Analysis of the Combined Effects of Finite Samples and Model Errors on Array Processing Performance

MATS VIBERG
Information Systems Laboratory
Stanford University, Stanford, CA 94305

A. LEE SWINDLEHURST
Dept of Elec. and Comp. Engineering
Brigham Young University, Provo, UT 84602

Abstract

This paper concerns the performance of the class of signal subspace fitting algorithms for signal parameter estimation using narrowband sensor array data. The principal sources of estimation error in such applications are the finite sample effects of additive noise and imprecise models for the antenna array and spatial noise statistics. The covariance matrix of the estimation error when all of these error sources are present is found to be the sum of the individual contributions from each component separately. This simplifying fact allows for the derivation of an overall optimal subspace weighting for a particular array and noise covariance error model. In fact, the resulting algorithm yields the lowest possible asymptotic estimation error variance of any method for the model in question.

1 Introduction

Since the end of the 70’s, a number of so-called “high-resolution” subspace based algorithms for array signal processing and parameter estimation have been introduced, see e.g. [1, 2, 3]. Most of these techniques are presented in the context of estimating the directions-of-arrival (DOAs) of multiple co-channel signals using an array of sensors. In any practical situation, the estimated DOAs deviate from their “true” values. The principal sources of errors are the finite sample effects of noise and imprecise models for the array response and noise covariance. The original MUSIC and ESPRIT algorithms [1, 2] are based on the properties of the exact covariance matrix of the array output. More recently, other techniques that take any of the above mentioned effects into account have been proposed. An optimal subspace based technique for finite sample errors only is derived in [3], whereas methods that are less sensitive to various kinds of modeling errors are proposed in [4, 5, 6, 7]. The statistical performance of the above mentioned methods have been studied in, e.g. [3, 8] (finite sample errors only) and [4, 5, 9, 10] (modeling errors only).

The goal of the present paper is to extend the analysis to the combined error case, and to propose an overall optimal estimation procedure. Such a procedure, based on a maximum a posteriori (MAP) approach, is presented in [11] (see also [12]). The technique proposed herein is based on Signal Subspace Fitting (SSF) [3], and is computationally more attractive than the method of [11]. However, it can be shown that the overall optimal SSF technique is asymptotically equivalent to the MAP estimator [13]. The price paid for this result is that a fairly simplistic model of the array response and noise covariance perturbations must be assumed. This model is based on (1) additive random array response errors that are statistically uniform from sensor to sensor, but possibly correlated between different DOAs (with known correlation), and (2) additive random perturbations to the noise covariance, that are independent from element to element.

2 Problem Formulation

Consider an array of m sensors, having arbitrary positions and characteristics. Impinging on the array are the waveforms of d far-field, narrowband point sources, where d < m. The vector of complex sensor outputs is denoted x(t), and is modeled by the following familiar equation:

\[ x(t) = A(B)s(t) + n(t) . \]  (1)

The columns of the \( m \times d \) transfer matrix \( A = [a(\theta_1), \ldots, a(\theta_d)] \) are the so-called array propagation
The vector \( a(\theta_i) \) describes the array response to a unit waveform with parameter(s) \( \theta_i \). Though not necessary, we shall assume that \( \theta_i \) is a real-valued scalar referred to as the \( i^{th} \) direction-of-arrival (DOA).

The \( d \)-vector \( s(t) \) is composed of the complex emitter waveforms as received by the first sensor (the reference) at time \( t \), and the \( m \)-vector \( n(t) \) accounts for additive measurement noise. The array output, is assumed to be sampled at \( N \) distinct time instants. Based on the measurements \( x(1), \ldots, x(N) \), the problem of interest is to determine the DOAs of all emitters. The number of signals, \( d \), is assumed to be known.

The signal waveforms are regarded as deterministic (i.e., fixed) sequences such that the following limit exists

\[
P = \lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} s(t)s^*(t) ,
\]

where \( X^* \) denotes the complex conjugate transpose of \( X \). On the contrary, the noise term, \( n(t) \), is modeled as a stationary, complex Gaussian random process, uncorrelated with the signals.

### 2.1 Perturbation Models

The exact parametrization of the array propagation vectors is unknown in any practical situation. Thus, the available model \( a(\theta) \) may differ from the "true" propagation vector. It will be assumed that the data have actually been generated by the equation

\[
x(t) = (A + \hat{A}) s(t) + n(t) ,
\]

where the columns of \( \hat{A} \) are assumed to be zero-mean random with second-order moments

\[
E[a(\theta_i)a^*(\theta_j)] = \nu_{ij} B \quad (4)
\]

\[
E[a(\theta_i)a^T(\theta_j)] = 0 . \quad (5)
\]

The sensor-to-sensor covariances are collected in the matrix

\[
\Sigma = \{\nu_{ij}\} . \quad (6)
\]

Both \( B \) and \( \Sigma \) are assumed to be available to the user, e.g., from system performance specifications. Note that the error model (4) allows for direction-dependent modeling errors, but the sensor-to-sensor correlation (if any) is independent of \( \theta \).

