IDENTIFICATION OF MULTIVARIABLE SYSTEMS

STRUCTURAL IDENTIFICATION AND PARAMETRIC ESTIMATION OF MULTIVARIABLE SYSTEMS

By

YORK-HON KWONG, B.Eng.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Engineering

McMaster University

February 1977

MASTER OF ENGINEERING (1976) McMASTER UNIVERSITY

(Electrical Engineering) Hamilton, Ontario

TITLE : Structural Identification and Parametric Estimation of Multivariable Systems

AUTHOR : York-Hon Kwong

B.Eng. (McMaster University)

SUPERVISOR : N.K. Sinha

B.Sc.Eng. (Banaras)

Ph.D. (University of Manchester)

NUMBER OF PAGES : viii, 70

ABSTRACT

The identification of discrete linear time-invariant multivariable systems has been investigated. This thesis consists of three interrelated parts. In the first part, linear systems have been described in canonical forms which are mostly suitable for system identification. Inter-relationships between the input-output difference canonical form due to Bonivento, Guidorzi and Marro (1973) and the state-space canonical form due to Luenberger (1967) have been studied. In the second part structures of linear multivariable systems have been studied. An efficient algorithm has been developed to identify a set of structural indices directly from the input-output data of a system. The order of the system is, then, determined in terms of these structural indices. In the last part, estimation of system parameters from the input-output data has been studied. Recursive algorithms have been developed for estimating the system parameters sequentially from noise-free and noise-corrupted data. Digital computer simulations with noise-free as well as noise-corrupted data have been utilized to illustrate the usefulness of the proposed procedures.

ACKNOWLEDGEMENTS

The author would like to express his sincere gratitude and appreciation to Dr. N.K. Sinha for constant guidance and spiritual encouragement throughout the preparation of this work.

Special thanks are due to Mr. J. Hickin and Mr. H. El-Sherief from whom the author received many helpful discussions and suggestions.

During the period of study at McMaster, the author is indebted to many of his former and present colleagues for ideas and encouragement. Since individual acknowledgement is impossible, they are collectively thanked.

Financial support of the work by the Department of Electrical Engineering of McMaster University is gratefully acknowledged.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

System identification has been under active study for the past decade. It has many interdisciplinary applications in the fields of engineering, physical sciences, and social sciences. Besides its important role in automatic control and systems engineering, it also finds many new applications in developing fields such as bioengineering and econometrics.

Classically, system identification is defined according to Zadeh as "the determination, on the basis of input and output, of a system within a specified class of systems, to which the system under test is equivalent." Basically, the problem of system identification consists of three parts: systems modelling, parameter estimation, and validation. Firstly, a certain class of models must be chosen to represent the system under test. This usually involves either qualitative or statistical determination of the structure of the system. Relationships between inputs and outputs of the system are investigated and defined. Secondly, the unknown parameters of the constructed model must be estimated from simulated or normal operating records. A predetermined objective is formulated and optimized by a suitable algorithm.

Thirdly, to see whether or not the estimated model is adequate, a validation test must be performed.

In recent years, hundreds of papers on system identification have been published. Many new algorithms and applications have been reported. Most of the work has, however, been concentrated on developing new algorithms for parametric estimation and finding new applications for system identification. Frequently, a system structure or system model is assumed a priori during the developments of new algorithms and only single-input single-output systems have been cited for examples. Little work has been done on multivariable systems.

With the rapid development of modern control theories, systems engineers must deal with sophisticated systems more often. Design and control of multivariable systems become commonplace and identification of multivariable systems has become a necessity.

It is only recently that identification of multivariable systems has drawn much attention. Early contributions in this field were made by Ho and Kalman (1966). Later approaches using Markov parameters were proposed by Ackermann (1971) and Mayne (1972). Gopinath (1969) and Budin (1971) have described a method utilizing a predetermined selector matrix to obtain a set of input-output equations to solve the problem. All these methods, however, require some a priori knowledge about the system and are useful only for

the case of noise-free observations. Tse and Weinert (1975) have considered the structure determination and parameter identification for multivariable stochastic linear systems, but only systems with zero control are considered. Sen and Sinha (1976) have developed an on-line method for identification of multivariable systems which identifies the transfer function matrix of the system, but poor results were obtained for noisy data. Moreover, there are too many parameters to be estimated. Guidorzi (1975) has presented an approach based on Luenberger's canonical form which minimizes the number of parameters to be estimated and also allows the determination of the structure of the system. However, an ordinary least-squares algorithm has been used for parameter estimation giving biased estimates in the case of noisy data and is not suitable for on-line applications.

In this thesis, a unified approach has been adopted to identify the structure and the parameters of a given system directly from normal operating records. No a priori assumptions about the system structure need be made nor need the input-output sequences of the system be severely restricted. A new canonical form has been chosen to represent multivariable systems. It consists of a set of differential/difference equations interconnecting the input and output sequences of a system. The constructed model is simple in structure and both the structure and parameters of the model can be identified directly from input-output sequences of

the Th ha

an alternative representation of multivariable systems. It has unique relationships with Luenberger's canonical form and the transfer function matrix representation of multivariable system.

The identification procedures considered in this thesis consist of first determining the structural invariants of a system from input-output sequences by statistical methods. The set of indices so determined will completely define the structure and the order of the system and is invariant under transformation of state variables (Popov (1972)). Subsequently the parameters of the model are identified recursively to increase the adaptive capability of the algorithm so that it can be used for on-line applications.

Chapter 2 will concern itself with a brief review of various classical and modern methods of representation of linear systems. Transfer function matrix, state-space and input-output descriptions have been discussed. Methods for transforming from one canonical form to another have been presented. Equivalence between Luenberger's canonical form and the input-output difference canonical form has also been discussed.

Chapter 3 considers the structural identification of linear multivariable systems. Invariants of linear systems have been introduced. A statistical method has been adopted to determine the structural indices of a given system from its normal operating records. Noise-free and noise-corrupted

measurements have been considered in the development of the algorithm.

Chapter 4 deals with the problem of parametric estimation. Systems described in input-output difference canonical forms have been decomposed into subsystems which as a result, may be separately identified. Recursive least-squares and recursive generalized least-squares algorithms have been adopted to estimate the parameters of each subsystem sequentially. Noise-free and noise-corrupted data have also been considered.

Chapter 5 is devoted to the testing of the structural identification and parametric estimation algorithms presented in Chapters 3 and 4. A fairly general class of two-input two-output system has been used for testing. Three examples in this class of systems have been simulated and the algorithms tested. One of the systems is of the 3rd-order and the other is of the 5th-order. The third example is of the 4th-order and has been used mainly to test the parametric estimation algorithms. Computational details of each example have been included to demonstrate the procedures involved.

Chapter 6 gives the conclusions and suggestions for future work in this area.

CHAPTER 2

MATHEMATICAL REPRESENTATION OF LINEAR SYSTEMS

2.1 Introduction

Linear time-invariant systems can be described in many different ways: by internal descriptions such as state-space models or by external descriptions such as impulse responses or transfer function matrices. Because of different analytical methods used and different objectives required, each description has its own importance and advantages. In the following, representations of linear discrete-time multivariable systems have first been expressed in the state-space forms and transfer function matrices. Later a new canonical description has been introduced which provides a simple relationship between the input and output sequences of a system and can be easily implemented for the purpose of identification. This input-output difference description of linear systems can be deduced uniquely from the block-companion form of the state-space model by linear transformation. Equivalence between these two classes of system representations has also been discussed.

2.2 The External Description - Input-output Difference Description

The external description of a system gives a mathematical

relation between the input and output of the system. No knowledge of the internal structure of the system is assumed and only the input and output terminals of the system are accessible for measurements.

Consider a linear multivariable system having m-inputs and r-outputs as represented by Figure 2.2.1. The system can be described by an r x m discrete transfer function matrix H(z), the elements of which are strictly proper rational functions of z, where z^{-1} is the unitary delay operator, that is,

$$Y(z) = H(z)\ U(z) \qquad (2.2.1)$$

where

$$Y(z) = \begin{bmatrix} y_1(z) \\ y_2(z) \\ \vdots \\ y_r(z) \end{bmatrix} \triangleq \text{output vector of dimension r.}$$

$$U(z) = \begin{bmatrix} u_1(z) \\ u_2(z) \\ \vdots \\ u_m(z) \end{bmatrix} \triangleq \text{input vector of dimension m}$$

The transfer function matrix H(z) can be decomposed into two matrices such that

$$H(z) = P^{-1}(z)\ Q(z) \qquad (2.2.2)$$

where P(z) and Q(z) are polynomial matrices of z and their orders are r x r and r x m respectively. P(z) is also

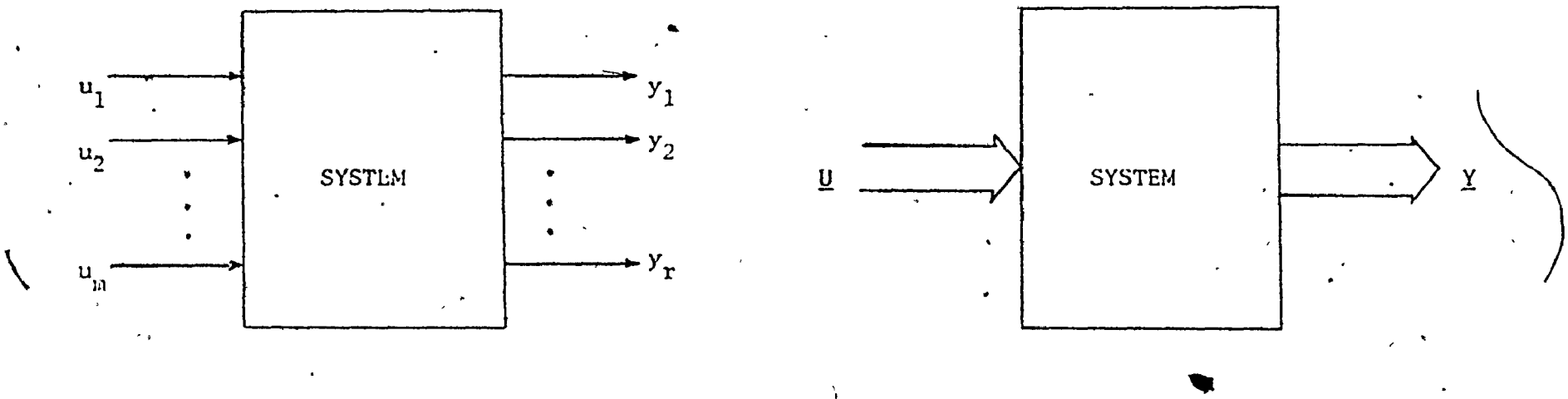

Figure 2.2.1 A multivariable system with m input terminals and r output terminals.

required to be non-singular. Substituting equation (2.2.2) in equation (2.2.1).

$$Y(z) = P^{-1}(z)\ Q(z)\ U(z) \qquad (2.2.3)$$

and premultiply both sides of equation (2.2.3) by P(z).

$$P(z)\ Y(z) = Q(z)\ U(z) \qquad (2.2.4)$$

Hence, systems described by transfer function matrix can be easily converted to input-output difference description. As one may note, the decomposition of H(z) into $P^{-1}(z)$ and Q(z) is not unique. Certain canonical forms for this input-output difference description have yet to be established.

2.3 The Internal Description - State-space Model

Consider the same system as shown in Figure 2.2.1, the system can be described by discrete state-space equations

$$X(k+1) = A\ X(k) + B\ U(k) \qquad (2.3.1a)$$

$$Y(k) = C\ X(k) \qquad (2.3.1b)$$

where X belongs to n-dimensional state vector space, U to the m-dimensional input vector space, Y to the r-dimensional output vector space. A, B and C are constant matrices and their orders are n x n, n x m and r x n, respectively, k is the sampling instant.

It is well known that this system representation S(A, B, C) is not unique. Any nonsingular matrix T can transform the system into another equivalent system $\bar{S}(TAT^{-1}, TB, CT^{-1})$ in the sense that both S and $\bar{S}$ have the

same input-output description. The fact that the input-output relation is invariant under a linear transformation of the state variables implies that not all parameters of the systems described by state-space model can be identified directly from input-output measurements.

In order to obtain unique solutions and to construct efficient algorithms for the purpose of system identification, it is of great interest to represent linear systems in some canonical forms. Different canonical forms for linear multivariable systems have been discussed by Luenberger (1967), Weinert and Anton (1972) and Rissanen (1974). Algorithms for transforming system dynamic equations to different canonical forms have also been developed by Sinha and Roza (1976). Besides having minimum number of parameters to be estimated, canonical forms also provide convenient ways for system decomposition. Of all the canonical forms discussed, the block-companion form due to Luenberger (1967) is the most important one in system identification and will be discussed in detail in the following section.

