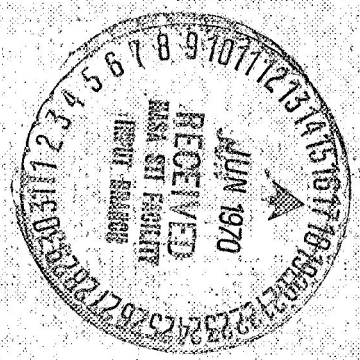## Department of AERONAUTICS and ASTRONAUTICS
STANFORD UNIVERSITY

Michael A. Budin

PARAMETER IDENTIFICATION AND
STATE ESTIMATION FOR LINEAR SYSTEMS

**January
1970**

**SUDAAR
NO. 392**

PARAMETER IDENTIFICATION AND STATE ESTIMATION

FOR LINEAR SYSTEMS

Michael A. Budin

Department of Aeronautics and Astronautics
Stanford University
Stanford, California

SUDAAR No. 392

January 1970

ACKNOWLEDGMENT

The author wishes to acknowledgment the help and guidance given to him by his advisor, Professor Benjamin Lange. The comments and suggestions of Professors Gene Franklin and John Breakwell are greatly appreciated. Special thanks are given to Dr. Paul Frost for his numerous suggestions and criticisms during many hours of discussion. For the trying job of typing the manuscript the author wishes to thank Miss Andrea Dinsmore. Finally, the author wishes to express his great appreciation to his wife, Beverly, for her constant support, understanding, and encouragement.

TABLE OF CONTENTS

LIST OF SYMBOLS

The following list of symbols is presented for reference purposes. A brief definition of each symbol is given along with the section where the symbol is introduced. Where there are two definitions for the same symbol the definition closest to the context is the relevant one.

Small Roman

Symbol	Definition	Section
$\bar{a}$	vector	4.1.1
$\bar{b}$	vector	4.1.1
c	vector	4.1.1
d	vector	4.1.1
d_i	root of characteristic equation	4.1.2
d^*	maximum of the d_i	4.1.2
k	discrete time	2.1
n	system dimension	2.2
n^*	dimension	3.1.1
n_a	dimension	3.1.1
$\hat{n}_N$	estimate of system dimension	3.2.3
p	number of outputs	2.1
r	number of inputs	2.1
t	continuous time	2.1
$u(i)$	input vector	2.1
$\bar{u}$	vector	3.1.1
$v(i)$	noise vector	3.2.1
$\bar{v}$	vector	3.2.2
$w(i)$	vector	3.2.1
$\bar{w}$	vector	3.2.2
$x(i)$	state vector	2.1
$\hat{x}$	estimate of x	2.3
$\tilde{x}$	error vector	4.1.1
$y(i)$	output vector	2.1
$\bar{y}$	vector	3.1.1
$y^{(i)}(t)$	ith derivative of $y(t)$	3.1.5
$z(i)$	observation vector	3.2.1
$\bar{z}$	vector	3.2.2

Capital Roman

Symbol	Description	Section
$\overline{A}$	matrix	4.1.1
A_{n*}	matrix	3.1.1
A_{n*_a}	matrix	3.1.1
$\tilde{A}_{n*_a}$	matrix	3.2.2
B_{n*}	matrix	3.1.1
B_{n*_a}	matrix	3.1.1
$\tilde{B}_{n*_a}$	matrix	3.2.2
$\tilde{\overline{B}}_{n*_a}$	matrix	3.2.2
C_{n*_a}	matrix	3.1.2
D_{n*_a}	matrix	3.1.2
E_{n*_a}	matrix	3.1.3
F	transition matrix	2.1
F_i	matrix	3.1.3
$\hat{F}$	estimate of F	2.3
G	input matrix	2.1
$\hat{G}$	estimate of G	2.3
H	observation matrix	2.1
$\hat{H}$	estimate of H	2.3
I	identity matrix	3.1.1
I_r	r dimensional identity matrix	3.1.1
K	discrete time space and gain matrix	2.1, 4.1.1
$\overline{K}$	matrix	3.1.3
K_i	submatrix of $\overline{K}$	3.1.3
L	delay bound	3.1.2
$\overline{M}$	matrix	4.1.1
$\overline{N}$	bound on system dimension	3.1.2
$Q_{N*,q}$	matrix	3.1.3
R	matrix	3.1.1

S	selector matrix	3.1.1
$\mathcal{S}$	selector matrix	3.1.1
$\bar{S}_i$	selector matrix	3.1.3
T	time domain and transformation matrix	2.1,2.2
U	input space and covariance matrix	2.1,4.1.2
V	covariance	4.1.2
X	state space and matrix	2.1,3.2.3
Y	output space	2.1
Z	matrix	3.2.2.c
$\bar{Z}_{n*}$	matrix	3.2.3

Small Greek

α_i	constant	4.1.1
β_i	constant	4.1.1
λ	eigenvalue	4.1.2
$\sigma_{x_{ij}}$	ijth element of Σ_x	4.1.2
$\mu(i)$	noise vector	3.2.1
$\bar{\mu}$	vector	3.2.2
ξ	vector	4.1.2
$\hat{\xi}$	estimate of ξ	4.1.2
$\tilde{\xi}$	error vector	4.1.2

Capital Greek

$\Gamma_{|i|}$	covariance matrix	3.2.1
Δ	matrix	4.1.2
$\Delta_{|i|}$	covariance matrix	3.2.1
Σ	system	2.1
Σ_ξ	covariance matrix	4.1.2
Σ_x	covariance matrix	4.1.2
Φ	space of transition matrices	2.1
Φ_i	covariance matrix	3.2.1
Ψ	vector	4.1.1
$\bar{\Psi}$	vector	4.1.1

Special Symbols and Operators

$\rho(A)$	rank of a matrix A	
$M^{\dagger}$	pseudo inverse of matrix M	
A^T	transpose of A	
$P(A)$	probability of event A	
$E\{A\}$	expected value of A	
$\chi(F)$	characteristic polynomial of F	4.1.1
$t_r(A)$	trace of a matrix A	
$_r C$	rotation of a matrix C	3.1.2
$\mathcal{J}_1(A)$	first identifier form of matrix A	3.1.2
$\|A\|$	Hilbert norm of a matrix A	
$K(\alpha, \beta, \gamma)$	function	3.1.5
$\bar{I}_n(t, \cdot)$	operator	3.1.5
$D(\tau_1, \tau_2, n, \cdot)$	operator	3.1.5
$L_n(t_1, t_2, \dots, t_n, \cdot)$	operator	3.1.5

1. INTRODUCTION

1.1 Statement of the Problem

This discussion is concerned with the problems of identifying the parameters and estimating the states of multiple input, multiple output, linear, time-invariant systems. The emphasis of the paper is not on obtaining expressions for optimal estimators which are difficult to implement or which rely on very restrictive conditions. Instead, various structural features of linear systems are utilized to obtain estimators which are computationally practical, have few restrictive conditions, and have satisfactory convergence behavior.

The need for system identification and state estimation is clear. To effectively predict the behavior of a system it is necessary to have an accurate model. If the intent is to track or control the state of a system it is also necessary to have a good idea of what the state is at each time instant. Much of systems theory can be applied only after these estimation problems are solved.

Although it would be desirable to have the ability to model and control arbitrary systems there are limitations to what can be done. Aside from the obvious limitation that not all systems can be controlled, one of the basic limitations is lack of information. This lack of information may take many forms — the statistical nature of the random phenomena involved may not be completely understood, the structural form of the system to be modeled may not be known, or the observations available may be insufficient to describe the system. To fill the gaps in information assumptions must be made, and the goodness of these assumptions will be reflected in the accuracy of the estimates.

Restricting discussion to linear, time-invariant systems is not as severe a limitation as it might seem. Many nonlinear systems can be analyzed effectively by linearizing about various regions of operation. Also, time variations may be sufficiently slow to allow the system to be adequately represented by a succession of fixed models with each model holding during a finite time interval. The developed mathematical framework available for the analysis of linear systems and the resulting simplifications make the choice of this structural form desirable. The

1

purpose of this paper is to explore various aspects of the structure of multiple input, multiple output, linear systems and to utilize the findings in the design of effective state and parameter estimators.

1.2 Previous Work

Much work has already been done in the areas of parameter and state estimation. At the foundation of the current state space approach to estimation and system description are the results of Kalman [Refs. 10,18]. He presented a linear state estimator which is optimal from the standpoint of minimizing the variance of the estimation error. The conditions for which the filter was derived are rather restrictive — the system model must be known, the statistics of the initial state must be known, and the density functions of the various random processes must be Gaussian and have known statistics. Various papers have been published dealing with the accuracy and limitations of the Kalman filter, while other papers have considered ways of extending the results.

Among the earlier efforts to extend Kalman's results are the work of Magill [Ref. 3] and Rauch [Ref. 6]. Magill considered the problem in which the system model is one of a finite set of possible models. The extension of his theory even to systems where parameters are limited to finite intervals is not computationally practical. Rauch addressed the problem of estimating the states of a system with random parameters whose statistics are known. Though an interesting problem it does not answer the question of what to do when a system with constant parameters is not known exactly.

Two different approaches have been taken to eliminate the inaccuracies of the Kalman filter which result from modeling errors. Farison [Refs. 4,5] and Kopp and Orford [Ref. 11] proposed an extended Kalman filter in which the uncertain system parameters are included as additional state elements and the inherently nonlinear equations are linearized about the expected value of this augmented state vector. The various conditions required for applicability of this solution are both restrictive and vague and, except for the special cases in which the nonlinearity disappears, results obtained using this method have

2

not been completely favorable. Schmidt [Refs. 13,14] and Neal [Ref. 22] proposed the addition of terms to the gain of the Kalman filter to account for system uncertainties and thereby prevent divergence of the filter. Although some good results have been obtained using this technique, it is essentially an ad hoc approach which may or may not work, depending on the size of the terms which are added to the gain.

Several papers have been concerned with the determination of the system model from input and output observations, without reference to the estimation of the system state. The papers of Saridis and Stein [Ref. 21] and Wong and Polak [Ref. 26] present several techniques for the identification of single input, single output systems of known size. The identification work of Gopinath [Ref. 15] is the most general, dealing with multiple input, multiple output systems of unknown size. However, Gopinath's identification procedure does not take full advantage of the structure available and is therefore not as efficient or flexible as possible.

1.3 Outline of New Results

The new results presented in this dissertation are as follows:

1) Two new methods, an algorithm and a direct procedure, for determining the minimal realization of a linear, time invariant system from input/output observations are derived.

2) The structure of the minimal realization obtained by use of the direct procedure is displayed.

3) A method is presented for obtaining the minimal realization of a continuous system without differentiating the inputs and outputs.

4) Consistent identifiers are presented which estimate the system parameters from noisy input/output observations for various conditions on the system, the input, and the noise.

5) The problem of obtaining a constant gain observer with arbitrary error dynamics is solved for a discrete system without recourse to canonical forms.

6) A bound is derived for the steady state mean square estimation error resulting from use of a constant gain state estimator. A procedure

3

which utilizes this bound is presented for choosing a constant gain estimator with satisfactory error dynamics and steady state error.

2. PROBLEM FORMULATION

2.1 Modeling the Process

A function f from A to B (f: A → B) is a mapping of each element of the set A into one and only one element of the set B. The function f may be specified in tabular form by listing all pairs (a,b) such that a∈A and b = f(a) ∈ B, or it may be specified by an algebraic expression associating a∈A and b∈B. An algebraic representation does not always exist, but when it does it is usually far more useful and practical than the tabular representation. Sometimes it is possible to deduce the algebraic representation from all or part of the tabular representation, depending on how much of the structure of the algebraic relationship is known.

These concepts are important when considering the state space representation of a dynamic system Σ. Suppose T is the time span of interest, U is the space of possible inputs, X is the space of possible system states, Φ is the set of transition functions, H is the space of output functions, and Y is the output space. Then Σ can be viewed as a function in the above sense and is specified by Σ: (T,U,X,Φ,H) → Y. Restricting attention to linear, time invariant systems the input and output can be related by the equations

$$x(k+1) = Fx(k) + Gu(k)$$

$$y(k) = Hx(k) \tag{2.1}$$

for a discrete system (k takes on integer values) or by the equations

$$\dot{x}(t) = Fx(t) + Gu(t)$$

$$y(t) = Hx(t) \tag{2.2}$$

for a continuous system, where

$x = nX1$ state vector

$y = pX1$ output vector

$u = rX1$ input vector

$F = nXn$ state transition matrix

$G = nXr$ input matrix

$H = pXn$ output matrix

5

and F, G, and H are time independent. To avoid repetition subsequent discussion shall be limited to the time invariant, discrete system of (2.1), indicating only the important results and differences for the continuous system of (2.2) in section 3.1.5.

The triplet (F,G,H) forms an internal description of Σ. Equations (2.1) constitute an algebraic description of the system. In functional notation we have $\Sigma: (K,U,X) \rightarrow Y$, where K is the set of discrete time instants of interest.

Suppose an internal description of Σ is not available, but a record of input-output pairs [u(k), y(k)] is. Since the specification of the state at some initial time instant k_o and the knowledge of all inputs for $k \geq k_o$ will determine the outputs y(k) for $k \geq k_o$ according to (2.1), this record of input-output pairs will be part of the tabular representation of Σ. We shall call the sequence of [u(k), y(k)] an external description of Σ — the term "external description" is often used to refer only to the special case of an impulse response, which seems to be an overly restrictive use of the term. The question now arises whether we can obtain an internal description from the external description. When the observations of input/output are corrupted by noise we would like to obtain an estimate of the internal description and state of Σ. These topics are the subject of Chapters 3 and 4.

2.2 Minimal Realization of the System

The internal description of a discrete dynamic system Σ which is characterized by (2.1) is not unique. It is well known that if (F_1, G_1, H_1) constitutes an internal description or realization of Σ then so will the triplet (F_2, G_2, H_2) where

$$F_2 = TF_1T^{-1}$$
$$G_2 = TG_1$$
$$H_2 = H_1T^{-1} \qquad\qquad (2.3)$$

The two realizations are said to be equivalent under the transformation T, indicated symbolically

$$(F_1, G_1, H_1) \overset{T}{\to} (F_2, G_2, H_2) \ .$$

Equivalent realizations are not the only kind. The number of states associated with a given realization is the number of elements of the state vector $x(k)$ and it is called the dimension of the realization. Two realizations of the same system can have different dimensions — they will have the same input/output behavior but they will not be equivalent.

Definition

A realization (F,G,H) of Σ is minimal if its dimension is less than or equal to the dimension of any other realization of Σ.

If (F,G,H) is an n-dimensional realization of a system having r inputs and p outputs then the dimensions of $F, G,$ and H are $n \times n$, $n \times r$, and $p \times n$, respectively.

Definition

An n-dimensional realization of Σ is completely observable iff it satisfies the rank condition

$$\rho \begin{bmatrix} H \\ HF \\ \vdots \\ HF^{n-1} \end{bmatrix} = n \ .$$

Definition

An n-dimensional realization of Σ is completely reachable iff it satisfies the rank condition

$$\rho [G \ FG \ \ldots \ F^{n-1}G] = n \ .$$

Lemma 2.1

If $\rho(H) = m$ the realization of Σ is completely observable iff

$$\rho \begin{bmatrix} H \\ HF \\ \vdots \\ HF^{n-m} \end{bmatrix} = n$$

7

Proof of Lemma 2.1

Obviously if the above rank condition is satisfied the system is completely observable. To show the necessity of this rank condition we introduce the following proposition.

Proposition 2.1

If

$$\rho \begin{bmatrix} H \\ HF \\ \cdot \\ \cdot \\ HF^q \end{bmatrix} = \rho \begin{bmatrix} H \\ HF \\ \cdot \\ \cdot \\ HF^{q+1} \end{bmatrix} = t$$

then

$$\rho \begin{bmatrix} H \\ HF \\ \cdot \\ \cdot \\ HF^s \end{bmatrix} = t \qquad \text{for all} \quad s \geq q .$$

Proof of Proposition 2.1

The fact that the rank is not increased by the additional rows HF^{q+1} implies

$$HF^{q+1} \subset \text{span}(H, HF, \ldots, HF^q)$$

where the span of $(H, HF, \ldots, HF^q)$ is the set of all linear combinations of the rows of $H, HF, \ldots, HF^q$. From this it follows that

$$HF^s \subset \text{span}(H, HF, \ldots, HF^q) \quad \text{for all} \quad s \geq q,$$

so that the additional rows HF^{q+1}, $HF^{q+2}, \ldots, HF^s$ will not change the rank and

$$\rho \begin{bmatrix} H \\ HF \\ \cdot \\ \cdot \\ HF^s \end{bmatrix} = t \qquad \text{for all} \quad s \geq q.$$

8

Returning to the proof of the lemma, if

$$\rho\begin{bmatrix} H \\ HF \\ \vdots \\ HF^{n-m} \end{bmatrix} = t < n$$

then, since $\rho(H) = m$, there exists a $q \leq n-m$ such that

$$HF^q \subset \text{span } (H,HF,\ldots,HF^{q-1})$$

(if not each set of p rows HF^i, $i \leq n-m$, would contribute a row not in span $(H,HF,\ldots,HF^{i-1})$ and

$$\rho\begin{bmatrix} H \\ HF \\ \vdots \\ HF^{n-m} \end{bmatrix}$$

would have to be n).
Then

$$\rho\begin{bmatrix} H \\ HF \\ \vdots \\ HF^{q-1} \end{bmatrix} = \rho\begin{bmatrix} H \\ HF \\ \vdots \\ HF^{q} \end{bmatrix}$$

and it follows that

$$\rho\begin{bmatrix} H \\ HF \\ \vdots \\ HF^{n-1} \end{bmatrix} = t < n$$

which contradicts the observability of the system. Therefore, if the system is completely observable

9

$$\rho \begin{bmatrix} H \\ HF \\ \circ \\ \circ \\ \circ \\ HF^{n-m} \end{bmatrix} = n$$

This completes the proof of Lemma 2.1[1]

Lemma 2.2

If $\rho(G) = \ell$ the realization of Σ is completely reachable iff $\rho[G \ FG \ \ldots \ F^{n-\ell}G] = n$.

The proof of this lemma is similar to that of lemma 2.1.

Properties of Minimal Realizations[2]

1) All minimal realizations of Σ are equivalent.

2) Any minimal realization fo Σ is completely reachable and completely observable.

3) If a realization of Σ is completely reachable and completely observable it is a minimal realization.

Since we cannot hope to identify states which are unobservable and since for all practical purposes there is no point in specifying more states than can be controlled, we shall be satisfied to obtain a minimal dimension realization of Σ from its external description. Since the minimal realizations are similar any one will suffice.

2.3 Separation of the Parameter Identifier and State Estimator

When noise free observations of the inputs and outputs of a system satisfying (2.1) are available it is possible, as shown in section 3.1, to obtain a realization of the completely observable and completely reachable portion of the system and to determine the state associated with this realization exactly. We shall deal only with finite dimensional systems so that only a finite portion of the external description is needed to yield the minimal realization.

