Large Random Matrices and Applications to Statistical Signal Processing

Jamal Najim
najim@univ-mlv.fr

CNRS & Université Paris Est

Summer school - June 2016 - Télécom Paristech

Large Random Matrices

Random matrices

It is a $N \times N$ matrix

$$\mathbf{Y}_N = \begin{bmatrix} Y_{11} & \cdots & Y_{1N} \\ \vdots & & \vdots \\ Y_{N1} & \cdots & Y_{NN} \end{bmatrix}$$

whose entries $(Y_{ij};\ 1 \le i,j \le N)$ are random variables.

Matrix features

Of interest are the following quantities

- $\mathbf{Y}_N$'s **spectrum** $(\lambda_i,\ 1 \le i \le N)$ in particular $\lambda_{\min}$ and $\lambda_{\max}$.
- linear statistics

$$\mathrm{Trace}\, f(\mathbf{Y}_N) = \sum_{i=1}^{N} f(\lambda_i)$$

- eigenvectors, etc.

Asymptotic regime

Often, the description of the previous features **takes a simplified form** as

$$\boxed{N \to \infty}$$

Moreover this regime is **of interest in many applications**.

Matrix model

Let $\mathbf{X}_N = (X_{ij})$ a symmetric $N \times N$ matrix with i.i.d. entries **on and above** the diagonal with

$$\mathbb{E} X_{ij} = 0 \text{ and } \mathbb{E} |X_{ij}|^2 = 1$$

and $X_{ij} = X_{ji}$ (for symmetry).

- consider the spectrum of **Wigner matrix** $\boxed{\mathbf{Y}_N = \frac{\mathbf{X}_N}{\sqrt{N}}}$

Large Random Matrices: Wigner Matrices

Matrix model

Let $\mathbf{X}_N = (X_{ij})$ a symmetric $N \times N$ matrix with i.i.d. entries **on and above** the diagonal with

$$\mathbb{E}X_{ij} = 0 \text{ and } \mathbb{E}\,|X_{ij}|^2 = 1$$

and $X_{ij} = X_{ji}$ (for symmetry).

- consider the spectrum of **Wigner matrix** $\boxed{\mathbf{Y}_N = \frac{\mathbf{X}_N}{\sqrt{N}}}$

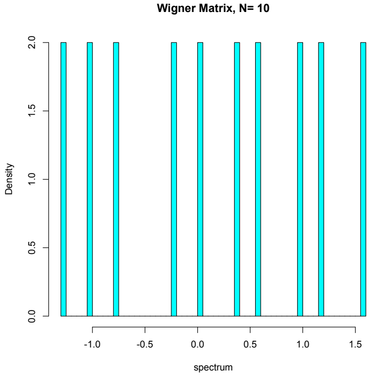

Figure: Histogram of the eigenvalues of $\mathbf{Y}_N$

Large Random Matrices: Wigner Matrices

Matrix model

Let $\mathbf{X}_N = (X_{ij})$ a symmetric $N \times N$ matrix with i.i.d. entries **on and above** the diagonal with

$$\mathbb{E}X_{ij} = 0 \text{ and } \mathbb{E}\,|X_{ij}|^2 = 1$$

and $X_{ij} = X_{ji}$ (for symmetry).

- consider the spectrum of **Wigner matrix** $\boxed{\mathbf{Y}_N = \frac{\mathbf{X}_N}{\sqrt{N}}}$

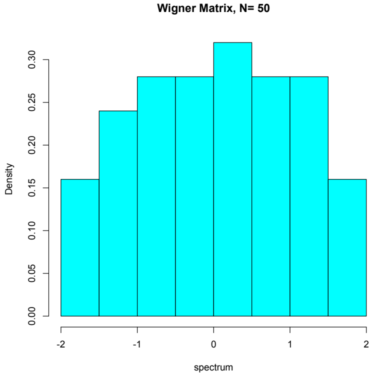

Figure: Histogram of the eigenvalues of $\mathbf{Y}_N$

Large Random Matrices: Wigner Matrices

Matrix model

Let $\mathbf{X}_N = (X_{ij})$ a symmetric $N \times N$ matrix with i.i.d. entries **on and above** the diagonal with

$$\mathbb{E}X_{ij} = 0 \text{ and } \mathbb{E}\,|X_{ij}|^2 = 1$$

and $X_{ij} = X_{ji}$ (for symmetry).

- consider the spectrum of **Wigner matrix** $\boxed{\mathbf{Y}_N = \frac{\mathbf{X}_N}{\sqrt{N}}}$

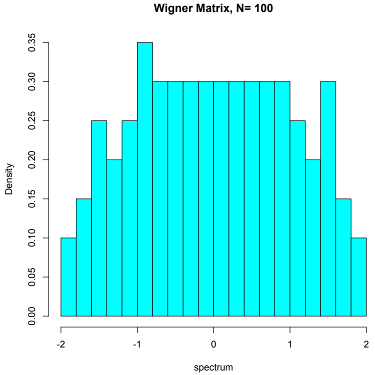

Figure: Histogram of the eigenvalues of $\mathbf{Y}_N$

Large Random Matrices: Wigner Matrices

Matrix model

Let $\mathbf{X}_N = (X_{ij})$ a symmetric $N \times N$ matrix with i.i.d. entries **on and above** the diagonal with

$$\mathbb{E}X_{ij} = 0 \text{ and } \mathbb{E}\left|X_{ij}\right|^2 = 1$$

and $X_{ij} = X_{ji}$ (for symmetry).

- consider the spectrum of **Wigner matrix** $\boxed{\mathbf{Y}_N = \frac{\mathbf{X}_N}{\sqrt{N}}}$

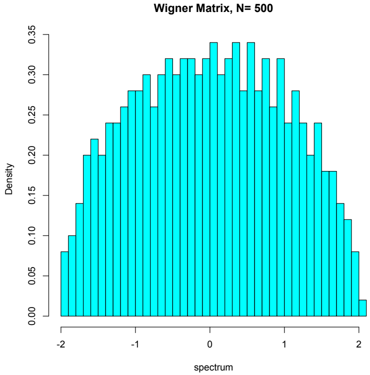

Figure: Histogram of the eigenvalues of $\mathbf{Y}_N$

Large Random Matrices: Wigner Matrices

Matrix model

Let $\mathbf{X}_N = (X_{ij})$ a symmetric $N \times N$ matrix with i.i.d. entries **on and above** the diagonal with

$$\mathbb{E} X_{ij} = 0 \text{ and } \mathbb{E}\left|X_{ij}\right|^2 = 1$$

and $X_{ij} = X_{ji}$ (for symmetry).

- consider the spectrum of **Wigner matrix** $\boxed{\mathbf{Y}_N = \frac{\mathbf{X}_N}{\sqrt{N}}}$

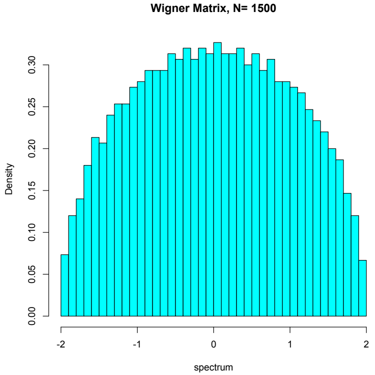

Figure: Histogram of the eigenvalues of $\mathbf{Y}_N$

Large Random Matrices: Wigner Matrices

Matrix model

Let $\mathbf{X}_N = (X_{ij})$ a symmetric $N \times N$ matrix with i.i.d. entries **on and above** the diagonal with

$$\mathbb{E}X_{ij} = 0 \text{ and } \mathbb{E}\,|X_{ij}|^2 = 1$$

and $X_{ij} = X_{ji}$ (for symmetry).

- consider the spectrum of **Wigner matrix** $\boxed{\mathbf{Y}_N = \frac{\mathbf{X}_N}{\sqrt{N}}}$

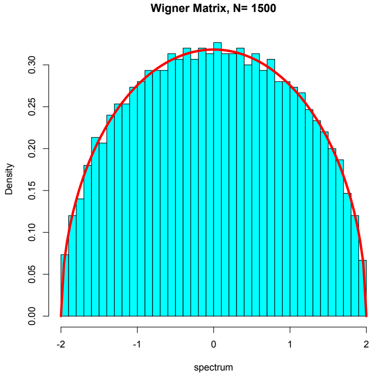

Figure: The semi-circular distribution (in red) with density $x \mapsto \frac{\sqrt{4-x^2}}{2\pi}$

Wigner's theorem (1948)

"The histogram of a Wigner matrix converges to the **semi-circular distribution**"

Large Covariance Matrices

Matrix model

Let $\mathbf{X}_N$ be a $N \times n$ matrix with i.i.d. entries

$$\mathbb{E}X_{ij} = 0 \ , \ \mathbb{E}|X_{ij}|^2 = 1$$

and consider the spectrum of $\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*$ in the regime where

$$N, n \to \infty \quad \text{and} \quad \frac{N}{n} \to c \in (0, \infty)$$

$\boxed{\text{dimensions of matrix } \mathbf{X}_N \text{ of the } \textbf{same order}}$

Large Covariance Matrices

Matrix model

Let $\mathbf{X}_N$ be a $N \times n$ matrix with i.i.d. entries

$$\mathbb{E}X_{ij} = 0 \ , \ \mathbb{E}|X_{ij}|^2 = 1$$

and consider the spectrum of $\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*$ in the regime where

$$N, n \to \infty \quad \text{and} \quad \frac{N}{n} \to c \in (0, \infty)$$

dimensions of matrix $\mathbf{X}_N$ of the **same order**

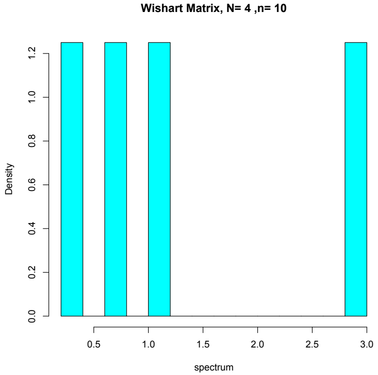

Figure: Spectrum's histogram - $\frac{N}{n} = 0.7$

Large Covariance Matrices

Matrix model

Let $\mathbf{X}_N$ be a $N \times n$ matrix with i.i.d. entries

$$\mathbb{E}X_{ij} = 0 \ , \ \mathbb{E}|X_{ij}|^2 = 1$$

and consider the spectrum of $\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*$ in the regime where

$$N, n \to \infty \quad \text{and} \quad \frac{N}{n} \to c \in (0, \infty)$$

$\boxed{\text{dimensions of matrix } \mathbf{X}_N \text{ of the } \textbf{same order}}$

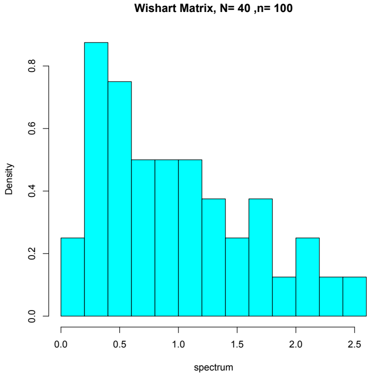

Figure: Spectrum's histogram - $\frac{N}{n} = 0.7$

Large Covariance Matrices

Matrix model

Let $\mathbf{X}_N$ be a $N \times n$ matrix with i.i.d. entries

$$\mathbb{E}X_{ij} = 0 \ , \ \mathbb{E}|X_{ij}|^2 = 1$$

and consider the spectrum of $\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*$ in the regime where

$$N, n \to \infty \quad \text{and} \quad \frac{N}{n} \to c \in (0, \infty)$$

$\boxed{\text{dimensions of matrix } \mathbf{X}_N \text{ of the } \textbf{same order}}$

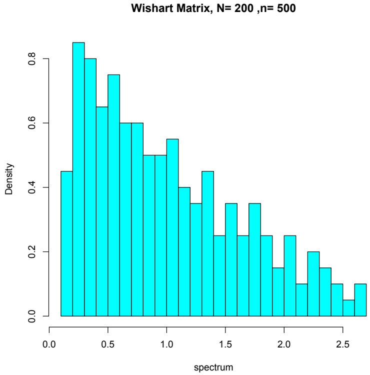

Figure: Spectrum's histogram - $\frac{N}{n} = 0.7$

Large Covariance Matrices

Matrix model

Let $\mathbf{X}_N$ be a $N \times n$ matrix with i.i.d. entries

$$\mathbb{E}X_{ij} = 0 \ , \ \mathbb{E}|X_{ij}|^2 = 1$$

and consider the spectrum of $\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*$ in the regime where

$$N, n \to \infty \quad \text{and} \quad \frac{N}{n} \to c \in (0, \infty)$$

$\boxed{\text{dimensions of matrix } \mathbf{X}_N \text{ of the } \textbf{same order}}$

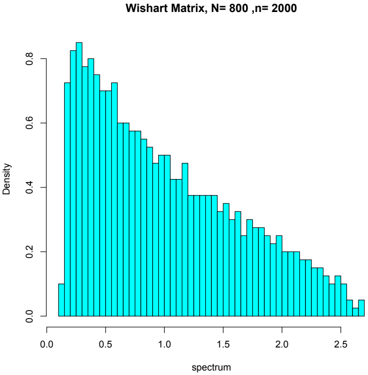

Figure: Spectrum's histogram - $\frac{N}{n} = 0.7$

Large Covariance Matrices

Matrix model

Let $\mathbf{X}_N$ be a $N \times n$ matrix with i.i.d. entries

$$\mathbb{E}X_{ij} = 0 \ , \ \mathbb{E}|X_{ij}|^2 = 1$$

and consider the spectrum of $\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*$ in the regime where

$$N, n \to \infty \quad \text{and} \quad \frac{N}{n} \to c \in (0, \infty)$$

$\boxed{\text{dimensions of matrix } \mathbf{X}_N \text{ of the } \textbf{same order}}$

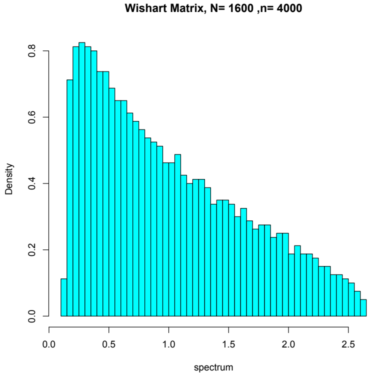

Figure: Spectrum's histogram - $\frac{N}{n} = 0.7$

Large Covariance Matrices : Marčenko-Pastur's theorem

Matrix model

Let $\mathbf{X}_N$ be a $N \times n$ matrix with i.i.d. entries

$$\mathbb{E}X_{ij} = 0 \ , \ \mathbb{E}|X_{ij}|^2 = 1$$

and consider the spectrum of $\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*$ in the regime where

$$N, n \to \infty \quad \text{and} \quad \frac{N}{n} \to c \in (0, \infty)$$

$\boxed{\text{dimensions of matrix } \mathbf{X}_N \text{ of the } \textbf{same order}}$

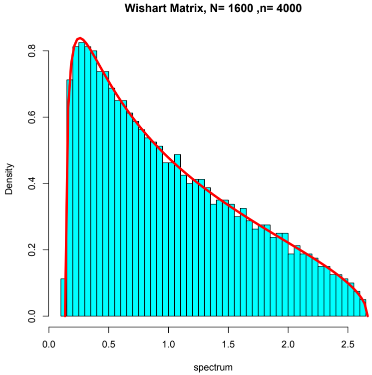

Figure: Marčenko-Pastur's distribution (in red)

Marčenko-Pastur's theorem (1967)

"The histogram of a **Large Covariance Matrix** converges to **Marčenko-Pastur distribution** with given parameter (here **0.7**)"

Matrix model

Let $\mathbf{X}_N$ be a $N \times N$ matrix with i.i.d. entries

$$\mathbb{E}X_{ij} = 0 \ , \ \mathbb{E}|X_{ij}|^2 = 1$$

and consider the spectrum of matrix $\mathbf{Y}_N = \frac{1}{\sqrt{N}}\mathbf{X}_N$ as $N \to \infty$

- In this case, the eigenvalues are **complex**!

Large Non-Hermitian Matrices

Matrix model

Let $\mathbf{X}_N$ be a $N \times N$ matrix with i.i.d. entries

$$\mathbb{E}X_{ij} = 0 \ , \ \mathbb{E}|X_{ij}|^2 = 1$$

and consider the spectrum of matrix $\mathbf{Y}_N = \frac{1}{\sqrt{N}}\mathbf{X}_N$ as $N \to \infty$

- In this case, the eigenvalues are **complex**!

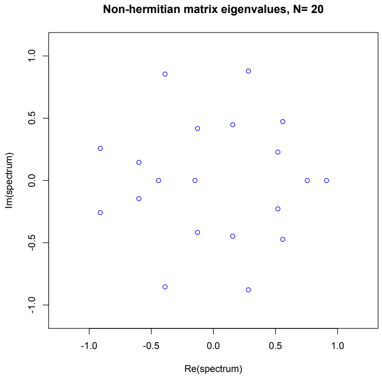

Figure: Distribution of $\mathbf{Y}_N$'s eigenvalues

Large Non-Hermitian Matrices

Matrix model

Let $\mathbf{X}_N$ be a $N \times N$ matrix with i.i.d. entries

$$\mathbb{E}X_{ij} = 0 \ , \ \mathbb{E}|X_{ij}|^2 = 1$$

and consider the spectrum of matrix $\mathbf{Y}_N = \frac{1}{\sqrt{N}}\mathbf{X}_N$ as $N \to \infty$

- In this case, the eigenvalues are **complex**!

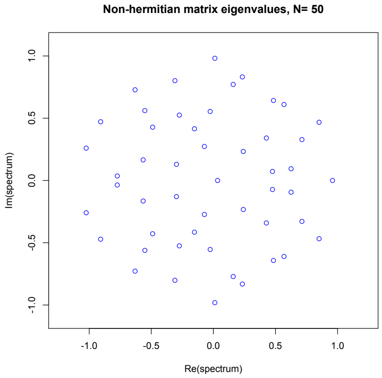

Figure: Distribution of $\mathbf{Y}_N$'s eigenvalues

Large Non-Hermitian Matrices

Matrix model

Let $\mathbf{X}_N$ be a $N \times N$ matrix with i.i.d. entries

$$\mathbb{E}X_{ij} = 0 \ , \ \mathbb{E}|X_{ij}|^2 = 1$$

and consider the spectrum of matrix $\mathbf{Y}_N = \frac{1}{\sqrt{N}}\mathbf{X}_N$ as $N \to \infty$

- In this case, the eigenvalues are **complex**!

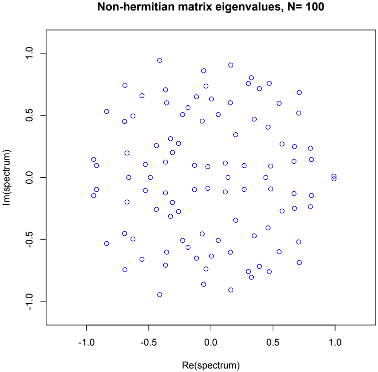

Figure: Distribution of $\mathbf{Y}_N$'s eigenvalues

Matrix model

Let $\mathbf{X}_N$ be a $N \times N$ matrix with i.i.d. entries

$$\mathbb{E}X_{ij} = 0 \ , \ \mathbb{E}|X_{ij}|^2 = 1$$

and consider the spectrum of matrix $\mathbf{Y}_N = \frac{1}{\sqrt{N}}\mathbf{X}_N$ as $N \to \infty$

- In this case, the eigenvalues are **complex**!

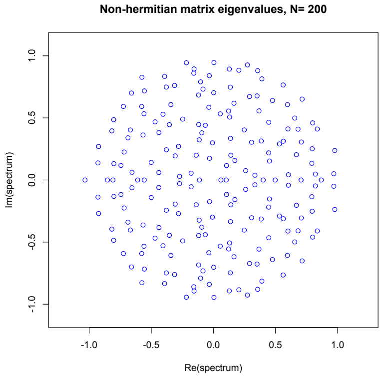

Figure: Distribution of $\mathbf{Y}_N$'s eigenvalues

Large Non-Hermitian Matrices

Matrix model

Let $\mathbf{X}_N$ be a $N \times N$ matrix with i.i.d. entries

$$\mathbb{E}X_{ij} = 0 \ , \ \mathbb{E}|X_{ij}|^2 = 1$$

and consider the spectrum of matrix $\mathbf{Y}_N = \frac{1}{\sqrt{N}}\mathbf{X}_N$ as $N \to \infty$

- In this case, the eigenvalues are **complex**!

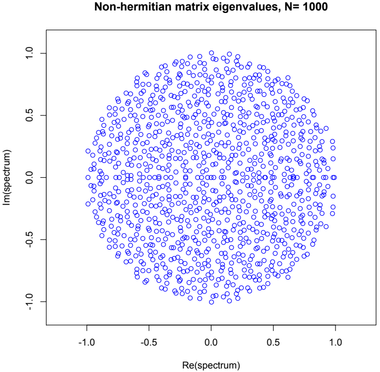

Figure: Distribution of $\mathbf{Y}_N$'s eigenvalues

Large Non-Hermitian Matrices

Matrix model

Let $\mathbf{X}_N$ be a $N \times N$ matrix with i.i.d. entries

$$\mathbb{E}X_{ij} = 0 \ , \ \mathbb{E}|X_{ij}|^2 = 1$$

and consider the spectrum of matrix $\mathbf{Y}_N = \frac{1}{\sqrt{N}}\mathbf{X}_N$ as $N \to \infty$

- In this case, the eigenvalues are **complex**!

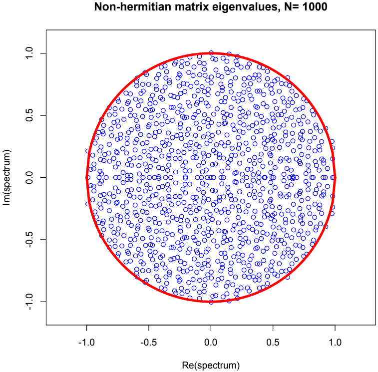

Figure: The circular law (in red)

Theorem: The Circular Law (Ginibre, Metha, Girko, Tao & Vu, etc.)

$\boxed{\text{The spectrum of } \mathbf{Y}_N \text{ converges to } \textbf{the uniform probability on the disc}}$

Motivations

An old history

- Data Analysis (Wishart, 1928)
- Theoretical Physics (from the '50s - Wigner, Dyson, Pastur, etc.)
- Pure mathematics (from the late '80s - non-commutative probability, free probability, operator algebra - Voiculescu, etc.)
- Graph theory (théorie spectrale des graphes)
- Wireless communication (Telatar, 1995 - Verdú, Tse, Shamai, Lévêque + important french group: Loubaton, Hachem, Debbah, Couillet, N., etc.)

Current trends

- Statistics in large dimension (El Karoui, Bickel & Levina, etc.)
- Pure mathematics: universality questions, operator algebra (Tao, Vu, Erdös, Guionnet, etc.)
- Social networks, communication networks
- Neuroscience (non-hermitian models - G. Wainrib)

Objective of this mini-course

Objective

- To present emblematic results and concepts in the theory of Large Random Matrices
- To give details on the technical means
- To present motivating applications of the theory

Also ..

- To demystify this theory because the technical price to enter it is substantial for a newcomer

The spectral theorem

Eigenvectors and eigenvalues

Given a $N \times N$ matrix $\mathbf{A}$ we are interested in its **eigenvalues** λ

$$\boxed{\mathbf{A}\vec{u} = \lambda\vec{u} \ , \quad (\vec{u} \neq 0)}$$

and its associated **eigenvectors** $\vec{u}$.

The spectral theorem - complex case

if $\mathbf{A}$ is hermitian:

$$\mathbf{A} = \mathbf{A}^* \quad \Leftrightarrow \quad [\mathbf{A}]_{ij} = \overline{[\mathbf{A}]}_{ji}$$

then $\mathbf{A}$ is diagonalizable with real eigenvalues:

$$\boxed{\mathbf{A} = \mathbf{U}^*\mathbf{\Lambda}\mathbf{U} \ , \quad \mathbf{U}\mathbf{U}^* = \mathbf{U}^*\mathbf{U} = \mathbf{I}_N}$$

with $\mathbf{U}$ unitary matrix and $\mathbf{\Lambda}$ real diagonal.

The spectral theorem - real case

If $\mathbf{A}$ is symmetric that is $\mathbf{A} = \mathbf{A}^T$, then

$$\boxed{\mathbf{A} = \mathbf{O}^T\mathbf{\Lambda}\mathbf{O} \ , \quad \mathbf{O}\mathbf{O}^T = \mathbf{O}^T\mathbf{O} = \mathbf{I}_N}$$

where $\mathbf{O}$ is (real) **orthogonal**.

The spectral measure of a matrix $\mathbf{A}$

.. also called the **empirical measure of the eigenvalues**

The Dirac measure

We define a **probability measure** δ_x over $\mathbb{R}$ by

$$\delta_x([a,b]) = \begin{cases} 1 & \text{if } x \in [a,b] \\ 0 & \text{else} \end{cases}$$

otherwise stated:

A set $[a,b]$ is assigned value 1 if $x \in [a,b]$ and value 0 else.

The spectral measure

If $\mathbf{A}$ is $N \times N$ hermitian with eigenvalues $\lambda_1, \cdots, \lambda_N$ then its **spectral measure** is:

$$L_N = \frac{1}{N} \sum_{i=1}^{N} \delta_{\lambda_i} \qquad \Rightarrow \qquad L_N([a,b]) = \frac{\#\{\lambda_i \in [a,b]\}}{N}$$

Otherwise stated

$L_N([a,b])$ is the **proportion** of eigenvalues of $\mathbf{A}$ in $[a,b]$.

Normalization

Given a matrix $\mathbf{Y}_N$ with random entries, we wish to find the **right normalization** of the entries so that the **eigenvalues** (λ_i) **are confined**.

