Spiked Models in Large Random Matrices and two statistical applications

Jamal Najim
najim@univ-mlv.fr

CNRS & Université Paris Est

Linstat - Linköping, Sweden - August 2014

Large covariance matrices I

The model

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

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

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

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

Large covariance matrices I

The model

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

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

- 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**.

Large covariance matrices I

The model

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

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

- 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

To understand the spectrum of $\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*$

as

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

Large covariance matrices II

Large covariance matrices II

Remarks

1. The asymptotic regime $N, n \to \infty$ and $\frac{N}{n} \to c \in (0, \infty)$ corresponds to the case where the data dimension **is of the same order** as the number of available samples.

Large covariance matrices II

Remarks

1. The asymptotic regime $N, n \to \infty$ and $\frac{N}{n} \to c \in (0, \infty)$ corresponds to the case where the data dimension **is of the same order** as the number of available samples.
2. If N fixed and $n \to \infty$ (small data, large samples) then

$$\boxed{\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^* \longrightarrow \mathbf{R}_N}$$

Large covariance matrices II

Remarks

1. The asymptotic regime $N, n \to \infty$ and $\frac{N}{n} \to c \in (0, \infty)$ corresponds to the case where the data dimension **is of the same order** as the number of available samples.
2. If N fixed and $n \to \infty$ (small data, large samples) then

$$\boxed{\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^* \longrightarrow \mathbf{R}_N}$$

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

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

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

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

Large covariance matrices II

Remarks

1. The asymptotic regime $N, n \to \infty$ and $\frac{N}{n} \to c \in (0, \infty)$ corresponds to the case where the data dimension **is of the same order** as the number of available samples.
2. If N fixed and $n \to \infty$ (small data, large samples) then

$$\boxed{\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^* \longrightarrow \mathbf{R}_N}$$

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

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

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

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

Large covariance matrices II

Remarks

1. The asymptotic regime $N, n \to \infty$ and $\frac{N}{n} \to c \in (0, \infty)$ corresponds to the case where the data dimension **is of the same order** as the number of available samples.
2. If N fixed and $n \to \infty$ (small data, large samples) then

$$\boxed{\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^* \longrightarrow \mathbf{R}_N}$$

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

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

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

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

Otherwise stated

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

Objectives of this talk

1. to describe the limiting spectral properties of the large covariance matrix

$$\frac{1}{n}\mathbf{Y}_n\mathbf{Y}_n^* = \frac{1}{n}\mathbf{R}_n^{1/2}\mathbf{X}_n\mathbf{X}_n^*\mathbf{R}_n^{1/2}$$

Objectives of this talk

1. to describe the limiting spectral properties of the large covariance matrix

$$\frac{1}{n}\mathbf{Y}_n\mathbf{Y}_n^* = \frac{1}{n}\mathbf{R}_n^{1/2}\mathbf{X}_n\mathbf{X}_n^*\mathbf{R}_n^{1/2}$$

2. to study a particular class of covariance matrix models: **spiked models**, for which one or several eigenvalues are clearly separated from the mass of the other eigenvalues.

Objectives of this talk

1. to describe the limiting spectral properties of the large covariance matrix

$$\frac{1}{n}\mathbf{Y}_n\mathbf{Y}_n^* = \frac{1}{n}\mathbf{R}_n^{1/2}\mathbf{X}_n\mathbf{X}_n^*\mathbf{R}_n^{1/2}$$

2. to study a particular class of covariance matrix models: **spiked models**, for which one or several eigenvalues are clearly separated from the mass of the other eigenvalues.
3. to present two applications of these results in **statistical signal processing**: signal detection and direction of arrival estimation.

Spectrum and eigenvectors analysis

The resolvent

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

Spectrum and eigenvectors analysis

The resolvent

- The resolvent of $\mathbf{A}$ is $\boxed{\mathbf{Q}(z) = (\mathbf{A} - z\mathbf{I})^{-1}}$
- its singularities are exactly **eigenvalues** of $\mathbf{A}$.

Spectrum and eigenvectors analysis

The resolvent

- The resolvent of $\mathbf{A}$ is $\boxed{\mathbf{Q}(z) = (\mathbf{A} - z\mathbf{I})^{-1}}$
- its singularities are exactly **eigenvalues** of $\mathbf{A}$.
- **Problem:** if size of $\mathbf{A}$ big, then size of $\mathbf{Q}$ big as well.

Spectrum and eigenvectors analysis

The resolvent

- The resolvent of $\mathbf{A}$ is $\boxed{\mathbf{Q}(z) = (\mathbf{A} - z\mathbf{I})^{-1}}$
- its singularities are exactly **eigenvalues** of $\mathbf{A}$.
- **Problem:** if size of $\mathbf{A}$ big, then size of $\mathbf{Q}$ big as well.

The normalized trace of the resolvent

- Function

$$\boxed{g_n(z) = \frac{1}{N}\mathrm{Trace}\,(\mathbf{A} - z\mathbf{I})^{-1}}$$

  provides information on the **spectrum** of $\mathbf{A}$.
- It is the **Stieltjes transform** of the spectral measure of $\mathbf{A}$ (cf. supra)

Spectrum analysis: The Stieltjes Transform

Given a probability $\mathbb{P}$, its **Stieltjes transform** is defined by

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

Spectrum analysis: The Stieltjes Transform

Given a probability $\mathbb{P}$, its **Stieltjes transform** is defined by

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

with inverse formula

$$\int f\, d\,\mathbb{P} \quad = \quad \frac{1}{\pi} \lim_{y \downarrow 0} \Im \int_{\mathbb{R}} f(x) g(x + \mathbf{i}y)\, dx\,,$$

for f bounded continuous.

Spectrum analysis: The Stieltjes Transform

Given a probability $\mathbb{P}$, its **Stieltjes transform** is defined by

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

with inverse formula

$$\int f \, d\,\mathbb{P} \quad = \quad \frac{1}{\pi} \lim_{y \downarrow 0} \Im \int_{\mathbb{R}} f(x) g(x + \mathbf{i}y) \, dx \,,$$

for f bounded continuous.

Properties

1. Convergence in distribution is characterized by pointwise convergence of Stieltjes transforms:

$$\mathbb{P}_n \xrightarrow[n\to\infty]{\mathcal{D}} \mathbb{P} \quad \Leftrightarrow \quad \forall z \in \mathbb{C}^+, \quad g_n(z) = \int \frac{\mathbb{P}_n(d\lambda)}{\lambda - z} \quad \xrightarrow[n\to\infty]{} \quad g(z) = \int \frac{\mathbb{P}(d\lambda)}{\lambda - z}$$

Spectrum analysis: The Stieltjes Transform

Given a probability $\mathbb{P}$, its **Stieltjes transform** is defined by

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

with inverse formula

$$\int f\, d\,\mathbb{P} \quad = \quad \frac{1}{\pi} \lim_{y \downarrow 0} \Im \int_{\mathbb{R}} f(x) g(x + \mathbf{i}y)\, dx\,,$$

for f bounded continuous.

Properties

1. Convergence in distribution is characterized by pointwise convergence of Stieltjes transforms:

$$\mathbb{P}_n \xrightarrow[n\to\infty]{\mathcal{D}} \mathbb{P} \quad \Leftrightarrow \quad \forall z \in \mathbb{C}^+, \quad g_n(z) = \int \frac{\mathbb{P}_n(d\lambda)}{\lambda - z} \quad \xrightarrow[n\to\infty]{} \quad g(z) = \int \frac{\mathbb{P}(d\lambda)}{\lambda - z}$$

2. Spectral measure:

$$\mathbb{P} = \frac{1}{N} \sum_{i=1}^{N} \delta_{\lambda_i(\mathbf{A})} \quad \Rightarrow \quad g_n(z) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\lambda_i(\mathbf{A}) - z} = \frac{1}{N} \mathrm{Trace}\,(\mathbf{A} - z\mathbf{I})^{-1}$$

Spectrum analysis: The Stieltjes Transform

Given a probability $\mathbb{P}$, its **Stieltjes transform** is defined by

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

with inverse formula

$$\int f \, d\,\mathbb{P} \quad = \quad \frac{1}{\pi} \lim_{y \downarrow 0} \Im \int_{\mathbb{R}} f(x) g(x + \mathbf{i}y) \, dx ,$$

for f bounded continuous.

Properties

1. Convergence in distribution is characterized by pointwise convergence of Stieltjes transforms:

$$\mathbb{P}_n \xrightarrow[n\to\infty]{\mathcal{D}} \mathbb{P} \quad \Leftrightarrow \quad \forall z \in \mathbb{C}^+, \quad g_n(z) = \int \frac{\mathbb{P}_n(d\lambda)}{\lambda - z} \quad \xrightarrow[n\to\infty]{} \quad g(z) = \int \frac{\mathbb{P}(d\lambda)}{\lambda - z}$$

2. Spectral measure:

$$\mathbb{P} = \frac{1}{N} \sum_{i=1}^{N} \delta_{\lambda_i(\mathbf{A})} \quad \Rightarrow \quad g_n(z) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\lambda_i(\mathbf{A}) - z} = \frac{1}{N} \mathrm{Trace}\,(\mathbf{A} - z\mathbf{I})^{-1}$$

The Stieltjes transfom g_n is the **normalized trace** of the resolvent $(\mathbf{A} - z\mathbf{I})^{-1}$

Wishart Matrices

The model

- We first focus on covariance matrices in the case where $\boxed{\mathbf{R}_N = \sigma^2 \mathbf{I}_N}$.

