Applications of Large Random Matrices to Digital Communications and Statistical Signal Processing

W. Hachem, Ph. Loubaton and J. Najim

EUSIPCO 2011
Barcelona

August / September 2011

1. Problem statement
   - Introduction to the Marcenko-Pastur distribution.
   - Some generalizations.
   - Short review of important previous works.
   - Brief overview of applications to digital communications
   - Introduction to the applications to statistical signal processing
2. $K = 0$: An overview of Marčenko and Pastur's results
3. K fixed: spiked models
4. K may scale with M. Application to the subspace method.
5. Some research prospects

Fundamental example.

$$\mathbf{V} = \begin{pmatrix} V_{11} & V_{12} & \dots & V_{1N} \\ V_{21} & V_{22} & \dots & V_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ V_{M1} & V_{M2} & \dots & V_{MN} \end{pmatrix}$$

$(V_{ij})_{1 \leq i \leq M, 1 \leq j \leq N}$ i.i.d. complex Gaussian random variables $\mathcal{CN}(0, \sigma^2)$.

$\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N$ columns of $\mathbf{V}$, $\mathbf{R} = \mathbb{E}(\mathbf{v}_n \mathbf{v}_n^H) = \sigma^2 \mathbf{I}_M$

Empirical covariance matrix:

$$\hat{\mathbf{R}} = \frac{1}{N} \mathbf{V}\mathbf{V}^H = \frac{1}{N} \sum_{n=1}^{N} \mathbf{v}_n \mathbf{v}_n^H$$

Behaviour of the empirical distribution of the eigenvalues of $\hat{\mathbf{R}}$ large M and N.

How behave the histograms of the eigenvalues $(\hat{\lambda}_i)_{i=1,\ldots,M}$ of $\hat{\mathbf{R}}$ when M and N increase.

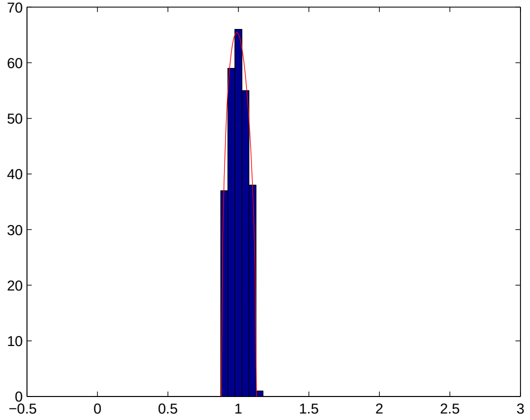
Well known case: M fixed, N increases i.e. $\frac{M}{N}$ small

$\frac{1}{N}\sum_{n=1}^{N}\mathbf{v}_n\mathbf{v}_n^H \simeq \mathbb{E}(\mathbf{v}_n\mathbf{v}_n^H) = \sigma^2\mathbf{I}_M$ by the law of large numbers.

If $N >> M$, the eigenvalues of $\frac{1}{N}\mathbf{V}\mathbf{V}^H$ are concentrated around σ^2.

Illustration.

$M = 256$, $\frac{M}{N} = \frac{1}{256}$, $\sigma^2 = 1$



If M et N are of the same order of magnitude.

$$M, N \to +\infty \text{ such that } \frac{M}{N} = c_N \in [a, b],\ a > 0, b < +\infty.$$

- $\hat{\mathbf{R}}_{i,j} \simeq \sigma^2 \delta_{i-j}$ but
- $\|\hat{\mathbf{R}} - \sigma^2 \mathbf{I}_M\|$ does not converge towards 0.

The histograms of the eigenvalues of $\hat{\mathbf{R}}$ tend to concentrate around the probability density of the so-called Marcenko-Pastur distribution:

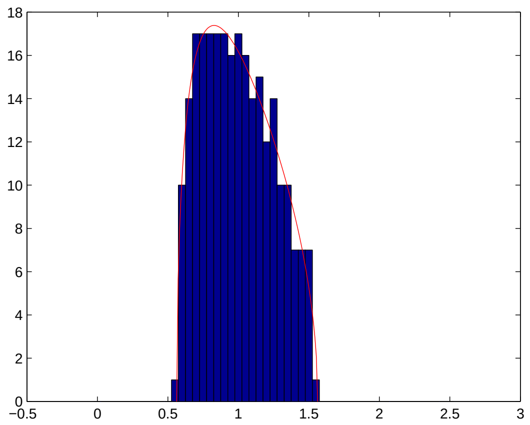
If $c_N \leq 1$,

$$\begin{aligned} p_{c_N}(\lambda) &= \frac{1}{2\pi c_N \lambda} \sqrt{[\sigma^2(1+\sqrt{c_N})^2 - \lambda][\lambda - \sigma^2(1-\sqrt{c_N})^2]} \\ & \quad \text{if } \lambda \in [\sigma^2(1-\sqrt{c_N})^2, \sigma^2(1+\sqrt{c_N})^2] \\ &= 0 \ \text{ otherwise} \end{aligned}$$

Result still true in the non Gaussian case

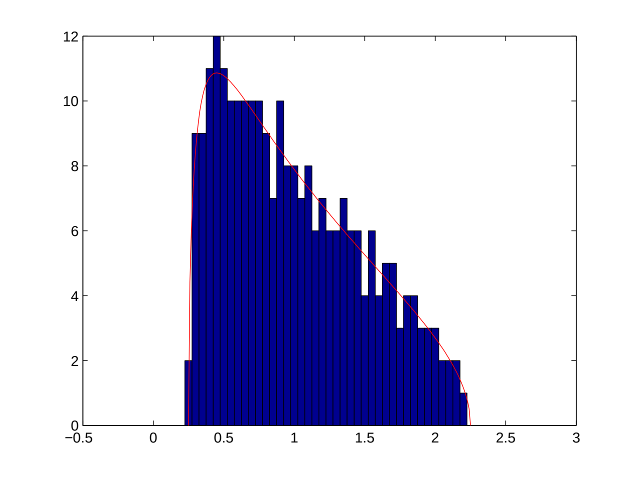
Illustrations I.

$M = 256$, $\frac{M}{N} = \frac{1}{16}$, $\sigma^2 = 1$



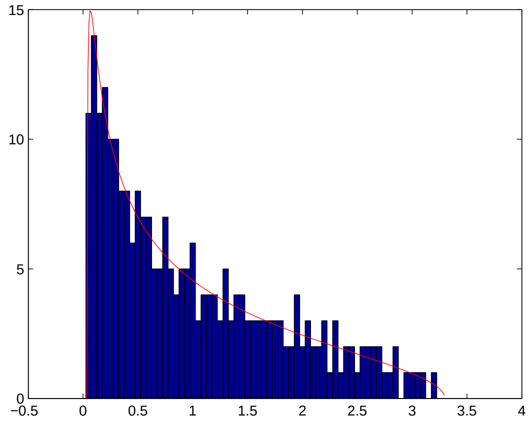
Illustrations II.

$M = 256$, $\frac{M}{N} = \frac{1}{4}$, $\sigma^2 = 1$



Illustrations III.

$M = 256$, $\frac{M}{N} = 2/3$, $\sigma^2 = 1$



Possible to evaluate the asymptotic behaviour of linear statistics

$$\frac{1}{M}\sum_{k=1}^{M} f(\hat{\lambda}_k) = \frac{1}{M}\mathrm{Trace}(f(\hat{\mathbf{R}})) \simeq \int f(\lambda) p_{c_N}(\lambda)\, d\lambda$$

Example 1: $f(\lambda) = \frac{1}{\rho^2+\lambda}$

- $\frac{1}{M}\mathrm{Trace}\left(\hat{\mathbf{R}} + \rho^2 \mathbf{I}\right)^{-1} \simeq \int \frac{p_{c_N}(\lambda)}{\rho^2+\sigma^2} d\lambda = m_N(-\rho^2)$

$m_N(-\rho^2)$ unique positive solution of the equation

$$m_N(-\rho^2) = \frac{1}{\rho^2 + \frac{\sigma^2}{1+\sigma^2 c_N m_N(-\rho^2)}}$$

Closed form solution (see below)

Example 2: $f(\lambda) = \log(1 + \frac{\lambda}{\rho^2})$

- $\frac{1}{M} \log \det \left(\mathbf{I}_M + \frac{\hat{\mathbf{R}}}{\rho^2} \right)$ nearly equal to

$$\frac{1}{c_N} \log \left(1 + \sigma^2 c_N m_N(-\rho^2)\right) + \log \left(1 + \sigma^2 c_N m_N(-\rho^2) + (1 - c_N)\frac{\sigma^2}{\rho^2} \right) \\ -\rho^2 \sigma^2 m_N(-\rho^2) \left(c_N m_N(-\rho^2) + \frac{1 - c_N}{\rho^2} \right)$$

Closed form formula

Fluctuations of the linear statistics.

The bias

$$\mathbb{E}\left[\frac{1}{M}\mathrm{Tr}\left(f(\hat{\mathbf{R}})\right)\right] = \int f(\lambda)\, p_{c_N}(\lambda) d\lambda + \mathcal{O}(\frac{1}{M^2})$$

The variance

$$M\left[\frac{1}{M}\mathrm{Tr}\left(f(\hat{\mathbf{R}})\right) - \int f(\lambda)\, p_{c_N}(\lambda) d\lambda\right] \to \mathcal{N}(0, \mathbf{\Delta}^2)$$

In other words:

$$\frac{1}{M}\mathrm{Tr}\left(f(\hat{\mathbf{R}})\right) - \int f(\lambda)\, p_{c_N}(\lambda) d\lambda \simeq \mathcal{N}(0, \frac{\mathbf{\Delta}^2}{M^2})$$

1. Problem statement
   - Introduction to the Marcenko-Pastur distribution.
   - Some generalizations.
   - Short review of important previous works.
   - Brief overview of applications to digital communications
   - Introduction to the applications to statistical signal processing

Generalizations of these behaviours, $\mathbf{W} = \frac{\mathbf{V}}{\sqrt{N}}$.

- $\mathbf{Y} = \mathbf{C}^{1/2}\mathbf{W}$, $\mathbf{C} \geq 0$ deterministic, zero mean correlated model.
- $\mathbf{Y} = \mathbf{C}^{1/2}\mathbf{W}\tilde{\mathbf{C}}^{1/2}$, $\mathbf{C} \geq 0, \tilde{\mathbf{C}} \geq 0$ deterministic, zero mean bi-correlated model also known as Kronecker model in the MIMO context.
- $\mathbf{Y} = \mathbf{A} + \mathbf{W}$, $\mathbf{A}$ deterministic, information plus noise model.
- $\mathbf{Y} = \mathbf{A} + \mathbf{C}^{1/2}\mathbf{W}\tilde{\mathbf{C}}^{1/2}$, Rician bi-correlated MIMO channel.
- $\mathbf{Y} = \mathbf{U}\left(\boldsymbol{\Delta} \odot \mathbf{W}\right)\mathbf{Q}^H$, $\mathbf{U}, \mathbf{Q}$ unitary deterministic matrices, $\boldsymbol{\Delta}$ deterministic, Sayeed model.
- $\mathbf{Y} = \mathbf{A} + \mathbf{U}\left(\boldsymbol{\Delta} \odot \mathbf{W}\right)\mathbf{Q}^H$, non zero mean Sayeed model.
- Replace i.i.d. matrix $\mathbf{W}$ by an isometric random Haar distributed matrix (obtained from a Gram-Schmidt orthogonalization of $\mathbf{W}$ when $c_N > 1$).

1 Problem statement

- Introduction to the Marcenko-Pastur distribution.
- Some generalizations.
- Short review of important previous works.
- Brief overview of applications to digital communications
- Introduction to the applications to statistical signal processing

Some important contributors.

In statistical physics.

Wigner (1950), Dyson, Mehta, Brézin, ...

In probability theory

- Marcenko, Pastur and colleagues from 1967, Girko from 1975, Bai, Silverstein from 1985.
- Voiculescu and the discovery of the free probability theory from 1993.
- From 1995, a large community using various techniques.

In our field

Digital communications: from 1997

- Seminal works of Tse and colleagues and Verdú and colleagues in 1997 on performance analysis of large CDMA systems
- Performance analysis of large MIMO systems
- Various applications to ressource allocation

Statistics and statistical signal processing

- Before 2007, some works of Girko who was the first to address parameter estimation problems in the context of large random matrices.
- El-Karoui (2008) followed by a number of other researchers addressed the population estimation: estimate the entries of diagonal matrix $\mathbf{P}$ from matrix $\frac{1}{N}\mathbf{V}\mathbf{P}\mathbf{V}^H$.
- Seminal works of Mestre-Lagunas (2008) and Mestre (2008) on the behaviour of the subspace method when the number of sensors and the number of snapshots converge toward ∞ at the same rate.
- More recent works on applications to source number estimation (Nadler 2010), to source detection (Bianchi et al. 2011), to power distribution estimation problems in the context of multiusers communication systems (Couillet et al. 2011).

1 Problem statement

- Introduction to the Marcenko-Pastur distribution.
- Some generalizations.
- Short review of important previous works.
- Brief overview of applications to digital communications
- Introduction to the applications to statistical signal processing

Performance analysis of large CDMA systems .

The simplest context: Tse and Hanly, Verdú and Shamai 1999

- M spreading factor, K number of users
- received M–dimensional vector $\mathbf{y} = h\mathbf{W}\mathbf{s} + \mathbf{n}$
- $\mathbf{s}$ K–dimensional vector of the transmitted symbols
- $\mathbf{n}$ additive white noise, $\mathbb{E}(\mathbf{n}\mathbf{n}^H) = \rho^2 \mathbf{I}_M$
- $\mathbf{W}$ $M \times K$ matrix of the codes allocated to the users, modelled as a realization of a zero mean i.i.d. matrix such that $\mathbb{E}|\mathbf{W}_{i,j}|^2 = \frac{1}{M}$
- h amplitude of the received signal

Performance of the MMSE receiver.

MMSE Estimation of s_1, $\mathbf{W} = (\mathbf{w}_1, \mathbf{W}_2)$

SINR $\beta_M = \mathbf{w}_1^H \left(\mathbf{W}_2 \mathbf{W}_2^H + \frac{\rho^2}{|h|^2} \right)^{-1} \mathbf{w}_1$

Analysis of β_M when $M, K \to \infty$, in such a way that $\frac{M}{K} \in [a, b]$

- $\beta_M \simeq \overline{\beta}_M = \frac{1}{M} \mathrm{Tr} \left(\mathbf{W}_2 \mathbf{W}_2^H + \frac{\rho^2}{|h|^2} \right)^{-1}$
- $\overline{\beta}_M \simeq \beta_{M,*}$ deterministic positive solution of the equation

$$\beta_{M,*} = \frac{1}{\frac{\rho^2}{|h|^2} + \frac{K-1}{M} \frac{1}{1+\beta_{M,*}}}$$

- Allows to have a better understanding of the MMSE receiver: find the loading factor for which $\beta_{M,*}$ is above a target SINR, find the loading factor maximizing the throughput $\frac{K}{M} \log(1 + \beta_{M,*})$, ...

