DOA Estimation using MUSIC and Root MUSIC Methods

EE602- Statistical signal Processing

4/13/2009

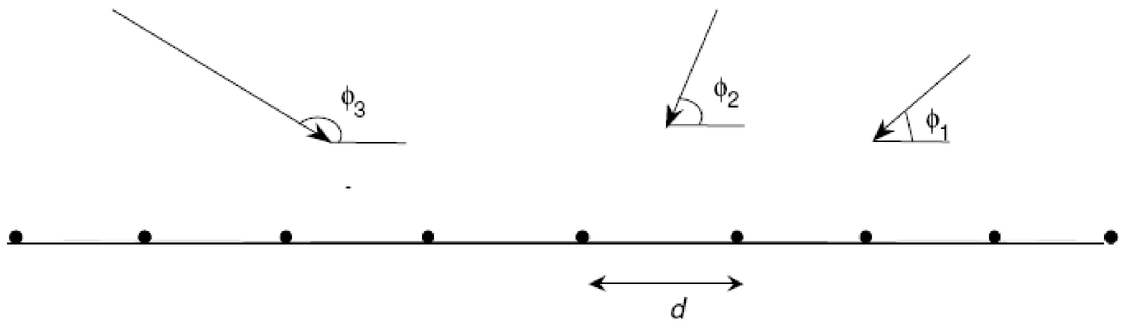
Presented By:

Chhavipreet Singh(Y515)
Siddharth Sahoo(Y5827447)

Table of Contents

1 Introduction

We know that there is a one-to-one relationship between the direction of a signal and the associated received steering vector. It should therefore be possible to invert the relationship and estimate the direction of a signal from the received signals. An antenna array therefore should be able to provide for direction of arrival estimation. There is a Fourier relationship between the beam pattern and the excitation at the array. This allows the direction of arrival (DOA) estimation problem to be treated as equivalent to spectral estimation.

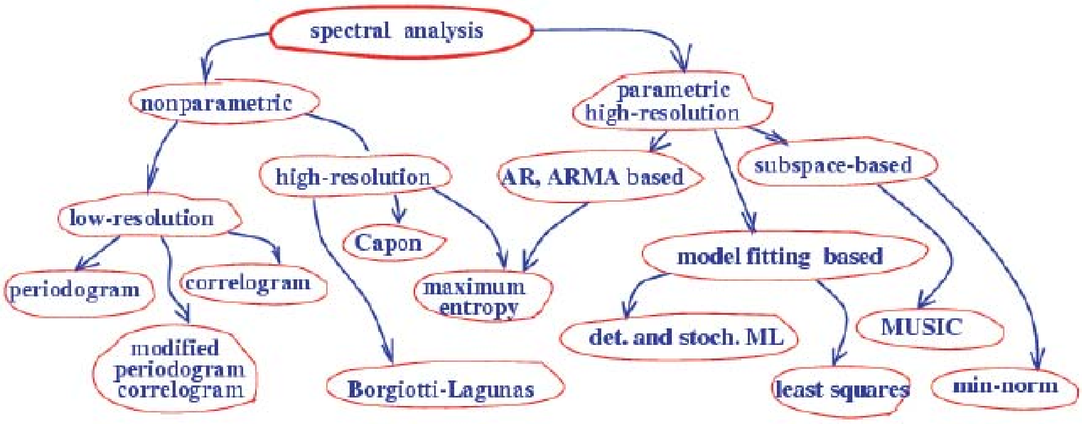


The DOA estimation problem.

The problem set up is shown in the figure above. Several (M) signals impinge on a linear, equispaced, array with N elements, each with direction i. The goal of DOA estimation is to use the data received at the array to estimate _i, i = 1, . . .M. It is generally assumed that M < N, though there exist approaches (such as maximum likelihood estimation) that do not place this constraint.

In practice, the estimation is made difficult by the fact that there are usually an unknown number of signals impinging on the array simultaneously, each from unknown directions and with unknown amplitudes. Also, the received signals are always corrupted by noise. Nevertheless, there are several methods to estimate the number of signals and their directions.

