Part 3. Spectrum Estimation
3.3 Subspace Approaches to Frequency Estimation

Electrical & Computer Engineering
University of Maryland, College Park

Acknowledgment: ENEE630 slides were based on class notes developed by Profs. K.J. Ray Liu and Min Wu. The slides were made by Prof. Min Wu, with updates from Mr. Wei-Hong Chuang. *Contact: minwu@umd.edu*

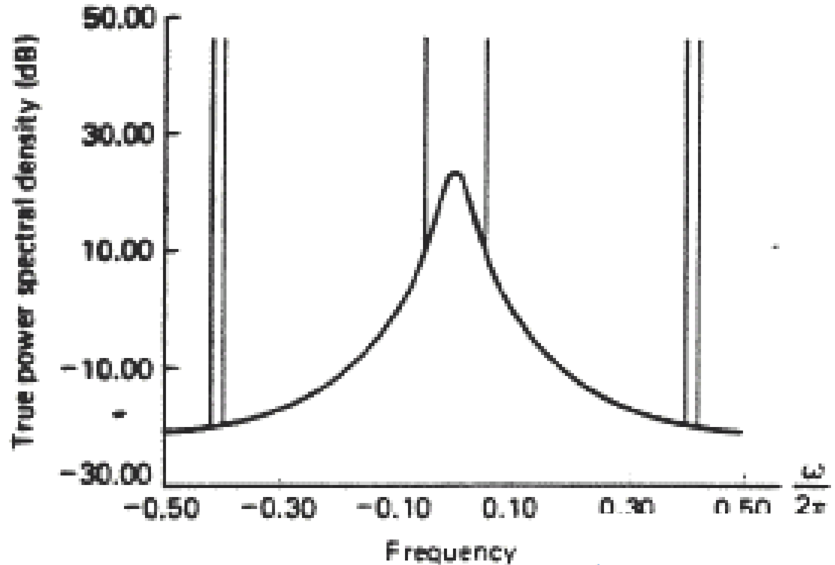
Logistics

- Final Exam: cover Part-II and III
  - Primary reference in your review: Lecture notes
  - Related readings (see a list of summary given)
  - Office hours will be posted
- Previous Lecture: Parametric approaches for spectral estimation
  - AR modeling and MESE
  - MA and ARMA modeling
- Today: (readings: Hayes 8.6)
  - Frequency estimation for complex exponential/sinusoid models

  * **Note**: Hayes book uses sig vector $\underline{x}$ = [x(n), x(n+1), … $]^T$ to define a correlation matrix, which is Hermitian w.r.t. the one per our convention with $\underline{x}$ = [x(n), x(n-1), x(n-2) … $]^T$

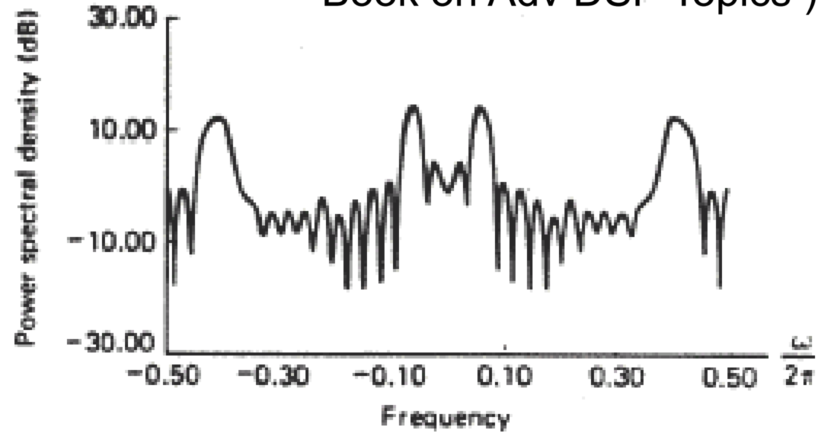
Recall: Limitations of Periodogram and ARMA

(Fig.2.17 from Lim/Oppenheim
Book on Adv DSP Topics)

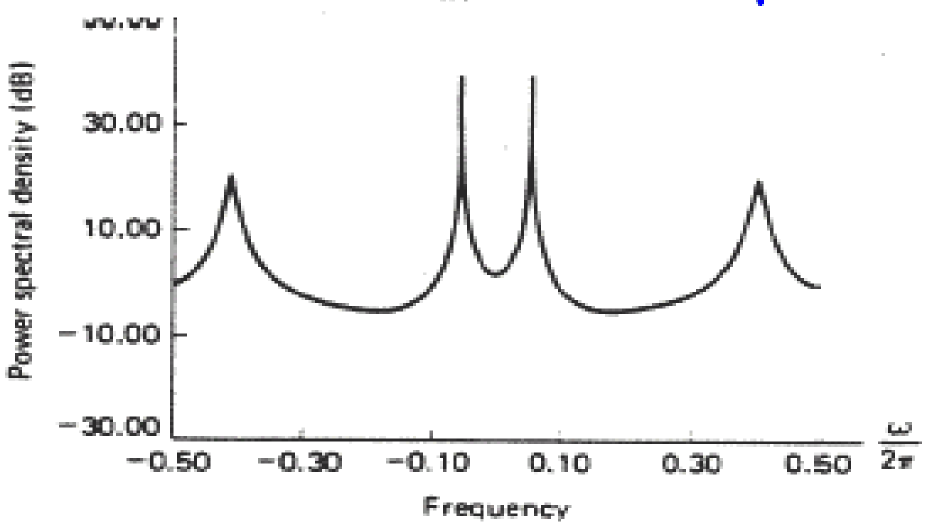


(a)