If the input/output observations are corrupted by noise the problem becomes one of estimation — we must estimate both the parameters and the

---

[1] Lemma 2.1 will be used in section 3.1 to reduce the number of output observations needed to obtain a realization of the system.

[2] Proofs of the properties appear in ref. [15].

states. By manipulating (2.1) it is possible to obtain an expression
relating the realization and the inputs and outputs which does not
involve the state vector x(k). No such capability exists for expressing
the state vector without involving the parameters since the state vector
is inherently parameter dependent by virtue of the difference equation.
These considerations suggest the possibility of breaking the problem
into two parts — obtaining an estimate of the parameters from the inputs
and outputs and using the estimated parameters together with the inputs
and outputs to estimate the state. This is shown diagramatically in
figure 2.1.
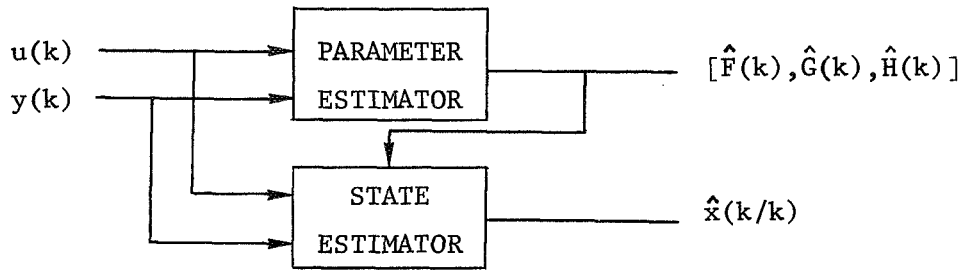
Figure 2.1. Parameter and State Estimator for Constant, Linear,
Discrete System.

There are several advantages to this separation. By estimating the
parameters from only input/output observations the nonlinearities result-
ing from coupling between the parameter and state estimates is avoided in
this phase of the identification (the nonlinearities will appear in the
estimation of the state vector). Since the state estimator operates when
the system estimate is fixed or time varying, it is not necessary to
update the system estimate after every observation — as the system estimate
converges to the true system description it is computationally practical
to update the model less frequently. It is often desired simply to obtain
a model of a given system without regard to its state so that using a
scheme in which the parameter estimate does not depend on the state
vector estimate precludes unnecessary computation.

Desirable features for the parameter and state estimators to have are
that

1) the estimators be computationally practical for a wide class of
noise statistics

11

2) the parameter estimator be consistent and have a satisfactory convergence rate

3) the state estimator have a satisfactory bound for the steady state error covariance.

It will be seen that the parameter estimators derived in Chapter 3 and the state estimator of Chapter 4 have these features.

3. PARAMETER IDENTIFICATION

3.1 Noise Free Input/Output Observations

3.1.1 Direct Relation Between Input and Output

The system equations (2.1) can be written as

$$x(k+1) = [F \ G] \begin{bmatrix} x(k) \\ u(k) \end{bmatrix} \tag{3.1}$$

$$y(k) = Hx(k)$$

If H is the identity matrix $y(k) = x(k)$ and the first of equations (3.1) becomes

$$y(k+1) = [F \ G] \begin{bmatrix} y(k) \\ u(k) \end{bmatrix} \tag{3.2}$$

which is a direct relation between the input and output—it does not have the state vector $x(k)$ appearing explicitly. Using (3.2) we can obtain [F,G] from the relation

$$\sum_{k=1}^{N} [y(k+1) \ . \ . \ . \ y(k+n+r)] = [F \ G] \sum_{k=1}^{N} \begin{bmatrix} y(k) \ . \ . \ . \ y(k+n+r-1) \\ u(k) \ . \ . \ . \ u(k+n+r-1) \end{bmatrix} \tag{3.3}$$

Whenever the matrix multiplying [F G] in (3.3) has an inverse there is a unique solution for [F G]. In the general case when the system is completely observable (H is not necessarily the identity), it is possible to obtain an expression similar to (3.2), involving the inputs, outputs, and system parameters, but not the state vectors. Before giving the more general expression (viz. eqn. (3.4)) which involves the selection of a set of output observations, it is necessary to introduce the concept of the selector matrix.

Definition

A selector matrix S is a $k \times \ell$ matrix $(k \leq \ell)$ with the property that when multiplying and $\ell \times m$ matrix A the resulting $k \times m$ matrix SA consists of k of the rows of A ordered as they are in A.

The above definition implies that $S = [s_{ij}]$ will have the properties

1) $s_{ij} = 0$ or $1 \qquad \forall i,j$

13

2) $\forall i$ there is one and only one value of j, j_i,   such that
$$s_{ij_i} = 1$$

3) $j_i < j_2 < \ldots < j_k$

Using the selector matrix we see that if a $p \times q$ matrix R has rank
r $(r \leq p, r \leq q)$ then there are two selector matrices S_1,   an $r \times p$
matrix, and S_2,   an $r \times q$ matrix, such that $S_1 R S_2^T$ is nonsingular.
If $r = q$ then $S_2 = I$,   and if $r = p$ then $S_1 = I$.

Notation

$S(i_1, i_2, \ldots, i_m)$ shall be used to specify that selector matrix
which deletes the rows $i_1, i_2, \ldots, i_m$ from the matrix it multiplies.

For a completely observable system there exists a selector matrix
S such that

$$S \, \bar{y}_{n*}(k+1) = [F \ R] \begin{bmatrix} S \, \bar{y}_{n*}(k) \\ \bar{u}_{n*}(k) \end{bmatrix} . \qquad (3.4)$$

where

$$\bar{y}_{n*}^T(k) \triangleq [y^T(k) \ y^T(k+1) \ \ldots \ y^T(k+n*-1)]$$

$$\bar{u}_{n*}^T(k) \triangleq [u^T(k) \ u^T(k+1) \ \ldots \ u^T(k+n*-1)]$$

$$R \triangleq -FSS(pn* + 1, \ldots, p(n*+1))R_{n*} + SS(1,2,\ldots,p)R_{n*}$$

$$R_{n*} \triangleq \begin{bmatrix} 0 & 0 & \ldots & \ldots & 0 \\ HG & 0 & \ldots & \ldots & 0 \\ HFG & HG & 0 & \ldots & 0 \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ HF^{n*-2}G & \ldots & & HG & 0 \\ HF^{n*-1}G & \ldots & & HFG & HG \end{bmatrix}$$

$$n* = n-\rho(H)+1$$

14

The derivation of (3.4) using n instead of n* is given in ref. [15], but for completeness and because of the differences in using n* the derivation will be given.

From (3.1) we obtain by enumeration the equations

$$y(k) = Hx(k)$$
$$y(k+1) = HFx(k) + HGu(k)$$
$$y(k+2) = HF^2x(k) + HFGu(k) + HGu(k+1)$$

$$\cdot$$
$$\cdot$$
$$\cdot$$

$$y(k+n*-1) = HF^{n*-1}x(k) + HF^{n*-2}Gu(k) + \ldots + HGu(k+n*-2)$$

(3.5)

Utilizing the definitions of $\bar{y}_{n*}(k)$, $\bar{u}_{n*}(k)$, R_{n*} and letting $S_1 = S(pn*+1, \ldots, p(n*+1))$ equations (3.5) can be written in the compact form

$$\bar{y}_{n*}(k) = \begin{bmatrix} H \\ HF \\ \cdot \\ \cdot \\ \cdot \\ HF^{n*-1} \end{bmatrix} x(k) + S_1 R_{n*} \bar{u}_{n*}(k)$$

(3.6)

Now since the system is completely observable it follows from Lemma 2.1 and the comment following the definiton of a selector matrix that there exists an n X n*p selector matrix S such that

$$S \begin{bmatrix} H \\ HF \\ \cdot \\ \cdot \\ \cdot \\ HF^{n*-1} \end{bmatrix} = T \text{ is nonsingular.}$$

(3.7)

It can be assumed that the basis of the internal description is such that T = I since a change of basis will not effect the external description. Multiplying (3.6) through by S under this assumption yields

15

$$S\bar{y}_{n*}(k) = x(k) + SS_1R_{n*}\bar{u}_{n*}(k) \qquad (3.8)$$

Using (3.1), (3.7), (3.8) and the definition of R   and letting
$S_2 = S(1,2,\ldots,p)$ we then obtain

$$S\bar{y}_{n*}(k+1) = \overset{*}{x}(k+1) + SS_1R_{n*}\bar{u}_{n*}(k+1)$$

$$= Fx(k) + Gu(k) + SS_1R_{n*}\bar{u}_{n*}(k+1)$$

$$= F[S\bar{y}_{n*}(k) - SS_1R_{n*}\bar{u}_{n*}(k)] + S \begin{bmatrix} H \\ HF \\ \cdot \\ \cdot \\ \cdot \\ HF^{n*-1} \end{bmatrix} Gu(k)$$

$$\qquad\qquad + SS_1R_{n*}\bar{u}_{n*}(k+1)$$

$$= FS\bar{y}_{n*}(k) - FSS_1R_{n*}\bar{u}_{n*}(k) + SS_2R_{n*}\bar{u}_{n*}(k)$$

$$= FS\bar{y}_{n*}(k) + R\bar{u}_{n*}(k)$$

which verifies equation (3.4).

Using (3.4) we have then

$$S[\bar{y}_{n*}(k+1) \ldots \bar{y}_{n*}(k+n+rn*)] = [F,R] \begin{bmatrix} S & | & 0 \\ \hline 0 & | & I_{rn*} \end{bmatrix} \begin{bmatrix} \bar{y}_{n*}(k)\ldots\bar{y}_{n*}(k+n+rn*-1) \\ \bar{u}_{n*}(k)\ldots\bar{u}_{n*}(k+n+rn*-1) \end{bmatrix}$$

$$(3.9)$$

We can write (3.9) compactly as

$$SA_{n*}(k+1) = [F\ R]\mathcal{S}B_{n*}(k) \qquad (3.10)$$

where the correspondences are obvious. Note that $\mathcal{S}$ is a $(rn* + n) \times$
$(r + p)n*$ selector matrix. A unique solution for $[F,R]$ exists when-
ever $\mathcal{S}B_{n*}(k)$ is nonsingular. For $\mathcal{S}B_{n*}(k)$ to be nonsingular $B_{n*}(k)$
must have rank $n + rn*$.

Lemma 3.1

For any selector matrix

$$\mathcal{S} = \left[\begin{array}{c|c} S & 0 \\ \hline 0 & I_{rn*} \end{array} \right]$$

if $\mathcal{S}B_{n*}(k)$ is nonsingular the associated S satisfies the property that

$$S \left[\begin{array}{c} H \\ HF \\ . \\ . \\ . \\ HF^{n*-1} \end{array} \right] = T \qquad (3.11)$$

is nonsingular (where T can be taken as the identity matrix without loss of generality).

Proof of Lemma 3.1

$$\mathcal{S}B_{n*}(k) = \left[\begin{array}{c|c} S & 0 \\ \hline 0 & I_{rn*} \end{array} \right] \left[\begin{array}{c|c} I_{pn*} & S_1 R_{n*} \\ \hline 0 & I_{rn*} \end{array} \right] \left[\begin{array}{cc|c} \left[\begin{array}{c} H \\ HF \\ . \\ . \\ . \\ HF^{n*-1} \end{array} \right] & & 0 \\ \hline 0 & & I_{rn*} \end{array} \right] \left[\begin{array}{c} x(k)\ldots x(k+n+rn*-1) \\ \bar{u}_{n*}(k)\ldots\bar{u}_{n*}(k+n+rn*-1) \end{array} \right]$$

$$= \left[\begin{array}{c|c} S \left[\begin{array}{c} H \\ HF \\ . \\ . \\ . \\ HF^{n*-1} \end{array} \right] & SS_1 R_{n*} \\ \hline 0 & I_{rn*} \end{array} \right] \left[\begin{array}{c} x(k) \; . \; . \; . \; x(k+n+rn*-1) \\ \bar{u}_{n*}(k) . \; . \; . \; \bar{u}_{n*}(k+n+rn*-1) \end{array} \right]$$

Since the rank of a product of two matrices is less than or equal to the rank of either factor

17

$$\rho[\not{S}B_{n*}(k)] \leq \rho\left(\begin{bmatrix} S \begin{bmatrix} H \\ HF \\ \cdot \\ \cdot \\ \cdot \\ HF^{n*-1} \end{bmatrix} & \Big| & SS_1R_{n*} \\ \hline 0 & \Big| & I_{rn*} \end{bmatrix}\right)$$

If $\not{S}B_{n*}(k)$ is nonsingular $\rho[SB_{n*}(k)] = n + rn*$. Therefore $\not{S}B_{n*}(k)$ nonsingular implies that

$$S\begin{bmatrix} H \\ HF \\ \cdot \\ \cdot \\ \cdot \\ HF^{n*-1} \end{bmatrix}$$

is nonsingular.

This completes the proof of Lemma 3.1

3.1.2 Algorithm for Determining a Minimal Realization

We shall now investigate various properties of the matrix $B_{n_a^*}(k)$ given by

$$B_{n_a^*}(k) = \begin{bmatrix} \bar{y}_{n_a^*}(k) & \bar{y}_{n_a^*}(k+1) & \cdots & \bar{y}_{n_a^*}(k+n_a+rn_a^* - 1) \\ \bar{u}_{n_a^*}(k) & \bar{u}_{n_a^*}(k+1) & \cdots & \bar{u}_{n_a^*}(k+n_a+rn_a^* - 1) \end{bmatrix} \qquad (3.12)$$

where n_a is an assumed minimal dimension of the system and does not necessarily equal the true minimal dimension. This investigation leads to an effective technique by means of which the true minimal dimension of the system and the matrices S and $[F,R]$ may be obtained.

Theorem 3.1

If the assumed minimal dimension of a system, n_a, is greater than the actual minimal dimension, n, then

$$\rho[B_{n_a^*}(k)] < n_a + rn_a^*$$

18

Proof of Theorem 3.1

$$
B_{n_a^*}(k) = \left[\begin{array}{c|c} I_{pn_a^*} & S_1 R_{n_a^*} \\ \hline 0 & I_{rn_a^*} \end{array}\right]
\left[\begin{array}{ccc}
\begin{bmatrix} H \\ HF \\ \cdot \\ \cdot \\ \cdot \\ HF^{n_a^*-1} \end{bmatrix} x(k) & \cdots & \begin{bmatrix} H \\ HF \\ \cdot \\ \cdot \\ \cdot \\ HF^{n_a^*-1} \end{bmatrix} x(k+n_a+rn_a^*-1) \\
\bar{u}_{n_a^*}(k) & \cdots\cdots & \bar{u}_{n_a^*}(k+n_a+rn_a^*-1)
\end{array}\right]
$$

where $S_1 = S(pn_a^* + 1, \ pn_a^* + 2, \ \ldots \ , \ p(n_a^* + 1))$. For simplicity in later discussions let

$$
C_{n_a^*} \triangleq \left[\begin{array}{c|c} I_{pn_a^*} & S_1 R_{n_a^*} \\ \hline 0 & I_{rn_a^*} \end{array}\right]
\tag{3.13}
$$

$$
D_{n_a^*}(k) \triangleq \left[\begin{array}{ccc}
\begin{bmatrix} H \\ HF \\ \cdot \\ \cdot \\ \cdot \\ HF^{n_a^*-1} \end{bmatrix} x(k) & \cdots & \begin{bmatrix} H \\ HF \\ \cdot \\ \cdot \\ \cdot \\ HF^{n_a^*-1} \end{bmatrix} x(k+n_a+rn_a^*-1) \\
\bar{u}_{n_a^*}(k) & \cdots\cdots & \bar{u}_{n_a^*}(k+n_a+rn_a^*-1)
\end{array}\right]
\tag{3.14}
$$

Since $C_{n_a^*}$ is nonsingular $\rho[B_{n_a^*}(k)] = \rho[D_{n_a^*}(k)]$. Now,

$$
\rho[D_{n_a^*}(k)] \leq rn_a^* + \rho\left(\begin{bmatrix} H \\ HF \\ \cdot \\ \cdot \\ \cdot \\ HF^{n_a^*-1} \end{bmatrix} [x(k) \ \ldots \ x(k+n_a+rn_a^*-1)]\right)
$$

and

19

$$\rho\left(\begin{bmatrix} H \\ HF \\ \circ \\ \cdot \\ \cdot \\ \cdot \\ HF^{n_a^*-1} \end{bmatrix} [x(k) \ldots x(k+n_a+rn_a^*-1)]\right) \leq n < n_a \; .$$

Therefore,

$$\rho[D_{n_a^*}(k)] \leq rn_a^* + n < rn_a^* + n_a$$

and thus

$$\rho[B_{n_a^*}(k)] < n_a + rn_a^* \; .$$

<u>This completes the proof of Theorem 3.1.</u>

<u>Theorem 3.2</u>

If the system Σ is completely observable and completely reachable and the assumed minimal dimension, n_a, equals the actual minimal dimension, n, then for almost all input sequences $\{u(i)\}$

$$\rho[B_{n_a^*}(k)] = n_a + rn_a^* \; .$$

<u>Proof of Theorem 3.2</u>

As in the proof of the preceding theorem we observe that

$$\rho[B_{n_a^*}(k)] = \rho[D_{n_a^*}(k)] \; .$$

Now,

$$D_{n_a^*}(k) = \begin{bmatrix} \begin{bmatrix} H \\ HF \\ \circ \\ \cdot \\ \cdot \\ HF^{n_a^*-1} \end{bmatrix} & 0 \\ 0 & I_{rn_a^*} \end{bmatrix} \begin{bmatrix} x(k) & \ldots & x(k+n_a+rn_a^*-1) \\ \bar{u}_{n_a^*}(k) & \ldots & \bar{u}_{n_a^*}(k+n_a+rn_a^*-1) \end{bmatrix}$$

Assuming the $u(i)$ are random variables with no concentrated probability mass in the joint distribution, Lemma A2 of reference [15] implies that

$$
\begin{bmatrix}
x(k) & \ldots & x(k+n+rn_a^*-1) \\
\bar{u}_{n_a^*}(k) & \ldots & \bar{u}_{n_a^*}(k+n+rn_a^*-1)
\end{bmatrix}
$$

is nonsingular for almost all input sequences. Thus, since $n_a = n$

$$
\begin{bmatrix}
x(k) & \ldots & x(k+n_a+rn_a^*-1) \\
\bar{u}_{n_a^*}(k) & \ldots & \bar{u}_{n_a^*}(k+n_a+rn_a^*-1)
\end{bmatrix}
$$

will be nonsingular for almost all input sequences. Therefore,

$$
\rho[D_{n_a^*}(k)] = \rho\left(\begin{bmatrix} \begin{bmatrix} H \\ HF \\ \cdot \\ \cdot \\ HF^{n_a^*-1} \end{bmatrix} & 0 \\ 0 & I_{rn_a^*} \end{bmatrix}\right)
$$

$$
= rn_a^* + \rho\left(\begin{bmatrix} H \\ HF \\ \cdot \\ \cdot \\ HF^{n_a^*-1} \end{bmatrix}\right)
$$

Since the system is completely observable and $n_a = n$

$$
\rho\left(\begin{bmatrix} H \\ HF \\ \cdot \\ \cdot \\ HF^{n_a^*-1} \end{bmatrix}\right) = n = n_a
$$

and

21

$$\rho[B_{n_a^*}(k)] = n_a + rn_a^* .$$

This completes the proof of Theorem 3.2.

