{2} Tr(H^2)} \tag{4}$$

- Symmetry properties and weight functions $P_{N\beta}(H)$ are invariant under orthogonal ($\beta = 1$), unitary ($\beta = 2$), and symplectic transformations ($\beta = 4$) of the Hamiltonian.
- These ensembles are known as *GOE*, *GUE* and *GSE*.
- Each member H_β can be represented in the form

$$H_\beta = V \Lambda V^{-1}, \tag{5}$$

where *Lambda* matrix of eigenvalues and V orthogonal, unitary or symplectic matrices (depending on whether $\beta = 1$, 2 or 4).

$\Rightarrow$ When measuring statistical properties of ensembles, the distribution of eigenvalues is independent of the eigenvectors.

Definition

- What is a random matrix? An matrix whose elements are random variables
- Random Matrices Theory (RMT) replaces deterministic matrices with random matrices
- When working with very complicated matrices RMT replaces the system matrix with a random one and calculates averages (as well as the statistical properties of interest).

Example

Be a $H_{N\times N}$ matrix with i.i.d. elements such that $H_{ij} \sim N(0,1)$, para $i,j = 1,\ldots,p$. Example($p = 3$):

$$H = \begin{pmatrix} 1.24 & 0.05 & -0.87 \\ -0.18 & 0.78 & -1.31 \\ -0.49 & -0.62 & 0.03 \end{pmatrix} \tag{6}$$

We can notice $H_{ij} \neq H_{ji}$.

We call each realization a sample of the ensemble.

In general, the eigenvalues of H are complex. To obtain real eigenvalues we symmetrize our matrix H. An example of symmetrization is

$$H_s = (H + H')/2 \rightarrow H_s = H_s' \tag{7}$$

GOE

In our example

$$H = \begin{pmatrix} 1.24 & -0.065 & -0.68 \\ -0.065 & 0.78 & -0.965 \\ -0.68 & -0.965 & 0.03 \end{pmatrix} \tag{8}$$

With the convenience that $\lambda \in \{-0.8, 1.1, 1.7\}$ are reals

Congratulations! you have created your first random matrix from the Gaussian Orthogonal Ensemble (GOE).

GUE

In the case of a Hermitian matrix, for example

$$H_{her} = \begin{pmatrix} 0.3252 & 0.3077 + 0.2803\mathrm{i} \\ 0.3077 - 0.2803\mathrm{i} & -1.7115 \end{pmatrix} \tag{9}$$

where $H_{her} = H_{her}^{H}$.

We now have a realization of the Unitary Gaussian Ensemble (GUE), with the characteristic that the inputs are complex, while the eigenvalues are real.

Specifically, in our example $\lambda \in \{-1.79, 0.406\}$

GSE

In the case of a symplectic matrix (relation to quaternions and systems with spin)

$$H_{2N\times 2N} = \begin{pmatrix} X & Y \\ -Y* & X* \end{pmatrix} \tag{10}$$

where $X, Y \in \mathbb{C}^{N\times N}$.

We symmetrize

$$H_{sc} = (H + H')/2 \tag{11}$$

It is obtained $2p$ eigenvalues, $\lambda \in \{\lambda_1, \lambda_1, \ldots, \lambda_p, \lambda_p\}$.

Then we have a recipe to build a random matrix from the Symplectic Gaussian Ensemble (GSE).

Average spectral density

How do you calculate the shape of the histograms of $N \times m$ eigenvalues from the jpdf $\rho(\lambda_1, \ldots, \lambda_N)$?

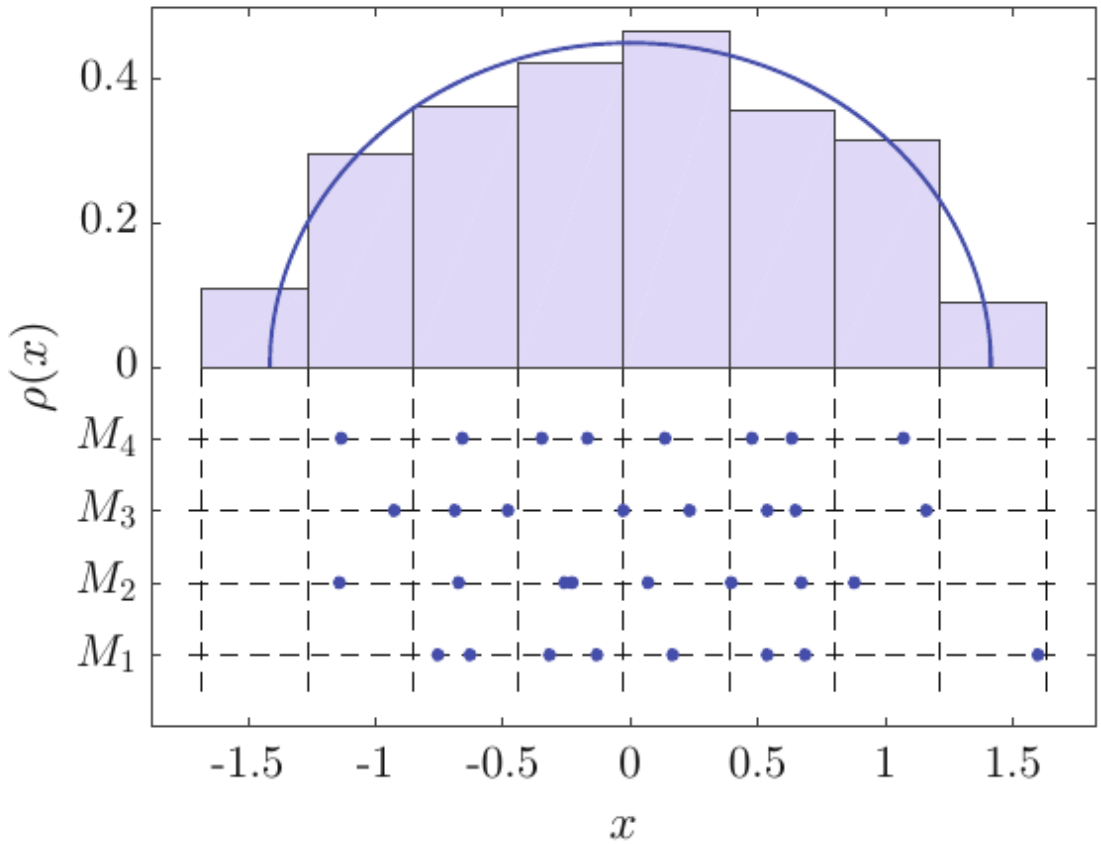


Figure (1) Source: Livan, G., Novaes, M., & Vivo, P. Introduction to random matrices theory and practice. Springer (2018).