2.4 Equivalence Between Input-output Description and State-space Model

Equivalence between a class of state-space models and the corresponding input-output difference description can be established by linear transformation and has been discussed by Guidorzi (1975). Once a certain canonical form is chosen to represent a system the corresponding input-output

differential/difference description will be unique. In the following, equivalence between Luenberger's canonical form and the corresponding input-output difference canonical form has been developed. Consider a linear multivariable system expressed in discrete state-space form

$$X(k+1) = \bar{A}\, X(k) + \bar{B}\, U(k) \qquad (2.4.1a)$$

$$Y(k) = \bar{C}\, X(k) \qquad (2.4.1b)$$

If the system is completely observable and controllable, a suitable nonsingular matrix T can be used to transform the system into an equivalent canonical form such that

$$A = T\, \bar{A}\, T^{-1} \qquad (2.4.2a)$$

$$B = T\, \bar{B} \qquad (2.4.2b)$$

$$C = \bar{C}\, T^{-1} \qquad (2.4.2c)$$

The resultant system S(A,B,C) is in the block row companion observable form and the matrices A, B and C have the following special structures:

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ \vdots & & & \\ A_{i1} & A_{i2} & \cdots & A_{in} \\ \vdots & & & \\ A_{n1} & \cdots & \cdots & A_{nn} \end{bmatrix} \qquad (2.4.3)$$

$$A_{ii} = \begin{bmatrix} 0 & & & \\ \vdots & & I_{V_i - 1} & \\ 0 & & & \\ a_{ii,1} & \cdots & \cdots & a_{ii,V_i} \end{bmatrix} \qquad (2.4.4a)$$

$$A_{ij} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \\ a_{ij,1} & \cdots & a_{ij,v_{ij}} \cdots 0 \end{bmatrix} \qquad (2.4.4b)$$

$$B = \begin{bmatrix} \bar{b}_{11} & \cdots & \bar{b}_{1m} \\ \vdots & & \vdots \\ \bar{b}_{n1} & \cdots & \bar{b}_{nm} \end{bmatrix} = \begin{bmatrix} b_1^T \\ b_2^T \\ \vdots \\ b_n^T \end{bmatrix} \qquad (2.4.5)$$

$$C = \begin{bmatrix} 1 & 0 & \cdots & & & & & 0 \\ 0 & & & 0 & 1 & & & 0 \\ \vdots & & & & & & & \vdots \\ 0 & \cdots & & & 1 & 0 & \cdots & 0 \end{bmatrix} \qquad (2.4.6)$$

$$\begin{matrix} \uparrow & & \uparrow & \uparrow \\ 1 & & (v_1+1) & \left(\sum_{i=1}^{r-1} v_i+1\right) \end{matrix}$$

where the integers $v_1, v_2, \ldots, v_r$ are the structural invariants of the system.

The order of the system is

$$n = \sum_{i=1}^{r} v_i \qquad (2.4.7a)$$

and

$$V_{ij} \leq v_i \qquad (2.4.7b)$$

Furthermore, systems described by this canonical form are completely observable and can be decomposed into their many observable subsystems. Hence conventional algorithms for the identification of single-input single-output systems can be applied to identify each of these subsystems individually. For example, the j-th subsystem in the

canonical representation is completely observable from the j-th component of the output.

Defining the following vectors:

$$X(k) = \left[x_{k,1}, x_{k,2} \cdots, x_{k,n} \right]^T \qquad (2.4.8)$$

$$Y(k) = \left[y_{k,1}, y_{k,2} \cdots, y_{k,r} \right]^T \qquad (2.4.9)$$

$$U(k) = \left[u_{k,1}, u_{k,2} \cdots, u_{k,m} \right]^T , \qquad (2.4.10)$$

The system described by equations (2.4.3) to (2.4.6) can be expressed in an alternative form

$$X(k) = V(z)\, Y(k) - W\, Z(z)\, U(k) \qquad (2.4.11)$$

where

$$V(z) = \begin{bmatrix} 1 & \cdots & 0 \\ z & & \vdots \\ \vdots & & \\ z^{v_1-1} & & \\ 0 & & 0 \\ \vdots & & \vdots \\ 0 & & 1 \\ & & z \\ & & \vdots \\ & & z^{v_r-1} \end{bmatrix} \qquad (2.4.12)$$

$$Z(z) = \begin{bmatrix} & & I \\ & & zI \\ & & \vdots \\ & & z^{v_M-1}I \end{bmatrix} \qquad (2.4.13a)$$

$$v_M = \max (v_i), \quad i = 1, 2, \ldots, r \qquad (2.4.13b)$$

$$W = \begin{bmatrix} 0 & \cdots & 0 \\ b_1^T & 0 & 0 \\ \vdots & & \vdots \\ b_{v_1-1}^T \cdots & b_1^T & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \\ b_{n-v_r-1}^T & 0 & 0 \\ \vdots & & \\ b_{n-1}^T \cdots & b_{n-v_r-1}^T & 0 \cdots 0 \end{bmatrix} \qquad (2.4.14)$$

The input-output difference canonical form can then be obtained by substitution of equation (2.4.11) into equation (2.4.1)

$$[(zI - A)\ V(z)]\ Y(k) = [(zI-A)WZ(z) + B]\ U(k) \qquad (2.4.15)$$

I is a unity matrix of appropriate order. Expanding equation (2.4.15), we can see immediately that only the v_1-th, $(v_1 + v_2)$-th, ..., v_r-th equations are really significant and the rest are just simple identities. Discarding these identities, equation (2.4.15) can be rewritten in a more compact concise form

$$P(z)\ Y(k) = Q(z)\ U(k) \qquad (2.4.16)$$

where

$$P(z) = \begin{bmatrix} p_{11}(z) & \cdots & p_{1r}(z) \\ \vdots & & \\ p_{r1}(z) & \cdots & p_{rr}(z) \end{bmatrix} \qquad (2.4.17)$$

$$Q(z) = \begin{bmatrix} q_{11}(z) & \cdots & q_{1m}(z) \\ \vdots & & \vdots \\ q_{r1}(z) & \cdots & q_{rm}(z) \end{bmatrix} \tag{2.4.18}$$

The polynomials $p_{ij}(z)$ and $q_{ij}(z)$ can be easily obtained by expansion of equation (2.4.15) and the following identities are obtained:

$$p_{ij}(z) = z^{v_i} - a_{ii,v_i} z^{v_i-1} \cdots - a_{ii,2} z^2 - a_{ii,1} \tag{2.4.19}$$

$$p_{ij}(z) = -a_{ij,v_{ij}} z^{v_{ij}-1} \cdots - a_{ij,2} z - a_{ij,1} \tag{2.4.20}$$

$$q_{ij}(z) = \beta_{(v_1+..v_i),j} z^{v_i-1} + \cdots + \beta_{(v_1+..v_{i-2}+2),j} z + \beta_{(v_1+..+v_{i-1}+1),j} \tag{2.4.21}$$

where the coefficients $\beta_{i,j}$ are the elements of the following matrix:

$$\overline{B} = MB = \begin{bmatrix} \beta_{11} & \cdots & \beta_{1m} \\ \vdots & & \vdots \\ \beta_{n1} & \cdots & \beta_{nm} \end{bmatrix} \tag{2.4.22}$$

and the matrix M is constructed as

$$M = \left[\begin{array}{cccc|c|ccc}
-a_{ll,2} & -a_{ll,3} & \cdots & -a_{ll,v_l} & 1 & -a_{lr,2} & \cdots & -a_{lr,v_{lr}} \\
-a_{ll,3} & -a_{ll,4} & \cdots & 1 & & -a_{lr,3} & \cdots & 0 \\
\vdots & & & & & \vdots & & \\
-a_{ll,v_l} & 1 & & & & -a_{ll,v_{lr}} & & \\
1 & & & & & 0 & & \\
\hline
\multicolumn{4}{c|}{} & & \multicolumn{3}{c}{} \\
\hline
-a_{rl,2} & \cdots & -a_{rl,v_{rl}} & 0 & & -a_{rr,2} & \cdots & -a_{rr,v_r} \\
-a_{rl,3} & & 0 & & & -a_{rr,3} & \cdots & \\
\vdots & & & & & \vdots & & \\
-a_{rl,v_{rl}} & 0 & & & & -a_{rr,v_r} & 1 & \\
0 & & & & & 1 & &
\end{array}\right] \qquad (2.4.23)$$

Because of the special structure of the matrix M it is always nonsingular and in fact $\det(M) = 1$.

It is useful to note that polynomials $p_{ij}(z)$ and $q_{ij}(z)$ in equation (2.4.16) also have the following properties:

$$\deg\{p_{ii}(z)\} > \deg\{p_{ij}(z)\} \quad \text{for } j > i \qquad (2.4.24)$$

$$\deg\{p_{ii}(z)\} \geq \deg\{p_{ij}(z)\} \quad \text{for } j < i \qquad (2.4.25)$$

$$\deg\{p_{ii}(z)\} > \deg\{p_{ji}(z)\} \quad \text{for } i \neq j \qquad (2.4.26)$$

$$\deg\{p_{ii}(z)\} > \deg\ q_{ij}(z) \qquad (2.4.27)$$

Thus the equivalence of the two classes of representations of linear systems has been established. It is useful to note that any input-output description of the form (2.4.16) has a unique equivalent state-space representation of the form (2.4.3) - (2.4.6) and vice-versa. Deduction of one form from the other can be accomplished by observing equations (2.4.19) to (2.4.23). From equations (2.4.19) to (2.4.21), we also see that the r x r matrix P(z) contains the same information about the system dynamics that is contained in the state-space matrix A in equation (2.4.3).

Systems described by the input-output difference canonical form have the advantage of being easily identified from normal operating records. If it is necessary, systems identified in input-output difference canonical form can be easily transformed to Luenberger's state-space canonical form by the forementioned procedures and will be discussed in detail in secton 4.5.

CHAPTER 3

STRUCTURAL IDENTIFICATION

3.1 Introduction

The first stage in system identification is to assume or find a suitable model or structure for the system. Once a particular model or structure is chosen the solution of the system identification problem will be straight forward because only the parameters of the model have yet to be estimated. The structure of a single-input single-output system is completely governed by its system order. Hence, the structural idenficiation of single-input single-output systems is the determination of an optimal order for the system. On the other hand, for multivariable systems, structural identification involves finding both the input-output structures as well as the system order.

Invariant description of linear time-invariant controllable systems was first considered by Popov (1972) and a complete set of invariants for such systems has been derived. Later, Weinert and Anton (1972) have extended Popov's work to develop invariant canonical forms which are suitable for system identification.

Statistical test (F-test) was first used by Astrom (1970) to find the order of linear systems. Later, Woodside

(1971) suggested three methods which are based on the product-moment matrix to estimate the system order of single-input single-output systems. Similar statistical techniques were used by Chow (1973) and Tse and Weinert (1975) to determine the structural indices of multivariable systems. Recently, a systematic procedure has been developed by Guidorzi (1975) to determine directly the structural indices of multivariable systems from input-output sequences.

In the following, conventional trial and error methods to find the structural indices of multivariable systems have been abandoned. Statistical methods due to Woodside (1971) and Guidorzi (1975) have been adopted to determine a set of structural indices directly from input-output data without requiring the estimation of the system parameters. This systematic procedure is simple in principle and has been proved to be computationally efficient in application. With little modifications the procedure can be adapted to handle noise-corrupted data too.

3.2 Structural Invariants of Multivariable Systems

Structural invariants are a set of indices which completely characterize the mapping of the input space into the output space of a system. They are invariant with respect to change of coordinates. In multivariable systems, structural invariants not only define the orders of the systems but also the individual structure of the input and output terminals.

Complete descriptions and proof of the basis of invariants and canonical forms for linear dynamical systems have been given by Rissanen (1974) and will not be repeated here. Instead, relationships between the structural invariants and the state-space and input-output difference canonical forms have been stressed.

Consider a discrete-time linear time-invariant system represented by the following discrete state-space equations:

$$X(k+1) = A\,X(k) + B\,U(k) \qquad (3.2.1a)$$

$$Y(k) = C\,X(k) \qquad (3.2.1b)$$

where $X \in R^n$, $Y \in R^r$, $U \in R^m$ and the matrices A, B and C are constant matrices with appropriate orders. k is the sampling instant.

For a stable and observable system, a set of structural indices will completely define the structure as well as the order of the system. It is well known that this set of structural indices is invariant with respect to state-variable transformations and certain kinds of state feedback in the general form of (A + KC), where K is the feedback matrix.

Since the system is completely observable, the following rows of the observability matrix can be constructed

$$\begin{matrix} C_1^T, & C_1^T A, & \ldots, & C_1^T A^{v_1} \\ C_2^T, & C_2^T A, & \ldots, & C_2^T A^{v_2} \\ \vdots & \vdots & & \vdots \\ C_r^T, & C_r^T A, & \ldots, & C_r^T A^{v_r} \end{matrix}$$

where C_i^T is the i-th row of C and v_i, i=1, ..., r is the smallest nonnegative integer such that $C_i^T A^{v_i+1}$ is linearly dependent on the previous one.

Immediately, we note that the order of the system is given by

$$n = \sum_{i=1}^{r} v_i \qquad (3.2.2)$$

If the following matrix is constructed from the rows of the observability matrix

$$T^T = \left[C_1, A^T C_1, \ldots, A^{T^{(v_1-1)}} C_1 | \ldots | C_r, \ldots, A^{T^{(v_r-1)}} C_r\right] \qquad (3.2.3)$$

such that its rank is n, then it can be used for state-transformation of equation (3.2.1). The observable Luenberger's canonical state-space form as described by equations (2.4.3) to (2.4.6) will be obtained. In this special case, the structural indices will correspond to the observability indices of the system.

Similarly with systems described by input-output difference canonical form (equation (2.4.6)), the orders of the diagonal polynomials of the P(z) matrix will be completely specified by the strucutral indices.