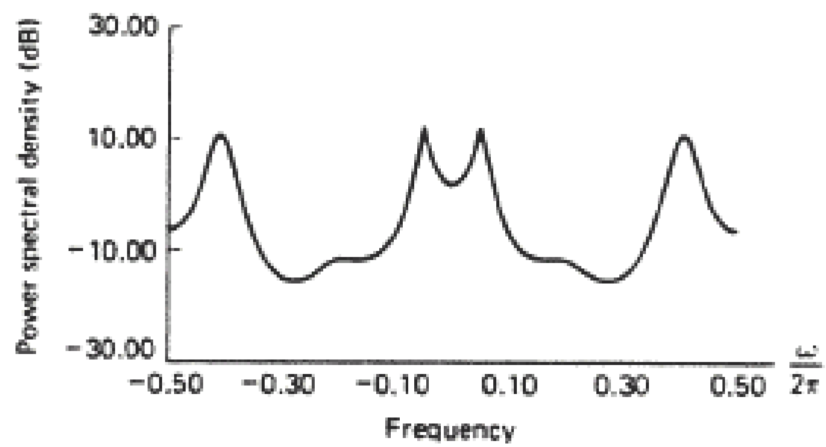
true p.s.d.



(b) Periodogram



(k) Least-squares modified Yule-Walker equations



(d) Minimum variance spectral estimator

Motivation

- Random process studied in the previous section:
  - w.s.s. process modeled as the output of a LTI filter driven by a white noise process ~ smooth p.s.d. over broad freq. range
  - Parametric spectral estimation: AR, MA, ARMA

- Another important class of random processes:
  A sum of several complex exponentials in white noise

$$x[n] = \sum_{i=1}^{p} A_i \exp[\, j(2\pi \; f_i n + \phi_i)] + w[n]$$

  - The amplitudes and p different frequencies of the complex exponentials are constant but unknown
    - *Frequencies contain desired info: velocity (sonar), formants (speech) ...*
  - Estimate the frequencies taking into account of the properties of such process

The Signal Model

$$x[n] = \sum_{i=1}^{p} A_i e^{j\phi_i} e^{j2\pi f_i n} + w[n]$$

$n = 0, 1, \ldots, N-1$ (observe N samples)

$w[n]$ white noise, zero mean, variance σ_w^2

A_i, f_i real, constant, unknown

➔ to be estimated

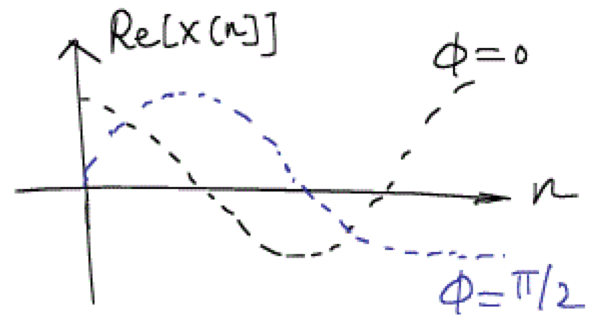
ϕ_i uniform distribution over [0, 2π); uncorrelated with w[n] and between different i

Correlation Matrix for the Process

Correlation Matrix for the Process

$$x[n] = A \exp[j(2\pi f_0 n + \phi)]$$

$$E[x[n]] = 0 \quad \forall n$$



$$E[x[n]\, x[n-k]]$$

$$= E[A \exp[j(2\pi f_0 n + \phi)] \cdot A \exp[-j(2\pi f_0 n - 2\pi f_0 k + \phi)]]$$

$$= A^2 \times \exp[j(2\pi f_0 k)]$$

$\therefore$ $x[n]$ is zero-mean w.s.s. with $r_x(k) = A^2 \exp(j 2\pi f_0 k)$.

$y[n] = x[n] + w[n]$ white noise: $E[w[n]\, w^*[n-k]] = \begin{cases} \sigma^2 & k=0 \\ 0 & \text{o.w.} \end{cases}$

$$r_y(k) = E[y[n]\, y^*[n-k]] = E[(x[n] + w[n])(x^*[n-k] + w^*[n-k])]$$

$$= r_x[k] + r_w[k] \quad (\because E[x[\cdot]\, w[\cdot]] = 0 \text{ uncorrelated})$$

$$= A^2 \exp[j 2\pi f_0 k] + \sigma^2 \delta[k]$$

E[x( ) w( )] = E[x( )] E[w( )] = 0
this crosscorr term vanish
because of uncorrelated *and*
zero mean for either x( ) or w( ).

