Direction of Arrival Estimation: Subspace Methods

Bhaskar D Rao
University of California, San Diego
Email: brao@ucsd.edu
1. Optimum Array Processing, H. L. Van Trees


Narrow-band Signals

\[
x(\omega_c, n) = \mathbf{V}(\omega_c, k_s)F_s[n] + \sum_{l=1}^{D-1} \mathbf{V}(\omega_c, k_l)F_l[n] + \mathbf{Z}[n]
\]

Assumptions

- \(F_s[n], F_l[n], l = 1, \ldots, D - 1\), and \(\mathbf{Z}[n]\) are zero mean
- \(E(\lvert F_s[n]\rvert^2) = p_s,\) and \(E(\lvert F_l[n]\rvert^2 = p_l, l = 1, \ldots, D - 1,\) and \(E(\mathbf{Z}[n]\mathbf{Z}^H[n]) = \sigma_z^2\mathbf{I}\)
- All the signals/sources are uncorrelated with each other and over time: \(E(F_l[n]F_m^*[p]) = p_l\delta[l - m]\delta[n - p]\) and \(E(F_l[n]F_s^*[p]) = 0\)
- The sources are uncorrelated with the noise: \(E(\mathbf{Z}[n]F_l^*[m]) = \mathbf{0}\)
Adaptive Beamforming

- MVDR, MPDR, and LMMSE beamformers and their relationship
- Generalized Sidelobe Canceller
- Transmit and Receive Beamforming
- MVDR Power Spectrum Estimation
Direction of Arrival (DOA) Estimation

\[ x[n] = \sum_{l=0}^{D-1} V_l F_l[n] + Z[n], \quad n = 0, 1, \ldots, L - 1 \]

Here we have chosen to denote \( V_s \) by \( V_0 \) for simplicity of notation. Note \( V_l = V(k_l) \). We will refer to the \( L \) vector measurements as \( L \) snapshots.

Goal: Estimate the direction of arrival \( k_l \) or equivalently \( (\theta_l, \phi_l) \), \( l = 0, \ldots, D - 1 \).
Can also consider estimating the source powers \( p_l, l = 0, 1, \ldots, D - 1 \), and noise variance \( \sigma_z^2 \).

ULA:
\[ V(\psi) = [e^{-j \frac{N-1}{2} \psi}, \ldots, e^{j \frac{N-1}{2} \psi}]^T = e^{-j \frac{N-1}{2} \psi} [1, e^j \psi, e^{j2\psi}, \ldots, e^{j(N-1)\psi}]^T \]
Note for a ULA, \( V(\psi) = JV^*(\psi) \), where \( J \) is all zeros except for ones along the anti-diagonal. Also note \( J = J^2 \).
ULA and Time Series

\[
\begin{bmatrix}
    x_0[n] \\
    x_1[n] \\
    \vdots \\
    x_{N-1}[n]
\end{bmatrix}
= \sum_{l=0}^{D-1} \begin{bmatrix}
    1 \\
    e^{j\psi_l} \\
    \vdots \\
    e^{j(N-1)\psi_l}
\end{bmatrix}
\begin{bmatrix}
    e^{-j\frac{N-1}{2}\psi_l} F_l[n] + Z[n] \end{bmatrix} , n = 0, 1, \ldots, L-1
\]

Considerable similarity with a time series problems composed of complex exponentials \( x[n] = \sum_{l=0}^{D-1} c_l e^{j\omega_l n} + z[n] \), \( n = 0, 1, \ldots, (M-1) \). In vector form

\[
\begin{bmatrix}
    x[0] \\
    x[1] \\
    \vdots \\
    x[M-1]
\end{bmatrix}
= \sum_{l=0}^{D-1} \begin{bmatrix}
    1 \\
    e^{j\omega_l} \\
    \vdots \\
    e^{j(M-1)\omega_l}
\end{bmatrix} \begin{bmatrix}
    c_l + \\
    z[0] \\
    z[1] \\
    \vdots
\end{bmatrix}
\]

The time series, sum of exponentials, problem with \( M \) samples is equivalent to a ULA with \( M \) sensors and one (vector) measurement or one snapshot.
In array processing we have a limited number of sensors $N$ but can compensate potentially with $L$ snapshots. The time series problem is equivalent to one snapshot but a potentially large array if the number of samples $M$ is large.

