IDENTIFICATION OF LINEAR MULTIVARIABLE DISCRETE-TIME SYSTEMS

by

HOSSNY E. EL-SHERIEF

A Thesis
Submitted to the School of Graduate Studies
in Partial Fulfilment of the Requirements
for the Degree
Doctor of Philosophy

McMaster University

May 1979

IDENTIFICATION OF LINEAR MULTIVARIABLE

DISCRETE-TIME SYSTEMS

DOCTOR OF PHILOSOPHY (1979)
(Electrical Engineering)

McMASTER UNIVERSITY
Hamilton, Ontario

TITLE: Identification of Linear Multivariable Discrete-Time Systems

AUTHOR: Hossny E. El-Sherief, B.Sc. (Elect. Eng., Cairo Univ.)
B.Sc. (Math., Ain Shams Univ.)
M.Eng. (Elect. Eng., McMaster Univ.)

SUPERVISOR: Naresh K. Sinha, Professor of Electrical Engineering
B.Sc. Eng. (Banaras)
Ph.D. (Manchester)

NUMBER OF PAGES: x, 152

ABSTRACT

The problem of on-line identification of linear multivariable discrete-time systems from input-output data is considered. A study has been made of the relative effectiveness of the four different models used in the area of identification of linear multivariable systems (transfer-function matrix, impulse response matrix, input-output difference equation and state space). The features of each model and its effect on the complexity of the identification algorithm as well as the bias of the parameter estimates while using the ordinary least-squares method have been studied. Different on-line algorithms have been proposed for the identification of the given system directly in each of the four different model representations. These algorithms estimate the parameters of the system from noisy measurements and no knowledge of the noise characteristics is required. The identification of a given multivariable system has been decomposed into the identification of m subsystems (where m is the number of outputs) and the parameters of each subsystem are estimated independently from each other. The problem of structure determination has been considered, and algorithms have been proposed for the estimation of the structural parameters of the transfer-function matrix and the state space representations from noise-free as well as noisy measurements. Also, a two-stage bootstrap algorithm has been derived for combined parameter and state estimation of linear multivariable systems.

ACKNOWLEDGEMENTS

The author wishes to express his sincere appreciation to Dr. N.K. Sinha for expert guidance and supervision throughout the course of this work. He also thanks Dr. J.W. Bandler, Dr. I.Z. Chorneyko and Dr. J.F. MacGregor, members of his supervisory committee, for their continuing interest.

It is the author's pleasure to acknowledge useful discussions with other colleagues at McMaster. Without being exhaustive they are S. Azim, Dr. J. Hickin, Z. El-Razaz, Dr. H.L. Abdel-Malek and M.R. Rizk.

Thanks are due to Miss P. Dillon (Word Processing Centre) for her cheerful and excellent typing of the manuscript.

The author especially wishes to express his appreciation for the encouragement and help from his wife, Hanaa.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction to the Problem

The subject of system identification has received much attention recently because of its importance 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. During the past decade several papers, books [1-6] and survey papers [7-23] have been published on the subject. Most of this work, however, deals with the problem of identification of single-input single-output systems and relatively little has been done on the identification of multivariable systems. The problem of identification of multivariable systems from input-output data is more complicated and involves several factors. This problem was first considered by Gopinath [24] and later by Budin [25]. In their work they assumed that the available data for identification is free from noise. In general, most of the practical systems have considerable measurement noise, hence a realistic identification algorithm should take into consideration the noise factor and assumes that the available data for identification is contaminated with noise.

Another factor in the identification problem is the choice of the model, and this is generally determined by the applications for which it will be used. Generally, every model has some effects on the properties of the identification algorithm, e.g. it affects the number of system parameters to be identified and the unbiasedness of the parameter estimates when the ordinary least-squares method is used. In general, there are four types of system models which have been studied and used in the area of identification of linear multivariable systems. These models are: the transfer-function matrix, the impulse response matrix, the input-output difference equation and the state space formulation. These four models are equivalent and transformations between them are possible.

The system identification problem can be defined according to Zadeh [26] as "the determination, on the basis of input and output, of a system model within a specified class of models to which the system under test is equivalent". Basically, the problem of system identification consists of two main steps: structure determination and parameter estimation. In the first step the structural parameters of the system have to be determined from the given input-output data. This step is difficult for the case of noisy data and depends on the type of model used for identification as every model has its own structural parameters. After determining or assuming the structural parameters, the model parameters can be estimated from simulated or normal operating records with a suitable estimation algorithm.

Computationally, identification algorithms can be divided into two main categories, namely off-line algorithms and on-line algorithms.

An off-line algorithm is a one-shot technique which normally requires a considerable amount of storage of the input-output data. On-line algorithm, on the other hand, employs an iterative scheme whereby the estimates of the parameters of the model are continuously updated as new measurements are made. For any identification algorithm to be of practical value in control applications it should be of a recursive nature. This idea has led to considerable work on the subject of on-line system identification.

1.2 Contents and Organization of the Thesis

The major effort in this thesis is directed towards the problem of identification of linear time-invariant discrete-time multivariable systems from input-output data. On-line algorithms have been developed for estimating the parameters of the four different types of model representations of linear multivariable systems from noise-free as well as noisy data. Also the problem of determination of the structural parameters has been considered.

A survey [27] of most of the existing literature on the problem of identification of linear discrete-time multivariable systems from input-output data has been included in the thesis, where the identification algorithms have been classified according to the type of model representation used.

In Chapter 2 the effect of model structure on the properties of the identification algorithm has been investigated. A study has been made of the relative effectiveness of the four models used for identification of linear discrete-time multivariable systems. The features of each model and its effect on the complexity of the identification algorithm as well as the bias of the parameter estimates while using the ordinary least-squares method have been studied.

Chapter 3 discusses the problem of identification of linear multivariable systems in the transfer-function matrix representation. An algorithm is proposed for estimating the parameters of a special transfer-function matrix representation from noise-free data. This algorithm is simpler than that of Sen and Sinha [28] with considerable reduction in computation time. Another algorithm has been developed for identifying a more general transfer-function matrix representation from noise-free as well as noisy data assuming the order of the system is unknown. This algorithm is based on determining the order of each row of the transfer-function matrix utilizing the residual error technique. This is followed by estimation of the parameters using a recursive adaptive least-squares algorithm.

In Chapter 4 the impulse response model has been considered for identification and two algorithms have been developed to estimate the Markov parameters of the system. The first algorithm uses the normalized stochastic approximation method which has been developed in [29] for estimating the Markov parameters of the system. A proof of the convergence and unbiasedness of the stochastic approximation algorithm has been obtained and is given in Appendix II. The second algorithm uses correlation techniques which eliminate the bias obtained in the first algorithm due to truncation of the infinite series of Markov parameters.

In Chapter 5 the state space representation has been considered for identification. First an algorithm has been developed for determining the structural parameters of a row-companion canonical form (the observability subindices) from noise-free as well as noisy data using the residual error technique. Then an algorithm has been developed for estimating the parameters of this canonical form from noisy data which combines stochastic approximation and pseudo-inverse.

Chapter 6 discusses the problem of combined parameter and state estimation of linear multivariable systems. A way has been obtained for representing the state space model in a nonparametric model form which causes the residual error to be uncorrelated with the forcing function. Hence unbiased estimate of system parameters can be obtained. The parameters and states of that model representation have been estimated in a bootstrap manner by least-squares and stochastic approximation algorithms, respectively.

Conclusions and suggestions for future investigation in the problem of identification of linear discrete-time multivariable systems are discussed in Chapter 7.

CHAPTER 2

EFFECT OF THE CHOICE OF MODEL ON THE PROPERTIES OF THE IDENTIFICATION ALGORITHM

2.1 Introduction

The type of the model used for a multivariable system has considerable effects on the properties of the corresponding algorithm for identification. Since the structural parameters and the number of parameters to be estimated for a multivariable system depend on the choice of the model, the complexity of the algorithm is determined by this choice. Another important effect is on the bias of the estimated system parameters as bias is often introduced into parameter estimation using least-squares method by choosing a model which causes the forcing function to include observations which are correlated with the error in the observed output. In general, there are four types of system models which have been studied and used in the area of identification of linear time-invariant discrete-time multivariable systems. These models are: the transfer-function matrix, the impulse response matrix, the input-output difference equation and the state space formulation. These four models are equivalent and transformations between them are possible.

In this chapter a study [30] is presented of the relative effectiveness of the four types of models which have been used in the area of identification of linear multivariable systems. The features of each model and its effect on the complexity of the identification algorithm are studied. The structural parameters required to characterize each model and the number of parameters to be estimated for each model are examined and compared. Also, the effect of each model representation on the bias of the parameter estimates of the system when using the ordinary least-squares method is investigated.

2.2 Statement of the Problem

Consider an nth order linear, time-invariant, multivariable, discrete-time system with p-inputs and m-outputs as shown in Fig. 2.1, where

u(k) is the p-dimensional input vector sequence

y(k) is the m-dimensional output vector sequence

z(k) is the m-dimensional measured output vector sequence

and v(k) is the m-dimensional noise vector sequence at the output.

The identification problem can be formulated as "to estimate a model from the available record of the input-output sequences which fits these sequences as closely as possible". This definition allows a number of different types of models to be used in the identification problem. In order to solve this problem one has to specify, first, a certain model to be identified then the structural parameters characterizing this model are to be determined from a record of the input-output data. Finally the parameters of this model have to be estimated from the input-output data by a suitable estimation algorithm.

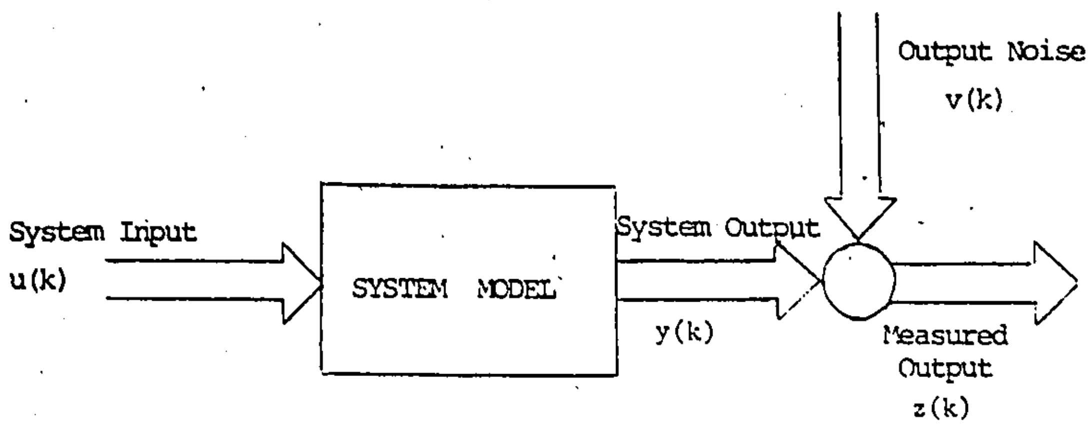

Fig. 2.1 Linear multivariable discrete-time system

2.3 Model Representation

In this section the four different types of models used for identification of linear multivariable systems will be studied.

2.3.1 Transfer-function Matrix

Consider the linear discrete-time system of Fig. 2.1. It can be represented by the following equation

$$y(k) = G(z)\, u(k) \tag{2.1}$$

where u(k) is the p-dimensional input vector sequence, y(k) is the m-dimensional output vector sequence, G(z) is the transfer-function matrix of the system and z is the unit advance operator.

Different forms of the transfer-function matrix have been used for identification. The following general form for G(z),

$$G(z) = \begin{bmatrix} \frac{A_{11}(z)}{B_{11}(z)} & \frac{A_{12}(z)}{B_{12}(z)} & \cdots & \frac{A_{1p}(z)}{B_{1p}(z)} \\ \frac{A_{21}(z)}{B_{21}(z)} & \frac{A_{22}(z)}{B_{22}(z)} & \cdots & \frac{A_{2p}(z)}{B_{2p}(z)} \\ \vdots & \vdots & & \vdots \\ \frac{A_{m1}(z)}{B_{m1}(z)} & \frac{A_{m2}(z)}{B_{m2}(z)} & \cdots & \frac{A_{mp}(z)}{B_{mp}(z)} \end{bmatrix} \tag{2.2}$$

where $A_{ij}(z)$ and $B_{ij}(z)$, $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, p$ are polynomials in z of degree less than or equal to n, has been considered for identification by Abaza [31] and Sinha and Caines [32]. Another form for G(z),

$$G(z) = \begin{bmatrix} \frac{A^*_{11}(z)}{D_1(z)} & \frac{A^*_{12}(z)}{D_1(z)} & \cdots & \frac{A^*_{1p}(z)}{D_1(z)} \\ \frac{A^*_{21}(z)}{D_2(z)} & \frac{A^*_{22}(z)}{D_2(z)} & \cdots & \frac{A^*_{2p}(z)}{D_2(z)} \\ \vdots & \vdots & & \vdots \\ \frac{A^*_{m1}(z)}{D_m(z)} & \frac{A^*_{m2}(z)}{D_m(z)} & \cdots & \frac{A^*_{mp}(z)}{D_m(z)} \end{bmatrix} \tag{2.3}$$

where $D_i(z)$ is defined as the least common denominator of the ith row of G(z) of equation (2.2) having the degree n_i (less than or equal to n) and $A^*_{ij}(z)$'s are polynomials in z of maximum degree n_i-1, has been considered for identification by Mital and Chen [33]. A third form for G(z),

$$G(z) = \frac{1}{D(z)} \begin{bmatrix} A'_{11}(z) & A'_{12}(z) & \cdots & A'_{1p}(z) \\ A'_{21}(z) & A'_{22}(z) & \cdots & A'_{2p}(z) \\ \vdots & \vdots & & \vdots \\ A'_{m1}(z) & A'_{m2}(z) & \cdots & A'_{mp}(z) \end{bmatrix} \tag{2.4}$$

where D(z) is the characteristic polynomial of the system (of degree n) defined as the least common monic denominator of all minors of G(z) of equation (2.2) and $A'_{ij}(z)$'s are polynomials in z of maximum degree n-1, has been considered for identification by Sen and Sinha [28].

The structural parameters required to characterize G(z) of equation (2.2) are the orders of the numerator and the denominator of each entry of G(z) and the number of parameters of the model to be estimated depends on these orders. The structural parameters for G(z) of equation (2.3) are the orders n_i's of each row of G(z) and the number of parameters to be estimated is $\Sigma_{i=1}^{m} n_i(p+1)$. The structural parameter for G(z) of equation (2.4) is the order of the system n and the number of parameters to be estimated is n(pm+1).

It can be noticed that the form of equation (2.2) for G(z) is unique and minimal, the forms of equations (2.3) and (2.4) are unique but not minimal.

2.3.2 Impulse Response

Consider the linear discrete-time system of Fig. 2.1. It can be represented by the following infinite series.

$$y(k) = [J_0 z^{-1} + J_1 z^{-2} + J_2 z^{-3} + \ldots] u(k) \qquad (2.5)$$

where J_0, J_1, ... are constant matrices called the Markov parameters of the system. These Markov parameters define the impulse response sequence for the system, hence it is possible to truncate the series after ℓ terms if the system is stable. The system can, therefore, be represented by the following equation

$$y(k) = [J_0 z^{-1} + J_1 z^{-2} + \ldots + J_{\ell-1} z^{-\ell}] u(k) \qquad (2.6)$$

The model of equation (2.6) has been considered for identification by Mehra [34], El-Sherief and Sinha [35] and Sinha et al. [36]. The structural parameter required to characterize the system in the model of equation (2.6) is the value ℓ, i.e. the minimum number of

Markov parameters required to describe the system completely. The value of ℓ is related to the observability and controllability subindices of the system. The number of parameters of the model required to be estimated depends on the choice of the value of ℓ and is given by ℓmp.

2.3.3 Input-output Difference Equation

Consider the linear discrete-time system of Fig. 2.1. It can be represented by the following difference equation

$$P(z^{-1})\ y(k) = Q(z^{-1})\ u(k) \tag{2.7}$$

where

$$P(z^{-1}) = I + P_1\ z^{-1} + \dots + P_{m_1}\ z^{-m_1} \tag{2.8}$$

$$Q(z^{-1}) = Q_1\ z^{-1} + Q_2\ z^{-2} + \dots + Q_{m_2}\ z^{-m_2} \tag{2.9}$$

and P_i's and Q_i's are constant matrices of proper dimension. The form of the matrices $P(z^{-1})$ and $Q(z^{-1})$ of equation (2.8) and (2.9) has been considered for identification by Kashyap and Nasburg [37]. Also a multivariable autoregressive moving average model for the input-output difference equation representations has been considered by Hannan [38], Akaike [39] and Dickinson et al. [40]. Another canonical input-output difference equation representation has been considered for identification by Guidorzi [41]. This representation has the following form

$$P(z) = \begin{bmatrix} p_{11}(z) & \dots & p_{1m}(z) \\ \vdots & & \vdots \\ p_{m1}(z) & \dots & p_{mm}(z) \end{bmatrix} \quad (2.10)$$

$$Q(z) = \begin{bmatrix} q_{11}(z) & \dots & q_{1p}(z) \\ \vdots & & \vdots \\ q_{m1}(z) & \dots & q_{mp}(z) \end{bmatrix} \quad (2.11)$$

where $p_{ii}(z)$ and $q_{ij}(z)$ are polynomials in z of the following form

$$p_{ii}(z) = z^{n_i} - a_{ii}(n_i)\, z^{n_i - 1} - \dots - a_{ii}(1) \quad (2.12)$$

$$p_{ij}(z) = - a_{ij}(n_{ij})\, z^{n_{ij} - 1} - \dots - a_{ij}(1) \quad (2.13)$$

$$q_{ij}(z) = b_{ij}(1)\, z^{n_i - 1} + \dots + b_{ij}(n_i) \quad (2.14)$$

and n_i's are the observability subindices of the system [41]. The canonical form of equations (2.10)-(2.14) has been also considered for identification by Sinha and Kwong [42] and El-Sherief and Sinha [43].

The structural parameters required to characterize the system model of equations (2.8) and (2.9) are the degrees m_1 and m_2 of the polynomial matrices $P(z^{-1})$ and $Q(z^{-1})$ and the number of parameters to be estimated is $m_1p^2+m_2mp$. The structural parameters characterizing the model of equations (2.10)-(2.14) are the observability subindices of the system (n_i, i = 1, 2, ..., m) and the number of parameters to be estimated is n(p+m).

2.3.4 State Space

Consider the linear discrete-time system of Fig. 2.1. It can be represented by the following equation

$$\begin{aligned} x(k+1) &= A\,x(k) + B\,u(k) \\ y(k) &= C\,x(k) \end{aligned} \qquad (2.15)$$

where x(k) is the n-dimensional state space vector and A, B and C are constant matrices of proper dimensions.

The form of matrices A, B and C of equation (2.15) is not unique and because of this nonuniqueness several canonical forms have been proposed for the purpose of identification. All these canonical forms aim to transform the parametric model of equation (2.15) into a nonparametric form suitable for identification and also to reduce the number of parameters to be estimated. Many algorithms have been proposed to identify the system in the state space representation from input-output data. Most of these algorithms start by identifying the system in one of the three models, transfer-function matrix [44], impulse response [45] or input-output difference equation [41], then the state space representation [i.e. matrices A, B and C of equation (2.15)] is obtained by different transformations.

Guidorzi [41] has obtained a unique relation between the state space representation in a certain canonical form and the input-output difference equation [equations (2.10)-(2.14)]. The matrices $\bar{A}$ and $\bar{C}$ in this canonical form have the following structure

$$\bar{A} = [\bar{A}_{ij}]$$

where

$$\bar{A}_{ii} = \begin{bmatrix} 0 & & \\ \vdots & I_{n_i-1} & \\ a_{ii}(1) & a_{ii}(2) \ \ldots & a_{ii}(n_i) \end{bmatrix}, \quad \bar{A}_{ij} = \begin{bmatrix} 0 & \ldots\ldots\ldots & 0 \\ \vdots & \cdot & \vdots \\ a_{ij}(1) & a_{ij}(2) \ \ldots & a_{ij}(n_{ij}) \end{bmatrix}$$

$$\bar{C} = \begin{bmatrix} e^1 \\ e^{n_1+1} \\ \vdots \\ e^{n_1+n_2+\ldots+n_{m-1}+1} \end{bmatrix} \qquad (2.16)$$

where e^i is the ith unit row vector of dimension n, n_i; i = 1, 2, ..., m are the observability subindices and the matrix $\bar{B}$ has no special form. Guidorzi [41] has showed that the parameters of the matrix P(z) of equation (2.10) are the same as those of the matrix $\bar{A}$ and the parameters of the matrix Q(z) of equation (2.11) are obtained from those of the matrices $\bar{A}$ and $\bar{B}$. He also showed that the structural parameters of the canonical input-output difference equation and those of the canonical state space representation are the same. The canonical form of (2.16) has been considered for identification also by Lobbia and Saridis [45] and El-Sherief and Sinha [29]. Other state space canonical forms have been proposed for identification, e.g. [46]-[49].

The structural parameter required to characterize the model of equation (2.15) is the order of the system n and the number of parameters of the model to be estimated is n(n+p+m). It is known that not all of these parameters are required to specify the system model and some of them can be put to zero or one by transforming the system matrices into special canonical forms; e.g. equation (2.15). The structural parameters required to characterize the state space representation of equation (2.16) are the observability subindices n_i's and the number of parameters to be estimated is, in general, n(p+m).

2.4 Effect of the Choice of the Model on the Bias of the Parameter Estimates Using the Ordinary Least-squares Method

Assume that the outputs of the system of Fig. 2.1 are measured with additive zero-mean white noise sequence v(k),

$$z_i(k) = y_i(k) + v_i(k) \qquad i = 1, 2, \ldots, m \qquad (2.17)$$

where $z_i(k)$ is the ith noisy output and $v_i(k)$ is the noise at the ith output.

Consider the following discrete-time observation representation which can be obtained for the ith output of the system

$$z_i(k) = U_i^T(k)\ \theta_i + e_i(k) \qquad (2.18)$$

where $z_i(k)$ is the ith observed output, $U_i(k)$ is a vector of forcing functions (in general function of the inputs and outputs), θ_i is a vector of parameters to be estimated characterizing the ith output and $e_i(k)$ is the residual error in the ith observed output. Using the ordinary least-squares method unbiased estimate of the parameter vector θ_i can be obtained if and only if the following two conditions are satisfied [50]:

a) The sequences $U_i(k)$ and $e_i(k)$ are statistically independent

b) The residual error sequence $e_i(k)$ has a zero-mean.

Thus bias is often introduced into parameter estimates using least-squares method by choosing a model which causes the forcing function to include observations which are correlated with the residual error in the measured output.

Next, the effect of choice of model representation for identification on the bias of the parameter estimates of the model from noisy data using the ordinary least-squares method will be discussed.

2.4.1 Transfer-function Matrix Representation

Consider the system in the transfer-function matrix form of equation (2.3) with the noise equation (2.17). The ith measured output of the system can be written as

$$z_i(k) = \sum_{j=1}^{p} \sum_{\ell=1}^{n_i} a_{ij}(\ell)\, u_j(k-\ell) - \sum_{\ell=1}^{n_i} d_i(\ell)\, z_i(k-\ell) + \sum_{\ell=1}^{n_i} d_i(\ell)\, v_i(k-\ell) + v_i(k) \tag{2.19}$$

where $a_{ij}(\ell)$ and $d_i(\ell)$ are the parameters of the polynomials $A^*_{ij}(z)$ and $D_i(z)$ respectively.

Equation (2.19) can be put into the form of equation (2.18) where

$$U_i(k) \triangleq [u_1(k-1) \ldots u_1(k-n_i) u_2(k-1) \ldots u_p(k-n_i) z_i(k-1) \ldots z_i(k-n_i)]^T \tag{2.20}$$

$$e_i(k) \triangleq \sum_{\ell=1}^{n_i} d_i(\ell)\, v_i(k-\ell) + v_i(k) \tag{2.21}$$

From equations (2.20) and (2.21) we can notice that the vector sequence $U_i(k)$ contains the outputs $z_i(k-1), \ldots, z_i(k-n_i)$ which are correlated with the residual error $e_i(k)$. Hence condition (a) is not satisified and the estimate of the vector θ_i is biased.

2.4.2 Impulse Response Representation

Consider the impulse response representation of equation (2.6) with the noise equation (2.17). The ith measured output of the system can be written as

$$z_i(k) = \sum_{j=0}^{\ell-1} J_{j,i}\, u(k-j-1) + v_i(k) \tag{2.22}$$

where J_s; $s = 0, 1, \ldots, \ell-1$ has been partioned as

$$J_s = \begin{bmatrix} J_{s,1} \\ J_{s,2} \\ \vdots \\ J_{s,m} \end{bmatrix} \tag{2.23}$$

Equation (2.22) can be put into the form of equation (2.18) where

$$U_i(k) \triangleq [u^T(k-1)\; u^T(k-2)\; \ldots\; u^T(k-\ell)]^T \tag{2.24}$$

$$e_i(k) \triangleq v_i(k) \tag{2.25}$$

From equations (2.24) and (2.25) we can see that the vector sequence $U_i(k)$ is uncorrelated with the residual error $e_i(k)$, hence condition (a) is satisfied. Moreover, if the sequence $v_i(k)$ has zero-mean then condition (b) is also satisfied. Therefore the estimate of the parameters of the impulse response model representation of the system is unbiased.

