

Mathematics Note

Note 77

May 1982

A DISCUSSION OF VARIOUS APPROACHES TO THE
LINEAR SYSTEM IDENTIFICATION PROBLEM

Tapan K. Sarkar
Soheil A. Dianat
Department of Electrical Engineering
Rochester Institute of Technology
Rochester, New York 14623

Donald D. Weiner
Department of Electrical & Computer Engineering
Syracuse University
Syracuse, New York 13210

ABSTRACT

This paper deals with the pole-zero identification of a linear system from a measured input-output record. One objective is to show that the pencil-of-function method minimizes a weighted version of the Kalman equation error. It follows that the pencil-of-function method is capable of yielding robust estimates for poles located in a given region of the complex s -plane. The second objective of this paper is to illustrate that identical sets of equations arise in three supposedly different analytical techniques for obtaining the impulse response of a system. The techniques investigated are: 1) the least squares technique based on the discrete Wiener-Hopf equation, 2) Pisarenko's eigenvalue method, and 3) Jain's pencil-of-function method. The proof of equivalence is valid only for the noise-free case when the system order is known. Instead of using the conventional differential-equation formulation, equivalence is shown with the integral form utilized in the pencil-of-function method.

This work has been supported by the Office of Naval Research under Contract N00014-79-C-0598

1. INTRODUCTION

In linear system identification, one is often interested in obtaining a pole-zero model of an unknown system from measured records of the input and output. If $x(t)$ and $y(t)$ are the respective time domain input and output to the system, then we are interested in characterizing the impulse response $h(t)$ by a sum of complex exponentials, i.e.

$$h(t) \approx \sum_{i=1}^n A_i \exp(s_i t). \quad (1)$$

Here n is referred to as the order of the system. s_i and A_i are the poles and the residues at the poles, respectively. In the Laplace domain, the problem is to model the transfer function $H(s)$ (which is the Laplace transform of $h(t)$) by a ratio of two rational functions as

$$\frac{Y(s)}{X(s)} = H(s) \approx \frac{b_0 + b_1 s + \dots + b_m s^m}{a_0 + a_1 s + \dots + a_n s^n} \triangleq \frac{B(s)}{A(s)} \quad (2)$$

where $Y(s)$ and $X(s)$ are the Laplace transforms of the input and output, respectively. Equality in (2) is attained when $y(t)$ and $x(t)$ are noise free and the system order n is exactly chosen.

Three basic approaches to solving the identification problem are:

- (A) Least Squares approach (based on the Wiener-Hopf technique) [1-3]
- (B) Eigenvalue method (based on Koopman's results [13] which were later applied by Levine [12] and Pisarenko [4-5])
- (c) Pencil-of-function method (based on the linear dependence/independence of a set of functions) [6-8].

In this paper, we show that the three techniques yield analytically equivalent equations when there is no noise in the measured waveforms $x(t)$ and $y(t)$ and the system order n is correctly chosen. However, in the presence of noise, performance differs from one technique to another.

2. THE CONCEPT OF ERROR IN THE VARIOUS TECHNIQUES

Given a specified input $x(t)$, one would like to minimize the mean squared error between the actual output $y(t)$ and the predicted output from the system model. In the Laplace domain, this is mathematically equivalent to minimization of $|E^1(s)|^2$ where

$$E^1(s) = Y(s) - \frac{B(s)}{A(s)} X(s) = \frac{Y(s)A(s) - B(s)X(s)}{A(s)} \triangleq \frac{E(s)}{A(s)} \quad (3)$$

and the unknowns a_i and b_j appear in A and B [as defined in (2)].

However, even though minimization of $|E^1(s)|^2$ with respect to b_j is a linear problem, the minimization of the squared error with respect to a_i is a nonlinear problem [9]. Hence, we tend to minimize $|E(s)|^2$ (where $E(s)$ is popularly known as the equation error, after Kalman [10]) rather than $|E^1(s)|^2$. This is because minimization of $|E(s)|^2$ with respect to a_i and b_j is a linear problem. In fact, beginning with Kalman [10] in 1958, almost all pole-zero modelling techniques utilize this error criterion. The first two techniques -- the least squares and the eigenvalue methods -- as implemented by present researchers, utilize the minimization of $|E(s)|^2$. On the other hand, the third technique -- the pencil-of-function method -- minimizes a weighted $|E(s)|^2$. This weighting is particularly useful when one is interested in very accurate locations of poles and zeros in a specified region of the complex s -plane.