Examples of extensions to more realistic models.

Downlink with frequency selective channel (Debbah *et.al.* 2003)

- $\mathbf{y} = \mathbf{HWs} + \mathbf{n}$, $\mathbf{H}$ Toeplitz matrix
- $$\beta_{M,*} = \frac{1}{M} \sum_{m=0}^{M-1} \frac{1}{\frac{\rho^2}{|h(e^{2i\pi m/M})|^2} + \frac{K-1}{M} \frac{1}{1+\beta_{M,*}}}$$

Downlink with frequency selective channel and random orthogonal Haar distributed code matrix (Debbah *et.al.* 2003)

- $$\beta_{M,*} = \frac{1}{M} \sum_{m=0}^{M-1} \frac{1}{\frac{\rho^2}{|h(e^{2i\pi m/M})|^2} \left(1 - \frac{K-1}{M} \frac{\beta_{M,*}}{1+\beta_{M,*}}\right) + \frac{K-1}{M} \frac{1}{1+\beta_{M,*}}}$$

Uplink with frequency selective channel (Li *et.al.* 2004)

- $\mathbf{y} = \sum_{k=1}^{K} \mathbf{H}_k \mathbf{w}_k s_k + \mathbf{n}$; the channel matrices $\mathbf{H}_k$ are Toeplitz.

Applications to optimal precoding of MIMO systems.

M receive antennas, N transmit antennas

$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$

- $\mathbf{H}$ MIMO channel, $M \times N$ non observable Gaussian random matrix with known (or well estimated) second order statistics
- $\mathbf{x}$ transmitted vector
- $\mathbf{n}$ additive white Gaussian noise, $\mathbb{E}(\mathbf{n}\mathbf{n}^H) = \rho^2 \mathbf{I}_M$

The optimum precoding problem

Find the covariance matrix $\mathbf{Q}$ of $\mathbf{x}$ so as to maximize some figure of merit of the system

Typical example: $I(\mathbf{Q}) = \mathbb{E}\left[\log\det\left(\mathbf{I}_M + \frac{\mathbf{H}\mathbf{Q}\mathbf{H}^H}{\rho^2}\right)\right]$

To be maximized w.r.t. $\mathbf{Q}$ on the convex domain $\mathbf{Q} \geq 0$ and $\frac{1}{M}\mathrm{Tr}(\mathbf{Q}) \leq 1$.

$\mathbf{Q} \rightarrow I(\mathbf{Q})$ is a concave function, but is in general difficult to evaluate in closed form its gradient and hessian. Have to be evaluated using Monte-Carlo simulations (Vu-Paulraj 2005).

A possible alternative: maximize a large system approximation of $I(\mathbf{Q})$

Example of bicorrelated Rician channels $\mathbf{H} = \mathbf{A} + \mathbf{C}^{1/2}\mathbf{W}\tilde{\mathbf{C}}^{1/2}$ (Dumont *et.al.* 2010)

- Eigenvectors of the optimum matrix $\mathbf{Q}_*$ have no closed form expression
- $\overline{I}(\mathbf{Q}) = I(\mathbf{Q}) + \mathcal{O}(\frac{1}{M})$, and $I(\mathbf{Q}_*) = I(\overline{\mathbf{Q}}_*) + \mathcal{O}(\frac{1}{M})$ where
- $\overline{I}(\mathbf{Q}) = \log\det\left(\mathbf{I}_M + \mathbf{Q} \times \mathbf{G}\left(\delta(\mathbf{Q}), \tilde{\delta}(\mathbf{Q})\right)\right) + j\left(\delta(\mathbf{Q}), \tilde{\delta}(\mathbf{Q})\right)$
  where $\left(\delta(\mathbf{Q}), \tilde{\delta}(\mathbf{Q})\right)$ are the unique solutions of a system of 2 non linear equations depending on $\mathbf{Q}, \mathbf{A}, \mathbf{C}, \tilde{\mathbf{C}}$,
  $\mathbf{G}$ is a matrix valued function of $\left(\delta(\mathbf{Q}), \tilde{\delta}(\mathbf{Q})\right)$ given in closed form,
  $j\left(\delta(\mathbf{Q}), \tilde{\delta}(\mathbf{Q})\right)$ is a function of $\left(\delta(\mathbf{Q}), \tilde{\delta}(\mathbf{Q})\right)$ given in closed form.

Maximization of $\overline{I}(\mathbf{Q})$ using an iterative waterfilling algorithm

$$\overline{I}(\mathbf{Q}) = \log\det\left[\mathbf{I}_M + \mathbf{Q} \times \mathbf{G}\left(\delta(\mathbf{Q}), \tilde{\delta}(\mathbf{Q})\right)\right] + j\left(\delta(\mathbf{Q}), \tilde{\delta}(\mathbf{Q})\right).$$

- $\mathbf{Q}^{(k-1)}$ available
- Compute $\left(\delta(\mathbf{Q}^{(k-1)}), \tilde{\delta}(\mathbf{Q}^{(k-1)})\right) = (\delta^{(k-1)}, \tilde{\delta}^{(k-1)})$
- $\mathbf{Q}^{(k)} = \mathrm{Argmax} \log\det\left(\mathbf{I}_M + \mathbf{Q} \times \mathbf{G}(\delta^{(k-1)}, \tilde{\delta}^{(k-1)})\right)$: waterfilling
- k=k+1
- If the algorithm converges, it converges towards $\overline{\mathbf{Q}}_*$

1 Problem statement

- Introduction to the Marcenko-Pastur distribution.
- Some generalizations.
- Short review of important previous works.
- Brief overview of applications to digital communications
- Introduction to the applications to statistical signal processing

The model considered in the following

Observation: M–dimensional time series $\mathbf{y}_n$ observed from $n = 1, \ldots, N$.

- $\mathbf{y}_n = \sum_{k=1}^{K} \mathbf{a}_k s_{k,n} + \mathbf{v}_n = \mathbf{A}\mathbf{s}_n + \mathbf{v}_n$
- $((s_{k,n})_{n \in \mathbb{Z}})_{k=1,K}$ are $K < M$ non observable "source signals", $\mathbf{s}_n = (s_{1,n}, \ldots, s_{K,n})^T$
- $\mathbf{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_K)$ deterministic unknown rank $K < M$ matrix
- $(\mathbf{v}_n)_{n \in \mathbb{Z}}$ additive complex white Gaussian noise such that $\mathbb{E}(\mathbf{v}_n \mathbf{v}_n^H) = \sigma^2 \mathbf{I}_M$

In matrix form

- $\mathbf{Y}_N = (\mathbf{y}_1, \ldots, \mathbf{y}_N)$ observation $M \times N$ matrix
- $\mathbf{Y}_N = \mathbf{A}\mathbf{S}_N + \mathbf{V}_N$
- $\mathbf{\Sigma}_N = \frac{\mathbf{Y}_N}{\sqrt{N}}$, $\mathbf{B}_N = \mathbf{A}\frac{\mathbf{S}_N}{\sqrt{N}}$, $\mathbf{W}_N = \frac{\mathbf{V}_N}{\sqrt{N}}$
- $\mathbf{\Sigma}_N = \mathbf{B}_N + \mathbf{W}_N$

The problems to be addressed.

Detection of the presence of signal(s) from matrix $\mathbf{\Sigma}_N$

- $K = 1$ versus $K = 0$ to simplify
- Various generalizations are possible

Estimation of direction of arrival (DOA) from matrix $\mathbf{\Sigma}_N$.

- $\mathbf{a}_k = \mathbf{a}(\varphi_k)$ where $\varphi \to \mathbf{a}(\varphi)$ is known
- Estimate the parameters $(\varphi_k)_{k=1,\dots,K}$

Problems addressed when M and N are of the same order of magnitude: $M, N \to \infty$ while the ratio $c_N = \frac{M}{N}$ is bounded away from 0 and upper bounded.

In the following

Study of the properties of $\mathbf{\Sigma}_N$ when

- $K = 0$, noise only
- K does not scale with M, i.e. $K \ll M$, **spiked model**: applications to the detection $K = 1$ versus $K = 0$, application to the subspace DOA estimation method
- K may scale with M, i.e. K is not much less than 0, application to the subspace DOA estimation method

The Stieltjes transform I (measure with density)

The **Stieltjes transform** is one of the numerous transforms associated to a measure. It is particularly well-suited to study Large Random Matrices and was introduced in this context by Marčenko and Pastur (1967).

Definition

If the measure μ admits a density f with support $\mathcal{S}$:

$$d\mu(\lambda) = f(\lambda) d\lambda \quad \text{on} \quad \mathcal{S} \ ,$$

then the Stieltjes transform $\Psi_\mu(z)$ is defined as:

$$\begin{aligned} \Psi_\mu(z) &= \int_{\mathcal{S}} \frac{f(\lambda)}{\lambda - z} d\lambda \ , \\ &= -\sum_{k=0}^{\infty} z^{-(k+1)} \left(\int_{\mathcal{S}} \lambda^k f(\lambda)\, d\lambda \right) \end{aligned}$$

The Stieltjes transform I (properties)

Let $\mathrm{im}(z)$ be the imaginary part of $z \in \mathbb{C}$.

Property 1 - identical sign for imaginary part

$$\mathrm{im}\,\Psi_\mu(z) \;=\; \mathrm{im}(z) \int_{\mathcal{S}} \frac{f(\lambda)}{(\lambda - x)^2} d\lambda$$

Property 2 - monotonicity

If $z = x \in \mathbb{R} \setminus \mathcal{S}$, then $\Psi_\mu(x)$ well-defined and:

$$\Psi'_\mu(x) \;=\; \int_{\mathcal{S}} \frac{f(\lambda)}{(\lambda - x)^2} d\lambda \;>\; 0 \qquad \Rightarrow \qquad \Psi_\mu(x) \nearrow \text{ on } \mathbb{R} \setminus \mathcal{S}\ .$$

Property 3 - Inverse formula

$$f(\lambda) = \frac{1}{\pi} \lim_{y \to 0^+} \mathrm{im}\,\Psi_\mu(\lambda + \iota y)\ ,$$

Note that if $\lambda \in \mathbb{R} \setminus \mathcal{S}$, then $\Psi_\mu(x) \in \mathbb{R} \quad \Rightarrow \quad f(\lambda) = 0$.

The Stieltjes transform II (measure with Dirac components)

Stieltjes transform for a Dirac measure

Let δ_x be the Dirac measure at x: $\delta_x(A) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{else.} \end{cases}$ Then

$$\Psi_{\delta_x}(z) = \frac{1}{x - z} \quad \text{in particular,} \quad \Psi_{\delta_0}(z) = -\frac{1}{z} \ .$$

- Important example:

$$L_M = \frac{1}{M} \sum_{k=1}^{M} \delta_{\lambda_k} \quad \Rightarrow \quad \Psi_{L_M}(z) = \frac{1}{M} \sum_{k=1}^{M} \frac{1}{\lambda_k - z} \ .$$

The Stieltjes transform III (link with the resolvent)

Let X be a $M \times M$ Hermitian matrix:

$$\mathbf{X} = \mathbf{U} \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_M \end{pmatrix} \mathbf{U}^*$$

and consider its resolvent $\mathbf{Q}(z)$ and spectral measure L_M:

$$\mathbf{Q}(z) = (\mathbf{X} - z\mathbf{I})^{-1} \ , \quad L_M = \frac{1}{M} \sum_{k=1}^{M} \delta_{\lambda_k} \ .$$

The Stieltjes transform of the spectral measure is the normalized trace of the resolvent:

$$\Psi_{L_M}(z) = \frac{1}{M} \operatorname{tr} \mathbf{Q}(z) \ .$$

2 $K = 0$: An overview of Marčenko and Pastur's results

Gaussian tools

Let the Z_i's be independent complex Gaussian random variables and denote by $\mathbf{z} = (Z_1, \cdots, Z_n)$. The two following results are extremely efficient when dealing with matrices with Gaussian entries (Pastur 2005).

Integration by part Formula

$$\mathbb{E}\left(Z_k \Phi(\mathbf{z}, \overline{\mathbf{z}})\right) = \mathbb{E}|Z_k|^2 \mathbb{E}\left(\frac{\partial \Phi}{\partial \overline{Z}_k}\right)$$

Poincaré-Nash Inequality

$$\mathrm{var}\left(\Phi(\mathbf{z}, \overline{\mathbf{z}})\right) \leq \sum_{k=1}^{n} \mathbb{E}|Z_k|^2 \left(\mathbb{E}\left|\frac{\partial \Phi}{\partial Z_k}\right|^2 + \mathbb{E}\left|\frac{\partial \Phi}{\partial \overline{Z}_k}\right|^2\right)$$

2 $K = 0$: An overview of Marčenko and Pastur's results

- The Stieltjes transform
- Gaussian tools
- Marčenko-Pastur Probability distribution
- A symmetric view of Marčenko-Pastur equation
- Behavior of the individual entries of the resolvent
- Finer convergence results

Marčenko - Pastur Probability distribution

We go back to Marčenko and Pastur framework and consider

$$\mathbf{W}_N = \frac{\mathbf{V}_N}{\sqrt{N}}$$

where $\mathbf{V}_N$ is a $M \times N$ matrix with i.i.d. complex Gaussian random variables $\mathcal{CN}(0, \sigma^2)$.

We are interested in the limiting spectral distribution of $\mathbf{W}_N \mathbf{W}_N^*$. Consider the associated resolvent and Stieltjes transform:

$$\mathbf{Q}(z) = \left(\mathbf{W}_N \mathbf{W}_N^* - z\mathbf{I}\right)^{-1}, \quad \hat{m}_N(z) = \frac{1}{M} \operatorname{tr} \mathbf{Q}(z) .$$

Marčenko - Pastur Probability distribution

We compute hereafter the equation satisfied by the Stieltjes transform associated to the limiting spectral distribution. Afterwards, we rely on the inverse formula for Stietjes transforms to get Marčenko - Pastur distribution.

Main assumption

The ratio $c_N = \frac{M}{N}$ is **bounded away from zero** and **upper bounded** as $M, N \to \infty$.

The three main steps are:

1. To prove that $\mathrm{var}\,(\hat{m}_N(z)) = \mathcal{O}(N^{-2})$. This enables to **replace** $\hat{m}_N(z)$ by $\mathbb{E}\hat{m}_N(z)$ in the computations .
2. To establish the limiting equation satisfied by $\mathbb{E}\hat{m}_N(z)$.
3. To recover the probability distribution with the help of the inverse formula for Stietjes transforms.

