REFERENCES


Asymptotic Distribution of the MUSIC Null Spectrum

Jinho Choi and Ickho Song

Abstract—In this correspondence we derive the asymptotic distribution of the MUSIC null spectrum, from which an exact expression of the asymptotic variance of the MUSIC null spectrum can be obtained. The result presented in this correspondence is a simpler alternative and a special case of a more general result recently obtained by Lee and Wengrovitz.

I. INTRODUCTION

For source localization purposes array sensors have been used widely and various high-resolution methods based on the eigenstructure have been proposed, e.g., multiple signal classification (MUSIC) [4] and minimum-norm [2] are among the typical techniques. In [1] and [6], the mean and variance of the MUSIC null spectrum have been obtained to the first-order approximation. It is pointed out in [6], however, that the variance of the MUSIC null spectrum cannot be exactly expressed by the first-order approximation. This is quite an important and reasonable observation, since in general the mean is expressed by a first-order approximation and the variance is usually expressed by a second-order approximation. In this correspondence we obtain the asymptotic distribution of the MUSIC null spectrum, from which we can immediately find a more exact expression of the variance.

II. BACKGROUND

Let us consider an L-element array of what output is $y(t) \in C^L \times 1$ with $C^L \times 1$ denoting the space of $L \times 1$ complex-valued vectors, and assume the standard model of observation

$$y(t) = Ax(t) + n(t), \quad 1, 2, \ldots, N, \quad (2.1)$$

In (2.1) it is assumed that the column vector $x(t)$ for $M$-signal sources is an $M \times 1$ zero mean complex normal random vector and the additive noise $n(t)$ is also a zero mean complex normal random vector with covariance matrix $\Omega$. The full-rank covariance matrix of $x(t)$ is $E[x(t)xH(t)] = R_x$ where $E$ denotes the statistical expectation and $H$ denotes the Hermitian transpose. The matrix $A$ is an $L \times M (L > M)$ complex matrix having the particular structure: $A = \{a(\theta_1), a(\theta_2), \ldots, a(\theta_M)\}$, where $\theta_i$ is the DOA of the $i$th signal source. Here $a(\theta) \in C^L \times 1$ is a steering or transfer vector. If we denote the covariance matrix of $y(t)$ by $R_y$, it is easy to see that

$$R_y = AR_xA^H + \Omega. \quad (2.2)$$

The eigenvalues and eigenvectors of $R_y$ are denoted by $\lambda_i \geq \lambda_2 \geq \cdots \geq \lambda_L$, and $e_1, e_2, \ldots, e_L$ respectively. It is noteworthy that $\lambda_{M+1} = \lambda_{M+2} = \cdots = \lambda_L = \sigma$. The ranges of the matrices $S = [e_1, e_2, \ldots, e_M]$ and $G = [e_{M+1}, e_{M+2}, \ldots, e_L]$ are called the signal and noise subspaces, respectively. We observe that

$$a^H(\theta)G = 0, \quad \text{for } \theta \in \Theta \quad (2.3)$$

where $\Theta = \{\theta_1, \theta_2, \ldots, \theta_M\}$, because the vectors $\{a(\theta_i), 1 \leq i \leq M\}$ are orthogonal to the noise subspace. If we define $f(\theta) = a^H(\theta)GG^H(\theta)$, the function $f(\theta)$ has zeros only at $\theta \in \Theta$ [5]. In practice, however, we can obtain only the estimates of $S$ and $G$, $\hat{S}$ and $\hat{G}$, from the estimate of $R_y$, $\hat{R}_y = (1/N) \Sigma_{i=1}^N y(\theta_i)yH(\theta_i)$. The MUSIC null-spectrum $D(\theta)$ is then defined by

$$D(\theta) = a^H(\theta)G\hat{G}^H(\theta) \quad (2.4)$$

and it is thus expected that $D(\theta)$ has minimum points at around $\theta \in \Theta$. Therefore, we can estimate the DOA by taking the local minimum points of $D(\theta)$.

III. ASYMPTOTIC DISTRIBUTION OF THE MUSIC NULL SPECTRUM

To establish the distribution of the MUSIC null spectrum, we first review the statistical properties of eigenvectors of the sample covariance matrix $\hat{R}_y$. Following [5, lemma 3.1], the orthogonal projections of $\{e_i\}$, $M \times 1 \leq i \leq L$, onto the column space of the signal subspace $S$ are asymptotically jointly Gaussian distributed

Manuscript received June 22, 1991; revised December 1, 1991.

The authors are with Department of Electrical Engineering, Korea Advanced Institute of Science and Technology (KAIST), Yuseong Gu, Daejeon 305-701, Korea.