One complex exponential (Noise free):

$$x[n] = c_0 e^{j\omega_0 n}, \quad n = 0, 1, \ldots, M - 1.$$  

Interested in the frequency $\omega_0$.
Potential solution: Take the FFT and look at the peak.

Another option is to note that $x[n] = c_0 e^{j\omega_0(n-1)} e^{j\omega_0} = e^{j\omega_0} x[n - 1]$.

In the noise free case, 2 samples is enough.

Can solve a least squares problem $\min_a \sum_{n=1}^{M-1} |x[n] - a x[n - 1]|^2$ and find $\omega_0$ from the estimated $a$. 
More than one exponentials

Suppose we have two exponentials:
\[ x[n] = c_0 e^{j\omega_0 n} + c_1 e^{j\omega_1 n}, n = 0, 1, \ldots, M - 1. \]

A FFT based approach may fail to resolve the frequencies if they are close and the number of samples \( M \) is small. Thumb rule \( |\omega_0 - \omega_1| > \frac{2\pi}{M} \).

Can show \( x[n] = a_1 x[n - 1] + a_2 x[n - 2] \) and so one can solve \( a_1 \) and \( a_2 \) using a least squares technique.

In the absence of noise we just need 4 samples to set up a system of equations to solve for the parameters.

\[
\begin{bmatrix}
  x[n] \\
  x[n - 1]
\end{bmatrix}
= 
\begin{bmatrix}
  x[n - 1] & x[n - 2] \\
  x[n - 2] & x[n - 3]
\end{bmatrix}
\begin{bmatrix}
  a_1 \\
  a_2
\end{bmatrix}
\]

How to get the frequencies from the parameters \( a_1 \) and \( a_2 \)?

Can show that the zeros of the filter \( H(z) = 1 - a_1 z^{-1} - a_2 z^{-2} \) are located at \( e^{j\omega_0} \) and \( e^{j\omega_1} \).
**D exponents**

Suppose we have a series composed of \( D \) exponentials:

\[
x[n] = \sum_{l=0}^{D-1} c_l e^{j\omega_l n}, \quad n = 0, 1, \ldots, M - 1.
\]

A FFT based approach can be used with drawbacks as before: resolution and large number of samples may be needed.

Can show \( x[n] = \sum_{k=1}^{D} a_k x[n - k] \) and so one can solve \( a_k, k = 1, \ldots, D \) using a least squares technique.

In the absence of noise we just need \( 2D \) samples to set up a system of equations to solve for the parameters.

\[
\begin{bmatrix}
x[n] \\
x[n-1] \\
\vdots \\
x[n-D+1]
\end{bmatrix}
= \begin{bmatrix}
x[n-1] & \ldots & x[n-D] \\
x[n-2] & \ldots & x[n-D-1] \\
\vdots & \ddots & \vdots \\
x[n-D] & \ldots & x[n-2D-1]
\end{bmatrix}
\begin{bmatrix}
a_1 \\
a_2 \\
\vdots \\
a_D
\end{bmatrix}
\]

How to get the frequencies from the parameters \( a_k, k = 1, \ldots, D \)?

Can show that the zeros of the filter \( H(z) = 1 - \sum_{k=1}^{D} a_k z^{-k} \) are located at \( e^{j\omega_l}, \quad l = 0, \ldots, D - 1 \).
Claim: A signal with $D$ exponentials, i.e. $x[n] = \sum_{l=0}^{D-1} c_l e^{j\omega_l n}$, is exactly predictable from its $D$ past samples, i.e. $x[n] = \sum_{k=1}^{D} a_k x[n-k]$ or $x[n] - \sum_{k=1}^{D} a_k x[n-k] = 0$

Consider a FIR filter $H(z) = 1 - \sum_{k=1}^{P} b_k x[n-k]$, $P \geq D$, with input $x[n]$ and denote its output by $e[n]$.

$$e[n] = x[n] - \sum_{k=1}^{P} b_k x[n-k] = \sum_{k=1}^{D} H(e^{\omega_k}) c_k e^{j\omega_k n}$$

If we minimize $\sum_n |e[n]|^2$, what is the result?

