Statistical Signal Processing
9. Subspace Approaches to Frequency Estimation

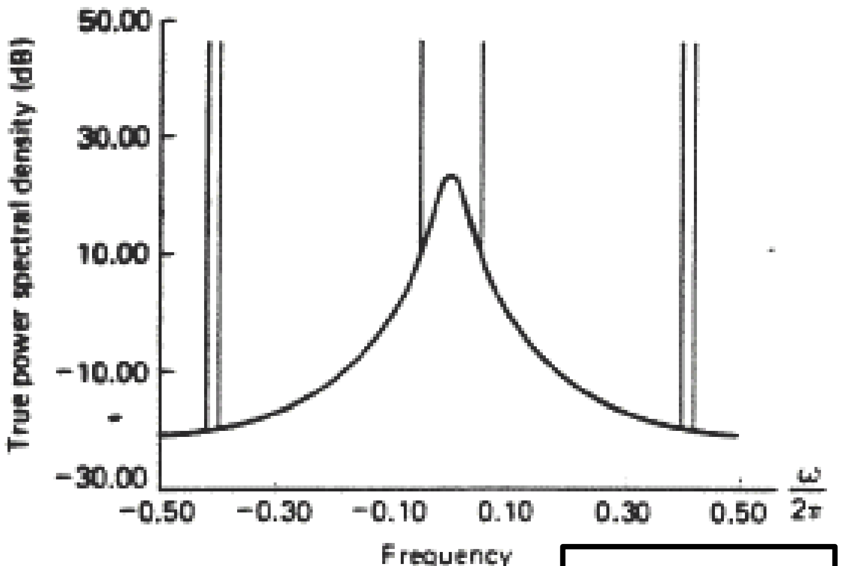
Electrical & Computer Engineering

North Carolina State University

Acknowledgment: ECE792-41 slides were adapted from ENEE630 slides developed by Profs. K.J. Ray Liu and Min Wu at the University of Maryland, College Park. Contact: chauwai.wong@ncsu.edu.

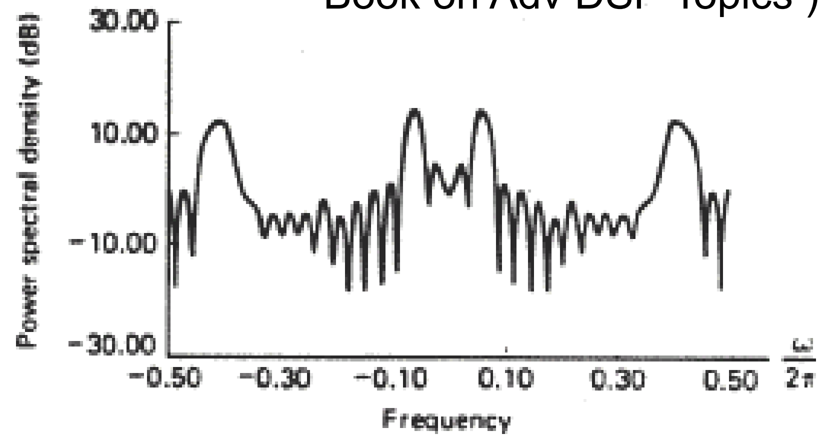
Recall: Limitations of Periodogram and ARMA

