April 1982

IDENTIFICATION OF MULTIVARIATE LINEAR SYSTEMS

James M. Griffith, Ph.D

U.S. Department of Energy

Idaho Operations Office • Idaho National Engineering Laboratory

This is an informal report intended for use as a preliminary or working document

Prepared for the
U. S. Department of Energy
Idaho Operations Office
Under DOE Contract No. DE-AC07-76ID01570

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

MASTER
EG&G Idaho

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

DISCLATMER

This book was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. References herein to any specific commerical product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United State Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

EGG-EE--5842

EGG-EE-5842

DE82 015201

IDENTIFICATION OF MULTIVARIATE LINEAR SYSTEMS

DISCLAIMER

This book was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

James M. Griffith, Ph.D

Published April 1982

EG&G Idaho, Inc.
Idaho Falls, Idaho 83415

Prepared for the
U.S. Department of Energy
Idaho Operations Office
Under DOE Contract No. DE-AC07-76ID01570

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

MBW

CONTENTS

ABSTRACT.....	iii
INTRODUCTION.....	1
SCALAR PROBLEMS.....	2
LS and Correlation Parameter Estimators.....	2
Measurement Noise and Parameter Estimation Bias.....	3
The Shifted-Least-Squares Parameter Estimator.....	6
Innovation Representation and Estimating Noise Variance.....	10
MULTIVARIATE PROBLEMS.....	14
Problem Formulation.....	14
Canonical Forms.....	15
Covariance Function.....	20
Parameter Estimation.....	30
Parameter Estimation Accuracy.....	32
Noise Considerations.....	37
Structure Estimation.....	37
RESULTS.....	29
Example 1.....	39
Example 2.....	42
Example 3.....	46
Example 4.....	47
DISCUSSION.....	52
REFERENCES.....	56

FIGURES

1. Autocorrelation function for measurement.....	5
2. Parameter estimate as a function of τ	9
3. Correlation function and one-standard-deviation error limits shown as solid and dashed lines respectively.....	12
4. Scalar system with white process and measurement noise.....	39
5. Estimated autocorrelation function.....	40
6. Scalar system with white noise as control signal.....	42

FIGURES CONT'D

7. Actual covariance functions.....	43
8. Covariance functions as estimated from 1000 data points.....	45
9. Scalar system with colored noise as the control signal.....	46
10. Tree diagrams for structure estimation.....	49

TABLES

1. Structure estimation.....	41
2. Parameter estimation.....	41
3. Structure estimation.....	44
4. Parameter estimation.....	44
5. Structure estimation.....	46
6. Parameter estimation.....	47
7. Parameter estimation for a 1, 3 structure.....	51

ABSTRACT

Multivariate identification problems are treated with a least-squares approach. A chapter on scalar problems focuses attention on the classical parameter-estimate bias problem caused by measurement noise and develops a straightforward and effective way to remove the bias. A chapter on multivariate problems generalizes the bias removal method and develops a form selection procedure. The form selection procedure generally ensures more accurate identification than is possible with identification methods which rely on a fixed form. The concept of a form selection procedure is new to this work.

A results chapter presents four example problems. Each example illustrates specific features of the identification technique. As a collection the examples emphasize the complexity of system identification and demonstrate that identification techniques will perform well when carefully applied.

INTRODUCTION

System identification is the modeling of an existing system from experimental data. The problem can be decomposed into smaller problems including; structure estimation, parameter estimation, and input evaluation. Identification papers often consider only one of the smaller problems and this has sometimes caused confusion about the ultimate goal of identification. This report attempts to present the various aspects of identification within a unified least-squares framework [1 - 9].

Identification methods are needed because investigators often need to control, predict the state of, or understand systems that are difficult or impossible to model by applying physical principles.

This report includes a chapter on scalar problems that are presented with minimal notation. Voiding the chapter of excess notation allows a simple treatment of several interesting points including: (a) Demonstration that the least-squares and correlation parameter estimators are the same. (b) Showing why measurement noise causes biased parameter estimates. (c) Finding a method to avoid the bias. (d) Illustration that the innovations-covariance matrix can be estimated within the least-squares framework. All of these points have been considered in the identification literature. The emphasis here is to provide simple explanations that lead naturally to multivariate methods.

The multivariate problem is analogous to the scalar problem with one major exception. For multivariate systems a canonical form must be specified. Essentially, a multivariate system has too many parameters to be estimated from input-output data. To eliminate this problem, the system equations are placed in a canonical form that contains a restricted number of parameters. The multivariate chapter discusses a class of forms and a technique for selecting a form which provides accurate parameter estimates from a given data set.

SCALAR PROBLEMS

This chapter illustrates several basic ideas concerning system identification. Hopefully these ideas will facilitate reading the following chapter on multivariate problems. Several scalar problems with varying degrees of complexity are analyzed to demonstrate the different ideas.

LS and Correlation Parameter Estimators

Least squares (LS) and correlation methods of parameter estimation are the same method. Although this fact is not widely known (except to specialists), a short example clarifies the situation.

Consider a system described by

$$x(k) = \phi x(k-1) + w(k-1) \quad (1)$$

where x , ϕ , and w respectively denote the system state, parameter, and a zero mean white noise. The parameter can be estimated from a set of measured x values available for $k = 1, 2, \dots, N$. By using the data Equation (1) can be written at a series of k values to yield

$$x(1) = \phi x(0) + w(0)$$

$$x(2) = \phi x(1) + w(1)$$

$$x(3) = \phi x(2) + w(2)$$

$$\vdots$$

or by collecting the equations into a matrix equation

$$\begin{bmatrix} x(1) \\ x(2) \\ x(3) \\ \vdots \\ x(N) \end{bmatrix} = \begin{bmatrix} x(0) \\ x(1) \\ x(2) \\ \vdots \\ x(N-1) \end{bmatrix} \phi + \begin{bmatrix} w(0) \\ w(1) \\ w(2) \\ \vdots \\ w(N-1) \end{bmatrix} \quad (2)$$

The least squares estimate of $\phi(\hat{\phi}_{LS})$ is obtained by premultiplying both sides of Equation (2) by $[x(0) \ x(1) \ \dots \ x(N-1)]$ to obtain

$$\sum_{k=1}^N x(k-1) x(k) = \phi \sum_{k=1}^N x(k-1) x(k-1) + \sum_{k=1}^N x(k-1) w(k-1) \quad (3)$$

and since no usable information is available to describe the last term on the right of Equation (3), let

$$\hat{\phi}_{LS} = \frac{\sum_{k=1}^N x(k-1) x(k)}{\sum_{k=1}^N x(k-1) x(k-1)} \quad (4)$$

From equation (4) the interpretation of the LS estimator as the correlation estimator is immediate. Correlation functions, r , are usually estimated as

$$\hat{r}(\tau) = \frac{1}{N} \sum_{k=1}^N x(k) x(k+\tau)$$

and thus Equation (4) can be rewritten as

$$\hat{\phi}_{LS} = \hat{r}(1)/\hat{r}(0) = \hat{\phi}_{CORR} \quad (5)$$

Where $\hat{\phi}_{CORR}$ denotes the parameter estimate from correlation values.

Measurement Noise and Parameter Estimation Bias

The least squares estimator discussed above is readily understood and applied and has been extensively employed. However, it results in biased parameter estimates if measurement noise is present. The following example illustrates the problem.

Consider a system and measurement described by

$$x(k) = \phi x(k-1) + w(k-1) \quad (6)$$

and

$$z(k) = x(k) + v(k) \quad (7)$$

where z and v denote measurement and measurement noise respectively. Combining Equations (6) and (7) to obtain a difference equation in the measurement yields

$$z(k) = \phi z(k-1) + w(k-1) + v(k) - \phi v(k-1) . \quad (8)$$

The LS estimator can be obtained from Equation (8) by writing out a series of equations and then forming a matrix equation as was done in the previous section. However, the notation is simpler if we directly start working in terms of correlations. Multiplying both sides of Equation (8) by $z(k-1)$ and taking expected values yields

$$r_z(1) = \phi r_z(0) + 0 + 0 - \phi \rho(0) \quad (9)$$

where $\rho(\tau) \equiv E\{z(k)v(k+\tau)\}$.

The LS estimator is still defined as $\hat{\phi}_{LS} \equiv \hat{r}(1)/\hat{r}(0)$, just as in the previous section, but from Equation (9) one might suspect that $\hat{\phi}_{LS}$ is a biased estimator because $E\{\rho(0)\} \neq 0$.

To calculate the bias let $\hat{r} = r + \tilde{r}$ and $\hat{\phi} = \phi + \tilde{\phi}$ where $(\hat{\cdot})$ denotes estimate and $(\tilde{\cdot})$ denotes estimation error. With these definitions Equation (9) can be manipulated to yield

$$\tilde{\phi} = [\tilde{r}_z(1) - \phi \tilde{r}_z(0) - \phi \rho(0)]/\hat{r}_z(0). \quad (10)$$

Assuming an unbiased $\hat{r}$ and enough data points (N) so that $\hat{r}(0) \approx r(0)$,

$$E[\tilde{\phi}] = -\phi \rho(0)/r_z(0) = -\phi \text{var}[v]/r_z(0). \quad (11)$$

The bias problem can also be understood from a graphical viewpoint. This approach helps to motivate the shifted-least-squares (SLS) estimator developed in the following section. The SLS estimator eliminates the bias problem.

For the graphical approach consider Figure 1.

Figure 1. Autocorrelation function for measurement.