The output will be zero. This is achieved by making $H(e^{j\omega_k}) = 0$ at the $D$ frequencies. Since the filter is of order $P \geq D$, we can select the filter zeros to meet this criteria and minimize the output.

The excess $(P - D)$ zeros can be placed arbitrarily. Can be a source of confusion when determining the frequency estimates.
Subspace Methods

\[ x(n) = \sum_{l=0}^{D-1} V(k_l) F_l[n] + Z[n] \]

\[ = [V_0, V_1, \ldots, V_{D-1}] \begin{bmatrix} F_s[n] \\ F_1[n] \\ \vdots \\ F_{D-1}[n] \end{bmatrix} + Z[n] \]

\[ = VF[n] + Z[n] \]

where \( V \in C^{N \times D} \), and \( F[n] \in C^{D \times 1} \).

Note that \( E(F[n]) = 0_{D \times 1} \), \( E(F[n]Z^H[n]) = 0_{D \times N} \).
Source Covariance matrix

$$S_F = E(F[n]F^H[n]) = S_F^H = \begin{bmatrix} S^H_s \\ S^H_1 \\ \vdots \\ S^H_{D-1} \end{bmatrix},$$

where $S_F \in C^{D \times D}$, $S_s^H \in C^{1 \times D}$ and the diagonal elements are $p_s, p_1, \ldots, p_{D-1}$.

For uncorrelated sources $S_F$ is a diagonal matrix, i.e. $S_F = \text{diag}(p_s, p_1, \ldots, p_{D-1})$. and $S_s^H = [p_s, 0, \ldots, 0]$

Data Covariance Matrix

$$S_x = S_s + \sigma_z^2 I = VS_FV^H + \sigma_z^2 I$$

$S_x \in C^{N \times N}$, $S_s \in C^{N \times N}$, $S_F \in C^{D \times D}$, and $V \in C^{N \times D}$
Signal and Noise Subspaces

Subspace methods: Based on eigen-decomposition of a covariance matrix

Eigen-decomposition of $S_s$: Note $S_s = V S_F V^H$

$S_s$ has rank $D$, is Hermitian symmetric and positive semi-definite. Hence $S_s$ has $D$ non-zero eigenvalues and corresponding orthogonal eigenvectors

$$S_s = [q_1, q_2, \ldots, q_D] \begin{bmatrix} \lambda_1^s & 0 & \ldots & 0 \\ 0 & \lambda_2^s & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_D^s \end{bmatrix} \begin{bmatrix} q_1^H \\ q_2^H \\ \vdots \\ q_D^H \end{bmatrix} = E_s \Lambda_s E_s^H$$

where $E_s \in \mathbb{C}^{N \times D}$, and $\Lambda_s = \text{diag}(\lambda_i)$. Note $E_s^H E_s = I_{D \times D}$ and $P_s = E_s E_s^H$ is an orthogonal projection matrix onto $\mathcal{R}(E_s)$, i.e. $P_s = P_s^H$ and $P_s = P_s^2$.

Important Observation: $\mathcal{R}(V) = \mathcal{R}(E_s)$ and $\mathcal{R}(E_s)$ is referred to as Signal subspace
Choose $E_n = [q_D, \ldots, q_N]$ such that $E_n^H E_n = I_{(N-D) \times (N-D)}$ and $E_n^H E_s = 0_{(N-D) \times D}$. Defining $E = [E_s, E_n] \in C^{N \times N}$, we have $E^H E = EE^H = I_{N \times N}$.

