Optimal Nonnegative Definite Approximations of Estimated Moving Average Covariance Sequences

Randolph L. Moses, *Senior Member, IEEE*, and Duixian Liu, *Student Member, IEEE*

***Abstract*—This paper considers the following problem: Given an estimated MA covariance sequence which may not be nonnegative definite (NND), find the closest NND sequence to it. Here, closeness is measured by the weighted Euclidean distance of the covariances. We provide a solution to this problem by considering a set of linear minimization problems which are parameterized by the zero frequencies of the optimal solution. Some properties of the optimal NND solutions are established, and these properties are used to simplify the minimization procedure.**

I. Introduction

There are many problems in which one is interested in obtaining a parametric model of the spectrum of a time series. The autoregressive (AR), moving average (MA), and autoregressive moving average (ARMA) models are widely used in many engineering problems. In obtaining MA and ARMA spectral estimates, a problem which often arises is that of ensuring that the resulting spectral estimate is nonnegative and real on the unit circle [1]. For example, a commonly used method of MA spectral estimation is to estimate the first $n + 1$ autocovariances $\{\gamma_k\}_0^n$ of a time series from some measurements of that time series. Depending on the estimator used for $\{\gamma_k\}$, the spectral estimate may not be nonnegative and real. A similar problem occurs in ARMA spectral estimation algorithms in which the AR parameters are estimated in a first step, and the MA part of the spectrum is estimated using the AR coefficient estimates [2]–[6].

It is well known that a necessary and sufficient condition for the MA spectrum to be nonnegative real is that the MA covariance sequence $\{\gamma_k\}_{k=-\infty}^{\infty}$ be nonnegative definite (NND). Here, the sequence is nonzero only for $|k| \leq n$ where n is the order of the MA process. If the MA covariance sequence estimate is not NND, there are various ways in which one can alter the estimate to make it NND. A common procedure entails multiplying the estimated autocovariances by some window sequence (such as the Bartlett window or an exponential window) [1], [5]. For some estimates, the window can be chosen in such a way as to guarantee NND estimates; however, such a window imposes a bias on the resulting estimate [1]. A second approach is to use a data adaptive window, in which a parameter in the window is chosen to ensure NND estimates, with a minimum of bias for that particular window. For example, an exponential window $w_k = \alpha^{|k|}$ can be used, where α is chosen by a one-dimensional search procedure to be as small as possible so that the sequence $\{w_k \gamma_k\}$ is NND. This variable window method imposes less bias than the fixed window method, but it requires iteration to find α.

In this paper we consider an optimum approach to obtaining a NND covariance sequence. Given an estimated covariance sequence $\{\gamma_k\}_{k=0}^n$ of an MA time series, we wish to find the closest NND sequence to that estimate, where closeness is measured in terms of a l_2 error norm in coefficient space. This is a nonlinear minimization problem. We discuss necessary and sufficient conditions for the solution to this problem. We then derive an algorithm for finding the global minimum.

This problem is closely related to the approximate stochastic realization problem as considered in [6]–[12]. In [6], [9]–[11], the approach taken is to parameterize the covariance sequence in terms of the parameters of an ARMA model which admits this sequence. Then the ARMA model which yields the closest covariance sequence to a given one is found by minimizing an error functional; this involves a nonlinear minimization procedures on the coefficients of the ARMA model. The functional dependence on the AR parameters is highly nonlinear, and convergence to local minima is a problem [6], [9], [10]. For the special case of a moving average model, the minimization is quartic in the MA parameters [11]. In either case, the minimization problem is of dimension equal to the number of ARMA (MA) parameters.

This work is also related to recent work by Steinhardt and others [13], [14]. In this work the authors have characterized the set of all partial covariance sequences $\{r_0, \cdots, r_n\}$ which admit an MA or ARMA model of given order. They also develop the expanding hull algorithm to find the smallest order model which can admit a given partial sequence $\{r_0, \cdots, r_n\}$ which is NND. The problem considered in this paper is somewhat different: given a sequence $\{\gamma_0, \cdots, \gamma_n\}$ which is not NND, we wish to

Manuscript received January 12, 1989; revised August 22, 1990. This work was supported by an Ohio State University Seed Grant.

The authors are with the Department of Electrical Engineering, Ohio State University, Columbus, OH 43210.

IEEE Log Number 9101341.

1053-587X/91/0900-2007$01.00 © 1991 IEEE

find the closest NND sequence $\{r_0, \cdots, r_n, 0, 0, \cdots\}$ (corresponding to an MA(n) model).

In this paper, the approach taken is to consider the minimization problem in the space of covariance parameters. One reason to consider the problem in covariance space is to gain insight on the geometry of the problem. Specifically, if a covariance sequence is not NND, then its Fourier transform (power spectral density function) is negative for some frequencies. The closest NND approximant to this sequence will have a Fourier transform which touches zero at some frequencies. We derive a minimization procedure which is based on locating these zero frequencies. These frequencies correspond to tangent hyperplanes in the set of solutions, and the minimization problem can be seen as a standard orthogonal projection problem onto these hyperplanes. The nonlinearity of this problem arises because these frequencies are not known *a priori*, and must be found using iterative minimization techniques. On the other hand, the number of zero frequencies is small; for γ_k sequences which are "close" to NND, the nonlinear minimization problem has low dimension. The worst case dimension of the minimization is $\lfloor n/2 \rfloor + 1$ where n is the MA order. Thus, iterative minimization is carried out in a space of lower dimension than if the solution is found in the space of MA parameters as in [6], [8]–[11].

An outline of this paper is as follows. In Section II we present a formal statement of the problem. In Section III some properties of the nonnegativity region are described. In Section IV we derive a solution to the minimization problem which uses the Lagrange method in terms of known zero-spectrum frequencies. In Section V we further characterize the optimal solution. Section VI presents examples of the algorithm, and Section VII concludes the paper.