Complex Exponential Vectors

$$\underline{e}(f) = \left[1, e^{-j2\pi f}, e^{-j4\pi f} \cdots, e^{-j2\pi(M-1)f} \right]^T$$

$$\underline{e}^H(f_1) \cdot \underline{e}(f_2) = \sum_{k=0}^{M-1} e^{j2\pi(f_1-f_2)k} = \frac{1-e^{j2\pi(f_1-f_2)M}}{1-e^{j2\pi(f_1-f_2)}} \text{ if } f_1 \neq f_2$$

If $f_1 - f_2 = a/M$ for some integer $a \Rightarrow \underline{e}^H(f_1) \cdot \underline{e}(f_2) = 0$

Correlation Matrix for the Process

$$r_x(k) = E\left[x[n]x^*[n-k]\right] =$$

An MxM correlation matrix for {x[n]} (M>p):

$$R_x = R_s + R_w$$

where $\underline{e}_i = [1, e^{-j2\pi f_i}, e^{-j4\pi f_i}, \ldots e^{-j2\pi f_i (M-1)}]^T$

Correlation Matrix for the Process

$$r_x(k) = E\left[x[n]x^*[n-k]\right] = \sum^{p} \underbrace{A_i^2}_{\triangleq P_i} e^{j2\pi f_i k} + \sigma_w^2 \delta(k)$$

An MxM correlation matrix for {x[n]} (M>p):

$$R_x = R_s + R_w$$

$$R_w = \sigma_w^2 I$$ ➔ full rank

$$R_s = \sum_{i=1}^{p} P_i \underline{e}_i \underline{e}_i^H$$

where $\underline{e}_i = [1, e^{-j2\pi f_i}, e^{-j4\pi f_i}, \ldots e^{-j2\pi f_i (M-1)}]^T$

Correlation Matrix for the Process (cont'd)

$$R_s = \sum_{i=1}^{p} P_i \underline{e}_i \underline{e}_i^H$$

$\underline{e}_i \underline{e}_i^H$ has rank

The MxM matrix R_s has rank

Correlation Matrix for the Process (cont'd)

$$R_s = \sum_{i=1}^{p} P_i \underline{e}_i \underline{e}_i^H$$

$$= \underbrace{[\underline{e}_1, \underline{e}_2, \ldots \underline{e}_p]}_{\triangleq S_{M \times p}} \underbrace{\begin{bmatrix} P_1 & & & \\ & P_2 & & \\ & & \ddots & \\ & & & P_p \end{bmatrix}}_{\triangleq D_{p \times p}} \begin{bmatrix} \underline{e}_1^H \\ \underline{e}_2^H \\ \vdots \\ \underline{e}_p^H \end{bmatrix}$$

$$= S D S^H$$

$\underline{e}_i \underline{e}_i^H$ has rank 1 (all columns are related by a factor)

The MxM matrix R_s has rank p, and has only *p* nonzero eigenvalues.

Review: Rank and Eigen Properties

- Multiplying a full rank matrix won't change the rank of a matrix

  i.e. $r(A) = r(PA) = r(AQ)$
  where A is mxn, P is mxm full rank, and Q is nxn full rank.

  – The rank of A is equal to the rank of $A A^H$ and $A^H A$.

  – Elementary operations (which can be characterized as multiplying by a full rank matrix) doesn't change matrix rank:

    - *including interchange 2 rows/cols; multiply a row/col by a nonzero factor; add a scaled version of one row/col to another.*

- Correlation matrix Rx in our model has full rank.

- Non-zero eigenvectors corresponding to distinct eigenvalues are linearly independent

- det(A) = product of all eigenvalues; so a matrix is invertible iff all eigenvalues are nonzero.

(see Hayes Sec.2.3 review of linear algebra)

Eigenvalues/vectors for Hermitian Matrix

- Multiplying A with a full rank matrix won't change rank(A)
- Eigenvalue decomposition
  - For an nxn matrix A having a set of n linearly independent eigenvectors, we can put together its eigenvectors as V s.t.

$$A\underline{v}_i = \lambda_i \underline{v}_i$$

$$A[\underline{v}_1, v_2 \ldots] = \underbrace{[v_1, \ldots v_n]}_{V} \begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}$$

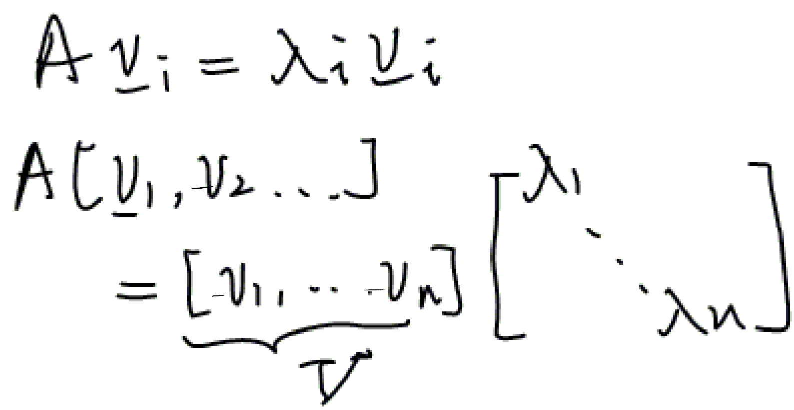
- For any nxn Hermitian matrix
  - There exists a set of n orthonormal eigenvectors
  - Thus V is unitary for Hermitian matrix A, and

(see Hayes Sec.2.3.9 review of linear algebra)

Eigenvalues/vectors for Hermitian Matrix

- Multiplying A with a full rank matrix won't change rank(A)
- Eigenvalue decomposition
  - For an nxn matrix A having a set of n linearly independent eigenvectors, we can put together its eigenvectors as V s.t.

