Applications of large random matrices to high dimensional statistical signal processing

Gia-Thuy Pham

LIGM, Université Paris-Est Marne la Vallée.

February 28, 2017

Overview

Marchenko-Pastur distribution

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

$(V_{ij})_{1 \leq i \leq M, 1 \leq j \leq N}$ i.i.d. complex Gaussian random variables $\mathcal{CN}(0, \sigma^2)$.
$\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N$ columns of $\mathbf{V}$, $\mathbb{E}(\mathbf{v}_n \mathbf{v}_n^*) = \sigma^2 \mathbf{I}_M$

Empirical covariance matrix:

$$\frac{\mathbf{V}\mathbf{V}^*}{N} = \frac{1}{N} \sum_{n=1}^{N} \mathbf{v}_n \mathbf{v}_n^*$$

Marchenko-Pastur distribution

Empirical distribution of the eigenvalues of $\frac{\mathbf{V}\mathbf{V}^*}{N}$

- $\hat{\lambda}_{1,N} \geq \hat{\lambda}_{2,N} \geq \ldots \geq \hat{\lambda}_{M,N}$ eigenvalues of $\frac{\mathbf{V}\mathbf{V}^*}{N}$
- Empirical eigenvalue distribution: $\hat{\mu}_N = \frac{1}{M} \sum_{i=1}^{M} \delta(\lambda - \hat{\lambda}_{i,N})$

Asymptotic behaviour of $\hat{\mu}_N \longleftrightarrow$ Behaviour of the histograms of the eigenvalues $(\hat{\lambda}_{i,N})_{i=1,\ldots,M}$

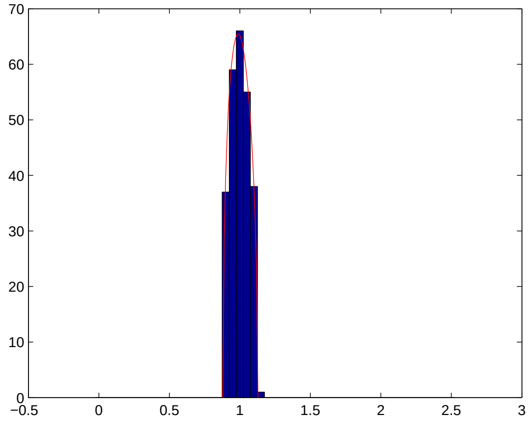
Well known case: M fixed, N increases i.e. $d_N = \frac{M}{N}$ small

- $\frac{\mathbf{V}\mathbf{V}^*}{N} \simeq \mathbb{E}(\mathbf{v}_n \mathbf{v}_n^*) = \sigma^2 \mathbf{I}_M$ by the law of large numbers
- $\hat{\mu}_N \xrightarrow{N \to +\infty} \delta(\sigma^2)$

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

Illustration

Histogram of the eigenvalues of $\frac{\mathbf{VV}^*}{N}$, $M = 256, d_N = \frac{M}{N} = \frac{1}{256}, \sigma^2 = 1$



Marchenko-Pastur distribution

M, N same order of magnitude, $d_N = \frac{M}{N} \to d$

$$\hat{\mu}_N \nrightarrow \delta(\sigma^2) \text{ because } \|\frac{\mathbf{V}\mathbf{V}^*}{N} - \sigma^2 \mathbf{I}_M\| \nrightarrow 0$$

Marchenko-Pastur distribution $\mathrm{MP}(\sigma^2, d)$: if $d \leq 1$

$$d\mu_{\sigma^2,d}(\lambda) = \frac{1}{2\pi\sigma^2 d\lambda}\sqrt{(\lambda^+ - \lambda)(\lambda - \lambda^-)}\mathbb{1}_{[\lambda^-,\lambda^+]}d\lambda$$

where $\lambda^{\pm} = \sigma^2(1 \pm \sqrt{d})^2$

Theorem (Marchenko-Pastur, 1967)

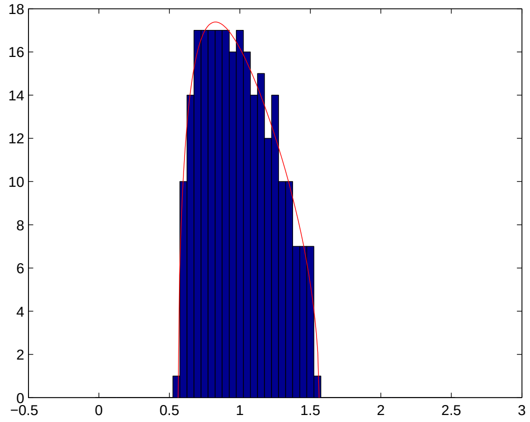
When $M, N \to +\infty$,$d_N = \frac{M}{N} \to d$, it holds that

$$\hat{\mu}_N \to \mu_{d,\sigma^2},\ a.s$$

Result still true in the non Gaussian case

Illustration

Histogram of the eigenvalues of $\frac{\mathbf{VV}^*}{N}$, $M = 256, d_N = \frac{M}{N} = \frac{1}{16}$, $\sigma^2 = 1$



Stieltjes transform

Definition

Let μ a measure (e.g a probability distribution) defined on $\mathbb{R}^+$, its Stieltjes transform is defined as

$$m_\mu(z) = \int_{\mathbf{R}^+} \frac{1}{\lambda - z} d\mu(\lambda),\ z \in \mathbb{C}\backslash\mathbb{R}^+$$

Remark

- $\mathbf{Q}_{\mathbf{V},N}(z) = \left(\frac{\mathbf{V}\mathbf{V}^*}{N} - z\mathbf{I}_M\right)^{-1}$ resolvent of $\frac{\mathbf{V}\mathbf{V}^*}{N}$
- $m_{\hat{\mu}_N}(z)$ coincides with $\frac{1}{M}\mathrm{Tr}\,\mathbf{Q}_{\mathbf{V},N}(z)$

Asymptotic regime: $d_N = \frac{M}{N} \to d$

It can be shown that $\lim_{N\to+\infty} m_{\hat{\mu}_N}(z) = m_{\mu_{d,\sigma^2}}(z)$ a.s, $z \in \mathbb{C}\backslash\mathbb{R}^+$. Thus it implies that

$$\hat{\mu}_N \to \mu_{d,\sigma^2},\ a.s$$

Important properties

- The eigenvalues of $\frac{\mathbf{V}\mathbf{V}^*}{N}$ concentrate in the neighbourhood of $[\sigma^2(1-\sqrt{d})^2, \sigma^2(1+\sqrt{d})^2] = [\lambda^-, \lambda^+]$

Denote $\mathbf{Q}_{\mathbf{V},N}(z) = \left(\frac{\mathbf{V}\mathbf{V}^*}{N} - z\mathbf{I}_M\right)^{-1}$, $\tilde{\mathbf{Q}}_{\mathbf{V},N}(z) = \left(\frac{\mathbf{V}^*\mathbf{V}}{N} - z\mathbf{I}_N\right)^{-1}$