Classification of Random Matrices Ensembles

$\Rightarrow$ No free lunch

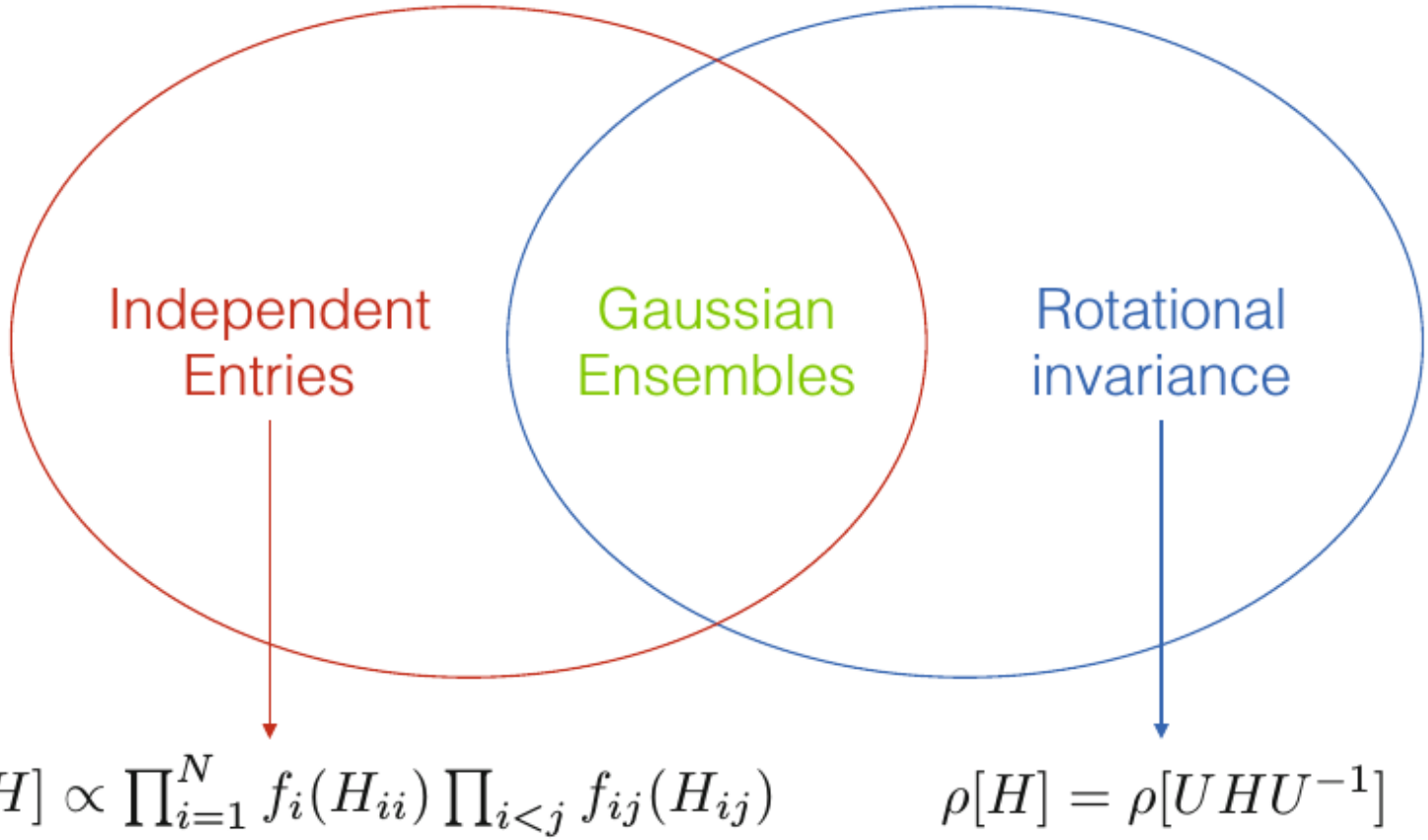


Figure (2) Source: Livan, G., Novaes, M., & Vivo, P. Introduction to random matrices theory and practice. Springer (2018).

Classification of Random Matrices Ensembles

- **Independent Entries**: the first group on the left brings together matrix models whose inputs are independent random variables, modulo the symmetry requirements. Random matrices of this type are often called Wigner matrices. Examples: adjacency matrices of random graphs or matrices with independent power-law entries.
- **Rotational Invariance**: the second group from the right is characterized by the so-called rotational invariance. In essence, this property means that any two matrices that are related by a similarity transformation: $H^{(2)} = UH^{(1)}U'$ occur in the ensemble with equal probability.
- **Intersección**: What about the intersection between the two classes? It turns out that it only contains the Gaussian Ensembles!

Normalized traces and sample averages

We can generalize the notion of expected value and moments of classical probability to *large* random matrices.

It turns out that the proper analog of expected value is the normalized trace operator defined for the random matrix A of dimension $N \times N$ as

$$\phi(A) = \frac{1}{N}\mathbb{E}[TrA] \tag{12}$$

The normalization by $1/N$ is to have a finite operator when $N \to \infty$. For example $\phi(\mathbb{1}) = 1$ independent of the dimension.

In the case of polynomial functions of the matrix A the trace of the function can be calculated on the eigenvalues [2]

$$\frac{1}{N} Tr(F(A)) = \frac{1}{N} \sum_{k=1}^{N} F(\lambda_k) \tag{13}$$

[2] Cyclic trace property

Normalized traces and sample averages

From here on we will denote $\langle \cdot \rangle$ as the average over the eigenvalues of a realization of A

$$\langle F(\lambda) \rangle := \frac{1}{N} \sum_{k=1}^{N} F(\lambda_k) \tag{14}$$

In the case of random matrices, many scalar quantities like $\phi(F(A))$ do not fluctuate from sample to sample, or in other words, such fluctuations tend to zero in the large N limit.

This phenomenon is known in physics as *self-averaging*, while mathematicians speak of *concetration of measure*

$$\phi(F(A)) = \frac{1}{N} \mathbb{E}[Tr(F(A))] \approx \frac{1}{N} Tr(F(A)) = \langle F(\lambda) \rangle, \tag{15}$$

for a single realization of A.

Normalized traces and sample averages

When the eigenvalues of a random matrix A converge to a well-defined density $\rho(\lambda)$ we can write

$$\phi(F(A)) = \int \rho(\lambda) F(\lambda) d\lambda \tag{16}$$

Using $F(A) = A^k$ it is possible to define the k-th moment of a random matrix

$$m_k = \phi(A^k) \tag{17}$$

The first moment is simply the normalized trace of A, while $m_2 = \frac{1}{N} \sum_{ij} A_{ij}^2$ is the normalized sum of squares of all items. In particular, the square root of m_2 satisfies the norm axioms and is known as the Frobenius norm of A

$$||A||_F := \sqrt{m_2} \tag{18}$$

Resolvent and Stieltjes Transform

Given a symmetric real matrix A of dimension $N \times N$, its resolvent is given by [3]

$$G_A(z) = (z\mathbb{1} - A)^{-1}, \tag{19}$$

where z is a complex variable defined outside the domain of the eigenvalues of A. Then the Stieltjes transform of A is given by (blackboard)

$$g_N^A(z) = \frac{1}{N} Tr(G_A(z)) = \frac{1}{N} \sum_{k=1}^{N} \frac{1}{z - \lambda_k}, \tag{20}$$

where λ_k are the eigenvalues of A.

On the other hand, for a random matrix A we can define its empirical spectral distribution (ESD) also called sample eigenvalue density

$$\rho_N(\lambda) = \frac{1}{N} \sum_{k=1}^{N} \delta(\lambda - \lambda_k), \tag{21}$$

where $\delta(x)$ is the Dirac delta function.

[3] The resolvent formalism is a technique for applying complex analysis concepts to the study of the spectrum of certain operators.

Resolvent and Stieltjes Transform

In this way, the Stieltjes transform can be written as (replace $\rho_N(\lambda)$ by eq. 21 and you recover eq. 20)

$$g_N(z) = \int_{-\infty}^{\infty} \frac{\rho_N(\lambda)}{z-\lambda} d\lambda \tag{22}$$

Note that $g_N(z)$ behaves well for any $z \notin \{\lambda_k : 1 \leq k \leq N\}$. In particular, it is well-defined at ∞:

$$\begin{aligned} g_N(z) &= \int \frac{\rho_N(\lambda)}{z-\lambda} d\lambda = \frac{1}{z} \int \frac{\rho_N(\lambda)}{1-\lambda/z} d\lambda \\ &= \frac{1}{z} \int \rho_N(\lambda) \left(\sum_{k=0}^{\infty} \left(\frac{\lambda}{z} \right)^k \right) d\lambda \quad \text{(geometric series)} \\ &= \frac{1}{z} \sum_{k=0}^{\infty} \int \rho_N(\lambda) \left(\frac{\lambda}{z} \right)^k d\lambda = \sum_{k=0}^{\infty} \frac{1}{z^{k+1}} \frac{1}{N} Tr(A^k), \quad \frac{1}{N} Tr(A^0) = 1. \end{aligned} \tag{23}$$

Resolvent and Stieltjes Transform

The cases of interest are the random matrices A such that for large values of N the normalized traces of the powers of A converge to their expected values

$$\lim_{N\to\infty} \frac{1}{N} Tr(A^k) = \phi(A^k) \tag{24}$$

It is found that for large values of z, the function $g_A(z)$ converges to a deterministic limit $g(z) = \lim_{N\to\infty} \mathbb{E}[g_N(z)]$, equivalently

$$g(z) = \sum_{k=0}^{\infty} \frac{1}{z^{k+1}} \phi(A^k). \tag{25}$$

Thus, $g(z)$ is the moment-generating function of A.

In random matrices it holds that the knowledge of all moments of A is equivalent to knowledge of the eigenvalue densities of A.

Resolvent and Stieltjes Transform

Technical argument:

In the limit $N \to \infty$ the poles of the ESD merge, so

$$\frac{1}{N}\sum_{k=1}^{N}\delta(\lambda-\lambda_k) \sim \rho(\lambda) \tag{26}$$

Thus, the density $\rho(\lambda)$ has an extended support and isolated points of measure zero

$$g(z) = \int_{supp[\rho]} \frac{\rho(\lambda)d\lambda}{z-\lambda} \tag{27}$$

which is the Stieltjes transform of the limit measure $\rho(\lambda)$

Limit spectral density

If we know the Stieltjes transform for a given ensemble it is possible to know its limiting spectral density

$$g(x-i\epsilon) = \int \frac{\rho(x')d\lambda}{z-x'} = \int dx' \frac{\rho(x')}{x-i\epsilon-x'}\left(\frac{x+i\epsilon-x'}{x+i\epsilon-x'}\right)$$
$$= \int dx' \frac{\rho(x')(x-x')}{(x-x')^2+\epsilon^2} + i\int dx'\rho(x')\frac{\epsilon}{(x-x')^2+\epsilon^2}. \tag{28}$$

Let us take the limit $\epsilon \to 0^+$ (▸ blackboard)

$$\lim_{\epsilon\to 0^+} g(x-i\epsilon) = Pr\left[\int dx' \frac{\rho(x')}{x-x'}\right] + i\pi\rho(x) \tag{29}$$

Thus, we can write

$$\rho(x) = \frac{1}{\pi}\lim_{\epsilon\to 0^+} Im\{g(x-i\epsilon)\} \tag{30}$$

This implies that if we can calculate the Stieltjes transform in the complex plane then it is possible to obtain the limiting spectral density.