Therefore, with the knowledge of the strucutral indices $v_1, \ldots, v_r$, the general structures of the matrices A, B, C, P(z) and Q(z) are known, and they can be constructed. Alternatively, structural indices can be deduced from either the A matrix or the p(z) matrix if it is known in advance. Furthermore, strucutral indices also specify the number of

parameters in the system model.

3.3 Determination of the Structural Invariants from the Input-Output Data

In this section, an algorithm to determine the structural indices directly from the input-output data without the intermediate construction of a parametric model for the system has been presented. This algorithm takes full advantage of the special structures of matrices A and $P(\bar{z})$. Literally, the algorithm is to find a set of indices such that all the input and output spaces of the system are spanned independently.

Let

$$L_i(y_i) = \left[y_j(k)\ y_j(k+1)\ \ldots\ y_j(k+i-1)\right] \tag{3.3.1}$$

$$L_i(u_j) = \left[u_j(k)\ u_j(k+1)\ \ldots\ u_j(k+i-1)\right] \tag{3.3.2}$$

then construct

$$R(\delta_1, \delta_2, \ldots, \delta_{m+r})$$

$$= \left[L_{\delta_1}(y_1) \ \ldots\ L_{\delta_r}(y_r) \ \ldots\ L_{\delta_{r+1}}(u_1) \ \ldots\ L_{\delta_{r+m}}(u_m)\right] \tag{3.3.3}$$

and define the product-moment matrix

$$S(\delta_1, \ldots, \delta_{m+r}) = R^T(\delta_1, \ldots, \delta_{m+r})\ R(\delta_1, \ldots, \delta_{m+r}) \tag{3.3.4}$$

Hence $S(\delta_1, \ldots, \delta_{m+r})$ is a square matrix whose dimension is $\delta_1 + \ldots + \delta_{m+r}$.

Construct the S matrices in increasing dimension sequence

$S(2, 1, \ldots, 1), S(2, 2, \ldots, 1), \ldots S(2, 2, \ldots 2), \ldots$

and test them for singularity. When a singular matrix is found, one of the indices v_i is determined. The procedure is repeated until all the indices associated with each output of the system have been determined. Let $S(\mu_1, \mu_2, \ldots, \mu_{m+r})$ be the singular matrix found and let the index μ_i be increased by one such that the previous matrix in the sequence is nonsingular. Then $v_i = \mu_i - 1$ and $v_{ij} = \mu_j$ $(j = 1, \ldots, m)$, $(j \neq i)$. The structural indices associated with the inputs can then, be determined by observing inequality (2.4.27).

Computational effort can be saved by noting that once one of the S matrices is found to be singular, the subsequent matrices containing elements of that matrix need not be tested again. Therefore only a limited number of matrices need be tested in the implementation of this algorithm.

With the knowledge of the values of the structural indices, the order of the system can be determined as

$$n = \sum_{i=1}^{r} v_i \tag{3.3.5}$$

The number of parameters in the canonical model needed to be identified is given by

$$P_{total} = n(r + m) \tag{3.3.6}$$

where m is the number of input and r is the number of output of the system.

It is not surprising to note that the forementioned

algorithm was developed chiefly for deterministic data. With data corrupted by additive noise, interdependences between the input-output sequences are weakened and to test for singularity of the product-moment matrix will be difficult. Consider the outputs of the system are corrupted with additive noise such that

$$Y_{k,j}^{*} = Y_{k,j} + d_{k,j} \tag{3.3.7}$$

where $d_{k,j}$ is the additive noise sequence.

For uncorrelated noise sequences and if large amounts of data is used, the product-moment matrix can be partioned into

$$S^{*} = S + N\,W \tag{3.3.8}$$

where W is the covariance matrix of the noise sequences and N is the number of data set used.

If the noise sequences are uncorrelated with zero means and equal amounts of them are added to the inputs and outputs of the system, the covariance matrix of the noise can then be approximated as

$$W = \sigma^{2} I \tag{3.3.9}$$

where σ^{2} is the variance of the noise and I is the identity matrix of appropriate order.

If no knowledge of the statistics of the noise is available, the variance of the noise can be assumed to be the least eigenvalue of the product-moment matrix S^{*}. Testing for singularity will then be performed on the enhanced product-moment matrix.

$$S = S^* - N \sigma^2 I \qquad (3.3.10)$$

When different amounts of noise are added to the various inputs and outputs of the system, the same technique can be applied by performing a suitable scaling on the data. Or in a more practical way, the covariance matrix of the noise can be approximated as

$$W = \text{diag}\left[\sigma_1 \; \sigma_2 \; \cdots \; \sigma_r \; \cdots \; \sigma_{r+m}\right] \qquad (3.3.11)$$

where σ_j is the variance of the j-th noise sequence. In this case, the variance of the noises have to be known in advance or be approximated by some suitable method as mentioned before.

CHAPTER 4

PARAMETRIC ESTIMATION

4.1 Introduction

Once the structure or model of a system is determined, the next step is to estimate the parameters of that model. Parametric estimation involves first establishing a reasonably prescribed criterion and then using a suitable algorithm to accomplish that criterion in an optimal way. With only the operating records available, the objective is usually formulated in terms of the residual errors between the measured plant outputs and the outputs predicted by the estimator. The objective is then minimized with respect to the parameters of the system by a suitable algorithm.

Numerous approaches to the parameter estimation problem have been proposed in the literature [Wong and polak (1967), Clarke (1967), Saridis and Stein (1968), Young (1969), Hasting-James and Sage (1969), Rowe (1970), Mehra (1971), Sinha and Pille (1971)] and their relative usefulness has been discussed in many survey papers [Åstrom and Eykhoff (1970), Isermann, Bamberger, Kneppo and Siebert (1973), Sinha and Sen (1974)]. Among these different algorithms and methods, least-squares estimation

is the one most commonly used. The reason is the simplicity in implementation and consistency in estimation. Recursive form of least-squares estimation can also be easily derived. Unfortunately, in the case of noisy measurements least-squares algorithm will normally give biased estimates. To overcome this drawback and avoid developing a new algorithm, the generalized least-squares algorithm can be applied to give an unbiased estimate.

In the following sections, parametric estimation of linear multivariable systems described by input-output difference canonical form have been developed. Systems are first decomposed into their corresponding subsystems. Each of these subsystems is then, very similar to a single-input single-output system and can be identified independently. Recursive least-squares algorithm and recursive generalized least-squares algorithm have been adapted to estimate the system parameters in the cases of low noise and high noise data respectively.

4.2 System Decomposition

Consider a linear multivariable system having m inputs and r outputs and described by the input-output difference canonical form

$$p(z)\ Y(k) = Q(z)\ U(k) \qquad (4.2.1)$$

where

$$Y(k) = \begin{bmatrix} y_1(k) \\ \cdot \\ \cdot \\ \cdot \\ y_r(k) \end{bmatrix} \triangleq \text{output vector}$$

$$U(k) = \begin{bmatrix} u_1(k) \\ \cdot \\ \cdot \\ \cdot \\ u_m(k) \end{bmatrix} \triangleq \text{input vector}$$

$P(z)$ and $Q(z)$ are polynomial matrices of z and are of order $r \times r$ and $r \times m$ respectively.

It will be assumed that the system is completely observable and controllable. It can, therefore, be conveniently decomposed into r observable subsystems each of which corresponds to one row of the matrices $P(z)$ and $Q(z)$. The j-th subsystem of the system is

$$\sum_{i=1}^{r} p_{j,i}(z)\, y_{i,k} = \sum_{i=1}^{m} q_{j,i}(z)\, u_{i,k} \qquad (4.2.2)$$

where $p_{j,i}(z)$ and $q_{j,i}(z)$ are polynomials of z and elements in the j-th row and i-th column of $P(z)$ and $Q(z)$ respectively. $y_{i,k}$ and $u_{i,k}$ are the i-th output and input at the k-th sampling instant, respectively.

Alternatively equation (4.2.2) can be rewritten explicitly as

$$y_{(k+v_j),j} = \sum_{i=1}^{r} \sum_{\ell=1}^{v_{ji}} \alpha_{ji,\ell}\, y_{(k+\ell-1),i}$$

$$+\sum_{i=1}^{m}\sum_{\ell=1}^{v_j} \beta_{(b_1+\ldots+v_{j-1}+\ell),i}\, u_{(k+\ell-1),i} \qquad (4.2.3)$$

where $\alpha_{ji,\ell}$ and $\beta_{r,i}$ are coefficients of the $p_{j,i}(z)$ and $q_{j,i}(z)$ polynomials and v_j and v_{ji} are structural indices of the system.

Each of these r subsystems can be considered as a multi-input single-output system and can be identified independently. The number of parameters to be identified in each subsystem is given by

$$p = n + \sum_{i=1}^{m} (v_i - 1) \qquad (4.2.4)$$

where n is the order of the system, m is the number of inputs of the system and v_i, i = 1, ..., m are part of the structural indices of the system.

Therefore, identification of the whole system can be accomplished in r steps while each subsystem is identified independently.

4.3 Recursive Least-squares Estimation

The input-output description for each subsystem of the main system can be concatenated in the form of a matrix equation. For example, the j-th subsystem can be concatenated as

$$A_{j,k}\, \phi_{j,k} = Y_{j,k} \qquad (4.3.1)$$

where

$$A_{j,k} = \begin{bmatrix} -y_{1,k-1} & -y_{2,k-1} & \cdots & -y_{r,k-1} & u_{1,k-1} & \cdots & u_{m,k-1} \\ -y_{1,k-2} & -y_{2,k-2} & \cdots & -y_{r,k-2} & u_{1,k-2} & \cdots & u_{m,k-2} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ -y_{1,k-N} & -y_{2,k-N} & \cdots & -y_{r,k-N} & u_{1,k-N} & \cdots & u_{m,k-N} \end{bmatrix} \tag{4.3.2}$$

$$\phi_{j,k} = \left[\alpha_{j1,1} \cdots \alpha_{ji,v_{ji}} \cdots \beta_{1,j} \cdots \beta_{i,j} \cdots\right]^T \tag{4.3.3}$$

$$Y_{j,k} \triangleq \left[y_{j,1} \quad y_{j,2} \cdots y_{j,k-N}\right]^T \tag{4.3.4}$$

If the residual errors between the measured plant output and the output predicted by the model are minimized with respect to the parameter vector $\phi_{j,k}$, a least-squares estimate of $\hat{\phi}_{j,k}$ for large N is given by

$$\hat{\phi}_{j,k} = (A_{j,k}^T A_{j,k})^{-1} A_{j,k}^T Y_{j,k} \tag{4.3.5}$$

Equation (4.3.5) involves matrix inversion and is not suitable for on-line application purposes. A recursive form of equation (4.3.5) can be easily derived as follows:

$$P_{j,k} = P_{j,k-1} + E \tag{4.3.6}$$

$$\hat{\phi}_{j,k+1} = \hat{\phi}_{j,k} + \frac{P_{j,k} a_{j,k+1} (y_{j,k+1} - a_{j,k+1}^T \hat{\phi}_{j,k})}{1 + a_{j,k+1}^T P_{j,k} a_{j,k+1}} \tag{4.3.7}$$

$$P_{j,k+1} = P_{j,k} - \frac{P_{j,k} a_{j,k+1} (P_{j,k} a_{j,k+1})^T}{1 + a_{j,k+1}^T P_{j,k} a_{j,k+1}} \tag{4.3.8}$$

where

$$A_{j,k+1} = \begin{bmatrix} A_{j,k} \\ a_{j,k+1} \end{bmatrix} \tag{4.3.9}$$

$$a_{j,k+1} = \begin{bmatrix} -y_{1,k} & -y_{2,k} & \cdots & -y_{r,k} & u_{1,k} & \cdots & u_{m,k} \end{bmatrix} \tag{4.3.10}$$

$$Y_{j,k+1} = \begin{bmatrix} Y_{j,k} \\ y_{j,k+1} \end{bmatrix} \tag{4.3.11}$$

$$Y_{j,k} = \begin{bmatrix} y_{j,1} & y_{j,2} & \cdots & y_{j,k} \end{bmatrix}^T \tag{4.3.12}$$

The estimate $\hat{\phi}_{j,k}$ is updated for each additional set of data received and for fast rate of convergence, initial conditions of the algorithm have been specified as

$$P_{j,o} = \gamma I \qquad \gamma >> 1 \tag{4.3.13}$$

$$\hat{\phi}_{j,o} = 0 \tag{4.3.14}$$

In equation (4.3.6), E is a positive definite matrix whose elements can be specified experimentally to offset the parameter variation between samples. Computationally, the E matrix limits the lower bound of the elements of matrix $P_{j,k}$ such that a continuous correction of $\hat{\phi}_{j,k}$ is possible.

4.4 Recursive Generalized Least-square Estimation

It is well known that the least-squares algorithm only works for noise-free or very little noise measurements. For very noisy measurements, a biased estimate is usually obtained because of the correlated residuals. In order to have an unbiased estimate in noisy environments, a generalized

least-squares algorithm proposed by Clarke (1967) and later the recursive form of this algorithm due to Hastings-James and Sage (1969) has been adapted to estimate the parameters of each subsystem. An autoregressive model of the noise has been developed to filter the input-output data such that the resulting residuals are uncorrelated with zero mean.