- Uniformly, for each z in a compact subset of $\mathbb{C} - [\lambda^-, \lambda^+]$, for each sequences of unit M–dimensional vectors $(\mathbf{a}_N), (\mathbf{b}_N)$ and each sequences of N-dimensional vectors $(\tilde{\mathbf{a}}_N), (\tilde{\mathbf{b}}_N)$, we have that

$$\mathbf{a}_N^*(\mathbf{Q}_{\mathbf{V},N}(z) - m_{d,\sigma^2}(z)\mathbf{I}_M)\mathbf{b}_N \to 0 \; a.s$$

$$\tilde{\mathbf{a}}_N^*(\tilde{\mathbf{Q}}_{\mathbf{V},N}(z) - \tilde{m}_{d,\sigma^2}(z)\mathbf{I}_N)\tilde{\mathbf{b}}_N \to 0 \; a.s$$

$$\mathbf{a}_N^*(\mathbf{Q}_{\mathbf{V},N}(z)\mathbf{V}_N)\tilde{\mathbf{b}}_N \to 0 \; a.s$$

To be used in 16, 22, 26

The additive spatial spiked model

Observations: M–dimensional vectors, N snapshots

- $\mathbf{y}_n = \mathbf{A}_N \mathbf{s}_n + \mathbf{v}_n, \;\; n = 1, ..., N$
- $\mathbf{Y}_N = (\mathbf{y}_1, ..., \mathbf{y}_N)$
- $\mathbf{Y}_N = \mathbf{A}_N \mathbf{S}_N + \mathbf{V}_N$
- $((\mathbf{V}_N)_{i,j})_{1 \leq i \leq M, 1 \leq j \leq N} \overset{i.i.d}{\sim} \mathcal{CN}(0, \sigma^2)$
- $\mathbf{A}_N$ a $M \times K$ matrix, $\mathbf{S}_N$ a $K \times N$ matrix, both deterministic
- $Rank(\mathbf{A}_N) = K$

Asymptotic regime: $N \to \infty$, $d_N = \frac{M}{N} \to d$, and K is fixed.

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

Behaviour of eigenvalues and eigenvectors of $\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N}$

Notations

Spectral factorizations:

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

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

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

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

Impact of the signal component on the eigenvalues and eigenvectors of $\frac{\mathbf{Y}_N\mathbf{Y}_N^*}{N}$

If M is fixed and $N \to +\infty, d_N = \frac{M}{N} \simeq 0$

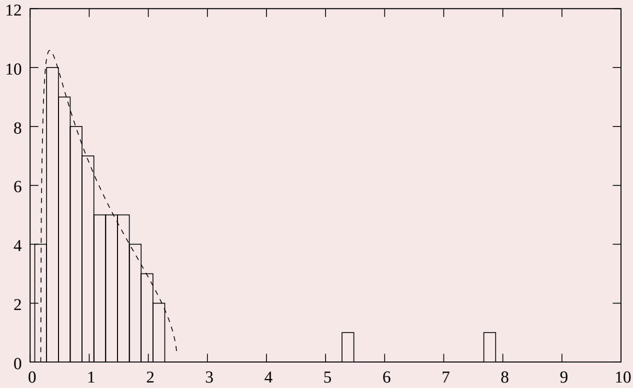
- $\frac{\mathbf{Y}_N\mathbf{Y}_N^*}{N} \simeq \mathbb{E}\left(\frac{\mathbf{Y}\mathbf{Y}^*}{N}\right) = \mathbf{A}_N \frac{\mathbf{S}_N\mathbf{S}_N^*}{N}\mathbf{A}_N^* + \sigma^2\mathbf{I}$
- $\hat{\lambda}_{k,N} \simeq \lambda_{k,N} + \sigma^2$ and $\hat{\mathbf{u}}_{k,N} \simeq \mathbf{u}_{k,N}$ if $1 \leq k \leq K$
- $\hat{\lambda}_{k,N} \simeq \sigma^2$ if $k > K$

In our asymptotic regime: $M, N \to +\infty \; d_N = \frac{M}{N} \to d$

- The asymptotic distribution of $M - K$ smallest eigenvalues of $\frac{\mathbf{Y}_N\mathbf{Y}_N^*}{N}$ is the Marchenko-Pastur
- Depending on the ratios $(\frac{\lambda_{k,N}}{\sigma^2})_{k=1,\ldots,K}$, at most K eigenvalues of $\frac{\mathbf{Y}_N\mathbf{Y}_N^*}{N}$ may escape from the support of the Marchenko Pastur and have a deterministic behaviour (more complicated than $\lambda_{k,N} + \sigma^2$)

Illustration

Histogram of the eigenvalues of $\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N}$, $d_N = \frac{M}{N} = 1/3$, $N = 192$, $K = 2$, $\lambda_1 = 6.25$, $\lambda_2 = 4$, $\sigma^2 = 1$



Main result on the eigenvalues and eigenvectors

Theorem : Benaych-Georges and Nadakuditi, 2011

- Assume that $\lambda_{k,N} \to \lambda_k$ for $k = 1, \ldots, K$.
- Let K_s the number of (λ_k) greater than $\sigma^2\sqrt{d}$.
  Then for $k = 1, \ldots, K_s$,

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

and for $K_s + 1 \leq k \leq K$

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

- Finally, for all deterministic sequences of unit vectors $(\mathbf{a}_N), (\mathbf{b}_N)$, for $k = 1, ..., K_s$

$$\mathbf{a}_N^* \hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^* \mathbf{b}_N = \frac{\lambda_k^2 - \sigma^4 d}{\lambda_k(\lambda_k + \sigma^2 d)} \mathbf{a}_N^* \mathbf{u}_{k,N} \mathbf{u}_{k,N}^* \mathbf{b}_N + o(1), \ a.s$$

$\lambda_{K_s} > \sigma^2\sqrt{d}$ "Signal Subspace Separation Condition"

Important remarks

It does not necessitate $\mathbf{V}_N$ i.i.d entries, the fundamental conditions are that

• The eigenvalues of $\frac{\mathbf{V}_N\mathbf{V}_N^*}{N}$ concentrate in the neighbourhood of $[\sigma^2(1-\sqrt{d})^2, \sigma^2(1+\sqrt{d})^2] = [\lambda^-, \lambda^+]$