2 Some of the Several Approaches to Spectral Estimation



3 The Cramer - Rao Bound

We begin by realizing that the DOA is a parameter estimated from the received data. The minimum variance in this estimate is given by the Cramer-Rao bound (CRB).

The CRB theorem: Given a length-N vector of received signals x dependent on a set of P parameters = $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_P]^T$, corrupted by additive noise,

$$\mathbf{x} = \mathbf{v}(\Theta) + \mathbf{n}$$

where v(Θ) is a known function of the parameters, the variance of an unbiased estimate of the p-th parameter, θ_p, is greater than the Cramer Rao bound

$$\mathrm{var}(\theta_p) \geq J^{-1}_{pp}$$

where J^{-1}_{pp} is the p-th diagonal entry of the inverse of the Fisher information matrix J whose (i, j)th is given by

$$\mathbf{J}_{ij} = \mathrm{E}\left\{ \frac{\partial^2}{\partial\theta_i \partial\theta_j} \left[\ln f_{\mathbf{X}}(\mathbf{x}/\theta)\right] \right\}$$

where, $f_X(x/\Theta)$ is the pdf of the received vector given the parameters Θ and E{ • } represents statistical expectation.

4 CRB for DOA Estimation

The model under consideration is a number of signals impinging at the array, corrupted by white noise. We will derive the CRB for a single signal corrupted by noise (M = 1). The data model is therefore

$$\mathbf{x} = \acute{\alpha}\mathbf{s}(\phi) + \mathbf{n}$$

where s(Ø) represents the steering vector of the signal whose direction (Ø) we are attempting to estimate. The noise vector n is zero-mean Gaussian with covariance $\boldsymbol{\sigma^2 I}$.. Here ά and ø are modelled as an unknown, but deterministic, constants, i.e., E{x} = άs(ø). In CRB literature, ά would be considered a nuisance parameter, which must be accounted for because it is unknown. Finally, ά itself represents two unknowns, its real and imaginary parts, or equivalently its magnitude and phase. Let _ = ae^{jb}. Therefore, ϴ = [a, b, ø $]^T$. In our case,

$$\begin{aligned} \mathbf{v}(\boldsymbol{\theta}) &= \alpha\,\mathbf{s}(\phi) \\ f_{\mathbf{X}}(\mathbf{x}/\boldsymbol{\theta}) &= Ce^{-(\mathbf{x}-\mathbf{v})^H\mathbf{R}^{-1}(\mathbf{x}-\mathbf{v})} \end{aligned}$$

where R = $\boldsymbol{\sigma^2 I}$ and C is a normalization constant.

$$\begin{aligned} \Rightarrow \ln f_{\mathbf{X}}(\mathbf{x}/\boldsymbol{\theta}) &= \ln C - \frac{(\mathbf{x}-\mathbf{v})^H(\mathbf{x}-\mathbf{v})}{\sigma^2} \\ &= \ln C + \frac{-\mathbf{x}^H\mathbf{x} + \mathbf{v}^H\mathbf{x} + \mathbf{x}^H\mathbf{v} - \mathbf{v}^H\mathbf{v}}{\sigma^2} \\ &= \ln C + \frac{-\mathbf{x}^H\mathbf{x} + \alpha^*\mathbf{s}^H(\phi)\mathbf{x} + \alpha\mathbf{x}^H\mathbf{s}(\phi) - |\alpha|^2\,\mathbf{s}^H(\phi)\,\mathbf{s}(\phi)}{\sigma^2} \end{aligned}$$

Since we are interested in taking derivatives of this expression, we can ignore these terms. Focusing on the important terms and writing the result in terms of the parameters ϴ = [a, b, ø $]^T$,

$$g(\boldsymbol{\theta}) = \frac{1}{\sigma^2}\left[ae^{-jb}\mathbf{s}^H(\phi)\mathbf{x} + ae^{jb}\mathbf{x}^H\mathbf{s}(\phi) - a^2\mathbf{s}^H(\phi)\mathbf{s}(\phi)\right]$$

