

STATISTICAL METHODS FOR THE
IDENTIFICATION AND CONTROL OF
MULTIVARIATE STOCHASTIC SYSTEMS

Ian Henry Rowe

A thesis submitted for the
degree of Doctor of Philosophy

November 1967

Centre for Computing and Automation,
Imperial College of Science and Technology,
University of London.

ABSTRACT

In this thesis, contributions are made to the general problem of identifying linear, multivariate systems (with more than one input and more than one output) from operating data for purposes of control. The problem is shown to be the proposal of a suitable mathematical model structure followed by the task of estimating parameters within the structure by statistical methods.

From the initial assumption that the systems of interest have a state-space description, deterministic canonical forms having fewer parameters than the original description are derived by means of linear transformations. The methods for construction of these transforms are related to the conditions for controllability and observability. The method for transforming an important new form, the A-canonical form, into a vector differential (or difference) equation in the input and output variables only is developed. The inverse problem, of transforming back into the A-canonical form is solved. Then both descriptions are extended to include structure for stochastic inputs that appear as correlated disturbances at the system output.

The problem of estimating the parameters of the multivariate stochastic difference equation is investigated. It is found that estimations by conventional least squares leads to unsuitable estimates. Consequently, a new, modified method of least squares is developed and shown to be asymptotically unbiased. Algorithms for estimation by maximum likelihood

methods are also developed. A number of examples are computed to illustrate the effectiveness of the algorithms and for comparison purposes. A simple example of stochastic control using a model, the parameters of which are estimated by running data, is shown.

In the final chapter, some possible extensions of the work are considered.

ACKNOWLEDGEMENTS

I am deeply indebted to my supervisor, Dr. D. Q. Mayne, for his encouragement, advice and constructive criticism relating to this research work. I would also like to thank Professor J. H. Westcott for his interest.

I wish to mention Professor G. R. Slemon and Dr. P. A. Lapp, whose enthusiasm and encouragement influenced me in the direction of postgraduate research.

I have benefitted considerably from useful discussions with other members of the Control Group. In particular, I should like to thank John Allwright, Stewart Cumming, Dr. David Jacobson, David Norris and Paul Turley.

This work would not have been possible without the financial support of an Athlone Fellowship, a Special Scholarship from the National Research Council of Canada, and a Management Committee Award from the de Havilland Aircraft of Canada, Limited.

Finally, I am most indebted to my wife, Shirley, for her understanding and encouragement.

I. H. Rowe

INDEX

STATISTICAL METHODS FOR THE IDENTIFICATION AND CONTROL
OF MULTIVARIATE STOCHASTIC SYSTEMS

Abstract	. . .	2
Acknowledgements	. . .	4
1. THE SYSTEM IDENTIFICATION PROBLEM	. . .	10
1.1 Introduction	. . .	10
1.2 Process Identification	. . .	15
1.3 Outline of the Thesis	. . .	18
1.4 Contributions of the Thesis	. . .	20
2. SOME STATE-VARIABLE CANONICAL FORMS OF A CONSTANT LINEAR, MULTIVARIATE, DETERMINISTIC SYSTEM	. . .	22
2.1 Introduction	. . .	22
2.2 The Minimum Number of Parameters in a State- Variable Canonical Form	. . .	24
1. The Problem	. . .	24
2. Transformation to a Minimum-Parameter Normal Form	. . .	25
2.3 The Conditions for Controllability and Observability of a Linear, Constant, Multivariate System	. . .	28
2.4 A Useful Aid to the Derivation of Canonical Forms	. . .	32
1. The Lemma	. . .	33
2.5 Multivariate State-Variable Canonical Forms Defined with the Conditions for Controllability	. . .	35
1. Companion Matrix (Transposed) Canonical Form	. . .	36
2. An Assertion on the Realizability of Minimum-Parameter Canonical Descriptions	. . .	39
3. Another Reduced Form	. . .	40

2.6	The A-Canonical Form Derived Using the Conditions for Observability	. . .	43
	1. The Transformation	. . .	44
	2. Algebraic Examples	. . .	47
3.	TRANSFORMATION BETWEEN A STATE-VARIABLE DESCRIPTION AND A DESCRIPTION IN THE INPUT AND OUTPUT VARIABLES	. . .	51
3.1	Introduction	. . .	51
3.2	The Transformation to Vector Differential (or Difference) Equation Description Σ_D	. . .	52
	1. The Computational Procedure	. . .	53
	2. Extension to Include the Vector Difference Equation Form	. . .	55
	3. A Numerical Example	. . .	56
3.3	Derivation of the Transformation from Description Σ to Σ_D	. . .	59
3.4	Modifications to Include Direct Control of the System Output	. . .	66
	1. Complete State-Variable Description	. . .	66
	2. Complete Description Σ_D with Direct Control	. . .	67
	3. A Numerical Example	. . .	68
3.5	Discussion of the Σ_D Description	. . .	71
	1. Historical Perspective	. . .	71
	2. Forms of the Σ_D Description	. . .	73
	3. Comments on the Existence and Uniqueness of the Description	. . .	76
3.6	Transformation of the Σ_D Description into State-Variable Formulation	. . .	80
	1. The Procedure for Transformation into State-Variable Form	. . .	81
	2. An Algebraic Example	. . .	85
3.7	Matrix Transfer Function Description Σ^z	. . .	89
	1. Derivation of the Transfer Function ^z	. . .	89
	2. Algebraic Example of a Matrix Transfer Function	. . .	91
	3. The Transfer Function Description as the Laplace Transform of a Vector Differential Equation of Order n	. . .	92
	4. Calculating the Matrix Transfer Function Corresponding to $\Sigma(F, G, H, D)$. . .	95
3.8	The Problem of Parametrization	. . .	97
	1. A Comparison of Some Descriptions	. . .	97
	2. An Algebraic and Geometric Interpretation of the Problem	. . .	99

4.	MULTIVARIATE STOCHASTIC MODELS FOR IDENTIFICATION, PREDICTION AND CONTROL	. . . 102
4.1	Introduction	. . . 102
4.2	A General Discrete-Time State-Variable Model of a Stochastic Process	. . . 103
4.3	Transformation to Stochastic Vector Difference Equation Form	. . . 109
	1. Introduction	. . . 109
	2. The Stochastic Difference Equation Form in Both Noise Variables	. . . 112
	3. Description Σ_{D_s}	. . . 115
	4. Some Special Cases	. . . 122
4.4	Transformation from the Σ_{D_s} Description to State-Variable Form	. . . 124
5.	STATISTICAL METHODS OF PARAMETER ESTIMATION	. . . 127
5.1	Introduction and Statement of the Problem	. . . 127
5.2	Linear Least Squares Estimator	. . . 131
	1. Linear Regression Model	. . . 131
	2. The LSE Algorithm	. . . 134
	3. Limitations of the LSE	. . . 140
5.3	Sequential Least Squares Estimator	. . . 143
5.4	Generalized Least Squares Estimator	. . . 146
	1. The Classical Viewpoint	. . . 146
	2. The Filtering Viewpoint ($m = 1$)	. . . 147
	3. GLSE When the Noise Process is Unidentified	. . . 150
	4. GLSE in the Vector Observation Case ($m > 1$)	. . . 150
	5. Filter Stability	. . . 153
	6. The Estimating Equation	. . . 164
	7. The Estimator for \wedge	. . . 164
5.5	Modified Least Squares Estimators	. . . 166
	1. Introduction	. . . 166
	2. The Rucker-Levadi Method	. . . 167
	3. Mayne's Method	. . . 169
	4. Bootstrap Least Squares Estimator	. . . 171
5.6	Maximum Likelihood Estimator	. . . 180
	1. Introduction	. . . 180
	2. Formulation of the MLE Problem	. . . 180
	3. Iterative Solution of the MLE Problem	. . . 189
	4. An Alternative Solution of the MLE Problem	. . . 193
	5. The Adaptive Filter	. . . 194
	6. Discussion of the MLE Method	. . . 196

6.	EXAMPLES OF IDENTIFICATION, PREDICTION AND CONTROL	. . .	200
6.1	Introduction	. . .	200
6.2	A Comparison of Some Sequential Least Squares Estimators of a Parameter of a First Order System	. . .	201
	1. Purpose and Procedure	. . .	201
	2. Observations and Discussion	. . .	202
6.3	The Sequential Estimation of the Parameters of a Second Order System with the "Bootstrap" Method	. . .	207
	1. Purpose and Procedure	. . .	207
	2. Observations and Discussion	. . .	209
6.4	A Comparison of the Estimation of a Second Order System by the "Bootstrap" and MLE Methods	. . .	213
	1. Purpose and Procedure	. . .	213
	2. Observations and Discussion	. . .	214
6.5	The Estimation of Parameters of Systems with Zeros of the Noise Process Outside the Unit Circle	. . .	224
	1. Statement of the Problem	. . .	224
	2. System with One Zero Outside the Unit Circle	. . .	225
	3. Observations and Discussion	. . .	225
	4. System with Two Complex Zeros Outside the Unit Circle	. . .	227
	5. Observations and Discussion	. . .	229
6.6	The Estimation of the Matrix of Second Derivatives of the Likelihood Function	. . .	232
	1. Purpose and Procedure	. . .	232
	2. Discussion	. . .	233
6.7	Testing the Order of the Model	. . .	237
	1. Purpose and Procedure	. . .	237
	2. Observations and Discussion	. . .	237
6.8	The Estimation of Parameters of a Multivariate System	. . .	240
	1. Statement of the Problem	. . .	240
	2. Estimation Procedure	. . .	242
	3. Observations and Discussion	. . .	244
6.9	Regulation of a Stochastic System	. . .	252
	1. Minimum Variance Control of the Output	. . .	252
	2. An Example	. . .	254
	3. State Control	. . .	258

7.	DISCUSSION AND EXTENSIONS	. . .260
7.1	Introduction	. . .260
7.2	Canonical Forms of Deterministic Systems	. . .260
7.3	Stochastic Models	. . .261
	1. Generalizing the Models	. . .261
	2. Applications of the Models	. . .262
7.4	Multivariate System Identification	. . .263
	1. The Problem of Choosing a Linear Structure	. . .263
	2. Simplifications	. . .265
	3. Spectral Factorization	. . .267
	4. Estimation of the Parameters of the Σ_s Description	. . .267
	References	. . .270

CHAPTER 1

THE SYSTEM IDENTIFICATION PROBLEM

1.1 Introduction

In this thesis, the general problem of identifying multivariable systems from operating data is considered. The types of variables involved in the processes of interest are shown in Figure 1.1 which has been adapted from [1].

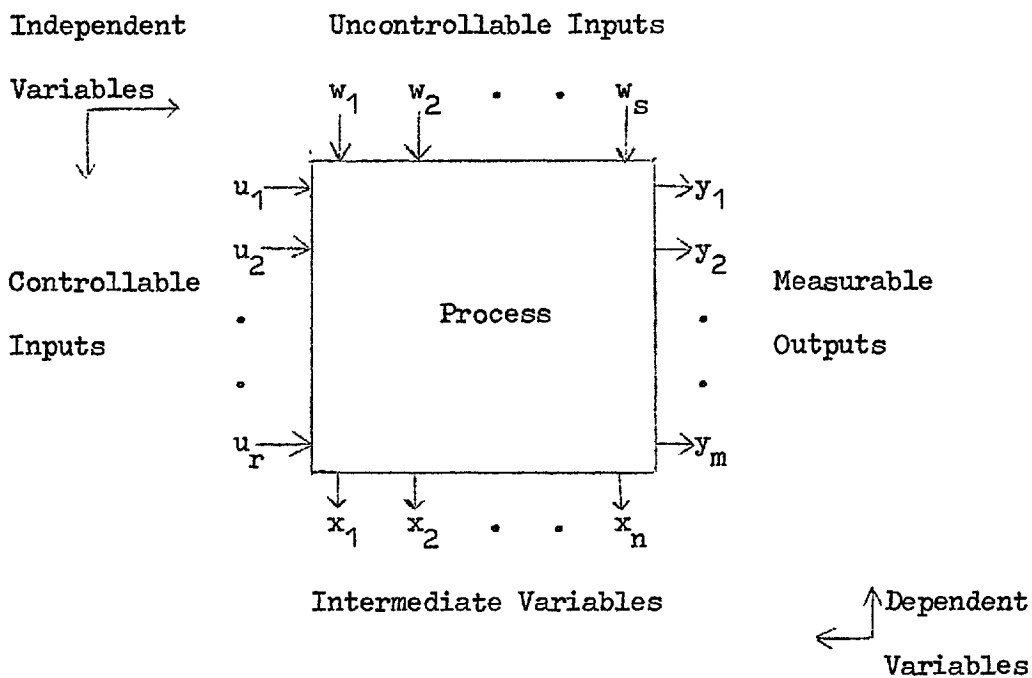


Figure 1.1 A Multivariable System

The identification problem is concerned with finding mathematical relations between the set of independent variables and the set of

dependent variables of the system.

The controllable inputs can be measured and manipulated whereas the uncontrollable inputs are disturbances which may or may not be measurable. The latter may also include changes originating within the system. In the following, distinction will be made between random processes, which cannot be controlled and control systems which have controllable inputs but may also have uncontrollable inputs.

The set of dependent variables includes, along with the measurable outputs, other convenient quantities that are part of the process but, possibly, cannot be measured. They are classed as intermediate variables.

The motivation for the identification of a control system is provided by the requirement to control the process in some optimal fashion. The implementation of control algorithms derived from modern control theory requires a knowledge of both the structure and parameters of the differential or difference equations describing the system. With this knowledge, and in the absence of disturbances, the optimal input from a class of inputs usually can be determined to achieve a desired response. In real situations, this objective may be obscured by the presence of disturbances that do not belong to the class of measurable inputs and affect the output in an undesirable manner. The control policy in such circumstances must then be modified to cope with the disturbances and minimize their influence in some sense. If the response to the uncontrolled inputs can be predicted, an improved control procedure can be devised; hence, the identification procedure must also determine the structure and parameters of the mathematical

relations connecting the uncontrollable inputs and the dependent variables of interest.

The identification of a random, uncontrollable process is motivated by the requirement to predict future behaviour given a recent history of past performance. Such processes may be encountered in economic planning situations as well as in the study of physical phenomena.

So far, we have implied that the task of system identification is undertaken in order to furnish an investigator with a mathematical description or model of the system for purposes of control or prediction. It is overambitious to suppose that an exact set of relations among the variables of the process of interest can be synthesized from a set of terminal measurements; hence, practical considerations force us to first propose a suitable mathematical model of the system specifying its structure (the functional relations among the dependent and independent variables) and then an estimation procedure to assign numerical values to the parameters in the model by experiment. Clearly, there is a risk involved in exercising judgement in proposing a model structure along with a parameter estimation procedure. Hence, a performance index is an essential element of any identification scheme as a measure of the "goodness" of a particular identification policy for which a model structure and an estimation procedure are assumed. However, there is no advantage in proposing an elegant model and a sophisticated experimental parameter estimation programme if the results cannot be fully utilized. Thus, it must be concluded that the

identification problem or model building task is strongly object-oriented [2] and the techniques employed will differ widely over the broad range of objects or systems investigated.

Let us now pose the identification problem in a very general mathematical form. We are required to estimate the elements in the vector of parameters $\Theta(t)$ appearing in the vector differential equation and observation equation

$$\dot{x}(t) = f[x(t), u(t), v_1(t), \Theta_f(t)] \quad (1.1-1)$$

$$y(t) = g[x(t), u(t), \Theta_g(t)] \\ + h[x(t), v_2(t), \Theta_h(t)] \quad (1.1-2)$$

which minimizes some cost or risk functional R where

$x(t)$ is the model state vector of dimension n

$u(t)$ is the r -dimensional control vector

$y(t)$ is an m -dimensional observation vector

$v_1(t)$ is an input noise vector

$v_2(t)$ is a measurement noise vector.

The model state vector $x(t)$ is equivalent to the vector of intermediate variables x of the general process shown in Figure 1.1 if the above equations are an exact description of the process of interest.

The major difficulty with this model is that it requires knowledge of the state-variable vector $x(t)$ in order that the parameter vector $\Theta(t)$ can be estimated. Thus, for this model, the identification

problem must be extended to include state estimation along with parameter estimation [2, 65]. The solution to this problem is difficult because it involves the joint estimation of parameters and variables appearing as products in the model. This is a nonlinear estimation problem. Unless the minimization of the risk function R specifically includes the requirement that an estimate be made of the process intermediate variables, it would be advisable to seek another model form in which the state variables have been eliminated. It is to this restricted class of identification problems that we now turn our attention.

The problem with which we shall presently be concerned is that of finding mathematical relations between the set of independent variables and the set of output variables of some noise-perturbed multivariate control systems or stochastic systems. We shall limit our main discussion to linear realizations with parameters that are invariant with time but will indicate how these constraints can be relaxed in some special circumstances.

For the purpose of this thesis, we shall define multivariate systems as processes for which, at any instant of time, it is necessary to represent the input and output variables by vector quantities of at least two dimensions respectively. The reference to single-input single-output systems which will be encountered throughout is self-explanatory.

The term stochastic process will be applied to the family of real time functions $f(t, \xi)$ defined for the outcomes ξ of an experiment

specified by the probabilities of the random events that make up ξ_1 . For every specific outcome ξ_{1i} , $f(t, \xi_{1i})$ is a single time function which, for our purposes, will be a real vector function when we discuss multivariate stochastic processes.

1.2 Process Identification

The introduction to the system identification problem in the last section leads to a natural subdivision of the problem into five parts:

- (1) the motivation to identify a system
- (2) the proposal of a model structure
- (3) an experiment to obtain data to which the model can be fitted
- (4) an evaluation of the results of the experiment
- (5) confirmation of the model or reproposal of the model structure.

Savas [1] makes an important distinction between models proposed for process design (or redesign) and models fitted to process data for prediction or control purposes. In the former, certain parameters and variables will correspond directly to such quantities as physical dimensions or material constants so that their influence on some performance criterion (or risk function) can be assessed directly. We are concerned with the latter requirement for which the model risk criterion should reflect the prediction capability only of the model

or models fitted. This specifically excludes the identification phase of the dual control problem considered by Feldbaum [66], Kwakernaak [67] and Riordon [68] because the cost function devised for their problems includes a measure of the controller performance.

Consider the selection and minimization problem:

$$\begin{array}{ll} \text{choose} & \text{minimize} \\ \Sigma_i & \Theta_i \end{array} R[l(y(t) - \hat{y}_i(t)), t; t_1] \quad (1.2-1)$$

where the loss function $l \in L$,

parameters $\Theta_i \in \omega$,

models $\Sigma_i \in \Omega$

and where the risk function R is defined as the expected value of the loss function l over the interval of time from t_0 to t_1 .

$\hat{y}_i(t)$ is the best prediction of the output $y(t)$ by the model Σ_i given all available prior knowledge of the system behaviour from measurements in the interval from t_0 to t_1 .

For example, for selecting and evaluating a discrete time model, we might choose to minimize the mean square output prediction error. Then the selection and minimization problem is (multivariate case):

$$\begin{array}{ll} \text{choose} & \text{minimize} \\ \Sigma_i & \Theta_i \end{array} \sum_{\tau=0}^t [y(\tau) - \hat{y}(\tau; \hat{\Theta}_i, \tau-1)]^T \Lambda^{-1} [y(\tau) - \hat{y}(\tau; \hat{\Theta}_i, \tau-1)] \quad (1.2-2)'$$

where Λ is a weighting matrix.

The purpose in introducing a selection procedure into the criterion is to allow a comparison of models that may differ in structure. For example, the order and degree of differential or difference equations,

the duration of lags, etc. Hence, the model Σ_i is chosen from the set of models Ω that the investigator may propose to compare by means of equation (1.2-1). Detailed examples are provided in the literature, in particular by Box and Jenkins, for systems with a single output [33]. Suitable models for the multivariate case are of current research interest.

Having chosen a model structure, how do we estimate the parameters Θ so that the risk function is minimized, given only a record of input and output measurements? In this thesis, we shall investigate a number of statistical methods for estimating the parameters of models. We shall show formally, and experimentally verify, that the classical method of least squares estimation leads to biased results when used to estimate the parameters in the dynamical models developed for multivariate systems. Thus we will be required to modify and extend existing statistical methods of parameter estimation to seek a solution to this problem. The results of the experiment will be evaluated by obtaining a measure of the mean square prediction error. This, we claim, provides the necessary model information for the design of a suitable controller. For example, if we can minimize the mean square prediction error of the output, we should be able to regulate the output within the same region of error.

1.3 Outline of the Thesis

Chapter 2 is a study of some state-variable canonical forms of constant, linear, multivariate, deterministic systems. The initial assumption is that all systems have a state-space description. This allows us to relate the well known conditions for controllability and observability [8] to the construction of some useful transformations. These are used to provide canonical state-space descriptions having fewer parameters than the original description. In particular, a new state-space description, the A-canonical form, is described.

In Chapter 3, it is shown how the A-canonical form leads to a description in the multivariate output and control variables. This description is called a vector differential (or difference) equation. It is related, in this chapter, to forms that have appeared in the literature a number of years ago and to matrix transfer functions. The procedure for transforming the vector difference (or differential) equation back into A-canonical form is given. A number of examples are supplied.

Systems with stochastic inputs are discussed in Chapter 4. It is shown that from a set of terminal measurements, systems with noise inputs to both states and outputs can be modelled by a description in which all the noise sources are combined into one vector noise process. The transforms relating stochastic state-space descriptions and stochastic vector difference equation descriptions are derived. This chapter provides a general model form suitable for the identification of multivariate systems.

Problems of parameter estimation are considered in Chapter 5. It is shown that estimation of the parameters of the models developed in Chapter 4 by least squares leads to incorrect estimates. A new method of modifying the least squares method is developed so that the estimates are asymptotically unbiased. The method of generalized least squares estimation is examined and passed over in favour of maximum likelihood estimation. The algorithms for estimating the parameters of the multivariate descriptions developed previously are provided.

In Chapter 6, examples of identification, prediction and control are provided. The superiority of the modified least squares method developed in Chapter 5 is demonstrated in comparison with other forms of least squares estimation. The criterion for comparison is the closeness of the estimated parameters to the true parameters of the systems simulated and also the variance of the prediction errors. The method is also compared with maximum likelihood estimation. The conclusion drawn is that the new method is suboptimal but considerable saving in computer effort is achieved. Finally, a simple regulator is designed for a multivariate system. New problems are shown to arise at this point.

The final chapter is a review of the contributions of the thesis and outlines the possibilities of future research.

1.4 Contributions of the Thesis

The principal contributions of this thesis, believed to be original, are the following:

- (1) New and practical techniques have been developed for transforming multivariate state-space descriptions into vector differential (or difference) equations in the input and output vector variables only. The core of the procedure is the transformation of a state-space description into the A-canonical form which is a new form. Attention is drawn to the fact that the transforms for reducing systems to Lur'e coordinates and phase-variable canonical form are special cases of the transform for reducing systems to the A-canonical form.
- (2) It is believed that the vector difference equation in the input and output variables has never been generalized for systems of any order and has never been applied to problems of system identification in the form presented here. The system description closest to the above appears to be the well known transfer function description [3, 29]. The chief distinction is that if a controllable and observable system can be described by a set of n first order difference equations, the vector difference equation requires fewer than n vector observations of the output if the output is of dimension greater than one.
- (3) The modified least squares method for parameter estimation (denoted here by the name "bootstrap" estimator) is believed to be a unique extension of methods introduced by Levadi [54] and Mayne [20, 56].

It is shown to perform remarkably well when compared to an optimal estimator and results in significant saving in computational effort.

(4) The derivation of algorithms for the identification of multivariate systems by maximum likelihood methods is thought to be an original extension of Astrom's [18, 19] elegant methods for the identification of single output systems.

In this thesis, an attempt has been made to illustrate the usefulness of the proposed algorithms by the computation of a number of examples.

CHAPTER 2

SOME STATE-VARIABLE CANONICAL FORMS OF A CONSTANT, LINEAR, MULTIVARIATE
DETERMINISTIC SYSTEM2.1 Introduction

In this chapter, the task of developing system models suitable for identification purposes is approached from the assumption that all systems have a state-space description. Some state-variable canonical forms will be developed that are useful from the identification point of view because they can be described with a minimum number of parameters in certain circumstances. The last form considered, the A-canonical form, will take us close to our ultimate objective of finding a suitable relation between the input and output vector variables only.

Assume that a linear, constant, multivariate, deterministic process has a known state-variable description Σ . The state variable equations of Σ are

$$\dot{x} = Fx + Gu \quad (2.1-1)$$

$$y = Hx \quad (2.1-2)$$

where x is an n -vector of state variables, u is an r -vector of controls, and y is an m -vector of measured variables. For convenience, the time dependence of the input vector $u = u(t)$, the state vector $x = x(t)$ and the output vector $y = y(t)$ will not be shown when they are attached to continuous-time descriptions of a system.

Suppose that there is a description Σ_1 defined by

$$\dot{x}_1 = F_1 x_1 + G_1 u \quad (2.1-3)$$

$$y_1 = H x_1 \quad (2.1-4)$$

such that the output of Σ_1 is identical to that of Σ for the same input and appropriate initial conditions. Then description Σ_1 is equivalent to Σ . This can be stated formally [27].

Definition: System description Σ_1 is equivalent to description Σ if there is a non-singular $n \times n$ matrix T with constant entries defined by

$$x_1 = T x \quad (2.1-5)$$

where x is the state vector of Σ and x_1 is the state vector of Σ_1 such that for the same input to Σ and Σ_1 , the output

$$y_1 \stackrel{\Delta}{=} H_1 x_1 \quad (2.1-6)$$

$$= y \quad (2.1-7)$$

Then the transformed system matrix F_1 , control matrix G_1 and observation matrix H_1 are defined by

$$F_1 = T F T^{-1} ; G_1 = T G ; H_1 = H T^{-1} \quad (2.1-8)$$

2.2 The Minimum Number of Parameters in a State-Variable Canonical Form

2.2.1 The Problem

In this Section, the problem of finding the minimum number of parameters in a state-variable model sufficient to describe a linear multivariate system is considered.

The state-variable equations of a linear, time-invariant, multivariate system $\Sigma = \Sigma(F, G, H)$ are given by (2.1-1) and (2.1-2). In general, $n^2 + n(m + r)$ parameters are required to specify Σ . The problem posed is that of finding a non-singular transformation T_m defined by

$$x_m = T_m x \quad (2.2-1)$$

so that the new description $\Sigma_m(F_m, G_m, H_m)$ has a minimum number of independent parameters. By definition, Σ_m is equivalent to Σ . We shall call Σ_m a minimum-parameter canonical form.

Kalman [3] has stated that the minimum number of independent parameters in some 'normal' form of Σ is

$$N_m = n(m + r) \quad (2.2-2)$$

In support of this statement, he outlined a scheme whereby F was reduced to a matrix F_m of n independent parameters. Then n parameters of either G_m or H_m were free to be chosen arbitrarily. As

Kalman did not provide the details for constructing the transformation, they are derived in the following Subsection.

2.2.2 Transformation to a Minimum-Parameter Normal Form

(1) In the first step, system matrix F is reduced by similarity transformation to any of its normal forms [4]. For example, a companion matrix or a diagonalized matrix is said to be a normal form. (An example of transformation to companion form is given in Subsection 2.4.2). Let us denote the transform by T_a and the resulting description by $\Sigma_a(F_a, G_a, H_a)$. If the degree of the minimal polynomial* of F is n_m , ($n_m \leq n$), then the transformed system matrix F_a has exactly n_m independent parameters. Hence the transformed description Σ_a has $n_m + n(m+r)$ independent parameters.

(2) Now we look for a transform that will allow us to arbitrarily choose some of the elements of the control and observation matrices without altering F_a . We are thus restricted to matrices of the following form which will commute with F_a . (That is, $T_b F_a = F_a T_b$.)

$$T_b = \alpha_0 I_n + \alpha_1 F_a + \dots + \alpha_{n_m-1} F_a^{n_m-1} \quad (2.2-3)$$

where I_n is a unit matrix of n rows and columns. We shall now

* Note: The minimal polynomial of a square matrix F is that polynomial $m(s)$ of least degree n_m such that $m(F) = 0$, the null matrix.

show how the coefficients $\alpha_0, \alpha_1, \dots, \alpha_{n_m-1}$ can be found to make certain elements of the control matrix take on arbitrary values. We could equally well have discussed the observation matrix. Setting

$$G_m = T_b G_a \quad (2.2-4)$$

and substituting for T_b , we obtain

$$G_m = [\alpha_0 I_n + \alpha_1 F_a + \dots + \alpha_{n_m-1} F_a^{n_m-1}] G_a \quad (2.2-5)$$

Let us define $g_{ja} = j^{\text{th}}$ column of G_a and substitute for the columns of G_a , and accordingly for G_m . In one additional step we write

$$\begin{bmatrix} g_{1m} \\ g_{2m} \\ \cdot \\ \cdot \\ g_{rm} \end{bmatrix} = \begin{bmatrix} g_{1a} & F_a g_{1a} & \cdot & \cdot & F_a^{n_m-1} g_{1a} \\ g_{2a} & F_a g_{2a} & \cdot & \cdot & F_a^{n_m-1} g_{2a} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ g_{ra} & F_a g_{ra} & \cdot & \cdot & F_a^{n_m-1} g_{ra} \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \cdot \\ \cdot \\ \alpha_{r-1} \\ \cdot \\ \alpha_{n_m-1} \end{bmatrix} \quad (2.2-6)$$

The array in (2.2-6) has rn_m rows and n_m columns. Each row corresponds to an element of G_m in the vector on the left. Clearly, we can solve

(2.2-6) for the vector of coefficients α_j ($j = 0, 1, \dots, n_m - 1$) by first constructing a non-singular matrix T_c with n_m rows taken from the array and used as in the following equation.

$$\left[\begin{array}{l} \text{An } n_m\text{-vector of} \\ \text{elements selected} \\ \text{from } G_m \text{ and set} \\ \text{to some desired} \\ \text{arbitrary value.} \end{array} \right] = \left[\begin{array}{l} T_c, \text{ a non-singular } n_m \times n_m \\ \text{matrix of rows taken from the} \\ n_m \times n_m \text{ array in (2.2-6);} \\ \text{one row for every element} \\ \text{selected in the } n\text{-vector on} \\ \text{the left.} \end{array} \right] \left[\begin{array}{l} \alpha_0 \\ \alpha_1 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_{n_m-1} \end{array} \right]$$

(2.2-7)

We are free to choose both the rows and the value of the corresponding element of G_m subject only to the restriction that rank $T_c = n_m$. Thus we specify the value of n_m parameters of G_m , solve for the vector of coefficients α_j and compute T_b . We then calculate the remaining coefficients of the minimum-parameter canonical form Σ_m . The number of parameters in Σ_m is

$$\begin{array}{rcl} n_m & \text{in} & F_m \\ r_n - n_m & \text{in} & G_m \\ mn & \text{in} & H_m \end{array}$$

which totals $n(m + r)$ in all.

A significant drawback is that we cannot decide, a priori, which elements of G_m (or H_m) can be chosen and set to some fixed value without first checking the rank of matrix T_c . For example, it can be shown that the elements of a column g_{jm} of G_m can be set to predetermined values if, and only if,

$$\text{rank} [g_{jm} \quad F_m g_{jm} \quad \dots \quad F_m^{n-1} g_{jm}] = n \quad (2.2-8)$$

This equation implies that the states of system Σ are completely controllable by the j^{th} element alone of the control vector.

In Section 2.5.2 it will be shown that complete controllability or complete observability is a sufficient condition to enable the derivation of a minimum-parameter canonical form in N_m independent parameters.

2.3 The Conditions for Controllability and Observability of a Linear, Constant, Multivariate System

Although the conditions for the controllability of a linear, constant, multivariate system are well known [8], a particularly simple proof is presented here because it provides an introduction to the construction of a useful transforming matrix. This proof has been influenced by the recent publication of Chen, Desoer and Niederlinski [9] who obtained the same simplified conditions for controllability by a somewhat different approach.

Theorem: Consider the state-variable equation (2.1-1) and (2.1-2) of the linear, constant system Σ in which the rank of G is r and the degree of the minimal polynomial of F is n_m . Then Σ is controllable if $\text{rank } c^P_p = n$ where c^P_p is defined by

$$c^P_p \triangleq [G \quad FG \quad \dots \quad F^{p-1}G] \quad (2.3-1)$$

and the controllability index by

$$p \leq \min(n_m, n-r+1) \quad (2.3-2)$$

Definition [10]: The controllability index p of the system Σ is defined as the smallest positive integer for which c^P_p has rank n .

Proof: The solution of (2.1-1), the state equations of system Σ is given by

$$x(t) = e^{Ft}x(0) + \int_0^t e^{F(t-\tau)}Gu(\tau)d\tau \quad (2.3-3)$$

The system is said to be controllable if the state of the system can be changed from any initial state $x(0)$ to the origin of the state space X in a finite time t_f by the application of an appropriate input u over this period. Hence, if Σ is controllable, any $x(0)$ may be expressed as

$$x(0) = - \int_0^{t_f} e^{-F\tau}Gu(\tau)d\tau \quad (2.3-4)$$

It is well known [27] that for an n^{th} order linear, constant system, coefficients a_k may be found that satisfy

$$e^{-F\tau} = \sum_{k=0}^{n-1} a_k(\tau) F^k \quad (2.3-5)$$

But, if the degree of the minimal polynomial of F is $n_m \leq n$, then every term F^k for all $k \geq n_m$ is a linear combination of terms of lower degree. Thus, in the general case, we can write

$$e^{-F\tau} = \sum_{k=0}^{n_m-1} a_k^*(\tau) F^k \quad (2.3-6)$$

On substituting (2.3-6) in (2.3-4) and reversing the order of summation and integration

$$x(0) = - \sum_{k=0}^{n_m-1} F^k G \int_0^{t_f} a_k^*(\tau) u(\tau) d\tau \quad (2.3-7)$$

Let
$$- \int_0^{t_f} a_k^*(\tau) u(\tau) d\tau = \mu_k \quad (2.3-8)$$

Then
$$x(0) = \sum_{k=0}^{n_m-1} F^k G \mu_k$$

$$= [G \quad FG \quad \dots \quad F^{n_m-1} G] \begin{bmatrix} \mu_0 \\ \mu_1 \\ \vdots \\ \mu_{n_m-1} \end{bmatrix} \quad (2.3-9)$$

Clearly, if ${}_c P_{n_m}$ has rank n , its range spans the entire vector space X including the point $x(0)$. Therefore, $\mu_0, \mu_1, \dots, \mu_{n_m-1}$ can be found to satisfy equation (2.3-9).^{*} This proves that Σ is controllable if ${}_c P_p$ has rank n and $p = n_m$.

The part of the proof to show that $p \leq n - r + 1$ follows from Chen et al. [9] who observed that the dimension of the range of ${}_c P_s$ ($0 \leq s \leq p$) is at least one greater than the dimension of the range of ${}_c P_{s-1}$ each time that a block of the form $F^{s-1}G$ is added to ${}_c P_{s-1}$. Now if the rank of G is r , then up to $n - r$ blocks of the form $F^s G$ are required in ${}_c P_p$. Hence, $p \leq \min(n_m, n-r+1)$. Finally, they show a special example in which ${}_c P_p$ has rank n and $p < \min(n_m, n-r+1)$. This concludes the proof.

Corollary: The system described by equations (2.2-1) and (2.2-2) is observable if rank ${}_o P_p = n$ where

$${}_o P_p = \begin{bmatrix} H \\ HF \\ \cdot \\ \cdot \\ HF^{p-1} \end{bmatrix} \quad (2.3-10)$$

We redefine p to be the observability index

$$p \leq \min(n_m, n - m + 1) \quad (2.3-11)$$

It will always be clear in which context p is used.

* See [26,27] for a treatment of the problem:
given μ , find $u(t)$.

Corollary: The results apply when the system equations are formulated in discrete time:

$$x(t + 1) = Fx(t) + Gu(t) \quad (2.3-12)$$

$$y(t) = Hx(t) \quad (2.3-13)$$

Corollary: Systems for which the degree of the minimal polynomial of F is less than n are neither controllable nor observable by a scalar input or scalar output respectively. This can be demonstrated in the observable case by making each matrix H a row vector. Then, if $p = n_m \ll n$, it follows that $\text{rank } \begin{matrix} P \\ O \end{matrix} < n$.

It should be observed that arrays $\begin{matrix} P \\ C \end{matrix}$ and $\begin{matrix} P \\ O \end{matrix}$ are not necessarily square. In the following sections, some important canonical forms are derived by constructing nonsingular transformations from these arrays. There may be more than one such matrix for each array depending on the number of combinations of n independent columns (rows) that can be found.

2.4 A Useful Aid to the Derivation of Canonical Forms

In this Section, a useful lemma is proven that provides simple rules whereby most of the elements of the transformed matrices in many canonical forms can be written down by inspection. The result is employed throughout the remainder of the Chapter where canonical forms are considered. An example is given to illustrate the use of the rules.

2.4.1 The Lemma

Given the i^{th} row vector a_i of matrix A which has n columns and at least one row and b_j , the j^{th} row vector of the nonsingular $n \times n$ matrix B, if

$$b_j = a_i \quad (2.4-1)$$

then the i^{th} row of the matrix E is

$$E = AB^{-1} \quad (2.4-2)$$

is given by $e_i = (0, \dots, \delta_{ij}, 0 \dots)$ (2.4-3)

That is, the j^{th} element of e_i is unity and all the other elements are zero.

The proof is obvious by inspection after postmultiplying both sides of equation (2.4-2) by B.

Corollary: A parallel result can be obtained for the form $E = B^{-1}A$.

If $b_j = a_i$ where b_j and a_i are the j^{th} and i^{th} columns of B and A respectively, then the i^{th} column of E has unity only in the j^{th} position and zeros elsewhere. That is

$$e_i = (0, \dots, \delta_{ji}, 0 \dots)^T \quad (2.4-4)$$

2.4.2 A Simple Example of a Canonical Form Obtained with the Lemma

Consider the $n \times n$ transform

$$T_m = [d \quad Fd \quad \dots \quad F^{n-1}d] \quad (2.4-5)$$

where d is any n -vector that satisfies the condition

$$\text{rank } T_m = n \quad (2.4-6)$$

Substituting $x_2 = T_m^{-1}x$ in (2.1-1) and employing the corollary above

$$\begin{aligned} F_2 &= [F^{-1}d \ d \ \dots \ F^{n-2}d]^{-1} [d \ Fd \ \dots \ F^{n-1}d] \\ &= \begin{bmatrix} 0 & 0 & \dots & 0 & -\omega_n \\ 1 & 0 & \dots & 0 & -\omega_{n-1} \\ 0 & 1 & \dots & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & 1 & -\omega_1 \end{bmatrix} \end{aligned} \quad (2.4-7)$$

which is the transpose of a normal form of F , the companion matrix [4]. Because the $n \omega_j$ terms satisfy the characteristic polynomial $f(\lambda)$ (2.4-8) of both F and F_2 (a criterion for their similarity [4]), they are unique and independent of the choice of the elements of the n -vector d .

$$f(\lambda) = (-)^n (\lambda^n + \omega_1 \lambda^{n-1} + \dots + \omega_n) \quad (2.4-8)$$

Thus d can be chosen, by the lemma, from columns of G to make n elements of G_2 arbitrary provided that the condition (2.4-6) is satisfied. There is no reduction in the number of independent parameters in H_2 for this choice. Hence, the transformed system Σ_2 is completely specified by $N_m = n(m+r)$ parameters if d

is so chosen.

In the special case of control by a single variable, the transforming matrix constructed with the single-column matrix G is that used for testing controllability [8] of a single input system. The basis of the state vector of Σ is transformed into the basis of Lur'e coordinates by operation with this matrix [13, 12, 5]. It is also basic to the method of Wonham and Johnson in constructing the phase-variable canonical form [6, 7].

2.5 Multivariate State-Variable Canonical Forms Defined with the Conditions for Controllability

In this section, the corollary of the last section will be used with the conditions for controllability to show how some simple canonical forms for multivariable systems can be derived. The transforming matrices T in the cases discussed will be made up of linearly independent column vectors taken from the array c_p^P shown in (2.3-1) and used for testing system controllability. Hence the functional dependence of T may be shown by

$$T = T(F,G) \quad (2.5-1)$$

We shall be interested to see if the canonical forms derived have N_m parameters, and are thus minimum-parameter realizations in the sense used by Kalman [3], (2.2-2).

Definition: If there is a system Σ_1 , completely specified by N_m parameters which is equivalent to system Σ , then Σ_1 is a minimum-parameter canonical description of Σ .

In the following, it will become evident that transformations to minimum-parameter realizations are not unique.

2.5.1 Companion Matrix (Transposed) Canonical Form

Consider the system Σ described by state-variable equation (2.1-1). If the system Σ is controllable, then array c^P_p (2.3-1) has n independent columns.

The feature of the construction of the transform for this form, for which the complete details are given in [10], is that the controllability of the system state by each control term is tested in turn. The linear independence of all vectors g_j ($j = 1, \dots, r$) of G is assumed. One of the columns of G (say g_1) is selected and entered in the first column of T . Then subsequent columns Fg_1, F^2g_1, \dots , etc. are added until a dependency is reached or $n \times n$ matrix T is full. In the first case the dependent vector $F^j g_1$ is discarded and the cycle is repeated for g_2, g_3 etc. until T is full. The resulting transform is

$$T = [g_1 \quad Fg_1 \quad \dots \quad F^{p_1-1} g_1 \quad g_2 \quad Fg_2 \quad \dots \quad F^{p_2-1} g_2 \quad \dots \quad g_v \quad Fg_v \quad \dots \quad F^{p_v-1} g_v] \quad (2.5-2)$$

where v is an integer such that $v \leq r$ and T has n independent columns.

On carrying out a change of coordinates under equivalence transformation from the basis of Σ to the basis of Σ_3 we get

$$\begin{aligned}
 F_3 &= T^{-1}FT \\
 &= (F^{-1}T)^{-1}T \\
 &= [F^{-1}g_1 \ g_1 \ Fg_1 \ \dots \ F^{p_1-2}g_1 \ F^{-1}g_2 \ \dots \ F^{p_v-2}g_v]^{-1} \\
 &\quad \cdot [g_1 \ Fg_1 \ F^2g_1 \ \dots \ F^{p_1-1}g_1 \ g_2 \ \dots \ F^{p_v-1}g_v] \quad (2.5-3)
 \end{aligned}$$

and

$$G_3 = [g_1 \ Fg_1 \ F^2g_1 \ \dots \ F^{p_1-1}g_1 \ g_2 \ \dots \ F^{p_v-1}g_v]^{-1}G \quad (2.5-4)$$

Now by carefully comparing columns in the matrix products according to the corollary of Subsection 2.4.1, we can write down the arbitrary columns of zeros and ones of F_3 and G_3 immediately by inspection. For example, the second column in the first matrix of the matrix product in (2.5-3) is the same as (can be paired with) the first column of the second matrix. Hence, the first column vector of F_3 has a one for its second element and zeros everywhere else. Continuing, we find that F_3 and G_3 have the form

$$F_3 = \begin{bmatrix} 0 & 0 & \cdot & 0 & x & 0 & 0 & & x & \cdot & 0 & & 0 & x \\ 1 & 0 & & \cdot & x & \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot & \cdot \\ 0 & 1 & & & x & & & & & & & & & \\ & & & & x & & & & & & & & & \\ & & & & x & & & & & & & & & \\ & & & & x & & & & & & & & & \\ 0 & \cdot & \cdot & 1 & \underline{x} & | & 0 & & \underline{x} & | & & & & \\ 0 & & & 0 & x & | & 0 & 0 & \cdot & \cdot & x & | & \cdot & \\ 0 & & & 0 & x & | & 1 & 0 & \cdot & \cdot & x & & & \\ \cdot & & & \cdot & \cdot & \cdot & \cdot & & \cdot & & \cdot & & & \cdot \\ \cdot & & & \cdot & \cdot & \cdot & 0 & \cdot & 1 & x & | & \cdot & & x \\ & & & & & & & & & & x & | & \cdot & x \\ & & & & & & & & & & \cdot & \cdot & | & 0 & \cdot & 0 & x \\ & & & & & & & & & & \cdot & \cdot & | & 1 & & x \\ \cdot & & & \cdot & \cdot & \cdot & \cdot & & \cdot & & \cdot & & & \cdot & & \cdot \\ 0 & \cdot & \cdot & 0 & x & 0 & 0 & \cdot & \cdot & x & \cdot & 0 & & 1 & x \end{bmatrix} \quad (2.5-5)$$

and

$$G_3 = \begin{bmatrix} 1 & 0 & \cdot & 0 & x & \cdot & x \\ 0 & \cdot & & & x & \cdot & x \\ \cdot & \cdot & & & \cdot & & \cdot \\ \cdot & \cdot & & & \cdot & & \cdot \\ \cdot & 0 & \cdot & \cdot & x & \cdot & x \\ \cdot & 1 & & & \cdot & & \cdot \\ \cdot & 0 & & 0 & & & \cdot \\ \cdot & \cdot & \cdot & 1 & & & \cdot \\ \cdot & \cdot & & \cdot & \cdot & & \cdot \\ 0 & \cdot & & 0 & x & & x \end{bmatrix} \quad (2.5-6)$$

where x's stand for the non-zero elements which characterize the system. The system matrix F_3 in the new coordinates is composed of v companion matrices (transposed) located in square blocks of dimension p_j along the main diagonal. The blocks can be considered

as coupled subsystems of order p_j . There will be v such columns corresponding to the v companion matrix (transposed) blocks resulting from a transformation constructed with just v columns of G . By similar reasoning, there will be $(r - v)$ nonarbitrary columns in G_3 corresponding to the $(r - v)$ columns of G_3 not employed in the construction of T .

Totalling the number of parameters that are not unity or zero in this form (and hence the parameters that specify the system), we find that there are nv in F_3 and $n(r-v)$ in G_3 . The number of independent parameters in H_3 remains unaltered under this transformation. The total is then

$$N_m = n(m + r) \quad (2.5-7)$$

hence, a minimum-parameter realization has been found.

As in the first minimum-parameter canonical form discussed, we cannot tell in advance which columns of the control matrix can have predetermined elements without first conducting some tests.

2.5.2 An Assertion on the Realizability of Minimum-Parameter

Canonical Descriptions

Assertion: Complete controllability or complete observability is a sufficient condition to show that the minimum number of parameters

required to describe Σ is given by

$$N_m = n(m + r) \quad (2.5-8)$$

Proof: In the last Subsection, a reduced canonical form in $N_m = n(m + r)$ parameters was found for any selection of column vectors of G , subject only to the condition that the rank of T must be n . If the system is controllable, then T can be formed by taking n independent columns from cP_p which has rank n . Then the rank of T is n and a canonical form in N_m parameters can be found. This proves the assertion for controllable systems. The assertion for observable systems can be shown similarly. Alternatively, the duality theorem of Kalman can be invoked directly [8].

2.5.3 Another Reduced Form

The reason for introducing the following transform is that its dual, derived using the conditions for observability, leads to a system description in the input and output variables.

Again consider the system Σ . This time, the construction of T will be undertaken by inserting blocks of the form $F^k G$ adjacent to each other. The steps are:

(1) An integer q in the range

$$p \leq q \leq n \quad (2.5-9)$$

is selected where p is the controllability index (which may not be known) and n is the dimension of the state space.

(2) The $n \times n$ transformation T is constructed as discussed below.

$$T = [F^{q-p}(G)_p \ F^{q-p+1}(G)_{p-1} \ \dots \ F^{q-2}(G)_2 \ F^{q-1}G_1] \quad (2.5-10)$$

The construction of transformation T is initiated by inserting the $n \times r$ block $F^{q-1}G$ in the last r columns of T . Then additional blocks in descending powers of q are inserted, deleting columns which are dependent upon column vectors previously inserted. When T is full, p is then known. The brackets and subscripts on a matrix $(G)_j$ are used to remind us that $(G)_j$ is comprised only of columns of G such that the columns of the product $F^{q-j}(G)_j$ are independent of the columns of $F^{q-k}G$ for all integers $k < j$. There is no special ordering of the columns. Then the n columns of T are independent. Clearly, T can now be simplified by premultiplying by F^{p-q} .

