Part 3. Spectrum Estimation
3.3 Subspace Approaches to Frequency Estimation

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 Sec.3.2: 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 \( \mathbf{x} = [x(n), x(n+1), \ldots]^T \) to define a correlation matrix, which is Hermitian w.r.t. the one per our convention with \( \mathbf{x} = [x(n), x(n-1), x(n-2), \ldots]^T \)

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 \( \sim \) 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
  \[
  x[n] = \sum_{i=1}^{p} A_i \exp\left[j(2\pi f_i n + \phi_i)\right] + w[n]
  \]
  - 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 \( \sigma_w^2 \)
- \( A_i, f_i \) real, constant, unknown \( \Rightarrow \) to be estimated
- \( \phi_i \) uniform distribution over \([0, 2\pi]\); uncorrelated with \( w[n] \) and between different \( i \)

**Recall: Single Complex Exponential Case**

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

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

\[ E\left[x[n]x^*[n-k]\right] = E\left[A \exp(\left[j(2\pi f_0 + k + \phi)\right]) \cdot A \exp(\left[j(2\pi f_0 n-k + \phi)\right])\right] \\
= A^2 \exp(\left[j2\pi f_0 k\right]) \quad \therefore \text{\( x[n] \) is zero-mean } \quad \text{m.s. with } r_x(k) = A^2 \exp(\left[j2\pi f_0 k\right]) + \sigma_w^2 \delta[k] \]

\[ y[n] = x[n] + \text{white noise; } E[\text{w}(n)\text{w}^*[n-k]] = \delta_{k=0}^\text{d.w.} \]

\[ r_y(k) = E[y[n]y^*[n-k]] = E[\left(x[n] + \text{w}(n)\right)\left(x^*[n-k] + \text{w}^*[n-k]\right)] \\
= r_x[k] + r_w[k] \quad (\because E[\text{w}(n)\text{w}(n)] = 0 \text{ uncorrelated}) \\
= A^2 \exp(\left[j2\pi f_0 k\right]) + \sigma_w^2 \delta[k] \]

**Deriving Autocorrelation Function**

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

\[ r_\alpha(k) = E[x[n]x^*[n-k]] = E\left[\sum_{i=1}^{p} s_i[n] + w[n]\right] \cdot \sum_{i=1}^{p} s_m^*[n-k] + w^*[n-k] \]

- \( E[s_i[n]s_m^*[n-k]] = E[s_i[n]]E[s_m^*[n-k]] = 0 \quad \text{for } l \neq m \)
- \( r_\alpha(k) = A_m^2 e^{j2\pi f_0 k} \quad \text{for } l = m \)
- \( E[s_i[n]w^*[n-k]] = E[s_i[n]]E[w[n-k]] = 0 \)
- \( E[w[n]w^*[n-k]] = \sigma_w^2 \cdot \delta[k] \)

\[ \Rightarrow r_\alpha(k) = E[x[n]x^*[n-k]] = \sum_{i=1}^{p} A_i^2 e^{j2\pi f_0 k} + \sigma_w^2 \delta[k] \]

**Deriving Correlation Matrix**

- May bring \( r_\alpha(k) \) into the correlation matrix
- Or from the expectation of vector’s outer product and use the correlation analysis from last page

\[ x[n] = \sum_{i=1}^{p} s_i[n] + w[n] \]

\[ R_x = E[x[n]x^H[n]] = E\left[\left[\sum_{i=1}^{p} s_i[n] + w[n]\right] \cdot \left[\sum_{m=1}^{p} s_m^H[n] + w^H[n]\right]\right] \]

\[ \Rightarrow R_x = \sum_{i=1}^{p} P_i s_i e_i^H + \sigma_w^2 I \]
Summary: Correlation Matrix for the Process

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

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

\[ R_x = R_s + R_w \]

where

\[ e_i = [1, e^{j2\pi f_i}, e^{j4\pi f_i}, \ldots, e^{j2\pi f_i(N-1)}]^T \]

Correlation Matrix for the Process (cont'd)

\[ R_s = \sum_{i=1}^{p} P_i e_i e_i^H \]

\( e_i e_i^H \) has rank 1 (all columns are related by a factor)

The MxM matrix \( R_s \) has rank \( p \), and has only \( p \) nonzero eigenvalues.