Theorems 3.1 and 3.2 provide a means for determining the minimal system dimension, an appropriate selector matrix, and the matrix [F,R] for almost any input sequence when an upper bound on the actual minimal dimension exists. If no upper bound on the actual minimal dimension exists it is not possible to determine n with certainty; more will be said about this after the procedure to be followed when an upper bound does exist is presented. Before describing the procedure for obtaining n, S, and [F,R] it is necessary to make the following definitions.

Definition

A matrix is said to be in row-echelon form if the first nonzero element appearing in row i is a 1 in column k_i and in column k_i the only nonzero element is the 1 in row i.

Any matrix can be put into a row-echelon form by repeatedly using the elementary operations of multiplying rows by nonzero constants and of adding multiples of rows to other rows.

Definition

A row-echelon form is normal if for all rows i,j containing nonzero elements

$$k_i < k_j \qquad \text{for} \qquad i < j$$

where k_i, k_j are the columns containing the first nonzero elements of rows i and j respectively.

Definition

If in obtaining a normal row-echelon form the restrictions that

1) until no further eleminations are possible a multiple of row k may be added to row ℓ only to eliminate the jth term in row ℓ, where the first nonzero term in row k is the jth term,

2) if row k is used to eliminate the jth term in row ℓ it

must be used to eliminate the jth term in every other row, and

3) nothing is added to a row of zeros

are imposed, the resulting form is an identifier form.

The rows of the identifier form which have nonzero elements identify the corresponding rows of the original matrix as comprising a linearly independent set (the largest independent set, though not necessarily unique). Since we shall use the identifier form in much of the work to follow we introduce the matrix operator $\mathcal{J}(.)$ which operates on a matrix to yield an identifier form.

If in obtaining the identifier form we impose the additional restriction that row k may be used to eliminate the terms in the jth column only if row k is the first row with a nonzero j^{th} term and all terms before the jth are zero then we obtain the first identifier form $\mathcal{J}_1(\cdot)$. If i_j is the jth independent row of the first identifier form and i'_j is the jth independent row of any other identifier form, $i_j \leq i'_j$. The first identifier form yields the earliest linearly independent set of rows.

Example

Consider the matrix

$$B = \begin{bmatrix} 0 & 2 & 0 \\ 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Then the matrices

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

are respectively a row-echelon form associated with B, an identifier form of B, and the first identifier form of B. Note that the independent rows of the row-echelon form with ones in the first three rows do not

23

correspond to a linearly independent set of rows in B.

Definition

Given any $p \times n$ matrix $C = [c_{ij}]$ the rotation of C, $_rC$, is specified by

$$_rC = [c_{j(p-i+1)}] \qquad (3.15)$$

The mth rotation of C, abbreviated as $_r^mC$, is given as

$$_r^mC = _r[_r[\cdots {}_r[_rC]] \cdots] \qquad (3.16)$$

where m is a positive integer. Clearly $_r^4C = C$ so that if $m = 4s + t$ with s any positive integer and $t = 0,1,2,$ or 3

$$_r^mC = _r^tC . \qquad (3.17)$$

Using the above definitons and Theorems 3.1 and 3.2 we now present a stepwise procedure for determining n, S, and $[F,R]$.

Procedure 3.1

1) Set $n_a = \overline{N}$, the upper bound on the actual minimal dimension.

2) Construct the matrix $B_{n_a^*}(k)$ from observations of the input and output vectors[1].

3) Obtain $\vartheta_1(_{r2}B_{n_a^*}(k))$.[2]

4) If $\rho[\vartheta_1(_{r2}B_{n_a^*}(k))] < n_a + rn_a^*$ reduce n_a by 1 and repeat
    steps 2) and 3). If $\rho[\vartheta_1(_{r2}B_{n_a^*}(k))] = n_a + rn_a^*$ — there will
    be a 1 in each column of $\vartheta_1[_{r2}B_{n_a^*}(k)]$ with no two 1's in
    the same row — then $n = n_a$ and

---

[1] If $\rho(H)$ is not known and cannot be determined, but it is known that $\rho(H) \geq q$, use $n-q+1$ in place of n_a^* — this will only cause the possible use of more data than was necessary.

[2] If the first rn_a^* elements on the main diagonal of $\vartheta_1(_{r2}B_{n_a^*}(k))$ are not all 1's the $u(i)$ sequence is inadequate and a new $B_{n_a^*}(k)$ must be constructed from a new input sequence.

24

$$\overline{S} = {}_{r^2}[(\mathcal{J}_1({}_{r^2}B_{n^*_a}(k)))^T]$$ (3.18)

from which we easily obtain S as

$$S = S(n+1,n+2, \ldots , n+rn^*)\overline{S}\ S^T(pn^*+1,pn^*+2, \ldots , (p+r)n^*)$$ (3.19)

and

$$[F\ R] = SA_{n^*}(k+1)[\overline{S}B_{n^*}(k)]^{-1}$$ (3.20)

where $A_{n^*}(k+1)$ is constructed from the input-output observations.

As shown in ref. [15] pp. 36-38 G can be obtained from [F,R] using the relation

$$G = \overline{R}_o + F\overline{R}_1 + \ldots + F^{n^*-1}\ \overline{R}_{n^*-1}$$ (3.21)

where R is partitioned as $[\overline{R}_o \overline{R}_1 \ldots \overline{R}_{n^*-1}]$ with each $\overline{R}_i$ an
n × m matrix. (The derivation in [15] uses n instead of n*, but
will be the same when n* is substituted.) The matrix H is found by
solving the n equations

$$S \begin{bmatrix} H \\ HF \\ \cdot \\ \cdot \\ \cdot \\ HF^{n^*-1} \end{bmatrix} = I$$ (3.22)

using the S and F matrices found above.

NOTE: If an upper bound on the minimal dimension is not known the above
procedure must be modified. In this case start with $n_a = \rho(H)$ and con-
tinue to increase it until the condition $\rho[B_{n^*_a}(k)] = n_a + rn^*_a$ is
satisfied; this n_a may be the minimal dimension. If n_a satisfies
the condition that $\rho[B_{n^*_a}(k)] = n_a + rn^*_a$ and for $n'_a = n_a + \ell$,
$\rho[B_{n'_a*}(k)] < n'_a + rn'_a*$ $\forall \ell \geq 1$ then n_a is the minimal dimension.

25

In practice it is only feasible to test the rank condition on n'_a for $\ell \leq L$, so that n_a will not be the minimal dimension if the system contains time delays greater than L. For the purpose of obtaining a realization of the system, if n_a satisfies $\rho[B_{n^*_a}(k)] = n_a + rn^*_a$ and $n'_a = n_a + \ell$ satifies $\rho[B_{n'^*_a}(k)] < n'_a + rn'^*_a$ $\forall^a \ell \leq L$ (where L is chosen sufficiently large), then n_a is used as the minimal dimension in the above procedure; the resulting F, G, and H matrices may not be correct, however.

Procedure 3.1 involves the manipulation of the matrix $B_{n^*_a}(k)$ to obtain a form from which it can be determined whether $n_a = n$ and from which an appropriate S is easily obtained when $n_a = n$. This procedure differs from Gopinath's procedure in which for each choice of n_a (starting with the upper bound for n) a set of selector matrices is searched for one which satisfies Gopinath's nonsingularity test and if one exists it is an appropriate S and $n_a = n$. Both procedures suffer from the fact that if the choice of n_a is not correct we must reduce it by one and repeat the procedure, and we must continue this reduction of n_a until $n_a = n$. Thus, both procedures will require a great deal of computation if the upper bound on the minimal dimension (the initial choice of n_a) is much larger than the actual minimal dimension.

3.1.3 Direct Procedure for Determining a Minimal Realization

A closer analysis of the matrix $B_{n^*_a}(k)$ leads to the following theorem which yields an identification procedure offering significant computational improvement over the methods discussed above.

Theorem 3.3

If the system Σ is completely observable and completely reachable and the assumed system dimension, n_a, is greater than the actual minimal dimension, n, then for almost all input sequences $\{u(i)\}$

$$\rho[B_{n^*_a}(k)] = n + rn^*_a .$$

26

Proof of Theorem 3.3

From the proof of Theorem 3.1 we have

$$\rho[B_{n_a^*}(k)] = \rho[D_{n_a^*}(k)] \leq n + rn_a^* \ .$$

Consider the submatrix $E_{n_a^*}(k)$ of $D_{n_a^*}(k)$ given by

$$E_{n_a^*}(k) = \left[\begin{bmatrix} H \\ HF \\ \cdot \\ \cdot \\ \cdot \\ HF^{n^*-1} \end{bmatrix} x(k) \ \ldots \ \begin{bmatrix} H \\ HF \\ \cdot \\ \cdot \\ \cdot \\ HF^{n^*-1} \end{bmatrix} x(k+n+rn_a^*-1) \right.$$

$$\left. \begin{array}{ccc} u(k) & \ldots & u(k+n+rn_a^*-1) \\ \cdot & & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ u(k+n_a^*-1) & \ldots & u[k+n+(r+1)n_a^*-2] \end{array} \right]$$

$$= \begin{bmatrix} \begin{bmatrix} H \\ HF \\ \cdot \\ \cdot \\ \cdot \\ HF^{n^*-1} \end{bmatrix} & 0 \\ \hline 0 & I_{rn_a^*} \end{bmatrix} \begin{bmatrix} x(k) & \ldots & x(k+n+rn_a^*-1) \\ \bar{u}_{n_a^*}(k) & \ldots & \bar{u}_{n_a^*}(k+rn_a^*+n-1) \end{bmatrix}$$

Since Σ is completely reachable, the system described by

$$\begin{bmatrix} x(k+) \\ u(k+1) \\ \cdot \\ \cdot \\ \cdot \\ u(k+n_a^*) \end{bmatrix} = \begin{bmatrix} F & G & 0 & & \ldots & 0 \\ 0 & 0 & I_r & 0 & \ldots & 0 \\ & & & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & & 0 & I_r \\ 0 & 0 & \ldots & & 0 & 0 \end{bmatrix} \begin{bmatrix} x(k) \\ u(k) \\ \cdot \\ \cdot \\ u(k+n_a^*-1) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ 0 \\ I_r \end{bmatrix} u(k+n_a^*)$$

is also completely reachable. Lemma A2 of ref. [15] then implies that

$$\begin{bmatrix} x(k) & \ldots & x(k+n+rn_a^*-1) \\ \bar{u}_{n_a^*}(k) & \ldots & \bar{u}_{n_a^*}(k+n+rn_a^*-1) \end{bmatrix}$$

is nonsingular for almost all input sequences. Therefore,

$$\rho[E_{n_a^*}(k)] = \rho\left(\begin{bmatrix} \begin{bmatrix} H \\ HF \\ \vdots \\ HF^{n^*-1} \end{bmatrix} & 0 \\ 0 & I_{rn_a^*} \end{bmatrix}\right) = n + rn_a^*$$

where the last equality follows from the complete observability of the system. Since $\rho[D_{n_a^*}(k)] \geq \rho[E_{n_a^*}(k)]$ it follows that

$$\rho[D_{n_a^*}(k)] = n + rn_a^*$$

<u>This completes the proof of Theorem 3.3.</u>

Now consider the matrix $\bar{B}_{n_a^*}(k)$ given as

$$\bar{B}_{n_a^*}(k) = \begin{bmatrix} u(k) & \cdot & \cdot & \cdot & u(k+n_a+rn_a^*-1) \\ y(k) & \cdot & \cdot & \cdot & y(k+n_a+rn_a^*-1) \\ u(k+1) & \cdot & \cdot & \cdot & u(k+n_a+rn_a^*) \\ y(k+1) & \cdot & & \cdot & y(k+n_a+rn_a^*) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ u(k+n_a^*-1) & \cdot & \cdot & \cdot & u[k+n_a+(r+1)n_a^*-2] \\ y(k+n_a^*-1) & & & & y[k+n_a+(r+1)n_a^*-2] \end{bmatrix} \quad (3.23)$$

By putting $\overline{B}_{n_a^*}(k)$ into first identifier form we can obtain n and an appropriate S. The procedure for determining n, S, and $[F\ R]$ is then

Procedure 3.2

1) Construct the matrix $\overline{B}_{\overline{N}^*}(k)$, where $\overline{N}^* = \overline{N} - \rho(H) + 1$ and $\overline{N}$ is the upper bound on the minimal dimension.

2) Obtain $\mathcal{J}_1[\overline{B}_{\overline{N}^*}(k)]$. [The $r\overline{N}^*$ rows of $\overline{B}_{\overline{N}^*}(K)$ consisting of input observations will be among the independent rows and n of the first pn^* rows of $\overline{B}_{\overline{N}^*}(K)$ consisting of output observations will complete the set of independent rows.]

3) $n = $ (number of independent rows of $\mathcal{J}_1[\overline{B}_{\overline{N}^*}(k)]$)$- r\overline{N}^*$.

4) Construct the $pn^* \times n$ submatrix $\overline{K}$ of $\mathcal{J}_1[\overline{B}_{\overline{N}^*}(k)]$ consisting of the first pn^* output rows and the first n columns not containing 1's associated with input rows.

$$S = \overline{K}^T \qquad (3.24)$$

5) Construct the matrices $A_{n^*}(k+1)$, $B_{n^*}(k)$, and $\mathcal{S}$ and obtain F and G using (3.20) and (3.21).

This procedure has the advantage that in one iteration n, F, and G are determined.

Because S given by (3.24) will have its 1's in the earliest set of columns much of the structure of H and F will be known once S is determined. Partition $\overline{K}$ as

$$\overline{K} = \begin{bmatrix} K_1 \\ K_2 \\ \cdot \\ \cdot \\ \cdot \\ K_{n^*} \end{bmatrix} \qquad (3.25)$$

where each K_i is $p \times n$. Let $q_i = \rho(K_i)$. Clearly, $\displaystyle\sum_{i=1}^{n^*} q_i = n$. If $\overline{S}_1$ is the $q_1 \times p$ selector matrix such that

$$\overline{S}_1 K_1 = [I_{q_1} \mid 0] \qquad (3.26]$$

29

then

$$\bar{S}_1 H = [I_{q_1} \mid 0] \qquad (3.27)$$

where H has the form

$$H = [H_1 \mid 0] \qquad (3.28)$$
$$ p \times q_1 \quad p \times (n-q_1)$$

Designating the rows of H which are selected by $\bar{S}_1$ as principal rows, it is obvious that each of the other rows is a linear combination of the principal rows preceding it; this linear combination can be determined quite simply from a record of the output. When $q_1 = p$, H will have the form

$$H = [I_p \mid 0] . \qquad (3.29)$$

Now let F be partitioned as

$$F = \begin{bmatrix} F_1 \\ F_2 \\ \cdot \\ \cdot \\ \cdot \\ F_\tau \end{bmatrix} \qquad (3.30)$$

where F_i is $q_i \times n$ and τ is the minimum integer such that $\sum\limits_{i=1}^{\tau} q_i = n$. We can construct the sequence of selector matrices $\{\bar{S}_i\}$ with $i = 1, 2, \ldots, \tau$ such that

$$\bar{S}_j \bar{S}_{j-1} \cdots \bar{S}_1 K_j = [\; \underset{j-1}{0} \mid I_{q_j} \mid 0] \;, \quad \forall_j \qquad (3.31)$$
$$q_j \times \sum\limits_{i=1}^{j-1} q_i$$

This sequence of selector matrices will then satisfy the property that

$$\bar{S}_{j+1}F_j = [\quad 0 \quad |I_{q_{j+1}}| \; 0], \forall_j < \tau \qquad (3.32)$$

$$q_{j+1} \times \sum_{i=1}^{j} q_i$$

This will specify $(n-q_1)$ rows of F. When $q_1 = 1$ the F matrix will be in companion form.

If an upper bound on the minimal dimension does not exist, procedure 3.2 must be modified. Clearly, if for all $N \geq N_o$, $\hat{n} = \rho[\bar{B}_{N*}(k)] - rN* = $ constant, then $\hat{n} = n$. In practice values for N would be chosen until $N - \hat{n} \geq L$ where L is chosen to account for possible time delays in the system. If L is not large enough $\hat{n}$ will not be the minimal dimension and the realization will not be correct.

NOTE: In obtaining $\mathcal{I}_I[\bar{B}_{N*}(k)]$ it will be necessary to choose a number ϵ such that any number whose absolute value is less than ϵ will be taken to be zero; this is to account for computer round off error. An appropriate choice of ϵ will allow the identification of a system which is nearly linear, i.e., we will obtain a model corresponding to the linearization of the actual system. This last point is an advantage which this technique has over others.

When $\bar{N}$ is large it may be advantageous to first find the minimal dimension, n, and then find the selector matrix from $\mathcal{I}_I[\bar{B}_{n*}(k)]$. For this purpose consider the matrix

$$Q_{\bar{N}*,q}(k) \triangleq \begin{bmatrix} \bar{y}_{\bar{N}*}(k) & \bar{y}_{\bar{N}*}(k+1) \; \ldots \; \bar{y}_{\bar{N}*}(k+q-1) \\ \bar{u}_{\bar{N}*}(k) & \bar{u}_{\bar{N}*}(k+1) \; \ldots \; \bar{u}_{\bar{N}*}(k+q-1) \end{bmatrix} .$$

Proposition 3.1

If the minimal dimension, n, is less than $\bar{N}$ then for almost all input sequences the matrix $Q_{\bar{N}*,q}^T(k) \; Q_{\bar{N}*,q}(k)$ will be nonsingular for $q \leq n + r\bar{N}*$ and will be singular for $q > n + r\bar{N}*$.

Proof of Proposition 3.1

In the proof of Theorem 3.3 it was shown that $\rho[E_{n_a^*}(k)] = n + rn_a^*$, where $n_a \geq n$. Now $E_{\bar{N}*}(k)$ is a submatrix of $Q_{\bar{N}*,(n+r\bar{N}*)}(k)$ so that

31

$Q_{\bar{N}*,(n+r\bar{N}*)}(k)$ has full rank. Thus any subset of the columns of $Q_{\bar{N}*,(n+r\bar{N}*)}(k)$ will be linearly independent and $\rho[Q_{\bar{N}*,q}(k)] = q$, $\forall q \leq r+r\bar{N}*$. Since for any matrix A, $\rho(AA^T) = \rho(A)$, it follows that $Q_{\bar{N}*,q}^T(k)Q_{\bar{N}*,q}(k)$ will have full rank and thus be nonsingular for $q \leq n + r\bar{N}*$.