From the figure

$$\phi = \frac{r(\tau+1)}{r(\tau)}, \quad \tau > 0$$

and thus the parameter can be calculated exactly from the correlation function. However, the LS estimator is calculated as $\hat{\phi}_{LS} = r(1)/r(0)$ and using $\hat{\phi}_{bias} = \hat{\phi} - \phi$ yields

$$\hat{\phi}_{bias} = -\phi \frac{\text{var}[v]}{r_z(0)} \quad (12)$$

which is identical to Equation (11). The LS parameter estimator is unbiased only for the case of zero measurement noise. However, the bias, caused by white measurement noise, can be eliminated by basing the estimator on correlation values with shifts (τ) greater than zero. The bias caused by other types of measurement noise can be similarly removed. The measurement autocorrelation function consists of contributions from the measurement noise and from the system dynamics; for a stable system both contributions approach zero as τ increases. An unbiased (or very nearly unbiased) parameter estimate

can be obtained by using correlation values from an area of the correlation function where the contribution from system dynamics is substantially greater than the contribution from measurement noise.

The following section provides a more formal and complete development of the SLS parameter estimator.

The Shifted-Least-Squares Parameter Estimator

Problem Formulation

Assume that a system and measurement can be described by Equations (13) and (14);

$$x(k+1) = \phi x(k) + w(k) \quad (13)$$

$$z(k) = x(k) + v(k) \quad (14)$$

where

x = system state

ϕ = parameter

z = measurement

w = process noise

v = measurement noise

and the noise sequences are zero mean, white, and uncorrelated. The system is stable and the parameter must be estimated from a set of measurements.

Parameter Estimator

Since the measurements are the only available data to describe the system, Equations (13) and (14) are converted to a difference equation in the measurement, $z(k) = \phi z(k-1) + w(k-1) + v(k) - \phi v(k-1)$, or

$$z[(k+\tau)+1] = \phi z(k+\tau) + w(k+\tau) + v[(k+\tau) + 1] - \phi v(k+\tau) . \quad (15)$$

Multiplying each term of Equation (15) by $z(k)$ and taking expected values yields

$$r(\tau+1) = \phi r(\tau) + 0 + 0 - \phi \rho(\tau) \quad (16)$$

where

$$\tau \geq 0.$$

The LS and SLS estimators can both be developed from Equation (16). Equation (16) can be rewritten as

$$\hat{r}(\tau+1) = \phi \hat{r}(\tau) + \tilde{r}(\tau+1) - \phi \tilde{r}(\tau) - \phi \rho(\tau) \quad (17)$$

by using

$$\hat{r} \equiv r + \tilde{r}$$

$(\hat{\cdot})$ = estimate

$(\tilde{\cdot})$ = estimation error.

In Equation (17), only the $\hat{r}$ terms can be estimated directly from experimental data. Thus the last three terms are considered noise (loosely speaking) and a parameter estimate is defined as

$$\hat{\phi}(\tau) = \frac{\hat{r}(\tau+1)}{\hat{r}(\tau)} \quad (18)$$

To determine the estimator accuracy we consider the mean and variance of the estimation error. Using $\hat{\phi} = \phi + \tilde{\phi}$ and $\hat{r}(\tau)\hat{\phi}(\tau) = \hat{r}(\tau+1)$ in Equation (17) yields

$$\tilde{\phi}(\tau) = [\tilde{r}(\tau+1) - \phi \tilde{r}(\tau) - \phi \rho(\tau)] / \hat{r}(\tau). \quad (19)$$

By using the maximum likelihood method to estimate the correlation functions [3], biases in the correlation estimates are avoided. Assuming that the number of data points (N) is large, replace $\hat{r}(\tau)$ with $r(\tau)$ to obtain

$$E\{\tilde{\phi}(\tau)\} = -\rho(\tau)/r(\tau). \quad (20)$$

Thus, the estimator is asymptotically unbiased only if $\rho(\tau) = 0$.

The LS estimator ($\hat{\phi}_{LS}$) is obtained from Equation (18) by letting $\tau = 0$; $\hat{\phi}_{LS} = \hat{r}(1)/\hat{r}(0)$. Even with white measurement noise $\rho(0) \neq 0$; $\hat{\phi}_{LS}$ is seen to be biased. However, with white measurement noise $\rho(1) = 0$, and for other measurement noises, ρ can be either zero or approximately zero if τ is sufficiently larger.

The SLS estimator ($\hat{\phi}_{SLS}$) is obtained by choosing a value of τ which "shifts" the estimator argument to a value such that $\rho(\tau) = 0$; this procedure ensures an unbiased parameter estimator.

Calculating the parameter estimate from correlation estimates with larger arguments causes some reduction in accuracy. From Equations (19) and (20)

$$\text{var}[\tilde{\phi}(\tau)] = E\{\tilde{r}^2(\tau+1) - 2\phi\tilde{r}(\tau+1)\tilde{r}(\tau) + \phi^2\tilde{r}^2(\tau)\}/r^2(\tau). \quad (21)$$

Equation (21) can be solved in closed form; unfortunately, the resulting expression is complex and does not provide much insight into how $\text{var}[\tilde{\phi}(\tau)]$ varies with τ . Equation (21) can however be evaluated numerically. Figure 2 illustrates the growth of $\text{var}[\tilde{\phi}(\tau)]$ as a function of τ . As the plot indicates, the best results are obtained when no measurement noise is present and $\tau = 0$ can be used. If white measurement noise is present τ should equal one, to eliminate the bias, and $\text{var}[\tilde{\phi}]$ increases slightly. If the measurement can be represented by a moving average model, then a larger τ must be employed and the accuracy continues to decrease. However, $\text{var}[\tilde{\phi}(4)]$ is only about twice $\text{var}[\tilde{\phi}(0)]$; thus, the bias caused by measurement noise can often be eliminated with an acceptable increase in variance. The shift, τ , should not be made large enough to cause $\hat{\phi}_{SLS}$ to depend upon correlation values which are approaching zero and are therefore estimated less accurately than other values.

Note: $\phi = 0.8$; 250 measurements are processed; $v(k) = 0$

INEL 2 0162

Figure 2. NOTE: $\phi = 0.8$; 250 measurements are processed; $v(k) = 0$.

Parameters estimate as a function of τ . The plot does not show the possible estimator bias which must be accounted for by employing a nonzero τ .

Innovation Representation and Estimating Noise Variance

For a system driven with both process and measurement noise it is not usually possible to estimate the covariance of both noise sources from measurement data [10]. However, by using an innovations representation (from Kalman filtering) the covariance of the innovations sequence can be estimated. This estimator fits in nicely with the SLS parameter estimator.

Consider a system and measurement described by the Kalman filter quantities

$$\hat{x}_k(-) = \phi \hat{x}_{k-1}(-) + \phi K v_{k-1} \quad (23a)$$

$$z_k = \hat{x}_k(-) + v_k \quad (23b)$$

where K and v denote Kalman gain and innovation respectively. From Equation (23b)

$$E\{z_k z_{k+\tau}\} = E\{\hat{x}_k(-) z_{k+\tau} + v_k z_{k+\tau}\} \quad (24a)$$

From Equations (23a) and (23b) it can be shown that

$$z_{k+\tau} = \phi^\tau x_k + \phi^\tau K v_k + \phi^{\tau-1} K v_{k+1} + \dots + \phi K v_{k+\tau-1} + v_{k+\tau}.$$

Using the above equation,

$$E\{\hat{x}_k(-) z_{k+\tau}\} = \phi^\tau E\{x_k^2(-)\}, \tau \geq 0. \quad (24b)$$

Using equation (23b)

$$E\{v_k z_{k+\tau}\} = \begin{cases} E\{v_k^2\} & , \tau = 0 \\ 0 & , \text{otherwise.} \end{cases} \quad (24c)$$

Combining Equations (24a), (24b), and (24c) yields

$$E\{z_k z_{k+\tau}\} = \begin{cases} \phi^\tau E\{\hat{x}_k^2(-)\} + E\{v_k^2\} & , \tau = 0 \\ \phi^\tau E\{\hat{x}_k^2(-)\} & , \tau > 0 \end{cases} \quad (24d)$$

Figure 3 illustrates the correlation function for the example problem and indicates the one-standard-deviation estimation error limits. In this case the errors tend to remain constant with increasing τ which explains why the parameter estimation errors increase with τ . The errors are calculated from the expression

$$E\{\tilde{r}(s)\tilde{r}(t)\} = \frac{\text{var}^2[w]}{N[1-\phi^2]^2} \sum_{\delta=-N}^N [\phi^{|\delta+s-t|} \phi^{|\delta|} + \phi^{|\delta+s|} \phi^{|\delta-t|}] \quad (22)$$

(see Equation 65).

It would be possible to make several different parameter estimates from different pairs of adjacent correlation function values. These estimates could then be averaged to form a new estimate. However, the validity of this approach is not clear to the author. The problem is that estimation errors for various points on the correlation function are crosscorrelated and therefore the various parameter estimates will have correlated errors. It is not clear that using more than two points on the correlation function will yield a more accurate parameter estimate than just using two points.

Figure 3. Correlation function and one-standard-deviation error limits shown as solid line and dashed lines respectively. Function is for the same example considered in Figure 3.