Also,

$$\mathbf{J} = \mathrm{E}\left\{ \begin{bmatrix} \frac{\partial^2 g}{\partial a^2} & \frac{\partial^2 g}{\partial a \partial b} & \frac{\partial^2 g}{\partial a \partial \phi} \\ \frac{\partial^2 g}{\partial b \partial a} & \frac{\partial^2 g}{\partial b^2} & \frac{\partial^2 g}{\partial b \partial \phi} \\ \frac{\partial^2 g}{\partial \phi \partial a} & \frac{\partial^2 g}{\partial \phi \partial b} & \frac{\partial^2 g}{\partial \phi^2} \end{bmatrix} \right\}$$

We choose any convenient form of the steering vector, including the form where the phase reference is at the center of the array, i.e., for an array with an odd number of elements

$$\mathrm{s}(\phi) = \left[z^{-(N-1)/2},\ z^{-(N-3)/2}\ \ldots,\ z^{-1},\ 1,\ z,\ \ldots,\ z^{(N-3)/2},\ z^{(N-1)/2} \right]^T$$

where

$$z = e^{jkd\cos\phi}$$

Consider the partial derivative of s(ø) as a function of ø. Denoting this vector to be s_1(ø),

$$\begin{aligned} \mathrm{s}_1(\phi) &= \frac{\partial \mathrm{s}(\phi)}{\partial \phi} \\ &= -jkd\sin\phi \left[\frac{-(N-1)}{2} z^{-(N-1)/2},\ \frac{-(N-3)}{2} z^{-(N-3)/2},\ \ldots \right. \\ &\qquad \left. -z^{-1},\ 0,\ z\ \ldots\ \frac{(N-3)}{2} z^{(N-3)/2},\ \frac{(N-1)}{2} z^{(N-1)/2} \right]^T \end{aligned}$$

$$\Rightarrow \mathrm{s}_1(\phi)_n = -jkdn\sin\phi\, z^n$$

where s_1(ø)n is the n-th element in the vector s_1. Similarly, denote as s_2(ø) the second derivative of s(ø) with respect to ø.

Using the definitions of s(ø), s1(ø), and z* = z−1 we can derive some terms that will be useful later:

$$
\begin{aligned}
\mathrm{E}[\mathbf{v}] &= \alpha \mathbf{s} = a e^{jb} \mathbf{s} \\
\mathbf{s}^H(\phi)\mathbf{s}(\phi) &= N \\
\mathbf{s}_1^H(\phi)\mathbf{s}(\phi) &= jkd\sin\phi \sum_{n=-(N-1)/2}^{(N-1)/2} n = 0 \\
\mathbf{s}_1^H(\phi)\mathbf{s}_1(\phi) &= (kd)^2 \sin^2\phi \sum_{n=-(N-1)/2}^{(N-1)/2} n^2 \triangleq B^2 (kd)^2 \sin^2\phi
\end{aligned}
$$

where B^2 represents the sum in the last eqn.