Wishart Matrices

The model

- We first focus on covariance matrices in the case where $\boxed{\mathbf{R}_N = \sigma^2 \mathbf{I}_N}$.
- Hence $\mathbf{Y}_N$ is a $N \times n$ matrix with i.i.d. entries $\mathbb{E} Y_{ij} = 0 \ , \quad \mathbb{E}|Y_{ij}|^2 = \sigma^2 \ .$

Wishart Matrices

The model

- We first focus on covariance matrices in the case where $\boxed{\mathbf{R}_N = \sigma^2 \mathbf{I}_N}$.
- Hence $\mathbf{Y}_N$ is a $N \times n$ matrix with i.i.d. entries $\mathbb{E}Y_{ij} = 0$, $\mathbb{E}|Y_{ij}|^2 = \sigma^2$.
- Matrix $\boxed{\frac{1}{n}\mathbf{Y}_N \mathbf{Y}_N^*}$ is a **Wishart matrix**.

Wishart Matrices

The model

- We first focus on covariance matrices in the case where $\boxed{\mathbf{R}_N = \sigma^2 \mathbf{I}_N}$.
- Hence $\mathbf{Y}_N$ is a $N \times n$ matrix with i.i.d. entries $\mathbb{E}Y_{ij} = 0$, $\mathbb{E}|Y_{ij}|^2 = \sigma^2$.
- Matrix $\boxed{\frac{1}{n}\mathbf{Y}_N \mathbf{Y}_N^*}$ is a **Wishart matrix**.

The standard case $N << n$

Assume N fixed and $n \to \infty$ (small data, large sample).

Wishart Matrices

The model

- We first focus on covariance matrices in the case where $\boxed{\mathbf{R}_N = \sigma^2 \mathbf{I}_N}$.
- Hence $\mathbf{Y}_N$ is a $N \times n$ matrix with i.i.d. entries $\mathbb{E}Y_{ij} = 0$, $\mathbb{E}|Y_{ij}|^2 = \sigma^2$.
- Matrix $\boxed{\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*}$ is a **Wishart matrix**.

The standard case $N << n$

Assume N fixed and $n \to \infty$ (small data, large sample). Since

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

L.L.N implies

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

Wishart Matrices

The model

- We first focus on covariance matrices in the case where $\boxed{\mathbf{R}_N = \sigma^2 \mathbf{I}_N}$.
- Hence $\mathbf{Y}_N$ is a $N \times n$ matrix with i.i.d. entries $\mathbb{E}Y_{ij} = 0 \ , \quad \mathbb{E}|Y_{ij}|^2 = \sigma^2 \ .$
- Matrix $\boxed{\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*}$ is a **Wishart matrix**.

The standard case $N << n$

Assume N fixed and $n \to \infty$ (small data, large sample). Since

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

L.L.N implies

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

In particular,

- all the eigenvalues of $\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*$ converge to σ^2,

Wishart Matrices

The model

- We first focus on covariance matrices in the case where $\boxed{\mathbf{R}_N = \sigma^2 \mathbf{I}_N}$.
- Hence $\mathbf{Y}_N$ is a $N \times n$ matrix with i.i.d. entries $\mathbb{E}Y_{ij} = 0$, $\quad \mathbb{E}|Y_{ij}|^2 = \sigma^2$.
- Matrix $\boxed{\frac{1}{n}\mathbf{Y}_N \mathbf{Y}_N^*}$ is a **Wishart matrix**.

The standard case $N << n$

Assume N fixed and $n \to \infty$ (small data, large sample). Since

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

L.L.N implies

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

In particular,

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

Marčenko-Pastur theorem

Theorem

- Consider the **spectral measure** L_N:

Marčenko-Pastur theorem

Theorem

- Consider the **spectral measure** L_N:

$$L_N = \frac{1}{N} \sum_{i=1}^{N} \delta_{\lambda_i} \ , \quad \lambda_i = \lambda_i \left(\frac{1}{n} \mathbf{Y}_N \mathbf{Y}_N^* \right)$$

Marčenko-Pastur theorem

Theorem

- Consider the **spectral measure** L_N:

$$L_N = \frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i}\ , \quad \lambda_i = \lambda_i\left(\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*\right)$$

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

$$L_N \xrightarrow[N,n\to\infty]{} \mathbb{P}_{\check{\mathrm{M}}\mathrm{P}} \quad \text{in distribution as } \frac{N}{n} \xrightarrow[n\to\infty]{} c \in (0,\infty)$$

Marčenko-Pastur theorem

Theorem

- Consider the **spectral measure** L_N:

$$L_N = \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i} \ , \quad \lambda_i = \lambda_i\left(\frac{1}{n}\mathbf{Y}_N \mathbf{Y}_N^*\right)$$

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

$$L_N \xrightarrow[N,n\to\infty]{} \mathbb{P}_{\check{\mathrm{M}}\mathrm{P}} \quad \text{in distribution as } \frac{N}{n} \xrightarrow[n\to\infty]{} c \in (0,\infty)$$

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

Marčenko-Pastur theorem

Theorem

- Consider the **spectral measure** L_N:

$$L_N = \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i} \ , \quad \lambda_i = \lambda_i\left(\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*\right)$$

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

$$L_N \xrightarrow[N,n\to\infty]{} \mathbb{P}_{\check{\mathrm{M}}\mathrm{P}} \quad \text{in distribution as } \frac{N}{n} \xrightarrow[n\to\infty]{} c \in (0,\infty)$$

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

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

with

$$\begin{cases} a & = & \sigma^2(1-\sqrt{c})^2 \\ b & = & \sigma^2(1+\sqrt{c})^2 \end{cases}$$

Histogram for Wishart matrices

Matrix model: Wishart matrix

Consider the spectrum of $\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*$ in the regime where

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

Plot the **histogram of its eigenvalues**.

Histogram for Wishart matrices

Matrix model: Wishart matrix

Consider the spectrum of $\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*$ in the regime where

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

Plot the **histogram of its eigenvalues**.

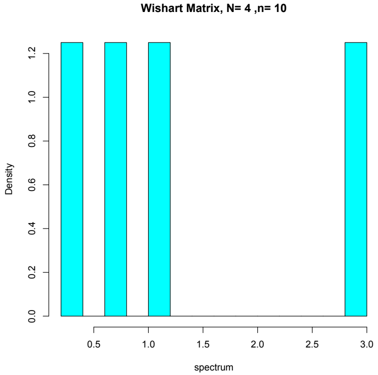

Figure : Spectrum's histogram - $\frac{N}{n} = 0.7$

Histogram for Wishart matrices

Matrix model: Wishart matrix

Consider the spectrum of $\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*$ in the regime where

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

Plot the **histogram of its eigenvalues**.

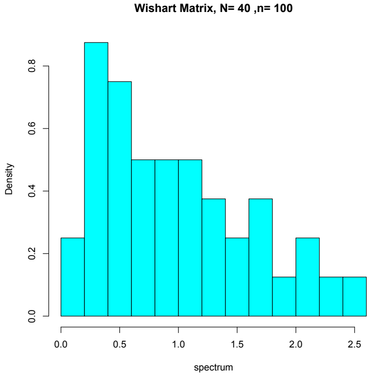

Figure : Spectrum's histogram - $\frac{N}{n} = 0.7$

Histogram for Wishart matrices

Matrix model: Wishart matrix

Consider the spectrum of $\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*$ in the regime where

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

Plot the **histogram of its eigenvalues**.

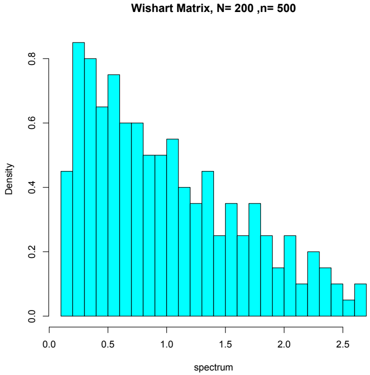

Figure : Spectrum's histogram - $\frac{N}{n} = 0.7$

Histogram for Wishart matrices

Matrix model: Wishart matrix

Consider the spectrum of $\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*$ in the regime where

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

Plot the **histogram of its eigenvalues**.

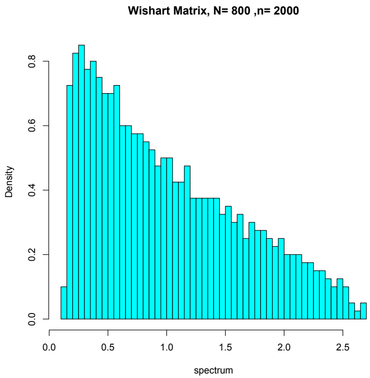

Figure : Spectrum's histogram - $\frac{N}{n} = 0.7$

Histogram for Wishart matrices

Matrix model: Wishart matrix

Consider the spectrum of $\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*$ in the regime where

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

Plot the **histogram of its eigenvalues**.

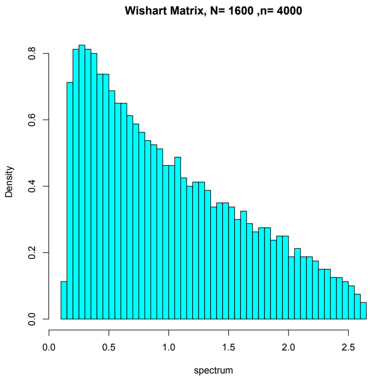

Figure : Spectrum's histogram - $\frac{N}{n} = 0.7$

Histogram for Wishart matrices: Marčenko-Pastur's theorem

Matrix model: Wishart matrix

Consider the spectrum of $\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*$ in the regime where

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

Plot the **histogram of its eigenvalues**.