The system parameter (ϕ) and the innovations covariance can be deduced by application of Equation (24d). To find the parameter use (for example)

$$r(1) = \phi^1 E\{\hat{x}_k^2(-)\}$$

$$r(2) = \phi^2 E\{\hat{x}_k^2(-)\}$$

and

$$\phi = \frac{r(2)}{r(1)} \quad (25a)$$

To find the innovations covariance use

$$r(0) = E\{\hat{x}_k^2(-)\} + E\{v_k^2\}$$

$$r(1) = \phi E\{\hat{x}_k^2(-)\} \quad (25b)$$

or

$$E\{\hat{x}_k^2(-)\} = \frac{1}{\phi} r(1)$$

and

$$E\{v_k^2\} = r(0) - \frac{1}{\phi} r(1). \quad (25c)$$

With the innovations representation the Kalman gain can also be determined. From Equation (23a) it can readily be shown that

$$E\{x_k^2(-)\} = \frac{\phi^2 k}{1 - \phi^2} E\{v_k^2\} \quad (26)$$

which can be solved for k if ϕ , $E\{\hat{x}_k^2(-)\}$, and $E\{v_k^2\}$ are determined from Equation (25a), (25b), and (25c) respectively.

Discussion

The SLS parameter estimator provides an effective but simple solution to the bias problem encountered with LS estimates. Estimator performance decreases somewhat as τ becomes larger and this decrease might be considered the penalty for using noisy measurements as opposed to noiseless measurements.

MULTIVARIATE PROBLEMS

The previous chapter considers parameter estimation in the scalar case. Since a scalar system is assumed, there is no need to estimate system order and hence the problem is not a full identification problem. For an identification problem the investigator must estimate (a) system order, or structure for single-input single-output or multi-input multi-output systems respectively, and (b) system parameters. Choosing a canonical form is the essential difficulty added by going from the scalar problem to the general multivariate problem [1, 15-20].

Canonical forms reduce the number of unknown system parameters to a number that can be uniquely estimated from input-output data. Many candidate forms exist. However, the parameter estimation accuracy, associated with a form, depends upon system dynamics. No single form can be expected to provide accurate parameter estimates for an arbitrary system. This report develops a method for choosing a form from a specified class of canonical forms. In addition, the SLS parameter estimator is extended for application to multivariate problems.

Problem Formulation

Consider a system and measurement described by Equations (27) and (28).

$$x(k+1) = \Phi x(k) + G u(k) \quad (27)$$

$$z(k) = H x(k) \quad (28)$$

where

x = $n \times 1$ state vector

Φ = $n \times n$ state transition matrix

G = $n \times q$ distribution matrix

u = $q \times 1$ control vector

z = $m \times 1$ measurement vector

H = $m \times n$ measurement matrix.

The system is stable and the parameter set (ϕ, G, H) must be estimated from a set of input-output measurements.

Canonical Forms

The observability matrix, O , is

$$O = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_m \\ h_1 \phi \\ h_2 \phi \\ \vdots \\ h_1 \phi^n \\ \vdots \\ h_m \phi^n \end{bmatrix} \quad (29)$$

where h_i denotes the i th row of H .

A canonical form can be specified by selecting a set of n linearly independent vectors from O and using those vectors to form a transformation matrix, T . T rotates the coordinate system so that the system and measurement equations are expressed in the associated canonical form. The essential restriction on our class of canonical forms is that no vector $h_i \phi^j$ is selected, as a member of the set of independent vectors, unless h_i times all lower powers of ϕ were previously selected [18]. Many different selection rules satisfy the restriction and thus many different canonical forms are allowed. A set of indices $v_1, v_2, \dots, v_m$ specifies the number of selected vectors containing $h_1, h_2, \dots, h_m$, respectively.

The selected independent vectors are arranged to form T as

$$T = \begin{bmatrix} h_1 \\ h_1 \phi \\ h_1 \phi^2 \\ \vdots \\ h_1 \phi^{v_1-1} \\ h_2 \\ \vdots \\ h_2 \phi^{v_2-1} \\ \vdots \\ h_m \phi^{v_m-1} \end{bmatrix} \quad (30)$$

For any given system, the selection rule determines the indices $v_1, v_2, \dots, v_m$ which are synonymous with system structure. Using T to rotate the coordinate system places the modified Φ matrix in a blocked form as defined by Equation (31) and (32) while the H matrix takes the form specified by Equation (33). Matrix G takes no special form.

$$\Phi_{ii} = \begin{bmatrix} 0 & & & \\ 0 & & & \\ \vdots & & I & \\ \phi_{ii,1} & \phi_{ii,2} & \dots & \phi_{ii,v_i} \end{bmatrix}_{v_i \times v_i} \quad (31)$$

$$\Phi_{ij} = \begin{bmatrix} 0 & & \dots & & 0 \\ \vdots & & & & \\ 0 & & \dots & & 0 \\ \phi_{ij,1} & \phi_{ij,2} & \dots & \phi_{ij,v_j} & 0 & \dots & 0 \end{bmatrix}_{v_i \times v_j} \quad (32)$$

$$H = \begin{bmatrix} 1 & 0 & \dots & & & & 0 \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ & \vdots & & & & & \vdots \\ 0 & \dots & & & 0 & 1 & 0 & \dots & 0 \end{bmatrix}_{m \times n} \quad (33)$$

$\overset{\uparrow}{1}$ $\overset{\uparrow}{v_1+1}$ $(\overset{\uparrow}{v_1+1} \dots + \overset{\uparrow}{v_m+1})$

Column Column Column

The number of nonzero coefficients (v_{ij}) in ϕ_{ij} is a function of the section rule. The relative values of the indices generally determine v_{ij} ; however, in certain cases v_{ij} can be zero.

Identically zero parameters in the significant--coefficient containing-- rows of Φ can be located by considering the selection rule and the transformation matrix. Consider a system in canonical form so that $T = I$. Inspection of the equation $T = I$ establishes H and all the rows of Φ which contain only zeros and a one. Consider an arbitrary selection procedure. Without loss of generality the vectors are selected in the following order

$$\dots, \overline{h_1}^{\nu_1}, \dots, \overline{h_2}^{\nu_2}, \dots, \overline{h_m}^{\nu_m}$$

where the overlined vectors $h_j \Phi^{v_j}$ are the first dependent vectors encountered that involve h_j . From definition of the selection rule $h_j \Phi^{v_j}$ is only dependent upon antecedent vectors. This fact specifies the identically zero coefficients in significant rows of Φ . For example, consider independent vectors chosen in the order

$$h_2, h_2 \phi, h_1, \overline{h_1 \phi^1}, h_2 \phi^2, h_2 \phi^3, \overline{h_2 \phi^4}.$$

Since overlined vectors are dependent on their antecedents, $v_1 = 1$ and $v_2 = 4$. For this example the canonical Φ and H are

$$\Phi = \begin{bmatrix} \phi_{11,1} & \phi_{12,1} & \phi_{12,2} & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ \phi_{21,1} & \phi_{22,1} & \phi_{22,2} & \phi_{22,3} & \phi_{22,4} \end{bmatrix}$$

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

With the chosen selection rule $\phi_{12,3} = \phi_{12,4} = 0$. Other selection rules would, generally, force other parameters to be identically zero. By including the identically zero coefficients in the parameter count, there are always $n \times m$ coefficients. However, the number of identically zero coefficients varies with the selection rule and the relative values of the indices.

Dependent vectors $h_i \phi^{v_i}$ are linear combinations of the independent vectors selected to form the T matrix and the coefficients are in the i th significant row of the canonical Φ matrix, i.e.,

$$\begin{aligned} h_i \phi^{v_i} &= \phi_{i1,1} h_1 + \phi_{i1,2} h_1 \phi + \dots + \phi_{i1, v_{i1}} h_1 \phi^{v_1-1} \\ &\quad + \phi_{i2,1} h_2 + \phi_{i2,2} h_2 \phi + \dots + \phi_{i2, v_{i2}} h_2 \phi^{v_2-1} \\ &\quad \vdots \\ &\quad + \phi_{im,1} h_m + \phi_{im,2} h_m \phi + \dots + \phi_{im, v_{im}} h_m \phi^{v_m-1}. \end{aligned} \quad (34)$$

where

$$v_{ii} \equiv v_i.$$

From the expressions for coordinate transformation the dependency relationship in Equation (34) can be shown valid even if the system is not in a canonical form. Let $\underline{\Phi}$ and $\underline{H}$ denote the state transition and measurement matrices in an arbitrary coordinate system while Φ and H denote the respective matrices in canonical form. With these definitions, $\Phi = T \underline{\Phi} T^{-1}$, and $H = \underline{H} T^{-1}$ it can be shown that

$$\begin{bmatrix}
 h_1 \\
 h_1 \Phi \\
 h_1 \Phi^2 \\
 \vdots \\
 h_1 \Phi^{v_1-1} \\
 h_2 \\
 \vdots \\
 h_2 \Phi^{v_2-1} \\
 \vdots \\
 h_m \Phi^{v_m-1} \\
 \hline
 h_1 \Phi^{v_1} \\
 h_2 \Phi^{v_2} \\
 \vdots \\
 h_m \Phi^{v_m}
 \end{bmatrix}
 =
 \begin{bmatrix}
 \underline{h}_1 \\
 \underline{h}_1 \underline{\Phi} \\
 \underline{h}_1 \underline{\Phi}^2 \\
 \vdots \\
 \underline{h}_1 \underline{\Phi}^{v_1-1} \\
 \underline{h}_2 \\
 \vdots \\
 \underline{h}_2 \underline{\Phi}^{v_2-1} \\
 \vdots \\
 \underline{h}_m \underline{\Phi}^{v_m-1} \\
 \hline
 \underline{h}_1 \underline{\Phi}^{v_1} \\
 \underline{h}_2 \underline{\Phi}^{v_2} \\
 \vdots \\
 \underline{h}_m \underline{\Phi}^{v_m}
 \end{bmatrix}
 T^{-1}
 \quad (35)$$

Inspection of Equation (35) reveals the validity of Equation (34) for an arbitrary coordinate system. Thus the indices and parameters in Equation (34) are said to form a complete set of system invariants [19].

