Line Spectrum Estimation - Methods

V. Rajbabu
rajbabu@ee.iitb.ac.in

Department of Electrical Engineering
Indian Institute of Technology Bombay

17 September 2012

Pisarenko Harmonic Decomposition (PHD)

- $x(n)$ is assumed to be a sum of p complex exponentials (uniformly distributed phase) in white noise and p is known

$$x(n) = \sum_{i=1}^{p} A_i e^{jn\omega_i} + w(n).$$

- Corresponding autocorrelation coefficients are

$$r_x(k) = \sum_{i=1}^{p} P_i e^{jk\omega_i} + \sigma_w^2 \delta(k), \quad k = 0, \ldots, M-1,$$

  where $P_i = |A_i|^2$ is the power in the i-th component
- Corresponding autocorrelation matrix $\mathbf{R}_x$ $((p+1) \times (p+1))$ is either known or can be estimated

$$\mathbf{R}_x = \mathbf{R}_{\text{sig}} + \mathbf{R}_{\text{noise}} = \sum_{i=1}^{p} P_i \mathbf{e}_i \mathbf{e}_i^H + \sigma_w^2 \mathbf{I},$$

Pisarenko Harmonic Decomposition (PHD)

- Dimension of noise subspace is one and is spanned by the eigenvector $\mathbf{v}_{min}$ corresponding to λ_{min}
- From orthogonality of signal subspace and noise eigenvector space

$$\mathbf{e}_i^H \mathbf{v}_{min} = \sum_{k=0}^{p} v_{min}(k) e^{-jk\omega} = 0; \qquad i = 1, \ldots, p.$$

where

$$\mathbf{e}_i = \left[1, e^{j\omega_i}, e^{j2\omega_i}, \ldots, e^{jp\omega_i}\right]^T$$

Pisarenko Harmonic Decomposition (PHD)

- DTFT of the noise vector

$$V_{min}(e^{j\omega}) = \sum_{k=0}^{p} v_{min}(k)e^{-jk\omega}$$

- This being zero at each of the complex exponential frequencies $\omega_i, i = 1, \dots, p$.
- Can be considered as an eigenfilter that has p zeros on the unit circle

$$V_{min}(z) = \sum_{k=0}^{p} v_{min}(k)z^{-k} = \prod_{k=1}^{p} \left(1 - e^{j\omega_k} z^{-1}\right).$$

and the roots of this provide the frequencies of the complex exponentials

Pisarenko Harmonic Decomposition (PHD)

- An alternative approach is to search for peak locations of the frequency estimation function ("pseudospectrum")

$$\hat{P}_{PHD}(e^{j\omega}) = \frac{1}{\left|\mathbf{e}^H \mathbf{v}_{min}\right|^2}.$$

Pisarenko Harmonic Decomposition (PHD)

- Having ω_i need to estimate corresponding power P_i
- Eigenvectors of $\mathbf{R}_x$ are normalized, i.e., $\mathbf{v}_i^H \mathbf{v}_i = 1$
- For $i = 1, \cdots, p$, we have

$$
\begin{aligned}
\mathbf{R}_x \mathbf{v}_i &= \lambda_i \mathbf{v}_i \\
\mathbf{v}_i^H \mathbf{R}_x \mathbf{v}_i &= \lambda_i \mathbf{v}_i^H \mathbf{v}_i = \lambda_i.
\end{aligned}
$$

- Using the signal and noise decomposition of $\mathbf{R}_x$, we have

$$
\mathbf{v}_i^H \mathbf{R}_x \mathbf{v}_i = \mathbf{v}_i^H \left\{ \sum_{k=1}^{p} P_k \mathbf{e}_k \mathbf{e}_k^H + \sigma_w^2 \mathbf{I}, \right\} \mathbf{v}_i,
$$

which

$$
\Longrightarrow \sum_{k=1}^{p} P_k \left| \mathbf{e}_k^H \mathbf{v}_i \right|^2 = \lambda_i - \sigma_w^2, \qquad i = 1, \ldots, p.
$$

Pisarenko Harmonic Decomposition (PHD)