$$A = V \,\mathrm{diag}(\lambda_1, \lambda_2, \ldots \lambda_n)\, V^{-1}$$



- For any nxn Hermitian matrix
  - There exists a set of n orthonormal eigenvectors
  - Thus V is unitary for Hermitian matrix A, i.e. $V^{-1} = V^H$

$$A = V \,\mathrm{diag}(\lambda_1, \lambda_2, \ldots \lambda_n)\, V^H = \lambda_1 \underline{v}_1 \underline{v}_1^H + \ldots + \lambda_n \underline{v}_n \underline{v}_n^H$$

(see Hayes Sec.2.3.9 review of linear algebra)

Eigen Analysis of the Correlation Matrix

Eigen Analysis of the Correlation Matrix

Let $\underline{v}_i$ be an eigenvector of R_x with the corresponding eigenvalue λ_i, i.e., $R_x \underline{v}_i = \lambda_i \underline{v}_i$

$$\because \; R_x \underline{v}_i = R_s \underline{v}_i + \sigma_w^2 \underline{v}_i = \lambda_i \underline{v}_i$$

$$\therefore \; R_s \underline{v}_i =$$

$$\therefore \; \lambda_i = \{$$

(R_s has p nonzero eigenvalues)

Eigen Analysis of the Correlation Matrix

Let $\underline{v}_i$ be an eigenvector of R_x with the corresponding eigenvalue λ_i, i.e., $R_x \underline{v}_i = \lambda_i \underline{v}_i$

$$\because R_x \underline{v}_i = R_s \underline{v}_i + \sigma_w^2 \underline{v}_i = \lambda_i \underline{v}_i$$

$$\therefore R_s \underline{v}_i = (\lambda_i - \sigma_w^2) \underline{v}_i$$

i.e., $\underline{v}_i$ is also an eigenvector for R_s, and the corresponding eigenvalue is

$$\lambda_i^{(s)} = \lambda_i - \sigma_w^2$$

$$\therefore \lambda_i = \begin{cases} \lambda_i^{(s)} + \sigma_w^2 > \sigma_w^2, & i = 1, 2, \cdots p \\ \sigma_w^2, & i = p+1, \cdots M \end{cases}$$

(R_s has p nonzero eigenvalues)

Signal Subspace and Noise Subspace

Signal Subspace and Noise Subspace

For $i = p+1, \cdots M$: $R_s \times \underline{v}_i = 0 \times \underline{v}_i$

Also, $R_s = S D S^H$;

$$\therefore \; S D S^H \underline{v}_i =$$

$\underbrace{S^H \underline{v}_i}_{p \times 1}$

$$\Rightarrow S^H \underline{v}_i =$$

Since $S = [\underline{e}_1, \cdots \underline{e}_p]$. $\Rightarrow$

Signal Subspace and Noise Subspace

For $i = p+1, \ldots M$: $R_s \times \underline{v}_i = 0 \times \underline{v}_i$

Also, $R_s = SDS^H$;

$\therefore \underbrace{S}_{M \times p,\ \text{full rank}=p} D \underbrace{S^H \underline{v}_i}_{p \times 1} = \underline{0}$ for i= p+1, …, M

M x p, full rank=p

i.e., the p column vectors are linearly independent

$\Rightarrow S^H \underline{v}_i = \underline{0}$

Since $S = [\underline{e}_1, \cdots \underline{e}_p]$ $\Rightarrow \underline{e}_l^H \underline{v}_i = 0,$ $l = 1,2,\ldots,p$ $i = p+1,\ldots,M$

$\therefore \text{span}\{\underline{e}_1, \cdots \underline{e}_p\} \perp \text{span}\{\underline{v}_{p+1}, \cdots \underline{v}_M\}$

SIGNAL SUBSPACE NOISE SUBSPACE correspond to eigenvalue = noise var

Relations Between Signal and Noise Subspaces

Since R_x and R_s are Hermitian matrices,

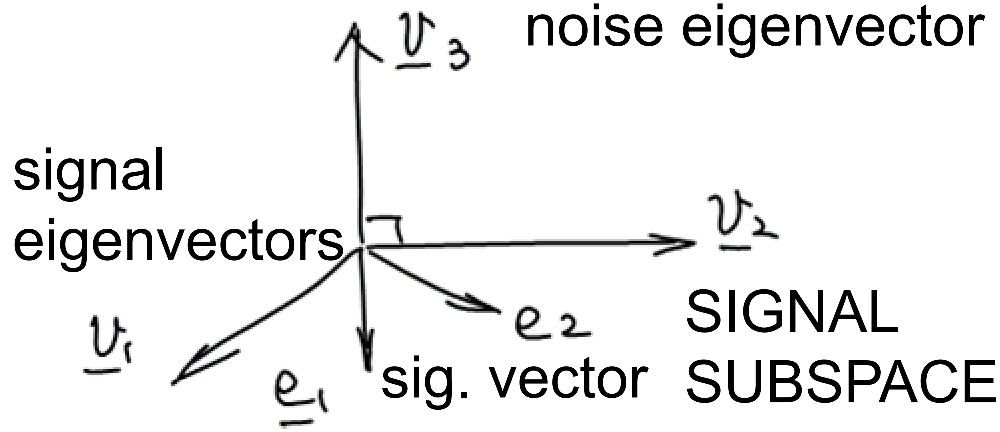
the eigenvectors are orthogonal to each other:

$$\underline{v}_i \perp \underline{v}_j \quad \forall\, i \neq j$$

$\Rightarrow$

Recall $\text{span}\{\underline{e}_1, \cdots \underline{e}_p\} \perp \text{span}\{\underline{v}_{p+1}, \cdots \underline{v}_M\}$,