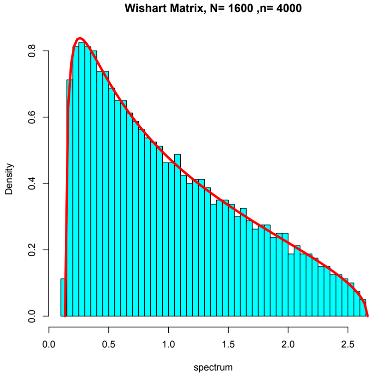

Figure : Marčenko-Pastur's distribution (in red)

Marčenko-Pastur's theorem (1967)

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

Elements of proof

1. Convergence of the Stieltjes transform.

Elements of proof

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.

Elements of proof

1. Convergence of the Stieltjes transform. Since

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

   we prove the convergence of g_n.

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

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

Elements of proof

1. Convergence of the Stieltjes transform. Since

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

   we prove the convergence of g_n.

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

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

3. Necessarily,

$$g_n \xrightarrow[N,n\to\infty]{} \mathbf{g}_{\check{\mathrm{MP}}}$$

   which satisfies **the fixed point equation**:

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

Elements of proof

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

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

3. Necessarily,

$$g_n \xrightarrow[N,n\to\infty]{} \mathbf{g}_{\check{\mathrm{M}}\mathrm{P}}$$

   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. Solving explicitely the previous equation, we identify

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

Theorem

Recall the notations

$$\mathbf{Y}_n = \mathbf{R}_N^{1/2}\mathbf{X}_N \quad \text{and} \quad g_n(z) = \frac{1}{N}\mathrm{Trace}\left(\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^* - z\mathbf{I}_N\right)^{-1}$$

We are interested in the limiting behaviour of

$$L_N = \frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i} \qquad \text{with} \qquad \lambda_i = \lambda_i\left(\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*\right)$$

Canonical equation

- Unknown $\mathbf{t}_N$ is a Stieltjes transform, solution of

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

Theorem

Recall the notations

$$\mathbf{Y}_n = \mathbf{R}_N^{1/2}\mathbf{X}_N \quad \text{and} \quad g_n(z) = \frac{1}{N}\mathrm{Trace}\left(\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^* - z\mathbf{I}_N\right)^{-1}$$

We are interested in the limiting behaviour of

$$L_N = \frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i} \qquad \text{with} \qquad \lambda_i = \lambda_i\left(\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*\right)$$

Canonical equation

- Unknown $\mathbf{t}_N$ is a Stieltjes transform, solution of

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

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

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

Theorem

Recall the notations

$$\mathbf{Y}_n = \mathbf{R}_N^{1/2}\mathbf{X}_N \quad \text{and} \quad g_n(z) = \frac{1}{N}\text{Trace}\left(\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^* - z\mathbf{I}_N\right)^{-1}$$

We are interested in the limiting behaviour of

$$L_N = \frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i} \qquad \text{with} \qquad \lambda_i = \lambda_i\left(\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*\right)$$

Canonical equation

- Unknown $\mathbf{t}_N$ is a Stieltjes transform, solution of

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

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

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

Convergence

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

Theorem

Recall the notations

$$\mathbf{Y}_n = \mathbf{R}_N^{1/2}\mathbf{X}_N \quad \text{and} \quad g_n(z) = \frac{1}{N}\mathrm{Trace}\left(\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^* - z\mathbf{I}_N\right)^{-1}$$

We are interested in the limiting behaviour of

$$L_N = \frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i} \qquad \text{with} \qquad \lambda_i = \lambda_i\left(\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*\right)$$

Canonical equation

- Unknown $\mathbf{t}_N$ is a Stieltjes transform, solution of

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

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

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

Convergence

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

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

Theorem

Recall the notations

$$\mathbf{Y}_n = \mathbf{R}_N^{1/2}\mathbf{X}_N \quad \text{and} \quad g_n(z) = \frac{1}{N}\mathrm{Trace}\left(\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^* - z\mathbf{I}_N\right)^{-1}$$

We are interested in the limiting behaviour of

$$L_N = \frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i} \qquad \text{with} \qquad \lambda_i = \lambda_i\left(\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*\right)$$

Canonical equation

- Unknown $\mathbf{t}_N$ is a Stieltjes transform, solution of

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

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

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

Convergence

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

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

Remark

Assume moreover that

$$\frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i(\mathbf{R_N})} \xrightarrow[N,n\to\infty]{} \mathbb{P}^{\mathbf{R}}$$

Remark

Assume moreover that

$$\frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i(\mathbf{R_N})} \xrightarrow[N,n\to\infty]{} \mathbb{P}^{\mathbf{R}}$$

Then instead of having a series of **canonical equations** depending on N:

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

Remark

Assume moreover that

$$\frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i(\mathbf{R_N})} \xrightarrow[N,n\to\infty]{} \mathbb{P}^{\mathbf{R}}$$

Then instead of having a series of **canonical equations** depending on N:

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

we can obtain a "limiting equation"

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

Remark

Assume moreover that

$$\frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i(\mathbf{R_N})} \xrightarrow[N,n\to\infty]{} \mathbb{P}^{\mathbf{R}}$$

Then instead of having a series of **canonical equations** depending on N:

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

we can obtain a "limiting 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 genuine limits

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

Remark

Assume moreover that

$$\frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i(\mathbf{R_N})} \xrightarrow[N,n\to\infty]{} \mathbb{P}^{\mathbf{R}}$$

Then instead of having a series of **canonical equations** depending on N:

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

we can obtain a "limiting 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 genuine limits

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

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

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

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 the limiting spectral distribution

$$\mathbb{P}_\infty$$

for different values of c.

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

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 the limiting spectral distribution

$$\mathbb{P}_\infty$$

for different values of c.

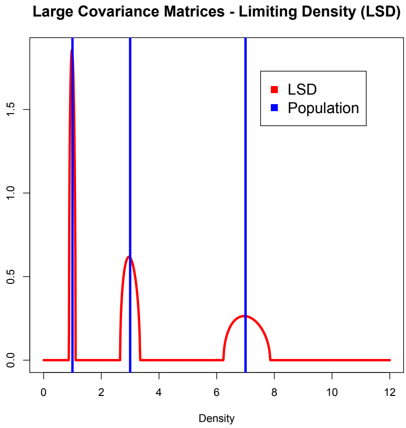

Figure : Plot of the Limiting Spectral Measure for $c = 0.01$

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

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 the limiting spectral distribution

$$\mathbb{P}_\infty$$

for different values of c.

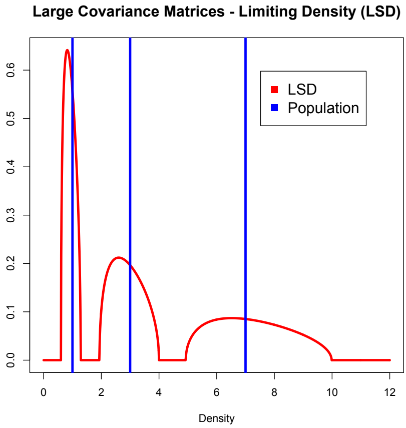

Figure : Plot of the Limiting Spectral Measure for $c = 0.1$

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

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 the limiting spectral distribution

$$\mathbb{P}_\infty$$

for different values of c.

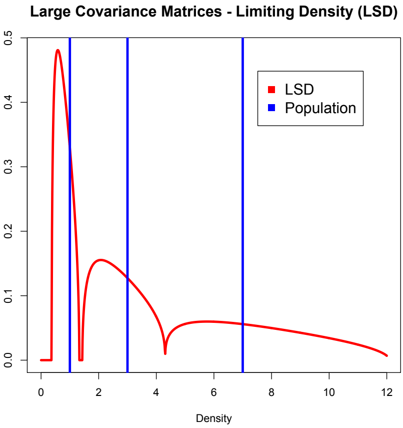

Figure : Plot of the Limiting Spectral Measure for $c = 0.25$

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

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 the limiting spectral distribution

$$\mathbb{P}_\infty$$

for different values of c.

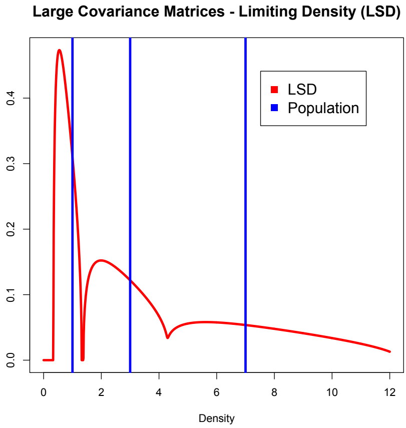

Figure : Plot of the Limiting Spectral Measure for $c = 0.275$

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

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 the limiting spectral distribution

$$\mathbb{P}_\infty$$

for different values of c.

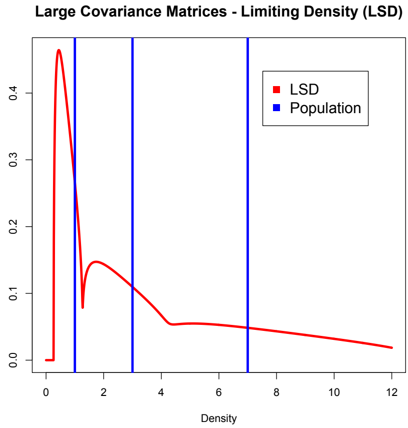

Figure : Plot of the Limiting Spectral Measure for $c = 0.35$

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

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 the limiting spectral distribution

$$\mathbb{P}_\infty$$

for different values of c.