The obvious question now raised is, "What guarantee does one have of obtaining a 'good' solution if $|E(s)|^2$ is minimized?" We show that when the data is noise free and the system order n is correctly chosen, minimization of $|E(s)|^2$ is equivalent to minimization of $|E^1(s)|^2$. We

observe that if the value of s does not coincide with a zero of $A(s)$ (i.e. $A(s) \neq 0$), then minimization of $|E(s)|$ is indeed equivalent to minimization of $|E^1(s)|$. However, it is not obvious that the same conclusion holds for $s = s_i$, where $A(s_i) = 0$. Since at $s = s_i$, minimization of $E^1(s)$ results in the indeterminate form $0/0$, we apply L'Hospital's rule to obtain

$$E^1(s_i) = \frac{\left\{ \frac{d}{ds} [E(s)] \right\}_{s=s_i}}{\left\{ \frac{d}{ds} [A(s)] \right\}_{s=s_i}} = 0, \quad (4)$$

where use has been made of the fact that $\frac{d}{ds} \{E(s)\}$ is zero independent of $s = s_i$. Hence, when there is no noise in the data and the system order is exactly chosen, minimization of $|E(s)|^2$ is indeed equivalent to minimization of $|E^1(s)|^2$.

However, if $y(t)$ is contaminated with noise, such that the noise contaminated output $Y_N(s)$ is

$$Y_N(s) = Y(s) + N(s), \quad (5)$$

then minimization of the error results in

$$E^1(s_i) = N(s_i). \quad (6)$$

Thus, the output noise plays a crucial role in computation of the system poles by minimizing $|E(s)|^2$. This is a well observed fact for Prony's method [3], which is similar to the least squares technique [3].

However, if one utilizes a noise correction scheme, then minimization of the equation error is quite meaningful and, as is shown later, may yield an efficient estimate of the true solution.

The pencil-of-function method utilizes the concept of linear dependence of a set of functions. Following the notations of Gantmacher [11], Jain has defined [6] a mathematical entity f , by combining two given functions defined on a common interval $[a,b]$ together with a scalar parameter as

$$f(t,\lambda) = \gamma g(t) + h(t). \quad (7)$$

We call f a pencil of functions $g(t)$ and $h(t)$ parameterized by γ .

In this paper we consider sets of pencils

$$\gamma g_1(t)+h_1(t); \gamma g_2(t)+h_2(t); \dots \gamma g_n(t)+h_n(t)$$

wherein the functions $g_i(t)$ and $h_i(t)$ for $i=1,2, \dots n$ span separately a common n -dimensional Hilbert space L_2 with the usual inner product

$$\langle f,g \rangle = \int_a^b f(t)g^*(t)dt \quad (8)$$

where $*$ denotes complex conjugate.

For a fixed set of values of parameters γ , the pencils obviously reduce to a set of functions, and the particular values of γ chosen determine properties such as the linear dependence or independence of the set. The main result concerning the linear dependence of pencil sets is derived in [6] and can be expressed as follows;

Theorem: Given that the pencil set is linearly dependent, the parameter γ must satisfy the polynomial equation

$$\begin{aligned} & \gamma^n \sqrt{G[g_1, g_2, \dots, g_n]} \pm \gamma^{n-1} \sqrt{G[g_{i_1}, \dots; h_{k_1}; \dots, g_{i_{n-1}}]} \pm \dots \\ & \pm \gamma \sqrt{G[h_{k_1}, \dots; g_{i_1}; \dots, h_{k_{n-1}}]} \pm \sqrt{G[h_1, h_2, \dots, h_n]} = 0. \end{aligned} \quad (9)$$

In every sum term here, the i 's and k 's form a complete complementary set of indices over the integers $1, 2, \dots, n$; furthermore, the notation $G[f_1, \dots, f_n]$ stands for the determinant of the n -dimensional Gram matrix of the functions f_1, \dots, f_n . Lastly, we remark that the sign of each sum term is to be determined as indicated in [8].

Translated to our problem, we then must look for the linear dependence of the pencil of set determined by [7]

$$y_0 + \lambda y_1, y_1 + \lambda y_2, \dots, y_{n-1} + \lambda y_n, x_1, x_2, \dots, x_n$$

where y_{i+1} and x_{i+1} are reverse time integrals of the functions y_i and x_i as defined in the next section. It is also shown there that by checking for linear dependence, we are actually minimizing a weighted version of $|E(s)|^2$.