So it follows that



Relations Between Signal and Noise Subspaces

Since R_x and R_s are Hermitian matrices,

the eigenvectors are orthogonal to each other:

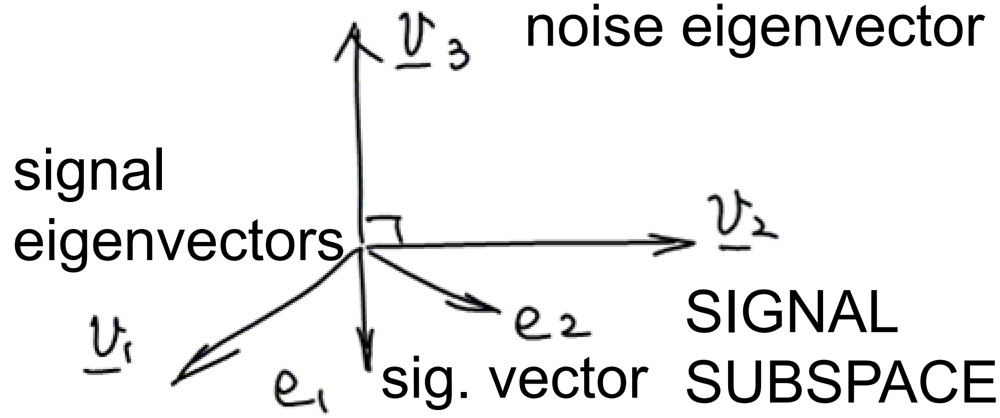
$$\underline{v}_i \perp \underline{v}_j \quad \forall\, i \neq j$$

$$\Rightarrow \quad \text{span}\{\underline{v}_1, \dots \underline{v}_p\} \perp \text{span}\{\underline{v}_{p+1}, \dots \underline{v}_M\}$$

Recall $\text{span}\{\underline{e}_1, \dots \underline{e}_p\} \perp \text{span}\{\underline{v}_{p+1}, \dots \underline{v}_M\}$,

So it follows that

$$\text{span}\{\underline{e}_1, \dots \underline{e}_p\} = \text{span}\{\underline{v}_1, \dots \underline{v}_p\}$$



Frequency Estimation Function: General Form

Recall $\underline{e}_l^H \underline{v}_i = 0$ *for l=1, … p; i = p+1, … M*

Knowing eigenvectors of correlation matrix R_x, we can use these orthogonal conditions to find the frequencies $\{f_l\}$:

$$\underline{e}^H(f)\underline{v}_i = 0 ?$$

We form a frequency estimation function

Here α_i are properly chosen constants (weights) for producing weighted average for projection power with all noise eigenvectors

Frequency Estimation Function: General Form

Recall $\underline{e}_l^H \underline{v}_i = 0$ *for l=1, … p; i = p+1, … M*

Knowing eigenvectors of correlation matrix R_x, we can use these orthogonal conditions to find the frequencies $\{f_l\}$:

$$\underline{e}^H(f)\underline{v}_i = 0\,?$$

We form a frequency estimation function

$$\hat{P}(f) = \frac{1}{\sum_{i=p+1}^{M} \alpha_i \left| \underline{e}(f)^H \underline{v}_i \right|^2}$$

$$\Rightarrow \hat{P}(f) \text{ is LARGE at } f_1, \ldots, f_p$$

Here α_i are properly chosen constants (weights) for producing weighted average for projection power with all noise eigenvectors

Pisarenko Method for Frequency Estimation (1973)

- This assumes the number of complex exponentials (*p*) and the first (p+1) lags of the autocorrelation function are known or have been estimated

r(0),…,r(P)

- The eigenvector corresponding to the smallest eigenvalue(s) of $R_{(p+1)x(p+1)}$ is in the noise subspace and can be used in the Pisarenko method.

- The equivalent frequency estimation function is:

Pisarenko Method for Frequency Estimation (1973)

- This assumes the number of complex exponentials (*p*) and the first (p+1) lags of the autocorrelation function are known or have been estimated

r(0),…,r(p)

- The eigenvector corresponding to the smallest eigenvalue(s) of $R_{(p+1)x(p+1)}$ is in the noise subspace and can be used in the Pisarenko method.