- We have $\left|\mathbf{e}_k^H \mathbf{v}_i\right|^2$ corresponding to the squared magnitude of the DTFT of the signal subspace eigenvector $\mathbf{v}_i$ evaluated at frequency ω_k,

$$\left|\mathbf{e}_k^H \mathbf{v}_i\right|^2 = \left|V_i(e^{j\omega_k})\right|^2.$$

- Using this we have

$$\sum_{k=1}^{p} P_k \left|V_i(e^{j\omega_k})\right|^2 = \lambda_i - \sigma_w^2, \qquad i = 1, \ldots, p.$$

- Which can be solved for powers P_k using p linear equations, as shown in matrix form

$$\begin{bmatrix} \left|V_1(e^{j\omega_1})\right|^2 & \left|V_1(e^{j\omega_2})\right|^2 & \cdots & \left|V_1(e^{j\omega_p})\right|^2 \\ \vdots & \vdots & \vdots & \vdots \\ \left|V_p(e^{j\omega_1})\right|^2 & \left|V_p(e^{j\omega_2})\right|^2 & \cdots & \left|V_p(e^{j\omega_p})\right|^2 \end{bmatrix} \begin{bmatrix} P_1 \\ \vdots \\ P_p \end{bmatrix} = \begin{bmatrix} \lambda_1 - \sigma_w^2 \\ \vdots \\ \lambda_p - \sigma_w^2 \end{bmatrix}.$$

MUSIC - MUltiple SIgnal Classification

- Improvement over PHD for frequency estimation
- In PHD - projection of signal vectors onto a single noise eigenvector
- MUSIC - projection of signal vectors onto the entire noise subspace
- Signal $x(n)$ is a mixture of p complex exponentials in white noise with variance σ_w^2 and has a $M \times M$ autocorrelation matrix $\mathbf{R}_x$ with $M > p + 1$
  - in contrast to PHD where $M = p + 1$

MUSIC - MUltiple SIgnal Classification

- We can divide the eigenvectors of $\mathbf{R}_x$ into two groups;
    - p signal EVecs corresponding to p largest EVals
    - $M - p$ noise EVecs that have an EVals of σ_w^2
- $M - p$ smallest EVals are approximately equal to σ_w^2

$$\hat{\sigma}_w^2 = \frac{1}{M-p} \sum_{k=p+1}^{M} \lambda_k.$$

MUSIC - MUltiple SIgnal Classification

- Taking the DTFT of the noise subspace eigenvectors, the eigenfilters can be obtained as

$$V_i(z) = \sum_{k=0}^{M-1} v_i(k) z^{-k}, \qquad i = p+1, \ldots, M.$$

- Has $M-1$ roots with p of them lying on the unit circle
- The location of those on the unit circle correspond to the frequencies of the complex exponentials.
- The pseudospectrum (eigen spectrum) of the signal

$$\left| V_i(e^{j\omega}) \right|^2 = \frac{1}{\left| \sum_i^{M-1} v_i(k) e^{-jk\omega} \right|^2}, \tag{1}$$

  has peaks at the ω_i corresponding to the complex exponentials.
- Uses only one of the noise eigenvectors $\mathbf{v}_i$.

MUSIC - MUltiple SIgnal Classification

- Issues with using single noise eigenvector
  - $M - p - 1$ zeros corresponding to noise might lie close to unit circle
  - p zeros corrsponding to signal might not lie exactly on unit circle
- This leads to spurious peaks
- Spurious peaks can be reduced by averaging

$$\hat{P}_{MU}(e^{j\omega}) = \frac{1}{\sum_{i=p+1}^{M} \left|\mathbf{e}^H \mathbf{v}_i\right|^2}.$$

- Location of the p largest peaks in $\hat{P}_{MU}(e^{j\omega})$ correspond to the frequencies of the complex exponentials
- Corresponding power P_i can be estimated as in PHD method

MUSIC - MUltiple SIgnal Classification

Alternate view

- Consider the projection matrix on to the noise space

$$\mathbf{P}_{\text{noise}} = \mathbf{V}_{\text{noise}}\mathbf{V}_{\text{noise}}^H$$