Loose conditions to control the eigenvalues is (for example):

$$(\mathbf{C_p}): \quad \frac{1}{N}\sum_{i=1}^{N}\lambda_i^{\mathbf{p}} \quad = \quad O(1) \ ,$$

Normalization: example of Wigner matrices

Consider a hermitian $N \times N$ matrix $\mathbf{X}_N = (X_{ij})$ with i.i.d. entries **on and above** the diagonal:

$$\begin{cases} X_{ii} & \text{real} \\ X_{ij} & \text{i.i.d. if } i < j \\ X_{ij} = \overline{X}_{ji} & \text{if } i > j\ . \end{cases} \qquad \text{with} \quad \mathbb{E}X_{ij} = 0 \text{ and } \mathbb{E}\,|X_{ij}|^2 = \sigma^2$$

Let $\mathbf{Y}_N = \alpha_N \mathbf{X}_N$, α_N to be determined so that $\mathbf{Y}_N$'s eigenvalues (λ_i) are confined.

$$(\mathbf{C_1}): \quad \frac{1}{N}\sum_{i=1}^N \lambda_i = \frac{1}{N}\operatorname{Trace}\mathbf{Y}_N = \frac{\alpha_N}{N}\sum_{i=1}^N X_{ii} \xrightarrow[N\to\infty]{LLN} 0 \qquad \text{if} \quad \boxed{\alpha_N = O(1)}$$

$$(\mathbf{C_2}): \quad \frac{1}{N}\sum_{i=1}^N \lambda_i^2 = \frac{1}{N}\operatorname{Trace}\mathbf{Y}_N^2 = \frac{{\alpha_N}^2}{N}\operatorname{Trace}\mathbf{X}_N^2$$

$$= \frac{{\alpha_N}^2}{N}\left\{\sum_{i=1}^N X_{ii}^2 + 2\sum_{i<j}|X_{ij}|^2\right\} = O(1) \qquad \text{if} \quad \boxed{\alpha_N \propto \frac{1}{\sqrt{N}}}$$

Definition: A **Wigner matrix** is a matrix $\mathbf{Y}_N = \dfrac{\mathbf{X}_N}{\sqrt{N}}$

Spectrum analysis: The historical proof of Wigner's theorem

1. **Compute the asymptotic moments** of the spectral distribution

$$L_N = \frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i} \quad \text{of} \quad \mathbf{Y}_N = \frac{\mathbf{X}_N}{\sqrt{N}}$$

   that is

$$m_p(N) = \int x^p\, L_n(dx) = \frac{1}{N}\sum_{i=1}^{N}\lambda_i^p = \frac{1}{N}\mathrm{Trace}\,\mathbf{X}_N^p$$

   and prove that

$$m_p(N) \xrightarrow[N\to\infty]{} \begin{cases} \frac{1}{k+1}\binom{2k}{k} & \text{if } p = 2k\ , \\ 0 & \text{if } p = 2k+1 \end{cases}$$

2. On the other hand, **compute the moments** of the semi-circular distribution:

$$\int_{-2}^{2}\lambda^k\,\frac{\sqrt{4-\lambda^2}}{2\pi}\,d\lambda \quad = \begin{cases} \frac{1}{k+1}\binom{2k}{k} & \text{if } p = 2k\ , \\ 0 & \text{if } p = 2k+1 \end{cases}$$

3. **Conclude**: convergence of moments + tightness implies the convergence of the spectral distribution.

$\Rightarrow$ Computation of empirical moments heavily relies on (sometimes difficult) **combinatorics**.

Spectrum analysis: The resolvent

- Consider the equation in $\vec{x}$:

$$\boxed{\mathbf{A}\,\vec{x} = z\,\vec{x} + \vec{b}} \quad \Leftrightarrow \quad (\mathbf{A} - z\mathbf{I})\vec{x} = \vec{b} \quad \Leftrightarrow \quad \vec{x} = (\mathbf{A} - z\mathbf{I})^{-1}\vec{b}$$

  if $z \notin \mathrm{spectrum}(\mathbf{A})$.
- The resolvent of $\mathbf{A}$ is

$$\boxed{\mathbf{Q}(z) = (\mathbf{A} - z\mathbf{I})^{-1}}$$

- its singularities are exactly **eigenvalues** of $\mathbf{A}$.
- Resolvent of a Hermitian matrix

$$\mathbf{A} = \mathbf{U}^*\mathbf{\Lambda}\mathbf{U} \quad \Rightarrow \quad \mathbf{Q}(z) = \mathbf{U}^*(\mathbf{\Lambda} - z\mathbf{I})^{-1}\mathbf{U}$$

$$\mathbf{A} = \mathbf{U}^* \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{bmatrix} \mathbf{U} \quad \Rightarrow \quad \mathbf{Q}(z) = \mathbf{U}^* \begin{bmatrix} \frac{1}{\lambda_1 - z} & & \\ & \ddots & \\ & & \frac{1}{\lambda_N - z} \end{bmatrix} \mathbf{U}$$

- **Problem:** if size of $\mathbf{A}$ big, then size of $\mathbf{Q}$ big as well.
- The right object to consider (cf. supra) is the **normalized trace of the resolvent**.

Spectrum Analysis: The Stieltjes Transform I

Given a probability measure $\mathbb{P}$, its **Stieltjes transform** is a function

$$g(z) = \int_{\mathbb{R}} \frac{\mathbb{P}(d\lambda)}{\lambda - z} \ , \quad z \in \mathbb{C}^+ \ ,$$

with inverse formulas

$$\begin{aligned} \mathbb{P}[a,b] &= \frac{1}{\pi} \lim_{y \downarrow 0} \Im \int_a^b g(x + \mathbf{i}y)\, dx \ , \quad \text{if } \mathbb{P}\{a\} = \mathbb{P}\{b\} = 0 \\ \int f \, d\,\mathbb{P} &= \frac{1}{\pi} \lim_{y \downarrow 0} \Im \int_{\mathbb{R}} f(x) g(x + \mathbf{i}y)\, dx \ , \end{aligned}$$

Examples

1. Dirac measure:

$$\mathbb{P} = \delta_{\lambda_0} \quad \Rightarrow \quad g(z) = \frac{1}{\lambda_0 - z}$$

2. Spectral measure:

$$\mathbb{P} = \frac{1}{N} \sum_{i=1}^{N} \delta_{\lambda_i} \quad \Rightarrow \quad g(z) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\lambda_i - z}$$

Spectrum Analysis: The Stieltjes Transform II

Relation with the resolvent of Large Random Matrices

Let $\mathbf{A}$ hermitian with eigenvalues (λ_i) and spectral measure $\frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i}$.

Then

$$
\begin{aligned}
g(z) &= \text{Stieltjes transform of } \left(\frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i}\right) \\
&= \frac{1}{N}\sum_{1}^{N}\frac{1}{\lambda_i - z} \\
&= \frac{1}{N}\mathrm{Trace}\begin{bmatrix} \frac{1}{\lambda_1 - z} & & \\ & \ddots & \\ & & \frac{1}{\lambda_N - z}\end{bmatrix} = \frac{1}{N}\mathrm{Trace}\,(\mathbf{A} - z\boldsymbol{I})^{-1}
\end{aligned}
$$

- $\boxed{\text{The Stieltjes transfom } g \text{ is the } \textbf{normalized trace} \text{ of the resolvent } (\mathbf{A} - z\boldsymbol{I})^{-1}}$
- Whatever size of $\mathbf{A}$, Stieltjes transform g remains a fonction $\mathbb{C} \to \mathbb{C}$.

Summary

Large Random Matrices

- Associated to a matrix $\mathbf{A}$ is its **spectral measure**: $L_N = \frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i(\mathbf{A})}$
- .. its resolvent: $\mathbf{Q}(z) = (\mathbf{A} - z\mathbf{I}_N)^{-1}$
- .. its Stieltjes transform

$$g_n(z) = \int \frac{L_n(d\,\lambda)}{\lambda - z} = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\lambda_i(\mathbf{A}) - z} = \frac{1}{N}\mathrm{Trace}\,\mathbf{Q}(z)$$

Normalizing a matrix

- In order to confine a matrix' eigenvalues, we consider the condition:

$$(\mathbf{C_p}) : \quad \frac{1}{N}\sum_{i=1}^{N}\lambda_i^{\mathbf{p}} \quad = \quad O(1)\ ,$$

Classical results

- Wigner's theorem, Marčenko-Pastur's theorem, The circular law.

Wishart Matrices I

The model

- Consider a $N \times n$ matrix $\mathbf{X}_N$ with i.i.d. entries

$$\mathbb{E}X_{ij} = 0 \ , \quad \mathbb{E}|X_{ij}|^2 = \sigma^2 \ .$$

Matrix $\mathbf{X}_N$ is a n-sample of N-dimensional vectors:

$$\mathbf{X}_N = [\mathbf{X}_{\cdot 1} \ \cdots \ \mathbf{X}_{\cdot n}] \quad \text{with} \quad \mathbb{E}\mathbf{X}_{\cdot 1}\mathbf{X}_{\cdot 1}^* = \sigma^2 \mathbf{I}_N \ .$$

Objective

- to describe **the limiting spectrum** of $\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*$ as

$$\frac{N}{n} \xrightarrow[n\to\infty]{} c \in (0,\infty) \ .$$

i.e. dimensions of matrix $\mathbf{X}_N$ are **of the same order**.

The usual case $N << n$

Assume N fixed and $n \to \infty$. Since

$$\mathbb{E}\mathbf{X}_{\cdot 1}\mathbf{X}_{\cdot 1}^{*} = \sigma^2 \mathbf{I}_N \ ,$$

L.L.N implies

$$\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^{*} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_{\cdot i}\mathbf{X}_{\cdot i}^{*} \quad \xrightarrow[n\to\infty]{a.s.} \quad \sigma^2 \mathbf{I}_N$$

In particular,

- all the eigenvalues of $\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^{*}$ converge to σ^2,
- equivalently, the spectral measure of $\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^{*}$ converges to δ_{σ^2}.

A priori observation # 1

If the ratio of dimensions $c \searrow 0$, then the spectral measure should look like a Dirac measure at point σ^2.

Wishart Matrices III

The case where $c > 1$

Recall that $\mathbf{X}_N$ is $N \times n$ matrix and $c = \lim \frac{N}{n}$.

If $N > n$, then $\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*$ is **rank-defficient** and has rank n;

- in this case, eigenvalue 0 has multiplicity $N - n$ and the spectral measure writes:

$$L_N = \frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i} = \frac{1}{N}\sum_{i=1}^{n}\delta_{\lambda_i} + \frac{N-n}{N}\delta_0$$

- The limiting spectral measure of L_N **necessarily features a Dirac measure at 0**:

$$\frac{N-n}{N}\delta_0 \longrightarrow \left(1 - \frac{1}{c}\right)\delta_0 \quad \text{as} \quad \frac{N}{n} \to c\ .$$

A priori observation #2

> If $c > 1$, then the limiting spectral measure will feature a Dirac measure at 0 with weight $1 - \frac{1}{c}$.

Simulations

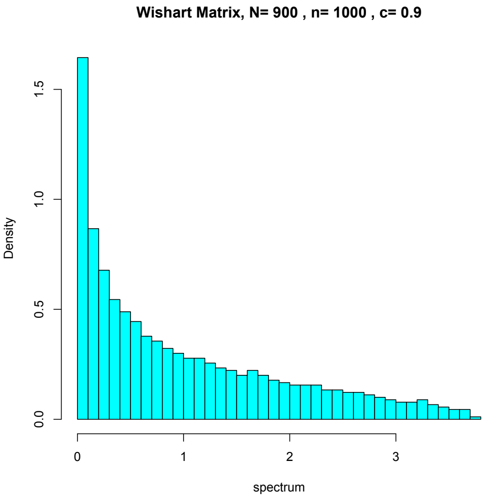

Figure: Histogram of $\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*$, $\sigma^2 = 1$

Simulations

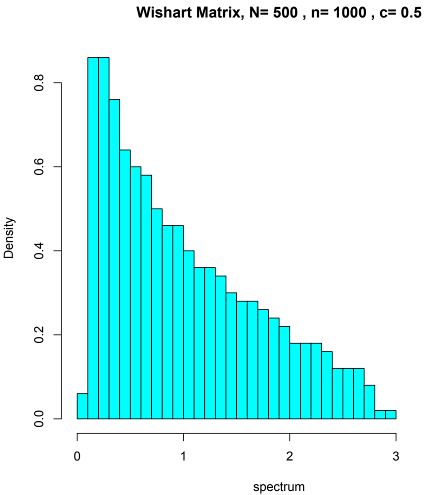

Figure: Histogram of $\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*$, $\sigma^2 = 1$

Simulations

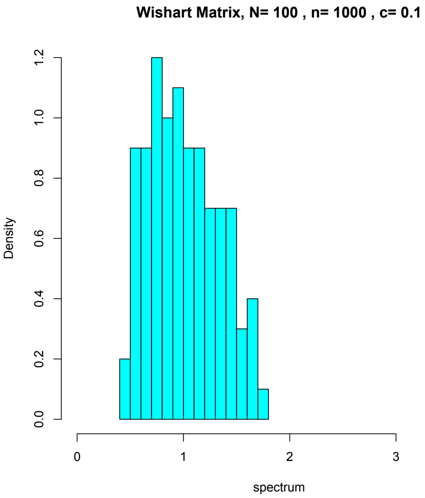

Figure: Histogram of $\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*$, $\sigma^2 = 1$

Simulations

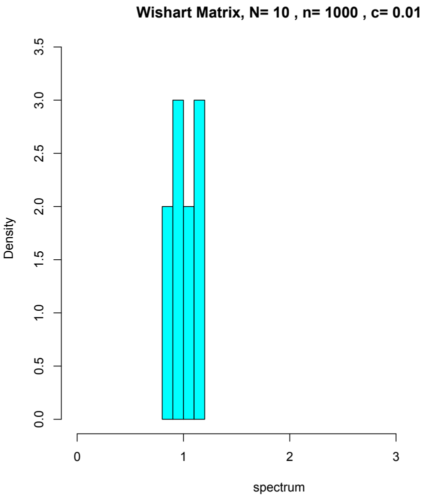

Figure: Histogram of $\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*$, $\sigma^2 = 1$

Marčenko-Pastur theorem

Theorem

- Consider a $N \times n$ matrix $\mathbf{X}_N$ with i.i.d. entries

$$\mathbb{E}X_{ij} = 0 \ , \quad \mathbb{E}|X_{ij}|^2 = \sigma^2 \ .$$

with N and n **of the same order** and L_N its **spectral measure**:

$$c_n \triangleq \frac{N}{n} \xrightarrow[n\to\infty]{} c \in (0,\infty) \ , \quad L_N = \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i} \ , \quad \lambda_i = \lambda_i\left(\frac{1}{n}\mathbf{X}_N \mathbf{X}_N^*\right)$$

- Then **almost surely** (= for almost every realization)

$$L_N \xrightarrow[N,n\to\infty]{} \mathbb{P}_{\check{\mathrm{M}}\mathrm{P}} \quad \text{in distribution}$$

where $\mathbb{P}_{\check{\mathrm{M}}\mathrm{P}}$ is **Marčenko-Pastur** distribution:

$$\begin{array}{rcl} \mathbb{P}_{\check{\mathrm{M}}\mathrm{P}}(dx) & = & \left(1-\frac{1}{c}\right)^+ \delta_0(dx) + \frac{\sqrt{(\lambda^+ - x)(x-\lambda^-)}}{2\pi\sigma^2 x c} 1_{[a,b]}(x)\, dx \\ \text{with} & & \left\{ \begin{array}{rcl} \lambda^- & = & \sigma^2(1-\sqrt{c})^2 \\ \lambda^+ & = & \sigma^2(1+\sqrt{c})^2 \end{array} \right. \end{array}$$

Simulations vs MP distribution

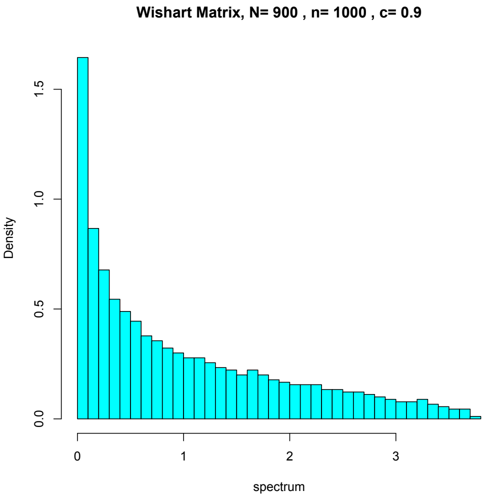

Figure: Histogram of $\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*$, $\sigma^2 = 1$

Simulations vs M̌P distribution

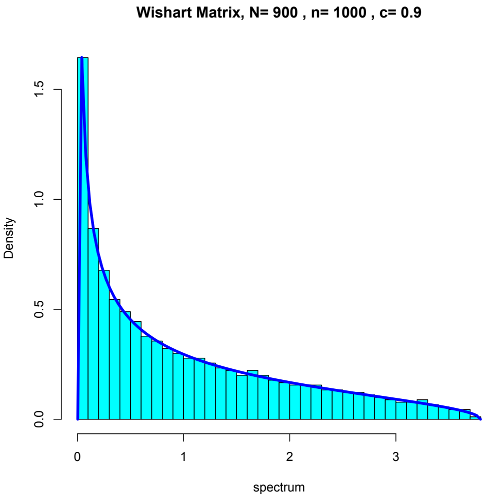

Figure: Marčenko-Pastur distribution for $c = 0.9$

Simulations vs MP distribution

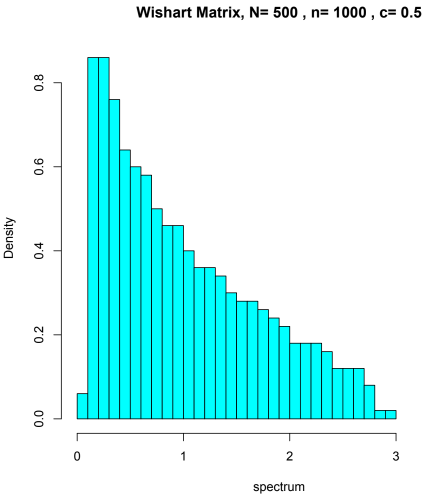

Figure: Histogram of $\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*$, $\sigma^2 = 1$

Simulations vs MP distribution

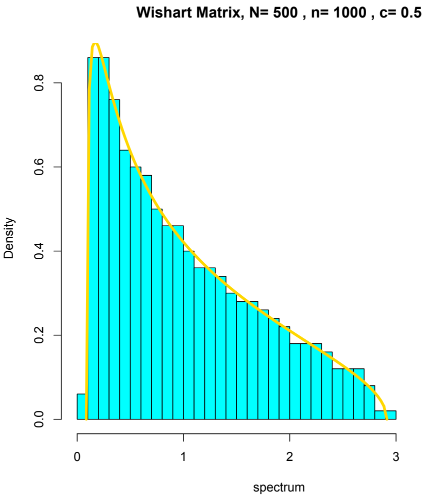

Figure: Marčenko-Pastur distribution for $c = 0.5$

Simulations vs MP distribution

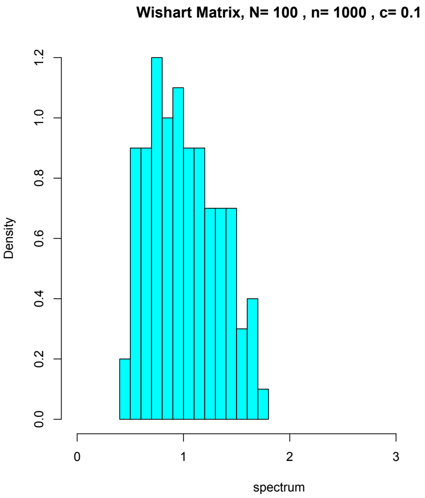

Figure: Histogram of $\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*$, $\sigma^2 = 1$

Simulations vs MP distribution

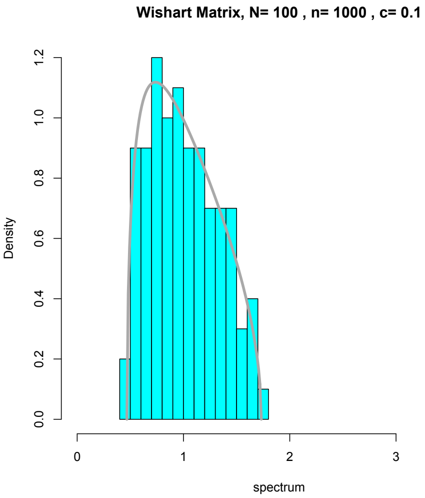

Figure: Marčenko-Pastur distribution for $c = 0.1$

Simulations vs $\check{\text{M}}$P distribution

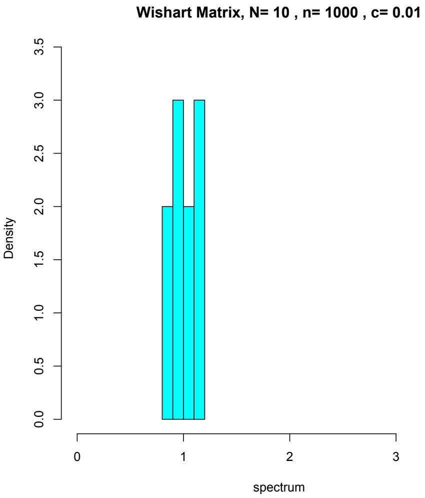

Figure: Histogram of $\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*$, $\sigma^2 = 1$

Simulations vs ̌MP distribution

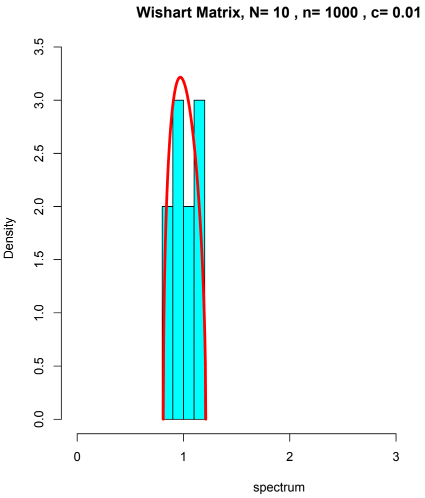

Figure: Marčenko-Pastur distribution for $c = 0.01$

Remarks I

- Marčenko-Pastur theorem describes the **global regime** of the spectrum.
- Convergence in distribution: For a given realization and every test function $\phi : \mathbb{R} \to \mathbb{R}$, the theorem states:

$$\frac{1}{N} \sum_{i=1}^{N} \phi(\lambda_i) \xrightarrow[N,n\to\infty]{} \int \phi(x) \mathbb{P}_{\check{\mathrm{M}}\mathrm{P}}(dx) \ .$$

- The Dirac measure at zero is **an artifact** due to the dimensions of the matrix if

$$N > n \quad \text{(cf. infra)} \ .$$

Remarks II

What if $c \searrow 0$?

- If $c \to 0$, that is $n >> N$, then typical from the usual regime **"small dimensional data vs large samples"**.
- the support of Marčenko-Pastur distribution

$$[\sigma^2(1-\sqrt{c})^2, \sigma^2(1+\sqrt{c})^2]$$

concentrates around $\{\sigma^2\}$ and

$$\mathbb{P}_{\check{\mathrm{M}}\mathrm{P}} \xrightarrow[c\to 0]{} \delta_{\sigma^2} \ .$$

- In accordance with **a priori information # 1**

Remarks II

What if $c \searrow 0$?

- If $c \to 0$, that is $n >> N$, then typical from the usual regime **"small dimensional data vs large samples"**.
- the support of Marčenko-Pastur distribution

$$[\sigma^2(1-\sqrt{c})^2, \sigma^2(1+\sqrt{c})^2]$$

concentrates around $\{\sigma^2\}$ and

$$\mathbb{P}_{\check{\mathrm{M}}\mathrm{P}} \xrightarrow[c \to 0]{} \delta_{\sigma^2} \quad .$$

- In accordance with **a priori information # 1**

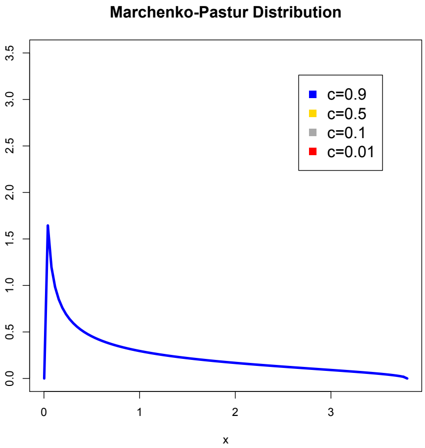

Figure: M̌P distribution as $c \searrow 0$

Remarks II

What if $c \searrow 0$?

- If $c \to 0$, that is $n >> N$, then typical from the usual regime **"small dimensional data vs large samples"**.
- the support of Marčenko-Pastur distribution

$$[\sigma^2(1-\sqrt{c})^2, \sigma^2(1+\sqrt{c})^2]$$

concentrates around $\{\sigma^2\}$ and

$$\mathbb{P}_{\check{\mathrm{M}}\mathrm{P}} \xrightarrow[c \to 0]{} \delta_{\sigma^2} \ .$$

- In accordance with **a priori information # 1**

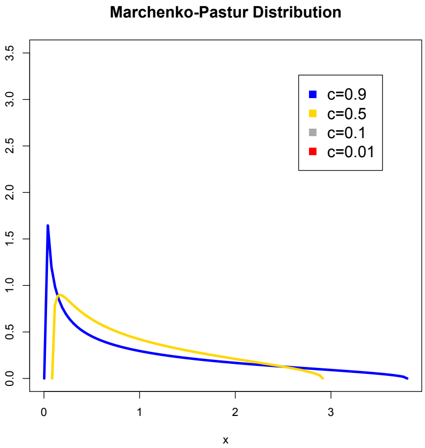

Figure: M̌P distribution as $c \searrow 0$

Remarks II

What if $c \searrow 0$?

- If $c \to 0$, that is $n >> N$, then typical from the usual regime **"small dimensional data vs large samples"**.
- the support of Marčenko-Pastur distribution

$$[\sigma^2(1-\sqrt{c})^2, \sigma^2(1+\sqrt{c})^2]$$

concentrates around $\{\sigma^2\}$ and

$$\mathbb{P}_{\check{\mathrm{M}}\mathrm{P}} \xrightarrow[c \to 0]{} \delta_{\sigma^2} \quad .$$

- In accordance with **a priori information # 1**

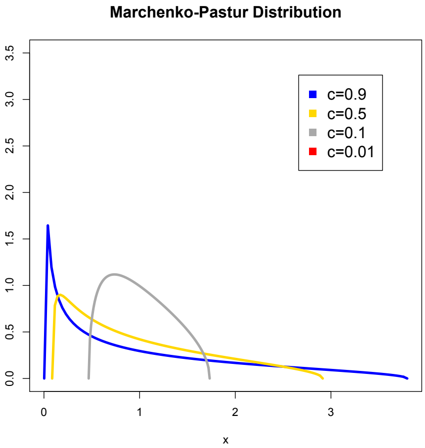

Figure: M̌P distribution as $c \searrow 0$

Remarks II

What if $c \searrow 0$?