- The equivalent frequency estimation function is:

$$\hat{P}(f) = \frac{1}{\left| \underline{e}(f)^H \underline{v}_{\min} \right|^2}$$

Estimating the Amplitudes

Once the frequencies of the complex exponentials are determined, the amplitudes can be found from the eigenvalues of R_x:

$$R_x \underline{v}_i = \lambda_i \underline{v}_i \quad (i = 1, 2, \ldots, p)$$

Normalize $\underline{v}_i$ s.t. $\underline{v}_i^H \underline{v}_i = 1$

Recall $R_x = \sum_{k=1}^{p} P_k \underline{e}_k \underline{e}_k^H + \sigma_w^2 I$

Estimating the Amplitudes

Once the frequencies of the complex exponentials are determined, the amplitudes can be found from the eigenvalues of R_x:

$$R_x \underline{v}_i = \lambda_i \underline{v}_i \quad (i = 1, 2, \ldots, p)$$

normalize $\underline{v}_i$ s.t. $\underline{v}_i^H \underline{v}_i = 1$

$$\Rightarrow \underline{v}_i^H R_x \underline{v}_i = \lambda_i \underline{v}_i^H \underline{v}_i = \lambda_i$$

Recall $R_x = \sum_{k=1}^{p} P_k \underline{e}_k \underline{e}_k^H + \sigma_w^2 I$

$$\Rightarrow \sum_{k=1}^{p} P_k \left| \underline{e}_k^H \underline{v}_i \right|^2 = \lambda_i - \sigma_w^2, \quad i = 1, \ldots, p$$

DTFT of sig eigvector $v_i(\cdot)$ at $-f_k$

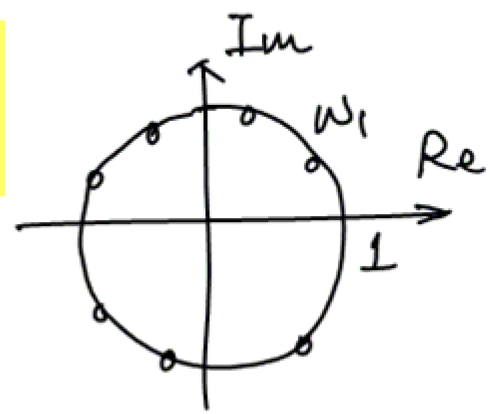
➔ Solve p equations for $\{ P_k \}$

Pisarenko Method and Interpretations

Interpretation of Pisarenko Method

Since $\underline{e}_l^H \underline{v}_i = 0, \quad \begin{matrix} l = 1,2,\ldots, p \\ i = p+1,\ldots, M \end{matrix},$ $\quad \underline{v}_i \triangleq \begin{bmatrix} v_i(0) \\ v_i(1) \\ \vdots \\ v_i(M-1) \end{bmatrix}$ noise eigvector

$$\Rightarrow \sum_{k=0}^{M-1} v_i(k) e^{j2\pi f_l k} = 0 \quad \text{for} \quad l = 1,2,\ldots, p$$



Thus given any $\underline{v}_i$, i=p+1,…,M, we can estimate the sinusoidal frequencies by finding the zeros on unit circle from

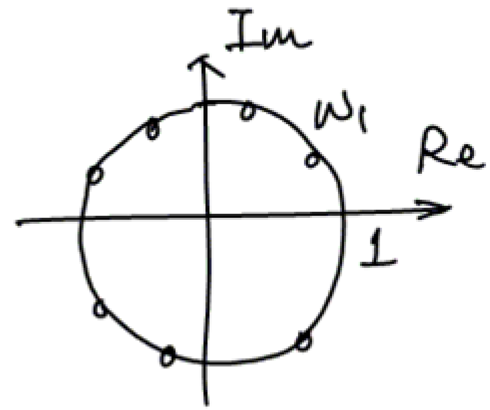
Interpretation of Pisarenko Method

Since $\underline{e}_l^H \underline{v}_i = 0, \quad \begin{matrix} l = 1,2,\ldots, p \\ i = p+1,\ldots, M \end{matrix}$, $\underline{v}_i \triangleq \begin{bmatrix} v_i(0) \\ v_i(1) \\ \vdots \\ v_i(M-1) \end{bmatrix}$

noise eigvector

$$\Rightarrow \sum_{k=0}^{M-1} v_i(k) e^{j2\pi f_l k} = 0 \quad \text{for} \quad l = 1,2,\ldots, p$$

i.e. $\text{DTFT}\{v_i(\cdot)\}\Big|_{f=-f_l} = 0$



Thus given any $\underline{v}_i$, i=p+1,…,M, we can estimate the sinusoidal frequencies by finding the zeros on unit circle from

$$Z[v_i(\cdot)] = \sum_{k=0}^{M-1} v_i(k) z^{-k}$$

the angle of zeros reflects the freq.

Improvement over Pisarenko Method

- Need to know or accurately estimate the # of sinusoids (p)
- Inaccurate estimation of autocorrelation values

  => Inaccurate eigen results of the (estimated) correlation matrix

  => p zeros on unit circle in frequency estimation function may not be on the right places

- What if we use larger MxM correlation matrix?
  - More than one eigen vectors to form the noise subspace: which of (M-p) eigen vectors shall we use to check orthogonality with $\underline{e}(f)$?

