An introduction to G-estimation with sample covariance matrices

Xavier Mestre

xavier.mestre@cttc.cat

Centre Tecnològic de Telecomunicacions de Catalunya (CTTC)

"Matrices aleatoires: applications aux communications numeriques et a l'estimation statistique"

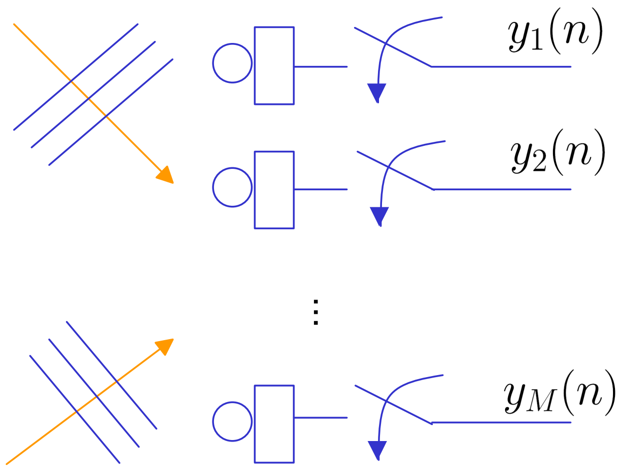
ENST (Paris), January 10, 2007

Outline

- Convergence of the eigenvalues and eigenvectors of the sample covariance matrix.
- Characterization of the clusters/support of the asymptotic sample eigenvalue distribution.
- The principle of G-estimation: M, N-consistency versus N-consistency. Some examples.
- An application in array signal processing: subspace-based estimation of directions-of-arrival (DoA).

The sample covariance matrix

We assume that we collect N independent samples (snapshots) from an array of M antennas: $\hat{\mathbf{R}}_M = \frac{1}{N}\sum_{n=1}^{N} \mathbf{y}(n)\mathbf{y}^H(n)$.

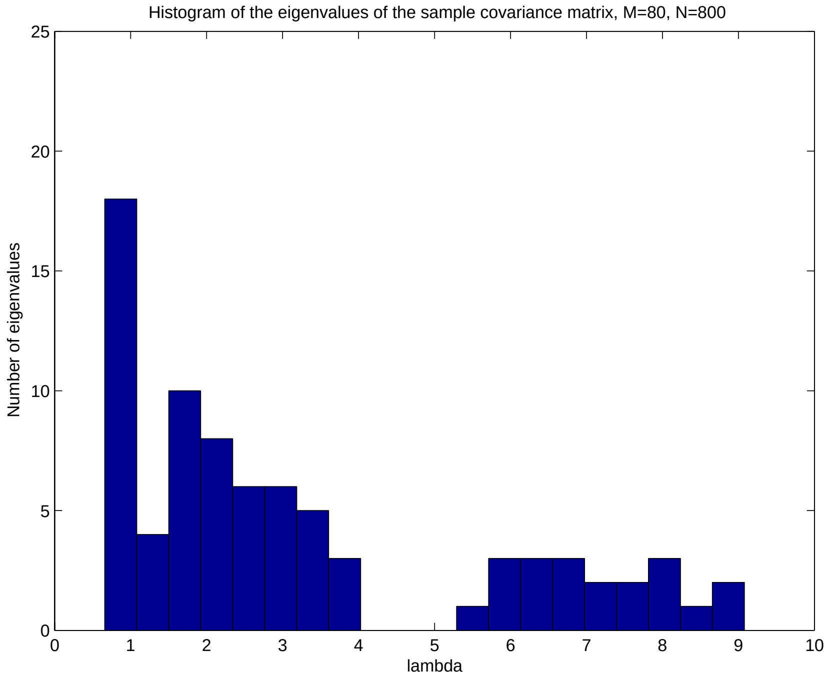


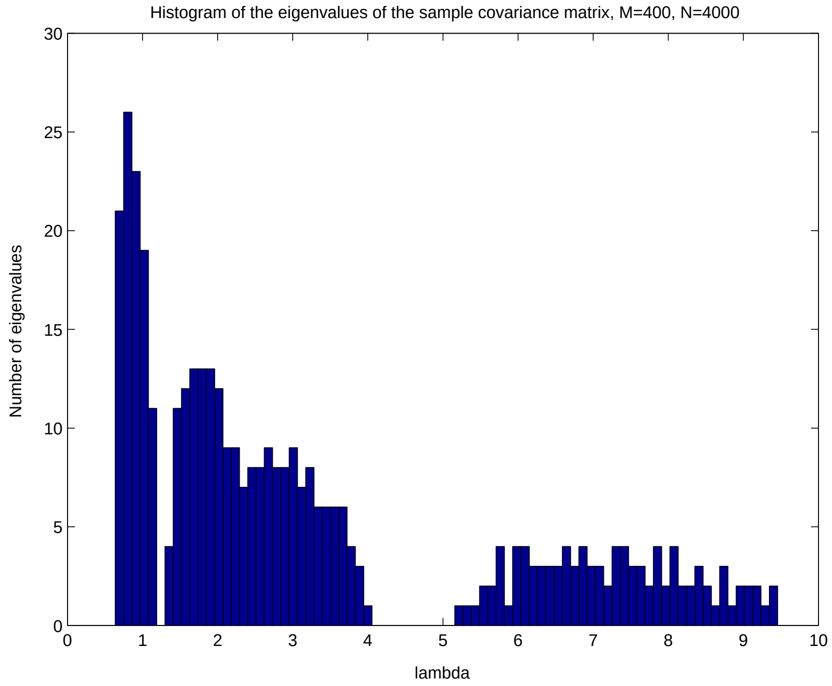
$$\mathbf{y}(n) = \begin{bmatrix} y_1(n) \\ y_2(n) \\ \vdots \\ y_M(n) \end{bmatrix}, n = 1 \ldots N$$

It is usual to model the observations $\{\mathbf{y}(n)\}$ as independent, identically distributed random vectors with zero mean and covariance $\mathbf{R}_M$.

The sample covariance matrix

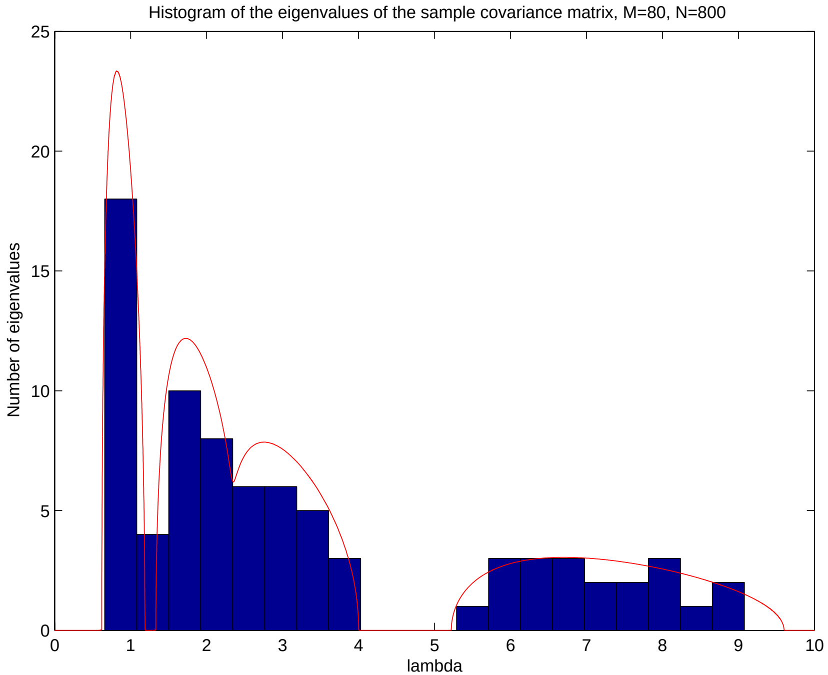
We assume that we collect N independent samples (snapshots) from an array of M antennas: $\hat{\mathbf{R}}_M = \frac{1}{N}\sum_{n=1}^{N} \mathbf{y}(n)\mathbf{y}^H(n)$. Example: $\mathbf{R}_M$ has 4 eigenvalues $\{1, 2, 3, 7\}$ with equal multiplicity.