On changing the coordinates from the basis of Σ to the basis of Σ_4 under equivalence transformation we get

$$\begin{aligned} F_4 &= T^{-1}FT \\ &= [F^{-1}(G)_p \ (G)_{p-1} \ F(G)_{p-2} \ \dots \ F^{p-3}(G)_2 \ F^{p-2}(G)_1]^{-1} \cdot \\ &\quad [(G)_p \ F(G)_{p-1} \ F^2(G)_{p-2} \ \dots \ F^{p-2}(G)_2 \ F^{p-1}(G)_1] \end{aligned} \quad (2.5-11)$$

and

$$\begin{aligned} G_4 &= T^{-1}G \\ &= [(G)_p \ F(G)_{p-1} \ \dots \ F^{p-2}(G)_2 \ F^{p-1}(G)_1]^{-1}G \end{aligned} \quad (2.5-12)$$

Again, employing the corollary of Subsection 2.4.1, we can write down the arbitrary columns of zeros and ones in F_4 and G_4 by inspection. We first note that all the columns in $(G)_j$ appear in $(G)_{j+1}$ because the independent columns were inserted sequentially from the right.

$$F_4 = \begin{bmatrix} (0)_{p,p} & & & \cdot & (0)_{p,2} & x & x & \cdot & x \\ [\delta]_{p-1,p} & & & & \cdot & x & x & \cdot & x \\ & [\delta]_{p-2,p-1} & & & & \cdot & & & \cdot \\ & & & \cdot & & & & & \cdot \\ & & & & & & x & x & \cdot & x \\ (0)_{1,p} & \cdot & & & [\delta]_{1,2} & x & x & \cdot & x \end{bmatrix} \quad (2.5-13)$$

The notation, admittedly difficult, is explained in this way. The x 's are elements which characterize the system. 0 is a null matrix. $(0)_{i,j}$ is a null matrix with as many rows as columns of $(G)_i$ and with the same number of columns as $(G)_j$. $[\delta]_{k-1,k}$ is a matrix of zeros and ones with as many rows as columns of $(G)_{k-1}$ and with the same number of columns as $(G)_k$. From the rule for paired columns, the matrix element δ_{ji} (of $[\delta]_{k-1,k}$) = 1 if the j^{th} column of $F^{p-k}(G)_{k-1}$ is the same as the i^{th} column of $F^{p-k}(G)_k$.

One possible permutation of G_4 , depending on the initial ordering of the columns in G is

$$G_4 = \begin{bmatrix} x & \cdot & x & [6]_{p,p} & x & \cdot & x \\ & & & & & & \\ & & & & & & \\ x & \cdot & x & 0 & x & \cdot & x \end{bmatrix} \quad (2.5-14)$$

By inspecting equation (2.5-11), it can be seen that the r columns of the last block $F^{p-1}(G)_1$ which cannot be paired result in r nonarbitrary columns in F_4 of nr elements. Similarly, the columns of G that are not paired with columns of $(G)_p$ in equation (2.5-12) result in nonarbitrary columns in G_4 .

2.6 The A-Canonical Form Derived Using the Conditions for Observability

The final canonical state description that we shall consider is that provided by operation with the dual of the transform (2.5-10) discussed in Subsection 2.5.3. We shall call it the A-canonical form. The reason for its importance is that it has many features common to a vector differential equation in the input and output variables. This will be made clear in the next Chapter.

In the following, we write down the transform, discuss the structure of the transformed description $\Sigma_A(F_A, G_A, H_A)$ and illustrate it with an example.

2.6.1 The Transformation

The transformation T_A is constructed from independent rows of o_p^P given by equation (2.3-10). T_A is the dual of the last transform (2.5-10) developed in Section 2.5.

$$T_A = \begin{bmatrix} (H)_p \\ (H)_{p-1}F \\ (H)_{p-2}F^2 \\ \cdot \\ (H)_{p-j}F^j \\ \cdot \\ HF^{p-1} \end{bmatrix} \quad (2.6-1)$$

By analogy to the previous notation, the brackets and subscripts on matrix $(H)_{p-k}$ now mean that $(H)_{p-k}$ is comprised only of rows of H such that the rows of the product $(H)_{p-k}F^k$ are independent of the rows of $(\frac{HF^j}{p-j})$ for all $j > k$.

Using the transformation

$$x_A = T_A x \quad (2.6-2)$$

the new system matrices are

$$F_A = \begin{bmatrix} (H)_p & (H)_{p^{F-1}} \\ (H)_{p^{-1}F} & (H)_{p^{-1}} \\ (H)_{p^{-2}F^2} & (H)_{p^{-2}F} \\ \cdot & \cdot \\ (H)_{2^F p^{-2}} & (H)_{2^F p^{-3}} \\ HF^{p-1} & HF^{p-2} \end{bmatrix}^{-1} \quad (2.6-3)$$

$$G_A = T_A G \quad (2.6-4)$$

and

$$H_A = H \begin{bmatrix} (H)_p \\ (H)_{p^{-1}F} \\ \cdot \\ \cdot \\ (H)_{2^F p^{-2}} \\ HF^{p-1} \end{bmatrix}^{-1} \quad (2.6-5)$$

For convenience, we can order the rows of H so that the non-arbitrary rows appear together in a block. With the lemma of Subsection 2.4.1, we write

$$F_A = \begin{bmatrix} (0)_{p,p} & [\delta]_{p,p-1} & \cdot & (0)_{p,\cdot} \\ & & [\xi]_{p-1,p-2} & \cdot \\ \cdot & & \cdot & \cdot \\ (0)_{2,p} & \cdot & \cdot & [\delta]_{2,\cdot} \\ -A_p & -A_{p-1} & -A_{p-2} & \cdot & -A_1 \end{bmatrix} \quad (2.6-6)$$

where $(0)_{i,j}$ is a null matrix with as many rows as $(H)_i$ and as many columns as rows of $(H)_j$. $[\delta]_{k,k-1}$ is a matrix of unity and zero elements with as many rows as $(H)_k$ and as many columns as rows of $(H)_{k-1}$. From the lemma, the matrix element δ_{ij} (of $[\delta]_{k,k-1}$) = 1 if the i^{th} row of $(H)_k F^{p-k}$ is the same as the j^{th} row of $(H)_{k-1} F^{p-k}$. The submatrices A_j have dimensions $m \times$ (the number of rows of $(H)_j$). The reason for the negative sign associated with A_j will be apparent later.

One permutation of H gives

$$H_A = \begin{bmatrix} x & . & . & x \\ x & . & . & x \\ (\mathcal{S})_{p,p} & & & 0 \end{bmatrix} \quad (2.6-7)$$

where the x 's are nonzero elements.

This system is not generally a minimal realization because the m rows of H_A are reduced to zeros and ones only if $(H)_p$ has m rows.

It will be shown in the next chapter that the submatrices A_j of F_A are identically the same as the matrix coefficients in the homogeneous part of a vector differential equation description of the original system Σ . For this reason, we note here that by comparing F_A in the product form of (2.6-3) and the partitioned form (2.6-6), we can write down the identity

$$HF^{p-1}(T_A F^{-1})^{-1} = HF^p T_A^{-1} \quad (2.6-8)$$

$$= -[A_p \ A_{p-1} \ \dots \ A_1] \quad (2.6-9)$$

2.6.2 Algebraic Examples

Consider a simple algebraic example of a system for which $n = 6$, $r = 2$ and $m = 3$. The three rows of H are designated by h_1 , h_2 and h_3 . Now suppose that as a result of discarding dependent rows, the nonsingular transforming matrix is given by

$$T_A = \begin{bmatrix} h_3 \\ h_2^F \\ h_3^F \\ h_1^{F^2} \\ h_2^{F^2} \\ h_3^{F^2} \end{bmatrix} \quad (2.6-10)$$

Then in our notation, $p = 3$ and

$$(H)_2 = \begin{bmatrix} h_2 \\ h_3 \end{bmatrix}; \quad (H)_3 = h_3 \quad (2.6-11)$$

Substituting in (2.6-3)

$$F_A = \begin{bmatrix} h_3 \\ h_2^F \\ h_3^F \\ h_1^{F^2} \\ h_2^{F^2} \\ h_3^{F^2} \end{bmatrix} \begin{bmatrix} | \\ | \\ | \end{bmatrix} \begin{bmatrix} h_3^{F^{-1}} \\ h_2 \\ h_3 \\ h_1^F \\ h_2^F \\ h_3^F \end{bmatrix}^{-1} \quad (2.6-12)$$

with the pairing shown. Using the lemma, we can write down the first three rows immediately. Then

$$F_A = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \hline & & & & & \\ & & & & & \\ & & & & & \\ \hline & & & & & \\ & & & & & \\ & & & & & \\ \hline -A_3 & & -A_2 & & & -A_1 \end{bmatrix} \quad (2.6-13)$$

where A_1 is 3×3 , A_2 is 3×2 and A_3 is 3×1 . Now for H_A there is only one pairing

$$H_A = \left[\begin{array}{c} h_1 \\ h_2 \\ h_3 \end{array} \right] \left[\begin{array}{c} h_3 \\ h_2^F \\ h_3^F \\ h_1^F{}^2 \\ h_2^F{}^2 \\ h_3^F{}^2 \end{array} \right]^{-1} \quad (2.6-14)$$

$$= \begin{bmatrix} x & x & x & x & x & x \\ x & x & x & x & x & x \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (2.6-15)$$

There is no simplification in G_A . The total number of parameters to be specified in this realization is: 3×6 for F_A , 6×2 for G_A and 2×6 for H_A , totalling 42 parameters in all. Clearly, these

parameters must be combinations of $N_m = 6(2 + 3) = 30$ basic parameters in some minimal realization.

For general interest, let us now look at a minimum-parameter realization for the same system using the information derived in the construction of T_A (2.6-10). The new transform T_m to be constructed is the dual of the companion matrix (transposed) canonical form discussed in general terms in Subsection 2.5.1 and shown in equation (2.5-2).

Since, in our example, $h_1 F$ was discarded in constructing T_A because $h_1 F$ and $h_1 F^2$ are linearly dependent, we discard the latter for T_m . Continuing in this manner, we get

$$T_m = \begin{bmatrix} h_1 \\ h_2 \\ h_2 F \\ h_3 \\ h_3 F \\ h_3 F^2 \end{bmatrix} \quad (2.6-16)$$

The system matrix in the new coordinates is shown with its pairs

$$F_5 = \begin{bmatrix} h_1 \\ h_2 \\ h_3 F \\ h_3 \\ h_3 F \\ h_3 F^2 \end{bmatrix} \begin{bmatrix} | \\ | \\ | \\ | \\ | \\ | \end{bmatrix} \begin{bmatrix} h_1 F^{-1} \\ h_2 F^{-1} \\ h_2 \\ h_3 F^{-1} \\ h_3 \\ h_3 F \end{bmatrix}^{-1} \quad (2.6-17)$$

$$= \begin{bmatrix} \boxed{x} & \boxed{x} & \boxed{x} & \boxed{x} & \boxed{x} & \boxed{x} \\ \hline 0 & 0 & 1 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \boxed{x} & \boxed{x} & \boxed{x} & \boxed{x} & \boxed{x} & \boxed{x} \\ \hline 0 & 0 & 0 & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \hline x & x & x & x & x & x \end{bmatrix} \quad (2.6-18)$$

Subsystems of order 1, 2 and 3 in companion form are set out in boxes along the principal diagonal. The nonzero entries outside the boxes can be considered as connections between the subsystems.

In the same way it can be shown that

$$H_5 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \quad (2.6-19)$$

but G_5 is not reduced. Totalling the number of parameters specific to the realization, we find that 3 x 6 belong to F_5 and 2 x 6 belong to G_5 which sums to $N_m = 30$. Thus the description is minimal* a significant reduction from the basic description in $6(6 + 2 + 3) = 66$ elements in all.

* It has recently come to the attention of the author that in [10], Luenberger mentions a special construction with which the matrix in (2.6-18) can be reduced further.

CHAPTER 3

TRANSFORMATION BETWEEN A STATE-VARIABLE DESCRIPTION AND A DESCRIPTION
IN THE INPUT AND OUTPUT VARIABLES3.1 Introduction

In this Chapter we shall develop a description in the multivariate output and control variables of a controllable and observable system. We would like to be able to write down such a description, given only a set of state-variable equations Σ and the requirement that the two descriptions respond identically to the same input. Conversely, given system equations in the input and output variables only, we would like to be able to transform them into a state-variable description.

The term vector differential (or difference) equation or description Σ_D will be applied to the set of equations in the input and output vector variables that can be separated into two parts as in the example of equation (3.1-1). For example, the Laplace transform with zero initial conditions of a vector differential equation of order 2 with the usual substitution $s = d/dt$ is

$$(s^2 I_m + sA_1 + A_2)y(s) = (s^2 B_0 + sB_1 + B_2)u(s) \quad (3.1-1)$$

Equation (3.1-2) below defines the Σ_D description in Laplace transform notation. Obtaining the Σ_D description is equivalent to determining the polynomial matrices $A(s)$ and $B(s)$ which satisfy equation (3.1-2)

$$A(s)y(s) = B(s)u(s) \quad (3.1-2)$$

The solution of the homogeneous part of (3.1-2), $A(s)y(s) = 0$, gives the free response of the dynamical system. The vector $B(s)u(s)$ is the forcing function. Although state-variable descriptions are first-order vector differential equations in the state variables, we shall not refer to them by this term.

In Section 3.2, the computational procedures for transforming between the state-variable description Σ and the Laplace transform of a vector differential equation description Σ_D are stated and then illustrated with algebraic and numerical examples. The procedures are later justified in Section 3.3.

The work of Duncan and Collar [14, 15] is related to this work and shown to be a restricted treatment of the general approach discussed here. A matrix transfer function description, Σ_Z , also in the vector input and output variables, is derived and shown to fit the Σ_D description.

The Chapter is concluded with a discussion based on a comparison of the number of parameters employed in some of the descriptions that have been considered.

3.2 The Transformation to Description Σ_D

In this Section, the procedure for deriving the Laplace transform of a vector differential (or difference) equation of order p , termed description Σ_D , from a state-variable description Σ is stated and

illustrated with a numerical example. The justification of the transformation is given in Section 3.3.

3.2.1 The Computational Procedure

(1) Construct the transform T_A used in the derivation of the A-canonical form and recalled here from (2.6-1) for convenience. It must have a nonzero determinant.

$$T_A = \begin{bmatrix} (H)_p \\ (H)_{p-1}F \\ (H)_{p-2}F^2 \\ \cdot \\ (H)_{p-j}F^j \\ \cdot \\ HF^{p-1} \end{bmatrix} \quad (3.2-1)$$

(2) Calculate the $m \times n$ matrix product below and partition according to the notation employed in deriving the A-canonical form. This identity follows immediately from equation (2.6-9).

$$HF^p T_A^{-1} = [-A_p \quad -A_{p-1} \quad -A_{p-2} \quad \dots \quad -A_1] \quad (3.2-2)$$

Recall from equation (2.6-6) that the submatrices A_k have the same number of columns as rows of $(H)_k$. They are the matrix coefficients of the matrix polynomial $A(s)$ of the Σ_D description defined by equation (3.1-2). Then $A(s)$ is given by

$$\begin{aligned}
 A(s) = & s^p I_m + s^{p-1} A_1 + s^{p-2} A_2 (I_m)_2 + \dots \\
 & \dots + s A_{p-1} (I_m)_{p-1} + A_p (I_m)_p
 \end{aligned} \tag{3.2-3}$$

The bracket notation, $(I_m)_k$ ($k = 2, 3, \dots, p$), is used in the following way. Recall that in Subsection 2.6.1, certain rows of H were deleted to form $(H)_k$. By this means, the reduced matrices were employed in making up T_A as in equation (3.2-1). For every row deleted from H to make $(H)_k$, we delete the corresponding row from the unit matrix I_m to make $(I_m)_k$. That is, if we delete the j^{th} row from H , we also eliminate the j^{th} row from I_m in constructing $(H)_k$ and $(I_m)_k$ respectively. Then $(H)_k$ and $(I_m)_k$ have the same number of rows.

(3) Construct the following lower triangular matrix with the $m \times m$ matrix products of the form $A_k (I_m)_k$ calculated in step (2) for (3.2-3).

$$T_P = \begin{bmatrix}
 I_m & 0 & 0 & \cdot & \cdot & 0 \\
 A_1 & I_m & 0 & \cdot & \cdot & \cdot \\
 A_2 (I_m)_2 & A_1 & I_m & & & \cdot \\
 \cdot & & \cdot & & & \cdot \\
 A_{p-j-1} (I_m)_{p-j-1} & A_{p-j-2} (I_m)_{p-j-2} & \cdot & & & 0 \\
 \cdot & \cdot & \cdot & & & \cdot \\
 \cdot & \cdot & \cdot & A_1 & I_m & 0 \\
 A_{p-1} (I_m)_{p-1} & A_{p-2} (I_m)_{p-2} & \cdot & A_2 (I_m)_2 & A_1 & I_m
 \end{bmatrix} \tag{3.2-4}$$

T_p is an invertible matrix of mp rows and columns. Note that the $m \times m$ submatrices are duplicated along lines parallel to the principal diagonal.

(4) Finally, calculate the elements of the array below

$$\begin{bmatrix} B_1 \\ B_2 \\ \cdot \\ B_j \\ \cdot \\ B_p \end{bmatrix} = T_p \circ_p G \quad (3.2-5)$$

where the B_j terms are the matrix coefficients of the matrix polynomial $B(s)$ below for the Σ_D description defined by (3.1-2). The calculation to include B_0 for direct control will be given in a later Section.

$$B(s) = s^{p-1}B_1 + s^{p-2}B_2 + \dots + sB_{p-1} + B_p \quad (3.2-6)$$

3.2.2 Extension to Include the Vector Difference Equation Form

By the simple substitution of the forward difference operator z for the differential operator s , the method may be used for discrete-time models. Then the Σ_D description is

$$A(z)y(t) = B(z)u(t) \quad (3.2-7)$$

$$\text{where } zy(t) \triangleq y(t+1) ; \quad zu(t) \triangleq u(t+1) \quad (3.2-8)$$

By comparison with (3.2-3) and (3.2-6), we write

$$A(z) = z^p I_m + z^{p-1} A_1 + z^{p-2} A_2 (I_m)_2 + \dots + z A_{p-1} (I_m)_{p-1} + A_p (I_m)_p$$

and (3.2-9)

$$B(z) = z^{p-1} B_1 + z^{p-2} B_2 + \dots + z B_{p-1} + B_p \quad (3.2-10)$$

We may prefer to write (3.2-7) in terms of past controls and observations. This is easily obtained by multiplying both sides of the equation by z^{-p} . Then, with the definition of (3.2-8), we have

$$y(t) = -A_1 y(t-1) - A_2 (y(t-2))_2 \dots - A_p (y(t-p))_p$$

$$+ B_1 u(t-1) + B_2 u(t-2) \dots + B_p u(t-p) \quad (3.2-11)$$

$$\text{where } (y(t-k))_k \triangleq (I_m)_k y(t-k) \quad (3.2-12)$$

The addition of a current control term $B_0 u(t)$ will be introduced later.

3.2.3 A Numerical Example

The system matrices used in this example are

$$F = \begin{bmatrix} 0.42 & -0.24 & -3.82 \\ 0.16 & 0.48 & 8.64 \\ 0.50 & 0.0 & -1.50 \end{bmatrix} ; \quad G = \begin{bmatrix} 0.4 & -0.7 \\ 0.9 & -0.2 \\ 0.1 & -0.3 \end{bmatrix}$$

$$H = \begin{bmatrix} 0.8 & -0.2 & 0.4 \\ 0.5 & 0.6 & -0.3 \end{bmatrix} \quad (3.2-13)$$

(1) Under transformation by the matrix T_A ,

$$\begin{aligned} T_A &= \begin{bmatrix} h_2 \\ h_1^F \\ h_2^F \end{bmatrix} \\ &= \begin{bmatrix} 0.500 & 0.600 & -0.300 \\ 0.504 & -0.288 & -5.384 \\ 0.156 & 0.168 & 3.724 \end{bmatrix} \end{aligned} \quad (3.2-14)$$

the matrices in the A-canonical form can be found. (This step is not necessary for the construction of the vector difference equation.)

$$\begin{aligned} F_A &= \begin{bmatrix} 0.0 & 0.0 & 1.0 \\ -0.915 & -3.031 & -3.472 \\ 0.604 & 2.526 & 2.431 \end{bmatrix}; \quad G_A = \begin{bmatrix} 0.710 & -0.380 \\ -0.596 & 1.320 \\ 0.586 & -1.260 \end{bmatrix} \\ H_A &= \begin{bmatrix} -0.263 & 1.258 & 1.905 \\ 1.0 & 0.0 & 0.0 \end{bmatrix} \end{aligned} \quad (3.2-15)$$

(2) From the last two rows of F_A , (or by multiplying out the product $HF^2T_A^{-1}$) we get

$$A_1 = \begin{bmatrix} 3.031 & 3.472 \\ -2.526 & -2.431 \end{bmatrix}; \quad A_2 = \begin{bmatrix} 0.915 \\ -0.604 \end{bmatrix} \quad (3.2-16)$$

(3,4) We now calculate the product

$$\begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} I_2 & 0 \\ A_1 & I_2 \end{bmatrix} \begin{bmatrix} H \\ HF \end{bmatrix} G \quad (3.2-17)$$

and obtain

$$B_1 = \begin{bmatrix} 0.180 & -0.640 \\ 0.710 & -0.380 \end{bmatrix} ; \quad B_2 = \begin{bmatrix} 2.415 & -1.939 \\ 1.595 & 1.280 \end{bmatrix} \quad (3.2-18)$$

To assist in this calculation, we note that the matrix belonging to the A-canonical form

$$G_A = \begin{bmatrix} h_2 \\ h_1 F \\ h_2 F \end{bmatrix} G \quad (3.2-19)$$

is contained in the last three rows of the product $\begin{bmatrix} H \\ HF \end{bmatrix} G$ used in (3.2-17).

Then, substituting (3.2-16) and (3.2-18) into (3.2-7)

$$A(z) = z^2 I_m + zA_1 + A_2 [0 \quad 1] \quad (3.2-20)$$

$$B(z) = zB_1 + B_2 \quad (3.2-21)$$

so that

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = -A_1 \begin{bmatrix} y_1(t-1) \\ y_2(t-1) \end{bmatrix} - A_2 y_2(t-2) + B_1 \begin{bmatrix} u_1(t-1) \\ u_2(t-2) \end{bmatrix} + B_2 \begin{bmatrix} u_1(t-2) \\ u_2(t-2) \end{bmatrix} \quad (3.2-22)$$

It is a simple additional step to find the characteristic equation for this system by transforming F into a companion matrix (transposed) by the method described in Subsection 2.5-1. We obtain the characteristic equation of F

$$\lambda^3 + 0.60\lambda^2 + 0.80\lambda + 0.48 = 0 \quad (3.2-23)$$

which has a real root at

$$\lambda_1 = -0.60 \quad (3.2-24)$$

and two imaginary roots

$$\lambda_2, \lambda_3 = \pm \sqrt{-0.80} \quad (3.2-25)$$

within the unit circle in the z plane.

The three system models, $\Sigma(F,G,H)$, $\Sigma_A(F_A,G_A,H_A)$ and $\Sigma_D(A,B)$ were simulated in discrete time on a digital computer. The output trajectories of the three were identical to eight decimal digits, starting from zero initial conditions, over ten time intervals.

3.3 Derivation of the Transformation from Description Σ to Σ_D

The method used in deriving the transformation is to find two independent sets of vector equations in the input, state and output vector variables. Subject to certain conditions, the state variables can be eliminated between the two equations giving us the required description in the input and output vector variables.

We proceed by differentiating system equations (2.1-2) $p - 1$ times and combining the result with equation (2.1-1). Let us assume that this operation can be carried out and that formal Laplace transforms with zero initial conditions can be taken at each stage ($s = d/dt$). Then we write

$$\begin{aligned}
 y(s) &= Hx(s) \\
 sy(s) &= HFx(s) + HGu(s) \\
 s^2y(s) &= HF^2x(s) + HFGu(s) + sHGu(s) \\
 &\quad \cdot \quad \quad \quad \cdot \quad \quad \quad \cdot \\
 s^jy(s) &= HF^jx(s) + HF^{j-1}Gu(s) + \dots + s^{j-1}HGu(s) \\
 &\quad \cdot \quad \quad \quad \cdot \quad \quad \quad \cdot \\
 s^{p-1}y(s) &= HF^{p-1}x(s) + HF^{p-2}Gu(s) + \dots + s^{p-2}HGu(s) \quad (3.3-1)
 \end{aligned}$$

We claim that the system states $x(t)$ are observable if, and only if, the array introduced as equation (2.3-10)

$${}_o^p P = \begin{bmatrix} H \\ HF \\ \cdot \\ \cdot \\ HF^{p-1} \end{bmatrix} \quad (3.3-2)$$

is of rank n . Recall that we selected n independent rows of ${}_o^p P$ to make up the nonsingular matrix T_A of equation (2.6-1). Here

we select the corresponding set of n equations from (3.3-1) so that every equation includes a row of T_A .

With our notation, we write n equations

$$\begin{aligned}
 (y(s))_p &= (H)_p x(s) \\
 s(y(s))_{p-1} &= (H)_{p-1} Fx(s) + (H)_{p-1} Gu(s) \\
 &\cdot \quad \cdot \quad \cdot \\
 s^j(y(s))_{p-j} &= (H)_{p-j} F^j x(s) + (H)_{p-j} F^{j-1} Gu(s) + s(H)_{p-j} F^{j-2} Gu(s) \\
 &\quad + \dots s^{j-1} (H)_{p-j} Gu(s) \\
 &\cdot \quad \cdot \quad \cdot \\
 s^{p-1} y(s) &= HF^{p-1} x(s) + HF^{p-2} Gu(s) + \dots s^{p-2} HGu(s) \quad (3.3-3)
 \end{aligned}$$

Note that there may be more than one way to select the n independent rows of T_A from rows of ${}_0^P P$ which corresponds to a choice in the outputs and successive derivatives of the outputs employed in the model. Regrouping the equation of (3.3-3) into vector form we get

$$S(s)y(s) = T_A x(s) + Q(s)u(s) \quad (3.3-4)$$

where $S(s)$ is an $n \times m$ matrix and $Q(s)$ is $n \times r$.

$$S(s) = \begin{bmatrix} (I_m)_p \\ s(I_m)_{p-1} \\ \cdot \\ s^j(I_m)_{p-j} \\ \cdot \\ s^{p-1}I_m \end{bmatrix} \quad (3.3-5)$$

$$Q(s) = \begin{bmatrix} (O)_p \\ (H)_{p-1}G \\ (H)_{p-2}FG \\ \cdot \\ (H)_{p-j}F^{j-1}G \\ \cdot \\ HF^{p-2}G \end{bmatrix} + s \begin{bmatrix} (O)_p \\ (O)_{p-1} \\ (H)_{p-2}FG \\ \cdot \\ (H)_{p-j}F^{j-2}G \\ \cdot \\ HF^{p-3}G \end{bmatrix} + \dots$$

$$+ \dots + s^j \begin{bmatrix} (O)_p \\ \cdot \\ \cdot \\ (O)_{p-j} \\ (H)_{p-j-1}G \\ \cdot \\ HF^{p-j-2}G \end{bmatrix} + \dots + s^{p-2} \begin{bmatrix} (O)_p \\ \cdot \\ \cdot \\ (O)_{p-j} \\ \cdot \\ (O)_2 \\ HG \end{bmatrix} \quad (3.3-6)$$

Now we take the Laplace transform of the p^{th} derivative of $y(t)$. This gives us m new equations.

$$s^p y(s) = HF^p x(s) + HF^{p-1} Gu(s) + \dots + s^{p-1} HGu(s) \quad (3.3-7)$$

Because the inverse of T_A exists, we can premultiply (3.3-4) by T_A^{-1} . Then

$$x(s) = T_A^{-1} S(s) y(s) - T_A^{-1} Q(s) u(s) \quad (3.3-8)$$

Eliminating $x(s)$ by substituting (3.3-8) into (3.3-7), we get

$$[s^p I_m - HF^p T_A^{-1} S(s)] y(s) = -HF^p T_A^{-1} Q(s) u(s) + HF^{p-1} Gu(s) + \dots + s^{p-1} HGu(s) \quad (3.3-9)$$

$$\triangleq [-HF^p T_A^{-1} Q(s) + R(s)] u(s) \quad (3.3-10)$$

Replacing the $m \times m$ matrix on the left by $A(s)$ and the $m \times r$ matrix on the right by $B(s)$ as defined by equation (3.1-2), we obtain

$$A(s) = s^p I_m - HF^p T_A^{-1} S(s) \quad (3.3-11)$$

$$B(s) = R(s) - HF^p T_A^{-1} Q(s) \quad (3.3-12)$$

In the expression for $A(s)$ we substitute for $HF^p T_A^{-1}$ from the identity (2.6-9) and for $S(s)$ from (3.3-5) and obtain the matrix polynomial $A(s)$ given by equation (3.2-3), repeated here for convenience.

$$A(s) = s^p I_m + s^{p-1} A_1 + s^{p-2} A_2 (I_m)_2 + \dots + s A_{p-1} (I_m)_{p-1} + A_p (I_m)_p \quad (3.2-3)$$

The derivation of the matrix coefficients in $B(s)$ is less direct. In equation (3.3-10) we substitute for $HF^p T_A^{-1}$ from (2.6-9), for $Q(s)$ from (3.3-6) and for $R(s)$ from (3.3-9). Noting that, by definition, the number of columns in F_{p-j} is the same as the number of rows of $(H)_{p-j}$, we calculate the term

$$-HF^p T_A^{-1} Q(s) = [A_p \dots A_1] Q(s) \quad (3.3-13)$$

$$\begin{aligned}
= & A_{p-1} [(H)_{p-1} G + s(O)_{p-1} + \dots + s^{p-2}(O)_{p-1}] \\
& + A_{p-2} [(H)_{p-2} FG + s(H)_{p-2} G + s^2(O)_{p-2} + \dots + s^{p-2}(O)_{p-2}] \\
& + \dots \\
& + A_{p-j} [(H)_{p-j} F^{j-1} G + s(H)_{p-j} F^{j-2} G + s^2(H)_{p-j} F^{j-3} G + \dots + s^{p-1}(H)_{p-j} G + s^{p-2}(O)_{p-j}] \\
& + \dots \\
& + A_1 [HF^{p-2} G + sHF^{p-3} G + \dots + s^{p-1} HF^{p-j-1} G + s^{p-2} HG]
\end{aligned} \quad (3.3-14)$$

We can now evaluate each of the matrix terms B_j because we know that B_j is the matrix coefficient that appears with the operator s^{p-j} . Collecting these terms, we write

$$\begin{aligned}
 B_1 &= HG \\
 B_2 &= A_1 HG + HFG \\
 B_3 &= A_1 HFG + A_2 (H)_2 G + HF^2 G \\
 &\cdot \quad \cdot \quad \cdot \quad \cdot \\
 B_{p-j} &= A_1 HF^{p-j-2} G + A_2 (H)_2 HF^{p-j-3} G + \dots + A_{p-j-1} (H)_{p-j-1} G + HF^{p-j-1} G \\
 &\cdot \quad \cdot \quad \cdot \quad \cdot \\
 B_p &= A_1 HF^{p-2} G + A_2 (H)_2 HF^{p-3} G + \dots + A_{p-1} (H)_{p-1} G + HF^{p-1} G
 \end{aligned} \tag{3.3-15}$$

The terms in $HF^j G$ can be removed if we insert a suitable operator that picks out the appropriate rows for every term. We can accommodate this requirement in our notation so that we write

$$\begin{aligned}
 \left[\begin{array}{c} B_1 \\ B_2 \\ B_3 \\ \cdot \\ B_{p-j} \\ \cdot \\ B_p \end{array} \right] &= \left[\begin{array}{cccccc} I_m & & & & & 0 \\ & A_1 & & & & I_m \\ & & A_2 (I_m)_2 & & & A_1 \\ \cdot & & \cdot & & & \cdot \\ & & & & & \cdot \\ & A_{p-j-1} (I_m)_{p-j-1} & & A_{p-j-2} (I_m)_{p-j-2} & & \cdot \\ \cdot & & & \cdot & & \cdot \\ & & & & & \cdot \\ & A_{p-1} (I_m)_{p-1} & & A_{p-2} (I_m)_{p-2} & & \cdot \\ & & & & & \cdot \\ & & & & & A_1 \\ & & & & & I_m \end{array} \right] \left[\begin{array}{c} HG \\ HFG \\ HF^2 G \\ \cdot \\ HF^{p-j-1} G \\ \cdot \\ HF^{p-1} G \end{array} \right]
 \end{aligned} \tag{3.3-16}$$

Finally, we note that the array on the right of the last equation can be separated into the matrix product $\begin{matrix} P \\ O \end{matrix} G$ giving us the identity (3.2-5) in step (4) of the computational procedure.

3.4 Modifications to Include Direct Control of the System Output

Suppose that a system which we wish to describe by a state-vector equation or vector differential equation in the input and output vector variables has a direct connection between the system input u and the system output y . The descriptions that we have discussed up to now can be extended to cope with this new requirement.

3.4.1 Complete State-Variable Description

The state-variable equations of a linear time-invariant, multi-variate system Σ , set out in (2.1-1) and (2.1-2) are modified by the addition of an output-control term Du . The new continuous-time equations are

$$\dot{x} = Fx + Gu \quad (3.4-1)$$

$$y = Hx + Du \quad (3.4-2)$$

The corresponding set of equations for a discrete-time formulation, replacing (2.3-12) and (2.3-13), is

$$x(t+1) = Fx(t) + Gu(t) \quad (3.4-3)$$

$$y(t) = Hx(t) + Du(t) \quad (3.4-4)$$

It can be seen that matrix D is unchanged if the basis of x in X undergoes transformation; hence the state-variable canonical forms of Chapter 2 are easily modified by the addition of the Du term as above.

3.4.2 Complete Description Σ_D with Direct Control

The addition of a term $B_0 u$ to the Σ_D description to correspond with the Du term in the state-variable form alters all the other matrix coefficients in the polynomial matrix $B(s)$. This affects steps 3 and 4 of the computational procedure of Section 3.2. It is a simple matter to insert the additional term in the derivation of the transform starting with equation (3.3-1). Following through the derivation, it is found that m additional columns are added to matrix T_P derived in equation (3.3-16) and (3.3-15) but first defined in step 3 of Subsection 3.2.1. The modified form, shown below with an asterisk, should be compared with equation 3.2-4.

$$T_P^* = \begin{bmatrix} I_m & & & & & & & 0 \\ & & & & & & & \vdots \\ A_1 & & & & & & & \vdots \\ & & & & & & & \vdots \\ A_2(I_m)_2 & & & & & & & \vdots \\ \cdot & & & & & & & \vdots \\ & & & & & & & \vdots \\ A_{p-j-1}(I_m)_{p-j-1} & & & & & & & \vdots \\ \cdot & & & & & & & \vdots \\ & & & & & & & \vdots \\ A_p(I_m)_p & & & & & & & \vdots \end{bmatrix} \quad T_P \quad (3.4-5)$$

T_P^* is a lower triangular, invertible matrix of $m(p+1)$ rows and columns with T_P imbedded in the lower right corner.

Equation (3.2-5) in Step 4 is replaced by

$$\begin{bmatrix} B_0 \\ B_1 \\ B_2 \\ \cdot \\ B_p \end{bmatrix} = T_P^* \begin{bmatrix} D \\ HG \\ HFG \\ \cdot \\ HF^{p-1}G \end{bmatrix} \quad (3.4-6)$$

Finally, we write in place of (3.2-6)

$$B(s) = s^p B_0 + s^{p-1} B_1 + \dots + s B_{p-1} + B_p \quad (3.4-7)$$

and in place of (3.2-10)

$$B(z) = z^p B_0 + z^{p-1} B_1 + \dots + z B_{p-1} + B_p \quad (3.4-8)$$

and
$$y(t) = -A_1 y(t-1) - A_2 (y(t-2))_2 \dots A_p (y(t-p))_p$$

$$+ B_0 u(t) + B_1 u(t-1) + \dots + B_p u(t-p) \quad (3.4-9)$$

in place of (3.2-11).

3.4.3 A Numerical Example

A discrete-time model of a first-order subsystem and a third order subsystem coupled one way was made up for this example. The system matrices were

$$\begin{aligned}
 F &= \begin{bmatrix} 0 & -0.5 & 1.0 & 0 \\ 0.25 & -0.5 & -0.5 & 0 \\ 0 & 0 & -0.5 & 1.0 \\ 0 & 0 & 0 & 0.75 \end{bmatrix} & G &= \begin{bmatrix} 1.0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1.0 \end{bmatrix} \\
 H &= \begin{bmatrix} 0 & 0 & 0 & 1.0 \\ 0 & 1.0 & 1.0 & 0 \end{bmatrix} & D &= \begin{bmatrix} 1.0 & 0.5 \\ 0.5 & 1.0 \end{bmatrix}
 \end{aligned}
 \tag{3.4-10}$$

The vector difference equation was computed by following the steps laid down before.

$$(1) \quad T_A = \begin{bmatrix} h_2 \\ h_2 F \\ h_1 F^2 \\ h_2 F^2 \end{bmatrix} \tag{3.4-11}$$

$$(2) \quad A_1 = \begin{bmatrix} -2.114 & 2.068 \\ -2.273 & 2.364 \end{bmatrix}; \quad A_2 = \begin{bmatrix} -0.148 \\ -0.080 \end{bmatrix}; \quad A_3 = \begin{bmatrix} -0.068 \\ -0.051 \end{bmatrix}$$

(3.4-12)

(3,4) The product below was calculated

$$\begin{bmatrix} B_0 \\ B_1 \\ B_2 \\ B_3 \end{bmatrix} = \begin{bmatrix} I_2 & 0 & \cdot & 0 \\ A_1 & I_2 & \cdot & \cdot \\ A_2 [0 \ 1] & A_1 & I_2 & \cdot \\ A_3 [0 \ 1] & A_2 [0 \ 1] & A_1 & I_2 \end{bmatrix} \begin{bmatrix} D \\ HG \\ HFG \\ HF^2G \end{bmatrix} \tag{3.4-13}$$

from which

$$\begin{aligned}
 B_0 = D &= \begin{bmatrix} 1.0 & 0.5 \\ 0.5 & 1.0 \end{bmatrix} ; & B_1 &= \begin{bmatrix} -1.080 & 2.011 \\ -1.091 & 1.277 \end{bmatrix} \\
 B_2 &= \begin{bmatrix} 0.176 & -1.511 \\ 0.210 & -1.352 \end{bmatrix} ; & B_3 &= \begin{bmatrix} -0.170 & 0.477 \\ -0.128 & 0.358 \end{bmatrix}
 \end{aligned}
 \tag{3.4-14}$$

Substituting in the form of (3.4-9):

$$\begin{aligned}
 \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} &= -A_1 \begin{bmatrix} y_1(t-1) \\ y_2(t-1) \end{bmatrix} - A_2 y_2(t-2) - A_3 y_3(t-3) \\
 &+ B_0 \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} + B_1 \begin{bmatrix} u_1(t-1) \\ u_2(t-1) \end{bmatrix} + B_2 \begin{bmatrix} u_1(t-2) \\ u_2(t-2) \end{bmatrix} + B_3 \begin{bmatrix} u_1(t-3) \\ u_2(t-3) \end{bmatrix}
 \end{aligned}
 \tag{3.4-15}$$

The roots of the characteristic equation were found at

$$\begin{aligned}
 \lambda_{1,2} &= -0.25 \pm j 0.25 \\
 \lambda_3 &= -0.5 ; \lambda_4 = 0.75
 \end{aligned}
 \tag{3.4-16}$$

They are distinct. The transformation method was recalculated without difficulty for a system with a pair of multiple roots. The models were then simulated on the computer. The output trajectories of the transformed models agreed with those of the initial description to at least seven decimal digits.

3.5 Discussion of the Σ_D Description

In this Section, some background for the Σ_D description is provided. The various forms in which the description may be useful are then summarized. The problems of existence and uniqueness of the description are then discussed.

3.5.1 Historical Perspective

In the introduction of his book, Pipes [17] credits Duncan and Collar [14, 15] with initiating the use of matrix algebra in engineering applications. One problem with which they were concerned is of interest to us because it has features common to the descriptions discussed in this Chapter.

Duncan and Collar dealt with solutions to the Lagrange equations of motion of a general system of m degrees of freedom oscillating in the neighbourhood of an equilibrium position. Their problem, formulated in our notation, was to solve the set of m second-order homogeneous differential equations

$$A_0 \ddot{x} + A_1 \dot{x} + A_2 x = 0 \quad (3.5-1)$$

which we write in the manner of (3.1-1) as

$$[s^2 I_m + sA_1^* + A_2^*]x = 0 \quad (3.5-2)$$

where x is an m -vector of generalized coordinates. The elements of A_0 , A_1 , and A_2 are inertia coefficients, damping coefficients and

stiffness coefficients respectively. The coefficient matrices are all $m \times m$ and A_0 is symmetric. Their method was to reduce the set of m second order equations to a set of $2m$ first-order equations in a space of $n = 2m$ dimensions in the vector variables x and \dot{x} . The reformulated system equations were

$$\dot{x}^* = \begin{bmatrix} 0 & I \\ -A_2^* & -A_1^* \end{bmatrix} x^* \quad (3.5-3)$$

where $x^* = \begin{bmatrix} \dot{x} \\ x \end{bmatrix}$; $A_1^* = A_0^{-1}A_1$; $A_2^* = A_0^{-1}A_2$ (3.5-4)

This is a special case of a transformation method to be described in Section 3.6 with $n = 2m$ and $p = 2$.

In [16], Frazer, Duncan and Collar also discuss the solution of a set of second order nonhomogeneous equations with a constant forcing term. This corresponds to the example of (3.1-1) with $B_2 \neq 0$ and $B_0 = B_1 = 0$.

Pipes [17] has revived interest in second-order vector differential equations, principally as a convenient way to derive the Lagrange equations of high-dimension for vibration analysis. However, he transforms the equations to first-order before he undertakes the formal analysis of the problem for the reason that the powerful methods of modal analysis can be employed directly.

Astrom and Bohlin [18] presented a method for the numerical identification of a process from single-input, single-output data. With Wensmark [19], they conducted a thorough investigation of a number of model forms. A difference equation description was found to be suitable for identification purposes. They remarked that they were restricted to studying systems with a single output because they did not have a suitable canonical form for multivariate systems.

In a recent paper, Box and Jenkins [23] proposed a multivariate model of the form

$$y(t+1) = y(t) + u(t+1) + B_1 u(t) \quad (3.5-5)$$

as an analog of a scalar process which they had investigated in detail. The components of $u(t)$ are taken from a sequence of uncorrelated disturbances.

Mayne proposed an ad hoc form [20] which was studied by Tzafestas [21]. Their model, which will be investigated in more detail in a later Chapter, is a vector difference equation with $n = 4$, $p = 2$ and $m = 2$.

3.5.2 Forms of the Σ_D Description

Let us list and name the forms of Σ_D by giving simple examples.

$$(1) \quad A_0 \frac{d^2 y(t)}{dt^2} + A_1 \frac{dy(t)}{dt} + A_2 y(t) = u(t) \quad (3.5-6)$$

is a nonhomogeneous equation of order 2 with constant coefficients first discussed by Frazer, Duncan and Collar [16] and referred to in the previous Subsection. The corresponding form, derived for control purposes in the last Section, is a normalized, more general version

$$\begin{aligned} & \frac{d^2 y(t)}{dt^2} + A_1 \frac{dy(t)}{dt} + A_2 y(t) \\ &= B_0 \frac{d^2 u(t)}{dt^2} + B_1 \frac{du(t)}{dt} + B_2 u(t) \end{aligned} \quad (3.5-7)$$

$$(2) \quad (s^2 I_m + sA_1 + A_2)y(s) = (s^2 B_0 + sB_1 + B_2)u(s) \quad (3.5-8)$$

is the Laplace transformed version of (3.5-7) with the substitution ($s = d/dt$).

$$(3) \quad (z^2 I_m + zA_1 + A_2)y(z) = (z^2 B_0 + zB_1 + B_2)u(z) \quad (3.5-9)$$

is the z-transformed version of (3.5-8).

(4) In preference to (3), we use the vector difference equation in operator form as in (3.2-7) where we treat z as a forward-shift operator. Then

$$(z^2 I_m + zA_1 + A_2)y(t) = z^2 B_0 + zB_1 + B_2)u(t) \quad (3.5-10)$$

(5) The discrete-time system equation in the input and output variables is

$$y(t) = -A_1 y(t-1) - A_2 y(t-2) + B_0 u(t) + B_1 u(t-1) + B_2 u(t-2) \quad (3.5-11)$$

(6) In a later section, we shall discuss the transfer function

$$Z(s) \triangleq H(sI - F)^{-1} G \quad (3.5-12)$$

where $y(s) = Z(s)u(s)$ (3.5-13)

By substitution from (3.1-2)

$$Z(s) = A^{-1}(s) B(s) \quad (3.5-14)$$

which, in this example, is

$$Z(s) = [s^2 I_m + sA_1 + A_2]^{-1} [s^2 B_0 + sB_1 + B_2] \quad (3.5-15)$$

We can see, by setting $m = r = 1$, that the equations listed are the familiar forms of the scalar-coefficient, single-input, single-output descriptions found in most textbooks on control.

The author believes that Kavanagh was the first to give the transfer function $Z(s)$, precisely as in (3.5-14), in terms of polynomial matrices $A(s)$ and $B(s)$ [35]. He gave a numerical example to show how a pair of simultaneous second-order equations could be

written both in terms of three first-order equations and as a transfer function matrix. However, he did not indicate a general method.

Reference to a multivariate transfer function first appeared in Western literature in a study of non-interacting controls by Boksenbom and Hood [36]. Their method, in a broad treatment, is more readily accessible in Tsien [37]. The problem of synthesis of the matrix transfer function was first considered by Amora [38]. Other early contributions are listed in the extensive bibliography of contributions up to 1962 prepared by Kavanagh [35].

3.5.3 Comments on the Existence and Uniqueness of the Description

We now summarize the conditions that must be met in order that we can obtain the Σ_D description. The following are restricted conditions because they pertain to and justify only the procedures of Sections 3.3 and 3.4 in deriving the description.

(1) The system has a known state-space description Σ or any other state-space description that is equivalent to Σ . This is required because the derivation is based on a given description Σ .

(2) The system is observable. The critical step in eliminating the intermediate variable x depends on being able to invert T_A which is comprised of n rows of the observability array o_p . T_A is non-singular if the system is observable.

(3) Formal Laplace transforms of the system equations can be carried out as in (3.3-1).

Now we inquire if the description Σ_D is unique. This question is of particular importance when parameters of a system are to be estimated. We shall investigate the problem for Σ_D written as a difference equation, the form that we shall use in later Chapters when we investigate methods of parameter estimation by statistical means.

Assertion: The description Σ_D is unique for a specified selection of observations that satisfy the conditions for state observability.

Proof: The n equations (3.3-3) in difference equation form can be represented by

$$y^* = f_1[x(0), u(0), zu(0), z^2u(0), \dots, z^{p-2}u(0)] \quad (3.5-16)$$

where

$$y^* = \begin{bmatrix} (y(0))_p \\ z(y(0))_{p-1} \\ \cdot \\ \cdot \\ z^{p-1}y(0) \end{bmatrix} = \begin{bmatrix} (y(0))_p \\ (y(1))_{p-1} \\ \cdot \\ \cdot \\ y(p-1) \end{bmatrix} \quad (3.5-17)$$

y^* is an n -vector made up of selected elements of the output over $p - 1$ time intervals beginning with $t = 0$. $f_1[\cdot]$ is a linear function of the system initial state vector $x(0)$, and $p - 1$ control inputs beginning with $u(0)$. For simplicity, the dependence on time of all

the variables is not shown in the following notation. The m equations (3.3-7) can be written

$$z^p y = f_2[x, u, zu, \dots, z^{p-2}u, z^{p-1}u] \quad (3.5-18)$$

Provided that the initial system states are observable with the choice y^* , there is a function g that can be obtained from (3.3-8) such that

$$x = g[y^*, u, zu, \dots, z^{p-2}u] \quad (3.5-19)$$

Hence, in (3.5-18)

$$\begin{aligned} z^p y &= f_2[g[y^*, u, zu, \dots, z^{p-2}u], u, zu, \dots, z^{p-2}u, z^{p-1}u] \\ &= f'_2[y^*, u, zu, \dots, z^{p-2}u, z^{p-1}u] \end{aligned} \quad (3.5-20)$$

because f_2 is a linear function of its variables. (3.5-20) is the difference equation form of (3.3-9) in the n selected elements of y over $p - 1$ intervals and u over the same number of intervals.