Covariance Function

In this section we use results from the canonical forms section to derive a difference equation for the covariance function. The equation is analogous to Equation (16) for a scalar problem.

By manipulating Equation (27) and (28),

$$\begin{aligned} z_i(k) &= h_i x(k) \\ z_i(k+1) &= h_i \Phi x(k) + h_i G u(k) \\ &\vdots \\ z_i(k+\sigma) &= h_i \Phi^\sigma x(k) + h_i \Phi^{\sigma-1} G u(k) + h_i \Phi^{\sigma-2} G u(k+1) + \dots + h_i G u(k+\sigma-1). \end{aligned} \tag{36}$$

For the class of canonical forms

$$\begin{aligned} h_i G &= g_i(1) \\ h_i \Phi G &= g_i(2) \\ &\vdots \\ h_i \Phi^{v_i-1} G &= g_i(v_i) \end{aligned} \tag{37}$$

where $g_i(j)$ denotes a row vector from G .

Using the first v_i-1 equations from (36) for $i = 1, 2, \dots, m$, plus Equation (37), and collecting the results yields

$$\begin{bmatrix} z_1(k) \\ z_1(k+1) \\ z_1(k+2) \\ \vdots \\ z_1(k+v_1-1) \\ \hline \vdots \\ \hline z_M(k) \\ \vdots \\ z_M(k+v_M-1) \\ \hline \vdots \\ \hline z_m(k) \\ \vdots \\ z_m(k+v_m-1) \end{bmatrix} = \begin{bmatrix} h_1 \\ h_1 \phi \\ h_1 \phi^2 \\ \vdots \\ h_1 \phi^{v_1-1} \\ \hline \vdots \\ \hline h_M \\ \vdots \\ h_M \phi^{v_M-1} \\ \hline \vdots \\ \hline h_m \\ \vdots \\ h_m \phi^{v_m-1} \end{bmatrix} x(k) + \begin{bmatrix} 0 & 0 & \cdots & 0 \\ g_1(1) & 0 & \cdots & 0 \\ g_1(2) & g_1(1) & 0 & \cdots & 0 \\ \vdots & & & & 0 \\ g_1(v_1-1) & g_1(v_1-2) & \cdots & g_1(1) & 0 & \cdots & 0 \\ \hline \vdots \\ \hline 0 & 0 & \cdots & 0 \\ g_M(1) & 0 & \cdots & 0 \\ \vdots & & & \\ g_M(v_M-1) & g_M(v_M-2) & \cdots & g_M(1) \\ \hline \vdots \\ \hline 0 & \cdots & 0 \\ \vdots & & \\ g_m(v_m-1) & g_m(v_m-2) & \cdots & g_m(1) & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} u(k) \\ u(k+1) \\ \vdots \\ u(k+v_M-2) \end{bmatrix} \quad (38)$$

where v_M denotes $\max (v_i)$.

From the properties of the class of canonical forms

$$\begin{bmatrix} h_1 \\ h_1 \phi \\ h_1 \phi^2 \\ \vdots \\ h_1 \phi^{v_1-1} \\ \hline \vdots \\ \hline h_M \\ \vdots \\ h_M \phi^{v_M-1} \\ \hline \vdots \\ \hline h_m \\ \vdots \\ h_m \phi^{v_m-1} \end{bmatrix} = I.$$

Therefore by using the obvious definitions Equation (38) can be rewritten as

$$\bar{z}(k) = I x(k) + \underline{x} \bar{u}(k)$$

or

$$x(k) = \bar{z}(k) - \underline{x} \bar{u}(k). \quad (39)$$

Using the v_{ith} equation from (36) for $i = 1, 2, \dots, m$, plus Equation (37), and collecting results yields

$$\begin{bmatrix} z_1(k+v_1) \\ \vdots \\ z_M(k+v_M) \\ \vdots \\ z_m(k+v_m) \end{bmatrix} = \begin{bmatrix} h_1 \phi^{v_1} \\ h_M \phi^{v_M} \\ h_m \phi^{v_m} \end{bmatrix} x(k) + \begin{bmatrix} g_1(v_1) & g_1(v_1-1) & \dots & g_1(1) & 0 & \dots & 0 \\ g_M(v_M) & \dots & & & & g_M(1) & \\ g_m(v_m) & \dots & & g_m(1) & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} u(k) \\ u(k+1) \\ \vdots \\ u(k+v_M-1) \end{bmatrix} \quad (40)$$

From the properties of the class of canonical forms

$$h_i \phi^{v_i} = \bar{\phi}_i \quad (41)$$

where $\bar{\phi}_i$ is the i th significant row of the ϕ matrix.

By using $x(k)$ from Equation (39), Equation (40) can be written as

$$\begin{bmatrix} z_1(k+v_1) \\ \vdots \\ z_M(k+v_M) \\ \vdots \\ z_m(k+v_m) \end{bmatrix} = \begin{bmatrix} \bar{\phi}_1 \\ \vdots \\ \bar{\phi}_M \\ \vdots \\ \bar{\phi}_m \end{bmatrix} \begin{bmatrix} z_1(k) \\ z_1(k+1) \\ \vdots \\ z_1(k+v_1-1) \\ \text{-----} \\ \vdots \\ \text{-----} \\ z_M(k) \\ \vdots \\ z_M(k+v_M-1) \\ \text{-----} \\ \vdots \\ \text{-----} \\ z_m(k) \\ \vdots \\ z_m(k+v_m-1) \end{bmatrix} + B \begin{bmatrix} u(k) \\ u(k+1) \\ \vdots \\ u(k+v_M-1) \end{bmatrix} \quad (42)$$

where

$$B = \begin{bmatrix} g_1(v_1) & g_1(v_1-1) & \dots & g_1(1) & 0 & 0 \\ g_M(v_M) & \dots & & & & g_M(1) \\ g_m(v_m) & \dots & g_m(1) & 0 & \dots & 0 \end{bmatrix}$$

$$\begin{bmatrix} \bar{\phi}_1 \\ \vdots \\ \bar{\phi}_M \\ \vdots \\ \bar{\phi}_m \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & \dots & 0 \\ g_1(1) & 0 & \dots & 0 \\ g_1(2) & g_1(1) & 0 & \dots & 0 \\ \vdots & & & & \\ g_1(v_1-1) & g_1(v_1-2) & \dots & g_1(1) & 0 & 0 \\ \hline \vdots & & & & & \\ \hline 0 & 0 & \dots & & & \\ g_M(1) & 0 & & & & \\ \vdots & & & & & \\ g_M(v_M-1) & g_M(v_M-2) & \dots & & g_M(1) & 0 \\ \hline \vdots & & & & & \\ \hline 0 & \dots & & & & \\ \vdots & & & & & \\ g_m(v_m-1) & g_m(v_m-2) & \dots & g_m(1) & 0 & \end{bmatrix}$$

Note:
↑

block column of zeros.

(43)

and G is in canonical form.

By using Equation (43), G can be determined from $\hat{\Phi}$ and $\hat{B}$.

Equation (42) is a difference equation relating measurements to past measurements and inputs. From this equation we can derive a difference equation for the correlation function. Consider the i th line of Equation (42) and change the argument from k to $k+\tau$,

$$z_i[(k+\tau)+v_i] = \bar{\phi}_i \begin{bmatrix} z_1(k+\tau) \\ z_1(k+\tau+1) \\ \vdots \\ z_1[(k+\tau)+v_1-1] \\ \hline \vdots \\ \hline z_M(k+\tau) \\ \vdots \\ z_M[(k+\tau)+v_M-1] \\ \hline \vdots \\ \hline z_m(k+\tau) \\ \vdots \\ z_m[(k+\tau)+v_m-1] \end{bmatrix} + b_i \begin{bmatrix} u_1(k+\tau) \\ u_2(k+\tau) \\ \vdots \\ u_q(k+\tau) \\ \hline u_1[(k+\tau)+1] \\ u_2[(k+\tau)+1] \\ \vdots \\ u_q[(k+\tau)+1] \\ \hline \vdots \\ \hline u_1[(k+\tau)+v_M-1] \\ u_2[(k+\tau)+v_M-1] \\ \vdots \\ u_q[(k+\tau)+v_M-1] \end{bmatrix} \quad (44)$$

By rearranging (notice that the control inputs are shuffled into the same order as the measurements) we obtain

$$\begin{aligned}
 z_i[(k+\tau)+v_i] &= \begin{bmatrix} z_1(k+\tau) \\ z_1(k+\tau+1) \\ \vdots \\ z_1[(k+\tau)+v_1-1] \\ \hline z_2(k+\tau) \\ \vdots \\ z_2[(k+\tau)+v_2-1] \\ \hline \vdots \\ \hline z_m[(k+\tau)+v_m-1] \\ \hline \hline u_1(k+\tau) \\ u_1[(k+\tau)+1] \\ \vdots \\ u_1[(k+\tau)+v_M-1] \\ \hline u_2(k+\tau) \\ u_2[(k+\tau)+1] \\ \vdots \\ u_2[(k+\tau)+v_M-1] \\ \hline \vdots \\ \hline u_q(k+\tau) \\ u_q[(k+\tau)+1] \\ \vdots \\ u_q[(k+\tau)+v_M-1] \end{bmatrix}^T \\
 &= \begin{bmatrix} \theta_{i1,1} \\ \theta_{i1,2} \\ \vdots \\ \theta_{i1,v_1} \\ \hline \theta_{i2,1} \\ \theta_{i2,2} \\ \vdots \\ \theta_{i2,v_2} \\ \hline \vdots \\ \hline \theta_{im,1} \\ \theta_{im,2} \\ \vdots \\ \theta_{im,v_m} \\ \hline \hline \theta_{i(m+1),1} \\ \theta_{i(m+1),2} \\ \vdots \\ \theta_{i(m+1),v_M} \\ \hline \hline \theta_{i(m+2),1} \\ \theta_{i(m+2),2} \\ \vdots \\ \theta_{i(m+2),v_M} \\ \hline \vdots \\ \hline \theta_{i(m+q),1} \\ \theta_{i(m+q),2} \\ \vdots \\ \theta_{i(m+q),v_M} \end{bmatrix}
 \end{aligned}
 \tag{45}$$

Where θ denotes a parameter from either Φ or G and the choice is specified by subscripts.