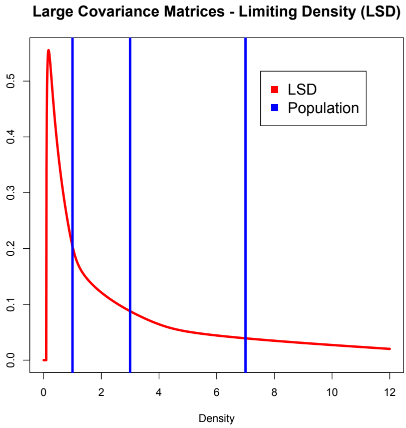

Figure : Plot of the Limiting Spectral Measure for $c = 0.6$

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

The largest eigenvalue in M̌P model

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

The largest eigenvalue in M̌P model

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

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

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

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

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

Introduction

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

Introduction

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:

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

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

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

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

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

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

(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.

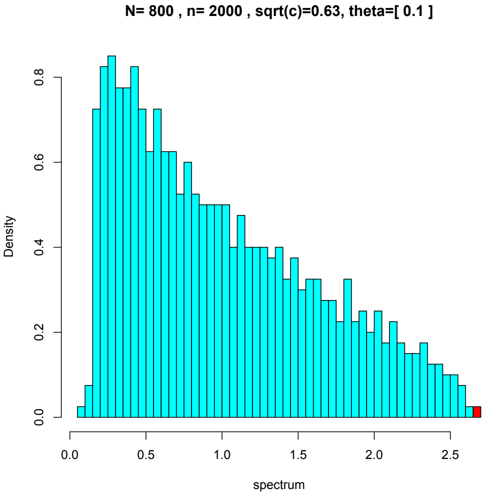

Figure : The largest eigenvalue (red) converges to the right edge of the bulk

Spiked Models I

Definition

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

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.
- and the $\vec{\mathbf{u}}_i$'s are orthonormal

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.
- and the $\vec{\mathbf{u}}_i$'s are orthonormal

Consider

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

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

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.
- and the $\vec{\mathbf{u}}_i$'s are orthonormal

Consider

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

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

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

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

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.
- and the $\vec{\mathbf{u}}_i$'s are orthonormal

Consider

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

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

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

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

Very important: The number k of perturbations is finite

Spiked Models II

Remarks

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

$$\mathbf{Y}_N = \frac{1}{n}\mathbf{R}_N^{1/2}\mathbf{X}_N \qquad \text{and} \qquad \mathbf{R}_N = \mathbf{I}_N + \sum_{\ell=1}^{k} \theta_\ell \vec{\mathbf{u}}_\ell \vec{\mathbf{u}}_\ell^*$$

Spiked Models II

Remarks

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

$$\mathbf{Y}_N = \frac{1}{n}\mathbf{R}_N^{1/2}\mathbf{X}_N \qquad \text{and} \qquad \mathbf{R}_N = \mathbf{I}_N + \sum_{\ell=1}^{k} \theta_\ell \vec{\mathbf{u}}_\ell \vec{\mathbf{u}}_\ell^*$$

- 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 II

Remarks

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

$$\mathbf{Y}_N = \frac{1}{n}\mathbf{R}_N^{1/2}\mathbf{X}_N \qquad \text{and} \qquad \mathbf{R}_N = \mathbf{I}_N + \sum_{\ell=1}^{k} \theta_\ell \vec{\mathbf{u}}_\ell \vec{\mathbf{u}}_\ell^*$$

- 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

Spiked Models II

Remarks

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

$$\mathbf{Y}_N = \frac{1}{n}\mathbf{R}_N^{1/2}\mathbf{X}_N \qquad \text{and} \qquad \mathbf{R}_N = \mathbf{I}_N + \sum_{\ell=1}^{k} \theta_\ell \vec{\mathbf{u}}_\ell \vec{\mathbf{u}}_\ell^*$$

- 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 the spectral limit of $L_N\left(\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*\right)$?

Spiked Models II

Remarks

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

$$\mathbf{Y}_N = \frac{1}{n}\mathbf{R}_N^{1/2}\mathbf{X}_N \qquad \text{and} \qquad \mathbf{R}_N = \mathbf{I}_N + \sum_{\ell=1}^{k} \theta_\ell \vec{\mathbf{u}}_\ell \vec{\mathbf{u}}_\ell^*$$

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

Spiked Models II

Remarks

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

$$\mathbf{Y}_N = \frac{1}{n}\mathbf{R}_N^{1/2}\mathbf{X}_N \qquad \text{and} \qquad \mathbf{R}_N = \mathbf{I}_N + \sum_{\ell=1}^{k} \theta_\ell \vec{\mathbf{u}}_\ell \vec{\mathbf{u}}_\ell^*$$

- 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 the spectral limit of $L_N\left(\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*\right)$?
- What is the influence of $\mathbf{\Pi}_N$ over $\lambda_{\max}\left(\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*\right)$?
- What is the influence of $\mathbf{\Pi}_N$ over the eigenvector $\vec{\mathbf{v}}_{\max}$ associated to $\lambda_{\max}$?

Simulations

Simulations

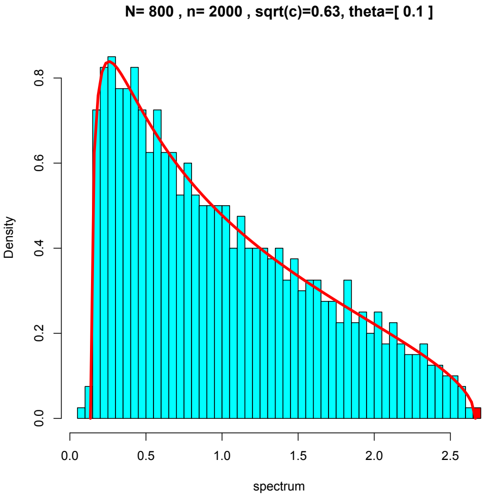

Figure : Spiked model - strength of the perturbation $\theta = 0.1$

Simulations

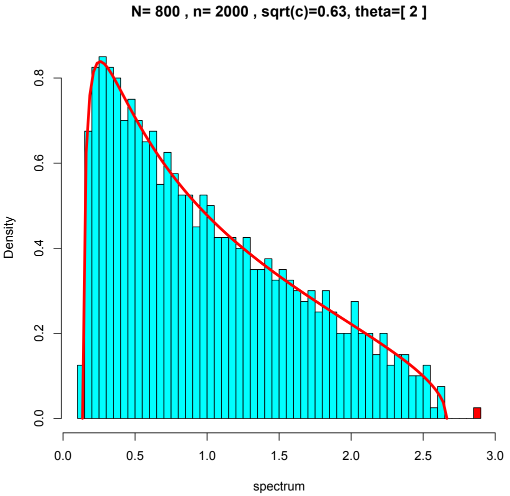

Figure : Spiked model - strength of the perturbation $\theta = 2$

Simulations

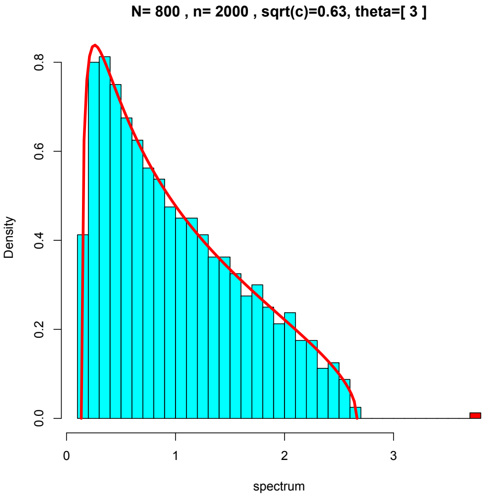

Figure : Spiked model - strength of the perturbation $\theta = 3$

Simulations

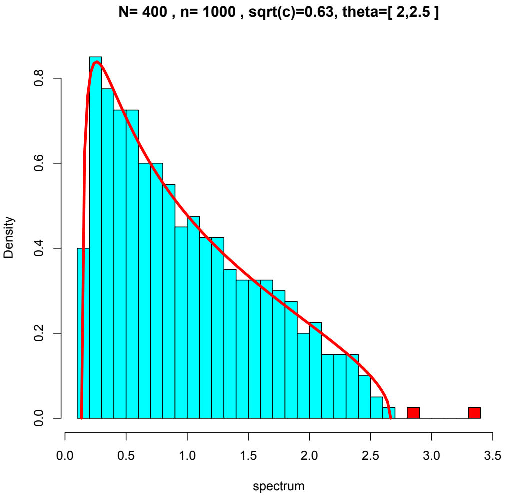

Figure : Spiked model - Two spikes

Simulations

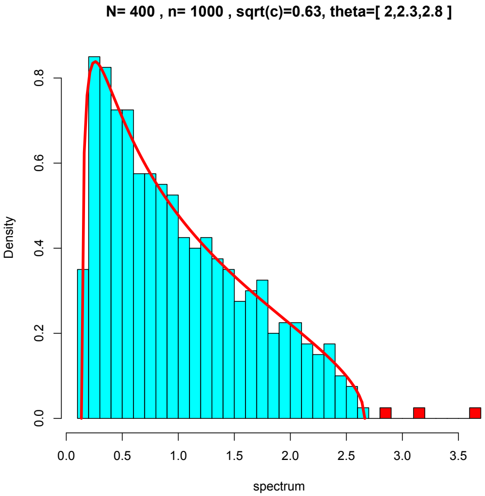

Figure : Spiked model - Three spikes

Simulations

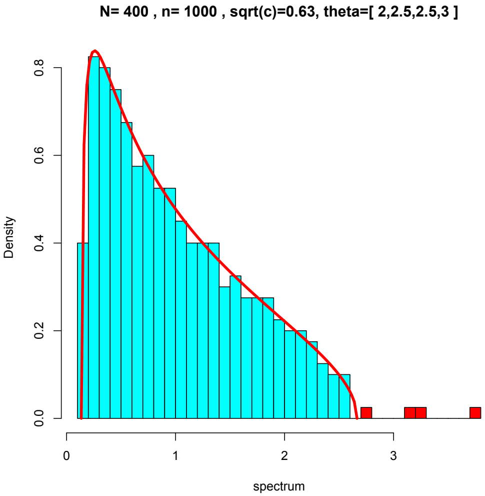

Figure : Spiked model - Multiple spikes

The limiting spectral measure

Theorem

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

The limiting spectral measure

Theorem

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

Remark

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

The largest eigenvalue

We consider the following spiked model:

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

The largest eigenvalue

We consider the following spiked model:

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

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

The largest eigenvalue

We consider the following spiked model:

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

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

Theorem

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

The largest eigenvalue

We consider the following spiked model:

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

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

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

The largest eigenvalue

We consider the following spiked model:

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

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

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

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

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

The largest eigenvalue

We consider the following spiked model:

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

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

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

- 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

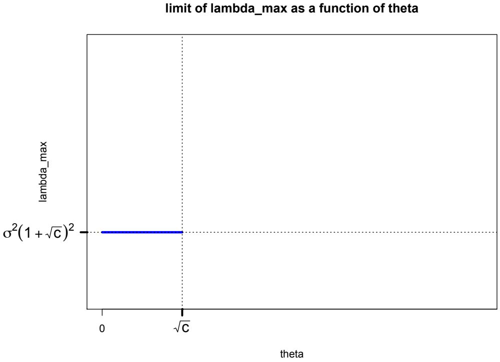

Figure : Limit of largest eigenvalue $\lambda_{\max}$ as a function of the perturbation θ

Phase transition Phenomenon

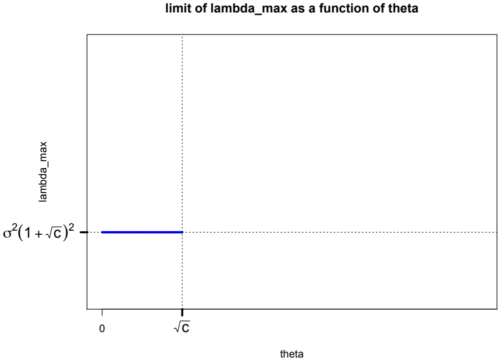

Figure : Limit of largest eigenvalue $\lambda_{\max}$ as a function of the perturbation θ

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

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

Phase transition Phenomenon

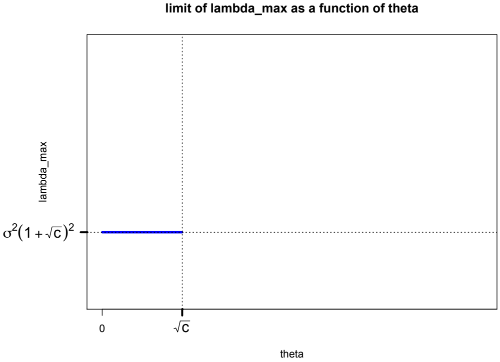

Figure : Limit of largest eigenvalue $\lambda_{\max}$ as a function of the perturbation θ

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

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

Below the threshold $\sqrt{c}$, $\lambda_{\max}\left(\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*\right)$ asymptotically **sticks to the bulk**.

Phase transition Phenomenon

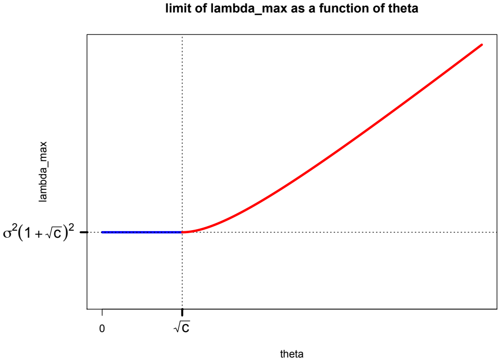

Figure : Limit of largest eigenvalue $\lambda_{\max}$ as a function of the perturbation θ

Phase transition Phenomenon

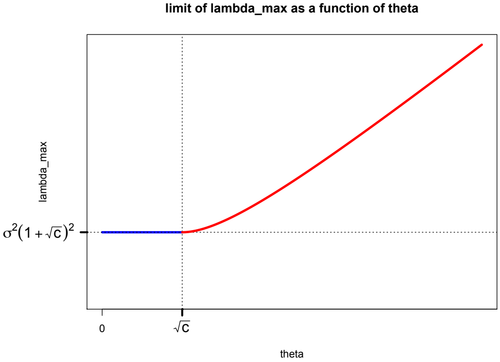

Figure : Limit of largest eigenvalue $\lambda_{\max}$ as a function of the perturbation θ

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

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

Phase transition Phenomenon

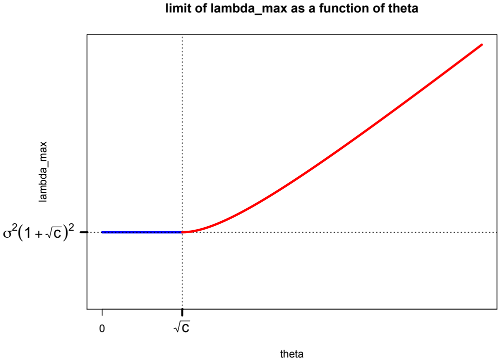

Figure : Limit of largest eigenvalue $\lambda_{\max}$ as a function of the perturbation θ

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

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

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

The eigenvector associated to $\lambda_{\max}$ I

- Let:

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

The eigenvector associated to $\lambda_{\max}$ I

- Let:

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

- Let $\vec{\boldsymbol{v}}_{\max}$ be the eigenvector associated to $\lambda_{\max}$:

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

The eigenvector associated to $\lambda_{\max}$ I

- Let:

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

- Let $\vec{\boldsymbol{v}}_{\max}$ be the eigenvector associated to $\lambda_{\max}$:

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

Question

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

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

The eigenvector associated to $\lambda_{\max}$ I

- Let:

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

- Let $\vec{\boldsymbol{v}}_{\max}$ be the eigenvector associated to $\lambda_{\max}$:

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

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:

The eigenvector associated to $\lambda_{\max}$ I

- Let:

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

- Let $\vec{\boldsymbol{v}}_{\max}$ be the eigenvector associated to $\lambda_{\max}$:

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

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 $\boxed{\theta \le \sqrt{c}}$ then $\lambda_{\max}$ converges to **the right edge of ṀP bulk.**

The eigenvector associated to $\lambda_{\max}$ I

- Let:

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

- Let $\vec{\boldsymbol{v}}_{\max}$ be the eigenvector associated to $\lambda_{\max}$:

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

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 $\boxed{\theta \le \sqrt{c}}$ then $\lambda_{\max}$ converges to **the right edge of M̌P bulk.**
- if $\boxed{\theta > \sqrt{c}}$ then $\lambda_{\max}$ **separates from the bulk.**

The eigenvector associated to $\lambda_{\max}$ II

Preliminary observations

The eigenvector associated to $\lambda_{\max}$ II

Preliminary observations

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

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

The eigenvector associated to $\lambda_{\max}$ II

Preliminary observations

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

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

As a consequence:

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

The eigenvector associated to $\lambda_{\max}$ II

Preliminary observations

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

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

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 $\boxed{\dim(\vec{\boldsymbol{v}}_{\max}) = N \nearrow \infty}$.

The eigenvector associated to $\lambda_{\max}$ II

Preliminary observations

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

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

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 $\vec{\boldsymbol{v}}_{\max}$ of a generic deterministic vector $\vec{\boldsymbol{a}}_N$

The eigenvector associated to $\lambda_{\max}$ II

Preliminary observations

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

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

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

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

The eigenvector associated to $\lambda_{\max}$ III

Theorem

Assume that $\theta > \sqrt{c}$ and let $\vec{\boldsymbol{a}}_N$ be a deterministic vector with norm 1, then

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

The eigenvector associated to $\lambda_{\max}$ III

Theorem

Assume that $\theta > \sqrt{c}$ and let $\vec{a}_N$ be a deterministic vector with norm 1, then

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

Remarks

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

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

The eigenvector associated to $\lambda_{\max}$ III

Theorem

Assume that $\theta > \sqrt{c}$ and 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

- 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 $c \to 0$.

The eigenvector associated to $\lambda_{\max}$ III

Theorem

Assume that $\theta > \sqrt{c}$ and 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

- 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 $c \to 0$.
- we recover the fact that if N is finite, $n \to \infty$ (small data, large samples), then

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

Summary

Summary

Spiked model

Let

- $\mathbf{\Pi}_N$ a small perturbation of the identity [Example: $\mathbf{\Pi}_N = \mathbf{I}_N + \theta \vec{\mathbf{u}}\vec{\mathbf{u}}^*$]

Summary

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

Summary

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

Summary

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

Global regime

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

Summary

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

Global regime

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

Largest eigenvalue (rank one perturbation)

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

Summary

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

Global regime

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

Largest eigenvalue (rank one perturbation)

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

Summary

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

Global regime

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

Largest eigenvalue (rank one perturbation)

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

Associated eigenvector

- In the large dimension setting, $\boxed{\vec{\mathbf{v}}_{\max} \approx \left(1 - \frac{c}{\theta^2}\right)\left(1 + \frac{c}{\theta}\right)^{-1}\vec{\mathbf{u}}}$

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,

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 (σ **unknown**):

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

The hypothesis testing problem

Statistical Setup

let

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

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 (σ **unknown**):

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

- $\vec{\mathbf{h}}$ is a $N \times 1$ **deterministic** and **unknown** vector and typically represents the propagation channel

The hypothesis testing problem

Statistical Setup

let

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

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 (σ **unknown**):

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

- $\vec{\mathbf{h}}$ 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

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 (σ **unknown**):

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

- $\vec{\mathbf{h}}$ 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 $(\vec{\mathbf{y}}(k), 1 \le k \le n)$, 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 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 (σ **unknown**):

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

- $\vec{\mathbf{h}}$ 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 $(\vec{\mathbf{y}}(k), 1 \le k \le n)$, 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

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

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 (σ **unknown**):

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

- $\vec{\mathbf{h}}$ 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 $(\vec{\mathbf{y}}(k), 1 \le k \le n)$, 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

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

The GLRT

Since σ^2 and $\vec{\mathbf{h}}$ are **unknown**, we cannot use the likelihood ratio test (which would have been optimal by Neyman-Pearson).