The matrix $Q_{\bar{N}*,(\bar{N}+r\bar{N}*)}(k)$ is simply $B_{\bar{N}*}(k)$ and it follows from Theorem 3.3 that $\rho[Q_{\bar{N}*,(\bar{N}+r\bar{N}*)}(k)] = n+r\bar{N}*$. Since $\rho[Q_{\bar{N}*,(n+r\bar{N}*)}(k)] = n+r\bar{N}*$ this means that the columns $(n+r\bar{N}*+1), (n+r\bar{N}*+2), \ldots,$ $(\bar{N}+r\bar{N}*)$ of $Q_{\bar{N}*,(\bar{N}+r\bar{N}*)}(k)$ are linearly dependent on the first $(n+r\bar{N}*)$ columns, and so $\rho[Q_{\bar{N}*,q}(k)] < q$ for $q > n+r\bar{N}*$. Therefore, $Q_{\bar{N}*,q}^T(k)Q_{\bar{N}*,q}(k)$ will be singular for $q > n+r\bar{N}*$.

This completes the proof of Proposition 3.1

Proposition 3.1 implies that the minimal dimension can be determined by finding that value of q such that $Q_{\bar{N}*,q}^T(k)Q_{\bar{N}*,q}(k)$ is nonsingular and $Q_{\bar{N}*,(q+1)}^T(k)Q_{\bar{N}*,(q+1)}(k)$ is singular. Then, letting $\bar{q}$ be the value of q satisfying this condition, n is obtained from the equation $n = \bar{q} - r\bar{N}*$. Methods of searching for $\bar{q}$ which minimize the number of computations required can easily be determined.

When n and S are determined directly from $\mathcal{J}_1[\bar{B}_{\bar{N}*}(k)]$ the property that $Q_{\bar{N}*,(n+r\bar{N}*)}^T(k)Q_{\bar{N}*,(n+r\bar{N}*)}(k)$ is nonsingular while $Q_{\bar{N}*,(n+1+r\bar{N}*)}^T(k)Q_{\bar{N}*,(n+1+r\bar{N}*)}(k)$ is singular can be used to check the minimal dimension obtained.

3.1.4 Example

Consider the system Σ given by

$$F = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \qquad G = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \qquad H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

with

$$x(0) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} .$$

Then for the input sequence

$$u(0) = -1, \quad u(1) = 2, \quad u(2) = 3, \quad u(3) = -6, \quad u(4) = -2, \quad u(5) = 1,$$

$$u(6) = -1, \quad u(7) = 7, \quad u(8) = 0, \quad u(9) = 1, \quad u(10) = -3$$

the output sequence will be

$$y(1) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \quad y(2) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad y(3) = \begin{bmatrix} -1 \\ 4 \end{bmatrix}, \quad y(4) = \begin{bmatrix} -1 \\ 2 \end{bmatrix},$$

$$y(5) = \begin{bmatrix} 3 \\ -2 \end{bmatrix}, \quad y(6) = \begin{bmatrix} 5 \\ -5 \end{bmatrix}, \quad y(7) \begin{bmatrix} 3 \\ -9 \end{bmatrix}, \quad y(8) = \begin{bmatrix} -2 \\ -6 \end{bmatrix},$$

$$y(9) = \begin{bmatrix} -11 \\ -3 \end{bmatrix}, \quad y(10) = \begin{bmatrix} -17 \\ 1 \end{bmatrix}, \quad y(11) = \begin{bmatrix} -20 \\ 2 \end{bmatrix}.$$

Assuming the upper bound on the minimal dimension of the system is $N = 4$ and assuming it is known that $\rho(H) = 2$, we have $N^* = 3$ and

$$\bar{B}_3(1) = \begin{bmatrix}
2 & 3 & -6 & -2 & 1 & -1 & 7 \\
0 & 0 & -1 & -1 & 3 & 5 & 3 \\
-1 & 0 & 4 & 2 & -2 & -5 & -9 \\
3 & -6 & -2 & 1 & -1 & 7 & 0 \\
0 & -1 & -1 & 3 & 5 & 3 & -2 \\
0 & 4 & 2 & -2 & -5 & -9 & -6 \\
-6 & -2 & 1 & -1 & 7 & 0 & 1 \\
-1 & -1 & 3 & 5 & 3 & -2 & -11 \\
4 & 2 & -2 & -5 & -9 & -6 & -3
\end{bmatrix}$$

Putting this into first identifier form we obtain

$$
\mathcal{J}_1[\overline{B}_3(1)] = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & x \\
0 & 1 & 0 & 0 & 0 & 0 & x \\
0 & 0 & 1 & 0 & 0 & 0 & x \\
0 & 0 & 0 & 1 & 0 & 0 & x \\
0 & 0 & 0 & 0 & 1 & 0 & x \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & x \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
$$

Then $n = [\rho(\mathcal{J}_1[\overline{B}_3(1)]) - 3r] = 3$ and

$$
S = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0
\end{bmatrix}
$$

Using this selector matrix we then obtain

$$
SA_2(2) = \begin{bmatrix}
0 & -1 & -1 & 3 & 5 \\
0 & 4 & 2 & -2 & -5 \\
-1 & -1 & 3 & 5 & 3
\end{bmatrix}
$$

$$
\mathcal{S}B_2(1) = \begin{bmatrix}
0 & 0 & -1 & -1 & 3 \\
-1 & 0 & 4 & 2 & -2 \\
0 & -1 & -1 & 3 & 5 \\
2 & 3 & -6 & -2 & 1 \\
3 & -6 & -2 & 1 & -1
\end{bmatrix}
$$

$$
[\mathcal{S}B_2(1)]^{-1} = \frac{1}{375} \begin{bmatrix}
460 & 535 & -84 & 248 & 138 \\
85 & 160 & -9 & 98 & -12 \\
275 & 275 & -60 & 70 & 45 \\
-155 & -5 & 87 & 11 & -9 \\
165 & 90 & 9 & 27 & 12
\end{bmatrix}
$$

Substituting these expressions into (3.20) yields

34

$$[F \ R] = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

and using (3.21) and (3.22) we obtain

$$F = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 2 & -1 \\ 0 & 1 & 1 \end{bmatrix}, \qquad G = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \qquad H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

This internal description of Σ is equivalent to the internal description given at the beginning of the example.

The matrices K_1 and K_2 are

$$K_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \qquad K_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

from which we find by inspection

$$\bar{S}_1 = I_2 \ , \qquad \bar{S}_2 = [1 \ \ 0] \ .$$

It is easily seen that $\bar{S}_1, \bar{S}_2$ together with the H and F matrices obtained satisfy (3.27) and (3.32).

3.1.5 Procedure for Determining the Minimal Realization of Continuous Systems

For the continuous time system of (2.2) the equations analagous to (3.5) are

$$y(t) \qquad = Hx(t)$$

$$y^{(1)}(t) = HFx(t) + HGu(t)$$

$$y^{(2)}(t) = HF^2x(t) + HFGu(t) + HGu^{(1)}(t) \qquad\qquad (3.33)$$

$$\vdots$$

$$y^{(n*-1)}(t) = HF^{n*-1}x(t) + HF^{n*-2}Gu(t) + \ldots + HGu^{(n*-2)}(t)$$

35

where $y^{(i)}(t) = \dfrac{d^i}{dt^i} y(t)$. The equation corresponding to (3.4) is then

$$S\bar{y}_{n*}^{(1)}(t) = [F \quad R]\begin{bmatrix} S\bar{y}_{n*}(t) \\ \bar{u}_{n*}(t) \end{bmatrix} \tag{3.34}$$

where

$$\bar{y}_{n*}^T(t) = [y^T(t) \quad y^{(1)^T}(t) \ldots y^{(n*-1)^T}(t)]$$

$$\bar{u}_{n*}^T(t) = [u^T(t) \quad u^{(1)^T}(t) \ldots u^{(n*-1)^T}(t)].$$

Mathematically it is now possible to obtain an internal description using either procedure 3.1 or 3.2 with $\bar{y}_{n*}(t)$ and $\bar{u}_{n*}(t)$ replacing $\bar{y}_{n*_a}(k)$ and $\bar{u}_{n*_a}(k)$. Because of the problems which are inherent in differentiating the signals $y(t)$ and $u(t)$ it is not feasible to use this solution. Fortunately it is possible to use a linear operator on both sides of (3.34) to yield an equation which involves no derivatives. We now present one such operator. To simplify the equations use will be made of the following definitions:

$$\bar{I}_n(t,\cdot) \triangleq \int_{t_o}^{t+t_o} \int_{t_o}^{t+t_o} \ldots \int_{t_o}^{t+t_o} (\cdot)\, d\tau_1 d\tau_2 \ldots d\tau_n \tag{3.35}$$

$$K(\alpha,\beta,\gamma) \triangleq \frac{1 - \alpha/\beta}{1 - \alpha/\gamma} \tag{3.36}$$

$$D(\tau_1,\tau_2,n,\cdot) \triangleq \bar{I}_n(\tau_1,\cdot) - (\tau_1/\tau_2)^{n-1} \bar{I}_n(\tau_2,\cdot) \tag{3.37}$$

It can then be shown that the operator

$$L_n(t_1,t_2,\ldots,t_n,\cdot) \triangleq \Big[\ldots \Big(\{[D(t_1,t_2,n,\cdot) - K(t_1,t_2,t_3) \cdot$$

$$D(t_1,t_3,n,\cdot)] - K(t_2,t_3,t_4)[D(t_1,t_2,n,\cdot) - K(t_1,t_2,t_4) \cdot$$

$$D(t_1,t_4,n,\cdot)]\} - K(t_3,t_4,t_5) \{[D(t_1,t_2,n,\cdot)$$

(continued)

36

$$- K(t_1,t_2,t_3)D(t_1,t_3,n,\cdot)] - K(t_2,t_3,t_5)[D(t_1,t_2,n,\cdot)$$

$$- K(t_1,t_2,t_5)D(t_1,t_5,n,\cdot)]\}\Big) - \ldots \Big] \tag{3.38}$$

when applied to $y^{(i)}(t)$ will yield an expression which involves no derivatives, for $i \leq n$.

Example

For $n = 3$ (3.38) becomes

$$L_3(t_1,t_2,t_3,\cdot) = D(t_1,t_2,3,\cdot) - K(t_1,t_2,t_3)D(t_1,t_3,3,\cdot) .$$

Taking the initial time as zero and applying this operator to $y^{(3)}(t)$, $y^{(2)}(t)$, $y^{(1)}(t)$, and $y(t)$ gives

$$L_3(t_1,t_2,t_3,y^{(3)}(t)) = y(t_1)[1-K(t_1,t_2,t_3)] - (t_1^2/t_2^2)y(t_2)$$

$$+ K(t_1,t_2,t_3)(t_1^2/t_3^2)y(t_3) - y(0)[1-(t_1^2/t_2^2)-K(t_1,t_2,t_3) \cdot$$

$$(1 - (t_1^2/t_3^2))]$$

$$L_3(t_1,t_2,t_3,y^{(2)}(t)) = \int_0^{t_1} y(t)dt - (t_1^2/t_2^2)\int_0^{t_2} y(t)dt$$

$$-K(t_1,t_2,t_3)\int_0^{t_1} y(t)dt + K(t_1,t_2,t_3)(t_1^2/t_3^2)\int_0^{t_3} y(t)dt$$

$$L_3(t_1,t_2,t_3,y^{(1)}(t)) = \int_0^{t_1}\int_0^{t_1} y(t)dt^2[1-K(t_1,t_2,t_3)]-(t_1^2/t_2^2)\int_0^{t_2}\int_0^{t_2} y(t)dt^2$$

$$+ K(t_1,t_2,t_3)(t_1^2/t_3^2)\int_0^{t_3}\int_0^{t_3} y(t)dt^2$$

$$L_3(t_1,t_2,t_3,y(t)) = \int_0^{t_1}\int_0^{t_1}\int_0^{t_1} y(t)dt^3[1-K(t_1,t_2,t_3)]-(t_1^2/t_2^2)\int_0^{t_2}\int_0^{t_2}\int_0^{t_2} y(t)dt^3$$

$$+ K(t_1,t_2,t_3)(t_1^2/t_3^2)\int_0^{t_3}\int_0^{t_3}\int_0^{t_3} y(t)dt^3$$

The expressions on the right are all free of derivatives.

Using the operator of (3.38) on both sides of (3.34) yields

$$SL_{n*}(t_1, t_2, \ldots, t_{n*}, \overline{y}_{n*}^{(1)}(t)) = [F,R] \begin{bmatrix} SL_{n*}(t_1, t_2, \ldots, t_{n*}, \overline{y}_{n*}(t)) \\ L_{n*}(t_1, t_2, \ldots, t_{n*}, \overline{u}_{n*}(t)) \end{bmatrix}$$

(3.39)

It is then possible to obtain an internal description using procedure 3.1 or 3.2 with $L_{n*}(t_1, \ldots, t_{n*}, \overline{y}_{n*}^{(1)}(t))$, $L_{n*}(t_1, \ldots, t_{n*}, \overline{y}_{n*}(t))$, and $L_{n*}(t_1, \ldots, t_{n*}, \overline{u}_{n*}(t))$ in derivative free form replacing $\overline{y}_{n*}(k+1)$, $\overline{y}_{n*}(k)$, and $\overline{u}_{n*}(k)$ respectively. This will avoid all problems involved with differentiation.

3.2 Identification from Noisy Input/Output Observations

3.2.1 Description of System with Noisy Observations

When the input and output measurements are corrupted by additive noise the identification procedure must be changed since noise terms will generally make the rank of $\overline{B}_{n*}(k)$, constructed from these noisy measurements, larger than it would be if constructed from noise free measurements. In the next section three distinct identification procedures will be presented. These procedures differ in the assumptions which can be made about the system, the input, and the observation noises.

The input/output description of the system will still be

$$S\overline{y}_{n*}(k+1) = [F \quad R] \begin{bmatrix} S\overline{y}_{n*}(k) \\ \overline{u}_{n*}(k) \end{bmatrix}$$

(3.40)

but instead of observing $y(k)$ and $u(k)$ we will have the measurements

$$z(k) = y(k) + v(k)$$

$$w(k) = u(k) + \mu(k)$$

(3.41)

where $v(k)$ and $\mu(k)$ are the additive noises. We shall assume that $v(k)$ and $\mu(k)$ are zero mean, are independent of the input and output, and that

38

$$E\{v(i)v^T(j)\} = \Gamma_{|i-j|}$$

$$E\{\mu(i)\mu^T(j)\} = \Delta_{|i-j|} \qquad\qquad (3.42)$$

$$E\{\mu(i)v^T(j)\} = \Phi_{|i-j|}$$

To simplify proofs in the following sections it will be assumed that all random processes are ergodic.

3.2.2 Consistent Identifiers

A drawback to many optimal identification schemes, e.g., maximum likelihood, is that they must assume the noise processes are Gaussian to make the solution computationally practical. In general this condition on the noises is not satisfied. Therefore it is desirable to have a computationally practical identification procedure which, though nonoptimal, is independent of the noise density functions and provides a convergence rate comparable to that provided by optimal identifiers. The consistent identifiers presented below satisfy these requirements.

Definition

An identifier yielding an internal representation $(\hat{F}_N, \hat{G}_N, \hat{H}_N)$ from N noisy observations of input and output is a consistent identifier if

$$P[\lim_{N \to \infty} (\hat{F}_N, \hat{G}_N, \hat{H}_N) = (F, G, H)] = 1$$

where (F, G, H) is a minimal realization of the system.

In discussing the consistent estimators presented in this section use will be made of the following definitions.

$$\bar{v}_n^T(k) \triangleq [v^T(k) \quad v^T(k+1) \quad \ldots \quad v^T(k+n-1)]$$

$$\bar{\mu}_n^T(k) \triangleq [\mu^T(k) \quad \mu^T(k+1) \quad \ldots \quad \mu^T(k+n-1)]$$

$$\bar{z}_n(k) \triangleq \bar{y}_n(k) + \bar{v}_n(k)$$

$$\bar{w}_n(k) \triangleq \bar{u}_n(k) + \bar{\mu}_n(k)$$

$$\widetilde{B}_{n^*_a}(k) \triangleq B_{n^*_a}(k) \Bigg|_{\substack{\bar{z}_{n^*_a}(k) \to \bar{y}_{n^*_a}(k) \\ \bar{w}_{n^*_a}(k) \to \bar{u}_{n^*_a}(k)}} = B_{n^*_a}(k) + B_{n^*_a}(k) \Bigg|_{\substack{\bar{v}_{n^*_a}(k) \to \bar{y}_{n^*_a}(k) \\ \bar{\mu}_{n^*_a}(k) \to \bar{u}_{n^*_a}(k)}}$$

$$\widetilde{\bar{B}}_{n^*_a}(k) \triangleq \bar{B}_{n^*_a}(k) \Bigg|_{\substack{\bar{z}_{n^*_a}(k) \to \bar{y}_{n^*_a}(k) \\ \bar{w}_{n^*_a}(k) \to \bar{u}_{n^*_a}(k)}}$$

$$\widetilde{A}_{n^*_a}(k+1) \triangleq A_{n^*_a}(k+1) \Bigg|_{\bar{z}_{n^*_a}(k) \to \bar{y}_{n^*_a}(k)} \tag{3.43}$$

It is obvious from (3.42) and the fact that the noises have zero mean that

$$E\{\widetilde{B}_{n^*_a}(k)\} = E\{B_{n^*_a}(k)\}$$

$$E\{\widetilde{\bar{B}}_{n^*_a}(k)\} = E\{\bar{B}_{n^*_a}(k)\}$$

$$E\{\widetilde{A}_{n^*_a}(k+1)\} = E\{A_{n^*_a}(k+1)\} \tag{3.44}$$

To avoid unnecessary complexity we shall in this section assume that the structure of the system is sufficiently well known that an adequate selector matrix is available. In section 3.2.3 we shall discuss the case when an adequate selector matrix is not known.

3.2.2.a Consistent Off-Line Identifier

Suppose that after sufficient observations have been taken to construct the matrices $[\widetilde{A}_{n^*}(1)]_i$ and $[\widetilde{B}_{n^*}(0)]_i$ the system is reinitialized and data is recorded to construct the matrices $[\widetilde{A}_{n^*}(1)]_{i+1}$ and $[\widetilde{B}_{n^*}(0)]_{i+1}$ corresponding to the $(i+1)^{th}$ iteration of this procedure. Consider then the equation

$$\frac{1}{N} \sum_{i=1}^{N} S[\widetilde{A}_{n*}(1)]_i = [\hat{F}_N \ \hat{R}_N] \frac{1}{N} \sum_{i=1}^{N} \mathcal{S}[\widetilde{B}_{n*}(0)]_i \ . \tag{3.45}$$

Now from the assumption of ergodicity and the fact that the noise processes have zero mean it follows that

$$\lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} \mathcal{S}[\widetilde{B}_{n*}(0)]_i = \mathcal{S}E\{B_{n*}(0)\} \ . \tag{3.46}$$

The matrix $\mathcal{S}E\{B_{n*}(0)\}$ will be nonsingular if

$$\rho\left(E\left\{ \begin{bmatrix} x(0) & \cdots & x(rn*+n-1) \\ \overline{u}_{n*}(0) & \cdots & \overline{u}_{n*}(rn*+n-1) \end{bmatrix} \right\} \right) = rn*+n \tag{3.47}$$

From the discussion in section 3.1.2 it is apparent that (3.47) will be satisfied almost surely provided that for each i $E\{u(i)\}$ be chosen from a nonlattice distribution.