Step 1: Marčenko - Pastur Equation

Proposition

$$\operatorname{var}\left(\hat{m}_N(z)\right) = \mathcal{O}\left(\frac{1}{N^2}\right) .$$

Proof:

$$\frac{\partial \mathbf{Q}_{r,r}}{\partial \overline{\mathbf{W}}_{ij}} = -(\mathbf{Q}\mathbf{w}_j)_r \mathbf{Q}_{i,r}$$

By summing over r, then over i and j, we obtain:

$$\begin{aligned} \sum_{i,j} \mathbb{E}\left|\frac{\partial \hat{m}_N(z)}{\partial \overline{\mathbf{W}}_{ij}}\right|^2 &= \mathbb{E}\left(\frac{1}{M^2} \operatorname{tr} \mathbf{Q}^2 \mathbf{W}\mathbf{W}^* \mathbf{Q}^{2*}\right) \\ &\leq \frac{1}{|\operatorname{im}(z)|^4} \mathbb{E}\left(\frac{1}{M^2} \operatorname{tr} \mathbf{W}\mathbf{W}^*\right) = \mathcal{O}\left(\frac{1}{M}\right) . \end{aligned}$$

Step 1: Marčenko - Pastur Equation (end of proof)

By Poincaré-Nash inequality

$$\begin{aligned}
\mathrm{var}\left(\hat{m}_N(z)\right) &\leq \sum_{i,j} \mathbb{E}|\mathbf{W}_{ij}|^2 \left(\mathbb{E}\left|\frac{\partial \hat{m}_N(z)}{\partial \overline{\mathbf{W}}_{ij}}\right|^2 + \mathbb{E}\left|\frac{\partial \hat{m}_N(z)}{\partial \mathbf{W}_{ij}}\right|^2 \right) \\
&= \frac{\sigma^2}{N} \times \left(\mathcal{O}\left(\frac{1}{M}\right) + \mathcal{O}\left(\frac{1}{M}\right)\right) \\
&= \mathcal{O}\left(\frac{1}{M^2}\right)
\end{aligned}$$

which ends the proof. ■

Step 2: Marčenko - Pastur Equation

Proposition

$\mathbb{E}\hat{m}_N(z) - m_N(z) \to 0$ where $m_N(z)$ satisfies:

$$m_N(z) = \frac{-1}{z\left[1 + \sigma^2 c_N m_N(z) - \frac{\sigma^2(1-c_N)}{z}\right]}, \quad c_N = \frac{M}{N} .$$

Proof: The mere definition of the resolvent yields

$$\mathbf{Q} = -\frac{\mathbf{I}}{z} + \frac{\mathbf{QWW}^*}{z} ,$$

hence

$$\mathbb{E}\mathbf{Q}_{r,i} = -\frac{\delta_{ri}}{z} + \frac{\mathbb{E}(\mathbf{QWW}^*)_{r,i}}{z} ,$$

where δ_{ri} stands for the Kronecker symbol.

Step 2: Marčenko - Pastur Equation (proof I)

Write

$$\mathbb{E}(\mathbf{QWW}^*)_{r,i} = \sum_{j=1}^{N}\sum_{s=1}^{M} \mathbb{E}\left(\mathbf{Q}_{r,s}\mathbf{W}_{s,j}\overline{\mathbf{W}}_{ij}\right)$$

Applying the integration by parts formula yields

$$\begin{aligned}
\mathbb{E}\left(\mathbf{Q}_{r,s}\mathbf{W}_{s,j}\overline{\mathbf{W}}_{ij}\right) &= \mathbb{E}\left|\mathbf{W}_{s,j}\right|^2 \mathbb{E}\left[\frac{\partial}{\partial \overline{\mathbf{W}}_{s,j}}\left(\mathbf{Q}_{r,s}\overline{\mathbf{W}}_{i,j}\right)\right] \\
&= \frac{\sigma^2}{N}\left[\delta_{si}\mathbb{E}(\mathbf{Q}_{r,s}) - \mathbb{E}\left(\left(\mathbf{Qw}_j\right)_r \mathbf{Q}_{s,s}\overline{\mathbf{W}}_{i,j}\right)\right]
\end{aligned}$$

Summing over s and then over j yields:

$$\mathbb{E}\left(\mathbf{QWW}^*\right)_{r,i} = \sigma^2 \mathbb{E}\mathbf{Q}_{r,i} - \sigma^2 c_N \mathbb{E}\left[\hat{m}_N \left(\mathbf{QWW}^*\right)_{r,i}\right]$$

Step 2: Marčenko - Pastur Equation (proof II)

Taking $r = i$, summing over i and dividing by M yields:

$$\mathbb{E}\left(\frac{1}{M}\operatorname{tr}\mathbf{QWW}^*\right) = \sigma^2\mathbb{E}\hat{m}_N - \sigma^2 c\mathbb{E}\left[\hat{m}_N\left(\frac{1}{M}\operatorname{tr}\mathbf{QWW}^*\right)\right]$$

As $\mathbf{QWW}^* = \mathbf{I} + z\mathbf{Q}$, we obtain:

$$1 + z\mathbb{E}\hat{m}_N = \sigma^2\mathbb{E}\hat{m}_N - \sigma^2 c\mathbb{E}\left[\hat{m}_N\left(1 + z\hat{m}_N\right)\right]$$

Using Poincaré-Nash inequality enables the following decorrelation:

$$\mathbb{E}\left[\hat{m}_N\left(1 + z\hat{m}_N\right)\right] = \left(\mathbb{E}\hat{m}_N\right)\left(1 + z\mathbb{E}\hat{m}_N\right) + \mathcal{O}\left(\frac{1}{M^2}\right)$$

Step 2: Marčenko - Pastur Equation (proof III)

Gathering the previous results yields:

$$\left(1 + z\mathbb{E}\hat{m}_N\right)\left(1 + \sigma^2 c_N \mathbb{E}\hat{m}_N\right) = \sigma^2 \mathbb{E}\hat{m}_N + \mathcal{O}\left(\frac{1}{M^2}\right)$$

Asymptotically, $\mathbb{E}\hat{m}_N - m_N \to 0$ which satisfies:

$$\left(1 + z m_N\right)\left(1 + \sigma^2 c_N m_N\right) = \sigma^2 m_N \ ,$$

which also writes:

$$m_N(z) = \frac{-1}{z\left[1 + \sigma^2 c_N m_N(z) - \frac{\sigma^2(1-c_N)}{z}\right]} \ .$$

This ends the proof. ■

Step 3: Marčenko - Pastur Probability distribution

Proposition

The probability distribution associated to the Stieltjes transform m_N admits the density p_{c_N} defined as:

$$p_{c_N}(\lambda) = \begin{cases} \frac{\sqrt{(\lambda-\lambda_-)(\lambda_+-\lambda)}}{2\pi\sigma^2 c_N} & \text{if } \lambda \in (\lambda_-, \lambda_+) \\ 0 & \text{else.} \end{cases} .$$

where $\lambda_- = \sigma^2(1-\sqrt{c_N})^2$ and $\lambda_+ = \sigma^2(1+\sqrt{c_N})^2$.

Step 3: Marčenko - Pastur Probability distribution (proof)

Proof: Solving the equation satisfied by m:

$$m_N(z) = -\left(z\left[1+\sigma^2 c_N m_N(z) - \frac{\sigma^2(1-c_N)}{z}\right]\right)^{-1}$$

yields

$$m_N(z) = \frac{-z+\sigma^2(1-c_N)+\sqrt{(z-\lambda_-)(z-\lambda_+)}}{2\sigma^2 c_N z}\ .$$

Using the inverse formula yields:

$$\begin{aligned} p_{c_N}(\lambda) &= \frac{1}{\pi}\lim_{y\to 0^+} \operatorname{im} m_N(\lambda+\iota y) \\ &= \begin{cases} \frac{\sqrt{(\lambda-\lambda_-)(\lambda_+-\lambda)}}{2\pi\sigma^2 c_N} & \text{if } \lambda\in(\lambda_-,\lambda_+) \\ 0 & \text{else.} \end{cases}, \end{aligned}$$

which is the desired result ■

Concluding remarks

- The fact that $\hat{m}_N - m_N \to 0$ implies that for f bounded and continuous,

$$\frac{1}{M}\sum_{i=1}^{M} f(\hat{\lambda}_{i,N}) - \int f(\lambda) p_{c_N}(\lambda) d\lambda \to 0 \ .$$

  p_{c_N} (resp. m_N) is a **deterministic equivalent** of the spectral measure L_N (resp. $\hat{m}_N$).

- if $c_N \to c_* \in (0, \infty)$, then $p_{c_N} \to p_{c_*}$ where p_{c_*} is obtained by replacing c_N by c_* and

$$\frac{1}{M}\sum_{i=1}^{M} f(\hat{\lambda}_{i,N}) \to \int f(\lambda) p_{c_*}(\lambda) d\lambda \ .$$

  in this case, the spectral measure **converges** to p_{c_*}.

2 $K = 0$: An overview of Marčenko and Pastur's results

- The Stieltjes transform
- Gaussian tools
- Marčenko-Pastur Probability distribution
- **A symmetric view of Marčenko-Pastur equation**
- Behavior of the individual entries of the resolvent
- Finer convergence results

A companion quantity in MP equation I

Instead of $\mathbf{WW}^*$, consider $\mathbf{W}^*\mathbf{W}$. Assuming $N \geq M$, both matrices have the same eigenvalues up to $N - M$ zeroes. The **associated Stieltjes transform** therefore writes:

$$\begin{aligned}\hat{\hat{m}}_N(z) &= \frac{1}{N}\operatorname{tr}\left(\mathbf{W}^*\mathbf{W} - z\mathbf{I}\right)^{-1} \\ &= \frac{1}{N}\left(\sum_{k=1}^{N}\frac{1}{\lambda_k - z} - \frac{N-M}{z}\right) = c_N \hat{m}_N(z) - (1-c_N)\frac{1}{z}\end{aligned}$$

A companion quantity in M̌P equation II

As $\hat{m}_N - m_N \to 0$, $\hat{\tilde{m}} - \tilde{m}_N \to 0$ which satisfies:

$$\tilde{m}_N(z) = c_N m_N(z) - (1 - c_N)\frac{1}{z} \ .$$

The inverse Stieltjes transform yields:

$$(ST)^{-1}(m_N) = p_{c_N}(\lambda) \quad \text{and} \quad (ST)^{-1}\left(-\frac{1}{z}\right) = \delta_0$$

Hence, we obtain

$$\tilde{p}_{c_N}(d\lambda) = c_N p_{c_N}(\lambda) d\lambda + (1 - c_N)\delta_0(d\lambda) \ ,$$

where δ_0 accounts for the null eigenvalues of $\mathbf{W}^*\mathbf{W}$.

A symmetric view of M̌P equation

As

$$
\begin{aligned}
m_N(z) &= -\left(z\left[1+\sigma^2 c_N m_N(z) - \frac{\sigma^2(1-c_N)}{z}\right]\right)^{-1} \\
\tilde{m}_N(z) &= c_N m_N(z) - (1-c_N)\frac{1}{z} \; .
\end{aligned}
$$

We readily obtain: $m_N(z) = \frac{-1}{z(1+\sigma^2 \tilde{m}_N(z))}$ Similarly, we can obtain the companion equation: $\tilde{m}_N(z) = \frac{-1}{z(1+\sigma^2 c_N m_N(z))}$. Hence a symmetric presentation of Marčenko-Pastur equation:

$$
\left\{
\begin{aligned}
m_N(z) &= \frac{-1}{z(1+\sigma^2 \tilde{m}_N(z))} \\
\tilde{m}_N(z) &= \frac{-1}{z(1+\sigma^2 c_N m_N(z))}
\end{aligned}
\right. \tag{1}
$$

2 $K = 0$: An overview of Marčenko and Pastur's results

Behavior of the individual entries of the resolvent

Proposition

$$\begin{array}{lrcl} \text{(diagonal)} & \mathbb{E}\,\mathbf{Q}_{i,i}(z) & = & m_N(z) + \mathcal{O}\left(\frac{1}{N^{3/2}}\right), \\ \text{(off-diagonal)} & \mathbb{E}\,\mathbf{Q}_{r,i}(z) & = & \mathcal{O}\left(\frac{1}{N^{3/2}}\right) \quad \text{for} \quad r \neq i\,. \\ \text{(quadratic form)} & \mathbb{E}\,\mathbf{u}^*\mathbf{Q}(z)\mathbf{v} & = & m_N(z)(\mathbf{u}^*\mathbf{v}) + \mathcal{O}\left(\frac{1}{N^{3/2}}\right) \end{array}$$

Behavior of the individual entries of the resolvent (proof I)

Proof: As previously, we have

$$\mathbb{E}\left(\mathbf{QWW}^*\right)_{r,i} = \sigma^2 \mathbb{E}\mathbf{Q}_{r,i} - \sigma^2 c_N \mathbb{E}\left[\hat{m}_N \left(\mathbf{QWW}^*\right)_{r,i}\right]$$

Since $\mathbf{QWW}^* = \mathbf{I} + z\mathbf{Q}$, we obtain: $\left(\mathbf{QWW}^*\right)_{r,i} = \delta_{ri} + z\mathbf{Q}_{r,i}$ Hence:

$$\begin{aligned}
\delta_{ri} + z\mathbb{E}\mathbf{Q}_{r,i} &= \sigma^2 \mathbb{E}\mathbf{Q}_{r,i} - \sigma^2 c_N \mathbb{E}\left[\hat{m}_N \left(\delta_{ri} + z\mathbf{Q}_{r,i}\right)\right] \\
&= \sigma^2 \mathbb{E}\mathbf{Q}_{r,i} - \sigma^2 c_N \delta_{ri} \mathbb{E}\hat{m}_N - z\sigma^2 c_N \mathbb{E}\left[\hat{m}_N \mathbf{Q}_{r,i}\right] \ .
\end{aligned}$$

Poincaré-Nash inequality yields

$$\mathbb{E}\left[\hat{m}_N \mathbf{Q}_{r,i}\right] = \mathbb{E}\hat{m}_N \mathbb{E}\mathbf{Q}_{r,i} + \mathcal{O}\left(\frac{1}{N^{3/2}}\right)$$

(follows from the fact that $\mathrm{var}\, \mathbf{Q}_{r,i} = \mathcal{O}(N^{-1})$).

Behavior of the individual entries of the resolvent (proof II)

- If $r \neq i$ then the result is obvious
- If $r = i$, then

$$1 + z\mathbb{E}\mathbf{Q}_{i,i} = \frac{\sigma^2 \mathbb{E}\mathbf{Q}_{i,i}}{1 + \sigma^2 c_N \mathbb{E}\hat{m}_N} + \mathcal{O}\left(\frac{1}{N^{3/2}}\right) .$$

Summing over i and dividing by M yields

$$1 + z\mathbb{E}\hat{m}_N = \frac{\sigma^2 \mathbb{E}\hat{m}_N}{1 + \sigma^2 c_N \mathbb{E}\hat{m}_N} + \mathcal{O}\left(\frac{1}{N^{3/2}}\right) ,$$

hence the required result.