We will also study the effects of errors in the noise model on algorithm performance. The conditional mean and covariances of the noise given the perturbation, \( \Sigma \), is assumed to be

\[
E[n(t) \mid \Sigma] = 0 \quad (7)
\]

\[
E[n(t)n^T(s) \mid \Sigma] = 0 \quad (8)
\]

\[
E[n(t)n^*(s) \mid \Sigma] = \sigma^2 I + \Sigma \delta_{i,s} , \quad (9)
\]

where \( \sigma^2 \) is the noise power and \( \delta_{i,s} \) is the Kronecker delta. Other than being Hermitean, \( \Sigma \) is treated as a random matrix with independent elements, \( \sigma_{ij} \), of equal variance, \( \mu^2 \):

\[
E[\sigma_{ij}] = 0 \quad (10)
\]

\[
E[\sigma_{ij}\sigma_{kl}] = \mu^2 \delta_{ij}\delta_{kl} . \quad (11)
\]

Thus, the real diagonal elements of \( \Sigma \) are independent of all other elements, whereas the off-diagonal terms are correlated only with their conjugate image. It is assumed that \( \mu^2 \) is known, whereas \( \sigma^2 \) may be unknown.

Since we are interested in studying the combined effects of finite sample errors and modeling errors, the size of the perturbations relative to the number of available snapshots plays a crucial role. To make the contributions of the two error sources of comparable magnitude, we make the artificial assumption that both \( \Sigma \) and \( \mu^2 \) are \( O(1/N) \). The so-obtained analysis is valid up to first order in \( 1/N, \langle \|\Sigma\| \rangle \) and \( \mu^2 \) simultaneously.

### 3 Signal Subspace Fitting Methods

Most parametric estimation methods depend on the measurements only through the sample covariance matrix

\[
\hat{R} = \frac{1}{N} \sum_{t=1}^{N} x(t)x^*(t) . \quad (12)
\]

Under the stated assumptions, \( \hat{R} \) converges (w.p.1) to the array covariance matrix

\[
R = APA^* + \sigma^2 I \quad (13)
\]

as \( N \to \infty \). Let the rank of the signal covariance matrix \( P \) be \( d' \), and let the columns of the matrix \( E_s \) be the \( d' \) principal eigenvectors of \( R \). It is then easy to see that the range space of \( E_s \) is contained in that of \( A(\theta) \), see e.g., [3]. This is the basis for all subspace based techniques for DOA estimation. An estimate of \( E_s \) is obtained from the eigendecomposition of \( \hat{R} \),

\[
\hat{R} = \sum_{i=1}^{m} \lambda_i \hat{e}_i \hat{e}_i^* = \hat{E}_s \hat{A}_s \hat{E}_s^* + \hat{E}_n \hat{A}_n \hat{E}_n^* . \quad (14)
\]

IV-373
The SSF class of methods then finds the DOA estimates by minimizing the following weighted least-squares criterion

$$\min_\mathbf{W} \| \mathbf{W}_r (\mathbf{E}_r - \mathbf{A} (\theta) \mathbf{T}) \mathbf{W}_e \|^2_2, \quad (15)$$

where $\| \cdot \|^2_2$ denotes the Frobenius norm. Here, $\mathbf{W}_r$ and $\mathbf{W}_e$ are user-specified, positive definite weighting matrices. Roughly speaking, the role of these is to whiten the equation error, $\mathbf{E}_r - \mathbf{A} \mathbf{T}$. The SSF estimates can be shown to give consistent estimates if the array is unambiguous and $d < (m + d')/2$.

### 3.1 Optimal SSF Method

The accuracy of the estimates obtained from (15) depends on the choice of $\mathbf{W}_r$ and $\mathbf{W}_e$. A natural selection rule for these weighting matrices is to try to minimize the variance of the DOA estimates. In the full version of this paper, it is shown that the asymptotic (for large $N$) covariance matrix of the estimation error is the sum of three terms; namely the finite sample effects of noise in the absence of modeling errors [3], the effect of array model perturbations assuming $N = \infty$ and $\mu^2 = 0$ [5], and finally the effect of the noise covariance perturbations assuming $N = \infty$ and $\mu^2 = 0$ [5]. Because of the nature of the resulting expression for the asymptotic estimation error covariance, it is possible to find an overall optimal weighting only for the special case where $\mathbf{B} = \mathbf{I}$. Thus, the case of array errors only is treated separately.

**Theorem 1** ([5]) Assume that $\mu^2 = 0$ and $\| \mathbf{Y} \|$ is large enough so that only array perturbations contribute to the estimation error. Then the SSF weightings

$$\mathbf{W}_r = \mathbf{B}^{-1/2} \quad \mathbf{W}_e = (\mathbf{T}^* \mathbf{Y}^* \mathbf{T})^{-1}, \quad (16)$$

where

$$\mathbf{T} = (\mathbf{A}^* \mathbf{A})^{-1} \mathbf{A}^* \mathbf{E}_r, \quad (17)$$

yield minimum variance DOA estimates in the class of SSF methods.