- If $c \to 0$, that is $n >> N$, then typical from the usual regime **"small dimensional data vs large samples"**.
- the support of Marčenko-Pastur distribution

$$[\sigma^2(1-\sqrt{c})^2, \sigma^2(1+\sqrt{c})^2]$$

concentrates around $\{\sigma^2\}$ and

$$\mathbb{P}_{\check{\mathrm{M}}\mathrm{P}} \xrightarrow[c \to 0]{} \delta_{\sigma^2} \ .$$

- In accordance with **a priori information # 1**

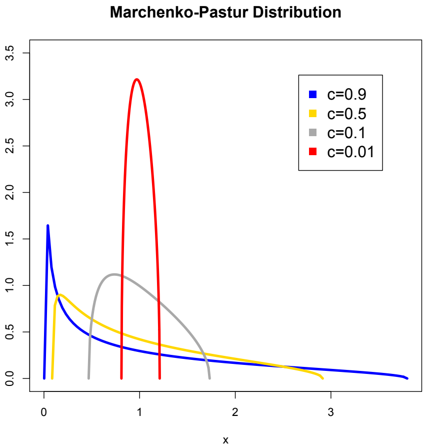

Figure: M̌P distribution as $c \searrow 0$

Results concerning the local regime for Wishart matrices

Convergence of extremal eigenvalues

Recall that $[\sigma^2(1-\sqrt{c})^2, \sigma^2(1+\sqrt{c})^2]$ is the support of M̌P distribution, then:

$$\lambda_{\max}\left(\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*\right) \xrightarrow[N,n\to\infty]{\text{almost surely}} \sigma^2(1+\sqrt{c})^2$$

$$\lambda_{\min}\left(\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*\right) \xrightarrow[N,n\to\infty]{\text{almost surely}} \sigma^2(1-\sqrt{c})^2$$

Fluctuations of $\lambda_{\max}$: Tracy-Widom distribution

We can fully describe the fluctuations of $\lambda_{\max}$:

$$\frac{N^{2/3}}{\Theta_N}\left\{\lambda_{\max}\left(\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*\right) - \sigma^2(1+\sqrt{c_n})^2\right\} \xrightarrow[N,n\to\infty]{\mathcal{L}} \mathbb{P}_{\mathrm{TW}}$$

where

$$c_n = \frac{N}{n} \quad \text{and} \quad \Theta_N = \sigma^2(1+\sqrt{c_n})\left(\frac{1}{\sqrt{c_n}}+1\right)^{1/3}$$

Strategy of proof

Recall definition of the **Stieltjes transform** g_n:

$$g_n(z) \quad = \quad \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\lambda_i - z} \quad = \quad \frac{1}{N}\mathrm{Trace}\left(\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^* - z\mathbf{I}_N\right)^{-1} \quad .$$

1. Convergence of the Stieltjes transform. Since

$$L_N = \frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i} \xrightarrow[N,n\to\infty]{} \mathbb{P}_{\check{\mathrm{M}}\mathrm{P}} \quad \Longleftrightarrow \quad g_n(z) \xrightarrow[N,n\to\infty]{} ST\left(\mathbb{P}_{\check{\mathrm{M}}\mathrm{P}}\right)$$

   we prove the convergence of g_n.

2. After algebraic manipulations and probabilistic arguments, we prove that

$$g_n(z) = \frac{1}{\sigma^2(1-c_n) - z - z\sigma^2 c_n g_n(z)} + \varepsilon_n(z) \quad \text{with} \quad \varepsilon_n(z) \xrightarrow[N,n\to\infty]{} 0 \ .$$

3. By stability of Marčenko-Pastur's equation, g_n converges to a function $\mathbf{g}_{\check{\mathrm{M}}\mathrm{P}}$ which satisfies **the fixed point equation**:

$$\mathbf{g}_{\check{\mathrm{M}}\mathrm{P}}(z) = \frac{1}{\sigma^2(1-c) - z - z\sigma^2 c\,\mathbf{g}_{\check{\mathrm{M}}\mathrm{P}}(z)}$$

4. We identify $\boxed{\mathbb{P}_{\check{\mathrm{M}}\mathrm{P}} = (\text{Stieltjes Transform})^{-1}(\mathbf{g}_{\check{\mathrm{M}}\mathrm{P}})}$.

Linear Algebra I: Diagonal element of the resolvent

Let $\mathbf{\Sigma}_N$ a $N \times n$ matrix with rows $\vec{\xi}_i$ and consider the resolvent of $\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*$:

$$\mathbf{\Sigma}_N = \begin{bmatrix} \vec{\xi}_1 \\ \vdots \\ \vec{\xi}_N \end{bmatrix} \quad \text{and} \quad \mathbf{Q}(z) = \left(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^* - z\mathbf{I}_N\right)^{-1} \ .$$

Proposition

The diagonal element $q_{ii} = [\mathbf{Q}]_{ii}$ expresses:

$$\boxed{q_{ii}(z) = \frac{1}{-z\left(1 + \vec{\xi}_i \left(\mathbf{\Sigma}_{(i)}^* \mathbf{\Sigma}_{(i)} - z\mathbf{I}_n\right)^{-1} \vec{\xi}_i^*\right)}}$$

where $\mathbf{\Sigma}_{(i)}$ is matrix $\mathbf{\Sigma}$ with row $\vec{\xi}_i$ removed:

$$\mathbf{\Sigma}_{(i)} = \begin{bmatrix} \vdots \\ \vec{\xi}_{i-1} \\ \vec{\xi}_{i+1} \\ \vdots \end{bmatrix}$$

Linear Algebra II: Rank-one perturbation

Let $\vec{u}$ a $N \times 1$ vector. Notice that $\boxed{\vec{u}\vec{u}^*}$ is a **rank-one** $N \times N$ matrix.

Proposition

- Let $\mathbf{A}$ be a $N \times N$ matrix then:

$$\left| \frac{1}{N}\mathrm{Trace}(\mathbf{A} + \vec{u}\vec{u}^* - z\mathbf{I}_N)^{-1} - \frac{1}{N}\mathrm{Trace}(\mathbf{A} - z\mathbf{I}_N)^{-1} \right| \leq \frac{1}{N\Im(z)}$$

Conclusion

Asymptotically, normalized trace of the resolvent **not sensitive** to **rank-one** perturbations.

Linear Algebra III: Stieltjes transform property

The Stieltjes transforms associated to $\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*$ and $\frac{1}{n}\mathbf{X}_N^*\mathbf{X}_N$ write

$$g_n(z) = \frac{1}{N}\mathrm{Trace}\left(\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^* - z\mathbf{I}_N\right)^{-1}, \quad \tilde{g}_n(z) = \frac{1}{n}\mathrm{Trace}\left(\frac{1}{n}\mathbf{X}_N^*\mathbf{X}_N - z\mathbf{I}_n\right)^{-1}.$$

Spectra of $\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*$ and $\frac{1}{n}\mathbf{X}_N^*\mathbf{X}_N$ coincide **up to the null eigenvalue**. As an important consequence

$$\boxed{\tilde{g}_n(z) = c_n g_n(z) + (1-c_n)\left(-\frac{1}{z}\right)} \qquad c_n = \frac{N}{n}$$

Proof: Let for example $n > N$ then

$$\mathrm{spectrum}\left(\frac{1}{n}\mathbf{X}_N^*\mathbf{X}_N\right) = \mathrm{spectrum}\left(\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*\right) \cup \{0\}$$

where 0 has multiplicity $n-N$. Hence

$$\begin{aligned}
\tilde{g}_n(z) &= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{\lambda_i - z} = \frac{1}{n}\sum_{i=1}^{N}\frac{1}{\lambda_i - z} + \frac{1}{n}\sum_{i=N+1}^{n}\left(-\frac{1}{z}\right) \\
&= \frac{N}{n}g_n(z) + \frac{n-N}{n}\left(-\frac{1}{z}\right) = c_n g_n(z) + (1-c_n)\left(-\frac{1}{z}\right)
\end{aligned}$$

Probability theory: Convergence of quadratic forms

Let

$$\vec{x} = (x_1, \cdots, x_N) \quad \text{with} \quad \mathbb{E}x_i = 0 \quad \mathbb{E}x_i^2 = \sigma^2$$

the x_i's being i.i.d. Consider the quadratic form $\boxed{\frac{1}{N}\vec{x}\mathbf{A}\vec{x}^*}$

Proposition

Let matrix $\mathbf{A}$ be **deterministic** or **independent from** $\vec{x}$

1. then

$$\mathbb{E}_{\vec{x}}\left\{\frac{1}{N}\vec{x}\mathbf{A}\vec{x}^*\right\} = \frac{\sigma^2}{N}\mathrm{Trace}\mathbf{A}$$

2. and

$$\frac{1}{N}\vec{x}\mathbf{A}\vec{x}^* - \frac{\sigma^2}{N}\mathrm{Trace}\mathbf{A} \xrightarrow[N\to\infty]{} 0 \ .$$

Let's summarize ..

In order to handle the normalized trace of the resolvent (= Stieltjes transform of the associated spectral measure), four important arguments are:

- Expression of the diagonal element of the resolvent

$$q_{ii}(z) = \frac{1}{-z\left(1 + \vec{\xi}_i\left(\mathbf{\Sigma}^*_{(i)}\mathbf{\Sigma}_{(i)} - z\mathbf{I}_n\right)^{-1}\vec{\xi}^*_i\right)}$$

- Robustness to rank-one perturbation

$$\frac{1}{N}\mathrm{Trace}(\mathbf{A} + \vec{u}\vec{u}^* - z\mathbf{I}_N)^{-1} \approx \frac{1}{N}\mathrm{Trace}(\mathbf{A} - z\mathbf{I}_N)^{-1} \quad \text{as } N \to \infty \ .$$

- Stieltjes transform property

$$\tilde{g}_n(z) = c_n g_n(z) + (1 - c_n)\left(-\frac{1}{z}\right)$$

- Approximation of quadratic forms

$$\frac{1}{N}\vec{x}\mathbf{A}\vec{x}^* \approx \frac{\sigma^2}{N}\mathrm{Trace}(\mathbf{A}) \quad \text{as } N \to \infty \ .$$

Approximate fixed-point equation I

Diagonal elements

Denote by

$$q_{ii}(z) = \left[\left(\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^* - z\mathbf{I}_N\right)^{-1}\right]_{ii} \quad \text{and} \quad \tilde{q}_{jj}(z) = \left[\left(\frac{1}{n}\mathbf{X}_N^*\mathbf{X}_N - z\mathbf{I}_n\right)^{-1}\right]_{jj}$$

the diagonal elements of the resolvents.

Stieltjes transforms

The Stieltjes transforms associated to $\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*$ and $\frac{1}{n}\mathbf{X}_N^*\mathbf{X}_N$ write

$$\begin{aligned} g_n(z) &= \frac{1}{N}\mathrm{Trace}\left(\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^* - z\mathbf{I}_N\right)^{-1} = \frac{1}{N}\sum_{i=1}^{N} q_{ii}(z)\ , \\ \tilde{g}_n(z) &= \frac{1}{n}\mathrm{Trace}\left(\frac{1}{n}\mathbf{X}_N^*\mathbf{X}_N - z\mathbf{I}_n\right)^{-1} = \frac{1}{n}\sum_{j=1}^{n} \tilde{q}_{jj}(z)\ . \end{aligned}$$

Approximate fixed-point equation II

For simplicity, denote by $\mathbf{Y}_N = \frac{\mathbf{X}_N}{\sqrt{n}}$ and recall

$$g_n(z) = \frac{1}{N}\mathrm{Trace}\left(\mathbf{Y}_N\mathbf{Y}_N^* - z\mathbf{I}_N\right)^{-1} = \frac{1}{N}\sum_{i=1}^{N} q_{ii}(z)$$

We have:

$$\begin{aligned}
q_{ii}(z) &\overset{(a)}{=} \frac{1}{-z\left(1+\vec{\xi}_i\left(\mathbf{Y}_{(i)}^*\mathbf{Y}_{(i)} - z\mathbf{I}_n\right)^{-1}\vec{\xi}_i^{\,*}\right)} \\
&\overset{(b)}{\approx} \frac{1}{-z\left(1+\frac{\sigma^2}{n}\mathrm{Trace}\left(\mathbf{Y}_{(i)}^*\mathbf{Y}_{(i)} - z\mathbf{I}_n\right)^{-1}\right)} \\
&\overset{(c)}{\approx} \frac{1}{-z\left(1+\frac{\sigma^2}{n}\mathrm{Trace}\left(\mathbf{Y}^*\mathbf{Y} - z\mathbf{I}_n\right)^{-1}\right)} = \frac{1}{-z\left(1+\frac{\sigma^2}{n}\sum_{j=1}^{n}\tilde{q}_{jj}(z)\right)}
\end{aligned}$$

- where (a) follows from the expression of the diagonal element of the resolvent,
- where (b) follows from **asymptotic behaviour of quadratic form**;
- where (c) follows from **rank-one perturbation argument**.

Approximate fixed-point equation III

We have

$$q_{ii}(z) \approx \frac{1}{-z\left(1+\frac{\sigma^2}{n}\sum_{j=1}^n \tilde{q}_{jj}(z)\right)} = -\frac{1}{z(1+\sigma^2\tilde{g}_n(z))}$$

Summing up,

$$\begin{aligned} g_n(z) &= \frac{1}{N}\sum_{i=1}^N q_{ii}(z) \approx -\frac{1}{z(1+\sigma^2\tilde{g}_n(z))} \\ &\overset{(d)}{=} \frac{1}{-z\left[1+\sigma^2\left\{c_n g_n(z)+(1-c_n)\left(-\frac{1}{z}\right)\right\}\right]} \\ &= \frac{1}{\sigma^2(1-c_n)-z-z\sigma^2 c_n g_n(z)} \end{aligned}$$

- where (d) follows from the fact that

$$\tilde{g}_n(z) = c_n g_n(z) + (1-c_n)\left(-\frac{1}{z}\right)$$

Approximate fixed-point equation IV

we finally obtain the approximate equation

$$\boxed{g_n(z) \approx \frac{1}{\sigma^2(1-c_n) - z - z\sigma^2 c_n g_n(z)}}$$

This method can be referred to as:

finding the limiting equation by approximating the diagonal elements of the resolvent

The limiting fixed-point equation

Theoretical arguments (tightness and compacity) yield the convergence

$$g_n(z) \xrightarrow[N,n\to\infty]{a.s.} \mathbf{g}_{\check{\mathrm{MP}}}(z)$$

where the Stieltjes transform $\mathbf{g}_{\check{\mathrm{MP}}}$ satisfies the **fixed-point** equation:

$$\boxed{\mathbf{g}_{\check{\mathrm{MP}}}(z) = \frac{1}{\sigma^2(1-c) - z - z\sigma^2 c\mathbf{g}_{\check{\mathrm{MP}}}(z)}}$$

or equivalently the following **second-degree polynomial**:

$$\boxed{zc\sigma^2\mathbf{g}_{\check{\mathrm{MP}}}^2 + [z - \sigma^2(1-c)]\mathbf{g}_{\check{\mathrm{MP}}} + 1 = 0}\ .$$

- We also refer to the fixed-point equation as the **canonical equation**.

Solving the limiting equation

Explicit Stieltjes transform

Given the second-degree polynomial $\boxed{zc\sigma^2\mathbf{g}_{\check{\mathrm{M}}\mathrm{P}}^2 + [z - \sigma^2(1-c)]\mathbf{g}_{\check{\mathrm{M}}\mathrm{P}} + 1 = 0}$, an explicit solution is given by

$$\boxed{\mathbf{g}_{\check{\mathrm{M}}\mathrm{P}}(z) = \frac{-(z + \sigma^2(c-1)) + \sqrt{(z-b)(z-a)}}{2zc\sigma^2}}$$

with $a = \sigma^2(1-\sqrt{c})^2$ and $b = \sigma^2(1+\sqrt{c})^2$ and where $\sqrt{(\cdot)}$ refers to the branch of the square root function for which $\mathbf{g}_{\check{\mathrm{M}}\mathrm{P}}$ is a Stieltjes transform.

Marčenko-Pastur's distribution

The inverse formula

$$\mathbb{P}_{\check{\mathrm{M}}\mathrm{P}}[a,b] = \frac{1}{\pi}\lim_{y\downarrow 0}\Im\int_a^b \mathbf{g}_{\check{\mathrm{M}}\mathrm{P}}(x + \mathbf{i}y)\,dx$$

can be used to find:

$$\boxed{\mathbb{P}_{\check{\mathrm{M}}\mathrm{P}}(dx) \;=\; \left(1 - \frac{1}{c}\right)^+ \delta_0(dx) + \frac{\sqrt{(b-x)(x-a)}}{2\pi\sigma^2 xc}1_{[a,b]}(x)\,dx}$$

Marčenko-Pastur's Theorem: Summary

- Consider the model $\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*$, then its spectral measure satisfies:

$$\boxed{\text{a. s.} \quad L_N = \frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i} \xrightarrow[N,n\to\infty]{\mathcal{L}} \mathbb{P}_{\check{\mathrm{MP}}} \ .}$$

- Instead of directly working on L_N, we consider its **Stieltjes tranform**

$$g_n(z) = \frac{1}{N}\mathrm{Trace}\left(\frac{1}{n}\mathbf{X}_N\mathbf{X}_n^* - z\mathbf{I}_N\right)^{-1} \ ,$$

then prove that it satisfies the **approximate fixed-point** equation

$$\boxed{g_n(z) \approx \frac{1}{\sigma^2(1-c_n) - z - z\sigma^2 c_n g_n(z)}}$$

and that it converges to the solution $\mathbf{g}_{\check{\mathrm{MP}}}$ of the **canonical equation**

$$\boxed{\mathbf{g}_{\check{\mathrm{MP}}}(z) = \frac{1}{\sigma^2(1-c) - z - z\sigma^2 c\,\mathbf{g}_{\check{\mathrm{MP}}}(z)}}$$

- Computing explicitely $\mathbf{g}_{\check{\mathrm{MP}}}$ and inverting it yields finally the formula for $\mathbb{P}_{\check{\mathrm{MP}}}$.

Complement: the isotropic Marčenko-Pastur theorem

- While proving M̌P's theorem, we have seen that

$$g_n(z) = \frac{1}{N}\mathrm{Trace}\,\mathbf{Q}_N(z) \xrightarrow[N,n\to\infty]{} \mathbf{g}_{\check{\mathrm{M}}\mathrm{P}}(z) \quad \text{for} \quad z \in \mathbb{C}\setminus\mathbb{R}^+ \ .$$

Isotropic M̌P theorem

- let $\vec{\mathbf{a}}_N$ and $\vec{\mathbf{b}}_N$ be $N \times 1$ deterministic vectors such that

$$\sup_{n\geq 1}\|\vec{\mathbf{a}}_N\| \ , \ \sup_{n\geq 1}\|\vec{\mathbf{b}}_N\| \ \leq K < \infty$$

then

$$\boxed{\vec{\mathbf{a}}_N^*\mathbf{Q}_N(z)\vec{\mathbf{b}}_N - \langle\vec{\mathbf{a}}_N,\vec{\mathbf{b}}_N\rangle\mathbf{g}_{\check{\mathrm{M}}\mathrm{P}}(z) \xrightarrow[N,n\to\infty]{} 0} \quad \text{for} \quad z\in\mathbb{C}\setminus\mathbb{R}^+ \ ,$$

where $\langle\vec{\mathbf{a}}_N,\vec{\mathbf{b}}_N\rangle = \vec{\mathbf{a}}_N^*\vec{\mathbf{b}}_N$.

- In particular, if $\vec{\mathbf{u}}_N$ is $N\times 1$ unitary, i.e.

$$\langle\vec{\mathbf{u}}_N,\vec{\mathbf{u}}_N\rangle = \|\vec{\mathbf{u}}_N\|^2 = 1$$

then

$$\boxed{\vec{\mathbf{u}}_N^*\mathbf{Q}_N(z)\vec{\mathbf{u}}_N \xrightarrow[N,n\to\infty]{} \mathbf{g}_{\check{\mathrm{M}}\mathrm{P}}(z)} \quad \text{for} \quad z\in\mathbb{C}\setminus\mathbb{R}^+ \ .$$

Hence the name isotropic: the limit does not depend on the direction $\vec{\mathbf{u}}_N$..

Large covariance matrices

The model

- Consider a $N \times n$ matrix $\mathbf{X}_N$ with i.i.d. entries

$$\mathbb{E}X_{ij} = 0 \ , \quad \mathbb{E}|X_{ij}|^2 = 1 \ .$$

- Let $\mathbf{R}_N$ be a **deterministic** $N \times N$ nonnegative definite hermitian matrix.
- Consider

$$\mathbf{Y}_N = \mathbf{R}_N^{1/2}\mathbf{X}_N \ .$$

  Matrix $\mathbf{Y}_N$ is a n-sample of N-dimensional vectors:

$$\mathbf{Y}_N = [\mathbf{Y}_{\cdot 1} \ \cdots \ \mathbf{Y}_{\cdot n}] \quad \text{with} \quad \mathbf{Y}_{\cdot 1} = \mathbf{R}_N^{1/2}\mathbf{X}_{\cdot 1} \quad \text{and} \quad \mathbb{E}\mathbf{Y}_{\cdot 1}\mathbf{Y}_{\cdot 1}^* = \mathbf{R}_N \ .$$

- $\mathbf{R}_N$ often called **Population covariance matrix**.

Objective

$$\boxed{\text{To describe the limiting spectrum of } \frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^* \text{ as } N, n \to \infty.}$$

Remark

- If N fixed and $n \to \infty$ then $\boxed{\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^* \longrightarrow \mathbf{R}_N}$

Guessing the canonical equation I: diagonal case

- Consider first the case where $\mathbf{R}_N$ is diagonal:

$$\mathbf{R}_N = \mathrm{diag}\left(\rho_i\,,\ i = 1 : N\right)\ .$$

Guessing the canonical equation I: diagonal case

- Consider first the case where $\mathbf{R}_N$ is diagonal:

$$\mathbf{R}_N = \operatorname{diag}(\rho_i\,,\ i = 1 : N)\ .$$

- The method of approximating the diagonal elements of the resolvent yields

$$\begin{aligned} g_n(z) &= \frac{1}{N}\sum_{i=1}^{N} q_{ii}(z) \\ &= \frac{1}{N}\sum_{i=1}^{N} \frac{1}{(1-c_n)\rho_i - z - zc_n\rho_i g_n(z)} + \varepsilon_N\ , \qquad c_n = \frac{N}{n} \\ &= \frac{1}{N}\mathrm{Trace}\left[(1-c_n)\mathbf{R}_N - z\mathbf{I}_N - zc_n g_n(z)\mathbf{R}_N\right]^{-1} + \varepsilon_N \end{aligned}$$

Guessing the canonical equation I: diagonal case

- Consider first the case where $\mathbf{R}_N$ is diagonal:

$$\mathbf{R}_N = \mathrm{diag}\left(\rho_i\,,\ i = 1:N\right)\ .$$

- The method of approximating the diagonal elements of the resolvent yields

$$\begin{aligned} g_n(z) &= \frac{1}{N}\sum_{i=1}^{N} q_{ii}(z) \\ &= \frac{1}{N}\sum_{i=1}^{N} \frac{1}{(1-c_n)\rho_i - z - zc_n\rho_i g_n(z)} + \varepsilon_N\ , \qquad c_n = \frac{N}{n} \\ &= \frac{1}{N}\mathrm{Trace}\left[(1-c_n)\mathbf{R}_N - z\mathbf{I}_N - zc_n g_n(z)\mathbf{R}_N\right]^{-1} + \varepsilon_N \end{aligned}$$

Hence the canonical equation (unknown $\mathbf{t}_N$):

$$\boxed{\mathbf{t}_N(z) = \frac{1}{N}\mathrm{Trace}\left[(1-c_n)\mathbf{R}_N - z\mathbf{I}_N - zc_n\mathbf{t}_N(z)\mathbf{R}_N\right]^{-1}}$$

Guessing the canonical equation II: non-diagonal cases

Gaussian entries

Assume that the entries of $\mathbf{X}_N$ are $\mathcal{N}(0,1)$ i.i.d. and consider the spectral decomposition of matrix

$$\mathbf{R}_N = \mathbf{O}_N^* \mathbf{\Lambda}_N \mathbf{O}_N$$

Due to Gaussian unitary invariance,

$$\text{spectrum}\left(\frac{1}{n}\mathbf{R}_N^{1/2}\mathbf{X}_N\mathbf{X}_N^*\mathbf{R}_N^{1/2}\right) \quad = \quad \text{spectrum}\left(\frac{1}{n}\mathbf{\Lambda}_N^{1/2}\widetilde{\mathbf{X}}_N\widetilde{\mathbf{X}}_N^*\mathbf{\Lambda}_N^{1/2}\right)$$

where $\widetilde{\mathbf{X}}_N$ has $\mathcal{N}(0,1)$ i.i.d. entries. Remember $\mathcal{L}(\mathbf{X}_N) = \mathcal{L}(\mathbf{O}_N\mathbf{X}_N)$!

For Gaussian entries, **sufficient** to consider **diagonal population covariance matrices** $\mathbf{R}_N$

Non-Gaussian entries

- Let $\mathbf{X}_N = [\vec{\mathbf{x}}_1, \cdots, \vec{\mathbf{x}}_n]$ the matrix with non-gaussian entries
- Let $\mathbf{X}_N^{\mathcal{N}} = [\vec{\mathbf{x}}_1^{\mathcal{N}}, \cdots, \vec{\mathbf{x}}_n^{\mathcal{N}}]$ the matrix with $\mathcal{N}(0,1)$ i.i.d. entries

Interpolate between $\mathbf{X}_N$ and $\mathbf{X}_N^{\mathcal{N}}$ by changing one column at a time

Guessing the canonical equation III

- We have proved so far that the Stieltjes transform $g_n(z)$ approximately satisfies the **canonical equation** (unknown $\mathbf{t}_N$):

Guessing the canonical equation III

- We have proved so far that the Stieltjes transform $g_n(z)$ approximately satisfies the **canonical equation** (unknown $\mathbf{t}_N$):

$$\boxed{\mathbf{t}_N(z) = \frac{1}{N}\mathrm{Trace}\left[(1-c_n)\mathbf{R}_N - z\mathbf{I}_N - zc_n\mathbf{t}_N(z)\mathbf{R}_N\right]^{-1}}$$

Guessing the canonical equation III

- We have proved so far that the Stieltjes transform $g_n(z)$ approximately satisfies the **canonical equation** (unknown $\mathbf{t}_N$):

$$\boxed{\mathbf{t}_N(z) = \frac{1}{N} \mathrm{Trace} \left[(1 - c_n) \mathbf{R}_N - z \mathbf{I}_N - z c_n \mathbf{t}_N(z) \mathbf{R}_N \right]^{-1}}$$

Problem: this equation depends on N!

Guessing the canonical equation III

- We have proved so far that the Stieltjes transform $g_n(z)$ approximately satisfies the **canonical equation** (unknown $\mathbf{t}_N$):

$$\boxed{\mathbf{t}_N(z) = \frac{1}{N}\mathrm{Trace}\left[(1-c_n)\mathbf{R}_N - z\mathbf{I}_N - zc_n\mathbf{t}_N(z)\mathbf{R}_N\right]^{-1}}$$

Problem: this equation depends on N!

Instead of having a single equation which describes the limit, we handle a sequence of equations whose solutions are refered to as **deterministic equivalents**.

Deterministic equivalents

- Let $\mathbf{t}_N$ be the Stieltjes transform solution of the canonical equation

$$\boxed{\mathbf{t}_N(z) = \frac{1}{N}\mathrm{Trace}\left[(1-c_n)\mathbf{R}_N - z\mathbf{I}_N - zc_n\mathbf{t}_N(z)\mathbf{R}_N\right]^{-1}}$$

  Consider associated probability $\mathbb{P}_N$ defined by

$$\mathbb{P}_N \quad = \quad (\text{Stieltjes transform})^{-1}(\mathbf{t}_N) \quad \text{i.e.} \quad \mathbf{t}_N(z) \; = \int \frac{\mathbb{P}_N(d\,\lambda)}{\lambda - z}$$

- Then $\mathbf{t}_N$ and $\mathbb{P}_N$ are the **determinitic equivalents** of g_n and L_N:

$$g_N(z) - \mathbf{t}_N(z) \quad \xrightarrow[N,n\to\infty]{a.s.} \quad 0\ ,$$

$$\frac{1}{N}\sum_{i=1}^{N} f(\lambda_i) - \int f(\lambda)\mathbb{P}_N(d\,\lambda) \quad \xrightarrow[N,n\to\infty]{a.s.} \quad 0\ ,$$

Genuine limits I

- In the case of Marčenko and Pastur, we have a single equation and a single limit

$$\boxed{\mathbf{g}_{\check{\mathrm{MP}}}(z) = \frac{1}{\sigma^2(1-c) - z - z\sigma^2 c\,\mathbf{g}_{\check{\mathrm{MP}}}(z)}}$$

- In the case of large covariance matrices, we have a **sequence** of equations:

$$\boxed{\mathbf{t}_N(z) = \frac{1}{N}\mathrm{Trace}\left[(1-c_n)\mathbf{R}_N - z\mathbf{I}_N - zc_n\mathbf{t}_N(z)\mathbf{R}_N\right]^{-1}}$$

  and we speak of deterministic equivalents rather than genuine limits.
- Notice that all these equations only depend on the **spectrum** of $\mathbf{R}_N$; denote by

$$L_N^{\mathbf{R}} = \frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i(\mathbf{R}_N)}$$

  the spectral measure of matrix $\mathbf{R}_N$; assume the following convergence:

$$\boxed{L_N^{\mathbf{R}} \xrightarrow[N,n\to\infty]{a.s.} \mathbb{P}^{\mathbf{R}}}$$

  where $\mathbb{P}^{\mathbf{R}}$ is a given probability distribution.