2 $K = 0$: An overview of Marčenko and Pastur's results

- The Stieltjes transform
- Gaussian tools
- Marčenko-Pastur Probability distribution
- A symmetric view of Marčenko-Pastur equation
- Behavior of the individual entries of the resolvent
- Finer convergence results

Convergence of the extreme eigenvalues

Denote by

$$\hat{\lambda}_{1,N} \geq \cdots \geq \hat{\lambda}_{N,N}$$

the **ordered** eigenvalues of $\mathbf{W}\mathbf{W}^*$ and recall that the support of Marčenko-Pastur distribution is $(\sigma^2(1-\sqrt{c_N})^2, \sigma^2(1+\sqrt{c_N})^2)$. Then:

Theorem

If $c_N \to c_*$, then the following convergences hold true:

$$\hat{\lambda}_{1,N} \xrightarrow[N,M\to\infty]{a.s.} \sigma^2(1+\sqrt{c_*})^2$$

$$\hat{\lambda}_{N,N} \xrightarrow[N,M\to\infty]{a.s.} \sigma^2(1-\sqrt{c_*})^2$$

Fluctuations of the extreme eigenvalues I

A Central Limit Theorem holds for the largest eigenvalue of matrix $\mathbf{W}\mathbf{W}^*$ as $N, M \to \infty$. The limiting distribution is known as **Tracy-Widom**'s distribution.

Fluctuations of $\hat{\lambda}_{1,N}$

Let $c_N \to c_*$. When correctly centered and rescaled, $\hat{\lambda}_{1,N}$ converges to a **Tracy-Widom** distribution:

$$\frac{N^{2/3}}{\sigma^2} \times \frac{\hat{\lambda}_{1,N} - \sigma^2(1+\sqrt{c_N})^2}{(1+\sqrt{c_N})\left(\frac{1}{\sqrt{c_N}}+1\right)^{1/3}} \xrightarrow[N,M\to\infty]{\mathcal{L}} F_{TW} \ .$$

The function F_{TW} stands for **Tracy-Widom** c.d.f. and is precisely described in the following slide.

A similar result holds for $\hat{\lambda}_{M,N}$, the smallest eigenvalue of matrix $\mathbf{W}\mathbf{W}^*$.

Fluctuations of the extreme eigenvalues II

Definition of Tracy-Widom's distribution

The c.d.f. F_{TW} is defined as:

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

where q solves the Painlevé II differential equation:

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

1. Problem statement

2. $K = 0$: An overview of Marčenko and Pastur's results

3. K fixed: spiked models
   - Problem Description
   - Main results
   - Some Applications
   - Proofs of main results: outline of the approach

4. K may scale with M. Application to the subspace method.

5. Some research prospects

Signal model

$$\begin{array}{ccccccc} \text{Rcv signal} & & \text{Channel} & \text{Src signal} & & \text{Noise} \\ \begin{bmatrix} \mathbf{y}_1 \cdots \mathbf{y}_N \end{bmatrix} & = & \begin{bmatrix} \mathbf{a}_1 \cdots \mathbf{a}_K \end{bmatrix} & \begin{bmatrix} \mathbf{s}^1 \\ \cdots \\ \mathbf{s}^K \end{bmatrix} & + & \begin{bmatrix} \mathbf{v}_1 \cdots \mathbf{v}_N \end{bmatrix} \\ \mathbf{Y}_N & = & \mathbf{A}_N & \mathbf{S}_N & + & \mathbf{V}_N \\ M \times N & & M \times K & K \times N & & M \times N \end{array}$$

$$\mathbf{\Sigma}_N = N^{-1/2}\mathbf{Y}_N = \mathbf{B}_N + \mathbf{W}_N$$

Recall that noise matrix $\mathbf{W}$ has independent $\mathcal{CN}(0, \sigma^2/N)$.

We assume here that the **number of sources** K **is** $\ll N$.

$\mathbf{\Sigma}_N$ = Matrix with Gaussian iid elements + fixed rank perturbation.

Asymptotic regime: $N \to \infty$, $M/N \to c_*$, and K is fixed.

Multiplicative Spiked Model

Assume $\mathbf{S}_N$ is a random matrix with independent $\mathcal{CN}(0,1)$ elements (Gaussian iid source signals), and $\mathbf{A}_N$ is deterministic. Then

$$\mathbf{\Sigma}_N = \left(\mathbf{A}_N\mathbf{A}_N^* + \sigma^2\mathbf{I}_M\right)^{1/2}\mathbf{X}_N$$

where $\mathbf{X}_N$ is $M \times N$ with independent $\mathcal{CN}(0,1/N)$ elements.

Consider a spectral factorization

$$\mathbf{A}_N\mathbf{A}_N^* = \mathbf{U}_N \begin{bmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda_K & & \\ & & & 0 & \\ & & & & \ddots \end{bmatrix} \mathbf{U}_N^*.$$

Multiplicative Spiked Model

Let $\mathbf{P}_N$ be the $M \times M$ matrix

$$\mathbf{P}_N = \text{diag}\left(\sqrt{\frac{\lambda_1 + \sigma^2}{\sigma^2}}, \ldots, \sqrt{\frac{\lambda_K + \sigma^2}{\sigma^2}}, 1, \ldots, 1\right).$$

Then

$$\mathbf{U}_N^* \mathbf{\Sigma}_N = \sigma \mathbf{P}_N \mathbf{U}_N^* \mathbf{X}_N \stackrel{\mathcal{D}}{=} \mathbf{P}_N \mathbf{W}_N$$

where $\mathbf{W}_N$ is $M \times N$ with independent $\mathcal{CN}(0, \sigma^2/N)$ elements as above.

$\mathbf{P}_N$ is a fixed rank perturbation of Identity.
$\Rightarrow$ **Multiplicative spiked model**:

$$\text{eigenvalues of } \mathbf{\Sigma}_N \mathbf{\Sigma}_N^* \quad \equiv \quad \text{eigenvalues of } \mathbf{P}_N \mathbf{W}_N \mathbf{W}_N^* \mathbf{P}_N^*.$$

Additive Spiked Model

Assume $\mathbf{S}_N$ is a deterministic matrix and $\mathbf{B}_N = N^{-1/2}\mathbf{A}_N\mathbf{S}_N$ is such rank$(\mathbf{B}_N) = K$ (fixed).
We call the model $\mathbf{\Sigma}_N = \mathbf{B}_N + \mathbf{W}_N$ an additive spiked model.

Impact of $\mathbf{B}_N$ on spectrum of $\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*$ in the asymptotic regime ?

Impact of $\mathbf{P}_N$ or $\mathbf{B}_N$?

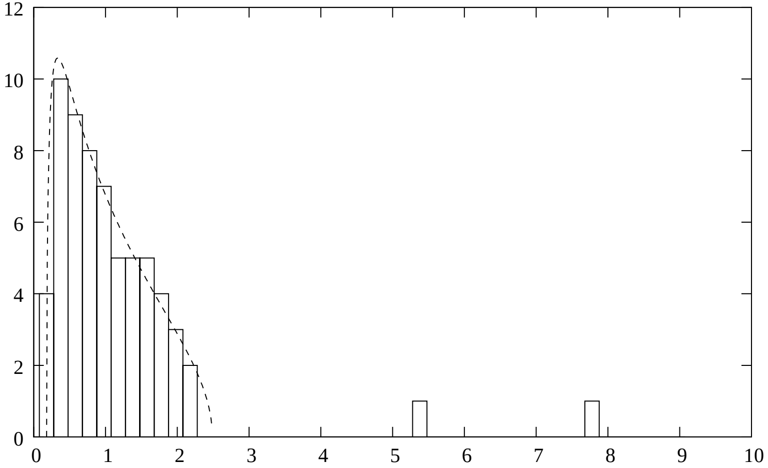
Let $\widetilde{F}_N$ and F_N be the distribution functions associated with the spectral measures of $\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*$ and $\mathbf{W}_N\mathbf{W}_N^*$ respectively. Then

$$\sup_x \left| F_N(x) - \widetilde{F}_N(x) \right| \leq \frac{1}{M} \operatorname{rank}\left(\mathbf{\Sigma}_N\mathbf{\Sigma}_N^* - \mathbf{W}_N\mathbf{W}_N^*\right) \xrightarrow[N\to\infty]{} 0$$

So $\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*$ and $\mathbf{W}_N\mathbf{W}_N^*$ have the **same** (Marčenko Pastur) **limit spectral measure**, either for the multiplicative or the additive spiked model.

However, $\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*$ might have **isolated eigenvalues**.
We shall restrict ourselves to the additive case and study these isolated eigenvalues as well as the projections on their eigenspaces.

Spectrum example for $\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*$



An eigenvalue histogram for $M = 64$, $N = 3M$, and $\sigma^2 = 1$.
$\mathbf{\Sigma}_N = \mathbf{B}_N + \mathbf{W}_N$ where $\mathbf{B}_N$ has rank 2 with singular values 2 and 2.5.

Notations

Spectral factorizations:

$$\mathbf{B}_N \mathbf{B}_N^* = \begin{bmatrix} \mathbf{u}_{1,N} & \cdots & \mathbf{u}_{K,N} \end{bmatrix} \begin{bmatrix} \lambda_{1,N} & & \\ & \ddots & \\ & & \lambda_{K,N} \end{bmatrix} \begin{bmatrix} \mathbf{u}_{1,N} & \cdots & \mathbf{u}_{K,N} \end{bmatrix}^*$$

where $\lambda_{1,N} \geq \cdots \geq \lambda_{K,N}$.

Assuming $N \geq M$

$$\boldsymbol{\Sigma}_N \boldsymbol{\Sigma}_N^* = \begin{bmatrix} \hat{\mathbf{u}}_{1,N} & \cdots & \hat{\mathbf{u}}_{M,N} \end{bmatrix} \begin{bmatrix} \hat{\lambda}_{1,N} & & \\ & \ddots & \\ & & \hat{\lambda}_{M,N} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{u}}_{1,N} & \cdots & \hat{\mathbf{u}}_{M,N} \end{bmatrix}^*$$

where $\hat{\lambda}_{1,N} \geq \cdots \geq \hat{\lambda}_{M,N}$.

Main result on the eigenvalues

Theorem 1

Model is $\mathbf{\Sigma}_N = \mathbf{B}_N + \mathbf{W}_N$ where

- $\mathbf{B}_N$ is a deterministic rank-K matrix such that $\lambda_{k,N} \to \rho_k$ for $k = 1, \ldots, K$,
- $\mathbf{W}_N$ is a $M \times N$ random matrix with independent $\mathcal{CN}(0, \sigma^2/N)$ elements.

Let $i \leq K$ be the maximum index for which $\rho_i > \sigma^2\sqrt{c_*}$. Then for $k = 1, \ldots, i$,

$$\hat{\lambda}_{k,N} \xrightarrow[N\to\infty]{\text{a.s.}} \gamma_k = \frac{\left(\sigma^2 c_* + \rho_k\right)\left(\rho_k + \sigma^2\right)}{\rho_k} > \sigma^2(1+\sqrt{c_*})^2$$

while

$$\hat{\lambda}_{i+1,N} \xrightarrow[N\to\infty]{\text{a.s.}} \sigma^2(1+\sqrt{c_*})^2.$$

Main result on the eigenvectors

Theorem 2

Assume the setting of Theorem 1. Assume in addition that $\rho_1 > \rho_1 > \cdots > \rho_i \;\; (> \sigma^2 \sqrt{c_*})$. For $k = 1, \ldots, i$, let

$$\mathbf{\Pi}_{k,N} = \mathbf{u}_{k,N} \mathbf{u}_{k,N}^* \quad \text{and} \quad \widehat{\mathbf{\Pi}}_{k,N} = \hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^* .$$

Then for any sequence $\mathbf{a}_N$ of deterministic $M \times 1$ vectors such that $\sup_N \|\mathbf{a}_N\| < \infty$,

$$\mathbf{a}_N^* \widehat{\mathbf{\Pi}}_{k,N} \mathbf{a}_N - h(\gamma_k) \mathbf{a}_N^* \mathbf{\Pi}_{k,N} \mathbf{a}_N \xrightarrow[N \to \infty]{\text{a.s.}} 0, \quad h(x) = \frac{x m(x)^2 \tilde{m}(x)}{(x m(x) \tilde{m}(x))'}$$

and m and $\tilde{m}$ are given by Equations (1) when c_N is replaced with c_*.

Generalization to the case of multiple limit eigenvalues ρ_k is possible.

K fixed: spiked models

Passive Signal Detection

- $\mathbf{\Sigma}_N = \mathbf{B}_N + \mathbf{W}_N$, non observable signal + AWGN.
- Assume $K = 1$ source:
  $\mathbf{B}_N = N^{-1/2}\mathbf{a}_{1,N}\mathbf{s}_N^1$, rank one matrix such that $\|\mathbf{B}_N\|^2 \xrightarrow[N\to\infty]{} \rho > 0$.

Hypothesis test: $\begin{cases} \textbf{H0} & : \quad \mathbf{\Sigma}_N = \mathbf{W}_N \quad \text{(Noise)} \\ \textbf{H1} & : \quad \mathbf{\Sigma}_N = \mathbf{B}_N + \mathbf{W}_N \quad \text{(Info+Noise)} \end{cases}$

Generalized Likelihood Ratio Test (GLRT)

$$T_N = \frac{\hat{\lambda}_{1,N}}{M^{-1}\,\mathrm{tr}\,(\mathbf{\Sigma}_N\mathbf{\Sigma}_N^*)}$$

Asymptotic behavior ?

Passive Signal Detection and Additive Spiked Models

- Under either **H0** or **H1**, $M^{-1}\,\mathrm{tr}\left(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*\right) \xrightarrow[N\to\infty]{\text{a.s.}} \sigma^2$.
- Under **H1** (consequence of main result on eigenvalues):
  - If $\rho > \sigma^2\sqrt{c_*}$, then

$$\hat{\lambda}_{1,N} \xrightarrow[N\to\infty]{\text{a.s.}} \gamma = \frac{\left(\sigma^2 c_* + \rho\right)\left(\rho + \sigma^2\right)}{\rho} > \sigma^2(1+\sqrt{c_*})^2,$$
$$\hat{\lambda}_{2,N} \xrightarrow[N\to\infty]{\text{a.s.}} \sigma^2(1+\sqrt{c_*})^2.$$

  - If $\rho \leq \sigma^2\sqrt{c_*}$, then

$$\hat{\lambda}_{1,N} \xrightarrow[N\to\infty]{\text{a.s.}} \sigma^2(1+\sqrt{c_*})^2.$$

Passive Signal Detection and Additive Spiked Models

We therefore have

- Under **H0**,

$$T_N \xrightarrow[N\to\infty]{\text{a.s.}} (1+\sqrt{c_*})^2.$$

- Under **H1**,
  - If $\rho > \sigma^2\sqrt{c_*}$, then

$$T_N \xrightarrow[N\to\infty]{\text{a.s.}} \frac{\left(\sigma^2 c_* + \rho\right)\left(\rho + \sigma^2\right)}{\sigma^2\rho} > (1+\sqrt{c_*})^2$$

  - If $\rho \leq \sigma^2\sqrt{c_*}$, then

$$T_N \xrightarrow[N\to\infty]{\text{a.s.}} (1+\sqrt{c_*})^2.$$

$\rho > \sigma^2\sqrt{c_*}$ provides the **limit of detectability** by the GLRT.