If f'_2 is not a unique representation of the system in the control and selected elements of the observation variables, then there exists another multivariate set of m linear relations $f'_3[\cdot] \neq f'_2[\cdot]$ such that

$$z^p y = f'_3[y^*, u, zu, \dots, z^{p-2}u, z^{p-1}u] \quad (3.5-21)$$

But, substituting from (3.5-16) for y^*

$$z^p y = f'_3[f_1[x, u, zu, \dots, z^{p-2}u], u, zu, \dots, z^{p-2}u, z^{p-1}u] \quad (3.5-22)$$

$$= f_3[x, u, zu, \dots, z^{p-2}u, z^{p-1}u] \quad (3.5-23)$$

Now we subtract equation (3.5-23) from equation (3.5-18) and obtain:

$$f_2[x, u, zu, \dots, z^{p-2}u, z^{p-1}u] - f_3[x, u, zu, \dots, z^{p-2}u, z^{p-1}u] = 0 \quad (3.5-24)$$

Clearly, to satisfy (3.5-24) (for any nontrivial choice of the initial state x and the sequence of p controls $u, zu, \dots, z^{p-1}u$) implies that $f_2[\cdot] = f_3[\cdot]$. If this is true, then

$$f_3'[y^*, u, zu, \dots, z^{p-1}u] = f_2'[y^*, u, zu, \dots, z^{p-1}u] \quad (3.5-25)$$

which contradicts our original statement that $f_3'[\cdot] \neq f_2'[\cdot]$. This proves the uniqueness of f_2' and f_2 for the choice of elements of y^* . By substituting for f_2' in (3.5-20), which has been proven unique, we get the required form

$$A(z)y(t) = B(z)u(t) \quad (3.5-26)$$

corresponding to the choice of $y^*(t)$.

We conclude that the pair $[A(z), B(z)]$ is unique for a specified selection of observations that satisfy the conditions of state observability, thereby proving the assertion.

Corollary: There are as many descriptions Σ_{D_i} that are input-output equivalent to a state-variable, observable system \mathcal{X} as there are ways in which n rows of independent vectors can be selected, without regard for order, from the observability array O_p^p for Σ .

Proof: The description Σ_D of the theorem was unique for a specified selection of outputs corresponding to the selection of n independent rows from the $m \times n$ array ${}_o P_p$ for the nonsingular transform T_A . Suppose that n rows are selected to form a nonsingular matrix T_{A1} . Corresponding to this selection, there will be a vector y_1^* of n elements (see equation (3.5-17)). If we make a new selection for another nonsingular matrix T_{A2} such that at least one row of T_{A2} is different than the previous selection for T_{A1} , this will define a new vector y_2^* of n elements. There will be as many elements of y_2^* different from y_1^* as rows of transforms T_{A2} that are different from rows of transform T_{A1} . Thus for every other such transform T_{Ai} , there is a unique vector y_i^* , disregarding all those cases in which the elements of y^* are reordered. Hence, by substituting the new vector y_i^* for y^* in (3.5-19) and subsequent equations, it is clear that there is a new description Σ_{D_i} for every unique vector y_i^* . This concludes the proof.

Note that in the single-output case ($m = 1$) there will be only one choice for T_A because ${}_o P_p$ has only n rows.

3.6 Transformation of the Σ_D Description into State-Variable Form

The requirement of this section is to establish a procedure for transforming the Σ_D description of order p into a set of state equations. The transformation from the state description Σ (in variables x , y and u) into the form Σ_D (in variables y and u) discussed in this

Chapter, resulted in a unique model. The problem is that the converse is not true; hence, it is necessary to propose a structure (or form) for the state-variable model so that the elements of the matrices in the proposed form can be determined uniquely from the parameters of the vector differential (difference) equation of order p .

It is most convenient to propose transformation into the A-canonical form of the last Chapter for the reason that simple relations have already been established in this Chapter to connect the matrix elements of the two descriptions. Then any other state-variable description can be derived from the A-canonical form by a suitable choice of a basis in the state space.

3.6.1 The Procedure for Transformation into A-Canonical Form

(1) Determine the integer n by totalling the number of columns in the matrix coefficients A_1, A_2, \dots, A_p . Then commence construction of the $n \times n$ matrix F_A by imbedding the above coefficients in the last m rows, observing the proper sign as in equation (2.6-6).

(2) Recall F_A in the form reproduced here from (2.6-3) for convenience.

$$F_A = \begin{bmatrix} (H)_p \\ (H)_{p-1}F \\ (H)_{p-2}F^2 \\ \cdot \\ (H)_2F^{p-2} \\ HF^{p-1} \end{bmatrix} \begin{bmatrix} (H)_p F^{-1} \\ (H)_{p-1} \\ (H)_{p-2}F \\ \cdot \\ (H)_2 F^{p-3} \\ HF^{p-2} \end{bmatrix}^{-1} \quad (2.6-3)$$

$$\triangleq T_A K^{-1} \quad (3.6-1)$$

Now the element δ_{ij} at the intersection of the i^{th} row and j^{th} column of F_A is unity if the i^{th} row of T_A and the j^{th} row of K are the same. (The rest of the i^{th} row is made up of zeros.) The components of vector y and the derivatives of y that are used in the vector differential equation are known from the Σ_D description. As a consequence, we know which rows of each submatrix HF^{p-k} have been eliminated as in Subsection (3.3.1) to form transform T_A although we do not know the row elements explicitly. We then also know which rows of every product HF^k have been used to construct matrix K in 3.6-1. Hence, we can write down the locations of the pairs that match, the rows of T_A and K , even though we cannot write down the quantitative value of each element. Finally, having paired $n - m$ rows in T_A and K , we can enter the appropriate zeros and ones in F_A according to the method previously discussed and shown in equation (2.6-6). This will be made clear in an example in the next Subsection.

(3) The matrix T_p^* of $m(p+1)$ rows and columns is constructed from the submatrices A_1, \dots, A_p as shown in (3.4-5). Its inverse exists because it is triangular with ones all along its principal diagonal.

Then we calculate the array on the left.

$$\begin{bmatrix} D \\ HG \\ HFG \\ \cdot \\ HF^{p-1}G \end{bmatrix} = T_p^{*-1} \begin{bmatrix} B_0 \\ B_1 \\ B_2 \\ \cdot \\ B_p \end{bmatrix} \quad (3.6-2)$$

This gives us D immediately. By the same method as in step (2), we eliminate all but n rows of the mp rows remaining in (3.6-2) after the removal of D , noting that

$$\begin{bmatrix} (HG)_p \\ (HFG)_{p-1} \\ \cdot \\ \cdot \\ (HF^{p-2}G)_2 \\ HF^{p-1}G \end{bmatrix} = \begin{bmatrix} (H)_p G \\ (H)_{p-1} FG \\ \cdot \\ \cdot \\ (H)_2 F^{p-2} G \\ HF^{p-1} G \end{bmatrix} \quad (3.6-3)$$

$$= T_A G \quad (3.6-4)$$

$$\triangleq G_A \quad (3.6-5)$$

from (2.6-4).

(4) To find the observation matrix H_A , we recall the following identities. Under equivalence transformation

$$F_A = T_A F T_A^{-1} \quad (3.6-6)$$

$$H_A = H T_A^{-1} \quad (3.6-7)$$

and from (2.6-9)

$$H F^p T_A^{-1} = [-A_p \quad -A_{p-1} \quad \dots \quad -A_1] \quad (3.6-8)$$

We define a new matrix of m rows and n columns

$$H_I \stackrel{\Delta}{=} \begin{bmatrix} 0 & I_m \end{bmatrix} \quad (3.6-9)$$

which has an $m \times m$ unit submatrix at the right and zeros elsewhere.

Comparing (3.6-8) and (2.6-6)

$$HF_A^p T_A^{-1} = H_I F_A \quad (3.6-10)$$

Now it is easy to establish that

$$F_A^p = T_A F_A^p T_A^{-1} \quad (3.6-11)$$

then
$$F_A^p = T_A^{-1} F_A^p T_A \quad (3.6-12)$$

Substituting (3.6-12) into (3.6-10) we find

$$H T_A^{-1} F_A^p T_A T_A^{-1} = H_I F_A \quad (3.6-13)$$

$$H T_A^{-1} = H_I F_A F_A^{-p} \quad (3.6-14)$$

Finally, with (3.6-7)

$$H_A = H_I F_A^{-p+1} \quad (3.6-15)$$

This completes the derivation of system matrices F_A , G_A and H_A in A-canonical form from description Σ_D .

3.6.2 An Algebraic Example

Consider the input-output description of a discrete-time system that corresponds to the algebraic example discussed in Subsection 2.6.2

$$\begin{aligned}
 \begin{bmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{bmatrix} &= -A_1 \begin{bmatrix} y_1(t-1) \\ y_2(t-1) \\ y_3(t-1) \end{bmatrix} - A_2 \begin{bmatrix} y_2(t-2) \\ y_3(t-2) \end{bmatrix} - A_3 y_3(t-3) \\
 &+ B_0 \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} + B_1 \begin{bmatrix} u_1(t-1) \\ u_2(t-1) \end{bmatrix} + B_2 \begin{bmatrix} u_1(t-2) \\ u_2(t-2) \end{bmatrix} + B_3 \begin{bmatrix} u_1(t-3) \\ u_2(t-3) \end{bmatrix}
 \end{aligned}
 \tag{3.6-16}$$

where A_1 is 3×3 , A_2 is 3×2 , A_3 is 3×1 and all of B_j are 3×2 .

(1) To find n , we total the columns in the matrix coefficients

A_1, A_2, A_3 :

$$n = 3 + 2 + 1 = 6 \tag{3.6-17}$$

We then proceed to imbed A_1, A_2 and A_3 in the last three rows of a 6×6 matrix F_A as in the example, equation (2.6-13).

(2) We know that the equations used to derive (3.6-16) originated with the set below which corresponds with the equation set (3.3-3) for the general case.

$$\begin{aligned}
y_3(t-3) &= h_3 x(t-3) \\
y_2(t-2) &= h_2 F x(t-3) + h_2 G u(t-3) \\
y_3(t-2) &= h_3 F x(t-3) + h_3 G u(t-3) \\
y_1(t-1) &= h_1 F^2 x(t-3) + h_1 F G u(t-3) + h_1 G u(t-2) \\
y_2(t-1) &= h_2 F^2 x(t-3) + h_2 F G u(t-3) + h_2 G u(t-2) \\
y_3(t-1) &= h_3 F^2 x(t-3) + h_3 F G u(t-3) + h_3 G u(t-2) \quad (3.6-18)
\end{aligned}$$

Therefore, we can immediately write down the algebraic form of T_A without being able to assign numerical values to its elements

$$T_A = \begin{bmatrix} h_3 \\ h_2 F \\ h_3 F \\ h_1 F^2 \\ h_2 F^2 \\ h_3 F^2 \end{bmatrix} \quad (3.6-19)$$

We can also write down the algebraic form of

$$F T_A^{-1} = [T_A F^{-1}]^{-1} \quad (3.6-20)$$

Finally, the algebraic expression for F_A can be written down exactly as in (2.6-12) with the pairs shown there. Then, with the lemma of

Subsection 2.4.1, the first three rows of F_A can be filled in as in (2.6-13), without ever calculating T_A explicitly.

(3) The lower triangular matrix T_P^* is constructed from the matrix coefficients A_j as in (3.4-5). It has $m(p+1) = 3(3+1) = 12$ rows and columns. Note how the submatrices repeat in T_P^* .

$$T_P^* = \begin{bmatrix} I_3 & 0 & 0 & 0 \\ A_1 & I_3 & 0 & 0 \\ A_2 \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & A_1 & I_3 & 0 \\ A_3 \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} & A_2 \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & A_1 & I_3 \end{bmatrix} \quad (3.6-21)$$

(4) We calculate the array

$$\begin{bmatrix} D \\ HG \\ HFG \\ HF^2G \end{bmatrix} = T_P^{*-1} \begin{bmatrix} B_0 \\ B_1 \\ B_2 \\ B_3 \end{bmatrix} \quad (3.6-22)$$

giving us D immediately. From the remaining array of 9 rows, we pick out rows 3, 5, 6, 7, 8 and 9 corresponding to the choice of rows for T_A in (3.6-19). Thus

$$G_A = \begin{bmatrix} h_3 \\ h_2^{FG} \\ h_3^{FG} \\ h_1^{F^2G} \\ h_2^{F^2G} \\ h_3^{F^2G} \end{bmatrix} \quad (3.6-23)$$

by elimination of rows calculated in (3.6-22) but without explicit calculation of H, F or G.

The method has been tested by construction and simulation on a digital computer. Models were constructed from each other in the following order:

$$\begin{array}{c} \Sigma(F,G,H) \\ \downarrow \\ \Sigma_A(F_A, G_A, H_A) \\ \downarrow \\ \Sigma_D(A,B) \\ \downarrow \\ \Sigma_A(F_A, G_A, H_A) \end{array}$$

The effect of round-off error was evident in the last model derived because the output trajectory observed sometimes showed departure in the seventh or eighth decimal place from that observed for the other models. This can be attributed to errors accumulated through the computation of a number of matrix inversions and the calculation of powers of matrices employed, all in single-precision arithmetic.

3.7 Matrix Transfer Function Description Σ_Z

3.7.1 Derivation of the Transfer Function

Given the Σ description (3.4-1 and 3.4-2) we can derive the matrix transfer function in the manner of Kalman [3] but with the addition of the direct control term D . Replacing d/dt by s , we write the formal Laplace transform (with zero initial conditions)

$$s\mathbf{x}(s) = \mathbf{F}\mathbf{x}(s) + \mathbf{G}u(s) \quad (3.7-1)$$

$$y(s) = \mathbf{H}\mathbf{x}(s) + Du(s) \quad (3.7-2)$$

From (3.7-1)

$$\mathbf{x}(s) = [s\mathbf{I}_n - \mathbf{F}]^{-1}\mathbf{G}u(s) \quad (3.7-3)$$

$$\mathbf{H}\mathbf{x}(s) = \mathbf{H}[s\mathbf{I}_n - \mathbf{F}]^{-1}\mathbf{G}u(s) - Du(s) + Du(s) \quad (3.7-4)$$

Then from (3.7-2)

$$y(s) = [\mathbf{H}[s\mathbf{I}_n - \mathbf{F}]^{-1}\mathbf{G} + D]u(s) \quad (3.7-5)$$

We shall call this set of equations description Σ_Z in the input and output variables. We define the $m \times r$ matrix transfer function

$$Z(s) \triangleq \mathbf{H}[s\mathbf{I}_n - \mathbf{F}]^{-1}\mathbf{G} + D \quad (3.7-6)$$

where only the D term is different from Kalman's [3]. (3.7-6) is the form considered by Rosenbrock in two recent papers on linear system transformations with the generalization, $D(s)$ in place of D [28, 29].

A direct way to find the poles and zeros of $Z(s)$ is to decompose the inverse matrix in (3.7-6) into its adjoint and determinant by the identity [22]

$$\begin{aligned} [sI_n - F][sI_n - F]^+ &= [sI_n - F]^+[sI_n - F] \\ &= |sI_n - F|I_n \end{aligned} \quad (3.7-7)$$

where $[sI_n - F]^+$ is the adjoint of $[sI_n - F]$ and $|sI_n - F|$ is its determinant. We can substitute for the inverse matrix in (3.7-6)

$$Z(s) = \frac{1}{|sI_n - F|} H[sI_n - F]^+G + D \quad (3.7-8)$$

Now [22]

$$|sI_n - F| = (-)^n |F - sI_n| \quad (3.7-9)$$

$$= (-)^n f(s) \quad (3.7-10)$$

where $f(s)$ is the characteristic polynomial of F defined by (2.4-8).

Substituting for $f(s)$, we get

$$|sI_n - F| = (-)^n (-)^n (s^n + \omega_1 s^{n-1} + \dots + \omega_n) \quad (3.7-11)$$

which factors for roots λ_j . Then

$$|sI_n - F| = \prod_{j=1}^n (s - \lambda_j) \quad (3.7-12)$$

Hence in (3.7-8)

$$Z(s) = \frac{1}{\prod_{j=1}^n (s - \lambda_j)} H[sI_n - F]^+G + D \quad (3.7-13)$$

3.7.2 Algebraic Example of a Matrix Transfer Function

It is not difficult to show that (3.7-13) is invariant for any basis of x in X . For our example, let us choose a system with three real and distinct roots and a basis such that F is in diagonal form. We assign some nonzero elements of G and H for simplicity, ensuring for this example that the system is both observable and controllable.

$$\begin{aligned}
 F &= \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}; & G &= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \\
 H &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}; & D &= \begin{bmatrix} 0 \\ 0 \end{bmatrix}
 \end{aligned} \tag{3.7-14}$$

The adjoint matrix is

$$[sI_3 - F]^+ = \begin{bmatrix} s-\lambda_1 & 0 & 0 \\ 0 & s-\lambda_2 & 0 \\ 0 & 0 & s-\lambda_3 \end{bmatrix}^+ \tag{3.7-15}$$

$$= \begin{bmatrix} (s-\lambda_2)(s-\lambda_3) & 0 & 0 \\ 0 & (s-\lambda_1)(s-\lambda_3) & 0 \\ 0 & 0 & (s-\lambda_1)(s-\lambda_2) \end{bmatrix} \tag{3.7-16}$$

Then

$$Z(s) = \frac{1}{(s-\lambda_1)(s-\lambda_2)(s-\lambda_3)} \begin{bmatrix} (s-\lambda_2)(s-\lambda_3) \\ 2(s-\lambda_1)(s-\frac{\lambda_2+\lambda_3}{2}) \end{bmatrix} \quad (3.7-17)$$

The system has a single pole between the input and the first output and a zero and two poles in the path to the second output. The calculation is tedious and not very well suited for routine computation on a digital computer because of the algebraic operations in finding the minors of the adjoint matrix. In general, Faddeev's modification of Leverrier's method [30] is preferable for expanding the adjoint and finding the coefficients of the characteristic polynomial, provided that the roots are distinct. A clear exposition is given by Zadeh and Desoer [25, p. 301].

3.7.3 The Transfer Function Description as the Laplace Transform of a Vector Differential Equation of Order n

Let us expand (3.7-5) as a matrix polynomial in s . Substituting for the inverse, we write

$$\begin{aligned} (s^n + \omega_1 s^{n-1} + \dots + \omega_n)y(s) \\ = [H[sI_n - F]^+G + (s^n + \omega_1 s^{n-1} + \dots + \omega_n)D]u(s) \end{aligned} \quad (3.7-18)$$

which is the Laplace transform of a nonhomogeneous vector differential equation of order n . To illustrate this in a simple case, let us

again consider a system with three real and distinct roots and a basis so that F is in diagonal form. We partition H into column vectors h_j and G into row vectors g_i^T .

$$F = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}; \quad G = \begin{bmatrix} g_1^T \\ g_2^T \\ g_3^T \end{bmatrix}$$

$$H = [h_1 \quad h_2 \quad h_3]; \quad D = D \quad (3.7-19)$$

Substituting for the adjoint matrix from (3.7-15)

$$H[sI_3 - F]^+G = (s - \lambda_2)(s - \lambda_3)h_1g_1^T$$

$$+ (s - \lambda_1)(s - \lambda_3)h_2g_2^T \quad (3.7-20)$$

$$+ (s - \lambda_1)(s - \lambda_2)h_3g_3^T$$

$$= s^2[h_1g_1^T + h_2g_2^T + h_3g_3^T]$$

$$- s[(\lambda_2 + \lambda_3)h_1g_1^T + (\lambda_1 + \lambda_3)h_2g_2^T + (\lambda_1 + \lambda_2)h_3g_3^T]$$

$$+ \lambda_2\lambda_3h_1g_1^T + \lambda_1\lambda_3h_2g_2^T + \lambda_1\lambda_2h_3g_3^T \quad (3.7-21)$$

We can now substitute with equations (3.7-18) and (3.7-21) for all the terms in the general matrix differential equation form for $n = 3$ and $p = 3$,

$$[s^3 + s^2A_1I_m + sA_2I_m + A_3I_m]y(s) = [s^3B_0 + s^2B_1 + sB_2 + B_3]u(s) \quad (3.7-22)$$

$$\begin{aligned}
 \text{where } A_1 &= \omega_1 = -(\lambda_1 + \lambda_2 + \lambda_3) \\
 A_2 &= \omega_2 = (\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1) \\
 A_3 &= \omega_3 = -\lambda_1\lambda_2\lambda_3
 \end{aligned} \tag{3.7-23}$$

$$B_0 = D$$

$$B_1 = HG + \omega_1 D$$

$$B_2 = -(\lambda_2 + \lambda_3)h_1g_1^T - (\lambda_1 + \lambda_3)h_2g_2^T - (\lambda_1 + \lambda_2)h_3g_3^T + \omega_2 D$$

$$B_3 = \lambda_2\lambda_3h_1g_1^T + \lambda_1\lambda_3h_2g_2^T + \lambda_1\lambda_2h_3g_3^T + \omega_3 D \tag{3.7-24}$$

In the above, the A_j and B_j terms were calculated by expanding the matrix transfer function without recourse to the earlier methods of the Chapter. Let us multiply out the following identity adapted from equation (3.4-6) of Subsection 3.4.2.

$$\begin{bmatrix} B_0 \\ B_1 \\ B_2 \\ B_3 \end{bmatrix} = \begin{bmatrix} I_m & 0 & \cdot & 0 \\ A_1 I_m & I_m & \cdot & \cdot \\ A_2 I_m & A_1 I_m & I_m & 0 \\ A_3 I_m & A_2 I_m & A_1 I_m & I_m \end{bmatrix} \begin{bmatrix} D \\ HG \\ HFG \\ HF^2G \end{bmatrix} \tag{3.7-25}$$

This yields

$$B_0 = D$$

$$B_1 = HG + \omega_1 D$$

(cont. over)

$$\begin{aligned}
B_2 &= \omega_2 D \\
&\quad - (\lambda_1 + \lambda_2 + \lambda_3)[h_1 g_1^T + h_2 g_2^T + h_3 g_3^T] \\
&\quad + \lambda_1 h_1 g_1^T + \lambda_2 h_2 g_2^T + \lambda_3 h_3 g_3^T \\
&= -(\lambda_2 + \lambda_3)h_1 g_1^T - (\lambda_1 + \lambda_3)h_2 g_2^T - (\lambda_1 + \lambda_2)h_3 g_3^T + \omega_2 D \\
B_3 &= \text{etc.} \tag{3.7-26}
\end{aligned}$$

The coefficients are the same as in (3.7-23).

Without completely formalizing this discussion, we tentatively conclude that

- (1) the matrix transfer function description Σ_Z is the Laplace transform of a nonhomogeneous vector differential equation of order n
- (2) the description Σ_Z can be calculated from a state variable description without requiring the algebraic expansion of the inverse of a matrix polynomial or the minors for an adjoint matrix. The steps are shown in the next subsection.

3.7.4 Calculating the Matrix Transfer Function Corresponding to Σ (F,G,H,D)

Given a description Σ with system matrices F , G , H and D , we seek the transfer function $Z(s)$ of description Σ_Z .

- (1) The companion matrix for F is found. (See, for example, Subsection 2.4.2). This provides us with the coefficients ω_j of

the characteristic equation of F . (This approach for finding the coefficients ω_j is not recommended for systems of high order because of the accumulated effect of numerical round-off when calculating the higher powers of F . The Leverrier-Faddeev method cited before [30] is preferable.)

(2) The array T_Z is calculated:

$$T_Z = \begin{bmatrix} I_m & 0 & \cdot & \cdot & 0 \\ \omega_1 I_m & I_m & \cdot & \cdot & \cdot \\ \omega_2 I_m & \omega_1 I_m & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0 \\ \omega_n I_m & \omega_{n-1} I_m & \cdot & \omega_1 I_m & I_m \end{bmatrix} \quad (3.7-27)$$

(3) The $m(n+1) \times r$ array below is calculated:

$$\begin{bmatrix} B_0 \\ B_1 \\ \cdot \\ \cdot \\ B_n \end{bmatrix} = T_Z \begin{bmatrix} D \\ HG \\ \cdot \\ \cdot \\ HF^{n-1}G \end{bmatrix} \quad (3.7-28)$$

(4) The terms are collected in the following expansion for

$Z(s)$:

$$Z(s) = \frac{1}{s^n + \omega_1 s^{n-1} + \dots + \omega_n} [s^n B_0 + s^{n-1} B_1 + \dots + B_n] \quad (3.7-29)$$

The general element $z_{ij}(s)$ of $Z(s)$ is:

$$z_{ij}(s) = \frac{1}{s^n + \omega_1 s^{n-1} + \dots + \omega_n} (s^n b_{0ij} + s^{n-1} b_{1ij} + \dots + b_{nij}) \quad (3.7-30)$$

where b_{kij} is the element at the intersection of the i^{th} row and j^{th} column of B_k .

If Σ is a minimal realization, there will be no cancellation of zeros with poles. Then, in the Σ_Z description there are $(n+1)mr$ elements b_{kij} and n elements ω_j totalling $n(mr+1) + mr$ parameters in all. The last mr terms are for the direct control term $B_0 = D$.

3.8 The Problem of Parametrization

3.8.1 A Comparison of Some Descriptions

By now the reader will be aware that the number of parameters in any one description may differ from the number required in another description, even though the descriptions are input-output equivalent and are derived for some abstract system Σ . For example, neglecting the D or B_0 term of mr parameters which is invariant for any description discussed, we can compare the number of parameters in the following Table for some descriptions of a completely observable system. A range in the number of parameters for some descriptions is shown to account for special situations described in notes with the Table.

TABLE 3.1 The Number of Parameters in Some System Descriptions

<u>Description</u>	<u>Number of Parameters</u>	<u>Notes</u>
Basic state-variable Σ	$n(m+r)$ to $n^2 + n(m+r)$	In the state coordinates of a physical system.
Minimum, normal state-variable Σ_m	$n(m+r)$	Obtained by a change of basis in the state space.
A-canonical form Σ_A	$n(m+r)$ to $n(2m+r-1)$	Minimum only for the case that $n = mp$. (p is the observability index.)
Matrix diff. equation Σ_D of order p	$mn + mpr$	$mn + mpr = n(m+r)$ in the case that $n = mp$.
Matrix transfer function Σ_Z	$n(mr+1)$	Σ is a minimal realization.

3.8.2 An Algebraic and Geometric Interpretation of the Problem

The problem of parameterization first discussed by Kalman [3] is concerned with finding an "efficient" description of a system. A system is not efficient in a quantitative sense if another description can be found with fewer parameters. Then the parameters of an inefficient description must be algebraic functions of the fewer parameters of a more efficient description. For example, we can find $n^2 + n(m+r)$ functions for the same number of parameters e of Σ ($D = 0$) in terms of the $N_m = n(m+r)$ parameters ϕ of Σ_m . Algebraically, this yields a degenerate [3] set of equations

$$\begin{aligned} e_1 &= e_1(\phi_1, \phi_2, \dots, \phi_{N_m}) \\ e_2 &= e_2(\phi_1, \phi_2, \dots, \phi_{N_m}) \\ \cdot &\quad \cdot \\ e_{n^2+n(m+r)} &= e_{n^2+n(m+r)}(\phi_1, \phi_2, \dots, \phi_{N_m}) \end{aligned} \quad (3.8-1)$$

$e \in \Theta$, a Euclidean space of $n^2 + n(m+r)$ dimensions

$\phi \in \Phi$, a Euclidean space of only $n(m+r)$ dimensions

The geometric interpretation in the space Θ , is that the system is defined for a point anywhere on a surface of $n(m+r)$ dimensions. The solution to the problem of parameterization in this case is that of moving on the surface to preferred coordinates. This can be accomplished for state-space models by a change in basis of the intermediate variables, the state variables. If we assume that the normal

form Σ_m is an efficient state-space description, then n^2 of the components of e on the left hand side of equation (3.8-1) are equal to a constant or zero.

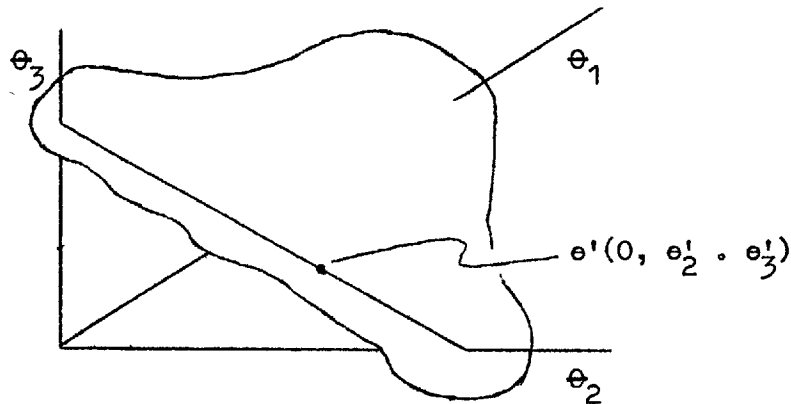


Figure 3.1 Coordinate Surface for e

Figure 3.1 illustrates a surface of two dimensions in the space Θ of three dimensions. If by some means we can make a component of e , $e_1 = 0$, then e is constrained to the line on which e' is shown. There is a one-to-one correspondence between any point e' on the line and ϕ in the two dimensioned space $\bar{\mathcal{D}}$.

The important point to be made here is that for a state-variable description Σ , e is a unique point in Θ only if a basis is chosen (which fixes the relations between parameters) or sufficient appropriate parameters are set to arbitrary quantities. Hence one or other of the above requirements must be satisfied in order to determine a unique e from a noise-free record of input and output data ($m < n$).

Now let us consider the Σ_D description. We have noted that this description is as efficient as the Σ_m description only if $n = mp$. For the case $n \neq mp$, e is still unique for a particular selection of components of the derivatives of the output in the continuous case (or specification of a particular set of past outputs in the discrete case) even though it is not efficient. Hence, we should be able to determine a unique e for Σ_D from a noise-free record of data.

CHAPTER 4

MULTIVARIATE STOCHASTIC MODELS FOR IDENTIFICATION, PREDICTION AND
CONTROL4.1 Introduction

In this Chapter, the deterministic system descriptions derived in the last two Chapters are extended to include the effect of additive stochastic disturbances. We continue to assume that the multivariate process to be identified and controlled is linear and add the basic premise that the vector observations on the process inputs and outputs are made at discrete instants of time. The theoretical and practical reasons for departing from the study of continuous time processes are:

- (1) the exact response of the system at the sampling instants can be described satisfactorily by difference equations instead of differential equations;
- (2) the problem of taking derivatives with respect to time of the stochastic process that is an input to the system is avoided;
- (3) there is an extensive body of applied statistical estimation theory that can be employed directly and conveniently in solving problems concerning physical experiments which are devised so that their outcomes are discrete random events;
- (4) in direct digital control practice, plant observation points are sampled periodically and control is adjusted in discrete steps.

In most industrial plants, the time constants are sufficiently long that a central digital processor can monitor and adjust a number of subsystems in sequence. Discrete-time models are ideally suited to being updated periodically and employed in control schemes in which corrective action is generated in steps.

4.2 A General Discrete-Time State-Variable Model of a Stochastic Process

Let us assume that a linear, constant, multivariate, stochastic process is modelled (at uniformly spaced, integer sampling instants) by the general discrete-time, state-variable description Σ_s below.

$$x(t+1) = Fx(t) + Gu(t) + \Gamma v_1(t) \quad (4.2-1)$$

$$y(t) = Hx(t) + Du(t) + v_2(t) \quad (4.2-2)$$

where the observation y is an m -vector and the control u is an r -vector as before. x is an n -vector of state variables of both the physical process and a "coloured"-noise process [24].

In this model, $v_1(t)$ is a ρ -vector of random variables drawn, at time t , from a ρ -variate normal distribution with statistical characteristics given below. The observation vector, $v_2(t)$ is an m -vector, of random noise variables, also drawn from a multivariate normal distribution. The noise processes have the following statistical characteristics.

$$E[v_1(t)] = 0 \quad (4.2-3)$$

$$E[v_1(t), v_1^T(\tau)] = V_{11}\delta(t-\tau) \quad (4.2-4)$$

$$E[v_2(t)] = 0 \quad (4.2-5)$$

$$E[v_2(t), v_2^T(\tau)] = V_{22}\delta(t-\tau) \quad (4.2-6)$$

$$E[v_1(t), v_2^T(\tau)] = V_{12}\delta(t-\tau) \quad (4.2-7)$$

where $E[\cdot] \triangleq$ expected value of $[\cdot]$

V_{11}, V_{22} are positive definite correlation matrices

$$\begin{aligned} \delta(t-\tau) &= 1 ; & t &= \tau \\ &= 0 ; & t &\neq \tau \end{aligned} \quad (4.2-8)$$

Let us now look at the structure of the stochastic model in more detail and distinguish pure measurement noise from noise disturbing states that are also under the influence of the control. Suppose that we could partition F as follows:

$$F = \begin{bmatrix} F_a & 0 \\ 0 & F_b \end{bmatrix} \quad (4.2-9)$$

where F_a is a square matrix of n_a rows and F_b is a square matrix of $n_b = n - n_a$ rows. If the model is a minimal realization (the dimension of the state vector cannot be reduced [3]), there will be no common factors between the characteristic polynomials of both F_a and F_b . In addition, we state that G has zero elements in its

last n_b rows and that there is at least one nonzero element in each of the last n_b rows of Γ . Then the final n_b states of the model belong to a measurement noise process which is distinct from the operation of the remainder of the plant model. This concept is illustrated in Figure 4.1. It is not uncommon to find this structure in industrial process situations where the measuring instruments are known to be inherently noisy.

The purpose in distinguishing measurement noise from the rest of the process is to draw attention to the fact that, in the general case, only some of the system states in the model are under the direct influence of the control. For purposes of plant identification, we are only concerned with

- (1) system states that are observable
- (2) system states that are controllable from either the control input or the noise process or both together.

The conditions for observability to which we referred in previous chapters are unaltered for use with this model but the conditions for controllability have to be modified because the model has more inputs than from the control alone. Thus we introduce two new assertions.

(a) Assertion: The system is controllable by the control input and noise input together if, and only if, the array

$$c_p^P = [\mathcal{Y} \quad F\mathcal{Y} \quad F^2\mathcal{Y} \quad \dots \quad F^{P-1}\mathcal{Y}] \quad (4.2-10)$$

is of rank n where

$$\mathcal{Y} \triangleq [G \quad \Gamma] \quad (4.2-11)$$

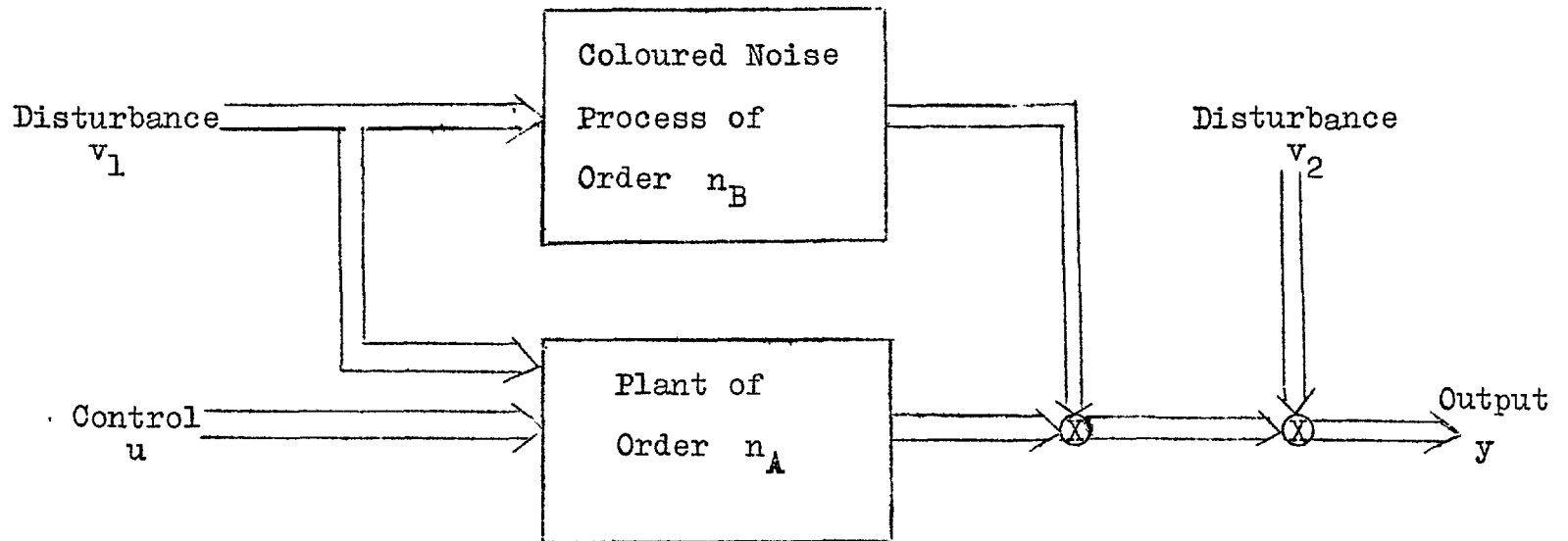


Figure 4.1 Plant With Coloured Measurement Noise

and the controllability index p is defined by

$$p \leq \min(n_m, n-r-\rho+1) \quad (4.2-12)$$

Proof: The assertion is sufficiently close to the theorem of Section 2.3 that a formal proof is not warranted. We first define a vector of dimension $(r + \rho)$

$$v = \begin{bmatrix} u \\ v_1 \end{bmatrix} \quad (4.2-13)$$

The problem then is to show that the image of the vector space $N(v \in N)$ of dimension $(r + \rho)$ spans the space of x provided that c_p^P is of rank n . By the Theorem of Section 2.3, the assertion is proved.

(b) Assertion: Given the system description Σ_s by equations (4.2-1) and (4.2-2) and the conditions of equations (4.2-3) to (4.2-7), then the system is output controllable [26] in a mean sense, if the array o_q^P defined by

$$o_q^P = [HG \quad HFG \quad \dots \quad HF^{q-1}G \quad D] \quad (4.2-14)$$

has rank m for some positive integer q of least magnitude such that

$$q \leq \min(n_m, n-r+1) \quad (4.2-15)$$

Definition: By output controllable in a mean sense, we imply that the expected value of any final output $y(k)$ for any $k \leq q$ can be

reached, starting with arbitrary initial conditions in the plant at $t = 0$.

Proof: We adapt the result of Kreindler and Sarachik [26] to the special model considered here and introduce a controllability index q into the proof.

The expected value of the state vector of the system Σ_s at every time interval, starting from $t = 0$, is given by

$$\begin{aligned} Ex(1) &= FEx(0) + Gu(0) \\ Ex(2) &= F^2Ex(0) + FGGu(0) + Gu(1) \\ &\quad \cdot \quad \cdot \quad \cdot \\ Ex(q) &= F^qEx(0) + F^{q-1}Gu(0) + \dots + FGGu(q-2) + Gu(q-1) \end{aligned} \tag{4.2-16}$$

where E is the expectation operator. Premultiplying (4.2-16) by H and combining it with (4.2-2), we obtain

$$Ey(q) - HF^qEx(0) = [HG \quad HFG \quad \dots \quad HF^{q-1}G \quad D] \begin{bmatrix} u(q-1) \\ u(q-2) \\ \cdot \\ \cdot \\ u(1) \\ u(0) \\ u(q) \end{bmatrix} \tag{4.2-17}$$

We want to determine m elements of the $r(q+1)$ vector of controls in (4.2-17) so that we can reach the expected value of any arbitrary

output $Ey(q)$ given the expected value of the initial state $Ex(0)$.

Clearly, the condition for a unique solution to exist is that rank

$$O^P_q = m \text{ for some smallest positive integer } q.$$

To establish a lower bound on q , we note that if rank $D = m$, it is sufficient that $q = 1$. For the upper bound, we take

$q = \min(n_m, n-r+1)$, the same upper bound employed in the theorem on controllability discussed in Section 2.3. This ends the proof.

Suppose that one is required to calculate the control sequence to take (the expected value of) the output to some specified point, given only observations on the input and output. We can see from the last proof that the (states of the) plant must be observable (but not necessarily controllable by the input) in order that a solution to the control problem can be found.

We conclude this section with the assumptions that description

Σ_s is

- (1) controllable by the control and noise inputs together
- (2) observable
- (3) output controllable by the control input u .

4.3 Transformation to Stochastic Vector Difference Equation Form

4.3.1 Introduction

In this section, we shall transform the stochastic state-variable description Σ_s , defined in the last section, into a stochastic vector difference equation description in the output variables y , control

variables u and noise variables v_1 and v_2 . This is a direct application of the procedures developed in the last chapter. We shall find that, given observations of the control and output only, we are unable to distinguish between the effects of v_1 and v_2 . This will lead us to a new description, Σ_{DS} , in the output y , control u and a single vector noise e .

Before we write down the description Σ_{DS} , it will be convenient to introduce a slight modification in the convention used in previous chapters for the difference equation description. We had (3.2-9)

$$A(z) \triangleq z^p I_m + z^{p-1} A_1 + \dots + A_p (I_m)_p \quad (4.3-1)$$

Then

$$A(z^{-1}) = z^{-p} I_m + z^{-p+1} A_1 + \dots + A_p (I_m)_p \quad (4.3-2)$$

Let us define

$$A^*(z^{-1}) \triangleq z^p A(z) \quad (4.3-3)$$

$$= I_m + z^{-1} A_1 + \dots + z^{-p} A_p (I_m)_p \quad (4.3-4)$$

and

$$B^*(z^{-1}) \triangleq z^p B(z) \quad (4.3-5)$$

$$= B_0 + z^{-1} B_1 + \dots + z^{-p} B_p \quad (4.3-6)$$

Then, if

$$A(z)y(t) = B(z)u(t) \quad (4.3-7)$$

it follows that

$$A^*(z^{-1})y(t) = B^*(z^{-1})u(t) \quad (4.3-8)$$

But for convenience we will drop the asterisk, that is, we mean equation (4.3-8) if we write

$$A(z^{-1})y(t) = B(z^{-1})u(t) \quad (4.3-9)$$

This convention will be used in the Σ_{DS} description that follows here.

$$A(z^{-1})y(t) = B(z^{-1})u(t) + C(z^{-1})\Lambda^{1/2}e(t) \quad (4.3-10)$$

where $y(t)$ is an m vector of the system output at time t , $u(t)$ is an r vector of controls and $e(t)$ is an m -vector of random variables from an m -variate normal distribution such that

$$E[e(t)] = 0 \quad (4.3-11)$$

$$E[e(t), e^T(t)] = I_m \quad (4.3-12)$$

$$E[e(t), e^T(\tau)] = I_m \delta(t-\tau) \quad (4.3-13)$$

where $\delta(t-\tau)$ was defined by equation (4.2-8).

$\Lambda^{1/2}$ is an $m \times m$ symmetric matrix that will be specified later in the derivation of the model. Finally,

$$A(z^{-1}) = I_m + z^{-1}A_1 + \dots + z^{-p}A_p (I_m)_p \quad (4.3-14)$$

$$B(z^{-1}) = B_0 + z^{-1}B_1 + \dots + z^{-p}B_p \quad (4.3-15)$$

$$C(z^{-1}) = I_m + z^{-1}C_1 + \dots + z^{-p}C_p \quad (4.3-16)$$

In the case that $m = r = 1$, this model reduces to the single-input, single-output model of Astrom, Bohlin and Wensmark [19]. Each of the coefficients A_j , B_j and C_j are then scalars and $\Lambda^{1/2}$ is a scaling constant so that the independent noise source has a unit variance. In the absence of control, this case is called a mixed autoregressive moving average model by Box and Jenkins [33].

4.3.2 The Stochastic Difference Equation Form in Both Noise Variables

The difference equation obtained by transformation with the procedures developed in Chapter 3 has the form

$$A(z^{-1})y(t) = B(z^{-1})u(t) + K(z^{-1})v_1(t) + L(z^{-1})v_2(t) \quad (4.3-17)$$

Note that (4.3-17) involves both v_1 and v_2 .

The computational procedure to find $A(z^{-1})$ and $B(z^{-1})$ is unaffected by the extension to include stochastic disturbances. That is, the matrix coefficients of $A(z^{-1})$ can be found by the procedure of Subsection 3.2.1 and the coefficients of $B(z^{-1})$ by the method of Subsection 3.4.2.

The coefficients of $K(z^{-1})$ are obtained by step (4) of Subsection 3.2.1.

$$K(z^{-1}) \triangleq z^{-1}K_1 + z^{-2}K_2 + \dots + z^{-p}K_p \quad (4.3-18)$$

where

$$\begin{bmatrix} K_1 \\ K_2 \\ \cdot \\ K_p \end{bmatrix} = T_p^* \sigma_p^p \Gamma \quad (4.3-19)$$

from (3.2-5). The coefficients of $L(z^{-1})$ can be obtained from (3.4-6) by setting

$$\begin{bmatrix} L_0 \\ L_1 \\ L_2 \\ \cdot \\ L_p \end{bmatrix} = T_p^* \begin{bmatrix} I_m \\ 0 \\ 0 \\ \cdot \\ 0 \end{bmatrix} \quad (4.3-20)$$

where the null matrices signify that the noise v_2 does not disturb the system states. Then

$$L(z^{-1}) \triangleq L_0 + z^{-1}L_1 + z^{-2}L_2 + \dots + z^{-p}L_p \quad (4.3-21)$$

$$= I_m + z^{-1}A_1 + z^{-2}A_2(I_m)_2 + \dots + z^{-p}A_p(I_m)_p \quad (4.3-22)$$

$$= A(z^{-1}) \quad (4.3-23)$$

by comparison with (4.3-14).

At time t , the disturbance in the difference equation form is obtained by subtracting $B(z^{-1})u(t)$ from both sides of equation (4.3-17).

$$v(t) \triangleq A(z^{-1})y(t) - B(z^{-1})u(t) \quad (4.3-24)$$

$$= K(z^{-1})v_1(t) + L(z^{-1})v_2(t) \quad (4.3-25)$$

$$= [K(z^{-1}) \ A(z^{-1})] \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} \quad (4.3-26)$$

$$= \begin{bmatrix} 0 & I_m \end{bmatrix} \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} + \begin{bmatrix} K_1 & A_1 \end{bmatrix} \begin{bmatrix} v_1(t-1) \\ v_2(t-1) \end{bmatrix} + \dots + \begin{bmatrix} K_p & A_p(I_m)_p \end{bmatrix} \begin{bmatrix} v_1(t-p) \\ v_2(t-p) \end{bmatrix} \quad (4.3-27)$$

We define the unnormalized matrix autocorrelation with delay τ , $R_v(\tau)$, of the stochastic process $v(t)$ by the following relation.

$$R_v(\tau) \triangleq E v(t) v^T(t-\tau) \quad (4.3-28)$$

where τ is an integer as is t . For example, $R_v(0)$ and $R_v(1)$ for the simple case in which $p = 1$ are given by

$$R_v(0) = E \begin{bmatrix} 0 & I_m & K_1 & A_1 \end{bmatrix} \begin{bmatrix} v_1(t) \\ v_2(t) \\ v_1(t-1) \\ v_2(t-1) \end{bmatrix} \begin{bmatrix} v_1(t) & v_2(t) & v_1(t-1) & v_2(t-1) \end{bmatrix} \begin{bmatrix} 0 \\ I_m \\ K_1^T \\ A_1^T \end{bmatrix}$$

$$= \begin{bmatrix} 0 & I_m & K_1 & A_1 \end{bmatrix} \begin{bmatrix} V_{11} & V_{12} & 0 & 0 \\ V_{21} & V_{22} & 0 & 0 \\ 0 & 0 & V_{11} & V_{12} \\ 0 & 0 & V_{21} & V_{22} \end{bmatrix} \begin{bmatrix} 0 \\ I_m \\ K_1^T \\ A_1^T \end{bmatrix} \quad (4.3-29)$$

where the covariance matrices V_{11} , V_{22} , V_{12} were defined in (4.2-4) to (4.2-7) and $V_{21} = V_{12}^T$.

$$R_v(1) = \begin{bmatrix} 0 & I_m & K_1 & A_1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ V_{11} & V_{12} & 0 & 0 \\ V_{21} & V_{22} & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ I_m \\ K_1^T \\ A_1^T \end{bmatrix} \quad (4.3-30)$$

In the general case, $R_v(0)$ is an $m \times m$ symmetric matrix and $R_v(\tau)$, ($1 \leq \tau \leq p$), is also $m \times m$ but not symmetric. The autocorrelation of the process $v(t)$ is completely specified by $m(m+1)/2$ parameters in $R_v(0)$ and m^2 parameters in each matrix $R_v(\tau)$ for $\tau = 1, 2, \dots, p$. $R_v(\tau) = 0$ for $\tau > p$. We shall refer to the sequence $R_v(0), R_v(1), \dots, R_v(p)$ as the serial autocorrelations of $v(t)$.

4.3.3 The Equivalence of Stochastic Processes in a Statistical Sense

The vector stochastic process $v(t)$ defined by equation (4.3-24) is a linear function of the vector random variables $\{v_1(\tau), v_2(\tau)\}$ for $\tau = t, t-1, \dots, t-p$. Since v_1 and v_2 are normally distributed, $v(t)$ is from a normal distribution. Now a Gaussian process is completely

specified by all of its first and second moments. We have noted that $v(t)$ has zero first moments and $m(m+1)/2 + pm^2$ second moments.