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

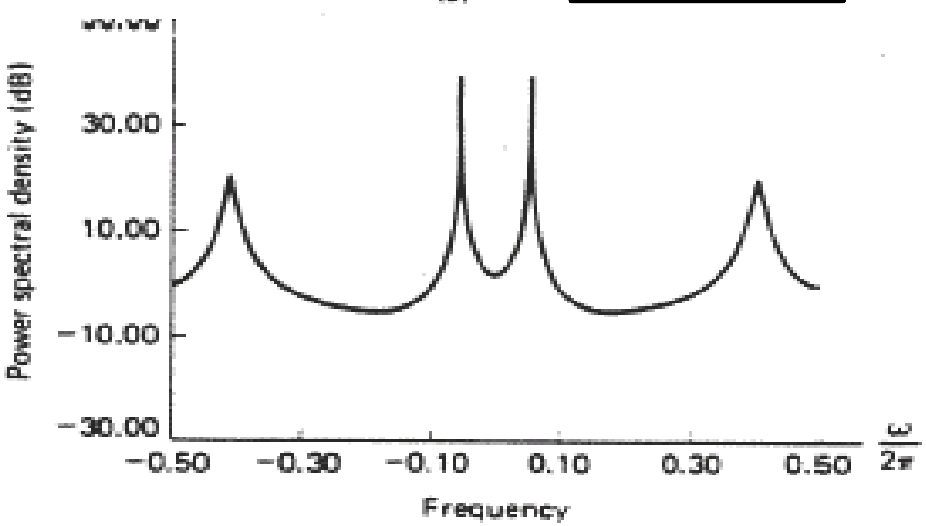


(a)

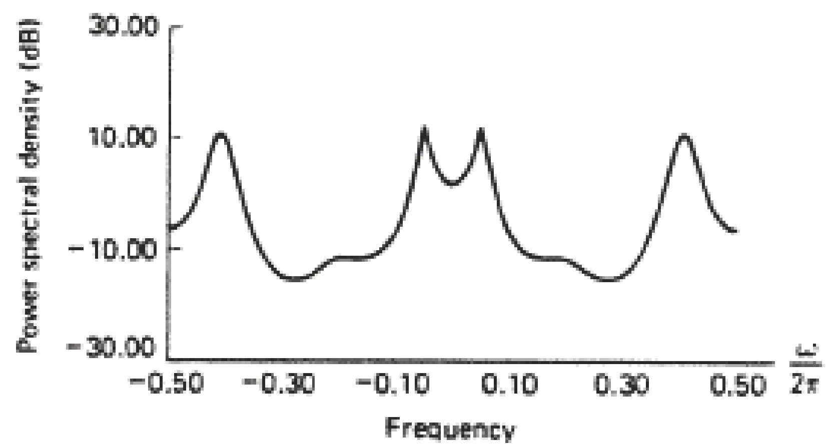
True p.s.d.



(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)

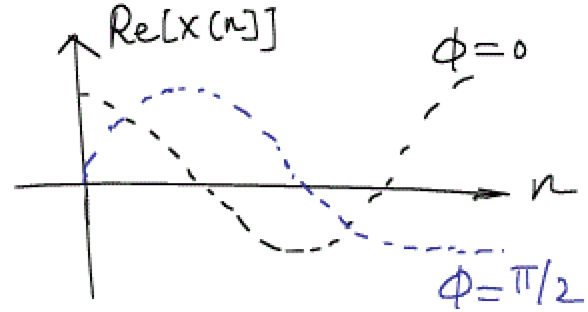
$w[n]$ white noise, zero mean, variance σ_w^2

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

Recall: Single Complex Exponential Case



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

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

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

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

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

$\therefore$ $x[n]$ is zero-mean w.s.s. with $r_x(k) = A^2 \exp(j2\pi f_0 k)$.

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

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

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

$$= A^2 \exp[j2\pi f_0 k] + \sigma^2 \delta[k]$$

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

Deriving Autocorrelation Function

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

$$r_x(k) = E\left[x[n]x^*[n-k]\right] = E\left[\left[\sum_{l=1}^{p} s_l[n] + w[n]\right] \cdot \left[\sum_{m=1}^{p} s_m{}^*[n-k] + w^*[n-k]\right]\right]$$

- $E\left[s_l[n]s_m{}^*[n-k]\right] = \begin{cases} E\left[s_l[n]\right]E\left[s_m[n-k]\right]^* = 0 & (\text{for } l \neq m) \\ \mathrm{r_{s_m}}(\mathrm{k}) = A_m^2 e^{j2\pi f_m k} & (\text{for } l = m) \end{cases}$
- $E\left[s_l[n]w^*[n-k]\right] = E\left[s_l[n]\right]E\left[w[n-k]\right]^* = 0$
- $E\left[w[n]w^*[n-k]\right] = \sigma_w^2 \cdot \delta[k]$

$$=> r_x(k) = E\left[x[n]x^*[n-k]\right] = \sum_{i=1}^{p} A_i^2 e^{j2\pi f_i k} + \sigma_w^2 \delta(k)$$