• Uniformly, for each z in a compact subset of $\mathbb{C} - [\lambda^-, \lambda^+]$, for each sequences of unit M–dimensional vectors $(\mathbf{a}_N), (\mathbf{b}_N)$ and each sequences of N-dimensional vectors $(\tilde{\mathbf{a}}_N), (\tilde{\mathbf{b}}_N)$, we have that

$$\mathbf{a}_N^*(\mathbf{Q}_{\mathbf{V},N}(z) - m_{d,\sigma^2}(z)\mathbf{I}_M)\mathbf{b}_N \to 0 \, a.s$$

$$\tilde{\mathbf{a}}_N^*(\tilde{\mathbf{Q}}_{\mathbf{V},N}(z) - \tilde{m}_{d,\sigma^2}(z)\mathbf{I}_N)\tilde{\mathbf{b}}_N \to 0 \, a.s$$

$$\mathbf{a}_N^*(\mathbf{Q}_{\mathbf{V},N}(z)\mathbf{V}_N)\tilde{b}_N \to 0 \, a.s$$

For $z \in \mathbb{C} - \mathbb{R}^+$

$$\mathbf{Q}_N(z) = \left(\frac{\mathbf{Y}_N\mathbf{Y}_N^*}{N} - z\mathbf{I}_M\right)^{-1}, \, \mathbf{F}_N(z) = \left(-z(1+\sigma^2\tilde{m}_{d,\sigma^2}(z)) + \frac{\frac{\mathbf{A}_N\mathbf{S}_N\mathbf{S}_N^*\mathbf{A}_N^*}{N}}{1+\sigma^2 d m_{d,\sigma^2}(z)}\right)^{-1}$$

$$\mathbf{a}_N^*(\mathbf{Q}_N(z) - \mathbf{F}_N(z))\mathbf{b}_N \to 0 \, a.s$$

10

The observed signal

Observations: M–dimensional vectors, N snapshots

- $\mathbf{y}_n = \sum_{p=0}^{P-1} \mathbf{h}_p s_{n-p} + \mathbf{v}_n = [\mathbf{h}(z)] s_n + \mathbf{v}_n$
- $(s_n)_{n \in \mathbb{Z}}$ scalar deterministic sequence
- $\mathbf{h}(z) = \sum_{p=0}^{P-1} \mathbf{h}_p z^{-p}$ unknown SIMO transfer function
- $(\mathbf{v}_n)_{n \in \mathbb{Z}}$ temporally and spatially white complex Gaussian noise with variance σ^2.

Associated spatial model with P sources

- $\mathbf{y}_n = \mathbf{A}\mathbf{s}_n + \mathbf{v}_n$
- $\mathbf{A} = (\mathbf{h}_{P-1}, \ldots, \mathbf{h}_0)$
- $\mathbf{s}_n = (s_{n-(P-1)}, s_{n-(P-1)+1}, \ldots, s_n)^T$
- $\mathbf{Y} = \mathbf{A}\mathbf{S} + \mathbf{V}$
- $\mathbf{S}$ is a Hankel matrix, not taken into account

The extended observed signal

$(y_{k,n})_{n\in\mathbb{Z}}$ scalar signal received on sensor k.

For L an integer, define for each n L-dimensional vector $\mathbf{y}_{k,n}^{(L)}$ by:

$\mathbf{y}_{k,n}^{(L)} = \left(y_{k,n}, y_{k,n+1}, \ldots, y_{k,n+L-1}\right)^T$ and ML–dimensional vector $\mathbf{y}_n^{(L)}$ by:

$$\mathbf{y}_n^{(L)} = \begin{pmatrix} \mathbf{y}_{1,n}^{(L)} \\ \vdots \\ \mathbf{y}_{M,n}^{(L)} \end{pmatrix}$$

Define $ML \times N$ matrix $\mathbf{Y}_N^{(L)}$ by:

$$\mathbf{Y}_N^{(L)} = \left(\mathbf{y}_1^{(L)}, \ldots, \mathbf{y}_N^{(L)}\right)$$

$\mathbf{Y}_N^{(L)}$ is a block-Hankel matrix

$\mathbf{Y}_N^{(L)}$ is given by:

- $$\mathbf{Y}_N^{(L)} = \begin{bmatrix} \mathbf{Y}_{1,N}^{(L)} \\ \vdots \\ \mathbf{Y}_{M,N}^{(L)} \end{bmatrix}$$

Where for each k, $\mathbf{Y}_{k,N}^{(L)}$ is the $L \times N$ Hankel matrix

$$\mathbf{Y}_{k,N}^{(L)} = \begin{pmatrix} y_{k,1} & y_{k,2} & \cdots & y_{k,N} \\ y_{k,2} & y_{k,3} & \cdots & y_{k,N+1} \\ y_{k,3} & \cdots & \cdots & y_{k,N+2} \\ \vdots & \vdots & \vdots & \vdots \\ y_{k,L} & y_{k,L+1} & \cdots & y_{k,N+L-1} \end{pmatrix}$$

Expression of $\mathbf{Y}_N^{(L)}$

- $\mathbf{Y}_{k,N}^{(L)} = \mathbf{H}_k^{(L)} \mathbf{S}_N^{(L)} + \mathbf{V}_{k,N}^{(L)}$
- where $\mathbf{H}_k^{(L)}$ is a $L \times (P+L-1)$ Toeplitz matrix and $\mathbf{S}_N^{(L)}$ is a $(P+L-1) \times N$ Hankel matrix
- $\mathbf{Y}_N^{(L)} = \begin{pmatrix} \mathbf{H}_1^{(L)} \\ \vdots \\ \mathbf{H}_M^{(L)} \end{pmatrix} \mathbf{S}_N^{(L)} + \mathbf{V}_N^{(L)} = \mathbf{H}^{(L)} \, \mathbf{S}_N^{(L)} + \mathbf{V}_N^{(L)}$
- $\mathbf{Y}_N^{(L)}$ block-Hankel Information plus Noise random matrix
- $Rank(\mathbf{H}^{(L)} \mathbf{S}_N^{(L)}) \leq P + L - 1$

Eigenvalues / eigenvectors of the empirical spatio-temporal covariance matrix $\frac{\mathbf{Y}_N^{(L)} \mathbf{Y}_N^{(L)*}}{N}$?

Asymptotic behaviour of the eigenvalues of $\frac{\mathbf{V}_N^{(L)}\mathbf{V}_N^{(L)*}}{N}$.