2.4.3 Input-output Difference Equation Representation

Consider the input-output difference equation (2.7)-(2.9) with the noise equation (2.17). The ith measured output of the system can be written as

$$z_i(k) = \sum_{j=1}^{m_2} Q_{j,i}\, u(k-j) - \sum_{j=1}^{m_1} P_{j,i}\, z(k-j) + \sum_{j=1}^{m_1} P_{j,i}\, v(k-j) + v_i(k) \tag{2.26}$$

where $P_{j,i}$ and $Q_{j,i}$ are the ith rows of P_j and Q_j respectively.

Equation (2.26) can be put into the form of equation (2.18) where

$$U_i(k) \triangleq [u^T(k-1) \ldots u^T(k-m_2)\ z^T(k-1) \ldots z^T(k-m_1)]^T \tag{2.27}$$

$$e_i(k) \triangleq \sum_{j=1}^{m_1} P_{j,i}\, v(k-j) + v_i(k) \tag{2.28}$$

From equations (2.27) and (2.28) we notice that the vector sequence $U_i(k)$ contains the outputs $z(k-1), \ldots, z(k-m_1)$ which are correlated with the residual $e_i(k)$ hence condition (a) is not satisfied and the estimate of the vector θ_i is biased.

2.4.4 State Space Representation

Consider the state space representation of equations (2.15) and (2.16) with the noise equation (2.17). The ith measured output can be written as follows (this is obtained using the relation between equation

(2.16) and equations (2.10)-(2.14))

$$z_i(k) = \sum_{j=1}^{p} \sum_{\ell=1}^{n_i} b_{ij}(\ell)\, u_j(k-\ell) + \sum_{j=1}^{m} \sum_{\ell=1}^{n_{ij}} a_{ij}(n_{ij}-\ell+1)\, z_j(k-\ell) + \qquad (2.29)$$

$$- \sum_{j=1}^{m} \sum_{\ell=1}^{n_{ij}} a_{ij}(n_{ij}-\ell+1)\, v_j(k-\ell) + v_i(k)$$

Equation (2.29) can be put into the form of equation (2.18) where

$$U_i(k) \triangleq [u_1(k-1) \ \dots \ u_1(k-n_i) \ \dots \ u_p(k-n_i) \ z_1(k-1) \ \dots \ z_1(k-n_i) \ \dots \ z_m(k-n_i)]^T \qquad (2.30)$$

$$e_i(k) \triangleq - \sum_{j=1}^{m} \sum_{\ell=1}^{n_{ij}} a_{ij}(n_{ij}-\ell+1)\, v_j(k-\ell) + v_i(k) \qquad (2.31)$$

From equations (2.30) and (2.31) we can notice that the vector sequence $U_i(k)$ contains the output sequence $z_1(k-1)$, ..., $z_m(k-n_i)$ which is correlated with the residual error $e_i(k)$, hence condition (a) is not satisfied and the estimate of the parameters obtained in this model representation is biased.

2.5 Results of Simulation

In this section the ideas discussed in sections 2.3 and 2.4 are demonstrated by estimating the parameters of a simulated 2-input 2-output 4th order system in the four different model representations by the recursive least-squares method.

The given system, in the transfer-function matrix form of equation (2.2), is as follows

$$G(z) = \begin{bmatrix} \frac{1}{z-.25} & \frac{2}{z-.4} \\ \frac{2}{z-.25} & \frac{4}{(z-.5)^2} \end{bmatrix} \tag{2.32}$$

This can be written in the form of equation (2.3) where

$$A^*_{11}(z) = z-.4$$

$$A^*_{12}(z) = 2(z-.25)$$

$$A^*_{21}(z) = 2(z-.5)^2$$

$$A^*_{22}(z) = 4(z-.25)$$

$$D_1(z) = (z-.25)(z-.4)$$

$$D_2(z) = (z-.25)(z-.5)^2 \tag{2.33}$$

The first five Markov parameters of the system are

$$J_0 = \begin{bmatrix} 1 & 2 \\ 2 & 0 \end{bmatrix},\ J_1 = \begin{bmatrix} .25 & .8 \\ .50 & 4 \end{bmatrix},\ J_2 = \begin{bmatrix} .0625 & .32 \\ .1250 & 4 \end{bmatrix},\ J_3 = \begin{bmatrix} .0156 & .1280 \\ .0313 & 3 \end{bmatrix}$$

$$J_4 = \begin{bmatrix} .0039 & .0512 \\ .0078 & 2 \end{bmatrix} \tag{2.34}$$

The input-output difference equation of the given system in the form of equations (2.10)-(2.14) is as follows

$$P(z) = \begin{bmatrix} z^2-.65z+.1 & 0 \\ (5/6)z-(1/3) & z^2-z+.25 \end{bmatrix},\ Q(z) = \begin{bmatrix} z-.4 & 2z-.5 \\ 2z-(2/3) & 17/3 \end{bmatrix} \tag{2.35}$$

The state space representation of this system in the canonical form of equation (2.16) is as follows

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -.1 & .65 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1/3 & -5/6 & -.25 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 \\ .25 & .8 \\ 2 & 0 \\ .5 & 4 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (2.36)$$

The above system was simulated on a CDC-6400 computer using the state space model representation of equation (2.39) with zero-initial states and the two inputs were taken as uncorrelated zero-mean white noise sequences with standard deviation of value 1. To each output a zero-mean white noise sequence was added with standard deviation adjusted to vary the noise level at that output. The parameters of the above system in the four different model representations were estimated using recursive least-squares method for different noise levels. The final estimate of the parameters of each model after 500 iterations is shown in Tables 2.1-2.4.

We can see from Tables 2.1-2.4 that good estimate of the parameters of the system in each of the four models has been obtained for the noise-free case. For the noisy case the estimate of the parameters of the three models, transfer-function matrix, the input-output difference equation and state space is biased and inconsistent. On the other hand we can notice from Table 2.2 that good estimates of the parameters of the impulse response model have been obtained even for high noise level, $\sigma_1 = \sigma_2 = .1$.

Table 2.1: Estimate of the parameters of the transfer-function matrix model

True parameters	Noise-free case	$\sigma_1 = .01$ $\sigma_2 = .01$	$\sigma_1 = .05$ $\sigma_2 = .05$	$\sigma_1 = .1$ $\sigma_2 = .1$
.65	.6500	.6454	.5432	.3791
.10	.1000	-.0982	-.0600	.0002
1.00	1.0000	.9994	.9972	.9970
-.40	-.4000	-.3954	-.2924	-.1277
2.00	2.0000	2.0001	2.0004	1.9996
-.50	-.5000	-.4912	-.2879	.0353
1.25	1.2500	1.2431	1.0223	.6903
-.50	-.5000	-.4932	-.2725	.0600
.0625	.0625	.0610	.0072	-.0742
2.00	2.0000	1.9997	1.9982	1.9983
-2.00	-2.0000	-1.9850	-1.5379	-.8688
.50	.50000	.4880	.1504	.3551
.00	.0000	.0015	.0076	.0038
4.00	4.0000	3.9976	3.9880	3.9886
-1.00	-1.0000	.9706	-.0818	1.2464
Error square of parameter estimates	.0000	.0407	1.3793	3.3167

Table 2.2: Estimate of the parameters of the impulse response model

True parameters	Noise-free case	$\sigma_1 = .01$ $\sigma_2 = .01$	$\sigma_1 = .05$ $\sigma_2 = .05$	$\sigma_1 = .1$ $\sigma_2 = .1$
1.00	1.0010	1.0014	1.0031	1.0052
2.00	1.9997	1.9993	1.9975	1.9953
.25	.2509	.2513	.2524	.2540
.80	.7989	.7985	.7969	.7950
.0625	.0619	.0619	.0619	.0619
.32	.3207	.3209	.3218	.3229
.0156	.0154	.0161	.0187	.0220
.128	.1282	.1279	.1270	.1258
.0039	.0043	.0044	.0049	.0056
.0512	.0509	.0514	.0531	.0552
2.00	2.0407	2.0473	2.0500	2.0534
.00	-.0423	-0.0414	-0.0378	-.0333
.50	.5632	.5648	.5709	.5785
4.00	3.9156	3.9145	3.9100	3.9045
.125	.1012	.1009	.0995	.0978
4.00	4.0180	4.0180	4.0183	4.0186
.0313	.0142	.0141	.0135	.0128
3.00	3.0096	3.0096	3.0098	3.0101
.0078	.0239	.0242	.0252	.0265
2.00	1.9821	1.9825	1.9843	1.9865
Error square of parameter estimates	.0164	.0174	.0191	.0214

Table 2.3: Estimate of the parameters of the input-output difference equation model

True parameters	Noise-free case	$\sigma_1 = .01$ $\sigma_2 = .01$	$\sigma_1 = .05$ $\sigma_2 = .05$	$\sigma_1 = .1$ $\sigma_2 = .1$
.65	.6497	.0792	-.0029	-.0071
-.10	-.1000	.0644	.0877	.0884
.00	.0000	.0319	.0368	.0375
.00	.0000	-.0187	-.0213	-.0213
1.00	1.0000	1.0003	1.0022	1.0045
-.40	-.3997	.1072	.1801	.1835
2.00	2.0000	1.9995	1.9975	1.9951
-.50	-.4994	.6412	.8035	.8098
-.83	-.8316	.0708	.1784	.1815
.33	.3328	.0721	.0383	.0350
1.00	.9999	.9503	.9467	.9487
-.25	-.2499	-.2207	-.2188	-.2201
2.00	2.0000	1.9996	1.9974	1.9947
-.67	-.6682	-1.4705	-1.5672	-1.5695
.00	.0000	.0016	.0079	.0157
5.67	5.6633	3.8566	3.6334	3.6177
Error square of parameter estimates	.0001	6.7243	8.2693	8.5007

Table 2.4: Estimate of the parameters of the state space model

True parameters	Noise-free case	$\sigma_1 = .01$ $\sigma_2 = .01$	$\sigma_1 = .05$ $\sigma_2 = .05$	$\sigma_1 = .1$ $\sigma_2 = .1$
-.10	-.1000	.0644	.0877	.0884
.65	.6497	.0792	-.0029	-.0071
.00	.0000	.0319	.0368	.0375
.00	.0000	-.0183	-.0213	-.0213
.33	.3328	.0721	.0383	.0350
-.83	-.8316	.0708	.1784	.1815
-.25	-.2499	-.2207	-.2188	-.2201
1.00	.9999	.9503	.9467	.9487
1.00	1.0000	1.0003	1.0022	1.0045
2.00	2.0000	1.9995	1.9975	1.9951
.25	.2500	.2502	.2507	.2512
.80	.8000	.7050	.7980	.7962
2.00	2.0000	1.9996	1.9974	1.9947
.00	.0000	.0165	.0079	.0157
.50	.5000	.5018	.3621	.3580
4.00	4.0001	4.0164	3.7174	3.7024
Error square of parameter estimates	.0000	1.3402	1.6681	1.6919

2.6 Concluding Remarks

Four different model representations have been used in the area of identification of linear multivariable systems and all these models are equivalent and transformations between them are possible. Each model has its own structural parameters which have to be determined in advance before parameter estimation.

Out of the four models used for identification the state space model is used much because of the smaller number of parameters needed in the model when canonical forms are used and also because of its practical use in control theory. In general to identify the system in state space form a nonparametric model for the system has to be estimated first, then estimates of the parameters of the state space model are obtained by a certain transformation.

When estimating the parameters of the system from noisy data using the ordinary least-squares method the parameter estimates are biased if the residual error is correlated with the forcing function. In general most of the identification algorithms concentrate mainly on the problem of removing this bias in the estimated parameters [e.g., generalized least-squares and maximum likelihood]. It has been shown in section 2.4 that identifying the system in the impulse response form has an advantage over the other three forms because it results in unbiased estimate of the parameters of the system when the least-squares method is used. This fact has been demonstrated by the simulation results of section 2.5.

CHAPTER 3

IDENTIFICATION OF THE TRANSFER-FUNCTION MATRIX MODEL

3.1 Introduction

The problem of identification of linear discrete-time multivariable system in the transfer-function matrix representation from input-output data has received less attention during the past years inspite of its use in many areas of control (Wolovich [51]). Mital and Chen [33] have proposed an off-line algorithm to estimate the parameters of the transfer-function matrix from noise-free data.

Furuta [44], has proposed an algorithm for the identification of the transfer-function matrix assuming that the inputs and outputs of the system are corrupted by additive white noise and the system order is not known. The algorithm starts by assuming a sufficiently large value for the order of the transfer-function matrix, then a set of models is identified to be equivalent to the original system. By taking the minimal realization of these equivalent models, the appropriate unique model is derived.

Sen and Sinha [28], have proposed a recursive algorithm which assumes noise-free measurements of the inputs and outputs, and the prior knowledge of the order of the transfer-function matrix. The algorithm proceeds as follows; the ith output of the system in the noise-free case, using the form of equation (2.4) for G(z), can be written as

$$y_i(k) = \frac{1}{D(z)} \sum_{j=1}^{p} A'_{ij}(z)\, u_j(k) \tag{3.1}$$

Let

$$D(z) = z^n + d(1)z^{n-1} + \dots + d(n)$$

$$A'_{ij}(z) = a_{ij}(1)z^{n-1} + a_{ij}(2)z^{n-2} + \dots + a_{ij}(n)$$

where n is the order of the system. Then, for i = 1, 2, ..., m equation (3.1) can be rewritten as

$$Y(k) = W(k)\,\phi \tag{3.2}$$

where

$$Y^T(k) = [y_1(k) \quad y_2(k) \quad \dots \quad y_m(k)]$$

$$W(k) = \begin{bmatrix} U^T(k) & 0 & \dots & 0 & -Y_1^T(k) \\ 0 & U^T(k) & \dots & 0 & -Y_2^T(k) \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & U^T(k) & -Y_m^T(k) \end{bmatrix}$$

$$U^T(k) = [u_1(k) \dots u_1(k-n+1) u_2(k) \dots u_p(k-n+1)]$$

$$Y_i^T(k) = [y_i(k-1) \quad y_i(k-2) \dots y_i(k-n)]$$

$$\phi^T = [a_{11}(1) \quad a_{11}(2) \dots a_{11}(n) \quad a_{12}(1) \dots a_{mp}(n) \; d(1) \dots d(n)]$$

and T denotes transposition. The parameter vector ϕ in equation (3.2) can be estimated using the following recursive pseudo-inverse algorithm

$$\hat{\phi}(k+1) = \hat{\phi}(k)+P(k)W^T(k+1)[I+W(k+1)P(k)W^T(k+1)]^{-1}[Y(k+1)-W(k+1)\hat{\phi}(k)]$$

$$P(k+1) = P(k)-P(k)W^T(k+1)[I+W(k+1)P(k)W^T(k+1)]^{-1}W(k+1)P(k).$$

Abaza [31], has proposed a two-step least-squares algorithm to identify the system in model (2.2). The ith output of equation (2.1) for the noisy case can be written as,

$$z_i(k) = \frac{A_{i1}(z)}{B_{i1}(z)} u_1(k) + \dots + \frac{A_{ip}(z)}{B_{ip}(z)} u_p(k) + \frac{D_{i1}(z)}{C_{i1}(z)} e_1(k)$$

$$+ \dots + \frac{D_{ip}(z)}{C_{ip}(z)} e_p(k) \qquad (3.3)$$

where

$$G_{ij}(z) = \frac{A_{ij}(z)}{B_{ij}(z)}$$

$$N_{ij}(z) = \frac{D_{ij}(z)}{C_{ij}(z)}, \quad \begin{array}{l} i = 1, 2, \dots, m; \\ j = 1, 2, \dots, p. \end{array}$$

and $N(z)$ is the noise transfer-function matrix and $e_i(k)$ is a white noise sequence.

By dividing $A_{ij}(z)$ into $B_{ij}(z)$ and replacing the noise term by the additive noise sequence $\xi_i(k)$, equation (3.3) is reduced to

$$z_i(k) = \nu_{i1}(z^{-1}) u_1(k) + \dots + \nu_{ip}(z^{-1}) u_p(k) + \xi_i(k) \qquad (3.4)$$

where the $\nu_{ij}(z^{-1})$ are polynomials in z^{-1}. By considering K measurements and truncating after s terms equation (3.4) can be written as

$$Y_i(K) = U_i(K) \nu_i + E_i(k)$$

where

$$\nu_i^T = [\nu_{i1}^1 \ldots \nu_{i1}^s \; \nu_{i2}^1 \ldots \nu_{ip}^1 \ldots \nu_{ip}^s]$$

$$Y_i^T(K) = [z_i(k) \;\; z_i(k+1) \;\ldots\; z_i(K+k-1)]$$

$$E_i^T(K) = [\xi_i(k) \;\; \xi_i(k+1) \;\ldots\; \xi_i(K+k-1)].$$

$U_i(K)$ is the data matrix made up of the system inputs and the ν_{ij}'s are the parameters of the polynomial $\nu_{ij}(z^{-1})$. The parameters of the transfer-function matrix can then be estimated using a 2-stage linear least-squares as follows; first an unbiased l.l.s. estimate of ν_i's is obtained and then used to obtain estimates of the noise-free outputs. Using these filtered data, unbiased estimates of the parameters in the $G_{ij}(z)$'s can be obtained using l.l.s. This procedure is repeated until convergence of the parameter estimates is obtained.

Recently, Sinha and Caines [32] have proposed an instrumental variable identification algorithm which uses binary shift register sequences as both system inputs and instrumental variables. They applied this algorithm to two cases of equation (2.1). In the first case they assumed that each row of the matrix G(z) has identical denominator polynomials. As the first case is not practical they applied the algorithm to the general case of equation (2.1) and estimated every term in equation (3.3) one at a time assuming all other terms as a coloured noise.

In this chapter two algorithms will be presented for identifying linear discrete-time multivariable systems in the transfer-function matrix representation. In section 3.2 a recursive algorithm [52] more efficient than Sen and Sinha's algorithm [28] is developed for estimating the parameters of a special transfer-function matrix model, equation (2.4), from noise-free data. In section 3.3 another algorithm is developed for identifying a more general transfer-function matrix representation of the system from noise-free as well as noisy data assuming that the order of the system is unknown.

3.2 Recursive Estimation of Model (2.4)

3.2.1 Formulation of the Problem

Consider a linear discrete-time multivariable system with p-inputs and m-outputs. It can be represented by an mxp transfer-function matrix, $G(z)$, with the following input-output relationship

$$y(k) = G(z)\,u(k) \tag{3.5}$$

where $u(k)$ is the p-dimensional input vector sequence, $y(k)$ is the m-dimensional output vector sequence and z is the unit advance operator.

The transfer-function matrix $G(z)$ can be written as, equation (2.4),

$$G(z) = \frac{1}{D(z)}\begin{bmatrix} A'_{11}(z) & A'_{12}(z) & \dots & A'_{1p}(z) \\ A'_{21}(z) & A'_{22}(z) & \dots & A'_{2p}(z) \\ \vdots & \vdots & & \vdots \\ A'_{m1}(z) & A'_{m2}(z) & \dots & A'_{mp}(z) \end{bmatrix} \tag{3.6}$$

where $D(z)$ is the characteristic polynomial of the system, defined as the least common monic denominator of all minors of $G(z)$ and $A'_{ij}(z)$'s are polynomials in z.

It will be assumed that $D(z)$ is a polynomial of order n, (order of the system), and $A'_{ij}(z)$'s are polynomials of order n-1. Hence one may write

$$D(z) = z^n + d(1)z^{n-1} + \dots + d(n-1)z + d(n) \qquad (3.7)$$

$$A'_{ij}(z) = a_{ij}(1)z^{n-1} + a_{ij}(2)z^{n-2} + \dots + a_{ij}(n-1)z + a_{ij}(n) \qquad (3.8)$$

[If the true order of any of the $A'_{ij}(z)$'s is less than n-1 the corresponding coefficient in equation (3.8) will be zero when estimated.]

The problem considered in this section is to estimate the parameters d(1), d(2), ..., d(n) and $a_{ij}(1)$, $a_{ij}(2)$, ..., $a_{ij}(n)$ for all i = 1, 2, ..., m and j = 1, 2, ..., p based on the measured input-output data sequence, u(k) and y(k) k = 1, 2, ...

3.2.2 System Decomposition

From the definition of G(z) given in equation (3.6) the outputs of the system can be written as

$$D(z)\ y_i(k) = \sum_{j=1}^{p} A'_{ij}(z)\ u_j(k) \qquad i = 1, 2, \dots, m \qquad (3.9)$$

Substituting for D(z) and $A'_{ij}(z)$ from equations (3.7) and (3.8), equation (3.9) can be written more explicitly in the following form

$$y_i(k) = \sum_{j=1}^{p} \sum_{\ell=1}^{n} a_{ij}(\ell)\ u_j(k-\ell) - \sum_{\ell=1}^{n} d(\ell)\ y_i(k-\ell) \quad i = 1, 2, \dots, m \qquad (3.10)$$

Now, the given system has been decomposed into m subsystems as shown in equation (3.10). Each of these subsystems corresponds to one row of the matrix G(z) and can be regarded as a single-output and multiple-input system. Hence the parameters of each subsystem, $a_{ij}(\ell)$'s and d(ℓ)'s, can be estimated independently and the identification of the whole system is accomplished in m separate steps.

3.2.3 The Identification Algorithm

Each subsystem of equation (3.10) can be written in the form of a matrix equation. For example, the jth subsystem may be expressed as

$$Y_j(k) = H_j(k)\ \phi_j \tag{3.11}$$

where

$$H_j(k) = \begin{bmatrix} u_1(0) & u_1(-1) \ldots u_1(1-n) & u_2(0) \ldots u_p(1-n) & -y_j(0) \ldots -y_j(1-n) \\ u_1(1) & u_1(0) \ldots u_1(2-n) & u_2(1) \ldots u_p(2-n) & -y_j(1) \ldots -y_j(2-n) \\ \vdots & \vdots \qquad \vdots & \vdots \qquad \vdots & \vdots \qquad \vdots \\ u_1(k-1) & u_1(k-2) \ldots u_1(k-n) & u_2(k-1) \ldots u_p(k-n) & -y_j(k-1) \ldots -y_j(k-n) \end{bmatrix} \tag{3.12}$$

$$\phi_j = [a_{j1}(1)\ \ a_{j1}(2)\ \ldots\ a_{j1}(n)\ a_{j2}(1)\ \ldots\ a_{jp}(n)\ d(1)\ \ldots\ d(n)]^T \tag{3.13}$$

$$Y_j(k) = [y_j(1)\ \ y_j(2)\ \ldots\ y_j(k)]^T$$

and the subscript T denotes transposition.

For large values of k, the least-squares estimate of ϕ_j is given by

$$\hat{\phi}_j(k) = (H_j^T(k)\ H_j(k))^{-1}\ H_j^T(k)\ Y_j(k) \tag{3.14}$$

where $\hat{\phi}_j(k)$ is the estimate of the parameter ϕ_j from k observations.

A recursive version of equation (3.14) for calculating $\hat{\phi}_j(k)$ was derived in [9] and [53] and is given by

$$\hat{\phi}_j(k+1) = \hat{\phi}_j(k) + \frac{P_j(k)h_j(k+1)[y_j(k+1)-h_j^T(k+1)\hat{\phi}_j(k)]}{1 + h_j^T(k+1)\ P_j(k)\ h_j(k+1)} \tag{3.15}$$

$$P_j(k+1) = P_j(k) - \frac{P_j(k)h_j(k+1)(P_j(k)h_j(k+1))^T}{1 + h_j^T(k+1)\ P_j(k)\ h_j(k+1)} \tag{3.16}$$

where

$$H_j(k+1) = \begin{bmatrix} H_j(k) \\ h_j^T(k+1) \end{bmatrix} \tag{3.17}$$

$$h_j(k+1) = [u_1(k)u_1(k-1)\ldots u_1(k-n)u_2(k)\ldots u_p(k-n)-y_j(k)\ldots -y_j(k-n)]^T$$

From equations (3.9) and (3.11) we can notice that the parameters of the characteristic polynomial, $D(z)$, are estimated m times during each iteration (i.e. for each one of the m subsystems). To reduce the computations in the proposed algorithm it is possible to avoid estimating the parameters of $D(z)$ more than one time for every iteration. This can be accomplished as follows. Using the first output an estimate of the parameters of $D(z)$ can be obtained, then using these estimated parameters and for $i = 2, 3, \ldots, m$ we may write

$$\hat{\bar{y}}_i(k) = y_i(k) + \sum_{\ell=1}^{n} \hat{d}(\ell)\ y_i(k-\ell) \tag{3.18}$$

where $\hat{d}(\ell)$ is an estimate of the parameter $d(\ell)$ and $\hat{\bar{y}}_i$ is an estimate of $\bar{y}_i$ defined as

$$\bar{y}_i(k) = y_i(k) + \sum_{\ell=1}^{n} d(\ell)\ y_i(k-\ell) \tag{3.19}$$

By using equation (3.18) the output equations (3.10) for $i = 2, 3, \ldots, m$ can be modified to

$$\hat{y}_i(k) = \sum_{j=1}^{p} \sum_{\ell=1}^{n} a_{ij}(\ell)\, u_j(k-\ell) \tag{3.20}$$

In the above equation we have eliminated the parameters of the polynomial $D(z)$ for the outputs $i = 2, 3, \ldots, m$. Hence the only parameters which need to be estimated for these outputs are the $a_{ij}(\ell)$'s which can be estimated by the algorithm (3.15)-(3.17) with the dimension of the parameter vector $\hat{\phi}_j(k)$, $j = 2, 3, \ldots, m$ reduced to np instead of $n(p+1)$. This results in considerable reduction in the number of computations.