To simplify the notation consider both the measurements and the inputs as data (d) and rewrite Equation (45) as

$$z_i[(k+\tau)+v_i] = \begin{bmatrix} d_1(k+\tau) \\ d_1[(k+\tau)+1] \\ \vdots \\ d_1[(k+\tau)+v_1-1] \\ \hline d_2(k+\tau) \\ d_2[(k+\tau)+1] \\ \vdots \\ d_2[(k+\tau)+v_2-1] \\ \hline \vdots \\ \hline d_m(k+\tau) \\ d_m[(k+\tau)+1] \\ \vdots \\ d_m[(k+\tau)+v_m-1] \\ \hline d_{m+1}(k+\tau) \\ d_{m+1}[(k+\tau)+1] \\ \vdots \\ d_{m+1}[(k+\tau)+v_M-1] \\ \hline d_{m+2}(k+\tau) \\ d_{m+2}[(k+\tau)+1] \\ \vdots \\ d_{m+2}[(k+\tau)+v_M-1] \\ \hline \vdots \\ \hline d_{m+q}(k+\tau) \\ d_{m+q}[(k+\tau)+1] \\ \vdots \\ d_{m+q}[(k+\tau)+v_M-1] \end{bmatrix}^T \begin{bmatrix} \theta_{i1,1} \\ \theta_{i1,2} \\ \vdots \\ \theta_{i1,v_1} \\ \hline \theta_{i2,1} \\ \theta_{i2,2} \\ \vdots \\ \theta_{i2,v_2} \\ \hline \vdots \\ \hline \theta_{im,1} \\ \theta_{im,2} \\ \vdots \\ \theta_{im,v_m} \\ \hline \theta_{i(m+1),1} \\ \theta_{i(m+1),2} \\ \vdots \\ \theta_{i(m+1),v_M} \\ \hline \theta_{i(m+2),1} \\ \theta_{i(m+2),2} \\ \vdots \\ \theta_{i(m+2),v_M} \\ \hline \vdots \\ \hline \theta_{i(m+q),1} \\ \theta_{i(m+q),2} \\ \vdots \\ \theta_{i(m+q),v_M} \end{bmatrix} \quad (46)$$

Premultiply each side of Equation (46) by $d_1(k)$ and take expected values to obtain,

$$r_{1i}(\tau+v_i) = \begin{bmatrix} r_{11}(\tau) \\ r_{11}(\tau+1) \\ \vdots \\ r_{11}(\tau+v_1-1) \\ \hline r_{12}(\tau) \\ r_{12}(\tau+1) \\ \vdots \\ r_{12}(\tau+v_2-1) \\ \hline \vdots \\ \hline r_{1m}(\tau) \\ r_{1m}(\tau+1) \\ \vdots \\ r_{1m}(\tau+v_m-1) \\ \hline r_{1,m+1}(\tau) \\ r_{1,m+1}(\tau+1) \\ \vdots \\ r_{1,m+1}(\tau+v_M-1) \\ \hline \vdots \\ \hline r_{1,m+q}(\tau) \\ r_{1,m+q}(\tau+1) \\ \vdots \\ r_{1,m+q}(\tau+v_M-1) \end{bmatrix}^T \quad \theta \quad (47)$$

Equation (47) can be written with τ replaced by $\tau+1, \tau+2, \dots$. In addition, similar equations can be obtained by premultiplying Equation (46) by d_i where $i = 1, 2, \dots, m+q$. Multiply by each d_i and write each equation with the variable $\tau, \tau+1, \tau+2, \dots$. A total of $N+mq$ independent equations can be written. The equations we choose to work with are collected in matrix Equation (48), see next page.

By condensing the notation of Equation (48)

$$R_i = R \theta_i \quad (49)$$

where R has a blocked structure defined by Equation (50).

$$R_{ij}(\tau) = \begin{bmatrix} r_{ij}(\tau) & r_{ij}(\tau+1) & \dots & r_{ij}(\tau+v_j-1) \\ r_{ij}(\tau+1) & r_{ij}(\tau+2) & \dots & \\ r_{ij}(\tau+2) & & & \\ \vdots & & & \\ r_{ij}(\tau+v_i-1) & \dots & & r_{ij}(\tau+v_i+v_j-2) \end{bmatrix}_{v_i \times v_j} \quad (50)$$

For $1 \leq m$ v_1 is determined by the system structure. For $l > m$, $v_l = v_M$ where v_M is the largest v_i ($i=1, 2, \dots, m$).

PARAMETER ESTIMATION

From the previous section (Equation (49)) the system parameters are related to the covariance function according to

$$R_i(\tau) = R(\tau) \theta_i.$$

To obtain a parameter estimator we let $\hat{R}_i = R_i + \tilde{R}_i$, $\hat{R} = R + \tilde{R}$, and $\hat{\theta}_i = \theta_i + \tilde{\theta}_i$. With these definitions,

$$\hat{R}_i - \tilde{R}_i = \hat{R} \theta_i - \tilde{R} \theta_i$$

$$\hat{R}_i = \hat{R} \theta_i + \tilde{R}_i - \tilde{R} \theta_i. \quad (50)$$

Since we cannot calculate the errors let

$$\boxed{\hat{R}_i = \hat{R} \hat{\theta}_i} \quad (51)$$

and then form Equation (50) and (51)

$$\hat{R}_i = \hat{R} \hat{\theta}_i - \hat{R} \tilde{\theta}_i + \tilde{R}_i - \tilde{R} \tilde{\theta}_i$$

and

$$\tilde{\theta}_i = \hat{R}^{-1} [\tilde{R}_i - \tilde{R} \theta_i]. \quad (52)$$

By letting

$$\epsilon_i \equiv \tilde{R}_i - \tilde{R} \theta_i, \quad (53)$$

$$\tilde{\theta}_i = \hat{R}^{-1} \epsilon_i. \quad (54)$$

Parameter Estimation Accuracy

Bias

By assuming a large number of data points (N) so that $\hat{R} \approx R$, one obtains

$$E\{\tilde{\theta}_i\} = R^{-1} E\{\tilde{R}_i - \tilde{R} \theta_i\}$$

Since the covariance estimates can be shown to be unbiased,

$$E\{\tilde{\theta}_i\} = 0. \quad (55)$$

Variance

$$E\{\tilde{\theta}_i \tilde{\theta}_i^T\} = R^{-1} E\{\epsilon_i \epsilon_i^T\} R^{-T} \quad (56)$$

and the problem is to evaluate the elements of $E\{\epsilon_i \epsilon_i^T\}$. Let

$$\epsilon_i = \begin{bmatrix} e_{i1,1} \\ e_{i1,2} \\ \vdots \\ e_{i1,v_1} \\ \hline \vdots \\ \hline e_{im,1} \\ e_{im,2} \\ \vdots \\ e_{im,v_m} \\ \hline e_{i(m+1),1} \\ e_{i1(m+1),2} \\ \vdots \\ e_{i1(m+1),v_M} \\ \hline \vdots \\ \hline \vdots \\ e_{im+q,v_M} \end{bmatrix} \quad (57)$$

With this definition Equation (53) can be expanded as

$$\begin{bmatrix} e_{i1,1} \\ e_{i1,2} \\ \vdots \\ e_{i1,v_1} \\ \hline e_{i2,1} \\ e_{i2,2} \\ \vdots \\ e_{i2,v_2} \\ \hline \vdots \\ \hline e_{im,v_m} \\ \hline e_{im+1,1} \\ e_{im+1,2} \\ \vdots \\ e_{im+1,v_1} \\ \hline \vdots \\ \hline e_{im+q,1} \\ e_{im+1,2} \\ \vdots \\ e_{im+q,v_M} \end{bmatrix} = \begin{bmatrix} \tilde{r}_{1i}(v_i+\tau) \\ \tilde{r}_{1i}(v_i+\tau+1) \\ \vdots \\ \tilde{r}_{1i}(v_i+\tau+v_1-1) \\ \hline \tilde{r}_{2i}(v_i+\tau) \\ \tilde{r}_{2i}(v_i+\tau+1) \\ \vdots \\ \tilde{r}_{2i}(v_i+\tau+v_2-1) \\ \hline \vdots \\ \hline \tilde{r}_{mi}(v_i+\tau) \\ \vdots \\ \tilde{r}_{mi}(v_i+\tau+v_m-1) \\ \hline \tilde{r}_{m+1,i}(v_i+\tau) \\ \tilde{r}_{m+1,i}(v_i+\tau+1) \\ \vdots \\ \tilde{r}_{m+1,i}(v_i+\tau+v_m-1) \\ \hline \tilde{r}_{m+q,i}(v_i+\tau) \\ \tilde{r}_{m+q,i}(v_i+\tau+1) \\ \vdots \\ \tilde{r}_{m+q,i}(v_i+\tau+v_M-1) \end{bmatrix} \begin{bmatrix} \tilde{r}_{11}(\tau) & \tilde{r}_{11}(\tau+1) & \dots & \tilde{r}_{11}(\tau+v_1-1) & \tilde{r}_{12}(\tau) & \tilde{r}_{12}(\tau+1) & \dots & \tilde{r}_{12}(\tau+v_2-1) & \dots \\ \tilde{r}_{11}(\tau+1) & \tilde{r}_{11}(\tau+2) & \dots & & & & & & \\ \tilde{r}_{11}(\tau+2) & & & & & & & & \\ \vdots & & & & & & & & \\ \hline \tilde{r}_{21}(\tau) & \tilde{r}_{21}(\tau+1) & & & \tilde{r}_{22}(\tau) & \tilde{r}_{22}(\tau+2) & \dots & & \\ \tilde{r}_{21}(\tau+1) & & & & \tilde{r}_{22}(\tau+1) & & & & \\ \vdots & & & & \vdots & & & & \\ \hline & & & & & & & & \end{bmatrix} \begin{bmatrix} \theta_{i1,1} \\ \theta_{i1,2} \\ \vdots \\ \theta_{i1,v_1} \\ \hline \theta_{i2,1} \\ \theta_{i2,2} \\ \vdots \\ \theta_{i2,v_2} \\ \hline \vdots \\ \hline \theta_{im,1} \\ \theta_{im,2} \\ \vdots \\ \theta_{im,v_m} \\ \hline \theta_{i(m+1),1} \\ \theta_{i(m+1),2} \\ \vdots \\ \theta_{i(m+1),v_M} \\ \hline \vdots \\ \hline \theta_{im+q,1} \\ \theta_{im+q,2} \\ \vdots \\ \theta_{im+q,v_M} \end{bmatrix} \quad (58)$$