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

\[ R_w = \sigma_w^2 I \quad \rightarrow \text{full rank} \]

\[ R_s = \sum_{i=1}^{p} P_i e_i e_i^H \]

where

\[ e_i = [1, e^{j2\pi f_i}, e^{j4\pi f_i}, \ldots, e^{j2\pi f_i(N-1)}]^T \]
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 \( R_x \) 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 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.
    \[ A = \sum_{i=1}^{n} \lambda_i \, v_i v_i^H = \sum_{i=1}^{n} \lambda_i v_i v_i^H = \sum_{i=1}^{n} \lambda_i v_i v_i^H \]
  - For any \( nxn \) Hermitian matrix
    - There exists a set of \( n \) orthonormal eigenvectors
    - Thus \( V \) is unitary for Hermitian
      matrix \( A \), and
      \[ A = \begin{bmatrix} \lambda_1 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} = \begin{bmatrix} \lambda_1 v_1 v_1^H & \cdots & \lambda_n v_n v_n^H \end{bmatrix} \]

(see Hayes Sec.2.3.9 review of linear algebra)

Eigen Analysis of the Correlation Matrix

Let \( v_i \) be an eigenvector of \( R_x \) with the corresponding eigenvalue \( \lambda_i \), i.e., \( R_x v_i = \lambda_i v_i \)

\[ R_x v_i = \sum_{j=1}^{p} \sigma_j v_j v_j^H = \lambda_i v_i \]

\[ R_s v_i = -\sum_{j=1}^{p} \sigma_j v_j v_j^H = \lambda_i v_i \]

\[ \lambda_i = \left\{ \begin{array}{c} \lambda_i \quad \text{(R_s has p nonzero eigenvalues)} \\
\end{array} \right. \]

(see Hayes Sec.2.3.9 review of linear algebra)
**Eigen Analysis of the Correlation Matrix**

Let \( v_i \) be an eigenvector of \( R_x \) with the corresponding eigenvalue \( \lambda_i \), i.e., \( R_x v_i = \lambda_i v_i \)

\[ R_x v_i = R_s v_i + \sigma_w^2 v_i = \lambda_i v_i \]

i.e., \( v_i \) is also an eigenvector for \( R_s \), and the corresponding eigenvalue is

\[ \lambda_i = \left( \lambda_i^{(s)} + \sigma_w^2 > 0, \quad i=1,2, \ldots, p \quad \text{if } R_s \text{ has } p \text{ nonzero eigenvalues} \right) \]

**Signal Subspace and Noise Subspace**

For \( i = p+1, \ldots, M \):

\[ R_s v_i = 0 \times v_i \]

Also,

\[ R_s = S D S^H \]

\[ S D S^H v_i = 0 \quad \text{for } i = p+1, \ldots, M \]

i.e., the \( p \) column vectors are linearly independent

\[ \Rightarrow S^H v_i = 0 \]

Since \( S = [e_1, \ldots, e_p] \Rightarrow e_i^H v_i = 0, \quad i = p+1, \ldots, M \)

**Relations Between Signal and Noise Subspaces**

Since \( R_x \) and \( R_s \) are Hermitian matrices, the eigenvectors are orthogonal to each other:

\[ v_i \perp v_j \quad \forall \ i \neq j \]

Recall

\[ \text{span} \{ e_1, \ldots, e_p \} \perp \text{span} \{ v_{p+1}, \ldots, v_M \} \]

So it follows that

\[ v_i \perp v_j \quad \forall \ i \neq j \]

**Signal Subspace and Noise Subspace**

For \( i = p+1, \ldots, M \):

\[ R_s v_i = 0 \times v_i \]