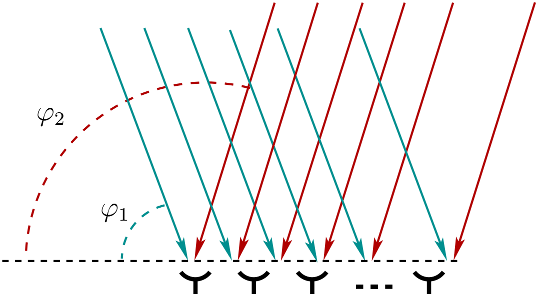
- False Alarm Probability can be evaluated with the help of the Tracy-Widom law.

Source localization

Problem

K radio sources send their signals to a uniform array of M antennas during N signal snapshots.

Estimate arrival angles $\varphi_1, \ldots, \varphi_K$



Example with two sources

Source localization with a subspace method (MUSIC)

Model: $\mathbf{\Sigma}_N = \underbrace{N^{-1/2}\mathbf{A}_N\mathbf{S}_N}_{\mathbf{B}_N} + \mathbf{W}_N$ with

- $\mathbf{A}_N = \left[\mathbf{a}_N(\varphi_1) \quad \cdots \quad \mathbf{a}_N(\varphi_K)\right]$ with $\mathbf{a}_N(\varphi) = \frac{1}{\sqrt{M}}\begin{bmatrix} 1 \\ e^{\imath\pi \sin\varphi} \\ \vdots \\ e^{\imath(M-1)\pi\sin\varphi}\end{bmatrix}$
- $\mathbf{S}_N$ is deterministic, $\mathrm{rank}(\mathbf{S}_N) = K$.

Let $\mathbf{\Pi}_N$ be the orthogonal projection matrix on the span of $\mathbf{A}\mathbf{A}^*$, or equivalently, on the eigenspace of $\mathbb{E}\mathbf{\Sigma}\mathbf{\Sigma}^* = \mathbf{B}\mathbf{B}^* + \sigma^2\mathbf{I}_M$ associated with the eigenvalues $> \sigma^2$ ("signal subspace"). Let $\mathbf{\Pi}_N^{\perp} = \mathbf{I}_M - \mathbf{\Pi}_N$ be the orthogonal projector on the "noise subspace".

MUSIC algorithm principle

$$\mathbf{a}_N(\varphi)^*\mathbf{\Pi}_N^{\perp}\mathbf{a}_N(\varphi) = 0 \quad \Leftrightarrow \quad \varphi \in \{\varphi_1, \ldots, \varphi_K\}.$$

MUSIC algorithm

Traditional MUSIC: angles are estimated as local minima of

$$\mathbf{a}_N(\varphi)^* \widehat{\mathbf{\Pi}}_N^{\perp} \mathbf{a}_N(\varphi)$$

where $\widehat{\mathbf{\Pi}}_N$ is the orthogonal projection matrix on the eigenspace associated with the K largest eigenvalues of $\mathbf{\Sigma\Sigma}^*$ and $\widehat{\mathbf{\Pi}}_N^{\perp} = \mathbf{I}_M - \widehat{\mathbf{\Pi}}_N$.

Asymptotic behavior of $\mathbf{a}_N(\varphi)^* \widehat{\mathbf{\Pi}}_N^{\perp} \mathbf{a}_N(\varphi)$ well known when M is fixed and $N \to \infty$.

- Behavior in our asymptotic regime ?
- Is it possible to **improve** the traditional estimator and to adapt it to our asymptotic regime ?

MUSIC algorithm and the spiked additive model

Modified MUSIC estimator: application of Theorem 2

Assume that $\liminf_N \lambda_{K,N} > \sigma^2 \sqrt{c_*}$. Then

$$\mathbf{a}_N(\varphi)^* \mathbf{\Pi}_N \mathbf{a}_N(\varphi) - \sum_{k=1}^{K} \frac{\left|\mathbf{a}_N(\varphi)^* \hat{\mathbf{u}}_{k,N}\right|^2}{h(\hat{\lambda}_{k,N})} \xrightarrow[N\to\infty]{\text{a.s.}} 0$$

uniformly on $\varphi \in [0, \pi]$.

Modification of the traditional estimator

$$\mathbf{a}(\varphi)^* \mathbf{\Pi}^{\perp} \mathbf{a}(\varphi) = \mathbf{a}(\varphi)^* \left(\sum_{k=1}^{M} \hat{\mathbf{u}}_k \hat{\mathbf{u}}_k^* - \mathbf{\Pi} \right) \mathbf{a}(\varphi)$$

$$\overset{N \text{ large}}{\simeq} \mathbf{a}(\varphi)^* \left(\sum_{k=1}^{K} \left(1 - \frac{1}{h(\hat{\lambda}_k)} \right) \hat{\mathbf{u}}_k \hat{\mathbf{u}}_k^* + \sum_{k=K+1}^{M} \hat{\mathbf{u}}_k \hat{\mathbf{u}}_k^* \right) \mathbf{a}(\varphi)$$

3 K fixed: spiked models

- Problem Description
- Main results
- Some Applications
- Proofs of main results: outline of the approach

Eigenvalues: principle of the proof of Theorem 1

We follow the approach of Benaych-Georges and Nadakuditi'2011.
We study the isolated eigenvalues of $\boldsymbol{\Sigma}\boldsymbol{\Sigma}^*$, or equivalently, the isolated singular values of $\boldsymbol{\Sigma}$.

A matrix algebraic lemma

Let $\mathbf{A}$ be a $M \times N$ matrix. Then $\sigma_1, \ldots, \sigma_{M \wedge N}$ are the singular values of $\mathbf{A}$ if and only if

$$\sigma_1, \ldots, \sigma_{M \wedge N}, -\sigma_1, \ldots, -\sigma_{M \wedge N}, \underbrace{0, \ldots, 0}_{|N-M|}$$

are the eigenvalues of

$$\underline{\mathbf{A}} = \begin{bmatrix} \mathbf{0} & \mathbf{A} \\ \mathbf{A}^* & \mathbf{0} \end{bmatrix}$$

Eigenvalues: principle of the proof of Theorem 1

Drop index N. Let $\mathbf{B} = \mathbf{U}\sqrt{\mathbf{\Lambda}}\mathbf{V}^*$, $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \ldots, \lambda_K)$ be a spectral factorisation of $\mathbf{B}$. Write

$$\underline{\mathbf{\Sigma}} = \begin{bmatrix} 0 & \mathbf{\Sigma} \\ \mathbf{\Sigma}^* & 0 \end{bmatrix} = \begin{bmatrix} 0 & \mathbf{W} \\ \mathbf{W}^* & 0 \end{bmatrix} + \begin{bmatrix} \mathbf{U} & 0 \\ 0 & \mathbf{V}\sqrt{\mathbf{\Lambda}} \end{bmatrix} \begin{bmatrix} 0 & \mathbf{I}_K \\ \mathbf{I}_K & 0 \end{bmatrix} \begin{bmatrix} \mathbf{U}^* & 0 \\ 0 & \sqrt{\mathbf{\Lambda}}\mathbf{V}^* \end{bmatrix}$$
$$= \underline{\mathbf{W}} + \mathbf{CJC}^*.$$

Assume $\hat{\lambda} \notin \text{spectrum}(\mathbf{WW}^*)$ and $\hat{\lambda} \in \text{spectrum}(\mathbf{\Sigma\Sigma}^*)$ or equivalently

$$\det\left(\underline{\mathbf{W}} - \sqrt{\hat{\lambda}}\mathbf{I}_{M+N}\right) \neq 0 \quad \text{and} \quad \det\left(\underline{\mathbf{\Sigma}} - \sqrt{\hat{\lambda}}\mathbf{I}_{M+N}\right) = 0.$$

We have

$$\det\left(\underline{\mathbf{\Sigma}} - x\mathbf{I}\right) = \det\left(\underline{\mathbf{W}} - x\mathbf{I} + \mathbf{CJC}^*\right)$$
$$= \det\left(\underline{\mathbf{W}} - x\mathbf{I}\right) \det\left(\mathbf{I}_{2K} + \mathbf{JC}^*\left(\underline{\mathbf{W}} - x\mathbf{I}\right)^{-1}\mathbf{C}\right)$$

Eigenvalues: principle of the proof of Theorem 1

Using inversion formula for partitioned matrices,

$$(\underline{\mathbf{W}} - x\mathbf{I})^{-1} = \begin{bmatrix} -x\mathbf{I} & \mathbf{W} \\ \mathbf{W}^* & -x\mathbf{I} \end{bmatrix}^{-1} = \begin{bmatrix} x\mathbf{Q}(x^2) & \mathbf{W}\widetilde{\mathbf{Q}}(x^2) \\ \widetilde{\mathbf{Q}}(x^2)\mathbf{W}^* & x\widetilde{\mathbf{Q}}(x^2) \end{bmatrix}$$

where we recall that $\mathbf{Q}(x) = (\mathbf{W}\mathbf{W}^* - x\mathbf{I})^{-1}$, and where we set $\widetilde{\mathbf{Q}}(x) = (\mathbf{W}^*\mathbf{W} - x\mathbf{I})^{-1}$.

Hence $\sqrt{\hat{\lambda}}$ is a zero of

$$\det\left(\mathbf{I}_{2K} + \mathbf{J}\mathbf{C}^*\left(\underline{\mathbf{W}} - x\mathbf{I}\right)^{-1}\mathbf{C}\right)$$
$$= (-1)^K \det \underbrace{\begin{bmatrix} x\mathbf{U}^*\mathbf{Q}(x^2)\mathbf{U} & \mathbf{I}_K + \mathbf{U}^*\mathbf{W}\widetilde{\mathbf{Q}}(x^2)\mathbf{V}\sqrt{\mathbf{\Lambda}} \\ \mathbf{I}_K + \sqrt{\mathbf{\Lambda}}\mathbf{V}^*\widetilde{\mathbf{Q}}(x^2)\mathbf{W}^*\mathbf{U} & x\sqrt{\mathbf{\Lambda}}\mathbf{V}^*\widetilde{\mathbf{Q}}(x^2)\mathbf{V}\sqrt{\mathbf{\Lambda}} \end{bmatrix}}_{\widehat{\mathbf{H}}(x)}$$

Eigenvalues: principle of the proof of Theorem 1

When $x^2 > \sigma^2(1+\sqrt{c_*})^2$, $\mathbf{Q}(x^2)$ and $\widetilde{\mathbf{Q}}(x^2)$ are well defined for N large, because $\|\mathbf{W}\mathbf{W}^*\| \xrightarrow[N\to\infty]{\text{a.s.}} \sigma^2(1+\sqrt{c_*})^2$.

By the approach developed in the previous chapter

$$\mathbf{U}^*\mathbf{Q}(x^2)\mathbf{U} \xrightarrow[N\to\infty]{\text{a.s.}} m(x^2)\mathbf{I}_K, \quad \mathbf{V}^*\widetilde{\mathbf{Q}}(x^2)\mathbf{V} \xrightarrow[N\to\infty]{\text{a.s.}} \tilde{m}(x^2)\mathbf{I}_K, \text{and}$$
$$\mathbf{V}^*\widetilde{\mathbf{Q}}(x^2)\mathbf{W}^*\mathbf{U} \xrightarrow[N\to\infty]{\text{a.s.}} \mathbf{0},$$

hence

$$\widehat{\mathbf{H}}(x) \xrightarrow[n\to\infty]{\text{a.s.}} \mathbf{H}(x) = \begin{bmatrix} xm(x^2)\mathbf{I}_K & \mathbf{I}_K \\ \mathbf{I}_K & x\tilde{m}(x^2)\mathbf{\Gamma} \end{bmatrix} \quad \text{where} \quad \mathbf{\Gamma} = \begin{bmatrix} \rho_1 & & \\ & \ddots & \\ & & \rho_K \end{bmatrix}$$

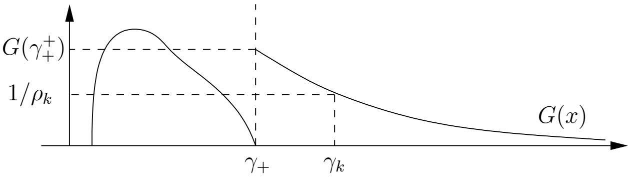
Eigenvalues: principle of the proof of Theorem 1

Consider the equation

$$\det \mathbf{H}(\sqrt{x}) = \prod_{k=1}^{K} (xm(x)\tilde{m}(x)\rho_k - 1) = 0 \ . \qquad (2)$$

- Let $\gamma_+ = \sigma^2(1+\sqrt{c_*})^2$. From the general properties of the Stieltjes Transforms, function $G(x) = xm(x)\tilde{m}(x)$ decreases from $G(\gamma_+^+)$ to zero for $x \in (\gamma_+, \infty)$.
- Recall the ρ_k's are arranged in decreasing order. Assume $\rho_k > 1/G(\gamma_+^+)$. Then the k^{th} largest zero γ_k of (2) (which satisfies $G(\gamma_k) = 1/\rho_k$) will satisfy $\gamma_k > \gamma_+$.
- In that situation, due to $\det \widehat{\mathbf{H}} \to_{\text{as}} \det \mathbf{H}$ outside the eigenvalue bulk, we infer that $\hat{\lambda}_k \to_{\text{as}} \gamma_k$. Otherwise, $\hat{\lambda}_k \to_{\text{as}} \gamma_+$.

Illustration



Exploiting the expressions of $m(z)$ and $\tilde{m}(z)$ (Stieltjes Transforms of M-P distributions), condition $\rho_k > 1/G(\gamma_+^+)$ can be rewritten $\rho_k > \sigma^2\sqrt{c_*}$.
In this case, solving $G(\gamma_k) = 1/\rho_k$ gives $\gamma_k = \left(\sigma^2 c_* + \rho_k\right)\left(\rho_k + \sigma^2\right)/\rho_k$.
Hence Theorem 1.

Eigenvectors: principle of the proof of Theorem 2

Matrix algebraic lemma (*cont'd*)

A pair $(\mathbf{u}, \mathbf{v})$ of unit norm vectors is a pair of (left,right) singular vectors of the $M \times N$ matrix $\mathbf{A}$ associated with the singular value σ if and only if $2^{-1/2} \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}$ is a unit norm eigenvector of

$$\underline{\mathbf{A}} = \begin{bmatrix} \mathbf{0} & \mathbf{A} \\ \mathbf{A}^* & \mathbf{0} \end{bmatrix}$$

associated with the eigenvalue σ.