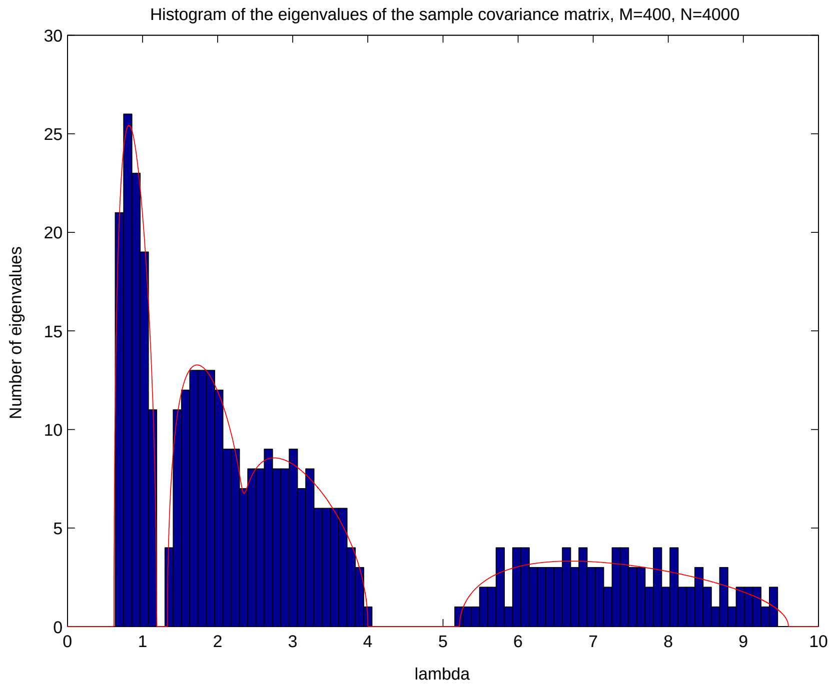




The sample covariance matrix

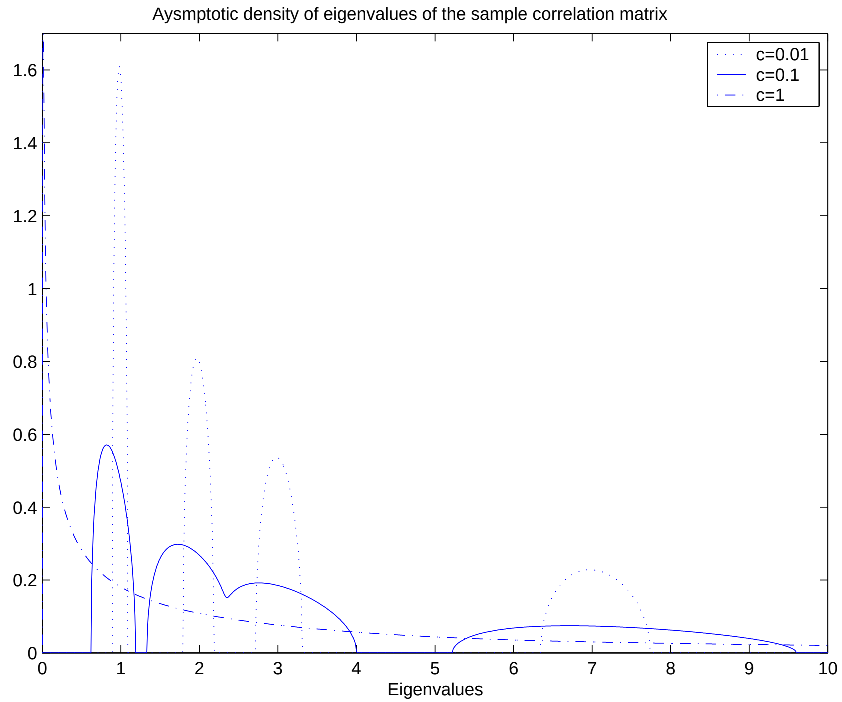
We assume that we collect N independent samples (snapshots) from an array of M antennas: $\hat{\mathbf{R}}_M = \frac{1}{N}\sum_{n=1}^{N} \mathbf{y}(n)\mathbf{y}^H(n)$. Example: $\mathbf{R}_M$ has 4 eigenvalues $\{1,2,3,7\}$ with equal multiplicity.





The sample covariance matrix: asymptotic properties

When both $M, N \to \infty$, $M/N \to c$, $0 < c < \infty$, the e.d.f. of the eigenvalues of $\hat{\mathbf{R}}_M$ tends to a deterministic density function. Example: $\mathbf{R}_M$ has 4 eigenvalues $\{1, 2, 3, 7\}$ with equal multiplicity.



Convergence of the eigenvalues and eigenvectors of the sample covariance matrix

- Can we characterize the asymptotic sample eigenvalue behavior analytically? What about the eigenvectors?
- We will consider the following two quantities:

$$\hat{b}_M(z) = \frac{1}{M}\operatorname{tr}\left[\left(\hat{\mathbf{R}}_M - z\mathbf{I}_M\right)^{-1}\right], \quad \hat{m}_M(z) = \mathbf{a}^H\left(\hat{\mathbf{R}}_M - z\mathbf{I}_M\right)^{-1}\mathbf{b}$$

  where $z \in \mathbb{C}$ and $\mathbf{a}, \mathbf{b} \in \mathbb{C}^M$.
- Note that $\hat{b}_M(z)$ depends only on the eigenvalues of $\hat{\mathbf{R}}_M$, whereas $\hat{m}_M(z)$ depends on both the eigenvalues and the eigenvectors of this matrix.
- Much of the asymptotic behavior of the eigenvalues and eigenvectors of $\hat{\mathbf{R}}_M$ can be inferred from this two quantities.

Stieltjes transforms

The function $\hat{b}_M(z)$ is the classical Stieltjes transform of the empirical distribution of eigenvalues:

$$\hat{b}_M(z) = \frac{1}{M}\sum_{m=1}^{M}\frac{1}{\hat{\lambda}_m - z} = \int \frac{1}{\lambda - z} d\mathbb{F}_{\hat{\mathbf{R}}_M}(\lambda)$$

where

$$\mathbb{F}_{\hat{\mathbf{R}}_M}(\lambda) = \frac{1}{M}\#\left\{m = 1 \ldots M : \hat{\lambda}_m \leq \lambda\right\} = \frac{1}{M}\sum_{m=1}^{M}\mathcal{I}\left(\hat{\lambda}_m \leq \lambda\right).$$

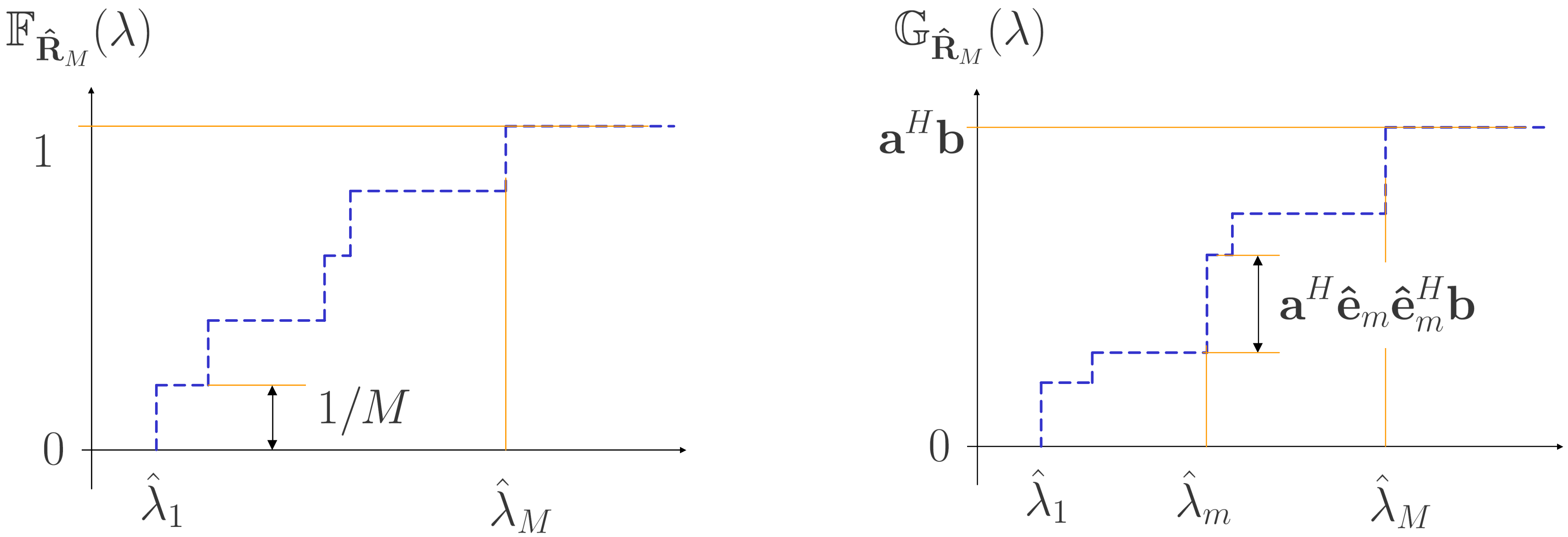
In the same way, $\hat{m}_M(z)$ is the Stieltjes transform of a different empirical distribution function

$$\hat{m}_M(z) = \sum_{m=1}^{M}\frac{\mathbf{a}^H\hat{\mathbf{e}}_m\hat{\mathbf{e}}_m^H\mathbf{b}}{\hat{\lambda}_m - z} = \int \frac{1}{\lambda - z} d\mathbb{G}_{\hat{\mathbf{R}}_M}(\lambda)$$

where

$$\mathbb{G}_{\hat{\mathbf{R}}_M}(\lambda) = \sum_{m=1}^{M}\mathbf{a}^H\hat{\mathbf{e}}_m\hat{\mathbf{e}}_m^H\mathbf{b}\,\mathcal{I}\left(\hat{\lambda}_m \leq \lambda\right).$$