By inspection of Equation (58) an arbitrary element of $E\{\epsilon_i \epsilon_i^T\}$ can be written as

$$E\{e_{ij,k} e_{iy,z}\} = E\{\tilde{r}_{ji}(v_i + \tau + k - 1) \tilde{r}_{yi}(v_i + \tau + z - 1)\} \quad (59)$$

$$\begin{aligned} & - \sum_{s=1}^{NB} \sum_{t=1}^{v_s} \theta_{is,t} E\{\tilde{r}_{ys}(\tau + t + z - 2) \tilde{r}_{ji}(v_i + \tau + k - 1)\} \\ & - \sum_{w=1}^{NB} \sum_{x=1}^{v_w} \theta_{iw,x} E\{\tilde{r}_{jw}(\tau + x + k - 2) \tilde{r}_{yi}(v_i + \tau + z - 1)\} \\ & + \sum_{s=1}^{NB} \sum_{t=1}^{v_s} \sum_{w=1}^{NB} \sum_{x=1}^{v_w} \theta_{is,t} \theta_{iw,x} E\{\tilde{r}_{ys}(\tau + t + z - 2) \tilde{r}_{jw}(\tau + x + k - 2)\} \end{aligned}$$

where $NB = m + 0.5$.

Equation (59) cannot yet be used to calculate elements of $E\{\epsilon_i \epsilon_i^T\}$ because (a) the parameter values are not known and (b) the expected value terms are not known. The first problem is solved by substituting estimated values for the true values. To solve the second problem we must develop an expression for $E\{\tilde{r}_{ij}(k) \tilde{r}_{st}(1)\}$.

By definition,

$$\begin{aligned} E\{\tilde{r}_{ij}(k) \tilde{r}_{st}(1)\} &= E\{(\hat{r}_{ij}(k) - r_{ij}(k))(\hat{r}_{st}(1) - r_{st}(1))\} \\ &= E\{\hat{r}_{ij}(k) \hat{r}_{st}(1)\} + E\{r_{ij}(k) r_{st}(1)\} \\ &\quad - E\{\hat{r}_{ij}(k) r_{st}(1)\} - E\{r_{ij}(k) \hat{r}_{st}(1)\} \end{aligned}$$

and since the estimates are unbiased

$$E\{\tilde{r}_{ij}(k) \tilde{r}_{st}(1)\} = E\{\hat{r}_{ij}(k) \hat{r}_{st}(1)\} - r_{ij}(k) r_{st}(1). \quad (60)$$

The first term on the right of Equation (60) can be evaluated by using the expression for $\hat{r}$,

$$\hat{r}_{ij}(k) = \frac{1}{N} \sum_{g=1}^N d_i(g) d_j(g+k) \quad (61)$$

$$\hat{r}_{st}(1) = \frac{1}{N} \sum_{h=1}^N d_s(h) d_t(h+1). \quad (62)$$

Using Equation (61) and (62) in (60) yields

$$\begin{aligned} E\{\tilde{r}_{ij}(k) \tilde{r}_{st}(1)\} &= \frac{1}{N^2} \sum_{g=1}^N \sum_{h=1}^N [E\{d_i(g) d_j(g+k) d_s(h) d_t(h+1)\} + \\ &\quad - r_{ij}(k) r_{st}(1)]. \end{aligned} \quad (63)$$

By using the normal assumption, $E\{d_i d_j d_s d_t\}$ can be expanded as

$$\begin{aligned} E\{d_i d_j d_s d_t\} &= E\{d_i(g) d_j(g+k)\} E\{d_s(h) d_t(h+1)\} \\ &\quad + E\{d_i(g) d_s(h)\} E\{d_j(g+k) d_t(h+1)\} \\ &\quad + E\{d_i(g) d_t(h+1)\} E\{d_j(g+k) d_s(h)\}. \end{aligned} \quad (64)$$

After some manipulation of Equation (64) and by assuming a large N , Equation (63) can be written as

$$E\{\tilde{r}_{i,j}(k) \tilde{r}_{s,t}(1)\} = \frac{1}{N} \sum_{\rho=-N}^N [r_{i,s}(\rho) r_{j,t}(\rho+1-k) + r_{i,t}(\rho+1) r_{j,s}(\rho-k)]. \quad (65)$$

Equation (65) becomes operational if we replace the exact covariance values with estimates obtained by using Equation (61) and experimental data.

The calculations needed to determine parameter estimation accuracy are:

(a) Equation (65): $E\{\tilde{r}_{i,j}(k) \tilde{r}_{s,t}(1)\}$

(b) Equation (59): $E\{e_{ij,k} e_{iy,z}\}$

(c) Collect results of applying Equation (59): $E\{\epsilon_i \epsilon_i^T\}$

(d) Equation 56): $E\{\tilde{\theta}_i \tilde{\theta}_i^T\}$.

Noise Considerations

Multivariate system and measurement equations that include both control inputs, process noise, and measurement noise are:

$$x(k+1) = \Phi X(k) + G u(k) + w(k) \quad (66)$$

$$z(k) = H X(k) + v(k). \quad (67)$$

By working with Equation (66) and (67) an equation similar to Equation (45) can be derived. The difference is that additional terms are involved to account for the measurement noise (process noise has no effect if $\tau \geq 0$). Fortunately, the measurement noise effects disappear if $\tau \geq 1$.

Therefore, the change (caused by noise) to Equation (45), can be eliminated by simply requiring $\tau \geq 1$.

If the noise sources were 2nd order moving average processes then their effects on the correlation function can be eliminated by setting $\tau \geq 3$.

Structure Estimation

Examination of Equation (48) provides an algorithm for estimating system structure, i.e., $v_1, v_2, \dots, v_m$. For convenience we rewrite Equation (48) using the compact notation;

$$R_i(\tau) = R(\tau) \theta_i.$$

The equation shows that the correlation vector, $R_i(\tau)$, depends upon a limited number of correlations with smaller arguments and located in $R(\tau)$. The system structure determines how many and which correlations are needed to form $R_i(\tau)$.

The details of how to select a canonical form can be varied. The essential question is how to find one such that $R(\tau)$ is strongly nonsingular. This will assure more accurate parameter estimates than would be likely with other forms. One reasonable way to select a structure is to start with $v_1 = v_2 = \dots = v_m = 1$ and increase each index by one and determine which index can be increased with the minimum reduction of R 's distance from singularity. Raise only this index by one. Search the resulting set of indices to determine which index can be increased with the smallest reduction of R 's distance to singularity. Continue increasing the indices in this way until no index can be increased without R becoming singular. The resulting set of indices define the system structure. The order in which indices are increased corresponds to the order of selecting independent vectors from the observability matrix to form T . Therefore, the order of selecting indices defines a canonical form.

Since many system structures can be quickly evaluated with the aid of an interactive computer program (without causing a larger computer burden), optimized procedures for selecting system structures are not required for modeling moderate sized systems.

RESULTS

Several systems were simulated on the INEL CYBER system and data from those simulations was used to test the identification methods [21]. The simulations are presented in this chapter as a series of examples with each example emphasizing one or more identification issues.

Example 1

Consider the system defined by Figure 4.

Figure 4. Scalar system with white process and measurement noise.

By processing a sequence of simulated measurements one can estimate the system structure (order for a univariate system) and the parameter ϕ .

Figure 5 illustrates an autocorrelation function estimate obtained from 1000 data points.

Table 1 shows results of structure estimation from the 1000 point data set. The structure was estimated with three different values of τ ; the matrix R is assumed singular when the ratio goes below one [14]. Because of the measurement noise, it should not be possible to correctly identify the structure with $\tau = 0$. However, larger τ values should provide proper structure estimation. Table 1 supports both hypotheses.

Figure 5. Estimated autocorrelation function.

Table 1. Structure Estimation

τ	STRUCTURE	RATIO
0	1	61.0
	2*	2.4
	3	0.16
1	1*	37.0
	2	0.58
	3	0.30
2	1*	2.9
	2	0.18
	3	0.24
*indicates estimated structure		

Table 2 shows results of estimating parameters from the 1000 point data set. Like structure, the parameter was estimated by using three different τ values; for each parameter estimate the correct structure (order = 1) was assumed.