Let the correlated residuals be generated as

$$\hat{e}_{j,k} = y_{j,k} - a_{j,k}^T \hat{\phi}_{j,k} \tag{4.4.1}$$

and approximated to fit an autoregressive model

$$\hat{e}_{j,k} = -f_{j,1} \hat{e}_{j,k-1} - f_{j,2} \hat{e}_{j,k-2} \cdots - f_{j,s} \hat{e}_{j,k-s} + W_{j,k} \tag{4.4.2}$$

where $W_{j,k}$ is an uncorrelated zero mean random sequence and s is the order of the autoregressive model.

Equation (4.4.2) can be concatenated as

$$\hat{e}_{j,k} = G_{j,k} \psi_{j,k} + W_{j,k} \tag{4.4.3}$$

where

$$\hat{e}_{j,k} = \left[\hat{e}_{j,1} \; \hat{e}_{j,2} \cdots \hat{e}_{j,k}\right]^T \tag{4.4.4}$$

$$\psi_{j,k} = \left[f_{j,1} \; f_{j,2} \cdots f_{j,s}\right]^T \tag{4.4.5}$$

$$W_{j,k} = \left[W_{j,1} \; W_{j,2} \cdots W_{j,k}\right]^T \tag{4.4.6}$$

and

$$G_{j,k} = \begin{bmatrix} -\hat{e}_{j,0} & -\hat{e}_{j,-1} & \cdots & -\hat{e}_{j,1-s} \\ -\hat{e}_{j,1} & & & -\hat{e}_{j,2-s} \\ \vdots & & & \vdots \\ -\hat{e}_{j,k-1} & \cdots & \cdots & -\hat{e}_{j,k-s} \end{bmatrix} \tag{4.4.7}$$

A least-squares estimate of the auxiliary parameter vector $\psi_{j,k}$ is given by

$$\hat{\psi}_{j,k} = (G_{j,k}^T G_{j,k})^{-1} G_{j,k}^T \hat{e}_{j,k} \qquad (4.4.8)$$

The noisy input and output data can then be filtered to be

$$u_{j,k}^* = u_{j,k} + \sum_{\ell=1}^{s} \hat{f}_{j,\ell}\, u_{j,k-\ell} \qquad (4.4.9)$$

$$y_{j,k}^* = y_{j,k} + \sum_{\ell=1}^{s} \hat{f}_{j,\ell}\, y_{j,k-\ell} \qquad (4.4.10)$$

After substituting these corrected data in equations (4.3.6) to (4.3.8), an unbiased estimate of the system parameter vector will be obtained. The entire procedure may have to be repeated many times in order that the residuals may be uncorrelated with zero mean.

To make the algorithm more adaptive and suitable for on-line applications, the estimation of the system parameters and auxiliary parameters can be both carried out recursively. The following equations are easily derived for the recursive generalized least-squares algorithm:

$$\hat{\phi}_{j,k+1} = \hat{\phi}_{j,k} + \frac{P_{j,k}\, a_{j,k+1}^* (y_{j,k+1} - a_{j,k+1}^{*T} \hat{\phi}_{j,k})}{1 + a_{j,k+1}^{*T} P_{j,k}\, a_{j,k+1}^*} \qquad (4.4.11)$$

$$P_{j,k} = P_{j,k-1} + E \qquad (4.4.12)$$

$$P_{j,k+1} = P_{j,k} - \frac{P_{j,k}\, a_{j,k+1}^{*}\, (P_{j,k}\, a_{j,k+1}^{*})^{T}}{1 + a_{j,k+1}^{*T}\, P_{j,k}\, a_{j,k+1}^{*}} \tag{4.4.13}$$

$$\hat{\psi}_{j,k+1} = \hat{\psi}_{j,k} + \frac{R_{j,k}\, g_{j,k+1}\, (\hat{e}_{j,k+1} - g_{j,k+1}^{T}\, \hat{\psi}_{j,k})}{1 + g_{j,k+1}^{T}\, R_{j,k}\, g_{j,k+1}} \tag{4.4.14}$$

$$R_{j,k+1} = R_{j,k} - \frac{R_{j,k}\, g_{j,k+1}\, (R_{j,k}\, g_{j,k+1})^{T}}{1 + g_{j,k+1}^{T}\, R_{j,k}\, g_{j,k+1}} \tag{4.4.15}$$

where

$$a_{j,k+1}^{*T} = \left[-y_{1,k}^{*} - y_{2,k}^{*} \cdots - y_{r,k}^{*}\ u_{1,k}^{*} \cdots u_{m,k}^{*} \right] \tag{4.4.16}$$

$$g_{j,k+1}^{T} = \left[-\hat{e}_{j,k} - \hat{e}_{j,k-1} \cdots - \hat{e}_{j,k-s} \right] \tag{4.4.17}$$

and the initial conditions are:

$$P_{j,0} = R_{j,0} = \alpha I \qquad \alpha >> 1 \tag{4.4.18}$$

$$\hat{\phi}_{j,0} = \hat{\psi}_{j,0} = 0 \tag{4.4.19}$$

$$E = \beta I \qquad 0 < \beta << 1 \tag{4.4.20}$$

Sufficient numbers of terms have to be included in the residual autoregressive model such that an unbiased system parameter estimation is obtained. Estimation of the auxiliary parameters can be started after sufficient residuals have been generated by using an ordinary recursive least-squares algorithm.

Although the forementioned algorithm was primarily developed for single-input single-output systems and the convergence has not yet been proved, all experiments have

been successful. Stable and consistent estimates have been obtained in all cases of simulations.

In the modelling of the uncorrelated residuals, several independent random processes have been combined to form a single random process. A brief proof of this theorem has been included in the appendix for reference.

4.5 From Input-output to State-space Model

After the system has been identified in the input-output difference canonical form (equation (2.4.16)), an equivalent state-space model of the form (equation (2.4.3) - (2.4.6)) can be easily derived. Because of the special structure of this kind of state-space model, the resultant system is completely observable. The state-space model can be constructed by the following procedures.

Matrix C can be constructed from the knowledge of the structural indices $v_1, \ldots, v_r$. Matrix A can be obtained directly from $P(z)$ since the elements of A are the coefficients of the polynomials in $P(z)$ (equation (2.4.19) - (2.4.20)). To compute matrix B, transformation matrix M has to be constructed according to equation (2.4.23) after A is known. Matrix $\bar{B}$ is obtained from $Q(z)$ since the elements of $\bar{B}$ are the coefficients of the polynomials in $Q(z)$ (equation (2.6.21)). Matrix B is then, given by

$$B = M^{-1} \bar{B} \tag{4.5.1}$$

The resultant triangle (A, B, C) is in the

Luenberger's observable state-space canonical form and has the same external behaviour as the input-output canonical description.

CHAPTER 5

RESULTS OF SIMULATIONS

5.1 Introduction

In this chapter, algorithms described in Chapter 3 for structural identification and in Chapter 4 for parametric estimation have been applied to simulated examples. For the purposes of simplicity and generality, a fairly general class of two-input two-output systems as shown in Figure 5.1.1 has been investigated. Disturbances and measurement noises are lumped together and modelled as additive noises to the outputs of the system. Computer programs have been written to estimate the structural indices and the system parameters separately. The structural indices of the system are first determined off-line from simulated data and the system parameters are then, estimated subsequently using the same input-output sequences. The original multivariable system has been first decomposed into its subsystems and each of these subsystems has been identified separately. Recursive least-squares algorithm has been used successfully to give an accurate estimate in the case of low noise-corrupted measurements; while in the case of noisy measurements, recursive generalized least-squares algorithm has been

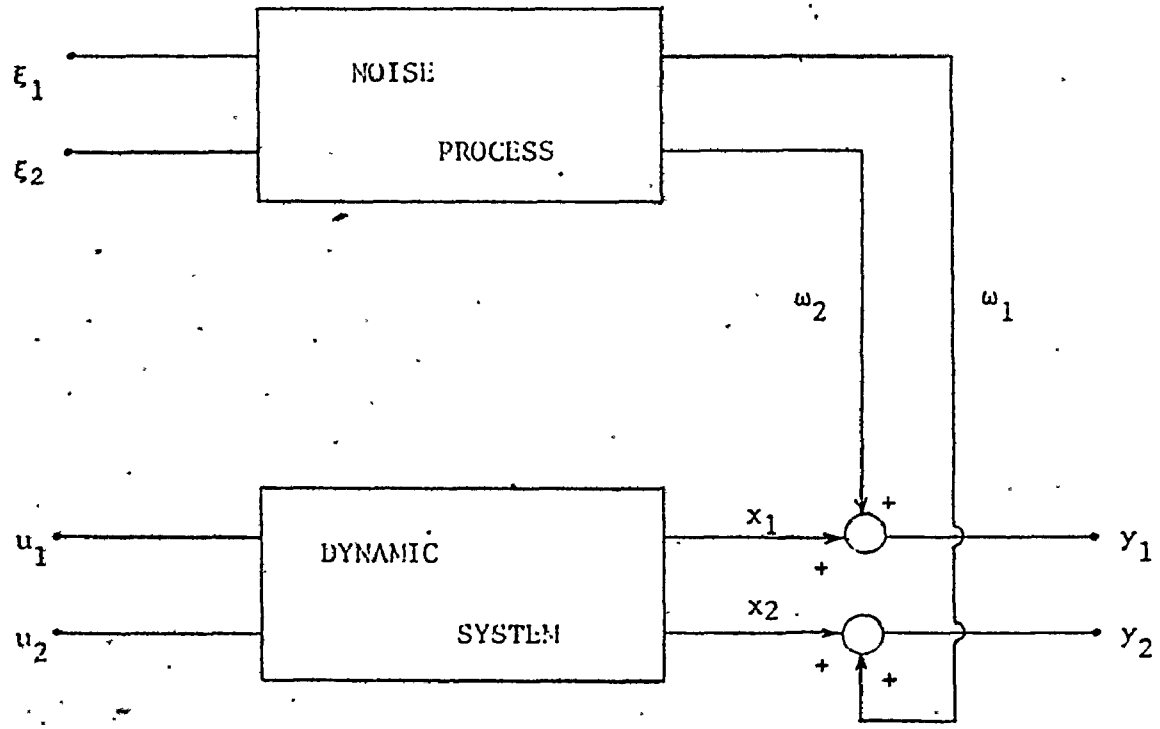

Figure 5.1.1 A general two-input two-output system.

adapted to give a consistent and unbiased estimate.

In the following sections, three examples have been included. Both structural and parametric estimation have been performed on the first two examples while the third example has been used exclusively to illustrate the usefulness of the parametric estimation algorithms.

5.2 Example 1--A Third Order System

The system used for simulation is described by the following transfer function matrix

$$H(z) = \begin{bmatrix} \frac{2}{z-0.5} & \frac{1}{z-0.25} \\ \frac{2}{z-0.25} & 0 \end{bmatrix}, \qquad (5.2.1)$$

Alternatively, the system can be expressed in the input-output difference canonical form

$$\begin{bmatrix} z^2-0.75z+0.125 & 0 \\ 0 & z-0.25 \end{bmatrix} X(k) = \begin{bmatrix} 2z-0.5 & z-0.5 \\ 2 & 0 \end{bmatrix} U(k) \qquad (5.2.2)$$

By inspection of equation (5.2.2), the structural indices of the two outputs of the system can be deduced immediately to be 2 and 1, respectively. Hence the order of the system is 3.

Due to imperfect measurements, the outputs of the system are corrupted by additive noise sequences such that the measured outputs of the system become

$$y_1(k) = X_1(k) + d_1\varepsilon(k) \qquad (5.2.3a)$$

$$y_2(k) = X_2(k) + d_2\varepsilon(k) \qquad (5.2.3b)$$

In order to observe all the modes of the system, uncorrelated unit-variance zero-mean white-noise sequences $\{U_1(k)\}$ and $\{U_2(k)\}$ have been used to excite the system. The additive output noise sequences $\{\xi_1(k)\}$ and $\{\xi_2(k)\}$ are coloured and are generated as the outputs of the following equations:

$$\xi_1(k) = 0.70\,\omega_1(k-1) + \omega_1(k) \qquad (5.2.4a)$$

$$\xi_2(k) = 0.85\,\xi_2(k-1) + \omega_2(k) \qquad (5.2.4b)$$

where $\{\omega_1(k)\}$ and $\{\omega_2(k)\}$ are unit-variance zero-mean sequences uncorrelated with $\{u_1(k)\}$ and $\{u_2(k)\}$, as well as with each other.

Addition of different levels of noise to the system has been accomplished by adjusting d_1 and d_2 of equation (5.2.3) for each case.