Now,

$$
\begin{aligned}
\frac{\partial g}{\partial \phi} &= \frac{1}{\sigma^2}\left[a e^{-jb}\mathbf{s}_1^H(\phi)\mathbf{x} + a e^{jb}\mathbf{x}^H \mathbf{s}_1(\phi) - a^2 \mathbf{s}_1^H(\phi)\mathbf{s}(\phi) - a^2 \mathbf{s}^H(\phi)\mathbf{s}_1(\phi)\right] \\
\Rightarrow \frac{\partial^2 g}{\partial \phi^2} &= \frac{1}{\sigma^2}\left[a e^{-jb}\mathbf{s}_2^H(\phi)\mathbf{x} + a e^{jb}\mathbf{x}^H \mathbf{s}_2(\phi) - a^2 \mathbf{s}_2^H(\phi)\mathbf{s}(\phi) - a^2 \mathbf{s}_1^H(\phi)\mathbf{s}_1(\phi) - \right. \\
&\qquad \left. a^2 \mathbf{s}_1^H(\phi)\mathbf{s}_1(\phi) - a^2 \mathbf{s}^H(\phi)\mathbf{s}_2(\phi)\right] \\
\Rightarrow \mathrm{E}\left[\frac{\partial^2 g}{\partial \phi^2}\right] &= \frac{1}{\sigma^2}\left[a^2 e^{-jb} e^{jb}\mathbf{s}_2^H(\phi)\mathbf{s} + a^2 e^{jb} e^{-jb}\mathbf{s}^H \mathbf{s}_2(\phi) - a^2 \mathbf{s}_2^H(\phi)\mathbf{s}(\phi) \right. \\
&\qquad \left. - a^2 \mathbf{s}_1^H(\phi)\mathbf{s}_1(\phi) - a^2 \mathbf{s}_1^H(\phi)\mathbf{s}_1(\phi) - a^2 \mathbf{s}^H(\phi)\mathbf{s}_2(\phi)\right] \\
\Rightarrow \mathrm{E}\left[\frac{\partial^2 g}{\partial \phi^2}\right] &= -2a^2 \frac{1}{\sigma^2}\mathbf{s}_1^H \mathbf{s}_1, \\
&= -\frac{2a^2 (kd\sin\phi)^2 B^2}{\sigma^2},
\end{aligned}
$$

Using definition of B,

$$
\begin{aligned}
B^2 &= \sum_{n=-\frac{N-1}{2}}^{\frac{N-1}{2}} n^2 = 2\sum_{n=1}^{\frac{N-1}{2}} n^2 \\
&= 2\left(\frac{N-1}{2}\right)\left(\frac{N+1}{2}\right)\frac{N}{6} = \frac{N(N^2-1)}{12}
\end{aligned}
$$

Using |ά| = a, the CRB for the DOA estimation problem is therefore

$$\begin{aligned} \mathrm{var}(\phi) &\geq \left[\mathrm{E}\left(\frac{\partial^2 g}{\partial \phi^2}\right)\right]^{-1} \\ &\geq \frac{6\sigma^2}{|\alpha|^2 N(N^2-1)(kd)^2 \sin^2\phi} \end{aligned}$$

5 MUSIC: MUltiple SIgnal Classification

MUSIC, as are many adaptive techniques, is dependent on the correlation matrix of the data.

$$\begin{aligned} \mathbf{x} &= \mathbf{S}\boldsymbol{\alpha} + \mathbf{n}. \\ \mathbf{S} &= [\mathbf{s}(\phi_1)\ \mathbf{s}(\phi_2)\ \ldots,\ \mathbf{s}(\phi_M)], \\ \boldsymbol{\alpha} &= [\alpha_1,\ \alpha_2\ \ldots\ \alpha_M]^T. \end{aligned}$$

The matrix S is a N ×M matrix of the M steering vectors. Assuming that the different signals to be uncorrelated, the correlation matrix of x can be written as

$$\begin{aligned} \mathbf{R} &= \mathrm{E}\left[\mathbf{x}\mathbf{x}^H\right], \\ &= \mathrm{E}\left[\mathbf{S}\boldsymbol{\alpha}\boldsymbol{\alpha}^H\mathbf{S}^H\right] + \mathrm{E}\left[\mathbf{n}\mathbf{n}^H\right], \\ &= \mathbf{SAS}^H + \sigma^2\mathbf{I}, \\ &= \mathbf{R}_s + \sigma^2\mathbf{I}, \end{aligned}$$

Where,

$$\begin{aligned} \mathbf{R}_s &= \mathbf{SAS}^H \\ \mathbf{A} &= \begin{bmatrix} \mathrm{E}\left[|\alpha_1|^2\right] & 0 & \cdots & 0 \\ 0 & \mathrm{E}\left[|\alpha_2|^2\right] & \cdots & 0 \\ 0 & 0 & \cdots & \mathrm{E}\left[|\alpha_M|^2\right] \end{bmatrix} \end{aligned}$$