Deriving Correlation Matrix

- May bring $r_x(k)$ into the correlation matrix
- Or from the expectation of vector's outer product and use the correlation analysis from last page

$$\underline{x}[n] = \sum_{i=1}^{p} \underline{s}_i[n] + \underline{w}[n]$$

$$R_x = E\left[\underline{x}[n]\underline{x}^H[n]\right] = E\left[\left[\sum_{l=1}^{p} \underline{s}_l[n] + \underline{w}[n]\right] \cdot \left[\sum_{m=1}^{p} \underline{s}_m{}^H[n] + \underline{w}^H[n]\right]\right]$$

$$\Rightarrow R_x = \sum_{i=1}^{p} P_i \, \underline{e}_i \underline{e}_i{}^H + \sigma_w^2 I$$

Summary: Correlation Matrix for the Process

$$r_x(k) = E\left[x[n]x^*[n-k]\right] = \sum_{i=1}^{p} \underbrace{A_i^2}_{\triangleq P_i} e^{j2\pi f_i k} + \sigma_w^2 \delta(k)$$

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

$$R_x = R_s + R_w$$

$$R_w = \sigma_w^2 I \ \text{(full rank)}$$

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

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

Correlation Matrix for the Process (cont'd)

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

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

$$= S D S^H$$

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

The $M \times M$ 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 $m \times n$, P is $m \times m$ full rank, and Q is $n \times n$ full rank.

  - The rank of A is equal to the rank of AA^H and A^HA.
  - 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 R_x 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 $\text{rank}(A)$
- Eigenvalue decomposition
  - For an $n \times n$ matrix A having a set of n linearly independent eigenvectors, we can put together its eigenvectors as V s.t.

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

$$Av_i = \lambda_i v_i$$

$$A\,[v_1, \ldots, v_n] = \underbrace{[v_1, \ldots, v_n]}_{V} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda n \end{bmatrix}$$

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

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

(see Hayes Sec.2.3.9 review of linear algebra)

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

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

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

$\therefore \underbrace{S}_{} D S^H \underbrace{\underline{v}_i}_{p\times 1} = \underline{0}$ 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 $S = [\underline{e}_1, \cdots \underline{e}_p]$. $\Rightarrow \underline{e}_l^H \underline{v}_i = 0$, $l = 1,2,\dots,p$; $i=p+1,\dots,M$

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

SIGNAL SUBSPACE NOISE SUBSPACE

eigenvalue = σ_e^2

Relations Between Signal and Noise Subspaces

Since R_x and R_s are Hermitian matrices,

the eigenvectors are orthogonal to each other:

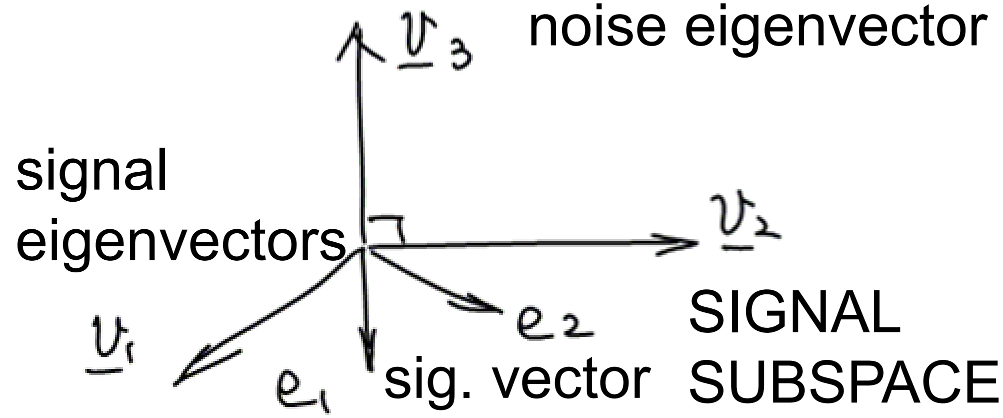
$$\underline{v}_i \perp \underline{v}_j \quad \forall\, i \neq j$$