Asymptotic regime

- $M \to +\infty$, $N \to +\infty$, $c_N = \frac{ML}{N} \to c$
- L may converge towards $+\infty$ but in such a way that $\frac{L}{N} \to 0$

Theorem [Loubaton, 2014]

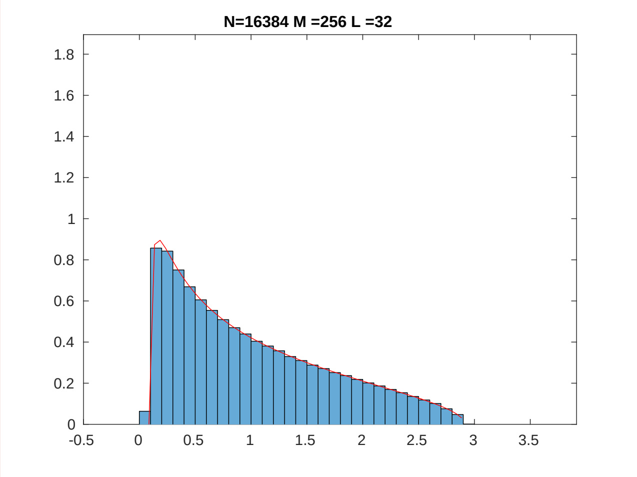
- The empirical eigenvalue distribution of $\frac{\mathbf{V}_N^{(L)}\mathbf{V}_N^{(L)*}}{N}$ has almost surely the same asymptotic behaviour than $\mathrm{MP}(\sigma^2, c)$
- If moreover $L = \mathcal{O}(N^\alpha)$ with $\alpha < 2/3$, nearly equivalent to $\frac{L}{M^2} \to 0$, then:
  - all the non zero eigenvalues of $\frac{\mathbf{V}_N^{(L)}\mathbf{V}_N^{(L)*}}{N}$ lie in a neighbourhood of $[\sigma^2(1-\sqrt{c})^2, \sigma^2(1+\sqrt{c})^2]$.

Moreover, we have proved that if $z \in \mathbb{C}\backslash[\sigma^2(1-\sqrt{c})^2, \sigma^2(1+\sqrt{c})^2]$, the bilinear forms of matrices $\mathbf{Q}_{\mathbf{V},N}(z) = (\frac{\mathbf{V}_N^{(L)}\mathbf{V}_N^{(L)*}}{N} - z\mathbf{I}_{ML})^{-1}$ and $\tilde{\mathbf{Q}}_{\mathbf{V},N}(z) = (\frac{\mathbf{V}_N^{(L)*}\mathbf{V}_N^{(L)}}{N} - z\mathbf{I}_N)^{-1}$ behave as if the entries of $\mathbf{V}_N^{(L)}$ were i.i.d.

10

Illustration

Histogram of the eigenvalues of $\frac{\mathbf{V}_N^{(L)}\mathbf{V}_N^{(L)*}}{N}$, $c_N = \frac{ML}{N} = \frac{1}{2}$, $N = 16384$, $M = 256, L = 32$ $\sigma^2 = 1$



Asymptotic behaviour of the largest eigenvalues and associated eigenvectors of $\frac{\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}}{N}$

Additive spatio-temporal spiked models asymptotic regime

- $M \to +\infty$, $N \to +\infty$, $d_N = \frac{M}{N} \to d$
- L and P do not scale with M and N

The rank $P + L - 1$ of signal matrix $\mathbf{H}^{(L)}\mathbf{S}_N^{(L)}$ does not scale with M and N

$$\mathbf{Y}_N^{(L)} = \mathbf{H}^{(L)}\mathbf{S}_N^{(L)} + \mathbf{V}_N^{(L)}$$

$\frac{\mathbf{V}_N^{(L)}\mathbf{V}_N^{(L)*}}{N}$ satisfies the properties that allow to use Benaych-Nadakuditi result.

Assumption

$(\lambda_{k,N}^{(L)})_{k=1,\dots,P+L-1}$ non zero eigenvalues of $\mathbf{H}^{(L)}\frac{\mathbf{S}_N^{(L)}\mathbf{S}_N^{(L)*}}{N}\mathbf{H}^{(L)*}$ converge towards $\lambda_1^{(L)} > \lambda_2^{(L)} > \dots > \lambda_{P+L-1}^{(L)}$ when $N \to +\infty$.

Notations

Spectral factorizations:

$$\frac{\mathbf{H}^{(L)}\mathbf{S}_N^{(L)}\mathbf{S}_N^{(L)*}\mathbf{H}^{(L)*}}{N} = \begin{bmatrix} \mathbf{u}_{1,N} & \cdots & \mathbf{u}_{P+L-1,N} \end{bmatrix} \begin{bmatrix} \lambda_{1,N}^{(L)} & & \\ & \ddots & \\ & & \lambda_{P+L-1,N}^{(L)} \end{bmatrix} \begin{bmatrix} \mathbf{u}_{1,N}^* \\ \vdots \\ \mathbf{u}_{P+L-1,N}^* \end{bmatrix}$$

where $\lambda_{1,N}^{(L)} \geq \cdots \geq \lambda_{P+L-1,N}^{(L)}$.

$$\frac{\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}}{N} = \begin{bmatrix} \hat{\mathbf{u}}_{1,N} & \cdots & \hat{\mathbf{u}}_{ML,N} \end{bmatrix} \begin{bmatrix} \hat{\lambda}_{1,N}^{(L)} & & \\ & \ddots & \\ & & \hat{\lambda}_{ML,N}^{(L)} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{u}}_{1,N}^* \\ \vdots \\ \hat{\mathbf{u}}_{ML,N}^* \end{bmatrix}$$

where $\hat{\lambda}_{1,N}^{(L)} \geq \cdots \geq \hat{\lambda}_{ML,N}^{(L)}$.

Results on eigenvalues of $\frac{\mathbf{V}_N^{(L)}\mathbf{V}_N^{(L)*}}{N}$ and bilinear forms of $\mathbf{Q}_{\mathbf{V},N}(z)$ and $\tilde{\mathbf{Q}}_{\mathbf{V},N}(z)$ allow to prove

Theorem

Let K_L the number of $\lambda_k^{(L)}$ greater than $\sigma^2\sqrt{dL}$

- For $k = 1, \dots, K_L$
  $\hat{\lambda}_{k,N}^{(L)} \xrightarrow[N\to\infty]{\text{a.s.}} \rho_k^{(L)} = \frac{\left(\lambda_k^{(L)}+\sigma^2\right)\left(\lambda_k^{(L)}+\sigma^2 dL\right)}{\lambda_k^{(L)}} > \sigma^2(1+\sqrt{dL})^2$
  while for $k = K_L + 1, ..., P + L - 1$

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

- For $k = 1, \dots, K_L$, for all deterministic sequences of ML–dimensional unit vectors $\mathbf{a}_N, \mathbf{b}_N$

$$\mathbf{a}_N^* \hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^* \mathbf{b}_N = \frac{\left(\lambda_k^{(L)}\right)^2 - \sigma^4 dL}{\lambda_k^{(L)}\left(\lambda_k^{(L)} + \sigma^2 dL\right)} \mathbf{a}_N^* \mathbf{u}_{k,N} \mathbf{u}_{k,N}^* \mathbf{b}_N + o(1)$$

10

Application to the analysis of subspace DoA estimation using spatial smoothing schemes