The signal covariance matrix, Rs, is clearly a N × N matrix with rank M. It therefore has N−M eigenvectors corresponding to the zero eigenvalue. Let qm be such an eigenvector. Therefore,

$$\begin{aligned} \mathbf{R}_s\mathbf{q}_m &= \mathbf{SAS}^H\mathbf{q}_m = 0, \\ \Rightarrow \mathbf{q}_m^H\mathbf{SAS}^H\mathbf{q}_m &= 0, \\ \Rightarrow \mathbf{S}^H\mathbf{q}_m &= 0 \end{aligned}$$

where this final equation is valid since the matrix A is clearly positive definite.It implies that all N −M eigenvectors (qm) of Rs corresponding to the zero eigenvalue are orthogonal to all M signal steering vectors. This is the basis for MUSIC. Call Qn the N ×(N −M) matrix of these eigenvectors. MUSIC plots the pseudo-spectrum

$$P_{\text{MUSIC}}(\phi) = \frac{1}{\sum_{m=1}^{N-M}\left|\mathbf{s}^H(\phi)\mathbf{q}_m\right|^2} = \frac{1}{\mathbf{s}^H(\phi)\mathbf{Q}_n\mathbf{Q}_n^H\mathbf{s}(\phi)} = \frac{1}{\left|\left|\mathbf{Q}_n^H\mathbf{s}(\phi)\right|\right|^2}$$

The key is that the eigenvectors in Qn can be estimated from the eigenvectors of R. For any eigenvector qm Є Q,

$$\mathbf{R}_s\mathbf{q}_m = \lambda\mathbf{q}_m$$

$$\begin{aligned} \Rightarrow \mathbf{R}\mathbf{q}_m &= \mathbf{R}_s\mathbf{q}_m + \sigma^2\mathbf{I}\mathbf{q}_m, \\ &= (\lambda_m + \sigma^2)\mathbf{q}_m, \end{aligned}$$

i.e. any eigenvector of Rs is also an eigenvector of R with corresponding eigenvalue $\lambda + \sigma^2$. Let
Rs = $Q\Lambda Q^H$. Therefore,

$$\mathbf{R} = \mathbf{Q}\left[\mathbf{\Lambda} + \sigma^2\mathbf{I}\right]\mathbf{Q}^H$$

$$= \mathbf{Q}\begin{bmatrix} \lambda_1+\sigma^2 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2+\sigma^2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_M^2+\sigma^2 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & \sigma^2 \end{bmatrix}\mathbf{Q}^H.$$

Based on this eigendecomposition, we can partition the eigenvector matrix Q into a signal matrix Qs with M columns, corresponding to the M signal eigenvalues, and a matrix Qn, with

(N − M) columns, corresponding the noise eigenvalues (σ2). Note that Qn, the N × (N − M) matrix of eigenvectors corresponding to the noise eigenvalue (σ2), is exactly the same as the matrix of eigenvectors of Rs corresponding to the zero-eigenvalue.Qs defines the signal subspace, while Qn defines the noise subspace.

There are few important observations to be made:

• The m-th signal eigenvalue is given by$\lambda m + \sigma^2 = N|\acute{\alpha}m|^2 + \sigma^2$.

• The smallest eigenvalues of R are the noise eigenvalues and are all equal to σ^2, i.e., one way of distinguishing between the signal and noise eigenvalues (equivalently the signal and noise subspaces) is to determine the number of small eigenvalues that are equal.

• By orthogonality of Q,

$$\mathbf{Q}_s \perp \mathbf{Q}_n$$

Using the final two observations, we see that all noise eigenvectors are orthogonal to the signal steering vectors. This is the basis for MUSIC. Consider the following function of ø:

$$P_{\text{MUSIC}}(\phi) = \frac{1}{\sum_{m=M+1}^{N} |\mathbf{q}_m^H \mathbf{s}(\phi)|^2} = \frac{1}{\mathbf{s}^H(\phi)\mathbf{Q}_n\mathbf{Q}_n^H\mathbf{s}(\phi)},$$

where qm is one of the (N − M) noise eigenvectors. If ø is equal to DOA one of the signals, $\mathbf{s}(\phi) \perp \mathbf{q}_m$ and the denominator is identically zero. MUSIC, therefore, identifies as the directions of arrival, the peaks of the function P_{MUSIC}(ø).