3. THE MATHEMATICAL DEVELOPMENT FOR THE NOISE FREE CASE

The fact different formulations yield the same poles for the noise free case is not widely appreciated. One of the objectives of this paper is to show that different formulations based on different assumptions result in an identical set of analysis equations.

(A) The Least Squares Approach

In the Laplace domain the equation error is given by

$$\begin{aligned} E(s) &= Y(s)A(s) - X(s)B(s) \\ &= \sum_{k=0}^n a_k s^k Y(s) - \sum_{k=0}^m b_k s^k X(s). \end{aligned} \quad (10)$$

In the conventional approaches, we generally deal with (10) and therefore work with the various higher-order derivatives of the functions $y(t)$ and

$x(t)$. Since it is difficult to obtain the derivatives numerically, we deal with the various successive integrals of $y(t)$ and $x(t)$ as is done in the pencil-of-function method. We also show that dealing with the integrals of $y(t)$ and $x(t)$ leads to a minimization of a weighted version of $|E(s)|^2$. We divide (10) by s^n (it is assumed $n > m$). Then

$$\tilde{E}(s) \triangleq \frac{E(s)}{s^n} = \sum_{k=0}^n a_k \frac{Y(s)}{s^{n-k}} - \sum_{k=0}^m b_k \frac{X(s)}{s^{n-k}} \quad (11)$$

We next define the functions $y_i(t)$ and $x_i(t)$ where

$$\left. \begin{aligned} y_0(t) &= y(t) \\ x_0(t) &= x(t) \end{aligned} \right\} \quad (12)$$

$$y_j(t) = \begin{cases} \underbrace{\int_{\infty}^t dt_1 \int_{\infty}^{t_1} dt_2 \dots \int_{\infty}^{t_{j-1}} y_0(t_j) dt_j}_{j \text{ times}} & \text{for } t \geq 0 \\ 0 & \text{for } t < 0 \end{cases} \quad (13)$$

and

$$x_j(t) = \begin{cases} \underbrace{\int_{\infty}^t dt_1 \int_{\infty}^{t_1} dt_2 \dots \int_{\infty}^{t_{j-1}} x_0(t_j) dt_j}_{j \text{ times}} & \text{for } t \geq 0 \\ 0 & \text{for } t < 0 \end{cases} \quad (14)$$

Observe that the integrals in (13) and (14) are from ∞ to t rather than from 0 to t . This is done because numerically stable results are obtained for the Gram matrix in (9) if one integrates from ∞ to t [7].

By taking two sided Laplace transforms, we have

$$\begin{aligned} Y_0(s) &= Y(s) \\ \text{and} \quad X_0(s) &= X(s) \end{aligned} \quad (15)$$

Moreover, using integration by parts,

$$Y_{k+1}(s) = \int_{-\infty}^{\infty} e^{-st} dt \int_{-\infty}^t y_k(\tau) d\tau = \frac{Y_k(s)}{s} = \frac{Y(s)}{s^{(k+1)}} \quad (16)$$

and

$$X_{k+1}(s) = \int_{-\infty}^{\infty} e^{-st} dt \int_{-\infty}^t x_k(\tau) d\tau = \frac{X_k(s)}{s} = \frac{X(s)}{s^{(k+1)}} \quad (17)$$

Substitution of (15), (16) and (17) in (11) yields

$$\tilde{E}(s) = \sum_{k=0}^n a_k Y_{n-k}(s) - \sum_{k=0}^n b_k X_{n-k}(s) \quad (18)$$

By taking the inverse Laplace transform, we get

$$\tilde{e}(t) = \sum_{k=0}^n a_k y_{n-k}(t) - \sum_{k=0}^n b_k x_{n-k}(t) \quad (19)$$

In matrix form (19) can be written as

$$\tilde{e}(t) = [y_n(t), y_{n-1}(t), \dots, y_0(t), -x_n(t), -x_{n-1}(t), \dots, -x_{n-m}(t)] \cdot \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \\ b_0 \\ b_1 \\ \vdots \\ b_m \end{bmatrix} \quad (20)$$