- The denominator term can be considered as

$$\mathbf{e}^H\mathbf{P}_{\text{noise}}\mathbf{e} = \mathbf{e}^H\mathbf{V}_{\text{noise}}\mathbf{V}_{\text{noise}}^H\mathbf{e}$$

- This goes to zero for $\mathbf{e}_i$ that are orthogonal to the noise subspace
- The MUSIC pseudospectrum can be represented as

$$\hat{P}_{MU}(e^{j\omega}) = \frac{1}{\mathbf{e}^H\mathbf{V}_{\text{noise}}\mathbf{V}_{\text{noise}}^H\mathbf{e}}.$$

Root MUSIC

- Obtain the roos of an equivalent polynomial (corresponding to the MUSIC method)

$$D(z) = \sum_{i=p+1}^{M} V_i(z) V_i^*(1/z^*),$$

- p roots that are closest to the unit circle correspond to the complex exponentials

Eigenvector Method

- Similar to MUSIC, differs only by a constant

$$\hat{P}_{EV}(e^{j\omega}) = \frac{\sigma_w^2}{\sum_{i=p+1}^{M} \left|\mathbf{e}^H \mathbf{v}_i\right|^2} = \frac{1}{\frac{1}{\lambda_i}\sum_{i=p+1}^{M} \left|\mathbf{e}^H \mathbf{v}_i\right|^2},$$

  where $\lambda_i = \sigma_w^2$ are the eigenvalues corresponding to the eigenvectors $\mathbf{v}_i, i = p+1, \cdots, M$
- Differs from MUSIC: in practise the estimated small λ_is all need not be equal

Minimum-norm Algorithm

- Choose a single noise vector **a** with minimum norm to define the pseudospectrum
- Compare to PHD or MUSIC
  - PHD uses single noise eigenvector
  - MUSIC uses all noise eigenvectors

Motivation for choice of **a**

- **a** lies in the noise subspace - ensures p roots of $A(z)$ lie on the unit circle
- **a** has minimum norm - ensures that spurious roots of $A(z)$ lie inside the unit circle, i.e., $|z_k| < 1$. i.e., $A(z)$ is a minimum-phase polynomial. Further these roots are approximately uniformly distributed around the unit-circle in sectors where the signal roots are not present
- the first element of **a** is unity - to ensure that the minimum norm solution is not the zero vector

Minimum-norm Algorithm

- The pseudospectrum is defined as

$$\hat{P}_{MN}(e^{j\omega}) = \frac{1}{\left|\mathbf{e}^H \mathbf{a}\right|^2}.$$

  where **a** is constrained to lie in the noise subspace and **e** is a complex exponential
- The peak locations of this function correspond to the frequencies of the complex exponentials
- Considering the z-domain polynomial corresponding to **a**

$$A(z) = \sum_{k=0}^{M-1} a(k) z^{-k} = \underbrace{\prod_{k=1}^{p} \left(1 - e^{j\omega_k} z^{-1}\right)}_{\text{roots on unit-circle}} \underbrace{\prod_{k=p+1}^{M-1} \left(1 - z_k z^{-1}\right)}_{\text{roots not on unit-circle}}$$

- Goal: determine the noise subspace vector **a** that minimizes the effect of the spurious zeros

Minimum-norm Algorithm

- The noise subspace vector **a** can be obtained by imposing three constraints
  - C1 **a** lies in the noise subspace
  - C2 **a** has minimum norm
  - C3 First element of **a** is unity
- C1 can be expressed using the noise projection matrix, $\mathbf{P}_{\text{n}} = \mathbf{V}_{\text{noise}}\mathbf{V}_{\text{noise}}^{H}$, which projects an arbitrary vector y onto the noise subspace

$$\mathbf{a} = \mathbf{P}_{\text{n}}\mathbf{y}.$$

- C3 can be represented using

$$\mathbf{a}^{H}\mathbf{u}_{1} = 1$$

where $\mathbf{u}_{1} = [1, 0, \cdots, 0]^{T}$.