The original model

- $\mathbf{y}_n = [\mathbf{a}_M(\varphi_1), ..., \mathbf{a}_M(\varphi_K)]\mathbf{s}_n + \mathbf{v}_n = \mathbf{A}_M \mathbf{s}_n + \mathbf{v}_n$
- $\mathbf{a}_M(\phi) = \frac{1}{\sqrt{M}}(1, e^{\imath\phi}, \cdots, e^{\imath(M-1)\phi})^T$
- $\mathbf{Y}_N = \mathbf{A}_M \mathbf{S}_N + \mathbf{V}_N$

Results known when $\frac{M}{N} \to 0$ and $\frac{M}{N} \to c > 0$

Context

- Source localization using subspace method when M, N large, but $N << M$
- Spatial smoothing can be used in this context

Spatial smoothing

$L < M$: artificially create NL snapshots of dimension $M - L + 1$.

$$\mathcal{Y}_n^{(L)} = \begin{pmatrix} \mathbf{y}_{1,n} & \mathbf{y}_{2,n} & \cdots & \cdots & \mathbf{y}_{L,n} \\ \mathbf{y}_{2,n} & \mathbf{y}_{3,n} & \cdots & \cdots & \mathbf{y}_{L+1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{y}_{M-L+1,n} & \mathbf{y}_{M-L+2,n} & \cdots & \cdots & \mathbf{y}_{M,n} \end{pmatrix}$$

$$\mathbf{Y}_N^{(L)} = \left(\mathcal{Y}_1^{(L)}, \ldots, \mathcal{Y}_N^{(L)}\right)$$

Properties of $\mathbf{Y}_N^{(L)}$

- $\mathbf{Y}_N^{(L)} = \mathbf{A}^{(L)}(\mathbf{S}_N \otimes \mathbf{I}_L) + \mathbf{V}_N^{(L)}$
- $\mathbf{A}^{(L)}(\mathbf{S}_N \otimes \mathbf{I}_L)$ is a rank K deterministic $(M - L + 1) \times NL$ matrix
- $\mathrm{Range}(\mathbf{A}^{(L)}) = \mathrm{sp}\{\mathbf{a}_{M-L+1}(\varphi_k), k = 1, \ldots, K\}$

Application to the analysis of subspace DoA estimation using spatial smoothing schemes

The asymptotic regime

- $M \to +\infty$, $N = \mathcal{O}(M^\beta)$, $1/3 < \beta \leq 1$, $L = \mathcal{O}(M^\alpha), 0 \leq \alpha < 2/3$
- $e_N = \frac{M-L+1}{NL} \simeq \frac{M}{NL} \to e$

Remark

The structure of $\mathbf{V}_N^{(L)*}$:

$$\mathcal{V}_n^{(L)*} = \begin{pmatrix} \mathbf{v}_{1,n}^* & \mathbf{v}_{2,n}^* & \dots & \dots & \mathbf{v}_{M-L+1,n}^* \\ \mathbf{v}_{2,n}^* & \mathbf{v}_{3,n}^* & \dots & \dots & \mathbf{v}_{M-L+2,n}^* \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{v}_{L,n}^* & \mathbf{v}_{L+1,n}^* & \dots & \dots & \mathbf{v}_{M,n}^* \end{pmatrix} \Longrightarrow \mathbf{V}_N^{(L)*} = \begin{bmatrix} \mathcal{V}_1^{(L)*} \\ \vdots \\ \mathcal{V}_N^{(L)*} \end{bmatrix}$$

Application to the analysis of subspace DoA estimation using spatial smoothing schemes

Properties of the eigenvalues of $\frac{\mathbf{V}_N^{(L)}\mathbf{V}_N^{(L)*}}{NL}$.

- Non zero eigenvalues of $\frac{\mathbf{V}_N^{(L)}\mathbf{V}_N^{(L)*}}{NL}$ = non zero eigenvalues of $\frac{\mathbf{V}_N^{(L)*}\mathbf{V}_N^{(L)}}{NL}$
- Properties of the eigenvalues of $\frac{\mathbf{V}_N^{(L)*}\mathbf{V}_N^{(L)}}{NL}$ already evaluated before
- Just exchange $N \Longleftrightarrow M - L + 1$ and $M \Longleftrightarrow N$

Possible to use Benaych-Georges/Nadakuditi results

Assumption

The K non zero eigenvalues $(\lambda_{k,N})_{k=1,\ldots,K}$ of matrix $\frac{\mathbf{A}^{(L)}(\mathbf{S}_N\mathbf{S}_N^* \otimes \mathbf{I}_L)\mathbf{A}^{(L)*}}{NL}$ converge towards $\lambda_1 > \ldots > \lambda_K > \sigma^2\sqrt{c}$

Results

G-MUSIC subspace method can be used and analysed from the statistical point of view in the high dimensional context

Subspace separation condition

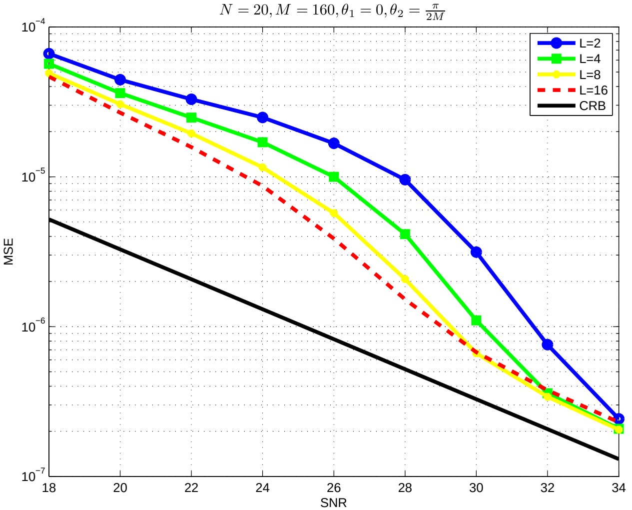
Comparison smoothed / unsmoothed when L does not converge $+\infty$.