Also,

\[ R_s = S D S^H \]

\[ S D S^H v_i = 0 \quad \text{for } i = p+1, \ldots, M \]

i.e., the \( p \) column vectors are linearly independent

\[ \Rightarrow S^H v_i = 0 \]

Since \( S = [e_1, \ldots, e_p] \Rightarrow e_i^H v_i = 0, \quad i = p+1, \ldots, M \)

**Signal Subspace and Noise Subspace**

For \( i = p+1, \ldots, M \):

\[ R_s v_i = 0 \times v_i \]

Also,

\[ R_s = S D S^H \]

\[ S D S^H v_i = 0 \quad \text{for } i = p+1, \ldots, M \]

i.e., the \( p \) column vectors are linearly independent

\[ \Rightarrow S^H v_i = 0 \]

Since \( S = [e_1, \ldots, e_p] \Rightarrow e_i^H v_i = 0, \quad i = p+1, \ldots, M \)

**Signal Subspace and Noise Subspace**

For \( i = p+1, \ldots, M \):

\[ R_s v_i = 0 \times v_i \]

Also,

\[ R_s = S D S^H \]

\[ S D S^H v_i = 0 \quad \text{for } i = p+1, \ldots, M \]

i.e., the \( p \) column vectors are linearly independent

\[ \Rightarrow S^H v_i = 0 \]

Since \( S = [e_1, \ldots, e_p] \Rightarrow e_i^H v_i = 0, \quad i = p+1, \ldots, M \)

**Signal Subspace and Noise Subspace**

For \( i = p+1, \ldots, M \):

\[ R_s v_i = 0 \times v_i \]

Also,

\[ R_s = S D S^H \]

\[ S D S^H v_i = 0 \quad \text{for } i = p+1, \ldots, M \]

i.e., the \( p \) column vectors are linearly independent

\[ \Rightarrow S^H v_i = 0 \]

Since \( S = [e_1, \ldots, e_p] \Rightarrow e_i^H v_i = 0, \quad i = p+1, \ldots, M \)
Relations Between Signal and Noise Subspaces

Since $R_x$ and $R_s$ are Hermitian matrices, the eigenvectors are orthogonal to each other:

$$v_i \perp v_j \quad \forall \ i \neq j$$

$\Rightarrow \text{span}\{v_1, \ldots, v_p\} \perp \text{span}\{v_{p+1}, \ldots, v_M\}$

Recall $\text{span}\{e_1, \ldots, e_p\} \perp \text{span}\{v_{p+1}, \ldots, v_M\}$,

So it follows that

$$\text{span}\{e_1, \ldots, e_p\} = \text{span}\{v_1, \ldots, v_p\}$$

Frequency Estimation Function: General Form

Recall $\mathbf{e}_i^H \mathbf{v}_i = 0$ for $l=1, \ldots, p; \ i = p+1, \ldots, M$

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

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

We form a frequency estimation function

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

Discussion: Complex Exponential Vectors

$$e(f) = [1, e^{-j2\pi f_1}, e^{-j4\pi f} \ldots, e^{-j2\pi(M-1)f}]^T$$

$$\mathbf{e}_i^H(f_1) \mathbf{e}(f_2) = \sum_{k=0}^{M-1} e^{j2\pi(f_1-f_2)k} = \frac{1-e^{j2\pi(f_1-f_2)M}}{1-e^{j2\pi(f_1-f_2)}}$$

If $f_1 - f_2 = \frac{a}{M}$ for some integer $a \Rightarrow \mathbf{e}_i^H(f_1) \mathbf{e}(f_2) = 0$

If $f_1 - f_2 = \frac{a}{M}$ for some integer $a \Rightarrow \mathbf{e}_i^H(f_1) \mathbf{e}(f_2) = 0$