  ZT[{ $v_i(0)$, … $v_i(M-1)$}] ~ $(M-1)^{th}$ order polynomial => (M-1) zeros

  - p zeros are on unit circle (corresponding to the freq. of sinusoids)
  - Other (M-1-p) zeros may lie anywhere and could be close to unit circle => may give false peaks

MUltiple SIgnal Classification (MUSIC) Algorithm

- Addressing issues with larger correlation matrix

  ZT[{ $v_i(0)$, … $v_i(M-1)$}] ~ $(M-1)^{th}$ order polynomial => (M-1) zeros

  – p zeros are on unit circle (corresponding to the freq. of sinusoids)

  – Other (M-1-p) zeros may lie anywhere and could be close to unit circle => may give false peaks

- Basic idea of MUSIC algorithm

  – Reduce spurious peaks of freq. estimation function by averaging over the results from (M-p) smallest eigenvalues of the correlation matrix

  => i.e. to find those freq. that give signal vectors **consistently orthogonal** to all noise eigen vectors

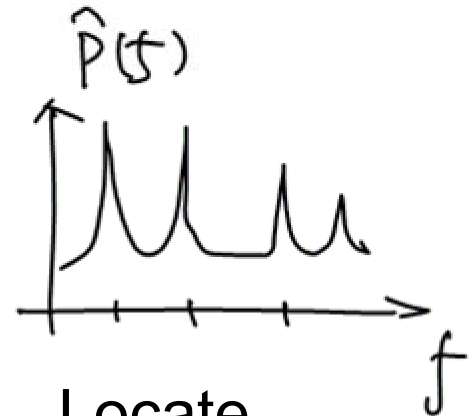
MUSIC Algorithm: Details

MUSIC Algorithm

The frequency estimation function

$$\hat{P}_{MUSIC}(f) = \frac{1}{\sum_{i=p+1}^{M} \left| \underline{e}^H(f)\, \underline{v}_i \right|^2}$$

$$= \frac{1}{\underline{e}^H(f)\, V\, V^H\, \underline{e}(f)}$$



Locate the peaks

where $\underline{e}(f) = \begin{bmatrix} 1 \\ e^{-j2\pi f} \\ \vdots \\ e^{-j2\pi f(M-1)} \end{bmatrix}$, $V = [\, \underline{v}_{p+1}, \cdots \underline{v}_M \,]$

Example-1

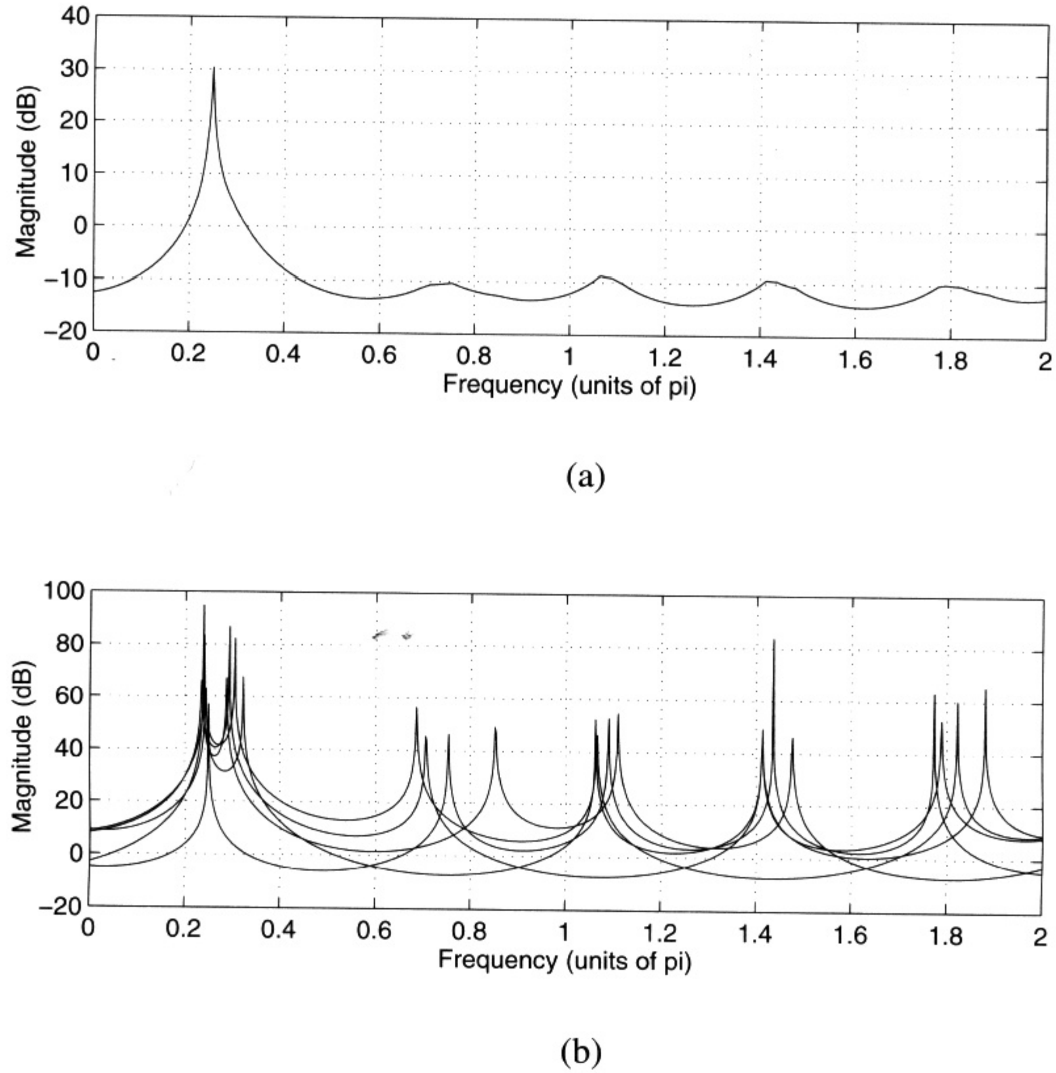
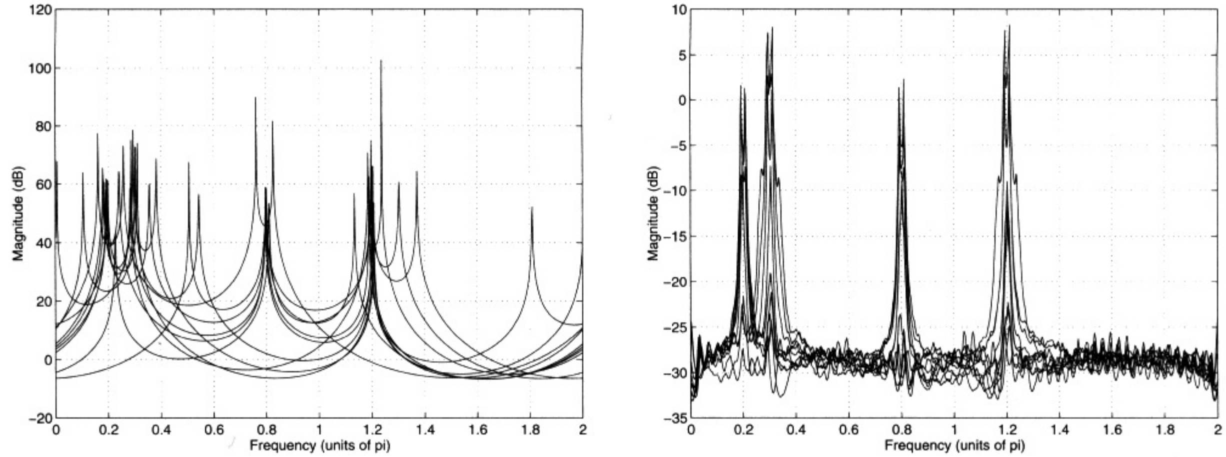
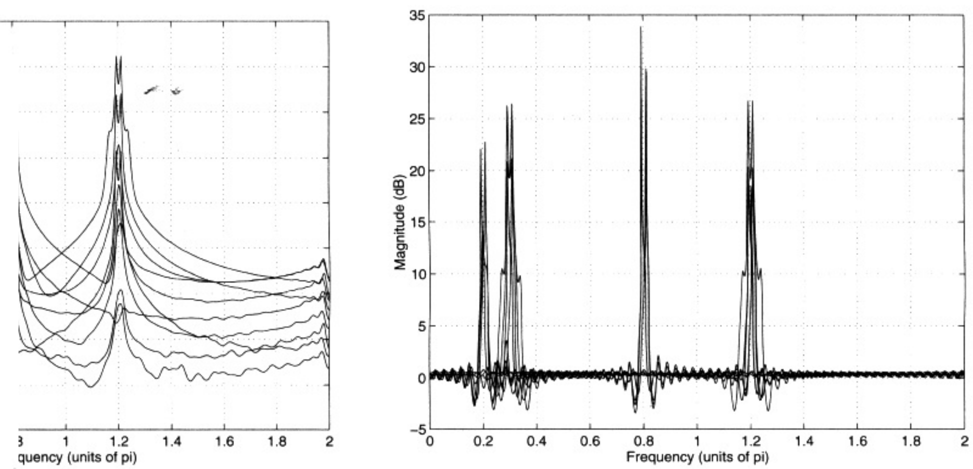