Genuine limits II

Theorem

If

$$L_N^{\mathbf{R}} \xrightarrow[N,n\to\infty]{a.s.} \mathbb{P}^{\mathbf{R}}$$

then the sequence of canonical equations "converges" to the following fixed-point equation

$$\boxed{\mathbf{t}(z) = \int \frac{\mathbb{P}^{\mathbf{R}}(d\,\lambda)}{(1-c)\lambda - z - zc\mathbf{t}(z)\lambda}} \quad \text{where} \quad \mathbf{t}(z) = \int \frac{\mathbb{P}_\infty(\,d\,\lambda)}{\lambda - z}$$

and the following convergences hold true

$$g_N(z) \quad \xrightarrow[N,n\to\infty]{a.s.} \quad \mathbf{t}(z)\ ,$$

$$\frac{1}{N}\sum_{i=1}^{N} f(\lambda_i) \quad \xrightarrow[N,n\to\infty]{a.s.} \quad \int f(\lambda)\mathbb{P}_\infty(d\,\lambda)\ ,$$

where the λ_i's are the eigenvalues of $\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*$.

Remarks

1. In general, there is no explicit solution to the equation

$$\boxed{\mathbf{t}(z) = \int \frac{\mathbb{P}^{\mathbf{R}}(d\,\lambda)}{(1-c)\lambda - z - zc\mathbf{t}(z)\lambda}}$$

2. In the theory of free probability, probability measure

$$\mathbb{P}_\infty = (ST)^{-1}(\mathbf{t})$$

is the **free multiplicative convolution** of $\mathbb{P}^{\mathbf{R}}$ with $\mathbb{P}_{\check{\mathrm{MP}}}$:

$$\boxed{\mathbb{P}_\infty = \mathbb{P}^{\mathbf{R}} \boxtimes \mathbb{P}_{\check{\mathrm{MP}}}}$$

Simulations

- Consider the distribution

$$\mathbb{P}^{\mathbf{R}} = \frac{1}{3}\delta_1 + \frac{1}{3}\delta_3 + \frac{1}{3}\delta_7$$

corresponding to a covariance matrix

$$\mathbf{R}_N = \mathrm{diag}(1, 3, 7)$$

each with multiplicity $\approx \frac{N}{3}$.

- We plot hereafter

$$\mathbb{P}_\infty = \mathbb{P}^{\mathbf{R}} \boxtimes \mathbb{P}_{\check{\mathrm{M}}\mathrm{P}}$$

for different values of c.

$$\boxed{\mathbf{t}(z) = \frac{\mathbf{1}}{\mathbf{3}}\left\{\frac{1}{(1-c)\boldsymbol{\lambda}_1 - z - zc\mathbf{t}(z)\boldsymbol{\lambda}_1} + \frac{1}{(1-c)\boldsymbol{\lambda}_2 - z - zc\mathbf{t}(z)\boldsymbol{\lambda}_2} + \frac{1}{(1-c)\boldsymbol{\lambda}_3 - z - zc\mathbf{t}(z)\boldsymbol{\lambda}_3}\right\}}$$

Simulations

- Consider the distribution
$$\mathbb{P}^{\mathbf{R}} = \frac{1}{3}\delta_1 + \frac{1}{3}\delta_3 + \frac{1}{3}\delta_7$$
corresponding to a covariance matrix
$$\mathbf{R}_N = \mathrm{diag}(1, 3, 7)$$
each with multiplicity $\approx \frac{N}{3}$.
- We plot hereafter
$$\mathbb{P}_\infty = \mathbb{P}^{\mathbf{R}} \boxtimes \mathbb{P}_{\check{\mathrm{MP}}}$$
for different values of c.

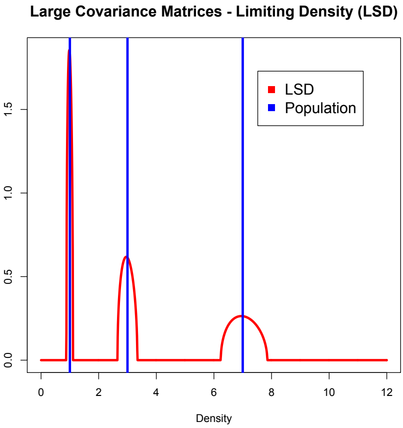

Figure: Plot of the Limiting Spectral Measure for $c = 0.01$

$$\boxed{\mathbf{t}(z) = \frac{\mathbf{1}}{\mathbf{3}}\left\{\frac{1}{(1-c)\boldsymbol{\lambda}_1 - z - zc\mathbf{t}(z)\boldsymbol{\lambda}_1} + \frac{1}{(1-c)\boldsymbol{\lambda}_2 - z - zc\mathbf{t}(z)\boldsymbol{\lambda}_2} + \frac{1}{(1-c)\boldsymbol{\lambda}_3 - z - zc\mathbf{t}(z)\boldsymbol{\lambda}_3}\right\}}$$

Simulations

- Consider the distribution

$$\mathbb{P}^{\mathbf{R}} = \frac{1}{3}\delta_1 + \frac{1}{3}\delta_3 + \frac{1}{3}\delta_7$$

corresponding to a covariance matrix

$$\mathbf{R}_N = \mathrm{diag}(1, 3, 7)$$

each with multiplicity $\approx \frac{N}{3}$.

- We plot hereafter

$$\mathbb{P}_\infty = \mathbb{P}^{\mathbf{R}} \boxtimes \mathbb{P}_{\check{\mathrm{MP}}}$$

for different values of c.

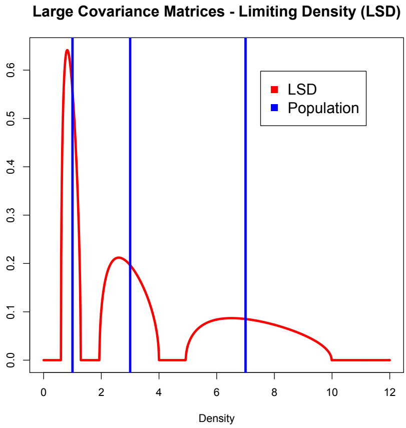

Figure: Plot of the Limiting Spectral Measure for $c = 0.1$

$$\boxed{\mathbf{t}(z) = \frac{\mathbf{1}}{\mathbf{3}}\left\{\frac{1}{(1-c)\boldsymbol{\lambda}_1 - z - zc\mathbf{t}(z)\boldsymbol{\lambda}_1} + \frac{1}{(1-c)\boldsymbol{\lambda}_2 - z - zc\mathbf{t}(z)\boldsymbol{\lambda}_2} + \frac{1}{(1-c)\boldsymbol{\lambda}_3 - z - zc\mathbf{t}(z)\boldsymbol{\lambda}_3}\right\}}$$

Simulations

- Consider the distribution
$$\mathbb{P}^{\mathbf{R}} = \frac{1}{3}\delta_1 + \frac{1}{3}\delta_3 + \frac{1}{3}\delta_7$$
corresponding to a covariance matrix
$$\mathbf{R}_N = \mathrm{diag}(1,3,7)$$
each with multiplicity $\approx \frac{N}{3}$.
- We plot hereafter
$$\mathbb{P}_\infty = \mathbb{P}^{\mathbf{R}} \boxtimes \mathbb{P}_{\check{\mathrm{MP}}}$$
for different values of c.

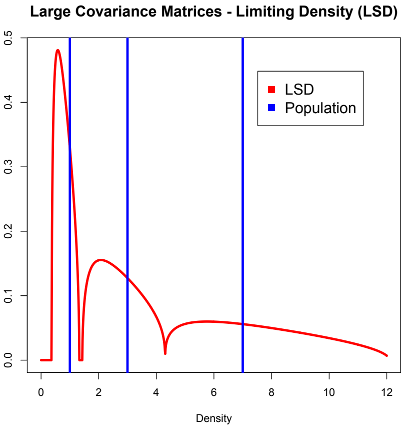

Figure: Plot of the Limiting Spectral Measure for $c = 0.25$

$$\boxed{\mathbf{t}(z) = \frac{\mathbf{1}}{\mathbf{3}}\left\{\frac{1}{(1-c)\boldsymbol{\lambda}_1 - z - zc\mathbf{t}(z)\boldsymbol{\lambda}_1} + \frac{1}{(1-c)\boldsymbol{\lambda}_2 - z - zc\mathbf{t}(z)\boldsymbol{\lambda}_2} + \frac{1}{(1-c)\boldsymbol{\lambda}_3 - z - zc\mathbf{t}(z)\boldsymbol{\lambda}_3}\right\}}$$

Simulations

- Consider the distribution
$$\mathbb{P}^{\mathbf{R}} = \frac{1}{3}\delta_1 + \frac{1}{3}\delta_3 + \frac{1}{3}\delta_7$$
corresponding to a covariance matrix
$$\mathbf{R}_N = \mathrm{diag}(1,3,7)$$
each with multiplicity $\approx \frac{N}{3}$.
- We plot hereafter
$$\mathbb{P}_\infty = \mathbb{P}^{\mathbf{R}} \boxtimes \mathbb{P}_{\check{\mathrm{MP}}}$$
for different values of c.

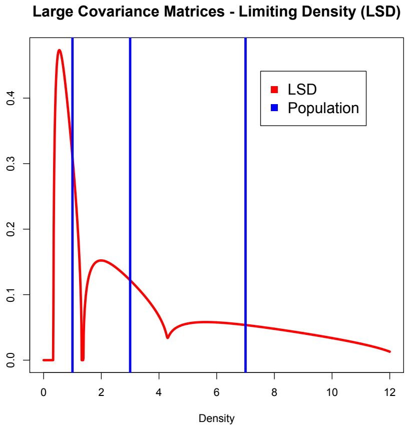

Figure: Plot of the Limiting Spectral Measure for $c = 0.275$

$$\boxed{\mathbf{t}(z) = \frac{\mathbf{1}}{\mathbf{3}}\left\{\frac{1}{(1-c)\boldsymbol{\lambda}_1 - z - zc\mathbf{t}(z)\boldsymbol{\lambda}_1} + \frac{1}{(1-c)\boldsymbol{\lambda}_2 - z - zc\mathbf{t}(z)\boldsymbol{\lambda}_2} + \frac{1}{(1-c)\boldsymbol{\lambda}_3 - z - zc\mathbf{t}(z)\boldsymbol{\lambda}_3}\right\}}$$

Simulations

- Consider the distribution

$$\mathbb{P}^{\mathbf{R}} = \frac{1}{3}\delta_1 + \frac{1}{3}\delta_3 + \frac{1}{3}\delta_7$$

corresponding to a covariance matrix

$$\mathbf{R}_N = \mathrm{diag}(1, 3, 7)$$

each with multiplicity $\approx \frac{N}{3}$.

- We plot hereafter

$$\mathbb{P}_\infty = \mathbb{P}^{\mathbf{R}} \boxtimes \mathbb{P}_{\check{\mathrm{MP}}}$$

for different values of c.

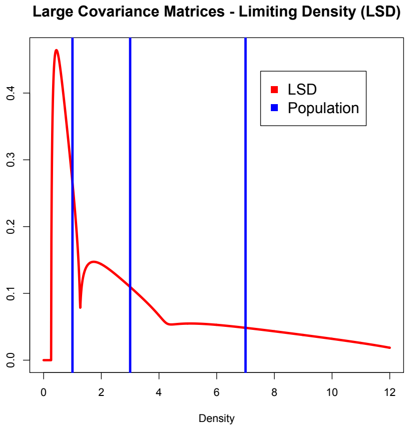

Figure: Plot of the Limiting Spectral Measure for $c = 0.35$

$$\boxed{\mathbf{t}(z) = \frac{\mathbf{1}}{\mathbf{3}}\left\{\frac{1}{(1-c)\boldsymbol{\lambda}_1 - z - zc\mathbf{t}(z)\boldsymbol{\lambda}_1} + \frac{1}{(1-c)\boldsymbol{\lambda}_2 - z - zc\mathbf{t}(z)\boldsymbol{\lambda}_2} + \frac{1}{(1-c)\boldsymbol{\lambda}_3 - z - zc\mathbf{t}(z)\boldsymbol{\lambda}_3}\right\}}$$

Simulations

- Consider the distribution

$$\mathbb{P}^{\mathbf{R}} = \frac{1}{3}\delta_1 + \frac{1}{3}\delta_3 + \frac{1}{3}\delta_7$$

corresponding to a covariance matrix

$$\mathbf{R}_N = \mathrm{diag}(1,3,7)$$

each with multiplicity $\approx \frac{N}{3}$.

- We plot hereafter

$$\mathbb{P}_\infty = \mathbb{P}^{\mathbf{R}} \boxtimes \mathbb{P}_{\check{\mathrm{MP}}}$$

for different values of c.

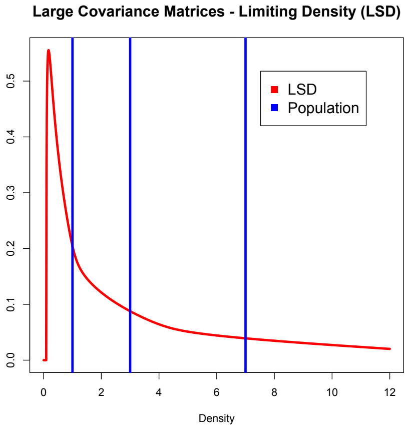

Figure: Plot of the Limiting Spectral Measure for $c = 0.6$

$$\boxed{\mathbf{t}(z) = \frac{\mathbf{1}}{\mathbf{3}}\left\{\frac{1}{(1-c)\boldsymbol{\lambda}_1 - z - zc\mathbf{t}(z)\boldsymbol{\lambda}_1} + \frac{1}{(1-c)\boldsymbol{\lambda}_2 - z - zc\mathbf{t}(z)\boldsymbol{\lambda}_2} + \frac{1}{(1-c)\boldsymbol{\lambda}_3 - z - zc\mathbf{t}(z)\boldsymbol{\lambda}_3}\right\}}$$

Example: Marčenko-Pastur's model

In the case of Marčenko-Pastur, $\mathbf{R}_N = \sigma^2 \mathbf{I}_N$ and many things simplify:

- The deterministic equivalent of $g_n(z)$ is $\mathbf{t}_N(z)$, solution of:

$$\boxed{\mathbf{t}_N(z) = \frac{1}{\sigma^2(1-c_N) - z - zc_N\sigma^2\mathbf{t}_N(z)}} \quad \text{with} \quad c_N = \frac{N}{n}\ .$$

- The deterministic equivalent of $L_N = \frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i}$ is

$$\boxed{\mathbb{P}^N_{\check{\mathrm{M}}\mathrm{P}}(dx) \ = \ \left(1 - \frac{1}{c_N}\right)^+ \delta_0(dx) + \frac{\sqrt{(b_N - x)(x - a_N)}}{2\pi\sigma^2 x c_N} 1_{[a,b]}(x)\, dx}$$

  with $a_N = \sigma^2(1-\sqrt{c_N})^2$ and $b_N = \sigma^2(1+\sqrt{c_N})^2$.

- Of course, all the genuine limits are obtained by replacing $c_N = \frac{N}{n}$ by $c = \lim \frac{N}{n}$.

Summary

- Consider Large Covariance Matrices

$$\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^* \quad \text{with} \quad \mathbf{Y}_N = \mathbf{R}_N^{1/2}\mathbf{X}_N$$

which model n samples of of N-dimensional observations $\mathbf{Y}_{\cdot i}$ with covariance

$$\mathrm{cov}(\mathbf{Y}_{\cdot i}) = \mathbf{R}_N \ .$$

in the **large dimensional regime** where $\boxed{N \propto n}$.

- The spectrum is described by a sequence of **fixed-point equations**

$$\boxed{\mathbf{t}_N(z) = \frac{1}{N}\mathrm{Trace}\left[(1-c_n)\mathbf{R}_N - z\mathbf{I}_N - zc_n\mathbf{t}_N(z)\mathbf{R}_N\right]^{-1}}$$

and we consider the associated **deterministic equivalents**

$$g_n(z) \sim \mathbf{t}_N(z)\ , \quad \mathbb{P}_N = (ST)^{-1}(\mathbf{t}_N) \sim L_N\ ,$$

- If the spectrum of $\mathbf{R}_N$ converges, we end up with a **single fixed-point equation**

$$\boxed{\mathbf{t}(z) = \int \frac{\mathbb{P}^{\mathbf{R}}(d\,\lambda)}{(1-c)\lambda - z - zc\mathbf{t}(z)\lambda}}$$

and genuine limits $g_n(z) \to \mathbf{t}(z)$ and $L_N \to (ST)^{-1}(\mathbf{t})$.

The largest eigenvalue in $\check{\text{M}}$P model

Given a $N \times n$ matrix $\mathbf{X}_N$ with i.i.d. entries $\mathbb{E}X_{ij} = 0$ and $\mathbb{E}|X_{ij}|^2 = \sigma^2$,

$$L_N \left(\frac{1}{n} \mathbf{X}_N \mathbf{X}_N^* \right) \xrightarrow[N,n\to\infty]{} \mathbb{P}_{\check{\text{M}}\text{P}}$$

where $\mathbb{P}_{\check{\text{M}}\text{P}}$ has support

$$\mathcal{S}_{\check{\text{M}}\text{P}} \;=\; \{0\} \;\cup\; \underbrace{\left[\sigma^2(1-\sqrt{c})^2 \;,\; \sigma^2(1+\sqrt{c})^2\right]}_{\textbf{bulk}}$$

(remove the set $\{0\}$ if $c < 1$)

Theorem

- Let $\mathbb{E}|X_{ij}|^4 < \infty$, then:

$$\lambda_{\max} \left(\frac{1}{n} \mathbf{X}_N \mathbf{X}_N^* \right) \xrightarrow[N,n\to\infty]{a.s.} \sigma^2(1+\sqrt{c})^2 \;.$$

Message: The largest eigenvalue converges to the right edge of the bulk.

Spiked Models I

Definition

Let $\mathbf{\Pi}_N$ be a **small perturbation of the identity**:

$$\mathbf{\Pi}_N = \mathbf{I}_N + \mathbf{P}_N \quad \text{where} \quad \mathbf{P}_N = \theta_1 \vec{\mathbf{u}}_1 \vec{\mathbf{u}}_1^* + \cdots + \theta_k \vec{\mathbf{u}}_k \vec{\mathbf{u}}_k^*$$

where k is **independent of the dimensions** N, n.

Consider

$$\boxed{\tilde{\mathbf{X}}_N = \mathbf{\Pi}_N^{1/2} \mathbf{X}_N}$$

This model will be refered to as a (multiplicative) **spiked model**.

Think of $\mathbf{\Pi}_N$ as

$$\mathbf{\Pi}_N = \begin{pmatrix} 1+\theta_1 & & & & \\ & \ddots & & & \\ & & 1+\theta_k & & \\ & & & 1 & \\ & & & & \ddots \end{pmatrix}$$

$\boxed{\textbf{Very important:} \text{ The number } k \text{ of perturbations is finite}}$

Spiked Models II

Remarks

- The spiked model is a particular case of **large covariance matrix** model with

$$\mathbf{R}_N = \mathbf{I}_N + \sum_{\ell=1}^{k} \theta_\ell \vec{\mathbf{u}}_\ell \vec{\mathbf{u}}_\ell^*$$

- There are additive spiked models: $\check{\mathbf{X}}_N = \mathbf{X}_N + \mathbf{A}_N$ where $\mathbf{A}_N$ is a matrix with finite rank.
- Spiked models have been introduced by Iain M. Johnstone in his paper

  On the distribution of the largest eigenvalue in principal components analysis, *Annals of Statistics*, **2001**.

  to take into account the fact that in many datasets, a small number of eigenvalues is "far away" the bulk of the other eigenvalues

Objective

- What is the influence of $\mathbf{\Pi}_N$ over $L_N\left(\frac{1}{n}\tilde{\mathbf{X}}_N\tilde{\mathbf{X}}_N^*\right)$?
- What is the influence of $\mathbf{\Pi}_N$ over $\lambda_{\max}\left(\frac{1}{n}\tilde{\mathbf{X}}_N\tilde{\mathbf{X}}_N^*\right)$?

Spiked Models II

Remarks

- The spiked model is a particular case of **large covariance matrix** model with

$$\mathbf{R}_N = \mathbf{I}_N + \sum_{\ell=1}^{k} \theta_\ell \vec{\mathbf{u}}_\ell \vec{\mathbf{u}}_\ell^*$$

- There are additive spiked models: $\check{\mathbf{X}}_N = \mathbf{X}_N + \mathbf{A}_N$ where $\mathbf{A}_N$ is a matrix with finite rank.
- Spiked models have been introduced by Iain M. Johnstone in his paper

  On the distribution of the largest eigenvalue in principal components analysis, *Annals of Statistics*, **2001**.

  to take into account the fact that in many datasets, a small number of eigenvalues is "far away" the bulk of the other eigenvalues

Objective

- What is the influence of $\mathbf{\Pi}_N$ over $L_N\left(\frac{1}{n}\tilde{\mathbf{X}}_N\tilde{\mathbf{X}}_N^*\right)$? None!
- What is the influence of $\mathbf{\Pi}_N$ over $\lambda_{\max}\left(\frac{1}{n}\tilde{\mathbf{X}}_N\tilde{\mathbf{X}}_N^*\right)$?

Spiked Models II

Remarks

- The spiked model is a particular case of **large covariance matrix** model with

$$\mathbf{R}_N = \mathbf{I}_N + \sum_{\ell=1}^{k} \theta_\ell \vec{\mathbf{u}}_\ell \vec{\mathbf{u}}_\ell^*$$

- There are additive spiked models: $\check{\mathbf{X}}_N = \mathbf{X}_N + \mathbf{A}_N$ where $\mathbf{A}_N$ is a matrix with finite rank.
- Spiked models have been introduced by Iain M. Johnstone in his paper

  On the distribution of the largest eigenvalue in principal components analysis, *Annals of Statistics*, **2001**.

  to take into account the fact that in many datasets, a small number of eigenvalues is "far away" the bulk of the other eigenvalues

Objective

- What is the influence of $\mathbf{\Pi}_N$ over $L_N\left(\frac{1}{n}\tilde{\mathbf{X}}_N\tilde{\mathbf{X}}_N^*\right)$? None!
- What is the influence of $\mathbf{\Pi}_N$ over $\lambda_{\max}\left(\frac{1}{n}\tilde{\mathbf{X}}_N\tilde{\mathbf{X}}_N^*\right)$? Well, it depends!

Simulations I: Single spikes

Simulations I: Single spikes

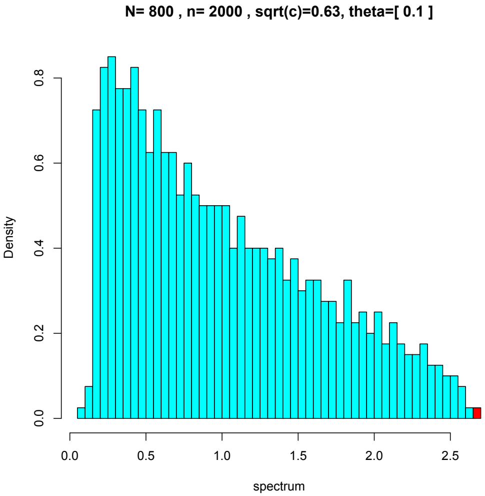

Figure: Spiked model - strength of the perturbation $\theta = 0.1$

Simulations I: Single spikes

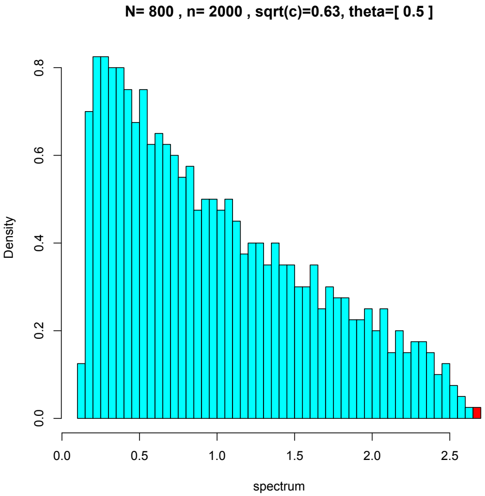

Figure: Spiked model - strength of the perturbation $\theta = 0.5$

Simulations I: Single spikes

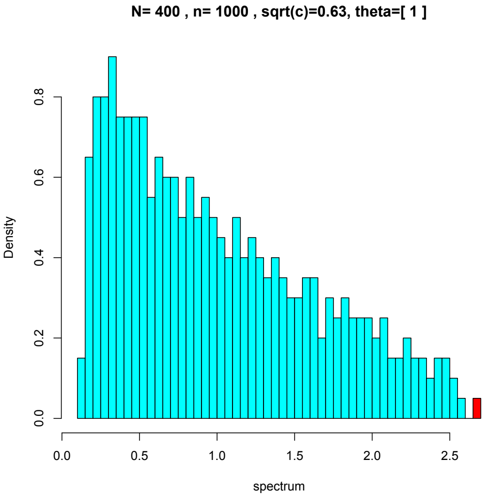

Figure: Spiked model - strength of the perturbation $\theta = 1$

Simulations I: Single spikes

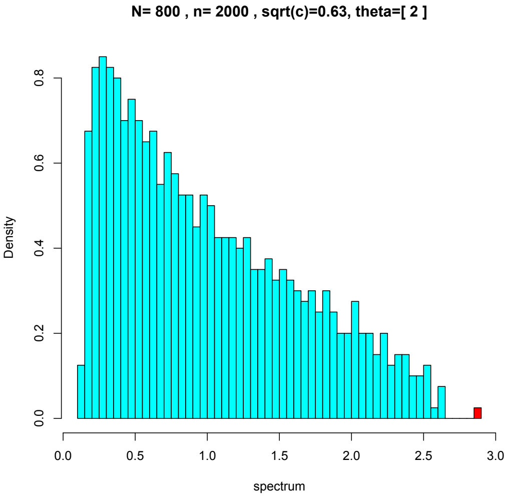

Figure: Spiked model - strength of the perturbation $\theta = 2$

Simulations I: Single spikes

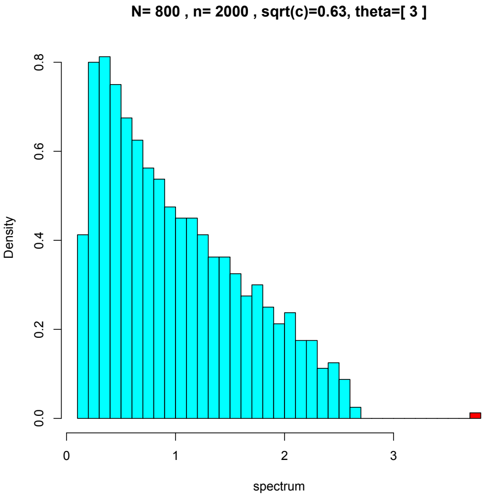

Figure: Spiked model - strength of the perturbation $\theta = 3$

Observation #1

If the **strength** θ of the perturbation $\mathbf{P}_N$ is large enough, then the limit of $\lambda_{\max}\left(\frac{1}{n}\tilde{\mathbf{X}}_N\tilde{\mathbf{X}}_N^*\right)$ is **strictly larger** than the right edge of the bulk.