Eigenvectors: principle of the proof of Theorem 2

Quadratic form $\mathbf{a}^* \widehat{\mathbf{\Pi}}_k \mathbf{a}$ can be written as a Cauchy-integral: using the previous lemma,

$$\mathbf{a}^* \widehat{\mathbf{\Pi}}_k \mathbf{a} = \frac{-1}{\imath\pi} \oint_{\mathcal{C}_k} \begin{bmatrix} \mathbf{a}^* & \mathbf{0} \end{bmatrix} \left(\begin{bmatrix} \mathbf{0} & \mathbf{\Sigma} \\ \mathbf{\Sigma}^* & \mathbf{0} \end{bmatrix} - z\mathbf{I}_{M+N} \right)^{-1} \begin{bmatrix} \mathbf{a} \\ \mathbf{0} \end{bmatrix} \, dz$$

where path $\mathcal{C}_k$ encloses eigenvalue $\sqrt{\hat{\lambda}_k}$.

Recalling that $\begin{bmatrix} \mathbf{0} & \mathbf{\Sigma} \\ \mathbf{\Sigma}^* & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{W} \\ \mathbf{W}^* & \mathbf{0} \end{bmatrix} + \mathbf{CJC}^*$, we obtain using the inversion formula for partitioned matrices

Eigenvectors: principle of the proof of Theorem 2

$$\mathbf{a}^*\widehat{\mathbf{\Pi}}_k\mathbf{a} = \underbrace{\frac{-1}{\imath\pi}\oint_{\mathcal{C}_k}\begin{bmatrix}\mathbf{a}^* & \mathbf{0}\end{bmatrix}\left(\begin{bmatrix}\mathbf{0} & \mathbf{W} \\ \mathbf{W}^* & \mathbf{0}\end{bmatrix} - z\mathbf{I}\right)^{-1}\begin{bmatrix}\mathbf{a} \\ \mathbf{0}\end{bmatrix}\, dz}_{=0 \text{ for large } N}$$

$$+ \frac{1}{\imath\pi}\oint_{\mathcal{C}_k}\hat{\mathbf{b}}^*(z)\widehat{\mathbf{H}}(z)^{-1}\hat{\mathbf{b}}(z)\, dz$$

where

$$\hat{\mathbf{b}}(z) = \begin{bmatrix} z\mathbf{U}^*\mathbf{Q}(z^2) \\ \sqrt{\mathbf{\Lambda}}\mathbf{V}^*\widetilde{\mathbf{Q}}(z^2)\mathbf{W}^*\end{bmatrix}\mathbf{a}\ ,$$

and recall that $\widehat{\mathbf{H}}(z) = \begin{bmatrix} z\mathbf{U}^*\mathbf{Q}(z^2)\mathbf{U} & \mathbf{I}_K + \mathbf{U}^*\mathbf{W}\widetilde{\mathbf{Q}}(z^2)\mathbf{V}\sqrt{\mathbf{\Lambda}} \\ \mathbf{I}_K + \sqrt{\mathbf{\Lambda}}\mathbf{V}^*\widetilde{\mathbf{Q}}(z^2)\mathbf{W}^*\mathbf{U} & z\sqrt{\mathbf{\Lambda}}\mathbf{V}^*\widetilde{\mathbf{Q}}(z^2)\mathbf{V}\sqrt{\mathbf{\Lambda}}\end{bmatrix}.$

Eigenvectors: principle of the proof of Theorem 2

Let

$$\mathbf{b}(z) = \begin{bmatrix} zm(z^2)\mathbf{U}^*\mathbf{a} \\ \mathbf{0} \end{bmatrix} \text{ and recall } \mathbf{H}(z) = \begin{bmatrix} zm(z^2)\mathbf{I}_K & \mathbf{I}_K \\ \mathbf{I}_K & z\tilde{m}(z^2)\mathbf{\Gamma} \end{bmatrix}$$

Since $\hat{\lambda}_k \rightarrow_{\text{a.s.}} \gamma_k = \left(\sigma^2 c_* + \rho_k\right)\left(\rho_k + \sigma^2\right)/\rho_k$, we replace $\mathcal{C}_k$ with a deterministic path C_k centered around γ_k, and

$$\begin{aligned} \mathbf{a}^*\widehat{\mathbf{\Pi}}_k\mathbf{a} &\overset{\text{large}N}{\simeq} \frac{1}{\imath\pi}\oint_{C_k} \mathbf{b}^*(z)\mathbf{H}(z)^{-1}\mathbf{b}(z)\,dz \\ &= \frac{\gamma_k m(\gamma_k)^2\tilde{m}(\gamma_k)}{(\gamma_k m(\gamma_k)\tilde{m}(\gamma_k))'}\mathbf{a}^*\mathbf{\Pi}_k\mathbf{a} \end{aligned}$$

using the residue theorem.

1. Problem statement

2. $K = 0$: An overview of Marčenko and Pastur's results

3. K fixed: spiked models

4. K may scale with M. Application to the subspace method.
   - Motivation.
   - The "asymptotic" limit eigenvalue distribution μ_N
   - Contours enclosing only the eigenvalue 0 of $\mathbf{B}_N \mathbf{B}_N^H$
   - The G-MUSIC algorithm.

5. Some research prospects

4. K may scale with M. Application to the subspace method.
    - Motivation.
    - The "asymptotic" limit eigenvalue distribution μ_N
    - Contours enclosing only the eigenvalue 0 of $\mathbf{B}_N \mathbf{B}_N^H$
    - The G-MUSIC algorithm.

$\mathbf{Y}_N = \mathbf{A}\mathbf{S}_N + \mathbf{V}_N$

- $\mathbf{A}$ $M \times K$ deterministic, the source $K \times N$ matrix $\mathbf{S}_N$ deterministic.
- K and M are possibly of the same order of magnitude: K may scale with N in contrast with the context of spiked models.
- After normalization by $\sqrt{N}$:

$\mathbf{\Sigma}_N = \mathbf{B}_N + \mathbf{W}_N$

- $\mathbf{B}_N = \frac{\mathbf{A}\mathbf{S}_N}{\sqrt{N}}$ deterministic, $\mathrm{Rank}(\mathbf{B}_N) = K = K(N) < M = M(N)$
- $\mathbf{W}_N$ complex Gaussian i.i.d. matrix, $\mathbb{E}|\mathbf{W}_{i,j}|^2 = \frac{\sigma^2}{N}$

$\mathbf{\Sigma}_N = \mathbf{B}_N + \mathbf{W}_N$

- Noise subspace: Orthogonal of the range of $\mathbf{B}_N$ = orthogonal of the range of $\mathbf{A}$ under mild conditions,
- Orthogonal projection matrix $\mathbf{\Pi}_N^{\perp}$
- Estimate consistently $\mathbf{a}^H \mathbf{\Pi}_N^{\perp} \mathbf{a}$ for each unit norm M–dimensional deterministic vector $\mathbf{a}$
- The conventional estimate $\mathbf{a}^H \hat{\mathbf{\Pi}}_N^{\perp} \mathbf{a}$ is not consistent:
- $\mathbf{a}^H \hat{\mathbf{\Pi}}_N^{\perp} \mathbf{a} - \mathbf{a}^H \mathbf{\Pi}_N^{\perp} \mathbf{a}$ does not converge to 0 if M and N converge to ∞ at the same rate

4. K may scale with M. Application to the subspace method.
   - Motivation.
   - The "asymptotic" limit eigenvalue distribution μ_N
   - Contours enclosing only the eigenvalue 0 of $\mathbf{B}_N \mathbf{B}_N^H$
   - The G-MUSIC algorithm.

Characterization of the limit eigenvalue distribution μ_N

Dozier-Silverstein 2007: It exists a deterministic probability measure μ_N carried by $\mathbb{R}^+$ such that

- $\frac{1}{M}\sum_{k=1}^{M} \delta(\lambda - \hat{\lambda}_{k,N}) - \mu_N \to 0$ weakly almost surely

Characterization of the limit eigenvalue distribution μ_N

Dozier-Silverstein 2007: It exists a deterministic probability measure μ_N carried by $\mathbb{R}^+$ such that

- $\frac{1}{M}\sum_{k=1}^{M}\delta(\lambda - \hat{\lambda}_{k,N}) - \mu_N \to 0$ weakly almost surely

How to characterize μ_N

- Stieltjes transform $m_N(z) = \int_{\mathbb{R}^+} \frac{\mu_N(d\lambda)}{\lambda - z}$ defined on $\mathbb{C} - \mathbb{R}^+$
- $m_N(z) := \frac{1}{M}\mathrm{Tr}\,\mathbf{T}_N(z)$ with
- $\mathbf{T}_N(z) = \left(\frac{\mathbf{B_N B}_N^*}{1+\sigma^2 c_N m_N(z)} - z(1+\sigma^2 c_N m_N(z))\mathbf{I}_M + \sigma^2(1-c_N)\mathbf{I}_M\right)^{-1}$.

Equivalent form of the equation

$m_N(z)$ is solution of the equation

$$\frac{m_N(z)}{1+\sigma^2 c_N m_N(z)} = \frac{1}{M}\mathrm{Trace}(\mathbf{B}_N\mathbf{B}_N^* - w_N(z)\mathbf{I}_M)^{-1} = f_N(w_N(z))$$

- $w_N(z) = z(1+\sigma^2 c_N m_N(z))^2 - \sigma^2(1-c_N)(1+\sigma^2 c_N m_N(z))$
- $f_N(w) = \frac{1}{M}\mathrm{Trace}(\mathbf{B}_N\mathbf{B}_N^* - w\mathbf{I}_M)^{-1} = \frac{1}{M}\sum_{k=1}^{M}\frac{1}{\lambda_{k,N}-w}$

Equivalent form of the equation

$m_N(z)$ is solution of the equation

$$\frac{m_N(z)}{1+\sigma^2 c_N m_N(z)} = \frac{1}{M}\mathrm{Trace}(\mathbf{B}_N\mathbf{B}_N^* - w_N(z)\mathbf{I}_M)^{-1} = f_N(w_N(z))$$

- $w_N(z) = z(1+\sigma^2 c_N m_N(z))^2 - \sigma^2(1-c_N)(1+\sigma^2 c_N m_N(z))$
- $f_N(w) = \frac{1}{M}\mathrm{Trace}(\mathbf{B}_N\mathbf{B}_N^* - w\mathbf{I}_M)^{-1} = \frac{1}{M}\sum_{k=1}^{M}\frac{1}{\lambda_{k,N}-w}$

Convergence results: $\mathbf{Q}_N(z) = (\mathbf{\Sigma}_N\mathbf{\Sigma}_N^* - z\mathbf{I}_M)^{-1}$

- $\frac{1}{M}\mathrm{Tr}\mathbf{Q}_N(z) = \hat{m}_N(z) \asymp m_N(z) = \frac{1}{M}\mathrm{Tr}\mathbf{T}_N(z)$

Equivalent form of the equation

$m_N(z)$ is solution of the equation

$$\frac{m_N(z)}{1+\sigma^2 c_N m_N(z)} = \frac{1}{M}\mathrm{Trace}(\mathbf{B}_N\mathbf{B}_N^* - w_N(z)\mathbf{I}_M)^{-1} = f_N(w_N(z))$$

- $w_N(z) = z(1+\sigma^2 c_N m_N(z))^2 - \sigma^2(1-c_N)(1+\sigma^2 c_N m_N(z))$
- $f_N(w) = \frac{1}{M}\mathrm{Trace}(\mathbf{B}_N\mathbf{B}_N^* - w\mathbf{I}_M)^{-1} = \frac{1}{M}\sum_{k=1}^{M}\frac{1}{\lambda_{k,N}-w}$

Convergence results: $\mathbf{Q}_N(z) = (\mathbf{\Sigma}_N\mathbf{\Sigma}_N^* - z\mathbf{I}_M)^{-1}$

- $\frac{1}{M}\mathrm{Tr}\mathbf{Q}_N(z) = \hat{m}_N(z) \asymp m_N(z) = \frac{1}{M}\mathrm{Tr}\mathbf{T}_N(z)$
- Hachem et al.(2010), for $\|\mathbf{d}_N\| = 1$,

$$\mathbf{d}_N^*\mathbf{Q}_N(z)\mathbf{d}_N \asymp \mathbf{d}_N^*\mathbf{T}_N(z)\mathbf{d}_N.$$

Properties of μ_N, $c_N = \frac{M}{N} < 1$

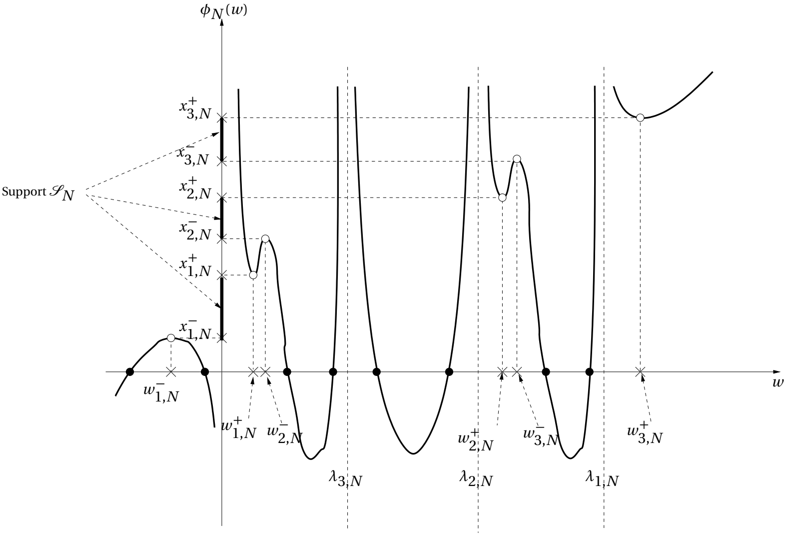
Dozier-Silverstein-2007

- For each $x \in \mathbb{R}$, $\lim_{z \to x, z \in \mathbb{C}^+} m_N(z) = m_N(x)$ exists
- $x \to m_N(x)$ continuous on $\mathbb{R}$, continuously differentiable on $\mathbb{R} \backslash \partial \mathcal{S}_N$
- $\mu_N(d\lambda)$ absolutely continuous, density $\frac{1}{\pi}\mathrm{Im}(m_N(x))$
- $\mathcal{S}_N$ support of μ_N. $\mathrm{Int}(\mathcal{S}_N) = \{x \in \mathbb{R}, \mathrm{Im}(m_N(x)) > 0\}$

Characterization of the support $\mathcal{S}_N$ of μ_N.