IEEE Log Number 9205117.
with zero mean and covariance matrices given by

\[ E[(SS^H \hat{e}_i)(SS^H \hat{e}_j)] = \sigma^2 \sum_{i=1}^{M} \frac{\lambda_i}{(\sigma - \lambda_i)^2} e_i e_j^H \delta_{ij}, \quad M + 1 \leq i, j \leq L \]  

and

\[ E[(SS^H \hat{e}_i)(SS^H \hat{e}_j)^T] = 0, \quad M + 1 \leq i, j \leq L \]

where \( \delta_{ij} \) is the Kronecker delta. For notational convenience let us define a scalar quantity

\[ r_i = a^H(\theta)\hat{e}_i, \quad \theta \in \Theta, \]  

(3.3)

Since \( a(\theta) \in \text{span} \{e_1, e_2, \ldots, e_M\} \) for \( \theta \in \Theta \), it is easy to see that

\[ a^H(\theta)SS^H = a^H(\theta) \]

(3.4)

from which the means of \( r_i, M + 1 \leq i \leq L \), can be shown to be

\[ E[r_i] = E[a^H(\theta)\hat{e}_i] \]

\[ = 0. \]  

(3.5)

Using (3.1)-(3.4) we have, for \( M + 1 \leq i, j \leq L \)

\[ E[r_i r_j^*] = E[(a^H(\theta)\hat{e}_i)(a^H(\theta)\hat{e}_j)^*) \]

\[ = a^H(\theta)E[(SS^H\hat{e}_i)(SS^H\hat{e}_j)^*)a(\theta) \]

\[ = \sigma^2 \sum_{i=1}^{M} \frac{\lambda_i}{(\sigma - \lambda_i)^2} a^H(\theta)e_i^H e_i a(\theta) \delta_{ij} \]  

(3.6)

and

\[ E[r_i r_j] = E[(a^H(\theta)\hat{e}_i)(a^H(\theta)\hat{e}_j)^*) \]

\[ = a^H(\theta)E[(SS^H\hat{e}_i)(SS^H\hat{e}_j)^*)a^*(\theta) \]

\[ = 0. \]  

(3.7)

Clearly the scalar quantities \( r_i, M + 1 \leq i \leq L \), being linear combinations of Gaussian random vectors, are asymptotically independent complex Gaussian random variables with zero mean and covariance given by (3.6). In addition from (3.7) \( E[r_i r_j] = 0 \), which implies that Re \((r_i)\) and Im \((r_i)\) are asymptotically independent and their variances are equal.

The variance of \( r_i \) denoted by \( \sigma_i^2 \) is from (3.6)

\[ E|r_i|^2 \sim \sigma_i^2 \]

\[ = \sigma^2 \sum_{i=1}^{M} \frac{\lambda_i}{(\sigma - \lambda_i)^2} a^H(\theta)e_i^H e_i a(\theta) \]  

(3.8)

for \( M + 1 \leq i \leq L \), with which we can now write \( r_i \sim N(0, \sigma_i^2) \), \( M + 1 \leq i \leq L \). The MUSIC null spectrum can be expressed as, from (2.4) and (3.3)

\[ D(\theta) = a^H(\theta) \left( \sum_{i=M+1}^{L} \hat{e}_i^H a(\theta) \right) \]

\[ = \sum_{i=M+1}^{L} |r_i|^2, \theta \in \Theta \]  

(3.9)

for which it is easy to see that the normalized MUSIC null spectrum, \( 2D(\theta)/\sigma_i^2, \theta \in \Theta \), has asymptotically \( \chi^2 \) distribution with degree of freedom \( 2(L - M) \). In other words

\[ \frac{2D(\theta)}{\sigma_i^2} \sim \chi_{2(L - M)}^2. \]

The asymptotic mean and variance of the MUSIC null spectrum can now be easily obtained at \( \theta \in \Theta \); we have

\[ E[D(\theta)] = (L - M)\sigma_i^2 \]

(3.10)

and

\[ \text{var}[D(\theta)] = (L - M)\sigma_i^2. \]

(3.11)

It is noteworthy that the variance of the MUSIC null spectrum has the order of \( N^{-1} \); note the order of \( N^{-1} \) as was approximated in [1], [6]. The result also agrees with the observation made in [6]. Empirical studies given in Section IV will also show that the asymptotic expressions (3.10) and (3.11) are correct.

Let us now consider the normalized standard deviation (NSD) [6], NSD \( (\theta) = \sqrt{\text{var}[D(\theta)]/E[D(\theta)]] \). It is easy to see that NSD \( \sim \sqrt{1/(L - M)} \). Note that the NSD is not dependent on the signal power, noise power, or number of snapshots, but is only dependent on the degree of freedom, \( 2(L - M) \), or on the numbers of sensors and sources.
IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 41, NO. 2, FEBRUARY 1993 987

0 10 20 30 40
SNR (dB)

Fig. 2. VAR[D(θ)] versus SNR.