Simulations II: Spectral measure

Simulations II: Spectral measure

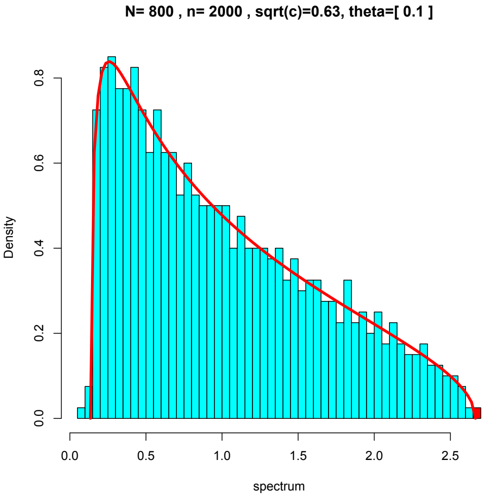

Figure: Spiked model - strength of the perturbation $\theta = 0.1$

Simulations II: Spectral measure

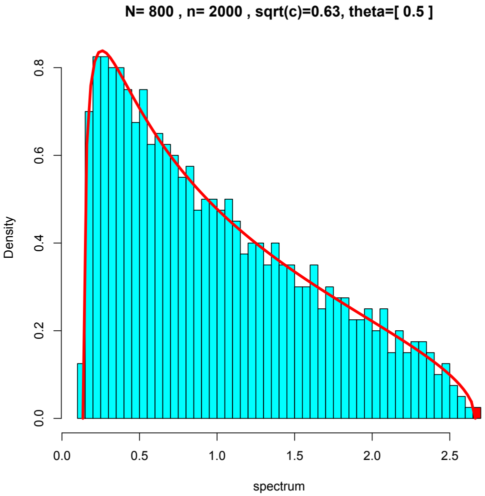

Figure: Spiked model - strength of the perturbation $\theta = 0.5$

Simulations II: Spectral measure

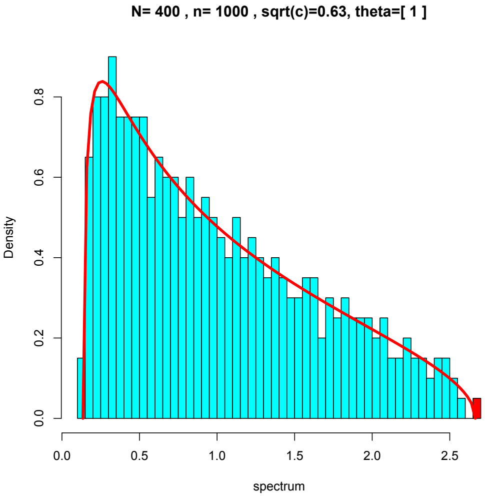

Figure: Spiked model - strength of the perturbation $\theta = 1$

Simulations II: Spectral measure

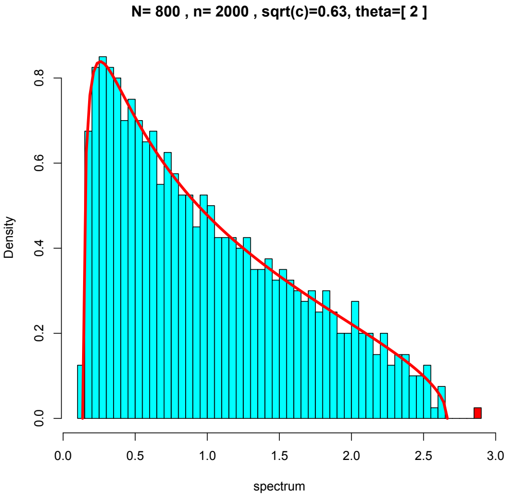

Figure: Spiked model - strength of the perturbation $\theta = 2$

Simulations II: Spectral measure

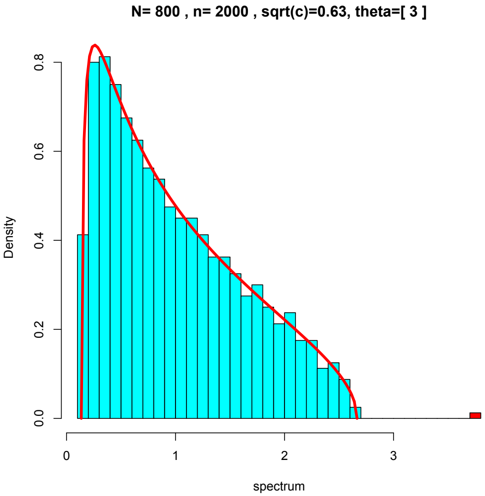

Figure: Spiked model - strength of the perturbation $\theta = 3$

Simulations III: Multiple Spikes

Simulations III: Multiple Spikes

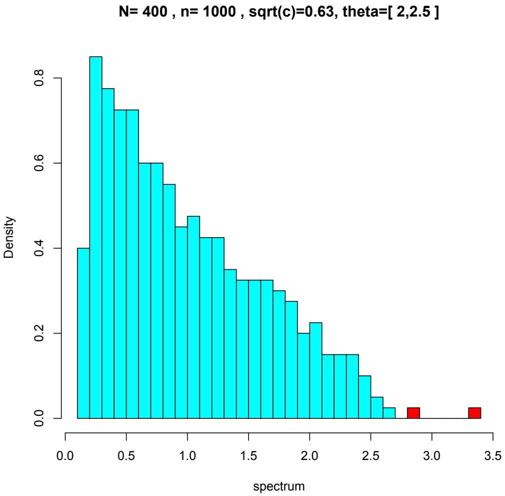

Figure: Spiked model - Two spikes

Simulations III: Multiple Spikes

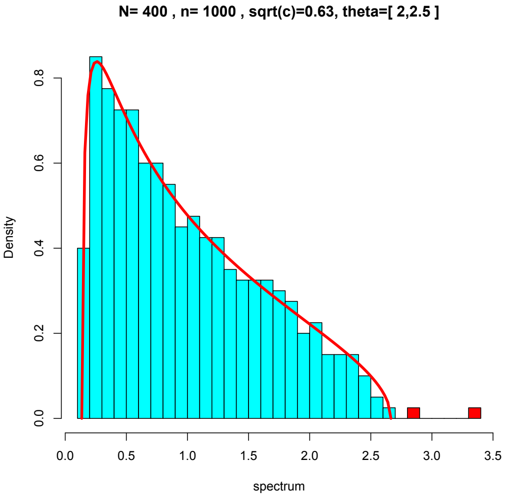

Figure: Spiked model - Two spikes

Simulations III: Multiple Spikes

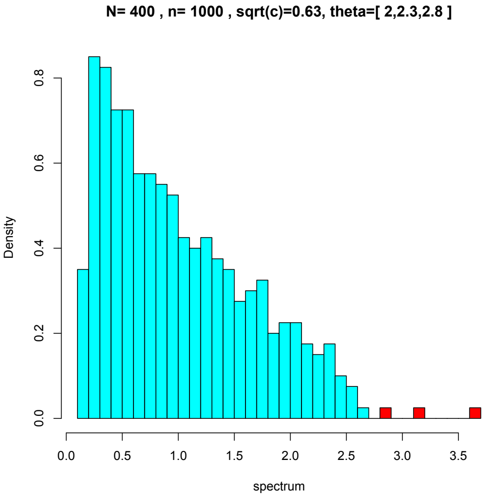

Figure: Spiked model - Three spikes

Simulations III: Multiple Spikes

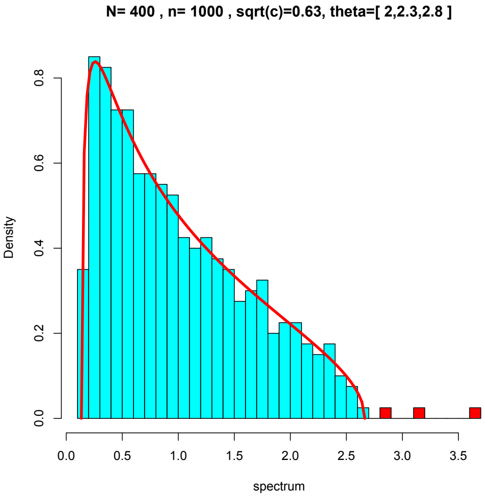

Figure: Spiked model - Three spikes

Simulations III: Multiple Spikes

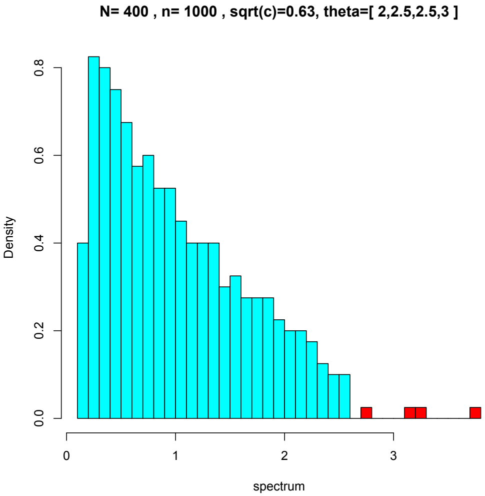

Figure: Spiked model - Multiple spikes

Simulations III: Multiple Spikes

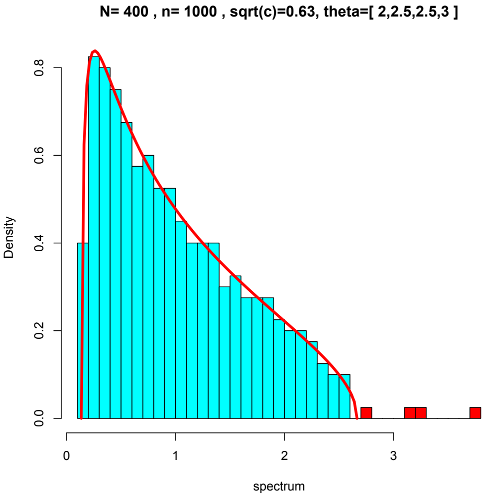

Figure: Spiked model - Multiple spikes

Observation # 2

Whathever the perturbations, the spectral measure converges toward Marčenko-Pastur distribution

The limiting spectral measure I

Theorem

The following convergence holds true: $\boxed{L_N\left(\frac{1}{n}\tilde{\mathbf{X}}_N\tilde{\mathbf{X}}_N^*\right) \xrightarrow[N,n\to\infty]{a.s.} \mathbb{P}_{\check{\mathrm{MP}}}\ .}$

Remark

The limiting spectral measure is not sensitive to the presence of spikes

The limiting spectral measure II

Proof

The spiked model is a particular case of **large covariance matrix** model with

$$\mathbf{R}_N = \mathbf{I}_N + \sum_{\ell=1}^{k} \theta_\ell \vec{\mathbf{u}}_\ell \vec{\mathbf{u}}_\ell^*$$

Consider the spectral measure of $\mathbf{R}_N$ (orthogonal eigenvectors for the perturbations assumed):

$$L_N^{\mathbf{R}} = \frac{1}{N}\sum_{i=1}^{k} \delta_{1+\theta_i} + \frac{1}{N}\sum_{i=k+1}^{N} \delta_1 \xrightarrow[N,n\to\infty]{} \mathbb{P}^{\mathbf{R}} = \delta_1$$

hence the limiting canonical equation

$$\begin{aligned} \mathbf{t}(z) &= \int \frac{\mathbb{P}^{\mathbf{R}}(d\,\lambda)}{(1-c)\lambda - z - zc\mathbf{t}(z)\lambda} = \frac{1}{(1-c) - z - zc\mathbf{t}(z)} \\ &\Leftrightarrow \boxed{zc\mathbf{t}^2 + [z-(1-c)]\mathbf{t} + 1 = 0} \end{aligned}$$

$\Rightarrow$ We recognize **Marčenko-Pastur canonical equation**.

Behaviour of the largest eigenvalue

We consider the following spiked model:

$$\tilde{\mathbf{X}}_N = (\mathbf{I}_N + \theta \vec{\mathbf{u}}\vec{\mathbf{u}}^*)^{1/2}\,\mathbf{X}_N \quad \text{with} \quad \|\vec{\mathbf{u}}\| = 1\ .$$

which corresponds to a **rank-one** perturbation.

Theorem

Recall that $c = \lim_{N,n\to\infty} \frac{N}{n}$.

- if $\boxed{\theta \leq \sqrt{c}}$ then

$$\lambda_{\max} = \lambda_{\max}\left(\frac{1}{n}\tilde{\mathbf{X}}_N \tilde{\mathbf{X}}_N^*\right) \xrightarrow[N,n\to\infty]{a.s.} \sigma^2(1+\sqrt{c})^2$$

- if $\boxed{\theta > \sqrt{c}}$ then

$$\lambda_{\max} \xrightarrow[N,n\to\infty]{a.s.} \sigma^2(1+\theta)\left(1+\frac{c}{\theta}\right) > \sigma^2(1+\sqrt{c})^2$$

Phase transition Phenomenon

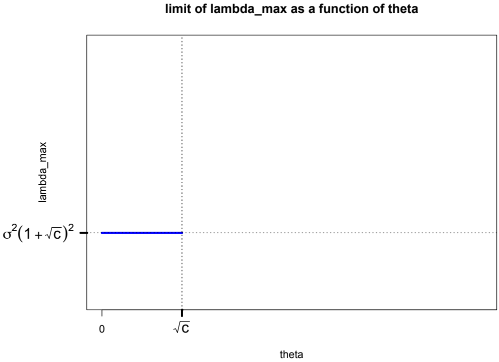

Figure: Limit of largest eigenvalue $\lambda_{\max}$ as a function of the perturbation θ

Phase transition Phenomenon

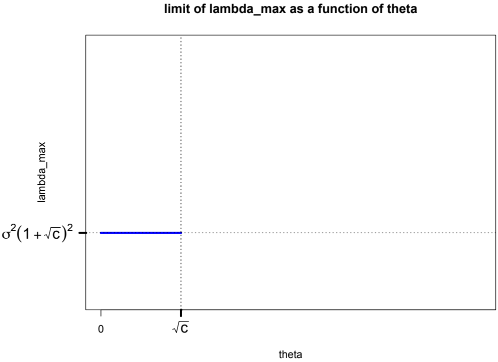

Figure: Limit of largest eigenvalue $\lambda_{\max}$ as a function of the perturbation θ

- If $\theta \leq \sqrt{c}$ then

$$\lambda_{\max}\left(\frac{1}{n}\tilde{\mathbf{X}}_N\tilde{\mathbf{X}}_N^*\right) \xrightarrow[N,n\to\infty]{} \sigma^2(1+\sqrt{c})^2 \ .$$

Phase transition Phenomenon

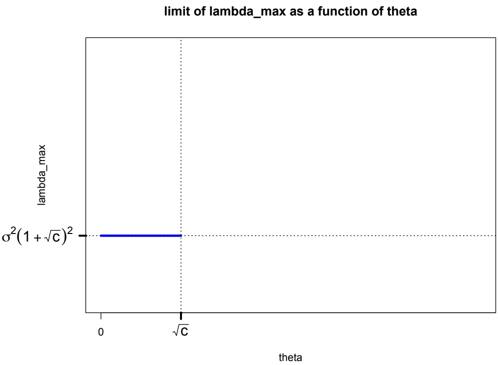

Figure: Limit of largest eigenvalue $\lambda_{\max}$ as a function of the perturbation θ

- If $\theta \leq \sqrt{c}$ then

$$\lambda_{\max}\left(\frac{1}{n}\tilde{\mathbf{X}}_N\tilde{\mathbf{X}}_N^*\right) \xrightarrow[N,n\to\infty]{} \sigma^2(1+\sqrt{c})^2 \ .$$

$\boxed{\text{Below the threshold } \sqrt{c},\ \lambda_{\max}\left(\tfrac{1}{n}\tilde{\mathbf{X}}_N\tilde{\mathbf{X}}_N^*\right) \text{ asymptotically } \textbf{sticks to the bulk}.}$

Phase transition Phenomenon

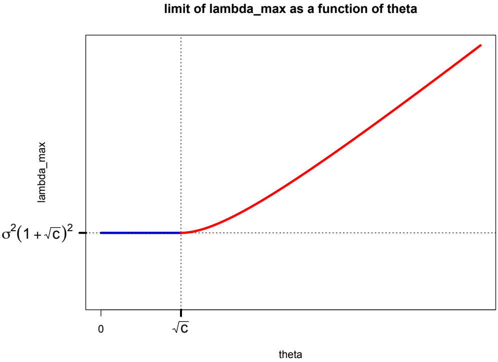

Figure: Limit of largest eigenvalue $\lambda_{\max}$ as a function of the perturbation θ

Phase transition Phenomenon

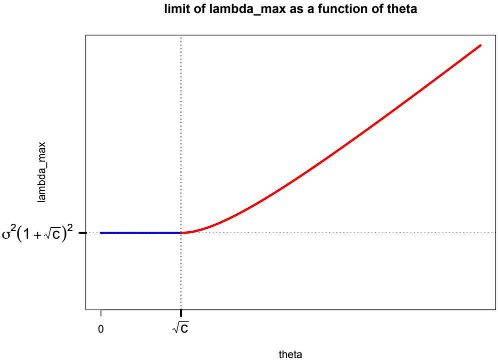

Figure: Limit of largest eigenvalue $\lambda_{\max}$ as a function of the perturbation θ

- if $\theta > \sqrt{c}$ then

$$\lim_{N,n} \lambda_{\max}\left(\frac{1}{n}\tilde{\mathbf{X}}_N\tilde{\mathbf{X}}_N^*\right) = \sigma^2(1+\theta)\left(1+\frac{c}{\theta}\right)$$

Phase transition Phenomenon

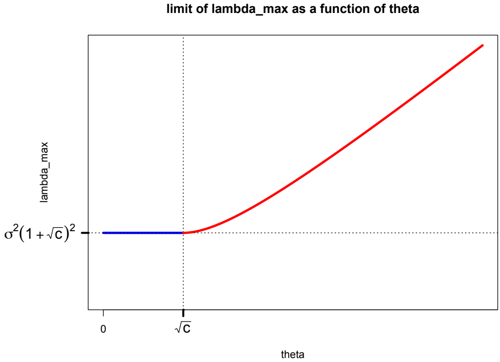

Figure: Limit of largest eigenvalue $\lambda_{\max}$ as a function of the perturbation θ

- if $\theta > \sqrt{c}$ then

$$\lim_{N,n} \lambda_{\max}\left(\frac{1}{n}\tilde{\mathbf{X}}_N\tilde{\mathbf{X}}_N^*\right) \quad = \quad \sigma^2(1+\theta)\left(1+\frac{c}{\theta}\right) > \sigma^2\left(1+\sqrt{c}\right)^2$$

Above the threshold $\sqrt{c}$, $\lambda_{\max}\left(\frac{1}{n}\tilde{\mathbf{X}}_N\tilde{\mathbf{X}}_N^*\right)$ asymptotically **separates from the bulk**.

Strategy of proof

1. We first express a condition for which

$$\lambda_{\max}\left(\frac{1}{n}\tilde{\mathbf{X}}_N\tilde{\mathbf{X}}_N^*\right)$$

separates from the bulk and refer to it as $\boxed{\textbf{the determinant condition}}$

2. Relying on Large Random Matrix theory, we simplify this condition and obtain

$$\boxed{\textbf{the asymptotic condition}}$$

3. We finally conclude, obtain the condition $\boxed{\theta > \sqrt{c}}$ for which the limit of $\lambda_{\max}\left(\frac{1}{n}\tilde{\mathbf{X}}_N\tilde{\mathbf{X}}_N^*\right)$ separates from the bulk, and compute this limit.

Notations

- Marčenko-Pastur model

$$\mathbf{Z}_N = \frac{1}{n}\mathbf{X}_N\mathbf{X}_N^* \qquad \text{and} \qquad \mathbf{Q}_N(z) = (-z\mathbf{I}_N + \mathbf{Y}_N)^{-1}$$

- Spiked model

$$\tilde{\mathbf{X}}_N = \mathbf{\Pi}^{1/2}\mathbf{X}_N = (\mathbf{I}_N + \theta\vec{\mathbf{u}}\vec{\mathbf{u}}^*)^{1/2}\,\mathbf{X}_N \qquad \text{and} \qquad \tilde{\mathbf{Z}}_N = \frac{1}{n}\tilde{\mathbf{X}}_N\tilde{\mathbf{X}}_N^*$$

The determinant condition I

We wish to find

- λ^θ eigenvalue of the spiked model

$$\tilde{\mathbf{Z}}_N = \frac{1}{n}\mathbf{\Pi}^{1/2}\mathbf{X}_N\mathbf{X}_N^*\mathbf{\Pi}^{1/2}$$

- λ^θ **not** an eigenvalue of M̌P model

$$\mathbf{Z}_N = \frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*$$

Otherwise stated

$$\boxed{\det\left(-\lambda^\theta\mathbf{I}_N + \tilde{\mathbf{Z}}_N\right) = 0} \quad \textbf{but} \quad \boxed{\det\left(-\lambda^\theta\mathbf{I}_N + \mathbf{Z}_N\right) \neq 0}$$

Inverse of a rank-one perturbation of the identity

Recall that

$$\mathbf{\Pi}_N = \mathbf{I}_N + \theta\vec{\mathbf{u}}\vec{\mathbf{u}}^*$$

Standard results from linear algebra yield

$$\mathbf{\Pi}_N^{-1} = (\mathbf{I}_N + \theta\vec{\mathbf{u}}\vec{\mathbf{u}}^*)^{-1} = \mathbf{I}_N - \frac{\theta}{1+\theta}\vec{\mathbf{u}}\vec{\mathbf{u}}^*$$

The determinant condition II

Let's go for **simple** computations:

$$\det\left(-\lambda^{\theta}\mathbf{I}_N + \tilde{\mathbf{Z}}_N\right) = 0 \quad \Leftrightarrow \quad \det\left(-\lambda^{\theta}\mathbf{I}_N + \mathbf{\Pi}_N^{1/2}\mathbf{Z}_N\mathbf{\Pi}_N^{1/2}\right) = 0$$

The determinant condition II

Let's go for **simple** computations:

$$\det\left(-\lambda^{\theta}\mathbf{I}_N + \tilde{\mathbf{Z}}_N\right) = 0 \quad \Leftrightarrow \quad \det\left(-\lambda^{\theta}\mathbf{I}_N + \mathbf{\Pi}_N^{1/2}\mathbf{Z}_N\mathbf{\Pi}_N^{1/2}\right) = 0$$

$$\Leftrightarrow \quad \det\left(-\lambda^{\theta}\mathbf{\Pi}_N^{-1} + \mathbf{Z}_N\right) = 0$$

The determinant condition II

Let's go for **simple** computations:

$$\begin{aligned}
\det\left(-\lambda^\theta \mathbf{I}_N + \tilde{\mathbf{Z}}_N\right) = 0 \quad &\Leftrightarrow \quad \det\left(-\lambda^\theta \mathbf{I}_N + \mathbf{\Pi}_N^{1/2}\mathbf{Z}_N\mathbf{\Pi}_N^{1/2}\right) = 0 \\
&\Leftrightarrow \quad \det\left(-\lambda^\theta \mathbf{\Pi}_N^{-1} + \mathbf{Z}_N\right) = 0 \\
&\Leftrightarrow \quad \det\left(-\lambda^\theta \left(\mathbf{I}_N - \frac{\theta}{1+\theta}\vec{\mathbf{u}}\vec{\mathbf{u}}^*\right) + \mathbf{Z}_N\right) = 0
\end{aligned}$$

The determinant condition II

Let's go for **simple** computations:

$$
\begin{aligned}
\det\left(-\lambda^\theta \mathbf{I}_N + \tilde{\mathbf{Z}}_N\right) = 0 \quad &\Leftrightarrow \quad \det\left(-\lambda^\theta \mathbf{I}_N + \mathbf{\Pi}_N^{1/2}\mathbf{Z}_N\mathbf{\Pi}_N^{1/2}\right) = 0 \\
&\Leftrightarrow \quad \det\left(-\lambda^\theta \mathbf{\Pi}_N^{-1} + \mathbf{Z}_N\right) = 0 \\
&\Leftrightarrow \quad \det\left(-\lambda^\theta\left(\mathbf{I}_N - \frac{\theta}{1+\theta}\vec{\mathbf{u}}\vec{\mathbf{u}}^*\right) + \mathbf{Z}_N\right) = 0 \\
&\Leftrightarrow \quad \det\left(-\lambda^\theta \mathbf{I}_N + \mathbf{Z}_N + \lambda^\theta \frac{\theta}{1+\theta}\vec{\mathbf{u}}\vec{\mathbf{u}}^*\right) = 0
\end{aligned}
$$

The determinant condition II

Let's go for **simple** computations:

$$
\begin{aligned}
\det\left(-\lambda^\theta \mathbf{I}_N + \tilde{\mathbf{Z}}_N\right) = 0 \quad &\Leftrightarrow \quad \det\left(-\lambda^\theta \mathbf{I}_N + \mathbf{\Pi}_N^{1/2}\mathbf{Z}_N\mathbf{\Pi}_N^{1/2}\right) = 0 \\
&\Leftrightarrow \quad \det\left(-\lambda^\theta \mathbf{\Pi}_N^{-1} + \mathbf{Z}_N\right) = 0 \\
&\Leftrightarrow \quad \det\left(-\lambda^\theta\left(\mathbf{I}_N - \frac{\theta}{1+\theta}\vec{\mathbf{u}}\vec{\mathbf{u}}^*\right) + \mathbf{Z}_N\right) = 0 \\
&\Leftrightarrow \quad \det\left(-\lambda^\theta \mathbf{I}_N + \mathbf{Z}_N + \lambda^\theta\frac{\theta}{1+\theta}\vec{\mathbf{u}}\vec{\mathbf{u}}^*\right) = 0 \\
&\Leftrightarrow \quad \det\left[\left(-\lambda^\theta \mathbf{I}_N + \mathbf{Z}_N\right)\left(\mathbf{I}_N + \lambda^\theta\frac{\theta}{1+\theta}\vec{\mathbf{u}}\vec{\mathbf{u}}^*\mathbf{Q}_N(\lambda^\theta)\right)\right] = 0
\end{aligned}
$$