II. Problem Statement

Let $\{\gamma_k\}_{k=0}^n$ denote a sequence of real numbers. This sequence can be thought of as estimates of the first $n + 1$ covariances of an MA(n) process. Consider the function

$$S_\gamma(z) = \sum_{k=-n}^{n} \gamma_{|k|} z^{-k}. \tag{1}$$

In order to ensure that $S_\gamma(z)$ is a valid spectral density function, we must have $S_\gamma(e^{j\omega}) \geq 0$, or, equivalently, that

$$g(\omega) = s_0 + s_1 \cos \omega + \cdots + s_n \cos n\omega \geq 0 \quad \forall \omega \in [0, \pi] \tag{2}$$

where $s_0 = \gamma_0$, and $s_k = 2\gamma_k$ for $k = 1, \cdots, n$. Nearly all covariance estimators guarantee that $\gamma_0 > 0$, but often do not guarantee that (2) is satisfied; thus, we will assume $s_0 > 0$ in the following discussion.

Assume condition (2) is not satisfied; that is $\{\gamma_k\}$ is not a NND sequence. In this case, we are interested in finding a covariance sequence which is NND and which is close to the given sequence. To this end, let $\rho = [\rho_0, \rho_1, \cdots, \rho_n]^T$ and define

$$f(\omega, \rho) = \rho_0 + \rho_1 \cos \omega + \cdots + \rho_n \cos n\omega. \tag{3}$$

Define the nonnegative definite set D by

$$D = \{\rho | f(\omega, \rho) \geq 0 \quad \text{for } \omega \in [0, \pi]\}.$$

Then the problem of finding the closest NND sequence can be stated as follows:

Problem P: Given a vector $s = [s_0, s_1, \cdots, s_n]^T \notin D$, find the vector $\rho^{\text{opt}} \in D$ such that $Q = \|\rho^{\text{opt}} - s\|^2 \leq \|\rho - s\|^2$ for all $\rho \in D$, where $\|\cdot\|$ is the l_2 (Euclidean) vector norm.

Note that even though s and ρ have the same dimension in the above problem statement, one can find the NND solution for a different order than that of s. If the closest NND sequence of order l is desired for $l < n$, s is replaced by $[s_0, \cdots, s_l]^T$. If the desired order l is greater than n, s is replaced by the $l + 1$-vector $[s_0, \cdots, s_n, 0, \cdots, 0]^T$.

Once ρ^{opt} is found, an MA(n) filter which realizes ρ^{opt} is obtained by performing a spectral factorization on the function

$$S_{\rho^{\text{opt}}}(z) = \sum_{k=-n}^{n} r_{|k|} z^{-k} \tag{4}$$

where $r_0 = \rho_0^{\text{opt}}$ and $r_k = \rho_k^{\text{opt}}/2$ for $1 \leq k \leq n$.

III. Description of the Admissible Set D

The minimization problem is nontrivial because the set D is a complicated function of the ρ vector. We first establish some properties of D.

Theorem 1:

a) $D \subset R^{n+1}$ is a closed convex cone with vertex at the origin.

b) Let ∂D denote the boundary of D. If $\rho \in \partial D$, there is at least one frequency $\omega_0 \in [0, \pi]$ such that $f(\omega_0, \rho) = 0$.

c) If $\rho^* \in \partial D$ and $f(\omega_0, \rho^*) = 0$, then the hyperplane

$$H_{\omega_0} = \{\rho | f(\omega_0, \rho) = 0\} \tag{5}$$

is tangent to D.

d) There is a unique solution ρ^{opt} to the minimization problem P.

e) Define the half-spaces H_ω^+

$$\begin{aligned} H_\omega^+ &= \{\rho \in R^{n+1} | f(\omega, \rho) \geq 0\} \\ &= \{\rho \in R^{n+1} | \langle b(\omega), \rho \rangle \leq 0\} \end{aligned} \tag{6}$$

where

$$b(\omega) = -[1, \cos \omega, \cos 2\omega, \cdots, \cos n\omega]$$

is the normal vector for H_ω. Then D is the intersection of these half spaces

$$D = \bigcap_{\omega \in [0, \pi]} H_\omega^+. \tag{7}$$

Proof:

a) It is readily verified that if $\rho_1, \rho_2 \in D$, $\alpha\rho_1 + (1 - \alpha)\rho_2 \in D$ for any $\alpha \in [0, 1]$, so D is convex. It is also clear from (3) that if $\rho \in D$, then $\alpha\rho \in D$ for all $\alpha \geq 0$.

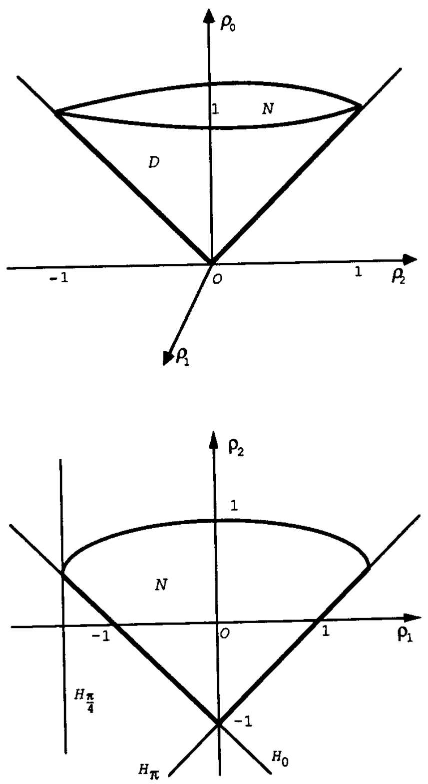


Fig. 1. Nonnegative definite regions N and D for $n = 2$.

Thus D is a convex cone with vertex at the origin. To show D is closed, consider any $\rho^0 \in D^c$, where D^c is the complement of D. Then $f(\omega, \rho^0) < 0$ for some $\omega \in [0, \pi]$. Since $f(\omega, \rho)$ is a continuous function of ρ for any fixed ω, there exists a neighborhood of ρ^0 such that $f(\rho, \omega) < 0$ for any ρ in that neighborhood. This implies that D^c is open, and thus D is closed.

b) This follows readily from the definition of $f(\omega, \rho)$.

c) Note that $\rho^* \in H_{\omega 0}$ so $H_{\omega 0}$ intersects D. Moreover, $H_{\omega 0}$ cannot contain an interior point of D, because every interior point ρ of D has the property that $f(\omega, \rho) > 0$ for all $\omega \in [0, \pi]$. Thus, H_ω is tangent to D.

d) This follows immediately from the fact that D is a closed and convex cone, and Q is a distance function [15, theorem 1, p. 69].