In order to reduce the amount of storage required in the structural identification, only 100 sets of input-output data have been used. In the absence of noise results of the structural identification are given in Figure 5.2.1. The structural indices of the outputs of the system have been determined, uniquely, to be $v_1 = 2$ and $v_2 = 1$. Results of structural identification of this example are tabled in Table 5.2.1 for different levels of noise. Experimental results indicate that no compensation to the product-moment matrices is necessary in the case of noise-free data. The

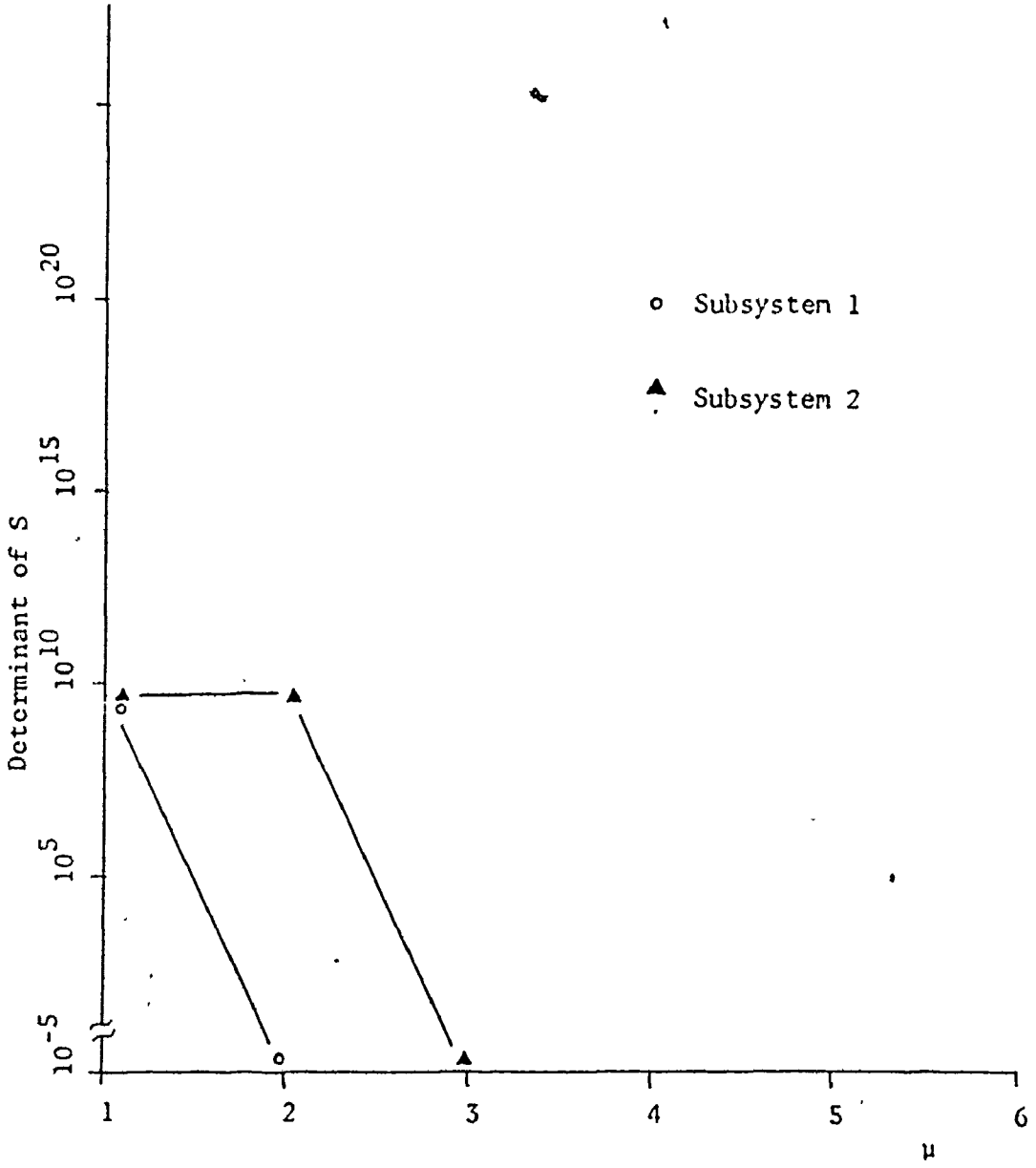

Figure 5.2.1 Results of structural identification of example 1 in the absence of noise.

Description	Noise level = 10%		Noise level = 20%	
	Output 1	Output 2	Output 1	Output 2
Under-compensation	3	2	3	-
Correct-compensation	2	1	2	1
Over-compensation	1	1	1	1

Table 5.2.1 Results of structural identification of example 1.

structural indices have been determined uniquely and consistently. As the levels of the additive noise increased, interdependences between the input-output sequences are deteriorated and difficulties have been encountered during the determination of the structural indices without any compensation to the product-moment matrices at all. The problem had been solved by including suitable amount of compensation to the product-moment matrices. Experience shows that an overestimate of the structural indices was obtained in the case of undercompensation, while an underestimate of the structural indices resulted in the case of overcompensation. Therefore, when the amount of additive noise to the system is not known in advance, compensation should be allowed to cover a wide predetermined range in the first run. In the subsequent runs, the range should be adjusted accordingly until the best compensation has been found.

After the structural indices have been determined, the order of the system as well as the number of parameters to be estimated in each subsystem can be calculated easily. For example, the structural indices of this example are 2 and 1, respectively. The order of the system will be 3 and the number of parameters to be estimated in each subsystem will be 7 and 5 accordingly.

Using 300 sets of input-output sequences, the

the parameters of each subsystem have been identified independently by the proposed recursive algorithms cited in section 4.3 and section 4.4. In the case of noise-free or very low noise data, the recursive least-squares algorithm has been used to estimate the parameters, while in the case of very noisy data, the recursive generalized least-squares algorithm has been adopted. Table 5.2.2 shows the final value of the parameters after 300 sets of input-output sequences for different levels of noise. As will be seen from Table 5.2.2, very accurate estimate of the parameters have been obtained in the case of noise-free data. Less accurate estimates have been obtained with the levels of the additive noises increased. In all cases, stable and consistent estimates have been obtained. Furthermore, the proposed algorithms have a very fast convergence rate even in the case of high noise. To show the convergence rate of the recursive generalized least-squares algorithm, the normalized error $(||\phi-\hat{\phi}||^2)/||\phi||^2$ of each subsystem has been plotted against the number of samples used. Fig. 5.2.2 and Figure 5.2.3 show the convergence rate for different levels of noise, respectively.

With properly defined state vectors, the Luenberger's canonical state-space form of this example can be constructed after the input-output difference canonical form has been identified. The following state-space equations can be easily derived:

Parameters \ Noise Level	True Value	Noise Free	y_1=12% y_2=25%	y_1=52% y_2=50%
$\alpha_{11,1}$	-0.7500	-0.7500	-0.6856	-0.6704
$\alpha_{11,2}$	0.1250	0.1250	0.1291	0.1298
$\alpha_{12,1}$	0	0	0.0060	-0.0415
$\beta_{2,1}$	2.0000	2.0000	2.0230	2.0090
$\beta_{1,1}$	-0.5000	-0.5000	-0.3399	-0.4622
$\beta_{4,1}$	1.0000	1.0000	0.9854	0.9377
$\beta_{3,1}$	-0.5000	-0.5000	-0.3886	-0.4389
$\alpha_{21,1}$	0	0	0.0797	0.0753
$\alpha_{21,2}$	0	0	-0.0259	-0.0355
$\alpha_{22,1}$	-0.2500	-0.2500	-0.3556	-0.3462
$\beta_{1,2}$	2.0000	2.0000	2.0130	2.0330
$\beta_{2,2}$	0	0	-0.0248	0.0107

Table 5.2.2 Results of parametric estimation of example 1.

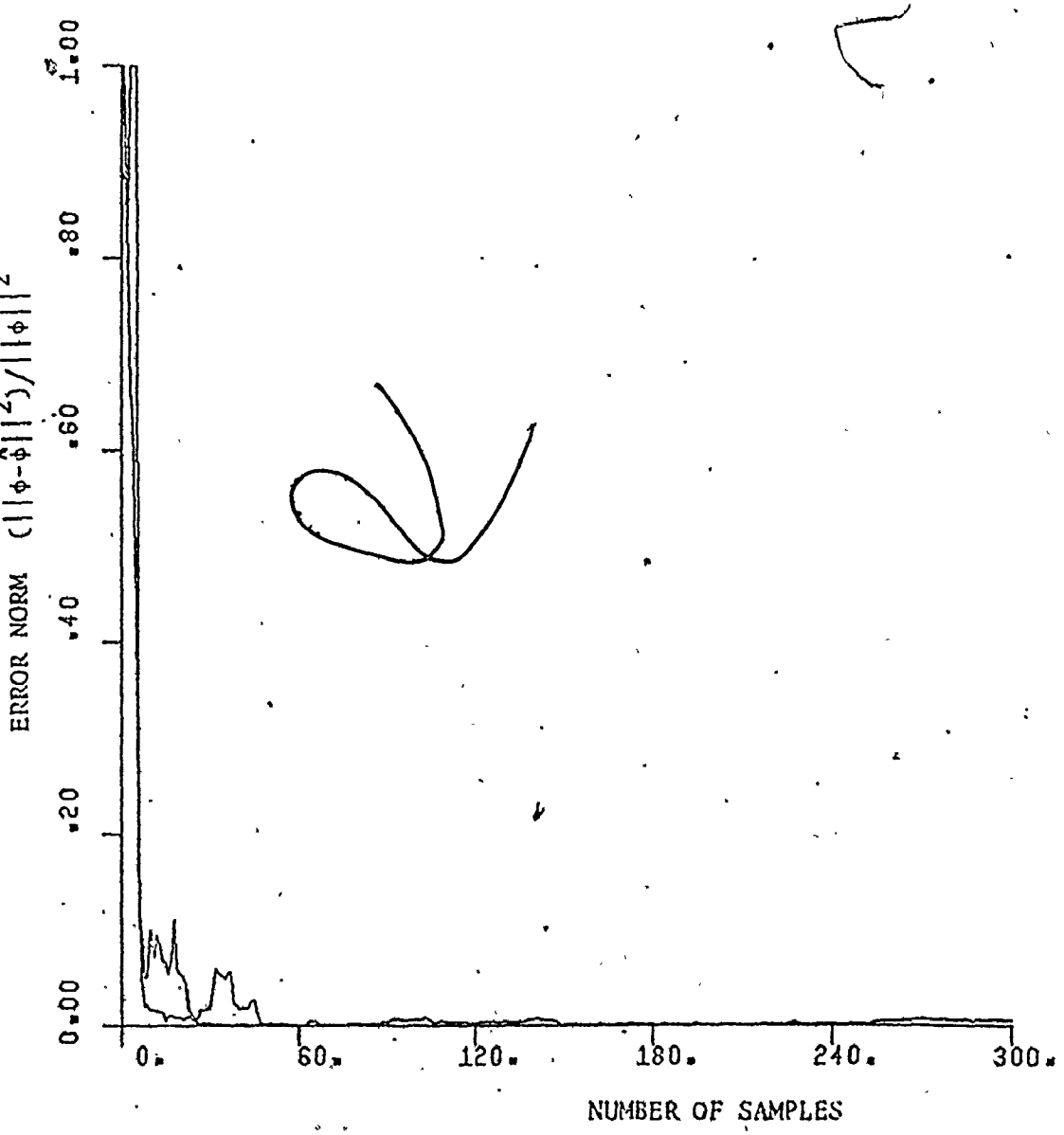

Figure 5.2.2 Convergence rate of example 1 with y_1=12%, y_2=25% noise levels.

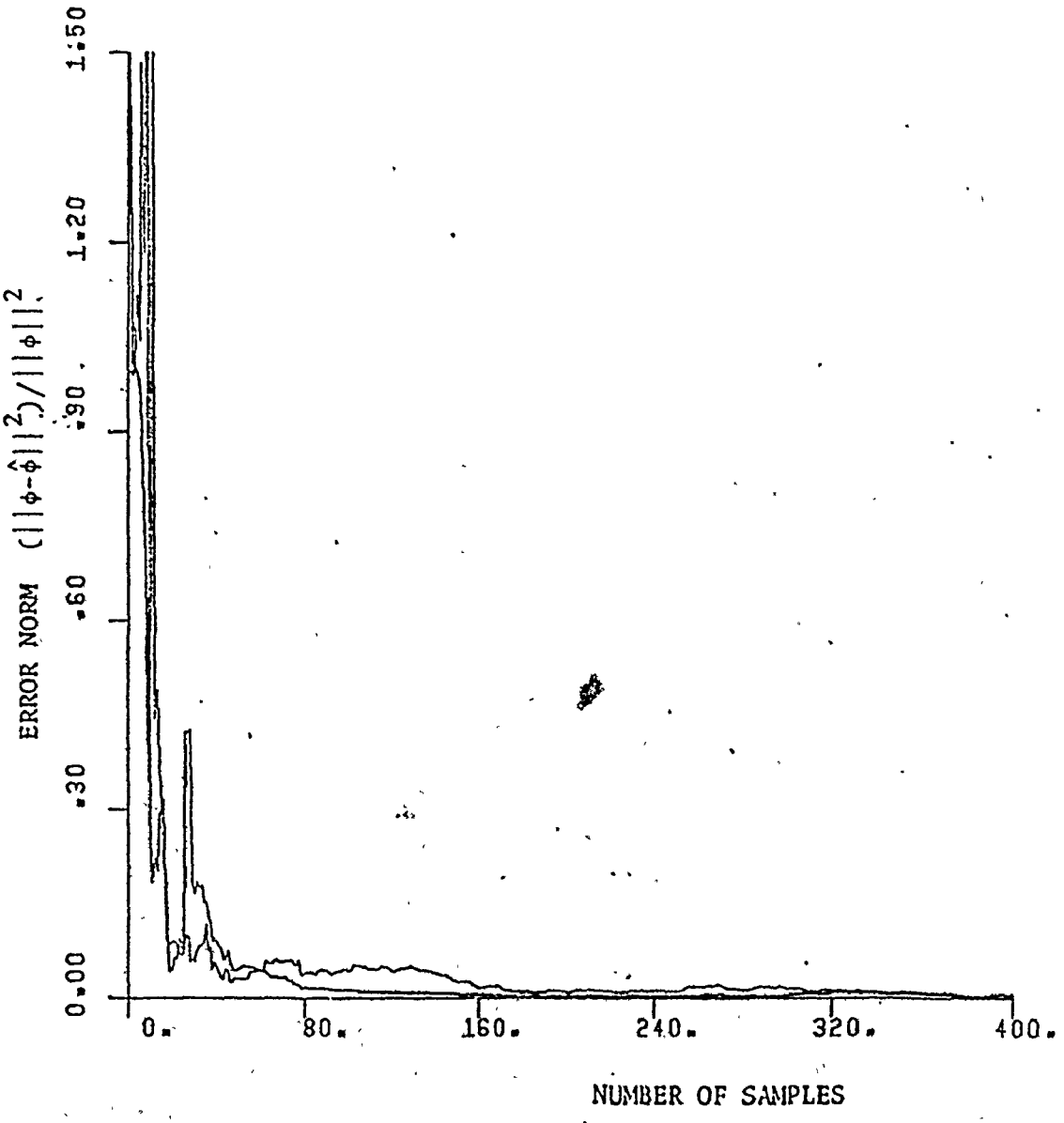

Figure 5.2.3 Convergence rate of example 1 with y_1=52%, y_2=50% noise levels.