- $\frac{\mathbf{S}_N \mathbf{S}_N^*}{N} \to \mathbf{D}$, $\mathbf{D}$ diagonal
- unsmoothed: $\lambda_K \left(\mathbf{A}_M^* \mathbf{A}_M \mathbf{D}\right) > \sigma^2 \sqrt{\frac{M}{N}}$
- smoothed: $\lambda_K \left(\mathbf{A}_{M-L}^* \mathbf{A}_{M-L} \mathbf{D}\right) > \frac{\sigma^2}{\sqrt{L}} \sqrt{\frac{M}{N}} = \sigma^2 \sqrt{\frac{M}{NL}}$

Discussion

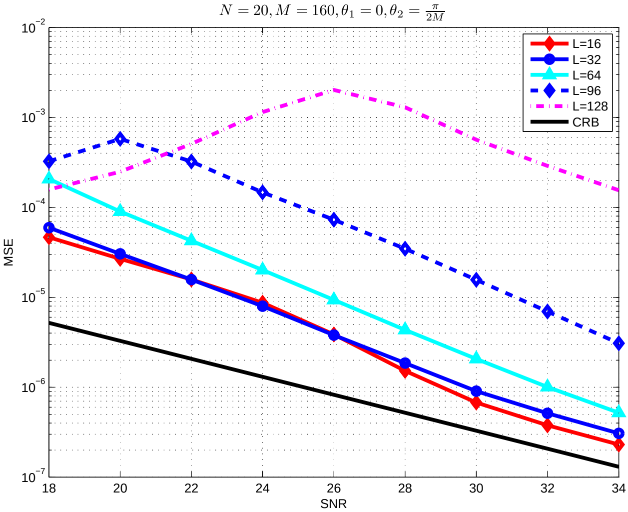
- If $L << M$, $\lambda_K \left(\mathbf{A}_{M-L}^* \mathbf{A}_{M-L} \mathbf{D}\right) \simeq \lambda_K \left(\mathbf{A}_M^* \mathbf{A}_M \mathbf{D}\right)$
- Clear improvement of the subspace separation condition if $L << M$
- If L increases too much, the diminution of the number of antennas due to the spatial smoothing becomes dominant.

Illustration I.



Empirical MSE of the improved subspace estimate of θ_1 for $L = 2, 4, 8, 16$ w.r.t. SNR.

Illustration II.



Empirical MSE of the improved subspace estimate of θ_1 for $L = 16, 32, 64, 128$ w.r.t. SNR.

Application to the loading factor estimation of trained spatio-temporal Wiener filters

Context

- $\mathbf{y}_n = \sum_{p=0}^{P-1} \mathbf{h}_p s_{n-p} + \mathbf{v}_n = [\mathbf{h}(z)] s_n + \mathbf{v}_n$, $[\mathbf{h}(z)]$ unknown
- Training sequence $(s_n)_{n=1,\ldots,N}$ available at the receiver side
- Estimate $\mathbf{g}^{(L)}$, ML–dimensional vector minimizing $\mathbb{E}|s_n - \mathbf{g}^{(L)*}\mathbf{y}_n^{(L)}|^2$
- Regularized least-squares estimate:

$$\hat{\mathbf{g}}_\lambda^{(L)} = \left(\frac{\mathbf{Y}_N^{(L)} \mathbf{Y}_N^{(L)*}}{N} + \lambda \mathbf{I} \right)^{-1} \left(\frac{1}{N} \sum_{n=1}^{N} \mathbf{y}_n^{(L)} s_n^* \right)$$

- Regularization necessary when $ML > N$, performance improved when $\frac{ML}{N}$ is not small enough
- Choose λ when M and N large and of the same order of magnitude
- Mestre-Lagunas IEEE SP 2006, $\mathbf{h}(z) = \mathbf{h}_0$ known (no training sequence), temporally white but spatially correlated noise + interference with unknown covariance matrix, $L = 1$

The SINR provided by filter $\hat{\mathbf{g}}_\lambda^{(L)}$

$\hat{\mathbf{g}}_\lambda^{(L)}$ is used to reconstruct s_n, for $n > N$

The signal to information plus noise ratio (SINR) mesures the performance of the reconstruction

$$\mathrm{SINR}(\hat{\mathbf{g}}_\lambda^{(L)}) = \frac{|\hat{\mathbf{g}}_\lambda^{(L)*}\mathbf{h}_P^{(L)}|^2}{\hat{\mathbf{g}}_\lambda^{(L)*}\mathbf{H}_{-P}^{(L)}\mathbf{H}_{-P}^{(L)*}\hat{\mathbf{g}}_\lambda^{(L)} + \sigma^2\|\hat{\mathbf{g}}_\lambda^{(L)}\|^2}$$

$\mathbf{h}_P^{(L)}$ column P of matrix $\mathbf{H}^{(L)}$, $\mathbf{H}_{-P}^{(L)}$ matrix obtained from $\mathbf{H}^{(L)}$ by deleting column P.

$\mathrm{SINR}(\hat{\mathbf{g}}_\lambda^{(L)})$ is a random variable because $\hat{\mathbf{g}}_\lambda^{(L)}$ depends on the noise corrupting the signal $(\mathbf{y}_n)_{n=1,,\ldots,N}$ received during the transmission of the training sequence.

Main results

$$\mathrm{SINR}(\hat{\mathbf{g}}_{\lambda}^{(L)}) = \frac{|\hat{\mathbf{g}}_{\lambda}^{(L)*}\mathbf{h}_P^{(L)}|^2}{\hat{\mathbf{g}}_{\lambda}^{(L)*}\mathbf{H}_{-P}^{(L)}\mathbf{H}_{-P}^{(L)*}\hat{\mathbf{g}}_{\lambda}^{(L)} + \sigma^2\|\hat{\mathbf{g}}_{\lambda}^{(L)}\|^2}$$

Main results: When M and N converge towards $+\infty$ at the same rate, and that P and L are fixed

- $\mathrm{SINR}(\hat{\mathbf{g}}_{\lambda}^{(L)})$ converges a.s. towards a deterministic term $\phi_L(\lambda)$ depending on λ and on $\sigma^2, \mathbf{H}^{(L)}$.
- While $\mathbf{H}^{(L)}$ is unknown at the receiver side, possible to estimate consistently $\phi_L(\lambda)$ for each $\lambda \geq 0$ by $\hat{\phi}_L(\lambda)$ from $(\mathbf{y}_n)_{n=1,\ldots,N}$
- λ_{opt} is estimated as the argmax of the consistent estimator $\lambda \to \hat{\phi}_L(\lambda)$.

Discussion

$$\text{Assume } \frac{\mathbf{S}_N^{(L)}\mathbf{S}_N^{(L)*}}{N} = \mathbf{I}_{P+L-1}$$

Assume $d_N L = \frac{ML}{N} < 1$ and $\lambda = 0$. Denote by γ the SINR provided by the true Wiener filter:

$$\gamma = \frac{\mathbf{h}_P^{(L)*}\left(\mathbf{H}^{(L)}\mathbf{H}^{(L)*} + \sigma^2\mathbf{I}\right)^{-1}\mathbf{h}_P^{(L)}}{1 - \mathbf{h}_P^{(L)*}\left(\mathbf{H}^{(L)}\mathbf{H}^{(L)*} + \sigma^2\mathbf{I}\right)^{-1}\mathbf{h}_P^{(L)}}$$