e) Equations (6) and (7) follow immediately from the definition of D. ■

We remark that a set which is related to D can found by considering the slice of D found by setting ρ_0 to a constant. For any $\rho \in D$ with $\rho_0 > 0$, we can write

$$f(\omega, \rho) = \rho_0[1 + \eta_1 \cos \omega + \cdots + \eta_n \cos n\omega]$$

where $\eta_i = \rho_i/\rho_0$ for $i = 1, \cdots, n$. We can then define a set $N \subset R^n$ by

$$N = \{\eta = [\eta_1, \cdots, \eta_n]^T | 1 + \eta_1 \cos \omega + \cdots + \eta_n \cos n\omega \geq 0 \quad \text{for } \omega \in [0, \pi]\}. \tag{8}$$

The set N is the $\rho_0 = 1$ slice of D. This set has been studied in [13]. In particular, it is shown in [13] that N is a compact, convex subset of R^n. One can define the minimization problem in terms of N instead of D, but this gives a different solution which has higher error. However, it is convenient to use N instead of D to visualize the geometry of the minimization problem. This is especially true when $n = 3$, because N is a subset of R^2 for this case.

Fig. 1 shows the regions $N \subset R^2$ and $D \subset R^3$ for $n = 2$. Also shown are three hyperplanes (lines) on N corresponding to three frequency values for which $f(\omega, \rho) = 0$. These three lines correspond to tangent planes of D in R^3, each defined by $\rho_0 + \rho_1 \cos \omega + \rho_2 \cos 2\omega = 0$.

IV. Solution to the Minimization Problem

One way to solve the minimization problem is to define a grid of points $\omega_1, \cdots, \omega_k$ in the interval $[0, \pi]$ and to find the solution to

$$\rho^* = \arg\min \|s - \rho\| \qquad \text{for } \rho \in H_k^+ = \bigcap_{i=1}^{k} H_{\omega_i}^+. \tag{9}$$

In this case, ρ^* is the closest point to s which lies on the supporting polygonal cone which contains D. As the number of grid points increases, ρ^* approaches the optimal solution ρ^{opt}. The minimization problem (9) is a quadratic minimization problem with k linear inequality constraints, and can be solved using standard techniques (see [16]).

The approach we take is based on this idea, but incorporates some structure of the problem to simplify the minimization. Because $\rho \in R^{n+1}$, $f(\omega, \rho) = 0$ for at most $\lfloor n/2 \rfloor + 1$ distinct frequencies in $[0, \pi]$. As a result, at most $\lfloor n/2 \rfloor + 1$ of the linear inequality constraints in (9) are active. The approach we take makes use of this fact by using k constraints in (9) for $k = 1, \cdots, \lfloor n/2 \rfloor + 1$, and by forcing all constraints to be active. By varying the frequencies corresponding to these constraints, we span over all possible points on the boundary of D, and thus span over all possible solutions to the minimization problem.

The minimization problem in (9) can be stated as a constrained minimization problem by defining

$$\boldsymbol{\omega} = [\omega_1, \cdots, \omega_k]$$

$$Q_k(\boldsymbol{\omega}) = \|\rho - s\|^2 - 2 \sum_{j=1}^{k} A_j f(\omega_j, \rho) \tag{10}$$

where each A_j is a Lagrange multiplier. Let $Q_k^*(\boldsymbol{\omega})$ denote the minimum of $Q_k(\boldsymbol{\omega})$ with respect to ρ for a given frequency vector $\boldsymbol{\omega}$, and let ρ^* denote the minimum point. Then minimization of $Q_k(\boldsymbol{\omega})$ gives the point ρ^* which is closest to s under the constraint that ρ^* lies on the hyperplane $H_k = H_{\omega_1} \cap \cdots \cap H_{\omega_k}$, and $Q_k^*(\boldsymbol{\omega}) = \|\rho^* - s\|^2$. The solution ρ^* is the orthogonal projection of the point s onto the hyperplane H_k, and can be found by solving a set of linear equations as given by

$$\begin{bmatrix} I_n & -C \\ -C^T & 0 \end{bmatrix} \begin{bmatrix} \rho^* \\ A^* \end{bmatrix} = \begin{bmatrix} s \\ 0 \end{bmatrix} \tag{11}$$

where

$$C^T = \begin{bmatrix} 1 & \cos \omega_1 & \cos 2\omega_1 & \cdots & \cos n\omega_1 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \cos \omega_k & \cos 2\omega_k & \cdots & \cos n\omega_k \end{bmatrix}$$

$$A^* = [A_1^*, \cdots, A_k^*]^T.$$

The solution to (11) is given by

$$\rho^* = s - C(C^TC)^{-1}g \tag{12}$$

$$A^* = -(C^TC)^{-1}g \tag{13}$$

$$Q_k^*(\boldsymbol{\omega}) = g^T(C^TC)^{-1}g \tag{14}$$

where

$$g = C^Ts = [g(\omega_1), \cdots, g(\omega_k)]^T. \tag{15}$$

Because each H_{ω_i} is tangent to D, any solution ρ^* lies either outside D or on the boundary ∂D. If the frequency vector $\boldsymbol{\omega}$ is chosen appropriately, the ρ^* lies in ∂D and coincides with the optimal solution ρ^{opt}. Below we develop conditions for which $\rho^* = \rho^{\text{opt}}$. From these conditions, we develop an algorithm for finding ρ^{opt} based on solving the projection problem in (10). To this end, the following theorems are of interest.

Theorem 2: Let ρ^* and A^* be the solutions to (12) and (13), and let $f(\omega, \rho^*) = 0$ for k distinct frequencies $\omega_1, \cdots, \omega_k$. Then $A_j^* \geq 0$ for $j = 1, \cdots, k$.

Remark: Theorem 2 states that $s - \rho^*$ is in the convex cone spanned by the normals of the hyperplanes H_{ω_i} for $i = 1, \cdots, k$.

Proof: This is a standard result from optimization theory (see, e.g., [15, pp. 213-270]); the conditions $A_j^* \geq 0$ for $j = 1, \cdots, k$ are the Kuhn-Tucker necessary conditions. ∎

Theorem 3: Let ρ^{opt} be the solution to the minimization problem P, and let $\{\omega_1^0, \cdots, \omega_k^0\}$ be the set of k distinct frequencies for which $f(\omega, \rho^{\text{opt}}) = 0$. Then the functional $Q_k^*(\boldsymbol{\omega})$ in (14) has a local maximum at the point $\boldsymbol{\omega} = (\omega_1^0, \cdots, \omega_k^0)$.