Consider a new vector stochastic process $v'(t)$. We define the process $v'(t)$ to be statistically equivalent to $v(t)$ if all the moments of $v'(t)$ are equal to the moments of $v(t)$.

The problem with which we are concerned can now be stated: $v(t)$ is a Gaussian process having rational (discrete) spectra; all of its (finite number of) moments are known. Can we find the parameters of a Gaussian process $v'(t)$ such that $v'(t)$ and $v(t)$ are equivalent in a statistical sense?

Since

$$Ev'(t) = Ev(t) \quad (4.3-31)$$

(by assumption), we satisfy the above requirement if

$$R_{v'}(\tau) = R_v(\tau) \quad (4.3-32)$$

for $\tau = 0, \pm 1, \pm 2, \dots, \pm p$ and

$$R_{v'}(\tau) = 0 \quad (4.3-33)$$

for $|\tau| > 0$.

Suppose that $v'(t)$ is given by :

$$v'(t) = (I_m + z^{-1}C_1 + \dots + z^{-p}C_p) \Lambda^{1/2} e(t) \quad (4.3-34)$$

which has the same number of degrees of freedom as moments in choosing the coefficients of C_j ($j = 1, 2, \dots, p$) and $\Lambda^{1/2}$, provided that

$\Lambda^{1/2}$ is symmetric. A convenient way of determining the serial auto-correlations for the process $v'(t)$ is to compute the function

$$\mathcal{R}_{v'}(z) \triangleq C(z)\Lambda C^T(z^{-1}) \quad (4.3-35)$$

$$= (I_m + zC_1 \dots + z^p C_p)\Lambda(I_m + z^{-1}C_1 \dots + z^{-p}C_p)^T \quad (4.3-36)$$

$$= R_{v'}(0) + zR_{v'}(1) + z^{-1}R_{v'}(-1) + \dots \\ + z^p R_{v'}(p) + z^{-p}R_{v'}(-p) \quad (4.3-37)$$

by the definition of equation (4.3-28). For example, if $p = 2$,

$$\mathcal{R}_{v'}(z) = (I_m + zC_1)\Lambda(I_m + z^{-1}C_1)^T \\ = \Lambda + C_1\Lambda C_1^T + zC_1\Lambda + z^{-1}(C_1\Lambda)^T \quad (4.3-38)$$

from which

$$R_{v'}(0) = \Lambda + C_1\Lambda C_1^T \\ R_{v'}(1) = C_1\Lambda \\ R_{v'}(-1) = \Lambda C_1^T \\ = R_{v'}^T(1) \quad (4.3-39)$$

To find $C(z)$ (or $C(z^{-1})$) and Λ , given $\mathcal{R}_{v'}(z)$ only, is the problem of spectral factorization. The usual difficulty of this problem is to find factors with stable roots (within the unit circle for discrete-time analysis). Recently, Ho and Kalman have described

an algorithm for spectral factorization (in the scalar, continuous-time case) using the Riccati equation [57]. Their iterative method commences with an arbitrary, stable initial estimate.

The problem in the multivariate case has been considered for continuous-time systems by Youla [58] and a general solution given by Davis [59]. More recently, Anderson has presented an algebraic solution to the multivariate spectral factorization problem [60; 61] and briefly compared the solution method with that obtained by using the Riccati equation.

Because the spectral factorization solutions referred to above are too lengthy to record here, we shall simply note the conditions under which $\mathcal{R}(z)$ can be factorized in the discrete-time case and then present a direct method for finding the parameters of the factors. The conditions are:

$\mathcal{R}(z)$ is a rational function in z subject to

$$(a) \quad \mathcal{R}(z) = \mathcal{R}^T(z^{-1}) \quad (4.3-40)$$

$$(b) \quad R_v(0) \geq 0 \quad (4.3-41)$$

Condition (a) is satisfied if the autocorrelation terms

$$R_v(\tau) = R_v^T(-\tau) \quad (4.3-42)$$

The second condition implies that the power gain of the process $v(t)$ is non-negative.

The most direct method of finding the parameters of $C(z^{-1})$ and $\Lambda^{1/2}$ is to employ the method of moments: Suppose that the

probability density of a random variable ϕ is a function of n parameters e_i , $f(\phi; e_1, \dots, e_n)$. The population moments are

$$M_j(e_1, \dots, e_n) = \int \phi^j f(e_1, \dots, e_n) d\phi \quad (4.3-43)$$

In this univariate case, we can solve for e by computing as many moments M_j as there are unknown parameters e_i , $i = 1, 2, \dots, n$.

We generalize this method to the multivariate case by choosing the coefficients of $C(z^{-1})$ and $\underline{\Lambda}$ to satisfy all the moment equations given by (4.3-32). (The first moments are all assumed zero.) Then $v'(t)$ is statistically equivalent to $v(t)$. Note that if a solution is found by the method of moments, the conditions of (4.3-40) and (4.3-41) are satisfied. That is, if we substitute z^{-1} for z in equation (4.3-37), equation (4.3-40) is satisfied. Also, if we satisfy

$$R_{v'}(\tau) = R_v(\tau) \quad (4.3-44)$$

for all τ , and $R_v(0) \geq 0$; then $R_{v'}(0) \geq 0$.

The task of calculating the coefficients of C_1, C_2, \dots, C_p , $\underline{\Lambda}^{1/2}$ given A_1, \dots, A_p and K_1, K_2, \dots, K_p is one of choosing the coefficients so that equations (4.3-31) to (4.3-33) are satisfied. There are $m(m+1)/2 + pm^2$ independent equations ($R_v(0)$ is symmetric) to be satisfied by the same number of coefficients. Unfortunately, because the equations are nonlinear, the solution by the method of moments is not unique.

The following iterative methods of determining the coefficients of C_1, C_2, \dots, C_p , and $\underline{\Lambda}^{1/2}$ that satisfy the above conditions has been

used on a number of occasions without difficulty.

(1) The elements of the autocorrelation matrices $R_v(\tau)$ of the process $v(t)$ are calculated for all integers $\tau = 0, 1, \dots, p$ as in the example following equation (4.3-28)

$$R_v(\tau) \triangleq E v(t)v^T(t-\tau) \quad (4.3-45)$$

(2) We define a process

$$v'(t) = [C'_0 + z^{-1}C'_1 + \dots + z^{-p}C'_p]e(t)$$

where C'_0 is symmetric. Then the algebraic relations between the autocorrelation matrices $R_{v'}(\tau)$ and the matrix coefficients C'_j , $j = 0, 1, \dots, p$ are found. For example, if $p = 2$ and $m = 2$,

$$R_{v'}(0) = C'_0{}^2 + C'_1C'_1{}^T + C'_2C'_2{}^T \quad (4.3-46)$$

$$R_{v'}(1) = C'_1C'_0 + C'_2C'_1{}^T \quad (4.3-47)$$

$$R_{v'}(2) = C'_2C'_0 \quad (4.3-48)$$

In this case, we have 11 equations in 11 coefficients of C'_j .

(3) A scalar cost function is devised of the form

$$\begin{aligned} fc = & \sum_{\substack{i=1,m \\ j=1,i}} [(r_{v'.ij}(0) - r_{v'.ij}(0))]^2 \\ & + \sum_{\substack{i=1,m \\ j=1,m \\ k=1,p}} [(r_{v'.ij}(k) - r_{v'.ij}(k))]^2 \end{aligned} \quad (4.3-49)$$

where $r_{v,ij}(k)$ is the element of the i^{th} row and j^{th} column of $R_v(k)$ and $r_{v',ij}(k)$ is an element of $R_{v'}(k)$ located at the same place.

(4) An iterative numerical method is employed for choosing the coefficients to minimize f_c . The algorithm used by the author was that of Fletcher and Powell [31]. It requires algebraic expressions for the gradient of the cost function with respect to each of the elements of C'_0, C'_1, \dots, C'_p . Recent experience indicates that the algorithm developed by Powell [32], requiring only function evaluations, is equally effective (both are second order methods) and easier to program for a digital computer.

To initiate the computation, it has been found convenient to choose $C'_0 = [R_v(0)]^{1/2}$ and $C'_j = 0$ for $j = 1, \dots, p$. In the examples considered, the zeros of $z^p C'(z^{-1})$ have, fortunately, settled inside the unit circle in the z plane for this starting condition. It is clearly implied by the references on spectral factorization that solutions with zeros outside the unit circle can occur, yet the moment relations are satisfied. Hence the solution obtained by this method is not unique.

(5) Finally, the required parameters are found by the following identities:

$$C_1 \triangleq C'_1 C'_0{}^{-1} \quad (4.3-50)$$

•
•
•

$$C_j \triangleq C'_j C'_0{}^{-1} \quad (4.3-51)$$

and
$$\Lambda^{1/2} = C'_0 \quad (4.3-52)$$

We then substitute the parameters C_j determined in the above procedure into the identity, equation (4.3-16) recalled here.

$$C(z^{-1}) = I_m + z^{-1}C_1 + \dots + z^{-p}C_p \quad (4.3-53)$$

This concludes the derivation of the Σ_{DS} description, a stochastic difference equation defined by

$$A(z^{-1})y(t) = B(z^{-1})u(t) + C(z^{-1})\Lambda^{1/2}e(t) \quad (4.3-54)$$

which has a noise process statistically equivalent to the noise process contained in equation (4.3-17).

4.3.4 Some Special Cases

In the development of the Σ_{DS} description, the presence of two disturbances v_1 and v_2 was assumed. v_1 corrupted the system states and v_2 appeared as measurement noise. Let us consider some special cases that may arise.

(1) $\Gamma = 0$; the system states are noise free. Then from (4.3-17) and (4.3-23)

$$A(z^{-1})y(t) = B(z^{-1})u(t) + A(z^{-1})v_2(t) \quad (4.3-55)$$

We define

$$A(z^{-1})v_2(t) \triangleq C(z^{-1})\Lambda^{1/2}e(t) \quad (4.3-56)$$

where

$$\begin{aligned}
 C_1 &= A_1 \Lambda^{-1/2} \\
 C_2 &= A_2 (I_m)_2 \Lambda^{-1/2} \\
 &\vdots \\
 C_p &= A_p (I_m)_p \Lambda^{-1/2}
 \end{aligned}
 \tag{4.3-57}$$

and

$$\Lambda^{1/2} = V_{22}^{1/2}
 \tag{4.3-58}$$

(2) Measurement noise $v_2 = 0$. Then from (4.3-17)

$$A(z^{-1})y(t) = B(z^{-1})u(t) + K(z^{-1})v_1(t)
 \tag{4.3-59}$$

and from (4.3-25) and (4.3-18)

$$v(t) = (z^{-1}K_1 + z^{-2}K_2 + \dots + z^{-p}K_p)v_1(t)
 \tag{4.3-60}$$

If we are given a section of record of observations of the process $v(t)$, we are unable to distinguish the sequence in (4.3-60) from

$$v(t) = (K_1 + z^{-1}K_2 + \dots + z^{-p+1}K_p)v_1(t)
 \tag{4.3-61}$$

because we will discover that

$$R_v(p) = 0
 \tag{4.3-62}$$

Hence, in this special case, when fitting matrix coefficients

C_1, C_2, \dots, C_p of $C(z^{-1})$ to the matrix autocorrelations $R_v(0), R_v(1), \dots, R_v(p-1)$ we set $C_p = 0$. Note that in this special case, if $q < m$, every autocorrelation term $R_v(\tau)$ will be singular (and of rank $q < m$). Hence fitting the coefficients will take special consideration.

4.4 Transformation from the Σ_{DS} Description to State-Variable Form

We recall that the stochastic process modelled by the state-variable description Σ_S defined by equations (4.2-1) and (4.2-2) involved two noise sources, v_1 and v_2 . These disturbances were combined into one when the Σ_{DS} description was derived after we discovered that, given only observations, we were unable to distinguish between v_1 and v_2 together and a vector random variable from an m -variate normal distribution. Thus, given only the description Σ_{DS} , we are unable to derive a state-variable description that includes both v_1 and v_2 .

Clearly, there must be a new description, Σ_{S1} , that is equivalent to the Σ_S description of (4.2-1) and (4.2-2) but in which the two noise variables are replaced by $e(t)$. Such a description appears below.

$$x(t+1) = Fx(t) + Gu(t) + \Gamma_1 e(t) \quad (4.4-1)$$

$$y(t) = Hx(t) + Du(t) + \sqrt{1/2} e(t) \quad (4.4-2)$$

where $e(t)$ is from an m -variate normal distribution $n(0, I_m)$. In this description, the system and control matrices F , G and D are identically the same as in the equations of the Σ_S description (4.2-1) and (4.2-2).

We are unable to derive the Σ_{S1} description of (4.4-1) and (4.4-2) directly. However, by applying the transformation procedures of Section 3.6, we can find a state-variable description of Σ_{SA} which is also equivalent to Σ_S . The new description Σ_{SA} is

$$x(t+1) = F_A x(t) + G_A u(t) + \Gamma_A e(t) \quad (4.4-3)$$

$$y(t) = H_A x(t) + D u(t) + \Lambda^{1/2} e(t) \quad (4.4-4)$$

By noting the correspondence between B_0 and $\Lambda^{1/2}$, B_1 and $C_1 \Lambda^{1/2}$, ... B_p and $C_p \Lambda^{1/2}$ we can substitute directly into (3.6-2) and obtain

$$\begin{bmatrix} \Lambda^{1/2} \\ H \Gamma_1 \\ H F \Gamma_1 \\ \cdot \\ H F^{p-1} \Gamma_1 \end{bmatrix} = T_p^{*-1} \begin{bmatrix} I_m \\ C_1 \\ \cdot \\ C_p \end{bmatrix} \Lambda^{1/2} \quad (4.4-5)$$

We then find Γ_A in the same way that G_A is found by selecting the appropriate n rows from (4.4-5). The procedure to find F_A , G_A and H_A from $A(z^{-1})$ and $B(z^{-1})$ is unaltered by the presence of the $C(z^{-1})$ terms.

Finally, consider the two special cases discussed in Subsection 4.3.4. In the first case, we found that for the noise process in the form

$$v(t) = (I_m + z^{-1}C_1 + \dots + z^{-p}C_p) \Lambda^{1/2} e(t) \quad (4.4-6)$$

$$C_1 = A_1$$

$$C_2 = A_2$$

$$\cdot \quad \cdot$$

$$C_p = A_p$$

$$(4.4-7)$$

This is the result when $\Gamma = 0$. Hence, $\Gamma_A = 0$.

In the second case, $R_v(p) = 0$ so that

$$\begin{bmatrix} H\Gamma_1 \\ HF\Gamma_1 \\ \cdot \\ \cdot \\ HF^{p-1}\Gamma_1 \end{bmatrix} = \Gamma_p^{-1} \begin{bmatrix} I_m \\ C_1 \\ \cdot \\ \cdot \\ C_{p-1} \end{bmatrix} \wedge^{1/2} \quad (4.4-8)$$

and there is no noise term in the observation equation (4.4-2). Again, Γ_A is found in the usual way. The above can be modified for the case in which $p < m$.

CHAPTER 5

STATISTICAL METHODS OF PARAMETER ESTIMATION5.1 Introduction and Statement of the Problem

In this chapter, the following problem is considered. Suppose that we are given the data record of

(1) a sequence of inputs $\{u(t), t = 1, 2, \dots, N\}$

and (2) the corresponding outputs $\{y(t), t = 1, 2, \dots, N\}$

of a dynamical stochastic system Σ_S . We are required to identify the plant in the sense that we fit a suitable plant model to the data for purposes of output prediction and control.

Recall that in Chapter 1, we divided the identification problem into two parts. The first was concerned with finding an algebraic structure between the input and output variables. The second dealt with estimating parameters in the assumed structure.

We shall presume that the first part of the problem is solved to the extent that we are justified in fitting the Σ_{DS} description to the data or, in other words, the mathematical relations between the input variables and the output variables are closely defined by the Σ_{DS} description. This statement implies that:

(1) the system is linear

(2) the system has finite dimensions

(3) the system is constant

(4) the system is stable

- (5) the system states are observable
- (6) every system state is controllable from either the control input or the noise process or both together
- (7) the disturbances are stationary, gaussian random processes with rational spectral density functions [19].

The first five parts above were assumptions leading to the derivation of the Σ_{DS} description. The sixth part relates the disturbance to physical processes with rational spectra that can be conveniently simulated with linear elements.

The second part of the problem, then, is the estimation of the parameters

- (a) n , the order of the system
- (b) p , the degree of the vector difference equation. (In one sense, (a) and (b) belong to the first part of the identification problem).
- (c) the parameters of $A(z^{-1})$, $B(z^{-1})$, $C(z^{-1})$ and $\wedge^{1/2}$.

Some special cases have been treated in the literature, a few of which are discussed below.

(1) $A_1 = A_2 = \dots = A_p = 0$, $B_1 = B_2 = \dots = B_p = 0$, $C_1 = C_2 = \dots = C_p = 0$ is the problem of multivariate regression considered by Anderson [40]. The sequence $\{u(k)\}$ is known.

(2) $B_1 = B_2 = \dots = B_p = 0$, $C_1 = C_2 = \dots = C_p = 0$ is the problem of identifying a multivariate autoregressive series.

The problem for the single-output case has received attention from a number of authors, beginning with Mann and Wald [41]. A particularly clear exposition is provided by Box and Jenkins [33].

(3) $B_0 = B_1 = \dots = B_p = 0$, $A_1 = A_2 = \dots = A_p = 0$ is a multivariate, finite moving average process. Again, [33] is a useful contribution because it discusses both this case and its relation to the autoregressive series in (2) above for the single-output case.

(4) $B_0 = B_1 = \dots = B_p = 0$ only. This is a combination of (2) and (3) and so is termed a mixed autoregressive moving average (ARMA) model in [33].

(5) $C_1 = C_2 = \dots = C_p = 0$ is the problem of a system with observable states (no measurement noise). It has been studied by Kalman [42] and Lee [43] in the single-output case.

To motivate the identification task, the reason for which the system model is required should be supplied. In this thesis, the requisite is to minimize the mean square error in predicting $y(t+1)$, given the sequence of data $\{y(t), u(t), y(t-1), u(t-1) \dots\}$. If we can solve this problem, we also claim that we can provide a control algorithm that will minimize the variance of the regulated output.

We shall consider and compare a number of parameter estimation methods for solving the multivariate identification problem posed here.

These are listed below:

- (1) classical least squares estimator (LSE)
- (2) modified least squares estimator (MLSE), developed to minimize the bias in the classical least squares estimator)
- (3) generalized least squares estimator (GLSE), (also known as the Markov estimator, minimum variance estimator or best linear unbiased estimator)
- (4) maximum likelihood estimator (MLE).

In order to compare the estimators, we list and define the usual statistical properties of estimators [34].

Definitions

- (1) Unbiased estimator: An estimator $\hat{\theta}$ is an unbiased estimator for θ if $E\hat{\theta} = \theta$ for all θ in Ω . That is, the expected value of the estimator $\hat{\theta}$ is equal to the true value θ .
- (2) Consistent estimator: For indefinitely increasing sample sizes N

$$\lim_{N \rightarrow \infty} \text{Prob} (\theta - \epsilon < \hat{\theta} < \theta + \epsilon) = 1$$

for every θ in Ω and every $\epsilon > 0$. (This is a definition of simple consistency [34]).

- (3) Efficient estimator: If $\hat{\theta}$ is an unbiased estimator for θ having finite variance, and no other unbiased estimator has a smaller variance, then $\hat{\theta}$ is an efficient estimator for θ . Because there may exist estimators that are efficient in a limiting sense for very large

sample sizes, we prefer to use the following criterion in place of (3).

(4) Minimum-variance unbiased estimator: \hat{e} , a function of a random sample from a distribution $f(x;e)$ is a minimum-variance unbiased estimator for e if

$$(a) \quad E(\hat{e}) = e$$

(b) $\text{var}(\hat{e})$ is less than the variance of any other unbiased estimator.

5.2 Linear Least Squares Estimator

5.2.1 Linear Regression Model

The stochastic description Σ_{DS} is not a suitable regression model for least squares estimation in the form in which it appears in equation (4.3-10). Before we rearrange (4.3-10) into a satisfactory form, let us briefly describe what we mean by a linear regression model and a regression function.

Suppose that the conditional expectation $E[y;x]$ is the random variable which minimizes some risk function $E[R[y - K(x); x]]$ where x is a vector random variable from an n -variate distribution and $K(x)$ is some function of x . The quantity $E[y;x]$, considered as a function of x , is called the linear regression function of y on x if we can write

$$E[y;x] = \mu + K_0 x \quad (5.2-1)$$

where μ is an n -vector and K_0 is $n \times n$ [39]. μ and K_0 are

called regression coefficients. In place of (5.2-1) we can write

$$[y;x] = \mu + K_0 x + e \quad (5.2-2)$$

This is called a regression model. The residual e is a random variable.

In the general case, we write

$$E[y_k; x_k, x_{k-1}, \dots, x_{k-p}] = \mu + K_0 x_k + K_1 x_{k-1} \dots K_p x_{k-p} \quad (5.2-3)$$

$$\triangleq K(\mu, \underline{x}) \quad (5.2-4)$$

where \underline{x} is a vector of vectors x_k, \dots, x_{k-p} .

The linear least squares regression problem is to determine μ and K_j ($j = 0, 1, \dots, p$), such that the risk function

$$R = E[y - K(\mu, \underline{x})]^T [y - K(\mu, \underline{x})] \quad (5.2-5)$$

is minimized. In classical regression, it is assumed that all the vectors x_j are random variables from an infinitely large population and the residuals e are unobservable, independent random variables from $n(0, I_m)$.

Let us write the stochastic difference equation description Σ_{DS} of equation (4.3-10) as a regression model (5.2-3). We have already indicated a possible structure in (3.5-11) when we discussed forms of the Σ_D description. By rearranging (4.3-10), we obtain the regression model (with $\mu = 0$):

$$\begin{aligned}
y(t) = & -A_1 y(t-1) - A_2 (I_m)_2 (t-2) \dots - A_p (I_m)_p y(t-p) \\
& + B_0 u(t) + B_1 u(t-1) \dots + B_p u(t-p) + v(t)
\end{aligned} \tag{5.2-6}$$

where

$$v(t) = \sqrt{\frac{1}{2}} e(t) + C_1 \sqrt{\frac{1}{2}} e(t-1) \dots + C_p \sqrt{\frac{1}{2}} e(t-p) \tag{5.2-7}$$

Because (5.2-6) is a linear relation in y , u and e , we observe that $y(t)$ is a random variable for all integers t if $u(t)$ and $e(t)$ are random variables. Clearly, we cannot claim that $y(t-1)$, $y(t-2)$, ... $y(t-p)$ are independent, nor that the residuals $v(t)$, $v(t-1)$... are independent.

There is no loss in generality in arranging the control sequence so that $E u(t) = 0$ and in assuming that there is no other constant input to the system. Then the constant term μ that appears in equations (5.2-1) to (5.2-5) is not required in the regression model (5.2-6). It will be convenient to make this assumption throughout the remainder of the thesis.

Lastly, it will be helpful to simplify the notation further by writing in place of (5.2-6), the regression model in general form

$$y(t) = \Theta x(t) + v(t) \tag{5.2-8}$$

where Θ is an $m \times s$ matrix ($s = n + (p+1)r$) of regression coefficients defined by

$$\Theta \stackrel{\Delta}{=} [-A_1 \quad -A_2 \quad \dots \quad -A_p \quad B_0 \quad B_1 \quad \dots \quad B_p] \tag{5.2-9}$$

and

$$x(t) \triangleq \begin{bmatrix} y(t-1) \\ (I_m)_2 y(t-2) \\ \cdot \\ \cdot \\ (I_m)_p y(t-p) \\ u(t) \\ u(t-1) \\ \cdot \\ \cdot \\ u(t-p) \end{bmatrix} \quad (5.2-10)$$

Here $x(t)$ is not a state variable. It should not be confused with the use of $x(t)$ in other places in the text where it is clear that it is the state-variable vector of a set of first order equations.

5.2.2 The LSE Algorithm

Let us rewrite the general regression model (5.2-8) in the form*

$$y(t) = X(t)\underline{e} + v(t)$$

$$\triangleq \begin{bmatrix} x^T(t) & 0 & 0 & \cdot & \cdot & 0 \\ 0 & x^T(t) & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & x^T(t) \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ \cdot \\ e_m \end{bmatrix} + v(t) \quad (5.2-11)$$

* Note: We introduce this form here, in preference to equation (5.2-8) because we shall be concerned with matrix operations on $y(t)$ and $X(t)$ directly in a later portion of this chapter.

where e_j^T is the j^{th} row of Θ . Suppose that we observe $y(t)$ and $x(t)$ for N time intervals, $N \geq s + m$. We can arrange the data in the form

$$\underline{y} = \underline{X}\underline{e} + \underline{v} \quad (5.2-12)$$

where

$$\underline{y} = \begin{bmatrix} y(1) \\ y(2) \\ \cdot \\ \cdot \\ \cdot \\ y(N) \end{bmatrix}_{mN \times 1} \quad \underline{e} = \begin{bmatrix} e_1 \\ e_2 \\ \cdot \\ e_j \\ \cdot \\ e_m \end{bmatrix}_{ms \times 1} \quad \underline{v} = \begin{bmatrix} v(1) \\ v(2) \\ \cdot \\ \cdot \\ \cdot \\ v(N) \end{bmatrix}_{mN \times 1} \quad (5.2-13)$$

and

$$\underline{X} \triangleq \begin{bmatrix} X(1) \\ \cdot \\ X(2) \\ \cdot \\ \cdot \\ X(N) \\ \cdot \\ \cdot \\ 0 \end{bmatrix} = \begin{bmatrix} x^T(1) & 0 & \cdot & 0 \\ 0 & x^T(1) & \cdot & 0 \\ \cdot & \cdot & \cdot & 0 \\ X(2) & 0 & \cdot & x^T(1) \\ x^T(2) & 0 & \cdot & 0 \\ \cdot & 0 & x^T(2) & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ x^T(N) & 0 & \cdot & 0 \\ X(N) & 0 & x^T(N) & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & x^T(N) \end{bmatrix}_{mN \times ms} \quad (5.2-14)$$

From (5.2-5), the linear least squares problem is to minimize

$$\underline{e}^T \underline{v}(t) = E[\underline{y} - \underline{X}\underline{e}]^T [\underline{y} - \underline{X}\underline{e}] \quad (5.2-15)$$

The minimizing LSE solution, on differentiating (5.2-15) with respect to \underline{e} is

$$\hat{\underline{e}} = [\underline{X}^T \underline{X}]^{-1} \underline{X}^T \underline{y} \quad (5.2-16)$$

As this requires the inversion of an $m \times m$ matrix, it is advantageous from a numerical viewpoint to decompose (5.2-16) into m separate estimators \hat{e}_j for each vector \underline{e}_j . We note that

$$\underline{X}^T \underline{X} = \begin{bmatrix} \sum_{t=1}^N X^T(t)X(t) & & & & \\ \sum_{t=1}^N x(t)x^T(t) & 0 & \cdot & & 0 \\ & 0 & \sum_{t=1}^N x(t)x^T(t) & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \sum_{t=1}^N x(t)x^T(t) & \cdot \end{bmatrix} \quad (5.2-17)$$

Similarly

$$\underline{X}^T \underline{y} = \sum_{t=1}^N X^T(t)y(t) \quad (\text{cont.})$$

$$= \begin{bmatrix} N \\ \sum_{t=1} x(t)y_1(t) \\ N \\ \sum_{t=1} x(t)y_2(t) \\ \vdots \\ N \\ \sum_{t=1} x(t)y_m(t) \end{bmatrix} \quad (5.2-18)$$

where $y_j(t)$ is the j^{th} element of $y(t)$.

This gives us the well known result [40]:

$$\hat{\theta}_j = \left[\sum_{t=1}^N x(t)x^T(t) \right]^{-1} \sum_{t=1}^N x(t)y_j(t) \quad (5.2-19)$$

whereby each row of $\hat{\theta}$ is estimated separately and only one $s \times s$ matrix inversion is required. Alternatively, by manipulating equation (5.2-19), we may write the matrix estimate

$$\hat{\theta} = \left[\sum_{t=1}^N y(t)x^T(t) \right] \left[\sum_{t=1}^N x(t)x^T(t) \right]^{-1} \quad (5.2-20)$$

The matrix $\left[\sum_{t=1}^N x(t)x^T(t) \right]$ of equations (5.2-19) and (5.2-20) can be partitioned in a meaningful way when N is sufficiently large.

Let us define the symmetric matrix

$$\mathcal{L}(N) \triangleq \sum_{t=1}^N x(t)x^T(t) \quad (5.2-21)$$

$$= \begin{bmatrix} \sum_{t=1}^N \begin{bmatrix} y(t-1) \\ (I_m)_2 y(t-2) \\ \vdots \\ (I_m)_p y(t-p) \end{bmatrix} \begin{bmatrix} y(t-1) \dots (I_m)_p y(t-p) \end{bmatrix} \\ \sum_{t=1}^N \begin{bmatrix} u(t) \\ u(t-1) \\ \vdots \\ u(t-p) \end{bmatrix} \begin{bmatrix} y(t-1) \dots (I_m)_p y(t-p) \end{bmatrix} \\ \sum_{t=1}^N \begin{bmatrix} u(t) \\ u(t-1) \\ \vdots \\ u(t-p) \end{bmatrix} \begin{bmatrix} u(t) \\ u(t-1) \\ \vdots \\ u(t-p) \end{bmatrix} \end{bmatrix} \begin{bmatrix} [u^T(t) \dots u^T(t-p)] \\ [u^T(t) \dots u^T(t-p)] \\ [u^T(t) \dots u^T(t-p)] \end{bmatrix} \quad (5.2-22)$$

As N approaches infinity,

$$\frac{1}{N} \Theta(N) = \begin{bmatrix} \text{Unnormalized autocorrelations (autocovariance) of system outputs} & \text{Unnormalized cross correlations (cross covariance) of system output to input} \\ \text{Transposed form of unnormalized cross correlations} & \text{Unnormalized auto-correlations (auto-covariance) of system inputs} \end{bmatrix} \quad (5.2-23)$$

The covariance between e_i^T and e_j^T , two rows of Θ , is defined by

$$\text{cov}(\hat{e}_i, \hat{e}_j) \triangleq E[\hat{e}_i - e_i][\hat{e}_j - e_j]^T \quad (5.2-24)$$

First let us examine just the term

$$\begin{aligned}
\hat{e}_j - e_j &= \left[\sum_{t=1}^N x(t)x^T(t) \right]^{-1} \sum_{t=1}^N x(t)y_j(t) - e_j \\
&= \left[\sum_{t=1}^N x(t)x^T(t) \right]^{-1} \sum_{t=1}^N x(t)[x^T(t)e_j + v_j(t)] - e_j \\
&= \mathcal{J}^{-1}(N) \sum_{t=1}^N x(t)v_j(t) \quad (5.2-25)
\end{aligned}$$

where $v_j(t)$ is the j^{th} element of $v(t)$. For convenience, we define

$$S(N) \triangleq \left[\frac{1}{N} \mathcal{J}(N) \right]^{-1} \quad (5.2-26)$$

$$= \left[\frac{1}{N} \sum_{t=1}^N x(t)x^T(t) \right]^{-1} \quad (5.2-27)$$

and note that $S(N)$ approaches a constant, positive definite, symmetric matrix for large N (the inverse of the matrix in equation (5.2-23)).

We also need to define the $m \times m$ symmetric matrix

$$E v(t)v^T(t) \triangleq V \quad (5.2-28)$$

with elements V_{ij} given by

$$E v_i(t)v_j(t) \triangleq V_{ij} \quad (5.2-29)$$

Now with equations (5.2-25) and (5.2-27) we can substitute in equation (5.2-24)

$$\begin{aligned}
\text{cov}(\hat{e}_i, \hat{e}_j) &= E \frac{1}{N} S(N) \sum_{t=1}^N x(t)v_i(t) \sum_{\tau=1}^N v_j(\tau)x^T(\tau) \frac{1}{N} S(N) \\
&\quad (5.2-30)
\end{aligned}$$

$$= \frac{1}{N} S(N) N \sum_{t, \tau=1}^N E v_i(t) v_j^T(\tau) \frac{1}{N} x(t) x^T(\tau) \frac{1}{N} S(N) \quad (5.2-31)$$

If, and only if,

$$E x(t) v_k^T(t) = 0 \quad \text{for } k = 1, 2, \dots, m \quad (5.2-32)$$

can we simplify (5.2-31) and obtain the minimum variance property [40]:

$$\text{cov}(\hat{\theta}_i, \hat{\theta}_j) = V_{ij} \frac{1}{N} S(N) \quad (5.2-33)$$

for all $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, m$.

Under what circumstances is this true for the regression model of equation (5.2-8)? Since the first m elements of $x(t)$ are fully correlated with $v(t-1)$ (see equations (5.2-8) and (5.2-10)), then equation (5.2-32) is satisfied if

$$R_v(\tau) \stackrel{\Delta}{=} E v(t) v^T(t-\tau) \quad (4.3-4)$$

$$= 0 \quad (5.2-34)$$

for all integers $\tau \neq 0$. This is the case when the stochastic system is an autoregressive process only and $C_j = 0$ for $j = 1, 2, \dots, p$. In this circumstance alone, the LSE for θ given by equations (5.2-19) or (5.2-20) is the maximum likelihood estimator for θ [40].

5.2.3 Limitations of the LSE

In the following, we consider only the complete mixed autoregressive moving average model (with control terms) given by equation (5.2-6).

The LSE was given by equations (5.2-19) and (5.2-20).

(1) Bias: The least squares estimator for the parameters of the regression model is biased. This is apparent on substituting equation (5.2-26) into equation (5.2-25) from which we obtain the bias term:

$$\hat{e}_j - e_j = \left[\sum_{t=1}^N x(t)x^T(t) \right]^{-1} \sum_{t=1}^N x(t)v_j(t) \quad (5.2-35)$$

$$= S(N) \frac{1}{N} \sum_{t=1}^N x(t)v_j(t) \quad (5.2-36)$$

$$= S(N) \frac{1}{N} \sum_{t=1}^N v_j(t) \begin{bmatrix} y(t-1) \\ (I_m)_2 y(t-2) \\ \vdots \\ (I_m)_p y(t-p) \\ u(t) \\ \vdots \\ u(t-p) \end{bmatrix} \quad (5.2-37)$$

Since every observation $y(t-k)$ is a linear combination of $e(t-k)$, $e(t-k-1)$, \dots , $e(t-k-p)$ for $k = 1, 2, \dots, p$ and $v_j(t)$ is similarly a linear combination of $e(t)$, $e(t-1)$, \dots , $e(t-p)$, we can see that $v_j(t)$ and $y(t-k)$ are correlated for $k = 1, 2, \dots, p$ ($t-k > 0$).

We also recall that $S(N)$ is a constant matrix for N large. Provided that the control sequence $\{u(t)\}$ is independent of the observations, it follows that:

$$E[\hat{\theta}_j - \theta_j] = \lim_{N \rightarrow \infty} \frac{S(N)}{N} E v_j(t) \begin{bmatrix} y(t-1) \\ (I_m)_2 y(t-2) \\ \vdots \\ (I_m)_p y(t-p) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (5.2-38)$$

$$= \lim_{N \rightarrow \infty} \frac{S(N)}{N} \begin{bmatrix} \text{An } n\text{-vector of nonzero constant elements, the unnormalized correlations of } v_j(t) \text{ with } v_i(t-\tau) \text{ for } \tau = 1, 2, \dots, p. \\ \text{An } (s-n) \text{ vector of zeros} \end{bmatrix} \neq 0 \quad (5.2-39)$$

Hence $\hat{\theta}_j$ is biased and $\hat{\theta}$ is biased.

(2) Consistency: Clearly, the LSE is not consistent for the case discussed because $\hat{\theta}$ converges to a biased value rather than its true value.

(3) Efficiency: Here we shall simply state that there is a more efficient estimator for θ than the LSE for the case discussed. This is well known for the single observation case and follows directly for the multivariate case.

We conclude that in the mixed autoregressive moving average (ARMA) model, the LSE is not an unbiased, minimum variance estimator.

5.3 Sequential Least Squares Estimator

In the previous section, the least squares estimator required the inversion of an $s \times s$ matrix ($s = n + (p+1)r$) with which each row of Θ was estimated in a separate operation. We shall now show that matrix Θ can be updated, row by row, each time that new data is made available. The convergence properties will be discussed briefly.

Recall the regression model, given by equation (5.2-8)

$$y(t) = \Theta x(t) + v(t) \quad (5.3-1)$$

Let us define $\hat{\Theta}(N)$ as the LSE for Θ given the set of data $\{y(1), y(2), \dots, y(N), x(1), x(2), \dots, x(N)\}$ and $\hat{\Theta}_j(N)$ as the LSE for $\Theta_j(N)$ for the set of data $\{y_j(1), y_j(2), \dots, y_j(N), x(1), x(2), \dots, x(N)\}$. The current estimate of the j^{th} row of Θ is:

$$\hat{\Theta}_j(N) = \left[\sum_{t=1}^N x(t)x^T(t) \right]^{-1} \sum_{t=1}^N x(t)y_j(t) \quad (5.3-2)$$

$$= \mathcal{J}^{-1}(N) \sum_{t=1}^N x(t)y_j(t) \quad (5.3-3)$$

by the definition of equation (5.2-21). Clearly

$$\hat{\Theta}_j(N+1) = \mathcal{J}^{-1}(N+1) \left[\sum_{t=1}^N x(t)y_j(t) + x(N+1)y_j(N+1) \right] \quad (5.3-4)$$

By the matrix inversion lemma

$$\mathcal{J}^{-1}(N+1) = \left[\mathcal{J}(N) + x(N+1)x^T(N+1) \right]^{-1} \quad (5.3-5)$$

$$= \mathcal{J}^{-1}(N) - \frac{\mathcal{J}^{-1}(N)x(N+1)x^T(N+1)\mathcal{J}^{-1}(N)}{x^T(N+1)\mathcal{J}^{-1}(N)x(N+1) + 1} \quad (5.3-6)$$

Then

$$\begin{aligned}
 \hat{\theta}_j(N+1) &= \mathcal{J}^{-1}(N) \sum_{t=1}^N x(t)y_j(t) - \frac{\mathcal{J}^{-1}(N)x(N+1)x^T(N+1)\mathcal{J}^{-1}(N) \sum_{t=1}^N x(t)y_j(t)}{x^T(N+1)\mathcal{J}^{-1}(N)x(N+1) + 1} \\
 &+ \mathcal{J}^{-1}(N)x(N+1)y_j(N+1) - \frac{\mathcal{J}^{-1}(N)x(N+1)x^T(N+1)\mathcal{J}^{-1}(N)x(N+1)y_j(N+1)}{x^T(N+1)\mathcal{J}^{-1}(N)x(N+1) + 1} \\
 &= \hat{\theta}_j(N) + \frac{\mathcal{J}^{-1}(N)x(N+1)[y_j(N+1) - x^T(N+1)\hat{\theta}_j(N)]}{x^T(N+1)\mathcal{J}^{-1}(N)x(N+1) + 1} \quad (5.3-7)
 \end{aligned}$$

This is a normalized form of the Kalman sequential estimating equation.

This algorithm is useful for the estimation of θ if immediate estimates are required in an on-line situation. Of course, because it is identical to the estimator given by equations (5.2-19) or (5.2-20), it is also subject to the limitations of the LSE set out in Subsection 5.2.3. The bias arises because of the correlation between $x(N+1)$ and $y_j(N+1)$ in the numerator of equation (5.3-7).

Let us rewrite (5.3-7) in terms of the matrix $S(N)$ which approaches a constant as $N \rightarrow \infty$:

$$\hat{\theta}_j(N+1) \underset{N \rightarrow \infty}{=} \hat{\theta}_j(N) + \frac{S(N)x(N+1)[y_j(N+1) - x^T(N+1)\hat{\theta}_j(N)]}{x^T(N+1)S(N)x(N+1) + N} \quad (5.3-8)$$

If the system is stable and the input is bounded, the inner product $x^T(N+1)S(N)x(N+1)$ has a constant upper bound. Clearly, the denominator term increases with N . Also, the numerator term

$$y_j(N+1) - x^T(N+1)\hat{\theta}_j(N) \quad (5.3-9)$$

$$= x^T(N+1)[\theta_j - \hat{\theta}_j(N)] + v_j(N) \quad (5.3-10)$$

tends to decrease as $\hat{\theta}_j(N)$ approaches constant value for increasing N . Convergence of the algorithm can be established from the equation obtained by combining (5.3-10) with (5.3-8).

$$\begin{aligned} \tilde{\theta}_j(N+1) &= \left[I - \frac{S(N)x(N+1)x^T(N+1)}{x^T(N+1)S(N)x(N+1) + N} \right] \tilde{\theta}_j(N) \\ &+ \frac{S(N)x(N+1)}{x^T(N+1)S(N)x(N+1) + N} v_j \end{aligned} \quad (5.3-11)$$

$$\text{where} \quad \tilde{\theta}_j(N) \triangleq \hat{\theta}_j(N) - \theta_j \quad (5.3-12)$$

The subject of convergence is discussed by Lee [43] and Mayne [44]. Lee shows that, as $N \rightarrow \infty$, the estimate becomes independent of the initial conditions assumed.

We note that if $x(N+1)$ is correlated with $y_j(N+1)$ in equation (5.3-8), the estimate will converge to an incorrect (biased) value for the reasons discussed in Subsection 5.2.1, unless

$$R_v(\tau) = 0 \quad (5.2-13)$$

for all integers $\tau \neq 0$. Again we conclude that the LSE is

- (1) a biased estimator
- (2) not a consistent estimator

for the parameters of the multivariate ARMA regression model of (5.2-6).

Clearly, a more suitable estimator is required.

5.4 Generalized Least Squares Estimator

5.4.1 The Classical Viewpoint

In the last section on linear least squares regression, the rows of Θ were determined so that the risk function

$$R = E[y(T) - \Theta x(T)]^T [y(t) - \Theta x(t)] \quad (5.4-1)$$

$$= E v^T(t) v(t) \quad (5.4-2)$$

was minimized. The difficulties encountered (biased, inconsistent and nonminimum variance estimates for Θ) were attributed to the correlation arising between the variables $x(t)$ of the regression model and $v(t)$ when the stochastic disturbance $v(t)$ exhibits serial correlation.

The method of generalized least squares estimation (GLSE) [46] is to transform the data so that the elements of the transformed noise sequence have zero autocorrelation for $\tau \neq 0$. The solution in the case of scalar ($m = 1$) observations $y(t)$ of a process in the variables $x(t)$ is to find a matrix T so that the variance-covariance matrix of the transformed noise process becomes:

$$E \underline{v v}^T T^T = \lambda^2 I_N \quad (5.4-3)$$

where λ^2 is the variance of the "whitened" noise process. To find this transformation T , it is necessary to have complete knowledge of the $N \times N$ variance-covariance matrix of the noise process, $\underline{v v}^T$. (For the scalar observation case, the variance-covariance matrix $\underline{v v}^T$ is a Toeplitz matrix.) The well known solution in this case is

(compare with equation (5.2-16)):

$$\hat{e} = [\underline{X}^T \underline{T}^T \underline{T} \underline{X}]^{-1} \underline{X}^T \underline{T}^T \underline{T} \underline{y} \quad (5.4-4)$$

in the notation of the last section. With

$$[\underline{T}^T \underline{T}] = \lambda^2 [\underline{E} \underline{v} \underline{v}^T]^{-1} \quad (5.4-5)$$

the GLSE is unbiased, efficient and consistent. However, in order to acquire these advantageous properties, it is obvious from equation (5.4-5) that a complete description of the noise process is required in order to compute the matrix $\underline{T}^T \underline{T}$. A clear exposition on the subject is given by Johnson [47] for scalar observations.

5.4.2 The Filtering Viewpoint ($m = 1$)

In this subsection, we review the method of generalized least squares estimation for the following conditions:

- (a) The system model in the form of the Σ_{DS} description is:

$$A(z^{-1})y(t) = B(z^{-1})u(t) + \lambda C(z^{-1})e(t) \quad (5.4-6)$$

- (b) The observations are scalar ($m = 1$).
- (c) The parameters of $C(z^{-1})$ are known.
- (d) Updated estimates of the parameters of $A(z^{-1})$ and $B(z^{-1})$ are required at every time interval.

Mayne [44] has suggested filtering the input and observed output sequences as in Figure 5.1 to obtain:

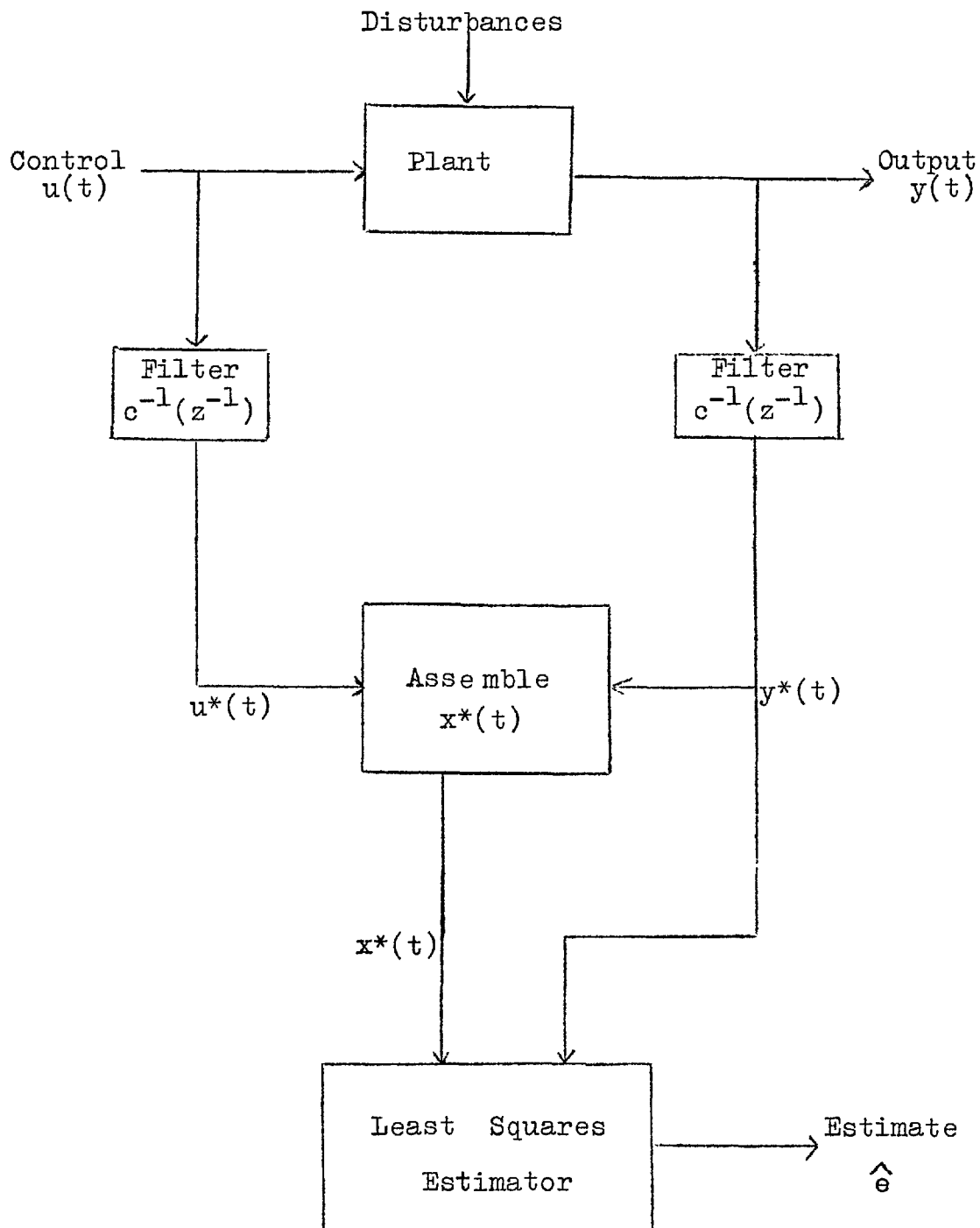


Figure 5.1 Generalized Least Squares Estimator

(m = 1)

$$y^*(t) = \frac{1}{C(z^{-1})} y(t) ; \quad u^*(t) = \frac{1}{C(z^{-1})} u(t) \quad (5.4-7)$$

Substituting for $y(t)$ and $u(t)$ in (5.4-6):

$$A(z^{-1})C(z^{-1})y^*(t) = B(z^{-1})C(z^{-1})u^*(t) + \lambda C(z^{-1})e(t) \quad (5.4-8)$$

Because $C(z^{-1})$ commutes with $A(z^{-1})$ and $B(z^{-1})$ (for $m = 1$ only), the transformed system model obtained by premultiplying (5.4-8) by $C^{-1}(z^{-1})$ is:

$$A(z^{-1})y^*(t) = B(z^{-1})u^*(t) + \lambda e(t) \quad (5.4-9)$$

In the transformed system model of (5.4-9) the noise sequence is independent and uncorrelated with its past. The parameters of $A(z^{-1})$ and $B(z^{-1})$ can then be obtained by applying the conventional least squares estimator given by equation (5.3-7) (for $j = 1$). The regression model is

$$y^*(t) = e^T x^*(t) + \lambda e(t)$$

where e is an s -vector and

$$x^*(t) = \begin{bmatrix} y^*(t-1) \\ y^*(t-2) \\ \vdots \\ y^*(t-n) \\ u^*(t) \\ u^*(t-1) \\ \vdots \\ u^*(t-n) \end{bmatrix} \quad (5.4-10)$$

Estimation by this method has also been discussed by Briggs, Clarke and Hammond [48] for the identification of control systems with disturbances that are serially correlated. Again, the disadvantage of the method is that the parameters of the noise process are required.