Graphical representation of the considered empirical distribution functions



Convergence of the eigenvalues and eigenvectors of the sample covariance matrix

Theorem. Under some statistical assumptions,

$$\left|\hat{m}_M(z) - \bar{m}_M(z)\right| \to 0, \quad \left|\hat{b}_M(z) - \bar{b}_M(z)\right| \to 0$$

almost surely for all $z \in \mathbb{C}^+ = \{z \in \mathbb{C} : \mathrm{Im}\,[z] > 0\}$ as $M, N \to \infty$ at the same rate, where $\bar{b}_M(z) = b$ is the unique solution to the following equation in the set $\left\{b \in \mathbb{C} : -(1-c)/z + cb \in \mathbb{C}^+\right\}$:

$$b = \frac{1}{M}\sum_{m=1}^{M} \frac{1}{\lambda_m\left(1 - c - czb\right) - z}$$

and

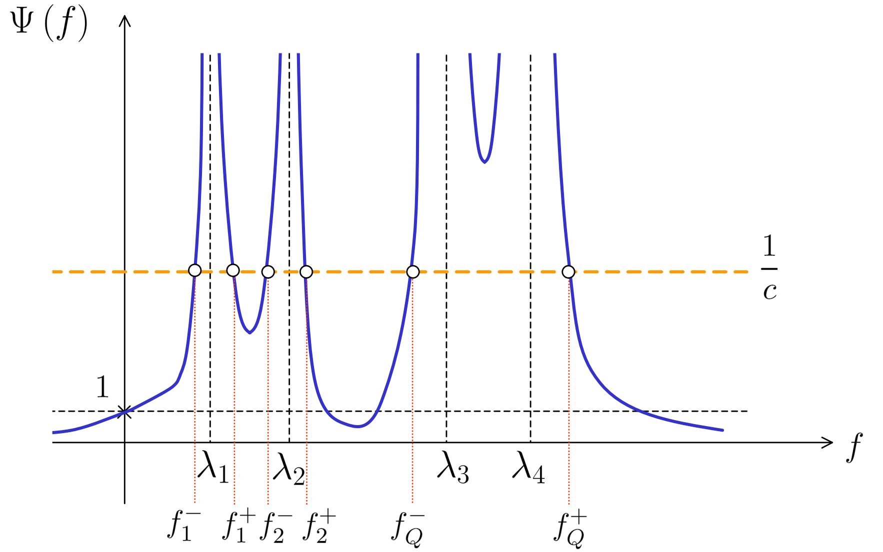
$$\bar{m}_M(z) = \frac{f_M(z)}{z}\mathbf{a}^H\left(\mathbf{R}_M - f_M(z)\mathbf{I}_M\right)^{-1}\mathbf{b}, \quad f_M(z) = \frac{z}{1 - c - cz\bar{b}_M(z)}.$$

Assuming that $\bar{b}(z) = \lim_{M\to\infty} \bar{b}_M(z)$ exists, then the asymptotic sample eigenvalue *density* can be obtained as $q(x) = \frac{1}{\pi}\lim_{y\to 0^+} \mathrm{Im}\left[\bar{b}(x + \mathrm{j}\,y)\right]$.

Determining the support of $q(x)$

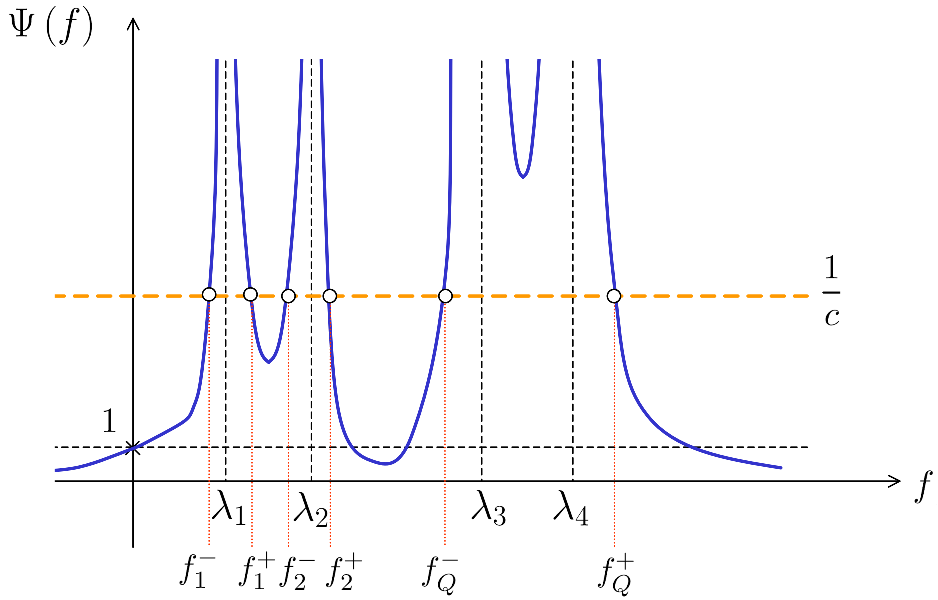
Consider the following equation in f, which has $2Q$ solutions (counting multiplicities)

$$\Psi(f) = \frac{1}{M}\sum_{m=1}^{M}\left(\frac{\lambda_m}{\lambda_m - f}\right)^2 = \frac{1}{c}.$$



Determining the support of $q(x)$

The solutions are denoted $\left\{f_1^-, f_1^+, \ldots, f_Q^-, f_Q^+\right\}$. For each eigenvalue λ_m, there exists a single $q \in \{1 \ldots Q\}$, such that $\lambda_m \in \left(f_q^-, f_q^+\right)$. In other words, each eigenvalue is associated with a single cluster, but one cluster may be associated with multiple eigenvalues.



Determining the support of $q(x)$

It turns out that Q is the number of clusters in $q(x)$, and the position of each cluster is given by

$$x_q^+ = \Phi(f_q^+), \quad x_q^- = \Phi(f_q^-)$$

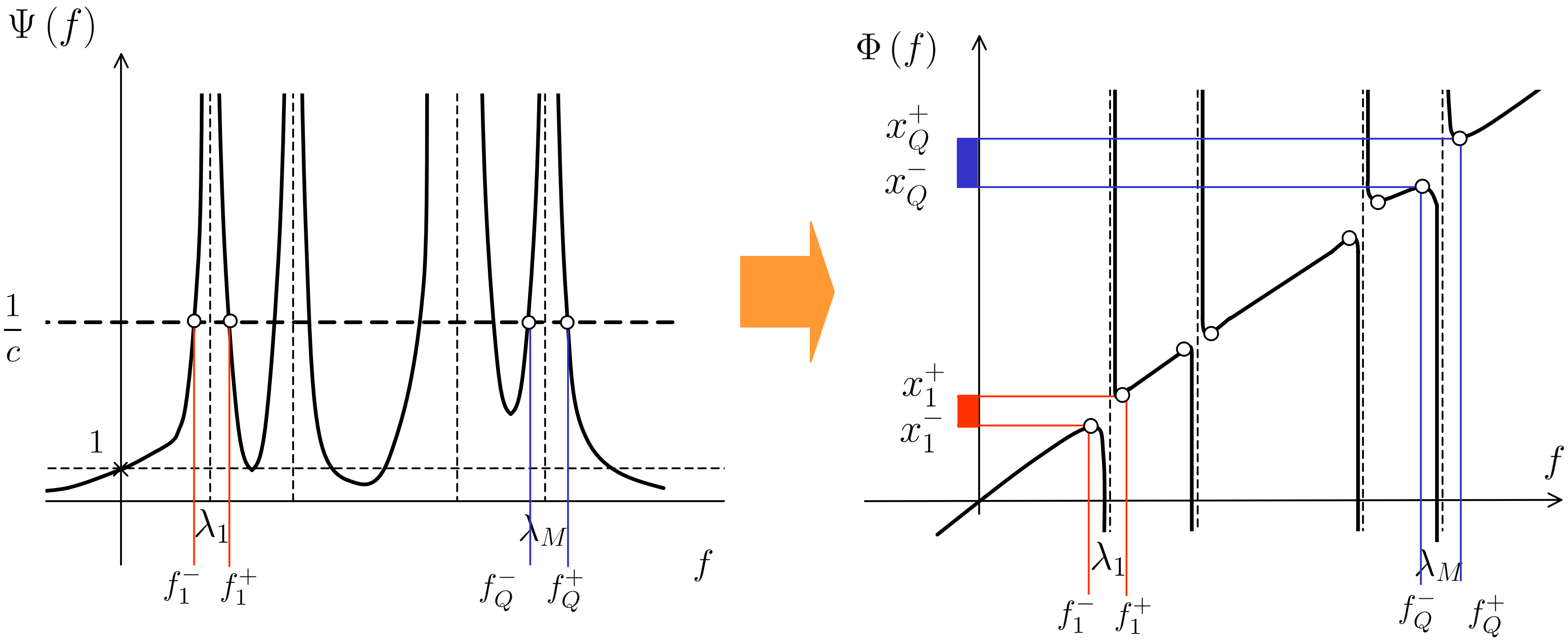
where the function $\Phi(f)$ is defined as

$$\Phi(f) = f\left(1 - \frac{c}{M}\sum_{m=1}^{M}\frac{\lambda_m}{\lambda_m - f}\right).$$