$$X(k+1) = \begin{bmatrix} 0 & 1 & 0 \\ -0.125 & 0.75 & 0 \\ 0 & 0 & 0.25 \end{bmatrix} X(k) + \begin{bmatrix} 2 & 1 \\ 1 & 0.25 \\ 8 & 0 \end{bmatrix} U(k) \qquad (5.2.5a)$$

$$Y(k) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} X(k) \qquad (5.2.5b)$$

where

X(k) is the state vector,

U(k) is the input vector,

Y(k) is the output vector,

5.3 Example 2--A Fifth Order System

As another illustration of the suggested algorithm the following system with structural indices $v_1 = 2$ and $v_2 = 3$ has beem simulated on a digital computer:

$$H(z) = \begin{bmatrix} \frac{1}{z-0.5} & \frac{2}{(z-0.5)(z-0.25)} \\ \frac{2}{z-0.25} & \frac{4}{(z-0.5)^2} \end{bmatrix} \qquad (5.3.1)$$

After a few algebraical manipulations, the system can be transformed to the following input-output difference canonical form

$$\begin{bmatrix} z^2-0.75z+0.125 & 0 \\ 0 & z^3-1.25z^2+0.5z-0.0625 \end{bmatrix} X(k) = \begin{bmatrix} z-0.25 & 2 \\ 2z^2-2z+0.5 & 4z-1 \end{bmatrix} U(k) \qquad (5.3.2)$$

By inspection, the structural indices of the system are 2 and 3, respectively and the order of the system is 5. The number of parameters to be estimated in each subsystem can be calculated from equation (4.2.4) and is 9 and 11 respectively in this case. Excitation and noise sequences added to the system were the same as described in the previous example. Simulated input-output sequences have been generated using equation (5.3.2) and have been used for structural identification and parametric estimation.

In the presence of noise, the structural indices of the system have been identified to be 2 and 3 respectively as given in Figure 5.3.1. Results of structural identification are shown in Table 5.3.1 for different levels of noise. Once again, unique and consistent determination of the structural indices had been obtained by the proposed method. Results of parametric estimation after 400 sets of input-output sequences are shown in Table 5.3.2 for different levels of noise. Rates of convergence for the two subsystems are shown in Figure 5.3.2 and Figure 5.3.3 respectively for different levels of noise. Experimental results indicate that in the presence of noise more data is required for the algorithm to converge to the neighbourhood of the true values of the system parameters in this example. This is due to the fact that more parameters have to be estimated in this case.

The Luenberger's observable canonical state-space form for this example is

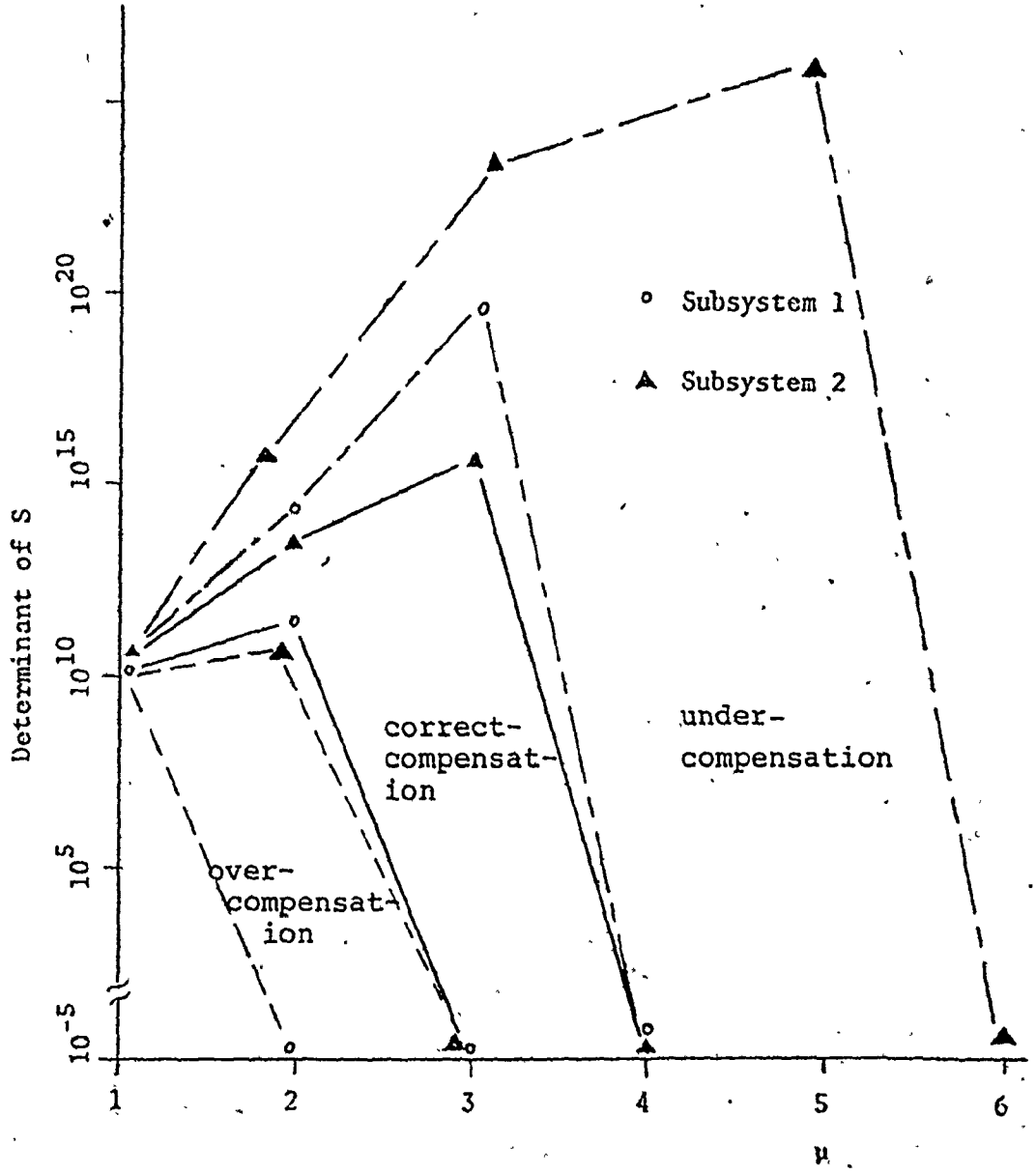

Fig. 5.3.1 Structural identification of example 2 in the presence of 20% noise.

Description	Noise level = 10%		Noise level = 20%	
	Output 1	Output 2	Output 1	Output 2
Under-compensation	3	4	3	5
Correct-compensation	2	3	2.	3
Over-compensation	1	2	1	2

Table 5.3.1 Results of structural identification of example 2.

Noise Level / Parameters	True Value	Noise Free	y_1 = 42.5% y_2 = 29.2%	y_1 = 85.0% y_2 = 58.4%
$\alpha_{11,1}$	-0.7500	-0.7500	-0.7194	-0.7322
$\alpha_{11,2}$	0.1250	0.1250	0.1871	0.2258
$\alpha_{12,1}$	0	0	-0.0557	-0.0328
$\alpha_{12,2}$	0	0	-0.0131	-0.0004
$\alpha_{12,3}$	0	0	-0.0543	-0.1304
$\beta_{2,1}$	1.0000	1.0000	1.0110	1.0010
$\beta_{1,1}$	-0.2500	-0.2500	-0.1864	-0.1125
$\beta_{4,1}$	0	0	-0.0297	-0.0921
$\beta_{3,1}$	2.0000	2.0000	2.0240	2.0220
$\alpha_{21,1}$	0	0	-0.0284	-0.0429
$\alpha_{21,2}$	0	0	0.0773	0.0856
$\alpha_{22,1}$	-1.2500	-1.2500	-1.1280	-1.0590
$\alpha_{22,2}$	0.5000	0.5000	0.3781	0.3141
$\alpha_{22,3}$	-0.0625	-0.0625	-0.0380	-0.0453
$\beta_{3,2}$	2.0000	2.0000	1.9960	2.0440
$\beta_{2,2}$	-2.0000	-2.0000	-1.7640	-1.5510
$\beta_{1,2}$	0.5000	0.5000	0.3581	0.3088
$\beta_{6,2}$	0	0	0.0478	0.1775
$\beta_{5,2}$	4.0000	4.0000	4.0430	4.1710
$\beta_{4,2}$	-1.0000	-1.0000	-0.7385	-0.6225

Table 5.3.2 Results of parametric estimation of example 2.

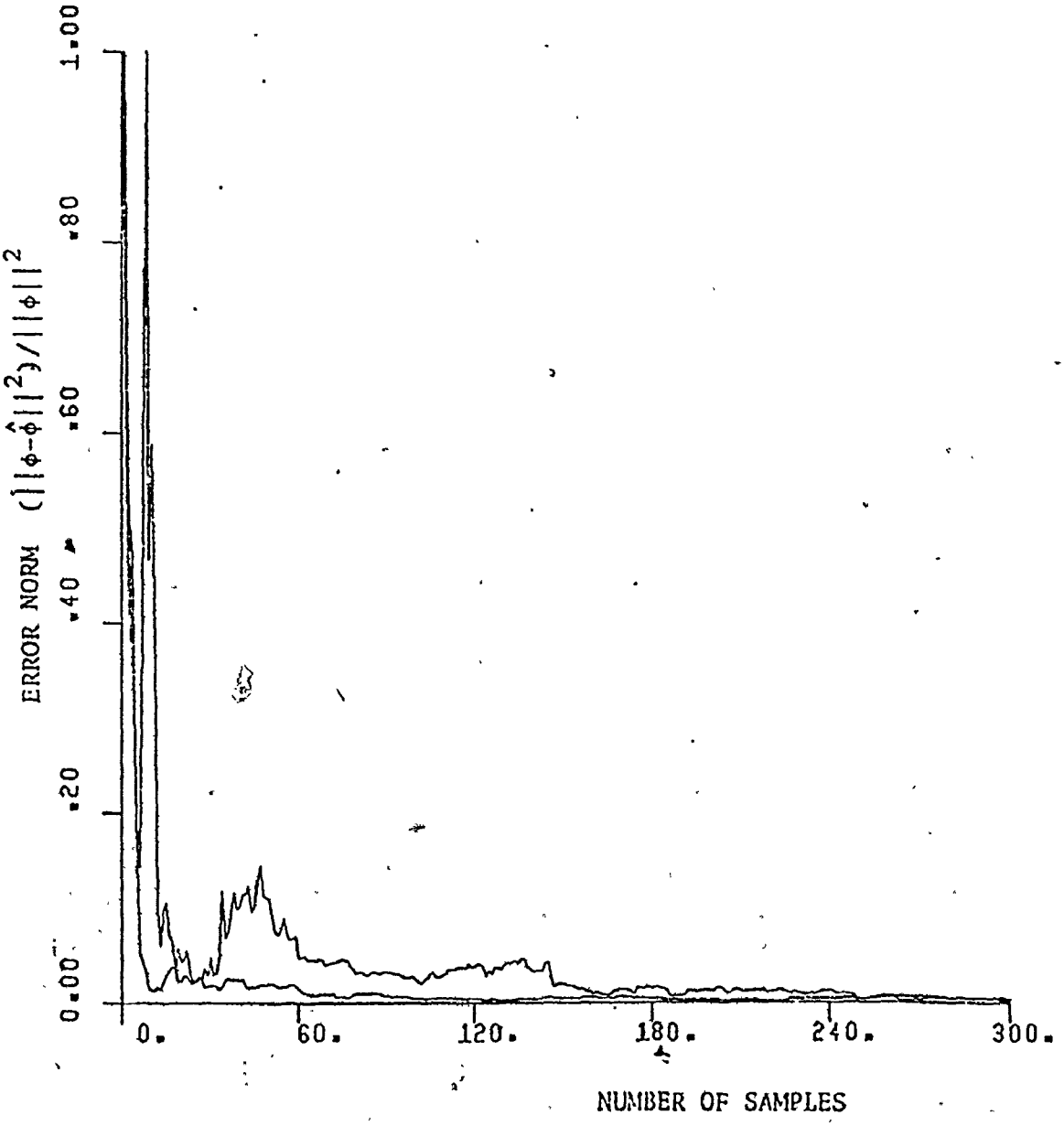

Figure 5.3.2 Convergence rate of example 2 with y_1=42.5%, y_2=29.2% noise levels.