Wigner Matrices

We have defined a Wigner matrix as a symmetric matrix (X=X') with Gaussian entries of mean zero (although they can be defined more generally).

The first few moments of the Wigner matrix X are given by

$$\phi(X) = \frac{1}{N}\mathbb{E}[TrX] = \frac{1}{N}Tr\mathbb{E}[X] = 0 \tag{31}$$

$$\phi(X^2) = \frac{1}{N}\mathbb{E}[TrXX'] = \frac{1}{N}\mathbb{E}\left[\sum_{ij=1}^{N} X_{ij}^2\right] = \frac{1}{N}[N(N-1)\sigma_{od}^2 + N\sigma_d^2] \tag{32}$$

$$\phi(X^k) = 0, \quad k > 2 \tag{33}$$

where $\sigma_d^2, \sigma_{od}^2$ are the diagonal and off-diagonal varianza, respectively.

To satisfy rotational invariance and keep the second finite moment is necessary to set

$$\sigma_d^2 = 2\sigma_{od}^2 = 2\sigma^2/N. \tag{34}$$

The above equations is another way to define the Orthogonal Gaussian Ensemble (GOE).

Rotational Invariance

We must remember that to rotate a vector v one applies a rotation matrix $O : w = Ov$, where O is an orthogonal matrix $O' = O^{-1}$ such that $OO' = 1$.

Observations:

- In general O is not symmetric.
- To rotate the base of X apply $\tilde{X} = OXO'$
- The eigenvalues of $\tilde{X}$ are the same as those of X.
- The corresponding eigenvectors are $\{Ov\}$, where $\{v\}$ are the eigenvectors of X.

A rotationally invariant ensemble of random matrices is such that the matrix $OXO'$is just as likely as X itself, that is, $OXO' \overset{\text{in law}}{=} X$

Rotational Invariance

One way to observe the rotational invariance of the Wigner ensemble is through the joint density of its matrix elements

$$P(\{X_{ij}\}) = \left(\frac{1}{2\pi\sigma_d^2}\right)^N \left(\frac{1}{2\pi\sigma_{od}^2}\right)^{N(N-1)/2} \exp\left\{ -\sum_{i=1}^{N} \frac{X_{ii}^2}{2\sigma_d^2} - \sum_{i<j}^{N} \frac{X_{ij}^2}{2\sigma_{od}^2} \right\}, \quad (35)$$

where only the diagonal and upper diagonal elements are independent variables.

Thus, with the election $\sigma_d^2 = 2\sigma_{od}^2 = 2\sigma^2/N$ we have

$$P(\{X_{ij}\}) \propto \exp\left\{ -\frac{N}{4\sigma^2} TrX^2 \right\} \quad (36)$$

Then, under the change of variable $X \to \tilde{X} = OXO'$ the argument of the exponential is invariant.

In general, a matrix will be rotationally invariant when its joint probability density of its elements can be written as $P(\{M_{ij}\}) \propto \exp\{-NTrV(M)\}$, where $V(\cdot)$ is an arbitrary function.

Density of Eigenvalues of a Wigner Matrix: self-consistent

To find the Stieltjes transform of the Wigner ensemble we can use the cavity method or *self-consistent equation*.

This method consists of finding a relation between the Stiltjes transform of a Wigner matrix of size N and one of size $N-1$

In the large N limit, both transformations must converge to the same limit, so we get a self-consistent equation that can be solved relatively easily.

So, we want to compute $g_N^X(z)$ when X is a Wigner matrix with $X_{ij} \sim N(0, \sigma^2/N)$ and $X_{ii} \sim N(0, 2\sigma^2/N)$. In the limit of N large, g_N^X converges to a well-defined function $g(z)$.

For this we need to remember the Schur's complement formula.

Density of Eigenvalues of a Wigner Matrix: self-consistent

Schur's complement relates the blocks of the inverse of a matrix to the inverse of the blocks of the original matrix.

Property

Let be an invertible M matrix which we divide into four blocks

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \quad y \quad M^{-1} = Q = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}, \tag{37}$$

where
$[M_{11}] = n \times n, [M_{12}] = n \times (N-n), [M_{21}] = (N-n) \times n, [M_{22}] = (N-n) \times (N-n)$,
y M_{22} is invertible. Also, the integer n can take any value between 1 and $N-1$. So, the upper left block of Q $(n \times n)$ is given by

$$Q_{11}^{-1} = M_{11} - M_{12}(M_{22})^{-1} M_{21}, \tag{38}$$

where the right hand side is known as the Schur's complement of the block M_{22} of the matrix M.

Density of Eigenvalues of a Wigner Matrix: self-consistent

Using Schur's complement formula we can calculate the elements of $M = z\mathbb{1} - X$. Así $M^{-1} = G_X$, y

$$\frac{1}{(G_X)_{11}} = M_{11} - \sum_{k,l=2}^{N} M_{1k}(M_{22})^{-1}_{kl} M_{l1}, \tag{39}$$

where M_{22} is the submatrix of size $(N-1)\times(N-1)$ of M such that the first row and column have been removed.

Assumptions: (i) $\phi(X) = 0$, (ii) $\phi(X^2) = \frac{1}{N}[N(N-1)\sigma_{od}^2 + N\sigma_d^2]$, (iii) $\sigma_d^2 = 2\sigma_{od}^2 = 2\sigma^2/N$. (iv) i.i.d. (v) $\phi(X^k) = 0, k > 2$

It is argued that for large N the right hand side is dominated by its expected values with small fluctuations of the order $\mathcal{O}(1/\sqrt{N})$. Therefore, only its expected value will be calculated, although it is also possible to estimate its fluctuations (see assumptions).

Density of Eigenvalues of a Wigner Matrix: self-consistent

Then, we need to compute the expected value of the terms:

1. $\frac{1}{(G_X)_{11}}$
2. M_{11}
3. $\sum_{k,l=2}^{N} M_{1k}(M_{22})^{-1}_{kl} M_{l1}$

The second term is direct: $\mathbb{E}[M_{11}] = \mathbb{E}[z - X_{11}] = z$ (assumption i)

The entries of M_{22} are independent of $M_{1i} = -X_{1i}$ (assumption iv).

Thus, we can take the partial expectation on the elements $\{X_{1i}\}, i = 2, \ldots, N$; first and obtain (assumption i and ii)

$$\mathbb{E}_{X_{1i}}\left[M_{1i}(M_{22})^{-1}_{ij} M_{j1}\right] = (M_{22})^{-1}_{ij}\mathbb{E}_{X_{1i}}\left[M_{1i}M_{j1}\right] = \tag{40}$$

$$(M_{22})^{-1}_{ii}\sigma^2_{od}\delta_{ij} = \frac{\sigma^2}{N}(M_{22})^{-1}_{ii}\delta_{ij} \tag{41}$$

Then

$$\mathbb{E}_{X_{1i}}\left[\sum_{k,l=2}^{N} M_{1k}(M_{22})^{-1}_{kl} M_{l1}\right] = \frac{\sigma^2}{N} Tr\left((M_{22})^{-1}\right) \tag{42}$$

Density of Eigenvalues of a Wigner Matrix: self-consistent

Further $\frac{1}{(N-1)} Tr\left((M_{22})^{-1}\right)$ is the Stieltjes transform of a Wigner matrix of size

$N-1$ and variance $\sigma^2(N-1)/N$.