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

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

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



Discussion: Complex Exponential Vectors

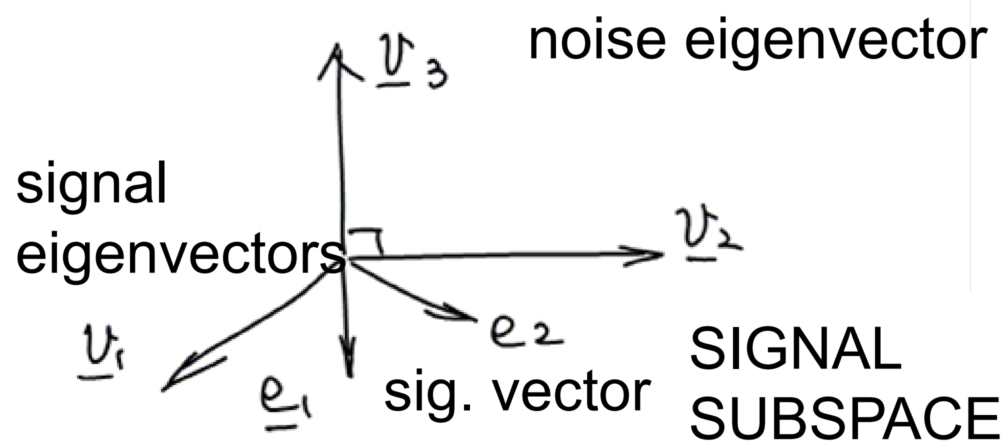
$$\underline{e}(f) = \left[1, e^{-j2\pi f}, e^{-j4\pi f}, \ldots, e^{-j2\pi(M-1)f} \right]^T$$

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

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

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

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



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

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

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

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)

- Assumes the number of complex exponentials, *p*, is known, and the first *p*+1 lags of the autocorrelation function, *r*(0), …, *r*(*p*), are known/have been estimated.

- The eigenvector corresponding to the smallest eigenvalue of $\mathbf{R}_{(p+1)\times(p+1)}$ is the sole component of the noise subspace.

- The equivalent frequency estimation function is:

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

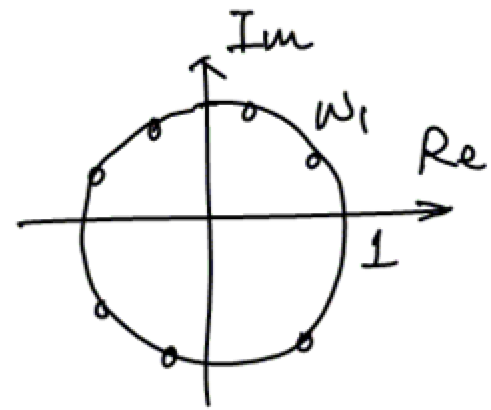
Interpretation of Pisarenko Method

Since $\underline{e}^H(f)\,\underline{v}_{\min} = 0$, where $\underline{v}_{\min} = [v(0),\ldots,v(p)]^T$

$$\Rightarrow \sum_{k=0}^{p} v_{\min}(k) e^{j2\pi fk} = 0$$

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

We can estimate the sinusoidal frequencies by finding the *p*−1 zeros on unit circle:



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

the angle of zeros reflects the freq.

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

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

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

➔ Solve *p* equations for { P_k }

Limitations of 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 a larger M×M correlation matrix?
  - More than one eigenvectors will form the noise subspace: Which of $M-p$ eigenvectors shall we use to check orthogonality with $\underline{e}(f)$?
  - For one particular eigenvector chosen, there are $M-1$ zeros:
    - p zeros correspond to the true frequency components, whereas
    - $M-1-p$ zeros lead to false peaks.