Then, the limit SINR $\phi_L(0)$ provided by $\hat{\mathbf{g}}_0^{(L)}$ is given by

$$\phi_L(0) = \gamma\,\frac{(1 - d_N L)\gamma}{\gamma + d_N}$$

SINR loss equal to $(1 - d_N L)\,\frac{\gamma}{\gamma + d_N}$

Some insights on the deterministic behaviour of the SINR

Expression of $\hat{\mathbf{g}}_\lambda^{(L)}$

$$\hat{\mathbf{g}}_\lambda^{(L)} = \left(\frac{\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}}{N} + \lambda \mathbf{I}\right)^{-1} \left(\frac{1}{N}\sum_{n=1}^{N} \mathbf{y}_n^{(L)} s_n^*\right)$$

$$\mathbf{Q}_N(-\lambda) = \left(\frac{\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}}{N} + \lambda \mathbf{I}\right)^{-1}, \mathbf{u}_N = \left(\frac{1}{\sqrt{N}}(s_1, \ldots, s_N)\right)^*,$$

$$\hat{\mathbf{g}}_\lambda^{(L)} = \mathbf{Q}_N(-\lambda)\, \frac{\mathbf{Y}_N^{(L)}}{\sqrt{N}}\, \mathbf{u}_N$$

Some insights on the deterministic behaviour of the SINR

$$\mathrm{SINR}(\hat{\mathbf{g}}_\lambda^{(L)}) = \frac{|\hat{\mathbf{g}}_\lambda^{(L)*}\mathbf{h}_P^{(L)}|^2}{\hat{\mathbf{g}}_\lambda^{(L)*}\mathbf{H}_{-P}^{(L)}\mathbf{H}_{-P}^{(L)*}\hat{\mathbf{g}}_\lambda^{(L)} + \sigma^2\|\hat{\mathbf{g}}_\lambda^{(L)}\|^2}$$

Evaluate the behaviour of

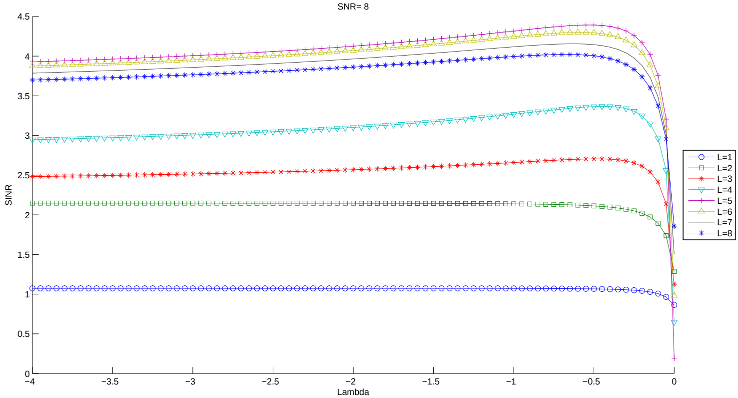
- $|\mathbf{a}_N^* \, \hat{\mathbf{g}}_\lambda^{(L)}|^2$ for each deterministic ML–dimensional vector $\mathbf{a}_N$.
- $\|\hat{\mathbf{g}}_\lambda^{(L)}\|^2$.

Equivalently

- $\mathbf{a}_N^*\mathbf{Q}_N(-\lambda)\frac{\mathbf{Y}_N^{(L)}}{\sqrt{N}}\mathbf{b}_N$
- $\mathbf{a}_N^*\frac{\mathbf{Y}_N^{(L)*}}{\sqrt{N}}\mathbf{Q}_N(-\lambda)^2\frac{\mathbf{Y}_N^{(L)}}{\sqrt{N}}\mathbf{b}_N = \mathbf{a}_N^*\left(\frac{d}{dz}\Big|_{z=-\lambda}(z\tilde{\mathbf{Q}}(z))\right)\mathbf{b}_N$

Illustration

$M = 40,\ N = 200, d_N = \frac{M}{N} = \frac{1}{5},\ P = 5,\ (\mathbf{h}_p)_{p=0,\ldots,4}$ random directional vectors



$\phi_L(\lambda)$ vs λ for various values of L.

Illustration

$M = 40,\ N = 200,\ P = 5,\ L = 5,\ (\mathbf{h}_p)_{p=0,\dots,4}$ random directional vectors

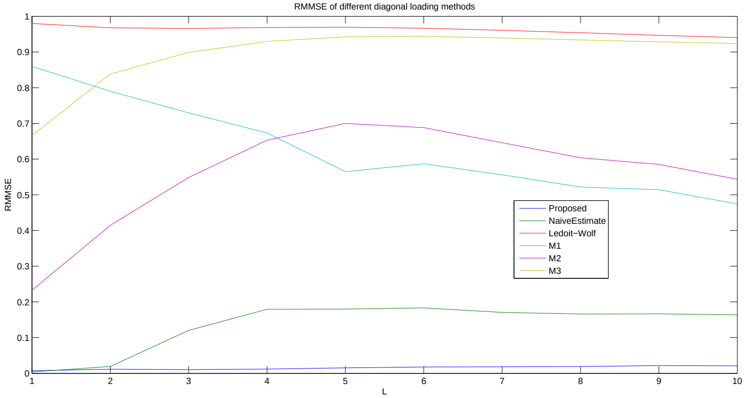


Figure: RRMSE (Root Relative Mean Square Error) of different diagonal loading methods versus L

General spatial-temporal information plus noise models

$$\mathbf{Y}_N^{(L)} = \mathbf{H}^{(L)}\mathbf{S}_N^{(L)} + \mathbf{V}_N^{(L)}$$

Extend the results to the case $P, L \to +\infty$

The general model

- $\mathbf{Y}_N^{(L)} = \mathbf{A}_N + \mathbf{V}_N^{(L)}$
- $\mathbf{A}_N$ deterministic, $\sup_N \left\| \frac{\mathbf{A}_N}{\sqrt{N}} \right\| < +\infty$, not necessary structured, $Rank(\mathbf{A}_N)$ not necessary finite.

Asymptotic regime

$N \to +\infty$, $c_N = \frac{ML}{N} \to c$, where $0 < c < +\infty$, $L = \mathcal{O}(N^\alpha)$, $\alpha < \frac{2}{3}$.