The determinant condition II

Let's go for **simple** computations:

$$
\begin{aligned}
\det\left(-\lambda^\theta \mathbf{I}_N + \tilde{\mathbf{Z}}_N\right) = 0 \quad &\Leftrightarrow \quad \det\left(-\lambda^\theta \mathbf{I}_N + \mathbf{\Pi}_N^{1/2}\mathbf{Z}_N\mathbf{\Pi}_N^{1/2}\right) = 0 \\
&\Leftrightarrow \quad \det\left(-\lambda^\theta \mathbf{\Pi}_N^{-1} + \mathbf{Z}_N\right) = 0 \\
&\Leftrightarrow \quad \det\left(-\lambda^\theta\left(\mathbf{I}_N - \frac{\theta}{1+\theta}\vec{\mathbf{u}}\vec{\mathbf{u}}^*\right) + \mathbf{Z}_N\right) = 0 \\
&\Leftrightarrow \quad \det\left(-\lambda^\theta \mathbf{I}_N + \mathbf{Z}_N + \lambda^\theta\frac{\theta}{1+\theta}\vec{\mathbf{u}}\vec{\mathbf{u}}^*\right) = 0 \\
&\Leftrightarrow \quad \det\left[\left(-\lambda^\theta \mathbf{I}_N + \mathbf{Z}_N\right)\left(\mathbf{I}_N + \lambda^\theta\frac{\theta}{1+\theta}\vec{\mathbf{u}}\vec{\mathbf{u}}^*\mathbf{Q}_N(\lambda^\theta)\right)\right] = 0 \\
&\Leftrightarrow \quad \boxed{\det\left[\mathbf{I}_N + \lambda^\theta\frac{\theta}{1+\theta}\vec{\mathbf{u}}\vec{\mathbf{u}}^*\mathbf{Q}_N(\lambda^\theta)\right] = 0}
\end{aligned}
$$

The determinant condition II

Let's go for **simple** computations:

$$
\begin{aligned}
\det\left(-\lambda^\theta \mathbf{I}_N + \tilde{\mathbf{Z}}_N\right) = 0 \quad \Leftrightarrow \quad & \det\left(-\lambda^\theta \mathbf{I}_N + \mathbf{\Pi}_N^{1/2} \mathbf{Z}_N \mathbf{\Pi}_N^{1/2}\right) = 0 \\
\Leftrightarrow \quad & \det\left(-\lambda^\theta \mathbf{\Pi}_N^{-1} + \mathbf{Z}_N\right) = 0 \\
\Leftrightarrow \quad & \det\left(-\lambda^\theta \left(\mathbf{I}_N - \frac{\theta}{1+\theta}\vec{\mathbf{u}}\vec{\mathbf{u}}^*\right) + \mathbf{Z}_N\right) = 0 \\
\Leftrightarrow \quad & \det\left(-\lambda^\theta \mathbf{I}_N + \mathbf{Z}_N + \lambda^\theta \frac{\theta}{1+\theta}\vec{\mathbf{u}}\vec{\mathbf{u}}^*\right) = 0 \\
\Leftrightarrow \quad & \det\left[\left(-\lambda^\theta \mathbf{I}_N + \mathbf{Z}_N\right)\left(\mathbf{I}_N + \lambda^\theta \frac{\theta}{1+\theta}\vec{\mathbf{u}}\vec{\mathbf{u}}^* \mathbf{Q}_N(\lambda^\theta)\right)\right] = 0 \\
\Leftrightarrow \quad & \boxed{\det\left[\mathbf{I}_N + \lambda^\theta \frac{\theta}{1+\theta}\vec{\mathbf{u}}\vec{\mathbf{u}}^* \mathbf{Q}_N(\lambda^\theta)\right] = 0}
\end{aligned}
$$

Interest of this expression

In this equation, perturbation features θ and $\vec{\mathbf{u}}$ are separated from the resolvent of M̌P model (non-spiked model)

The determinant condition III

Recall the condition

$$\boxed{\det\left[\mathbf{I}_N + \lambda^\theta \frac{\theta}{1+\theta}\vec{\mathbf{u}}\vec{\mathbf{u}}^* \mathbf{Q}_N(\lambda^\theta)\right] = 0}$$

Matrix

$$\lambda^\theta \frac{\theta}{1+\theta}\vec{\mathbf{u}}\vec{\mathbf{u}}^* \mathbf{Q}_N(\lambda^\theta)$$

- has **rank one**,
- admits necessarily eigenvalue -1
  (and eigenvalue 0 with multiplicity $N-1$)

Hence the determinant condition writes

$$\begin{aligned} & \det\left[\mathbf{I}_N + \lambda^\theta \frac{\theta}{1+\theta}\vec{\mathbf{u}}\vec{\mathbf{u}}^* \mathbf{Q}_N(\lambda^\theta)\right] = 0 \\ \Leftrightarrow \quad & \operatorname{Trace}\left\{\lambda^\theta \frac{\theta}{1+\theta}\vec{\mathbf{u}}\vec{\mathbf{u}}^* \mathbf{Q}_N(\lambda^\theta)\right\} = -1 \\ \Leftrightarrow \quad & \boxed{\lambda^\theta\, \vec{\mathbf{u}}^* \mathbf{Q}_N(\lambda^\theta)\vec{\mathbf{u}} = -\frac{1+\theta}{\theta}} \end{aligned}$$

The asymptotic condition I

Recall the condition

$$\boxed{\lambda^\theta \vec{\mathbf{u}}^* \mathbf{Q}_N(\lambda^\theta)\vec{\mathbf{u}} = -\frac{1+\theta}{\theta}}$$

Asymptotic simplification

$$\vec{\mathbf{u}}^* \mathbf{Q}_N(\lambda^\theta)\vec{\mathbf{u}} \xrightarrow[N,n\to\infty]{} \mathbf{g}_{\check{\mathrm{MP}}}\left(\lambda^\theta\right) \ .$$

Hence the final form of the condition

$$\boxed{\lambda^\theta \mathbf{g}_{\check{\mathrm{MP}}}\left(\lambda^\theta\right) = -\frac{1+\theta}{\theta}}$$

The asymptotic condition II

- We introduce the following function $\rho(z)$:

$$\boxed{\rho(z) = 1 + z\, g(z)}$$

- Let $\rho_{\check{\mathrm{M}}\mathrm{P}}$ associated to the Stieltjes transform:

$$\rho_{\check{\mathrm{M}}\mathrm{P}}(z) = 1 + z\mathbf{g}_{\check{\mathrm{M}}\mathrm{P}}(z) \ .$$

Then the condition over λ^θ writes:

$$\lambda^\theta \mathbf{g}_{\check{\mathrm{M}}\mathrm{P}}\left(\lambda^\theta\right) = -\frac{1+\theta}{\theta} \quad \Leftrightarrow \quad \rho_{\check{\mathrm{M}}\mathrm{P}}(\lambda^\theta) - 1 = -\frac{1+\theta}{\theta}$$

$$\Leftrightarrow \quad \boxed{\rho_{\check{\mathrm{M}}\mathrm{P}}\left(\lambda^\theta\right) = -\frac{1}{\theta}}$$

The asymptotic condition III

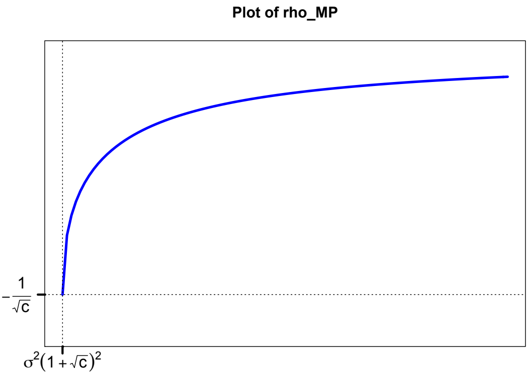

Figure: Plot of $\rho_{\check{\mathrm{M}}\mathrm{P}}$ on $(\sigma^2(1+\sqrt{c})^2, \infty)$

The function $\rho_{\check{\mathrm{M}}\mathrm{P}}$ admits an explicit expression on $(\sigma^2(1+\sqrt{c})^2, \infty)$

$$\rho_{\check{\mathrm{M}}\mathrm{P}}(x) = 1 + \frac{1}{2c}\left\{(1-x-c) + \sqrt{(1-x-c)^2 - 4cx}\right\} \quad (\sigma^2 = 1)$$

The asymptotic condition III

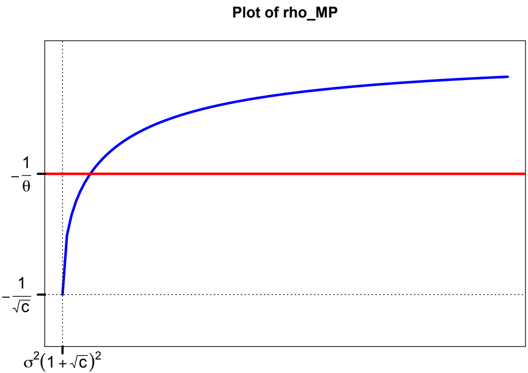

Figure: Plot of $\rho_{\check{\mathrm{M}}\mathrm{P}}$ on $(\sigma^2(1+\sqrt{c})^2, \infty)$

The asymptotic condition is satisfied if

$$\rho_{\check{\mathrm{M}}\mathrm{P}}\left(\lambda^\theta\right) = -\frac{1}{\theta} \quad \Leftrightarrow \quad -\frac{1}{\theta} > -\frac{1}{\sqrt{c}} \quad \Leftrightarrow \quad \boxed{\theta > \sqrt{c}}$$

Computing the limit λ^θ

We have

$$\rho_{\check{\mathrm{MP}}}\left(\lambda^\theta\right) = -\frac{1}{\theta} \quad \Leftrightarrow \quad \boxed{\lambda^\theta = \rho_{\check{\mathrm{MP}}}^{-1}\left(-\frac{1}{\theta}\right)}$$

We therefore need to inverse $\rho_{\check{\mathrm{MP}}}$.

- Using Marčenko-Pastur equation and the relation between $\mathbf{g}_{\check{\mathrm{MP}}}$ and $\rho_{\check{\mathrm{MP}}}$

$$\begin{aligned} \mathbf{g}_{\check{\mathrm{MP}}}(z) &= \frac{1}{\sigma^2(1-c) - z - z\sigma^2 c\,\mathbf{g}_{\check{\mathrm{MP}}}(z)} \\ \rho_{\check{\mathrm{MP}}}(z) &= 1 + z\mathbf{g}_{\check{\mathrm{MP}}}(z) \end{aligned}$$

  we get

$$z = \frac{\sigma^2}{\rho_{\check{\mathrm{MP}}}(z)}\left(\rho_{\check{\mathrm{MP}}}(z) - 1\right)\left(1 - c\rho_{\check{\mathrm{MP}}}(z)\right)$$

- Replacing now $z = \rho_{\check{\mathrm{MP}}}^{-1}\left(-\frac{1}{\theta}\right)$ into the equation yields:

$$\boxed{\lambda^\theta = \rho_{\check{\mathrm{MP}}}^{-1}\left(-\frac{1}{\theta}\right) = \sigma^2(1+\theta)\left(1 + \frac{c}{\theta}\right)}$$

Spiked model eigenvectors I

- Consider the following $N \times n$ spiked model:

$$\begin{aligned}\tilde{\mathbf{X}}_N &= (\mathbf{I}_N + \theta \vec{\mathbf{u}}\vec{\mathbf{u}}^*)^{1/2}\,\mathbf{X}_N \quad \text{with} \quad \|\vec{\mathbf{u}}\| = 1\ , \\ &= \mathbf{\Pi}^{1/2}\mathbf{X}_N\end{aligned}$$

  where $\mathbf{X}_N$ has i.i.d. $0/\sigma^2$ entries.
- Let $\vec{\boldsymbol{v}}_{\max}$ be the eigenvector associated to $\lambda_{\max}$, the largest eigenvalue of the covariance matrix associated to $\tilde{\mathbf{X}}_N$:

$$\left(\frac{1}{n}\tilde{\mathbf{X}}_N\tilde{\mathbf{X}}_N^*\right)\vec{\boldsymbol{v}}_{\max} = \lambda_{\max}\vec{\boldsymbol{v}}_{\max}$$

Question

- What is the behavior of $\vec{\boldsymbol{v}}_{\max}$ as $N, n \to \infty$ in the regime where

$$\frac{N}{n} \to c \in (0, \infty)?$$

Reminder

Behaviour of largest eigenvalue $\lambda_{\max}$ well-understood:

- if $\boxed{\theta \leq \sqrt{c}}$ then $\lambda_{\max} \to \sigma^2(1+\sqrt{c})^2$, **the right edge of MP bulk.**
- if $\boxed{\theta > \sqrt{c}}$ then $\lambda_{\max} \to \sigma^2(1+\theta)(1+c/\theta)$, i.e. $\lambda_{\max}$ **separates from the bulk.**

Spiked model eigenvectors II

Preliminary observations

1. Let N finite, $n \to \infty$, then

$$\frac{1}{n}\tilde{\mathbf{X}}_N\tilde{\mathbf{X}}_N^* = \mathbf{\Pi}^{1/2}\left(\frac{1}{n}\mathbf{X}_N\mathbf{X}_N^*\right)\mathbf{\Pi}^{1/2} \xrightarrow[n\to\infty]{} \mathbf{\Pi}$$

   Largest eigenvalue of $\mathbf{\Pi}$ is $1+\theta$; associated eigenvector is $\vec{\mathbf{u}}$:

$$\mathbf{\Pi}\vec{\mathbf{u}} = (\mathbf{I}_N + \theta\vec{\mathbf{u}}\vec{\mathbf{u}}^*)\,\vec{\mathbf{u}} = (1+\theta)\vec{\mathbf{u}}\ .$$

   As a consequence:

$$\vec{\boldsymbol{v}}_{\max} \xrightarrow[n\to\infty]{} \vec{\mathbf{u}}\ .$$

2. If

$$N, n \to \infty\ , \qquad \frac{N}{n} \to c\ ,$$

   then $\boxed{\dim(\vec{\boldsymbol{v}}_{\max}) = N \nearrow \infty}$. We therefore consider the projection

$$\vec{\boldsymbol{v}}_{\max}\vec{\boldsymbol{v}}_{\max}^*$$

   on $\vec{\boldsymbol{v}}_{\max}$ on a generic deterministic vector $\vec{\boldsymbol{a}}_N$, i.e.

$$\boxed{\vec{\boldsymbol{a}}_N^*\vec{\boldsymbol{v}}_{\max}\vec{\boldsymbol{v}}_{\max}^*\vec{\boldsymbol{a}}_N}$$

Spiked model eigenvectors III

Theorem

Let $\vec{\boldsymbol{a}}_N$ be a deterministic vector with norm 1, then

$$\vec{\boldsymbol{a}}_N^* \vec{\boldsymbol{v}}_{\max} \vec{\boldsymbol{v}}_{\max}^* \vec{\boldsymbol{a}}_N - \left(1 - \frac{c}{\theta^2}\right)\left(1 + \frac{c}{\theta}\right)^{-1} \vec{\boldsymbol{a}}_N^* \vec{\mathbf{u}} \vec{\mathbf{u}}^* \vec{\boldsymbol{a}}_N \xrightarrow[N,n\to\infty]{a.s.} 0 \ .$$

Remarks

- If N finite, $n \to \infty$, then

$$\vec{\boldsymbol{a}}_N^* \vec{\boldsymbol{v}}_{\max} \vec{\boldsymbol{v}}_{\max}^* \vec{\boldsymbol{a}}_N - \vec{\boldsymbol{a}}_N^* \vec{\mathbf{u}} \vec{\mathbf{u}}^* \vec{\boldsymbol{a}}_N \xrightarrow[N,n\to\infty]{a.s.} 0 \ .$$

- The large dimension $\frac{N}{n} \to c$ induces a correction factor:

$$\boxed{\kappa(c) = \left(1 - \frac{c}{\theta^2}\right)\left(1 + \frac{c}{\theta}\right)^{-1}}$$

- Of course $\kappa(c) \to 1$ if $c \to 0$.

Proof I

Reminder from complex analysis

We need a simple result from complex analysis:

$$\frac{1}{2\boldsymbol{i}\pi}\oint_{\mathcal{C}^-}\frac{dz}{z}=1$$

if $\mathcal{C}^-$ is a contour (take a circle of radius 1) enclosing counterclockwise 0.

- Proof:

$$\text{let } z=e^{i\theta}: \qquad \frac{1}{2\boldsymbol{i}\pi}\oint_{\mathcal{C}^-}\frac{dz}{z}=\frac{1}{2\boldsymbol{i}\pi}\int_0^{2\pi}\frac{d(e^{\boldsymbol{i}\theta})}{e^{\boldsymbol{i}\theta}}=\frac{1}{2\boldsymbol{i}\pi}\int_0^{2\pi}\frac{\boldsymbol{i}e^{\boldsymbol{i}\theta}d\theta}{e^{\boldsymbol{i}\theta}}=1$$

In particular, if $\mathcal{C}^+$ is a contour enclosing λ **clockwise**, then:

$$\boxed{\frac{1}{2\boldsymbol{i}\pi}\oint_{\mathcal{C}^+}\frac{dz}{\lambda-z}=1}$$

(let $\mathcal{C}^+$ be a circle $(\lambda+\rho e^{\boldsymbol{i}\theta}; 0\leq\theta\leq 2\pi)$ and perform a change of variable).

If $\mathcal{C}^+$ does not enclose λ, then the integral **equals zero**.

Proof II

Our objective

To express $\vec{\boldsymbol{v}}_{\max}$ with the help of the **resolvent** $\tilde{\mathbf{Q}}_n(z) = \left(\frac{1}{n}\tilde{\mathbf{X}}_N\tilde{\mathbf{X}}_N^* - z\mathbf{I}_N\right)^{-1}$

By the spectral theorem,

$$\begin{aligned} \frac{1}{n}\tilde{\mathbf{X}}_N\tilde{\mathbf{X}}_N^* &= \boldsymbol{O}_N \begin{pmatrix} \lambda_{\max} & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix} \boldsymbol{O}_N^* \\ &= [\vec{\boldsymbol{v}}_{\max}\ \boldsymbol{O}_{N-1}] \begin{pmatrix} \lambda_{\max} & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix} \begin{bmatrix} \vec{\boldsymbol{v}}_{\max}^* \\ \boldsymbol{O}_{N-1}^* \end{bmatrix} \end{aligned}$$

In particular,

$$\left(\frac{1}{n}\tilde{\mathbf{X}}_N\tilde{\mathbf{X}}_N^* - z\mathbf{I}_N\right)^{-1} = [\vec{\boldsymbol{v}}_{\max}\ \boldsymbol{O}_{N-1}] \begin{pmatrix} \frac{1}{\lambda_{\max}-z} & & \\ & \ddots & \\ & & \frac{1}{\lambda_N - z} \end{pmatrix} \begin{bmatrix} \vec{\boldsymbol{v}}_{\max}^* \\ \boldsymbol{O}_{N-1}^* \end{bmatrix}$$

Recall that

- if $\theta > \sqrt{c}$, $\lambda_{\max}$ separates from the bulk

and consider a contour $\mathcal{C}^+$ **exclusively enclosing** the eigenvalue $\lambda_{\max}$.

Proof III

We have

$$\boxed{\vec{\boldsymbol{a}}_N^* \vec{\boldsymbol{v}}_{\max} \vec{\boldsymbol{v}}_{\max}^* \vec{\boldsymbol{a}}_N = \frac{1}{2\boldsymbol{i}\pi} \oint_{\mathcal{C}+} \vec{\boldsymbol{a}}_N^* \tilde{Q}_n(z) \vec{\boldsymbol{a}}_N \, dz}$$

Indeed,

$$\begin{aligned}
&\frac{1}{2\boldsymbol{i}\pi} \oint_{\mathcal{C}+} \vec{\boldsymbol{a}}_N^* \tilde{Q}_n(z) \vec{\boldsymbol{a}}_N \, dz \\
&= \frac{1}{2\boldsymbol{i}\pi} \oint_{\mathcal{C}+} \vec{\boldsymbol{a}}_N^* [\vec{\boldsymbol{v}}_{\max} \ \boldsymbol{O}_{N-1}] \begin{pmatrix} \frac{1}{\lambda_{\max}-z} & & \\ & \ddots & \\ & & \frac{1}{\lambda_N - z} \end{pmatrix} \begin{bmatrix} \vec{\boldsymbol{v}}_{\max}^* \\ \boldsymbol{O}_{N-1}^* \end{bmatrix} \vec{\boldsymbol{a}}_N \, dz \\
&= \vec{\boldsymbol{a}}_N^* [\vec{\boldsymbol{v}}_{\max} \ \boldsymbol{O}_{N-1}] \begin{pmatrix} \frac{1}{2\boldsymbol{i}\pi} \oint \frac{1}{\lambda_{\max}-z} dz & & \\ & \ddots & \\ & & \frac{1}{2\boldsymbol{i}\pi} \oint \frac{1}{\lambda_N - z} dz \end{pmatrix} \begin{bmatrix} \vec{\boldsymbol{v}}_{\max}^* \\ \boldsymbol{O}_{N-1}^* \end{bmatrix} \vec{\boldsymbol{a}}_N \\
&= \vec{\boldsymbol{a}}_N^* [\vec{\boldsymbol{v}}_{\max} \ \boldsymbol{O}_{N-1}] \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 0 \end{pmatrix} \begin{bmatrix} \vec{\boldsymbol{v}}_{\max}^* \\ \boldsymbol{O}_{N-1}^* \end{bmatrix} \vec{\boldsymbol{a}}_N \\
&= \vec{\boldsymbol{a}}_N^* \vec{\boldsymbol{v}}_{\max} \vec{\boldsymbol{v}}_{\max}^* \vec{\boldsymbol{a}}_N \ .
\end{aligned}$$

Proof IV

Recall

$$\frac{1}{2\boldsymbol{i}\pi}\oint_{\mathcal{C}+}\vec{\boldsymbol{a}}_N^*\tilde{\mathbf{Q}}_n(z)\vec{\boldsymbol{a}}_N\,dz$$

and temporarily forget about the integral. Our objective now is:

to find a new formulation of $\vec{\boldsymbol{a}}_N^*\tilde{Q}_n(z)\vec{\boldsymbol{a}}_N$ and clearly separate the contribution from the perturbation ($\vec{\mathbf{u}}$ and θ) and the resolvent $\mathbf{Q}_n(z)$ from the non-pertubated model.

Introduce the notations

$$\mathbf{Z}_N = \frac{1}{n}\mathbf{X}_N\mathbf{X}_N^* \quad \text{and} \quad \tilde{\mathbf{Z}}_N = \frac{1}{n}\tilde{\mathbf{X}}_N\tilde{\mathbf{X}}_N^*$$

and recall the formula for the inverse of a rank-one perturbation:

$$(\mathbf{A}+\vec{\mathbf{u}}\vec{\mathbf{u}}^*)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\vec{\mathbf{u}}\vec{\mathbf{u}}^*\mathbf{A}^{-1}}{1+\vec{\mathbf{u}}\mathbf{A}\vec{\mathbf{u}}^*},$$

In particular

$$\mathbf{\Pi}^{-1} = (\mathbf{I}_N+\theta\vec{\mathbf{u}}\vec{\mathbf{u}}^*)^{-1} = \mathbf{I}_N - \frac{\theta}{1+\theta}\vec{\mathbf{u}}\vec{\mathbf{u}}^*$$

Proof V

$$\tilde{\mathbf{Q}}_n(z) \quad = \quad \left(\mathbf{\Pi}^{1/2}\tilde{\mathbf{Z}}_N\mathbf{\Pi}^{1/2} - z\mathbf{I}_N\right)^{-1}$$

Proof V

$$
\begin{aligned}
\tilde{\mathbf{Q}}_n(z) &= \left(\mathbf{\Pi}^{1/2}\tilde{\mathbf{Z}}_N\mathbf{\Pi}^{1/2} - z\mathbf{I}_N\right)^{-1} \\
&= \mathbf{\Pi}^{-1/2}\left(\mathbf{Z}_N - z\mathbf{\Pi}^{-1}\right)^{-1}\mathbf{\Pi}^{-1/2}
\end{aligned}
$$

Proof V

$$
\begin{aligned}
\tilde{\mathbf{Q}}_n(z) &= \left(\mathbf{\Pi}^{1/2}\tilde{\mathbf{Z}}_N\mathbf{\Pi}^{1/2} - z\mathbf{I}_N\right)^{-1} \\
&= \mathbf{\Pi}^{-1/2}\left(\mathbf{Z}_N - z\mathbf{\Pi}^{-1}\right)^{-1}\mathbf{\Pi}^{-1/2} \\
&= \mathbf{\Pi}^{-1/2}\left(\mathbf{Z}_N - z(\mathbf{I}_N + \theta\vec{\mathbf{u}}\vec{\mathbf{u}}^*)^{-1}\right)^{-1}\mathbf{\Pi}^{-1/2}
\end{aligned}
$$

Proof V

$$
\begin{aligned}
\tilde{\mathbf{Q}}_n(z) &= \left(\mathbf{\Pi}^{1/2}\tilde{\mathbf{Z}}_N\mathbf{\Pi}^{1/2} - z\mathbf{I}_N\right)^{-1} \\
&= \mathbf{\Pi}^{-1/2}\left(\mathbf{Z}_N - z\mathbf{\Pi}^{-1}\right)^{-1}\mathbf{\Pi}^{-1/2} \\
&= \mathbf{\Pi}^{-1/2}\left(\mathbf{Z}_N - z(\mathbf{I}_N + \theta\vec{\mathbf{u}}\vec{\mathbf{u}}^*)^{-1}\right)^{-1}\mathbf{\Pi}^{-1/2} \\
&= \mathbf{\Pi}^{-1/2}\left(\mathbf{Z}_N - z\left(\mathbf{I}_N - \frac{\theta}{1+\theta}\vec{\mathbf{u}}\vec{\mathbf{u}}^*\right)\right)^{-1}\mathbf{\Pi}^{-1/2}
\end{aligned}
$$

Proof V

$$\begin{aligned}
\tilde{\mathbf{Q}}_n(z) &= \left(\mathbf{\Pi}^{1/2}\tilde{\mathbf{Z}}_N\mathbf{\Pi}^{1/2} - z\mathbf{I}_N\right)^{-1} \\
&= \mathbf{\Pi}^{-1/2}\left(\mathbf{Z}_N - z\mathbf{\Pi}^{-1}\right)^{-1}\mathbf{\Pi}^{-1/2} \\
&= \mathbf{\Pi}^{-1/2}\left(\mathbf{Z}_N - z(\mathbf{I}_N + \theta\vec{\mathbf{u}}\vec{\mathbf{u}}^*)^{-1}\right)^{-1}\mathbf{\Pi}^{-1/2} \\
&= \mathbf{\Pi}^{-1/2}\left(\mathbf{Z}_N - z\left(\mathbf{I}_N - \frac{\theta}{1+\theta}\vec{\mathbf{u}}\vec{\mathbf{u}}^*\right)\right)^{-1}\mathbf{\Pi}^{-1/2} \\
&= \mathbf{\Pi}^{-1/2}(\mathbf{Z}_N - z\mathbf{I}_N + \xi\vec{\mathbf{u}}\vec{\mathbf{u}}^*)^{-1}\mathbf{\Pi}^{-1/2} \quad \text{where } \xi = z\frac{\theta}{1+\theta}
\end{aligned}$$