The main advantage of the proposed algorithm over Sen and Sinha's algorithm is that it requires less computation and avoids matrix inversion. This is possible because it decomposes the identification problem into m separate single-output multiple-input problems. To show this, the number of arithmetic operations per iteration needed in the proposed algorithm, its reduced form and Sen and Sinha's algorithm are given in Table 3.1.

Algorithm	Total number of additions and subtractions	Total number of multiplications and divisions	Matrix inversion
Sen & Sinha's algorithm	$2m[n^2(mp+1)^2+mn(mp+1)+1-m]$	$2mn[n(mp+1)^2+m(mp+1)+(mp+1)]$	One (mxm) matrix
Proposed alg. without reduction	$2mn(p+1)(np+n+1)$	$2nm(p+1)(np+n+2)+m$	None
Proposed alg. with reduction	$2n[(p+1)(np+n+1)+p(m-1)(np+1)]$	$2n^2(p+1)^2+4npm+2n^2p^2(m-1)+4n+m$	None

Table 3.1 Comparison of the arithmetic operations per iteration

3.2.4 Simulation Results

As an example the proposed algorithm was applied to the identification of a simulated system with two-inputs and two-outputs.

As in equation (3.6), the transfer-function matrix of such a system is of the following form

$$G(z) = \frac{1}{D(z)} \begin{bmatrix} A'_{11}(z) & A'_{12}(z) \\ A'_{21}(z) & A'_{22}(z) \end{bmatrix}$$

with

$$D(z) = z^3 + d(1)z^2 + d(2)z + d(3)$$
$$A'_{11}(z) = a_{11}(1)z^2 + a_{11}(2)z + a_{11}(3)$$
$$A'_{12}(z) = a_{12}(1)z^2 + a_{12}(2)z + a_{12}(3)$$
$$A'_{21}(z) = a_{21}(1)z^2 + a_{21}(2)z + a_{21}(3)$$
$$A'_{22}(z) = a_{22}(1)z^2 + a_{22}(2)z + a_{22}(3)$$

where

$a_{11}(1) = 3$	$a_{11}(2) = -3.5$	$a_{11}(3) = -1.5$
$a_{12}(1) = 1$	$a_{12}(2) = -0.167$	$a_{12}(3) = -0.167$
$a_{21}(1) = -4$	$a_{21}(2) = -2$	$a_{21}(3) = -1$
$a_{22}(1) = 1$	$a_{22}(2) = -0.167$	$a_{22}(3) = -0.083$
$d(1) = 0.833$	$d(2) = 0.417$	$d(3) = 0.083$

The two inputs to the system were taken as uncorrelated unit variance white Gaussian sequences. Two cases were considered. First, the above system was identified from noise-free data. Next, the outputs were corrupted by a white noise with noise-to-signal ratio of 1%. The proposed algorithm, its reduced form and Sen and Sinha's algorithm were used to estimate the parameters of the system for these two cases. For the proposed algorithm with reduction an initial estimate of the parameters of the polynomial D(z) was obtained from the first output for 20 iterations and then used in identifying the second output. The final estimates of the system parameters after 150 iterations are shown in Table 3.2. Also, the total number of arithmetic operations and the total computation time per iteration are compared for the proposed algorithm and its reduced form with Sen and Sinha's algorithm and these are given in Table 3.3.

3.3 Identification of a General Transfer-function Matrix

3.3.1 Introduction

In this section an algorithm will be presented for identifying linear, discrete-time, multivariable systems in a more general transfer-function matrix (TFM) representation from noise-free as well as noisy data without prior knowledge of the order of the system. First the system is decomposed into m subsystems; each corresponds to one row of the transfer-function matrix. Then the order of each row is estimated using a proposed recursive algorithm for the noise-free case. The parameters of each subsystem are then estimated by a multivariable recursive least-squares algorithm. When the outputs of the system are corrupted by additive noise the order of each subsystem is estimated by an algorithm based on the residual error technique which has been used by Suen and Liu [54]. The parameters of each row are then estimated using a multivariable adaptive least-squares algorithm similar to that proposed by Panuska [55] for the case of single-input single-output

Parameters	True Value	All Alg. Noise-free case	Sen and Sinha's Alg. N.S.R. = 1%	Proposed Alg. N.S.R. = 1%	Proposed Reduced Alg. N.S.R. = 1%
$a_{11}(1)$	3.000	3.000	3.006	3.008	3.004
$a_{11}(2)$	-3.500	-3.500	-3.521	-3.576	-3.534
$a_{11}(3)$	-1.500	-1.500	-1.472	-1.419	-1.435
$a_{12}(1)$	1.000	1.000	1.003	1.002	1.002
$a_{12}(2)$	-0.167	-0.167	-0.168	-0.169	-0.468
$a_{12}(3)$	-0.167	-0.167	-0.146	-0.135	-0.139
$a_{21}(1)$	-4.000	-4.000	-3.989	-3.989	-3.998
$a_{21}(2)$	-2.000	-2.000	-1.958	-1.923	-1.940
$a_{21}(3)$	-1.000	-1.000	-0.989	-0.969	-0.971
$a_{22}(1)$	1.000	1.000	0.999	1.000	0.997
$a_{22}(2)$	-0.167	-0.167	-0.176	-0.177	-0.189
$a_{22}(3)$	-0.083	-0.083	-0.080	-0.088	-0.078
d(1)	0.833	0.833	0.825	0.814	0.821
d(2)	0.417	0.417	0.413	0.404	0.411
d(3)	0.083	0.083	0.081	0.081	0.085

Table 3.2 Results of simulation of the given example after 150 iterations

Algorithm	Total number of additions & subtractions	Total number of multiplications & divisions	Matrix inversion	Total computation time (sec.)	Reduction in computation time
Sen & Sinha's algorithm	1016	1080	2x2	0.0372	-
Proposed algorithm without reduction	360	398	None	0.0111	70.13%
Proposed algorithm with reduction	264	296	None	0.0078	78.81%

Table 3.3 Comparison of the computation effort per iteration for the given example

systems.

3.3.2 Problem Formulation

Consider a linear discrete-time multivariable system with p-inputs and m-outputs. It can be represented by an mxp TFM, G(z), with the following input-output relationship

$$y(k) = G(z)\ u(k) \tag{3.21}$$

where u(k) is the p-dimensional input vector sequence, y(k) is the m-dimensional output vector sequence and z is the unit advance operator.

In general, the TFM G(z) can be written as follows (equation (2.3))

$$G(z) = \begin{bmatrix} \frac{A_{11}(z)}{B_{11}(z)} & \frac{A_{12}(z)}{B_{12}(z)} & \cdots & \frac{A_{1p}(z)}{B_{1p}(z)} \\ \frac{A_{21}(z)}{B_{21}(z)} & \frac{A_{22}(z)}{B_{22}(z)} & \cdots & \frac{A_{2p}(z)}{B_{2p}(z)} \\ \vdots & \vdots & & \vdots \\ \frac{A_{m1}(z)}{B_{m1}(z)} & \frac{A_{m2}(z)}{B_{m2}(z)} & \cdots & \frac{A_{mp}(z)}{B_{mp}(z)} \end{bmatrix} \tag{3.22}$$

where $A_{ij}(z)$ and $B_{ij}(z)$ $i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, p$ are polynomials in z.

The TFM G(z) can also be expressed as

$$G(z) \doteq \begin{bmatrix} \frac{A^*_{11}(z)}{D_1(z)} & \frac{A^*_{12}(z)}{D_1(z)} & \cdots & \frac{A^*_{1p}(z)}{D_1(z)} \\ \frac{A^*_{21}(z)}{D_2(z)} & \frac{A^*_{22}(z)}{D_2(z)} & \cdots & \frac{A^*_{2p}(z)}{D_2(z)} \\ \vdots & & & \\ \frac{A^*_{m1}(z)}{D_m(z)} & \frac{A^*_{m2}(z)}{D_m(z)} & \cdots & \frac{A^*_{mp}(z)}{D_m(z)} \end{bmatrix} \quad (3.23)$$

where $D_i(z)$ is defined as the least common denominator of the ith row of $G(z)$ of equation (3.22) having the degree n_i and can be expressed as

$$D_i(z) = z^{n_i} + d_i(1)\, z^{n_i-1} + \ldots + d_i(n_i-1)\, z + d_i(n_i) \quad (3.24)$$

$A^*_{ij}(z)$ $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, p$ are polynomials in z of maximum degree n_i-1, $i = 1, 2, \ldots, m$ and can be expressed as

$$A^*_{ij}(z) = a_{ij}(1)\, z^{n_i-1} + a_{ij}(2)\, z^{n_i-2} + \ldots + a_{ij}(n_i-1)\, z + a_{ij}(n_i) \quad (3.25)$$

It can be noticed that the values of the orders n_i's (order of each row of $G(z)$) are in general less than or equal to the order of the system.

The system identification problem requires the determination of $G(z)$, i.e.

$$\{A_{ij}(z),\ D_i(z)\} :\ i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, p$$

from measurements of the input-output data sequence $u(k)$ and $y(k)$; $k = 1, 2, \ldots$ and without prior knowledge of the structural parameters of the system.

3.3.3 System Decomposition

From the representation of G(z) given in equation (3.23) the outputs of the system can be written as

$$D_i(z)\; y_i(k) = \sum_{j=1}^{p} A^*_{ij}(z)\; u_j(k) \qquad i = 1, 2, \ldots, m \quad (3.26)$$

where $y_i(k)$ is the ith output and $u_j(k)$ is the jth input of the system.

Substituting for $D_i(z)$ and $A^*_{ij}(z)$ from equations (3.24) and (3.25), equation (3.26) can be rewritten more explicitly as follows

$$y_i(k) = \sum_{j=1}^{p} \sum_{\ell=1}^{n_i} a_{ij}(\ell)\; u_j(k-\ell) - \sum_{\ell=1}^{n_i} d_i(\ell)\; y_i(k-\ell) \qquad i = 1, 2, \ldots, m \quad (3.27)$$

Now, the given system has been decomposed into m subsystems as shown in equation (3.27) where each subsystem corresponds to one row of the TFM and can be regarded as a single-output multi-input system of order n_i.

In the next subsections, the order n_i and the parameters $a_{ij}(\ell)$'s and $d_i(\ell)$'s of the ith subsystem (row) will be estimated independently from those of the other subsystems for the noise-free as well as noisy data cases. The identification of the whole system will be accomplished in m separate steps.

3.3.4 Noise-free Case

3.3.4.1 Parametric Identification

In this subsection the parameters of each subsystem will be estimated recursively from noise-free data assuming a prior knowledge of the orders n_i's. For the ith subsystem equation (3.27) can be concatenated as

$$Y_i(k) = H_i(k)\ \phi_i \tag{3.28}$$

where

$$H_i(k) = \begin{bmatrix} u_1(0) & u_1(-1) & \cdots & u_1(1-n_i) & u_2(0) & \cdots & u_p(1-n_i) & -y_i(0) & \cdots & -y_i(1-n_i) \\ u_1(1) & u_1(0) & \cdots & u_1(2-n_i) & u_2(1) & \cdots & u_p(2-n_i) & -y_i(1) & \cdots & -y_i(2-n_i) \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ u_1(k-1) & u_1(k-2) & \cdots & u_1(k-n_i) & u_2(k-1) & \cdots & u_p(k-n_i) & -y_i(k-1) & \cdots & -y_i(k-n_i) \end{bmatrix} \tag{3.29}$$

$$\phi_i = [a_{i1}(1)\quad a_{i1}(2)\ \ldots\ a_{i1}(n_i)\quad a_{i2}(1)\ \ldots\ a_{ip}(n_i)\quad d_i(1)\ \ldots\ d_i(n_i)]^T \tag{3.30}$$

$$Y_i(k) = [y_i(1)\quad y_i(2)\ \ldots\ y_i(k)]^T \tag{3.31}$$

and the subscript T denotes transposition. Note that the number of parameters to be estimated for each subsystem is $n_i(p+1)$.

For large values of k, the least-squares estimate of ϕ_i is given by

$$\hat{\phi}_i(k) = (H_i^T(k)\, H_i(k))^{-1}\, H_i^T(k)\, Y_i(k) \tag{3.32}$$

where $\hat{\phi}_i(k)$ is the estimate of ϕ_i from k input-output sequences. A recursive version of equation (3.32) for estimating $\hat{\phi}_i(k)$ was given by Astrom and Eykhoff [9] and Sinha and Pille [53] as follows

$$\hat{\phi}_i(k+1) = \hat{\phi}_i(k) + \frac{P_i(k)\, h_i(k+1)\, [y_i(k+1) - h_i^T(k+1)\, \hat{\phi}_i(k)]}{1 + h_i^T(k+1)\, P_i(k)\, h_i(k+1)} \tag{3.33}$$

$$P_i(k+1) = P_i(k) - \frac{P_i(k)\, h_i(k+1)\, (P_i(k)\, h_i(k+1))^T}{1 + h_i^T(k+1)\, P_i(k)\, h_i(k+1)} \tag{3.34}$$

where

$$H_i(k+1) = \begin{bmatrix} H_i(k) \\ \\ h_i^T(k+1) \end{bmatrix} \tag{3.35}$$

and

$$h_i(k+1) = [u_1(k)\; u_1(k-1)\; \ldots\; u_1(k-n_i)\; u_2(k)\; \ldots\; u_p(k-n_i)\; -y_i(k)\; \ldots\; -y_i(k-n_i)\,]^T \tag{3.36}$$

3.3.4.2 Recursive Estimation of n_i from Noise-free Data

For the ith subsystem of equation (3.21), let the square matrix $Q_i(k)$ of dimension $n_i(p+1)$ be defined as follows

$$Q_i(k) = I - H_i^+(k)\, H_i(k) \tag{3.37}$$

where $H_i^+(k)$ is the pseudo-inverse of the matrix $H_i(k)$ (equation (3.29)) and I is the unit matrix. A recursive version of equation (3.37) was given by Sinha and Pille [53] as follows

$$Q_i(k+1) = Q_i(k) - \frac{Q_i(k)\ h_i(k+1)\ (Q_i(k)\ h_i(k+1))^T}{h_i^T(k+1)\ Q_i(k)\ h_i(k+1)} \qquad (3.38)$$

where $$Q_i(0) = I \qquad (3.39)$$

and $h_i(k+1)$ is as defined in equation (3.36).

The algorithm to determine the order of each subsystem n_i is based on the following theorem.

Theorem

Consider the ith row (subsystem) of the matrix G(z) of equation (3.23) which can be written as

$$y_i(k) = \sum_{j=1}^{p} \frac{A^*_{ij}(z)}{D_i(z)}\ u_j(k) \qquad (3.40)$$

Assume that the order of each subsystem is N_i, where N_i is an integer and may be chosen arbitrarily large and let

$$q_i = \text{tr}\ Q_i(k) \qquad (3.41)$$

where $Q_i(k)$ is obtained recursively as in equations (3.38) and (3.39). If k is incremented from 1 to M_i (where $M_i \leq N_i(p+1)$) until q_i becomes constant, then the true order of each subsystem n_i is given by

$$n_i = N_i - q_i \qquad (3.42)$$

Proof:

It is known that (Albert, [56])

$$\text{Rank}\ H_i(k) = \text{tr}\ (H_i^+(k)\ H_i(k))$$

Then from equation (3.37) and (3.41) we have

$$\text{Rank } H_i(k) = \text{tr } (I - Q_i(k))$$

$$= N_i(p+1) - \text{tr } Q_i(k)$$

$$= N_i(p+1) - q_i$$

The maximum rank of $H_i(k)$ is $N_i p+n_i$ since $H_i(k)$ has $n_i(p+1)$ + $(N_i-n_i)p$ degrees of freedom. Thus, when $H_i(k)$ has attained the maximum rank, q_i becomes a constant and we have

$$N_i p+n_i = N_i(p+1) - q_i$$

and hence

$$n_i = N_i - q_i$$

3.3.5 Noisy Case

3.3.5.1 Parametric Identification in the Presence of Noise

In this subsection the algorithm previously considered for estimating the parameters of each subsystem will be extended to the noisy case. It will be assumed that the output sequence is corrupted by an additive, uncorrelated, zero-mean noise sequence as follows

$$z_i(k) = y_i(k) + v_i(k) \tag{3.43}$$

where $z_i(k)$ is the ith noisy output and $v_i(k)$ is the noise at the ith output.

Substituting for $y_i(k)$ from equation (3.43) into equation (3.27), the system outputs for the noisy case can be represented by the following equations

$$z_i(k) = \sum_{j=1}^{p} \sum_{\ell=1}^{n_i} a_{ij}(\ell)\, u_j(k-\ell) - \sum_{\ell=1}^{n_i} d_i(\ell)\, z_i(k-\ell) + \sum_{\ell=1}^{n_i} d_i(\ell)\, v_i(k-\ell)$$

$$+ v_i(k) \qquad i = 1, 2, \ldots, m \qquad (3.44)$$

Equation (3.44) can be written in a vector form as follows

$$z_i(k) = h_i^{*T}(k)\, \phi_i^* + v_i(k) \qquad i = 1, 2, \ldots, m \qquad (3.45)$$

where

$$h_i^*(k+1) = [u_1(k)\ u_1(k-1)\ \ldots\ u_1(k-n_i)\ \ldots\ u_p(k-n_i)\ -z_i(k)\ \ldots$$

$$\ldots\ -z_i(k-n_i)\ v_i(k)\ \ldots\ v_i(k-n_i)]^T \qquad (3.46)$$

$$\phi_i^* = [a_{i1}(1)\ a_{i1}(2)\ \ldots\ a_{i1}(n_i)\ \ldots\ a_{ip}(n_i)\ d_i(1)\ \ldots\ d_i(n_i)\ d_i(1)\ \ldots\ d_i(n_i)]^T \qquad (3.47)$$

The extended parameter vector ϕ_i^* for each subsystem can be estimated by the recursive least-squares algorithm (equations (3.33)-(3.36)) but the residuals $v_i(k)$'s are not known. However, a reasonable estimate of the residuals can be obtained as follows

$$\hat{v}_i(k) = z_i(k) - \hat{h}_i^{*T}\, \hat{\phi}_i^*(k) \qquad i = 1, 2, \ldots, m \qquad (3.48)$$

where $\hat{\phi}_i^*(k)$ is the estimate of ϕ_i^* at the kth iteration which is obtained by the recursive least-squares algorithm (equations (3.33)-(3.36)), and $\hat{h}_i^{*T}$ is constructed as in equations (3.46) but with the values of $v_i(k)$'s are substituted with their current estimates from equation (3.48). Assuming that the predictor of equation (3.48) is stable the convergence of the identification algorithm can be obtained in the same way as Panuska [55].

3.3.5.2 Estimation of n_i from Noisy Data by the Residual Error Technique

For the case of noisy data the method of subsection 3.3.4.2 will not work because the matrix $H_i(k)$ will be a full rank matrix and hence q = 0. In this subsection a nonrecursive method will be represented, which uses the residual error technique (Suen and [illegible], [54]), for estimating the orders n_i's for the case of noisy data. This residual error technique is described in Appendix I.

In this subsection the residual error technique discussed in Appendix I will be applied to determine the orders n_i's from noisy data. Using k input-output sequences and assuming the order of the ith subsystem to be ℓ_i and using equation (3.44) an expression similar to (3.28) can be obtained as

$$Y_i^*(k) = H_i^*(k, \ell_i)\ \phi_i(\ell_i) + W_i(k) \tag{3.49}$$

where

$$W_i(k) \triangleq [w_i(1) \quad w_i(2) \ \dots \ w_i(k)]^T \tag{3.50}$$

$$w_i(k) \triangleq \sum_{\ell=1}^{\ell_i} d_i(\ell)\ v_i(k-\ell) + v_i(k) \tag{3.51}$$

and $Y_i^*(k)$, $H_i^*(k, \ell_i)$ and $\phi_i(\ell_i)$ are defined as in equations (3.29)-(3.31) but with $y_i(k)$ and n_i replaced by $z_i(k)$ and ℓ_i respectively. Defining

$$Z_i(k) \triangleq Y_i^*(k) - W_i(k) \tag{3.52}$$

Then from equations (3.44) and (3.49) and the definition of $H_i^*(k, \ell_i)$ we can observe that the vector $Z_i(k)$ is a linear combination of the vectors of $H_i^*(k, \ell_i)$ if $\ell_i \geq n_i$ and is not if $\ell_i < n_i$. Following lemma 2 (Appendix I) we can obtain the following results

$$E\{e_i^o(\ell_i) \mid H_i^*(k, \ell_i)\} = g_i(\ell_i) \qquad \text{if } \ell_i \geq n_i \quad (3.53)$$

$$= g_i(\ell_i) + \Delta_i(\ell_i) \qquad \text{if } \ell_i < n_i \quad (3.54)$$

where

$$e_i^o(\ell_i) \triangleq Y_i^{*T}(k)\, [I - H_i^*(k, \ell_i)\, H_i^{*+}(k, \ell_i)]\, Y_i^*(k) \quad (3.55)$$

$$g_i(\ell_i) \triangleq E\, \{W_i^T(k)\, [I - H_i^*(k, \ell_i)\, H_i^{*+}(k, \ell_i)]\, W_i(k)\} \quad (3.56)$$

$$\Delta_i(\ell_i) \triangleq Z_i^T(k)\, [I - H_i^*(k, \ell_i)\, H_i^{*+}(k, \ell_i)]\, Z_i(k) \quad (3.57)$$

Let $\hat{e}_i^o(\ell_i)$ be the estimate of $E\{e_i^o(\ell_i)|H_i^*(k,\ell_i)\}$ where $e_i^o(\ell_i)$ is evaluated by equation (3.55) then equations (3.53) and (3.54) can be rewritten as follows

$$\hat{e}_i^o(\ell_i) = E\, \{e_i^o(\ell_i) \mid H_i^*(k, \ell_i)\} = f_i(\ell_i) + g_i(\ell_i) \quad (3.58)$$

where

$$f_i(\ell_i) = 0 \qquad \text{if } \ell_i \geq n_i \quad (3.59)$$

$$= \hat{\Delta}_i(\ell_i) > 0 \qquad \text{if } \ell_i < n_i \quad (3.60)$$

From equation (3.56) we can see that $g_i(\ell_i)$ is nearly constant if k is large enough (Suen and Liu, [54]) hence the order n_i can be estimated as follows:

For the ith output the residual error $\hat{e}_i^o(\ell_i)$ is plotted versus ℓ_i and from this plot n_i is obtained as the smallest integer ℓ_i for which the part of the plot is almost flat for $\ell_i \geq n_i$. In practice, instead of plotting the residual error $\hat{e}_i^o(\ell_i)$ it is better to plot, $e_i^*(\ell_i)$, the difference in the residual error defined as follows

$$e_i^*(\ell_i) \triangleq \hat{e}_i^o(\ell_i) - \hat{e}_i^o(\ell_i+1) \tag{3.61}$$

3.3.6 Results of Simulation

The proposed algorithm was applied to the identification of the following 2-input 2-output 4th order system

$$x(k+1) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -0.1 & 0.65 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 4 & -10 & -0.25 & 1 \end{bmatrix} x(k) + \begin{bmatrix} 1 & 2 \\ 0.25 & 0.8 \\ 0 & 0 \\ 0.9 & 0.9 \end{bmatrix} u(k)$$

$$z(k) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} x(k) + v(k) \tag{3.62}$$

The TFM representation, equation (3.22), of this system is

$$G(z) = \begin{bmatrix} \dfrac{1}{z-0.25} & \dfrac{2}{z-0.4} \\ \dfrac{0.9z+1.275}{(z-0.25)(z-0.5)^2} & \dfrac{0.9}{(z-0.5)^2} \end{bmatrix} \tag{3.63}$$

As in equation (3.23) the above TFM can be expressed in the following form

$$G(z) = \begin{bmatrix} \dfrac{A_{11}^*(z)}{D_1(z)} & \dfrac{A_{12}^*(z)}{D_1(z)} \\ \dfrac{A_{21}^*(z)}{D_2(z)} & \dfrac{A_{22}^*(z)}{D_2(z)} \end{bmatrix} \tag{3.64}$$

where

$$A^*_{11}(z) = z - 0.4$$

$$A^*_{12}(z) = 2z - 0.5$$

$$A^*_{21}(z) = 0.9z + 1.275$$

$$A^*_{22}(z) = 0.9z - 0.225$$

$$D_1(z) = z^2 - 0.65\,z + 0.1$$

$$D_2(z) = z^3 - 1.25z^2 + 0.5z - 0.0625$$

From the representation of G(z) in equation (3.64) we can see that the order of each subsystem (row) is

$$n_1 = 2 \quad \text{(order of } D_1(z))$$

$$n_2 = 3 \quad \text{(order of } D_2(z))$$

The above system was simulated using equation (3.62) with zero initial states and the system input was taken as uncorrelated zero-mean white noise sequence with unit variance. Each of the two outputs was contaminated with a zero-mean white noise sequence with standard deviation (σ) varies to vary the noise level at each output.