Reformulation of Dozier-Silverstein 2007 in Vallet-Loubaton-Mestre-2010

- Function $\phi_N(w)$ defined on $\mathbb{R}$ by
  $\phi_N(w) = w(1 - \sigma^2 c_N f_N(w))^2 + \sigma^2(1 - c_N)(1 - \sigma^2 c_N f_N(w))$
- ϕ_N has $2Q$ positive extrema with preimages $w_{1,-}^{(N)} < w_{1,+}^{(N)} < w_{2,-}^{(N)} < \dots w_{Q,-}^{(N)} < w_{Q,+}^{(N)}$. These extrema verify $x_{1,-}^{(N)} < x_{1,+}^{(N)} < x_{2,-}^{(N)} < \dots x_{Q,-}^{(N)} < x_{Q,+}^{(N)}$.
- $\mathcal{S}_N = [x_{1,-}^{(N)}, x_{1,+}^{(N)}] \cup \dots [x_{Q,-}^{(N)}, x_{Q,+}^{(N)}]$
- Each eigenvalue $\lambda_{l,N}$ of $\mathbf{B}_N \mathbf{B}_N^*$ belongs to an interval $(w_{k,-}^{(N)}, w_{k,+}^{(N)})$


$\phi_N(w)$
$x^+_{3,N}$
$x^-_{3,N}$
$x^+_{2,N}$
$x^-_{2,N}$
$x^+_{1,N}$
$x^-_{1,N}$
Support $\mathscr{S}_N$
$w^-_{1,N}$
$w^+_{1,N}$
$w^-_{2,N}$
$w^+_{2,N}$
$w^-_{3,N}$
$w^+_{3,N}$
$\lambda_{3,N}$
$\lambda_{2,N}$
$\lambda_{1,N}$
w

- If c_N is small enough or σ^2 small enough, there are $Q = K+1$ clusters nearly centered around σ^2 and $(\lambda_k + \sigma^2)_{k=1,\ldots,K}$.
- If c_N or σ^2 increases, certain clusters merge, and $Q < K+1$.
- An eigenvalue $\lambda_{k,N}$ of $\mathbf{B}_N\mathbf{B}_N^*$ is said to be associated to the cluster $[x_{q,N}^-, x_{q,N}^+]$ if $\lambda_{k,N} \in]w_{q,N}^-, w_{q,N}^+[$.

Illustration (I).

The parameters.

- $\sigma^2 = 2$
- Eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$: 0 and 5 with multiplicity $\frac{M}{2}$
- Eigenvalues of $\mathbf{B}_N \mathbf{B}_N^* + \sigma^2 \mathbf{I}$: 2 and 7 with multiplicity $\frac{M}{2}$

Illustration (I).

The parameters.

- $\sigma^2 = 2$
- Eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$: 0 and 5 with multiplicity $\frac{M}{2}$
- Eigenvalues of $\mathbf{B}_N \mathbf{B}_N^* + \sigma^2 \mathbf{I}$: 2 and 7 with multiplicity $\frac{M}{2}$

Remark

- $f_N(w) = \frac{1}{2}\left(-\frac{1}{w} + \frac{1}{5-w}\right)$ independent of M, N
- μ_N does not depend on M, N if $c_N = \frac{M}{N} = c$ independent of M, N

Illustration (II).

$c = \frac{M}{N} = 0.05$

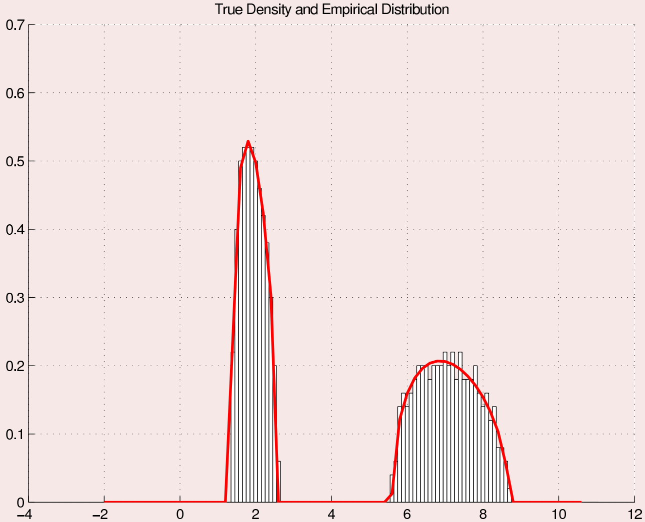


Illustration (III).

$c = \frac{M}{N} = 0.2$

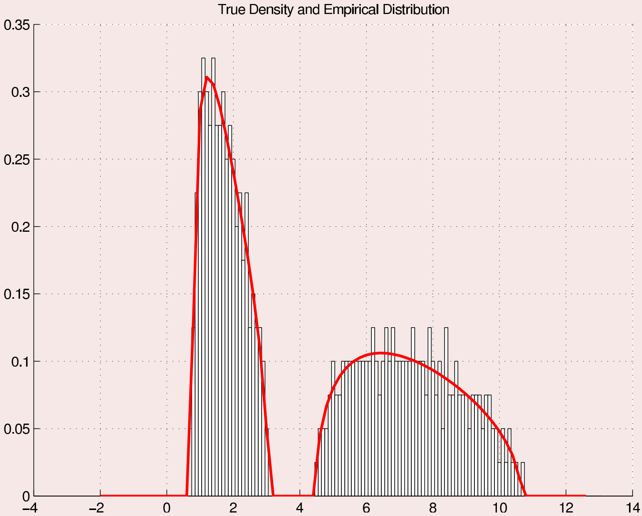
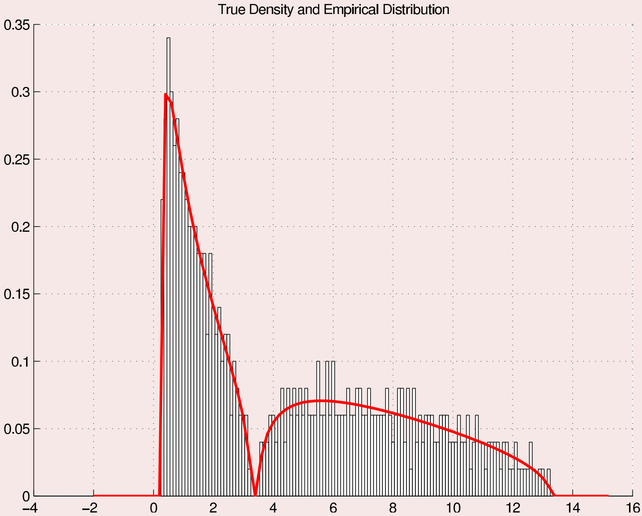


Illustration (IV).

$c = \frac{M}{N} = 0.5$



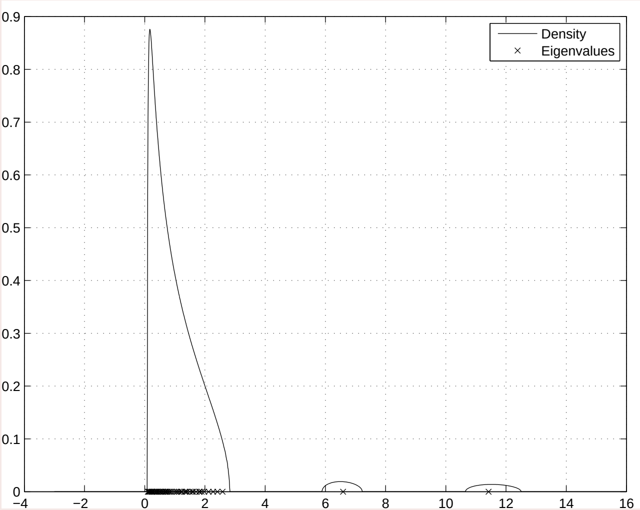
$\mathcal{S}_N$ in the context of spiked models.

Assumptions: $c_N \to c_*$, $\lambda_{k,N} \to \rho_k > \sigma^2\sqrt{c}_*$ for $k = 1, \ldots, K$, $\rho_k \neq \rho_l$.

- $\hat{\lambda}_{k,N} \to \gamma_k = \frac{(\sigma^2 c_* + \rho_k)(\rho_k + \sigma^2)}{\rho_k}$ for $k = 1, \ldots, K$
- $Q = K + 1$ clusters
- $[x^-_{1,N}, x^+_{1,N}] = [\sigma^2(1 - \sqrt{c}_N)^2 - \mathcal{O}(\frac{1}{N}), \sigma^2(1 + \sqrt{c}_N)^2 - \mathcal{O}(\frac{1}{N})]$
- $[x^-_{k,N}, x^+_{k,N}] = [\psi(\lambda_{K+2-k,N}, c_N) - \mathcal{O}(\frac{1}{\sqrt{N}}), \psi(\lambda_{K+2-k,N}, c_N) + \mathcal{O}(\frac{1}{\sqrt{N}})]$ for $k = 2, \ldots, K+1$
- $\psi(\lambda, c) = \frac{(\sigma^2 c + \lambda)(\lambda + \sigma^2)}{\lambda}$ so that $\psi(\lambda_{K+2-k,N}, c_N)$ close from $\psi(\rho_{K+2-k}, c_*) = \gamma_{K+2-k}$.

Illustration

$c = \frac{M}{N} = 0.5, N = 100,\ K = 2,\ \sigma^2 = 1$



4. K may scale with M. Application to the subspace method.
   - Motivation.
   - The "asymptotic" limit eigenvalue distribution μ_N
   - Contours enclosing only the eigenvalue 0 of $\mathbf{B}_N \mathbf{B}_N^H$
   - The G-MUSIC algorithm.

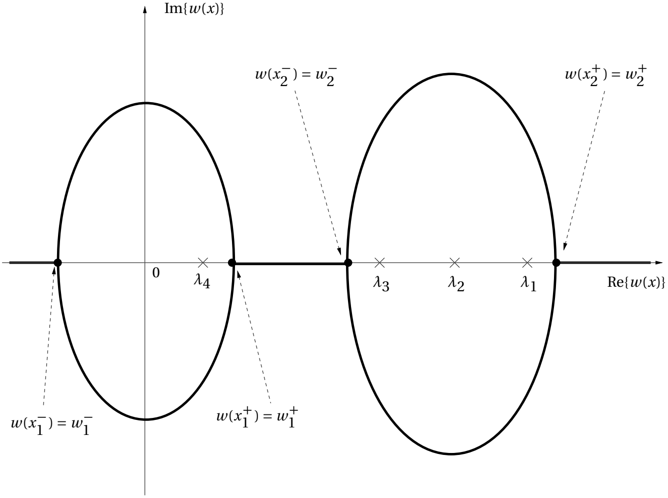
Some useful properties of $w_N(z)$

$$w_N(z) = z(1+\sigma^2 c_N m_N(z))^2 - \sigma^2(1-c_N)(1+\sigma^2 c_N m_N(z)).$$

- $\mathrm{Im}(w_N(z)) > 0$ if $\mathrm{Im}(z) > 0$
- $\mathrm{Int}(\mathcal{S}_N) = \{x, \mathrm{Im}(w_N(x)) > 0\}$
- $w_N(x)$ is real and increasing on each component of $\mathcal{S}_N^c$
- $w_N(x_{q,N}^-) = w_{q,N}^-, w_N(x_{q,N}^+) = w_{q,N}^+$
- $w_N(x)$ is continuous on $\mathbb{R}$ and continuously differentiable on $\mathbb{R} \backslash \partial\mathcal{S}_N$

Illustration of the behaviour of $x \to w_N(x)$

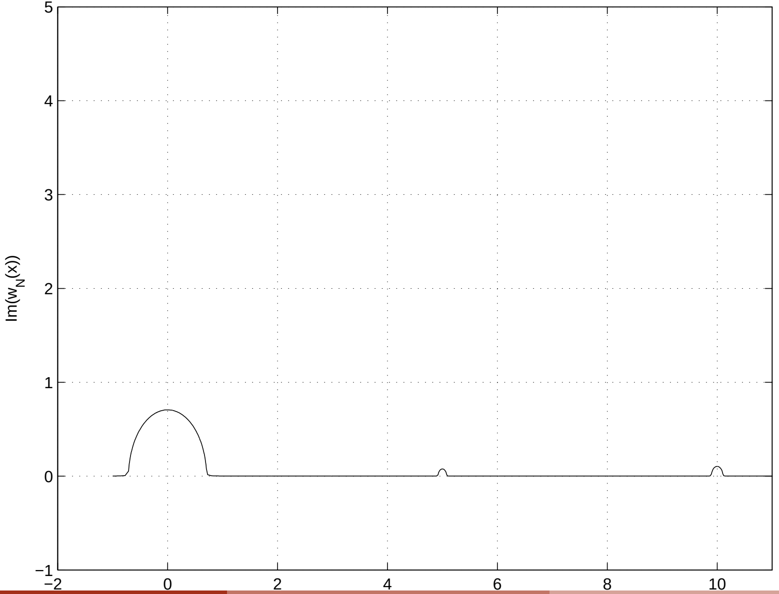
Illustration 2 clusters.



In the case of the MP distribution, $\mathbf{B}_N = 0$

- $w_N(x)$ is real and increasing on $(-\infty, \sigma^2(1 - \sqrt{c_N})^2)$
- $w_N(\sigma^2(1 - \sqrt{c_N})^2) = -\sigma^2\sqrt{c_N}$
- $|w_N(x)| = \sigma^2\sqrt{c_N}$ if $x \in [\sigma^2(1 - \sqrt{c_N})^2, \sigma^2(1 + \sqrt{c_N})^2$
- $w_N(1 + \sqrt{c_N})^2) = \sigma^2\sqrt{c_N}$
- $w_N(x)$ is real and increasing on $(\sigma^2(1 + \sqrt{c_N})^2, +\infty)$

Illustration in the spiked case $K = 2$, $N = 100$, $M = 50$, $\sigma^2 = 1$



4 K may scale with M. Application to the subspace method.

- Motivation.
- The "asymptotic" limit eigenvalue distribution μ_N
- Contours enclosing only the eigenvalue 0 of $\mathbf{B}_N \mathbf{B}_N^H$
- The G-MUSIC algorithm.

Valid under the following hypotheses.

Assumptions.

- 0 is the unique eigenvalue associated with $[x_{1,N}^-, x_{1,N}^+]$ for each N large enough,
- $0 < \liminf_N x_{1,N}^- < \limsup_N x_{1,N}^+ < \liminf_N x_{2,N}^-$

- for all N large enough , t_1^-, t_1^+, t_2^- independent of N

Consequences of the assumptions

- almost surely for N large enough

$$\hat{\lambda}_{K+1,N}, \ldots, \hat{\lambda}_{M,N} \in (t_1^-, t_1^+) \quad \text{and} \quad \hat{\lambda}_{1,N}, \ldots, \hat{\lambda}_{K,N} > t_2^-,$$

Consequences of the assumptions

- almost surely for N large enough

$$\hat{\lambda}_{K+1,N}, \ldots, \hat{\lambda}_{M,N} \in (t_1^-, t_1^+) \quad \text{and} \quad \hat{\lambda}_{1,N}, \ldots, \hat{\lambda}_{K,N} > t_2^-,$$

- almost surely for N large enough,

$$\hat{\omega}_{K+1,N}, \ldots, \hat{\omega}_{M,N} \in (t_1^-, t_1^+) \quad \text{and} \quad \hat{\omega}_{1,N}, \ldots, \hat{\omega}_{K,N} > t_2^-$$

  with $\hat{\omega}_{1,N} \geq \ldots \geq \hat{\omega}_{M,N}$ the solutions of the equation $1 + \sigma^2 c_N \hat{m}_N(z) = 0$ with $\hat{m}_N(z) = \frac{1}{M}\mathrm{Tr}\mathbf{Q}_N(z)$

Consequences of the assumptions

- For $y > 0$, we define the domain

$$\mathcal{R}_y = \left\{ u + iv : u \in [t_1^- - \delta, t_1^+ + \delta], v \in [-y, y] \right\}.$$

Then, if $t_1^+ + \delta < t_2^-$, $\mathcal{C}_y = w_N(\partial \mathcal{R}_y)$ encloses 0 and no other eigenvalue of $\mathbf{B}_N \mathbf{B}_N^*$ for N large enough.