Minimum-norm Algorithm

- Combining C1 and C3, we have

$$\mathbf{y}^H(\mathbf{P}_n^H \mathbf{u}_1) = 1.$$

- Objective is to minimize norm of **a**

$$\begin{aligned} \|\mathbf{a}\|^2 &= \|\mathbf{P}_n \mathbf{y}\|^2 \\ &= \mathbf{y}^H \left(\mathbf{P}_n^H \mathbf{P}_n\right) \mathbf{y} \\ &= \mathbf{y}^H \mathbf{P}_n \mathbf{y} \quad (\because \mathbf{P}_n \text{ is Hermitian, idempotent }, \mathbf{P}_n^2 = \mathbf{P}_n). \end{aligned}$$

- Minimizing the norm of **a** can be formulated as a constrained optimization problem

$$\min \mathbf{y}^H \mathbf{P}_n \mathbf{y} \quad \text{subject to} \quad \mathbf{y}^H \left(\mathbf{P}_n^H \mathbf{u}_1\right) = 1.$$

and the solution **y** can be used to obtain **a**

Gradient of a Scalar Quantity

- Let $\mathcal{Q}$ be a scalar quantity depending on the real vector **a**. The gradient is defined as

$$\nabla_{\mathbf{a}} \mathcal{Q} \overset{\text{def}}{=} \begin{bmatrix} \frac{\partial \mathcal{Q}}{\partial a_1} \\ \frac{\partial \mathcal{Q}}{\partial a_2} \\ \vdots \\ \frac{\partial \mathcal{Q}}{\partial a_N} \end{bmatrix}$$

Table : Gradient relation for real vectors and symmetric **B**

Quantity $\mathcal{Q}$	$\mathbf{a}^T\mathbf{b}$	$\mathbf{b}^T\mathbf{a}$	$\mathbf{a}^T\mathbf{B}\mathbf{a}$
Gradient $\nabla_{\mathbf{a}}\mathcal{Q}$	$\mathbf{b}$	$\mathbf{b}$	$2\mathbf{B}\mathbf{a}$

Gradient of a Scalar Quantity

- Let $\mathcal{Q}$ be a scalar quantity depending on the complex vector $\mathbf{a} = \mathbf{a}_r + j\mathbf{a}_i$. The gradient is defined as

$$\nabla_{\mathbf{a}} \mathcal{Q} \stackrel{\text{def}}{=} \frac{1}{2}(\nabla_{\mathbf{a}_r} \mathcal{Q} - j \nabla_{\mathbf{a}_i} \mathcal{Q})$$

and

$$\nabla^*_{\mathbf{a}} \mathcal{Q} \stackrel{\text{def}}{=} \frac{1}{2}(\nabla_{\mathbf{a}_r} \mathcal{Q} + j \nabla_{\mathbf{a}_i} \mathcal{Q})$$

Table : Gradient relation for complex vectors and Hermitian **B**

Quantity $\mathcal{Q}$	$\mathbf{a}^H\mathbf{b}$	$\mathbf{b}^H\mathbf{a}$	$\mathbf{a}^H\mathbf{B}\mathbf{a}$
Gradient $\nabla_{\mathbf{a}} \mathcal{Q}$	0	$\mathbf{b}^*$	$(\mathbf{B}\mathbf{a})^*$
Gradient $\nabla_{\mathbf{a}^*} \mathcal{Q}$	$\mathbf{b}$	0	$\mathbf{B}\mathbf{a}$

Minimum-norm Algorithm

- Solution to the constrained optimization problem can be obtained using a Lagrange multiplier λ and minimizing

$$Q(\mathbf{y}, \lambda) = \mathbf{y}^H \mathbf{P}_\mathrm{n} \mathbf{y} + \lambda(1 - \mathbf{y}^H(\mathbf{P}_\mathrm{n}^H \mathbf{u}_1)).$$

- Setting the gradient of $Q(\mathbf{v}, \lambda)$ w.r.t $\mathbf{y}^*$ equal to zero

$$\nabla_{\mathbf{y}^*} Q(\mathbf{y}, \lambda) = \mathbf{P}_n \mathbf{y} - \lambda \left(\mathbf{P}_\mathrm{n}^H \mathbf{u}_1\right) = 0$$
$$\implies \mathbf{y} = \lambda \mathbf{P}_n^{-1} \left(\mathbf{P}_\mathrm{n}^H \mathbf{u}_1\right) = \lambda \mathbf{u}_1.$$