First the order of each subsystem was estimated from noise-free data using the algorithm of subsection 3.3.4.2. Then the algorithm of subsection 3.3.5.2 was applied to estimate the orders n_i's from noisy data for the two cases $\sigma = 0.1$ and $\sigma = 0.3$. The plot of the difference in the residual error $e^*_i(\ell_i)$ for each subsystem and for the two noise cases is shown in Figures 3.1 and 3.2. It is clearly determined from the Figures that $n_1 = 2$ and $n_2 = 3$ and we can notice that increasing the

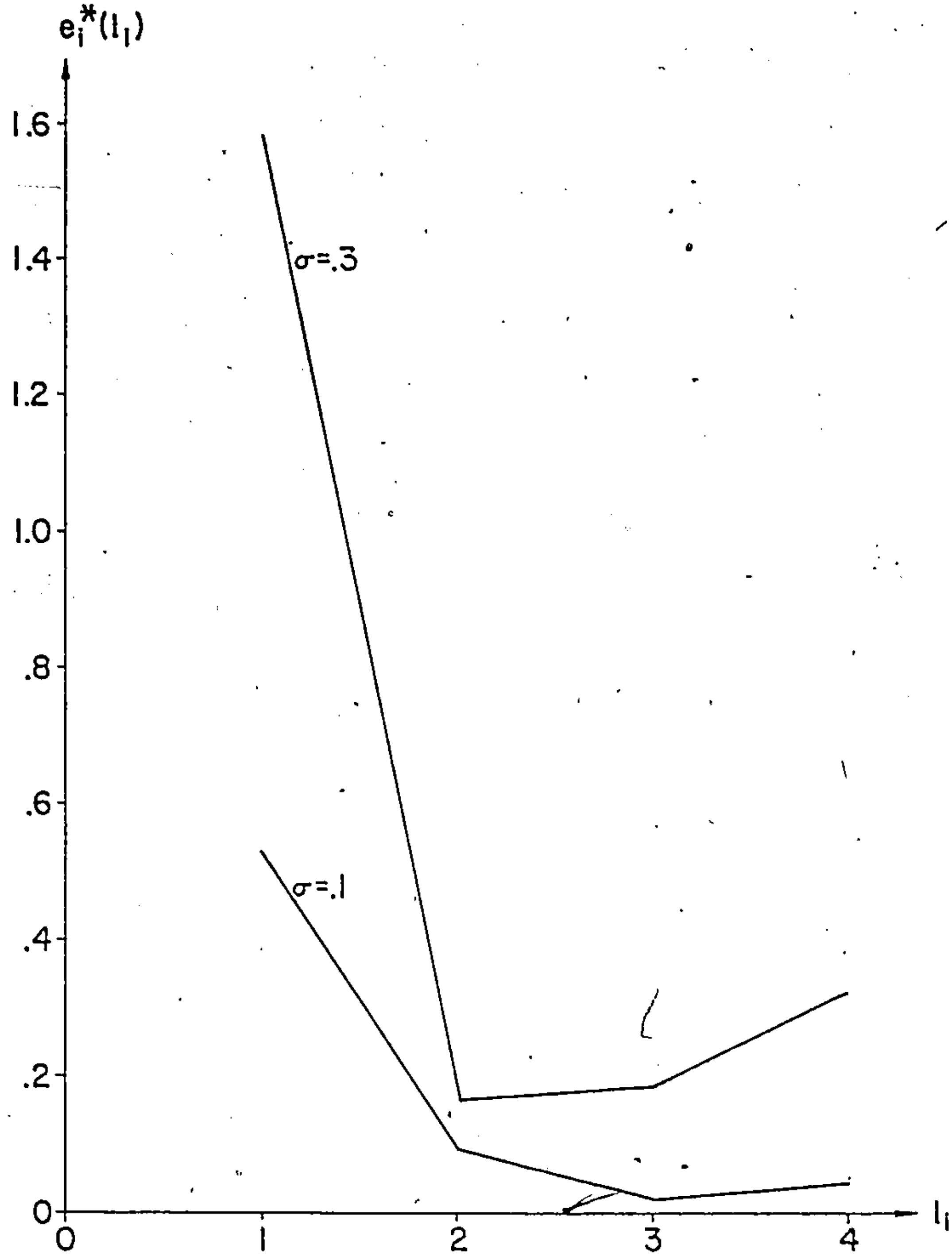

Fig. 3.1 Estimation of the structural index n_1 for the cases σ = 0.1 and σ = 0.3 .

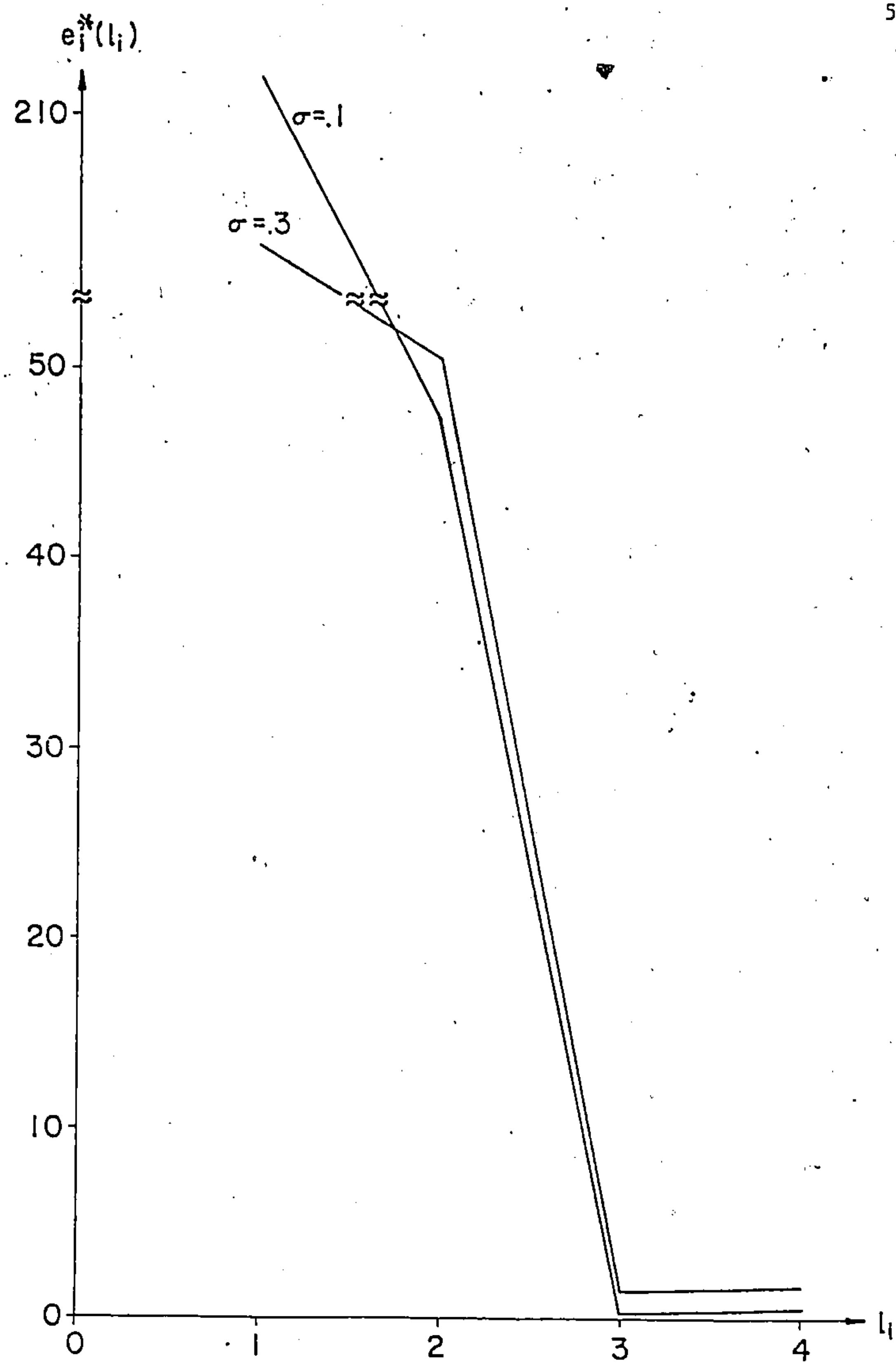

Fig. 3.2 Estimation of the structural index n_2 for the cases $\sigma = 0.1$ and $\sigma = 0.3$.

noise level will increase the residual difference error.

The parameters of the TFM were estimated by the algorithm of subsection 3.3.4.1 for the noise-free case and using the algorithm of subsection 3.3.5.1 for the two noisy cases of $\sigma = 0.1$ and $\sigma = 0.3$. The final estimate of the parameters after 900 iterations is shown in Table 3.4. To show the rate of convergence of the proposed algorithm of subsection 3.3.5.1 the normalized error $||\phi_i-\hat{\phi}_i(k)||^2/||\phi_i||^2$ of each subsystem has been plotted against the number of samples (k) for the two noise levels and is shown in Figures 3.3 and 3.4. As is seen from Table 3.4 exact estimates of the parameters have been obtained for noise-free case (and was obtained after few iterations). Increasing the noise level the estimates of the parameters are less accurate but stable and consistent estimates have been obtained for the two cases $\sigma = 0.1$ and $\sigma = 0.3$.

3.4 Concluding Remarks

In this chapter two different transfer-function matrix representations have been considered for identification. In section 3.2 a special transfer-function matrix has been identified where the order of each row was assumed to be equal to the order of the system. On the other hand, a general transfer-function matrix was identified in section 3.3 where the order of each row (in general less than or equal to the order of the system) was determined before parameter estimation for noise-free as well as the noisy case. The proposed algorithm of section 3.2 estimates the parameters of each row of the transfer-function matrix

Table 3.4 Estimate of the parameters after 900 iterations

Parameters	True value	Noise-Free Case	Case of σ=0.1	Case of σ=0.3
$a_{11}(1)$	1.0000	1.0000	1.0052	1.0168
$a_{11}(2)$	-0.4000	-0.4000	-0.3829	-0.2064
$a_{12}(1)$	2.0000	2.0000	1.9981	1.9936
$a_{12}(2)$	-0.5000	-0.5000	-0.4662	-0.1219
$d_1(1)$	-0.6500	-0.6500	-0.6315	-0.4568
$d_1(2)$	0.1000	0.1000	0.0929	0.0298
$a_{21}(1)$	0.0000	0.0000	0.0038	0.0130
$a_{21}(2)$	0.9000	0.9000	0.9054	0.9125
$a_{21}(3)$	1.2750	1.2750	1.2607	1.2370
$a_{22}(1)$	0.0000	0.0000	-0.0009	-0.0034
$a_{22}(2)$	0.9000	0.9000	0.8983	0.8961
$a_{22}(3)$	-0.2250	-0.2250	-0.2269	-0.2275
$d_2(1)$	-1.2500	-1.2500	-1.2543	-1.2598
$d_2(2)$	0.5000	0.5000	0.5050	0.5091
$d_2(3)$	-0.0625	-0.0625	-0.0639	-0.0631

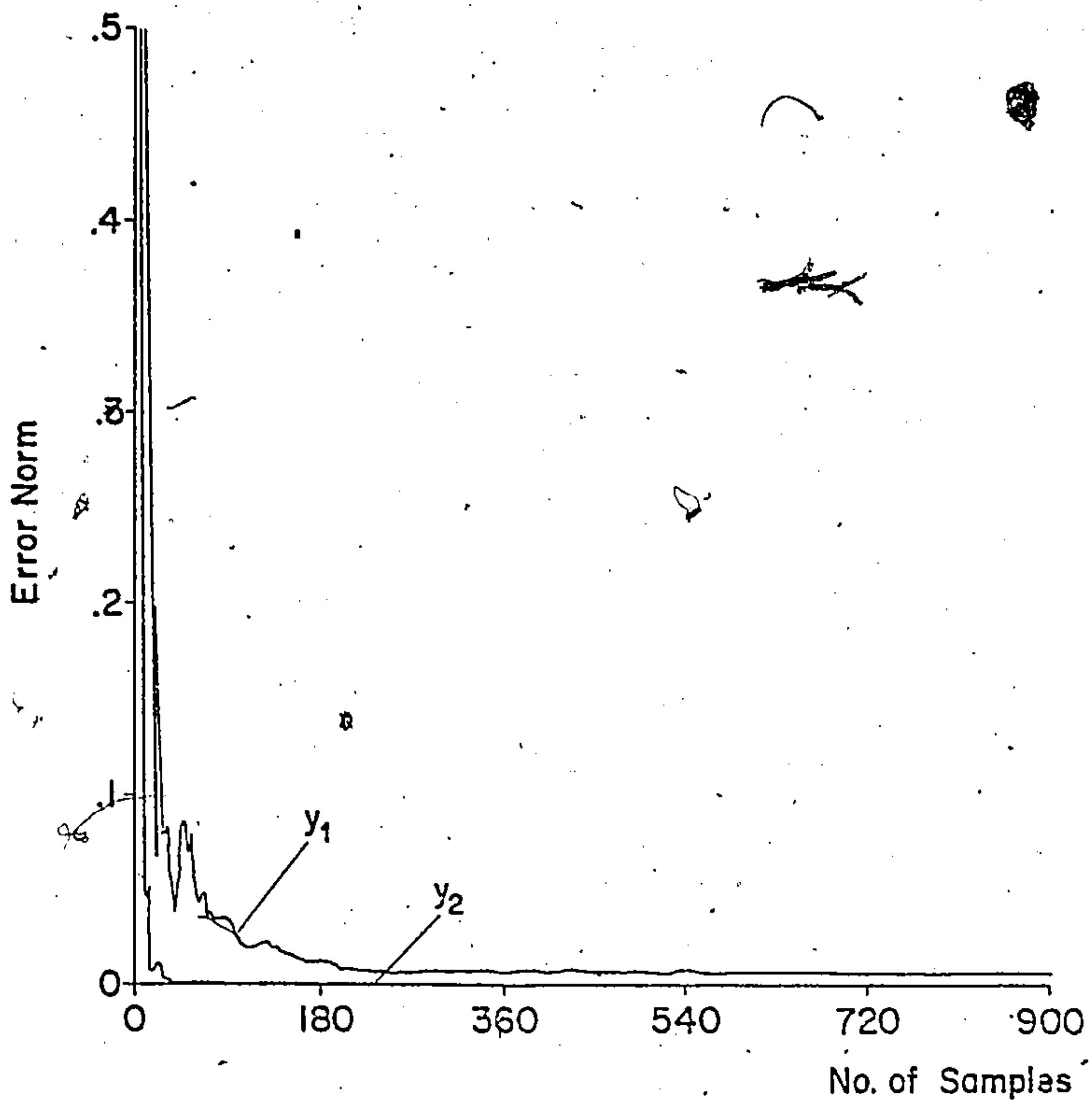

Fig. 3.3 Rate of convergence of the parameter estimates for the case $\sigma = 0.1$.

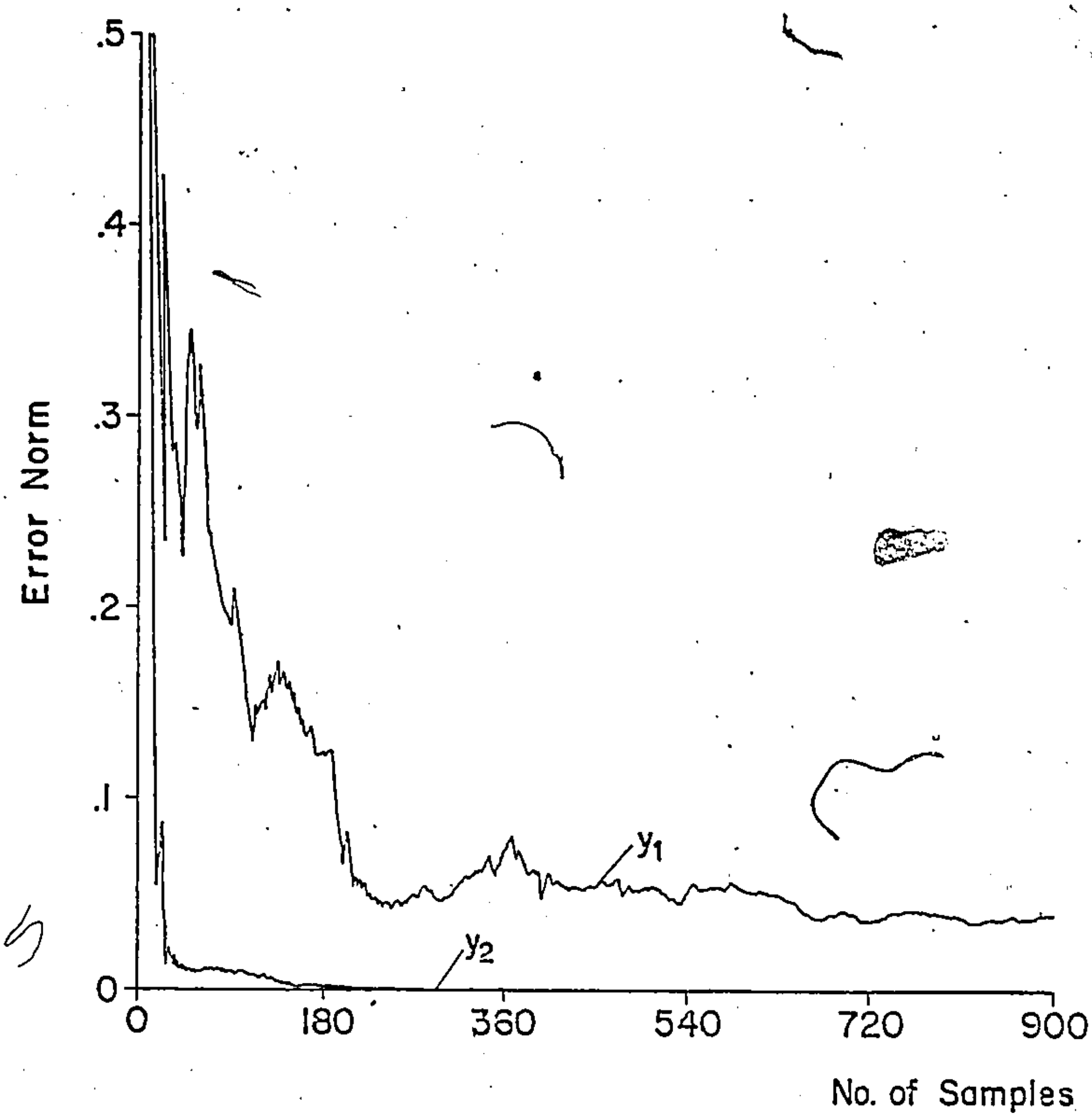

Fig. 3.4 Rate of convergence of the parameter estimates for the case $\sigma = 0.3$.

separately from other rows hence the matrix inversion required by Sen and Sinha's algorithm [28] is avoided and therefore a reduction in the computation time is achieved (Table 3.3).

As we can see from Tables 3.2 and 3.4, exact estimates of the parameters of the system were obtained for the case of noise-free data using the ordinary least-squares method. For the noisy case the adaptive least-squares method was used in section 3.3 to estimate the system parameters and good estimates were obtained for different noise levels (Table 3.4).

CHAPTER 4

IDENTIFICATION OF THE IMPULSE RESPONSE SEQUENCE

4.1 Introduction

The problem of identification of the impulse response sequence of the system has not received much attention. This is because most of the control theory uses the state space representation of the system, and the identified system in its impulse response has to be transformed into the state space form. The latter problem has been solved completely, first by Ho and Kalman [57] and later by Rozsa and Sinha [58], for the case where the Markov parameters of the system are exactly known. Bar-Shalom and Schwartz [59] have proposed a stochastic approximation algorithm for estimating the Markov parameters of single-input single-output linear systems. In their technique they used deterministic orthognal sequence of inputs. Mehra [34] has developed an on-line scheme for the identification of the impulse response sequence for multivariable systems, similar to the stochastic approximation scheme of Ho and Lee [60], from the Yule-Walker equations [61]. The weighting coefficients in this scheme are chosen recursively by solving a matrix equation. Sinha and Sen [62], have proposed an algorithm for the identification of the Markov parameters of equation (2.6) using a certain value for the parameter ℓ. Equation (2.6) can be rewritten in the form

$$z(k) = \sum_{i=0}^{\ell-1} J_i\, u(k-i-1) + \sum_{i=0}^{\ell-1} M_i\, e(k-i-1) \qquad (4.1)$$

where M_i's are the Markov parameters of the noise model and e(k) is a white noise sequence. They estimated the Markov parameters of the system J_i's using an on-line algorithm combining stochastic approximation and pseudo-inverse methods.

In this chapter two on-line algorithms will be developed for estimating the Markov parameters of linear discrete-time multivariable systems. In section 4.2 an algorithm is proposed for estimating the Markov parameters of the system from the measured input-output data, which are contaminated with additive noise, using a normalized stochastic approximation algorithm. In section 4.3 another algorithm is proposed to estimate the Markov parameters of the system by cross-correlation between the outputs and a white noise inputs and can be used with the system under operation if a dither signal can be added for identification. A state space representation of the system will be obtained from the estimated Markov parameters using an efficient algorithm for minimal realization [63].

4.2 Stochastic Approximation for the Estimation of the Markov Parameters

Consider a linear discrete-time multivariable system described by the following impulse response representation

$$z(k+\ell) = \sum_{j=1}^{\ell} J_j\, u(k+\ell-j) + v(k+\ell) \qquad (4.2)$$

where $z(k)$ is an m-dimensional measured output vector, $u(k)$ is an p-dimensional input vector sequence of independent random variables with zero-mean and $v(k)$ is an m-dimensional noise vector sequence of additive (not directly measured) random variables with zero-mean and finite variances, uncorrelated with $u(k)$, i.e.

$$E\{u(k)\} = 0, \quad E\{v(k)\} = 0$$

and

$$E\{u(k)\, v^T(k)\} = 0 \tag{4.3}$$

Defining the following ℓp dimensional vectors

$$\theta_i^T = [J_{1,i} \quad J_{2,i} \quad \cdots \quad J_{\ell,i}] \tag{4.4}$$

$$U^T(k+\ell-1) = [u^T(k+\ell-1) \quad u^T(k+\ell-2) \quad \cdots \quad u^T(k)] \tag{4.5}$$

where J_j, $j = 1, 2, \ldots, \ell$ has been partitioned as

$$J_j = \begin{bmatrix} J_{j,1} \\ J_{j,2} \\ \vdots \\ J_{j,m} \end{bmatrix} \tag{4.6}$$

Then the ith output of the system, $z_i(k+\ell)$, can be represented by the following equation

$$z_i(k+\ell) = U^T(k+\ell-1)\, \theta_i + v_i(k+\ell) \tag{4.7}$$

Now, assuming that the value of ℓ is known and defining $\hat{\theta}_i$ as the estimate of the unknown parameter vector θ_i of equation (4.4) which minimizes the following set of normalized mean-square error criterion

$$f(\hat{\theta}_i) = E\left\{\frac{v_i^2(k+\ell)}{||U(k+\ell-1)||^2}\right\} \quad i = 1, 2, \ldots, m$$

where

$$v_i(k+\ell) = z_i(k+\ell) - U^T(k+\ell-1)\,\theta_i \tag{4.8}$$

Then the parameters θ_i, $i = 1, 2, \ldots, m$ can be estimated recursively by means of the following normalized stochastic approximation algorithm

$$\hat{\theta}_i(k+\ell) = \hat{\theta}_i(k-1) + \nu(k)\,\frac{U(k+\ell-1)}{||U(k+\ell-1)||^2}\,[z_i(k+\ell) - U^T(k+\ell-1)\,\hat{\theta}_i(k-1)]$$

$$i = 1, 2, \ldots, m;\ k=1,\ \ell+2,\ 2\ell+3,\ \ldots \tag{4.9}$$

where the sequence $\nu(k)$ and the initial estimates satisfy the following conditions

$$\lim_{k\to\infty} \nu(k) = 0,\quad \sum_{k=1}^{\infty} \nu(k) = \infty,\quad \sum_{k=1}^{\infty} \nu^2(k) < \infty$$

and

$$E\{||\hat{\theta}_i(0)||^2\} < \infty \qquad i = 1, 2, \ldots, m \tag{4.10}$$

A proof of the unbiasedness of the estimate of the parameter vectors $\hat{\theta}_i$, $i = 1, 2, \ldots, m$ and their convergence in the mean-square sense to the true values θ_i's has been obtained and is given in Appendix II.

After estimating the unknown parameter vectors θ_i's, an estimate of Markov parameters of the system is obtained directly from relations (4.4) and (4.6) as follows

$$\hat{J}_j = \begin{bmatrix} \hat{\theta}_{j,1} \\ \hat{\theta}_{j,2} \\ \vdots \\ \hat{\theta}_{j,m} \end{bmatrix}$$

where $\hat{J}_j$ is the estimate of J_j and $\hat{\theta}_i$ has been partitioned as

$$\hat{\theta}_i^T = [\hat{\theta}_{1,i} \quad \hat{\theta}_{2,i} \quad \cdots \quad \hat{\theta}_{\ell,i}]$$

4.2.1 State Space Realization

After the Markov parameters of the system have been estimated, by the algorithm described above, a state space realization of the system (estimate of the matrices A, B and C, equation (2.15)) can be obtained from these estimated Markov parameters. A canonical realization of the matrices A, B and C can be obtained by an efficient algorithm for minimal realization proposed by Rozsa and Sinha [63] from the Hankel matrix of the system which consists of the estimated Markov parameters. Since the order of the system model is equal to the rank of the Hankel matrix, some difficulties may arise due to the fact that even a small perturbation in the Markov parameters may change the rank considerably. Hence, either the order of the system should be known in advance or one may obtain a partial realization of an arbitrary selected order from the estimated Hankel matrix.