The integral squared error can be expressed as

$$\begin{aligned} \text{ERR} &= \int_0^{\infty} \tilde{e}^2(t) dt = \{[A] \ [B]\}^T \cdot \begin{bmatrix} [YY] & -[YX] \\ -[XY] & [XX] \end{bmatrix} \cdot \begin{bmatrix} [A] \\ [B] \end{bmatrix} \\ &\triangleq \{[A] \ [B]\}^T \cdot [P] \cdot \begin{Bmatrix} [A] \\ [B] \end{Bmatrix} \end{aligned} \quad (21)$$

where the above matrices are defined as

$$[A]^T = [a_0 \ a_1 \ \dots \ a_n]_{1 \times (n+1)} \quad (22)$$

$$[B]^T = [b_0 \ b_1 \ \dots \ b_m]_{1 \times (m+1)} \quad (23)$$

$$[YY] = \begin{bmatrix} \int_0^\infty y_n(t)y_n(t)dt & \int_0^\infty y_n(t)y_{n-1}(t)dt & \dots & \int_0^\infty y_n(t)y_0(t)dt \\ \vdots & \vdots & & \vdots \\ \int_0^\infty y_0(t)y_n(t)dt & \int_0^\infty y_0(t)y_{n-1}(t)dt & \dots & \int_0^\infty y_0(t)y_0(t)dt \end{bmatrix} \quad (24)$$

$$[YX] = \begin{bmatrix} \int_0^\infty y_n(t)x_n(t)dt & \int_0^\infty y_n(t)x_{n-1}(t)dt & \dots & \int_0^\infty y_n(t)x_{n-m}(t)dt \\ \vdots & \vdots & & \vdots \\ \int_0^\infty y_0(t)x_n(t)dt & \int_0^\infty y_0(t)x_{n-1}(t)dt & \dots & \int_0^\infty y_0(t)x_{n-m}(t)dt \end{bmatrix} \quad (25)$$

$$[XX] = \begin{bmatrix} \int_0^\infty x_n(t)x_n(t)dt & \int_0^\infty x_n(t)x_{n-1}(t)dt & \dots & \int_0^\infty x_n(t)x_{n-1}(t)dt \\ \vdots & \vdots & & \vdots \\ \int_0^\infty x_{n-m}(t)x_n(t)dt & \int_0^\infty x_{n-m}(t)x_{n-1}(t)dt & \dots & \int_0^\infty x_{n-m}(t)x_{n-m}(t)dt \end{bmatrix} \quad (26)$$

$$[YX] = [XY]^T \quad (27)$$

and T denotes the transpose of a matrix. Observe that the matrix [P] is a symmetric positive semidefinite matrix. Also, note that the squared integral of $\tilde{e}(t)$ is being minimized rather than that of $e(t)$. The basic difference between $\tilde{e}(t)$ and $e(t)$ is that $\tilde{e}(t)$ is the nth integral of $e(t)$ [equation (11)] - in other words $e(t)$ is passed through n integrators

in cascade to get $\tilde{e}(t)$. In the Laplace domain this is equivalent to minimizing $\left| \frac{E(s)}{s^n} \right|^2$ instead of $|E(s)|^2$. Thus, a weighted least squares error is being minimized. The error $E(s)$ is weighted by the function $1/s^n$ to yield $\tilde{E}(s)$. This implies that with respect to error $\tilde{E}(s)$, the low frequency poles are weighted more heavily than the high frequency poles. Hence, it is expected that minimization of $\tilde{E}(s)$ yields a stable result when the noise is modelled by a high frequency phenomenon. This approach improves the solution for low frequency poles. This was observed when analyzing transient data in [7].

It is important to point out that instead of using successive integrations in (13) and (14) where the functions y_{i+1} and x_{i+1} are defined, one could have utilized a series of band-pass filters instead. This would have amounted to reducing the integrated squared error of $\bar{e}(t)$ where $\bar{e}(t)$ is obtained by passing $e(t)$ through n band-pass filters, with properly adjusted cut-off frequencies. It is conjectured that such a technique would yield very robust estimates for poles in a certain pre-determined region of the complex s -plane.