The following result is the main contribution of this paper.

**Theorem 2** Assume that $\mathbf{B} = \mathbf{I}$. Then, the choice

$$\mathbf{W}_r = \mathbf{I} \quad (18)$$

$$\mathbf{W}_e = \left( \mathbf{T}^* \mathbf{Y}^* \mathbf{T} + \mu^2 \mathbf{\Lambda}^{-2} + \frac{\sigma^2}{N} \mathbf{\Lambda}^{-2} \mathbf{\Lambda}_e \right)^{-1} \quad (19)$$

provide minimum variance DOA estimates. Here,

$$\mathbf{\Lambda} = \mathbf{\Lambda}_e - \sigma^2 \mathbf{I}, \quad (20)$$

and the diagonal matrix $\mathbf{\Lambda}_e$ contains the $d'$ principal eigenvalues of $\mathbf{R}$.

No overall optimal choice of weighting matrices is known in the general case for arbitrary array and noise model errors. However, one may still suggest reasonable choices that normally lead to improved performance. One possibility is to try

$$\mathbf{W}_r = (\mathbf{B} + \alpha \mathbf{I})^{-1/2}, \quad (21)$$

where $\alpha$ is a scaling that controls the relative contribution of the AP optimal row-weighting, $\mathbf{W}_r = \mathbf{B}^{-1/2}$, and the NP+FS choice $\mathbf{W}_r = \mathbf{I}$. Briefly, $\alpha$ should be small if array perturbations are the major sources of error (high signal-to-noise ratio and/or large $N$), whereas it should be chosen large if noise modeling errors and/or finite sample effects dominate.

As seen above, the optimal weighting matrices depend on unknown quantities. However, it is easy to show that these can be replaced by consistent estimates without affecting the asymptotic optimality. Furthermore, the overall optimally weighted SSF technique has been shown to coincide, in large samples, with the MAP estimator for the problem at hand. Thus, no other technique could give lower estimation error variance, at least not in large samples.

### 4 An Example

Assume that the wavefield of two Gaussian signal sources is recorded using a perturbed uniform linear array (ULA) of $m = 6$ sensors with half-wavelength interelement spacing. The emitters are located at $\theta = [0^\circ, 5^\circ]$ relative to array broadside, which corresponds to an angle separation of a quarter of the Rayleigh beamwidth. The array response is perturbed according to the model (4)-(6), with $\mathbf{B} = \mathbf{I}$ and

$$\mathbf{Y} = 0.001 \begin{bmatrix} 1.0 & 0.7 \\ 0.7 & 1.0 \end{bmatrix}.$$

A non-diagonal $\mathbf{Y}$ is used here since the DOAs are closely spaced, so some correlation between the perturbations is expected (it should be mentioned that this does not drastically affect the performance of any of the methods). The baseband signals are zero-mean Gaussian, 90% correlated, and of equal power.

The covariance of the additive noise is slightly perturbed from the nominal spatially white case. The noise covariance perturbations are generated according to the model (10)-(11), with $\mu^2 = 0.001$.

A batch of $N = 100$ snapshots is generated for a variety of signal-to-noise ratios (SNR), and the SSF
technique is applied to each data set using $W_c = I$ and three different column weightings:

- $W_c = A^2 A_c^{-1}$ (finite sample errors only), referred to as weighted subspace fitting (WSF) [3];
- $W_c = (T^* Y^T Y)^{-1}$ (array perturbations only), referred to as robust subspace fitting (RSF) [5];
- $W_c = \left( T^* Y^T T + \mu^2 A^2 + \frac{1}{N} \bar{A}^{-2} \bar{A} \right)^{-1}$ (overall optimal weighting), referred to as optimal subspace fitting (OSF).

Figure 1 depicts the results of a Monte-Carlo simulation involving 256 independent trials for each SNR. The lines represent the error predicted by the theoretical expressions, whereas the symbols indicate the empirically calculated RMS errors. Only the results for the estimation of $\theta_1 = 0^\circ$ are shown, the results for $\theta_2$ being similar. The empirical RMS errors agree well with the theoretically predicted values for SNR $\geq 5$ (WSF, OSF) and SNR $\geq 20$ (RSF). Note also that, as expected, WSF is optimal for low SNR, RSF is optimal for high SNR, whereas OSF provides overall optimal estimates.

5 Conclusions

We have presented a performance analysis of subspace fitting algorithms that accounts for errors introduced by additive noise and modeling errors simultaneously. For the special case of random unstructured perturbations to the array and noise models, we have derived optimal row and column subspace weightings that yield minimum variance DOA estimates. For more complicated perturbation models, heuristic subspace weightings based on our analysis will often lead to significantly improved performance as compared to not taking the model errors into account.

References