4.2.2 Simulation Results

The proposed algorithm of this section was applied to the identification of a simulated 3rd order system with one input and two outputs described by the following equations

$$x(k+1) = \begin{bmatrix} 0 & 0 & 0.025 \\ 1 & 0 & -0.1 \\ 0 & 1 & 1.0 \end{bmatrix} x(k) + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u(k)$$

$$z(k) = \begin{bmatrix} 0.12 & 0.36 & 0.14 \\ 0.20 & 0.29 & 0.559 \end{bmatrix} x(k) + v(k)$$

The scalar input u(k) was taken as a zero-mean white noise sequence with unit variance. Each of the two outputs was contaminated with an additive zero-mean white noise of standard deviation of 0.3 and 0.85 for the first and second outputs respectively.

The Markov parameters of the above system were estimated using the algorithm of equation (4.9) and the value of the parameter ℓ was taken as $\ell = 10$. The final estimates of the first four Markov parameters obtained from 900 samples of the input-output data are shown in Table 4.1. To show the rate of convergence of the proposed algorithm the error norm of the Markov parameters has been plotted against the number of samples and is shown in Figure 4.1. The error norm used is defined as

$$\text{Error norm} \triangleq ||\hat{\theta}(k)-\theta||^2 \;/\; ||\theta||^2$$

where θ is a vector formulated from the Markov parameters to be estimated. From the estimated Markov parameters and assuming that the

Markov Parameter	True Value	Estimated Value
J_0	0.12	0.113
	0.20	0.204
J_1	0.36	0.361
	0.29	0.310
J_2	0.14	0.141
	0.559	0.586
J_3	0.107	0.112
	0.535	0.537

Table 4.1 Final estimate of the Markov parameters using stochastic approximation

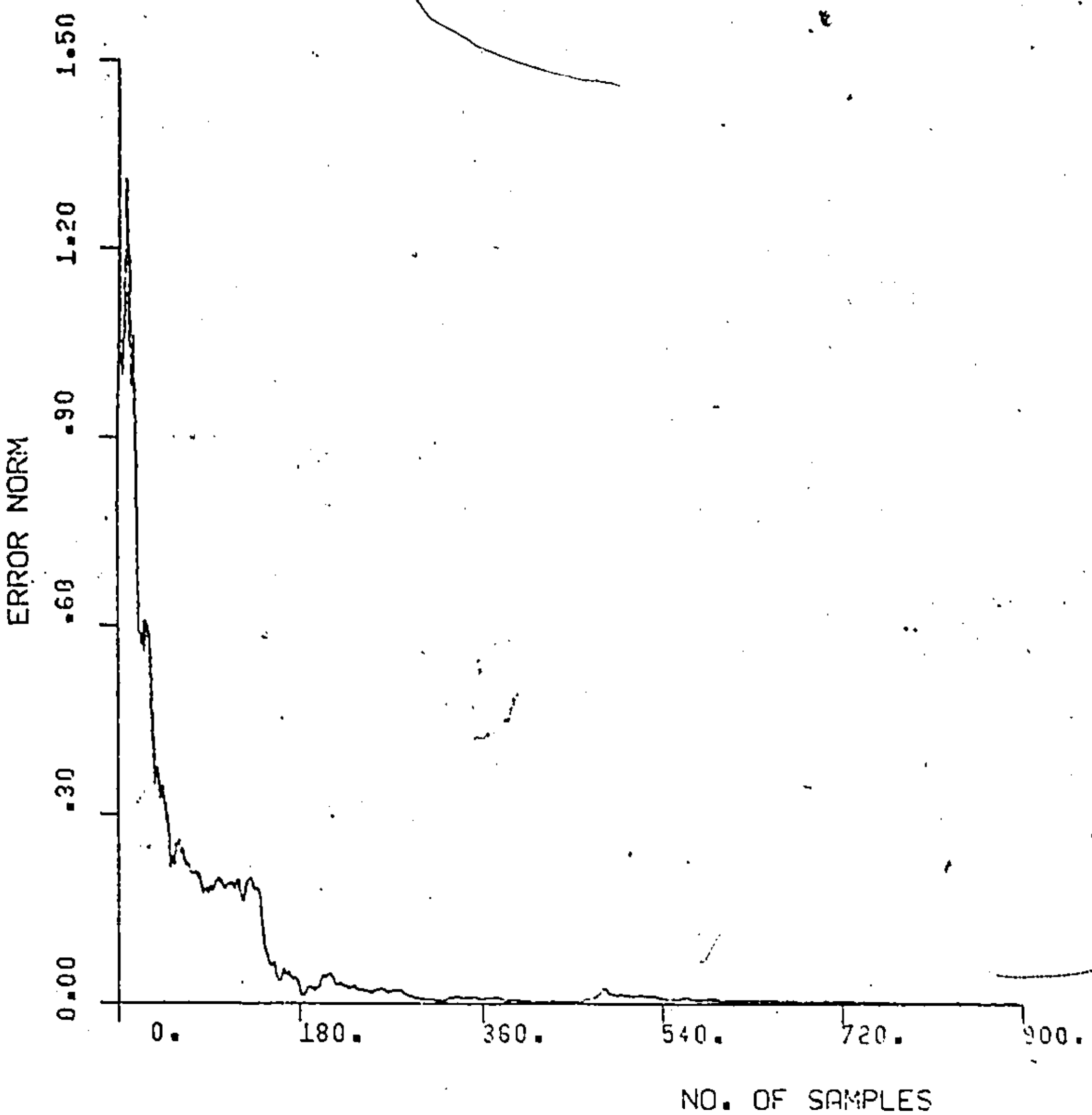

Fig. 4.1 Rate of convergence of the proposed stochastic approximation algorithm .

order of the system is known a state space representation of the system is obtained using the minimal realization algorithm proposed by Rozsa and Sinha [63]. Table 4.2 shows the obtained estimates of the matrices A and C.

4.3 Estimation of the Markov Parameters by Cross-correlation

Consider a linear discrete-time multivariable system described by the following impulse response representation

$$z(k) = \sum_{i=0}^{\ell-1} J_i \, u(k-i-1) + v(k) \tag{4.11}$$

Post-multiplying both sides of equation (4.11) by $u^T(j)$, where the superscript T denotes transposition, and taking expectation, we get

$$E\,\{z(k)\; u^T(j)\} = \sum_{i=0}^{\ell-1} J_i \, E\,\{u(k-i-1)\; u^T(j)\} + E\,\{v(k)\; u^T(j)\} \tag{4.12}$$

where E{•} denotes expected value.

Since the noise sequence v(k) is a zero-mean sequence uncorrelated with u(k), i.e.

$$E\,\{v(k)\; u^T(j)\} = 0 \tag{4.13}$$

then equation (4.12) can be reduced to

$$E\,\{z(k)\; u^T(j)\} = \sum_{i=0}^{\ell-1} J_i \, E\,\{u(k-i-1)\; u^T(j)\} \tag{4.14}$$

Now consider the case when the input sequence is uncorrelated, so that

$$E\,\{u(i)\; u^T(j)\} = I\,\delta_{ij}$$

Matrix	True Value			Estimated Value		
	0.0	0.0	0.025	0.0	0.0	0.015
A	1.0	0.0	-0.1	1.0	0.0	-0.089
	0.0	1.0	1.0	0.0	1.0	1.020
C	0.12	0.36	0.14	0.113	0.361	0.141
	0.20	0.29	0.559	0.204	0.310	0.586

Table 4.2 Estimate of the state space matrices

where I is the identity matrix and δ_{ij} is the kronecker delta. For this case equation (4.14) is reduced to

$$J_{k-j-1} = E\ \{z(k)\ u^T(j)\} \qquad (4.15)$$

Equation (4.15) indicates that the Markov parameter J_{k-j-1} is equal to the cross-correlation between the jth samples of the input sequence and the kth sample of the observed output sequence provided that the input sequence is a unit-variance white-noise sequence uncorrelated with the measurement noise. The expression for J_{k-j-1} of equation (4.15) is not suitable for practical application since it requires an ensemble average. Assuming ergodic process one may use the time average to obtain

$$J_{k-1} = \lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} z(k+j)\ u^T(j) \qquad (4.16)$$

For a good approximation, however, it is necessary that N, the number of samples be very large. This would normally require storing a large amount of data. The following recursive algorithm may be used for improving the estimates as more data arrives and may therefore be used for on-line estimation

$$J_{k-1}(N+1) = J_{k-1}(N) - \frac{1}{N+1}\,[J_{k-1}(N) - z(k+N+1)\ u^T(N+1)] \qquad (4.17)$$

where $J_k(N)$ is the estimate of J_k obtained from N samples of the input-output data.

It may be pointed out that although this proposed algorithm requires that the input sequence be white noise, this does not present any difficulty for those cases where it is necessary to identify the system under actual operation with some other inputs. In such cases it is usually permissible to add a white noise dither signal from a pseudo-random noise generator to the actual input. The cross-correlation between this signal and the observed output can be utilized as before, provided that this dither signal is uncorrelated with the other input signal and the observation noise.

After the Markov parameters have been estimated by the algorithm of equation (4.17) a state space realization can be obtained from these estimates as described in subsection 4.2.1.

4.3.1 Results of Simulation

The proposed algorithm of this section was applied to the identification of the simulated example of subsection 4.2.2. The input to the system was taken as a zero-mean unit variance white noise sequence. To each of the outputs an uncorrelated zero-mean white noise sequence with standard deviations of 0.3 and 0.85 was added to the first and second outputs respectively.

The final estimates of the Markov parameters from 1500 samples of the input-output data are given in Table 4.3. From these estimated Markov parameters and assuming that the order of the system is known, a state space representation of the system was obtained using the minimal realization algorithm of Rozsa and Sinha [63] and is shown in Table 4.4. As we can see from Tables 4.3 and 4.4 good estimates of the Markov parameters of the system and of the matrices A and C of the state space

Markov Parameter	True Value	Estimated value
J_0	0.12	0.121
	0.20	0.230
J_1	0.36	0.346
	0.29	0.315
J_2	0.14	0.140
	0.559	0.563
J_3	0.107	0.106
	0.535	0.539

Table 4.3 Final estimate of the Markov parameters using cross-correlation

Matrix	True Value			Estimated Value		
	0.0	0.0	0.025	0.0	0.0	0.027
A	1.0	0.0	-0.1	1.0	0.0	-0.095
	0.0	1.0	1.0	0.0	1.0	0.966
C	0.12	0.36	0.14	0.121	0.346	0.140
	0.20	0.29	0.559	0.230	0.315	0.563

Table 4.4 Estimate of the state space matrices

representation of the system have been obtained from the noisy measurements.

4.4 Concluding Remarks

In this chapter two algorithms have been developed for estimating the Markov parameters of multivariable systems. The algorithm of section 3.2 truncates the impulse response sequence after ℓ terms and a normalized stochastic approximation method is used to estimate the Markov parameters. As ℓ terms only are being used in the impulse response representation a bias will be introduced into the estimated Markov parameters due to this truncation. In order to reduce the effect of this truncation the value of ℓ may be increased [29]. The algorithm used in section 3.3 uses the correlation technique which does not require any truncation of the impulse response sequence and hence the problem of bias will not appear. On the other hand this algorithm requires a special type of input sequence for the purpose of identification.

After an estimate of the Markov parameters of the system has been obtained the matrices of the state space representation of the system can be calculated, if required, from the resulting Hankel matrix for the system. An efficient algorithm for this purpose has been proposed by Rozsa and Sinha [63].

CHAPTER 5

IDENTIFICATION OF THE STATE SPACE MODEL

5.1 Introduction

Due to the practical importance of state space representation, especially in control theory, much work has been done on the problem of identification of multivariable systems in the state space representation. In general, this problem can be divided into two main steps: structure determination and parameter estimation of the system matrices. It is well known that the representation in the state space form is not unique and any non-singular transformation of the state vector will lead to similar representation. Because of this non-uniqueness several canonical forms of the state space representation have been developed for the identification problem which reduce the number of parameters to be estimated in the system matrices (Weinert and Anton [46], Mayne [64] and Irwin and Roberts [65]).

In this chapter the problem of identification of linear multivariable discrete-time systems in the state space representation will be considered. A survey of some of the work done on the problem of identification of linear multivariable systems from input-output data in the state space representation will be presented in section 5.2. In section 5.3 an algorithm will be developed for determining the structural parameters of a canonical state space representation from noise-free as well as noisy data. The algorithm utilizes the residual error technique and does not need the knowledge of the noise characteristics. In section 5.4 an algorithm [43] will be developed for on-line estimation of the parameters of the canonical state space form used in section 5.3 from noisy data. This algorithm combines stochastic approximation with the pseudo-inverse method.

5.2 A Survey of State Space Identification Algorithms

First, consider the identification of multivariable systems when the input-output data is free from noise. The pioneering work in this area was done by Gopinath [24] and later by Budin [25]. Their methods present a direct procedure for minimal realization in a well-defined structure, from input-output data, which uses a selector matrix. Although this approach also gives the order of the system besides estimating the system matrices, it is computationally very involved.

Guidorzi [41], has identified the system in state space form by first estimating the parameters of an equivalent input-output difference equation form. Then the system matrices are recovered from these estimated parameters by a direct substitution.

Bingulac and Farias [48], have proposed an algorithm based on an identification identity relating input-output data to state space realization. By solving this identification identity it is possible in addition to system parameter estimation to determine observability indices and minimal order as well.

Liu and Suen [66], have proposed a simple algorithm by which a minimal dimension realization from input-output data can be obtained. The algorithm starts by constructing a sequential selector matrix by a minimal dimension algorithm, then a minimal dimension realization is obtained by direct substitution. This algorithm can also give a minimal dimension realization from an input-output sequence which may not be identifiable.

The identification of multivariable systems in state space model from noisy data is more involved than for the noise-free case, discussed above. For the noisy case there are two approaches, in the first the noise-free algorithms are modified, the second considers directly the identification from noisy data. Guidorzi [41], has extended his algorithm, discussed above to identify systems from noisy data. He proposed obtaining least-squares estimators in a manner similar to the instrumental variable approach. Recently, Sinha and Kwong [42] have developed a recursive algorithm which utilizes the canonical difference equation description due to Guidorzi [41]. This method uses the generalized least-squares algorithm to identify the parameters of each subsystem, which is a decomposition of the given system.

Mehra [34], has proposed an on-line algorithm for the identification of the state space model. First the autocovariance matrices of the system are estimated and then using the Ho and Kalman algorithm [67] a minimal realization is obtained. Valis [68], has obtained a vector difference equation of higher order from the state space description. A canonical form of the vector difference equation was given to enable the estimation of the order of the system if unknown. The order of the system was obtained by evaluating and testing (n+1) models, and the proper order is the one that gives a significant decrease of a certain loss function. The parameters of the system are estimated from the vector difference equation by least-squares.

Tse and Weinert [47], have proposed a procedure for structure determination and parameter estimation for multivariable stochastic systems where the control is identically zero. Their algorithm proceeds as follows; consider the following system model

$$\begin{aligned} x(k+1) &= Ax(k) + w(k) \\ z(k) &= Cx(k) + v(k) \end{aligned} \tag{5.1}$$

where w(k) and v(k) are zero-mean Gaussian noises.

System (5.1) can be represented by the following equation

$$\begin{aligned} x(k+1) &= Ax(k) + G\nu(k) \\ z(k) &= Cx(k) + \nu(k) \end{aligned} \tag{5.2}$$

where $\nu(k)$ is the zero-mean innovations with unknown covariance Q. By using a particular canonical form and defining the set of structural parameters of the system $\{p_i\}^m$, there exist a unique set $\{\beta_{ijk}\}$ such that for i = 1, 2, ..., m

$$\begin{aligned} c_i^T A^{p_i} &= \sum_{j=1}^{i} \sum_{k=0}^{p_j-1} \beta_{ijk} c_j^T A^k \quad \text{if} \quad p_i > 0 \\ c_i^T &= \sum_{j=1}^{i-1} \sum_{k=0}^{p_j-1} \beta_{ijk} c_j^T A^k \quad \text{if} \quad p_i = 0 \end{aligned} \tag{5.3}$$

where c_i^T is the ith row of C. Then the matrices A and C in the canonical form can be constructed from the set $\{p_i, \beta_{ijk}\}$. The structural parameters $\{p_i\}^m$ of the system are identified in the following way. Let P denote the covariance matrix of the states in (5.2), and let

$$R(\sigma) \equiv E\{z_{k+\sigma} z_k^T\}$$

Then (5.2) implies

$$\begin{aligned} P &= A P A^T + G Q G^T \\ R(0) &= C P C^T + Q \\ R(\sigma) &= C A^{\sigma-1} S \qquad \sigma > 0 \end{aligned} \tag{5.4}$$

where

$$S = A P C^T + G Q \tag{5.5}$$

Let $r_{ij}(\sigma)$ be the i, jth element of $R(\sigma)$, then (5.3) and (5.5) imply

$$r_{ij}(p_i+\tau) = \begin{cases} \sum_{\ell=1}^{i} \sum_{k=0}^{p_\ell-1} \beta_{i\ell k} r_{\ell j}(k+\tau) & p_i > 0 \\ \sum_{\ell=1}^{i-1} \sum_{k=0}^{p_\ell-1} \beta_{i\ell k} r_{\ell j}(k+\tau) & p_i = 0 \end{cases} \tag{5.6}$$

where $\tau = 1, 2, \ldots$. Now for the case i = 1 and $\tau = 1, 2, \ldots, p_1$ we have from (5.6)

$$r_1 = \phi_1(p_1)\beta_1$$

where

$$r_1^T = [r_{ij}(p_1+1) \ \ldots \ r_{ij}(2p_1)]$$

$$\beta_1^T = [\beta_{11,0} \ \cdots \ \beta_{11,p_1-1}]$$

$$\phi_1(k) = \begin{bmatrix} r_{ij}(1) & r_{ij}(2) & \cdots & r_{ij}(k) \\ r_{ij}(2) & r_{ij}(3) & \cdots & r_{ij}(k+1) \\ \vdots & \vdots & & \vdots \\ r_{ij}(k) & \cdots & & r_{ij}(2k-1) \end{bmatrix}$$

Let

$$d_1(k) = |\det \phi_1(k)|$$

Then we can see that

$$d_1(k) > 0 \quad \text{for} \quad k = 1, 2, \ldots, p_1$$

$$d_1(k) = 0 \quad \text{for} \quad k > p_1$$

Using an estimate $\hat{R}(\sigma)$ of $R(\sigma)$ as,

$$\hat{R}(\sigma) = \frac{1}{N} \sum_{k=1}^{N-\sigma} z_{k+\sigma} \, z_k^T$$

p_1 can be estimated as follows; if the first sharp decrease in $d_1(k)$ occurs at $k = k^*$, then $\hat{p}_1$ is chosen as k^*. The estimate of β_1 is obtained from the equation

$$\hat{r}_1 = \hat{\phi}_1(\hat{p}_1)\hat{\beta}_1$$

For $i = 2, 3, \ldots, m$; $\hat{p}_i$ and $\hat{\beta}_i$ are computed in an analogous manner. Finally the matrices G and Q are estimated by an algorithm similar to that of Carew and Belanger [69].

Nelson and Stear [70], have proposed a scheme for combined parameter and state estimation for multivariable systems. The parameters of the system are estimated separately by a linear Kalman filter. The separation is possible because of the use of a canonical form of the state space equations of the system. Each component of the output,

$z_i(k)$ can be written in terms of the system inputs, outputs, innovations and a set of parameters arranged in a vector denoted by θ_i. This relationship has the form

$$z_i(k) = E^T(k)\ \theta_i(k) + \nu_i(k) \tag{5.7}$$

where

$$E^T(k) = [u^T(k-1),\ldots,u^T(k-q),z^T(k-1),\ldots,z^T(k-q),\nu^T(k-1),\ldots,\nu^T(k-q)]$$

Assuming that any time variability of the parameters $\theta_i(k)$ can be modeled by a random walk, that is

$$\theta_i(k+1) = \theta_i(k) + s_i(k) \tag{5.8}$$

where $s_i(k)$ is assumed to be a zero-mean, white, Gaussian process with covariance $S(k)$. By applying Kalman's results [71] to equations (5.7) and (5.8) we obtain the following estimate for the parameter θ_i.

$$\hat{\theta}_i(k) = \hat{\theta}_i(k-1) + K_i(k-1)\ [z_i(k) - E^T(k-1)\ \hat{\theta}_i(k-1)]$$

$$K_i(k) = P_i(k-1)\ E(k)\ [E^T(k)\ P_i(k-1)\ E(k) + 1]^{-1}$$

$$P_i(k) = [I - K_i(k)\ E^T(k)]\ P_i(k-1) + S(k)$$

The innovations sequence can be estimated by the following algorithm, which is due to Panuska [55],

$$\hat{\nu}_i(k) = z_i(k) - \hat{E}^T(k-1)\ \hat{\theta}_i(k)$$

After estimating the parameters θ_i's, the coefficients of the matrices A, B and C can be obtained by a certain transformation.

Martin and Stubberud [72], have proposed an uncoupling method for the identification of the system matrices and the noise covariance matrices. A Kalman filter, predicated on the best available knowledge of system parameters is constructed. The matrices A and B are identified by requiring that the mean of the measurement residual sequence be zero. An adaptive stochastic approximation algorithm is used to iteratively adjust the system parameters so that the above requirement is satisfied.

Bohn and DeBeer [73], have proposed an algebraic approach to obtain a consistent parameter estimate of the state space model. They used a special canonical structure for the system matrices which reduces the dimension of the data matrix and decouples the parameter estimation equations into independent subsystem equations. Due to the decoupling the residual equations take a very simple algebraic form. The solution of this residual equations in conjunction with least-squares estimates yields consistent parameter estimates.

Blessing [74], has proposed a procedure for structure determination and parameter estimation. The structural parameters of the system are first estimated by an eigenvalue test. The system parameters are estimated by first identifying the Markov parameters of the system by correlation analysis, using quaternary input sequences simultaneously, and then a least-squares fit.

DeLarminat and Doncarli [75], have proposed a real time generalized least-squares method for estimating the parameters and the optimal filter gain of the system. They assumed that the structural parameters of the system are known in advance. Their method starts by transforming the state space model into the input-output model. Under a special canonical form an explicit passing formula between the two models has been obtained by Salut and Gavier [76], and Doncarli [77]. The parameters of the input-output model were identified recursively using an estimation of the output errors in a manner similar to the stochastic approximation method. Finally, a canonical state space form of the system is obtained from the identified parameters of the input-output model.

5.3 A Proposed Algorithm for Structure Determination of a Canonical State Space Model

In this section an algorithm [78] is presented for determining the structural parameters of a certain canonical state space model which has been used much for identification (Lobbia and Saridis [45], Beghelli and Guidorzi [79] and El-Sherief and Sinha [80]). The structural parameters are obtained for the noise-free case as well as the noisy case. The structural parameters of this canonical form have been obtained by Guidorzi [41] using the determinant test for the noise-free case. He has also extended the algorithm to the case of noisy data using the enhanced information matrix with some assumptions about the noise characteristics. The proposed algorithm utilizes the residual error technique which has been used by Suen and Liu [54] for determining the structural parameters of a different canonical form without knowing the noise characteristics.

The use of the residual error technique for determining the structural parameters has the following main advantages over the determinant test method used by Lobbia and Saridis [45] and Guidorzi [41].

a) The problem of determining the structural parameters is decoupled where every parameter is determined independently from the others. Hence an error in estimating one parameter will not affect the estimate of the others.

b) It will not encounter the difficulty of rank checking required by the determinant test method.