The next step is to minimize the error ERR given by (21) with respect to a_i and b_j . This is equivalent to taking the derivative of the functional ERR with respect to a_i and b_j and setting the first derivative equal to zero. This leads to the set of equations

$$[P] \cdot \begin{Bmatrix} [A] \\ [B] \end{Bmatrix} = 0. \quad (28)$$

Observe that one solution of this equation is the trivial solution. Hence, we fix $a_0 = 1$ and minimize the functional ERR with this constraint. As pointed out by Koopmans [13] and Levine [12], the minimization process

leads to the set of equations

$$[P_{1R}]_{(n+m+1) \times (n+m+2)} \cdot \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ \vdots \\ a_n \\ b_0 \\ b_1 \\ \vdots \\ b_m \end{bmatrix}_{(n+m+2) \times 1} = 0 \quad (29)$$

which is identical to the discrete Wiener-Hopf equations. Here the matrix $[P_{1R}]$ is the same as the matrix $[P]$ with the first row deleted.

The solution of (29) can be written in the compact form

$$[u] = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \\ b_0 \\ b_1 \\ \vdots \\ b_m \end{bmatrix}_{(n+m+1) \times 1} = \begin{bmatrix} \Delta_{12}/\Delta_{11} \\ \Delta_{13}/\Delta_{11} \\ \vdots \\ \Delta_{1(n+1)}/\Delta_{11} \\ \Delta_{1(n+2)}/\Delta_{11} \\ \Delta_{1(n+3)}/\Delta_{11} \\ \vdots \\ \Delta_{1(n+m+2)}/\Delta_{11} \end{bmatrix}_{(n+m+1) \times 1} \quad (30)$$

where Δ_{ij} is the ij th cofactor of the matrix $[P]$ defined by (21), i.e. Δ_{ij} is the determinant of the matrix remaining after the i th row and j th column are deleted from $[P]$.

The solutions given by (30) yield the global minimum for the functional ERR under the constraint $a_0 = 1$. However, a major objection to this technique is that there is no compelling reason for choosing $a_0 = 1$ as opposed to selecting any a_i or b_j to be unity. When there is no noise and the system order is correctly chosen, the

solutions obtained, as different coefficients are set equal to unity, are identical except for a scalar multiplier. However, when the data are contaminated by noise, each of the choices may lead to a completely different result [3].

[B] The Eigenvector Method

Instead of constraining one of the coefficients to be unity, we may require

$$\sum_{i=0}^n a_i^2 + \sum_{j=0}^m b_j^2 = \{[A] \ [B]\}^T \cdot \begin{Bmatrix} [A] \\ [B] \end{Bmatrix} = 1. \quad (31)$$

Observe that minimization of the functional ERR in (21) is now equivalent to finding the minimum eigenvalue of the matrix [P]. By using (21) and (31), it follows that we need to find

$$\min. \text{ of } [ERR = \{[A] \ [B]\}^T \cdot [P] \cdot \begin{Bmatrix} [A] \\ [B] \end{Bmatrix}] \quad (32)$$

The solution $\{[A] \ [B]\}^T$ is given by the normalized eigenvector corresponding to the minimum eigenvalue of [P].

[C] The Pencil-of-Function Method

In this case, we also minimize the weighted version of the equation error $|E(s)|^2$ given by $|\tilde{E}(s)|^2$. This is equivalent to checking for linear dependence among the following pencil-of-functions:

$$\begin{aligned} & y_n(t) + \lambda y_{n-1}(t); y_{n-1}(t) + \lambda y_{n-2}(t); \dots; y_1(t) + \lambda y_0(t); \\ & x_n(t) + \lambda x_{n-1}(t); x_{n-2}(t) + \lambda x_{n-3}(t); \dots; x_{n+m+1}(t) + \lambda x_{n-m}(t) \end{aligned} \quad (33)$$

As shown in [7] the values of λ for which (33) becomes linearly dependent

are the roots of the polynomial equation

$$\lambda^{n+m+2} a_0 + \lambda^{n+m+1} a_1 + \dots + \lambda b_{m-1} + b_m = 0.$$

The solutions for the coefficients of the polynomial equation are given by

$$\begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \\ b_0 \\ b_1 \\ \vdots \\ b_m \end{bmatrix}_{(n+m+2) \times 1} = \begin{bmatrix} \sqrt{\Delta_{11}} \\ \sqrt{\Delta_{22}} \\ \vdots \\ \sqrt{\Delta_{n+1,n+1}} \\ \sqrt{\Delta_{n+2,n+2}} \\ \sqrt{\Delta_{n+3,n+3}} \\ \vdots \\ \sqrt{\Delta_{n+m+2,n+m+2}} \end{bmatrix}$$

where Δ_{ii} is the i th diagonal cofactor of the matrix $[P]$.