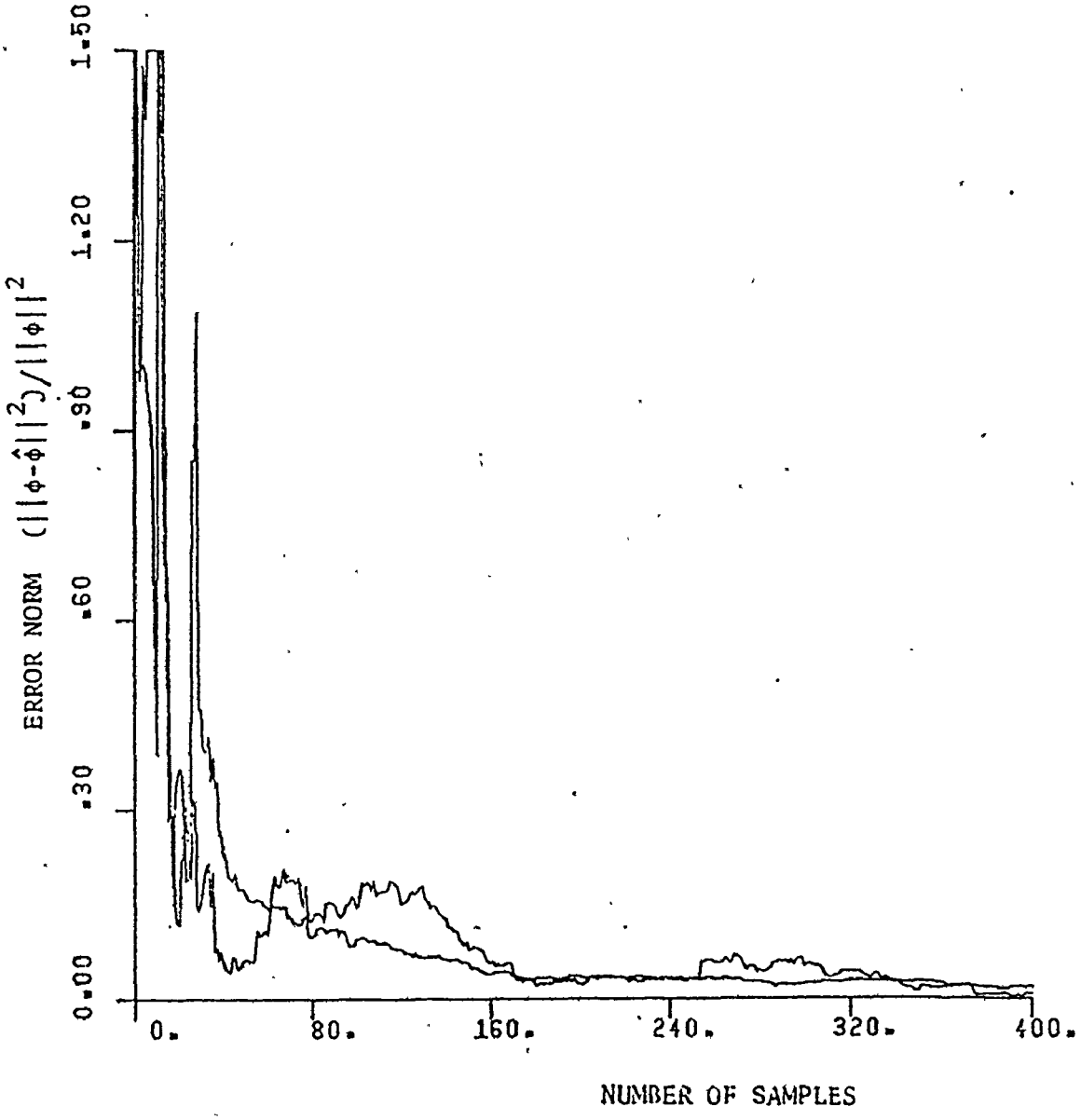

Figure 5.3.3 Convergence rate of example 2 with y_1=85%, y_2=58.4% noise levels.

$$X(k+1) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ -0.125 & 0.75 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0.0625 & -0.5 & 1.25 \end{bmatrix} X(k) + \begin{bmatrix} 1 & 0 \\ 0.5 & 2 \\ 2 & 0 \\ 0.5 & 4 \\ 0.125 & 4 \end{bmatrix} U(k) \qquad (5.3.3a)$$

$$Y(k) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} X(k) \qquad (5.3.3b)$$

where X(k), U(k), Y(k) are the same as defined in last section.

5.4 Example 3--A Fourth Order System

To further the testing of the parameter estimation algorithm, the following system has been considered.

$$\begin{bmatrix} z^2-0.75z+0.125 & z-0.25 \\ 0.5z-0.6 & z^2-z+0.24 \end{bmatrix} X(k) = \begin{bmatrix} z-0.2 & 0.1 \\ 0.3 & z-0.5 \end{bmatrix} U(k) \qquad (5.4.1)$$

Noise is added to the digitally simulated input-output data and the noise was generated as the output of the following system, with white noise excitation.

$$\begin{bmatrix} \xi_1(k) \\ \xi_2(k) \end{bmatrix} = \begin{bmatrix} 0.70 & 0 \\ 0 & 0.80 \end{bmatrix} \begin{bmatrix} \xi_1(k-1) \\ \xi_2(k-1) \end{bmatrix} + \begin{bmatrix} \omega_1(k) \\ \omega_2(k) \end{bmatrix} \qquad (5.4.2)$$

Parametric estimation has been performed for different cases of noise levels. Fxperimental results are shown in Table 5.4.1. It is evident that unbiased and consistent estimates have been obtained in all cases. Figure 5.4.1 and Figure 5.4.2 show the convergence rates of this example with different levels of noise.

The Luenberger's observable canonical form of this

Noise Level / Parameters	True Value	Noise Free	y_1=9.8% y_2=10.2%	y_1=29.4% y_2=30.6%
$\alpha_{11,1}$	-0.7500	-0.7500	-0.7257	-0.7004
$\alpha_{11,2}$	0.1250	0.1250	0.1088	0.0946
$\alpha_{12,1}$	0.5000	0.5000	0.4704	0.3492
$\alpha_{12,2}$	-0.2500	-0.2500	-0.2177	-0.1310
$\beta_{2,1}$	1.0000	1.0000	0.9962	0.9751
$\beta_{1,1}$	-0.2000	-0.2000	-0.1790	-0.1556
$\beta_{4,1}$	0	0	0.0009	-0.0091
$\beta_{3,1}$	0.1000	0.1000	0.0593	-0.0643
$\alpha_{21,1}$	0.5000	0.5000	0.3916	0.1759
$\alpha_{21,2}$	-0.6000	-0.6000	-0.5394	-0.4035
$\alpha_{22,1}$	-1.0000	-1.0000	-1.0220	-1.0200
$\alpha_{22,2}$	0.2400	0.2400	0.2278	0.1736
$\beta_{2,2}$	0	0	0.0009	0.0096
$\beta_{1,2}$	0.3000	0.3000	0.1983	-0.0180
$\beta_{4,2}$	1.0000	1.0000	0.9961	0.9851
$\beta_{3,2}$	-0.5000	-0.5000	-0.5159	-0.4987

Table 5.4.1 Results of parametric estimation of example 3.

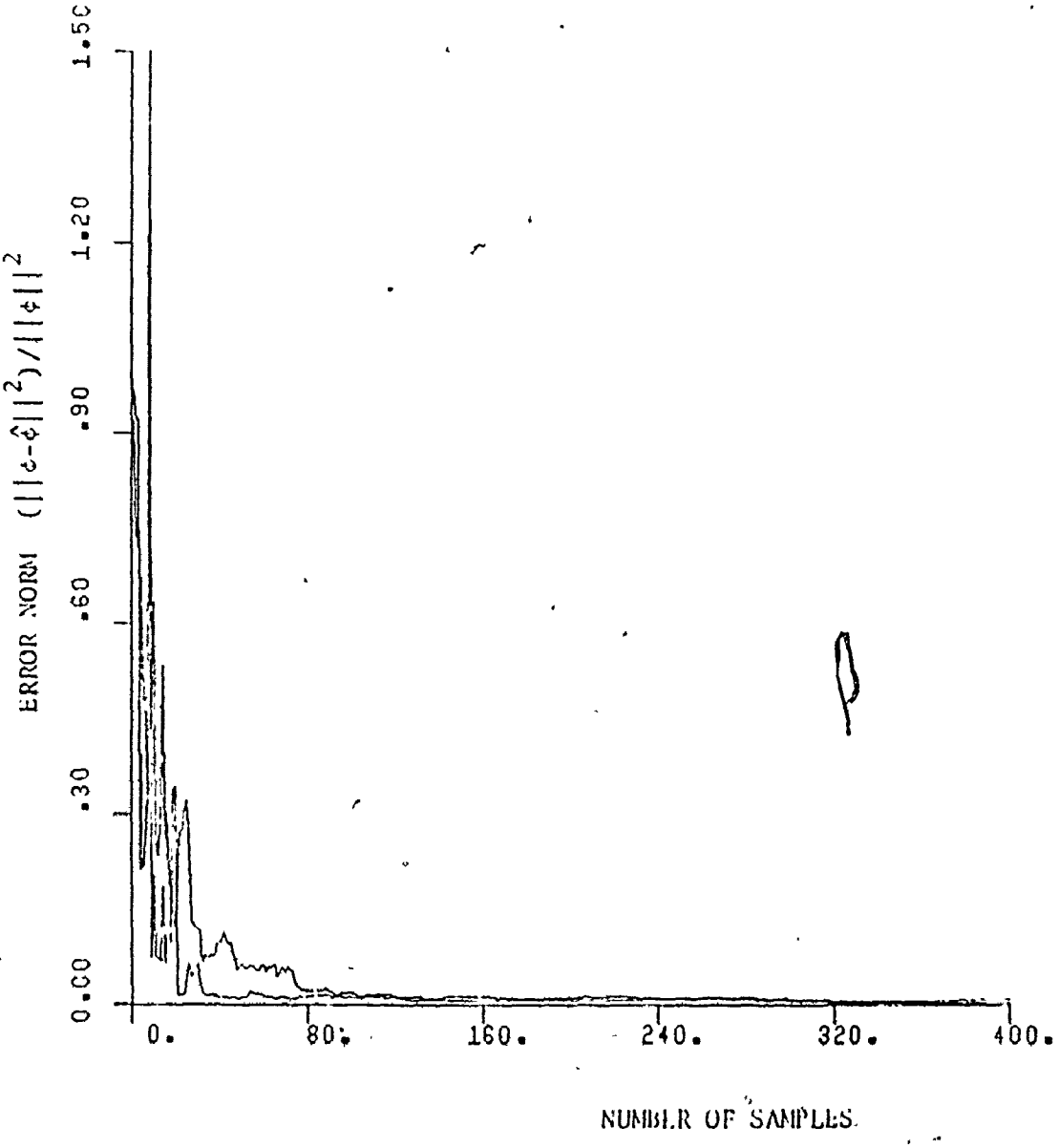

Figure 5.4.1 Convergence rate of example 3 with y_1=9.8%, y_2=10.2% noise levels.