Important Observation: $\mathcal{R}(E_n) \perp \mathcal{R}(V) = \mathcal{R}(V_s)$. $\mathcal{R}(E_n)$ is referred as Noise subspace.

$$ S_s = E_s \Lambda_s E_s^H = [E_s, E_n] \begin{bmatrix} \Lambda_s & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} E_s^H \\ E_n^H \end{bmatrix} = E \begin{bmatrix} \Lambda_s & 0 \\ 0 & 0 \end{bmatrix} E^H $$

$$ \sigma_z^2 I = \sigma_z^2 EE^H = E (\sigma_z^2 I) E^H = E \begin{bmatrix} \sigma_z^2 & 0 & \cdots & 0 \\ 0 & \sigma_z^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_z^2 \end{bmatrix} E^H $$

In fact $E$ can be replaced by any orthonormal matrix in the expansion for $\sigma_z^2 I$. For subspace methods, choosing a matrix compatible with $S_s$ is more useful.
Eigen decomposition of Data Covariance Matrix

\[ S_x = S_s + \sigma_z^2 I \]

\[ = E \begin{bmatrix} \Lambda_s & 0 \\ 0 & 0 \end{bmatrix} E^H + E \begin{bmatrix} \sigma_z^2 & 0 & \ldots & 0 \\ 0 & \sigma_z^2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma_z^2 \end{bmatrix} E^H \]

\[ = E \begin{bmatrix} \Lambda_s + \sigma_z^2 I_{D \times D} & 0 \\ 0 & \sigma_z^2 I_{(N-D) \times (N-D)} \end{bmatrix} E^H \]

\[ = E_s (\Lambda_s + \sigma_z^2 I_{D \times D}) E_s^H + \sigma_z^2 E_n E_n^H \]

\[ = \sum_{l=1}^{D} (\lambda_l^s + \sigma_z^2) q_l q_l^H + \sigma_z^2 \sum_{l=(D+1)}^{N} q_l q_l^H \]
Observations

\[ S_x = S_s + \sigma_z^2 I \]

\[ = E_s (\Lambda_s + \sigma_z^2 I_{D \times D}) E_s^H + \sigma_z^2 E_n E_n^H \]

\[ = \sum_{l=1}^{D} (\lambda_l^s + \sigma_z^2) q_l q_l^H + \sigma_z^2 \sum_{l=(D+1)}^{N} q_l q_l^H \]

- \( S_x \) has eigenvalues \( \lambda_l = \lambda_l^s + \sigma_z^2, l = 1, \ldots, D \), and \( \lambda_l = \sigma_z^2, l = (D + 1), \ldots, N \)

- The number of sources \( D \) can be identified by examining the multiplicity of the smallest eigenvalue which is expected to be \( \sigma_z^2 \) with multiplicity \( (N - D) \)

- The eigenvectors corresponding to the signal subspaces and noise subspaces are readily available from the eigen decomposition of \( S_x \), i.e. \( E_s \) and \( E_n \) are easily extracted.
MUltiple SIgnal Classification (MUSIC)

MUSIC uses the noise subspace spanned by the "noise" eigenvectors $E_n$. Since $\mathcal{R}(V) = \mathcal{R}(E_s) \perp \mathcal{R}(E_n)$, we have $V_l \perp \mathcal{R}(E_n), l = 0, 1, \ldots, (D - 1)$ or $V_l^H E_n = 0, l = 0, 1, \ldots, (D - 1)$ Many options to use this orthogonality property. All methods start with $S_x$, usually estimated from the data.

MUSIC Algorithm (also known as Spectral-MUSIC)

- Perform an eigen-decomposition of $S_x$
- Estimate $D$ and the signal subspace $E_s$ and the noise subspace $E_n$
- Compute the spatial MUSIC spectrum as follows:

$$P_{MUSIC}(k) = \frac{1}{|V^H(k)E_n|^2}$$

Note that $|V^H(k)E_n|^2 = V^H(k)E_nE_n^H V(k) = V^H(k)P_n V(k)$ where $P_n = E_nE_n^H$

- The peaks in $P_{MUSIC}(k)$ gives us the DOA.
Fig. 3. Example of azimuth-only DF performance.
ULA: \( \mathbf{V}(\psi) = e^{-j\frac{N-1}{2}\psi} [1, e^{j\psi}, e^{j2\psi}, \ldots, e^{j(N-1)\psi}]^T \). Will drop \( e^{-j\frac{N-1}{2}\psi} \) for this discussion for simplicity and because it does not matter.