TABLE I

EMPIRICAL NSD FOR SEVERAL VALUES OF SNR, NUMBER OF SOURCES, AND NUMBER OF SENSORS (NUMBER OF SNAPSHOTS = 100, NUMBER OF TRIALS = 100)

<table>
<thead>
<tr>
<th>Degree of freedom</th>
<th>Number of sources (L)</th>
<th>Number of sources (M)</th>
<th>SNR (dB)</th>
<th>Theoretical NSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1</td>
<td>0.544 0.612 0.525 0.643 0.561 0.642 0.581 0.576 0.619</td>
<td>0.577</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>0.537 0.465 0.670 0.640 0.607 0.499 0.491 0.705 0.557</td>
<td>0.577</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>0.584 0.629 0.591 0.587 0.633 0.576 0.486 0.543 0.618</td>
<td>0.447</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>0.496 0.485 0.422 0.521 0.427 0.448 0.462 0.427 0.442</td>
<td>0.353</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>0.387 0.457 0.427 0.434 0.427 0.420 0.505 0.475 0.450</td>
<td>0.353</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>0.437 0.461 0.497 0.464 0.460 0.400 0.400 0.400 0.400</td>
<td>0.353</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>0.330 0.364 0.359 0.361 0.400 0.333 0.434 0.423 0.357</td>
<td>0.353</td>
<td></td>
</tr>
</tbody>
</table>

IV. SIMULATION RESULTS

In this section, the asymptotic mean and variance of the MUSIC null spectrum are evaluated for several values of signal-to-noise ratio, numbers of sensors and signals, and number of snapshots. Let us consider a uniform linear array of L sensors with sensor spacing half the wavelength. Let SNR = P/ρ, where P denotes the power of the signal and ρ is the noise power. We assume that all signal sources are uncorrelated with equal power and that the signal sources and the noise sources are uncorrelated. The signal sources are assumed to be located at 15° and 20° for M = 2 and the numbers of sensors are 5, 7, and 10. The numbers of snapshots used in the simulation are 10, 50, and 100. The results of a set of simulations are given in Figs. 1 and 2.

In Figs. 1 and 2, we show some plots of E[D(θ)] versus SNR and of var[D(θ)] versus SNR respectively. The lines in Figs. 1 and 2 represent the theoretical values of the asymptotic mean and variance calculated by (3.10) and (3.11), respectively, and the '+'s denote the values obtained from Monte Carlo simulation with 100 trials. From the figures it is shown that the simulation results are quite close to the theoretical results when the number of sensors are sufficiently large compared to that of signals and when the SNR is large. It is noteworthy that at high SNR the theoretical asymptotic mean and variance agree well with the simulation results even when N = 10, despite the theoretical results are derived for large values of N. From further simulations, it can be seen that when L - M is small, the simulation and theoretical results tend to be different from each other at low SNR, and that when L - M is large they tend to have the same values.

The effect of SNR on NSD is shown in Table I for various cases. It is clearly shown that the NSD is independent of the SNR.

V. CONCLUDING REMARKS

The asymptotic distribution of the MUSIC null spectrum at the source location is derived. It is shown that the asymptotic distribution of the MUSIC null spectrum has a $\chi^2$ distribution: the asymptotic mean and variance of the MUSIC null spectrum depend on the degree of freedom that is two times the difference between the number of sensors and that of signals. From this result we can obtain more explicit expressions of the asymptotic variance of the MUSIC null spectrum is the order of $N^{-1}$, where N is the number of snapshots.

It should be noted that the result presented in this correspondence is simpler and a special case of the result obtained independently by Lee and Wengrovitz [3].

REFERENCES

[1] M. Kaveh and A. J. Barabell, "The statistical performance of the MUSIC and the minimum-norm algorithms in resolving plane waves in
A Necessary and Sufficient Condition for the Existence of the Maximum Likelihood Estimate in Autoregressive Models

Serge Degerine

Abstract—This correspondence concerns the existence problem of a maximum likelihood estimate (MLE) for the parameters of a \( p \)-th order autoregressive (AR) model from \( n \geq 1 \) independent records of length \( m \) of a complex time series. It is shown that, for almost all such sets of observations, the MLE exists if and only if the \( n \) records cannot be exactly fitted by complex undamped sinusoids using the same set of \( p \) distinct frequencies.