- λ can be obtained by setting $\nabla_\lambda Q(\mathbf{y}, \lambda)$ equal to zero

$$1 - \mathbf{y}^H \left(\mathbf{P}_n^H \mathbf{u}_1\right) = 0 \implies \lambda = \frac{1}{\mathbf{u}_1^H \mathbf{P}_n \mathbf{u}_1}.$$

- Hence the minimum-norm solution is

$$\mathbf{a} = \mathbf{P}_n \mathbf{y} = \lambda \mathbf{P}_n \mathbf{u}_1 = \frac{\mathbf{P}_n \mathbf{u}_1}{\mathbf{u}_1^H \mathbf{P}_n \mathbf{u}_1} = \frac{\left(\mathbf{V}_n \mathbf{V}_n^H\right) \mathbf{u}_1}{\mathbf{u}_1^H \left(\mathbf{V}_n \mathbf{V}_n^H\right) \mathbf{u}_1}.$$

ESPRIT Method

ESTimation of Signal Parameters by Rotational Invariance Technique

- Exploits invariance principle that exists for discrete sequences
- Consider the multiple complex signal model

$$x(n) = \sum_{i=1}^{p} A_i e^{jn\omega_i} + w(n)$$

and considering $n = 0, \ldots, M-1$, we have

$$\mathbf{x} = \sum_{i=1}^{p} A_i \mathbf{e}_i + \mathbf{w},$$

where $\mathbf{e}_i = \begin{bmatrix} 1 & e^{j\omega_i} & \cdots & e^{j(M-1)\omega_i} \end{bmatrix}^T$.

ESPRIT Method

- Consider $\mathbf{x}'$ a translated version of $\mathbf{x}$ i.e.,

$$\mathbf{x} = \begin{bmatrix} x(0) & x(1) & \cdots & x(M-1) \end{bmatrix}$$
$$\mathbf{x}' = \begin{bmatrix} x(1) & x(2) & \cdots & x(M) \end{bmatrix}$$
$$\Longrightarrow \mathbf{x}' = \sum_{i=1}^{p} A_i \mathbf{e}'_i + \mathbf{w}',$$

  where

$$\mathbf{e}'_i = \begin{bmatrix} e^{j\omega_i} \\ e^{j2\omega_i} \\ \dots \\ e^{j(N)\omega_i} \end{bmatrix} = e^{j\omega_i}\mathbf{e}.$$

- Translational invariance between the two data sets leads to a rotatinal invariance between the two corresponding signal subspaces

ESPRIT Method

- The cross-correlation matrix for the vectors $\mathbf{x}$ and $\mathbf{x}'$ is

$$\mathbf{R}_{\mathbf{xx}'} = \sum_{i=1}^{p} P_i e^{-j\omega_i} \mathbf{e}_i \mathbf{e}_i^H + \sigma_w^2 \mathbf{D}_{-1}$$

or

$$\mathbf{R}_{\mathbf{xx}'} = \mathbf{E}\mathbf{P}\mathbf{\Phi}^H \mathbf{E}^H + \sigma_w^2 \mathbf{D}_{-1}$$

where

$$\mathbf{\Phi} = \begin{bmatrix} e^{j\omega_1} & 0 & \cdots & 0 \\ 0 & e^{j\omega_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{j\omega_p} \end{bmatrix}, \mathbf{D}_{-1} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ -1 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & -1 & 0 \end{bmatrix}$$

- Obtaining $\mathbf{\Phi}$ will give us the required frequencies

ESPRIT Method

- ESPRIT begins by forming the matrices

$$\mathbf{R}_s \stackrel{\text{def}}{=} \mathbf{R}_x - \sigma_w^2 \mathbf{I} = \mathbf{E}\mathbf{P}\mathbf{E}^H$$

  and

$$\mathbf{R}_{ss'} \stackrel{\text{def}}{=} \mathbf{R}_{xx'} - \sigma_w^2 \mathbf{D}_{-1} = \mathbf{E}\mathbf{P}\mathbf{\Phi}^H \mathbf{E}^H$$