Proof: According to the duality principle [15], the minimum distance between a convex set and a point outside this set is equal to the maximum of the distances from the point to supporting hyperplanes separating the point and the convex set. In addition, the point ρ^* (and consequently, the distance $Q_k^*(\boldsymbol{\omega})$ between s and ρ^*) is completely parameterized by the zero frequencies of $f(\omega, \rho^*)$. Thus, varying ρ to minimize $\|\rho - s\|^2$ is equivalent to varying the zero frequencies $\omega_1, \cdots, \omega_k$ to find local maxima of $Q_k^*(\boldsymbol{\omega})$. ∎

We illustrate the above theorems by means of a simple example on the set N as shown in Fig. 2. Consider first the point a as the given covariance vector s. For any frequency $\omega \in [0, \pi]$, the point $\rho^*(\omega)$ is the orthogonal projection of a onto the hyperplane H_ω, which in this case is a line. The hyperplanes H_0 and H_π are shown, and for $\omega \in (0, \pi)$ the hyperplane H_ω is the tangent line to D, which rotates clockwise from H_0 around to H_π as ω increases. The value of $Q^*(\omega)$ is the squared distance between a and the hyperplane H_ω. For the point a, this distance achieves local maxima at $\omega = \pi/4$ and at $\omega = 3\pi/4$; however, $A^*(\pi/4) > 0$ and $A^*(3\pi/4) < 0$ because a and D are on opposite sides of the hyperplane $H_{\pi/4}$ but on the same side of $H_{3\pi/4}$. The projection $\rho^*(\pi/4)$ is the point b, and this is also ρ^{opt} for this problem.

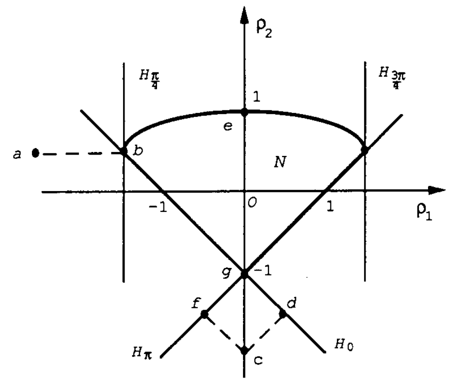


Fig. 2. Solutions on N for $n = 2$.

Next, consider the point c. For this point $Q^*(\omega)$ has local maxima at $\omega = 0$, $\pi/2$, and π, and the corresponding ρ^* points are marked d, e, and f, respectively. For points d and f, the corresponding $A^*(\omega) \geq 0$, but they are outside D; the point e is on the boundary ∂D, but the corresponding $A^*(\omega) < 0$, so none of these points are admissible solutions to problem P. Thus, we must consider the two constraint problem. The only possible two-constraint set for this order is $\omega = 0$, $\omega = \pi$, and $H_0 \cap H_\pi$ gives the point g, which is the optimum solution ρ^{opt} for this case.

The above theorems and example motivate a method for finding the optimum solution to the minimization problem P via projections onto the tangent subspaces H_k, as given by the following theorem.

Theorem 4: Let $\boldsymbol{\omega}^* = [\omega_1, \cdots, \omega_k]$ be a set of k distinct frequencies, each in the interval $[0, \pi]$. Let ρ^*, A^*, and $Q^*(\boldsymbol{\omega})$ be the corresponding solutions to (12)–(14). Assume

1) $Q_k(\boldsymbol{\omega})$ achieves a local maximum at $\boldsymbol{\omega}^*$;
2) $A_j^* \geq 0$ for $j = 1, \cdots, k$; and
3) $\rho^* \in D$.

Then $\rho^* = \rho^{\text{opt}}$.

Proof: We know that

$$\|s - \rho^*\| \leq \|s - \rho\| \qquad \forall \rho \in H_k^+ = \bigcap_{i=1}^{k} H_{\omega_i}^+$$

with equality if and only if $\rho = \rho^*$. We also know that $D \subset H_k^+$. These two statements imply

$$\|s - \rho^*\| \leq \|s - \rho\| \qquad \forall \rho \in D.$$

But the above statement is the definition of ρ^{opt}, and since there is a unique solution to the minimization problem, $\rho^* = \rho^{opt}$. ■

A. Description of the Algorithm

The above theorems provide a means for finding ρ^{opt}. Start with one frequency constraint ($k = 1$) in (10). Find local maxima of $Q^*(\omega_1)$ in (14) and corresponding ρ^* and A^* in (12) and (13) (note that this involves a nonlinear maximization in one variable $\omega_1 \in [0, \pi]$). For each local maximum, check conditions 2) and 3) of Theorem 4. If both are satisfied, $\rho^* = \rho^{opt}$. If not, increment k and continue.

Note that if n is even $f(\omega, \rho) = 0$ for at most $\lfloor n/2 \rfloor$ distinct frequencies in $(0, \pi)$. There can be up to $\lfloor n/2 \rfloor + 1$ distinct zeros in $[0, \pi]$, but only if $f(\omega, \rho) = 0$ at both $\omega = 0$ and $\omega = \pi$. Thus, the maximum value of k is $\lfloor n/2 \rfloor + 1$, and the dimension of the nonlinear maximization problem is at most $\lfloor n/2 \rfloor$. Similarly, if n is odd there are at most $\lfloor n/2 \rfloor$ zeros of $f(\omega, \rho)$ in $(0, \pi)$, and k is at most $\lfloor n/2 \rfloor + 1$. In either case, then, $k \leq \lfloor n/2 \rfloor + 1$. Also, because of the symmetry of $Q_k^*(\boldsymbol{\omega})$ with respect to interchanging of two frequencies, it suffices to find maxima of $Q_k^*(\boldsymbol{\omega})$ on the set

$$I_k = \{\boldsymbol{\omega} \in [0, \pi]^k | \omega_1 \in [0, \pi], \omega_2 \in [0, \omega_1), \cdots, \omega_k \in [0, \omega_{k-1})\}. \tag{16}$$

The set I_k can be further restricted as described later in (25). Finally, we mention that in our implementation of the above procedure we used the alternative maximization method as described in [17] to perform the nonlinear maximization step, although other methods could be used.