4. PROOF OF EQUIVALENCE AMONG THE THREE TECHNIQUES

To show that the three techniques yield equivalent solutions for the noise free case and when the system order is correctly chosen, we perform the following arithmetic manipulations.

If u_i represents the i th element of the column matrix $[u]$ given by the least squares approach in (30), then

$$u_i = \frac{\Delta_{1i}}{\Delta_{11}} = \frac{\{\det[P]\} \cdot [P^{-1}]_{1i}}{\{\det[P]\} \cdot [P^{-1}]_{11}} \quad (35)$$

where $[P]_{1i}^{-1}$ represents the element belonging to the 1 th row and the i th column of $[P]^{-1}$ and $\det[P]$ is the determinant of the matrix $[P]$.

ERR must be zero. This implies that the smallest eigenvalue of [P] must be zero. In that case, we have from (40)

$$u_i = \frac{v_{ji}}{v_{j1}} \quad (41)$$

where v_{ji} is the i th element of the eigenvector \bar{v}_{ji} [as defined in (37)] corresponding to the eigenvalue λ_j , which is zero. Observe that the result given by (41) is the normalized eigenvector corresponding to the zero eigenvalue of matrix [P] where the elements of the eigenvector have been normalized with respect to the first element. Thus, we have shown that the least squares solution of (35) can be expressed in terms of the normalized eigenvector given by (41).

To establish the equivalence between the least squares technique and the pencil-of-function method, we utilize the cofactor identity

$$\Delta_{ij}^2 = \Delta_{ii} \cdot \Delta_{jj} \quad \text{when } \det[P] = 0. \quad (42)$$

Substitution of (42) into (35) yields

$$u_i = \frac{\Delta_{1i}}{\Delta_{11}} = \frac{\sqrt{\Delta_{11}} \cdot \sqrt{\Delta_{1i}}}{\Delta_{11}} = \frac{\sqrt{\Delta_{1i}}}{\sqrt{\Delta_{11}}} \quad (43)$$

By comparing (34) and (43) we find that (43) is identical to (34) except that the solution given by (34),

$$[W] = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \\ b_0 \\ b_1 \\ \vdots \\ b_m \end{bmatrix},$$

has been scaled by the quantity a_0 .

In conclusion, all three techniques yield equivalent results when the data are noise free and the order of the system has been correctly chosen. However, when noise is present in the data, one of the techniques may have superior performance. In general, which technique performs best depends on the type of data and the assumed order of approximation. The performance of each of these techniques can be improved for noisy data if noise correction is applied to the signal matrix.

5. NOISE CORRECTION FOR THE THREE TECHNIQUES

If it is now assumed that both the input and the output contain some noise in the data, then the expected value of the Gram matrix [P] yields

$$E[P] = [S] + \sigma^2 [Z] \quad (44)$$

where [S] is the noise free signal matrix and [Z] is the unit noise vector covariance matrix and σ^2 is the scale factor. It is assumed signal and noise are uncorrelated and the noise is zero mean. The noise covariance matrix not only contains information about the autocorrelation matrix of the input and output noise, but also about the cross-correlation between the input noise and the output noise components.

The crux of implementing a noise correction procedure is to have a good estimate of [Z]. The scale factor σ^2 for a particular problem can be obtained from the mathematics of the problem.

For the least squares method, the noise correction takes the form of the Markov estimates [12]. For a general case, it is

extremely difficult to incorporate the information of the [Z] matrix in this formulation. However, in certain special cases, particularly when the noise is white, certain simplifications are possible.

The problem of noise correction is much alleviated for the eigenvalue method. It has been shown by Koopmans [13] that to obtain a minimum variance unbiased estimate for the parameters a_i, b_i one solves the generalized eigenvalue problem instead of (32), i.e.

$$\{[P] - \sigma^2[Z]\} \cdot \begin{Bmatrix} [A] \\ [B] \end{Bmatrix} = 0. \quad (45)$$

One has to find σ^2 first and then obtain the eigenvector, corresponding to the zero eigenvalue of the matrix $\{[P] - \sigma^2[Z]\}$. If the noise is Gaussian, in addition to the parameters being unbiased, the Cramer-Rao inequality holds with the equal sign, i.e. the parameters estimates are efficient [12]. Even though iterative methods exist for the solution of the generalized eigenvalue problem, the computation becomes quite time consuming.