MUltiple SIgnal Classification (MUSIC) Algorithm

- 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

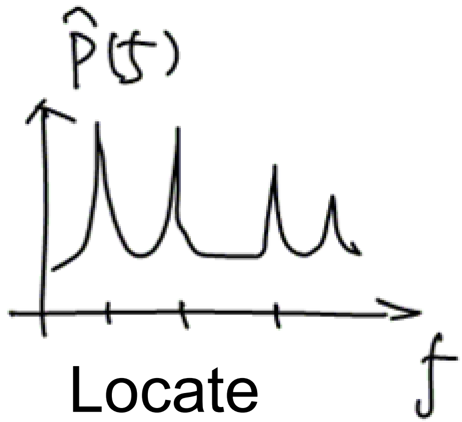
  => i.e., to find those freq. that give signal vectors **consistently orthogonal** to all noise eigenvectors.

MUSIC Algorithm

The frequency estimation function

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

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

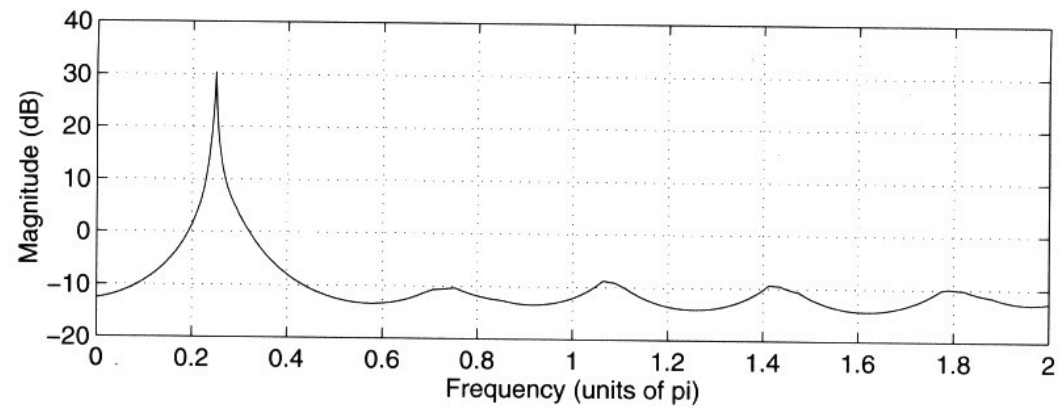


Locate the peaks

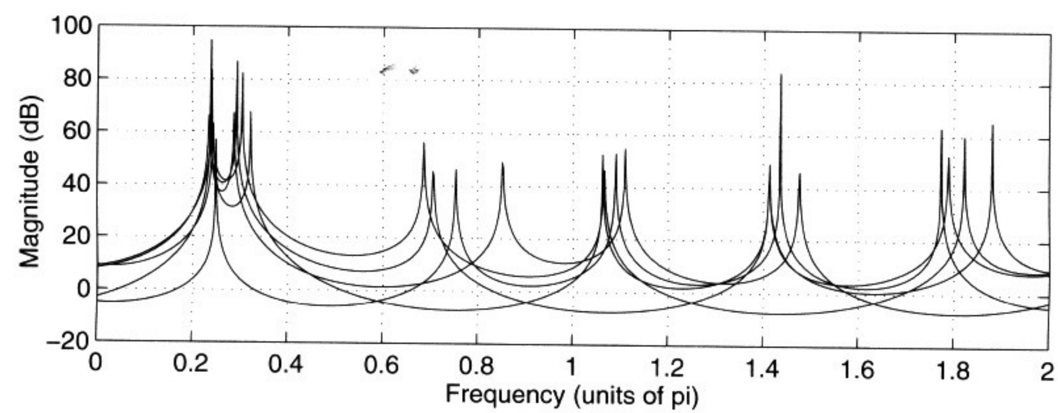
where $\underline{e}(f) = \left[1, e^{-j2\pi f}, e^{-j4\pi f}, \ldots, e^{-j2\pi f(M-1)}\right]^T$

$$V = \left[\underline{v}_{p+1}, \ldots, \underline{v}_M\right]$$

Example-1



(a)



(b)