\[
|\mathbf{V}^H(\psi)\mathbf{E}_n|^2 = \sum_{l=D+1}^{N} |\mathbf{V}^H(\psi)\mathbf{q}_l|^2 = \sum_{l=D+1}^{N} |\mathbf{q}_l(e^{j\psi})|^2
\]

where \( \mathbf{q}_l(e^{j\psi}) = \mathbf{V}^H(\psi)\mathbf{q}_l = \sum_{m=0}^{N-1} \mathbf{q}_l(m)e^{-jm\psi} \), with \( \mathbf{q}_l = [\mathbf{q}_l[0], \mathbf{q}_l[1], \ldots, \mathbf{q}_l[N-1]]^T \)

\( \mathbf{q}_l(e^{j\psi}) \) is the Fourier transform of the impulse response of a filter formed from the eigenvector \( \mathbf{q}_l \)

Defining \( B(e^{j\psi}) = \sum_{l=D+1}^{N} |\mathbf{q}_l(e^{j\psi})|^2 \). The MUSIC spectrum is

\[
P_{\text{MUSIC}}(\psi) = \frac{1}{B(e^{j\psi})}, \text{ and it will have peaks at } \psi_l, l = 0, 1, \ldots, (D - 1) \text{ corresponding to the DOA}
\]

Then \( B(e^{j\psi_l}) = 0, l = 0, 1, \ldots, (D - 1) \)
With $B(e^{j\psi}) = \sum_{l=D+1}^{N} |q_l(e^{j\psi})|^2$, in the z-domain we have

$B(z) = \sum_{l=D}^{N} q_l(z)q_l^*(\frac{1}{z^*})$. Alternately

$B(z) = \sum_{l=D+1}^{N} B_l(z)$ where

$B_l(z) = q_l(z)q_l^*(\frac{1}{z^*}) = \sum_{l=-(N-1)}^{N-1} r_{ql}[m]z^{-m}$,

Each of the $B_l(z)$ represents a Moving Average Spectrum and the sum $B(z)$ is also a Moving Average spectrum

$$B(z) = \sum_{l=-(N-1)}^{N-1} r_{q}[m]z^{-m} = H_{rm}(z)H_{rm}^*(\frac{1}{z^*}),$$

where $r_{q}[m]$ is an autocorrelation sequence and $H_{rm}(z)$ is obtained by spectral factorization

$B(e^{j\psi_l}) = 0, l = 0, 1, \ldots, (D - 1)$ implies $H_{rm}(z)$ will have $D$ of it’s zeros at $e^{j\psi_l}$, where $\psi_l$ corresponds to the DOA

The remaining $(N - 1 - D)$ zeros will be inside the unit circle assuming we use a minimum-phase spectral factor.
Root-MUSIC Algorithm

- Perform an eigen-decomposition of $S_x$
- Estimate $D$ and the signal subspace $E_s$ and the noise subspace $E_n$
- Compute $q_l(z) = \sum_{m=0}^{N-1} q_l[m] z^{-m}, l = D + 1, \ldots, N - 1$.
- Compute $B(z) = \sum_{l=D+1}^{N} B_l(z)$ where $B_l(z) = q_l(z) q_l^* \left( \frac{1}{z^*} \right) = \sum_{m=-(N-1)}^{N-1} r_{ql}[m] z^{-m}$
- Perform a spectral factorization to find a minimum phase factor $H_{rm}(z)$, i.e. $B(z) = H_{rm}(z) H_{rm}^* \left( \frac{1}{z^*} \right)$.
- The $D$ roots on/close to the unit circle are identified and their argument used to estimate the DOA
Root-MUSIC and Spectral-MUSIC Example

Spectral MUSIC: 3 trials

Spectral MUSIC: 3 trials

Root-MUSIC: 6 trials

Two Sources at 18 and 22 degrees, Array
Size 8, 100 Snapshots, SNR 6 db