In the large N limit, the Stieltjes transform must be independent of the size of the matrix, so the difference between N and $N-1$ is negligible.
From the above we have

$$\mathbb{E}\left[\sum_{k,l=2}^{N} M_{1k}(M_{22})^{-1}_{kl} M_{l1}\right] = \mathbb{E}\left[\frac{\sigma^2}{N} Tr\left((M_{22})^{-1}\right)\right] \to \sigma^2 g(z) \tag{43}$$

Thus $1/(G_X)_{11}$ approaches a deterministic number with negligible fluctuations. Its expectation can be approximated by the Jensen inequality

Theorem

If X is a random variable and ϕ is a convex function then $\phi(\mathbb{E}[X]) \leq \mathbb{E}[\phi(X)]$, where $\mathbb{E}[X]$ represents the expected value of X

Density of Eigenvalues of a Wigner Matrix: self-consistent

We have

$$\frac{1}{\mathbb{E}\left[(G_X)_{11}\right]} \leq \mathbb{E}\left[\frac{1}{(G_X)_{11}}\right] \tag{44}$$

The equality holds when $N \to \infty$ (negligible variance)

Now, from the rotational invariance of X and consequently of G_X, all diagonal elements of G_X must have the same expected value (see derivation of 20)

$$\mathbb{E}[(G_X)_{11}] = \frac{1}{N}(\mathbb{E}[(G_X)_{11}] + \cdots + \mathbb{E}[(G_X)_{11}]) = \tag{45}$$

$$\frac{1}{N}\mathbb{E}[(G_X)_{11} + \cdots + (G_X)_{NN}] = \tag{46}$$

$$\frac{1}{N}\mathbb{E}[Tr(G_X)] = \mathbb{E}[g_N] \to g \tag{47}$$

Then, for $N \to \infty$

$$\mathbb{E}\left[\frac{1}{(G_X)_{11}}\right] \to \frac{1}{g(z)} \tag{48}$$

Density of Eigenvalues of a Wigner Matrix: self-consistent

Considering all the previous arguments, the following quadratic expression for g(z) is finally obtained

$$\frac{1}{g(z)} = z - \sigma^2 g(z) \tag{49}$$

Then the limiting spectral density is given by

$$\rho(x) = \frac{1}{\pi} \lim_{\epsilon \to 0^+} Im\{g(x - i\epsilon)\} \tag{50}$$

With a little extra algebra we can finally get the well-known semicircle law(▸ blackboard) ...

$$\rho(x) = \frac{1}{\pi}\sqrt{2 - x^2} \tag{51}$$

Density of Eigenvalues of a Wigner Matrix: Coulomb gas

Our intention is to derive an expression for the joint probability distribution function of eigenvalues of the symmetric matrices belonging to the GOE ensemble.

The first step in finding the jpdf is to perform the change of variables $H \to \{\Lambda, O\}$

$$\rho(H_{11}, \ldots, H_{pp}) \prod_{i \leq j} dH_{ij} = \rho(H_{11}(\Lambda, O), \ldots, H_{pp}(\Lambda, O)) |J(H \to \{\Lambda, O\}| dO \prod_{i=1}^{p} d\lambda_i \tag{52}$$

Density of Eigenvalues of a Wigner Matrix: Coulomb gas

For the GOE ($\beta = 1$) we have

$$\rho(H) = \prod_{i=1}^{p} \frac{1}{\sqrt{2\pi}} \exp\left(-H_{ii}^2/2\right) \prod_{i<j} \frac{1}{\sqrt{\pi}} \exp\left(-H_{ij}^2\right), \tag{53}$$

for the entries of the upper triangular matrix of H_s (including the diagonal).

In addition, due to rotational invariance, we had arrived at the following expression

$$\begin{aligned} \rho(H) &= \frac{1}{(2\pi)^{p/2}\pi^{\frac{p^2-p}{4}}} \exp\left(-\frac{1}{2} Tr(H^2)\right) \\ &= \frac{1}{(2\pi)^{p/2}\pi^{\frac{p^2-p}{4}}} \exp\left(-\frac{1}{2}\sum_{i=1}^{p} \lambda_i^2\right), \end{aligned} \tag{54}$$

where the last equality is due to the cyclic property of the trace.

Density of Eigenvalues of a Wigner Matrix: Coulomb gas

On the other hand, it can be shown that the Jacobian is a function only of the eigenvalues

$$|det(J)| = \prod_{j<k}^{p} |\lambda_j - \lambda_k| = \Delta_p(\Lambda) \tag{55}$$

which is known as the Vandermonde determinant. (Blackboard)

Density of Eigenvalues of a Wigner Matrix: Coulomb gas

So how do we get the jpdf of the eigenvalues? Integrating the components of $\{O\}$, i.e., through the marginal

$$\rho(\lambda_1, \ldots, \lambda_p) \prod_{i=1}^{p} d\lambda_i = \left(\int_{\mathbb{V}_p} dO \rho(\lambda_1, \ldots, \lambda_p, O) \right) \prod_{i=1}^{p} d\lambda_i \tag{56}$$

This integral is possible only if the inputs, once expressed in terms of the eigenvalues and eigenvectors, are themselves independent of the eigenvectors. Fortunately this is our case

$$\begin{aligned} \rho(\lambda_1, \ldots, \lambda_p, O) &= \rho(H_{11}(\Lambda, O), \ldots, H_{pp}(\Lambda, O)) |J(H \to \{\Lambda, O\}| \\ &= \frac{1}{(2\pi)^{p/2} \pi^{\frac{p^2-p}{4}}} \exp\left(-\frac{1}{2} \sum_{i=1}^{p} \lambda_i^2 \right) \prod_{j<k}^{p} |\lambda_j - \lambda_k| \\ &= \text{función unicamente de } \Lambda. \end{aligned} \tag{57}$$

Density of Eigenvalues of a Wigner Matrix: Coulomb gas

$$\rho(\lambda_1,\ldots,\lambda_p) = \frac{1}{(2\pi)^{p/2}\pi^{\frac{p^2-p}{4}}} \exp\left(-\frac{1}{2}\sum_{i=1}^{p}\lambda_i^2\right)\prod_{j<k}^{p}|\lambda_j-\lambda_k|\int_{\mathbb{V}_p} dO \tag{58}$$

and all that remains for us is to compute the integral over the elements of $\{O\}$.

However, this integral is not trivial, it requires integration over the Stiefel manifold denoted as $\mathbb{V}_p$ of dimension $p(p-1)/2$, where dO is a volume element within this variety.

In order not to make life more difficult, you have to[4]

$$Vol(\mathbb{V}_p) = \int_{\mathbb{V}_p} dO = \frac{2^p\pi^{p^2/2}}{\Gamma_p(p/2)}, \tag{59}$$

[4] Muirhead, Wiley, 2005

Density of Eigenvalues of a Wigner Matrix: Coulomb gas

where

$$\Gamma_m(a) = \pi^{m(m-1)/4} \prod_{i=1}^{m} \Gamma(a - (i-1)/2), \tag{60}$$

and we call

$$\mathcal{D}O = \frac{dO}{Vol(\mathbb{V}_p)} \tag{61}$$

The Haar measure over $O(p)$.

This measure is invariant under orthogonal conjugation and defines a probability space over the orthogonal matrices.

Density of Eigenvalues of a Wigner Matrix: Coulomb gas

Before having the final expression, two aspects must be considered in the derivation:

1. The above expression was obtained for a specific sequence or ordering of the eigenvalues, so the factor $p!$ must be added to account for all possible permutations.

2. The change of variables $H \to O\Lambda O'$ must be one-one (bijective). But the eigenvectors are defined except for one phase. To guarantee the uniqueness of the eigendecomposition, it is enough to fix the sign of the first row of the matrix O. This reduces the volume of the integral

$$\int_{\mathbb{V}_p} dO, \tag{62}$$

by a factor of 2^p in the orthogonal case.

Density of Eigenvalues of a Wigner Matrix: Coulomb gas

Taking into account all of the above, an expression is reached for the jpdf of the eigenvalues of the symmetric matrices belonging to the *GOE*

$$\rho(\lambda_1, \ldots, \lambda_p) = \frac{1}{Z} \exp\left(-\frac{1}{2}\sum_{i=1}^{p} \lambda_i^2\right) \prod_{j<k}^{p} |\lambda_j - \lambda_k|, \tag{63}$$

where

$$Z = \frac{(2\pi)^{p/2} \pi^{\frac{p^2-p}{4}} \Gamma_p(p/2)}{p!} \tag{64}$$

Is normalization constant

Density of Eigenvalues of a Wigner Matrix: Coulomb gas

- The term

$$\exp\left(-\frac{1}{2}\sum_{i=1}^{p}\lambda_i^2\right) \tag{65}$$

  tells us that eigenvalues don't like to be too far from the origin: *confinement*

- On the other hand, the term

$$\prod_{j<k}^{p}|\lambda_j-\lambda_k| \tag{66}$$

  It tells us that the eigenvalues are not independent (they are correlated): *repulsion*

This game between CONFINEMENT and REPULSION is the physical mechanism of many results in random matrices (at least from the interpretation of the physical sciences).

Density of Eigenvalues of a Wigner Matrix: Coulomb gas

The jpdf for Gaussian ensembles (in general) is given by

$$\rho(\lambda_1, \ldots, \lambda_p) = \frac{1}{Z_\beta} \exp\left(-\frac{1}{2}\sum_{i=1}^{p} \lambda_i^2\right) \prod_{j<k}^{p} |\lambda_j - \lambda_k|^\beta, \tag{67}$$

para $\beta = 1, 2, 4$.