Taking the limit of the left side of (3.45) yields

$$\lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} S[\widetilde{A}_{n*}(1)]_i = SE\{A_{n*}(1)\} \ . $$

Thus, provided $E\{u(i)\}$ are chosen from a nonlattice distribution, the equation

$$SE\{A_{n*}(1)\} = [F_\infty \ R_\infty] \mathcal{S}E\{B_{n*}(0)\} \tag{3.48}$$

which is the limiting form of (3.45), will have a unique solution for $[\hat{F}_\infty \ \hat{R}_\infty]$. It is apparent from taking the expectation on both sides of (3.10) that

$$[F_\infty \ R_\infty] = [F \ R] \ .$$

Therefore the *identifier* of (3.45) *is consistent.* For any N such that

41

$$\frac{1}{N} \sum_{i=1}^{N} \rho[\widetilde{B}_{n*}(0)]_i$$

is singular the estimate $[\hat{F}_N \ \hat{R}_N]$ is not made.

If no input is applied or if $E\{u(i)\} = 0 \quad \forall i$, it is possible to identify F consistently using

$$\frac{1}{N} \sum_{i=1}^{N} S[\widetilde{A}_{n*}(1)]_i = \hat{F}_N \frac{1}{N} \sum_{i=1}^{N} S[\widetilde{A}_{n*}(0)]_i \qquad (3.49)$$

provided $\rho([E\{x(0)\} \ FE\{x(0)\} \ \ldots \ F^{n-1}E\{x(0)\}]) = n$, which is satisfied for almost all $E\{x(0)\}$ if F is cyclic.

The off-line identifier has the advantage that it is not necessary to know the variances of the noise processes. The obvious disadvantage is that it is an off-line procedure and it is often desired to identify a system which is running or one which cannot be conveniently reinitialized.

3.2.2.b Consistent On-Line Identifier

Assume the covariances in (3.42) are known and define

$$E\{\overline{v}_n(i)\overline{v}_n^T(j)\} \triangleq \overline{\Gamma}_n(i-j)$$

$$E\{\overline{\mu}_n(i)\overline{\mu}_n^T(j)\} \triangleq \overline{\Delta}_n(i-j)$$

$$E\{\overline{\mu}_n(i)\overline{v}_n^T(j)\} \triangleq \overline{\Phi}_n(i-j) \qquad (3.50)$$

Now consider the equation

$$\left\{ \frac{1}{N} \sum_{k=1}^{N} S\widetilde{A}_{n*}(k+1)\widetilde{B}_{n*}^T(k)\mathbf{\rho}^T - (n+rn*)[S\overline{\Gamma}_{n*}(1)S^T \quad S\overline{\Phi}_{n*}^T(-1)] \right\} = \qquad (3.51)$$

$$[\hat{F}_N \ \hat{R}_N] \left\{ \frac{1}{N} \sum_{k=1}^{N} \rho\widetilde{B}_{n*}(k)\widetilde{B}_{n*}^T(k)\mathbf{\rho}^T - (n+rn*)\begin{bmatrix} S\overline{\Gamma}_{n*}(0)S^T & S\overline{\Phi}_{n*}^T(0) \\ \overline{\Phi}_{n*}(0)S^T & \overline{\Delta}_{n*}(0) \end{bmatrix} \right\}$$

From the assumptions of ergodicity and independence of signals and noise it is easily shown that the limiting form of (3.51) as N tends to infinity will be

$$E\{SA_{n*}(k+1)B_{n*}^T(k)\mathcal{S}^T\} = [\hat{F}_\infty \ \hat{R}_\infty]E\{\mathcal{S}B_{n*}(k)B_{n*}^T(k)\mathcal{S}^T\} \tag{3.52}$$

Provided the input sequence is such that

$$\rho[E\{\bar{u}_{n*}(k)\bar{u}_{n*}^T(k)\}] = rn* \tag{3.53}$$

the matrix multiplying $[\hat{F}_\infty \ \hat{R}_\infty]$ will be nonsingular and a unique solution will exist for $[\hat{F}_\infty \ \hat{R}_\infty]$. Multiplying both sides of (3.10) from the right by $B_{n*}^T(k)\mathcal{S}^T$ it is apparent that

$$[\hat{F}_\infty \ \hat{R}_\infty] = [F \ R]$$

so that the identifier of (3.51) is consistent.

This identifier has the advantage that it uses the data from a continuously running system, but it has the serious drawback that the noise covariances must be known. Errors in the covariances will result in bias in the parameter estimates.

3.2.2.c Variance Free On-Line Identifier

If the inputs are observed without noise and the observation noise on the outputs is such that for some finite M, $\bar{\Gamma}_{n*}(i) = 0$ $\forall i \ni |i| \geq M-1$ it is possible in some cases to identify the system on-line without knowing the noise covariance. Define

$$Z_{n*}(k,M) \triangleq \begin{bmatrix} \bar{y}_{n*}(k+M) & \cdots & \bar{y}_{n*}(k+M+n+rn*-1) \\ \bar{u}_{n*}(k) & \cdots & \bar{u}_{n*}(k+n+rn*-1) \end{bmatrix}$$

$$\tilde{Z}_{n*}(k,M) \triangleq Z_{n*}(k,M) \Big|_{\bar{z}_{n*}(k) \to \bar{y}_{n*}(k)} \tag{3.54}$$

and consider the equation

$$\frac{1}{N} \sum_{k=1}^{N} S\widetilde{A}_{n*}(k+1)\widetilde{Z}_{n*}^{T}(k,M)\mathcal{S}^{T} = [\hat{F}_{N} \; \hat{R}_{N}] \frac{1}{N} \sum_{k=1}^{N} \mathcal{S}\widetilde{B}_{n*}(k)\widetilde{Z}_{n*}^{T}(k,M)\mathcal{S}^{T} \tag{3.55}$$

As N tends to infinity we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} S\widetilde{B}_{n*}(k)\widetilde{Z}_{n*}^{T}(k,M)\mathcal{S}^{T} = E\{\mathcal{S}B_{n*}(k)Z_{n*}^{T}(k,M)\mathcal{S}^{T}\}$$

$$= (n+rn*)E\left\{ \begin{bmatrix} S\bar{y}_{n*}(k) \\ \bar{u}_{n*}(k) \end{bmatrix} [\bar{y}_{n*}^{T}(k+M)S^{T} \; \bar{u}_{n*}^{T}(k)] \right\}$$

and

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} S\widetilde{A}_{n*}(k+1)\widetilde{Z}_{n*}^{T}(k,M)\mathcal{S}^{T} = E\{SA_{n*}(k+1)Z_{n*}^{T}(k,M)\mathcal{S}^{T}\} \; .$$

Now from (3.4) it follows that

$$S\bar{y}_{n*}(k+M) = F^{M}S\bar{y}_{n*}(k) + F^{M-1}R\bar{u}_{n*}(k) + \ldots + R\bar{u}_{n*}(k+M-1) \tag{3.56}$$

so that

$$E\{S\bar{y}_{n*}(k)\bar{y}_{n*}^{T}(k+M)S^{T}\} = E\{S\bar{y}_{n*}(k)\bar{y}_{n*}^{T}(k)S^{T}F^{M^{T}}\} +$$

$$E\{S\bar{y}_{n*}(k)\bar{u}_{n*}^{T}(k)R^{T}F^{M-1^{T}}\} +$$

$$\cdot \quad \cdot \quad \cdot \quad +$$

$$E\{S\bar{y}_{n*}(k)\bar{u}_{n*}^{T}(k+M-1)R^{T}\} \; . \tag{3.57}$$

Therefore, if F is nonsingular and the input satisfies (3.53) the matrix

$$E\{\mathcal{S}B_{n*}(k)Z_{n*}^{T}(k,M)\mathcal{S}^{T}\}$$

will be nonsingular, so that $[F_{\infty} \; R_{\infty}]$ is uniquely determined. For F nonsingular the matrix $\mathcal{S}Z_{n*}(k,M)$ will almost surely have full rank, and multiplying (3.4) on both sides by $Z_{n*}^{T}(k,M)\mathcal{S}^{T}$ from the right

44

it follows that $[F_\infty \; R_\infty] = [F \; R]$.

When the inputs are observed with noise it is still possible to identify a system with a nonsingular F matrix provided that 1) for some P $\bar{\Delta}_{n*}(i) = 0$ $\forall i \ni |i| \geq P$ and $\bar{\Phi}_{n*}(j) = 0$ $\forall j \ni |j| \geq P-1$, and 2) the input sequence satisfies the condition $\rho[E\{\bar{u}_{n*}(k)\bar{u}_{n*}^{-T}(k-P)\}] = rn*$. Under these conditions consistent on line identification can be obtained using

$$\frac{1}{N} \sum_{k=1}^{N} S\tilde{A}_{n*}(k+1)\tilde{V}_{n*}^{T}(k,M,P)S^{T} = [\hat{F}_N \; \hat{R}_N] \frac{1}{N} \sum_{k=1}^{N} S\tilde{B}_{n*}(k)\tilde{V}_{n*}^{T}(k,M,P)S^{T} \qquad (3.58)$$

where

$$V_{n*}(k,M,P) \triangleq \begin{bmatrix} \bar{y}_{n*}(k+M) & \cdots & \bar{y}_{n*}(k+M+n+rn*-1) \\ \bar{u}_{n*}(k-P) & \cdots & \bar{u}_{n*}(k-P+n+rn*-1) \end{bmatrix}$$

and

$$\tilde{V}_{n*}(k,M,P) = V_{n*}(k,M,P) \Bigg|_{\substack{\bar{z}_{n*}\rightarrow\bar{y}_{n*} \\ \bar{w}_{n*}\rightarrow\bar{u}_{n*}}}$$

The choice of which consistent estimator to use depends on 1) the information available about the noise processes, 2) the singularity of the F matrix, 3) the ability to reinitiate the system, and 4) the amount of control over the inputs. Each of the identifiers presented has the advantage that it is computationally practical for a wide class of noise processes and not just gaussian noise processes.

3.2.3 Consistent Determination of the Selector Matrix

In the preceding section it was assumed that the system dimension and an appropriate selector matrix were known. When this is not the case n and S must be estimated from the noisy observations and these estimates, $\hat{n}_N$ and $\hat{S}_N$, would be used in (3.45), (3.51), or (3.55). Consistent methods for determining n and S will now be given for each condition in the previous section. It shall be assumed throughout

this discussion that an upper bound $\bar{N}$ on the minimal dimension is known.

Because of the noise the rank of $\tilde{B}_{\bar{N}*}(k)$ will almost surely be $\bar{N} + r\bar{N}*$. Suppose that $[\bar{B}_{\bar{N}*}(0)]_i$ corresponds to the ith reinitializing of the system. Then almost surely $\rho([\bar{B}_{\bar{N}*}(0)]_i) = \bar{N} + r\bar{N}*$ for each i. However, averaging over i as the number of initializations becomes infinite, we have

$$\rho \left\{ \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} [\tilde{B}_{\bar{N}*}(0)]_i \right\} = \rho[E\{\bar{B}_{\bar{N}*}(0)\}] \qquad (3.59)$$

When the conditions under which (3.45) is consistent are satisfied

$$\rho[E\{\bar{B}_{\bar{N}*}(0)\}] = n + r\bar{N}* . \qquad (3.60)$$

For any finite value of N the rank of

$$\frac{1}{N} \sum_{i=1}^{N} [\tilde{B}_{\bar{N}*}(0)]_i$$

would almost surely be $\bar{N} + r\bar{N}*$. However, it can easily be shown that the variance of each term of the matrix

$$\frac{1}{N} \sum_{i=1}^{N} [\tilde{B}_{\bar{N}*}(0)]_i$$

decreases as $1/N$. Therefore, for an appropriate choice of the tolerance used in obtaining the identifier form

$$\rho \left(\mathcal{J}_1 \left(\frac{1}{N} \sum_{i=1}^{N} [\tilde{B}_{\bar{N}*}(0)]_i \right) \right) = n + r\bar{N}* \quad \text{for almost all } N \geq N_o \qquad (3.61)$$

for some constant N_o. Theoretically the tolerance ϵ could be chosen as a decreasing function of N, converging to zero in the limit. In

practice ϵ must always be greater than zero to account for the computer round off. For use in (3.45) $\hat{n}_N$ and $\hat{S}_N$ would thus be obtained using procedure 3.2 with

$$\frac{1}{N} \sum_{i=1}^{N} [\tilde{\bar{B}}_{\bar{N}*}(0)]_i$$

replacing $\bar{B}_{\bar{N}*}(k)$ and the tolerance chosen appropriately. From (3.61) it follows that

$$\hat{n}_N = n$$
$$\qquad\qquad \text{for almost all } N \qquad\qquad\qquad (3.62)$$
$$\hat{S}_N = S$$

when ϵ is chosen appropriately and from (3.59) and (3.60)

$$P\{\hat{n}_\infty = n\} = P\{\hat{S}_\infty = S\} = 1 . \qquad\qquad (3.63)$$

When (3.50) is used for identification $\hat{n}_N$ and $\hat{S}_N$ should be determined from

$$\mathcal{J}_1 \left(\frac{1}{N} \sum_{k=1}^{N} \tilde{\bar{B}}_{\bar{N}*}(k)\tilde{\bar{B}}_{\bar{N}*}^T(k) - (p+r)\bar{N}*C \right) \qquad (3.64)$$

where

$$C = E\{\bar{v}_{\bar{N}*}(k)\bar{v}_{\bar{N}*}^T(k)\}$$

$$\bar{v}_{n*}^T(k) \triangleq [v^T(k) \ \mu^T(k) \ v^T(k+1) \ \ldots \ v^T(k+n+rn*-1) \ \mu^T(k+n+rn*-1)]$$

The limit of (3.64) as N tends to infinity is

$$\mathcal{J}_1(E\{\bar{B}_{\bar{N}*}(k) \ \bar{B}_{\bar{N}*}^T(k)\}) . \qquad\qquad (3.65)$$

Lemma 3.2

For any matrix X, if S^* is the selector matrix with $\rho(S^*) = \rho(X)$ and $\rho[S^* \mathcal{I}_1(X)] = \rho(X)$, then $\rho[S^* \mathcal{I}_1(XX^T)] = \rho(X)$.

Proof of Lemma 3.2

Suppose X is $p \times q$ and has rank r. The statement that $\rho(S^*) = \rho(X)$ simply implies that S^* is an $r \times p$ selector matrix, and the fact that $\rho[S^* \mathcal{I}_1(X)] = \rho(X)$ means that S^* is that $r \times p$ selector matrix which picks out the r independent rows of $\mathcal{I}_1(X)$.

Since $\rho(XX^T) = \rho(X)$, $\mathcal{I}_1(XX^T)$ will have r independent rows. We want to show that these rows are the same as the independent rows of $\mathcal{I}_1(X)$. Now if the ith row of X is linearly dependent on the rows preceding it, it will not be an independent row in $\mathcal{I}_1(X)$; and designating the ith row of X by x_i it follows that the ith row of XX^T, which is $x_i X^T$, will be linearly dependent on the preceding rows of XX^T. Thus, all zero rows in $\mathcal{I}_1(X)$ will be zero rows in $\mathcal{I}_1(XX^T)$, and the independent rows must also correspond.

This completes the proof of Lemma 3.2

Lemma 3.2 implies that the selector matrix obtained from $\mathcal{I}_1(\overline{B}_{\overline{N}*}(k) \, \overline{B}_{\overline{N}*}^T(k))$ will be the same as that obtained from $\mathcal{I}_1(\overline{B}_{\overline{N}*}(k))$. Therefore, by appropriately choosing the tolerance $\epsilon(N)$, the estimates of n and S obtained using (3.64) should satisfy (3.62) and in view of (3.65) will satisfy (3.63).

Similar considerations apply for estimating n and S when (3.55) is used for identification. In this case $\hat{n}_N$ and $\hat{S}_N$ would be obtained from

$$\mathcal{I}_1 \left(\frac{1}{N} \sum_{k=1}^{N} \widetilde{B}_{\overline{N}*}(k) \, \widetilde{Z}_{\overline{N}*}^T(k,M) \right)$$

where

$$\bar{Z}_{n*}(k,M) \triangleq \begin{bmatrix} y(k+M) & \cdot\ \cdot\ \cdot & y(k+M+n+rn*-1) \\ u(k) & \cdot\ \cdot\ \cdot & u(k+n+rn*-1) \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ y(k+M+n*-1) & \cdot\ \cdot\ \cdot & y(k+M+n+(r+1)n*-1) \\ u(k+n*-1) & \cdot\ \cdot\ \cdot & u(k+n+(r+1)n*-1) \end{bmatrix}$$

$$\widetilde{\bar{Z}}_{n*}(k,M) = \left. \bar{Z}_{n*}(k,M) \right|_{z_{n*} \rightarrow y_{n*}} \cdot$$

The tolerance $\epsilon(N)$ has the effect of determing with what precision the model $(\hat{F}_N,\ \hat{G}_N,\ \hat{H}_N)$ would match the output of the actual system (F,G,H).

NOTE: Considerations completely analogous to those presented in sections 3.2.2 and 3.2.3 are also valid for continuous systems.

4. STATE ESTIMATION

4.1 Constant Gain Observer

To estimate the state of a constant linear system Σ described by

$$x(k+1) = Fx(k) + Gu(k)$$

$$y(k) = Hx(k) \tag{4.1}$$

from observations of input and output a convenient method is to construct another linear system $\hat{\Sigma}$ given by

$$\hat{x}(k+1|k) = \hat{F}\hat{x}(k|k) + \hat{G}w(k)$$

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + K(k+1)[z(k+1) - \hat{H}\hat{x}(k+1|k)] \tag{4.2}$$

where

$\hat{x}(k+1|j) \triangleq$ estimate of $x(k+1)$ based on observations of input and output up to time j

$(\hat{F},\hat{G},\hat{H}) =$ model of the actual system

$w(k) =$ noisy observation of input $u(k)$

$z(k) =$ noisy observation of output $y(k)$.

The state of $\hat{\Sigma}$ can be observed. The problem is to choose the gain $K(k+1)$ so that the error $\tilde{x}(k+1|k+1) \triangleq x(k+1) - \hat{x}(k+1|k+1)$ is made as small as possible — for noisy input/output observations $K(k+1)$ should be chosen to minimize the trace of the error covariance matrix.