One step in the procedure requires testing if $\rho^* \in D$. This test can be implemented using a Schur-Cohn algorithm as we now describe. From ρ^* form the $2n$ degree polynomial $z^n S_{\rho^*}(z)$ as defined in (4). This polynomial has zeros at $e^{\pm j\omega_i}$ for $i = 1, 2, \cdots, k$, so we divide this polynomial by

$$C(z) = \prod_{i=1}^{k} (z^2 - 2 \cos \omega_i z + 1) \tag{17}$$

to form the remainder polynomial $R(z)$. Now, $\rho^* \in D$ if $R(z)$ has no zeros of odd multiplicity on the unit circle. (In fact, $R(z)$ will have no zeros on the unit circle except in the rare case that the optimum solution is found using k frequency constraints, when in fact there are more than k zero frequencies; this is a degenerate case, as it occurs for points s on a set of measure zero.) The needed zero test for $R(z)$ can be implemented using a Schur-Cohn test; see [18], [19] for details.

We summarize the above discussion by a concise statement of the algorithm.

1) Set $k = 1$.
2) Form $Q^*(\boldsymbol{\omega})$ in (14). Find local maxima of $Q^*(\boldsymbol{\omega})$ on the region I_k defined in (16) (using, for example, the alternative maximization method described in [17]). For each local maximum:
   a) Find A^* using (13) and check condition 2) of Theorem 4.
   b) Find ρ^* using (12) and check if $\rho^* \in D$ using the Schur-Cohn procedure described above.
3) If an admissible solution is found in the previous step, it is the optimal solution ρ^{opt}. If not, increment k and go to step 2.

One important feature of the above algorithm is that the nonlinear maximizations are carried out in low dimensional space. This is in contrast to methods which parameterize the covariances in terms of MA or ARMA parameters, where the dimension of the minimization problem is fixed at n. The procedure presented above starts on with $Q^*(\omega_1)$; that is, we maximize a function over a single dimensional variable ω_1, which itself lies in a compact region $\omega_1 \in [0, \pi]$. If an admissible solution is found using k constraints, this solution is the optimum one (by Theorem 4) and we need not search for a solution using a larger number of constraints. For many problems, admissible solutions are found for small numbers of constraints k, and thus the nonlinear maximization is a low dimensional problem. For an MA order of n, maximization of $Q_k^*(\boldsymbol{\omega})$ is a maximization in over at most $\lfloor n/2 \rfloor + 1$ variables, compared to an $n + 1$ dimensional minimization if MA parameters are used as in [6], [9], [11].

V. Further Characterization of Solutions

In this section we consider some geometrical properties of the solution to the minimization problem P. These properties provide some insight on the spectral properties of the optimum solution with respect to s, or equivalently, to $g(\omega)$. These properties lead to additional necessary conditions on the solution to the minimization problem, and these conditions can be used to reduce the regions over which one needs to find local maxima of $Q^*(\boldsymbol{\omega})$.

A. Variance Bounds

An immediate result of Theorem 2 is given below:

Corollary 1: Assume $s \notin D$ and ρ^{opt} is the optimal solution to the minimization problem P, then $\rho_0^{opt} > s_0$.

Proof: From (12) and (13) we have $\rho_0^* - s_0 = \Sigma_{i=1}^k A_i^*$. Since $A_i^* \geq 0$ for $i = 1, \cdots, k$, and not all $A_i^* = 0$, the result follows. ■

The above corollary states that the estimated variance of the MA(n) process is always increased to arrive at the closest NND covariance sequence. An upper bound on this variance can also be obtained. If $M = -\min_{0 \leq \omega \leq \pi} g(\omega)$, then $\rho = [s_0 + M, s_1, \cdots, s_n]$ is an admissible solution with error $Q = M^2$; thus we have

$$s_0 < \rho_0^{opt} \leq s_0 + M.$$

B. Characterization of Zero Frequencies

It is of interest to obtain properties of the solution to the minimization problem in terms of the original given vector s or, equivalently, in terms of $g(\omega)$. In particular, the solution to the minimization problem can be simplified if we find restrictions on the regions of frequencies for which $f(\omega, \rho^{opt}) = 0$. To this end, we consider the following two conjectures:

Conjecture C1: If $f(\omega_0, \rho^{opt}) = 0$, then $g(\omega_0) \leq 0$.

Conjecture C2: If $g(\omega) < 0$ for $\omega \in (a, b)$, and if $g(a) = g(b) = 0$, then $f(\omega, \rho^{opt}) = 0$ for at least one $\omega \in [a, b]$.

Simply stated, these conjectures say that the spectral density function corresponding to the optimum solution to the minimization problem is zero at ω_0 if and only if $g(\omega)$ is negative there. Conjectures C1 and C2 seem reasonable from an approximation point of view. In fact, as the MA order $n \rightarrow \infty$, the solution to the minimization problem becomes

$$f_\infty(\omega, \rho^{opt}) = \max\{g(\omega), 0\} \qquad \text{for } 0 \leq \omega \leq \pi \tag{18}$$

and in this case, both conjectures are satisfied. It turns out that for finite MA order n, neither conjecture is true in general. We discuss each conjecture below.

1) Conjecture C1: In general, it is not true that if $f(\omega, \rho^{opt}) = 0$ at some frequency ω_0, then $g(\omega_0) < 0$. A counterexample for $n = 3$ is given by

$$s = [3.8, -6.16, 3.34, -1.]^T.$$

In this case, optimal NDD solution is found to be

$$\rho^{opt} = [3.80538, -6.15485, 3.34475, -0.99528]^T.$$

It is readily verified that $f(\omega, \rho^{opt}) = 0$ for $\omega = 70.126°$, but $g(70.126°) > 0$.

The reason that conjecture C1 does not hold in general results from acute angles on the boundary of D. A simplified sketch of the problem is shown in Fig. 3. In this figure, H_{ω_1} is the hyperplane corresponding to frequency ω_1; that is, H_{ω_1} is the set of all points ρ for which $f(\rho, \omega_1) = 0$. H_{ω_2} is similarly defined. It can be seen that ρ^{opt} is the closest point in D to s, and that $f(\omega, \rho^{opt}) = 0$ at frequencies ω_1 and ω_2. Also, $g(\omega_1) < 0$ and $g(\omega_2) > 0$, which follows from the fact that s is on the same side of H_{ω_2} as D, but on the opposite side of H_{ω_1} from D.