If we rescale the eigenvalues: $\lambda_i \to \lambda_i \sqrt{\beta p}$, we get

$$\begin{aligned} &\int \prod_{j=1}^{p} d\lambda_j \rho(\lambda_1, \ldots, \lambda_p) = \\ &\frac{1}{Z_\beta} \int_{\mathbb{R}^p} (\beta p)^{p/2} \prod_{j=1}^{p} d\lambda_j e^{\left(-\frac{\beta p}{2} \sum_{i=1}^{p} \lambda_i^2\right)} \prod_{j<k}^{p} |\sqrt{\beta p}(\lambda_j - \lambda_k)|^\beta = 1 \end{aligned} \tag{68}$$

Density of Eigenvalues of a Wigner Matrix: Coulomb gas

In this way

$$
\begin{aligned}
Z_\beta &= \left(\sqrt{\beta p}\right)^{p+\beta p(p-1)/2} \int_{\mathbb{R}^p} \prod_{j=1}^{p} d\lambda_j \exp\left(-\frac{\beta p}{2}\sum_{i=1}^{p}\lambda_i^2 + \ln\left(\prod_{j<k}^{p}|\lambda_j-\lambda_k|^\beta\right)\right) \\
&= C_\beta \int_{\mathbb{R}^p} \prod_{j=1}^{p} d\lambda_j \exp\left[-\beta p\left(\frac{1}{2}\sum_{i=1}^{p}\lambda_i^2 - \frac{1}{2p}\sum_{i\neq j}^{p}\ln|\lambda_i-\lambda_j|\right)\right] \\
&= C_\beta \int_{\mathbb{R}^p} \prod_{j=1}^{p} d\lambda_j e^{-\beta p\mathcal{V}[\lambda]}
\end{aligned}
\tag{69}
$$

where

$$
\mathcal{V}[\lambda] = \frac{1}{2}\sum_{i}\lambda_i^2 - \frac{1}{2p}\sum_{i\neq j}\ln|\lambda_i-\lambda_j| \tag{70}
$$

it is associated in physics with the term of potential energy of a gas charged particles type Gibbs-Boltzmann.

Density of Eigenvalues of a Wigner Matrix: Coulomb gas

We can solve the above equation using the saddle point method

In its simplest form ($x \in \mathbb{R}$) the saddle point method consists of approximating integrals of the form

$$I = \int_{-\infty}^{\infty} dx e^{-f(x)} \tag{71}$$

The idea is to keep the contribution of $f(x)$ when it is at its minimum value.

Let's say $f(x)$ is minimum at x_0, so we can approximate $f(x)$ as

$$f(x) \approx f(x_0) + \frac{1}{2}(x - x_0)^2 f''(x_0) + \dots \tag{72}$$

si $x \in (x_0 - \epsilon, x_0 + \epsilon)$

Density of Eigenvalues of a Wigner Matrix: Coulomb gas

Under this approximation

$$\begin{aligned} I &\approx \int_{-\infty}^{\infty} dx e^{-f(x_0)-\frac{1}{2}(x-x_0)^2 f''(x_0)} \\ &= e^{-f(x_0)} \int_{-\infty}^{\infty} dx e^{-\frac{1}{2}(x-x_0)^2 f''(x_0)} \\ &= e^{-f(x_0)} \sqrt{\frac{2\pi}{f''(x_0)}} \end{aligned} \tag{73}$$

In our problem, the condition $f'(x_0) = 0$ allows us to derive Wigner's semicircle law.

The 2nd order approximation is sufficient when $p \to \infty$.

Density of Eigenvalues of a Wigner Matrix: Coulomb gas

To find the value of the integral Z_β we have the condition (given by the saddle point method)

$$\frac{\partial \mathcal{V}[\lambda]}{\partial \lambda_i} = 0 \tag{74}$$

this implies ([▸ after one more hour of algebra]) ...

$$g^2(z) - 2zg(z) + 2 = 0 \tag{75}$$

which is the same expression given by the self-consistent method (with $\sigma^2 = 1/2$).

Consequently

$$\rho(\lambda) = \frac{1}{\pi}\sqrt{2 - \lambda^2} \tag{76}$$

Semicircle law

Theorem

Suppose $H = (A + A^{'})/2$ *is a* $N \times N$ *matrix, where the elements* $A_{i,j}$ *are real random variables i.i.d.* $\sim \mathcal{N}(0, 1)$. *Then, when* $N \to \infty$ *the spectral density function of* H *converges (a.s.) to Wigner's semicircle law:*

$$\rho(x) = \frac{1}{\pi}\sqrt{2 - x^2} \tag{77}$$

Density of Eigenvalues of a Wishart Matrix

Equivalently, a quadratic equation in g(z) can be found for the Wishart ensemble

$$\frac{1}{g(z)} = z - 1 + q - qzg(z) \tag{78}$$

from where we get the density of eigenvalues

$$\rho(\lambda) = \frac{\sqrt{(\lambda_{max} - \lambda)(\lambda - \lambda_{min})}}{2\pi q \lambda}, \tag{79}$$

where

$$\lambda_{min}^{max} = (1 \pm \sqrt{q})^2, \quad q = p/n \tag{80}$$

Marchenko-Pastur law

Theorem

Let X be a matrix of dimensions $p \times n$, where the elements $X_{i,j}$ are i.i.d. $\mathcal{N}(0,1)$. Then, when $p, n \to \infty$, such that $\frac{p}{n} \to q \in (0, \infty)$, the eingenvalue spectral density of the Wishart matrix $W = \frac{1}{T} XX'$ converges (a.s.) to the Marcenko-Pastur law [a]

$$\rho(\lambda) = \frac{\sqrt{(\lambda_{max} - \lambda)(\lambda - \lambda_{min})}}{2\pi q \lambda}, \tag{81}$$

where

$$\lambda_{min}^{max} = (1 \pm \sqrt{q})^2. \tag{82}$$

[a] Marchenko VA, Pastur LA. Distribution of eigenvalues for some sets of random matrices. Sb. Math. 1967;114(4):507-36.

Example: semicircle law

We generate $m = 10000$ matrices for $N = 8$, and the normalized histograms of the full sample of $m \times N$ eigenvalues are computed.

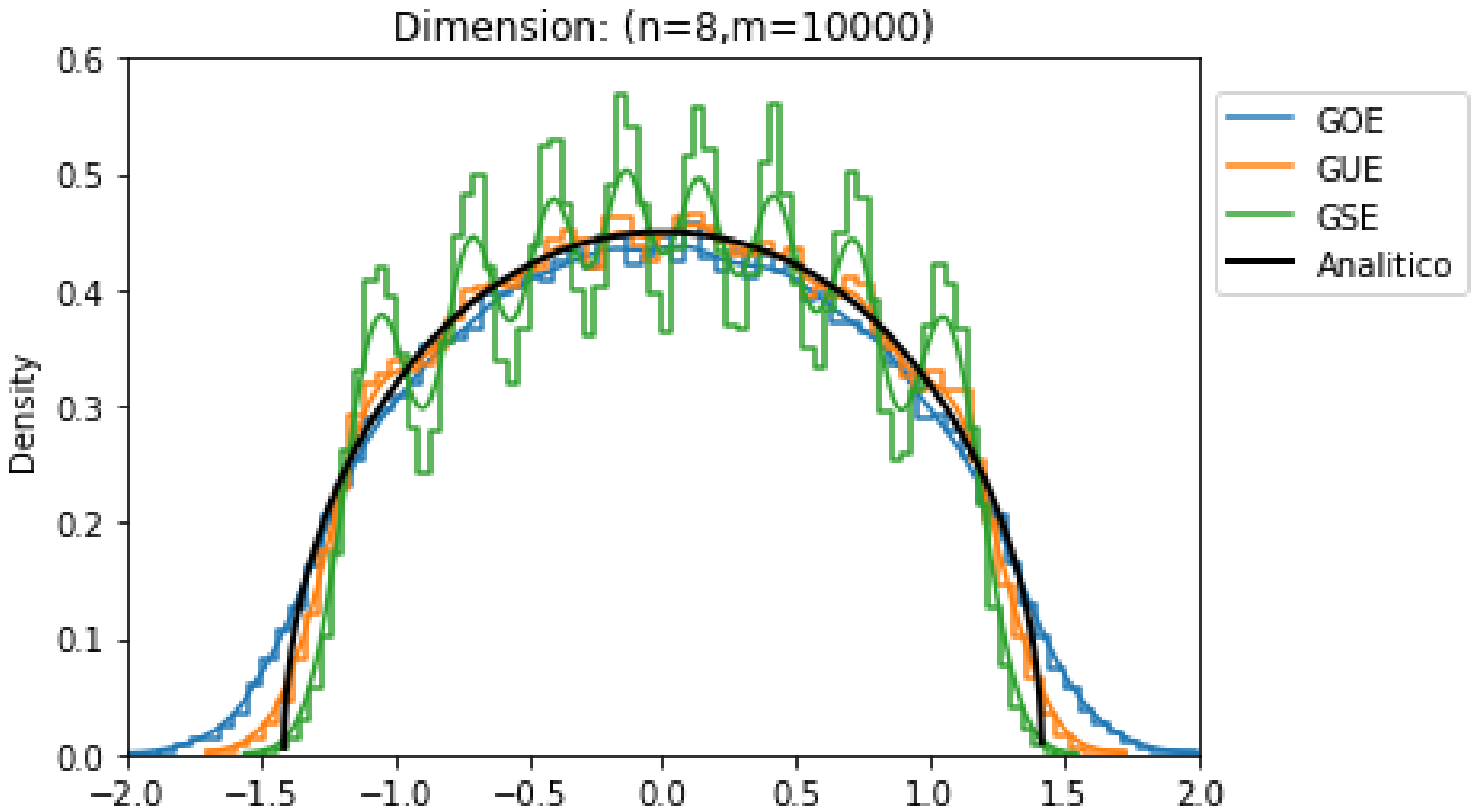
To obtain the normalized histograms, they must be scaled by the factor $\frac{1}{\sqrt{\beta N}}$, where $\beta = 1, 2, 4$, for the case *GOE*, *GUE*, and *GSE*; respectively.