5.3.1 Formulation of the Problem

Consider the following multivariable discrete-time, completely observable system

$$\begin{aligned} x^*(k+1) &= A^* x^*(k) + B^* u(k) \\ y(k) &= C^* x^*(k) \end{aligned} \tag{5.9}$$

where $x^*(k) \in R^n$, $u(k) \in R^p$ and $y(k) \in R^m$ are the state, input and output vectors respectively. Let,

$$C^* \triangleq \begin{bmatrix} c_1^{*T} \\ c_2^{*T} \\ \vdots \\ c_m^{*T} \end{bmatrix} \tag{5.10}$$

Construct the vector sequences

$$C_1^*,\ A^{*T} C_1^*,\ A^{*T^2} C_1^*,\ \ldots,\ \ldots$$
$$\vdots$$
$$C_m^*,\ A^{*T} C_m^*,\ \ldots \qquad (5.11)$$

and select them in the following order

$$C_1^*,\ C_2^*,\ \ldots,\ C_m^*,\ A^{*T} C_1^*,\ \ldots,\ A^{*T} C_m^*,\ A^{*T^2} C_1^*,\ \ldots \qquad (5.12)$$

retaining a vector $A^{*T^i} C_j^*$ if and only if it is independent from all previously selected ones. Let $n_1, n_2, \ldots, n_m$ be the number of vectors selected from the first, second, ..., mth sequence in (5.11) [called the structural parameters of the system]. Because of the complete observability of the system it follows that

$$n_1 + n_2 + \ldots + n_m = n \qquad (5.13)$$

Using the following state transformation [81] $x^*(k) = S\ x(k)$ where

$$S = [C_1^*,\ A^{*T} C_1^*,\ \ldots,\ A^{*T^{n_1-1}} C_1^*,\ \ldots,\ A^{*T^{n_m-1}} C_m^*] \qquad (5.14)$$

system (5.9) is transformed to the following canonical form, Guidorzi [41]

$$\begin{aligned} x(k+1) &= A\ x(k) + B\ u(k) \\ y(k) &= C\ x(k) \end{aligned} \qquad (5.15)$$

where $\quad A = S\ A^*\ S^{-1} = \{A_{ij}\} \qquad i,j = 1,\ 2,\ \ldots,\ m$

$$A_{ii} = \begin{bmatrix} 0 & I_{n_i-1} \\ \vdots & \\ a_{ii}(1) & \dots \; a_{ii}(n_i) \end{bmatrix} \qquad \underset{i \neq j}{A_{ij}} = \begin{bmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ a_{ij}(1) & \dots & a_{ij}(n_{ij}) \end{bmatrix}$$

$$C = \begin{bmatrix} e^{1} \\ e^{n_1+1} \\ \vdots \\ e^{n_1+n_2+\dots+n_{m-1}+1} \end{bmatrix} \qquad B = S\,B^{*} \triangleq \begin{bmatrix} b^{1} \\ b^{2} \\ \vdots \\ b^{n} \end{bmatrix}$$

and e^{i} is the unit row vector and n_{ij} is the number of nonzero coefficients in A_{ij} which because of the order followed in the selection of the vectors in (5.12) and the construction of the matrix S is at most equal to (Guidorzi [41])

$$\begin{aligned} n_{ij} &\leq n_i+1 \qquad \text{if} \qquad i > j \\ n_{ij} &\leq n_i \qquad \text{if} \qquad i < j \end{aligned} \tag{5.16}$$

The above state space canonical form of the system is equivalent to the following canonical input-output difference equation representation (Guidorzi [41])

$$P(z)\, y(k) = Q(z)\, u(k) \tag{5.17}$$

where

$$p_{ii}(z) = z^{n_i} - a_{ii}(n_i)\, z^{n_i-1} - \dots - a_{ii}(1)$$

$$p_{ij}(z) = -\, a_{ij}(n_{ij})\, z^{n_{ij}-1} - \dots - a_{ij}(1)$$

$$q_{ij}(z) = d_{n_1+\ldots+n_i,j}\, z^{n_i-1} + \ldots + d_{n_1+\ldots+n_{i-1}+2,j}\, z + d_{n_1+\ldots+n_{i-1}+1,j}$$

and the parameters $d_{i,j}$'s can be obtained from the matrices A and B (Guidorzi [41]).

Remark 1:

It can be noticed from equation (5.17) and relations (5.16) that the orders of the polynomials P(z) and Q(z) satisfy the following relations

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

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

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

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

From equation (5.17) the ith subsystem [row of the matrices P(z) and Q(z)] of the main system is described by the following equation

$$\sum_{j=1}^{m} p_{ij}(z)\, y_j(k) = \sum_{j=1}^{p} q_{ij}(z)\, u_j(k) \tag{5.18}$$

which can be rewritten more explicitly as

$$y_i(k+n_i) = \sum_{j=1}^{m} \sum_{\ell=1}^{n_{ij}} a_{ij}(\ell)\, y_j(k+\ell-1) + \sum_{j=1}^{p} \sum_{\ell=1}^{n_i} d_{(n_1+\ldots+n_{i-1}+\ell),j}\, u_j(k+\ell-1) \tag{5.19}$$

where the number $n_{ii} \triangleq n_i$.

Remark 2:

It can be noticed that n_i is the order of equation (5.19) and $y_i(k+n_i)$ for $k > 1$ is linearly dependent on

$$y_1(k+n_{i1}-1),\ y_1(k+n_{i1}-2),\ \ldots,\ y_1(k),\ y_2(k+n_{i2}-1),\ \ldots,\ y_i(k+n_i-1),\ \ldots,$$
$$y_m(k),\ u_1(k+n_i-1),\ u_1(k+n_i-2),\ \ldots,\ u_p(k).$$

this can be seen from equation (5.19) and remark 1.

Our objective is to determine the structural parameters n_1, n_2, ..., n_m from noise-free as well as noisy measurements without knowing the noise characteristics.

5.3.2 The Noise-free Case

Suppose the set of vectors $\{y, x_1, x_2, \ldots, x_n\}$ is given and we want to study the problem whether or not the vector y is a linear combination of the set of vectors $\{x_1, x_2, \ldots, x_n\}$. This problem is discussed in Appendix I.

In this subsection the residual error technique discussed in Appendix I will be applied to determine the structural parameters of the system (5.9) from noise-free data. For the ith subsystem, equation (5.19), and from K input-output sequences and assuming the order of the ith subsystem to be ℓ_i we get

$$Y_i(K) = H_i(\ell_i, K)\ \theta_i(\ell_i) \tag{5.20}$$

where

$$H_i(\ell_i,K) = \begin{bmatrix} y_1(k) & .. & y_1(k-\ell_i) & y_2(k) & .. & y_i(k-1) & .. & y_i(k-\ell_i) & .. \\ y_1(k+1) & .. & y_1(k-\ell_i+1) & y_2(k+1) & .. & y_i(k) & .. & y_i(k-\ell_i+1) & .. \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots & \\ y_1(k+K) & .. & y_1(k+K-\ell_i) & y_2(k+K) & .. & y_i(k+K-1) & . & y_i(k+K-\ell_i) & . \end{bmatrix}$$

$$\begin{matrix} .. & y_{i+j}(k-1) & .. & u_1(k-1) & .. & u_p(k-\ell_i) \\ .. & y_{i+j}(k) & .. & u_1(k) & .. & u_p(k-\ell_{i+1}) \\ & \vdots & & \vdots & & \vdots \\ .. & y_{i+j}(k+K-\ell) & .. & u_1(k+K-1) & .. & u_p(k+K-\ell_i) \end{matrix}$$

$$Y_i(K) = [y_i(k) \; y_i(k+1) \; \ldots \; y_i(k+K)]^T$$

and $\theta_i(\ell_i)$ is a vector of parameters characterizing the ith subsystem.

From remark 2 and equation (5.19) we can see that the vector $Y_i(K)$ is a linear combination of the vectors of $H_i(\ell_i,K)$ if $\ell_i \geq n_i$ and it is not if $\ell_i < n_i$. From equations (A1.6) (Appendix I) and (5.20) we can get

$$e_i^o(\ell_i) = Y_i^T(K) \; [I - H_i(\ell_i,K) \; H_i^+(\ell_i,K)] \; Y_i(K) \triangleq \Delta_i(\ell_i) \tag{5.21}$$

Following lemma 1 (Appendix I) we get

$$\begin{aligned} e_i^o(\ell_i) &= 0 & \ell_i \geq n_i \\ &= \Delta_i(\ell_i) > 0 & \ell_i < n_i \end{aligned} \tag{5.22}$$

Hence we have

<u>Estimation rule 1</u>

For the ith output (subsystem) the residual $e_i^o(\ell_i)$ is plotted versus ℓ_i. From this plot n_i is obtained as the smallest integer ℓ_i for which $e_i^o(\ell_i) = 0$.

5.3.3 The Noisy Case

Let a vector y be corrupted by a zero-mean noise vector v and let y^* be the noisy observation

$$y^* = y + v \tag{5.23}$$

From a given set of vectors $\{y^*, x_1, x_2, \ldots, x_n\}$ we want to study the problem whether or not y is a linear combination of the set of vectors $\{x_1, x_2, \ldots, x_n\}$. This problem is discussed in Appendix I.

In this subsection the residual error technique discussed in Appendix I will be applied to determine the structural parameters of system (5.9) from noisy data. First, consider system (5.9) for noisy measurements where

$$z_i(k) = y_i(k) + v_i(k) \qquad i = 1, 2, \ldots, m \tag{5.24}$$

where $z_i(k)$ is the ith noisy output and $v_i(k)$ is the noise at the ith output which is a zero-mean white noise sequence. Then substituting for $y_i(k)$ from equation (5.24) into equation (5.19) we get

$$z_i(k+n_i) = \sum_{j=1}^{m} \sum_{\ell=1}^{n_{ij}} a_{ij}(\ell)\, z_j(k+\ell-1) + \sum_{j=1}^{p} \sum_{\ell=1}^{n_i} d_{(n_1+..+n_{i-1}+\ell),j}\, u_j(k+\ell-1) + w_i(k+n_i) \tag{5.25}$$

where

$$w_i(k+n_i) = -\sum_{j=1}^{m} \sum_{\ell=1}^{n_{ij}} a_{ij}(\ell)\, v_j(k+\ell-1) + v_i(k+n_i)$$

and $w_i(k)$ is a zero-mean noise sequence.

Using K input-output sequences and assuming the order of the ith subsystem [equation (5.25)] is ℓ_i and using equation (5.25) an expression similar to (5.20) can be obtained as follows

$$Y_i^*(K) = H_i^*(\ell_i,K)\ \theta_i(\ell_i) + W_i(K) \tag{5.26}$$

where

$$W_i(K) \triangleq [w_i(k)\ \dots\ w_i(k+1)\ \dots\ w_i(k+K)]^T$$

$Y_i^*(K)$ and $H_i^*(\ell_i,K)$ are defined as in equation (5.20) where $y_i(k)$ is replaced by $z_i(k)$. Defining

$$Z_i(K) \triangleq Y_i^*(K) - W_i(K) \tag{5.27}$$

then from equation (5.25) and the definition of $H_i^*(\ell_i,K)$ we can observe that the vector $Z_i(K)$ is a linear combination of the vectors of $H_i^*(\ell_i,K)$ if $\ell_i \geq n_i$ and is not if $\ell_i < n_i$. Following lemma 2 (Appendix I) we can obtain the following results

$$E\{e_i^o(\ell_i)\ |\ H_i^*(\ell_i,K)\} = g_i(\ell_i) \qquad \text{if } \ell_i \geq n_i \tag{5.28}$$

$$= g_i(\ell_i) + \Delta_i(\ell_i) \qquad \text{if } \ell_i < n_i \tag{5.29}$$

where

$$e_i^o(\ell_i) \triangleq Y_i^{*T}(K)\ [I - H_i^*(\ell_i,K)\ H_i^{*+}(\ell_i,K)]\ Y_i^*(K) \tag{5.30}$$

$$g_i(\ell_i) \triangleq E\ \{W_i^T(K)\ [I - H_i^*(\ell_i,K)\ H_i^{*+}(\ell_i,K)]\ W_i(K)\} \tag{5.31}$$

$$\Delta_i(\ell_i) \triangleq Z_i^T(K)\ [I - H_i^*(\ell_i,K)\ H_i^{*+}(\ell_i,K)]\ Z_i(K) \tag{5.32}$$

Let $\hat{e}_i^o(\ell_i)$ be the estimate of $E\{e_i^o(\ell_i) \mid H_i^*(\ell_i,K)\}$ where $e_i^o(\ell_i)$ is evaluated by equation (5.30) from the given input-output sequences. Then equations (5.28) and (5.29) can be rewritten as follows

$$\hat{e}_i^o(\ell_i) = E\{e_i^o(\ell_i) \mid H_i^*(\ell_i,K)\} = f_i(\ell_i) + g_i(\ell_i) \tag{5.33}$$

where

$$f_i(\ell_i) = 0 \qquad \text{if } \ell_i \geq n_i \tag{5.34}$$

$$= \hat{\Delta}_i(\ell_i) > 0 \qquad \text{if } \ell_i < n_i \tag{5.35}$$

From equation (5.31) we can see that $g_i(\ell_i)$ is nearly constant if K is large enough (Suen and Liu [54]). Therefore, the residual plot $\hat{e}_i^o(\ell_i)$ versus ℓ_i for the noisy case has the same shape as the residual plot for the noise-free case except that it is raised by a nearly constant value.

Estimation rule 2

From the ith output the residual error $\hat{e}_i^o(\ell_i)$ is plotted versus ℓ_i. From this plot n_i is obtained as the smallest integer ℓ_i for which the part of the residual plot for $\ell_i \geq n_i$ is almost flat.

In practice, instead of plotting the residual error $\hat{e}_i^o(\ell_i)$ it is better to plot the difference in the residual error $e_i^*(\ell_i)$ versus ℓ_i where

$$e_i^*(\ell_i) \triangleq \hat{e}_i^o(\ell_i) - \hat{e}_i^o(\ell_i+1) \tag{5.36}$$

5.3.4 Results of Simulation

Example 1

The proposed algorithm was applied to determine the structural parameters of the following 4th order two-input two-output system

$$x(k+1) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -0.1 & 0.65 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -2/3 & 5/3 & -0.25 & 1 \end{bmatrix} x(k) + \begin{bmatrix} 0 & 2 \\ 0.25 & 0.8 \\ 0 & 0 \\ 1 & 1 \end{bmatrix} u(k)$$

$$z(k) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} x(k) + v(k) \qquad (5.37)$$

The above system was simulated on a CDC 6400 where the input vector u(k) was taken as a zero-mean white noise sequence with unit variance and the noise v(k) was taken as a zero-mean uncorrelated noise sequence with standard deviation (σ) varies to vary the noise level at the output.

The above system is in the canonical form of (5.15) where the structural parameters are

$$n_1 = 2 \text{ and } n_2 = 2$$

The input-output representation of the above system [equation (5.17)] is as follows

$$\begin{bmatrix} z^2 - 0.65\,z + 0.1 & 0 \\ -5/3\,z + 2/3 & z^2 - z + 0.25 \end{bmatrix} y(k) = \begin{bmatrix} z - 0.4 & 2z - 0.5 \\ -2/3 & -7/3 \end{bmatrix} u(k) \qquad (5.38)$$

The proposed algorithm was applied to determine the structural parameters of the above system for three different noise levels with σ = 0.1, 0.2 and 0.3 and a sequence length K = 30 which is very short compared with similar studies. The residual plots, e_i's, for the two outputs are given in Figures 5.1a and 5.2a. It is clearly seen from the figures that $n_1 = n_2 = 2$ and with increasing the noise level the

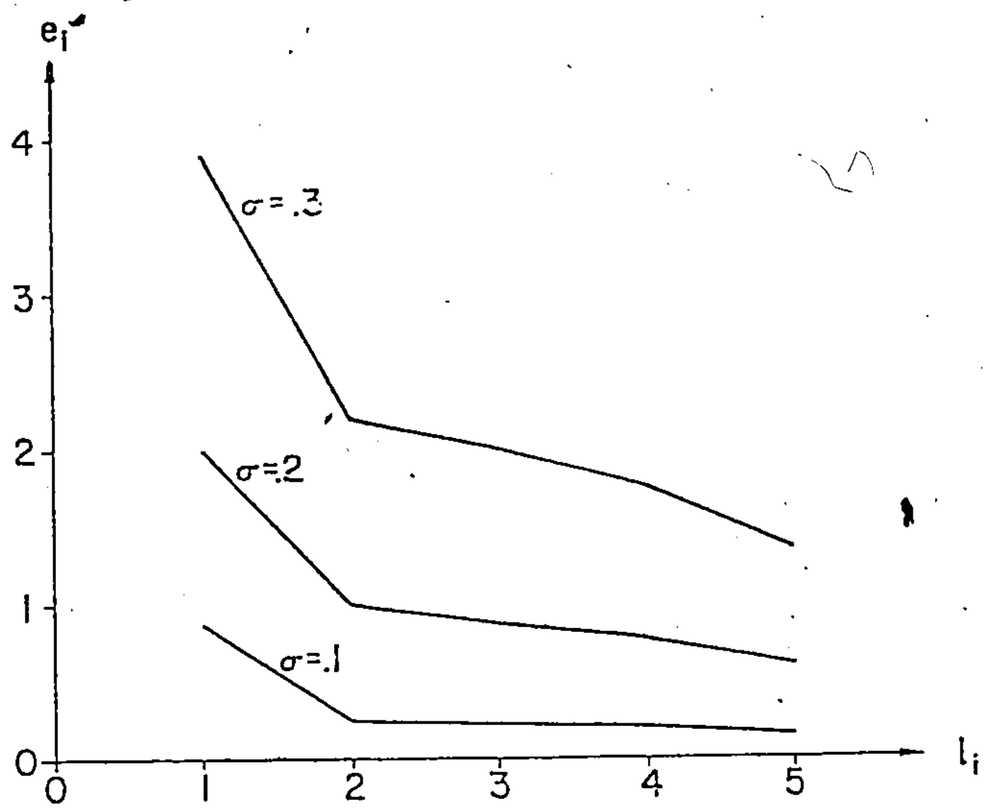

Fig. 5.1a Residual error plot of the 1st output of example 1 .

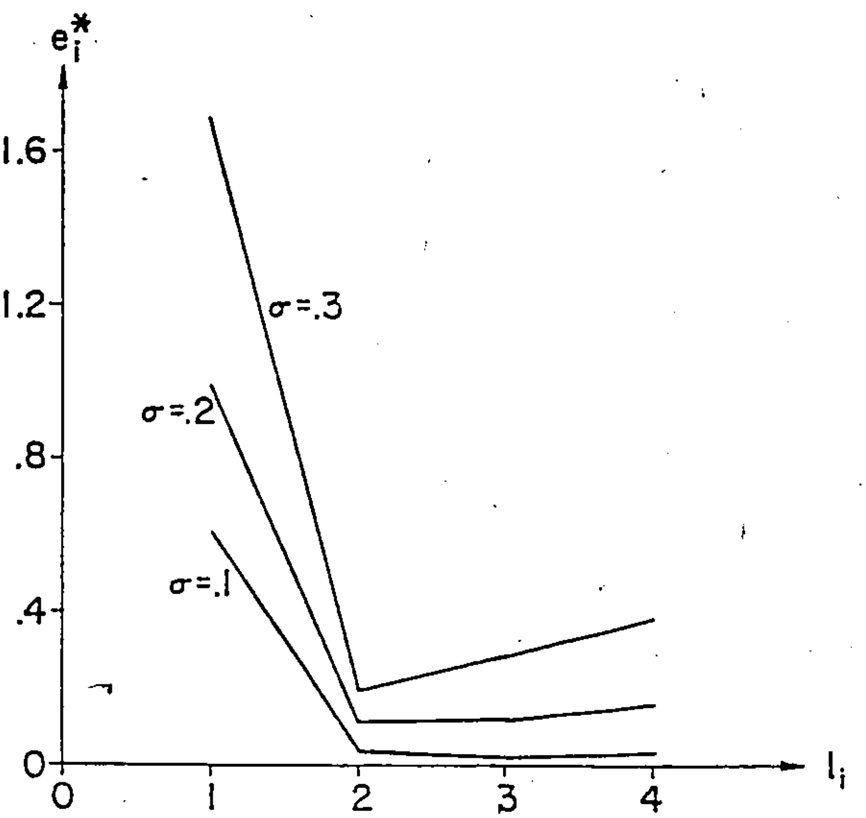

Fig. 5.1b Residual difference plot of the 1st output of example 1 .

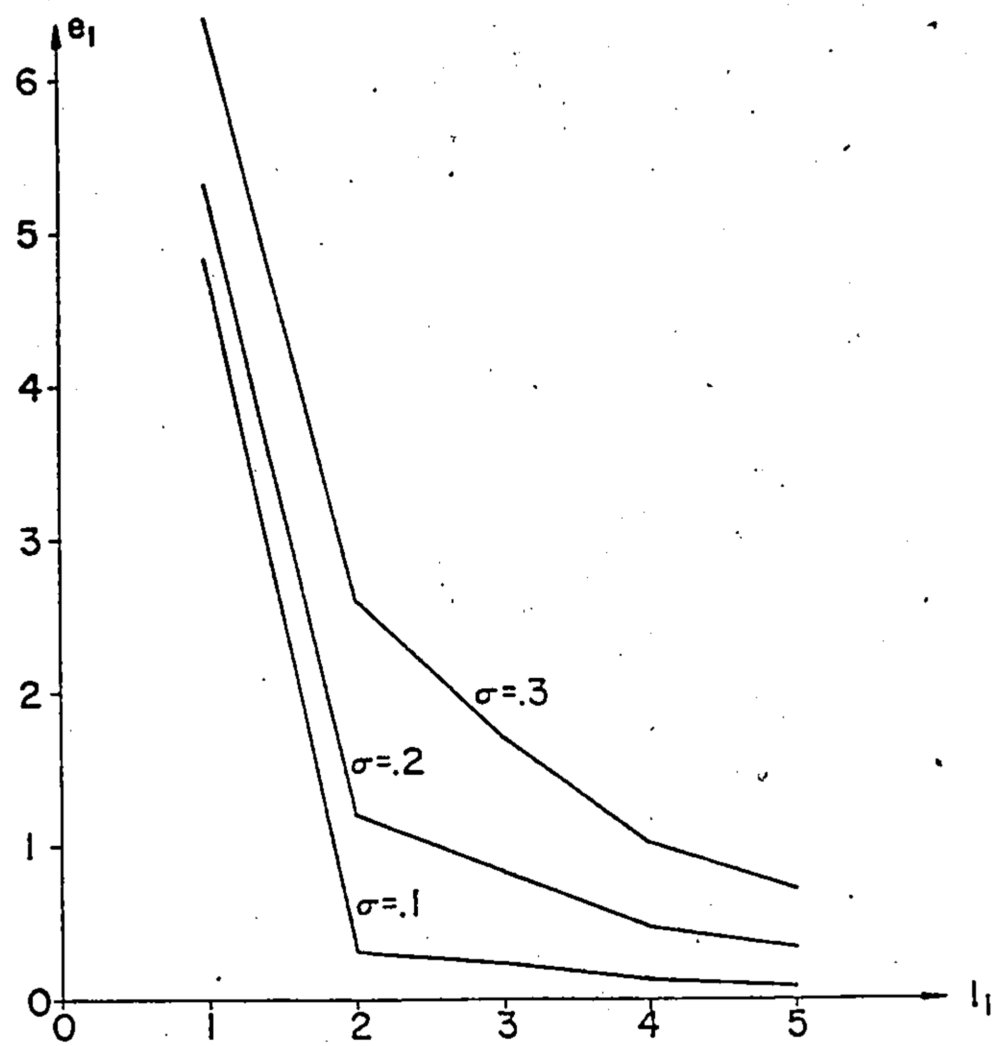

Fig. 5.2a Residual error plot of the 2nd output of example 1 .

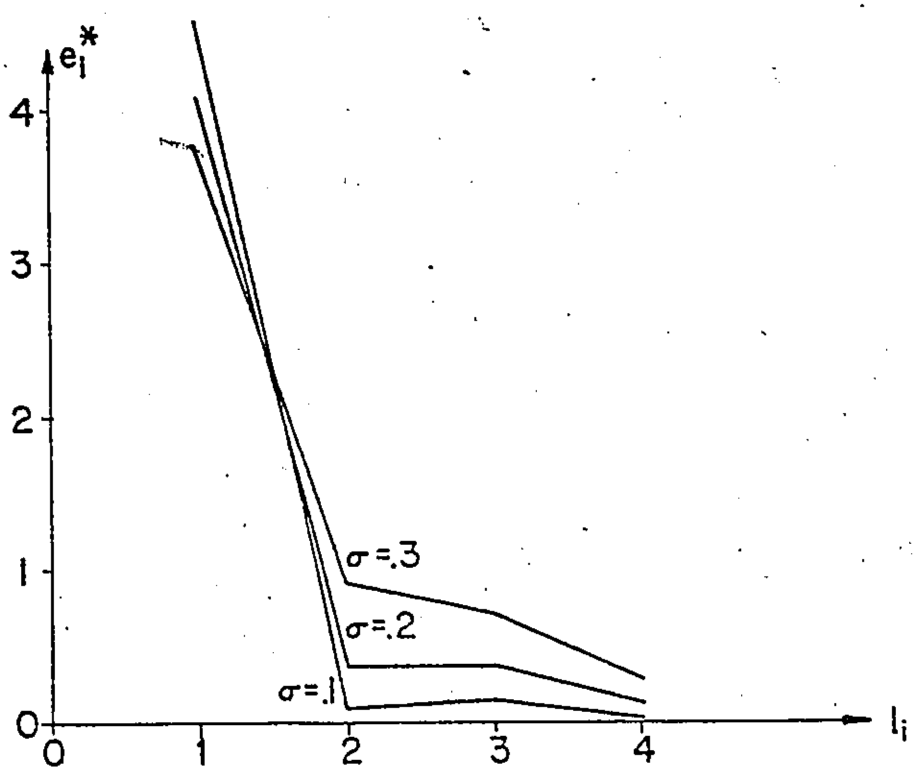

Fig. 5.2b Residual difference plot of the 2nd output of example 1 .

residual error increases. Instead of plotting the residual error e_i the residual difference e_i^* has been plotted in Figures 5.1b and 5.2b. It can be seen that these plots give a better indication of the value of the structural parameters than the plots in Figures 5.1a and 5.2a.