The angle between two hyperplanes H_{ω_1} and H_{ω_2} cannot be greater than a certain amount, and this angle bound can be used to restrict the regions of zero frequencies, as the following theorem shows.

Theorem 5: Let ρ^{opt} be the solution to the minimization problem P, and assume that $f(\omega, \rho^{opt}) = 0$ for k distinct frequencies $\omega_1, \cdots, \omega_k$. Define

$$M = \max_{0 \leq \omega \leq \pi} - g(\omega)$$

$$\delta_n = \min_{0 \leq \omega_1, \omega_2 \leq \pi} \left\langle \frac{b(\omega_1)}{\|b(\omega_1)\|}, \frac{b(\omega_2)}{\|b(\omega_2)\|} \right\rangle. \tag{19}$$

(Note that $\delta_n \leq 0$ and depends only on the model order n.) Then

$$g(\omega_i) \leq -M\delta_n \|b(\omega_i)\|. \tag{20}$$

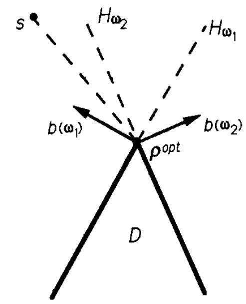


Fig. 3. Geometric description explaining the counterexample for conjecture C1.

Proof: Equation (19) states that the cosine of the angle between the $b(w_i)$ vectors is no less than δ_n. From Theorem 2 we know that the vector $s - \rho^{opt}$ is in the convex cone spanned by the $b(\omega_i)$ vectors. These two statements imply

$$\left\langle \frac{s - \rho^{opt}}{\|s - \rho^{opt}\|}, \frac{b(\omega_i)}{\|b(\omega_i)\|} \right\rangle \geq \delta_n \tag{21}$$

$$\Rightarrow -g(\omega_i) + f(\rho^{opt}, \omega_i) \geq \delta_n \|s - \rho^{opt}\| \, \|b(\omega_i)\|$$

$$\Rightarrow g(\omega_i) \leq -\delta_n M \|b(\omega_i)\|. \tag{22}$$

■

If the zero frequencies are such that the angle between their corresponding normal hyperplanes is always acute, then conjecture 1 holds. This is stated below.

Corollary 2: Let ρ^{opt} be the solution to the minimization problem P, and assume that $f(\omega, \rho^{opt}) = 0$ for k distinct frequencies $\omega_1, \cdots, \omega_k$. If $\langle b(\omega_i), b(\omega_j) \rangle \geq 0$ for $j = 1, \cdots, k$, then $g(\omega_i) \leq 0$.

Proof: In this case the cosine of the angle between the $b(\omega_i)$ vectors is greater than or equal to zero, so $\delta_n \geq 0$, and the result follows immediately. ■

Theorem 6: Under the assumptions of Theorem 2, there is at least one zero frequency ω_i such that $g(\omega_i) < 0$. Thus, if $k = 1$, then $g(\omega_1) < 0$.

Proof: Equations (12) and (13) give

$$s - \rho^* = \sum_{i=1}^{k} A_i^* b(\omega_i) \tag{23}$$

with $A_i^* \geq 0$ (here $\rho^* = \rho^{opt}$ is the corresponding solution to the constrained minimization problem (10)). Equation (23) implies

$$\langle s, s - \rho^{opt} \rangle = \sum_{i=1}^{k} A_i^* \langle s, b(\omega_i) \rangle.$$

Using the formulas $Q^*(\boldsymbol{\omega}) = \langle s - \rho^*, s - \rho^* \rangle = \langle s, s - \rho^* \rangle$ and $g(\omega_i) = -\langle s, b(\omega_i) \rangle$, we have

$$Q^*(\boldsymbol{\omega}) = \sum_{i=1}^{k} A_i^* [-g(\omega_i)].$$

Since $Q^*(\boldsymbol{\omega}) > 0$ and $A_i^* \geq 0$, we have the result. ■

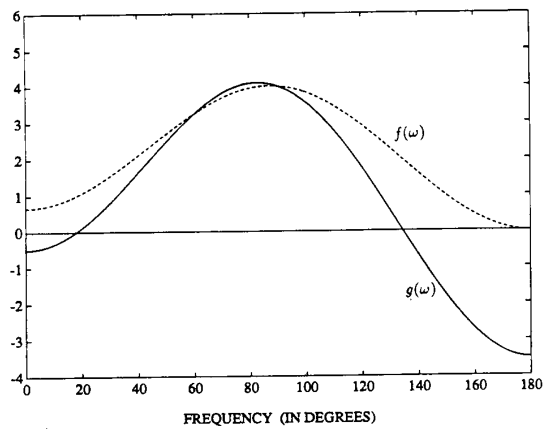


Fig. 4. The plots of $g(\omega)$ and $f(\omega, \rho^{\rm opt})$ for $n = 2$ and $s = [1, 1.5, -3]^T$. Note that if $g(\omega) < 0$ on an interval, it is not necessarily true that $f(\omega, \rho^{\rm opt}) = 0$ on that interval.

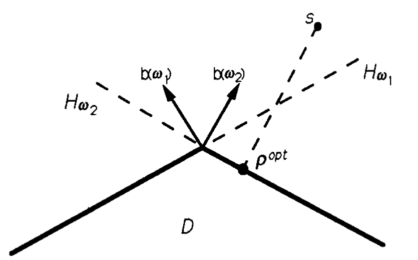


Fig. 5. Geometric description explaining the counterexample to conjecture C2.

Theorem 5 provides an equation which restricts the frequency intervals over which $Q^*(\boldsymbol{\omega})$ must be maximized. Specifically, define

$$J = \{\omega \in [0, \pi] \,|\, g(\omega_i) \leq -M\delta_n \|b(\omega_i)\| \} \tag{24}$$

Then the search region I_k in the maximization algorithm can be replaced by

$$I_k' = I_k \cap J^k \tag{25}$$

where J^k is the Cartesian product of J, k times. Note that J depends only on n and s, so can be computed in the beginning of the algorithm (at step 1).