If $(\hat{F},\hat{G},\hat{H}) = (F,G,H)$ and the mean and covariance of the state are known at some initial time k_o, then the Kalman filter solution will be optimal. When the model differs from the actual system, use of the Kalman solution would yield incorrect expressions for the covariance, and the gain sequence, which depends on the covariance, would in no way be optimal—in fact, use of this gain sequence may result in divergence of the actual and estimated states. Also, initial errors in the mean or covariance of the state cause the Kalman filter to yield incorrect expressions which might result in divergence. These limitations are

quite important since model errors will generally exist and since accurate statistics on the state are usually unavailable.

A constant gain state estimator, though sub-optimal, does not have these limitations since knowledge of the covariance is not required. This filter is less sensitive to model errors, though divergence may still result if the errors are large. In section 4.1.1 the case of noise free observations is considered, and it is shown how for a minimal system the constant gain may be chosen to yield almost arbitrary dynamics for the estimation error (arbitrary dynamics when F is nonsingular). Section 4.1.2 deals with the case of noisy observations, and criteria for choosing K are established. In both sections model errors will be neglected — these will be considered in section 4.2.

4.1.1 Gain Selection for Noise Free Observations of Input/Output

For noise free observations, no model errors, and constant gain equations (4.2) become

$$\hat{x}(k+1|k) = F\hat{x}(k|k) + Gu(k)$$

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + K[y(k+1) - H\hat{x}(k+1|k)] .$$ (4.3)

From (4.1) and (4.3) it is easily seen that

$$\tilde{x}(k+1|k+1) = (F-KHF)\tilde{x}(k|k) .$$ (4.4)

Theorem 4.1

If the system is completely observable and F is cyclic and non-singular the gain K can be chosen so that (F - KHF) has a prescribed characteristic polynomial.

Proof of Theorem 4.1

Suppose the characteristic polynomial $\chi(F)$ is

$$\chi(F) = s^n + \sum_{i=1}^{n} \alpha_i s^{n-i}$$ (4.5)

51

and the desired characteristic polynomial $\chi(F-KHF)$ is

$$\chi(F-KHF) = s^n + \sum_{i=1}^{n} \beta_i s^{n-i} \tag{4.6}$$

Now,

$$\chi(F-KHF) = \det(sI-F + KHF)$$

$$= \det[(sI-F)(I+(sI-F)^{-1}KHF)]$$

$$= \chi(F)\det[I+(sI-F)^{-1}KHF]$$

If the choice of K is restricted to those gain matrices of rank one (i.e., $K = cd^T$ where c is an $n\times 1$ matrix and d is a $p\times 1$ matrix) then

$$\det[I+(sI-F)^{-1}KHF] = 1 + \text{tr}[(sI-F)^{-1}KHF]$$

and

$$\chi(F-KHF) = \chi(F) + \text{tr}[\chi(F)(sI-F)^{-1}KHF] \tag{4.7}$$

$$= \chi(F) + \text{tr}[\chi(F) \sum_{i=0}^{\infty} \frac{F^i}{s^{i+1}} KHF]$$

Using (4.5) it follows that

$$\chi(F) \sum_{i=0}^{\infty} \frac{F^i}{s^{i+1}} = \sum_{i=0}^{\infty} F^i s^{n-i-1} + \alpha_1 \sum_{i=0}^{\infty} F^i s^{n-i-2} + \ldots + \alpha_n \sum_{i=0}^{\infty} F^i s^{-i-1}$$

$$= (F^n + \alpha_1 F^{n-1} + \ldots + \alpha_n) \sum_{i=0}^{\infty} F^i s^{-i-1}$$

$$+ (\alpha_{n-1} + \alpha_{n-2}F + \ldots + F^{n-1})$$

$$+ (\alpha_{n-2} + \alpha_{n-3}F + \ldots + F^{n-2})s$$

$$+ \ldots$$

$$+ (\alpha_1 + F)s^{n-2} + s^{n-1} \tag{4.8}$$

From the Cayley-Hamilton theorem we have

$$F^n + \sum_{i=1}^{n} \alpha_i F^{n-i} = 0$$

so that (4.8) becomes

$$\chi(F) \sum_{i=0}^{\infty} \frac{F^i}{s^{i+1}} = (\alpha_{n-1} + \alpha_{n-2} F + \ldots + F^{n-1})$$

$$+ (\alpha_{n-2} + \alpha_{n-3} F + \ldots + F^{n-2}) s$$

$$+ \ldots$$

$$+ (\alpha_1 + F) s^{n-2} + s^{n-1} \qquad (4.9)$$

and substituting this into (4.7) yields

$$\chi(F-KHF) = \chi(F) + [\alpha_{n-1} \, tr(KHF) + \alpha_{n-2} \, tr(FKHF) + \ldots$$

$$+ tr(F^{n-1} KHF)] + [\alpha_{n-2} \, tr(KHF) + \ldots + tr(F^{n-2} KHF)] s$$

$$+ \ldots + [tr(KHF)] s^{n-1} \qquad (4.10)$$

Equating the coefficients of the various powers of s in (4.6) and (4.10) we obtain the following equations

$$\beta_1 = \alpha_1 + tr(KHF)$$

$$\beta_2 = \alpha_2 + \alpha_1 \, tr \, (KHF) + tr(FKHF)$$

$$.$$
$$. \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (4.11)$$
$$.$$

$$\beta_n = \alpha_n + \alpha_{n-1} \, tr(KHF) + \ldots + tr(F^{n-1} KHF)$$

Equations (4.11) can be written more compactly as

$$b = a + A\Psi \qquad (4.12)$$

where

$$b^T \triangleq [\beta_1 \quad \beta_2 \, \cdot \, \cdot \, \cdot \, \beta_n]$$

$$a^T \triangleq [\alpha_1 \quad \alpha_2 \, \cdot \, \cdot \, \cdot \, \alpha_n]$$

$$A \triangleq \begin{bmatrix} 1 & 0 & \cdot & \cdot & \cdot & 0 \\ \alpha_1 & 1 & 0 & \cdot & \cdot & \cdot & 0 \\ \alpha_2 & \alpha_1 & 1 & 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & & & & \\ \cdot & \cdot & & & & \\ \cdot & \cdot & & & & \\ \alpha_{n-1} & \alpha_{n-2} & \cdot & \cdot & \cdot & \alpha_1 & 1 \end{bmatrix}$$

$$\Psi \triangleq \begin{bmatrix} \mathrm{tr}(KHF) \\ \mathrm{tr}(FKHF) \\ \cdot \\ \cdot \\ \cdot \\ \mathrm{tr}(F^{n-1}KHF) \end{bmatrix}$$

Since A is nonsingular

$$\Psi = A^{-1}(b-a) \tag{4.13}$$

Now, for $K = cd^T$ we obtain

$$\mathrm{tr}(F^i KHF) = \mathrm{tr}(HF^{i+1}K)$$

$$= \mathrm{tr}(HF^{i+1}cd^T)$$

$$= d^T HF^{i+1}c \tag{4.14}$$

so that

$$\begin{bmatrix} d^T H \\ d^T HF \\ \cdot \\ \cdot \\ \cdot \\ d^T HF^{n-1} \end{bmatrix} Fc = A^{-1}(b-a) \qquad (4.15)$$

For almost any $p \times 1$ matrix d the matrix

$$M \triangleq \begin{bmatrix} d^T H \\ d^T HF \\ \cdot \\ \cdot \\ \cdot \\ d^T HF^{n-1} \end{bmatrix}$$

will be nonsingular. Therefore, after choosing an appropriate d, one for which M is nonsingular, c is found from the equation

$$c = F^{-1}M^{-1}A^{-1}(b-a) \qquad (4.16)$$

and the gain $K = cd^T$, with c and d determined in this manner will produce the prescribed characteristic polynomial.
<u>This completes the proof of Theorem 4.1.</u>

<u>Theorem 4.2</u>

If the system is completely observable and F is cyclic and singular, the characteristic polynomial of $(F-KHF)$ will have no constant term but the gain K can be chosen to yield an otherwise arbitrary characteristic equation.

<u>Proof of Theorem 4.2</u>

The constant term β_n in (4.6) is given by

$$\beta_n = \det(F-KHF) = \det(I-KH)\det(F)$$

so that if F is singular $(\det(F) = 0)$ $\beta_n = 0$. In this case equations (4.11) reduce to

$$\beta_1 = \alpha_1 + \text{tr}(KHF)$$

$$\beta_2 = \alpha_2 + \alpha_1 \; \text{tr}(KHF) + \text{tr}(FKHF)$$

$$\cdot$$
$$\cdot \qquad\qquad\qquad\qquad\qquad\qquad\qquad (4.17)$$
$$\cdot$$

$$\beta_{n-1} = \alpha_{n-1} + \alpha_{n-2} \; \text{tr}(KHF) + \ldots + \text{tr}(F^{n-2}KHF)$$

which can be written compactly as

$$\bar{b} = \bar{a} + \bar{A}\,\bar{\Psi} \qquad\qquad\qquad\qquad (4.18)$$

where

$$\bar{b}^T \triangleq [\beta_1 \; \beta_2 \; \cdot \; \cdot \; \cdot \; \beta_{n-1}]$$

$$\bar{a}^T \triangleq [\alpha_1 \; \alpha_2 \; \cdot \; \cdot \; \cdot \; \alpha_{n-1}]$$

$$\bar{A} \triangleq \begin{bmatrix} 1 & 0 & \cdot & \cdot & \cdot & 0 \\ \alpha_1 & 1 & 0 & \cdot & \cdot & 0 \\ \cdot & & \cdot & & & \\ \cdot & & \cdot & & & \\ \cdot & & \cdot & & & \\ \alpha_{n-2} & \alpha_{n-3} & \cdot & \cdot & \cdot & \alpha_1 & 1 \end{bmatrix}$$

$$\bar{\Psi} \triangleq \begin{bmatrix} \text{tr}(KHF) \\ \text{tr}(FKHF) \\ \cdot \\ \cdot \\ \cdot \\ \text{tr}(F^{n-2}KHF) \end{bmatrix} = \begin{bmatrix} d^THF \\ d^THF^2 \\ \cdot \\ \cdot \\ \cdot \\ d^THF^{n-1} \end{bmatrix} c\;.$$

For almost any choice of d the matrix

$$\overline{M} \triangleq \begin{bmatrix} d^T HF \\ d^T HF^2 \\ \cdot \\ \cdot \\ \cdot \\ d^T HF^{n-1} \end{bmatrix}$$

will have full rank (i.e., $\rho(\overline{M}) = n-1$). Then, since $\overline{A}$ is nonsingular, (4.18) represents (n-1) linear, independent equations in n unknowns, the elements of c. Therefore, (4.18) will have many solutions for c. One such solution, which minimizes the Euclidean norm of c, is

$$c = \overline{M}^\dagger \overline{A}^{-1}(\overline{b} - \overline{a}) \qquad (4.20)$$

Thus, after finding a choice of d such that $\rho(\overline{M}) = n-1$, an $n \times 1$ matrix c can be found using (4.20) such that $K = cd^T$ will yield the desired characteristic polynomial.

<u>This completes the proof of Theroem 4.2</u>

Theorems 4.1 and 4.2 show that for a completely observable system with a cyclic F matrix the dynamics of the error in (4.4) can be made arbitrary if F is nonsingular and almost arbitrary if F is singular by an appropriate choice of the gain K. In particular K can be chosen such that the characteristic roots of (F-KHF) are all zero — this will cause the error to go to zero in at most n steps. The convergence of the constant gain observer does not depend on the accuracy of the initial statistics of $\tilde{x}$.

4.1.2 Gain Selection for Noisy Observations of Input/Output

When the observations are corrupted by noise equations (4.3) become

$$\hat{x}(k+1|k) \quad = F\,\hat{x}(k|k) + G\,w(k)$$

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + K[z(k+1) - H\,\hat{x}(k+1|k)] \qquad (4.21)$$

and (4.4) becomes

$$\tilde{x}(k+1|k+1) = (F-KHF)\tilde{x}(k|k)-(G-KHG)\mu(k) - Kv(k+1) \qquad (4.22)$$

In this case it will not be possible to drive the error to zero. The gain K will be chosen to yield a sufficiently small trace of the steady state covariance matrix, however.

The characteristic polynomial of $(F-KHF)$ can be written in factored form as

$$\chi(F-KHF) = (s-d_1)(s-d_2) \ldots (s-d_n) . \tag{4.23}$$

Theorems 4.1 and 4.2 show that for a completely observable system with a cyclic F matrix the gain K can be chosen to yield any real set of d_i, with the restriction that $d_n = 0$ if F is singular. Suppose the d_i are chosen real and distinct (with $d_n = 0$ for F singular) and an appropriate K is determined. Then the eigenvectors p_i of $(F-KHF)$ can be found and the matrix $P \triangleq [p_1 \; p_2 \; \ldots \; p_n]$ will satisfy the condition

$$P^{-1}(F-KHF)P = \begin{bmatrix} d_1 & & & \\ & d_2 & & \text{O} \\ & & \ddots & \\ \text{O} & & & d_n \end{bmatrix} \triangleq \Delta . \tag{4.24}$$

If the transformation $\xi(k) = P^{-1}x(k)$ is made, equations (4.1), (4.21), and (4.22) become

$$\xi(k+1) = P^{-1}FP\xi(k) + P^{-1}Gu(k)$$

$$y(k) = HP\xi(k)$$

$$\hat{\xi}(k+1|k) = P^{-1}FP\hat{\xi}(k|k) + P^{-1}Gw(k)$$

$$\hat{\xi}(k+1|k+1) = \hat{\xi}(k+1|k) + P^{-1}K[z(k+1) - HP\hat{\xi}(k+1|k)]$$

$$\tilde{\xi}(k+1|k+1) = \Delta\tilde{\xi}(k|k) - P^{-1}(G-KHG)\mu(k) - P^{-1}K \, v(k+1) . \tag{4.25}$$

Assuming $v(k)$ and $\mu(k)$ are zero mean, mutually independent random variables with variances V and U, respectively, the equation for the covariance $\Sigma_\xi(k+1|k+1) \triangleq E\{\tilde{\xi}(k+1|k+1)\tilde{\xi}^T(k+1|k+1)\}$ is

$$\Sigma_\xi(k+1|k+1) = \Delta\Sigma_\xi(k|k)\Delta^T + P^{-1}(G-KHG)U(G^T-G^TH^TK^T)P^{-1^T} + P^{-1}KVK^TP^{-1^T} . \tag{4.26}$$

If the d_i are restricted such that $|d_i| < 1$ for all i, a steady state covariance $\Sigma_\xi(\infty)$ exists which is specified by

$$\Sigma_\xi(\infty) = \Delta\Sigma_\xi(\infty)\Delta^T + P^{-1}(G-KHG)U(G^T-G^TH^TK^T)P^{-1^T} + P^{-1}KVK^TP^{-1^T} \,.$$

$$(4.27)$$

Now,

$$tr[\Sigma_\xi(\infty)] = tr[\Delta\Sigma_\xi(\infty)\Delta^T] + tr[P^{-1}(G-KHG)U(G^T-G^TH^TK^T)P^{-1^T}]$$

$$+ tr[P^{-1}KVK^TP^{-1T}]$$

and letting $d^* = \max\limits_i \{d_i\}$

$$tr[\Delta\Sigma_\xi(\infty)\Delta^T] = tr[\Delta^T\Delta\Sigma_\xi(\infty)] \leq d^{*2} \, tr[\Sigma_\xi(\infty)]$$

so that

$$tr[\Sigma_\xi(\infty)] \leq \frac{tr[P^{-1}(G-KHG)U(G^T-G^TH^TK^T)P^{-1^T}] + tr(P^{-1}KVK^TP^{-1^T})}{1-d^{*2}} \quad (4.28)$$

The covariances $\Sigma_\xi(k+1|k+1)$ and $\Sigma_x(k+1|k+1) = E\{\tilde{x}(k+1|k+1)\tilde{x}^T(k+1|k+1)\}$ are related by

$$\Sigma_x(k+1|k+1) = P\Sigma_\xi(k+1|k+1)P^T \qquad (4.29)$$

Taking the Hilbert norm of both sides of (4.29) gives

$$\|\Sigma_x(k+1|k+1)\| = \|P\Sigma_\xi(k+1|k+1)P^T\|$$

$$\leq \|P\|^2\|\Sigma_\xi(k+1|k+1)\| \qquad (4.30)$$

and since $\|A\| = \lambda_{max}(A)$, the largest eigenvalue of A, if A is symmetric and positive semidefinite equation (4.30) becomes

$$\lambda_{max}(\Sigma_x(k+1|k+1)) \leq \|P\|^2 \, \lambda_{max}(\Sigma_\xi(k+1|k+1)) \,. \qquad (4.31)$$

For a square matrix A of dimension n

59

$$tr(A) = \sum_{i=1}^{n} \lambda_i(A) \tag{4.32}$$

and if A is positive semidefinite $\lambda_i(A) \geq 0$ $\forall i$ so that

$$tr(A) \geq \lambda_{max}(A) . \tag{4.33}$$

Using (4.28), (4.31), and (4.33) it is easily seen that

$$\lambda_{max}[\Sigma_x(\infty)] \leq \|P\|^2 \frac{tr[P^{-1}(G-KHG)U(G^T-G^TH^TK^T)P^{-1^T}]+tr(P^{-1}KVK^TP^{-1^T})}{1-d*^2} \tag{4.34}$$

Now,

$$\lambda_{max}[\Sigma_x(\infty)] \geq \sigma_{x_{ii}} \qquad \forall i$$

where $\sigma_{x_{ii}}(\infty)$ is the ith diagonal element of $\Sigma_x(\infty)$ and, as such, is the steady state variance of the error in estimating the ith element of the state vector. Therefore the criterion to be used for choosing the gain K will be minimization of the bound in (4.34). The procedure for choosing K is then

1) Choose a set of d_i with adequate convergence properties, observing the constraint for singular F.

2) Find a gain K yielding the corresponding characteristic equation using the results of Theorems 4.1 or 4.2.

3) Obtain the matrix of eigenvectors, P.

4) Compute the upper bound in (4.34)

5) If this bound is satisfactorily small the gain K can be used. If the bound is too large, iteration can be performed either on steps 2)-5) or on 1)-5) until a satisfactory bound is obtained.

NOTE: For a continuous system equations (4.22), (4.16) and (4.34) become

$$\dot{\tilde{x}}(t) = (F-KH)\tilde{x}(t)-(G-KHG)\mu(t) - Kv(t) \tag{4.35}$$

$$c = M^{-1}A^{-1}(b-a) \tag{4.36}$$

$$\lambda_{max}[\Sigma_x(\infty)] \le \|P\|^2 \; \frac{tr[P^{-1}(G-KHG)U(G^T-G^TH^TK^T)P^{-1^T}]+tr(P^{-1}KVK^TP^{-1^T})}{2|d^*|} \tag{4.37}$$

where the d_i are chosen as negative real numbers (there is no restriction for F singular) and $d^* = \max_i \{d_i\}$. The p_i are eigenvectors of (F-KH) for the continuous case. The procedure given above can then be used in conjunction with (4.35), (4.36), and (4.37) to choose the gain K for a continuous observer.