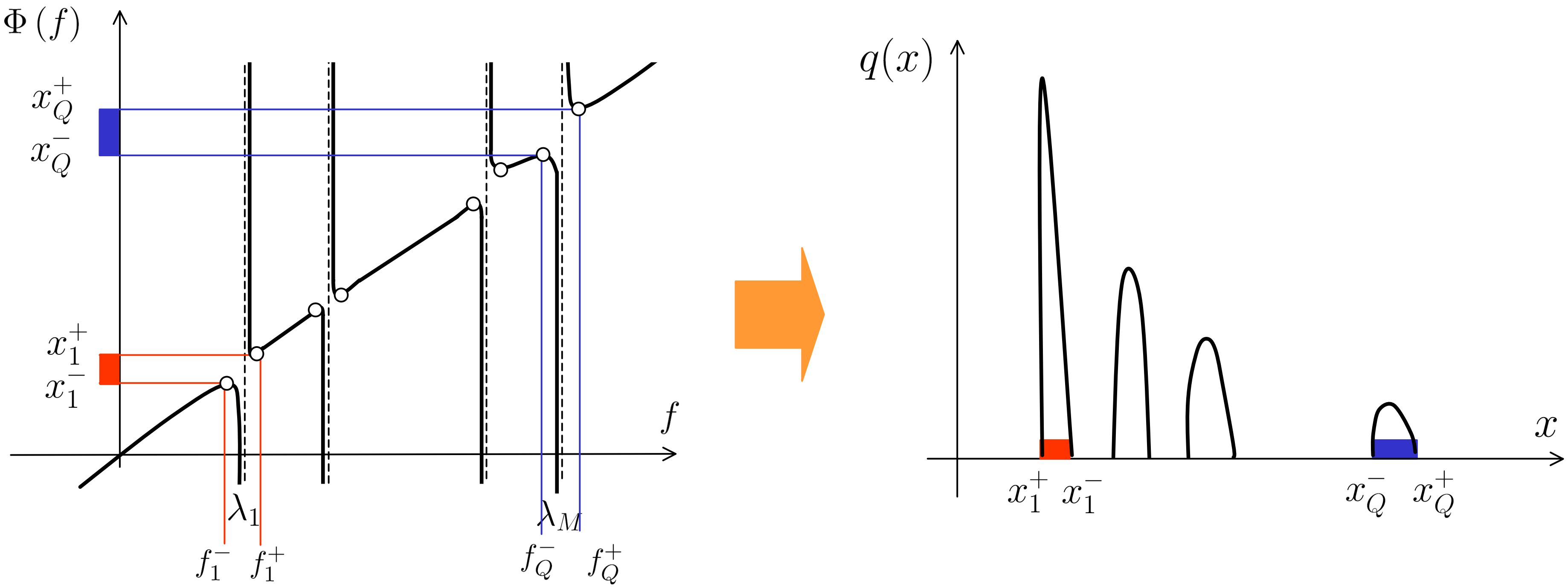
Hence, the support of $q(x)$ can be expressed as $\mathcal{S} = \left[x_1^-, x_1^+\right] \cup \ldots \cup \left[x_Q^-, x_Q^+\right]$.

Note: It may happen that $\lambda_m \notin \left(x_q^-, x_q^+\right)$. even when $\lambda_m \in \left(f_q^-, f_q^+\right)$.

Determining the support of $q(x)$



Determining the support of $q(x)$

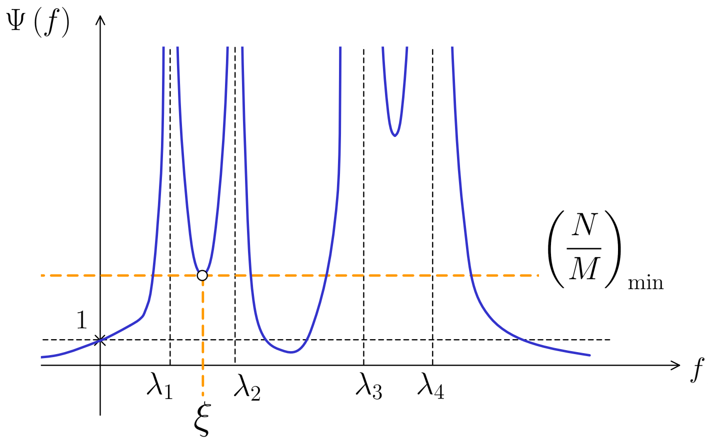


Splitting number for two clusters in $q(x)$

Given two consecutive eigenvalues $\{\lambda_m, \lambda_{m+1}\}$ there exists a minimum number of samples per antenna to guarantee that the corresponding eigenvalue clusters split.

$$\left(\frac{N}{M}\right)_{\min} > \frac{1}{M}\sum_{k=1}^{M}\left(\frac{\lambda_k}{\lambda_k - \xi}\right)^2$$

where ξ is the mth real valued solution to $\Psi'(f) = 0$.



Modeling the finite sample size effect using random matrix theory

Random Matrix Theory offers the possibility of analyzing the behavior of different quantities depending on $\hat{\mathbf{R}}_M$ when the sample size and the number of sensors/antennas have the same order of magnitude.

Assume that we want to analyze the asymptotic behavior of a certain scalar function of $\hat{\mathbf{R}}_M$, namely $\phi\left(\hat{\mathbf{R}}_M\right)$.

- **Traditional Approach:** Assuming that the number of samples is high, we might establish that $\phi\left(\hat{\mathbf{R}}_M\right) \to \phi\left(\mathbf{R}_M\right)$ in some stochastic sense as $N \to \infty$ while M remains fixed.
- **New Approach:** In order to characterize the situation where M, N have the same order of magnitude, one might consider the limit $N, M \to \infty$, $M/N \to c$, $0 < c < \infty$. Note that, in general,

$$\left|\phi\left(\hat{\mathbf{R}}_M\right) - \phi\left(\mathbf{R}_M\right)\right| \nrightarrow 0, \quad N, M \to \infty, M/N \to c$$

For example:

$$\left|\frac{1}{M}\operatorname{tr}\left[\hat{\mathbf{R}}_M^{-1}\right] - \frac{1}{1-c}\frac{1}{M}\operatorname{tr}\left[\mathbf{R}_M^{-1}\right]\right| \to 0, \quad c < 1.$$

Using random matrix theory to analyze the finite sample size effect

Assume that the quantity that we need to characterize can be expressed in terms of a Stieltjes transform:

$$\frac{1}{M}\operatorname{tr}\left[\hat{\mathbf{R}}_M^{-1}\right] = \frac{1}{M}\sum_{m=1}^{M}\frac{1}{\hat{\lambda}_m} = \frac{1}{M}\sum_{m=1}^{M}\left.\frac{1}{\hat{\lambda}_m - z}\right|_{z=0} = \hat{b}_M(0)$$

$$\left\{\hat{\mathbf{R}}_M^{-1}\right\}_{i,j} = \mathbf{u}_i^H\hat{\mathbf{R}}_M^{-1}\mathbf{u}_j = \mathbf{u}_i^H\left(\sum_{m=1}^{M}\frac{\hat{\mathbf{e}}_m\hat{\mathbf{e}}_m^H}{\hat{\lambda}_m}\right)\mathbf{u}_j = \sum_{m=1}^{M}\left.\frac{\mathbf{u}_i^H\hat{\mathbf{e}}_m\hat{\mathbf{e}}_m^H\mathbf{u}_j}{\hat{\lambda}_m - z}\right|_{z=0} = \hat{m}_M(0)$$

where $\mathbf{u}_i = \left[\begin{array}{ccc}\mathbf{0}^T & \underbrace{1}_{i\text{th position}} & \mathbf{0}^T\end{array}\right]^T$. Now,

$$\left|\hat{b}_M(z) - \bar{b}_M(z)\right| \;\to\; 0 \text{ “}\Longrightarrow\text{”} \left|\frac{1}{M}\operatorname{tr}\left[\hat{\mathbf{R}}_M^{-1}\right] - \bar{b}_M(0)\right| \to 0, \quad \bar{b}_M(0) = \frac{1}{1-c}\frac{1}{M}\operatorname{tr}\left[\mathbf{R}_M^{-1}\right]$$

$$\left|\hat{m}_M(z) - \bar{m}_M(z)\right| \;\to\; 0 \text{ “}\Longrightarrow\text{”} \left|\left\{\hat{\mathbf{R}}_M^{-1}\right\}_{i,j} - \bar{m}_M(0)\right| \to 0, \quad \bar{m}_M(0) = \frac{1}{1-c}\left\{\hat{\mathbf{R}}_M^{-1}\right\}_{i,j}$$

Estimation under low sample support with random matrix theory

When designing an estimator of a certain scalar function of $\mathbf{R}_M$, namely $\varphi(\mathbf{R}_M)$, one can distinguish between:

- **Traditional N-consistency:** Consistency when $N \to \infty$ while M remains fixed.
- **M, N-consistency:** Consistency when $M, N \to \infty$ at the same rate.

We observe that M, N-consistency guarantees a good behavior when the number of samples N has the same order of magnitude as the observation dimension M.