2) Conjecture C2: Conjecture C2 does not hold in general, as the following example shows. For $n = 2$ and $s = [1, 1.5, -3]^T$, it is readily verified that $g(\omega) < 0$ for $\omega \in [0, 18.0°)$ and $\omega \in (134.5°, 180°]$ (see Fig. 4). The optimal solution is found to be $\rho^{\rm opt} = [13/6, 1/3, -11/6]^T$; $f(\omega, \rho^{\rm opt}) = 0$ has a zero at $\omega_1 = 180°$. The loss function $Q^* = 4.0833$ and $A^* = 7/6$. The solution obtained by enforcing two constraints at $\omega_1 = 0$ and $\omega_2 = 180°$ gives $\rho = [2, 0, -2]^T$ with $Q^* = 4.25$, $A_1^* = -0.25$, and $A_2^* = 1.25$, so this solution is not optimal.

Fig. 5 shows the geometric situation which allows conjecture C2 to be violated. Here $g(\omega_1) < 0$ and $g(\omega_2) < 0$ because s is on the opposite side of the hyperplanes H_{ω_1} and H_{ω_2} from D. However, the projection onto D is such that $\rho^{\rm opt}$ is on the other side of H_{ω_1}, so $f(\omega_1, \rho^{\rm opt}) > 0$. We note that by considering this problem in N, the optimum solution would have zeros at $\omega = 0$ and $\omega = \pi$ (as this problem corresponds to point d in Fig. 2). Thus, the optimal solutions using D or N as the constraint set give results with different geometrical properties.

VI. Examples

Below we consider three examples which illustrate the algorithm for finding the solution to the minimization problem P.

Example 1: $n = 2$, $s = [1, 2, 3]^T$.

For $n = 2$, $f(\omega, \rho^{\rm opt})$ has at most two distinct zero frequencies. We first consider $k = 1$ constraint. A plot of $Q_1^*(\omega)$ is shown in Fig. 6. It can be seen that this function has a local maximum at $\omega_1 = 102.85°$. From (12)–(14) we find that

$$\rho^* = [2.154, 1.743, 1.960]^T$$
$$A^* = 1.154, \qquad Q_1^*(\omega_1) = 2.479.$$

Since $A_1^* \geq 0$ and $f(\omega, \rho^*) \geq 0$, it follows from Theorem 4 that $\rho^* = \rho^{\rm opt}$. The functions $g(\omega)$ and $f(\omega, \rho^{\rm opt})$ are shown in Fig. 6.

Note that $g(\omega)$ has only one negative interval, given by $\omega \in (64.26°, 140.14°)$, and it follows from Theorem 6 that the zero frequency corresponding to the optimal solution must lie in this region.

Example 2: $n = 3$, $s = [1, -2, 0, 0]^T$.

In example 2, we have appended two zeros to s to find the closest third-order NND sequence to a given first-order sequence which is not NND. For $n = 3$, $f(\omega, \rho^{\rm opt})$ has at most two distinct zero frequencies.

From s we find $g(\omega)$ as shown in Fig. 7. We first consider $k = 1$ constraint, and $Q_1^*(\omega)$ is also shown in Fig. 7. Note that $g(\omega)$ is negative for $\omega \in [0, 60°)$, so by Theorem 6, the frequency ω satisfying $f(\omega, \rho^{\rm opt}) = 0$ must lie in this interval. Maximization of $Q_1^*(\omega)$ over the interval $\omega \in [0, 60°)$ gives $\omega_1 = 23.47°$, with corresponding

$$\rho^* = [1.345, -1.684, 0.235, 0.116]^T$$
$$A_1^* = 0.345, \qquad Q_1^*(\omega_1) = 0.288.$$

This solution is admissible ($A_1^* \geq 0$) and in D, so it is the optimal solution $\rho^{\rm opt}$. $f(\omega, \rho^{\rm opt})$ is also shown in Fig. 7.

Example 3: $n = 4$, $s = [1, 2, 3, 4, 5]^T$.

This case is shown in Fig. 8. For $n = 4$, $f(\omega, \rho^{\rm opt})$ has at most three distinct zero frequencies. Using $k = 1$ constraint gives no admissible solutions at maxima of $Q_1^*(\omega)$. With $k = 2$ constraints, a local maximum of $Q_2^*(\boldsymbol{\omega})$ is found at $(\omega_1, \omega_2) = (55.6°, 138.3°)$, with corresponding

$$\rho^* = [3.556, 2.872, 2.283, 2.187, 3.011]^T$$
$$A^* = [2.120, 0.436]^T, \qquad Q^*(\boldsymbol{\omega}) = 33.75.$$

Again according Theorem 4 this is the optimal solution.

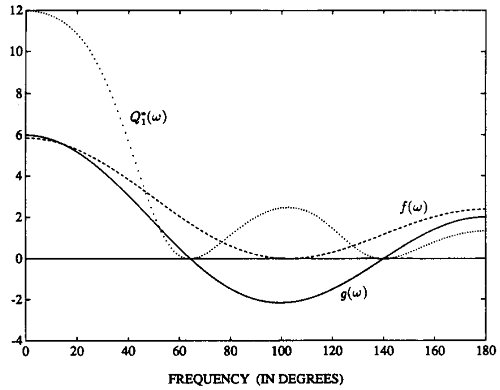


Fig. 6. Example 1 original, optimal NND, and distance measure functions.

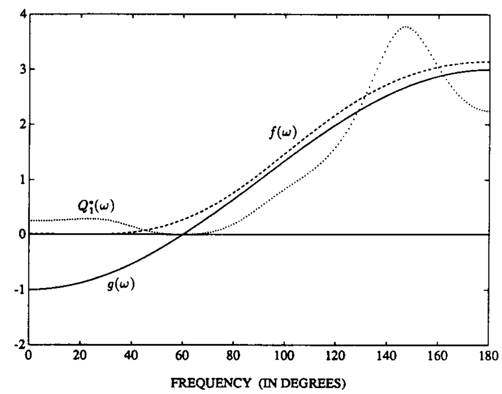


Fig. 7. Example 2 original, optimal NND, and distance measure functions.