The GLRT

Since σ^2 and $\vec{\mathbf{h}}$ are **unknown**, we cannot use the likelihood ratio test (which would have been optimal 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.

The GLRT

Since σ^2 and $\vec{\mathbf{h}}$ are **unknown**, we cannot use the likelihood ratio test (which would have been optimal 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

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

The test statistics T_n

Denote by $\widetilde{\mathbf{X}}_N$ a matrix with i.i.d. 0/1 entries and let

$$\mathbf{\Pi}_N = \left(\mathbf{I}_N + \frac{\|\vec{\mathbf{h}}\|^2}{\sigma^2} \vec{\mathbf{u}}\vec{\mathbf{u}}^* \right) \quad \text{where} \quad \vec{\mathbf{u}} = \frac{\vec{\mathbf{h}}}{\|\vec{\mathbf{h}}\|} \ .$$

The test statistics T_n

Denote by $\widetilde{\mathbf{X}}_N$ a matrix with i.i.d. 0/1 entries and let

$$\mathbf{\Pi}_N = \left(\mathbf{I}_N + \frac{\|\vec{\mathbf{h}}\|^2}{\sigma^2} \vec{\mathbf{u}}\vec{\mathbf{u}}^* \right) \quad \text{where} \quad \vec{\mathbf{u}} = \frac{\vec{\mathbf{h}}}{\|\vec{\mathbf{h}}\|} \ .$$

After some massaging, it turns out that

The test statistics T_n

Denote by $\widetilde{\mathbf{X}}_N$ a matrix with i.i.d. 0/1 entries and let

$$\mathbf{\Pi}_N = \left(\mathbf{I}_N + \frac{\|\vec{\mathbf{h}}\|^2}{\sigma^2}\vec{\mathbf{u}}\vec{\mathbf{u}}^*\right) \quad \text{where} \quad \vec{\mathbf{u}} = \frac{\vec{\mathbf{h}}}{\|\vec{\mathbf{h}}\|} \ .$$

After some massaging, it turns out that

- Under H_0,

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

T_N is simply the largest eigenvalue of a Wishart matrix

The test statistics T_n

Denote by $\widetilde{\mathbf{X}}_N$ a matrix with i.i.d. 0/1 entries and let

$$\mathbf{\Pi}_N = \left(\mathbf{I}_N + \frac{\|\vec{\mathbf{h}}\|^2}{\sigma^2} \vec{\mathbf{u}}\vec{\mathbf{u}}^* \right) \quad \text{where} \quad \vec{\mathbf{u}} = \frac{\vec{\mathbf{h}}}{\|\vec{\mathbf{h}}\|} \ .$$

After some massaging, it turns out that

- Under H_0,

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

T_N is simply the largest eigenvalue of a Wishart matrix

- Under H_1,

$$T_n = \lambda_{\max} \left(\frac{1}{n} \mathbf{\Pi}_N^{1/2} \widetilde{\mathbf{X}}_N \widetilde{\mathbf{X}}_N^* \mathbf{\Pi}_N^{1/2} \right)$$

T_N is the largest eigenvalue of a rank-one perturbated Wishart matrix

The test statistics T_n

Denote by $\widetilde{\mathbf{X}}_N$ a matrix with i.i.d. 0/1 entries and let

$$\mathbf{\Pi}_N = \left(\mathbf{I}_N + \frac{\|\vec{\mathbf{h}}\|^2}{\sigma^2} \vec{\mathbf{u}}\vec{\mathbf{u}}^* \right) \quad \text{where} \quad \vec{\mathbf{u}} = \frac{\vec{\mathbf{h}}}{\|\vec{\mathbf{h}}\|} \ .$$

After some massaging, it turns out that

- Under H_0,

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

T_N is simply the largest eigenvalue of a Wishart matrix

- Under H_1,

$$T_n = \lambda_{\max} \left(\frac{1}{n} \mathbf{\Pi}_N^{1/2} \widetilde{\mathbf{X}}_N \widetilde{\mathbf{X}}_N^* \mathbf{\Pi}_N^{1/2} \right)$$

T_N is the largest eigenvalue of a rank-one perturbated Wishart matrix

The good news is that in both case, we can describe the limit.

Limits of the test statistics T_n

Limits of the test statistics T_n

Under H_0

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

Limits of the test statistics T_n

Under H_0

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

Under H_1

Denote by

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

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

Limits of the test statistics T_n

Under H_0

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

Under H_1

Denote by

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

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

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

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

Limits of the test statistics T_n

Under H_0

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

Under H_1

Denote by

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

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

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

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

Limits of the test statistics T_n

Under H_0

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

Under H_1

Denote by

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

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

- If $\boxed{\mathbf{snr} \leq \sqrt{c}}$ then the test statistics does not discriminate between the two hypotheses.

Limit of the test statistics T_n III

Remarks

- If $\boxed{\mathbf{snr} \leq \sqrt{c}}$ then the test statistics does not discriminate between the two hypotheses.
- Condition $\boxed{\mathbf{snr} > \sqrt{c}}$ is **automatically fulfilled** in the standard regime where

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

Limit of the test statistics T_n III

Remarks

- If $\boxed{\mathbf{snr} \leq \sqrt{c}}$ then the test statistics does not discriminate between the two hypotheses.
- Condition $\boxed{\mathbf{snr} > \sqrt{c}}$ is **automatically fulfilled** in the standard regime where

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

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

Limit of the test statistics T_n III

Remarks

- If $\boxed{\mathbf{snr} \le \sqrt{c}}$ then the test statistics does not discriminate between the two hypotheses.
- Condition $\boxed{\mathbf{snr} > \sqrt{c}}$ is **automatically fulfilled** in the standard regime where

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

- 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

Detection occurs if $\mathbf{snr}$ higher than **asymptotic data noise**.