- As in MUSIC σ_w^2 the smallest eigenvalue of $\mathbf{R}_x$ can be computed and used in the subtraction
- The values $e^{j\omega_k}$ are the eigenvalues of the generalized eigenvalue problem

$$\mathbf{R}_s \dot{\mathbf{e}} = \dot{\lambda} \mathbf{R}_{ss'} \dot{\mathbf{e}}$$

ESPRIT Method

- To show $e^{j\omega_k}$ are eigenvalues of $\mathbf{R}_s\dot{\mathbf{e}} = \dot{\lambda}\mathbf{R}_{ss'}\dot{\mathbf{e}}$

$$
\begin{aligned}
\mathbf{R}_s\dot{\mathbf{e}} - \dot{\lambda}\mathbf{R}_{ss'}\dot{\mathbf{e}} &= \mathbf{EP}(\mathbf{I} - \dot{\lambda}\mathbf{\Phi})\mathbf{E}^H\dot{\mathbf{e}} \\
&= \mathbf{EP}\begin{bmatrix} 1-\dot{\lambda}e^{-j\omega_1} & 0 & \cdots & 0 \\ 0 & 1-\dot{\lambda}e^{-j\omega_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1-\dot{\lambda}e^{-j\omega_p} \end{bmatrix}\mathbf{E}^H\dot{\mathbf{e}}
\end{aligned}
$$

- Since $\dot{\lambda} = e^{j\omega_k}$ reduces the rank of the matrix, they are the generalized eigenvalues

ESPRIT Method

- Let us consider the p complex exponential signal matrix $\mathbf{E}$
- We can form two matrices $\mathbf{E}_1$ and $\mathbf{E}_2$ that are related by a rotation

$$\begin{aligned} \mathbf{E}_1 &= [\mathbf{I}_{M-1} \quad \mathbf{0}]\mathbf{E} \qquad (M-1 \times p) \\ \mathbf{E}_2 &= [\mathbf{0} \quad \mathbf{I}_{M-1}]\mathbf{E} \qquad (M-1 \times p) \end{aligned}$$

  where $\mathbf{I}_{N-1}$ is $(M-1) \times (M-1)$ identity matrix
- We have $\mathbf{E}_2 = \mathbf{E}_1 \mathbf{D}$ where

$$\mathbf{D} = \begin{bmatrix} e^{j\omega_1} & 0 & \cdots & 0 \\ 0 & e^{j\omega_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{j\omega_p} \end{bmatrix}$$

  is a unitary matrix and causes a rotation

ESPRIT Method

- Now lets consider the signal eigenvectors $\mathbf{V}_\mathrm{S}$ and

$$\mathbf{V}_{\mathrm{S}1} = [\mathbf{I}_{M-1} \quad \mathbf{0}]\mathbf{V}_\mathrm{S} \quad \text{and} \quad \mathbf{V}_{\mathrm{S}2} = [\mathbf{I}_{M-1} \quad \mathbf{0}]\mathbf{V}_\mathrm{S}$$

- Using the autocorrelation matrix $\mathbf{R}_x$,

$$\mathbf{R}_x\mathbf{V}_\mathrm{S} = \mathbf{V}_\mathrm{S}\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_p \end{bmatrix} = \mathbf{E}\mathbf{P}\mathbf{E}^H\mathbf{V}_\mathrm{S} + \sigma_w^2\mathbf{V}_\mathrm{S}$$

- Accordingly we have, $\mathbf{V}_\mathrm{S} = \mathbf{E}(\mathbf{P}\mathbf{E}^H\mathbf{V}_\mathrm{S}\boldsymbol{\Lambda}^{-1})$ where $\boldsymbol{\Lambda} = diag\left(\lambda_1 - \sigma_w^2, \quad \cdots, \quad \lambda_p - \sigma_w^2\right)$
- We have $\mathbf{V}_\mathrm{S} = \mathbf{E}\mathbf{C}$, where $\mathbf{C} = \mathbf{P}\mathbf{E}^H\mathbf{V}_\mathrm{S}\boldsymbol{\Lambda}^{-1}$ is non-singular