4.2 Open-Closed Loop Observer

If the system is not known exactly and $[\hat{F}(k+1), \hat{G}(k+1), \hat{H}(k+1)]$ is the model used for estimation at the $k+1^{th}$ time instant, equations (4.2) become

$$\hat{x}(k+1|k) = \hat{F}(k+1)\hat{x}(k|k) + \hat{G}(k+1)w(k)$$

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + K(k+1)[z(k+1) - \hat{H}(k+1)\hat{x}(k+1|k)] \tag{4.38}$$

and the error $\tilde{x}(k+1|k+1)$ is given by

$$\begin{aligned}
\tilde{x}(k+1|k+1) = &[\hat{F}(k+1) - K(k+1)\hat{H}(k+1)\hat{F}(k+1)]\tilde{x}(k|k)-[\hat{G}(k+1)\\
& - K(k+1)\hat{H}(k+1)\hat{G}(k+1)]\mu(k) - K(k+1)v(k+1)\\
& + [\tilde{F}(k+1) - K(k+1)\hat{H}(k+1)\tilde{F}(k+1)]x(k) + [\tilde{G}(k+1)\\
& - K(k+1)\hat{H}(k+1)\tilde{G}(k+1)]u(k) - K(k+1)\tilde{H}(k+1)x(k+1) \tag{4.39}
\end{aligned}$$

where $\tilde{F}(k+1)$, $\tilde{G}(k+1)$, and $\tilde{H}(k+1)$ are the unknown errors in F, G, and H respectively. Lacking any information about these modeling errors a reasonable choice for $K(k+1)$ is the constant gain that would be used if $(\hat{F}(k+1), \hat{G}(k+1), \hat{H}(k+1))$ were the actual system. This is an open-closed loop observer. It is open loop in the sense that the gain at each instant is computed as though the model of the system would not change, but it is closed loop because the gain is recomputed at each instant to take account of changes in the model. If the model at time (k+1) is the same as that at time k then $K(k+1) = K(k)$.

Suppose the model is obtained from a consistent identifier. Then

61

$$\lim_{k \to \infty} [\hat{F}(k), \hat{G}(k), \hat{H}(k)] = (F, G, H)$$

$$\lim_{k \to \infty} [\tilde{F}(k), \tilde{G}(k), \tilde{H}(k)] = (0, 0, 0)$$

and in the limit (4.39) becomes

$$\tilde{x}(k+1|k+1) = (F - K^*HF)\tilde{x}(k|k) - (G - K^*HG)\mu(k) - K^*v(k+1) \qquad (4.41)$$

where K^* is the gain that would be computed for (F, G, H) by the procedure in section 4.1.2. The steady state covariance obtained using this open-closed loop observer in conjunction with the consistent identifier will satisfy (4.34). Note that using the consistent identifier the model does not have to be updated after each input/output measurement, but can be updated at any desired rate. Clearly, as the model converges to the true parameters the rate of updating can be decreased. Between time instants at which the model is recomputed, the gain of the estimator will remain constant. These features represent a great computational saving over schemes in which an approximate covariance matrix must be computed at each time instant.

In summary, the combination of a consistent identifier and an open-closed loop observer is recommended for solving the problem of parameter identification and state estimation because of

1)   consistent parameter identification,

2)   computational savings,

3)   bounding of the steady state covariance matrix.

The state estimation will be asymptotically efficient in that a bound will exist for the maximum mean square estimation error, and the estimation will be asymptotically optimum when K* actually minimizes the trace of the steady state covariance matrix.

5. COMPUTATIONAL RESULTS

To demonstrate the application of the estimators presented in this paper two simple examples are considered. The first example involves only the problem of parameter identification and demonstrates the convergence rates of the three consistent parameter identifiers proposed in Chapter 3. The second example presents a specific case in which the open-closed loop observer of Chapter 4 yields better results than the extended Kalman filter of Farison [4].

5.1 Example 1 - Convergence Rates of Consistent Parameter Identifiers

For comparison of the convergence rates of the consistent parameter identifiers of Chapter 3 with the convergence rate of an estimator which requires information about the initial state and the noise distribution the fourth order discrete system presented in section 4.6 of Reference [27] was considered. The system is described by the matrices

$$
F = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ a_1 & a_2 & a_3 & a_4 \end{bmatrix} \qquad G = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \qquad H = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}
$$

with $a_1 = -0.656$, $a_2 = 0.784$, $a_3 = -0.18$, and $a_4 = 1.0$.

As pointed out in [27] the equivalent z-transfer function is

$$
G(z) = \frac{z^3}{(z^2 - 1.8z + 0.8)(z^2 + 0.8z + 0.8)}
$$

which, except for the numerator, could be considered as a hypothetical missile with the short period and first bending mode included.

The problem is to identify the parameters a_1, a_2, a_3, and a_4. Since F is nonsingular each of the identification procedures of Chapter 3 can be applied. Computations were performed using an IBM 360 computer. The initial state of the system was chosen as zero in all cases. Tables 5.1, 5.2, and 5.3 show the results of using the identifiers of

Table 5.1

Number of Samples	S/N Ratio	$\hat{a}_1$	$\hat{a}_2$	$\hat{a}_3$	$\hat{a}_4$
10	5	-.61871	.74622	-.13465	.94386
20		-.62132	.71972	-.14975	.98330
40		-.67475	.76967	-.14042	.98806
70		-.65777	.78078	-.17209	1.00620
100		-.67117	.79758	-.18008	1.00598
150		-.65796	.76809	-.17575	1.00331
200		-.65833	.77732	-.16900	.99629
10	2	-.54664	.67158	-.06093	.86558
20		-.55864	.61782	-.09792	.96476
40		-.68416	.74600	-.10354	.98236
70		-.67717	.74593	-.15180	1.01122
100		-.68760	.81193	-.17961	1.01341
150		-.65236	.73935	-.15039	.99760
200		-.66162	.76870	-.15184	.98993
10	1	-.35733	.46797	.09088	.73543
20		-.43247	.43566	-.00937	.94280
40		-.68564	.70826	-.05821	.98170
70		-.72002	.69083	-.06974	.98933
100		-.70507	.82560	-.17830	1.02343
150		-.63477	.68840	-.09717	.98301
200		-.66632	.75542	-.12210	.97844
1	∞	-.65600	.78399	-.18000	1.00001

Table 5.2

Number of Samples	S/N Ratio	$\hat{a}_1$	$\hat{a}_2$	$\hat{a}_3$	$\hat{a}_4$
10	5	-.75381	.98018	-.36200	1.06755
20		-.71687	.88474	-.17404	.94500
40		-.60332	.69874	-.12968	.97300
70		-.60682	.70240	-.13137	.97772
100		-.61077	.72178	-.17683	1.01035
150		-.63216	.74779	-.16420	.99449
200		-.63889	.76582	-.18761	1.00712
10	2	-.93273	.77086	.59831	.42505
20		-.09601	-.52864	1.29846	.22615
40		-.49969	.50117	.04100	.88435
70		-.50574	.51703	.01068	.91039
100		-.51407	.56738	-.11166	.99726
150		-.59100	.68052	-.12901	.98181
200		-.59062	.67612	-.12112	.97907
10	1	-.45403	-.11971	.62848	.75416
20		.16403	-.12263	-.68309	1.54398
40		-1.20981	1.22876	1.07037	-.23569
70		-.24380	-.12720	.73067	.54325
100		-.24617	.06768	.23425	.86865
150		-.43011	.35667	.13178	.87719
200		-.43971	.33459	.22726	.81612
1	∞	-.65625	.78540	-.18214	1.00084

Table 5.3

Number of Samples	S/N Ratio	$\hat{a}_1$	$\hat{a}_2$	$\hat{a}_3$	$\hat{a}_4$
10	5	-.69009	.86722	-.26875	1.03464
20		-.69853	.89274	-.24852	.99675
40		-.60864	.64691	-.05688	.96387
70		-.63072	.66182	-.04608	.96086
100		-.64618	.72998	-.13518	.99886
150		-.65741	.76717	-.17021	1.00722
200		-.65375	.76782	-.18405	1.01648
10	2	-.82304	1.08218	-.37869	1.05067
20		-.92197	1.35181	-.49124	.99950
40		-.48903	.36778	.14923	.91866
70		-.56576	.44138	.14693	.92145
100		-.58957	.57003	-.03333	.99798
150		-.62833	.66540	-.11423	1.02252
200		-.62006	.67268	-.16364	1.06056
10	1	-1.20331	1.60579	-.45515	.95444
20		2.92358	-5.22095	1.62671	1.65619
40		-.29253	-.11536	.57866	.77359
70		-.47436	.10945	.47281	.83593
100		-.49057	.25693	.22080	.95620
150		-.54094	.35681	.08825	1.04121
200		-.52029	.32918	.01751	1.12331
1	∞	-.65918	.78979	-.18506	1.00146

66

sections 3.2.2.a, 3.2.2.b, and 3.2.2.c, respectively, for various signal
to noise ratios. In figures 5.1, 5.2, and 5.3 the normalized errors

$$\frac{\|\phi - \hat{\phi}\|}{\|\phi\|} = \frac{\sum\limits_{i}(a_i - \hat{a}_i)^2}{\sum\limits_{i} a_i^2}$$

associated with these tables are plotted. In each figure it is apparent
that the larger the signal to noise ratio the faster the convergence.
The convergence rates obtained compare favorably with those in Reference
[27].

5.2 Example 2 - Open-Closed Loop Observer

To study the operation of the open-closed loop observer the simple
scalar system

$$x(k+1) = ax(k)$$

$$z(k) = x(k) + v(k)$$

was considered. The statistics used in simulations on the IBM 360
computer were $\hat{x}(0|0) = 10.0$, $\hat{a}(0) = 1.2$, $\sigma_x(0|0) = .5$, $\sigma_a(0) = .1$,
$E\{v(i)\} = 0$, and $\sigma_v = 10.0$ with the distributions of the initial
state, of the parameter a, and of the noise being Gaussian (Gaussian
statistics were chosen so that comparison with Farison's solution would
be possible). Note that $\sigma_a(0)\sigma_x(0|0) \ll \hat{a}(0)\hat{x}(0|0)$ so that Farison's
result should be applicable — it is not a requirement of the open-closed
loop observer. The gain of the open-closed loop observer was not chosen
as zero when $\hat{a}_N < 1$, but was chosen to yield a small steady state
variance in the state estimate. The parameter identifier of the open-
closed loop observer was started after three observations so that initial
noise in the observations would be less important. In general the perform-
ance of the observer of Chapter 4 and Farison's observer were comparable;
however, in several instances use of Farison's filter results in
erroneous estimates. In figures 5.4 and 5.5 results of a particular
computer run are presented which shows the poorer performance of the
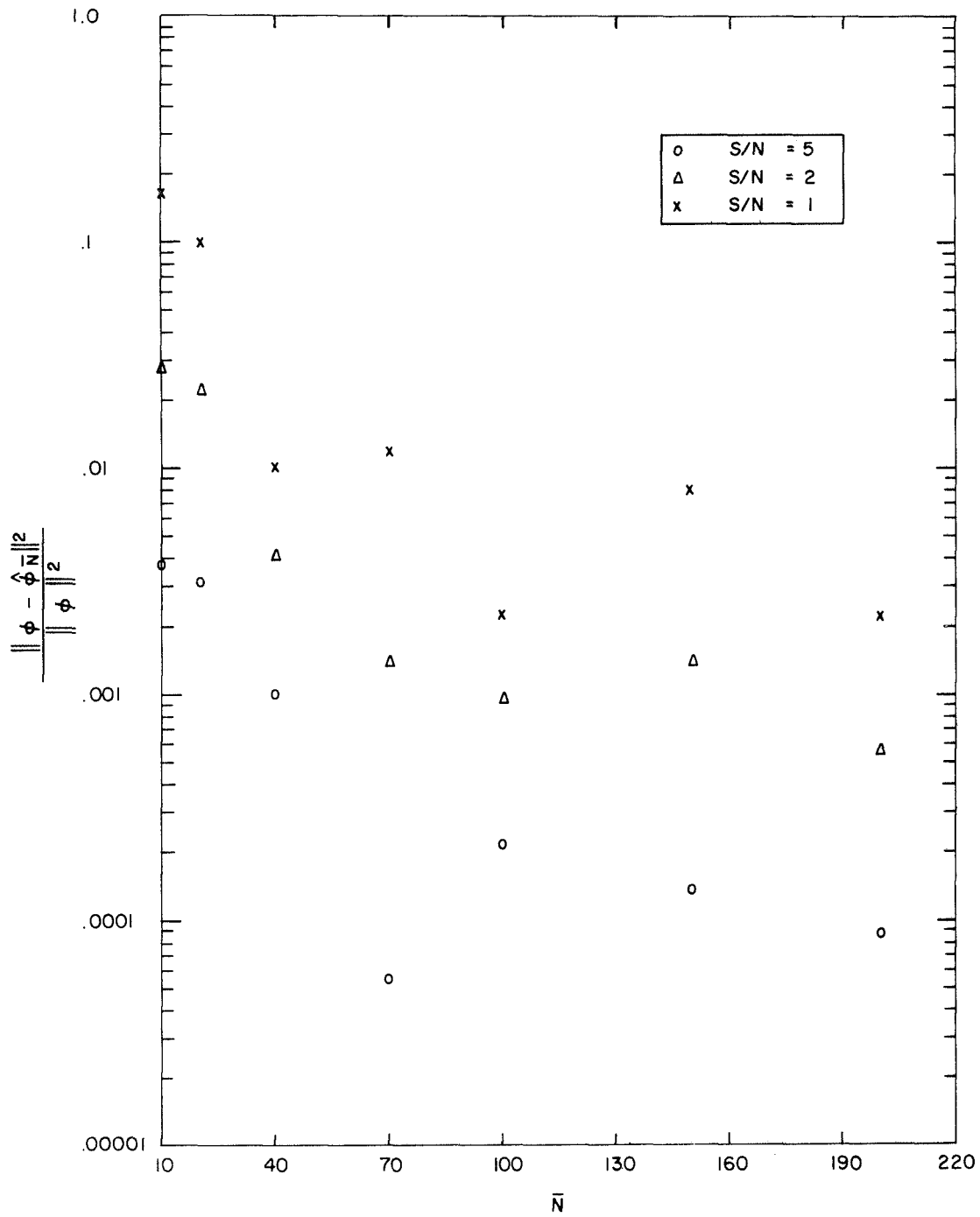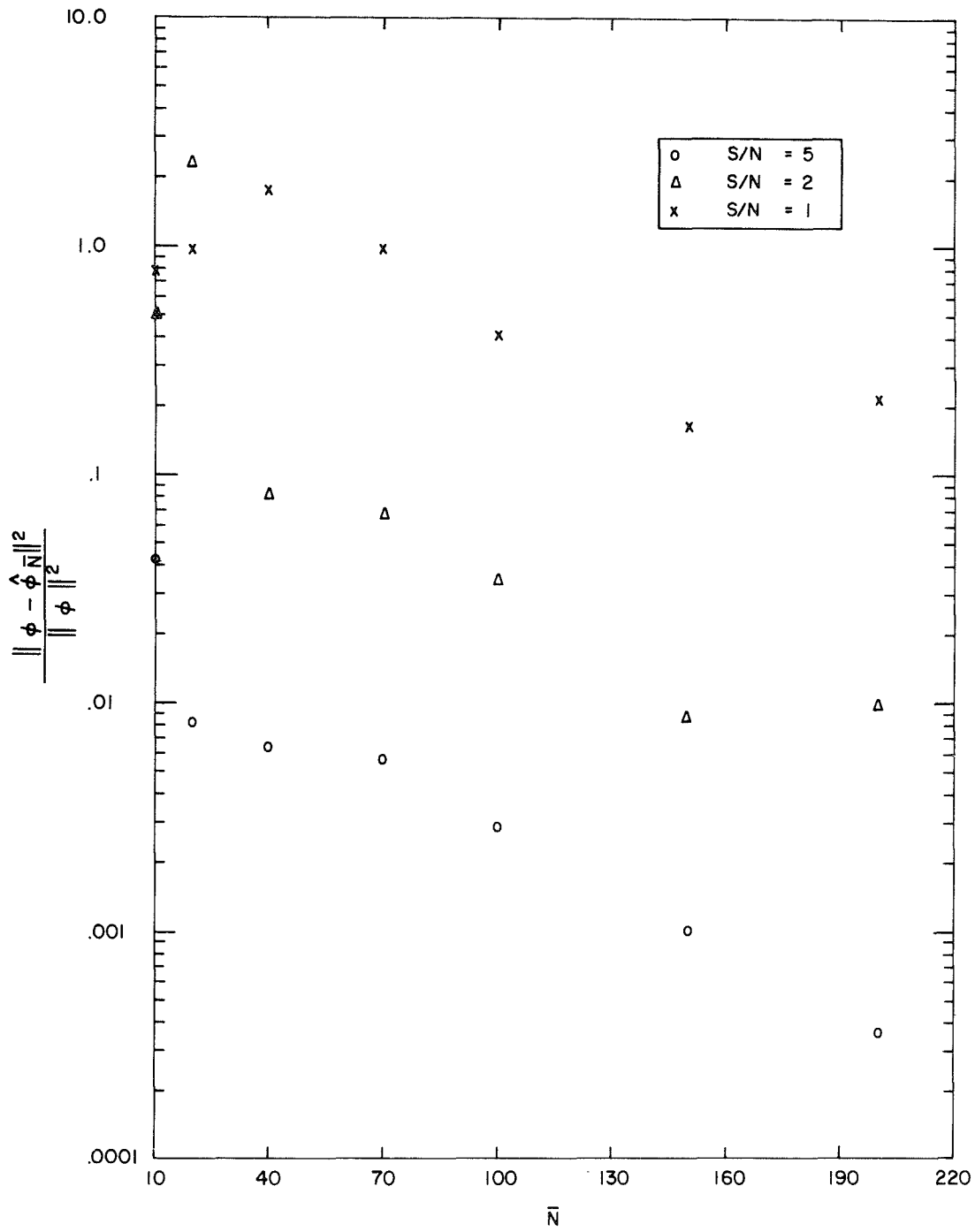Farison filter — the data is presented in table 5.4.