5.4.3 GLSE When the Noise Process is Unidentified

A number of authors have sought the advantages of GLSE by estimating the parameters of the noise process separately from the remaining parameters of the system model ($m = 1$).

Durbin [49] appears to have been the first to suggest a two-stage estimation procedure. Tretter and Steiglitz [50] obtained the parameters of a mixed autoregressive moving average stochastic system by least squares estimation of the parameters of $A(z^{-1})$ and direct search for the parameters of $C(z^{-1})$ to minimize the variance of the residuals. Clarke [51], adjusted the parameters of an all-pole filter approximating the function of the ideal filter discussed in the last subsection for "whitening" the residual noise sequence. Norris has investigated special cases in which the latter two methods may fail [52].

5.4.4 GLSE in the Vector Observation Case ($m > 1$)

Filtering the observation and control sequence in the vector observation case ($m > 1$) is not a simple extension of the method for the scalar observation case. The reason is that the algebraic operation between equations (5.4-8) and (5.4-9) (commuting $C(z^{-1})$ with $A(z^{-1})$ and $B(z^{-1})$) is invalid if $A(z^{-1})$, $B(z^{-1})$ and $C(z^{-1})$

are matrix polynomials for which $m > 1$. In the following, we show that the principle of removing serial correlation by filtering can be extended to the multivariate observation case. The design equations for a suitable multivariate filter are developed.

The regression model in general form (see equations (5.2-8) to (5.2-13)) is:

$$y(t) = \Theta x(t) + v(t) \quad (5.4-11)$$

$$= X(t)\underline{e} + v(t) \quad (5.4-12)$$

where \underline{e} is defined by equation (5.2-13) and

$$X(t) \triangleq \begin{bmatrix} x^T(t) & 0 & 0 & \cdot & \cdot & 0 \\ 0 & x^T(t) & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & x^T(t) \end{bmatrix} \quad (5.4-13)$$

is a matrix of m rows and ms columns. Recall

$$v(t) = C(z^{-1})\sqrt{1/2}e(t) \quad (5.4-14)$$

Substituting in (5.4-12):

$$y(t) = X(t)\underline{e} + C(z^{-1})\sqrt{1/2}e(t) \quad (5.4-15)$$

$$\text{Then } C^{-1}(z^{-1})y(t) = C^{-1}(z^{-1})X(t)\underline{e} + \sqrt{1/2}e(t) \quad (5.4-16)$$

It will be convenient to rewrite (5.4-16) as

$$y^*(t) = X^*(t)\underline{e} + \sqrt{1/2}e(t) \quad (5.4-17)$$

which should be compared with the model of equation (5.4-12) in which the stochastic process $v(t)$ is serially correlated. Also, (5.4-17) should be compared with the filtered equation (5.4-9) for the simple case ($m = 1$). Note that $X^*(t)$ is uncorrelated with $e(t)$ in the transformed model of equation (5.4-17). This is a necessary condition for the unbiased, least squares estimation of Θ from the filtered data.

The remaining problem is to design the filters in terms of the parameters of $C(z^{-1})$.

(a) The filter to obtain y^* :

From equations (5.4-16) and (5.4-17)

$$y^*(t) = C^{-1}(z^{-1})y(t) \quad (5.4-18)$$

That is
$$[I + z^{-1}C_1 + z^{-2}C_2 + \dots + z^{-p}C_p]y^*(t) = y(t) \quad (5.4-19)$$

Writing out equation (5.4-19) in full and rearranging, we obtain:

$$y^*(t) = -[C_1y^*(t-1) + C_2y^*(t-2) \dots + C_p y^*(t-p)] + y(t) \quad (5.4-20)$$

which is simply the recursion relation of a multivariate dynamical process of order p . This is the design equation of the multivariate filter to find $y^*(t)$.

(b) The filter to obtain $X^*(t)$:

Following the same derivation as in (a), we obtain:

$$[I + z^{-1}C_1 + z^{-2}C_2 + \dots + z^{-p}C_p]X^*(t) = X(t) \quad (5.4-21)$$

That is:

$$X^*(t) = - [C_1 X^*(t-1) + C_2 X^*(t-2) + \dots + C_p X^*(t-p)] + X(t) \quad (5.4-22)$$

Note that (5.4-22) is the matrix difference equation of a dynamical process. Although $X(t)$ is a matrix comprised mostly of zeros and rows that contain the same information (see equation (5.4-13)), $X^*(t)$ is a full matrix of rows that do not resemble each other. Provided that the sequence of inputs $\{u(t)\}$ is from an r-variate distribution of random variables, we are assured that the elements of $X^*(t)$ are random variables because only linear operations are involved in deriving $X(t)$ and $X^*(t)$ from $\{u(t)\}$ and $\{e(t)\}$.

Equations (5.4-20) and (5.4-22) are the design equations of the digital filters employed to filter $y(t)$ and $X(t)$ as shown in Figure 5.2. Note that the filters are identical. The fact that one filter is excited by a vector process and the other by a matrix process is irrelevant. It is easy to see this after partitioning $X(t)$ and $X^*(t)$ into s column vectors of dimension m . We note that the k^{th} column of $X^*(t)$ is dependent (or excited by) the k^{th} column only of $X(t)$. Hence the same algorithm (or computing subroutine) can be used to update $y^*(t)$ and each column of $X^*(t)$ in turn.

5.4.5 Filter Stability

In designing filters, the question of stability arises. Will the output of the filter remain bounded over any finite interval of time when excited by any arbitrary sequence of bounded inputs? To answer this question, we must evaluate all the natural modes [25]

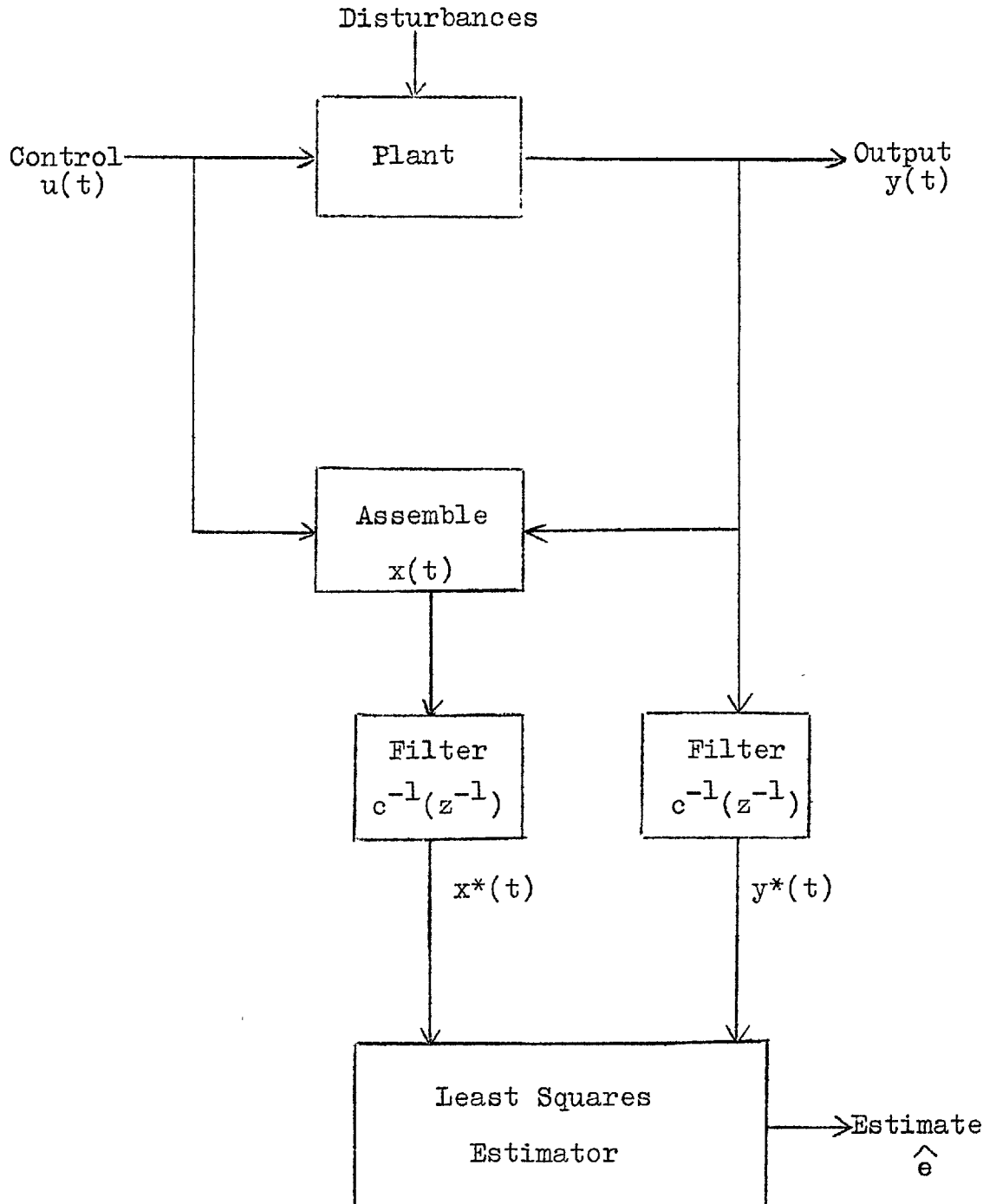


Figure 5.2 Generalized Least Squares Estimator

 $(m > 1)$

of the filter.

The filtering equation may be written as

$$C(z^{-1})y^*(t) = y(t) \quad (5.4-23)$$

This is a vector difference equation of order p in which each of the matrix coefficients C_j ($j = 1, 2, \dots, p$) is $m \times m$. This fact facilitates the reduction of equation (5.4-23) to a set of mp first order difference equations by the methods of Section 3.6. We obtain the discrete-time state transition matrix F_A in canonical form by inspection:

$$F_A = \begin{bmatrix} 0 & I_m & 0 & \cdot & 0 \\ 0 & 0 & I_m & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & I_m \\ -C_p & -C_{p-1} & -C_{p-2} & \cdot & -C_1 \end{bmatrix} \quad (5.4-24)$$

The filter of equation (5.4-23) is stable provided that the eigenvalues of F_A fall within the unit circle in the z plane.

In what circumstances can some or all of the eigenvalues of F_A fall outside the unit circle in the z plane? This is the case when zeros of the stochastic process $C(z)e^{j\omega t}$ are found outside the unit circle, corresponding to zeros of a continuous-time process in the right-half s -plane. Such conditions are encountered in real situations that exhibit nonminimum phase characteristics.

We shall consider two possible solutions to the problem if the above conditions are encountered and illustrate the discussion with examples for the case $m = 1$.

(1) Reduce the order of the filter, so as to exclude compensation for the "troublesome" zeros of the noise process.

Suppose that the noise process is given by

$$v(t) = (1 - 1.5z^{-1})e(t) \quad (5.4-25)$$

which has a zero at $z = 1.5$. The corresponding filter to "whiten" this process has a pole at the same location. Clearly, it is unstable. The solution proposed above is unsatisfactory because the filter of reduced order is zero order and no compensation is offered. Least squares estimation of the parameters of a system driven by this noise would lead to biased estimates.

(2) Find the best stable filter that will "whiten" the noise process. The solution to this problem is a special case of the solution given by Wiener for the optimum realizable linear filter that minimizes the mean-square error between a "desired" signal corrupted by additive white noise and the observed output signal [62]. Because the "desired" signal is zero and the noise source $e(t)$ is white, the problem is greatly simplified.

Consider the following filtering equation in which the filtered output is given by $y(t)$ and the correlated disturbance by $v(t)$.

$$A(z^{-1})y(t) = v(t) \quad (5.4-26)$$

where
$$v(t) = C(z^{-1})e(t) \quad (5.4-27)$$

is a moving average stochastic process of order n . The sequence $\{e(t)\}$ is from $n(0, \lambda)$.

The problem is to find the polynomial coefficients of $A(z^{-1})$ so that:

$$(1) \quad E y(t)y(t-\tau) = 0 \quad \text{for all } \tau \neq 0. \quad (5.4-28)$$

($y(t)$ is then a white noise process).

(2) The roots of $A(z^{-1})$ are within the unit circle in the z plane; that is, the filter is stable. Clearly, $A(z^{-1})$ and $C(z^{-1})$ must have the same number of real zeros and complex zero pairs to satisfy equation (5.4-28).

Let us write $C(z^{-1})$ in factored form as the product of real and complex roots to be annihilated or compensated by real or complex (but stable) roots of $A(z^{-1})$. The roots are enumerated as follows.

There are:

- (a) p_1 real roots within the unit circle; that is $(1 + {}_1a_i z^{-1})$ is a factor of $C(z^{-1})$ to be annihilated by $(1 + {}_1\alpha_i z^{-1})$, a factor of $A(z^{-1})$ for $i = 1, 2, \dots p_1$.
- (b) p_2 real roots outside the unit circle; $(1 + {}_2a_i z^{-1})$ is a factor of $C(z^{-1})$ to be compensated by $(1 + {}_2\alpha_i z^{-1})$, a factor of $A(z^{-1})$ for $i = 1, 2, \dots p_2$.
- (c) q_1 complex root pairs inside the unit circle; $(1 + {}_1b_i z^{-1} + {}_1c_i z^{-2})$ is a factor of $C(z^{-1})$ to be annihilated by $(1 + {}_1\beta_i z^{-1} + {}_1\gamma_i z^{-2})$,

a factor of $A(z^{-1})$ for $i = 1, 2, \dots, q_1$.

- (d) q_2 complex root pairs outside the unit circle; $(1 + {}_2b_i z^{-1} + {}_2c_i z^{-2})$ is a factor of $C(z^{-1})$ to be compensated by $(1 + {}_2\beta_i z^{-1} + {}_2\gamma_i z^{-2})$, a factor of $A(z^{-1})$ for $i = 1, 2, \dots, q_2$.

Then

$$C(z^{-1}) = \prod_{i=1}^{p_1} (1 + {}_1a_i z^{-1}) \prod_{i=1}^{p_2} (1 + {}_2a_i z^{-1}) \\ \cdot \prod_{i=1}^{q_1} (1 + {}_1b_i z^{-1} + {}_1c_i z^{-2}) \prod_{i=1}^{q_2} (1 + {}_2b_i z^{-1} + {}_2c_i z^{-2}) \quad (5.4-29)$$

An identical equation can be written for $A(z^{-1})$ with the replacement:

$$\alpha \text{ for } a, \beta \text{ for } b \text{ and } \gamma \text{ for } c.$$

The problem can now be restated:

Find α_i , β_i and γ_i so that $Ey(t)y(t-\tau) = 0$ for $\tau \neq 0$. The problem is solved if the (discrete) spectrum of the output $y(t)$ can be demonstrated to be representative of a white noise sequence.

We assume that the original disturbances of the system have rational spectra so that the white noise source $e(t)$ also has a rational (discrete) spectra given by $I_{ee}(z)$. The filter is linear and α_i , β_i and γ_i are chosen so that the filter is stable. The discrete-process spectral density of $y(t)$ is given by

$$\Phi_{yy}(z) = \frac{C(z)C(z^{-1})}{A(z)A(z^{-1})} I_{ee}(z) \quad (5.4-30)$$

We require:

$$\Phi_{yy}(z) = (\text{constant}) \cdot I_{ee}(z) \quad (5.4-31)$$

We choose: ${}_1\alpha_i = {}_1a_i$ for $i = 1, 2, \dots, p_1$

${}_1\beta_i = {}_1b_i$ for $i = 1, 2, \dots, q_1$

${}_1\gamma_i = {}_1c_i$ for $i = 1, 2, \dots, q_1$

With this substitution, the poles of $A(z^{-1})$ within the unit circle cancel the zeros of $C(z^{-1})$ within the unit circle, leaving:

$$\begin{aligned} \Phi_{yy}(z) &= \frac{\prod_{i=1}^{p_2} (1 + {}_2a_i z)(1 + {}_2a_i z^{-1})}{\prod_{i=1}^{p_2} (1 + {}_2\alpha_i z)(1 + {}_2\alpha_i z^{-1})} \\ &\cdot \frac{\prod_{i=1}^{q_2} (1 + {}_2b_i z + {}_2c_i z^2)(1 + {}_2b_i z^{-1} + {}_2c_i z^{-2})}{\prod_{i=1}^{q_2} (1 + {}_2\beta_i z + {}_2\gamma_i z^2)(1 + {}_2\beta_i z^{-1} + {}_2\gamma_i z^{-2})} \Phi_{ee} \end{aligned} \quad (5.4-32)$$

The solution to satisfy (5.4-31) is:

$${}_2\alpha_i = 1/{}_2a_i \quad (5.4-33)$$

for $i = 1, 2, \dots, p_2$, and

$${}_2\beta_i = {}_2b_i/{}_2c_i \quad (5.4-34)$$

$${}_2\gamma_i = 1/{}_2c_i \quad (5.4-35)$$

for $i = 1, 2, \dots, q_2$. Substituting the above values in equation (5.4-32) we obtain:

$$\Phi_{yy}(z) = \prod_{i=1}^{p_2} a_i \prod_{i=1}^{q_2} c_i \Phi_{ee} \quad (5.4-36)$$

That is, if the variance of the white noise source $e(t)$ is λ^2 , the variance of the white noise of the output $y(t)$ is given by

$$\prod_{i=1}^{p_2} a_i \prod_{i=1}^{q_2} c_i \lambda^2. \text{ It can be shown that the filter defined by } A(z^{-1})$$

is the optimum (minimum mean-squared error) stable filter given by the Wiener theory.

Example (1):

$$v(t) = (1 - 2z^{-1})e(t) \quad (5.4-37)$$

From equation (5.4-33), we choose

$$A(z^{-1}) = (1 - 0.5z^{-1}) \quad (5.4-38)$$

Then

$$\begin{aligned} \Phi_{yy}(z) &= \frac{(1 - 2z^{-1})(1 - 2z)}{(1 - 0.5z^{-1})(1 - 0.5z)} \Phi_{ee}(z) \\ &= \frac{1 - 2z - 2z^{-1} + 4}{1 - 0.5z - 0.5z^{-1} + 0.25} \Phi_{ee}(z) \\ &= 4 \frac{1 - 2z - 2z^{-1} + 4}{4 - 2z - 2z^{-1} + 1} \Phi_{ee}(z) \\ &= 4 \Phi_{ee}(z) \end{aligned}$$

and $Ey^2(t) = 4\lambda^2 \quad (5.4-39)$

Example (2):

$$v(t) = (1 + 2z^{-1} + 2z^{-2})e(t)$$

From equations (5.4-34) and (5.4-35)

$$A(z^{-1}) = 1 + z^{-1} + 0.5z^{-2} \quad (5.4-40)$$

Then

$$\Phi_{yy}(z) = \frac{(1 + 2z + 2z^2)(1 + 2z^{-1} + 2z^{-2})}{(1 + z + 0.5z^2)(1 + z^{-1} + 0.5z^{-2})} \Phi_{ee}(z) \quad (5.4-41)$$

$$= \frac{9 + 6z + 2z^2 + 6z^{-1} + 2z^{-2}}{2.25 + 1.5z + 0.5z^2 + 1.5z^{-1} + 0.5z^{-2}} \Phi_{ee}(z) \quad (5.4-42)$$

$$= 4 \Phi_{ee}(z)$$

$$\text{and} \quad E y^2(t) = 4 \lambda^2 \quad (5.4-43)$$

We conclude that, given a description of a stochastic process with a finite number of nonzero serial autocorrelation terms (a finite moving average process), we can always find a stable filter to effectively remove the correlation. Although reported here in detail only for the case in which $m = 1$, the writer has found that the principle can be extended to the multivariate case for $m > 1$. To establish the principle (but not the details of finding the parameters of the filter), we prove the following:

Given the finite moving average stochastic process of order p

$$v(t) = C(z^{-1})e^*(t) \quad (5.4-44)$$

such that

$$(1) \quad Ee^*(t) = 0 \quad (5.4-45)$$

$$\begin{aligned} Ee^*(t)e^{*\top}(t-\tau) &= \begin{cases} \Delta & \text{for } \tau = 0 \\ 0 & \text{for } \tau \neq 0 \end{cases} \end{aligned} \quad (5.4-46)$$

$$\begin{aligned} (2) \quad \mathcal{R}(z) &\stackrel{\Delta}{=} C(z) \Delta C^\top(z^{-1}) \quad (5.4-47) \\ &\stackrel{\Delta}{=} R_V(0) + zR_V(1) + z^{-1}R_V(-1) \\ &\quad \dots + z^p R_V(p) + z^{-p} R_V(-p) \\ &= \mathcal{R}^\top(z^{-1}) \end{aligned} \quad (5.4-48)$$

$$(3) \quad R_V(0) \succ 0 \quad (5.4-49)$$

Then we can find a stable filtering equation

$$y(t) = A^{-1}(z^{-1})v(t) \quad (5.4-50)$$

such that the roots of the characteristic equation for $A^{-1}(z^{-1})$ are within the unit circle in the z plane and

$$Ey(t)y^\top(t-\tau) = 0 \quad \text{for } \tau \neq 0 \quad (5.4-51)$$

The proof is based on being able to factorize $\mathcal{R}(z)$ by performing the discrete-time procedure analogous to the continuous-time method for spectral factorization given by Anderson [61] and referred to in Subsection 4.3-3. The factorization of $\mathcal{R}(z)$ requires assumptions (2) and (3) above. Now

$$\Phi_{yy}(z) = A^{-1}(z)\Phi_{vv}(z)[A^{-1}(z^{-1})]^T \quad (5.4-52)$$

where

$$\begin{aligned} \Phi_{vv}(z) &= C(z)\Phi_{e^*e^*}(z)C^T(z^{-1}) \\ &= C(z)\Delta C^T(z^{-1}) \end{aligned} \quad (5.4-53)$$

$$\Delta \triangleq R(z) \quad (5.4-54)$$

By assumptions (2) and (3) above, we can find a matrix polynomial $A(z^{-1})$ such that

$$A(z)A^T(z^{-1}) = R(z) \quad (5.4-55)$$

Then substituting equations (5.4-54) and (5.4-55) into (5.4-52), we find

$$\Phi_{yy}(z) = I_m \quad (5.4-56)$$

Performing the inverse transform, then

$$\begin{aligned} E y(t)y^T(t-\tau) &= I_m \quad \text{for } \tau = 0 \\ &= 0 \quad \text{for } \tau \neq 0 \end{aligned} \quad (5.4-57)$$

as required.

We conclude that a generalized least squares estimate can be found for the parameters of the system model (equation (5.4-6) without imposing restrictions on the zeros of $C(z^{-1})$. However, in such circumstances, the estimates obtained are not minimum variance. This point is of little practical interest because the minimum variance estimator requires an unstable filter.

5.4.6 The Estimating Equation

Finally, we consider the least squares estimating equation for Θ , given filtered data $y^*(t)$ and $X^*(t)$. Substituting directly into equation (5.2-16) for $\hat{\underline{\theta}}$, we obtain

$$\hat{\underline{\theta}} = [\underline{X}^{*\top} \underline{X}^*]^{-1} \underline{X}^{*\top} y^* \quad (5.4-58)$$

from which the rows of $\hat{\Theta}$ can be worked out from equation (5.2-13). Alternatively, $\hat{\Theta}$ may be updated with each new set of observations $y^*(N)$ and $X^*(N)$ by updating the complete vector $\hat{\Theta}^{(N-1)}$, m times, taking each transformed observation $y_j^*(N)$ in turn for $j = 1, 2, \dots, m$.

The variance-covariance of $\hat{\underline{\theta}}$ is given by:

$$E[\hat{\underline{\theta}} - \underline{\theta}][\hat{\underline{\theta}} - \underline{\theta}]^{\top} = [\underline{X}^{*\top} \underline{X}^*]^{-1} \underline{X}^{*\top} \underline{\Lambda} \underline{X}^* [\underline{X}^{*\top} \underline{X}^*] \quad (5.4-59)$$

It is quite likely that (5.4-59) can be simplified.

5.4.7 The Estimator for $\underline{\Lambda}$

To estimate $\underline{\Lambda}$, we observe from equation (5.4-17) that

$$y^*(t) = X^*(t) \underline{\theta} + \underline{\Lambda}^{1/2} e(t) \quad (5.4-60)$$

where $\underline{\Lambda}^{1/2}$ is assumed symmetric. If we proceed directly, using (5.4-60), we obtain

$$\begin{aligned} E \underline{\Lambda} &= E[y^*(t) - X^*(t) \underline{\theta}][y^*(t) - X^*(t) \underline{\theta}]^{\top} \\ &= E[y^*(t) y^{*\top}(t)] - E[y^*(t) \underline{\theta}^{\top} X^{*\top}(t) + X^*(t) \underline{\theta} y^{*\top}(t)] \\ &\quad + E[X^*(t) \underline{\theta} \underline{\theta}^{\top} X^{*\top}(t)] \end{aligned} \quad (5.4-61)$$

This equation, although correct, is not acceptable for computing purposes because it requires complete storage of both \underline{X}^* and \underline{y}^* . Hence, a method that minimizes storage is desired. This is the feature of the following solution to the problem of estimating $\hat{\underline{\Lambda}}$.

Equation (5.4-60) can be rewritten

$$\underline{y}^*(t) = \begin{bmatrix} \underline{e}^T & 0 & \cdot & 0 \\ 0 & \underline{e}^T & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \underline{e}^T \end{bmatrix} \begin{bmatrix} x_1^*(t) \\ x_2^*(t) \\ \cdot \\ x_m^*(t) \end{bmatrix} + \underline{\Lambda}^{1/2} \underline{e}(t) \quad (5.4-62)$$

where $x_j^*(t)$ is the j^{th} row of $\underline{X}^*(t)$. Replacing equation (5.4-62) by the following new notation:

$$\underline{y}^*(t) = [\underline{e}^T] \underline{x}^*(t) + \underline{\Lambda}^{1/2} \underline{e}(t) \quad (5.4-63)$$

we rederive the expression for

$$\begin{aligned} E \underline{\Lambda} = E[\underline{y}^*(t) \underline{y}^{*T}(t)] - E[\underline{y}^*(t) \underline{x}^{*T}(t) [\underline{e}^T]^T - [\underline{e}^T] \underline{x}^*(t) \underline{y}^{*T}(t)] \\ + E[[\underline{e}^T] \underline{x}^*(t) \underline{x}^{*T}(t) [\underline{e}^T]^T] \end{aligned} \quad (5.4-64)$$

Now we substitute for $\underline{y}^*(t)$ in this last equation, using the fact that $E \underline{e}(t) \underline{x}^{*T}(t) = 0$. Thus

$$\begin{aligned} E \underline{\Lambda} = E[\underline{y}^*(t) \underline{y}^{*T}(t)] - E[[\underline{e}^T] \underline{x}^*(t) \underline{x}^{*T}(t) [\underline{e}^T]^T] \\ - [\underline{e}^T] \underline{x}^*(t) \underline{x}^{*T}(t) [\underline{e}^T]^T + E[[\underline{e}^T] \underline{x}^*(t) \underline{x}^{*T}(t) [\underline{e}^T]^T] \end{aligned} \quad (5.4-65)$$

$$= E[\underline{y}^*(t) \underline{y}^{*T}(t)] - [\underline{e}^T] \underline{x}^*(t) \underline{x}^{*T}(t) [\underline{e}^T]^T \quad (5.4-66)$$

The feature of equation (5.4-66) is that it is sufficient to store the following two arrays:

$$S_1 = \sum_{t=1}^N y^*(t)y^{*\top}(t) \quad (5.4-67)$$

which is $m \times m$ and

$$S_2 = \sum_{t=1}^N \underline{x}^*(t)\underline{x}^{*\top}(t) \quad (5.4-68)$$

which is $m_s \times m_s$, but being symmetric, can be stored in compact form.

The unbiased [40] estimator for $\hat{\theta}$ is then given by

$$\hat{\theta} = \frac{1}{N-m} [S_1 - [e^\top] S_2 [e^\top]^\top] \quad (5.4-69)$$

5.5 Modified Least Squares Estimators

5.5.1 Introduction

Again, we recall that the difficulties encountered in obtaining an unbiased, minimum variance estimator were attributed to the correlation arising between $x(t)$ and $v(t)$ of the regression model given by equation (5.2-8) and restated here:

$$y(t) = \theta x(t) + v(t) \quad (5.5-1)$$

Suppose that, in place of postmultiplying both sides of (5.5-1) by $x(t)$ and summing as below for the first part of finding the LSE for θ ,

$$\sum_{t=1}^N y(t)x^\top(t) = \theta \sum_{t=1}^N x(t)x^\top(t) + \sum_{t=1}^N v(t)x^\top(t) \quad (5.5-2)$$

we postmultiply by the vector $\hat{x}(t)$, an estimate of $x(t)$ subject to the condition that

$$E v(t) \hat{x}(t) = 0 \quad (5.5-3)$$

This is the underlying principle of the method proposed by Rucker [53] and generalized by Levadi [54, 55] for the continuous time estimation of parameters in systems in which $m = 1$.

The remainder of this section is devoted to ways of determining $\hat{x}(t)$ from current running data so that unbiased estimates for θ can be updated on-line for real-time identification, prediction and control purposes.

5.5.2 The Rucker-Levadi Method

The main principle of the method proposed by Rucker [53] and improved by Levadi [54, 55] can be stated in our notation for discrete-time problems as follows. (Our notation extends the method to include cases for which $m > 1$). The (noise-free) system model with output $\bar{y}(t)$ is given by (in our notation)

$$A(z^{-1})\bar{y}(t) = B(z^{-1})u(t) \quad (5.5-4)$$

The measurement noise, $\tilde{y}(t)$, is additive at the output, that is:

$$y(t) = \bar{y}(t) + \tilde{y}(t) \quad (5.5-5)$$

Then
$$A(z^{-1})y(t) = B(z^{-1})u(t) + A(z^{-1})\tilde{y}(t) \quad (5.5-6)$$

The regression model and the system model (before the measurement noise is added) are respectively given by

$$y(t) = \Theta x(t) + v(t) \quad (5.5-7)$$

$$\bar{y}(t) = \Theta \bar{x}(t) \quad (5.5-8)$$

where

$$x(t) = \begin{bmatrix} y(t-1) \\ y(t-2) \\ \cdot \\ y(t-p) \\ u(t) \\ \cdot \\ \cdot \\ u(t-p) \end{bmatrix} \quad \bar{x}(t) = \begin{bmatrix} \bar{y}(t-1) \\ \bar{y}(t-2) \\ \cdot \\ \bar{y}(t-p) \\ u(t) \\ \cdot \\ \cdot \\ u(t-p) \end{bmatrix} \quad (5.5-9)$$

The estimating method is iterative. At each new scalar observation, the following steps are taken.

- (1) A system model output $\hat{y}(t)$ is synthesized by the model

$$\hat{y}(t) = \hat{\Theta}(t-1; y(t-1), x(t-1), \hat{x}(t-1))\hat{x}(t) \quad (5.5-10)$$

- where:
- (a) $\hat{y}(t)$ is an estimate of $\bar{y}(t)$;
 - (b) $\hat{x}(t)$ is an estimate of $\bar{x}(t)$, updated by the inclusion of $\hat{y}(t-1)$ (from the last iteration) and $u(t)$;
 - (c) $\hat{\Theta}(t-1; y(t-1), x(t-1), \hat{x}(t-1))$ is the most recent estimate for Θ , conditional on the observation $y(t-1)$, $x(t-1)$ and the estimate $\hat{x}(t-1)$.

(2) The estimate $\hat{\theta}(t; y(t), x(t), \hat{x}(t))$ is obtained by a sequential LSE algorithm. See reference [54] for the details of the sequential estimator used by Levadi for the continuous-time updating of $\hat{\theta}$ for the case in which θ is a vector of three elements.

Rucker used $\hat{x}(t)$ everywhere in the estimating equation. Levadi [54] shows why this method can fail. Because $\hat{x}(t)$ is correlated with $y(t)$ (hence not correlated with $v(t)$), it is claimed that the estimator is unbiased and consistent. It is not a minimum variance estimator.

Clearly, the method is not generally applicable to the systems considered in this thesis because the method identifies only that part of the system persistently excited by $\{u(t)\}$. However, the principle has merit which can be exploited further.

5.5.3 Mayne's Method

Mayne has published an independent method of obtaining $\hat{x}(t)$ by a two-stage estimating scheme so that equation (5.5-3) is satisfied [20, 56]. We shall review his method for the case $m = 1$ and $\{u(t)\}$ assumed independent.

Suppose that the serial autocorrelations $R_v(\tau)$ defined by equation (4.3-41) are zero for $\tau > n$. Mayne regresses $y(t)$ on to

$$\mu(t) \triangleq \begin{bmatrix} y(t-n-1) \\ y(t-n-2) \\ \cdot \\ \cdot \\ y(t-2n) \\ u(t) \\ u(t-1) \\ \cdot \\ \cdot \\ u(t-2n) \end{bmatrix} \quad (5.5-11)$$

The corresponding regression model is:

$$y(t) = \beta^T \mu(t) + v(t) \quad (5.5-12)$$

where $y(t)$ is a linear combination of $v(t)$, $v(t-n-1)$, $v(t-n-2)$... $v(t-2n)$ and β is a vector of regression coefficients. Clearly, $\mu(t)$ is not correlated with $v(t)$. Although $\hat{\beta}$ is biased when obtained by a least squares estimator, the sequence of predictions

$$\begin{aligned} \hat{y}(t-1) &= \hat{\beta}^T \mu(t-1) \\ \hat{y}(t-2) &= \hat{\beta}^T \mu(t-2) \\ &\cdot \\ &\cdot \\ \hat{y}(t-n) &= \hat{\beta}^T \mu(t-n) \end{aligned} \quad (5.5-13)$$

is not. Hence

$$\hat{x}(t) \triangleq \begin{bmatrix} \hat{y}(t-1) \\ \hat{y}(t-2) \\ \cdot \\ \cdot \\ \hat{y}(t-n) \\ u(t) \\ \cdot \\ \cdot \\ u(t-n) \end{bmatrix} \quad (5.5-14)$$

is uncorrelated with $v(t)$.

In a real-time estimation situation, $\hat{\beta}$ and $\hat{\theta}$ are updated together. The model in β is used to provide up-to-date running estimates of $x(t)$ employed in a sequential LSE algorithm.

The estimate for θ obtained by this method is unbiased and consistent although not minimum variance if $\hat{\beta}$ is obtained by a previous experiment on different data. If $\hat{\beta}$ is obtained concurrently with $\hat{\theta}$, the method is only asymptotically unbiased and consistent.

The method has been extended to systems of two output variables with success [20, 21, 56].

5.5.4 Bootstrap Least Squares Estimator

The writer has developed a sequential estimator which (in hindsight) combines some of the features of the methods of both Levadi and Mayne. As previously, the purpose is to provide an estimate $\hat{x}(t)$ for $x(t)$ using the current running data.

The method is most easily understood by referring to the flow diagram of Figure 5.3 and the following explanation. Suppose that the serial correlations $R_V(\tau) = 0$ for $\tau > p$.

(1) The n -vector $x(t-p)$ is stored. (See equation (5.2-10) for a definition of $x(t)$ and replace t by $t-p$.)

(2) The most recent estimate,

$$\hat{\theta}(t-1; y(t-1), x(t-1), \hat{x}(t-1)) \quad (5.5-15)$$

is used to estimate $\hat{y}(t-p)$ by the model relation below:

$$\hat{y}(t-p) = \hat{\theta}(t-1; \dots)x(t-p) \quad (5.5-16)$$

(3) The vector $\hat{x}(t-p+1)$ is assembled. It is different from $x(t-p+1)$ only where $\hat{y}(t-p)$ replaces $y(t-p)$.

(4) By the same means, $\{\hat{y}(t-p-1), \dots, \hat{y}(t-1)\}$ is estimated. Thereby $\hat{x}(t)$ is obtained.

(5) A new estimate $\hat{\theta}(t; y(t), x(t), \hat{x}(t))$ is obtained by the following modified sequential least squares estimator (compare with the sequential LSE of equation (5.3-7)):

$$\hat{\theta}_j(t) = \hat{\theta}_j(t-1) + \frac{\hat{\mathcal{J}}^{-1}(t)\hat{x}(t)(y_j(t) - x^T(t)\hat{\theta}_j(t-1))}{x^T(t)\hat{\mathcal{J}}^{-1}(t)\hat{x}(t) + 1}$$

where

$$\hat{\mathcal{J}}(t) \triangleq \sum_{k=1}^t \hat{x}(k)x(k) \quad (5.5-17)$$

so that

$$\hat{\mathcal{J}}^{-1}(t) = \hat{\mathcal{J}}^{-1}(t-1) - \frac{\hat{\mathcal{J}}^{-1}(t-1)\hat{x}(t)x^T(t)\hat{\mathcal{J}}^{-1}(t-1)}{x^T(t)\hat{\mathcal{J}}^{-1}(t)\hat{x}(t) + 1} \quad (5.5-18)$$

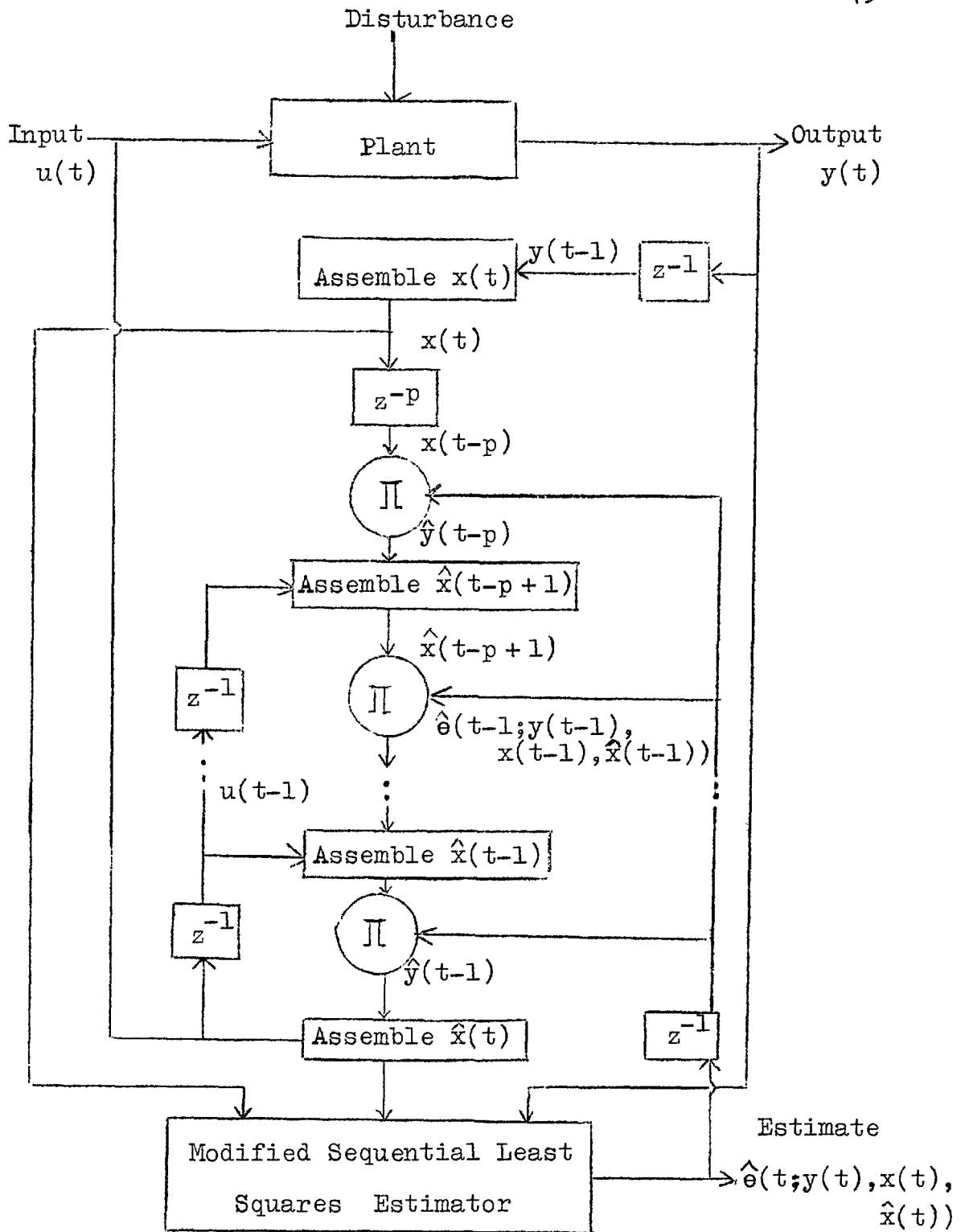


Figure 5.3 Bootstrap Least Squares Estimator

The new estimate is used to predict $\hat{x}(t+1)$ in preparation for the next estimate; hence the name "bootstrap estimator".

(6) To initiate the estimating scheme, the writer has usually obtained a minimal data set estimate after setting $\hat{x}(t) = x(t)$ for the first s intervals. Thereafter, $\hat{x}(t)$ is estimated and incorporated into the algorithm of equation (5.5-18).

We cannot claim that the bootstrap estimator is a minimum variance estimator, but we can attempt to show that it is asymptotically unbiased, hence consistent, by a proof similar to that given by Mayne in [56].

To the seven assumptions concerning the system stated in Section 5.1, we add the following:

- (a) $\{u(t)\}$ is a stationary process, uncorrelated with $\{v(t)\}$.
- (b) The sequential estimating equation given by equation (5.5-17) is stable so that for N large, $\hat{\Theta}$ converges asymptotically to some constant value independent of the initial (or early) estimates.

Hence we claim that:

(1) By including assumption (a) above, $\{x(t)\}$ is a stationary process.

(2) Because the model to find $\hat{y}(\cdot)$ given by equation (5.5-10) involves operations on $x(t-p)$, $\hat{x}(t)$ is a random process. Further, by assumption (b) above, $\hat{x}(t)$ becomes a stationary process as $\hat{\Theta}$ approaches a constant value. Hence $\hat{x}(t)$ tends to a stationary process that is correlated with $x(t)$.

From (2) above, it follows that the matrix $\frac{1}{N} \sum_{t=1}^N \hat{x}(t)x^T(t)$ is a nonsingular matrix (for N suitably large). We shall designate

its inverse by

$$\hat{S}(N) \triangleq \left[\frac{1}{N} \sum_{t=1}^N \hat{\mathbf{x}}(t) \hat{\mathbf{x}}^T(t) \right]^{-1} \quad (5.5-19)$$

In the limit, $\hat{S}(N)$ approaches a constant matrix with inverse given by

$$\hat{S}^{-1}(\infty) = E \hat{\mathbf{x}}(t) \hat{\mathbf{x}}^T(t) \quad (5.5-20)$$

By analogy with equation (5.2-36), the estimation error term is

$$\tilde{\hat{\theta}}_j(N) \triangleq \hat{\theta}_j(N) - \theta_j \quad (5.5-21)$$

$$= \hat{S}(N) \frac{1}{N} \sum_{t=1}^N \hat{\mathbf{x}}(t) v_j(t) \quad (5.5-22)$$

Since $\hat{\mathbf{x}}(t)$ is uncorrelated with $v(t)$ for large N (from assumption (2) and by construction), the asymptotic error term is

$$E \tilde{\hat{\theta}}_j = 0 \quad (5.5-23)$$

As before, $\hat{\theta}$ is made up from the rows $\hat{\theta}_i^T$ for $i = 1, 2, \dots, m$.

By analogy with equations (5.2-30) and (5.2-31), we obtain the covariance given by:

$$\text{cov}(\hat{\theta}_i, \hat{\theta}_j) \triangleq E \frac{1}{N} \hat{S}(N) \sum_{t=1}^N \hat{\mathbf{x}}(t) v_i(t) \sum_{\tau=1}^N v_j(\tau) \hat{\mathbf{x}}^T(\tau) \frac{1}{N} \hat{S}^T(N) \quad (5.5-24)$$

$$= \frac{1}{N} \hat{S}(N) \left[\frac{1}{N} \sum_{t, \tau=1}^N E v_i(t) v_j(\tau) \frac{1}{N} \hat{\mathbf{x}}(t) \hat{\mathbf{x}}^T(\tau) \right] \frac{1}{N} \hat{S}^T(N) \quad (5.5-25)$$

$$= V_{ij} \left[\sum_{t=1}^N \hat{\mathbf{x}}(t) \hat{\mathbf{x}}^T(t) \right]^{-1} \left[\sum_{t=1}^N \hat{\mathbf{x}}(t) \hat{\mathbf{x}}^T(t) \right] \left[\sum_{t=1}^N \mathbf{x}(t) \mathbf{x}^T(t) \right]^{-1} \quad (5.5-26)$$

which is symmetric.

Let us now compare the above covariance with the usual least squares estimator. For convenience, we define:

$$(a) \quad \sum_{t=1}^N x(t)x^T(t) = [x(1) \dots x(N)] \begin{bmatrix} x^T(1) \\ \vdots \\ x^T(N) \end{bmatrix} \\ \triangleq P^T P \quad (5.5-27)$$

$$(b) \quad \sum_{t=1}^N \hat{x}(t)x^T(t) = [\hat{x}(1) \dots \hat{x}(N)] \begin{bmatrix} x^T(1) \\ \vdots \\ x^T(N) \end{bmatrix} \\ \triangleq [(x(1) - \tilde{x}(1)) \dots (x(N) - \tilde{x}(N))] \begin{bmatrix} x^T(1) \\ \vdots \\ x^T(N) \end{bmatrix} \\ \triangleq [P + Q]^T P \quad (5.5-28)$$

$$(c) \quad \sum_{t=1}^N \hat{x}(t)\hat{x}^T(t) = [P + Q]^T [P + Q] \quad (5.5-29)$$

With these substitutions, we replace the modified least-squares covariance of equation (5.5-26) by:

$$\widehat{\text{cov}}(\hat{\theta}_i, \hat{\theta}_j) = v_{ij} [(P+Q)^T P]^{-1} [(P+Q)^T (P+Q)] [P^T (P+Q)]^{-1} \quad (5.5-30)$$

and the usual least squares covariance (for uncorrelated noise) by:

$$\text{cov}(\hat{\theta}_i, \hat{\theta}_j) = V_{ij} [P^T P]^{-1} \quad (5.5-31)$$

Now assume the simplified form:

$$[(P+Q)^T P]^{-1} [(P+Q)^T (P+Q)] [P^T (P+Q)]^{-1} = [P^T P]^{-1} \quad (5.5-32)$$

Premultiplying and postmultiplying by the appropriate matrices, we find that

$$[P+Q]^T [P+Q] = [P+Q]^T P [P^T P]^{-1} P^T [P+Q] \quad (5.5-33)$$

$$= [P+Q]^T [P+Q] \quad (5.5-34)$$

which is true. Hence equation (5.5-32) is true so that for the "bootstrap" estimator, in the limit,

$$\widehat{\text{cov}}(\hat{\theta}_i, \hat{\theta}_j) = \lim_{N \rightarrow \infty} V_{ij} \left[\sum_{t=1}^N x(t) x^T(t) \right]^{-1} \quad (5.5-35)$$

$$= \text{cov}(\hat{\theta}_i, \hat{\theta}_j) \quad (5.5-36)$$

That is, in the limit, the "bootstrap" covariance equation and the covariance given by equation (5.5-31) are equal.