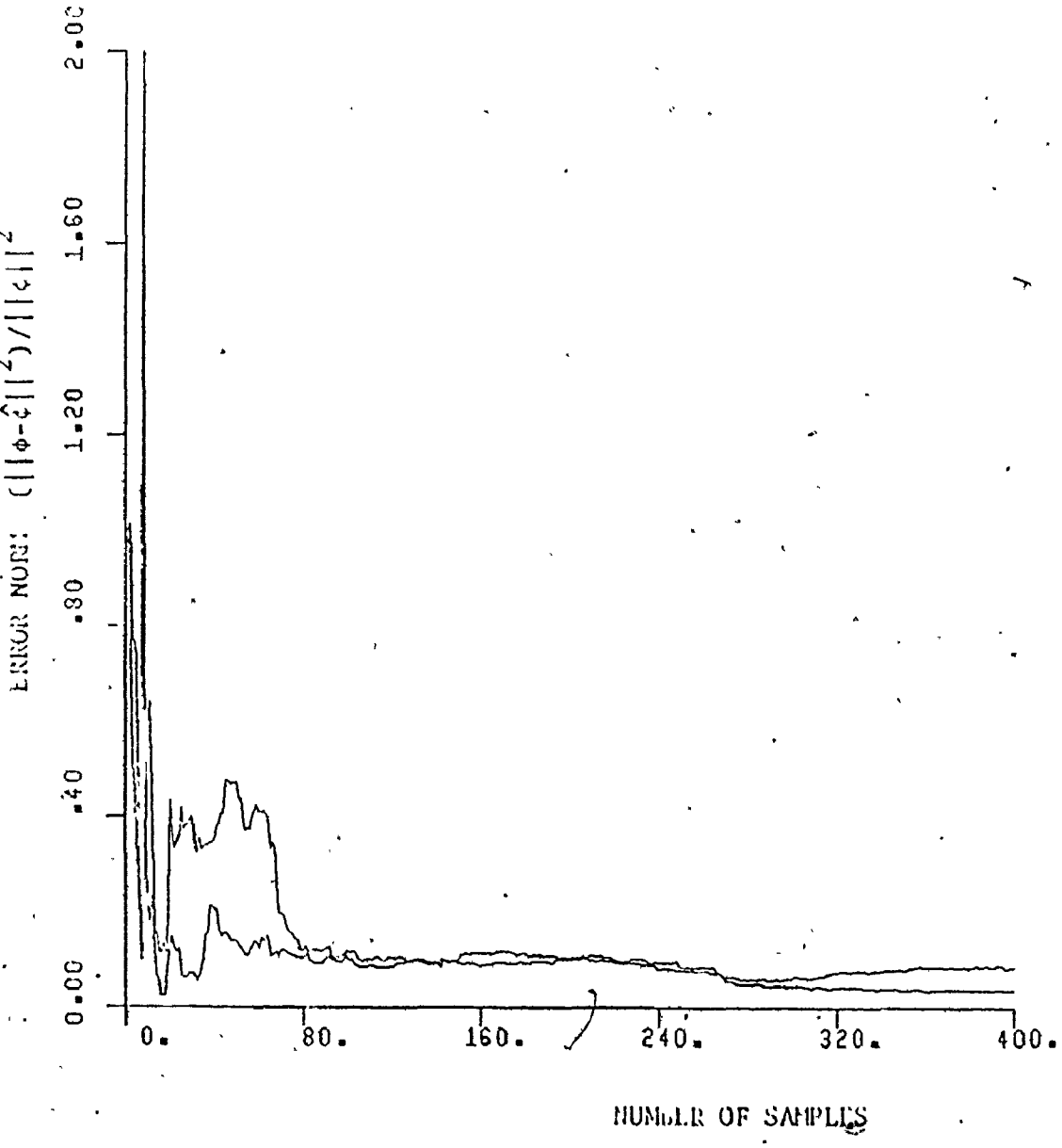

Figure 5.4.2 Convergence rate of example 3 with y_1=29.4%, y_2=30.6% noise levels.

example is

$$X(k+1) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ -0.125 & 0.75 & 0.25 & -1 \\ 0 & 0 & 0 & 1 \\ 0.6 & -0.5 & -0.24 & 1.0 \end{bmatrix} X(k) + \begin{bmatrix} 1 & 0 \\ 0.55 & -0.9 \\ 0 & 1 \\ 0.2 & 0.5 \end{bmatrix} U(k) \qquad (5.4.3a)$$

$$Y(k) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} X(k) \qquad (5.4.3b)$$

Where X(k) is the state vector at the k-th sampling instant,

U(k) is the input vector at the k-th sampling instant,

Y(k) is the output vector at the k-th sampling instant.

5.5 Summary

As will be seen from the results of these simulated examples, stable and consistent estimates of the structural indices and the system parameters have been obtained by the proposed method. In the case of low noise data, nearly exact estimates of the structural and system parameters have been obtained. Fast convergence rate for parametric estimation has been attained by using the proposed recursive generalized least-squares algorithm. In the case of high noise, the proposed method also gives unbiased estimates with a fast rate of convergence.

It is fair to note that, as far as parametric estimation is concerned, the suggested algorithm seems to work better with systems without output interaction (examples 1 and 2, where the off diagonal elements of the P(z) matrices are zero). For other general systems (example 3), the algorithm works up to certain critical limit of noise to signal ratio.

This phenomenon has not been completely understood and will be under further investigation.

CHAPTER 6

CONCLUSIONS

In this thesis a method to solve the identification problem for general multivaraible linear discrete-time dynamic systems has been suggested. A unified approach has been adopted for estimating the system structure as well as the system parameters. The suggested identification procedure starts from normal operating records which may contain additive zero-mean measurement noise. A systematic and computationally efficient algorithm has been suggested for estimating the system structure off-line, while system parameters have been estimated by recursive algorithms.

Linear discrete-time multivariable systems have been described in the input-output difference canonical form. This canonical form has a unique equivalence with the block-companion state-space representation and is completely observable and controllable. Systems described in this canonical form have the unique advantage of being conveniently decomposed into subsystems. Each of these subsystems is observable and can, therefore, be identified independently.

A statistical algorithm has been developed to determine a set of structural indices, which completely define the

structure as well as the order of the system, directly from input-output sequences such that intermediate construction of a parametric model is avoided. This structural identification algorithm is remarkable efficient because a well-defined model among many possible structure is selected. Trial and error, based on the optimality test of an assumed order or structure, which requires the parametric estimation of a large number of models, is avoided.

The problem of parametric estimation has been solved by identifying each of the subsystems, which is a decomposition of the main system, independently. A recursive least-squares algorithm has been used to estimate the system parameters in the case of low noise data, and a recursive generalized least-squares algorithm has been adapted in the case of noisy data.

Realistic digital computer simulated noise-free and noise-corrupted input-output sequences have been generated and used to test the merits of the proposed algorithm. The results of experiments indicate that in the absence of noise nearly exact estimates have been obtained for structural identification and parametric estimation. In noisy environment, the problem of structural identification has been resolved by including a suitable amount of compensation to the algorithm. The chief purpose of this compensation is to offset the disturbances produced by the noise. To reflect the efficiency of the algorithm the computer time spent on the structural identification of example 2 in a noisy environment has been recorded. Using a

CDC 6400 machine the computational time, which included the time spent on input-output handling, was just under 15 seconds. As far as parametric estimation is concerned, the proposed recursive generalized least-squares algorithm gives consistent and unbiased estimates in the presence of noise, and the algorithm is computationally efficient. Matrix inversion is avoided and data storage is saved by using the recursive algorithm. The number of arithmetic operations carried out in each iteration has been kept to a minimum and the estimate has been updated sequentially. Because of the simplicity of this algorithm it can be easily implemented on a mini-computer and used for on-line applications. The actual computational time for each iteration is highly dependent on the order of the parameter vector as well as the particular machine used. For example, the computational time for each iteration of example 2 was determined to be less than 0.05 second on a CDC 6400 machine.

Although a rigorous proof of the convergence properties of the proposed algorithm has not been included, all experiments have been successful. Stable and consistent estimates have been obtained.

Further work in this area can be continued in many ways. For example, a rigorous proof of the convergence properties of the proposed algorithm has yet to be established. Furthermore, the algorithm needs more tests to establish its general usefulness in realistic situations. It may be pointed out, however, that the noise model used in parametric estimation

of this work may be inadequate or unrealistic in some practical applications, because only additive noise has been considered. Disturbances and measurement noise have been lumped together and modelled as additive noise to the output of the system. Interaction in the noise at different inputs and outputs has been ignored. Perhaps it would be more desirable to consider an interactive noise model, as proposed by Meerov, Pajziev and Rajbman (1969). This is a topic for future investigation.

APPENDIX

In this appendix, the sum of several independent moving average processes has been shown to be another moving average process (Box and Jenkins, 1970).

consider a stochastic process w(k), which is the sum of n independent moving average processes of orders $q_1, q_2, \ldots, q_n$ respectively. That is

$$w(k) = \sum_{i=1}^{n} \theta_i(z) a_i(k) \qquad \text{(A-1)}$$

where $\theta_i(z)$, i=1, . . . , n, are polynomials in z, z^{-1} is the unit delay operator, and $a_i(k)$, i=1, . . . , n, are mutually independent, zero-mean white noise processes; k is the sampling instant.

Suppose $q = \max(q_1, q_2, \ldots, q_n)$; then it is clear that the autocovariance function γ_j for w(k) must be zero for j> q. It follows that there exists a representation of w(k) as a single moving average process of order q.

$$w(k) = \theta(z) u(k) \qquad \text{(A-2)}$$

where $\theta(z)$ is a polynomial in z and u(k) is a white noise process with zero mean.

Thus, the sum of several independent moving average processes is another moving average process, whose order is the same as that of the component process of highest order.

BIBLIOGRAPHY

(A1) Ackermann, J.E. and Bucy, R.S. (1971), "Canonical minimal realization of a matrix of impulse response systems." Inform. Contr. 19, pp. 224-231.

(A2) Anderson, B.D.O. and Jury, E.I. (1976), "Generalized Bezoutain and Sylvester Matrices in Multivariable Linear Control." IEEE Trans. Automat. Contr. Vol. AC-21, pp. 551-556.

(A3) Astrom, K.J. and Eykoff, R. (1970), "System identification--a survey." Proc. IFAC Symp. on Identification and Process Parameter Estimation, Prague.

(B1) Bonivento, C., Guidorzi R. and Marro, G., (1973), "Irreducible canonical realizations from external data sequences." Int. J. Control, Vol. 17 No. 3, pp. 553-563.

(B2) Box, G.E.P. and Jenkins, G.M. (1970), Time Series Analysis--forecasting and control, Holden-Day, San Francisco.

(B3) Budin, M.A. (1971), "Minimal realization of discrete linear systems from input-output observations." IEEE Trans. Automat. Contr. Vol. AC-16, pp. 395-401.

(C1) Chen, C.T. (1970), Introduction to Linear System Theory, Holt Rinehart Winston, Inc., New York.

The Hague, pp. 1061-1082.

(L1) Luenberger, D.G. (1967), "Canonical forms for linear multivariable systems," IEEE Trans. Automat. Contr., Vol. AC-12, pp. 290-293.

(M1) Mayne, D.Q. (1972), "Parametrization and identification of linear multivariable systems," Stability of Stochastic Dynamical Systems 294, Lecture Notes in Mathematics, Springer-Verlag.

(M2) Meerov, M.V., Pajziev, E. and Rajbman, N.S. (1969), "Identification of stationary multivariable plants with the aid of partial correlation and variance functions," Automation and Remote Control, pp. 691-696.

(M3) Mehra, M.K. (1971), "On-line identification of linear dynamic systems with applications to Kalman filtering," IEEE Trans, Automat. Contr., Vol. AC-16, No.1, pp. 12-21.

(P1) Popov, V. (1972), "Invariant description of linear time-invariant controllable systems." SIAM J. Contr., Vol. 10, pp. 252-264.

(R1) Rissanen, J. (1974), "Basis of invariants and canonical forms of linear dynamic systems." Automatica, Vol. 10, pp. 361-374.

(R2) Rosenbrock, H. (1970), State-space and Multivariable Theory, John Wiley and Sons, Inc., New York.

(R3) Rowe, I.H. (1970), "A bootstrap method for the statistical estimation of model parameters," Int. J. of Control,

Vol. 12, pp. 721-738.

(S1) Sage, A.P. and Melsa J.L. (1971), System Identification, Academic Press, Inc., New York.

(S2) Saridis, G.N. and Stein, G. (1968), "Stochastic approximation algorithms for linear system identification," IEEE Trans. Automat. Contr., Vol. AC-13, pp. 512-523.

(S3) Sen, A. and Sinha, N.K. (1975), "A generalized pseudo-inverse algorithm for unbiased parameter estimation," Int. J. of Systems Science, Vol. 6, No. 12, pp. 1103-1109.

(S4) Sen, A. and Sinha, N.K. (1976), "On-line estimation of the parameters of a multivariable system using matrix pseudoinverse," Int. J. of Systems Science, Vol. 7, No. 4, pp. 461-471.

(S5) Sinha, N.K. and Pille, W. (1971), "Online parameter estimation using matrix pseudoinverse," Proc. IEE, Vol. 118, No.8, pp. 1041-1046.

(S6) Sinha, N.K. and Rozsa, P. (1976), "Some canonical forms for linear multivariable systems," Int. J. of Control, Vol. 23, No. 6, pp. 865-883.

(S7) Sinha, N.K. and Sen, A. (1974), "Comparison of some on-line identification methods for a simulated first order process," Automatic Control Theory and Applications, Vol. 2, No. 2, pp. 37-43.

(S8) Soderstrom, T. (1972), "Convergence properties of the generalized least squares identification method," Report 7228, Division of Automatic Control, Lund Institute of

Tech.

(T1) Tse, E. and Weinert, H.L. (1975), "Structure determination and parameter identification for multivariable stochastic linear systems," IEEE Trans. Automat. Contr., Vol. AC-20, No. 5, pp. 603-613.

(W1) Weinert, H. and Anton, J. (1972), "Canonical forms for multivariable system identification," Proc. IEEE Conf. on Decision and Control and 11th Symp. on Adaptive processes New Orleans, pp. 37-39.

(W2) Wolovich, W.A. (1974), Linear Multivariable Systems, Springer-Verlag, New York.

(W3) Wolovich, W.A. and Falb, P.L. (1969), "On the structure of multivariable systems," SIAM J. Control. Vol. 7, No. 3, pp. 437-451.

(W4) Wong, K.Y., and Polak, E. (1967), "Identification of linear discrete-time systems using the instrumental variable method," IEEE Trans. on Automat. Contr. Vol. AC-12, pp. 707-718.

(W5) Woodside, C.M. (1971), "Estimation of the order of linear systems," Automatica, Vol. 7, pp. 727-733.

(Y1) Young, P.C. (1969), "An instrumental variable method for real-time identification of noisy processes," Proc. IFAC Congress, Warsaw.