Figure 5.1

68

Figure 5.2
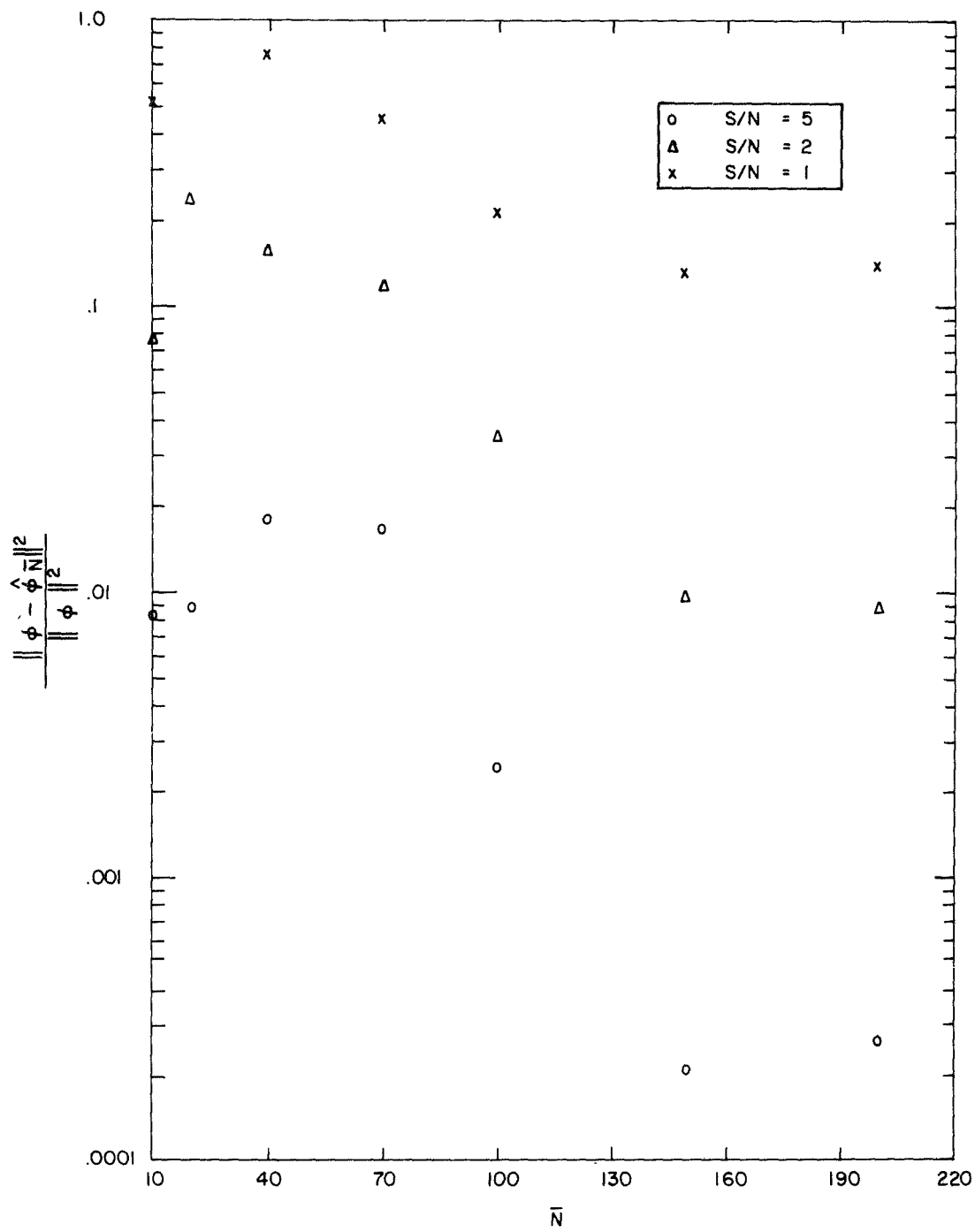
Figure 5.3

The failure of the Farison filter may be attributed to the high
noise level and the fact that there is no driving noise to keep the
filter gain from going to zero; but these conditions are not precluded
in the general theory of the extended Kalman filter. Further study of
the relative advantages of the extended filter and the open-closed
loop observer seem desirable.

Table 5.4 [1]

N	a	$\hat{a}_{N_C}$	$\hat{a}_{N_F}$	x(N)	$\hat{x}(N/N)_{OCL}$	$\hat{x}(N/N)_F$
0	1.0388	1.2000	1.2000	10.221	10.000	10.000
1		1.2000	.0980	10.624	5.023	.186
2		1.2000	.1647	11.035	7.795	.099
3		.2303	.1870	11.463	3.161	.022
4		.6142	.1928	11.908	3.245	.005
5		.3911	.1926	12.369	1.008	.001
6		.3426	.1933	12.849	2.987	.000
7		.5934	.1934	13.347	3.806	
8		.7063	.1934	13.864	4.638	
9		.7883		14.402	5.690	
10		.7962		14.960	6.032	
11		.7907		15.540	5.715	
12		.8149		16.142	5.850	
13		.8908		16.768	7.700	
14		.8927		17.418	8.869	
15		.8889		18.093	9.380	
16		.9117		18.795	10.272	
17		.9144		19.523	10.834	
18		.9338		20.280	11.818	
19		.9885		21.066	14.583	
20		.8988		21.883	13.715	

[1]

$\hat{a}_{N_C}$ = consistent estimate of a

$\hat{a}_{N_F}$ = estimate of a using the Farison filter

$\hat{x}(N/N)_{OCL}$ = open-closed loop estimate of x(N) from N observations

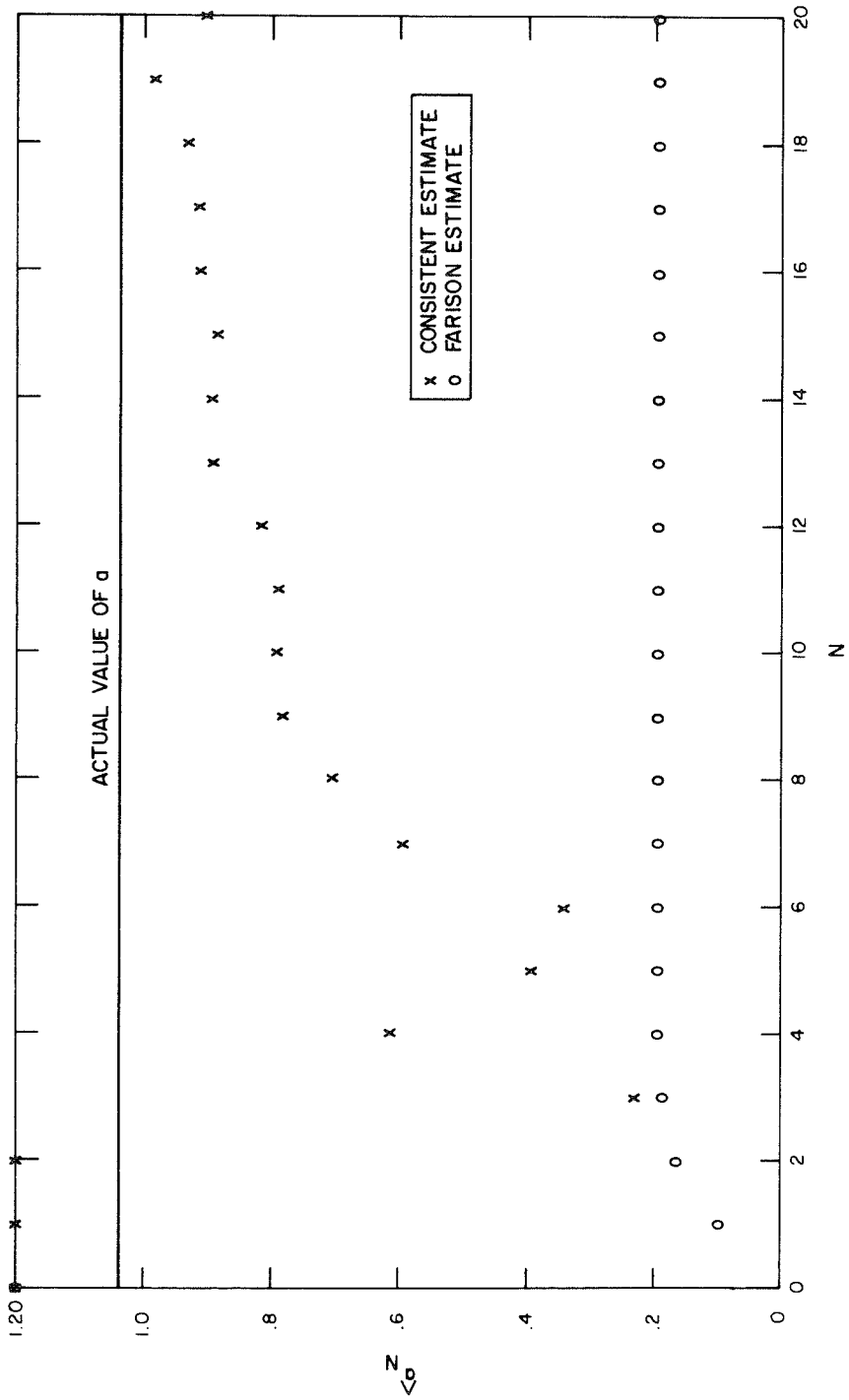$\hat{x}(N/N)_F$ = estimate of x(N) from N observations using the Farison filter

72

Figure 5.4

73

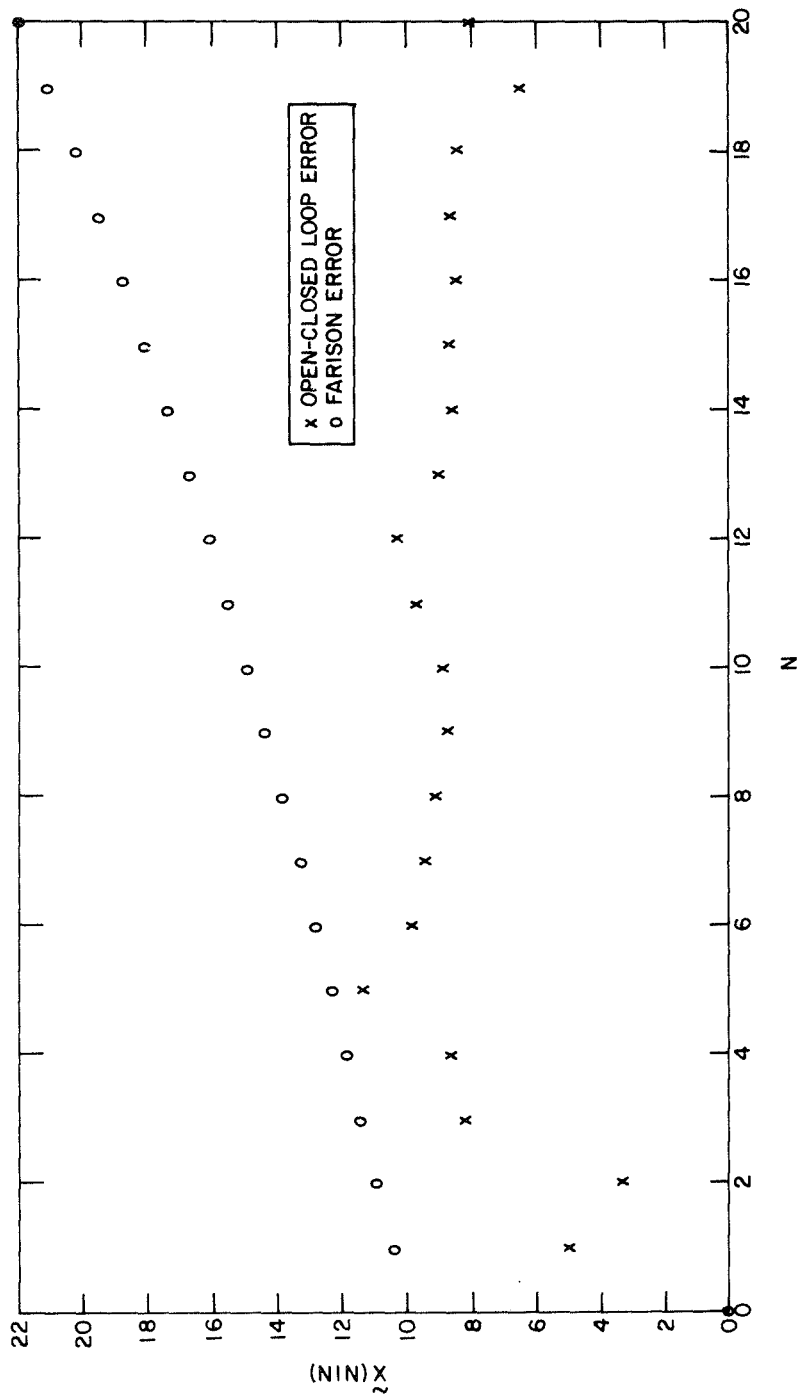x OPEN-CLOSED LOOP ERROR
o FARISON ERROR

N

$\tilde{X}(N|N)$

Figure 5.5

6. CONCLUSIONS

6.1 Summary

A solution having desirable asymptotic behavior has been obtained for the problem of simultaneously estimating the parameters and states of linear systems. This solution has been obtained by separating the inherently nonlinear problem into two parts, a parameter estimator which does not depend on estimates of the state and a state estimator whose dynamics can be chosen to yield a satisfactory bound on the steady state mean square error when the system model is known exactly.

By analyzing the structure of minimal systems it was possible to obtain a direct procedure for determining the minimal realizations of a linear system from noise free input/output observations. Three procedures, depending on different assumptions about the system structure and the noise, were then presented which yield consistent estimates of the parameters when the input/output observations are corrupted by noise. These estimators do not require state estimation or recursive calculation of any parameter associated covariance matrix for implementation.

The advantages of using constant gain observers are important— they do not require knowledge of the initial mean or variance of the state; they are less sensitive to modeling errors; and they give a significant computational saving. When estimating the states of a minimal system it is possible to choose the gain of the constant gain observer to give prescribed error dynamics. A procedure for choosing this constant gain has been established, which gives an upper bound for the steady state mean square estimation error if the system model is exact.

When the gain of the observer is changed each time the model is changed the observer is termed open-closed loop. If the model is obtained from a consistent identifier the open-closed loop observer will be asymptotically efficient — the steady state bound computed for the mean square error will be correct.

6.2 Suggestions for Future Research

Some areas where further investigation might prove fruitful are the following:

1) Determine the effect of the computer tolerance ϵ on the identification procedure. In particular investigate the possibility of obtaining the linearized model of a nonlinear system which is operating in a nearly linear region by appropriately choosing ϵ.

2) Study the effects of various choices of the times t_i to be used in the linear operator for continuous systems.

3) Evaluate the performance of a controller which uses the output of the combined parameter and state estimator developed in this dissertation.

4) Study the effects of changes in the rate at which the model is updated.

5) Investigate the performance of the on-line identifiers when the plant dynamics are slowly varying.

REFERENCES

[1] A. Papoulis, <u>Probability, Random Variables, and Stochastic Processes</u>, McGraw-Hill Book Co., Inc., New York, 1965.

[2] J. H. Laning and R. H. Battin, <u>Random Processes in Automatic Control</u>, McGraw-Hill Book Co., Inc., New York, 1956.

[3] T. Magill, "Optimal Adaptive Estimation of Sampled Stochastic Processes", Ph.D. Dissertation, Stanford University, Stanford, Calif., Jan. 1964.

[4] J. B. Farison, "Identification and Control of Random Parameter, Discrete Systems", Ph.D. Dissertation, Stanford University, Stanford, Calif., Feb. 1964.

[5] J. B. Farison, R. E. Graham and R. C. Shelton Jr., "Identification and Control of Linear Discrete Systems", Trans. IEEE (Automatic Control), AC-12, Aug. 1967, pp. 438-442.

[6] H. E. Rauch, "Linear Estimation of Sampled Stochastic Processes with Random Parameters", Ph.D. Dissertation, Stanford University, Stanford, Calif., Oct. 1962.

[7] R. Bellman and R. Kalaba, <u>Dynamic Programming and Modern Control Theory</u>, Academic Press, New York, 1965.

[8] W. J. Murphy, "Optimal Stochastic Control of Discrete Linear Systems with Unknown Gain", Joint Automatic Control Conference of the AACC, University of Michigan Press, Ann Arbor, Michigan, June 1968.

[9] R. B. McGhee and R. B. Walford, "A Monte Carlo Approach to the Evaluation of Conditional Expectational Parameter Estimates for Nonlinear Dynamic Systems", Trans. IEEE (Automatic Control), AC-13, Feb. 1968.

[10] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems", Trans. ASME, 82D, (<u>J. Basic Engr.</u>), Mar. 1960, pp. 35-45.

[11] R. E. Kopp and R. J. Orford, "Linear Regression Applied to System Identification for Adaptive Control Systems", <u>AIAA Journal</u>, Oct. 1963, pp. 2300-2306.

[12] F. H. Schlee, C. J. Standish, and N. F. Toda, "Divergence in
the Kalman Filter", *AIAA Journal*, Vol. 5, No. 6, June 1967.

[13] S. F. Schmidt, "Estimation of State With Acceptable Accuracy
Constraints", Report 67-4, Analytical Mechanics Associates, Inc.,
Westbury, New York, Jan. 1967.

[14] S. F. Schmidt, "Compensation for Modeling Errors in Orbit
Determination Problems", Report 67-16 Analytical Mechanics
Associates, Westbury, N. Y., Nov. 1967.

[15] B. Gopinath, "On the Identification and Control of Linear Systems",
Ph.D. Dissertation, Stanford University, Stanford, Calif., June 1968.

[16] T. L. Gunkel and G. Franklin, "A General Solution for Linear,
Sampled Data Control", Trans. ASME, Vol. 85, Series D, June 1963.

[17] W. B. Davenport and W. L. Root, *An Introduction to the Theory of
Random Signals and Noise*, McGraw-Hill Book Co., Inc., New York,
1958.

[18] R. E. Kalman and R. S. Bucy, "New Results in Linear Filtering and
Prediction Theory", Trans. ASME, 83D, (*J. Basic Engr.*), Mar. 1961,
pp. 95-108.

[19] H. Cox, "On the Estimation of State Variables and Parameters for
Noisy Dynamic Systems", Trans. IEEE (Automatic Control), AC-9,
Jan. 1964, pp. 5-12.

[20] M. J. Levin, "Estimation of System Pulse Transfer Function in the
Presence of Noise", IEEE Transactions on Automatic Control, Vol.
AC-9, No. 3, July 1964, pp. 229-35.

[21] G. N. Saridis and G. Stein, "Stochastic Approximation Algorithms for
Linear Discrete-Time System Identification", IEEE Transactions on
Automatic Control, Vol. AC-13, No. 5, Oct. 1968, pp. 515-523.

[22] S. R. Neal, "Linear Estimation in the Presence of Errors in
Assumed Plant Dynamics", IEEE Transactions on Automatic Control,
Vol. AC-12, No. 5, Oct. 1967, pp. 592-594.

[23] S. L. Fagin, "Recursive Linear Regression Theory, Optimal Filter Theory, and Error Analysis of Optimal Systems", 1964 IEEE Conv. Rec., pp. 216-240.

[24] C. R. Rao, <u>Linear Statistical Inference and Its Applications</u>, John Wiley and Sons, 1965.

[25] Y. W. Lee, <u>Statistical Theory of Communication</u>, John Wiley and Sons, 1960.

[26] K. Y. Wong and E. Polak, "Identification of Linear Discrete Time Systems Using the Instrumental Variable Method", IEEE Transactions on Automatic Control, Vol. AC-12, No. 6, Dec. 1967, pp. 707-18.

[27] R. C. K. Lee, <u>Optimal Estimation, Identification, and Control</u>, The M.I.T. Press, 1964.