Behaviour of $\mathbf{Q}_N(z) = \left(\frac{\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}}{N} - z\mathbf{I}_{ML} \right)^{-1}$ and $\tilde{\mathbf{Q}}_N(z) = \left(\frac{\mathbf{Y}_N^{(L)*}\mathbf{Y}_N^{(L)}}{N} - z\mathbf{I}_N \right)^{-1}$

Deterministic equivalent matrices

Theorem

The resolvents $\mathbf{Q}_N(z) = (\mathbf{Y}_N\mathbf{Y}_N^* - \mathbf{I}_{ML})^{-1}$ and $\tilde{\mathbf{Q}}_N(z) = (\mathbf{Y}_N^*\mathbf{Y}_N - \mathbf{I}_N)^{-1}$ have the same behaviour than the deterministic matrices $\mathbf{T}_N(z)$, $\tilde{\mathbf{T}}_N(z)$ defined as

$$\begin{cases} \mathbf{T}_N(z) = \left[-z\left(\mathbf{I}_{ML} + \sigma^2\mathbf{I}_M \otimes \mathcal{T}_{L,L}(\tilde{\mathbf{T}}_N^T(z))\right) + \mathbf{A}_N\left(\mathbf{I}_N + \sigma^2 c_N \mathcal{T}_{N,L}^{(M)}(\mathbf{T}_N^T(z))\right)^{-1}\mathbf{A}_N^*\right]^{-1} \\ \tilde{\mathbf{T}}_N(z) = \left[-z\left(\mathbf{I}_N + \sigma^2 c_N \mathcal{T}_{N,L}^{(M)}(\mathbf{T}_N^T(z))\right) + \mathbf{A}_N^*\left(\mathbf{I}_{ML} + \sigma^2\mathbf{I}_M \otimes \mathcal{T}_{L,L}(\tilde{\mathbf{T}}_N^T(z))\right)^{-1}\mathbf{A}_N\right]^{-1} \end{cases}$$

For $z \in \mathbb{C}\backslash\mathbb{R}^+$,

- $\frac{1}{ML}\mathrm{Tr}\left[(\mathbf{Q}_N(z) - \mathbf{T}_N(z))\mathbf{B}_N\right] \to 0$, $\frac{1}{N}\mathrm{Tr}\left[(\tilde{\mathbf{Q}}_N(z) - \tilde{\mathbf{T}}_N(z))\tilde{\mathbf{B}}_N\right] \to 0$, $a.s$
- If $\alpha < \frac{2}{3}$, $\frac{L}{M^2} \to 0$, $\|\mathbf{Q}_N(z) - \mathbf{T}_N(z)\| \to 0$, $\|\tilde{\mathbf{Q}}_N(z) - \tilde{\mathbf{T}}_N(z)\| \to 0$, $a.s$

Application to the initial model

- If $\mathbf{A}_N = \mathbf{H}^{(L)}\mathbf{S}_N^{(L)}$
- $Rank(\mathbf{A}_N) = P + L - 1 = \mathcal{O}\left(N^{\alpha}\right)$

Signal assumption

- $\sup_N \|\mathbf{H}^{(L)}\| < +\infty \iff \sup_{M,L,\nu} \|\mathbf{h}(e^{2\imath\pi\nu})\|^2 < +\infty$
- $(s_n)_{n\in\mathbb{Z}}$ a real i.i.d sequence

Application to the initial model

Theorem

For $z \in \mathbb{C} \backslash \mathbb{R}^+$, it holds that

$$\|\mathbf{T}_N(z) - \mathbf{F}_N(z)\| \to 0$$
$$\|\tilde{\mathbf{T}}_N(z) - \tilde{\mathbf{F}}_N(z)\| \to 0$$

where

$$\mathbf{F}_N(z) = \left(-z(1 + \sigma^2 \tilde{m}_{c,\sigma^2}(z))\mathbf{I}_{ML} + \frac{\mathbf{H}^{(L)}\mathbf{H}^{(L)*}}{1 + \sigma^2 c m_{c,\sigma^2}(z)} \right)^{-1}$$
$$\tilde{\mathbf{F}}_N(z) = \left(-z(1 + \sigma^2 c m_{c,\sigma^2}(z))\mathbf{I}_N + \frac{\frac{\mathbf{S}_N^{(L)*}}{\sqrt{N}} \mathbf{H}^{(L)*}\mathbf{H}^{(L)} \frac{\mathbf{S}_N^{(L)}}{\sqrt{N}}}{1 + \sigma^2 \tilde{m}_{c,\sigma^2}(z)} \right)^{-1}$$

As a consequence,

$$\|\mathbf{Q}_N(z) - \mathbf{F}_N(z)\| \to 0, \ a.s$$
$$\|\tilde{\mathbf{Q}}_N(z) - \tilde{\mathbf{F}}_N(z)\| \to 0, \ a.s$$

Application to the loading factor estimation of trained spatio-temporal Wiener filters

$$\mathrm{SINR}(\hat{\mathbf{g}}_\lambda^{(L)}) = \frac{|\hat{\mathbf{g}}_\lambda^{(L)*}\mathbf{h}_P^{(L)}|^2}{\hat{\mathbf{g}}_\lambda^{(L)*}\mathbf{H}_{-P}^{(L)}\mathbf{H}_{-P}^{(L)*}\hat{\mathbf{g}}_\lambda^{(L)} + \sigma^2\|\hat{\mathbf{g}}_\lambda^{(L)}\|^2}$$

Asymptotic regime

$M, N \to +\infty$, $c_N = \frac{ML}{N} \to c$, $P, L = \mathcal{O}(N^\alpha)$, $0 < \alpha < \frac{1}{2}$, $\frac{L}{M} \to 0$.

Evaluate the behaviour of

- $|\mathbf{a}^* \mathbf{Q}_N(-\lambda) \frac{\mathbf{Y}_N^{(L)}}{\sqrt{N}} \mathbf{b}_N|^2$
- $\mathbf{a}_N^* \frac{\mathbf{Y}_N^{(L)*}}{\sqrt{N}} \mathbf{Q}_N(-\lambda) \mathbf{H}_{-P}^{(L)*} \mathbf{H}_{-P}^{(L)} \mathbf{Q}_N(-\lambda) \frac{\mathbf{Y}_N^{(L)}}{\sqrt{N}} \mathbf{b}_N$
- $\mathbf{a}_N^* \frac{\mathbf{Y}_N^{(L)*}}{\sqrt{N}} \mathbf{Q}_N(-\lambda)^2 \frac{\mathbf{Y}_N^{(L)}}{\sqrt{N}} \mathbf{b}_N = \mathbf{a}_N^* \left(\frac{d}{dz}\Big|_{z=-\lambda} (z\tilde{\mathbf{Q}}_N(z)) \right) \mathbf{b}_N$

Same results as the case where P, L are fixed

Conclusion

Spatial-temporal spiked models

- Behaviour of the largest eigenvalues and eigenvectors of $\frac{\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}}{N}$
- Application to detection of a wideband signal
- Loading factor estimation for trained regularized spatio-temporal Wiener filtering
- Analysis of spatial smoothing schemes in narrow band array processing

General spatial-temporal information plus noise models

- Behaviour of resolvent and co-resolvent of $\frac{\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}}{N}$
- Loading factor estimation for trained regularized spatio-temporal Wiener filtering

Perspectives

- Convergence rate of normalized trace, bilinear forms and spectral norms of the resolvents towards deterministic equivalents.
- Improvement of the convergence conditions for the $SINR$ ($\frac{L}{N} \to 0$, $\frac{L}{M^2} \to 0$)
- Second order of the detection test.


MERCI!
THANK YOU!
FRAPAR.


PhD student
First year !!
Last year !!