Jain [8] has developed a numerically efficient method for noise corrections in the pencil-of-function method provided σ^2 is small. Since, in the pencil of function method, $\det [S]$ is required to be zero, Jain asserts that,

$$\det \{[P] - \sigma^2[Z]\} = 0. \quad (46)$$

when σ^2 is small, Jain expands (39) utilizing the concept of determinant expansion for a sum of two matrices. He obtains

$$\begin{aligned} \det \{[P] - \sigma^2[Z]\} = & \det [P] - \sigma^2 \sum_i \det \{[P|z_i]\} \\ & + \sigma^4 \sum_i \sum_j \det \{[P|z_i z_j]\} + \dots \quad (47) \end{aligned}$$

When σ^2 is small, we have

$$\sigma^2 = \frac{\det [P]}{\sum_i \det \{[P|Z_i]\}} \quad (48)$$

where the symbol $[P|Z_i]$ represents the matrix $[P]$ for which the i th column of $[P]$ has been replaced by the i th column of $[Z]$. Mathematically, (48) is equivalent to

$$\sigma^2 = \frac{1}{[P^{-1}] \odot [Z]} \triangleq \frac{1}{\sum_i \sum_j [P^{-1}]_{ij} [Z]_{ij}}$$

where the symbol \odot represents a matrix kronecker product. Since the matrix $[Z]$ and $[P]$ are known and $\det [P] \neq 0$, this procedure for the solution of σ^2 is quite straightforward. Computationally, this has been found to yield good results [8].

6. CONCLUSION

It has been shown that all the three techniques (least squares method, eigenvalue method and pencil-of-function method) yield mathematically equivalent results for the noise free case when the system order is known. However, when noise is present in the data, the three techniques may yield different solutions. The degree of difference depends on the amount of noise in the data and the choice of the system order n . However, a noise correction technique can be implemented to enhance the resolution of these three techniques. Finally, it is shown that the pencil-of-function method is amenable to a computationally efficient way of implementing noise correction to the signal matrix.

7. REFERENCES

1. J.D. Markel and A.H. Gray, Linear Prediction of Speech, Berlin: Springer-Verlag, 1976.
2. J. Makhoul, "Linear Prediction: A Tutorial Review," Proc. IEEE, Vol. 63, pp. 561-580, April 1980.
3. T.K. Sarkar, J. Nebat, D.D. Weiner, and V.K. Jain, "A Comparison of the Pencil of Function Method with Prony's Method, Wiener Filters and Other Identification Techniques," IEEE Trans. on Ant. and Prop., Vol. AP-30, pp. 89-98, January 1982.
4. V.F. Pisarenko, "The Retrieval of Harmonics from a Covariance Function," Geophysics J. of the Royal Astronomical Society, Vol. 33, pp. 347-366, 1973.
5. C. Geugen, Y. Grenier, and F. Gianella, "Factorial Linear Modelling, Algorithms and Applications," Proc. of Intl. Conf. on ASSP, pp. 618-621, 1981.
6. V.K. Jain, "Filter Analysis by use of Pencil of Functions, Part 1," IEEE Trans. on Circuits and Systems, Vol. CAS-21, pp. 574-579, Sept. 1974.
7. T.K. Sarkar, J. Nebat, D.D. Wiener and V.K. Jain, "Suboptimal Approximation/Identification of Transient Waveforms from Electromagnetic Systems by Pencil of Functions Method," IEEE Trans. on Ant. and Prop., Vol. AP-28, pp. 928-933, November 1980.
8. V.K. Jain, T.K. Sarkar and D.D. Weiner, "Rational Modelling by Pencil of Functions Method," TR 5 on Contract N00014-79-C-0598; Rochester, New York, August 1981.
9. K. Steiglitz, "On the Simultaneous Estimation of Poles and Zeros in Speech Analysis," IEEE Trans. on Acoustics, Speech and Signal Processing, Vol. ASSP-25, pp. 229-234, June 1977.
10. R.E. Kalman, "Design of a Self-Optimizing Control System," Trans. ASME, Vol. 80, pp. 468-474, Feb. 1958.
11. F.R. Gantmacher, Theory of Matrices, Vol. 1, New York: Chelsea 1960.
12. M.J. Levine, "Estimation of the Characteristics of Linear Systems in the Presence of Noise," Ph.D. Thesis, Columbia University, 1959.
13. T. Koopmans, "Linear Regression Analysis of Economic Time Series," Netherlands Economic Institute, Haarlem-DeErren F. Bohn N.V. - 1937.