ESPRIT Method

- The term $\mathbf{V}_\mathrm{S} = \mathbf{EC}$ indicates the relation between the signal space $\mathbf{E}$ and signal eigenvectors $\mathbf{V}_\mathrm{S}$
- From the earlier partitioning, and using the above relation we have

$$\mathbf{V}_{\mathrm{S}2} = \mathbf{E}_2\mathbf{C}$$

- Using the rotational relation

$$\mathbf{V}_{\mathrm{S}2} = \mathbf{E}_2\mathbf{C} = \mathbf{E}_1\mathbf{DC} = \mathbf{V}_{\mathrm{S}1}\mathbf{C}^{-1}\mathbf{DC} = \mathbf{V}_{\mathrm{S}1}\mathbf{\Phi}.$$

- Since $\mathbf{V}_{\mathrm{S}1}$ and $\mathbf{V}_{\mathrm{S}2}$ are known, we can solve for $\mathbf{\Phi}$ using least squares

$$\mathbf{\Phi} = (\mathbf{V}_{\mathrm{S}1}^H\mathbf{V}_{\mathrm{S}1})^{-1}\mathbf{V}_{\mathrm{S}1}^H\mathbf{V}_{\mathrm{S}2}$$

- From $\mathbf{C}^{-1}\mathbf{DC} = \mathbf{\Phi}$, it can be seen that $\mathbf{D}$ and $\mathbf{\Phi}$ have the same eigen values

ESPRIT Method

- After estimating $\mathbf{\Phi}$, it's eigenvalues can be obtained
- These are also the eigenvalues of $\mathbf{D}$, and hence $\{\omega_k\}_{k=1}^{p}$ can be found from

$$arg(\gamma_k)$$

where $\{\gamma_k\}_{k=1}^{p}$ are eigenvalues of the estimated $\mathbf{\Phi}$

- Here $\mathbf{\Phi}$ is obtained using least squares (LS) approach
- LS solves for $\mathbf{X}$ from $\mathbf{AX} \approx \mathbf{B}$ or

$$\min \|\mathbf{AX} - \mathbf{B}\|^2$$

or

$$\min \|\Delta\mathbf{B}\|^2 \text{ subject to } \mathbf{AX} = \mathbf{B} + \Delta\mathbf{B}$$

- LS solution $\mathbf{X} = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{B} = \mathbf{A}^\dagger\mathbf{B}$

ESPRIT Method using Total Least Squares (TLS)

- Total least squares (TLS) error both in **A** and **B**
- TLS solves for X from

$$\min \| \mathbf{AX} - \mathbf{B} \|^2$$

or

$$\min \| \Delta\mathbf{A}\ \Delta\mathbf{B} \|^2 \text{ subject to } (\mathbf{A} + \Delta\mathbf{A})\mathbf{X} = \mathbf{B} + \Delta\mathbf{B}$$

- This can be solved using Singular Value Decomposition (SVD)

Singular Value Decomposition (SVD)

- SVD provides better numerical precision while computing eigenvalues and eigenvectors
- Uses the data $\mathbf{X}$ instead of the autocorrelation matrix $\mathbf{R}_x$

Singular Value Decomposition (SVD)

- SVD theorem states that any $K \times N$ matrix $\mathbf{X}$ can be decomposed as

$$\mathbf{X} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H$$

where $\mathbf{U}$ is the $K \times K$ unitary matrix of *left singular vectors*

$$\mathbf{U} = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots \mathbf{u}_K \end{bmatrix},$$

where $\mathbf{V}$ is the $N \times N$ unitary matrix of *right singular vectors*

$$\mathbf{V} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots \mathbf{v}_N \end{bmatrix},$$

$\boldsymbol{\Sigma}$ is the $K \times N$ nonnegative real singular values ($K > N$)

$$\boldsymbol{\Sigma} = \begin{bmatrix} \mathbf{S}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \text{ with } \mathbf{S}_1 = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \sigma_N. \end{bmatrix}$$

Singular Value Decomposition (SVD)

- Left singular vectors $\mathbf{u}_k$ are the eigenvectors of $\mathbf{X}\mathbf{X}^H$
- Right singular vectors $\mathbf{v}_k$ are the eigenvectors of $\mathbf{X}^H\mathbf{X}$