Proof V

$$
\begin{aligned}
\tilde{\mathbf{Q}}_n(z) &= \left(\mathbf{\Pi}^{1/2}\tilde{\mathbf{Z}}_N\mathbf{\Pi}^{1/2} - z\mathbf{I}_N\right)^{-1} \\
&= \mathbf{\Pi}^{-1/2}\left(\mathbf{Z}_N - z\mathbf{\Pi}^{-1}\right)^{-1}\mathbf{\Pi}^{-1/2} \\
&= \mathbf{\Pi}^{-1/2}\left(\mathbf{Z}_N - z(\mathbf{I}_N + \theta\vec{\mathbf{u}}\vec{\mathbf{u}}^*)^{-1}\right)^{-1}\mathbf{\Pi}^{-1/2} \\
&= \mathbf{\Pi}^{-1/2}\left(\mathbf{Z}_N - z\left(\mathbf{I}_N - \frac{\theta}{1+\theta}\vec{\mathbf{u}}\vec{\mathbf{u}}^*\right)\right)^{-1}\mathbf{\Pi}^{-1/2} \\
&= \mathbf{\Pi}^{-1/2}(\mathbf{Z}_N - z\mathbf{I}_N + \xi\vec{\mathbf{u}}\vec{\mathbf{u}}^*)^{-1}\mathbf{\Pi}^{-1/2} \quad \text{where } \xi = z\frac{\theta}{1+\theta} \\
&= \mathbf{\Pi}^{-1/2}\left(\mathbf{Q}_n - \frac{\mathbf{Q}_n\xi\vec{\mathbf{u}}\vec{\mathbf{u}}^*\mathbf{Q}_n}{1+\xi\vec{\mathbf{u}}^*\mathbf{Q}_n\vec{\mathbf{u}}}\right)\mathbf{\Pi}^{-1/2}
\end{aligned}
$$

Proof V

$$\begin{aligned}
\tilde{\mathbf{Q}}_n(z) &= \left(\mathbf{\Pi}^{1/2}\tilde{\mathbf{Z}}_N\mathbf{\Pi}^{1/2} - z\mathbf{I}_N\right)^{-1} \\
&= \mathbf{\Pi}^{-1/2}\left(\mathbf{Z}_N - z\mathbf{\Pi}^{-1}\right)^{-1}\mathbf{\Pi}^{-1/2} \\
&= \mathbf{\Pi}^{-1/2}\left(\mathbf{Z}_N - z(\mathbf{I}_N + \theta\vec{\mathbf{u}}\vec{\mathbf{u}}^*)^{-1}\right)^{-1}\mathbf{\Pi}^{-1/2} \\
&= \mathbf{\Pi}^{-1/2}\left(\mathbf{Z}_N - z\left(\mathbf{I}_N - \frac{\theta}{1+\theta}\vec{\mathbf{u}}\vec{\mathbf{u}}^*\right)\right)^{-1}\mathbf{\Pi}^{-1/2} \\
&= \mathbf{\Pi}^{-1/2}(\mathbf{Z}_N - z\mathbf{I}_N + \xi\vec{\mathbf{u}}\vec{\mathbf{u}}^*)^{-1}\mathbf{\Pi}^{-1/2} \quad \text{where } \xi = z\frac{\theta}{1+\theta} \\
&= \mathbf{\Pi}^{-1/2}\left(\mathbf{Q}_n - \frac{\mathbf{Q}_n\xi\vec{\mathbf{u}}\vec{\mathbf{u}}^*\mathbf{Q}_n}{1+\xi\vec{\mathbf{u}}^*\mathbf{Q}_n\vec{\mathbf{u}}}\right)\mathbf{\Pi}^{-1/2}
\end{aligned}$$

Hence

$$\vec{\boldsymbol{a}}_N^*\tilde{\mathbf{Q}}_n(z)\vec{\boldsymbol{a}}_N = \vec{\boldsymbol{a}}_N^*\mathbf{\Pi}^{1/2}\mathbf{Q}_n(z)\mathbf{\Pi}^{1/2}\vec{\boldsymbol{a}}_N - \xi\frac{\vec{\boldsymbol{a}}_N^*\mathbf{\Pi}^{1/2}\mathbf{Q}_n\vec{\mathbf{u}}\vec{\mathbf{u}}^*\mathbf{Q}_n\mathbf{\Pi}^{1/2}\vec{\boldsymbol{a}}_N}{1+\xi\vec{\mathbf{u}}^*\mathbf{Q}_n\vec{\mathbf{u}}}$$

Proof V

$$
\begin{aligned}
\tilde{\mathbf{Q}}_n(z) &= \left(\mathbf{\Pi}^{1/2}\tilde{\mathbf{Z}}_N\mathbf{\Pi}^{1/2} - z\mathbf{I}_N\right)^{-1} \\
&= \mathbf{\Pi}^{-1/2}\left(\mathbf{Z}_N - z\mathbf{\Pi}^{-1}\right)^{-1}\mathbf{\Pi}^{-1/2} \\
&= \mathbf{\Pi}^{-1/2}\left(\mathbf{Z}_N - z(\mathbf{I}_N + \theta\vec{\mathbf{u}}\vec{\mathbf{u}}^*)^{-1}\right)^{-1}\mathbf{\Pi}^{-1/2} \\
&= \mathbf{\Pi}^{-1/2}\left(\mathbf{Z}_N - z\left(\mathbf{I}_N - \frac{\theta}{1+\theta}\vec{\mathbf{u}}\vec{\mathbf{u}}^*\right)\right)^{-1}\mathbf{\Pi}^{-1/2} \\
&= \mathbf{\Pi}^{-1/2}(\mathbf{Z}_N - z\mathbf{I}_N + \xi\vec{\mathbf{u}}\vec{\mathbf{u}}^*)^{-1}\mathbf{\Pi}^{-1/2} \quad \text{where } \xi = z\frac{\theta}{1+\theta} \\
&= \mathbf{\Pi}^{-1/2}\left(\mathbf{Q}_n - \frac{\mathbf{Q}_n\xi\vec{\mathbf{u}}\vec{\mathbf{u}}^*\mathbf{Q}_n}{1+\xi\vec{\mathbf{u}}^*\mathbf{Q}_n\vec{\mathbf{u}}}\right)\mathbf{\Pi}^{-1/2}
\end{aligned}
$$

Hence

$$
\vec{\boldsymbol{a}}_N^*\tilde{\mathbf{Q}}_n(z)\vec{\boldsymbol{a}}_N = \vec{\boldsymbol{a}}_N^*\mathbf{\Pi}^{1/2}\mathbf{Q}_n(z)\mathbf{\Pi}^{1/2}\vec{\boldsymbol{a}}_N - \xi\frac{\vec{\boldsymbol{a}}_N^*\mathbf{\Pi}^{1/2}\mathbf{Q}_n\vec{\mathbf{u}}\vec{\mathbf{u}}^*\mathbf{Q}_n\mathbf{\Pi}^{1/2}\vec{\boldsymbol{a}}_N}{1+\xi\vec{\mathbf{u}}^*\mathbf{Q}_n\vec{\mathbf{u}}}
$$

Not so ugly!

Proof V

$$
\begin{aligned}
\tilde{\mathbf{Q}}_n(z) &= \left(\mathbf{\Pi}^{1/2}\tilde{\mathbf{Z}}_N\mathbf{\Pi}^{1/2} - z\mathbf{I}_N\right)^{-1} \\
&= \mathbf{\Pi}^{-1/2}\left(\mathbf{Z}_N - z\mathbf{\Pi}^{-1}\right)^{-1}\mathbf{\Pi}^{-1/2} \\
&= \mathbf{\Pi}^{-1/2}\left(\mathbf{Z}_N - z(\mathbf{I}_N + \theta\vec{\mathbf{u}}\vec{\mathbf{u}}^*)^{-1}\right)^{-1}\mathbf{\Pi}^{-1/2} \\
&= \mathbf{\Pi}^{-1/2}\left(\mathbf{Z}_N - z\left(\mathbf{I}_N - \frac{\theta}{1+\theta}\vec{\mathbf{u}}\vec{\mathbf{u}}^*\right)\right)^{-1}\mathbf{\Pi}^{-1/2} \\
&= \mathbf{\Pi}^{-1/2}(\mathbf{Z}_N - z\mathbf{I}_N + \xi\vec{\mathbf{u}}\vec{\mathbf{u}}^*)^{-1}\mathbf{\Pi}^{-1/2} \quad \text{where } \xi = z\frac{\theta}{1+\theta} \\
&= \mathbf{\Pi}^{-1/2}\left(\mathbf{Q}_n - \frac{\mathbf{Q}_n\xi\vec{\mathbf{u}}\vec{\mathbf{u}}^*\mathbf{Q}_n}{1+\xi\vec{\mathbf{u}}^*\mathbf{Q}_n\vec{\mathbf{u}}}\right)\mathbf{\Pi}^{-1/2}
\end{aligned}
$$

Hence

$$
\vec{\boldsymbol{a}}_N^*\tilde{\mathbf{Q}}_n(z)\vec{\boldsymbol{a}}_N = \vec{\boldsymbol{a}}_N^*\mathbf{\Pi}^{1/2}\mathbf{Q}_n(z)\mathbf{\Pi}^{1/2}\vec{\boldsymbol{a}}_N - \xi\frac{\vec{\boldsymbol{a}}_N^*\mathbf{\Pi}^{1/2}\mathbf{Q}_n\vec{\mathbf{u}}\vec{\mathbf{u}}^*\mathbf{Q}_n\mathbf{\Pi}^{1/2}\vec{\boldsymbol{a}}_N}{1+\xi\vec{\mathbf{u}}^*\mathbf{Q}_n\vec{\mathbf{u}}}
$$

Not so ugly! And we have separated the contribution of the perturbation from the non-perturbated model.

Proof VI

Recall

$$\vec{\boldsymbol{a}}_N^* \tilde{\mathbf{Q}}_n(z)\vec{\boldsymbol{a}}_N = \vec{\boldsymbol{a}}_N^* \mathbf{\Pi}^{1/2}\mathbf{Q}_n(z)\mathbf{\Pi}^{1/2}\vec{\boldsymbol{a}}_N - \xi \frac{\vec{\boldsymbol{a}}_N^* \mathbf{\Pi}^{1/2}\mathbf{Q}_n \vec{\mathbf{u}}\vec{\mathbf{u}}^* \mathbf{Q}_n \mathbf{\Pi}^{1/2}\vec{\boldsymbol{a}}_N}{1 + \xi \vec{\mathbf{u}}^* \mathbf{Q}_n \vec{\mathbf{u}}}$$

Proof VI

Recall

$$\vec{\boldsymbol{a}}_N^* \tilde{\mathbf{Q}}_n(z) \vec{\boldsymbol{a}}_N = \vec{\boldsymbol{a}}_N^* \mathbf{\Pi}^{1/2} \mathbf{Q}_n(z) \mathbf{\Pi}^{1/2} \vec{\boldsymbol{a}}_N - \xi \frac{\vec{\boldsymbol{a}}_N^* \mathbf{\Pi}^{1/2} \mathbf{Q}_n \vec{\mathbf{u}} \vec{\mathbf{u}}^* \mathbf{Q}_n \mathbf{\Pi}^{1/2} \vec{\boldsymbol{a}}_N}{1 + \xi \vec{\mathbf{u}}^* \mathbf{Q}_n \vec{\mathbf{u}}}$$

and integrate the first term

$$\frac{1}{2\boldsymbol{i}\pi} \oint_{\mathcal{C}+} \vec{\boldsymbol{a}}_N^* \mathbf{\Pi}^{1/2} \mathbf{Q}_n(z) \mathbf{\Pi}^{1/2} \vec{\boldsymbol{a}}_N = ??$$

Proof VI

Recall

$$\vec{\boldsymbol{a}}_N^* \tilde{\mathbf{Q}}_n(z) \vec{\boldsymbol{a}}_N = \vec{\boldsymbol{a}}_N^* \mathbf{\Pi}^{1/2} \mathbf{Q}_n(z) \mathbf{\Pi}^{1/2} \vec{\boldsymbol{a}}_N - \xi \frac{\vec{\boldsymbol{a}}_N^* \mathbf{\Pi}^{1/2} \mathbf{Q}_n \vec{\mathbf{u}} \vec{\mathbf{u}}^* \mathbf{Q}_n \mathbf{\Pi}^{1/2} \vec{\boldsymbol{a}}_N}{1 + \xi \vec{\mathbf{u}}^* \mathbf{Q}_n \vec{\mathbf{u}}}$$

and integrate the first term

$$\frac{1}{2\boldsymbol{i}\pi} \oint_{\mathcal{C}+} \vec{\boldsymbol{a}}_N^* \mathbf{\Pi}^{1/2} \mathbf{Q}_n(z) \mathbf{\Pi}^{1/2} \vec{\boldsymbol{a}}_N = \mathbf{0}$$

Why?

Proof VI

Recall

$$\vec{\boldsymbol{a}}_N^* \tilde{\mathbf{Q}}_n(z) \vec{\boldsymbol{a}}_N = \vec{\boldsymbol{a}}_N^* \mathbf{\Pi}^{1/2} \mathbf{Q}_n(z) \mathbf{\Pi}^{1/2} \vec{\boldsymbol{a}}_N - \xi \frac{\vec{\boldsymbol{a}}_N^* \mathbf{\Pi}^{1/2} \mathbf{Q}_n \vec{\mathbf{u}} \vec{\mathbf{u}}^* \mathbf{Q}_n \mathbf{\Pi}^{1/2} \vec{\boldsymbol{a}}_N}{1 + \xi \vec{\mathbf{u}}^* \mathbf{Q}_n \vec{\mathbf{u}}}$$

and integrate the first term

$$\frac{1}{2\boldsymbol{i}\pi} \oint_{\mathcal{C}+} \vec{\boldsymbol{a}}_N^* \mathbf{\Pi}^{1/2} \mathbf{Q}_n(z) \mathbf{\Pi}^{1/2} \vec{\boldsymbol{a}}_N = 0$$

Why? Because

1. the contour only encloses $\lambda_{\max}(\tilde{\mathbf{Z}}_n)$ which is away from the bulk,
2. but all the eigenvalues of $\mathbf{Z}_n$ are **in the bulk**. Hence:

$$\frac{1}{2\boldsymbol{i}\pi} \oint_{\mathcal{C}+} \frac{1}{\lambda_i(\mathbf{Z}_n) - z} \, dz = 0.$$

Proof VI

Recall

$$\vec{a}_N^* \tilde{\mathbf{Q}}_n(z)\vec{a}_N = \vec{a}_N^* \mathbf{\Pi}^{1/2}\mathbf{Q}_n(z)\mathbf{\Pi}^{1/2}\vec{a}_N - \xi \frac{\vec{a}_N^* \mathbf{\Pi}^{1/2}\mathbf{Q}_n \vec{\mathbf{u}}\vec{\mathbf{u}}^* \mathbf{Q}_n \mathbf{\Pi}^{1/2}\vec{a}_N}{1 + \xi \vec{\mathbf{u}}^* \mathbf{Q}_n \vec{\mathbf{u}}}$$

and integrate the first term

$$\frac{1}{2\boldsymbol{i}\pi} \oint_{\mathcal{C}+} \vec{a}_N^* \mathbf{\Pi}^{1/2}\mathbf{Q}_n(z)\mathbf{\Pi}^{1/2}\vec{a}_N = 0$$

Why? Because

1. the contour only encloses $\lambda_{\max}(\tilde{\mathbf{Z}}_n)$ which is away from the bulk,
2. but all the eigenvalues of $\mathbf{Z}_n$ are **in the bulk**. Hence:

$$\frac{1}{2\boldsymbol{i}\pi} \oint_{\mathcal{C}+} \frac{1}{\lambda_i(\mathbf{Z}_n) - z}\, dz = 0.$$

Last step is to simplify the remaining expression by systematically use the large N, n quadratic form approximation:

$$\boxed{\vec{c}^* \mathbf{Q}_n(z)\vec{d} - \vec{c}^* \vec{d}\, \mathbf{g}_{\check{\mathrm{MP}}}(z) \xrightarrow[N,n\to\infty]{a.s.} 0}$$

Proof VI

Recall

$$\vec{\boldsymbol{a}}_N^* \tilde{\mathbf{Q}}_n(z) \vec{\boldsymbol{a}}_N = \vec{\boldsymbol{a}}_N^* \mathbf{\Pi}^{1/2} \mathbf{Q}_n(z) \mathbf{\Pi}^{1/2} \vec{\boldsymbol{a}}_N - \xi \frac{\vec{\boldsymbol{a}}_N^* \mathbf{\Pi}^{1/2} \mathbf{Q}_n \vec{\mathbf{u}} \vec{\mathbf{u}}^* \mathbf{Q}_n \mathbf{\Pi}^{1/2} \vec{\boldsymbol{a}}_N}{1 + \xi \vec{\mathbf{u}}^* \mathbf{Q}_n \vec{\mathbf{u}}}$$

and integrate the first term

$$\frac{1}{2\boldsymbol{i}\pi} \oint_{\mathcal{C}+} \vec{\boldsymbol{a}}_N^* \mathbf{\Pi}^{1/2} \mathbf{Q}_n(z) \mathbf{\Pi}^{1/2} \vec{\boldsymbol{a}}_N = 0$$

Why? Because

1. the contour only encloses $\lambda_{\max}(\tilde{\mathbf{Z}}_n)$ which is away from the bulk,
2. but all the eigenvalues of $\mathbf{Z}_n$ are **in the bulk**. Hence:

$$\frac{1}{2\boldsymbol{i}\pi} \oint_{\mathcal{C}+} \frac{1}{\lambda_i(\mathbf{Z}_n) - z} \, dz = 0.$$

Last step is to simplify the remaining expression by systematically use the large N, n quadratic form approximation:

$$\boxed{\vec{c}^* \mathbf{Q}_n(z) \vec{d} - \vec{c}^* \vec{d} \, \mathbf{g}_{\check{\mathrm{MP}}}(z) \xrightarrow[N,n\to\infty]{a.s.} 0} \Rightarrow \begin{cases} \vec{\boldsymbol{a}}_N^* \mathbf{\Pi}^{1/2} \mathbf{Q}_n \vec{\mathbf{u}} & \approx & \vec{\boldsymbol{a}}_N^* \mathbf{\Pi}^{1/2} \vec{\mathbf{u}} \, \mathbf{g}_{\check{\mathrm{MP}}}(z) \\ \vec{\mathbf{u}}^* \mathbf{Q}_n \mathbf{\Pi}^{1/2} \vec{\boldsymbol{a}}_N & \approx & \vec{\mathbf{u}}^* \mathbf{\Pi}^{1/2} \vec{\boldsymbol{a}}_N \, \mathbf{g}_{\check{\mathrm{MP}}}(z) \\ \vec{\mathbf{u}}^* \mathbf{Q}_n \vec{\mathbf{u}} & \approx & \mathbf{g}_{\check{\mathrm{MP}}}(z) \end{cases}$$

Proof VII

After simplifications,

$$\begin{aligned}\vec{\boldsymbol{a}}_N^* \vec{\boldsymbol{v}}_{\max} \vec{\boldsymbol{v}}_{\max}^* \vec{\boldsymbol{a}}_N &\approx -\frac{1}{2\boldsymbol{i}\pi} \oint_{\mathcal{C}+} |\vec{\boldsymbol{a}}_N^* \boldsymbol{\Pi}^{1/2} \vec{\mathbf{u}}|^2 \frac{\mathbf{g}_{\check{\mathrm{MP}}}^2(z)}{\xi^{-1} + \mathbf{g}_{\check{\mathrm{MP}}}(z)}\, dz \\ &= -\frac{\vec{\boldsymbol{a}}_N^* \vec{\mathbf{u}} \vec{\mathbf{u}}^* \vec{\boldsymbol{a}}_N}{1+\theta} \oint_{\mathcal{C}+} \frac{\mathbf{g}_{\check{\mathrm{MP}}}^2(z)}{\xi^{-1} + \mathbf{g}_{\check{\mathrm{MP}}}(z)}\, dz\end{aligned}$$

It remains to compute the correction factor

$$-\frac{1}{1+\theta} \oint_{\mathcal{C}+} \frac{\mathbf{g}_{\check{\mathrm{MP}}}^2(z)}{\xi^{-1} + \mathbf{g}_{\check{\mathrm{MP}}}(z)}\, dz$$

by residue calculus (not that difficult).

A **minor miracle** occurs: **This factor admits a closed form formula!**

$$-\frac{1}{1+\theta} \oint_{\mathcal{C}+} \frac{\mathbf{g}_{\check{\mathrm{MP}}}^2(z)}{\xi^{-1} + \mathbf{g}_{\check{\mathrm{MP}}}(z)}\, dz = \left(1 - \frac{c}{\theta^2}\right)\left(1 + \frac{c}{\theta}\right)^{-1}$$

Finally:

$$\boxed{\vec{\boldsymbol{a}}_N^* \vec{\boldsymbol{v}}_{\max} \vec{\boldsymbol{v}}_{\max}^* \vec{\boldsymbol{a}}_N - \left(1 - \frac{c}{\theta^2}\right)\left(1 + \frac{c}{\theta}\right)^{-1} \vec{\boldsymbol{a}}_N^* \vec{\mathbf{u}} \vec{\mathbf{u}}^* \vec{\boldsymbol{a}}_N \xrightarrow[N,n\to\infty]{a.s.} 0\ .}$$

Summary I

Spiked model

Let

- $\mathbf{\Pi}_N$ a small perturbation of the identity [Example: $\mathbf{\Pi}_N = \mathbf{I}_N + \theta \vec{\mathbf{u}}\vec{\mathbf{u}}^*$]
- $\mathbf{X}_N$ a $N \times n$ matrix with i.i.d. entries

then $\boxed{\widetilde{\mathbf{X}}_N = \mathbf{\Pi}_N^{1/2} \mathbf{X}_N}$ is a (multiplicative) spiked model

Global regime

The spectral measure $L_N\left(\frac{1}{N}\widetilde{\mathbf{X}}_N\widetilde{\mathbf{X}}_N^*\right)$ converges to Marčenko-Pastur distribution:

$$\text{almost surely,} \;\; L_N\left(\frac{1}{N}\widetilde{\mathbf{X}}_N\widetilde{\mathbf{X}}_N^*\right) \xrightarrow[N,n\to\infty]{\mathcal{L}} \mathbb{P}_{\check{\mathrm{MP}}}$$

Largest eigenvalue

- if $\boxed{\theta \leq \sqrt{c}}$, then $\lambda_{\max}\left(\frac{1}{N}\widetilde{\mathbf{X}}_N\widetilde{\mathbf{X}}_N^*\right)$ converges to the right edge of the bulk
- if $\boxed{\theta > \sqrt{c}}$, then $\lambda_{\max}\left(\frac{1}{N}\widetilde{\mathbf{X}}_N\widetilde{\mathbf{X}}_N^*\right)$ separates from the bulk

$$\lambda_{\max}\left(\frac{1}{N}\widetilde{\mathbf{X}}_N\widetilde{\mathbf{X}}_N^*\right) \to \sigma^2(1+\theta)\left(1+\frac{c}{\theta}\right) > \sigma^2(1+\sqrt{c})^2$$

Summary II

1. Expression of $\vec{\boldsymbol{v}}_{\max}$ with the help of the resolvent

$$\vec{\boldsymbol{a}}_N^* \vec{\boldsymbol{v}}_{\max} \vec{\boldsymbol{v}}_{\max}^* \vec{\boldsymbol{a}}_N = \frac{1}{2\boldsymbol{i}\pi} \oint_{\mathcal{C}+} \vec{\boldsymbol{a}}_N^* \tilde{Q}_n(z) \vec{\boldsymbol{a}}_N \, dz$$

2. Convenient expression of $\vec{\boldsymbol{v}}_{\max}$ where the contribution of the perturbation is separated from the resolvent of the non-perturbated model ($\check{\text{M}}$P)

$$\vec{\boldsymbol{a}}_N^* \vec{\boldsymbol{v}}_{\max} \vec{\boldsymbol{v}}_{\max}^* \vec{\boldsymbol{a}}_N \quad \approx \quad -\frac{\vec{\boldsymbol{a}}_N^* \vec{\mathbf{u}} \vec{\mathbf{u}}^* \vec{\boldsymbol{a}}_N}{1+\theta} \oint_{\mathcal{C}+} \frac{\mathbf{g}^2_{\check{\text{M}}\text{P}}(z)}{\xi^{-1} + \mathbf{g}_{\check{\text{M}}\text{P}}(z)} \, dz$$

3. Residue calculus to find the final form

$$\vec{\boldsymbol{a}}_N^* \vec{\boldsymbol{v}}_{\max} \vec{\boldsymbol{v}}_{\max}^* \vec{\boldsymbol{a}}_N - \left(1 - \frac{c}{\theta^2}\right) \left(1 + \frac{c}{\theta}\right)^{-1} \vec{\boldsymbol{a}}_N^* \vec{\mathbf{u}} \vec{\mathbf{u}}^* \vec{\boldsymbol{a}}_N \xrightarrow[N,n\to\infty]{a.s.} 0 \; .$$

The hypothesis testing problem

Statistical Setup

let

$$\vec{\mathbf{y}}(k) = \begin{cases} \sigma \vec{\mathbf{w}}(k) & \text{under } H_0 \\ \vec{\mathbf{h}}\, s(k) + \sigma \vec{\mathbf{w}}(k) & \text{under } H_1 \end{cases} \qquad \text{for } k = 1:n$$

The $\vec{\mathbf{y}}(k)$'s are n observations all either drawn under H_0 or H_1. Here,

- $\vec{\mathbf{w}}(k)$ is a $N \times 1$ complex gaussian white noise process:

$$\vec{\mathbf{w}}(k) \ \sim \ \mathcal{C}N(0, \mathbf{I}_N)$$

- $\vec{\mathbf{h}}$ is a $N \times 1$ **deterministic** and **unknown** vector and typically represents the propagation channel
- $s(k)$ represent the signal; it is a scalar complex gaussian i.i.d. process

Objective

Given n observations $(\vec{\mathbf{y}}(k), 1 \leq k \leq n)$, and the associated **sample covariance matrix**

$$\hat{\mathbf{R}}_n = \frac{1}{n} \mathbf{Y}_n \mathbf{Y}_n^* \quad \text{where} \quad \mathbf{Y}_n = [\vec{\mathbf{y}}(1), \cdots, \vec{\mathbf{y}}(n)] \quad \text{is } N \times n\ ,$$

the aim is to decide H_0 (no signal) or H_1 (single-source detection) in the case where

$$\frac{N}{n} \to c \in (0,1) \quad \text{i.e.} \quad \boxed{\text{Dimension } N \text{ of observations} \propto \text{size } n \text{ of sample}}$$

Neyman-Pearson procedure

Likelihood functions

Notice that $\mathbf{Y}_n$ is a $N \times n$ matrix whose columns are i.i.d. vectors with covariance matrix defined by

$$\mathbf{\Sigma}_N = \begin{cases} \sigma^2 \mathbf{I}_N & \text{under } H_0 \ , \\ \vec{\mathbf{h}}\vec{\mathbf{h}}^* + \sigma^2 \mathbf{I}_N & \text{under } H_1 \end{cases}$$

hence the likelihood functions write

$$\begin{aligned} p_0(\mathbf{Y}_N; \sigma^2) &= \frac{1}{(\pi\sigma^2)^{Nn}} \exp\left(-\frac{n}{\sigma^2} \mathrm{Trace}\, \hat{\mathbf{R}}_N\right) \\ p_1(\mathbf{Y}_N; \vec{\mathbf{h}}; \sigma^2) &= \frac{1}{\left[\pi^N \det\left(\vec{\mathbf{h}}\vec{\mathbf{h}}^* + \sigma^2 \mathbf{I}_N\right)\right]^n} \exp\left(-\frac{n}{\sigma^2} \mathrm{Trace}\, \hat{\mathbf{R}}_N \left(\vec{\mathbf{h}}\vec{\mathbf{h}}^* + \sigma^2 \mathbf{I}_N\right)^{-1}\right) \end{aligned}$$

Neyman-Pearson

In case where σ^2 and $\vec{\mathbf{h}}$ are **known**, the **Likelihood Ratio Statistics**

$$\frac{p_1(\mathbf{Y}_N; \vec{\mathbf{h}}; \sigma^2)}{p_0(\mathbf{Y}_N; \sigma^2)}$$

provides a **uniformly most powerful test**:

- Fix a given level $\alpha \in (0, 1)$
- The condition over the **Probability of False Alarm** $\mathbb{P}(H_1 \mid H_0) \leq \alpha$ sets the threshold
- the **maximum achievable power**

$$1 - \mathbb{P}(H_0 \mid H_1)$$

is guaranteed by Neyman-Pearson.