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

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

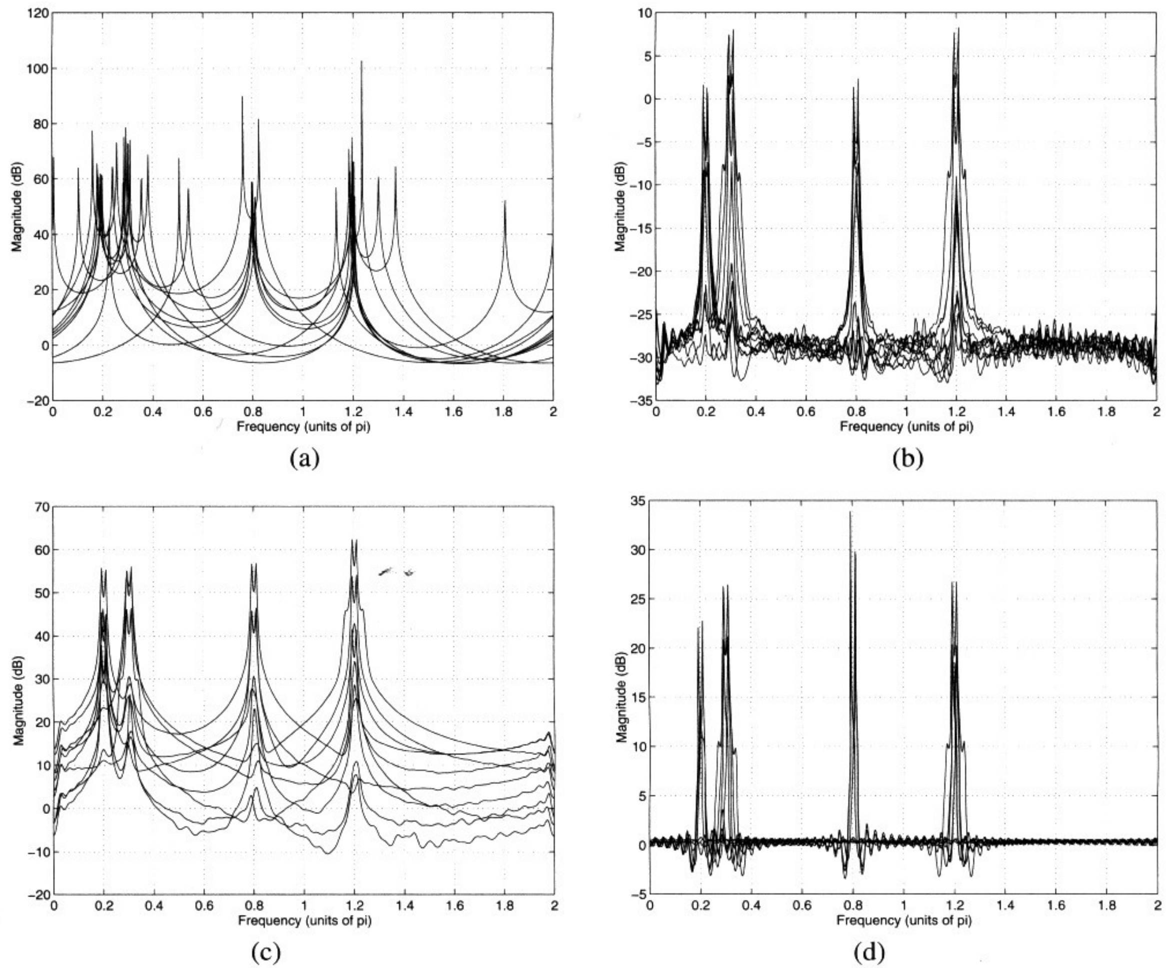


Figure 8.37 *The frequency estimation functions for a process consisting of four complex exponentials in white noise using (a) the Pisarenko harmonic decomposition, (b) the MUSIC algorithm, (c) the eigenvector method and (d) the minimum norm algorithm.*

(Fig.8.37 & Table 8.10 from M. Hayes Book; overlaying results of 10 realizations with 64 observed signal points each.)