Having estimated $\hat{\theta}$ (recall equation (5.2-8) defining θ),

$$\hat{\theta} = [-\hat{A}_1 \quad -\hat{A}_2 \quad \dots \quad \hat{A}_p \quad \hat{B}_0 \quad \hat{B}_1 \quad \dots \quad \hat{B}_p] \quad (5.5-37)$$

we have the remaining problem of estimating the parameters of $C(z^{-1})$ and $\sqrt{1/2}$. We proceed by obtaining estimates of the serial autocorrelations $\hat{R}_v(0), \hat{R}_v(1), \dots, \hat{R}_v(p)$ by the following relations:

$$R_v(\tau) \triangleq E[v(t)v^T(t-\tau)] \quad (4.3-37)$$

$$= E[y(t) - \theta x(t)][y(t-\tau) - \theta x(t-\tau)]^T \quad (5.5-38)$$

The best unbiased estimate for $R_v(\tau)$ is given by:

$$\hat{R}_v(\tau) = \frac{1}{N-m-\tau} \sum_{t=\tau+1}^N [y(t) - \theta x(t)][y(t-\tau) - \theta x(t-\tau)]^T \quad (5.5-39)$$

$$= \frac{1}{N-m-\tau} \sum_{t=\tau+1}^N [y(t)y^T(t-\tau) - \theta x(t)y^T(t-\tau) - y(t)x^T(t-\tau)\theta^T + \theta x(t)x^T(t-\tau)\theta^T] \quad (5.5-40)$$

By storing the arrays

$$S_1(\tau) = \sum_{t=\tau+1}^N y(t)y^T(t-\tau)$$

$$S_2(\tau) = \sum_{t=\tau+1}^N y(t)x^T(t-\tau)$$

$$S_3(\tau) = \sum_{t=\tau+1}^N x(t)y^T(t-\tau)$$

$$S_4(\tau) = \sum_{t=\tau+1}^N x(t)x^T(t-\tau) \quad (5.5-41)$$

we obtain an estimate for $R_v(\tau)$ employing our (nonminimum variance) estimate for θ . Then

$$\hat{R}_v(\tau) = \frac{1}{N-m-\tau} [s_1(\tau) - s_2(\tau)\hat{\theta}^T - \hat{\theta}s_3(\tau) + \hat{\theta}s_4(\tau)\hat{\theta}^T] \quad (5.5-42)$$

The procedure of fitting moments for finding $\hat{C}_1, \hat{C}_2, \dots, \hat{C}_p$ and $\hat{\Lambda}^{1/2}$ is the same as that set out in Subsection 4.3.3.

This concludes the description of the method of "bootstrap" least squares estimation. Numerical examples are given in the next chapter. We have observed that the "bootstrap" estimator:

- (1) is asymptotically unbiased
- (2) is an improvement of Levadi's method because it makes use of available data. The reason is that the prediction $\hat{x}(t)$, conditional on $x(t-p)$, is superior to the prediction $\hat{x}(t)$ based on past predictions.
- (3) is easier to implement in on-line situations than Mayne's method because it is unnecessary to estimate the parameters of a second regression model.
- (4) The simplicity of the method is achieved at the expense of increased variance in the estimates.

5.6 Maximum Likelihood Estimator

5.6.1 Introduction

Maximum likelihood estimation, by making the most optimal use of available data, is said to provide the "best" estimating method when judged by the properties listed at the beginning of this chapter. Cramer states: "From a theoretical point of view, the most important general method of estimation so far known is the method of maximum likelihood".

In this chapter, we shall extend the elegant method of Astrom et al [18, 19] for the MLE of parameters of single output systems to the general multivariate case ($m > 1$). Astrom anticipated that possibly this might be accomplished when a suitable canonical form was available. The vector stochastic difference equation model (from equation (4.3-10))

$$A(z^{-1})y(t) = B(z^{-1})u(t) + C(z^{-1})\sqrt{1/2}e(t) \quad (5.6-1)$$

fulfills this requirement precisely.

The contribution here is in proposing and incorporating a suitable multivariate model into the MLE method and solving the new problems that arise in doing this.

5.6.2 Formulation of the MLE problem

The formulation of the multivariate MLE problem is similar to that given by Astrom and Bohlin for the scalar observation case [18]. We define the one-step-ahead prediction error $e^*(t)$ by the equation obtained from (5.6-1):

$$e^*(t) = C(z^{-1})[A(z^{-1})y(t) - B(z^{-1})u(t)] \quad (5.6-2)$$

where $y(t)$ and $u(t)$ are observed values from the stochastic sequences $\{y(t)\}$ and $\{u(t)\}$. The numbers $e^*(t)$ are from an independent, normal, m -variate distribution $n(0, \Lambda)$. The negative logarithm of the probability density function (likelihood function) of the sequence $\{e^*(t)\}$ is [40]:

$$-\ln L(\Theta, \Lambda) = \frac{mN}{2} \ln(2\pi) + \frac{N}{2} \ln |\Lambda| + \frac{1}{2} \sum_{t=1}^N e^{*T}(t) \Lambda^{-1} e^*(t) \quad (5.6-3)$$

$$\triangleq L^*(\Theta, \Lambda) \quad (5.6-4)$$

where Θ stands for the parameters in $A(z^{-1})$, $B(z^{-1})$ and $C(z^{-1})$. The likelihood function L is dependent upon Θ and Λ taken as variables. Since $\ln L$ is a monotonically increasing function of L , its maximum is at the same point in the space of Θ and Λ as the maximum of L . In the problem formulated here, we seek to minimize $L^*(\Theta, \Lambda)$ by choosing an optimal Θ^0 and Λ^0 . That is, the maximum likelihood estimates of Θ and Λ are those parameters Θ^0 and Λ^0 such that $L^*(\Theta^0, \Lambda^0)$ is a minimum.

In the scalar case ($m = 1$ and e is a vector), the negative logarithm of the likelihood function is simply

$$L^*(\Theta, \lambda^2) = \frac{N}{2} \ln(2\pi) + N \ln \lambda + \frac{1}{2\lambda^2} \sum_{t=1}^N e^*(t)^2 \quad (5.6-5)$$

for the scalar model:

$$A(z^{-1})y(t) = B(z^{-1})u(t) + C(z^{-1})\lambda e(t) \quad (5.6-6)$$

The optimum estimate e^0 is obtained by minimizing the risk function

$$R(e) \triangleq \frac{1}{2} \sum_{t=1}^N e^*(t)^2 \quad (5.6-7)$$

with respect to e by an iterative method.

The optimum value of λ^2 is obtained by minimizing the function, $L^*(e^0, \lambda^2)$, with respect to λ . Then the optimum estimate

$$\lambda^{o2} = \frac{1}{N} \sum_{t=1}^N (e^*(e^0, t))^2 \quad (5.6-8)$$

$$= \frac{2}{N} R(e^0) \quad (5.6-9)$$

The procedure is less direct in the multivariate case. For convenience in minimizing $L_N(\theta, \Lambda)$, we show five identities:

Identity A

The determinant of the product of two square matrices P and Q is given by [22]:

$$|PQ| = |P| |Q| \quad (5.6-10)$$

If we define a new matrix Φ such that $\Phi^{-1} \triangleq \Lambda$ (5.6-11)

then $|\Phi| |\Lambda| = 1$ (5.6-12)

and $\ln |\Lambda| = -\ln |\Phi|$ (5.6-13)

Identity B

The inner product with a vector variable $x(t)$ given by:

$$\sum_{t=1}^N x^T(t) P x(t) = \sum_{i,j=1}^m \sum_{t=1}^N p_{ij} x_i(t) x_j(t) \quad (5.6-14)$$

From the properties of the trace of a matrix:

$$\text{trace } PQ = \sum_{i,j=1}^m p_{ij}q_{ji} \quad (5.6-15)$$

where p_{ij} and q_{ji} are elements of P and Q . Then

$$\sum_{t=1}^N x^T(t)Px(t) = \text{trace } PQ \quad (5.6-16)$$

where

$$Q \triangleq \sum_{t=1}^N x(t)x^T(t) \quad (5.6-17)$$

Identity C

Let us define a nonsingular $m \times m$ matrix P by

$$P^{-1} = \frac{1}{N} \sum_{t=1}^N x(t)x^T(t) \quad (5.6-18)$$

Substituting in equation (5.6-16) of Identity B:

$$\begin{aligned} \sum_{t=1}^N x^T(t)Px(t) &= \text{trace } N[PP^{-1}] \\ &= mN \end{aligned} \quad (5.6-19)$$

Identity D

Consider the following function of the nonsingular, symmetric matrix \mathbb{F} .

$$\frac{d}{d\mathbb{F}} \ln |\mathbb{F}| = \frac{1}{|\mathbb{F}|} \frac{d}{d\mathbb{F}} |\mathbb{F}| \quad (5.6-20)$$

$$\text{Since } |\mathbb{F}| = \sum_{t=1}^n \Psi_{ij}(\mathbb{F})_{ij} \quad (5.6-21)$$

where $(\mathbb{F})_{ij}$ is the cofactor of element Ψ_{ij} of \mathbb{F} :

$$\frac{d|\Psi|}{d\psi_{ij}} = (\Psi)_{ij} \quad (5.6-22)$$

Then

$$\frac{d}{d\mathbb{I}} \ln |\Psi| = \frac{1}{|\Psi|} \left[\frac{d|\Psi|}{d\psi_{ij}} \right]_{m \times m} \quad (5.6-23)$$

$$= \frac{1}{|\Psi|} \left[(\Psi)_{ij} \right]_{m \times m} \quad (5.6-24)$$

$$= \Psi^{-1} \quad (5.6-25)$$

Identity E

Consider the inner product

$$x^T \Psi x = \sum_{i,j=1}^m x_i \psi_{ij} x_j \quad (5.6-26)$$

Clearly

$$\frac{\partial}{\partial \psi_{ij}} x^T \Psi x = x_i x_j \quad (5.6-27)$$

Let us partition Ψ by columns

$$\Psi \triangleq [\psi_1 \ \psi_2 \ \dots \ \psi_m] \quad (5.6-28)$$

where the element ψ_{ij} is the i^{th} element of column ψ_j . We define the derivative of a scalar with respect to a column vector to be a column vector.

$$\frac{\partial}{\partial \Psi_j} x^T \mathbb{F} x = \begin{bmatrix} \frac{\partial}{\partial \Psi_{1j}} \\ \cdot \\ \cdot \\ \frac{\partial}{\partial \Psi_{mj}} \end{bmatrix} \quad x^T \mathbb{F} x = \begin{bmatrix} x_1 x_j \\ \cdot \\ \cdot \\ x_m x_j \end{bmatrix} \quad (5.6-29)$$

and

$$\begin{aligned} \frac{d}{d\mathbb{F}} x^T \mathbb{F} x &= \begin{bmatrix} x_1^2 & \cdot & x_1 x_m \\ \cdot & x_i x_j & \cdot \\ x_m x_1 & \cdot & x_m^2 \end{bmatrix} \\ &= xx^T \end{aligned} \quad (5.6-30)$$

Similarly

$$\frac{d}{d\mathbb{F}} \sum_{t=1}^N x^T(t) \mathbb{F} x(t) = \sum_{t=1}^N x(t) x^T(t) \quad (5.6-31)$$

Now let us consider minimizing the function $L^*(\Theta, \Lambda)$ (equation (5.6-4)).

(1) The optimum value of Λ for a given Θ is obtained by minimizing $L^*(\Theta, \Lambda)$ with respect to Λ . As we find it more convenient to minimize $L^*(\Theta, \Lambda)$ with respect to \mathbb{F} , we note that at the optimum $\Lambda^{\circ} = (\mathbb{F}^{\circ})^{-1}$.

Incorporating identity A, we obtain:

$$L^*(\Theta, \mathbb{F}) = \frac{mN}{2} \ln(2\pi) - \frac{N}{2} \ln(\mathbb{F}) + \frac{1}{2} \sum_{t=1}^N e^{*T}(t) \mathbb{F} e^*(t) \quad (5.6-32)$$

Then with identities D and E:

$$\begin{aligned} \frac{d}{d\mathbb{F}} L^*(\Theta, \mathbb{F}) &= -\frac{N}{2} \mathbb{F}^{-1} + \frac{1}{2} \sum_{t=1}^N e^*(t) e^{*\mathbb{T}}(t) \\ &= 0 \end{aligned} \quad (5.6-33)$$

for

$$\mathbb{F}^{-1} = \underline{\Lambda} = \frac{1}{N} \sum_{t=1}^N e^*(t) e^{*\mathbb{T}}(t) \quad (5.6-34)$$

$$\stackrel{\Delta}{=} \hat{\Lambda}(\Theta) \quad (5.6-35)$$

That is, for some choice of Θ , we can generate the sequence $\{e^*(t)\}$.

With this choice of Θ , the function $L^*(\Theta, \underline{\Lambda})$ can be minimized with respect to $\underline{\Lambda}$ by setting $\underline{\Lambda} = \hat{\Lambda}(\Theta)$ according to equations (5.6-34) and (5.6-35). The resultant value of $L^*(\Theta, \hat{\Lambda}(\Theta))$ is:

$$\begin{aligned} L^*(\Theta, \hat{\Lambda}(\Theta)) &= \frac{mN}{2} \ln(2\pi) + \frac{N}{2} \ln |\hat{\Lambda}| + \frac{1}{2} \sum_{t=1}^N e^*(t) \hat{\Lambda}^{-1} e^*(t) \\ & \quad (5.6-36) \end{aligned}$$

$$= \frac{mN}{2} \ln(2\pi) + \frac{N}{2} \ln |\hat{\Lambda}| + \frac{mN}{2} \quad (5.6-37)$$

by applying Identity C.

(2) The optimum value of Θ is obtained by minimizing $L^*(\Theta, \hat{\Lambda}(\Theta))$ with respect to Θ . It can be seen from equation (5.6-37) that the optimization problem is equivalent to minimizing $\ln |\hat{\Lambda}(\Theta)|$ with respect to Θ for $\hat{\Lambda}$ given by equations (5.6-34) and (5.6-35). At the optimum

$$L^0(\Theta^0, \hat{\Lambda}^0(\Theta^0)) = \frac{mN}{2} (\ln(2\pi) + 1) + \frac{N}{2} \ln |\hat{\Lambda}^0(\Theta^0)| \quad (5.6-38)$$

Let us now differentiate $\ln |\hat{\Lambda}(\theta)|$ with respect to e_{kl} , an element of θ . We define $\hat{\lambda}_{ij}$ to be an element of $\hat{\Lambda}$ and $\hat{\psi}_{ij}$ to be an element of $\hat{\Psi} = \hat{\Lambda}^{-1}$.

$$\frac{\partial}{\partial e_{kl}} L^*(\theta, \hat{\Lambda}(\theta)) = \frac{\partial}{\partial e_{kl}} \frac{N}{2} \ln |\hat{\Lambda}(\theta)| \quad (5.6-39)$$

$$= \frac{N}{2} \sum_{i,j=1}^m \frac{\partial}{\partial \hat{\lambda}_{ij}} \ln |\hat{\Lambda}(\theta)| \frac{\partial \hat{\lambda}_{ij}}{\partial e_{kl}} \quad (5.6-40)$$

From Identity D, we have by substitution in equation (5.6-25):

$$\frac{\partial}{\partial \hat{\lambda}_{ij}} \ln |\hat{\Lambda}(\theta)| = \hat{\psi}_{ij} \quad (5.6-41)$$

Substituting (5.6-41) into (5.6-40):

$$\begin{aligned} \frac{\partial}{\partial e_{kl}} L^*(\theta, \hat{\Lambda}(\theta)) &= \frac{N}{2} \sum_{i,j=1}^m \hat{\psi}_{ij} \frac{\partial \hat{\lambda}_{ij}}{\partial e_{kl}} \\ &= \sum_{i,j=1}^m \hat{\psi}_{ij} \frac{\partial}{\partial e_{kl}} \left[\frac{1}{2} \sum_{t=1}^N e_i^*(t) e_j^*(t) \right] \\ &= \sum_{i,j=1}^m \hat{\psi}_{ij} \sum_{t=1}^N e_i^*(t) \frac{\partial e_j^*(t)}{\partial e_{kl}} \quad (5.6-42) \end{aligned}$$

$$\begin{aligned} &= \sum_{t=1}^N e^{*T}(t) \hat{\Psi} \frac{\partial e^*(t)}{\partial e_{kl}} \\ &= \sum_{t=1}^N e^{*T} \hat{\Lambda}^{-1}(\theta) \frac{\partial e^*(t)}{\partial e_{kl}} \quad (5.6-43) \end{aligned}$$

$$\triangleq L_{e_{kl}}^*(\theta, \hat{\Lambda}(\theta)) \quad (5.6-44)$$

When calculating derivatives of L^* for all elements of Θ , a notational difficulty arises. The gradient of L^* with respect to Θ is a matrix and the array of second derivatives is a three dimensional array. For this reason, we prefer to decompose Θ into m row vectors e_i^T for $i = 1, 2, \dots, m$ as in previous sections. Then e_i is a vector of $n+(p+1)r+pm$ elements from the i^{th} rows of $A_1, \dots, A_p, B_0, \dots, B_p, C_1, \dots, C_p$. As before, we designate \underline{e} as the long vector of vectors e_i :

$$\underline{e} = \begin{bmatrix} e_1 \\ e_2 \\ \cdot \\ \cdot \\ e_m \end{bmatrix}_{s \times 1} \quad (5.6-45)$$

where s is now

$$s = m(n + (p + 1)r + pm) \quad (5.6-46)$$

Then the vector form of equation (5.6-44) is

$$\underline{L}_e^*(\underline{e}, \hat{\Lambda}(\underline{e})) = \sum_{t=1}^N e^{*T}(t) \hat{\Lambda}^{-1}(\underline{e}) \frac{\partial e^*(t)}{\partial \underline{e}} \quad (5.6-47)$$

The matrix of second derivatives of L^* with respect to elements e_{ik} and e_{jl} of \underline{e} is obtained from equation (5.6-43). The elements of the matrix are given by:

$$\frac{\partial^2}{\partial e_{ik} \partial e_{jl}} L^*(\theta, \hat{\lambda}(\theta)) = \frac{\partial}{\partial e_{ik}} L_{e_{jl}}^*(\theta, \hat{\lambda}(\theta)) \quad (5.6-48)$$

$$= \sum_{t=1}^N \frac{\partial e^{*\top}(t)}{\partial e_{ik}} \hat{\lambda}^{-1} \frac{\partial e^*(t)}{\partial e_{jl}} + \sum_{t=1}^N e^{*\top}(t) \hat{\lambda}^{-1} \frac{\partial^2 e^*(t)}{\partial e_{ik} \partial e_{jl}} \quad (5.6-49)$$

$$\stackrel{\Delta}{=} L_{e_{ik} e_{jl}}^*(\theta, \hat{\lambda}(\theta))$$

One particularly useful aspect of this decomposition is that the gradient and the matrix of second derivatives of $L^*(\theta)$ with respect to elements of θ have precisely the same structure (and perform the same function) as their counterparts in the single-observation case. This will be discussed later.

5.6.3 The Iterative Solution of the MLE Problem

The solution in the single-output case is notably simpler than in the multivariate case. The reason is that it has been possible to find a risk function ($R(\theta)$ given by equation (5.6-7)) that is easier to minimize than the negative logarithm of the likelihood function as in the multivariate case.

In the single-output case discussed in [19], a second order method is employed for updating e . The gradient of $R(e)$ is obtained by

$$R_{\mathbf{e}}(\mathbf{e}) \triangleq \frac{\partial R(\mathbf{e})}{\partial \mathbf{e}_i} \quad (5.6-50)$$

$$= \sum_{t=1}^N e^*(t) \frac{\partial e^*(t)}{\partial \mathbf{e}_i} \quad (5.6-51)$$

where \mathbf{e}_i is an element of vector \mathbf{e} . The elements of the gradient vector are calculated by the relations obtained from (5.6-51):

$$\begin{aligned} C(z^{-1}) \frac{\partial e^*(t)}{\partial A_i} &= z^{-i} y(t) \\ C(z^{-1}) \frac{\partial e^*(t)}{\partial B_i} &= -z^{-i} u(t) \\ C(z^{-1}) \frac{\partial e^*(t)}{\partial C_i} &= -z^i e^*(t) \end{aligned} \quad (5.6-52)$$

where A_i , B_i and C_i are scalar parameters of $A(z^{-1})$, $B(z^{-1})$ and $C(z^{-1})$. The second order partial derivatives of $e^*(t)$ are obtained from

$$R_{\mathbf{e}\mathbf{e}}(\mathbf{e}) \triangleq \frac{\partial^2 R(\mathbf{e})}{\partial \mathbf{e}_i \partial \mathbf{e}_j} \quad (5.6-53)$$

$$= \sum_{t=1}^N \left(\frac{\partial e^*(t)}{\partial \mathbf{e}_i} \cdot \frac{\partial e^*(t)}{\partial \mathbf{e}_j} + e^*(t) \frac{\partial^2 e^*(t)}{\partial \mathbf{e}_i \partial \mathbf{e}_j} \right) \quad (5.6-54)$$

To minimize the risk function, the following Newton-Raphson algorithm is used at the $k + 1$ iteration.

$$\mathbf{e}(k+1) = \mathbf{e}(k) - R_{\mathbf{e}\mathbf{e}}^{-1}(\mathbf{e}(k)) R_{\mathbf{e}}(\mathbf{e}(k)) \quad (5.6-55)$$

In the multivariate case, we use the following procedure:

(1) Given $\Theta(k)$ where $\Theta(k)$ stands for the model parameters in

$A(z^{-1})$, $B(z^{-1})$ and $C(z^{-1})$, we calculate the sequence $\{e^*(t; \theta(k))\}$ by the equation:

$$e^*(t; \theta(k)) = C(z^{-1})[A(z^{-1})y(t) - B(z^{-1})u(t)] \quad (5.6-56)$$

To avoid complication at this point, we shall assume that the observations are available from the time of zero initial conditions in the plant.

(2) $\hat{\Lambda}(\theta(k))$ is calculated by

$$\hat{\Lambda}(\theta(k)) = \frac{1}{N} \sum_{t=1}^N e^*(t; \theta(k)) e^{*\top}(t; \theta(k)) \quad (5.6-57)$$

(3) The derivatives of $e^*(t)$ are obtained by differentiating the difference equation written as

$$C(z^{-1})e^*(t) = A(z^{-1})y(t) - B(z^{-1})u(t) \quad (5.6-58)$$

For example, if A_{kij} is the element of the i^{th} row and j^{th} column of A_k , then

$$C(z^{-1}) \frac{\partial e^*(t)}{\partial A_{kij}} = \begin{bmatrix} 0 \\ \cdot \\ z^{-k} y_j(t) \\ \cdot \\ 0 \end{bmatrix} \leftarrow i^{\text{th}} \text{ element} \quad (5.6-59)$$

$m \times 1$

Similarly:

$$C(z^{-1}) \frac{\partial e^*(t)}{\partial B_{kij}} = \begin{bmatrix} 0 \\ \cdot \\ -z^{-k} u_j(t) \\ \cdot \\ 0 \end{bmatrix} \leftarrow i^{\text{th}} \text{ element} \quad (5.6-60)$$

$m \times 1$

$$C(z^{-1}) \frac{\partial e^*(t)}{\partial C_{kij}} = \begin{bmatrix} 0 \\ \cdot \\ -z^{-k} e_j(t) \\ \cdot \\ 0 \end{bmatrix} \leftarrow i^{\text{th}} \text{ element} \quad (5.6-61)$$

$m \times 1$

Considerable simplification can be achieved by shifting [19]. We note that for $k \leq t + 1$

$$\frac{\partial e(t)}{\partial A_{kij}} = \frac{\partial e(t-i+1)}{\partial A_{1ij}} \quad (5.6-62)$$

hence, it is only necessary to calculate equations (5.6-59) and (5.6-61) for $k = 1$ and equation (5.6-60) for $k = 0$. The remaining terms are obtained by shifting. The second order partial derivatives for $L_{\underline{e}\underline{e}}^*(\underline{\theta}, \hat{\Lambda}(\underline{\theta}))^{\text{arc}}$ computed in a similar fashion. Considerable saving in numerical effort is again realized by shifting.

The algorithm for minimizing the risk function in the scalar case ($m = 1$) is [18, 19]:

- (1) Set $e(1) =$ [least squares (biased) estimate for the parameters of $A(z^{-1})$ and $B(z^{-1})$ with $C_i = 0$ for $i > 0$].
- (2) Evaluate $R_e(e(k))$ and $R_{\underline{e}\underline{e}}(e(k))$ with equations (5.6-51) to (5.6-54) where k is an iteration index.
- (3) Calculate $e(k + 1)$ by equation (5.6-55) and repeat from (2).

The algorithm for minimizing the risk function in the multivariate case is essentially the same.

- (1) Set $\underline{e}(1) = [\text{LSE for the parameters of } A(z^{-1}) \text{ and } B(z^{-1})$
with $C_i = 0$ for $i > 0]$.
- (2) Evaluate $L_{\underline{e}}^*(\theta(k))$ and $L_{\underline{ee}}^*(\theta(k))$ with equations (5.6-47) and
(5.6-49).
- (3) Calculate

$$\underline{e}(k+1) = \underline{e}(k) - L_{\underline{ee}}^{*-1}(\theta(k))L_{\underline{e}}^*(\theta(k)) \quad (5.6-63)$$

Even for very small systems, $L_{\underline{ee}}^*(\theta)$ is a (symmetric) matrix of formidable size. For example, if $n = 4$, $m = 2$ and $r = 2$, the dimension of $L_{\underline{ee}}^*(\theta)$ is $2(4 + 3 \times 2 + 2 \times 2) = 28$.

5.6.4 An Alternate Solution of the MLE Problem

To ease the programming and computational burden, the decision was made to develop an alternate algorithm for the multivariate case, without explicitly calculating $L_{\underline{ee}}^*(\theta)$ or its inverse, but preserving the superior convergence capability of a second order method. This decision was based on the following experience:

(1) Studies in the estimation of the parameters of simulated plant models (with $m = 1$) indicated that the LSE "bootstrap estimator" discussed in a previous section provided estimates close to the true values but the estimated standard deviations did not instil confidence in the results.

(2) Experience derived in using the second-order numerical descent methods of Fletcher and Powell [31] and Powell [32] for function minimization led the writer to believe that application of one or other of these methods would rapidly converge on the optimum solution provided

that a good initial estimate for θ was supplied.

In the Fletcher and Powell routine, the algorithm for function minimization with respect to \underline{e} is (compare with equation 5.6-63):

$$\underline{e}(k+1) = \underline{e}(k) - H(k)L_{\underline{e}}^*(\theta(k)) \quad (5.6-64)$$

where $H(k)$ is a positive definite symmetric matrix updated with the most recent gradient information. The algorithm ensures that as the procedure converges, H tends to the inverse of the matrix of second derivatives evaluated at the minimum. The initial condition for the inverse matrix of second derivatives is taken to be

$$H(0) = I_s \quad (5.6-65)$$

On the assumption that the risk-function surface is nearly quadratic near to the optimum and after (at least) s iterations, $H(k \geq s)$ is taken to be an estimate for $L_{\underline{e}\underline{e}}^{*-1}(\theta(k))$.

5.6.5 The Adaptive Filter

It is constructive to compare the maximum likelihood method discussed here with a generalized least-squares estimator in which the parameters of the filter in the GLSE are adjusted to minimize a minimum-variance risk function. Recall from the discussion of Sub-section 5.4.3 that Tretter and Steiglitz [50], and Clarke [51], adjusted the parameters of $C(z^{-1})$ (scalar case only) to minimize the variance of the residuals. Note that in the MLE method, the computation of $L(\theta, \hat{\Lambda}(\theta))$, the gradient $L_{\underline{e}}^*(\theta)$ and the matrix of second derivatives

all require numerical filtering by operation with $C^{-1}(z^{-1})$.

For example

$$\hat{\lambda}(\theta) \triangleq \sum_{t=1}^N e^*(t) e^{*T}(t) \quad (5.6-66)$$

requires the calculation of

$$e^*(t) = C^{-1}(z^{-1})A(z^{-1})y(t) - C^{-1}(z^{-1})B(z^{-1})u(t) \quad (5.6-67)$$

$$= C^{-1}(z^{-1})v^*(t) \quad (5.6-68)$$

$$\text{where } v^*(t) \triangleq A(z^{-1})y(t) - B(z^{-1})u(t) \quad (5.6-69)$$

by computing

$$\begin{aligned} e^*(t) = & -C_1 e^*(t-1) \dots -C_p e^*(t-p) \\ & + y(t) + A_1 y(t-1) \dots A_p (I_m)_p y(t-p) \\ & - B_0 u(t) - B_1 u(t-1) \dots - B_p u(t-p) \end{aligned} \quad (5.6-70)$$

Similar expressions can be obtained for the partial derivatives of $e^*(t)$. The following important points can be made:

- (1) The process of finding $e^*(t)$ (and its derivatives) is one of filtering by operation with $C^{-1}(z^{-1})$. Clearly the filter must be stable. (The roots of the characteristic equation for $C^{-1}(z^{-1})$ must lie within the unit circle in the z plane).
- (2) The problem of stability in the MLE case is identical with that of the GLSE case. (See the discussion of Subsection 5.4.5).
- (3) An important distinction can be drawn between the MLE methods

and the adaptive GLSE method. In the former, the parameters are adjusted together; they have a direct effect on the likelihood function which can be gauged by observing the reduction in the gradient of the likelihood function. In the adaptive GLSE method, adjustment of the filter parameters has a somewhat indirect effect because of the sequencing of operations in the two-stage procedure. That is, after altering the parameters of $C(z^{-1})$ according to past performance and filtering the data, the parameters of $A(z^{-1})$ and $B(z^{-1})$ are adjusted to minimize the squares of the residuals. The intuitive impression is that the two-stage procedure of the adaptive GLSE approach couples the parameters of $A(z^{-1})$ and $B(z^{-1})$ to $C(z^{-1})$ rather more tightly than in the MLE method which is a one-stage procedure. This may make minimization of the risk function more difficult.

5.6.6 Discussion of the MLE Method

We have already noted that by decomposing θ into the vector $\underline{\theta}$, functions $L_{\underline{\theta}}^*(\theta)$ and $L_{\underline{\theta}\underline{\theta}}^*(\theta)$ resemble their counterparts in the single-output case. Summarizing, the models are:

$$e^*(t) = C^{-1}(z^{-1})[A(z^{-1})y(t) - B(z^{-1})u(t)]$$

where in the multivariate case, $e^*(t)$ is the m -variate distribution $n(0, \Lambda)$ and in the single-output case, $e^*(t)$ is distributed according to $n(0, \lambda^2)$.

The multivariate case likelihood function is:

$$L^*(\theta, \Lambda(\theta)) = \frac{mN}{2} \ln(2\pi) + \frac{N}{2} \ln |\Lambda| + \frac{1}{2} \sum_{t=1}^N e^{*\top}(t) \Lambda^{-1} e^*(t) \quad (5.6-71)$$

and in the scalar case:

$$L^*(\mathbf{e}, \lambda^2(\mathbf{e})) = \frac{N}{2} \ln(2\pi) + \frac{N}{2} \ln \lambda^2 + \frac{1}{2\lambda^2} \sum_{t=1}^N (e^*(t))^2 \quad (5.6-72)$$

For

$$\hat{\Lambda}(\boldsymbol{\theta}) = \sum_{t=1}^N e^*(t) e^{*\text{T}}(t) \quad (5.6-73)$$

$$\hat{\lambda}^2(\boldsymbol{\theta}) = \sum_{t=1}^N (e^*(t))^2 \quad (5.6-74)$$

$$L^*(\boldsymbol{\theta}, \hat{\Lambda}(\boldsymbol{\theta})) = \frac{mN}{2} (\ln(2\pi) + 1) + \frac{N}{2} \ln |\hat{\Lambda}|(\boldsymbol{\theta}) \quad (5.6-75)$$

$$L^*(\mathbf{e}, \hat{\lambda}^2(\mathbf{e})) = \frac{N}{2} (\ln(2\pi) + 1) + \frac{N}{2} \ln \hat{\lambda}^2(\mathbf{e}) \quad (5.6-76)$$

Note that the effective term to be minimized in (5.6-75) is $\ln |\hat{\Lambda}|(\boldsymbol{\theta})$ whereas in (5.6-76) it is simply $\hat{\lambda}^2(\mathbf{e})$.

The gradient of L^* is:

$$\underline{\frac{\partial}{\partial \mathbf{e}}} L^*(\boldsymbol{\theta}, \hat{\Lambda}(\boldsymbol{\theta})) = \sum_{t=1}^N e^{*\text{T}}(t) \hat{\Lambda}^{-1}(\boldsymbol{\theta}) \underline{\frac{\partial e^*(t)}{\partial \mathbf{e}}} \quad (5.6-77)$$

whereas

$$\underline{\frac{\partial}{\partial \mathbf{e}}} L^*(\mathbf{e}, \hat{\lambda}^2(\mathbf{e})) = \frac{1}{\hat{\lambda}^2} \sum_{t=1}^N e^*(t) \underline{\frac{\partial e^*(t)}{\partial \mathbf{e}}} \quad (5.6-78)$$

Finally, the elements of the matrix $\underline{\underline{\frac{\partial^2}{\partial \mathbf{e} \partial \mathbf{e}}}} L^*(\boldsymbol{\theta}, \hat{\Lambda}(\boldsymbol{\theta}))$ are given by

$$\begin{aligned} \frac{\partial^2}{\partial e_{ik} \partial e_{jl}} L^*(\boldsymbol{\theta}, \hat{\Lambda}(\boldsymbol{\theta})) &= \sum_{t=1}^N \frac{\partial e^{*\text{T}}(t)}{\partial e_{ik}} \hat{\Lambda}^{-1} \frac{\partial e^*(t)}{\partial e_{jl}} \\ &+ \sum_{t=1}^N e^{*\text{T}}(t) \hat{\Lambda}^{-1} \frac{\partial^2 e^*(t)}{\partial e_{ik} \partial e_{jl}} \end{aligned} \quad (5.6-79)$$

and of $L_{ee}^*(e, \hat{\lambda}^2(e))$ by

$$\begin{aligned} \frac{\partial^2}{\partial e_i \partial e_j} L^*(e, \hat{\lambda}^2(e)) &= \frac{1}{\hat{\lambda}^2} \sum_{t=1}^N \frac{\partial e^*(t)}{\partial e_i} \frac{\partial e^*(t)}{\partial e_j} \\ &+ \frac{1}{\hat{\lambda}^2} \sum_{t=1}^N e^*(t) \frac{\partial^2 e^*(t)}{\partial e_i \partial e_j} \end{aligned} \quad (5.6-80)$$

In [18, 19], Astrom and Bohlin discuss the large sample properties of the maximum likelihood estimator for the single output model above.

They show that:

(1) for large N , the estimate e^0 obtained by the method is unique and asymptotically consistent subject to the conditions that:

(a) the system, the model and the optimal predictor of $e^*(t)$ are all stable (that is, the roots of $A(z^{-1})$ and $C(z^{-1})$ are all within the unit circle in the plane)

(b) the input $u(t)$ meets certain boundedness conditions and is persistently exciting. In practice, the conditions can be met if the system is excited by an observed random process having finite moments. For example, a pseudorandom binary sequence (PRBS) is a satisfactory source of excitation.

(c) Every state of the system is controllable from either $u(t)$ or $e(t)$.

(2) The estimate e^0 is a stochastic variable that is asymptotically normal with mean being the true value of e . One important consequence of this point is that calculation of the matrix $L_{ee}^{*-1}(e^0)$ gives the covariances of the parameter estimates directly. Another is that

confidence regions can be determined and significance tests performed.

(3) The estimates are asymptotically efficient. This means that we cannot expect to find an estimator with greater accuracy for long samples.

The formal derivation by Astrom and Bohlin [18, 19], the results of which are summarized above, also includes the multivariate case with only a minor extension. We require that the elements of the vector input sequence $\{u(t)\}$ are uncorrelated with each other. That is

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=\tau}^N u_i(t)u_j(t-\tau) \longrightarrow 0 \quad \text{for every } i, j = 1, 2, \dots, m$$

and for every positive integer τ . Then:

(1) for large N , the estimate $\hat{\theta}^0$ is unique and asymptotically unbiased subject to the conditions given previously;

(2) $\hat{\theta}^0$ is a stochastic variable that is asymptotically normal. The mean of the distribution is the true value of θ . The parameter covariance matrix is given by $L_{\theta\theta}^{*-1}(\theta^0, \hat{\Lambda}(\theta^0))$.

(3) The estimates are efficient.

CHAPTER 6

EXAMPLES OF IDENTIFICATION, PREDICTION AND CONTROL

6.1 Introduction

The purpose of this chapter is to demonstrate how the modelling techniques developed in the previous five chapters can be employed in problems of identification, prediction and control. In addition to the restrictions that were assumed in earlier chapters (with respect to linearity, controllability, observability, stationarity, etc.) we assume the following conditions for reasons of simplicity:

- (1) the systems are initially at rest
- (2) the systems have zero offset.

These limitations can be removed in practice without difficulty.

The first few sections are devoted to contrasting the "bootstrap" method with other methods for estimating the parameters of systems of known order. The problem of identifying systems of unknown order is then briefly reviewed. The thirty-one parameters of a multivariate fourth order system are estimated by both the MLE and "bootstrap" estimators.

In the later examples, the importance of minimizing the variance of the one-step-ahead prediction error is emphasized because this is the crucial step in deriving minimum variance control algorithms. To conclude, a simple multivariate control example is shown.

6.2 A Comparison of Some Sequential Least Squares Estimators of a Parameter of a First Order System

6.2.1 Purpose and Procedure

To illustrate the features and failings of

- (a) the conventional least squares estimator (Section 5.3)
- (b) a discrete-time version of Levadi's modified least squares estimator (Subsection 5.5.2)
- (c) the author's "bootstrap" modified least squares estimator (Subsection 5.5.4)

the estimators were compared in the following problem.

The simple system below was simulated on the digital computer.

$$x(t+1) = Fx(t) + u(t) + I v_1(t) \quad (6.2-1)$$

$$y(t) = x(t) + v_2(t) \quad (6.2-2)$$

where x , y , u , v_1 and v_2 were all scalar quantities and $F = 0.8$.

The three sequential estimators were required to estimate the parameter F and the standard deviation of the estimate for the two cases of $\Gamma = 1$ (the state disturbed by v_1) and $\Gamma = 0$.

The difference equation of the process for estimating purposes is given by

$$w(t+1) = Fy(t) + v(t+1) \quad (6.2-3)$$

$$\text{where } w(t+1) = y(t+1) - u(t) \quad (6.2-4)$$

$$v(t+1) = v_2(t+1) - Fv_2(t) + \Gamma v_1(t) \quad (6.2-5)$$

Zero initial conditions were taken each time. The input excitation was provided by a 63 bit pseudorandom binary sequence (PRBS) of \pm unit amplitude adjusted for zero mean. The stochastic variables $v_1(t)$ and $v_2(t)$ were generated by a machine algorithm providing an $n(0,1)$ distribution of random numbers.

6.2.2 Observations and Discussion

A comparison is shown in Table 6.1 which displays estimates \hat{F} for F along with their standard deviations calculated from the variance estimated by the relation

$$\text{s.d.} = [\hat{V}(\sum_{t=1}^N y(t)\hat{y}(t))^{-1}]^{1/2} \quad (6.2-6)$$

where $\hat{y}(t) = y(t)$ for the conventional estimator but $\hat{y}(t)$ is obtained by a linear predictor in the case of the modified estimators. \hat{V} is the estimated variance of the noise process $v(t)$ based on the last computed estimate \hat{F} . The variance expression within the square brackets in the above expression was derived in Subsection 5.5.4.

There, the true variance V of the noise process was assumed known. In addition, it was assumed that N was large enough for $\frac{1}{N} \sum_{t=1}^N y(t)\hat{y}(t)$ to be equal to $E y(t)\hat{y}(t)$ with negligible error.

For comparison purposes, we have calculated the "standard deviations" for the conventional estimator as in (6.2-6) but with the replacement of $y(t)$ for $\hat{y}(t)$. This gives a conservative but imperical measure.

Trial	Run Length	True F	Conventional Least Squares		Levadi's Method		Bootstrap Method	
			\hat{F}	s.d.	\hat{F}	s.d.	\hat{F}	s.d.
$\Gamma = 0$	189	0.8	0.647	0.0433	<u>*0.826</u>	0.0525	<u>0.831</u>	0.0500
$\Gamma = 0$	190	0.8	0.555	0.0452	<u>0.756</u>	0.0568	<u>*0.757</u>	0.0552
Continue	567	0.8	0.517	0.0243	0.749	0.0326	<u>*0.774</u>	0.0359
$\Gamma = 1$	189	0.8	0.687	0.0453	<u>0.777</u>	0.0692	<u>*0.814</u>	0.0562
$\Gamma = 1$	189	0.8	0.659	0.0484	<u>0.746</u>	0.0790	<u>*0.792</u>	0.0381
$\Gamma = 1$	190	0.8	0.701	0.0465	<u>0.871</u>	0.0727	*0.862	0.0554
Continue	567	0.8	0.680	0.0270	0.845	0.0406	<u>*0.811</u>	0.0322

* Best estimate of trial

___ Within one s.d. of true

Table 6.1 Comparison of Conventional and Modified Least Squares Estimators

The best estimates are marked (*) in Table 6.1. Estimates within one standard deviation of the true parameter are underlined.

The bias in the conventional least squares estimator shows up strongly in both the table and the graphs of Figures 6.1 and 6.2. These display the convergence of the estimators for the two longer runs summarized in the table.

In the examples shown we observe that:

- (1) The conventional LSE is strongly biased in this example.
- (2) The "bootstrap" estimate is closer (in all but one case) to the true parameter value than the discrete-time version of Levadi's estimator.
- (3) The standard deviations calculated for the "bootstrap" estimator are smaller (in all but one case) than the standard deviations calculated for the other modified LSE.
- (4) In spite of the observation in (3) above, the true parameter is within one standard deviation of the estimated parameter in one more case with the "bootstrap" estimator than the other. All estimates obtained by the modified estimators are within two standard deviations of the true parameter.

We conclude that although the discrete-form of the Levadi estimator is considerably improved over the conventional LSE, the "bootstrap" estimator is superior and merits further consideration and evaluation as a useful method of estimating parameters of system models when conventional LSE methods lead to biased estimates.

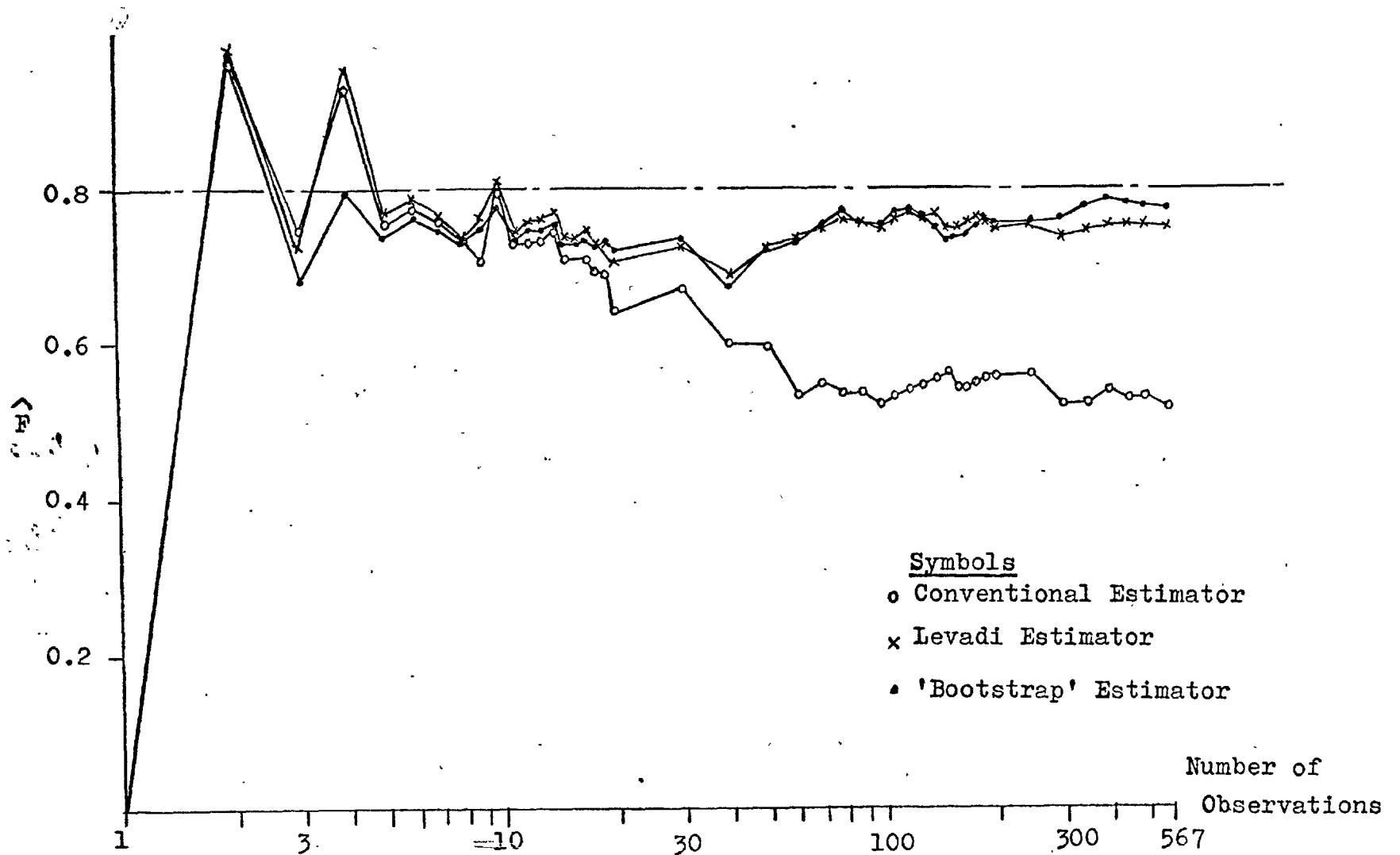


Figure 6.1 Comparison of Estimates ($F = 0$)

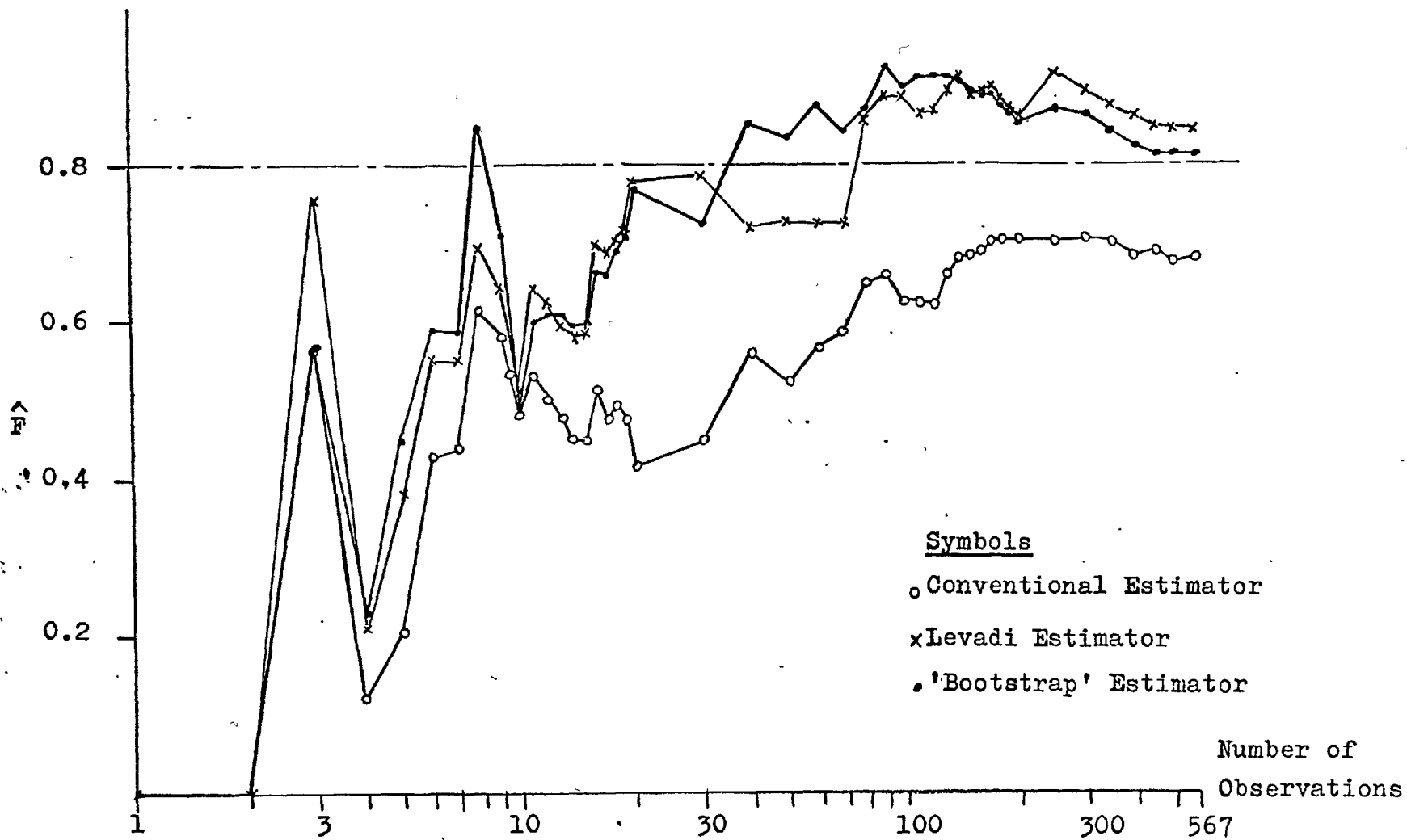


Figure 6.2 Comparison of Estimates ($T = 1$)

6.3 The Sequential Estimation of the Parameters of a Second Order System with the "Bootstrap" Method

6.3.1 Purpose and Procedure

In this section, the "bootstrap" method is used to estimate the parameters of a second-order discrete-time system during simulation. The model employed is that used by Mayne for illustrating his method for estimating discrete-time transfer functions [20].

$$(1 - 1.5z^{-1} + 0.7z^{-2})y(t) = (1 + z^{-1} + z^{-2})u(t) + 1(1 - z^{-1} + 0.2z^{-2})e(t) \quad (6.3-1)$$

The regression model form is

$$y(t) = e_1 y(t-1) + e_2 y(t-2) + e_3 u(t) + e_4 u(t-1) + e_5 u(t-2) + v(t) \\ \triangleq \mathbf{x}^T(t)\mathbf{e} + v(t) \quad (6.3-2)$$

$$\text{where } v(t) \triangleq \lambda(e(t) + c_1 e(t-1) + c_2 e(t-2)) \quad (6.3-3)$$

Both the control sequence and the disturbance sequence were from a distribution of normal random variables of unit variance.