I. INTRODUCTION

Numerous situations in signal processing fall into the following general framework: estimate the parameters of an AR (\( p \)) model from \( n \geq 1 \) independent records of length \( m \) of a complex time series. Autoregressive spectral estimation generally concerns the case in which only one record is observed (\( n = 1 \)). The problem of estimating a Toeplitz covariance matrix (\( p = m - 1 \)) occurs essentially in array processing [1]. Well-known usual methods (Yule-Walker or correlation method, modified covariance method, and Burg’s technique) are designed for spectral estimation extend to the general situation [2]. For small lengths \( m \) the exact MLE [3], [4] must be preferably used. However, the likelihood function can present local maxima [5]. Here we look at the existence problem. From a practical point of view we might be satisfied with the sufficient conditions \( p < 2m/3 \) when \( n = 1 \) or \( n \geq m/2 \) when \( p = m - 1 \) established in [6] for a real-time series but obviously true in the complex case. The interest of our result is at first to give a necessary and sufficient condition covering all situations and, moreover, to show that this condition is clearly related to the sinusoids fitting problem. Actually, for one sequence the MLE exists provided that the order \( p \) of the fitted AR model is less than the smallest number of undamped sinusoids fitting the data exactly.

This gives rise to a (very expensive) method to determine this number and compute a set of fitted sinusoids. Notice that this last result differs from the original Prony’s concept in which the \( m \) available data points arise from \( p \) sinusoids of sufficient length \( (m > 3p/2) \) and then the sinusoids are given by a linear system. In estimating a Toeplitz matrix (\( p = m - 1 \)) our result states precisely the condition, by excluding a null set, and rectifies the arguments given in [7].

II. THE MAXIMUM LIKELIHOOD METHOD

Let \( x_1(1), \ldots, x_r(m), r = 1, \ldots, n \) be \( n \) independent records of length \( m \) arising from an AR (\( p \)) model

\[
\sum_{k=0}^{p} a_k x(t - k) = e(t), \quad a_0 = 1
\]

where the \( e(t) \) are i.i.d. zero-mean complex Gaussian random variables with variance \( \sigma^2 \). The probability density of these observations is

\[
f(S; \Lambda_p) = |\det (\pi \Lambda_p)|^{-n} \exp \left\{ -n \text{tr} (\Lambda_p^{-1} S) \right\}
\]

where \( \Lambda_p = E(xx^*) \) is the covariance matrix of \( X = [x_1(1), \ldots, x_r(m)]^T \) and \( S \) is the sample covariance matrix:

\[
S = \frac{1}{n} \sum_{r=1}^{n} [x_r(1), \ldots, x_r(m)] [x_r(1), \ldots, x_r(m)]^T.
\]

Writing \( \Lambda_p = \sigma^2 \Gamma_p \), where \( \sigma^2 = E(x(t)x(t)) \) is the variance of \( X(t) \), the log likelihood function becomes

\[
L(\Gamma_p; S) = \ln \text{det} \Gamma_p + m \ln \sigma^2 + \frac{1}{\sigma^2} \text{tr} (\Gamma_p^{-1} S).
\]

Its derivative with respect to \( \sigma^2 \) shows that the MLE is given by \( \tilde{\sigma}^2 = \text{tr} (\Gamma_p^{-1} S)/m \) where \( \tilde{\sigma}^2 \) is the correlation matrix that minimizes

\[
L(\Gamma_p; S) = \ln \det \Gamma_p + m \ln \sigma^2 + \frac{1}{\sigma^2} \text{tr} (\Gamma_p^{-1} S).
\]

The correlation matrix \( \Gamma_p \) depends only on the AR parameters \( a_1, \ldots, a_p \) which must satisfy the condition

\[
\phi_p(z) = \sum_{k=0}^{p} a_k z^{-k} \neq 0 \quad \text{in} \ |z| \geq 1.
\]

So the minimization of (3) by gradient or relaxation methods in the AR parameters domain is a difficult numerical task. It is the reason why the relaxation method used in [3] and [4] is conducted in the reflection coefficients domain. For these coefficients are a set of parameters \( \beta_1, \ldots, \beta_p \) subject to the common restraint \( |\beta_k| < 1, k = 1, \ldots, p \). The one-to-one correspondence between these two sets of parameters is realized by the well-known Levinson–Durbin recursion:

\[
a_0(k) = \beta_1; \quad a_i(k) = a_i(k - 1) + \beta_i a_{i-k}(k - 1),
\]

\[
1 \leq i < k; \quad k = 1, \ldots, p
\]

where \( a_i = a_i(\rho) \).

III. SINUSOIDS FITTING THE DATA

Let \( p(S) \) be the smallest integer \( p < m \) for which the \( n \) records satisfy the difference equation

\[
\sum_{k=0}^{p} a_k x(t - k) = 0, \quad t = p + 1, \ldots, m, \quad r = 1, \ldots, n
\]