We form a frequency estimation function

$$\hat{P}(f) = \frac{1}{\sum_{i=p+1}^{M} \alpha_i |e(f)^H v_i|^2}$$

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

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

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

- The equivalent frequency estimation function is:

\[
\hat{P}(f) = \frac{1}{\left| e(f)^H v_{\min} \right|^2}
\]

**Estimating the Amplitudes**

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

\[
\begin{align*}
R_x v_i &= \lambda_i v_i \quad (i = 1, 2, \ldots, p) \\
\text{normalize } v_i &\text{ s.t. } v_i^H v_i = 1
\end{align*}
\]

Recall

\[
R_x = \sum_{k=1}^{p} P_k e_k e_k^H + \sigma_w^2 I
\]

**Estimating the Amplitudes**

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

\[
\begin{align*}
R_x v_i &= \lambda_i v_i \quad (i = 1, 2, \ldots, p) \\
\text{normalize } v_i &\text{ s.t. } v_i^H v_i = 1
\end{align*}
\]

Recall

\[
R_x = \sum_{k=1}^{p} P_k e_k e_k^H + \sigma_w^2 I
\]

\[
\Rightarrow \sum_{k=1}^{p} P_k \left| e_k^H v_i \right|^2 = \lambda_i - \sigma_w^2, \quad i = 1, \ldots, p
\]

DTFT of sig eigenvector \( v_i(\cdot) \) at \( -f_k \)  \(\Rightarrow\) Solve \( p \) equations for \( \{ P_k \} \)
**Interpretation of Pisarenko Method**

Since \( e_i^H v_i = 0, \quad l = 1, 2, \ldots, p \)
\( i = p+1, \ldots, M \), \( v_i \) noise eigenvector
\[
\begin{bmatrix}
v_i(0) \\
v_i(1) \\
\vdots \\
v_i(M-1)
\end{bmatrix} = e_i^H v_i \
\]
\[
\Rightarrow \sum_{k=0}^{M-1} v_i(k)e^{i2\pi f/k} = 0 \quad \text{for} \quad l = 1, 2, \ldots, p
\]

Thus given any \( v_i, i=p+1, \ldots, M \), we can estimate the sinusoidal frequencies by finding the zeros on unit circle from

**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 eigenvectors to form the noise subspace: which of (M-p) eigenvectors shall we use to check orthogonality with \( g(f) \) ?
  
  \[
  ZT[ \{ v_i(0), \ldots, v_i(M-1) \} ] \sim (M-1)^{th} \text{ order polynomial} \Rightarrow (M-1) \text{ zeros}
  \]
  
  - \( p \) 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
  
  \[
  ZT[ \{ v_i(0), \ldots, v_i(M-1) \} ] \sim (M-1)^{th} \text{ order polynomial} \Rightarrow (M-1) \text{ zeros}
  \]
  
  - \( p \) 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 eigenvectors
**MUSIC Algorithm**

The frequency estimation function

\[
\hat{\mathbf{P}}_{\text{music}}(f) = \frac{1}{\sum_{k=1}^{M} |e(f) v_k|^2} e^H(f) \mathbf{V} e(f)
\]

where \( e(f) = \left[ \begin{array}{c} e^{j2\pi f} \\ \vdots \\ e^{-j2\pi f(M-1)} \end{array} \right] \), \( \mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_M] \)

**Example-1**

(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)

| Method       | \( \hat{p}_{\text{pnl}}(\omega) = \frac{1}{|\mathbf{v}_{\text{null}}|^2} \) |
|--------------|----------------------------------------------------------------------------------|
| MUSIC        | \( \hat{p}_{\text{music}}(\omega) = \frac{1}{\sum_{k=1}^{M} |\mathbf{v}_k|^2} \) |
| Eigenvector  | \( \hat{p}_{\text{eig}}(\omega) = \frac{1}{\sum_{k=1}^{M} |\mathbf{e}_k|^2} \) |
| Minimum Norm | \( \hat{p}_{\text{mn}}(\omega) = \frac{1}{|\mathbf{a}|^2} \); \( a = \mathbf{v}_{\text{null}} \) |

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