The Generalized Likelihood Ratio Test

In the case where $\vec{\mathbf{h}}$ and σ^2 are unknown, we use instead:

$$L_n = \frac{\sup_{\sigma^2, \vec{\mathbf{h}}} \; p_1(\mathbf{Y}_n, \sigma^2, \vec{\mathbf{h}})}{\sup_{\sigma^2} \; p_0(\mathbf{Y}_n, \sigma^2)}$$

which is no longer uniformily most powerful.

Expression of the GLRT

The GLRT statistics writes

$$L_n = \frac{\left(1 - \frac{1}{N}\right)^{(1-N)n}}{\left(\frac{\lambda_{\max}(\hat{\mathbf{R}}_n)}{\frac{1}{N}\mathrm{Trace}\,\hat{\mathbf{R}}_n}\right)^n \left(1 - \frac{1}{N}\frac{\lambda_{\max}(\hat{\mathbf{R}}_n)}{\frac{1}{N}\mathrm{Trace}\,\hat{\mathbf{R}}_n}\right)^{(N-1)n}}$$

and is a **deterministic function** of $\boxed{T_n = \frac{\lambda_{\max}(\hat{\mathbf{R}}_n)}{\frac{1}{N}\mathrm{Trace}\,\hat{\mathbf{R}}_n}}$

Limit of the test statistics T_n I

Under H_0

Recall $T_n = \frac{\lambda_{\max}(\hat{\mathbf{R}}_n)}{\frac{1}{N}\mathrm{Trace}\,\hat{\mathbf{R}}_n}$. We have:

$$\lambda_{\max}(\hat{\mathbf{R}}_n) \xrightarrow[N,n\to\infty]{a.s.} \sigma^2(1+\sqrt{c})^2$$

$$\frac{1}{N}\mathrm{Trace}\,\hat{\mathbf{R}}_n = \frac{1}{Nn}\sum_{i,j}|Y_{ij}|^2 \xrightarrow[N,n\to\infty]{a.s.} \sigma^2$$

hence

$$\boxed{T_n = \frac{\lambda_{\max}(\hat{\mathbf{R}}_n)}{\frac{1}{N}\mathrm{Trace}\,\hat{\mathbf{R}}_n} \xrightarrow[N,n\to\infty]{a.s.} (1+\sqrt{c})^2}$$

Limit of the test statistics T_n II

Under H_1

Let

$$\boxed{\mathbf{snr} = \frac{\|\vec{\mathbf{h}}\|^2}{\sigma^2}}$$

the **Signal-to-Noise (SNR)** ratio.

- if $\mathbf{snr} > \sqrt{c}$ then

$$T_n \xrightarrow[N,n\to\infty]{a.s.} (1+\mathbf{snr})\left(1+\frac{c}{\mathbf{snr}}\right) > (1+\sqrt{c})^2$$

- if $\mathbf{snr} \leq \sqrt{c}$ then

$$T_n \xrightarrow[N,n\to\infty]{a.s.} (1+\sqrt{c})^2$$

Limit of the test statistics T_n III

Remarks

- Condition $\boxed{\mathbf{snr} > \sqrt{c}}$ is **automatically fulfilled** in the standard regime where

$$N \text{ fixed} \quad \text{and} \quad n \to \infty \quad \text{as} \quad c = \lim_{n \to \infty} \frac{N}{n} = 0 \; .$$

- In the case $N, n \to \infty$, recall that the support of Marčenko-Pastur distribution is

$$[(1 - \sqrt{c})^2, (1 + \sqrt{c})^2] \; ,$$

i.e.

$$\boxed{\text{The higher } \sqrt{c}\text{, the larger the support}}$$

One can interpret $\sqrt{c}$ as a level of the **asymptotic noise induced by the data dimension** (=asymptotic data noise).

Hence the rule of thumb

$$\boxed{\text{Detection occurs if } \mathbf{snr} \text{ higher than } \textbf{asymptotic data noise}.}$$

Simulations

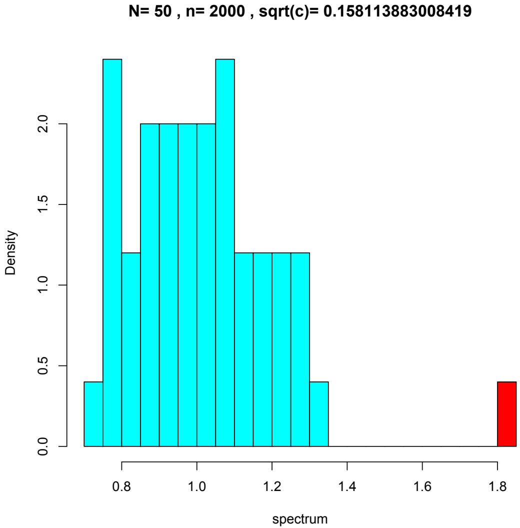

Figure: Influence of **asymptotic data noise** as $\sqrt{c}$ increases

Simulations

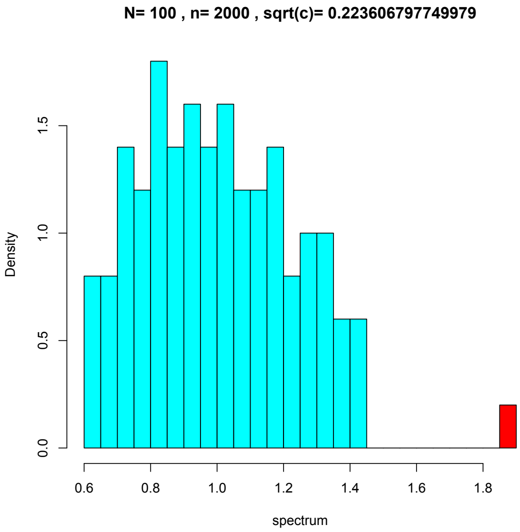

Figure: Influence of **asymptotic data noise** as $\sqrt{c}$ increases

Simulations

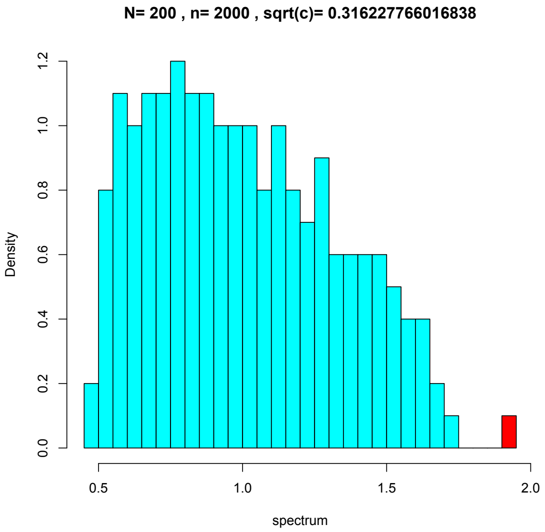

Figure: Influence of **asymptotic data noise** as $\sqrt{c}$ increases

Simulations

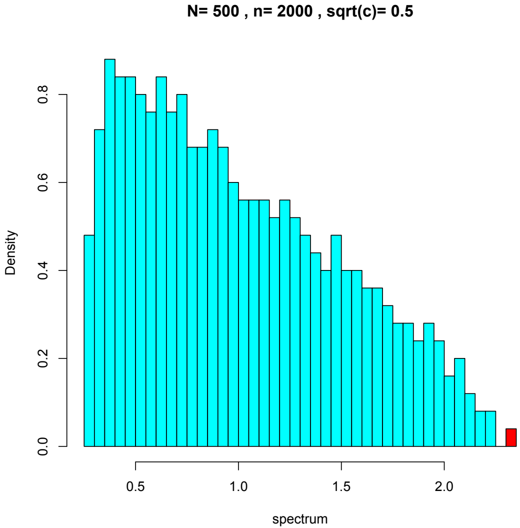

Figure: Influence of **asymptotic data noise** as $\sqrt{c}$ increases

Simulations

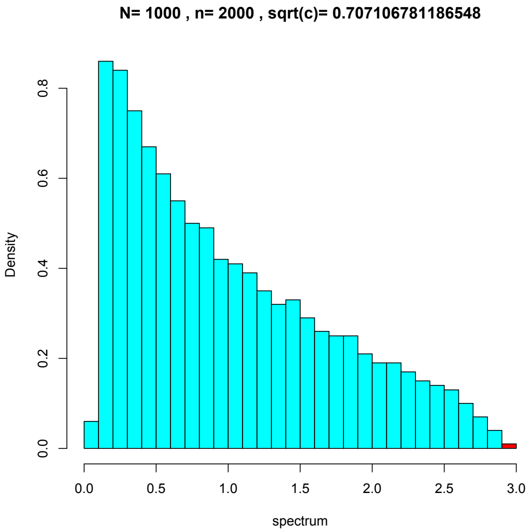

Figure: Influence of **asymptotic data noise** as $\sqrt{c}$ increases

Elements of proof I

- We are interested in the largest eigenvalue of the matrix model

$$\frac{\frac{1}{n}\mathbf{Y}_n\mathbf{Y}_n^*}{\frac{1}{N}\mathrm{Trace}(\hat{\mathbf{R}}_n)}$$

asymptotically equivalent to

$$\frac{1}{n}\frac{\mathbf{Y}_n\mathbf{Y}_n^*}{\sigma^2} \quad \text{as} \quad \frac{1}{N}\mathrm{Trace}(\hat{\mathbf{R}}_n) \xrightarrow[N,n\to\infty]{a.s.} \sigma^2$$

- Notice that

$$\mathbf{Y}_n = [\vec{\mathbf{y}}_1, \cdots, \vec{\mathbf{y}}_n] \quad \text{with} \quad \vec{\mathbf{y}}_i \sim \mathcal{C}N(0, \vec{\mathbf{h}}\vec{\mathbf{h}}^* + \sigma^2\mathbf{I}_N)$$

Hence

$$\mathbf{Y}_N = \left(\vec{\mathbf{h}}\vec{\mathbf{h}}^* + \sigma^2\mathbf{I}_N\right)^{1/2}\mathbf{X}_N \quad \Rightarrow \quad \frac{\mathbf{Y}_N}{\sigma} = \left(\mathbf{I}_N + \frac{\vec{\mathbf{h}}\vec{\mathbf{h}}^*}{\sigma^2}\right)^{1/2}\mathbf{X}_N$$

$$= \left(\mathbf{I}_N + \frac{\|\vec{\mathbf{h}}\|^2}{\sigma^2}\vec{\mathbf{u}}\vec{\mathbf{u}}^*\right)^{1/2}\mathbf{X}_N$$

with $\mathbf{X}_N$ a $N \times n$ matrix having i.i.d. entries $\mathcal{C}N(0,1)$ and $\vec{\mathbf{u}} = \frac{\vec{\mathbf{h}}}{\|\vec{\mathbf{h}}\|}$

Conclusion

$$\boxed{\text{Spectrum of } \tfrac{1}{n}\mathbf{Y}_n\mathbf{Y}_n^* \text{ follows a spiked model with } \textbf{rank-one perturbation}}$$

Elements of proof II

We can now conclude:

- If $\boxed{\mathbf{snr} > \sqrt{c}}$ then

$$\frac{\lambda_{\max}\left(\hat{\mathbf{R}}_n\right)}{\frac{1}{N}\mathrm{Trace}(\hat{\mathbf{R}}_n)} \xrightarrow[N,n\to\infty]{(H_1)} (1+\mathbf{snr})\left(1+\frac{c}{\mathbf{snr}}\right) > (1+\sqrt{c})^2$$

  and the test statistics discriminates between the hypotheses H_0 and H_1.

- If $\boxed{\mathbf{snr} \leq \sqrt{c}}$ then

$$\frac{\lambda_{\max}\left(\hat{\mathbf{R}}_n\right)}{\frac{1}{N}\mathrm{Trace}(\hat{\mathbf{R}}_n)} \xrightarrow[N,n\to\infty]{(H_1)} (1+\sqrt{c})^2$$

  Same limit as under H_0. The test statistics does not discriminate between the two hypotheses.

Fluctuations of the GLRT under H_0 - I

- The exact distribution of the statistics

$$L_n = \frac{\lambda_{\max}(\hat{\mathbf{R}}_n)}{\frac{1}{N}\operatorname{Trace}\hat{\mathbf{R}}_n}$$

is needed to set the threshold of the test fo a given confidence level $\alpha \in (0,1)$:

$$\mathbb{P}_{H_0}\left(L_N > \boldsymbol{t_\alpha}\right) = \alpha \ ,$$

but **hard to obtain**.

- We rather study the asymptotic fluctuations of L_n under the regime

$$N, n \to \infty \ , \quad \frac{N}{n} \to c \in (0,1) \ .$$

- L_N is the ratio of **two random variables**. We need to understand
  - the fluctuations of $\lambda_{\max}(\hat{\mathbf{R}}_n)$ under H_0,
  - the fluctuations of $\frac{1}{N}\operatorname{Trace}\hat{\mathbf{R}}_n$ under H_0.

Fluctuations of the GLRT under H_0 - II

Fluctuations of $\lambda_{\max}(\hat{\mathbf{R}}_n)$: Tracy-Widom distribution **at rate $N^{2/3}$**

$$\frac{N^{2/3}}{\Theta_N} \left\{ \lambda_{\max}\left(\hat{\mathbf{R}}_n\right) - \sigma^2(1+\sqrt{c_n})^2 \right\} \xrightarrow[N,n\to\infty]{\mathcal{L}} \mathbb{P}_{\mathrm{TW}}$$

where

$$c_n = \frac{N}{n} \quad \text{and} \quad \Theta_N = \sigma^2(1+\sqrt{c_n})\left(\frac{1}{\sqrt{c_n}}+1\right)^{1/3}$$

Otherwise stated,

$$\boxed{\lambda_{\max}\left(\hat{\mathbf{R}}_n\right) = \sigma^2(1+\sqrt{c_n})^2 + \frac{\Theta_N}{N^{2/3}}\boldsymbol{X}_{TW} + \varepsilon_n}$$

where $\boldsymbol{X}_{TW}$ is a random variable with Tracy-Widom distribution.

Details on Tracy-Widom distribution

Tracy-Widom distribution is defined by

- its cumulative distribution function

$$F_{TW}(x) = \exp\left\{ -\int_x^{\infty} (u-x)^2 q^2(u)\, du \right\}$$

- where

$$q''(x) = xq(x) + 2q^3(x) \quad \text{and} \quad q(x) \sim \mathrm{Ai}(x) \text{ as } x \to \infty\ .$$

  $x \mapsto \mathrm{Ai}(x)$ being the Airy function.

Don't bother .. just download it

- For simulations, cf. `R Package 'RMTstat'`, by Johnstone et al.
- Also, Folkmar Bornemann (TU München) has developed fast `matlab` code

Tracy-Widom curve

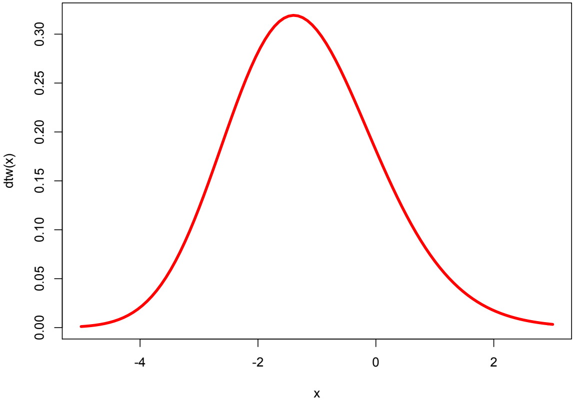

Figure: Tracy-Widom density

Tracy-Widom curve

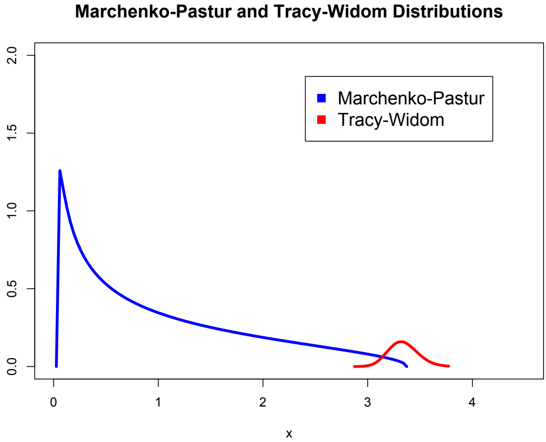

Figure: Fluctuations of the largest eigenvalue $\lambda_{\max}(\hat{\mathbf{R}}_n)$ under H_0

Fluctuations of the GLRT under H_0 - III

Fluctuations of $\frac{1}{N}\mathrm{Trace}\,(\hat{\mathbf{R}}_n)$: Gaussian distributions **at rate N**

$$N\left\{\frac{1}{N}\sum_{i=1}^{N}\lambda_i(\hat{\mathbf{R}}_n)-\sigma^2\right\}\xrightarrow[N,n\to\infty]{\mathcal{L}}\mathcal{N}(0,\Gamma)\ ,$$

Otherwise stated:

$$\boxed{\frac{1}{N}\mathrm{Trace}\,(\hat{\mathbf{R}}_n)=\frac{1}{N}\sum_{i=1}^{N}\lambda_i(\hat{\mathbf{R}}_n)=\sigma^2+\frac{\sqrt{\Gamma}}{N}\boldsymbol{Z}+\varepsilon_n}$$

where $\boldsymbol{Z}$ is a random variable with distribution $\mathcal{N}(0,1)$.

Fluctuations of the GLRT under H_0 - IV

Conclusion

- Fluctuations of $L_n = \frac{\lambda_{\max}(\hat{\mathbf{R}}_n)}{\frac{1}{N}\operatorname{Trace}\hat{\mathbf{R}}_n}$ are driven by $\lambda_{\max}(\hat{\mathbf{R}}_n)$:

$$\frac{N^{2/3}}{\widetilde{\Theta}_N}\left\{L_N - (1+\sqrt{c_n})^2\right\} \xrightarrow[N,n\to\infty]{\mathcal{L}} \mathbb{P}_{\mathrm{TW}} \quad \text{with} \quad \widetilde{\Theta}_N = (1+\sqrt{c_n})\left(\frac{1}{\sqrt{c_n}}+1\right)^{1/3}$$

- In order to set the threshold $\boldsymbol{\alpha}$, we choose $\boldsymbol{t_\alpha^n}$ as

$$\boldsymbol{t_\alpha^n} = (1+\sqrt{c_n})^2 + \frac{\widetilde{\Theta}_N}{N^{2/3}}\boldsymbol{t_\alpha}^{\textbf{Tracy-Widom}}$$

where $\boldsymbol{t_\alpha}^{\textbf{Tracy-Widom}}$ is the corresponding quantile for a Tracy-Widom random variable:

$$\mathbb{P}\{\boldsymbol{X}_{TW} > \boldsymbol{t_\alpha}^{\textbf{Tracy-Widom}}\} \leq \alpha.$$

Power of the GLRT I

Type II error and Power of the test

Given a level of confidence $\alpha \in (0,1)$, the **type I error** defines the associate quantile $\boldsymbol{t_\alpha}$

$$\mathbb{P}_{H_0}\left(L_N > \boldsymbol{t^n_\alpha}\right) \quad \leq \quad \alpha \ .$$

The **type II error** is defined as

$$\mathbb{P}_{H_1}\left(L_N < \boldsymbol{t^n_\alpha}\right) \ ,$$

and the associated **power of the test** is defined as

$$\mathbb{P}_{H_1}\left(L_N \geq \boldsymbol{t^n_\alpha}\right) = 1 - \mathbb{P}_{H_1}\left(L_N < \boldsymbol{t^n_\alpha}\right) \ .$$

No optimality

Contrary to Neyman-Pearson procedure, there is **no theoretical guarantee** that the GLRT is a uniformily most powerful test.

- It is therefore important to be able to compute the power of the GLRT

Power of the GLRT II

- For fixed level of confidence α

$$\mathbb{P}_{H_0}\left(L_N > \boldsymbol{t_\alpha^n}\right) \le \alpha \ ,$$

  the type II error **exponentially decreases to 0**.

- Indeed, we want to evaluate

$$\begin{aligned}
&\mathbb{P}_{H_1}\left(L_N \le \boldsymbol{t_\alpha^n}\right) \\
&\quad = \quad \mathbb{P}_{H_1}\left(L_N \le (1+\sqrt{c})^2 + \frac{\widetilde{\Theta}_N}{N^{2/3}} \boldsymbol{t_\alpha}^{\textbf{Tracy-Widom}}\right) \\
&\quad = \quad \mathbb{P}_{H_1}\left(L_N \le (1+\sqrt{c})^2 + \mathcal{O}\left(\frac{1}{N^{2/3}}\right)\right)
\end{aligned}$$

  but

$$L_N \quad \xrightarrow[N,n\to\infty]{H_1} \quad (1+\mathbf{snr})\left(1+\frac{c}{\mathbf{snr}}\right) > (1+\sqrt{c})^2 + \mathcal{O}\left(\frac{1}{N^{2/3}}\right)$$

Power of the GLRT III: How to compute type II error?

Asymptotically, of course .. **but how exactly?**

Easy but risky: The fluctuations of $\lambda_{\max}(\hat{\mathbf{R}}_n)$ under H_1

Theorem

$$\sqrt{N}\left(\lambda_{\max}(\hat{\mathbf{R}}_n) - (1+\mathbf{snr})\left(1+\frac{c_n}{\mathbf{snr}}\right)\right) \xrightarrow[N,n\to\infty]{\mathcal{L}} \mathcal{N}(0,\Gamma)$$

Then evaluate $\mathbb{P}_{H_1}\left(L_N \leq \boldsymbol{t_\alpha^n}\right)$ by Gaussian quantiles. Well .. **not a very good idea** ..

Be serious, compute the large deviations!

- one can define its error exponent $\boldsymbol{\mathcal{E}}$ as:

$$\boldsymbol{\mathcal{E}} = \lim_{N,n\to\infty} -\frac{1}{n}\log \mathbb{P}_{H_1}\left(L_N < \boldsymbol{t_\alpha^n}\right) .$$

- Hence, the type II error writes:

$$\mathbb{P}_{H_1}(L_N < t(\alpha)) \approx_{N,n\to\infty} e^{-n\boldsymbol{\mathcal{E}}}$$

Power of the GLRT IV: The error exponent

Theorem

- The type II error writes:

$$\mathbb{P}_{H_1}(L_N < \boldsymbol{t}_{\boldsymbol{\alpha}}^{\boldsymbol{n}}) \approx_{N,n\to\infty} e^{-n\boldsymbol{\mathcal{E}}}$$

in the sense that

$$\boldsymbol{\mathcal{E}} = \lim_{N,n\to\infty} -\frac{1}{n} \log \mathbb{P}_{H_1}(L_N < \boldsymbol{t}_{\boldsymbol{\alpha}}) \ .$$

- The error exponent $\boldsymbol{\mathcal{E}}$ is fully explicit

$$\boxed{\boldsymbol{\mathcal{E}} = \frac{\lambda^+ - \lambda_{spk}^\infty}{1 + \mathbf{snr}} - (1-c)\log\left(\frac{\lambda^+}{\lambda_{spk}^\infty} - 2c\left[F^+(\lambda^+) - F^+(\lambda_{spk}^\infty)\right]\right)}$$

Elements of proof

- Proof essentially based on the large deviations of $\lambda_{\max}(\hat{\mathbf{R}}_n)$ under H_1

The error exponent curve

Instead of letting the type I error fixed, it is of interest to let it go exponentially to zero:

$$\mathbb{P}_{H_0}\left(L_N > t(\mathbf{a})\right) \approx_{N,n\to\infty} e^{-n\mathbf{a}}$$

and to compute the corresponding type II error (or its error component $\boldsymbol{\mathcal{E}}(\mathbf{a})$)

$$\mathbb{P}_{H_1}\left(L_N < t(\mathbf{a})\right) \approx_{N,n\to\infty} e^{-n\boldsymbol{\mathcal{E}}(\mathbf{a})}$$

- The set $(\mathbf{a}, \boldsymbol{\mathcal{E}}(\mathbf{a}))$ such that

$$\begin{array}{lcl} \mathbb{P}_{H_0}\left(L_N > t(\mathbf{a})\right) & \approx_{N,n\to\infty} & e^{-n\mathbf{a}} \\ \mathbb{P}_{H_1}\left(L_N < t(\mathbf{a})\right) & \approx_{N,n\to\infty} & e^{-n\boldsymbol{\mathcal{E}}(\mathbf{a})} \end{array}$$

  is called the **error exponent curve** and can be fully described in terms of large deviation rate functions.

Summary

- Consider the following hypothesis

$$\vec{\mathbf{y}}(k) = \begin{cases} \sigma \vec{\mathbf{w}}(k) & \text{under } H_0 \\ \vec{\mathbf{h}}\, s(k) + \sigma \vec{\mathbf{w}}(k) & \text{under } H_1 \end{cases} \qquad \text{for } k = 1 : n$$

then the GLRT amounts to study

$$\boxed{T_n = \frac{\lambda_{\max}(\hat{\mathbf{R}}_n)}{\frac{1}{N} \operatorname{Trace} \hat{\mathbf{R}}_n}}$$

- The test statistics T_n discriminates between H_0 and H_1 if $\boxed{\mathbf{snr} = \frac{\|\vec{\mathbf{h}}\|^2}{\sigma^2} > \sqrt{c}}$
- The threshold can be asymptotically determined by Tracy-Widom quantiles.
- The type II error (equivalentlty power of the test) can be analyzed via the error exponent of the test

$$\boldsymbol{\mathcal{E}} = \lim_{N,n \to \infty} -\frac{1}{n} \log \mathbb{P}_{H_1}(L_N < \boldsymbol{t_\alpha}) \ .$$

.. But **more to come**!!!