Simulations

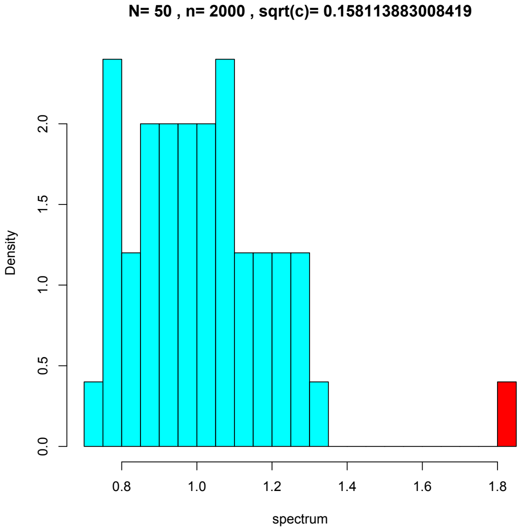

Figure : Influence of **asymptotic data noise** as $\sqrt{c}$ increases

Simulations

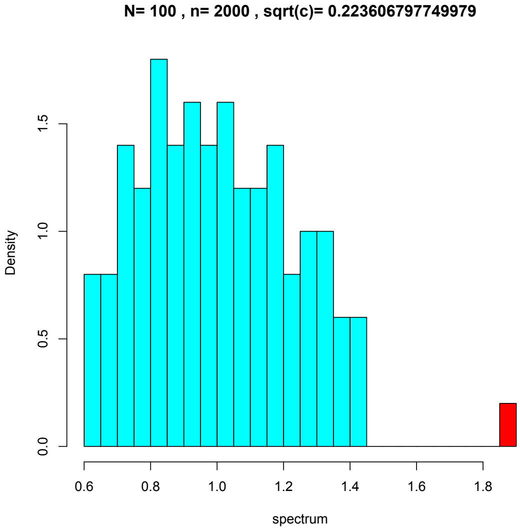

Figure : Influence of **asymptotic data noise** as $\sqrt{c}$ increases

Simulations

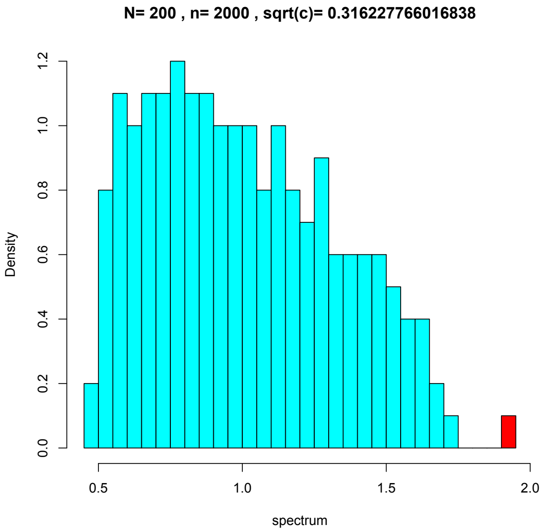

Figure : Influence of **asymptotic data noise** as $\sqrt{c}$ increases

Simulations

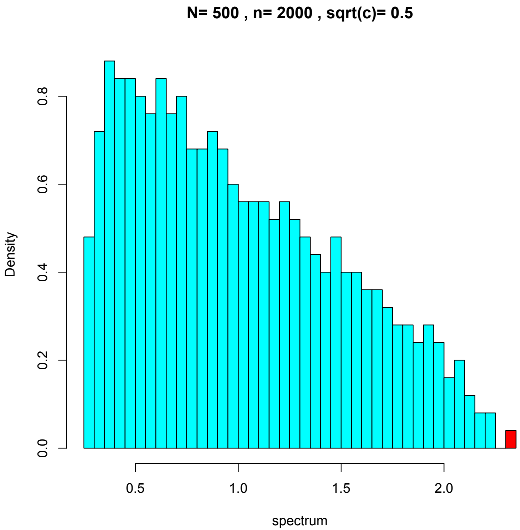

Figure : Influence of **asymptotic data noise** as $\sqrt{c}$ increases

Simulations

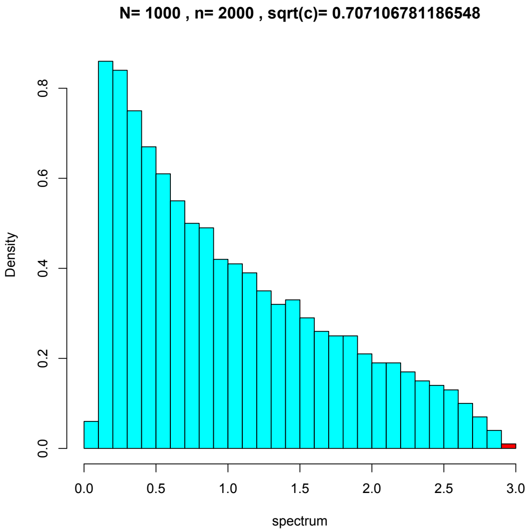

Figure : Influence of **asymptotic data noise** as $\sqrt{c}$ increases

Threshold of the test I

- The exact distribution of the statistics L_n is needed to set the threshold of the test for a given confidence level $\alpha \in (0, 1)$:

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

Threshold of the test I

- The exact distribution of the statistics L_n is needed to set the threshold of the test for a given confidence level $\alpha \in (0, 1)$:

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

  but **hard to obtain**.

Threshold of the test I

- The exact distribution of the statistics L_n is needed to set the threshold of the test for 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

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

Threshold of the test I

- The exact distribution of the statistics L_n is needed to set the threshold of the test for 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

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

- Recall that L_N is the largest eigenvalue of a Whishat matrix $\frac{1}{n}\widetilde{\mathbf{X}}_N\widetilde{\mathbf{X}}_N^*$.

Threshold of the test I

- The exact distribution of the statistics L_n is needed to set the threshold of the test for 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

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

- Recall that L_N is the largest eigenvalue of a Whishat matrix $\frac{1}{n}\widetilde{\mathbf{X}}_N\widetilde{\mathbf{X}}_N^*$.

◦ We need to understand the fluctuations of $\lambda_{\max}\left(\frac{1}{n}\widetilde{\mathbf{X}}_N\widetilde{\mathbf{X}}_N^*\right)$ under H_0,

Threshold of the test II

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

Threshold of the test II

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

Theorem (Tracy-Widom)

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

Threshold of the test II

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

Theorem (Tracy-Widom)

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

where

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

Threshold of the test II

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

Theorem (Tracy-Widom)

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

where

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

Otherwise stated,

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

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

Threshold of the test II

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

Theorem (Tracy-Widom)

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

where

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

Otherwise stated,

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

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

- Definition of Tracy-Widom distribution complicated ..

Threshold of the test II

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

Theorem (Tracy-Widom)

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

where

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

Otherwise stated,

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

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

- Definition of Tracy-Widom distribution complicated ..

Don't bother .. just download it

- For simulations, cf. `R Package 'RMTstat'`, by Johnstone et al.

Tracy-Widom curve

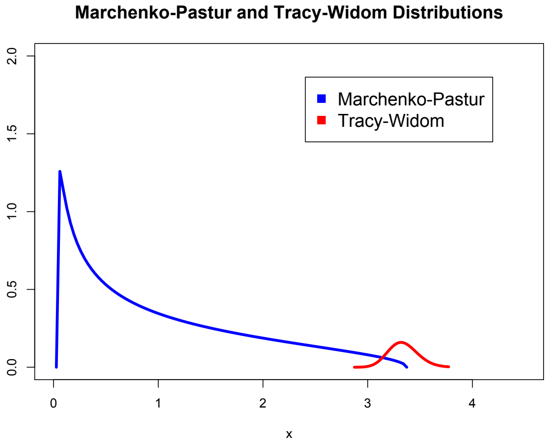

Figure : Fluctuations of the largest eigenvalue $\lambda_{\max}(\hat{\mathbf{R}}_n)$ under H_0

Power of the GLRT I

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 I

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

Large deviations

- Using large deviation techniques, one can compute the error exponent $\boldsymbol{\mathcal{E}}$ as:

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

Power of the GLRT I

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

Large deviations

- Using large deviation techniques, one can compute the error exponent $\boldsymbol{\mathcal{E}}$ as:

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

- Hence, the type II error writes:

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

Summary

- Consider the following hypothesis

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

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

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

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

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.

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

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

which relies on the study of large deviations of $\lambda_{\max}$ under H_1.

Estimation problem

The aim of the problem is:

Estimation problem

The aim of the problem is:

- given n observations $\vec{\boldsymbol{y}}_1, \cdots, \vec{\boldsymbol{y}}_n$ each of dimension N with

$$\vec{\boldsymbol{y}} = \vec{\boldsymbol{y}}(\boldsymbol{\varphi}_1, \cdots, \boldsymbol{\varphi}_r)$$

Estimation problem

The aim of the problem is:

- given n observations $\vec{\boldsymbol{y}}_1, \cdots, \vec{\boldsymbol{y}}_n$ each of dimension N with

$$\vec{\boldsymbol{y}} = \vec{\boldsymbol{y}}(\boldsymbol{\varphi}_1, \cdots, \boldsymbol{\varphi}_r)$$

- to estimate $\boldsymbol{r}$ scalar parameters $\boldsymbol{\varphi}_1, \cdots, \boldsymbol{\varphi}_r$

Otherwise stated, the goal is to produce the following estimators:

$$[\vec{\boldsymbol{y}}_1, \cdots, \vec{\boldsymbol{y}}_n] \longrightarrow \boxed{\text{estimation}} \longrightarrow (\hat{\boldsymbol{\varphi}}_1, \cdots, \hat{\boldsymbol{\varphi}}_r)$$

Estimation problem

The aim of the problem is:

- given n observations $\vec{\boldsymbol{y}}_1, \cdots, \vec{\boldsymbol{y}}_n$ each of dimension N with

$$\vec{\boldsymbol{y}} = \vec{\boldsymbol{y}}(\boldsymbol{\varphi}_1, \cdots, \boldsymbol{\varphi}_r)$$

- to estimate $\boldsymbol{r}$ scalar parameters $\boldsymbol{\varphi}_1, \cdots, \boldsymbol{\varphi}_r$

Otherwise stated, the goal is to produce the following estimators:

$$[\vec{\boldsymbol{y}}_1, \cdots, \vec{\boldsymbol{y}}_n] \longrightarrow \boxed{\text{estimation}} \longrightarrow (\hat{\boldsymbol{\varphi}}_1, \cdots, \hat{\boldsymbol{\varphi}}_r)$$

Regime of interest

- N, n of the same order and **large**. Formally: $N, n \to \infty$ and $\frac{N}{n} \to c \in (0, \infty)$
- r finite

Source localization

Problem

$\boldsymbol{r}$ radio sources send their signal to a uniform array of N antennas during n signal snapshots.

Problem: estimate arrival angles $\boldsymbol{\varphi}_1, \cdots, \boldsymbol{\varphi}_r$

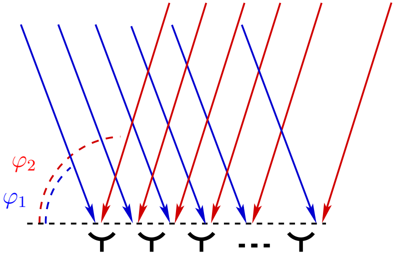

Figure : Two sources $\boldsymbol{\varphi}_1$ and $\boldsymbol{\varphi}_2$ to be estimated

Signal model

The generic observation writes

$$\boxed{\vec{\boldsymbol{y}} = \sum_{\ell=1}^{\boldsymbol{r}} \vec{\boldsymbol{a}}(\boldsymbol{\varphi}_\ell) s_\ell + \sigma \vec{\boldsymbol{w}}} \quad \text{with} \quad \vec{\boldsymbol{a}}(\boldsymbol{\varphi}) = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 \\ e^{\boldsymbol{i}\boldsymbol{\varphi}} \\ \vdots \\ e^{\boldsymbol{i}(N-1)\boldsymbol{\varphi}} \end{pmatrix} \quad \text{and } \vec{\boldsymbol{w}} \sim \mathcal{C}N(0, \mathbf{I}_N)\ .$$

where

- s_ℓ is the scalar source signal associated to DoA $\boldsymbol{\varphi}_\ell$
- $\vec{\boldsymbol{w}}$ is the white noise with variance σ^2