The simulations are compared with the analytical result $\rho(x) = \frac{1}{\pi}\sqrt{2 - x^2}$, called *Wigner's semicircle law*. The following analytical results are obtained for the bounds of the support of the distributions

- $\pm\sqrt{2N}$ (GOE)
- $\pm\sqrt{4N}$ (GUE)
- $\pm\sqrt{8N}$ (GSE)

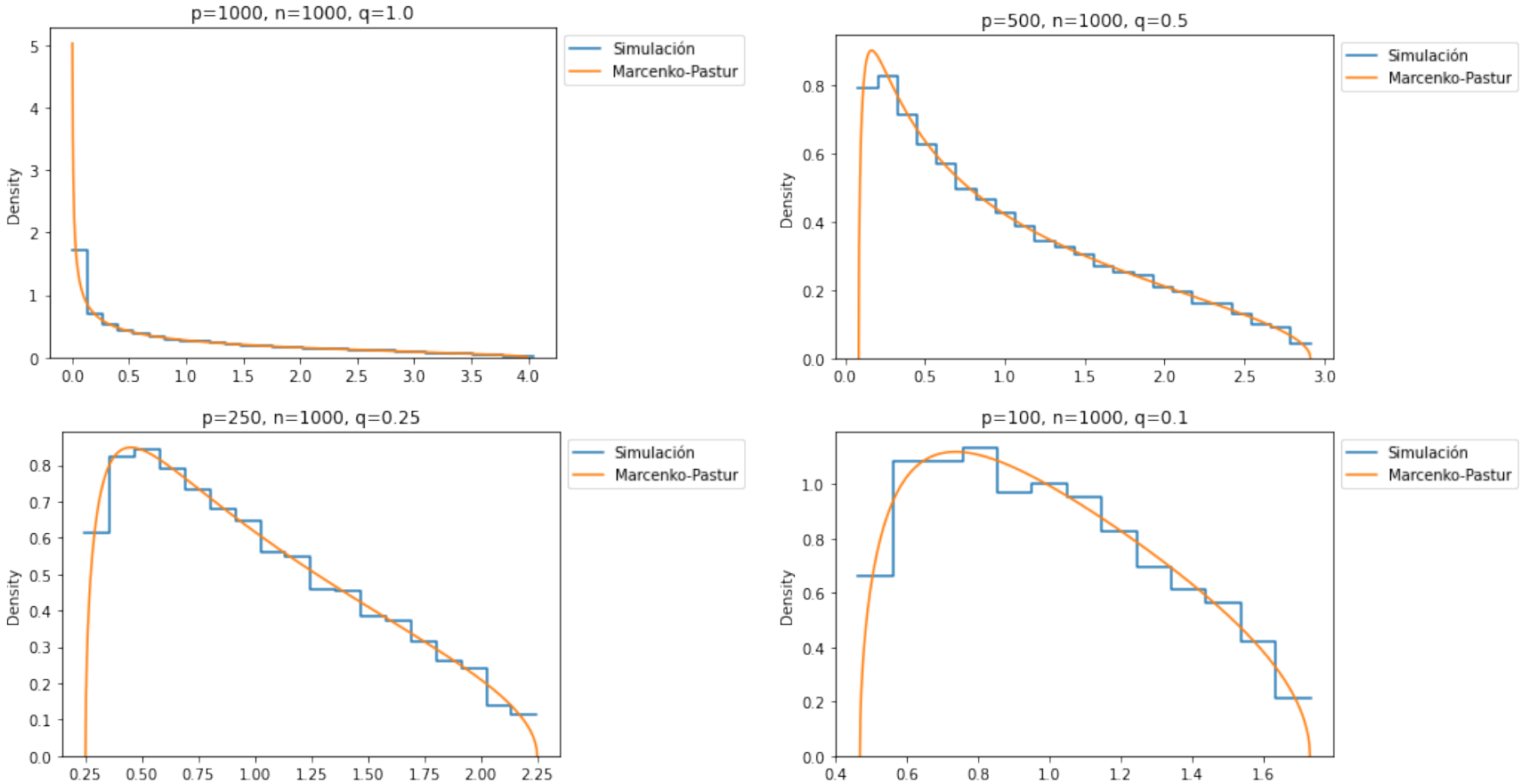
Example: semicircle law

Notebook



Example: Marchenko-Pastur law

Notebook



Exercises and homework

- Exercise: Reproduce the density for the GUE and GSE cases
- Exercise: The Marchenko-Pastur law is valid for $N(0, \sigma^2)$ processes. In this case the density is divided by σ^2 and the bounds are extended by the same factor. Simulate processes with $\sigma^2 = 0.1, 2, 10$. What do you observe?
- Homework: read *Krbálek, M., & Seba, P. (2000). The statistical properties of the city transport in Cuernavaca (Mexico) and random matrix ensembles. Journal of Physics A: Mathematical and General, 33(26), L229.*
- Material: `github.com/agarciam/curso-de-matrices-aleatorias-y-portafolios`

Wigner's Surmise

Consider the matrix

$$H_{s_{2\times}} = \begin{pmatrix} x_1 & x_3 \\ x_3 & x_2 \end{pmatrix} \tag{83}$$

such that $H_s \in GOE$, y $x_1, x_2 \sim N(0,1)$, $x_3 \sim N(0,1/2)$

What is the pdf $p(s)$ of the spacing $s = \lambda_2 - \lambda_1$ between its two eigenvalues ($\lambda_2 > \lambda_1$)?

▸ Blackboard

Wigner's Surmise

It is convenient to rescale this result to have a density (pdf) and define $\bar{p}(s) = \langle s \rangle p(\langle s \rangle s)$, where

$$\langle s \rangle = \int_0^\infty s p(s) ds \tag{84}$$

This last expression is known as *average level spacing*.

For the GOE case, we have

$$\bar{p}(s) = \frac{\pi s}{2} e^{\frac{-\pi s^2}{4}} \tag{85}$$

which is the well-known form of the Wigner surmise

In general, this result tells us that the probability of finding two very close eigenvalues ($s \to 0$) is very small, as if each eigenvalue feel the presence of the other and tried to avoid it.

Exercises

- Plot the contiguous spacing distribution (ordering the eigenvalues) for an ensemble of $m = 1000$ symmetric matrices of dimension 100×100 ($H_{s_{100\times 100}}$).

- Overlap the simulation with the result of the Wigner conjecture for a *GOE* matrix of 2×2.

- Another application of the Wigner's Surmise for *GUE* is Kikkawa, A. (2018). Random matrix analysis for gene interaction networks in cancer cells. Scientific reports, 8(1), 1-12.

Portfolio Theory

The implementation of portfolio theory is not simple. It requires the estimation of the covariance matrix and the forecast of the return on assets.

One problem is that optimal portfolios are very sensitive to inputs.

Some stability problems make the practice of investment portfolio optimization less attractive than the theory.

In this case, regularization techniques can be used to alleviate these problems.

This approach is widely supported by Ledoit and Wolf (2003) [5], who propose combining different covariance matrix estimators to stabilize the solution.

[5] Improved Estimation of the Covari- ance Matrix of Stock Returns With an Application to Portfolio Selection, Journal of Empirical Finance, 10(5), pp. 603-621.

Portfolio Theory

Nowadays, one of the most promising approaches consists in interpreting the optimal portfolios as the solution of a linear regression problem and using the *Lasso* or *Ridge* penalty.

However, regularization is not enough to obtain satisfactory solutions, so practitioners introduce some constraints in the optimization problem.