The sequential estimation procedure to find the five parameters in \mathbf{e} is given in Subsection 5.5.4. The estimating equations are given by:

$$\hat{\mathbf{e}}(t) = \hat{\mathbf{e}}(t-1) + \frac{P(t)\hat{\mathbf{x}}(t)(y(t) - \mathbf{x}^T(t)\hat{\mathbf{e}}(t-1))}{\mathbf{x}^T(t)P(t)\hat{\mathbf{x}}(t) + 1} \quad (6.3-4)$$

$$P(t) = P(t-1) - \frac{P(t-1)\hat{\mathbf{x}}(t)\mathbf{x}^T(t)P(t-1)}{\mathbf{x}^T(t)P(t)\hat{\mathbf{x}}(t) + 1} \quad (6.3-5)$$

The above equations were employed following the fifth iteration after obtaining the minimum data set solution at the fifth iteration by matrix inversion. The prediction procedure for finding $\hat{x}(t)$ was employed for the first time at the sixth iteration and used the estimate $\hat{e}(5)$ from the fifth iteration.

Estimates of the serial autocorrelations $\hat{R}_v(0)$, $\hat{R}_v(1)$ and $\hat{R}_v(2)$ were obtained by the procedure given in Subsection 5.5.4 by storing the quantities

$$\begin{aligned} S_1(\tau) &= \sum_{t=\tau+1}^N y(t)y(t-\tau) \\ S_2(\tau) &= \sum_{t=\tau+1}^N y(t)x(t-\tau) \\ S_3(\tau) &= \sum_{t=\tau+1}^N x(t)y(t-\tau) \\ S_4(\tau) &= \sum_{t=\tau+1}^N x(t)x^T(t-\tau) \end{aligned} \quad (6.3-6)$$

for $\tau = 0, 1$ and 2 . Then

$$\hat{R}_v(\tau) = \frac{1}{N-\tau-1} [S_1(\tau) - S_2^T(\tau)\hat{e}(N) - S_3^T(\tau)\hat{e}(N) + \hat{e}^T(N)S_4\hat{e}(N)] \quad (6.3-7)$$

etc. For the coefficients in $v(t)$, the parameters $c_0^* \triangleq \lambda$, $c_1^* = c_1/\lambda$ and $c_2^* = c_2/\lambda$ were found by the method of moments described in Subsection 4.3.3 with the initial estimate

$$\hat{c}_0^* = (\hat{R}_v(0))^{1/2}; \quad \hat{c}_1^* = 0; \quad \hat{c}_2^* = 0$$

6.3.2 Observations and Discussion

The estimates obtained from an ensemble of ten trials, each commencing from zero initial conditions are given in Table 6.2. Standard deviations computed from the variances of the estimates are shown alongside the parameters $\hat{\theta}$. If the true value falls within the calculated standard deviation from the estimate, the estimate is shown underlined. Three out of fifty estimates are outside two standard deviations. They are followed by an asterisk.

In Table 6.3, the ensemble estimates $\hat{\theta}$ for the ten trials are summarized and compared with the ensemble results of ten similar trials carried out by Tzafestas when testing Mayne's method [21]. It is not clear from the referenced report if the parameters of the predictor equation were derived from the current running data or from a separate record. (In [20], Mayne estimated the parameters for the linear predictor from a separate record using normal regression analysis.)

The number of trials is insufficient to conclude positively that one estimator is superior to the other. Because e is known, it is possible to calculate the true error which is displayed as a percentage of the true value. By this comparison, the "bootstrap" estimator is superior.

Note that for both estimators, the true value of e falls within one standard deviation from the estimated mean.

The estimates of both the autocorrelation coefficients and the parameters of the stochastic process $v(t)$ are given in Table 6.4.

Table 6.2 The Parameters of a Second Order System Estimated by the "Bootstrap" Method

True Values		$\theta_1 = 1.5$	$\theta_2 = -0.7$	$\theta_3 = 1.0$	$\theta_4 = 1.0$	$\theta_5 = 1.0$
Trial	Iterations	$\hat{\theta}_1$ s.d.	$\hat{\theta}_2$ s.d.	$\hat{\theta}_3$ s.d.	$\hat{\theta}_4$ s.d.	$\hat{\theta}_5$ s.d.
1	240	1.475	-0.678	0.973	1.169	0.838
	1,000	<u>1.497</u> 0.0181	<u>-0.695</u> 0.0166	<u>1.023</u> 0.0479	<u>0.987</u> 0.0508	<u>0.944</u> 0.0559
2	240	1.463	-0.675	1.054	1.032	1.186
	1,000	1.483 0.0156	<u>-0.688</u> 0.0149	<u>1.043</u> 0.0462	<u>1.036</u> 0.0488	1.097 0.0527
3	240	1.463	-0.673	0.897	1.007	1.214
	1,000	<u>1.495</u> 0.0141	<u>-0.698</u> 0.0134	<u>1.004</u> 0.0428	<u>1.027</u> 0.0461	<u>1.027</u> 0.0494
4	240	1.465	-0.672	0.908	0.997	1.220
	1,000	<u>1.487</u> 0.0166	<u>-0.692</u> 0.0155	<u>0.990</u> 0.0464	<u>1.070</u> 0.0491	<u>1.004</u> 0.0539
5	240	1.486	-0.690	0.908	1.068	1.037
	1,000	<u>1.509</u> 0.0150	<u>-0.710</u> 0.0145	0.920 0.0462	<u>1.043</u> 0.0491	0.944 0.0540
6	240	1.499	-0.700	1.038	1.125	0.871
	1,000	<u>1.508</u> 0.0163	<u>-0.707</u> 0.0154	0.953 0.0461	1.094 0.0493	0.888* 0.0545
7	240	1.484	-0.681	0.993	1.127	0.964
	1,000	1.481 0.0154	-0.683 0.0147	<u>0.996</u> 0.0453	<u>1.047</u> 0.0474	<u>1.041</u> 0.0527
8	240	1.495	-0.696	1.047	0.997	1.009
	1,000	1.516 0.0149	<u>-0.712</u> 0.0142	<u>1.039</u> 0.0448	<u>0.979</u> 0.0472	0.927 0.0522
9	240	1.462	-0.673	1.013	1.133	0.937
	1,000	<u>1.498</u> 0.0175	-0.698 0.0165	<u>1.027</u> 0.0475	<u>1.025</u> 0.0504	<u>0.952</u> 0.0553
10	240	1.484	-0.685	0.930	1.004	0.977
	1,000	<u>1.493</u> 0.0157	<u>-0.695</u> 0.0150	<u>0.998</u> 0.0469	0.891* 0.0493	1.109* 0.0536

Method	Parameter	θ_1	θ_2	θ_3	θ_4	θ_5
	True Value	1.5	-0.7	1.0	1.0	1.0
"Bootstrap" Method	Mean	1.4966	-0.6979	0.9993	1.0198	0.9933
	Per Cent Error	-0.2251	-0.3015	-0.0734	1.9796	-0.6749
	Standard Deviation	0.0111	0.0095	0.0384	0.0532	0.0744
Mayne's Method (Tzafestas)	Mean	1.5082	-0.7082	0.9981	0.9919	0.9744
	Per Cent Error	0.5487	1.1786	-0.1930	-0.8110	-2.5620
	Standard Deviation	0.0212	0.0289	0.0307	0.0258	0.0752

Table 6.3 Comparison of the Estimates of θ by the "Bootstrap" Method and
Mayne's Method

Parameters		$R_V(0)$	$R_V(1)$	$R_V(2)$	λ	C_1	C_2
True Values		2.04	-1.2	0.2	1.0	-1.0	0.2
Trial	Iterations	$\hat{R}_V(0)$	$\hat{R}_V(1)$	$\hat{R}_V(2)$	$\hat{\lambda}$	\hat{C}_1	\hat{C}_2
1	1000	2.264	-1.328	0.221	1.061	-1.000	0.196
2	1000	2.164	-1.281	0.231	1.045	-0.968	0.211
3	1000	1.906	-1.121	0.182	0.955	-1.025	0.200
4	1000	2.085	-1.228	0.239	1.057	-0.905	0.214
5	1000	2.104	-1.341	0.438	1.043	-0.877	0.401
6	1000	2.178	-1.340	0.333	1.064	-0.913	0.293
7	1000	2.000	-1.171	0.209	1.024	-0.929	0.200
8	1000	2.001	-1.140	0.154	1.003	-0.980	0.152
9	1000	2.165	-1.338	0.351	1.067	-0.899	0.308
10	1000	2.167	-1.290	0.282	1.088	-0.877	0.238
Mean		2.103	-1.258	0.274	1.041	-0.937	0.241

Table 6.4 The Parameters of a Noise Process Estimated by the Method of Moments

It is not possible to determine the variance of these estimates at each trial. In the next and later sections, the merit of this method of finding the stochastic process parameters will be assessed by comparing its performance with a maximum likelihood method.

The entire set of ten trials, including the simulation, estimation and data reduction parts took 4.5 minutes of machine time for compilation and execution.

6.4 A Comparison of the Estimation of a Second Order System by the "Bootstrap" and MLE Methods

6.4.1 Purpose and Procedure

The estimation problem is:

given the input sequence $\{u(t), t = 1, \dots, N\}$

and the output sequence $\{y(t), t = 1, \dots, N\}$

of the process discussed in the last section and described by equation (6.3-1), estimate the parameters A_1 , A_2 , B_0 , B_1 , B_2 , λ , C_1 and C_2 . We shall compare the "bootstrap" estimates with the MLE estimates in this task. We shall also compare the one-step-ahead prediction capability of the models found.

The estimating equations for the "bootstrap" estimator are given in the previous section. The data is processed sequentially. The estimating equations for the MLE method are summarized in Subsection 5.6.3 of the last chapter.

The control excitation was provided by a 511 bit PRBS sequence (unrepeated) of unit amplitude adjusted for zero mean. Zero initial

conditions were assumed.

The machine algorithm for producing random numbers came under suspicion when it was observed that the MLE algorithm was capable of generating a sequence of noise estimates $\{e^*(t)\}$ with variance noticeably less than unity. Chi squared tests were conducted on a number of sequences produced by the computer. It was found that the mode of the chi squared variates for 19 degrees of freedom was in the region of the seventy per cent significance level. It was concluded that the computer algorithm was a satisfactory source of normally distributed random numbers. However, for the relatively short run lengths in subsequent simulations, the random number sequences were adjusted to zero mean and unit variance (without altering their chi squared values) so that parameters determined by the estimation procedures could be compared sensibly with their true values. In the parameter estimation experiment described here, the sample variance was 0.9791 and the sample mean was -0.0571 before shifting and rescaling. A chi squared value of 8.421 for nineteen degrees of freedom was calculated. The significance level is > 98 per cent.

6.4.2 Observations and Discussion

The parameters found by the two estimating methods are shown in Table 6.5. In addition, the MLE method produced an estimate of the constant component, D , in the output record. Standard deviations are shown for all the elements of the parameter vector (excluding λ) estimated by MLE. This is one distinct advantage of the MLE method.

Parameters		A_1	A_2	B_0	B_1	B_2	λ	C_1	C_2	D
True Values		-1.5	0.7	1.0	1.0	1.0	1.0	-1.0	0.2	0.0
Maximum Likelihood	Estimated	<u>-1.5042</u>	<u>0.7030</u>	* <u>0.9970</u>	<u>0.9905</u>	* <u>0.9478</u>	*0.9928	-1.0617	<u>0.2116</u>	<u>-0.0005</u>
	True Error	-0.0042	0.0030	-0.0030	-0.0095	-0.0522	-0.0072	-0.0617	0.0116	-0.0005
	Std. Dev.	0.0076	0.0061	0.0440	0.0638	0.0592	-	0.0434	0.0443	0.0066
"Bootstrap" Estimator	Estimated	* <u>-1.5023</u>	* <u>0.7007</u>	<u>1.0159</u>	* <u>0.9942</u>	<u>0.9314</u>	1.0289	*-0.9888	*0.2082	-
	True Error	-0.0023	0.0007	0.0159	-0.0058	-0.0686	0.0289	0.0112	0.0082	-
	Std. Dev.	0.0242	0.0229	0.0648	0.0693	0.0759	-	-	-	-

* Best estimate of this parameter
 — Within one s.d. of true

Table 6.5 Comparison of the Estimate of the Parameters of a Second Order System by MLE and the "Bootstrap" Method

As expected, the standard deviations obtained for the "bootstrap" method estimates are much higher than their counterparts of the MLE estimates. However, note that the true error is well within one standard deviation.

The autocorrelation terms calculated with the parameters derived by the "bootstrap" method were

$$R_V(0) = 2.400 ; R_V(1) = -1.265 ; R_V(2) = 0.222$$

which led to the estimates:

$$C_0^* = 1.0289 ; C_1^* = -1.0174 ; C_2^* = 0.2142$$

From these, the estimates $\hat{\lambda}$, \hat{C}_1 and \hat{C}_2 shown in Table 6.5 were derived.

Previously, we have proposed that the models identified by employing one of the parameter estimation schemes might be used in a control algorithm that is required to minimize the mean square deviation of the output from a desired target value. An important step in finding the control algorithm is to demonstrate that the model is a minimum variance predictor. We can write (in our previous notation)

$$e^*(t+1) = C^{-1}(z^{-1})A(z^{-1})y(t+1) - C^{-1}(z^{-1})B(z^{-1})u(t+1) \quad (6.4-1)$$

$$\text{Let us define} \quad y(t+1) = zy(t) \quad (6.4-2)$$

Taking the difference between (6.4-2) and (6.4-1) we obtain

$$\begin{aligned}
y(t+1) &= e^*(t+1) + zy(t) - C^{-1}(z^{-1})A(z^{-1})zy(t) + C^{-1}(z^{-1})B(z^{-1})u(t+1) \\
&= \lambda e(t+1) + C^{-1}(z^{-1})B(z^{-1})u(t+1) + C^{-1}(z^{-1})[C(z^{-1})-A(z^{-1})]zy(t)
\end{aligned}
\tag{6.4-3}$$

$$\stackrel{\Delta}{=} e^*(t+1) + y^o(t+1; y(t), \dots, u(t+1), \dots)
\tag{6.4-4}$$

The term $y^o(t+1; y(t), \dots, u(t+1), \dots)$ in equation (6.4-4) depends only on the data $y(t), y(t-1), \dots, u(t), u(t-1), \dots$ and the control exercised at $u(t+1)$. We interpret $y^o(t+1; y(t), \dots, u(t+1), \dots)$ as the minimum mean square prediction of $y(t+1)$ given this data with prediction error $\lambda e(t+1) \stackrel{\Delta}{=} e^*(t+1)$.

For our comparison, we write the following recursive relation corresponding to equation (6.4-1):

$$\begin{aligned}
\hat{e}^*(t+1) &= y(t+1) + \hat{A}_1 y(t) + \hat{A}_2 y(t-1) \\
&\quad - \hat{B}_0 u(t+1) - \hat{B}_1 u(t) - \hat{B}_2 u(t-1) \\
&\quad - \hat{C}_1 \hat{e}^*(t) - \hat{C}_2 \hat{e}^*(t-1)
\end{aligned}
\tag{6.4-5}$$

where the parameters \hat{A}_j , \hat{B}_j , \hat{C}_j are the estimated parameters of the model and the numbers $\hat{e}^*(t)$ are generated recursively. Then, given observation $y(t+1)$, we can compare the prediction error $\hat{e}^*(t+1)$ obtained by equation (6.4-5) with the true disturbance $e^*(t+1)$ generated for the simulation. We can calculate the standard deviation of the prediction error by measuring the variance of the sequence $\{\hat{e}^*(t)\}$. Also, we can calculate the autocorrelations for the series.

True Disturbance	Prediction Error	
	for Maximum Likelihood Estimator Model	for "Bootstrap" Method Model
-1.0845	*-1.0870	-1.0686
1.5137	1.4411	* 1.5566
0.0656	0.0354	* 0.0426
1.0264	* 0.9540	0.9377
-0.5668	*-0.6094	-0.7324
0.8993	* 0.8016	0.6903
0.6115	* 0.6539	0.5139
0.6953	0.8353	* 0.6857
-0.9302	-0.7283	*-0.8570
1.5417	1.6812	* 1.6480
-0.2628	-0.1235	*-0.2892
-1.4769	-1.1343	*-1.4994
0.4130	* 0.4835	0.4847
0.4155	* 0.5311	0.5571
0.7804	0.8470	* 0.8281
-1.1894	*-1.1442	-1.2393
-0.1937	*-0.2593	-0.3147
2.0517	* 1.9620	1.9344
0.2986	0.4089	* 0.2561
-0.6942	-0.5944	*-0.7927
0.6839	0.7761	* 0.7005
-0.1019	* 0.0455	0.0079
-0.9271	-0.8744	*-0.9230
-0.5241	*-0.5975	-0.6144
-1.6078	*-1.6661	-1.1600
0.0859	*-0.0101	0.1768
2.0659	1.9500	* 2.0786
-1.0464	-0.9820	*-1.0008
-1.0601	*-1.0127	-0.9851
0.2660	0.2063	* 0.2644
	15	15

* Closest to true

Table 6.6 Comparison of Prediction Errors

The prediction errors are compared in Figure 6.3. This graph clearly indicates that the predictor error for each model is very close to the true disturbance. The prediction errors determined by the two models are listed in Table 6.6. For comparison purposes, the prediction error closest to the true value is shown with an asterisk. By this measure, the model obtained by the "bootstrap" method compares favourably with the model obtained by the MLE method.

The standard deviation of the prediction error sequence generated by each model over the record length of 511 observations is given in Table 6.7.

	Standard Deviation
True noise sequence	1.00000
Prediction error sequence generated by the MLE method model	0.992796
Prediction error sequence generated by "bootstrap" method model	1.00049

Table 6.7 Comparison of the Standard Deviation of the Prediction Errors

The first nineteen normalized autocorrelation coefficients, $R_e(\tau)$ were found for the true noise sequence $\{e(t)\}$ and the prediction error sequences generated by the two models. The defining equation is

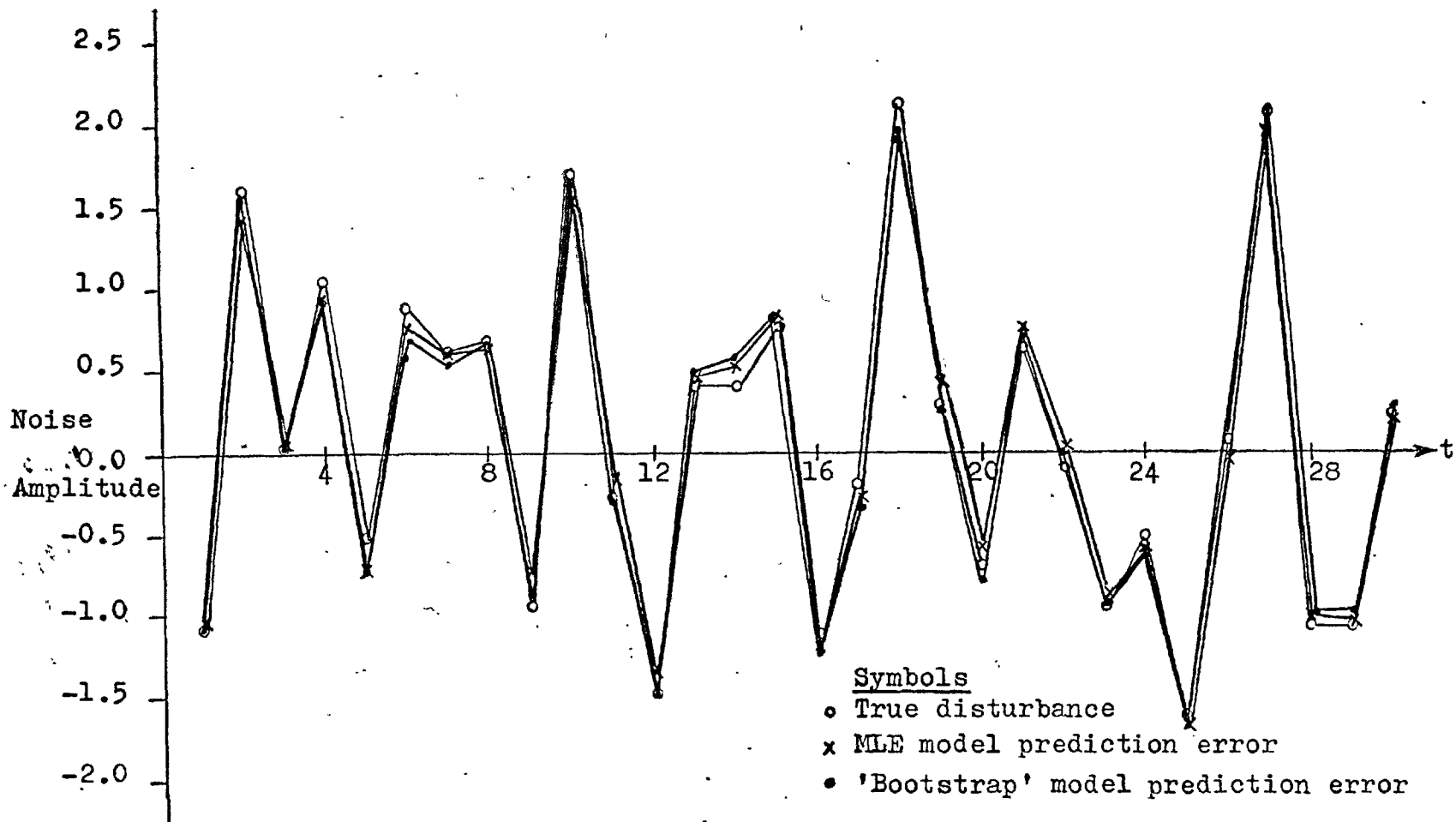


Figure 6.3 Comparison of Prediction Errors

$$R_e(\tau) = \frac{N}{N-\tau} \cdot \frac{\sum_{t=\tau+1}^N e(t)e(t-\tau)}{\sum_{t=1}^N e(t)^2} \quad (6.4-6)$$

where $N = 511$ is the record length. The three sets of autocorrelation coefficients for the true noise and prediction errors are shown plotted against time shift τ in Figure 6.4.

We reject the hypothesis at the five per cent level that the sequences are correlated if not more than one in twenty autocorrelation terms exceed $2\sqrt{511} = 0.0883$. (In this experiment, we should have recorded the autocorrelations for twenty shifts; we show nineteen.) Only one spike occurs at the fourteenth shift. Hence we claim that the input stochastic disturbance is uncorrelated.

Observe that the autocorrelation coefficients for the prediction error sequence generated by the model with parameters determined by MLE are particularly small for the first few lags. The reason is that in the MLE method, the model parameters are adjusted until the variance of the prediction error sequence $\{\hat{e}^*(t)\}$ is minimized. Hence any correlation of the input sequence for short runs (however slight) is removed and incorporated into the model parameters. That is, the model includes the combined second order dynamics of the process and input noise. Thus the prediction error sequence is more truly independent than the input noise sequence! This explains why the standard deviation of the prediction error obtained by the MLE method is less than unity in the last example.

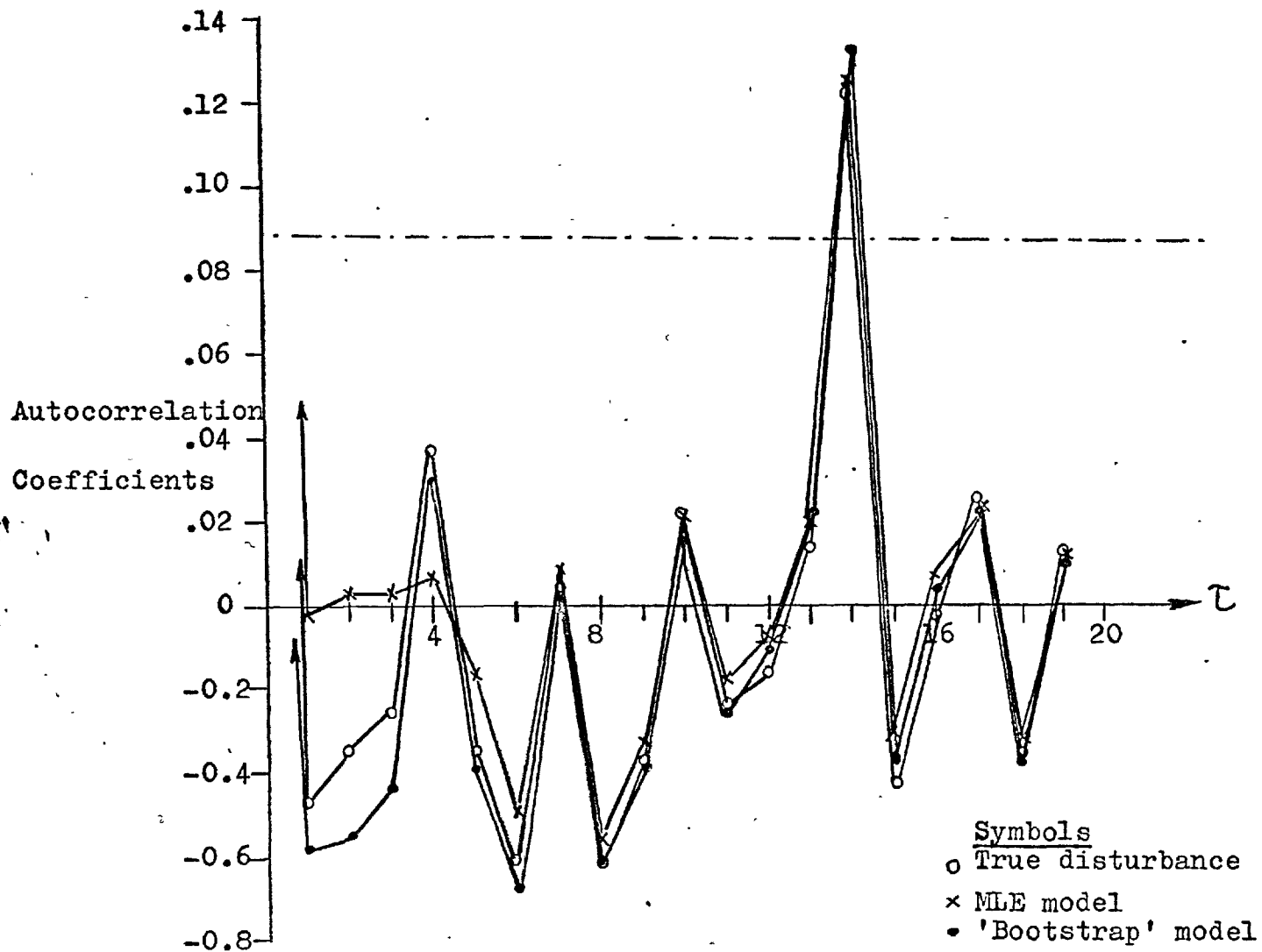


Figure 6.4 Comparison of Autocorrelation Coefficients

The autocorrelation coefficients for the prediction error sequence generated by the other model (with parameters determined by the "bootstrap" method) do not exhibit the effect referred to above. However, on the basis of the evidence presented, we would reject the hypothesis (at the five per cent level) that the prediction error sequence generated by the model with parameters determined by the "bootstrap" method is correlated.

The prediction error sequences were subjected to a Chi Squared test to ascertain the closeness of fit of their distributions to a normal distribution divided into twenty classes. The results of this test are summarized in Table 6.8.

	χ^2	Significance Level (per cent)
True noise sequence	8.421	98
Prediction error sequence generated by MLE method model	5.778	> 99
Prediction error sequence generated by "bootstrap" method model	8.735	97.5

Table 6.8 χ^2 Test (19 degrees of freedom)

The Chi Squared test indicates that the prediction error sequence generated by the MLE method model fits a normal distribution more closely than either the true noise sequence or the prediction error sequence generated by the "bootstrap" method model. We present this

as an observation for which we are not prepared to offer an explanation without further investigation.

The time taken to compile and execute the MLE method, including data processing, was 2.1 minutes. The corresponding time for the estimation of the parameters by the "bootstrap" method, along with all the other functions, was 1.1 minutes.

We conclude that while the MLE method represents the optimal solution to the parameter estimation problem, the performance of the model determined by the "bootstrap" method was only slightly suboptimal by comparison. We have demonstrated that the latter generates a prediction error sequence very near to minimum variance that is statistically uncorrelated. It has accomplished this with a significant reduction in computation time.

6.5 The Estimation of Parameters of Systems with Zeros of the Noise Process Outside the Unit Circle

6.5.1 Statement of the Problem

In Section 5.4, it was shown that for a stochastic process with a finite number of nonzero serial autocorrelation terms (a finite moving average process), we can always find a stable filter to remove the correlation. Two examples were given showing filters designed to "whiten" a correlated noise process described by equations with one or more zeros outside the unit circle.

In this section, we show the results of simulating two systems with the noise processes discussed in Subsection 5.4.5 and referred to above. We shall require the MLE and "bootstrap" estimators to find the parameters of the systems.

6.5.2 System with One Zero Outside the Unit Circle

The equation of the system simulated is:

$$(1 - 1.5z^{-1} + 0.7z^{-2})y(t) = (1 + z^{-1} + z^{-2})u(t) + 1(1 + 1.5z^{-1} - z^{-2})e(t) \quad (6.5-1)$$

The stochastic term $v(t)$ factors as:

$$v(t) = 1(1 + 2z^{-1})(1 - 0.5z^{-1})e(t) \quad (6.5-2)$$

which has roots in the z plane at -2 and 0.5 .

By analogy to example (1) (equations (5.4-37)) we expect to identify the process

$$v(t) = 2(1 + 0.5z^{-1})(1 - 0.5z^{-1})e(t) \quad (6.5-3)$$

The simulation of the process in equation (6.5-1) and the estimation of the process parameters by the MLE method and "bootstrap" method was carried out as in the last section. The results are reported below.

6.5.3 Observations and Discussion

The parameters found by the two estimating methods are shown in Table 6.9. The MLE method finds only the solution with zeros within

Parameters		A_1	A_2	B_0	B_1	B_2	λ	C_1	C_2	D
True Values		-1.5	0.7	1.0	1.0	1.0	1.0	1.5	-1.0	0.0
Stable Values		"	"	"	"	"	2.0	0.0	-0.25	0.0
Maximum Likelihood	Estimated (Stable Values)	* <u>-1.4967</u>	<u>0.7137</u>	<u>0.9460</u>	* <u>1.0721</u>	* <u>1.0659</u>	1.9808	0.0501	-0.3063	-0.0009
	True Error	0.0033	0.0137	-0.0540	0.0721	0.0659	-	-	-	-
	Std. Dev.	0.0261	0.0229	0.0882	0.0952	0.1022	-	0.0521	0.0489	0.0653
"Bootstrap" Estimator	Estimated	<u>-1.4805</u>	* <u>0.6994</u>	<u>0.9689</u>	1.1408	<u>1.0966</u>	1.251	1.055	-0.798	-
	True Error	0.0195	-0.0006	*-0.0311	0.1408	0.0966	0.251	-0.445	0.202	-
	Std. Dev.	0.0274	0.0256	0.0920	0.0956	0.1044	-	-	-	-
	Estimated (Stable Values)	-	-	-	-	-	1.9618	0.1234	-0.3280	-

* Best estimate of this parameter
 - Within one s.d. of true

Table 6.9 Comparison of the Estimates of the Parameters of a System with One Zero of the Noise Process Outside the Unit Circle

the unit circle. The reason is that the value of the number

$$e^*(t) = C^{-1}(z^{-1})[A(z^{-1})y(t) - B(z^{-1})u(t)] \quad (6.5-4)$$

employed in the iterative estimation algorithm grows quickly if the poles of $C(z^{-1})$ are outside the unit circle. We call this the "stable" solution for brevity.

Two sets of parameters are shown for the estimates of λ , C_1 and C_2 for the model obtained by the "bootstrap" estimator because the two stochastic processes have the same autocorrelation coefficients.

The autocorrelation terms calculated for the one-step-ahead prediction error sequence $\{\hat{e}^*(t)\}$ (computed with the "stable" parameters) are listed in Table 6.10. For both models, we claim at a 95 per cent confidence level that the prediction error sequences are uncorrelated.

This simulation has demonstrated that a stable filter can be found by both the MLE method and the "bootstrap" methods to eliminate effectively the correlation of a stochastic process described by an equation with one zero outside the unit circle.

6.5.4 System with Two Complex Zeros Outside the Unit Circle

The equation of the system simulated is:

$$(1 - 1.5z^{-1} + 0.7z^{-2})y(t) = (1 + z^{-1} + z^{-2})u(t) + 1(1 + 2z^{-1} + 2z^{-2})e(t) \quad (6.5-5)$$

The stochastic term $v(t)$ factors to give two roots outside the unit

Autocorrelation Terms for

<u>Shift</u>	<u>MLE Model</u>	<u>"Bootstrap" Model</u>
1	0.0210	-0.0405
2	-0.0150	0.0354
3	-0.0577	-0.0820
4	0.0511	0.0654
5	-0.0052	-0.0297
6	-0.0391	-0.0344
7	-0.0774	-0.0847
8	-0.0469	-0.0402
9	-0.0716	-0.0709
10	0.0107	0.0113
11	0.0525	0.0568
12	-0.0333	-0.0336
13	0.0737	0.0840
14	-0.0097	-0.0146
15	0.0352	0.0386
16	0.0352	0.0306
17	-0.0604	-0.0531
18	-0.0513	-0.0544
19	0.0367	0.0379

Table 6.10 Autocorrelation Terms for Prediction Error Sequences of
Models Found by the MLE and "Bootstrap" Methods

circle at $z = -1 \pm \sqrt{-1}$. As shown by equation (5.4-40) of example (2) in Subsection 5.4.5, we expect to identify the process

$$v(t) = 2(1 + z^{-1} + 0.5z^{-2})e(t) \quad (6.5-6)$$

The simulation of the process in equation (6.5-5) and the estimation of the process parameters were carried out as in the previous subsections.

6.5.5 Observations and Discussion

The parameters found by the MLE method and the "bootstrap" method are shown in Table 6.11. The standard deviations shown for the parameters of the model obtained by the "bootstrap" estimator which are lower than their counterparts for the model parameters obtained by MLE are not recorded in error.

The autocorrelation terms for the prediction error sequences of the models estimated are shown in Table 6.12. We accept the hypothesis that the sequences are uncorrelated at the 95 per cent confidence level.

This simulation has demonstrated that a stable filter can be found by both the MLE method and the "bootstrap" method to effectively eliminate the correlation of a stochastic process described by an equation with a pair of complex zeros outside the unit circle. This example and the last example together demonstrate that the stochastic processes described by equations with real zeros or complex zero pairs outside the unit circle can be compensated by stable filters so that

Parameters		A_1	A_2	B_0	B_1	B_2	λ	C_1	C_2	D
True Values		-1.5	0.7	1.0	1.0	1.0	1.0	2.0	2.0	0.0
Stable Values		"	"	"	"	"	2.0	1.0	0.5	0.0
Maximum Likelihood	Estimated (Stable Values)	-1.4659	<u>0.6735</u>	*0.9899	<u>1.0790</u>	*1.0429	1.9895	0.9696	0.5179	-0.0027
	True Error	0.0341	-0.0265	-0.0101	0.0790	0.0429	-	-	-	-
	Std. Dev.	0.0321	0.0319	0.0883	0.1155	0.0885	-	0.0421	0.0426	0.2178
"Bootstrap" Estimator	Estimated	* <u>-1.4759</u>	* <u>0.6795</u>	0.8353	* <u>1.0199</u>	<u>1.0789</u>	0.860	2.289	2.335	-
	True Error	0.0241	-0.0205	-0.1647	0.0199	0.0789	-0.140	0.287	0.335	-
	Std. Dev.	0.0299	0.0272	0.1301	0.1330	0.1390	-	-	-	-
	Estimated (Stable Values)	-	-	-	-	-	2.017	0.9752	0.4283	-

* Best estimate of this parameter

- Within one s.d. of true

Table 6.11 Comparison of the estimates of the Parameters of a System with Two Complex Zeros of the Noise Process Outside the Unit Circle

Autocorrelation Terms for

<u>Shift</u>	<u>MLE Model</u>	<u>"Bootstrap" Model</u>
1	0.0072	-0.0293
2	-0.0031	0.0813
3	-0.0124	-0.0956
4	-0.0303	0.0061
5	0.0010	-0.0018
6	-0.0155	-0.0425
7	-0.0260	-0.0015
8	-0.0024	-0.0026
9	0.0337	0.0369
10	0.0144	0.0168
11	-0.0920	-0.0877
12	-0.0537	-0.0527
13	0.0162	0.0879
14	-0.0059	-0.0029
15	0.0091	0.0028
16	-0.0525	-0.0473
17	-0.0031	-0.0075
18	-0.0031	-0.0176
19	0.0251	0.0464

Table 6.12 Autocorrelation Terms for Prediction Error Sequences of
Models Found by the MLE and "Bootstrap" Methods

the prediction error sequence is uncorrelated. The consequence of this demonstration is that a physical process that may have non-minimum phase characteristics in its stochastic part will be identified as a system with minimum phase characteristics in its stochastic part.

6.6 The Estimation of the Matrix of Second Derivatives of the Likelihood Function

6.6.1 Purpose and Procedure

In Subsection 5.6.4, it was proposed that a second-order minimization algorithm such as the Fletcher and Powell routine [31] could be used to minimize the negative log likelihood function without requiring the explicit calculation of the matrix of partial second derivatives. Recall that the Newton-Raphson iterative algorithm that uses the matrix of calculated partial second derivatives is ($m = 1$)

$$e(k+1) = e(k) - L_{ee}^{*-1}(e(k))L_e^*(e(k)) \quad (6.6-1)$$

where the elements of $L_{ee}^*(e(k))$ are given by

$$\begin{aligned} \frac{\partial^2}{\partial e_i \partial e_j} L^*(e(k)) &= \frac{1}{\lambda^2} \left[\sum_{t=1}^N \frac{\partial e^*(t)}{\partial e_i} \cdot \frac{\partial e^*(t)}{\partial e_j} \right. \\ &\quad \left. + \sum_{t=1}^N e^*(t) \frac{\partial^2 e^*(t)}{\partial e_i \partial e_j} \right] \end{aligned} \quad (6.6-2)$$

The alternative solution proposed was to employ the iterative algorithm

$$\mathbf{e}(k+1) = \mathbf{e}(k) - \mathbf{H}(k)\mathbf{L}_e^*(\mathbf{e}(k)) \quad (6.6-3)$$

where $\mathbf{H}(k)$ is a positive definite symmetric matrix updated with the most recent gradient information. (See [31] for a complete description in which it is shown that $\mathbf{H}(k)$ tends to the matrix of partial second derivatives evaluated at the minimum.)

One proposed advantage of this alternative method is that it simplifies the programming effort. A possible disadvantage is that it may not converge to the same solution obtained by calculating the matrix of partial second derivatives and inverting it at every iteration.

The purpose of the following experiment is to compare the two algorithms in estimating the parameters of the second order system previously discussed in Section 6.3. Plant data (from zero initial conditions) was generated and stored. The proper MLE method employed was that used in the preceding sections. The alternative method using the Fletcher-Powell (FP) algorithm was initiated with a unit matrix for $\mathbf{H}(0)$.

6.6.2 Discussion

The first experiments with the FP algorithm failed. It was found that the parameters obtained in the first iteration resulted in overflow values for $\mathbf{L}_e^*(\mathbf{e}(1))$ from which it was unable to recover.

Further investigation showed that Davidon's cubic interpolation procedure recommended by Fletcher and Powell, which employs function and gradient values for minimization along a line, would not work in this application. The first step away from the working point in the direction of steepest descent invariably resulted in a choice of parameters such that the zeros of the polynomial $C(z^{-1})$ were outside the unit circle in the z plane.

To constrain the search to points for which the zeros of $C(z^{-1})$ were kept within the unit circle, the following procedure was developed.

- (1) A search point was found along the descent direction chosen by the FP algorithm.
- (2) The estimate $\hat{C}(z^{-1})$ of the polynomial $C(z^{-1})$ was factorized.
- (3) If the absolute value of the distance from the origin to any of the roots found in (2) exceeded 0.975, the step size in the descent direction was halved.
- (4) The procedure was repeated from (1) until the roots were all within the constraint circle of radius 0.975.
- (5) The resulting parameters were used to generate the prediction error sequence $\{\hat{e}^*(t)\}$.

As a result of this modification, no further difficulty was experienced. The function and gradient values all remained finite.

Table 6.13 compares the estimated parameters and the standard deviations of the estimates for the two alternative MLE methods. The

Parameters		A_1	A_2	B_0	B_1	B_2	λ	C_1	C_2	D
True Values		-1.5	0.7	1.0	1.0	1.0	1.0	-1.0	0.2	0.0
MLE (Calculated covariance)	Estimated	-1.4917	0.6932	-0.8558	1.0849	0.4618	1.0705	-0.9668	0.1291	0.0015
	Std. Dev.	0.0144	0.0121	0.0443	0.0615	0.0582	-	0.0425	0.0418	0.0072
MLE (Estimated covariance)	Estimated	-1.4917	0.6932	-0.8432	1.0833	0.4627	1.0705	-0.9665	0.1286	0.0015
	Std. Dev.	0.0146	0.0125	0.0346	0.0464	0.0516	-	0.0432	0.0414	0.0067

Table 6.13 Comparison of the Estimates of the Parameters of a Second Order System by Alternate

MLE Methods

two sets of estimated parameters are quite close to each other. The observation that two of the standard deviations derived from the estimated covariance matrix are twenty per cent low is cause for concern because it suggests that the associated parameters are estimated more accurately than is justified in the circumstances.

More computation was involved in using the FP algorithm than the direct method. This is shown in the following table.

	MLE (Newton-Raphson)	MLE (Fletcher-Powell)
Number of Iterations	10	11
Total Time (minutes)	2.1	3.8

Table 6.14 Comparison of Computation Effort

Possibly the main reason for increased computation time is that unstable roots were encountered in the minimization-along-a-line section of the FP algorithm 51 times during the first two iterations.

Since carrying out the above comparison, D. W. Norris has modified the Fletcher-Powell algorithm by replacing Davidon's minimization-along-a-line procedure by cubic interpolation with a quadratic minimization method. The important advantage to the above problem of this modification is that the steps in the descent direction are initially smaller than in the original procedure.

6.7 Testing the Order of the Model

6.7.1 Purpose and Procedure

The consequences of changing the order of the model in an identification problem is well known. If the order is assumed too low, it is impossible to remove the correlation from the prediction error sequence $\{e^*(t)\}$. On the other hand, Astrom and Bohlin [18] note that the matrix of partial second derivatives, L_{ee}^* , may become singular on choosing a model of too high order.

In this section, we show the results of having investigated changing the order of the model fitted to data generated by a second order system. The parameters are fitted by the MLE method. However, Mayne's pseudo-inverse routine [64] is used to invert the matrix L_{ee}^* for reasons that will become clear.

The second order model used for simulation is the same as that employed in the last and other sections. Zero initial conditions were assumed. The record length was 511 observations as in previous trials.

6.7.2 Observations and Discussion

A summary of the parameters found and their standard deviations for models of different order is shown in Table 6.15. The autocorrelation terms for the prediction error sequence $\{e^*(t)\}$ of each model was computed for 19 shifts as in previous sections. The number of terms $R_e(\tau)$ that exceed $2/\sqrt{511}$ is displayed in the table.

Order	p = 1		p = 2		p = 3	
Param.	Est.	s.d.	Est.	s.d.	Est.	s.d.
A ₁	-0.867	0.021	-1.494	0.008	-0.593	0.140
A ₂	-	-	0.695	0.006	-0.649	0.211
A ₃	-	-	-	-	0.625	0.099
B ₀	0.973	0.114	0.995	0.044	0.993	0.044
B ₁	0.832	0.118	1.017	0.063	1.918	0.143
B ₂	-	-	1.045	0.060	1.954	0.163
B ₃	-	-	-	-	0.956	0.139
λ	2.563	-	0.996	-	0.995	-
C ₁	0.554	0.036	-1.035	0.043	-0.128	0.144
C ₂	-	-	0.256	0.043	-0.694	0.141
C ₃	-	-	-	-	0.247	0.051
Number of Large* Terms R _e (τ)	9		0		0	
Number of Iterations	6		10		15	

* Exceeding $2/\sqrt{511}$

Table 6.15 Testing the Order of a Model

We reject the model of order $p = 1$ because the magnitude of the first nine autocorrelation terms indicates that the prediction error sequence for this model is strongly correlated. In addition, the estimate $\hat{\lambda}$ is more than two times larger than for the other models.

We reject the model of order $p = 3$ because the estimate $\hat{\lambda}$ is not significantly lower than for the model of order $p = 2$ and because the ratio $\frac{\text{standard deviation of parameter estimate}}{\text{parameter estimate}}$ is much higher for the third-order model than for the second-order model. That is, the estimates for the parameters of the third-order model do not merit the confidence that we have in the estimates of the parameters of the second-order model.

There are two advantages in using the pseudo-inverse routine referred to above to invert the matrix L_{ee}^* of partial second derivatives:

(1) When the matrix L_{ee}^* is of full rank, its pseudo-inverse and true inverse are identical.

(2) The evidence from experiments in fitting models of up to order $p = 4$ to a second-order system indicate that the iteration procedure is able to converge even when the matrix L_{ee}^* is singular (or tends towards singularity). This is useful because comparison of the models is then facilitated. (The possible alternative is a set of meaningless parameter estimates.)

The prediction error sequence $\{\hat{e}^*(t)\}$ estimated by the model of order $p = 3$ was observed to fit the true sequence $\{e^*(t)\}$ quite well. In fact, the sequence for $p = 4$ was a fair fit.

6.8 The Estimation of Parameters of a Multivariate System

6.8.1 Statement of the Problem

In this section, the MLE and "bootstrap" estimation methods are used to find the parameters of a fourth order system with two inputs and two outputs. The following equations of the system considered are written in the A-canonical form developed in Chapters 2 and 4.

$$x(t+1) = F_A x(t) + G_A u(t) + \Gamma_A e(t) \quad (6.8-1)$$

$$y(t) = H_A x(t) + D_A u(t) + \sqrt{1/2} e(t) \quad (6.8-2)$$

where $x(t)$ is the system state vector ($n = 4$), u is the control vector ($r = 2$) and $e(t)$ is an independent random variable from $n(0, I_2)$.

The system parameters were selected to be representative of system situations that might be encountered in practice. The derivation began with the cross-coupling of two second-order continuous-time systems with poles on the real axis (of the s -plane) at -0.5 , -0.1 , -2.0 and -0.4 . Measurement noise was introduced and feedback loops were added from the noise perturbed outputs, moving the system poles to $\lambda_1, \lambda_2 = -0.3 \pm 0.66j$ and $\lambda_3, \lambda_4 = -1.2 \pm 1.4j$.