Consistent estimation of $\eta_N = \mathbf{a}_N \mathbf{\Pi}_N^{\perp} \mathbf{a}_N$.

From residues theorem:

$$\eta_N = \frac{1}{2\pi\imath} \oint_{\mathcal{C}_y^-} \mathbf{a}_N^* \left(\mathbf{B}_N \mathbf{B}_N^* - \lambda \mathbf{I}_M \right)^{-1} \mathbf{a}_N \mathrm{d}\lambda,$$

$$\eta_N = \frac{1}{2\pi\imath} \oint_{\partial\mathcal{R}_y^-} \mathbf{a}_N^* \left(\mathbf{B}_N \mathbf{B}_N^* - w_N(z) \mathbf{I}_M \right)^{-1} \mathbf{a}_N w_N'(z) \mathrm{d}z$$

$$\eta_N = \frac{1}{2\pi\imath} \oint_{\partial\mathcal{R}_y^-} \mathbf{a}_N^* \mathbf{T}_N(z) \mathbf{a}_N \frac{w_N'(z)}{1 + \sigma^2 c_N m_N(z)} \mathrm{d}z$$

The integrand can be estimated consistently.

$g_N(z) = \mathbf{a}_N^* \mathbf{T}_N(z)\mathbf{a}_N \frac{w_N'(z)}{1+\sigma^2 c_N m_N(z)}$

- From the previous result, we have the following convergence on $\mathbb{C} - \mathcal{S}_N$

$$m_N(z) \asymp \hat{m}_N(z) = \frac{1}{M}\mathrm{Tr}\mathbf{Q}_N(z) \quad \text{and} \quad \mathbf{a}_N^* \mathbf{T}_N(z)\mathbf{a}_N \asymp \mathbf{a}_N^* \mathbf{Q}_N(z)\mathbf{a}_N$$

  with $\mathbf{Q}_N(z) = (\mathbf{\Sigma}_N \mathbf{\Sigma}_N^* - z\mathbf{I}_M)^{-1}$.

- Let $\hat{g}_N(z) := \mathbf{a}_N^* \mathbf{Q}_N(z)\mathbf{a}_N \frac{\hat{w}_N'(z)}{1+\sigma^2 c_N \hat{m}_N(z)}$ with $\hat{w}_N(z) = z(1+\sigma^2 c_N \hat{m}_N(z))^2 - \sigma^2 c_N (1+\sigma^2 c_N \hat{m}_N(z))$. $\hat{g}_N(z)$ has no pole on $\partial\mathcal{R}_y$ and

$$\left| \frac{1}{2\pi\imath} \oint_{\partial\mathcal{R}_y^-} (g_N(z) - \hat{g}_N(z))\,\mathrm{d}z \right| \to 0 \text{ a.s.},$$

The new consistent estimator.

$$\hat{\eta}_{N,new} = \frac{1}{2\pi \imath} \oint_{\partial \mathcal{R}_y^-} \mathbf{a}_N^* \mathbf{Q}_N(z) \mathbf{a}_N \frac{\hat{w}_N'(z)}{1+\sigma^2 c_N \hat{m}_N(z)} \mathrm{d}z$$

- Integral can be solved using the residue's theorem
- $\hat{\eta}_{N,new} = \mathbf{a}_N^* \left(\sum_{k=1}^{M} \hat{\xi}_{k,N} \hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^* \right) \mathbf{a}_N$ with $(\hat{\xi}_{k,N})$ depending on $\hat{\lambda}_{1,N}, \ldots, \hat{\lambda}_{M,N}$ and $\hat{\omega}_{1,N}, \ldots, \hat{\omega}_{M,N}$.
- $\hat{\eta}_{N,new}$ depend on the $(\hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^*)_{k=K+1,\ldots,M}$ and on the $(\hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^*)_{k=1,\ldots,K}$

Numerical evaluations.

Comparisons between:

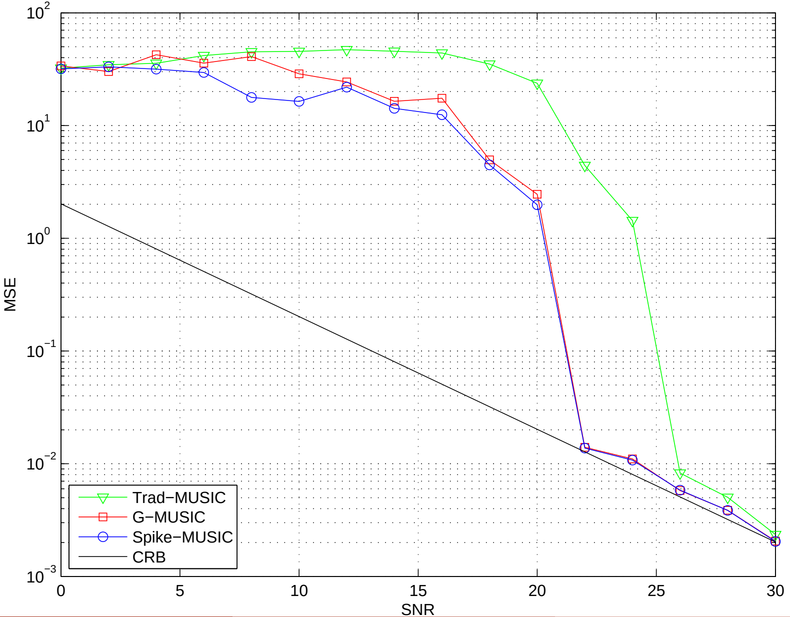
- The traditional subspace method
- The spike subspace method
- The improved subspace method

Experiment 1

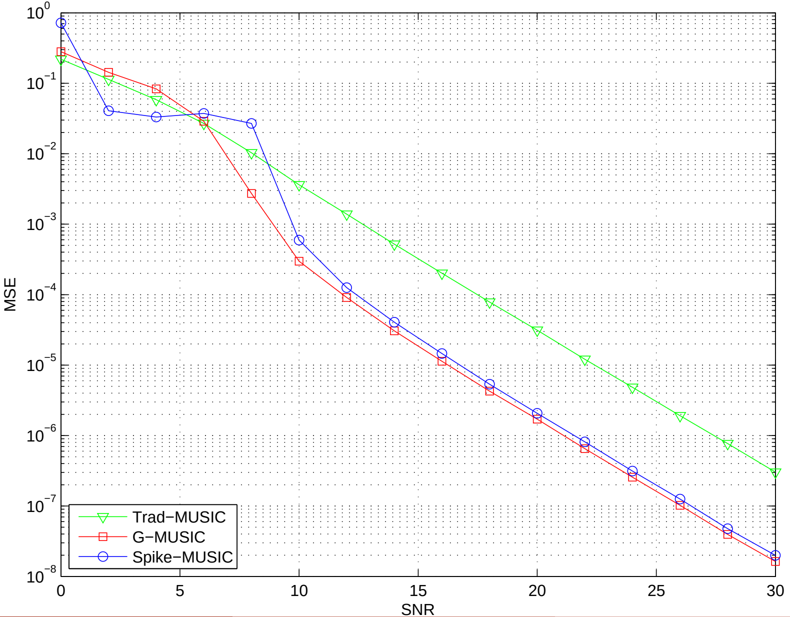
Parameters

- $\mathbf{a}(\varphi) = \frac{1}{\sqrt{M}}[1, \exp^{\imath\pi\sin(\varphi)}, \ldots, \exp^{\imath(M-1)\pi\sin(\varphi)}]^T$
- source signals are AR(1) processes with correlation coefficient of 0.9
- $K = 2, M = 20, N = 40,\ \varphi_1 = 16,\ \varphi_2 = 18$

Mean of the MSE of $\hat{\varphi}_1$ and $\hat{\varphi}_2$ versus SNR.



Mean of the MSE of the $\mathbf{a}(\varphi_i)^H \hat{\mathbf{\Pi}}_N^{\perp} \mathbf{a}(\varphi_i)$ versus SNR

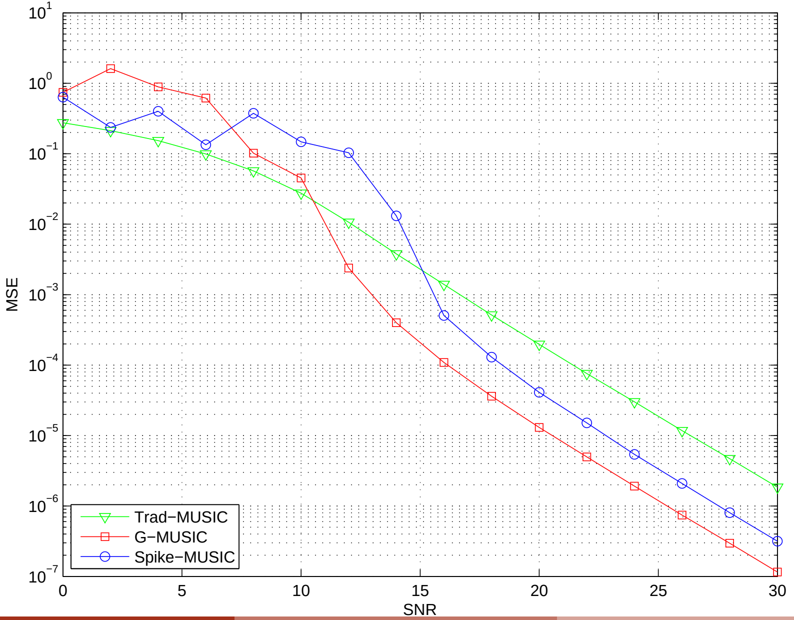


Experiment 2

Parameters

- $K = 5, M = 20, N = 40$
- angles equal to $-20, -10, 0, 10, 20$

Mean of the MSE of the $\mathbf{a}(\varphi_i)^H \hat{\boldsymbol{\Pi}}_N^{\perp} \mathbf{a}(\varphi_i)$ versus SNR



1. Problem statement
2. $K = 0$: An overview of Marčenko and Pastur's results
3. K fixed: spiked models
4. K may scale with M. Application to the subspace method.
5. Some research prospects

Future applications

- G-estimation of other parameters: number of sources, power distribution, ...
  Applications: cognitive radio or passive network metrology.
- Application of the spiked models for local failure detection/diagnosis in large data or power networks.

Methodological future research

- Spiked models:
  - Performance of tests for isolated eigenvalues, *e.g.* with the help of large deviations theory.
  - Design and evaluation of sphericity tests.
- G-estimation:
  - Extension of the G-estimation techniques to other matrix models.
  - Consistency and fluctuations of estimates.

Bibliography

- F. Benaych-Georges and R. R. Nadakuditi, "The singular values and vectors of low rank perturbations of large rectangular random matrices". ArXiv e-prints, 2011.
- P. Bianchi, M. Debbah, M. Maida, J. Najim, "Performance of Statistical Tests for Source Detection using Random Matrix Theory", *IEEE Trans. IT*, 57 (4), 2011.
- R. Couillet, J. W. Silverstein, Z. Bai, M. Debbah, "Eigen-Inference for Energy Estimation of Multiple Sources", *IEEE Trans. IT.*, 57(4), 2011.
- M. Debbah, W. Hachem, Ph. Loubaton and M. de Courville, "MMSE Analysis of Certain Large Isometric Random Precoded Systems", *IEEE Trans. IT*, 49 (5), 2003.
- R.B. Dozier and J.W. Silverstein, "On the empirical distribution of eigenvalues of large dimensional information plus noise type matrices", *JMVA*, 98 (4), 2007.

Bibliography

- R.B. Dozier and J.W. Silverstein, "Analysis of the limiting spectral distribution of large dimensional information plus noise type matrices", *JMVA*, 98 (6), 2007.
- J. Dumont, W. Hachem, S. Lasaulce, Ph. Loubaton and J. Najim, "On the Capacity Achieving Covariance Matrix for Rician MIMO Channels: An Asymptotic Approach", *IEEE Trans. IT*, 56 (3), 2010.
- W. Hachem, Ph. Loubaton, J. Najim and P. Vallet, "On bilinear forms based on the resolvent of large random matrices", ArXiv e-prints, 2010.
- N. El Karoui, "Spectrum estimation for large dimensional covariance matrices using random matrix theory", *Ann. Statist.*, 36 (6), 2008.
- L. Li, A.M. Tulino, and S. Verdú, "Design of reduced-rank MMSE multiuser detectors using random matrix methods", *IEEE Trans. IT*, 50 (6), 2004.

Bibliography

- X. Mestre, "Improved Estimation of Eigenvalues and Eigenvectors of Covariance Matrices Using Their Sample Estimates", *IEEE Trans. IT*, 54 (11), 2008.
- X. Mestre and M.A. Lagunas, "Modified subspace algorithms for DoA estimation with large arrays", *IEEE Trans. SP*, 56 (2), 2008.
- B. Nadler, "Nonparametric Detection of Signals by Information Theoretic Criteria: Performance Analysis and an Improved Estimator", *IEEE Trans. SP*, 58 (5), 2010.
- L.A. Pastur, "A Simple Approach to the Global Regime of Gaussian Ensembles of Random Matrices", *Ukrainian Math. J.*, 57 (6), 2005.
- I.E. Telatar, "Capacity of multi-antenna Gaussian channel", *European Trans. on Telecom.*, 1999.

Bibliography

- D. N. C. Tse and S.V. Hanly, "Linear multiuser receivers: effective interference, effective bandwidth and user capacity", *IEEE Trans. IT*, 45 (2), 1999.
- P. Vallet, Ph. Loubaton, and X. Mestre, "Improved Subspace Estimation for Multivariate Observations of High Dimension: The Deterministic Signal Case", accepted to *IEEE Trans. IT*, arXiv e-prints, 2010.
- S. Verdú and Sh. Shamai, "Spectral efficiency of CDMA with random spreading", *IEEE Trans. IT*, 45 (2), 1999.
- M. Vu and A. Paulraj, "Capacity optimization for Rician correlated MIMO wireless channels", in *Proc. Asilomar Conf.*, 2005.