Example 2

The proposed algorithm was applied to determine the structural parameters of the following 5th order two-input two-output system

$$x(k+1) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0.0625 & -0.5 & 1.25 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ -0.1 & 0.65 & 1 & -0.49 & 1.4 \end{bmatrix} x(k) + \begin{bmatrix} 0.5 & -0.35 \\ 1 & 1 \\ 0 & 0 \\ -0.4 & -0.5 \\ 1 & 2 \end{bmatrix} u(k)$$

$$z(k) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} x(k) + v(k)$$

The above system was simulated as in example 1. The structural parameters of this system are

$$n_1 = 3 \text{ and } n_2 = 2$$

The proposed algorithm was applied to determine the structural parameters of the above system for the case $\sigma = 0.3$ and a sequence length $K = 30$. The residual difference e_i^* is plotted for both outputs and is shown in Figure 5.3. It is clearly determined from the figure that $n_1 = 3$ and $n_2 = 2$.

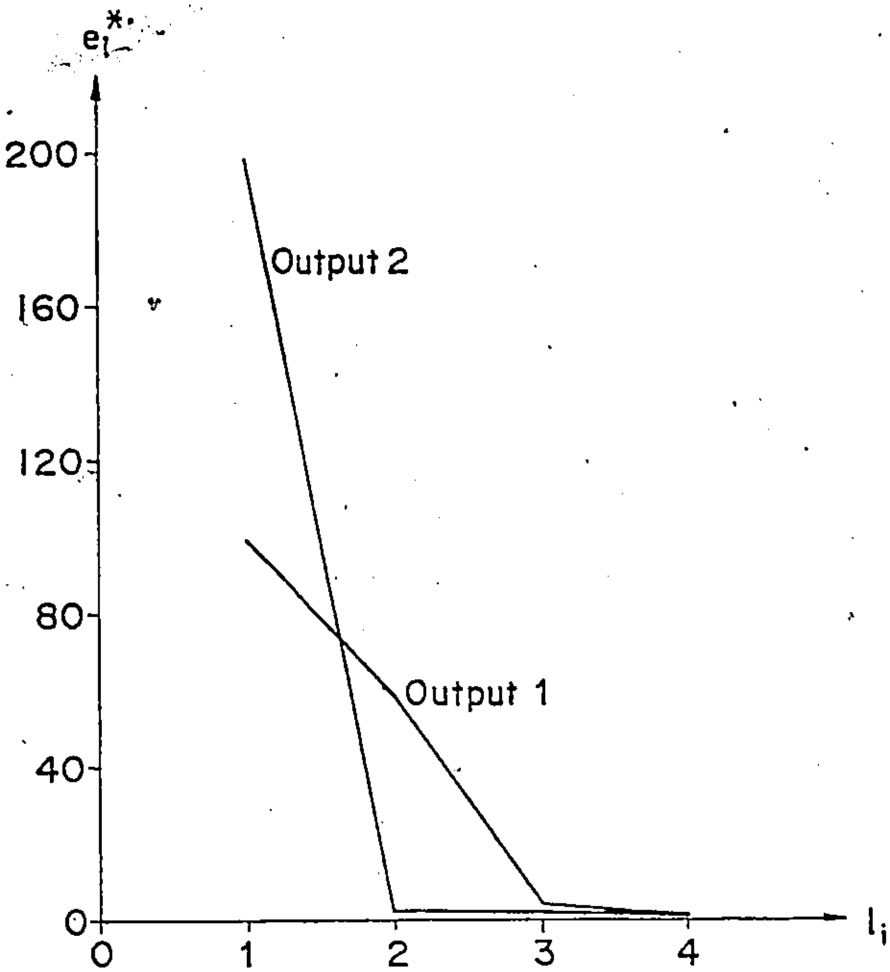

Fig. 5.3 Residual difference plot of the two outputs of example 2 .

5.4 Algorithm Combining Stochastic Approximation and Pseudo-inverse

5.4.1 Introduction

In a recent paper (Sinha and Kwong [42]) an algorithm was presented for the identification of linear multivariable discrete-time systems from noisy data. This algorithm uses the recursive generalized least-squares method and utilizes the special canonical representation of the system presented in subsection 5.3.1 and was proposed by Guidorzi [41]. In this section an algorithm is developed based on the idea of Sinha and Kwong [42]. The main difference between the proposed algorithm and the one in [42] is that a normalized stochastic approximation method (Kwatny [82]) is used for estimating the parameters of the noise model instead of the pseudo-inverse method. This idea was used by Sen and Sinha [83] for the identification of single-input single-output systems and as was mentioned in [83] the algorithm combining stochastic approximation and pseudo-inverse retains the advantages of both methods while doing away with their disadvantages.

5.4.2 Statement of the Problem

Consider a linear, time-invariant, discrete-time system. Following Guidorzi [41], it can be represented by an input-output difference equation of the type

$$P(z)\ y(k) = Q(z)\ u(k) \tag{5.39}$$

where P(z) and Q(z) are polynomial matrices in z, z is the unit advance operator, u(k) is the p-dimensional input vector and y(k) is the noise-free m-dimensional output vector. It has been shown [41] that there is a complete equivalence between the representation in equation (5.39) and the state space representation using the row companion form [81],

$$x(k+1) = A\, x(k) + B\, u(k)$$

$$y(k) = C\, x(k)$$

where the matrices A, B and C have special canonical structures (equation (5.15)).

The identification problem is to estimate the coefficients of the polynomial matrices P(z) and Q(z) from a given record of input-output observations. The state space representation (matrices A, B and C) can be obtained from these estimated parameters by the relations given by Guidorzi [41].

5.4.3 Development of the Algorithm

For noisy output measurements, equation (5.39) can be rewritten as follows

$$P(z)\, z(k) = Q(z)\, u(k) + P(z)\, v(k) \tag{5.40}$$

where z(k) is the noisy output vector and v(k) is the output noise vector.

Assuming that the system (5.40) is completely controllable and observable it can be decomposed into m observable subsystems. Each of these subsystems corresponds to one row of the matrices P(z) and Q(z). The jth row (subsystem) of equation (5.40) can be written as

$$\sum_{i=1}^{m} p_{ji}(z)\, z_i(k) = \sum_{i=1}^{p} q_{ji}(z)\, u_i(k) + \sum_{i=1}^{m} p_{ji}(z)\, v_i(k) \tag{5.41}$$

where $p_{ji}(z)$ and $q_{ji}(z)$ are polynomials in z and elements in the jth row and ith column of P and Q, respectively (equation (5.17)).

Equation (5.41) may be written more explicitly in the following form

$$z_j(k+n_j) = \sum_{i=1}^{m} \sum_{\ell=1}^{n_{ji}} a_{ji}(\ell)\, z_i(k+\ell-1) + $$

$$+ \sum_{i=1}^{p} \sum_{\ell=1}^{n_j} d_{(n_1+\ldots n_{j-1}+\ell),i}\, u_i(k+\ell-1)+e_j(k+n_j) \tag{5.42}$$

where

$$e_j(k+n_j) = v_j(k+n_j) - \sum_{i=1}^{m} \sum_{\ell=1}^{n_{ji}} a_{ji}(\ell)\, v_i(k+\ell-1),$$

n_j and n_{ji} are the structural parameters defined in [41], $a_{ji}(\ell)$ and $d_{i,j}$ are the coefficients of the polynomials $p_{ji}(z)$ and $q_{ij}(z)$ (equation (5.17)).

Now, our purpose is to estimate the parameters $a_{ji}(\ell)$ and $d_{i,j}$ of each subsystem (row), independently, through subsequent observations of the variables u(k) and z(k).

The jth subsystem (row), equation (5.42), can be rewritten as

$$z_j(k) = a_j^T(k)\, \phi_j + e_j(k) \tag{5.43}$$

where,

$$a_j(k) = [z_1(k)\ z_1(k+1)\ \ldots\ z_1(k+n_{j1}-1)\ z_2(k)\ \ldots$$

$$\ldots\ z_m(k+n_{jm}-1)\ u_1(k)\ u_1(k+1)\ \ldots\ u_p(k+n_j-1)]^T$$

and

$$\phi_j = [a_{j1}(1)\ a_{j1}(2)\ \ldots\ a_{j1}(n_{j1}-1)\ \ a_{j2}(1)\ \ldots$$

$$\ldots\ a_{jm}(n_{jm})\ \ d_{(n_1+\ldots+n_{j-1}+1),1}\ \ldots\ d_{(n_1+\ldots+n_{j-1}+n_j),p}]^T$$

A recursive pseudo-inverse algorithm can be used for estimating ϕ_j as follows

$$\phi_j(k+1) = \phi_j(k) + \frac{P_j(k)\ a_j(k+1)\ [z_j(k+1) - a_j(k+1)^T\ \phi_j(k)]}{1 + a_j^T(k+1)\ P_j(k)\ a_j(k+1)}$$

$$P_j(k+1) = P_j(k) - \frac{P_j(k)\ a_j(k+1)\ (P_j(k)\ a_j(k+1))^T}{1 + a_j^T(k+1)\ P_j(k)\ a_j(k+1)} \tag{5.44}$$

where $\phi_j(k)$ is the estimate of the parameter vector ϕ_j at the kth iteration.

It is known that the least-squares method gives biased estimates if the residuals are correlated. One way of overcoming this difficulty is to introduce filters [84], such that the resulting residuals are uncorrelated. The correlated residuals may be estimated by assuming the autoregressive model

$$e_j(k) = -\sum_{i=1}^{s} f_{j,i}\ e_j(k-i) + w_j(k) \tag{5.45}$$

where

$$e_j(k) = z_j(k) - a_j^T(k)\ \phi_j(k)$$

and $w_j(k)$ is an uncorrelated zero-mean random sequence. Equation (5.45) can be rewritten as

$$e_j(k) = \psi_j^T \epsilon_j(k) + w_j(k)$$

where

$$\psi_j = [f_{j,1} \ \ f_{j,2} \ \cdots \ f_{j,s}]^T$$

$$\epsilon_j(k) = [-e_j(k-1) \ \ -e_j(k-2) \ \ldots \ -e_j(k-s)]^T$$

In [42] the estimates of ψ_j were obtained by using the recursive least-squares algorithm. It is proposed to use the normalized stochastic approximation algorithm [82] for estimating the parameter vector ψ_j. i.e.

$$\psi_j(k+1) = \psi_j(k) + \frac{\nu}{k+1} \frac{[e_j(k) - \psi_j^T(k)\ \epsilon_j(k)]}{||\epsilon_j(k)||2} \epsilon_j(k) \quad (5.46)$$

where ν is a positive gain constant and $\psi_j(k)$ is the kth estimate of ψ_j, i.e.,

$$\psi_j(k) = [f_{j,1}(k) \ f_{j,2}(k) \ \ldots \ f_{j,s}(k)]^T$$

Utilizing this estimate of $\psi_j(k)$, the input-output sequences can be filtered according to the equations

$$u_j^*(k) = u_j(k) + \sum_{i=1}^{s} f_{ji} u_j(k-i)$$

$$z_j^*(k) = z_j(k) + \sum_{i=1}^{s} f_{ji} z_j(k-i) \quad (5.47)$$

In algorithm (5.44) the filtered sequences $u_j^*(k)$ and $z_j^*(k)$ are used in place of $u_j(k)$ and $z_j(k)$ respectively which results in uncorrelated residuals. Thus, the proposed method consists of using the normalized stochastic approximation algorithm to obtain the auto-regressive noise model parameters, whereas the pseudo-inverse algorithm is used to determine the process model parameters after the input-output data are suitably corrected utilizing the noise model (5.45).

5.4.4 Results of Simulation

The proposed algorithm was applied to on-line identification of a simulated 2-input 2-output 4th order system;

$$x(k+1) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -.125 & .75 & .25 & -1 \\ 0 & 0 & 0 & 1 \\ 0.6 & -.5 & -.24 & 1 \end{bmatrix} x(k) + \begin{bmatrix} 1 & 0 \\ .55 & -.9 \\ 0 & 1 \\ -.2 & .5 \end{bmatrix} u(k)$$

$$y(k) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} x(k)$$

After some algebraic manipulations the system can be transformed to the following input-output difference equation form

$$\begin{bmatrix} z^2-.75z+.125 & z-.25 \\ .5z-.6 & z^2-z+.24 \end{bmatrix} y(k) = \begin{bmatrix} z-.2 & .1 \\ .3 & z-.5 \end{bmatrix} u(k)$$

The number of parameters to be estimated in each subsystem is 8. The outputs are contaminated with additive noise, i.e.

$$z_1(k) = y_1(k) + v_1(k)$$

$$z_2(k) = y_2(k) + v_2(k)$$

where $v_1(k)$ and $v_2(k)$ are coloured noise sequences generated as the outputs of the following equations;

$$v_1(k+1) = .7\ v_1(k) + w_1(k)$$

$$v_2(k+1) = .8\ v_2(k) + w_2(k)$$

where $w_1(k)$ and $w_2(k)$ are zero-mean white-noise sequences uncorrelated with $u_1(k)$ and $u_2(k)$ and have standard deviation varies to vary the noise level at each output.

Using 400 samples of the input-output data the parameters of the input-output form were estimated for two cases of noise levels at the outputs using the proposed algorithm. In the first case the standard deviation of the noise was taken as $\sigma = 0.2$ and in the second case as $\sigma = 0.3$. For the sake of comparison the estimate of the system parameters were also obtained using the pseudo-inverse method and the generalized pseudo-inverse algorithm [42].

A comparison of the final estimates of the parameters of the input-output representation after 400 iterations and for the two noise cases is shown in Table 5.1. Estimate of the system matrices A and B is obtained from the estimated parameters of the input-output representation (Table 5.1) for the second case of $\sigma = 0.3$ and is shown in Table 5.2. The total computation time for each algorithm after 400 iterations is shown in Table 5.3. To show the relative convergence rate of each of the three algorithms, the normalized error $||\phi_j - \phi(k)||^2 / ||\phi_j||^2$ for each of the two subsystems (outputs) and for the second case ($\sigma = 0.3$) has been plotted against the number of iterations and is shown in Figures 5.4 and 5.5.

True Value	Ord. pseudo-inverse Case (1)	Case (2)	G.L.S. Case (1)	Case (2)	The proposed alg. Case (1)	Case (2)
0.75	1.078	1.114	0.726	0.700	0.766	0.732
-0.125	-0.302	-0.333	-0.109	-0.094	-0.131	-0.115
-1.00	-0.965	-0.928	-0.970	1.021	-1.004	-1.002
0.25	0.528	0.500	0.218	0.131	0.271	0.235
1.00	1.039	1.044	0.996	0.975	1.022	1.062
-0.20	-0.515	-0.572	-0.179	-0.155	-0.233	-0.181
0.00	0.029	0.045	0.001	0.009	0.024	0.049
0.10	0.070	0.008	0.059	-0.064	0.105	0.093
-0.50	-0.401	-0.388	-0.392	-0.175	-0.427	-0.429
0.60	0.554	0.542	0.539	0.403	0.560	0.548
1.00	0.990	1.005	1.022	1.020	1.009	1.044
-0.24	-0.127	-0.135	-0.228	-0.173	-0.197	-0.233
0.00	0.002	0.046	0.001	0.009	-0.002	0.002
0.30	0.201	0.210	0.198	-0.018	0.297	0.293
1.00	0.984	0.944	0.996	0.985	0.985	0.949
-0.50	-0.508	-0.525	-0.516	-0.498	-0.507	-0.535

Table 5.1 Comparison of the estimate of the paramters of the input-ouput model after 400 iterations

Algorithm	Matrix A				Matrix B	
Ord. pseudo-inverse	0	1	0	0	1.044	0.045
	-0.333	1.114	0.500	-0.928	0.548	-0.818
	0	0	0	1	0.046	0.944
	0.542	-0.388	-0.135	1.005	-0.149	0.406
G.L.S.	0	1	0	0	0.975	0.009
	-0.094	0.700	0.131	-1.021	0.518	-1.063
	0	0	0	1	0.009	0.985
	0.403	-0.175	-0.173	1.020	-0.179	0.505
The proposed algorithm	0	1	0	0	1.062	0.049
	-0.115	0.732	0.235	-1.002	0.594	-0.822
	0	0	0	1	0.002	0.949
	0.548	-0.429	-0.233	1.044	-0.161	0.435

Table 5.2 Comparison of the estimates of system matrices for case 2 ($\sigma = 0.3$)

Table 5.3 Comparison of the total computation time for 400 iterations

Algorithm	Ord. pseudo-inverse	G.L.S.	The proposed alg.
Computation time (sec.)	3.488	4.360	3.810

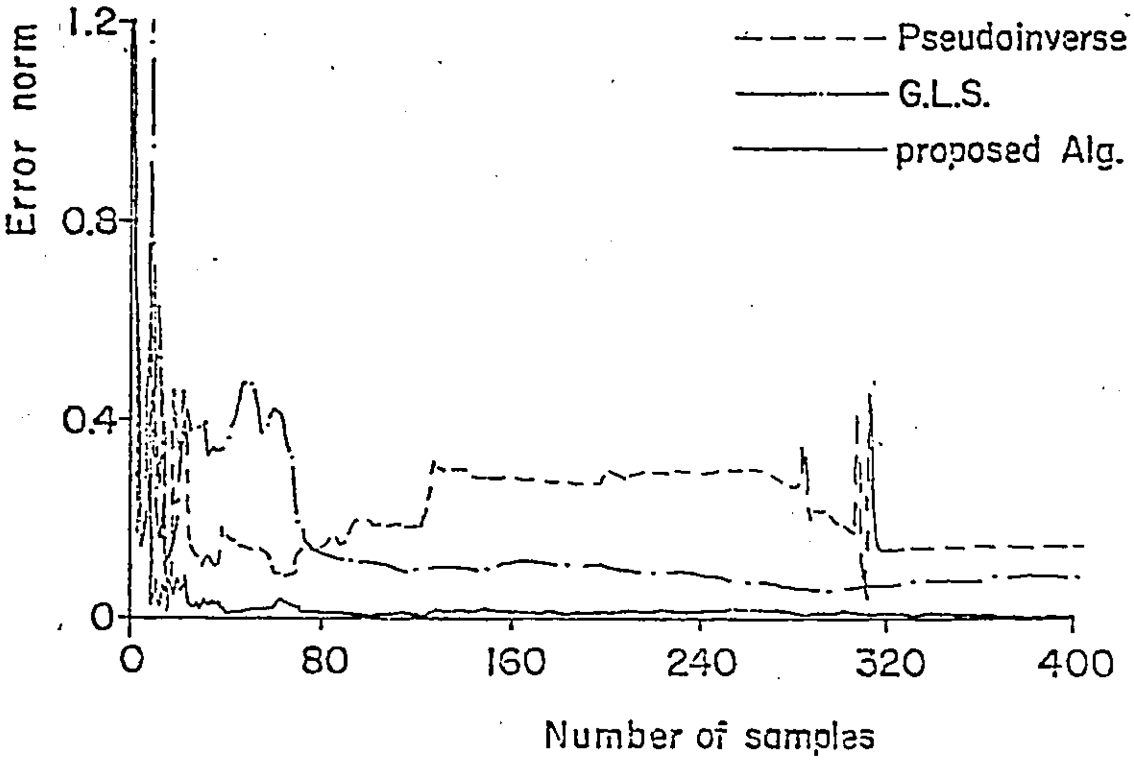

Fig. 5.4 Rate of convergence for subsystem 1 for case 2 (σ = 0.3)

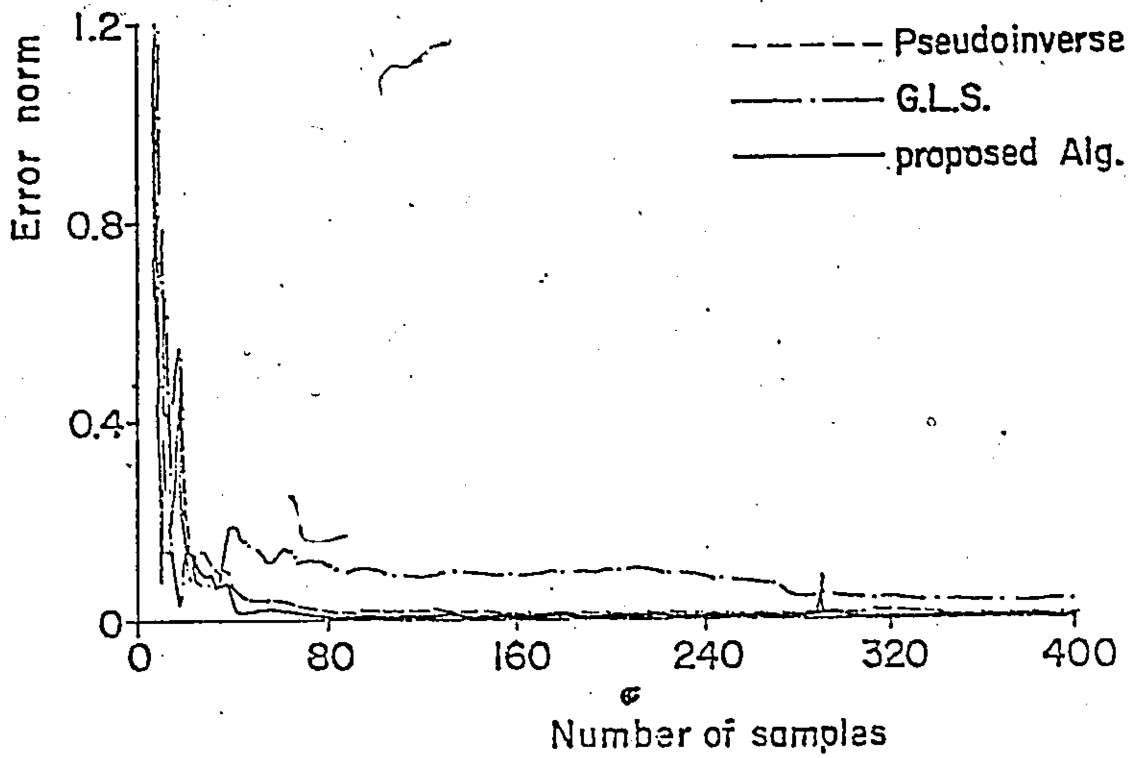

Fig. 5.5 Rate of convergence for subsystem 2 for case 2 ($\sigma = 0.3$)

As we can see from Table 5.1 the proposed algorithm gives better estimates, of the parameters of the input-output representation, than the ordinary pseudo-inverse and the generalized least-squares algorithms especially when the noise level at the outputs increases. In Table 5.2 the parameters of the A matrix are the same as those in Table 5.1 but the parameters of the B matrix have been obtained from those of Table 5.1 by applying a transformation formula given by Guidorzi [41]. From the comparison of the total computation time given in Table 5.3 we can see that the proposed algorithm takes more computation time than ordinary pseudo-inverse algorithm but less time than the generalized pseudo-inverse algorithm. This is because it uses the stochastic approximation method instead of the least-squares method used by the generalized pseudo-inverse algorithm. Figures 5.4 and 5.5 show that the proposed algorithm has a better convergence rate than the other two algorithms.

CHAPTER 6

COMBINED PARAMETER AND STATE ESTIMATION

6.1 Introduction

The problem of combined parameter and state estimation was originally posed as a nonlinear estimation problem by augmenting the state vector with the parameter vector and the extended Kalman filter was used [85] and [86]. This led to problems of divergence and excessive computation for multivariable systems. To avoid these difficulties two approaches have been proposed recently for suboptimal estimation ([65], [70] and [87]). The first approach ([65] and [70]) estimates the system parameters from the input-output data; these estimates are then used for state estimation through a Kalman filter. A particular canonical form has been used in [70], but it has been shown recently [88] that this form cannot be obtained in the general case. The second approach [87] estimates the parameters and the states of the system in two stages in a bootstrap manner. In the first stage the states are estimated with assumed nominal values of the parameters. In the second stage the parameters of the system are estimated from a pseudo-parameter measurement equation which contains the recent estimates of the states from stage one in addition to the input-output data. These two stages are coupled in a bootstrap manner.

In this chapter an on-line two-stage bootstrap algorithm will be developed for combined parameter and state estimation of linear discrete-time multivariable systems. This algorithm utilizes the idea of the second approach [87] with the following improvements

(a) A canonical form of the state equations is used which allows the parameters of the pseudo-parameter measurement equation to be related directly to the parameters of the canonical state space model.

(b) The use of this canonical form simplifies the parameter estimation problem by decomposing the system into m subsystems (where m is the number of outputs) so that the parameters of each subsystem can be estimated independently.

(c) It is shown that using the pseudo-parameter measurement equation for estimating the parameters causes the residual errors to be uncorrelated with the forcing function. Hence, unbiased estimates of the system parameters can be obtained using ordinary least-squares [89] without requiring any knowledge of the noise characteristics.

The proposed algorithm starts by transforming the state equations to the row-companion form [81] and [90]. Then assuming an initial estimate of the states, a recursive least-squares algorithm is used for estimating the parameters of the pseudo-parameter measurement equation. From these estimates, the state equations are obtained directly and these are then utilized for estimating the states of the system by a stochastic approximation algorithm. This procedure is continued in a bootstrap manner.

6.2 Formulation of the Problem

Consider the following multivariable discrete-time system

$$x^*(k+1) = A^* x^*(k) + B^* u(k)$$

$$z(k) = C x^*(k) + v(k) \qquad (6.1)$$

where $x^*(k) \in R^n$, $u(k) \in R^p$ and $y(k) \in R^m$ are the state, input and measured output vectors respectively. Also, $v(k) \in R^m$ is the noise at the output which is a zero-mean white noise sequence.