These restrictions can be interpreted as a shrinkage method.

By imposing weight restrictions, the portfolio manager implicitly changes the covariance matrix.

This approach is then equivalent to having some *a priori* assumptions and is therefore related to the model of Black and Litterman (1992) [6].

[6] Black, F., & Litterman, R. (1992). Global portfolio optimization. Financial analysts journal, 48(5), 28-43.

Efficient frontier

- In 1952 Markowitz introduced the concept of the efficient frontier. This was the first mathematical formulation about portfolio optimization.
- The idea behind it is that investors see the expected return as a good thing, while the variance of returns as a bad thing.
- What Markowitz showed is that there is a set of optimal portfolios, which maximize the expected return, given a value at risk. This was what he named *efficient frontier*.

Basic definitions

Consider a universe of p assets, and denote as $s_{i,t}$ the price of asset i at time t. Thus, the return or performance of the asset i over a unit of time (days) is defined as

$$r_{i,t} = \frac{s_{i,t} - s_{i,t-1}}{s_{i,t-1}} \tag{86}$$

where $i = 1, \ldots, p$.

and the vector of portfolio weights as

$$\mathrm{w} = \{w_1, \ldots, w_p\} \tag{87}$$

where w is the amount of money (dollars) invested in the asset i, which can be positive or negative.

Basic definitions

The total amount of capital invested will be expressed as $\mathcal{C}$, where usually $\mathcal{C} = \sum_i w_i$.

More generally, the following scenarios exist:

- $\sum_i w_i > \mathcal{C}$: Borrow and pay a risk-free rate (interest) r_0 on the credit
- $\sum_i w_i < \mathcal{C}$: An amount of capital is reserved to invest at the risk-free rate (bonds) r_0
- In general, the rates are different in each case, but we will assume them to be the same for simplicity

Basic definitions

Thus, the return on the portfolio at time t is given

$$R_t = \sum_{i=1}^{p} w_i r_{i,t} + (\mathcal{C} - \sum_i w_i) r_0, \tag{88}$$

and is defined as *excess return* (over the risk-free rate) the expression

$$R_t - \mathcal{C} r_0 = \sum_i w_i (r_{i,t} - r_0) = w' e, \tag{89}$$

where e is the excess return vector (we omit the time from here on for simplicity in notation).
The expected return (excess return) of the portfolio is defined as

$$\mathcal{G} = \mathbb{E}(w' e) = w' \mathbb{E}(e) = w' g, \tag{90}$$

where g is the vector of expected profits (excess return).

Basic definitions

Portfolio risk is traditionally defined as the variance of returns. Here it will be defined in relation to the excess return

$$\begin{aligned}\mathcal{R}^2 &= Var[R] = Var[w'e] = \mathbb{E}\left[(w'e - w'g)(w'e - w'g)'\right] \\ &= \mathbb{E}\left[\mathrm{w}'(\mathrm{e} - \mathrm{g})(\mathrm{e} - \mathrm{g})'\mathrm{w}\right] \\ &\mathrm{w}'\mathbb{E}\left[(\mathrm{e} - \mathrm{g})(\mathrm{e} - \mathrm{g})'\right]\mathrm{w} \\ &= \mathrm{w}'\Sigma\mathrm{w}\end{aligned} \tag{91}$$

Optimal portfolios

Under the Markowitz approach, the idea is to solve the following quadratic optimization problem

$$\begin{cases} \underset{\mathrm{w}\in\mathbb{R}^p}{min}\ \frac{1}{2}\mathrm{w}'\Sigma\mathrm{w} \\ \text{s.t. } \mathrm{w}'\mathrm{g} \geq \mathcal{G} \end{cases} \tag{92}$$

This mathematical problem can be easily solved by introducing a Lagrangian multiplier γ to rewrite the system as an unconstrained optimization problem

$$\underset{\mathrm{w}\in\mathbb{R}^p}{min}\ \frac{1}{2}\mathrm{w}'\Sigma\mathrm{w} - \gamma\mathrm{w}'\mathrm{g} \tag{93}$$

Solution

Writing the above with a Lagrange function

$$\mathcal{L}(\mathrm{w}) = \frac{1}{2}\mathrm{w}'\Sigma\mathrm{w} - \gamma\mathrm{w}'\mathrm{g} \tag{94}$$

Hence it must be fulfilled

$$\frac{d\mathcal{L}}{d\mathrm{w}} = \Sigma\mathrm{w} - \gamma\mathrm{g} = 0 \tag{95}$$

Assuming that the inverse of Σ exists, the optimal portfolio is found to be given by

$$\hat{\mathrm{w}} = \gamma\Sigma^{-1}\mathrm{g} \tag{96}$$

On the other hand, by substituting (96) in (90) the expression for γ is found

$$\gamma = \frac{\mathcal{G}}{\mathrm{g}'\Sigma^{-1}\mathrm{g}} \tag{97}$$

In this way, the optimal solution is given by

$$\hat{\mathrm{w}} = \mathcal{G}\frac{\Sigma^{-1}\mathrm{g}}{\mathrm{g}'\Sigma^{-1}\mathrm{g}} \tag{98}$$

This requires knowledge of Σ and g, which is not known *a priori*.

We will leave the job of creating expectations about the future behavior of returns to the investor or financial analyst.

Here we will assume that g is given, and estimate Σ in the high-dimensional regime, i.e., when $p, n \to \infty$ where $q = p/n$ constant.

Which is natural to consider in the context of *"Big Data"*, and particularly if we seek to avoid the non-stationary effect of the financial time series

True risk

What is the minimum risk associated with this asset allocation when risk is measured as the variance of portfolio returns?

To answer this, let's substitute (98) in (91) to obtain

$$\mathcal{R}^2_{true} = \frac{\mathcal{G}^2}{\mathrm{g}'\Sigma^{-1}\mathrm{g}} \tag{99}$$

However, the optimal strategy is not achievable in practice since Σ is not known in principle.

What can we do then? How different is the *realized risk* or out-sample from the *true risk* of the optimal model?

Predicted risk and realized risk.

A naive approximation is to apply the empirical covariance matrix E instead of Σ.

However, we can anticipate that this strategy will suffer from a strong bias as long as n is not too large compared to p.

Despite the above, it is easy to see that the optimal investment weights are given in an analogous way in this case by

$$\mathrm{w}_E = \mathcal{G}\frac{\mathrm{E}^{-1}\mathrm{g}}{\mathrm{g}'\mathrm{E}^{-1}\mathrm{g}}, \tag{100}$$

Therefore, the minimum risk associated with this portfolio is given by

$$\mathcal{R}^2_{in} = \frac{\mathcal{G}^2}{\mathrm{g}'\mathrm{E}^{-1}\mathrm{g}}, \tag{101}$$

which is known as *predicted risk* or in-sample risk, because the available information is used.

Predicted risk and realized risk.

Remembering the Jensen's inequality we have that

$$\mathbb{E}[\mathrm{g}' E^{-1} \mathrm{g}] \geq \mathrm{g}' \mathbb{E}[E^{-1}] \mathrm{g} = \mathrm{g}' \Sigma^{-1} \mathrm{g} \tag{102}$$

- Since E is positive semidefinite in general its inverse is also positive and implies convexity.
- Thus, the first inequality holds for fixed vectors g fixed.
- The equality holds since E is an unbiased estimator of Σ.

Thus $\mathcal{R}^2_{in} \leq \mathcal{R}^2_{true}$, and the optimal portfolio w_E underestimates the true risk of the investment $\mathcal{R}^2_{true}$.

Predicted risk and realized risk.

The realized risk is known as out-sample since it uses new data once the weights have been optimized.

Let us denote by $\tilde{\mathrm{E}}$ the out-sample covariance matrix. In this case, the associated risk is given by the expression

$$\mathcal{R}^2_{out} = \mathrm{w}'_E \tilde{\mathrm{E}} \mathrm{w}_E \tag{103}$$

It can be assumed that the in-sample noise (contained in w_E) is independent from the out-of-sample noise, so w_E and $\tilde{E}$ will not be correlated [7]. We get for large p

$$\mathcal{R}^2_{out} = \mathrm{w}'_E \tilde{\mathrm{E}} \mathrm{w}_E \approx \mathrm{w}'_E \Sigma \mathrm{w}_E = \frac{\mathcal{G}^2 \mathrm{g}' \mathrm{E}^{-1} \Sigma \mathrm{E}^{-1} \mathrm{g}}{(\mathrm{g}' \mathrm{E}^{-1} \mathrm{g})^2} \tag{104}$$

[7] S. Pafka, I. Kondor, Noisy covariance matrices and portfolio optimization II, Physica A 319 (2003) 487–494.

Predicted risk and realized risk.