The objective of G-estimation (V.L. Girko) is to provide a systematic approach for the derivation of M, N-consistent estimators of different scalar functions of the true covariance matrix. For example, the G-estimator of $\frac{1}{M}\operatorname{tr}\left[\mathbf{R}_M^{-1}\right]$ will be

$$\frac{(1-c)}{M}\operatorname{tr}\left[\hat{\mathbf{R}}_M^{-1}\right].$$

The principle of G-estimation

We will assume that the quantity that we need to estimate, namely $\varphi(\mathbf{R}_M)$, can be expressed in terms of the Stieltjes transforms associated with $\mathbf{R}_M$, namely

$$b_M(\omega) = \frac{1}{M}\operatorname{tr}\left[(\mathbf{R}_M - \omega\mathbf{I}_M)^{-1}\right], \quad m_M(\omega) = \mathbf{a}^H (\mathbf{R}_M - \omega\mathbf{I}_M)^{-1}\mathbf{b}.$$

For example,

$$\left\{\mathbf{R}_M^{-1}\right\}_{i,j} = \mathbf{u}_i^H\mathbf{R}_M^{-1}\mathbf{u}_j = \mathbf{u}_i^H\left(\sum_{m=1}^{M}\frac{\mathbf{e}_m\mathbf{e}_m^H}{\lambda_m}\right)\mathbf{u}_j = m_M(0).$$

If we find M,N-consistent estimators of these two functions $b_M(\omega)$, $m_M(\omega)$, we will implicitly find (under some regularity conditions) M,N-consistent estimators of $\varphi(\mathbf{R}_M)$. The problem is, we don't have access to $b_M(\omega)$ and $m_M(\omega)$ because the true covariance matrix $\mathbf{R}_M$ is unknown.

We only have access to the sample covariance matrix, and hence

$$\hat{b}_M(z) = \frac{1}{M}\operatorname{tr}\left[\left(\hat{\mathbf{R}}_M - z\mathbf{I}_M\right)^{-1}\right], \quad \hat{m}_M(z) = \mathbf{a}^H\left(\hat{\mathbf{R}}_M - z\mathbf{I}_M\right)^{-1}\mathbf{b}.$$

The principle of G-estimation (ii)

Recall that, by definition, $\hat{b}_M(z)$ and $\hat{m}_M(z)$ are M, N-consistent estimators of $\bar{b}_M(z)$ and $\bar{m}_M(z)$ respectively when $z \in \mathbb{C}^+$. Now $\bar{b}_M(z)$ is solution to the equation

$$\bar{b}_M(z) = \frac{1}{M} \sum_{m=1}^{M} \frac{1}{\lambda_m \left(1 - c - cz\bar{b}_M(z)\right) - z}$$

which can also be expressed as

$$\bar{b}_M(z) = \frac{f_M(z)}{z} \frac{1}{M} \sum_{m=1}^{M} \frac{1}{\lambda_m - f_M(z)}, \qquad f_M(z) = \frac{z}{1 - c - cz\bar{b}_M(z)}.$$

Observe the similarity with $b_M(\omega)$, the quantity that we need to estimate

$$b_M(\omega) = \frac{1}{M} \sum_{m=1}^{M} \frac{1}{\lambda_m - \omega}.$$

The principle of G-estimation (iii)

Compare the quantity that needs to be estimated $b_M(\omega)$ with $z\bar{b}_M(z)/f_M(z)$:

$$b_M(\omega) = \frac{1}{M}\sum_{m=1}^{M}\frac{1}{\lambda_m - \omega} \longleftrightarrow \frac{1}{M}\sum_{m=1}^{M}\frac{1}{\lambda_m - f_M(z)} = z\frac{\bar{b}_M(z)}{f_M(z)}.$$

Now, if the equation in z

$$\boxed{\omega = f_M(z)}$$

has a single solution, namely $z = f_M^{-1}(\omega)$, then

$$\frac{1}{M}\sum_{m=1}^{M}\frac{1}{\lambda_m - \omega} = \frac{1}{M}\sum_{m=1}^{M}\frac{1}{\lambda_m - f_M(z)}\Bigg|_{z=f_M^{-1}(\omega)} = z\frac{\bar{b}_M(z)}{f_M(z)}\Bigg|_{z=f_M^{-1}(\omega)} = \bar{b}_M(z)\left(1 - c - cz\bar{b}_M(z)\right)\Big|_{z=f_M^{-1}(\omega)}.$$

As a conclusion, $b_M(\omega)$ is M, N-consistently estimated by

$$G_M(\omega) = \hat{b}_M(z)\left(1 - c - cz\hat{b}_M(z)\right)\Big|_{z=\hat{f}_M^{-1}(\omega)}.$$

An example of G-estimator

Assume that we want to estimate $\varphi(\mathbf{R}_M) = \frac{1}{M}\operatorname{tr}\left[\mathbf{R}_M^{-2}\right]$.

- We express $\varphi(\mathbf{R}_M)$ as a function of $b_M(\omega)$, namely

$$\varphi(\mathbf{R}_M) = \frac{1}{M}\operatorname{tr}\left[\mathbf{R}_M^{-2}\right] = \frac{1}{M}\sum_{m=1}^{M}\frac{1}{\lambda_m^2} = \frac{d}{d\omega}\left[\frac{1}{M}\sum_{m=1}^{M}\frac{1}{\lambda_m-\omega}\right]_{\omega=0} = \left.\frac{db_M(\omega)}{d\omega}\right|_{\omega=0}.$$

- The G-estimator is constructing replacing $b_M(\omega)$ with the G-estimator $G(\omega)$, namely

$$\hat{\varphi}(\mathbf{R}_M) = \left.\frac{dG_M(\omega)}{d\omega}\right|_{\omega=0} = \frac{d}{d\omega}\left[\hat{b}_M(z)\left(1-c-cz\hat{b}_M(z)\right)\Big|_{z=\hat{f}_M^{-1}(\omega)}\right]_{\omega=0} =$$

$$= (1-c)^2\frac{1}{M}\operatorname{tr}\left[\hat{\mathbf{R}}_M^{-2}\right] - c(1-c)\left(\frac{1}{M}\operatorname{tr}\left[\hat{\mathbf{R}}_M^{-1}\right]\right)^2.$$

where we have used the chain rule, the inverse function theorem and the fact that $0 = f_M(z)$ has a single solution $z = 0$ whenever $c < 1$.

G-estimation of functions of $m_M(\omega)$

Assume that the function that we need to estimate $\varphi(\mathbf{R}_M)$ can be expressed in terms of the Stieltjes transform $m_M(\omega)$. In this case, the procedure to derive a G-estimator of $\varphi(\mathbf{R}_M)$ boils down to finding a G-estimator of $m_M(\omega)$.

We recall that $|\bar{m}_M(z) - \hat{m}_M(z)| \to 0$ as $M, N \to \infty$ at the same rate, where

$$\bar{m}_M(z) = \frac{f_M(z)}{z}\mathbf{a}^H \left(\mathbf{R}_M - f_M(z)\mathbf{I}_M\right)^{-1} \mathbf{b} = \frac{f_M(z)}{z}\sum_{m=1}^{M} \frac{\mathbf{a}^H \mathbf{e}_m \mathbf{e}_m^H \mathbf{b}}{\lambda_m - f_M(z)}$$

Compare now the above equation with the quantity that needs to be estimated, namely

$$m_M(\omega) = \sum_{m=1}^{M} \frac{\mathbf{a}^H \mathbf{e}_m \mathbf{e}_m^H \mathbf{b}}{\lambda_m - \omega} \longleftrightarrow \sum_{m=1}^{M} \frac{\mathbf{a}^H \mathbf{e}_m \mathbf{e}_m^H \mathbf{b}}{\lambda_m - f_M(z)} = z\frac{\bar{m}_M(z)}{f_M(z)} = \bar{m}_M(z)\left(1 - c - cz\bar{b}_M(z)\right).$$

Hence, if $\omega = f_M(z)$ has a unique solution, then $m_M(\omega)$ is M,N-consistently estimated by

$$H(\omega) = \hat{m}_M(z)\left(1 - c - cz\hat{b}_M(z)\right)\Big|_{z=\hat{f}_M^{-1}(\omega)}.$$

Another example of G-estimator

Assume that we want to estimate $\varphi(\mathbf{R}_M) = \{\mathbf{R}_M^{-1}\}_{i,j}$.

- We express $\varphi(\mathbf{R}_M)$ as a function of $m_M(\omega)$, namely

$$\varphi(\mathbf{R}_M) = \mathbf{u}_i^H \mathbf{R}_M^{-1} \mathbf{u}_j = \sum_{m=1}^{M} \frac{\mathbf{u}_i^H \mathbf{e}_m \mathbf{e}_m^H \mathbf{u}_j}{\lambda_m} = \left[\sum_{m=1}^{M} \frac{\mathbf{u}_i^H \mathbf{e}_m \mathbf{e}_m^H \mathbf{u}_j}{\lambda_m - \omega}\right]_{\omega=0} = m_M(\omega)\big|_{\omega=0}.$$

- The G-estimator is constructing replacing $m_M(\omega)$ with the G-estimator $H(\omega)$, namely

$$\hat{\varphi}(\mathbf{R}_M) = H(\omega)\big|_{\omega=0} = \hat{m}_M(z)\left(1 - c - cz\hat{b}_M(z)\right)\Big|_{z=\hat{f}_M^{-1}(0)} = (1-c)\,\hat{m}_M(0) = (1-c)\left\{\hat{\mathbf{R}}_M^{-1}\right\}_{i,j}.$$

where we have used the fact that $0 = f_M(z)$ has a single solution $z = 0$ whenever $c < 1$.