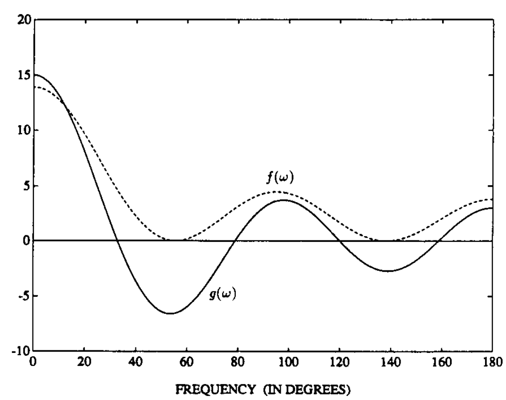


Fig. 8. Example 3 original and optimal NND functions.

VII. Conclusions

We have considered the problem of finding the closest nonnegative definite MA covariance sequence to a given estimate which may not be nonnegative definite. We developed an algorithm which is based on a set of constrained minimization problems, each parameterized by the zero frequencies of the spectral density function corresponding to the optimal solution. The algorithm entails first solving a simple minimization problem with linear constraints whose closed-form solution is given by a projection onto a subspace. These solutions lie either outside the set of NND sequences, or on its boundary; if the solution lies on the boundary, it is the optimal solution.

One property of this algorithm is that we consider the problem directly in the space of covariance sequence elements. As a result, the nonlinear maximization step is performed on sets of low dimension (up to $\lfloor n/2 \rfloor$, where n is the MA order). In addition, by considering the minimization problem in this space, we were able to characterize some of the geometrical properties of the optimal solution in terms of the locations of its zero frequencies.

References

[1] R. L. Moses and A. A. Beex, "Instrumental variable adaptive array processing," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, no. 2, pp. 192-202, Mar. 1988.

[2] M. Kaveh, "High resolution spectral estimation for noisy signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-27, no. 3, pp. 286-287, June 1979.

[3] S. Kay, "A new ARMA spectral estimator," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, no. 5, pp. 585-588, Oct. 1980.

[4] J. A. Cadzow, "ARMA spectral estimation: A model equation error procedure," in *Proc. Int. Conf. Acoust., Speech, Signal Processing* (Denver, CO), Apr. 9-11, 1980, pp. 598-602.

[5] R. L. Moses, J. A. Cadzow, and A. A. Beex, "A recursive procedure for ARMA modeling," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, no. 4, pp. 1188-1196, Oct. 1985.

[6] A. A. Beex, "ARMA modeling of covariance sequences," in *Third ASSP Workshop Spectrum Estimat. Modeling* (Boston, MA), Nov. 17-18, 1986, pp. 41-44.

[7] A. A. Beex and L. L. Scharf, "Covariance sequence approximation for parametric spectrum modeling," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, no. 5, pp. 1042-1052, Oct. 1981.

[8] A. A. Beex, "ARMA covariance realization from noisy data," in *Proc. 1986 Int. Conf. Acoust., Speech, Signal Processing* (Tokyo, Japan), 1986, pp. 2747-2750.

[9] A. A. Beex, "Robust approximate stochastic modeling and partial realization," in *8th IFAC Symp. Identification Syst. Parameter Estimat.*, Aug. 27-31, 1988, pp. 1399-1404.

[10] R. L. Moses, "Optimal approximate stochastic partial realization," in *Proc. 1985 Symp. Math. Theory Networks Syst.* (Stockholm, Sweden), June 10-14, 1985.

[11] G. Feyh and C. T. Mullis, "Moving average separation," in *Proc. 1988 Int. Conf. Acoust., Speech, Signal Processing*, Apr. 11-14, 1988, pp. 2280-2283.

[12] R. L. Moses and D. Liu, "On nonnegative definiteness of estimated moving average autocovariance sequences," in *Proc. 1988 Int. Conf. Acoust., Speech, Signal Processing*, Apr. 11-14, 1988, pp. 2340-2343.

[13] A. O. Steinhardt, "Correlation matching by finite length sequences," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-36, no. 4, pp. 545-559, Apr. 1988.

[14] A. O. Steinhardt, "Reconstructing a finite sequence from several of its correlation lags," in *Proc. Int. Conf. Acoust., Speech, Signal Processing* (Dallas, TX), Apr. 6-9, 1987, pp. 2019-2022.

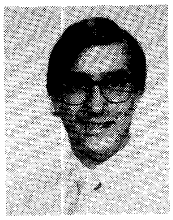
[15] D. G. Luenberger, *Optimization by Vector Space Methods.* New York: Wiley, 1969.

[16] D. Goldfarb and A. Idnani, "A numerically stable dual method for solving strictly convex quadratic programs," *Math. Programming*, vol. 27, pp. 1-33, 1983.

[17] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-36, no. 10, pp. 1553-1560, Oct. 1988.

[18] E. I. Jury, *Inners and Stability of Dynamic Systems.* New York: Wiley, 1974.

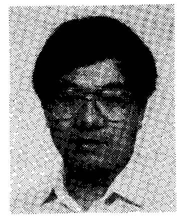
[19] P. Stoica and R. Moses, "On the unit circle problem: the Schur-Cohn procedure revisited," Tech. Rep. SAMPL-87-06, Dep. Elec. Eng., Ohio State Univ., 1987.



Randolph L. Moses (S'78–M'83–SM'90) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Virginia Polytechnic Institute and State University in 1979, 1980, and 1984, respectively.

During the summer of 1983 he was a SCEEE Summer Faculty Research Fellow at Rome Air Development Center, Rome, NY. From 1984 to 1985 he was with the Eindhoven University of Technology, Eindhoven, The Netherlands, as a NATO postdoctoral fellow. Since 1985 he has been an Assistant Professor in the Department of Electrical Engineering, Ohio State University. His research interests are in digital signal processing, and include parametric time series analysis, system identification, and model reduction.

Dr. Moses is a member of Eta Kappa Nu, Tau Beta Pi, Phi Kappa Phi, and Sigma Xi.



Duixian Liu (S'84) was born in Shaanxi, China, on September 13, 1960. He received the B.S. degree in electrical engineering from Northwest Telecommunications Engineering Institute, Xi'an, China, and the M.S. and Ph.D. degrees in electrical engineering from Ohio State University, Columbus, in 1986 and 1990, respectively.

Presently he is with Valor Enterprises, Inc., in West Milton, OH. His current research interests are in electromagnetic theory and scattering, antennas, digital signal processing, spectral estimation, and time series analysis.