Table 2. Parameter Estimates

τ	$\hat{\phi}$	S.D.	$\tilde{\phi}$
0	0.577	0.03	-0.223
1	0.758	0.04	-0.042
2	0.882	0.05	0.082
S.D. denotes standard deviation of $\hat{\phi}$ as calculated from Equation (56)			

Because the data comes from a simulation, it was possible to determine the actual $\tilde{\phi}$ and check whether $\tilde{\phi}$ is reasonably predicted by the calculated standard deviation. From Table 2 S.D. fails to predict the magnitude of $\tilde{\phi}$ only when $\tau = 0$; that failure occurs because of the measurement noise. The expected bias (with $\tau = 0$) can be calculated from Equation (20) as

$$E\{\tilde{\phi}(0)\} = - (0)/r(0) = -(0.8)(1.0)/3.78 = -0.212$$

which is very close to the actual bias shown in Table 2.

Example 2

Consider the system defined by Figure 6.

Figure 6. Scalar system with white noise as control signal.

This example is very similar to Example 1 except that the driving signal is known and an additional parameter must be estimated. This example emphasizes a surprising fact. While white process provides an excellent excitation for identification, a white noise control signal may or may not be a suitable excitation. The answer depends upon which τ value used in the identification algorithms. To clarify the situation, examine the correlation functions for the system; the functions are summarized in Figure 7.

$R(\tau)$ matrices for a few values of τ follow:

$$R(0) = \begin{bmatrix} r_{zz}(0) & r_{zu}(0) \\ r_{uz}(0) & r_{uu}(0) \end{bmatrix} = \begin{bmatrix} 6.25 & 0 \\ 0 & 1.0 \end{bmatrix} \quad (68a)$$

$$R(1) = \begin{bmatrix} r_{zz}(1) & r_{zu}(1) \\ r_{uz}(1) & r_{uu}(1) \end{bmatrix} = \begin{bmatrix} 5.0 & 0 \\ 1.5 & 0 \end{bmatrix} \quad (68b)$$

$$R(2) = \begin{bmatrix} r_{zz}(2) & r_{zu}(2) \\ r_{uz}(2) & r_{uu}(2) \end{bmatrix} = \begin{bmatrix} 4.0 & 0 \\ 1.2 & 0 \end{bmatrix} \quad (68c)$$

From Equation (68) and Figure 7, $R(\tau)$ is singular for $\tau > 0$. Therefore, a white noise control signal makes a very poor excitation except in the special case of zero measurement noise where $\tau = 0$ can be employed.

Figure 7. Actual covariance functions.

Figure 8 illustrates the covariance function as estimated from 1000 data points. The estimated function was used to estimate structure and Table 3 presents the results.

Table 3. Structure Estimation

τ	STRUCTURE	RATIO
0	1*	7.8
	2	0.36×10^{-4}
	3	0.25×10^{-5}
1	1	0.034
	2	0.82×10^{-4}
2	1	0.054
*Indicates estimated structure		

Table 3 confirms the hypothesis of white noise being a poor choice for control input (except with zero measurement noise). Table 4 summarizes the parameter estimation results.

Table 4. Parameter Estimates

τ	$\hat{\theta}_{11,1}$	S.D.	$\tilde{\theta}_{11,1}$	$\hat{\theta}_{12,1}$	S.D.	$\tilde{\theta}_{12,1}$
0	0.80	0.0017	0	1.5	0.0042	0
1	0.80	0.01	0	1.42	1.4	0.08
2	0.80	0.0052	0	1.53	0.82	0.03
S.D. denotes standard deviation as calculated from Equation (56).						

The parameters are estimated with much better accuracy than one would expect (for $\tau = 1$ or $\tau = 2$) from Equation (68). This is because (a) the $R(1)$ and $R(2)$ are not actually singular as they would ideally be and (b) because all the covariances are calculated from exactly the same data (which is not necessary for parameter estimation). However, in most applications some measure noise and uncertainty about the controls can be expected to degrade the parameter estimates when $\tau \neq 0$.

INEL 2 0165

Figure 8. Covariance functions are estimated from 1000 data points.

Example 3

Consider the system defined by Figure 9. The control signal is the only difference from the system in Example 2.

Figure 9. Scalar system with colored noise as the control signal.

For this example the colored noise control signal prevents $R(1)$ and $R(2)$ from being (approximately) singular and thus the structure and parameter estimation routines can be more effective as shown by Tables 5 and 6.

Table 5. Structure Estimation

τ	STRUCTURE	RATIO
0	1*	0.51
	2	0.43×10^{-4}
	3	0.18×10^{-4}
1	1*	0.35
	2	0.42×10^{-5}
	3	0.11×10^{-4}
2	1*	0.25
	2	0.39×10^{-4}
	3	0.16×10^{-4}
*indicates estimated structure		

Table 6. Parameter Estimates

τ	$\hat{\theta}_{11,1}$	S.D.	$\tilde{\theta}_{11,1}$	$\hat{\theta}_{12,1}$	S.D.	$\tilde{\theta}_{12,1}$
0	0.80	0.0018	0	1.50	0.010	0
1	0.80	0.0023	0	1.50	0.015	0
2	0.80	0.0028	0	1.50	0.023	0
S.D. denotes standard deviation as calculated from Equation (56).						

For this example $\hat{R}$ is technically singular even for structure = 1. However, the ratio drops by orders of magnitude when a higher order is tested; thus the system is clearly first order. The low ratio (when structure = 1) could result from the lack of high frequencies in the excitation. The example emphasizes the partly intuitive nature of the singularity test [14].

Example 4

In this example the form selection method (FSM) is used to select structures. Let the system and measurement equations for a physical system be defined by

$$\Phi = \begin{bmatrix} -0.5 & 1 & 0 & 0 \\ -0.445 & -0.8 & 0 & 0 \\ 0 & 1-\alpha & 0.7 & 0 \\ 0 & \alpha & 0 & 0.7 \end{bmatrix}, \quad (12)$$

$$H = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (13)$$

$QK = I$, and $RK = 0.1 I$. No control inputs are present. The first two states represent a damped oscillator with eigenvalues values at $-0.65 \pm j 0.65$; the third and fourth states represent separate low-pass filters with eigenvalues at 0.7. The filters connect to the oscillator through

gains of α and $1-\alpha$. Intuitively it seems reasonable to place heavy reliance on the first measurement when $\alpha \approx 0$ and heavy reliance on the second measurement when $\alpha \approx 1$. The subsequently described simulations and analysis verify this intuitive feeling.

Simulations were performed and data sets of 1000 measurement vectors collected for three values of α . Figure 10 summarizes the results of estimating system structure from the three data sets. For the first data set $\alpha = 0.1$; thus the first measurement contains a larger signal from the damped oscillator and intuitively the system parameters should be more easily observed through the first measurement than the second. Figure 10a confirms the intuition. Each dot in the tree diagram specifies an assumed structure; the pair of numbers above a dot designates the structure and the number below the dot denotes the distance of the associated R from singularity. The normed sense singular (NSS) test for singularity [6] is used and R declared singular when the distance becomes less than one. (Some properties of the NSS test are given in the discussion.)

From Figure 10a the system structure is estimated as $v_1 = 3, v_2 = 1$ or just 3,1. Examination of Equation (48) shows that this structure forces the parameter estimation to rely most heavily on the first measurement.

The second data set was generated with $\alpha = 0.3$ and Figure 10b again shows that a 3,1 structure should be used to represent the data. The third data set was generated with $\alpha = 0.5$ and Figure 10c shows that either a 3,1 or a 1,3 structure will represent this data set. The applicability of either structure is reassuring because the same strength of signal from the damped oscillator is fed into both measurements when $\alpha = 0.5$. The inability to select a structure when $\alpha = 0.5$ is a recognition of the physical situation rather than a weakness of the method.

INEL 2 0166

Figure 10. Structure estimation.

From Figure 10c one may note that 2,2 is not indicated as a candidate structure even though 1,3 and 3,1 are shown as possible structures and all three sets of indices add up to a system order, four. At first this seems unreasonable; however, for the example

$$\begin{bmatrix} h_1 \\ h_{1\phi} \\ h_2 \\ h_{2\phi} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1-\alpha & 0.7 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & \alpha & 0 & 0.7 \end{bmatrix}$$

which is singular and hence 2,2 is not a valid structure.

Very importantly the 1,3 and 3,1 structures are both valid representations for each data set studied. An identification technique based on a fixed structure would ideally select a particular structure regardless of the value of α . A fixed structure procedure does not yield the most accurate parameter estimates. Table 7 confirms that the 1,3 structure produces relatively poor parameter estimates for small α and much better estimates for large α . The 3,1 structure was verified, by simulation, to perform in the opposite fashion.

The 1,3 structure can be obtained by selecting independent vectors from the observability matrix in the following order;

$$h_1, h_2, h_{2\phi}, \overline{h_{1\phi}}, \overline{h_{2\phi}^2}, \overline{h_{2\phi}^3}$$

where the overlined vectors, containing h_i , are the first dependent vectors encountered and containing h_i . From this selection order $\phi_{21,1} = 0$ and thus does not really need to be estimated. The parameter $\phi_{12,3}$ is also zero but this fact cannot be deduced from the given selection procedure. Parameter $\phi_{12,3}$ being zero indicates that a second selection procedure could be followed in determining the 1,3 structure.