For simulation purposes, the continuous-time feedback model was made into a sampled-data model and transformed into the A-canonical form shown in equation (6.8-3).

$$F_A = \begin{bmatrix} 0.0 & 0.0 & 1.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 1.0 \\ -0.6249 & -0.0093 & 1.36240 & 0.0221 \\ 0.3107 & -0.1406 & -0.28132 & 0.4587 \end{bmatrix}$$

$$G_A = \begin{bmatrix} 3.0580 & 6.2429 \\ -1.4081 & -6.4913 \\ 1.1748 & 5.5693 \\ -0.3864 & -1.9434 \end{bmatrix}; \quad \Gamma_A = \begin{bmatrix} 0.1021 & 0.6551 \\ -0.6553 & -0.4926 \\ -0.1164 & 0.4678 \\ -0.0043 & -0.2228 \end{bmatrix}$$

$$H_A = \begin{bmatrix} 1.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 1.0 & 0.0 & 0.0 \end{bmatrix}$$

$$D_A = \begin{bmatrix} 10.0 & 1.0 \\ 1.0 & 10.0 \end{bmatrix}; \quad \sqrt{\Lambda}^{1/2} = \begin{bmatrix} 1.0 & 0.6 \\ 0.6 & 1.0 \end{bmatrix} \quad (6.8-3)$$

The equivalent vector difference equation of the above system is obtained by carrying out the transformation procedures outlined in Chapters 3 and 4. The vector difference equation is

$$A(z^{-1})y(t) = B(z^{-1})u(t) + v(t) \quad (6.8-4)$$

where, for later convenience, we may write $v(t)$ in either form below:

$$v(t) = C(z^{-1})\Lambda^{1/2}e(t) \quad (6.8-5)$$

$$= C'(z^{-1})e(t) \quad (6.8-6)$$

For this problem, $p = 2$ so that

$$A(z^{-1}) = I_2 + z^{-1}A_1 + z^{-2}A_2 \quad (6.8-7)$$

$$B(z^{-1}) = B_0 + z^{-1}B_1 + z^{-2}B_2 \quad (6.8-8)$$

$$C(z^{-1})_{\Delta}^{1/2} = (I_2 + z^{-1}C_1 + z^{-2}C_2)_{\Delta}^{1/2} \quad (6.8-9)$$

$$C'(z^{-1}) = C'_0 + z^{-1}C'_1 + z^{-2}C'_2 \quad (6.8-10)$$

The problem is one of estimating the thirty-one parameters of the above model ($C'_0 \stackrel{\Delta}{=} \Delta^{1/2}$ is symmetric). We assume that the order is known ($p = 2$), the experiment is commenced from zero initial conditions and there is no delay in the control terms.

6.8.2 Estimation Procedure

For comparison purposes, a series of experiments were carried out in the following order:

(1) Operating data was simulated by running the state model for a total of 889 time intervals (from zero initial conditions). Control excitation for the simulation was provided by 127 bit PRBS sequences adjusted for zero mean and repeated seven times. Feedback was taken from the fourth register of the PRBS generator for one input and from the third register for the other. Thus the chain code produced for one input in forward time was the same as the code produced for the other input in reverse time. Preferred methods for which the cross-correlation between channels is reduced are discussed in Chapter 7.

(2) The "bootstrap" estimator was used to estimate the twenty parameters of $A(z^{-1})$ and $B(z^{-1})$. The equations for the sequential estimator in the multivariate case are given by (5.5-17) and (5.5-18) in Subsection 5.5.4.

The serial autocorrelations $\hat{R}_v(0)$, $\hat{R}_v(1)$ and $\hat{R}_v(2)$ were obtained by the methods described beginning equation (5.5-38) in Subsection 5.5.4. From these, the matrix coefficients \hat{C}'_0 , \hat{C}'_1 and \hat{C}'_2 were obtained by the method of moments set out in Subsection 4.3.3 exactly as for the example shown beginning with equation (4.3-46). These led to estimates of C_1 , C_2 and \dots .

(3) The MLE algorithm was provided with the estimate obtained in (2) as the initial starting value. The purpose of this was to obtain the value of the likelihood function for the parameters obtained by the "bootstrap" method. Then the MLE method was required to improve the estimate by minimizing the likelihood function. The Fletcher-Powell algorithm was employed in this task.

Two problems were encountered in carrying out this plan. It was found that the record length of 889 intervals was insufficient to provide estimates by the "bootstrap" method that could be used in the simple control scheme to be described later. The controller was found to be unstable. To solve this problem, it was sufficient to prime the first "bootstrap" prediction with an estimate obtained from a previous run. However, the covariance matrix was not altered at the same time; thus the variance of the final estimate was conditional on the 889 samples only.

The other problem encountered was the effect of numerical round-off. We are warned (IBM 360 System Scientific Subroutine Package Manual) that multiple regression in more than five variables can lead to erroneous results if single precision arithmetic is used. The problem here can be formulated as that of finding the parameters of the regression model in ten variables

$$y(t) = \Theta \xi(t) + v(t) \quad (6.8-11)$$

where

$$\xi(t) = \begin{bmatrix} y(t-1) \\ y(t-2) \\ u(t) \\ u(t-1) \\ u(t-2) \end{bmatrix} \quad (6.8-12)$$

The parameters found to be the most sensitive to round-off were the autocorrelation coefficients $R_v(\tau)$ which were affected in the third significant digit when computed by single precision arithmetic.

6.8.3 Observations and Discussion

A comparison of the parameters obtained by the MLE and "bootstrap" estimators is given by Table 6.16 along with the true values. For instance, the biased estimates obtained by conventional least squares have also been displayed to indicate the necessity for estimating by other means. The significant features in the table are summarized as follows:

Table 6.16 Comparison of Parameter Estimates for a Multivariate System


Parameter Matrix	Estimator	Estimated Parameters		Standard Deviation	
A_1	True	-1.3624 0.2813	-0.0221 -0.4587		
	LSE	-1.0253 0.5135	0.2994 -0.1548		
	MLE	<u>-1.3272</u> <u>0.2776</u>	<u>0.0302</u> <u>-0.4600</u>	0.0465 0.0542	0.0534 0.0532
	"Bootstrap"	<u>-1.3501</u> <u>0.2837</u>	<u>0.0107</u> <u>-0.4384</u>	0.0264 0.0230	0.0378 0.0318
A_2	True	0.6249 -0.3107	0.0093 0.1406		
	LSE	0.4007 -0.4840	-0.0032 0.0874		
	MLE	<u>0.5935</u> <u>-0.3110</u>	<u>-0.0040</u> <u>0.1398</u>	0.0309 0.0351	0.0476 0.0312
	"Bootstrap"	<u>0.6081</u> <u>0.3217</u>	<u>0.0015</u> <u>0.1272</u>	0.0190 0.0159	0.0227 0.0190
B_0	True	10.0000 1.0000	1.0000 10.0000		
	LSE	9.9449 1.0175	1.0559 10.0131		
	MLE	<u>10.0244</u> <u>1.0308</u>	<u>0.9890</u> <u>9.9793</u>	0.5250 0.5130	0.0720 0.0480
	"Bootstrap"	<u>9.9995</u> <u>1.0427</u>	<u>1.0274</u> <u>10.0073</u>	0.0498 0.0417	0.0489 0.0412

— within one s.d. of true

Table 6.16 (continued)

Parameter Matrix	Estimator	Estimated Parameters	Standard Deviation	
B_1	True	-10.5881 4.6591 0.9465 -10.7966		
	LSE	-6.8108 8.1297 3.6034 -7.6182		
	MLE	<u>-10.1332</u> <u>5.1077</u> <u>0.9317</u> <u>-10.9081</u>	0.528 0.528 0.556 0.533	
	"Bootstrap"	<u>-10.4054</u> <u>4.9315</u> <u>0.9949</u> <u>-10.6670</u>	0.298 0.389 0.248 0.326	
B_2	True	3.2977 -2.0743 -1.8468 3.8850		
	LSE	1.6056 -2.4159 -3.3526 2.6867		
	MLE	<u>2.9409</u> <u>-2.3185</u> <u>-1.8614</u> <u>3.9157</u>	0.532 0.767 0.476 0.500	
	"Bootstrap"	<u>3.0668</u> <u>-2.3127</u> <u>-2.0167</u> <u>3.6767</u>	0.194 0.447 0.163 0.375	
C_1	True	-1.8171 0.9058 -0.2808 -0.6140		
	MLE	<u>-1.9171</u> <u>1.1425</u> <u>-0.4196</u> <u>-0.4109</u>	0.340 0.356 0.298 0.342	
	"Bootstrap"	-1.8207 0.9880 -0.2238 -0.5588		

Table 6.16 (Continued)

Parameter Matrix	Estimator	Estimated Parameters		Standard Deviation	
C_2	True	0.6364	-0.4115		
		0.0213	0.1288		
	MLE	<u>0.5996</u>	<u>-0.4367</u>	0.355	0.394
		<u>0.0877</u>	<u>0.0395</u>	0.320	0.371
	"Bootstrap"	0.7044	-0.5418		
		0.0866	0.0266		
	True	1.3600	1.2000		
		1.2000	1.3600		
	MLE	1.3585	1.1950		
		1.1950	1.3533		
	"Bootstrap"	1.3611	1.2010		
		1.2010	1.3659		
	"Bootstrap" (see text)	2.1226	1.3668		
		1.3668	1.4000		

(1) The estimates obtained by conventional least squares are strongly biased.

(2) All the MLE estimates are within one standard deviation of the true parameter.

(3) 17 out of 20 "bootstrap" estimates for parameters of $A(z^{-1})$ and $B(z^{-1})$ are within one standard deviation of the true parameter. All are within two standard deviations.

(4) The standard deviations calculated for the "bootstrap" method are smaller than for the MLE method. There is strong evidence that this is a limitation imposed by the method employed for determining the matrix of partial second derivatives (that is, by using the Fletcher-Powell algorithm in place of direct calculation of the matrix). There are two reasons for stating this:

(a) The evidence of the simple experiment carried out in Section 6.6 indicated that, in the main, the standard deviations calculated by the Fletcher-Powell algorithm were larger than those calculated directly.

(b) The Fletcher-Powell routine terminated after 24 iterations and 226 function evaluations when it found that the function values were being affected by round-off errors. At least 28 iterations are required to satisfy a quadratic search. In view of the low likelihood function value achieved, it is doubtful that further searching would have had significant effect on the parameters. A careful study of the search points during convergence showed that in the last six iterations, the dominant terms moved about .001 per cent while the smaller

terms moved about .1 per cent; the function value fell from 2127.0577 to 2126.9687 but the elements on the diagonal of the covariance matrix were reduced by factors in the order of 50 to 100 per cent! Hence, it is concluded that single precision computation has limited the capability of the algorithm.

(5) The covariance matrix $\hat{\Sigma}$ was calculated from the prediction error sequence $\{\hat{e}^*(t)\}$ using the parameters calculated by the "bootstrap" method. The result is the last entry in Table 6.16. It indicates that the estimate $\hat{\Sigma}$ obtained by the method of moments is optimistic; that is, the true variance of the prediction error exceeds that calculated by fitting moments.

(6) The computing time expended to provide the data for Table 6.16 was divided as follows:

(a) Compilation and execution of the program to find - (i) the parameter estimates of $A(z^{-1})$ and $B(z^{-1})$ by the LSE method (ii) the parameter estimates of $A(z^{-1})$ and $B(z^{-1})$ by the "bootstrap" method (iii) the parameter estimates of $C(z^{-1})$ by the method of moments - took 2.3 minutes. The time to find the parameters by the MLE method, starting with the solution obtained by the "bootstrap" method took 20.6 minutes.

The (negative log) likelihood function $L^*(\theta)$ was evaluated for the true parameters and for the parameters estimated by the two methods. These are shown in Table 6.17 below.

	$L^*(\theta)$
True Parameters	2126.1
Parameters by MLE Estimator	2126.9687
Parameters by "Bootstrap" Estimator	2543.9190

Table 6.17 Comparison of Likelihood Values for MLE and "Bootstrap" Estimators

From this comparison, it is clear that the MLE estimates are close to the optimum values to minimize $L^*(\theta)$. However, this improvement over the "bootstrap" estimates required significant computing effort.

It is an easy task to transform the vector difference equation models in the input and output variables into the A-canonical form as state-variable models. The equations of the system for simulation were given in A-canonical form at the beginning of this section. The method of transformation was discussed in the stochastic case in Section 4.4. In the following equation (6.8-13), we show the state-variable model transformed from the model identified by MLE. This is followed by the transformed model identified by the "bootstrap" method. The observation matrix H_A is fixed.

$$\hat{F}_A = \begin{bmatrix} 0.0 & 0.0 & 1.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 1.0 \\ -0.5935 & 0.0039 & 1.3272 & -0.0302 \\ 0.3110 & -0.1398 & -0.2776 & 0.4601 \end{bmatrix}$$

$$\hat{G}_A = \begin{bmatrix} 3.1403 & 6.1192 \\ -1.3770 & -6.5917 \\ 1.2052 & 5.4547 \\ -0.3929 & -1.9028 \end{bmatrix}; \quad \hat{T}_A = \begin{bmatrix} 0.0776 & 0.7584 \\ -0.6677 & -0.3692 \\ -0.0666 & 0.4631 \\ 0.0616 & -0.2407 \end{bmatrix}$$

$$\hat{D}_A = \begin{bmatrix} 10.0244 & 0.9890 \\ 1.0308 & 9.9793 \end{bmatrix} \quad (6.8-13)$$

$$\hat{F}_A = \begin{bmatrix} 0.0 & 0.0 & 1.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 1.0 \\ -0.6081 & 0.0015 & 1.3501 & -0.0107 \\ 0.3217 & -0.1272 & -0.2837 & 0.4384 \end{bmatrix}$$

$$\hat{G}_A = \begin{bmatrix} 3.0834 & 6.2120 \\ -1.3847 & -6.5714 \\ 1.1656 & 5.5341 \\ -0.4144 & -1.9085 \end{bmatrix}; \quad \hat{T}_A = \begin{bmatrix} 0.1145 & 0.6984 \\ -0.5802 & -0.4250 \\ -0.0666 & 0.4631 \\ 0.0616 & -0.2410 \end{bmatrix}$$

$$\hat{D}_A = \begin{bmatrix} 9.9952 & 1.0274 \\ 1.0427 & 10.0073 \end{bmatrix} \quad (6.8-14)$$

Note that in transforming the models from the difference equation description to the state-space description, the product of various parameters is found a number of times in the procedure. Because all the parameters are correlated with each other to some extent, (that is, the covariance elements are nonzero) these operations must lead to biased parameters in the state model. It would be interesting to simulate the transformed model along with the difference equation model to ascertain the extent of the bias.

To conclude this section, we note that:

- (1) The MLE method has proven to be a powerful tool for the estimation of the thirty-one parameters of a multivariate system. However, the method has been demonstrated to be expensive in computing time.
- (2) The "bootstrap" estimator has provided a quick but suboptimal estimate of the parameters of a system when compared with the MLE method. However, the solution may very well be adequate for the estimation of parameters in real situations.

6.9 Regulation of a Multivariate Stochastic System

6.9.1 Minimum Variance Control of the Output

The multivariate form of the identity that leads to the minimum-variance predictor equation given in Section 6.4 is also

$$\begin{aligned}
 y(t+1) &= e^*(t+1) + C^{-1}(z^{-1})B(z^{-1})u(t+1) \\
 &\quad + C^{-1}(z^{-1})[C(z^{-1}) - A(z^{-1})]zy(t) \\
 &= e^*(t+1) + y^o(t+1; y(t), \dots, u(t+1), \dots) \quad (6.9-1)
 \end{aligned}$$

where $y^0(t+1); y(t), \dots, u(t+1), \dots$) is the minimum mean square prediction of $y(t+1)$ given the observations $y(t), y(t-1), \dots, u(t), u(t-1), \dots$ and the control exercised at $u(t+1)$. The variance of the prediction error is $\hat{\sigma}_e^2$ if $A(z^{-1}), B(z^{-1})$ and $C(z^{-1})$ are known perfectly. In the special case of $B_0 = 0$ and $B_1 \neq 0$, the prediction equation $y^0(t+1; \dots)$ is conditional only on observations and controls up to time t . We observe (in the manner of Astrom et al. for the scalar observation case [18]) that

$$E[y(t+1)y^T(t+1)] = \hat{\sigma}_e^2 \quad (6.9-2)$$

for the minimum variance controlling equation

$$\begin{aligned} B(z^{-1})u(t+1) &= (A_1(z^{-1}) - C_1(z^{-1}))y(t) \\ &+ (A_2(z^{-1}) - C_2(z^{-1}))y(t-1) \\ &\quad \cdot \\ &\quad \cdot \\ &+ (A_p(z^{-1}) - C_p(z^{-1}))y(t-p+1) \end{aligned} \quad (6.9-3)$$

The above equation for the controller is the equation of a dynamical system of the form

$$B(z^{-1})u(t+1) = \hat{\mathcal{A}}(z^{-1})y(t) \quad (6.9-4)$$

which is physically realizable only if equation (6.9-4) is stable. It is not stable if zeros of $B(z^{-1})$ are outside the unit circle (corresponding to a nonminimum phase system). Hence the minimum variance controller derived above is suited only to a subclass of minimum phase systems.

6.9.2 An Example

Let us consider the problem of regulating the system discussed in the last section. The theoretical controller equation is

$$B_0 u(t+1) + B_1 u(t) + B_2 u(t-1) = (A_1 - C_1)y(t) + (A_2 - C_2)y(t-1) \quad (6.9-5)$$

which has to be solved for $u(t+1)$. In this example, it is particularly convenient that B_0 is nonsingular. Then we can check the stability of equation (6.9-5) with the estimates \hat{B}_0 , \hat{B}_1 and \hat{B}_2 by determining the eigenvalues of the matrix

$$F_B^{\hat{}} = \begin{bmatrix} 0 & I_2 \\ \hat{B}_0^{-1}\hat{B}_2 & \hat{B}_0^{-1}\hat{B}_1 \end{bmatrix} \quad (6.9-6)$$

This method of checking the stability of vector difference equations was discussed in Subsection 5.4.5.

For the parameters estimated by the "bootstrap" method, it was found that the eigenvalues of $F_B^{\hat{}}$ were within the unit circle. Thus the controller appeared practical.

The system equations were simulated on the computer excited by noise with the same variance as that used when estimating the parameters of the system. However, no attempt was made to make the noise sequence zero mean. The controller equation obtained from (6.9-5)

$$u(t+1) = -\hat{B}_0^{-1} [\hat{B}_1 u(t) + \hat{B}_2 u(t-1) + (C_1 - A_1)y(t) + (C_2 - A_2)y(t-1)] \quad (6.9-7)$$

was operated with the true parameters. It was demonstrated that

the deviation of both outputs at every time instance was the same (to more than three significant figures) as the theoretical minimum prediction and control error. That is, at every t

$$y(t) = e^*(t) \quad (6.9-8)$$

Then the controller was operated with the parameters obtained by the "bootstrap" estimator and tabulated in the last section. In Figure 6.5, one of the two outputs of the system regulated by the controller with estimated parameters is compared with the theoretical minimum control error. It can be seen that the performance is remarkably good. However, after operation in the order of one hundred time intervals, it was observed that the tracking performance had deteriorated. The following analysis explains this observation.

For simplicity, let us assume that $B(z^{-1})$ has an inverse. The system equation is

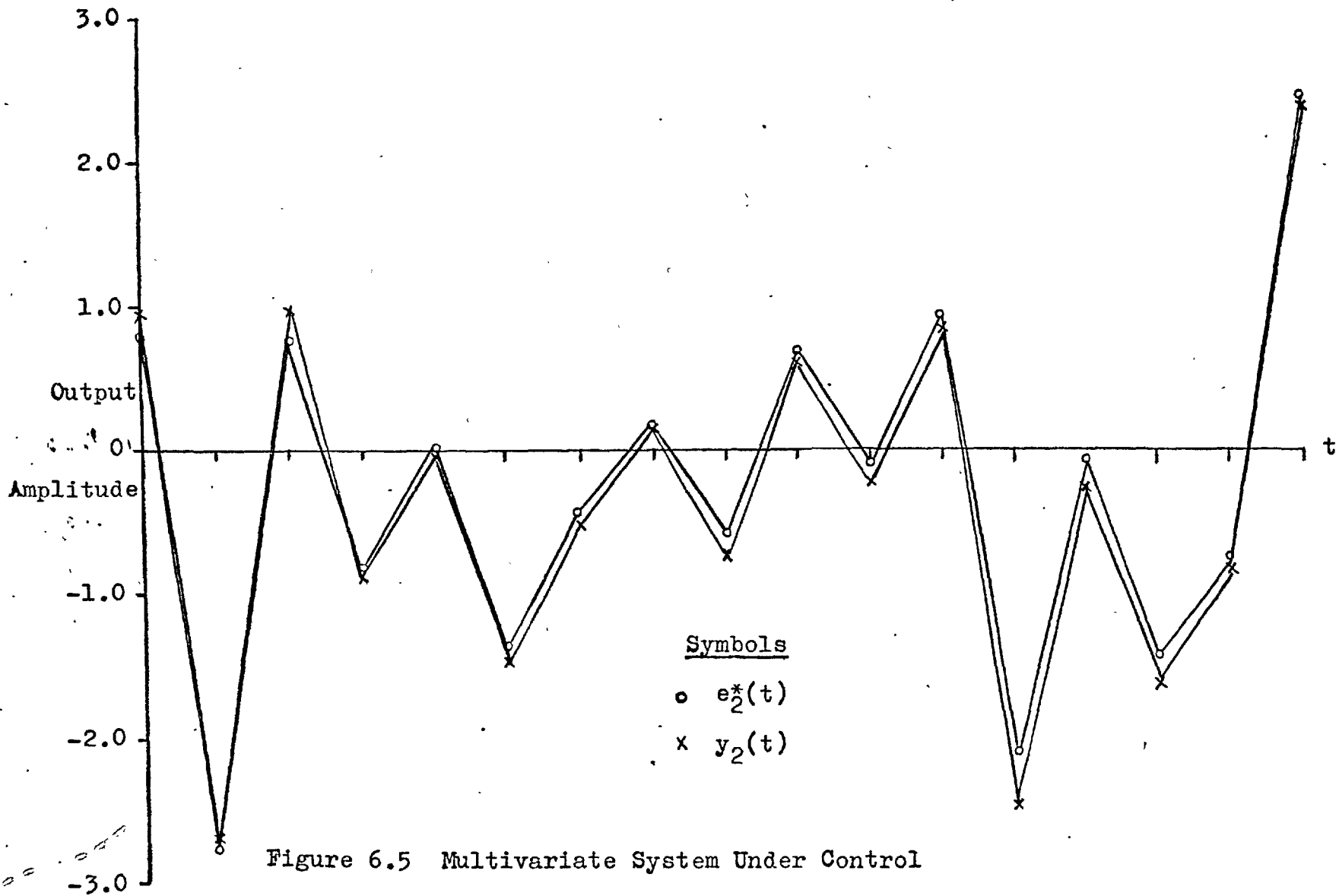
$$A(z^{-1})y(t) = B(z^{-1})u(t) + C(z^{-1})e^*(t) \quad (6.9-9)$$

The control equation

$$u(t) = \hat{B}^{-1}(z^{-1})[\hat{A}(z^{-1}) - \hat{C}(z^{-1})]y(t) \quad (6.9-10)$$

is a function of $y(t-1), y(t-2) \dots$ but not of $y(t)$ because I_m is the first term of both polynomials $A(z^{-1})$ and $C(z^{-1})$. Hence we may combine (6.9-10) with (6.9-9) and obtain

$$[A(z^{-1}) - B(z^{-1})\hat{B}^{-1}(z^{-1})[\hat{A}(z^{-1}) - \hat{C}(z^{-1})]]y(t) = C(z^{-1})e^*(t) \quad (6.9-11)$$



Clearly, if the estimates are exact, equation (6.9-11) reduces to the desired relation

$$y(t) = e^*(t) \quad (6.9-12)$$

Otherwise, equation (6.9-11) is the equation of a dynamical system. There is no guarantee that equation (6.9-11) is even stable! Analysis of this equation is difficult, because both the system equation and the controller equation are fourth order.

It was observed that over several runs, the system behaved like a random walk, the direction of drift being unpredictable.

To test the hypothesis that the closed loop stochastic system was operating as if it had a pole near the origin (in the s -plane), the impulse response of the homogeneous part of equation (6.9-11) was found by two simulation experiments. The factor $C(z^{-1})e^*(t)$ on the right was replaced by the term $\mu(1) = \begin{bmatrix} 100 \\ 0 \end{bmatrix}$ for the first experiment and $\mu(1) = \begin{bmatrix} 0 \\ 100 \end{bmatrix}$ for the second. The first experiment dominated the results. It showed that the response of each output could be fitted approximately by the response to a pole at about -0.1 in the s -plane.

The engineering solution proposed for this problem was to restore the nominal operating point of the output to zero by integral-type feedback. A moving average over the past 30 samples, fed back from each output to the input, restored the tracking performance.

The conclusion is that in a real control situation where the

control algorithm used in the above incorporates estimates of the parameters of the system, the controller has to be supplemented by a reset actuator if it is required to regulate about a set point. This is well known in practice.

6.9.3 State Control

In Section 6.8, it was demonstrated that state-variable models in the A-canonical form can be constructed from the difference equation models derived by the estimating schemes discussed earlier. The state-variable system can be written with $e^*(t)$ as follows:

$$x(t+1) = F_A x(t) + G_A u(t) + I_A^* e^*(t) \quad (6.9-13)$$

$$y(t) = H_A x(t) + D_A u(t) + e^*(t) \quad (6.9-14)$$

But an estimate of the prediction error $e^*(t)$ can be obtained at every observation instant. Therefore an estimate of the next state can be made after every observation. The procedure is simply:

- (1) Assume $x(t)$ is known.
- (2) $e^*(t) = y(t) - H_A x(t) - D_A u(t)$
- (3) Calculate $x(t+1)$ by equation (6.9-13) and repeat.

The expectation of the state of this model is related by a linear transform to the expectation of the state of the system modelled.

The state-variable model here does not supply any more information than can be obtained from the vector difference equation. However,

it may be more convenient to design a controller with this model than with the vector difference equation model.

For example, a controller based on the solution of the steady state Ricatti-equation can be designed to minimize some integral criterion based on costing both control and deviation of the state from some target value. When used as a regulator, such a controller should also cost the integral of the deviation of the state from the set point to avoid the "drift" situation encountered in the last example.

The important point here is that by employing either the state-variable model or the vector difference equation model, the parameters of which have been estimated by fitting the model to input and output data, the minimum variance prediction error $e^*(t)$ can be obtained. The design of a controller using this information adds new problems of stability, etc. which extend beyond the scope of this thesis.

CHAPTER 7

DISCUSSION AND EXTENSIONS

7.1 Introduction

In this final chapter, we shall briefly summarize the contributions of this thesis and indicate areas of engineering application. We shall also point out some of the main difficulties in identifying systems and briefly indicate how to counter them. Finally, we shall suggest directions of extended research.

7.2 Canonical Forms of Deterministic Systems

In this thesis, the A-canonical form has been developed to add to the list of canonical forms of a state-space description. Each has its use; they are complementary to each other:

- (1) Jordan form [25]:- displays system eigenvalues (real and complex, including repeated roots) along the diagonal of the system matrix F .
- useful for stability analysis
- (2) Companion form [4]:- displays the characteristic polynomial of the system matrix F in the bottom row of F .
- useful for system simulation and analysis.
- (3) Multivariate forms due to Luenberger [10] or Luenberger and Anderson [11], (also Section 2.5.1):
- system decomposed into a series of coupled subsystems, each in

companion form

- useful for designing feedback configurations[11].

- (4) A-canonical form:- provides a direct means of transforming a state-space description $\Sigma(F, G, H)$ into a scalar or vector differential (or difference) equation description $\Sigma_D(A, B)$
- provides a direct means of synthesizing a state-space description $\Sigma(F, G, H)$ from a description Σ_D .

It is easy to lose sight of the fact that the A-canonical form and its associated transforms are not limited to multivariate system analysis. We recall that for the case of single outputs, each matrix coefficient is simply a scalar.

7.3 Stochastic Models

7.3.1 Generalizing the Models

Two new, multivariate stochastic models have been derived in this thesis; the difference equation model (the Σ_{DS} description)

$$A(z^{-1})y(t) = B(z^{-1})u(t) + C(z^{-1})\frac{1}{\sqrt{2}}e(t) \quad (7.3-1)$$

and its counterpart, the Σ_{SA} description which was given by equations (4.4-3) and (4.4-4). The important feature of these models is the combination of the noise sources into one vector process. That is, the noise process $v_1(t)$ that excited the system states and the measurement noise $v_2(t)$ were combined into one stochastic process in Chapter 4.

To cope with transport delays and nonzero mean processes that are to be anticipated in practice, the difference equation model can be made more general by the following additions:

$$A(z^{-1})y(t) = z^{-k}B(z^{-1})u(t) + C(z^{-1})\Lambda^{1/2}e(t) + \mu_0 \quad (7.3-2)$$

where z^{-k} represents a delay as a multiple of the sampling period because of transport lag. The m -vector μ_0 is an offset term. The addition of these factors does not affect any of the results reported earlier; it simply makes the model more general for use in real situations. Similar extensions can be added to the state model.

7.3.2 Applications of the Models

It has been shown that the stochastic models are suitable for output prediction and output control. We have also demonstrated that the parameters of these models can be estimated for the same purposes. An important ^{point} is that for control purposes, the models provide information suitable for feedforward as well as feedback control when correlated disturbances are encountered.

We believe that a single output form of the Σ_{DS} description shown in equation (7.3-2) has been used to model part of the process of kraft paper making at the Billerud Company in Sweden for a number of years [69]. The control requirement is the regulation of basis weight of the paper within certain narrow limits. With the introduction of a multivariate model, which is a generalization of the single output model, it would be possible to simultaneously predict and control

other output variables which interact with basis weight (such as moisture control).

The use of the Σ_{DS} description is not limited to control situations. In Chapter 5, it was shown that by setting various matrix coefficients to zero, the equations can be used to model multivariate autoregressive series, moving average processes and combined mixed autoregressive moving average processes that arise in time series analysis. In their comprehensive series of technical reports (including [23] and [33] which we referenced previously), Box and Jenkins give a number of examples of interesting time series such as stock market trends, weekly business or sales results, etc. The Σ_{DS} description for $m = 1$ (single output) is essentially the same as the difference equation models that they have considered. Hence, we suggest that the multivariate stochastic models may have application in broader problems in business and economic forecasting.

7.4 Multivariate System Identification

7.4.1 The Problem of Choosing a Linear Structure

In the first chapter, when introducing the subject, we subdivided the identification problem into two phases:

- (1) the choice of a model structure
- (2) the estimation of the parameters of the structure.

Only the latter phase has been considered in detail. We are concerned only with linear models for the former.

In Section 6.7, we demonstrated how to choose the order of the model in the simple case of a single output. In the general case, the problem of selecting a structure is much more difficult. The problem involves:

- (1) Choosing p , the order of the vector difference equation.
- (2) Choosing k , the exponent of the delay term.
- (3) Choosing n (the order of the system matrix F).

For example, consider a system with:

- (a) two inputs, $r = 2$
- (b) two outputs, $m = 2$
- (c) three states, $n = 3$

We should be prepared to compare the following models:

- (1) $n = 2, p = 1$ so that the homogeneous term

$$A(z^{-1})y(t) = y(t) + A_1y(t-1) \quad (7.4-1)$$

- (2) $n = 3, p = 2$ so that either

$$A(z^{-1})y(t) = y(t) + A_1y(t-1) + A_2y_1(t-2)$$

$$\text{or} \quad = y(t) + A_1y(t-1) + A_2'y_2(t-2) \quad (7.4-2)$$

where A_2 or A_2' are 2×1 .

(3) We carry on until we are convinced that the estimated sequence of residuals $\{\hat{e}^*(t)\}$ is uncorrelated and that the likelihood function $L^*(\Theta)$ is minimized. In addition, we should also test for delayed versions of the input sequence by altering the exponent k .

Clearly, the problem of finding the structure of a system (and a linear system at that) is time consuming. There is considerable scope for ingenuity in simplifying the search. Note, however, that this problem is common to the identification of single output systems also.

7.4.2 Simplifications

Recently, Briggs and Williams have demonstrated a method for the simultaneous determination of the cross-correlations from several inputs to the output of a system as an application of modified pseudo-random binary sequences [70]. For our application, the most feasible method of determining the delay exponent k is to study the first few terms of the cross-correlations. No additional measurements are needed because the input excitation requirements for cross-correlations by the above method and parameter estimation by the methods of this thesis are the same. In fact, the input sequences devised by Briggs and Williams are superior to the sequences used by the author in the example of Section 6.8 because the cross-correlation between the inputs is bound to be lower for the former than for the latter.

The following method of determining the order p of a multivariate system model is suggested as a topic of further research. Suppose that we fit the following two models by regression methods (using the "bootstrap" estimator, for example) to a long record of output observations obtained from a system that can be described by a model of order p :

Model (1) of order p is $A(z^{-1})y(t) = v(t)$ (7.4-3)

Model (2) of order $q < p$ is $Q(z^{-1})y(t) = v'(t)$ (7.4-4)

If $A(z^{-1})$ is exact, there are no more than p nonzero autocorrelation terms $R_v(0), R_v(1) \dots R_v(p)$. Now consider the following operations on equation (7.4-2).

$$\begin{aligned} A(z^{-1})Q^{-1}(z^{-1})Q(z^{-1})y(t) &= A(z^{-1})Q^{-1}(z^{-1})v'(t) \\ &= v(t) \\ &= C(z^{-1})e^*(t) \end{aligned} \quad (7.4-5)$$

$$\text{Hence } v'(t) = Q(z^{-1})A^{-1}(z^{-1})C(z^{-1})e^*(t) \quad (7.4-6)$$

which, for an infinitely long record, has an infinite number of terms, although $C(z^{-1})e^*(t)$ has no more than p terms. Hence, there will be more than q nonzero autocorrelation terms.

The questions to consider are:

- (1) Under what circumstances, if any, are poles of $A^{-1}(z^{-1})$ annihilated by zeros of $Q(z^{-1})$ and $C(z^{-1})$ in (7.4-6)?
- (2) Under what circumstances can we reject the hypothesis that model (2) is the correct model? (The likely hypothesis will be based on tests of the autocorrelation coefficients $R_{v'}(\tau)$ for $\tau \geq q$.)

For example, suppose that we fit the first order process

$$A(z^{-1})y(t) = e^*(t) \quad (7.4-7)$$

by the model
$$y(t) = v'(t) \quad (7.4-8)$$

Then
$$v'(t) = A^{-1}(z^{-1})e^*(t) \quad (7.4-9)$$

which has nonzero autocorrelation coefficients $R_{v'}(\tau)$ for $\tau \neq 0$.

Hence we reject the model.

7.4.3 Spectral Factorization

In this thesis, the problem of finding the coefficients C_1, C_2, \dots, C_p and σ^2 of the process $C(z^{-1})e^*(t)$ was solved by moment fitting (see Subsection 4.3.3). The drawback of this method is that an "unstable" solution can be obtained. Hence, as a research topic, the methods of spectral factorization recently considered by Ho and Kalman [57] and by Anderson [60, 61] are worthy of investigation.

The problem is: given the function (from Subsection 4.3.3)

$$R(z) \triangleq C(z) \Lambda C^T(z^{-1}) \quad (4.3-35)$$

find $C(z^{-1})$ and Λ . There is a distinct possibility that the solution proposed by Ho and Kalman that employs a solution to the Riccati equation can be extended to the multivariate case.

7.4.4 Estimation of the Parameters of the Σ_s Description

As a direct result of the derivation of the Σ_s description in the form

$$x(t+1) = Fx(t) + Gu(t) + \Gamma e^*(t) \quad (7.4-10)$$

$$y(t) = Hx(t) + Du(t) + e^*(t) \quad (7.4-11)$$

we propose as a new research problem, the estimation of the parameters of the above model by MLE methods. It will be necessary to assume a suitable canonical form. Then the equations to generate the error sequence $\{e^*(t)\}$ for a given set of parameters Θ are:

(1) Assume $x(t)$ is known. (It may be convenient to assume $x(0) = 0$.)

$$(2) \text{ Calculate } e^*(t) = y(t) - Hx(t) - Du(t) \quad (7.4-12)$$

(3) Calculate

$$x(t+1) = (F - IH)x(t) + (B - ID)u(t) + Iy(t) \quad (7.4-13)$$

and repeat from (1).

The (negative log) likelihood function defined for this problem is exactly the same as that formulated in Section 5.6.

$$L^*(\Theta) = \frac{mN}{2} \ln(2\pi) + \frac{N}{2} \ln |\Lambda| + \frac{1}{2} \sum_{t=1}^N e^{*T}(t) \Lambda^{-1} e^*(t) \quad (5.6-71)$$

The method of minimizing $L^*(\Theta)$ is essentially the same as that explained in Section 5.6. The main problem is to assume a suitable canonical form. In this regard, a form in which $H = [I_m \ 0]$ is of some advantage because it minimizes the number of elements in equation (7.4-13) that appear as products of system parameters. For example, if $D = 0$, then equation (7.4-13) is simply

$$x(t+1) = [F - I[I_m \ 0]]x(t) + Bu(t) + Iy(t) \quad (7.4-14)$$

Hence, when we differentiate $x(t+1)$ with respect to any parameter

of the system in this special case, the resulting function is independent of other system parameters.

An effective solution to the above problem would be of significant assistance in the identification of multivariate linear systems.

REFERENCES

- [1] E. S. SAVAS
"Mathematical models"; Computer Control of Industrial Processes, McGraw Hill, New York, 1965, p. 43
- [2] P. EYKHOFF, P. van der GRINTEN, H. KWAKERNAAK and B. VELTMAN
"Systems modelling and identification"; Survey paper, Third Congress of IFAC, London, June 1966
- [3] R. E. KALMAN
"On structural properties of linear constant multivariable systems"; Paper 6A, Third Congress of IFAC, London, 1966
- [4] F. R. GANTMACHER
"The Theory of Matrices"; Chelsea Publishing Co. (New York), 1959
- [5] R. W. BASS and I. GURA
"Canonical forms for controllable systems with applications to optimal nonlinear feedback"; Paper 7A, Third Congress of IFAC, London, 1966
- [6] W. M. WONHAM and C. D. JOHNSON
"Optimal bang-bang control with quadratic performance"; Proc. 4th JACC, 1963, pp. 101-112
- [7] C. D. JOHNSON and W. M. WONHAM
"Another note on the transformation to canonical (phase-variable) form"; IEEE Trans. on Auto. Control, Vol. AC-11, 1966, pp. 609-610
- [8] R. E. KALMAN
"On the general theory of control"; First Congress of IFAC, Butterworths (London), 1961
- [9] C. T. CHEN, C. A. DESOER and A. NIEDERLINSKI
"Simplified conditions for controllability and observability of linear time-invariant systems"; IEEE Trans. on Auto. Control, Vol. AC-11, 1966, pp. 613-614
- [10] D. G. LUENBERGER
"Canonical forms for linear multivariable systems"; IEEE Trans. on Auto. Control, Vol. AC-12, 1967, pp. 290-293

- [11] B. D. O. ANDERSON and D. G. LUENBERGER
"Design of multivariable feedback systems"; Proc. IEE,
Vol. 114, 1967, pp. 395-399
- [12] R. E. KALMAN
"Mathematical description of linear dynamical systems";
Jnl. SIAM on Control, Ser. A., Vol. 1, 1963, pp. 152-192
- [13] A. M. LEPOV
"Stability in nonlinear control systems"; Princeton U.P.,
1961
- [14] W. J. DUNCAN and A. R. COLLAR
"A method for the solution of oscillation problems by
matrices"; Philosophical Magazine, Ser. 7, Vol. 17, 1934,
pp. 865
- [15] W. J. DUNCAN and A. R. COLLAR
"Matrices applied to the motions of damped systems";
Philosophical Magazine, Ser. 7, Vol. 19, 1935, pp. 197
- [16] R. A. FRAZER, W. J. DUNCAN and A. R. COLLAR
"Elementary Matrices"; Cambridge University Press, 1938
- [17] L. A. PIPES
"Matrix methods for engineering"; Prentice-Hall (New Jersey),
1963
- [18] K. J. ASTROM and T. BOHLIN
"Numerical identification of linear dynamic systems from
normal operating records"; IFAC Symposium on "The theory
of self-adaptive systems", Teddington, 1965
- [19] K. J. ASTROM, T. BOHLIN and S. WENSMARK
"Automatic construction of linear stochastic dynamic models
for stationary industrial processes using operating records";
IBM Nordic Laboratory (Sweden), Tech. paper No. 18.150, 1965
- [20] D. Q. MAYNE
"A method for estimating discrete-time transfer functions";
Paper C2, Second U.K.A.C. Control Convention, Bristol,
April 1967
- [21] S. G. TZAFESTAS
"Estimation of parameters of multi-input and multi-output
linear dynamic systems"; Report, Dept. of Electrical
Engineering, Imperial College of Science and Technology,
Sept. 1966

- [22] G. HADLEY
"Linear Algebra"; Addison-Wesley (1961)
- [23] G. E. P. BOX and G. M. JENKINS
"Recent advances in forecasting and control"; Technical Report No. 5, Dept. of Systems Engineering, University of Lancaster, July 1966
- [24] D. E. JOHANSEN and A. E. BRYSON
"Linear filtering for time-varying systems using measurements containing coloured noise"; IEEE Trans. on Auto. Control, Vol. AC-10, No. 1, 1965, pp. 4-10
- [25] L. A. ZADEH and C. A. DESOER
"Linear System Theory"; McGraw Hill Book Co. (1963)
- [26] E. KREINDLER and P. E. SARACHIK
"On the concepts of controllability and observability of linear systems"; IEEE Trans. on Auto. Control, Vol. AC-9, No. 2, 1964, pp. 129-136
- [27] M. ATHANS and P. L. FALB
"Optimal Control"; McGraw Hill Book Co. (1966)
- [28] H. H. ROSENBROCK
"Transformation of linear constant system equations"; Proc. IEE, Vol. 114, pp. 541-544
- [29] H. H. ROSENBROCK
"On linear system theory"; Proc. IEE, Vol. 114, pp. 1353-1359
- [30] V. N. FADDEEVA
"Computational Methods of Linear Algebra"; Dover Publications, (1959)
- [31] R. FLETCHER and M. J. D. POWELL
"A rapidly convergent descent method for minimization"; The Computer Journal, Vol. 6, 1963, pp. 163-168
- [32] M. J. D. POWELL
"An efficient method of finding the minimum of a function of several variables without calculating derivatives"; The Computer Journal, Vol. 7, 1964, pp. 155-162
- [33] G. E. P. BOX and G. M. JENKINS
"Models for prediction and control; II linear stationary models"; Technical Report No. 2, Dept. of Systems Engineering, University of Lancaster

- [34] A. M. MOOD and F. A. GRAYBILL
"Introduction to the Theory of Statistics"; McGraw Hill
Book Co., 1963
- [35] R. J. KAVANAGH
"Control of multivariable systems"; National Research
Council of Canada proceedings of the seminar on optimal
control theory, Montreal, Sept. 1962
- [36] A. S. BOKSENBOM and R. HOOD
"General algebraic method applied to control analysis of
complex engine types"; NACA Tech. Report 980, April 1949
- [37] H. S. TSIEN
"Engineering Cybernetics"; McGraw Hill Book Co., 1954
- [38] R. C. AMARA
"Linear least squares synthesis of multivariable control
systems"; Trans. AIEE, pt. 2, May 1957, p. 115
- [39] R. DEUTSCH
"Estimation Theory"; Prentice-Hall Inc. (1965)
- [40] T. W. ANDERSON
"Introduction to Multivariate Statistical Analysis";
John Wiley and Sons (1957)
- [41] H. MANN and A. WALD
"On the statistical treatment of linear stochastic differ-
ence equations"; Econometrica, Vol. 11, 1943, pp. 173-220
- [42] R. E. KALMAN
"Design of a self-optimizing control system"; Trans. ASME,
Vol. 80, 1958, pp. 468-478
- [43] R. C. K. LEE
"Optimal Estimation, Identification and Control";
MIT Press (1964)
- [44] D. Q. MAYNE
"Parameter estimation"; Automatica, Vol. 3, 1966, pp. 245-255
- [45] R. E. KALMAN and R. S. BUCY
"New results in linear filtering and prediction theory";
Trans. ASME (J. Basic Eng.), March 1961, pp. 95-108

- [46] A. C. AITKEN
"On least squares and linear combinations of observations";
Proc. Royal Soc. (Edinburgh), A, Vol. 55, 1934, p. 42
- [47] J. JOHNSTON
"Econometric Methods"; McGraw Hill Book Co., (1963)
- [48] P. A. N. BRIGGS, D. W. CLARKE and P. H. HAMMOND
"An introduction to statistical identification methods in
control systems"; National Physical Laboratory (Teddington),
Autonomics Division Report No. Auto 24, Dec. 1966
- [49] J. DURBIN
"Estimation of parameters in time-series regression models";
Jnl. Royal Stat. Soc., B, Vol. 22, 1960, pp. 139-153
- [50] S. A. TRETTER and K. STEIGLITZ
"Power spectrum identification in terms of rational models";
IEEE Trans. on Auto. Control, Vol. AC-12, 1967, pp. 185-188
- [51] D. W. CLARKE
"Generalized least squares estimation of the parameters of a
dynamic model"; National Physical Laboratory (Teddington),
Autonomics Division Report No. Auto 26, Nov. 1966; also paper
3-17 in Proceedings of the IFAC Symposium on Identification
in Auto. Control Systems, Prague, June 1967
- [52] D. W. NORRIS
"Investigations into practical parameter estimation problems";
Centre for Computing and Automation Research Report, Imperial
College of Science and Technology (London), Aug. 1967
- [53] R. A. RUCKER
"Real time system identification in the presence of noise";
IEEE Wescon Convention Record, Aug. 1963
- [54] V. S. LEVADI
"Parameter estimation of linear systems in the presence of
noise"; Honeywell Document M-247, Aug. 1964
- [55] V. S. LEVADI
"Parameter estimation of linear systems in the presence of
noise"; Proceedings of International Conference on Microwaves,
Circuit Theory and Information Theory, Tokyo, Sept. 1964
- [56] D. Q. MAYNE
"Estimation and Control of Stochastic Systems"; Ph.D. Thesis,
University of London, 1967

- [57] B. L. HO and R. E. KALMAN
"Spectral factorization using the Riccati equation"; 4th Annual Allerton Conference on Circuit and System Theory, University of Illinois, Oct. 1966, pp. 388-399
- [58] D. C. YOULA
"On the factorization of rational matrices"; IEEE Trans. on Info. Theory, Vol. IT-7, 1961, pp. 172-189
- [59] M. C. DAVIS
"Factorizing the spectral matrix"; IEEE Trans. on Auto. Control, Vol. AC-8, 1963, pp. 296-305
- [60] B. D. O. ANDERSON
"Development and application of a system theory criterion for rational positive real matrices"; 4th Annual Allerton Conference on Circuit and System Theory, University of Illinois, Oct. 1966, pp. 400-407
- [61] B. D. O. ANDERSON
"An algebraic solution to the spectral factorization problem"; IEEE Trans. on Auto. Control, Vol. AC-12, 1967, pp. 410-414
- [62] N. WIENER
"The Extrapolation, Interpolation and Smoothing of Stationary Time Series"; John Wiley and Sons, 1958
- [63] H. CRAMER
"Mathematical methods of statistics"; Princeton University Press, 1946
- [64] D. Q. MAYNE
"An algorithm for the calculation of the pseudo-inverse of a singular matrix"; The Computer Journal, Vol. 9, Nov. 1966
- [65] D. M. DETCHMENDY and R. SRIDHAR
"Sequential estimation of states and parameters in noisy nonlinear dynamical systems"; Proc. JACC, Session VI, Paper 4, pp. 56-63
- [66] A. A. FELDBAUM
"Dual control theory"; Automatic and Remote Control; Vol. 21, 1960, p. 874; Vol. 21, 1960, p. 1033; Vol. 22, 1961, p. 1; Vol. 22, 1961, p. 109
- [67] H. KWAKERNAAK,
"On-line dynamic optimization of stochastic control systems"; Report, Technological University, Delft, Netherlands, December 1964

- [68] J. S. RIORDON
"Adaptive control of discrete state Markov processes"
Ph.D. Thesis, Imperial College, University of London,
February, 1967
- [69] K. J. ASTROM
"Control problems in paper making"; presented at IBM Scientific Computing Symposium: Control Theory and Applications, Yorktown Heights, N.Y., October 1964
- [70] P. A. N. BRIGGS and WILLIAMS, B. J.
"Application of modified pseudo-random sequences to multi-input processes"; IEE Colloquium on Pseudo-Random Sequences Applied to Control Systems, January 1967