- The singular values of $\mathbf{X}^H\mathbf{X}$ are σ_k^2 the squared singular values of $\mathbf{X}$
- If $K > N$, then there are atmost N singular values, i.e., $K - N$ singular values are zero
- The pseudoinverse (used in LS) can be defined as

$$\mathbf{X}^\dagger \stackrel{\text{def}}{=} \mathbf{V}\boldsymbol{\Sigma}^\dagger\mathbf{U}^H$$

where

$$\boldsymbol{\Sigma}^\dagger = \begin{bmatrix} \mathbf{S}_1^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

ESPRIT Method using Total Least Squares (TLS)

- SVD of matrix in ESPRIT method

$$\underbrace{\begin{bmatrix} \mathbf{V}_{S1} & \mathbf{V}_{S2} \end{bmatrix}}_{M-1\times 2p} = \underbrace{\begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix}}_{M-1\times M-1} \underbrace{\begin{bmatrix} \mathbf{\Sigma}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_2 \end{bmatrix}}_{M-1\times 2p} \underbrace{\begin{bmatrix} \mathbf{V}_1^H \\ \mathbf{V}_2^H \end{bmatrix}}_{2p\times 2p}$$

- Partitioning $\mathbf{V}_2$ as

$$\mathbf{V}_2^H = \begin{bmatrix} \underbrace{\mathbf{V}_{21}^H}_{p} \underbrace{\mathbf{V}_{22}^H}_{p} \end{bmatrix}$$

- The TLS soultion can be obtained as

$$\mathbf{\Phi}_{\mathrm{TLS}} = -\mathbf{V}_{21}\mathbf{V}_{22}^{-1}$$

Principal Components Spectrum Estimation

- Earlier methods made of use of the noise subspace vectors to estimate the complex exponential frequencies

- Here we use the signal subspace vectors to estimate the power spectrum
- From eigendecomposition of $\mathbf{R}_x$ we have

$$\mathbf{R}_x = \sum_{i=1}^{M} \lambda_i \mathbf{v}_i \mathbf{v}_i^H = \sum_{i=1}^{p} \lambda_i \mathbf{v}_i \mathbf{v}_i^H + \sum_{i=p+1}^{M} \lambda_i \mathbf{v}_i \mathbf{v}_i^H,$$

  where the eigenvalues of $\mathbf{R}_x$ have been arranged in decreasing order, $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_M$.
- We can obtain a reduced rank approximation to the signal autocorrelation matrix $\mathbf{R}_s$, by retaining the principal eigenvectors of $\mathbf{R}_x$,

$$\hat{\mathbf{R}}_s = \sum_{i=1}^{p} \lambda_i \mathbf{v}_i \mathbf{v}_i^H.$$

Principal Components Spectrum Estimation

- We use the reduced rank approximation instead of $\mathbf{R}_x$ in other spectral estimators
- Effect of using principal components is to filter out noise components and hence better spectral estimation of signal components
- Using this in Blackman-Tukey estimate

$$\hat{P}_{\mathrm{BT}}(e^{j\omega}) = \sum_{k=-M}^{M} \hat{r}_x(k) w(k) e^{-jk\omega}$$

- With Bartlett window, we have

$$\hat{P}_{\mathrm{BT}}(e^{j\omega}) = \frac{1}{M} \sum_{k=-M}^{M} (M - |k|)\, \hat{r}_x(k) e^{-jk\omega} = \frac{1}{M} \mathbf{e}^H \mathbf{R}_x \mathbf{e}.$$

Principal Components Spectrum Estimation

- Using Eigendecomposition

$$\hat{P}_{\mathrm{BT}}(e^{j\omega}) = \frac{1}{M}\sum_{i=1}^{M}\left|\mathbf{e}^H\mathbf{v}_i\right|^2.$$

- With principal component approximation for $\mathbf{R}_x$, we have

$$\hat{P}_{\mathrm{PC\text{-}BT}}(e^{j\omega}) = \frac{1}{M}\mathbf{e}^H\hat{\mathbf{R}}_s\mathbf{e} = \frac{1}{M}\sum_{i=1}^{p}\left|\mathbf{e}^H\mathbf{v}_i\right|^2.$$

- Similar approximations can be used in other techniques that use the autocorrelation matrix $\mathbf{R}_x$