Furthermore, given the optimality of $\tilde{w} = w_\Sigma$ we know that

$$w_\Sigma' \Sigma w_\Sigma \leq w_E' \Sigma w_E \tag{105}$$

Hence we get the following general inequality

$$\mathcal{R}^2_{in} \leq \mathcal{R}^2_{true} \leq \mathcal{R}^2_{out} \tag{106}$$

In short, w_E is overly optimistic and can lead to disastrous results in practice.

This conclusion is valid even for other risk measures assuming Gaussianity

Efficient frontier inequality

Notebook

The following figure shows the optimal frontier for the true, in-sample, and out-sample process in the case of a Wishart matrix W_q.

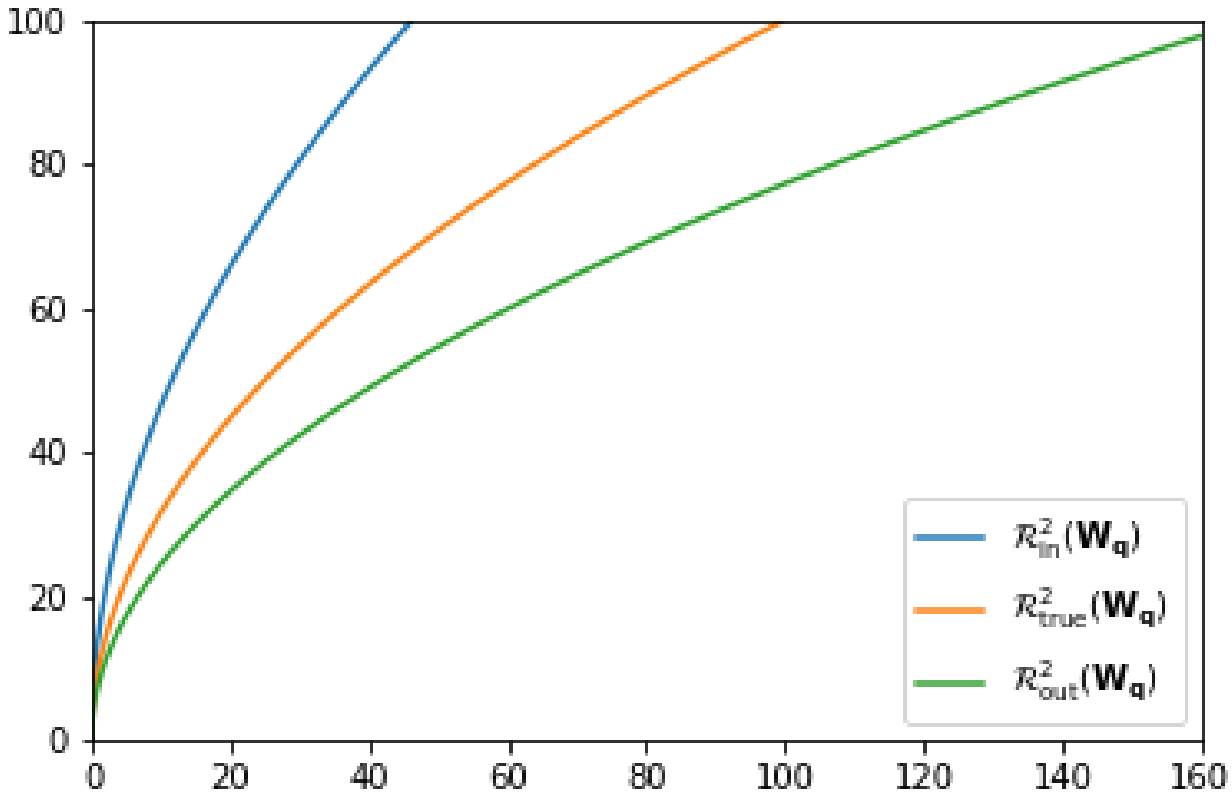


Figure (3) $X \sim (0, \sigma^2 I)$, $q = 1/2$, $p = 100$, $\sigma^2 = 0.2$ y g $= 1$

Exercise

- What is the effect of varying the parameters σ, p, q, and the vector g?
- Material: `github.com/agarciam/curso-de-matrices-aleatorias-y-portafolios`

RMT denoising technique

Procedure to filter covariance matrix proposed by Bouchaud et. al. (2002):

- $\hat{\lambda}_k^B = \begin{cases} \bar{\lambda} & if \quad \lambda_k < (1+\sqrt{q})^2 \\ \lambda_k & \text{otro caso} \end{cases}$
- $Q^B = V\hat{\Lambda}^B V'$
- $\Xi_{ij}^B = \frac{q_{ij}^B}{\sqrt{q_{ii}^B q_{jj}^B}}$

Efficient frontier inequality applying RMT filter

Notebook

`github.com/agarciam/curso-de-matrices-aleatorias-y-portafolios`

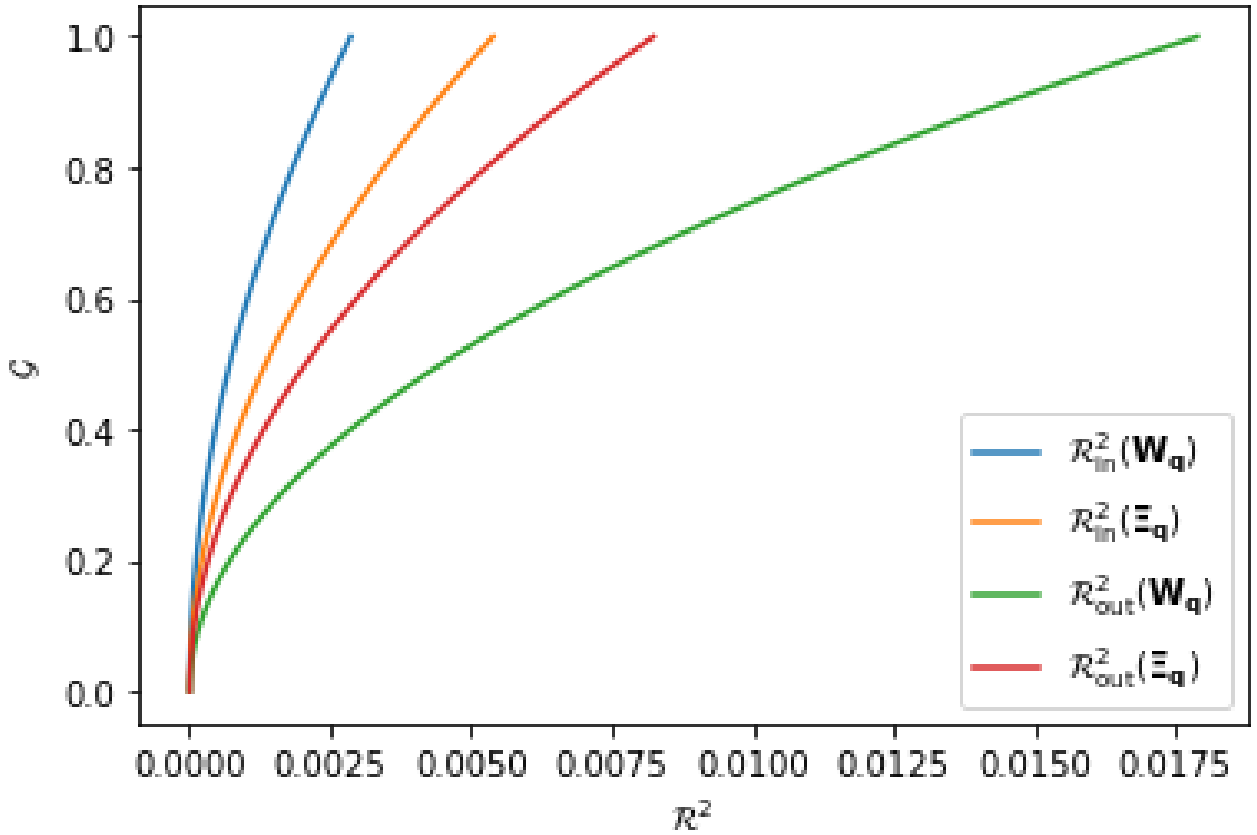


Figure (4) System with 10 artificial signals, $q = 1/2$, $p = 100$

References

1. Potters, M., & Bouchaud, J. P. (2020). A First Course in Random Matrix Theory: For Physicists, Engineers and Data Scientists. Cambridge University Press.
2. Livan, G., Novaes, M., & Vivo, P. Introduction to random matrices theory and practice. Springer (2018).
3. Bun, J., Bouchaud, J. P., & Potters, M. (2017). Cleaning large correlation matrices: tools from random matrix theory. Physics Reports, 666, 1-109.
4. Mehta, M. L. (2004). Random matrices. Elsevier.
5. Mingo, J. A., & Speicher, R. (2017). Free probability and random matrices (Vol. 35). New York: Springer.