Yet another example of G-estimator

Assume that we want to estimate $\varphi(\mathbf{R}_M) = \left\{\mathbf{R}_M^{-1/2}\right\}_{i,j}$ (whitening filter).

- Using the integral
$$\frac{1}{\sqrt{x}} = \frac{2}{\pi}\int_0^\infty \frac{1}{x+t^2}dt, \quad x>0,$$
We express $\varphi(\mathbf{R}_M)$ as a function of $m_M(\omega)$, namely
$$\left\{\mathbf{R}_M^{-1/2}\right\}_{i,j} = \frac{2}{\pi}\int_0^\infty \mathbf{u}_i^H\left[\mathbf{R}_M + t^2\mathbf{I}_M\right]^{-1}\mathbf{u}_j dt = \frac{2}{\pi}\int_0^\infty m_M(-t^2)dt$$

- If $c<1$, the equation $f_M(z) = -t^2$ has a unique solution for any $t \in \mathbb{R}$. The G-estimator is constructing replacing $m_M(\omega)$ with the G-estimator $H(\omega)$, namely

$$\hat{\varphi}(\mathbf{R}_M) = \frac{2}{\pi}\int_0^\infty \hat{m}_M(z)\left(1-c-cz\hat{b}_M(z)\right)\Big|_{z=\hat{f}_M^{-1}(-t^2)} dt =$$
$$= \frac{2}{\pi}\int_0^\infty \sum_{m=1}^M \frac{\mathbf{u}_i^H\hat{\mathbf{e}}_m\hat{\mathbf{e}}_m^H\mathbf{u}_j}{\hat{\lambda}_m - \hat{f}_M^{-1}(-t^2)}\left(1-c-cz\frac{1}{M}\sum_{k=1}^M \frac{1}{\hat{\lambda}_k - \hat{f}_M^{-1}(-t^2)}\right)dt$$

Solutions to the equation $\omega = f_M(z)$

It turns out that $f_M(z)$ can be obtained as a particular root of the polynomial equation in f

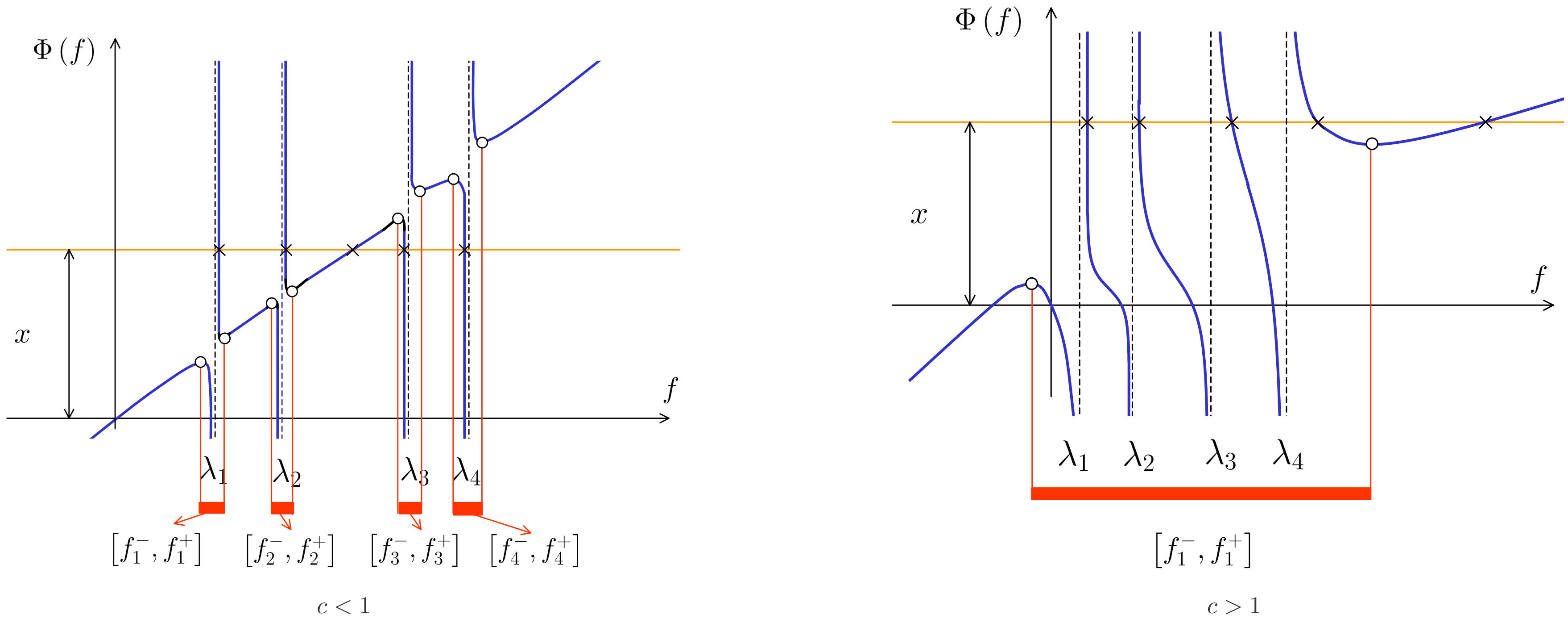
$$\Phi(f) = f\left(1 - \frac{c}{M}\sum_{m=1}^{M}\frac{\lambda_m}{\lambda_m - f}\right) = z.$$

- If there exists a root $f \in \mathbb{C}^+$, it is unique and we choose $f_M(z) = f$.
- Otherwise, all the roots are real, but there is only one root such that $\Phi'(f) > 0$ or, equivalently,

$$\frac{1}{M}\sum_{m=1}^{M}\left(\frac{\lambda_m}{\lambda_m - f}\right)^2 \leq \frac{1}{c}$$

  and we choose $f_M(z) = f$.

Solutions to the equation $\omega = f_M(x)$ for $\omega \in \mathbb{R}$



The equation has a unique solution when $\omega \in \mathbb{R} \backslash \left[f_1^-, f_1^+\right] \cup \ldots \cup \left[f_Q^-, f_Q^+\right]$.

Application to subspace-based DoA estimation. Introduction and signal model

We consider DoA detection based on subspace approaches (MUSIC), that exploit the orthogonality between signal and noise subspaces.

Consider a set of K sources impinging on an array of M sensors/antennas. We work with a fixed number of snapshots N,

$$\{\mathbf{y}(1), \ldots, \mathbf{y}(N)\}$$

assumed i.i.d., with zero mean and covariance $\mathbf{R}_M$.

The true spatial covariance matrix can be described as

$$\mathbf{R}_M = \mathbf{S}(\Theta)\,\Phi_S \mathbf{S}(\Theta)^H + \sigma^2 \mathbf{I}_M$$

where $\mathbf{S}(\Theta)$ is an $M \times K$ matrix that contains the steering vectors corresponding to the K different sources,

$$\mathbf{S}(\Theta) = \left[\begin{array}{cccc} \mathbf{s}(\theta_1) & \mathbf{s}(\theta_2) & \cdots & \mathbf{s}(\theta_K) \end{array}\right]$$

and σ^2 is the background noise power.

Introduction and signal model (ii)

The eigendecomposition of $\mathbf{R}_M$ allows us to differentiate between signal and noise subspaces:

$$\mathbf{R}_M = \left[\begin{array}{cc} \mathbf{E}_S & \mathbf{E}_N \end{array}\right] \left[\begin{array}{cc} \mathbf{\Lambda}_S & \mathbf{0} \\ \mathbf{0} & \sigma^2 \mathbf{I}_{M-K} \end{array}\right] \left[\begin{array}{cc} \mathbf{E}_S & \mathbf{E}_N \end{array}\right]^H .$$

It turns out that $\mathbf{E}_N^H \mathbf{s}(\theta_k) = \mathbf{0}$, $k = 1 \ldots K$.

Since $\mathbf{R}_M$ is unknown, one must work with the sample covariance matrix

$$\hat{\mathbf{R}}_M = \frac{1}{N} \sum_{n=1}^{N} \mathbf{y}(n) \mathbf{y}^H(n).$$

The MUSIC algorithm uses the sample noise eigenvectors, and searches for the deepest local minima of the cost function

$$\eta_{\text{MUSIC}}(\theta) = \mathbf{s}^H(\theta) \hat{\mathbf{E}}_N \hat{\mathbf{E}}_N^H \mathbf{s}(\theta) .$$

It is interesting to investigate the behavior of $\eta_{\text{MUSIC}}(\theta)$ when M, N have the same order of magnitude.

Noise sample eigenvalue cluster separation assumption

In our asymptotic approach, we assume that there exists separation between noise and signal eigenvalue clusters in the asymptotic sample eigenvalue distribution.