6 Root MUSIC: Model Based Parameter Estimation

There is a significant problem with MUSIC as described above. The accuracy is limited by the discretization at which the MUSIC function P_{MUSIC}(ø) is evaluated. More importantly, it requires
either human interaction to decide on the largest M peaks or a comprehensive search algorithm to determine these peaks. This is an extremely computationally intensive process. Therefore, MUSIC by itself is not very practical - we require a methodology that results directly in numeric values for the estimated directions. This is where Root-MUSIC comes in.

We use a model of the received signal as a function of the DOA - here, the model is the steering vector. The DOA, ø, is a parameter in this model. Based on this model and the received data, we will estimate this parameter.

We define,

$z = e^{jkd \cos ø}$

Then, assuming the receiving antenna is a linear array of equispaced, isotropic, elements,

$$\mathbf{s}(\phi) = \left[1,\ z,\ z^2,\ \ldots,\ z^{N-1}\right]^T,$$
$$\Rightarrow \mathbf{q}_m^H \mathbf{s} = \sum_{n=0}^{N-1} q_{mn}^* z^n = q_m(z),$$

i.e., the inner product of the eigenvector qm and the steering vector s(ø) is equivalent to a polynomial in z. Since we are looking for the directions (ø) where $\mathbf{q}_m \perp \mathbf{s}(\phi)$, m = (M + 1), . . . ,N, we are looking for the roots of a polynomial.

To find the polynomial whose roots we wish to evaluate, we use

$$P_{\text{MUSIC}}^{-1}(\phi) = \mathbf{s}^H(\phi)\mathbf{Q}_n\mathbf{Q}_n^H\mathbf{s}(\phi)$$
$$= \mathbf{s}^H(\phi)\mathbf{C}\mathbf{s}(\phi)$$

where,

$$\mathbf{C} = \mathbf{Q}_n\mathbf{Q}_n^H$$

$$\Rightarrow P_{\text{MUSIC}}^{-1}(\phi) = \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} z^n C_{mn} z^{-m} = \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} z^{(n-m)} C_{mn}$$

The final double summation can be simplified by rewriting it as a single sum by setting l = n−m. The range on l is set by the limits on n and m, i.e. −(N − 1) ≤ l ≤ (N − 1) and

$$\Rightarrow P_{\text{MUSIC}}^{-1}(\phi) = \sum_{l=-(N-1)}^{(N+1)} C_l z^l,$$

$$C_l = \sum_{n-m=l} C_{mn},$$

i.e., C_l is the sum of the elements of C on the n-th diagonal. The eqn. polynomial of degree (2N − 2) with (2N − 2) zeros. However, we can show that not all zeros are independent. If z is a zero of the above polynomial, and of $P^{-1}_{\text{MUSIC}}(ø)$), $1/z^*$ is also a zero of the polynomial. The zeros of $P^{-1}_{\text{MUSIC}}(ø)$ therefore come in pairs.

The steps of Root-MUSIC are:
1. Estimate the correlation matrix R, find its eigen decomposition $R = Q\Lambda Q^H$.
2. Partition Q to obtain Q_n, corresponds to the (N −M) smallest eigenvalues of Q, which spans
the noise subspace. Find $C = Q_n Q^H{}_n$.
3. Obtain C_l by summing the l-th diagonal of C.
4. Find the zeros of the resulting polynomial in terms of (N − 1) pairs.
5. Of the (N − 1) roots within the unit circle, choose the M closest to the unit circle (z_m, m = 1, . . . M).
6. Obtain the directions of arrival using

$$\phi_m = \cos^{-1}\left[\frac{\Im \ln(z_m)}{kd}\right], \quad m = 1, \ \ldots \ M$$