Figure 8.31 *Frequency estimation functions of a single complex exponential in white noise. (a) The frequency estimation function that uses all of the noise eigenvectors with a weighting $\alpha_i = 1$. (b) An overlay plot of the frequency estimation functions $V_i(e^{j\omega}) = 1/|\mathbf{e}^H \mathbf{v}_i|^2$ that are derived from each noise eigenvector.*

(Fig.8.31 from M. Hayes Book; examples are for 6x6 correlation matrix estimated from 64-value observations)

Example-2





quency estimation functions for a process consisting of four complex exponen- sing (a) the Pisarenko harmonic decomposition, (b) the MUSIC algorithm, (c) od and (d) the minimum norm algorithm.

Table 8.10 ***Noise Subspace Methods for Frequency Estimation***

Pisarenko	$\hat{P}_{PHD}(e^{j\omega}) = \dfrac{1}{\lvert \mathbf{e}^H \mathbf{v}_{\min} \rvert^2}$
MUSIC	$\hat{P}_{MU}(e^{j\omega}) = \dfrac{1}{\sum_{i=p+1}^{M} \lvert \mathbf{e}^H \mathbf{v}_i \rvert^2}$
Eigenvector Method	$\hat{P}_{EV}(e^{j\omega}) = \dfrac{1}{\sum_{i=p+1}^{M} \frac{1}{\lambda_i} \lvert \mathbf{e}^H \mathbf{v}_i \rvert^2}$
Minimum Norm	$\hat{P}_{MN}(e^{j\omega}) = \dfrac{1}{\lvert \mathbf{e}^H \mathbf{a} \rvert^2}$; $\mathbf{a} = \lambda \mathbf{P}_n \mathbf{u}_1$

(Fig.8.37 & Table 8.10 from M. Hayes Book; overlaying results of10 realizations with 64 observed signal points each.)