Table 7. Parameter estimates for a 1,3 structure

	$\phi_{11,1}$	$\phi_{12,1}$	$\phi_{12,2}$	$\phi_{12,3}$	$\phi_{21,1}$	$\phi_{22,1}$	$\phi_{22,2}$	$\phi_{22,3}$
$\alpha = 0.1$								
Actual Value	0.7	-6.3	9.0	0	0	0.59	0.065	-0.6
Estimated Value	0.14	-1.1	0.84	0.97	0.039	-0.03	0.36	0.29
$\alpha = 0.5$								
Actual Value	0.7	-0.7	1.0	0	0	0.59	0.065	-0.6
Estimated Value	0.63	-0.63	0.89	0.01	0.08	0.47	0.31	-0.6
$\alpha = 0.9$								
Actual Value	0.7	-0.08	0.11	0	0	0.59	0.065	-0.6
Estimated Value	0.67	-0.08	0.12	-0.02	0.02	0.59	0.18	-0.63

DISCUSSION

This report treats system identification from a least-squares point-of-view. The classic limitation of least-squares parameter estimation is the parameter-estimation bias caused by measurement noise. Previously reported methods such as maximum likelihood and generalized least squares effectively eliminate the bias but considerably complicate the numerical procedure. This report avoids biases by employing a new variation of the instrumental variable approach. The new variation is analytically and numerically more straight forward than certain other approaches.

Choosing a suitable canonical form is the essential difficulty added by going from univariate to multivariate problems. Previously reported identification algorithms depend upon a fixed form that is used for all systems. However, no single form ensures "accurate" parameter estimates for an arbitrary system because of variations in the coupling between segments of systems. This report at least partially resolves the form selection problem by using a class of canonical forms. An algorithm is employed to find a form (from within the class) which allows "accurate" parameter estimates. Explicitly searching for a suitable canonical form seems to be new to this work. Unfortunately, there is no assurance that our class of forms contains the "best" form. The class is analytically tractable but further work might reveal better classes.

Identification methods developed herein are strictly applicable only to linear time-invariant systems. In addition the system must already exist so that operating data can be obtained. However, the restrictions can often be relaxed with very little performance degradation. For example, with slight modifications the parameter estimator can age-weight the data to provide on-line parameter tracking. For non-operating systems, physically based computer models might generate data from which simple mathematical models could be developed.

A reliable test for singularity is very important for the structure determination, and the NSS test has worked well in our experience. The NSS test has three important properties. First, a calculated singularity threshold accounts for R departures from singularity due to small data sets and the associated errors in $\hat{r}_{ij}(k)$. Second, the threshold varies with matrix dimension to compensate for the inherent growth of singular values with matrix size. Third, data scaling, which changes the singular values, is accounted for with the calculated threshold.

This work does not consider evaluating or choosing excitation signals. However, Example 2 emphasizes the importance of a properly chosen signal and from [4, pp 102] it appears that the portion of R affected by the inputs should be examined to establish excitation quality.

During the study that led to this report, an alternate parameter estimator was examined which seemed to suggest a slightly different structure estimator. (Only the details differ from the algorithms given in the report.) The alternate estimator "looks" very attractive but can lead to incorrect results; thus the alternate is documented below for a simple case.

Consider a single output system, ignore the inputs, and rewrite Equation (44).

$$z[(k+\tau)+v] = \overline{\phi} \begin{bmatrix} z(k+\tau) \\ z(k+\tau+1) \\ \vdots \\ z[(k+\tau) + v-1] \end{bmatrix}$$

or

$$z[(k+\tau)+v] = [z(k+\tau) \ z(k+\tau+1) \ \dots \ z[(k+\tau) + v-1]] \overline{\phi}^T.$$

Premultiplying both sides by $\begin{bmatrix} z(k) \\ z(k+1) \\ \vdots \\ z(k+v-1) \end{bmatrix}$ yields

$$\begin{bmatrix} r(\tau+v) \\ r(\tau+v-1) \\ r(\tau+v-2) \\ \vdots \\ r(\tau+1) \end{bmatrix} = \begin{bmatrix} r(\tau) & r(\tau+1) & \dots & r(\tau+v-1) \\ r(\tau-1) & r(\tau) & \dots & \\ r(\tau-2) & & & \\ \vdots & & & \\ & & & \end{bmatrix} \bar{\phi}^T \quad (71)$$

The matrix in Equation (71) resembles $R(\tau)$ and one might be tempted to try estimating structure by increasing the dimension of this new matrix until it becomes singular. An example demonstrates that the procedure is not valid.

Consider a scalar system defined by

$$x(k) = \phi x(k-1) + w(k-1)$$

$$z(k) = x(k).$$

The measurement covariance function can be shown to be

$$E\{z(k) z(k+\tau)\} = r(\tau) = \phi^{|\tau|} \left[\frac{\text{var}[w]}{1-\phi^2} \right]; \quad (72)$$

for simplicity let $\text{var}[w]/(1-\phi^2) = 1.0$, which yields

$$r(\tau) = \phi^{|\tau|}. \quad (73)$$

By using (Equation (73)) the matrix in Equation (71) can be expressed (assuming $v=2$) for several values of τ and the determinant evaluated.

Let $\tau=0$,

$$\begin{bmatrix} r(0) & r(1) \\ r(-1) & r(0) \end{bmatrix} = \begin{bmatrix} 1 & \phi \\ \phi & 1 \end{bmatrix}, \text{Det}(\cdot) = 1 - \phi^2 \neq 0.$$

Let $\tau=1$,

$$\begin{bmatrix} r(1) & r(2) \\ r(0) & r(1) \end{bmatrix} = \begin{bmatrix} \phi & \phi^2 \\ 1 & \phi \end{bmatrix}, \text{Det}(\cdot) = \phi^2 - \phi^2 = 0.$$

The new matrix is not singular (as it should be) when $\tau=0$ and thus the matrix in Equation (71) should not be used for estimating system structure. Depending upon the type of measurement and process noise, the new matrix also fails (to provide correct structure estimates) for other values of τ .

The utility of new methods can be fully judged only after extensive application. However, the examples studied to date (not all of which are reported herein) indicate that the described identification methods are reliable and can be used to model a wide variety of systems or processes.

REFERENCES

1. K. J. Åström and P. Eykhoff, "System Identification - a Survey," *Automatica*, Volume 7, 1971, pp 123-162.
2. G. N. Saridis, "Comparison of Six On-Line Identification Algorithms," *Automatica*, Volume 10, 1974, pp 69-79.
3. D. Graupe, V. K. Jain, and J. Salarki, "A Comparative Analysis of Various Least-Square Identification Algorithms," *Automatica*, Volume 16, 1980, pp 663-681.
4. K. J. Åström and T. Bohlin, "Numerical Identification of Linear Dynamic Systems from Normal Operating Records," *Proceedings of the Second IFAC Symposium on the Theory of Self-Adaptive Control Systems*, September 14-17, 1965, National Physical Laboratory, Teddington, England, pp 96-111.
5. K. Y. Wong and E. Polak, "Identification of Linear Discrete Time Systems Using the Instrumental Variable Method," *IEEE Trans on A-C*, December 1967, pp 707-718.
6. P. Young, A. Jakeman, and R. McMurtrie, "An Instrumental Variable Method for Model Order Identification," *Automatica*, Volume 16, 1980, pp 281-294.
7. S. Gentil, J. P. Sandraz, and C. Foulard, "Different Methods for Dynamic Identification of an Experimental Paper Machine," *Identification and System Parameter Estimation (Part 1)*, *Proceedings of the Third IFAC Symposium*, The Hague/Delft, the Netherlands, June 12-15, 1973.
8. T. C. Hsia, System Identification, Lexington Books, Lexington, Massachusetts, 1977.
9. K. J. Åström, Lectures on the Identification Problem - The Least Squares Method, Report 6806, September 1968, Lund Institute of Technology.
10. E. Tse and H. L. Weinert, "Structure Determination and Parameter Identification for Multivariate Stochastic Linear Systems," 1973 *Proceedings of JACC*, Columbus, Ohio, pp 604-610.
11. B. Gopinath, "On the Identification of Linear Time-Invariant Systems from Input-Output Data," *Bell System Technical Journal*, May-June 1969, pp 1101-1113.
12. M. A. Budin, "Minimal Realization of Discrete Linear Systems from Input-Output Observations," *IEEE Trans-AC*, October 1971, pp 395-401.
13. R. Guidorzi, "Canonical Structures in the Identification of Multivariate Systems," *Automatica*, Volume 11, 1975, pp 361-374.
14. J. E. Purviance, Structure Identification of Linear Stationary Multivariate Systems from Input-Output Data, Ph.D. Dissertation, University of Idaho, 1980.

15. V. L. Shrikhande, D. P. Mital, L. M. Ray, "On Minimal Canonical Realization from Input-Output Data Sequences," IEEE Trans-AC, April 1980, pp 309-312.
16. R. K. Mehra, "On-Line Identification of Linear Dynamic Systems with Applications to Kalman Filtering," Proceedings of JACC, Atlanta, Georgia, June 24-26, 1970, pp 373-382.
17. J. M. Griffith, "Identification of Multivariate Linear Systems," Proceedings of the 27th ISA International Instrumentation Symposium, Indianapolis, Indiana, April 1981.
18. D. G. Leunberger, "Canonical Forms for Linear Multivariable Systems," IEEE Trans-AC, June 1967, pp 290-293.
19. V. M. Popov, "Invariant Description of Linear, Time-Invariant Controllable Systems," SIAM J. Control, May 1972, pp 252-264.
20. M. J. Denham, "Canonical Forms for the Identification of Multivariable Linear Systems," IEEE Trans-AC, December 1974, pp 646-656.
21. J. M. Griffith, "User's Guide for the State-Variable Systems-Identification Programs," EG&G Idaho, Internal Technical Report, ED-E3-81-010, October 1981.