In matrix form

$$\boxed{\mathbf{Y}_N = \mathbf{A}_N(\vec{\boldsymbol{\varphi}})\mathbf{S}_N + \sigma \mathbf{W}_N}$$

with

- $\mathbf{A}_N(\vec{\boldsymbol{\varphi}}) = [\vec{\boldsymbol{a}}(\boldsymbol{\varphi}_1), \cdots, \vec{\boldsymbol{a}}(\boldsymbol{\varphi}_{\boldsymbol{r}})]$ **deterministic** of size $N \times \boldsymbol{r}$
- $\mathbf{W}_N$ **random** with i.i.d. entries of size $N \times n$
- $\mathbf{S}_N$ of size $\boldsymbol{r} \times n$ either deterministic or random

Signal model

The generic observation writes

$$\boxed{\vec{\boldsymbol{y}} = \sum_{\ell=1}^{\boldsymbol{r}} \vec{\boldsymbol{a}}(\boldsymbol{\varphi}_\ell) s_\ell + \sigma \vec{\boldsymbol{w}}} \quad \text{with} \quad \vec{\boldsymbol{a}}(\boldsymbol{\varphi}) = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 \\ e^{\boldsymbol{i}\boldsymbol{\varphi}} \\ \vdots \\ e^{\boldsymbol{i}(N-1)\boldsymbol{\varphi}} \end{pmatrix} \quad \text{and } \vec{\boldsymbol{w}} \sim \mathcal{C}N(0, \mathbf{I}_N) \; .$$

where

- s_ℓ is the scalar source signal associated to DoA $\boldsymbol{\varphi}_\ell$
- $\vec{\boldsymbol{w}}$ is the white noise with variance σ^2

In matrix form

$$\boxed{\mathbf{Y}_N = \mathbf{A}_N(\vec{\boldsymbol{\varphi}})\mathbf{S}_N + \sigma \mathbf{W}_N}$$

with

- $\mathbf{A}_N(\vec{\boldsymbol{\varphi}}) = [\vec{\boldsymbol{a}}(\boldsymbol{\varphi}_1), \cdots, \vec{\boldsymbol{a}}(\boldsymbol{\varphi}_{\boldsymbol{r}})]$ **deterministic** of size $N \times \boldsymbol{r}$
- $\mathbf{W}_N$ **random** with i.i.d. entries of size $N \times n$
- $\mathbf{S}_N$ of size $\boldsymbol{r} \times n$ either deterministic or random

In a nutshell

$\mathbf{Y}_N$ is a (multiplicative) spiked model with a perturbation of rank $\boldsymbol{r}$.

Source localization with subspace estimation

Method known as **MUSIC** for {**MU**}ltiple {**SI**}gnal {**C**}lassification (Schmidt '86)

Source localization with subspace estimation

Method known as **MUSIC** for {**MU**}ltiple {**SI**}gnal {**C**}lassification (Schmidt '86)

Subspace estimation

- The estimation of the angles $\varphi_1, \cdots, \varphi_r$ relies on the estimation of the orthogonal projection $\mathbf{\Pi}_N$ of the **eigenspace of the** r **largest eigenvalues** of

$$\frac{1}{n}\mathbb{E}\mathbf{Y}_n\mathbf{Y}_n^*$$

Source localization with subspace estimation

Method known as **MUSIC** for {**MU**}ltiple {**SI**}gnal {**C**}lassification (Schmidt '86)

Subspace estimation

- The estimation of the angles $\varphi_1, \cdots, \varphi_r$ relies on the estimation of the orthogonal projection $\mathbf{\Pi}_N$ of the **eigenspace of the** r **largest eigenvalues** of

$$\frac{1}{n}\mathbb{E}\mathbf{Y}_n\mathbf{Y}_n^*$$

Small data, large samples: standard estimator

Consider $\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*$, the **empirical counterpart** of $\frac{1}{n}\mathbb{E}\mathbf{Y}_N\mathbf{Y}_N^*$ and its $\boldsymbol{r}$ eigenvectors

$$(\vec{\boldsymbol{u}}_{\boldsymbol{i}}, \cdots, \vec{\boldsymbol{u}}_{\boldsymbol{r}})$$

associated to its $\boldsymbol{r}$ largest (empirical) eigenvalues.

Source localization with subspace estimation

Method known as **MUSIC** for {**MU**}ltiple {**SI**}gnal {**C**}lassification (Schmidt '86)

Subspace estimation

- The estimation of the angles $\varphi_1, \cdots, \varphi_r$ relies on the estimation of the orthogonal projection $\mathbf{\Pi}_N$ of the **eigenspace of the** r **largest eigenvalues** of

$$\frac{1}{n}\mathbb{E}\mathbf{Y}_n\mathbf{Y}_n^*$$

Small data, large samples: standard estimator

Consider $\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*$, the **empirical counterpart** of $\frac{1}{n}\mathbb{E}\mathbf{Y}_N\mathbf{Y}_N^*$ and its $\boldsymbol{r}$ eigenvectors

$$(\vec{\boldsymbol{u}}_{\boldsymbol{i}}, \cdots, \vec{\boldsymbol{u}}_{\boldsymbol{r}})$$

associated to its $\boldsymbol{r}$ largest (empirical) eigenvalues.

- Then the orthogonal projector associated to the r largest eigenvalues of $\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*$ is

$$\widehat{\mathbf{\Pi}}_N = \sum_{\ell=1}^{\boldsymbol{r}} \vec{\boldsymbol{u}}_\ell \vec{\boldsymbol{u}}_\ell^*$$

The large dimension

If N, n of the same order

$$\boxed{\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^* \text{ no longer a good estimator of } \frac{1}{n}\mathbb{E}\mathbf{Y}_N\mathbf{Y}_N^*.}$$

The large dimension

If N, n of the same order

$\frac{1}{n}\mathbf{Y}_N\mathbf{Y}_N^*$ no longer a good estimator of $\frac{1}{n}\mathbb{E}\mathbf{Y}_N\mathbf{Y}_N^*$.

Large data, large sample

- The consistent estimator or $\mathbf{\Pi}_N$ is given by

$$\hat{\mathbf{\Pi}}_N = \sum_{k=1}^{r} \left(1 + \frac{c}{\hat{\theta}_k}\right)\left(1 - \frac{c}{\hat{\theta}_k^2}\right)^{-1} \vec{\mathbf{u}}_k \vec{\mathbf{u}}_k^*$$

  where the $\hat{\theta}_k$'s are the estimated perturbations associated to the kth largest eigenvalue.
- notice the **correction terms** with respect to the standard estimator.

Simulation results I (courtesy from Romain Couillet)

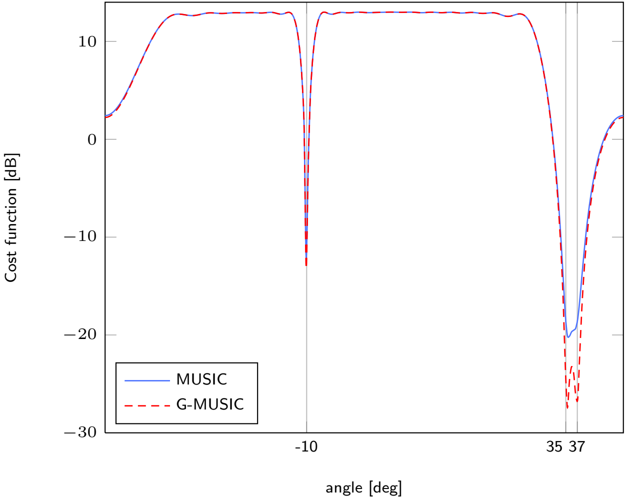

Figure : MUSIC against G-MUSIC for DoA detection of $K = 3$ signal sources, $N = 20$ sensors, $M = 150$ samples, SNR of 10 dB. Angles of arrival of 10°, 35°, and 37°.

Simulation results II

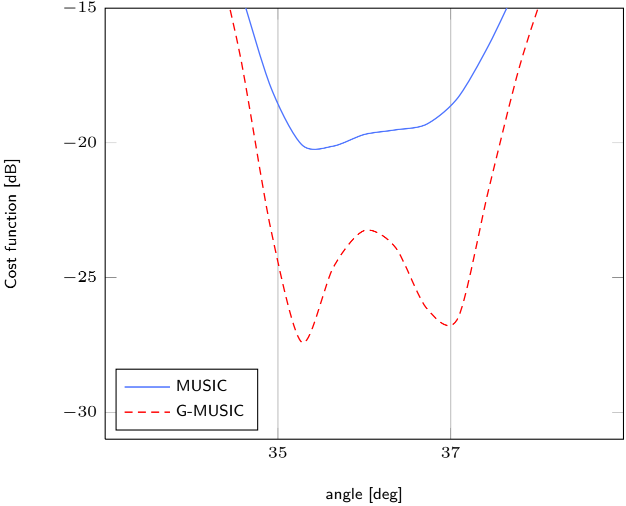

Figure : MUSIC against G-MUSIC for DoA detection of $K = 3$ signal sources, $N = 20$ sensors, $M = 150$ samples, SNR of 10 dB. Angles of arrival of 10°, 35°, and 37°.

Conclusion

Large random matrix theory provides a number of methods which might be of interest for the statistician in particular if one has to handle large data sets.

References

- W. Hachem, P. Loubaton, X. Mestre, J. Najim and P. Vallet.
  " A subspace estimator for fixed rank perturbations of large random matrices",
  Journal of Multivariate Analysis, pp. 427-447, Vol. 114, Feb. 2013 .
- P. Bianchi, M. Debbah, M. Maïda and J. Najim.
  "Performance of Statistical Tests for Single Source Detection using Random Matrix Theory"
  IEEE Trans. Inf. Theory, Vol. 57 (4), april 2011 , 2400–2419.