This implies that there is a minimum number of samples per antenna:

$$\left(\frac{N}{M}\right) > \frac{1}{M}\sum_{m=1}^{M}\left(\frac{\lambda_m}{\lambda_m - \xi}\right)^2$$

where ξ is the smallest real valued solution to the following equation:

$$\frac{1}{M}\sum_{m=1}^{M}\frac{\lambda_m^2}{(\lambda_m - \xi)^3} = 0.$$

Asymptotic behavior of MUSIC

The MUSIC algorithm suffers from the **breakdown effect**. The performance breaks down when the number of samples or the SNR falls below a **certain threshold**. Cause: $\hat{\mathbf{E}}_N$ is not a very good estimator of $\mathbf{E}_N$ when M, N have the same order of magnitude.

The performance breakdown effect can be easily analyzed using random matrix theory, especially under a noise eigenvalue separation assumption: $|\eta_{\text{MUSIC}}(\theta) - \bar{\eta}_{\text{MUSIC}}(\theta)| \to 0$

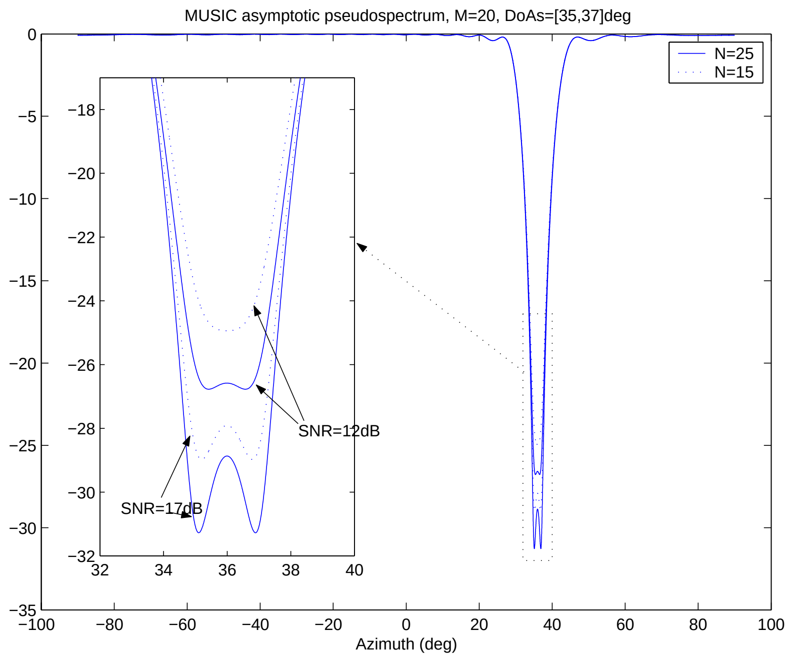
$$\bar{\eta}_{\text{MUSIC}}(\theta) = \mathbf{s}^H(\theta) \left(\sum_{k=1}^{M} w(k) \mathbf{e}_k \mathbf{e}_k^H \right) \mathbf{s}(\theta)$$

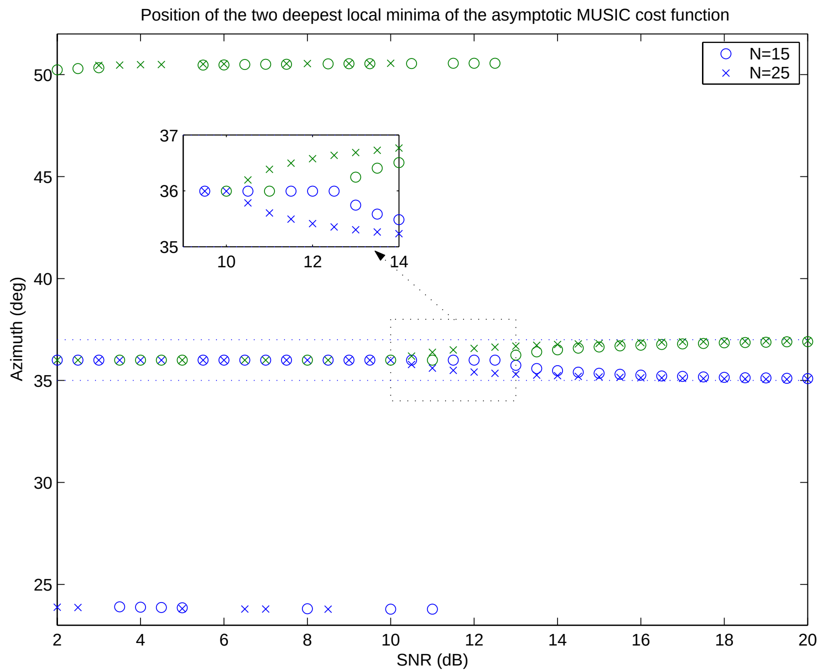
$$w(k) = \begin{cases} 1 - \frac{1}{M-K} \sum_{r=M-K+1}^{M} \left(\frac{\sigma^2}{\lambda_r - \sigma^2} - \frac{\mu_1}{\lambda_r - \mu_1} \right) & k \leq M-K \\ \frac{\sigma^2}{\lambda_k - \sigma^2} - \frac{\mu_1}{\lambda_k - \mu_1} & k > M-K \end{cases}$$

where $\{\mu_r,\ r = 1, \ldots, M\}$ are the solutions to $\frac{1}{M} \sum_{r=1}^{M} \frac{\lambda_r}{\lambda_r - \mu} = \frac{1}{c}$.

Asymptotic behavior of MUSIC: an example

We consider a scenario with two sources impinging on a ULA ($d/\lambda_c = 0.5$, $M = 20$) from DoAs: $35°$, $37°$.





M, N-consistent subspace detection: G-MUSIC

We propose to use an M, N-consistent estimator of the cost function $\eta(\theta) = \mathbf{s}^H(\theta)\mathbf{E}_N\mathbf{E}_N^H\mathbf{s}(\theta)$:

$$\eta_{\text{G-MUSIC}}(\theta) = \mathbf{s}^H(\theta)\left(\sum_{k=1}^{M}\phi(k)\hat{\mathbf{e}}_k\hat{\mathbf{e}}_k^H\right)\mathbf{s}(\theta)$$

$$\phi(k) = \begin{cases} 1 + \sum_{r=M-K+1}^{M}\left(\frac{\hat{\lambda}_r}{\hat{\lambda}_k - \hat{\lambda}_r} - \frac{\hat{\mu}_r}{\hat{\lambda}_k - \hat{\mu}_r}\right) & k \le M-K \\ -\sum_{r=1}^{M-K}\left(\frac{\hat{\lambda}_r}{\hat{\lambda}_k - \hat{\lambda}_r} - \frac{\hat{\mu}_r}{\hat{\lambda}_k - \hat{\mu}_r}\right) & k > M-K \end{cases}$$

where now $\hat{\mu}_1, \ldots, \hat{\mu}_M$ are the solutions to the equation

$$\frac{1}{M}\sum_{k=1}^{M}\frac{\hat{\lambda}_k}{\hat{\lambda}_k - \hat{\mu}} = \frac{1}{c}.$$

Sketch of the derivation

The quantity that needs to be estimated can be written as

$$\eta(\theta) = \mathbf{s}^H(\theta)\,\mathbf{E}_N\mathbf{E}_N^H\mathbf{s}(\theta) = \|\mathbf{s}(\theta)\|^2 - \mathbf{s}^H(\theta)\,\mathbf{E}_S\mathbf{E}_S^H\mathbf{s}(\theta)$$
$$\mathbf{s}^H(\theta)\,\mathbf{E}_S\mathbf{E}_S^H\mathbf{s}(\theta) = \frac{1}{2\pi\,\mathrm{j}}\oint_{\mathcal{C}^-} m_M(\omega)d\omega$$

where

$$m_M(\omega) = \mathbf{s}^H(\theta)\left(\mathbf{R}_M - \omega\mathbf{I}_M\right)^{-1}\mathbf{s}(\theta)$$

and $\mathcal{C}^-$ is a negatively oriented contour enclosing only the signal eigenvalues $\{\lambda_{M-K+1},\ldots,\lambda_M\}$.

The idea of the derivation is based on obtaining a proper parametrization of the contour $\mathcal{C}^-$.

Sketch of the derivation (ii)

The derivation is based on the following random matrix theory result. Let

$$\hat{m}_M(z) = \mathbf{s}^H(\theta)\left(\hat{\mathbf{R}}_M - z\mathbf{I}_M\right)^{-1}\mathbf{s}(\theta), \quad \hat{b}_M(z) = \frac{1}{M}\operatorname{tr}\left[\left(\hat{\mathbf{R}}_M - z\mathbf{I}_M\right)^{-1}\right]$$

It turns out that

$$\left|\hat{m}_M(z) - \bar{m}_M(z)\right| \to 0, \quad \left|\hat{b}_M(z) - \bar{b}_M(z)\right| \to 0$$

almost surely for all $z \in \mathbb{C}^+$ as $M, N \to \infty$ at the same rate, where $\bar{b}_M(z) = b$ is the unique solution to the following equation in the set $\{b \in \mathbb{C} : -(1-c)/z + cb \in \mathbb{C}^+\}$:

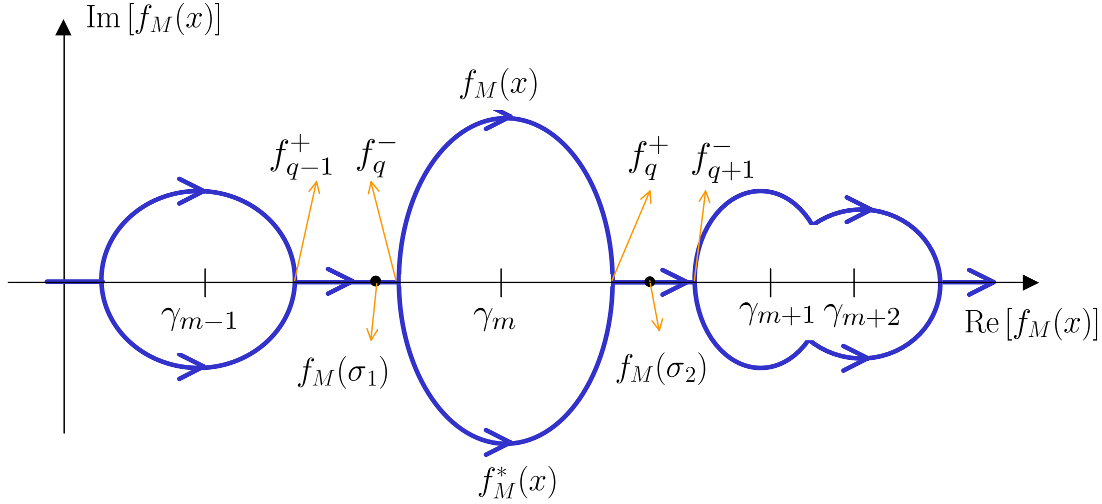
$$b = \frac{1}{M}\sum_{m=1}^{M}\frac{1}{\lambda_m(1-c-czb)-z}$$

and

$$\bar{m}_M(z) = \frac{f_M(z)}{z}\mathbf{s}^H(\theta)\left(\mathbf{R}_M - f_M(z)\mathbf{I}_M\right)^{-1}\mathbf{s}(\theta), \quad f_M(z) = \frac{z}{1-c-cz\bar{b}_M(z)}.$$

Sketch of the derivation (iii)

It turns out that, surprisingly enough, $f_M(x)$ gives us a valid parametrization of $\mathcal{C}^-$:



$$\eta(\theta) = \mathbf{s}^H(\theta)\,\mathbf{E}_N\mathbf{E}_N^H\mathbf{s}(\theta) = \|\mathbf{s}(\theta)\|^2 - \frac{1}{\pi}\,\mathrm{Im}\left[\int_{\sigma_1}^{\sigma_2} m_M(f_M(x))f_M'(x)dx\right]$$

Sketch of the derivation (iv)

From this point, we can express the integral in terms of $\bar{m}_M(z)$ and $\bar{b}_M(z)$, namely

$$\eta(\theta) = \|\mathbf{s}(\theta)\|^2 - \frac{1}{2\pi\mathrm{j}} \oint_{\mathcal{C}^-} m_M(\omega) d\omega = \|\mathbf{s}(\theta)\|^2 - \frac{1}{\pi} \mathrm{Im}\left[\int_{\sigma_1}^{\sigma_2} \bar{m}_M(x) \frac{1 - c + cx^2 \bar{b}_M'(x)}{1 - c - cx\bar{b}_M(x)} dx \right]$$

and we only need to replace $\bar{m}(z)$ and $\bar{b}(z)$ with their M, N-consistent estimates:

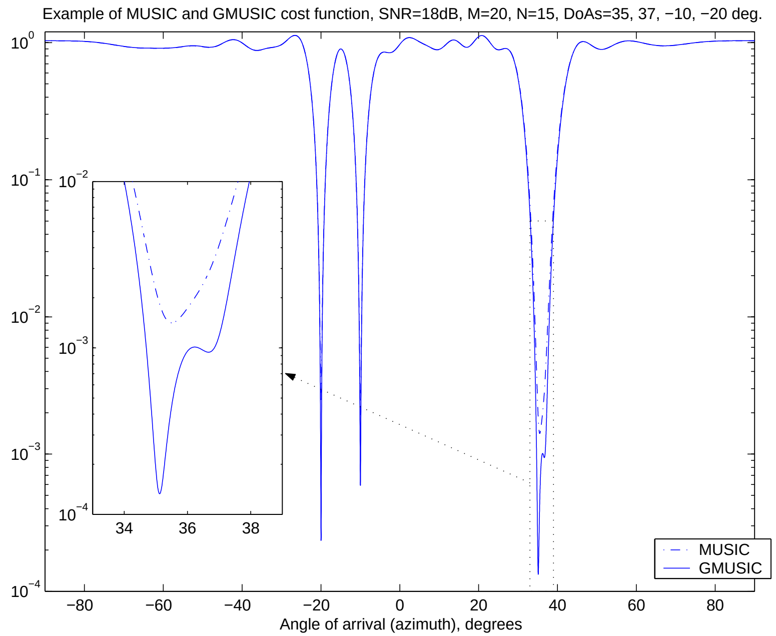
$$\eta_{\text{G-MUSIC}}(\theta) = \|\mathbf{s}(\theta)\|^2 - \frac{1}{\pi} \mathrm{Im}\left[\int_{\sigma_1}^{\sigma_2} \hat{m}_M(x) \frac{1 - c + cx^2 \hat{b}_M'(x)}{1 - c - cx\hat{b}_M(x)} dx \right].$$

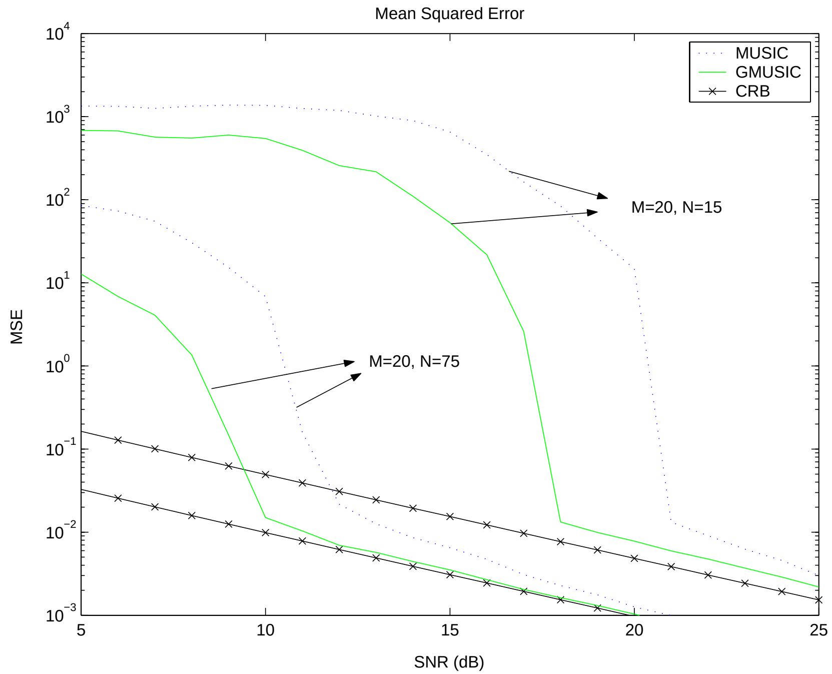
Integrating this expression, we get to the proposed estimator:

$$\eta_{\text{G-MUSIC}}(\theta) = \mathbf{s}^H(\theta) \left(\sum_{k=1}^{M} \phi(k) \hat{\mathbf{e}}_k \hat{\mathbf{e}}_k^H \right) \mathbf{s}(\theta)$$

Performance evaluation MUSIC vs. G-MUSIC

Comparative evaluation of MUSIC and G-MUSIC via simulations. Scenarios with four $(-20°, -10°, 35°, 37°)$ and two $(35°, 37°)$ sources respectively, ULA $(M = 20,\ d/\lambda_c = 0.5)$.





Some references

J. Silverstein, "Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices," Journal of Multivariate Analysis, vol. 5, pp. 331–339, 1995.

V. Girko, An Introduction to Statistical Analysis of Random Arrays. The Netherlands: VSP, 1998.

J. Silverstein and P. Combettes, "Signal detection via spectral theory of large dimensional random matrices," IEEE Trans. on Signal Processing, vol. 40, pp. 2100–2105, Aug. 1992.

X. Mestre, "Improved estimation of eigenvalues of covariance matrices and their associated subspaces using their sample estimates," 2006. Preprint. Available at http://www.cttc.es/drafts/cttc-rc-2006-004.pdf.

Z. Bai and J. Silverstein, "No eigenvalues outside the support of the limiting spectral distribution of large dimensional sample covariance matrices," Annals of probability, vol. 26, pp. 316–345, 1998.

Z. Bai and J. Silverstein, "Exact Separation of Eigenvalues of Large Dimensional Sample Covariance Matrices," Annals of probability, vol. 27, pp. 1536-1555, 1999.

Thank you for your attention!!!