Assuming that the system is completely observable, it can be transformed to the row-companion form through the transformation $x = Sx^*$ (section 5.3.1, Chapter 5) to obtain

$$x(k+1) = A x(k) + B u(k)$$

$$z(k) = C x(k) + v(k) \qquad (6.2)$$

where $A = S A^* S^{-1}$ and $B = S B^*$. The matrix A can be written as a block matrix $\{A_{ij}\}$, $i,j = 1, 2, \ldots, m$ where

$$A_{ii} = \begin{bmatrix} 0 & & & \\ \vdots & & I_{n_i-1} & \\ a_{ii}(1) & a_{ii}(2) & \ldots & a_{ii}(n_i) \end{bmatrix}$$

$$\underset{i \neq j}{A_{ij}} = \begin{bmatrix} 0 & \ldots & 0 \\ \vdots & & \vdots \\ a_{ij}(1) & \ldots & a_{ij}(n_{ij}) \end{bmatrix} \qquad (6.3)$$

The matrix C consists of unit row vectors only, and can be written as

$$C = \begin{bmatrix} e^{1} \\ e^{n_1+1} \\ \vdots \\ e^{n_1+n_2+\ldots+n_{m-1}+1} \end{bmatrix} \tag{6.4}$$

where e^i is the ith unit row vector of dimension n and the integers n_i's are called the observability subindices of the system.

Our problem is to obtain consistent estimates of the parameters as well as the states of the system (6.2) from the given input-output data without knowing the noise characteristics. The structural parameters n_i's are assumed to be known.

6.3 Parameter Estimation

Define the following matrices

$$B = \begin{bmatrix} b^{1} \\ b^{2} \\ \vdots \\ b^{n} \end{bmatrix} \tag{6.5}$$

and

$$\bar{A} = \begin{bmatrix} a^{1} \\ a^{2} \\ \vdots \\ a^{m} \end{bmatrix} \tag{6.6}$$

where

$$a^j = [a_{j1}(1) \ldots a_{j1}(n_1) \; a_{j2}(1) \ldots a_{jn_{jm}}(n_{jm})] \tag{6.7}$$

for $j = 1, 2, \ldots, m$. The system parameters to be estimated are the parameters of the matrices $\bar{A}$ and B.

Because of the canonical structure of the matrices A and C, the following equations can be obtained for the jth subsystem.

$$\begin{aligned}
x_{n_1+\ldots+n_{j-1}+1}(k+1) &= x_{n_1+\ldots+n_{j-1}+2}(k) + b^{n_1+\ldots+n_{j-1}+1}\,u(k) \\
x_{n_1+\ldots+n_{j-1}+2}(k+1) &= x_{n_1+\ldots+n_{j-1}+3}(k) + b^{n_1+\ldots+n_{j-1}+2}\,u(k) \\
&\vdots \\
x_{n_1+\ldots+n_{j-1}+\ell}(k+1) &= x_{n_1+\ldots+n_{j-1}+\ell+1}(k) + b^{n_1+\ldots+n_{j-1}+\ell}\,u(k) \\
&\vdots \\
x_{n_1+\ldots+n_j}(k+1) &= \sum_{i=1}^{m}\sum_{\ell=1}^{n_{ji}} a_{ji}(\ell)\, x_{n_1+\ldots+n_{i-1}+\ell}(k) + b^{n_1+\ldots+n_j}\,u(k)
\end{aligned} \tag{6.8}$$

and

$$z_j(k) = x_{n_1+\ldots+n_{j-1}}(k) + v_j(k) \tag{6.9}$$

Hence substituting for $x_{n_1+\ldots+n_{j-1}+1}(k)$ in equation (6.9) from equations (6.8) we get the following expression which characterizes the jth output (subsystem) of the system.

$$z_j(k+n_j) = Z_j^T(k+n_j-1)\,\theta_j + v_j(k+n_j) \tag{6.10}$$

where

$$Z_j(k+n_j-1) = [x^T(k) \quad u^T(k) \quad u^T(k+1) \ \dots \ u^T(k+n_j-1)]^T \tag{6.11}$$

and

$$\theta_j = [a^j \quad b^{n_1+\dots+n_j} \ \dots \ b^{n_1+\dots+n_{j-1}+1}]^T \tag{6.12}$$

where n_{ji} is assumed to be equal to n_i.

From the above equations we can see that the parameter identification problem can be decoupled where the parameter vector characterizing each subsystem can be estimated independently of other subsystems. The parameter vector θ_j of equation (6.10) can be estimated recursively by the following least-squares algorithm

$$\hat{\theta}_j(k+1) = \hat{\theta}_j(k) + \frac{P_j(k)\ Z_j(k+n_j-1)\ [z_j(k+n_j) - Z_j^T(k+n_j-1)\ \hat{\theta}_j(k)]}{1 + Z_j^T(k+n_j-1)\ P_j(k)\ Z_j(k+n_j-1)}$$

$$P_j(k+1) = P_j(k) - \frac{P_j(k)\ Z_j(k+n_j-1)\ [P_j(k)\ Z_j(k+n_j-1)]^T}{1 + Z_j^T(k+n_j-1)\ P_j(k)\ Z_j(k+n_j-1)} \tag{6.13}$$

where $\hat{\theta}_j(k)$ is defined as the estimate of θ_j at the kth iteration.

6.3.1 Proof of the Unbiasedness of the Parameter Estimates

Equation (6.10) can be put into the form of equation (2.18) (Chapter 2) where $U_j(k)$ and $e_j(k)$ are defined as follows

$$U_j(k) \triangleq Z_j(k) = [x^T(k-n_j+1) \quad u^T(k-n_j+1) \quad u^T(k-n_j+2) \ \dots \ u^T(k)]^T \tag{6.14}$$

and

$$e_j(k) \triangleq v_j(k) \tag{6.15}$$

From equations (6.14) and (6.15) we can see that the vector sequence $U_j(k)$ is uncorrelated with the residual error $e_j(k)$ hence condition (a) (subsection 2.4, Chapter 2) is satisfied. Moreover, if the noise $v_j(k)$ has zero-mean, then condition (b) (subsection 2.4) is also satisfied and hence the parameter estimates $\hat{\theta}_j$, $j = 1, 2, \ldots, m$ obtained by equation (6.13) are unbiased.

6.4 State Estimation

From equation (6.11) we can see that the vector Z_j includes the states of the system which are not known. But if an estimate of the states is known then we can estimate the parameter vectors using the algorithm of equation (6.13). Hence an estimate of system matrices $\hat{A}(k)$ and $\hat{B}(k)$ at the kth iteration is obtained directly from the estimated system parameters $\hat{\theta}_j$, $j = 1, 2, \ldots, m$ through relations (6.5)-(6.7) and (6.12).

After having an estimate of system matrices at the kth iteration, they can be used for estimating the state vectors of the system using the following stochastic approximation algorithm

$$\hat{x}(k+1|k) = \hat{A}(k)\ \hat{x}(k|k) + \hat{B}(k)\ u(k)$$

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + \gamma(k)\ C^T[z(k+1) - C\ \hat{x}(k+1|k)] \qquad (6.16)$$

where $\hat{x}(k|k)$ is the estimate of the state vector $x(k)$ at the kth iteration and $\gamma(k)$ is a scalar sequence satisfying Dvoretzky's conditions [91].

After this state estimate is obtained, the expression for $x(k)$ in equation (6.11) may be replaced by $\hat{x}(k|k)$ to obtain a new estimate of the parameter vectors θ_j, j = 1, 2, ..., m. The procedure is then repeated between the two stages in a bootstrap manner.

6.5 Results of Simulation

The proposed algorithm was applied to the following simulated 3rd-order two-output one-input system

$$x(k+1) = \begin{bmatrix} 0.00 & 1.0 & 0.0 \\ 0.10 & 0.3 & 0.1 \\ 0.95 & 0.1 & 0.7 \end{bmatrix} x(k) + \begin{bmatrix} 0.12 \\ 0.36 \\ 0.20 \end{bmatrix} u(k)$$

$$y(k) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} x(k) + v(k)$$

The input to the system u(k) was taken as a zero-mean white noise sequence with unit variance. The noise sequences $v_1(k)$ and $v_2(k)$ were taken as uncorrelated zero-mean white noise sequences with standard deviations of 0.1 and 0.3 respectively.

The proposed combined parameter identification and state estimation algorithm was applied to the above example with zero initial states and assuming zero initial values for the parameters. The final parameter estimates after 1000 iterations are given in Table 6.1. Also, the rate of convergence of the parameter estimates is shown in Fig. 6.1 which is a plot of the normalized squared error $||\theta_j-\hat{\theta}_j(k)||^2/||\theta_j||^2$ against the number of samples k. The convergence of the state estimation is shown in Fig. 6.2, where the squared error of the state

Parameter	$a_{11}(1)$	$a_{11}(2)$	$a_{12}(1)$	$a_{21}(1)$	$a_{21}(2)$	$a_{22}(1)$	b(1)	b(2)	b(3)
True value	0.10	0.30	0.10	0.95	0.10	0.70	0.12	0.36	0.20
Estimated value	0.10	0.29	0.10	0.89	0.12	0.71	0.12	0.36	0.20

Table 6.1: Final estimates of system parameters after 1000 iterations

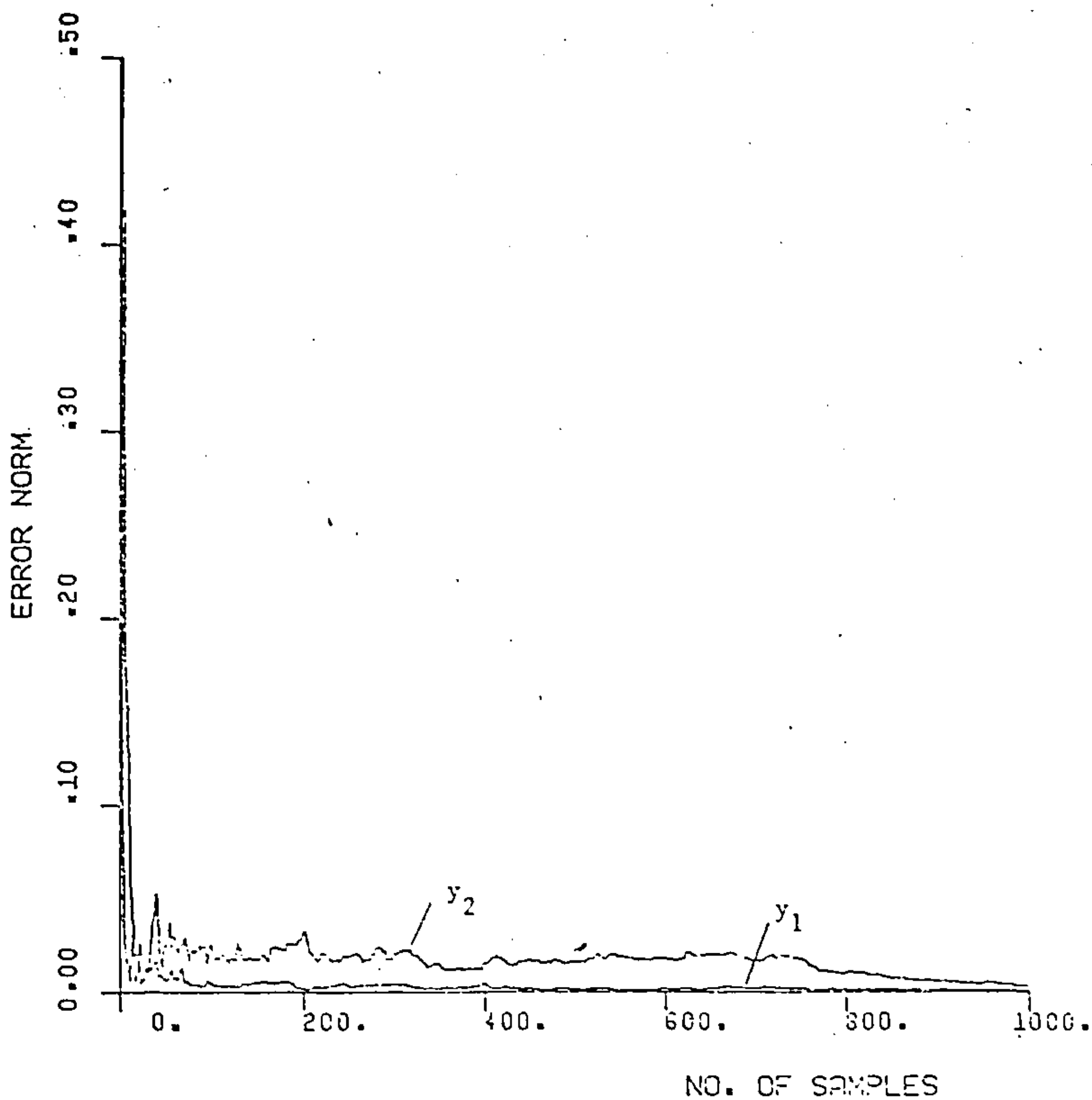

Fig. 6.1 Rate of convergence of the parameter estimates for 1000 iterations .

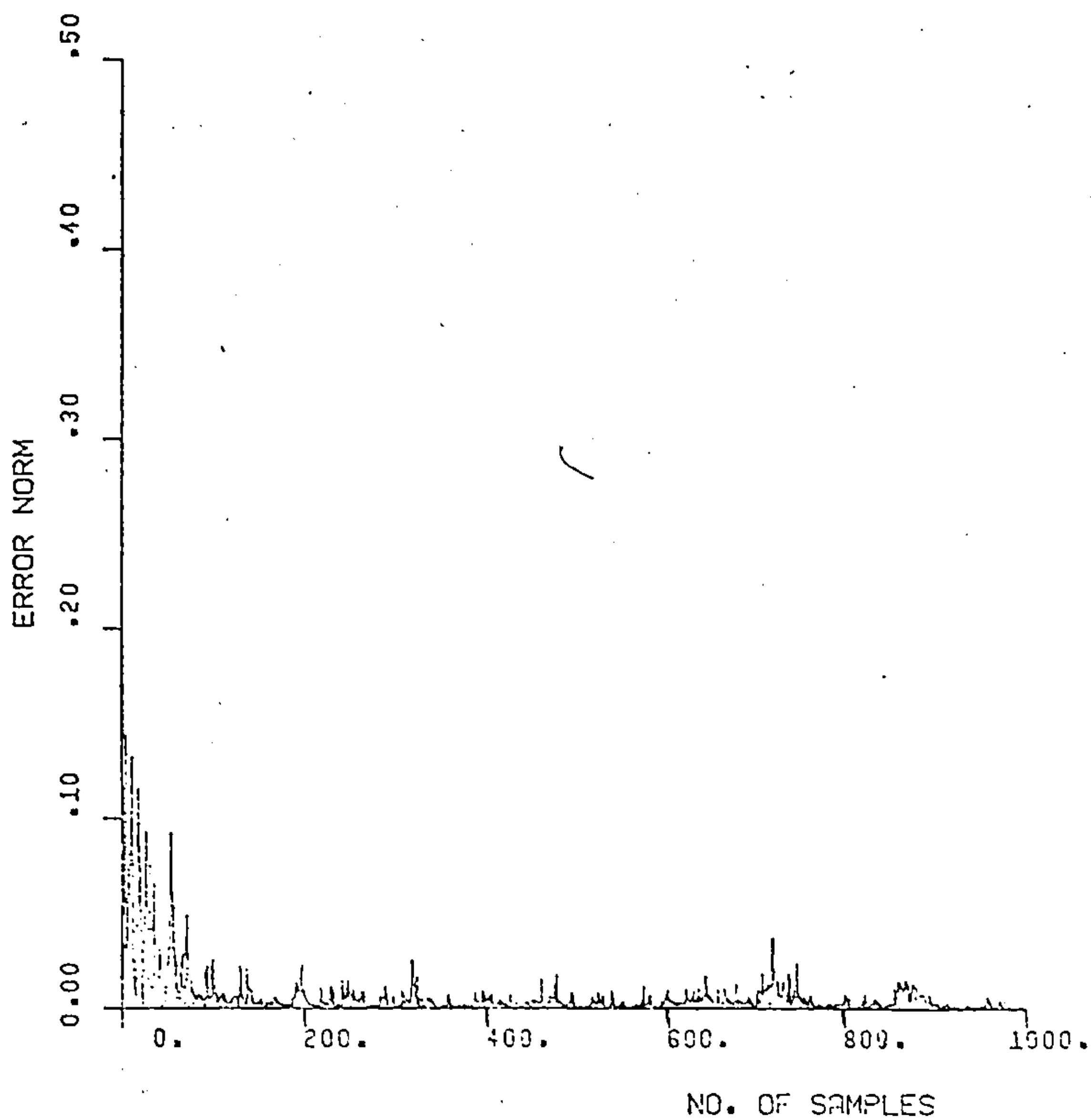

Fig. 6.2 Error norm of the state estimates for 1000 iterations .

estimates $||x(k)-\hat{x}(k)||^2$ is plotted against k.

As we can see from Table 6.1, good estimates of the parameters of the system $a_{ij}(\ell)$'s and b_{ω}^{i}'s have been obtained after 1000 iterations from the noisy measurements. Figure 6.1 shows that the rate of convergence of the parameter estimates is good and the parameter estimates converge after a small number of iterations. Also Figure 6.2 shows that the error between the estimated states and their true values is small and the state estimator converges.

CHAPTER 7

CONCLUSIONS

The major effort in this thesis has been directed towards the problem of system identification in the presence of additive noise. Attention has been directed towards the identification of multivariable discrete-time linear models from input-output data which is contaminated with noise and without the knowledge of the noise characteristics. It has been shown that based on a given set of input-output data different models can be identified from this data to represent the system. In general, there are four types of models which have been used in the area of identification of linear multivariable systems. These models are: transfer-function matrix, impulse response, input-output difference equation and state space. These four models are equivalent and transformations between them are possible. The problem of system identification, normally, starts by determining the structural parameters characterizing each model from a record of the input-output data, otherwise they have to be assumed from practical considerations. After determining the structural parameters the system parameters can be estimated by a suitable algorithm. It has been shown in Chapter 2 that the type of model used for a multivariable system has some effects on the properties of the corresponding algorithm for identification. Since the number of parameters to be estimated depends on the choice of the model, the complexity of the algorithm is determined by this choice. Another important effect is on the bias of the estimated system parameters as bias is often introduced into parameter estimation using least-squares method by choosing a model form which causes the forcing function to include observations correlated with the error in the observed output. In general, it is recommended to identify a given system in that model which is required by the application for which it will be used and also to avoid any transformation of the estimated model parameters to obtain another representation. In this thesis different algorithms have been developed to identify a given system directly in the four system models discussed before.

It has been shown that the identification of a multivariable system in the impulse response representation, out of the other three models, gives unbiased estimates of the system parameters when the ordinary least-squares method is used. Two different algorithms have been proposed in Chapter 3 to estimate the Markov parameters of the system from noisy measurements. The first algorithm uses a normalized stochastic approximation method which is computationally simple but a bias is introduced into the parameter estimates since a truncation of the impulse response sequence is used. The second algorithm uses correlation techniques which avoid the problem of truncation but on the other hand a special type of inputs is required for identification. In general, the identification of the system in the impulse response representation has a certain disadvantage because most of the control theory uses the state space representation of the system and hence the identified impulse response representation has to be transformed into the state space representation. Often if the Markov parameters are first estimated even with a very small bias and then used for minimal realization, we may obtain a wrong estimate of the system model which may even be unstable for a stable system. This problem can be thought of as a partial realization of the system, having an infinite number of Markov parameters, which retains some of its Markov parameters (Hickin and Sinha [92]).

The identification of multivariable systems in the transfer-function matrix form from the input-output data has not received much attention as this model is overparameterized. In Chapter 4 two algorithms have been proposed for identifying two different transfer-function matrix representations. The first algorithm estimates the parameters of a matrix representation where the order of each of its rows is assumed to be equal to the order of the system. On the other hand the second algorithm utilizes a general matrix representation where the order of each of its rows is estimated in advance by the residual error technique from noisy measurements. In both algorithms the parameters of each row of the transfer-function matrix have been estimated separately from the other rows. The second algorithm has been modified to estimate the parameters of the system when the outputs are contaminated with additive noise using a multivariable version of the adaptive least-squares method proposed by Panuska [55].

Due to the practical importance of the state space representation, especially in control theory, many algorithms have been proposed for the identification of multivariable systems in state space form from input-output data. Most of these algorithms start by identifying the system in a nonparametric model representation and then the state space representation can be obtained by different transformations. For example a state space representation has been obtained from the impulse response representation in Chapter 3 and from the input-output difference equation in Chapter 5. In general, the identification of systems in the state space representation requires first the determination of the structural indices of the system (which implies knowing the system order) and then the estimation of the parameters of the state space matrices A, B and C. Because of the non-uniqueness of these matrices some canonical forms have been suggested for the identification problem which aim to simplify the transformation of the state space model into a nonparametric model suitable for identification and also to reduce the number of parameters to be estimated. In Chapter 5 an algorithm has been proposed to estimate the structural parameters of a row companion state space canonical form (the observability indices). The main advantage of using this canonical form is that a unique relation between this state space representation and a canonical input-output difference equation has been obtained by Guidorzi [41] where he showed that the two models have the same structural parameters. The parameters of the equivalent input-output difference equation representation have been estimated from noisy data by a proposed recursive algorithm which combines stochastic approximation and pseudo-inverse. In Chapter 6 a way has been derived for representing the row-companion canonical state space model (used in Chapter 5) in a nonparametric model representation which causes the residual error to be uncorrelated with the forcing function. Hence, according to section 2.4, unbiased estimates of system parameters can be obtained using ordinary least-squares.

7.1 Suggestions for Further Research

1. Throughout this thesis it was assumed that the input sequence to the system u(k) was functionally independent of the system output z(k); i.e. the system to be identified had an open loop structure. However, despite the fundamental importance of closed loop identification, most parameter estimation algorithms suffer severe difficulties in the presence of feedback. These difficulties arise from the fact that the presence of the feedback control signal produces additional correlations in existing identification algorithms that were intended for open loop systems only. There appears to be little work done in the area of identification of closed loop systems, especially for the multivariable case. The single-input single-output problem has been considered by several authors (e.g. Graupe [93], Saridis and Lobbia [94] and Box and MacGregor [95]). The multivariable problem has been considered by Ljung et al. [96] and Lobbia and Saridis [45]. The problem of identification of closed loop systems is an area where there is much scope for further work. One way of solving this problem is to add a small independent noise sequence or dither signal to the input during collection of the closed loop data.

2. The convergence of the algorithms of the combined stochastic approximation and pseudo-inverse and of the combined state and parameter estimation has not been theoretically justified. However, when each stage of these two stage algorithms is treated separately the convergence of each stage can be argued if the other one satisfies certain properties. Generally, the convergence of the overall algorithm is not obvious but it may be possible to study its convergence in the same way as Ljung [97] and [98] and Soderstrom et al. [99]. This aspect has to be further studied.

3. The algorithm for combined parameter and state estimation proposed in Chapter 6 can be modified and applied to power system analysis. For example it can be used in the area of modeling and short term prediction of the active load demand of inter-connected power systems. Also it can be used in the area of dynamic state estimation of power systems.

4. The major appeal of the proposed on-line identification algorithms is their simplicity. Since the storage requirements for data are small and the identification algorithms require few arithmetic operations, they may be implemented in real time on a minicomputer, in a manner similar to that in a recent paper [100], or on a microcomputer.

5. The problem of structure estimation for the state space model from noisy data is of great importance in system identification and more work should be done about it. One way of solving this problem is the generalization of the algebraic approach of Tse and Weinert [47] (section 5.2) to handle the case of nonzero control. Consider the

following system equations instead of equation (5.1), Chapter 5

$$x(k+1) = Ax(k) + Bu(k)$$

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

where u(k) is the input vector sequence of zero-mean and covariance U and v(k) is the output noise vector sequence of zero-mean and covariance V. Then equations (5.4) and (5.5) can be modified to

$$P = A\ P\ A^T + B\ R\ B^T$$

$$R(0) = C\ P\ C^T + V$$

$$R(\sigma) = C\ A^{\sigma-1}\ S$$

where

$$S = A\ P\ C^T$$

Proceeding in the same way as Tse and Weinert (section 5.2), the set $\{p_i,\ \beta_{ijk}\}$ can be estimated, without any modification of the algorithm, hence the system matrices A and C in their canonical form can be obtained. The identification of the input matrix B in its canonical form can be obtained as follows. Let B be written as

$$B^T = [b_{11} \cdots\ b_{1p_1}\ \ b_{21} \cdots b_{s1} \cdots b_{sp_s}]$$

where

$$b_{ij}^T = [b_{ij1}\ \ b_{ij2} \cdots b_{ijp}]$$

If the Markov parameters of the system are given by

$$J_\ell = \begin{bmatrix} J^T_{\ell,1} \\ J^T_{\ell,2} \\ \vdots \\ J^T_{\ell,m} \end{bmatrix} \qquad \ell = 1, 2, \ldots$$

then due to the special canonical form of the matrices A and C, (Tse and Weinert [47]), the input matrix B can be written as

$$B = \begin{bmatrix} J^T_{1,1} \\ J^T_{2,1} \\ \vdots \\ J^T_{p_1,1} \\ \vdots \\ J^T_{1,s} \\ \vdots \\ J^T_{n_s,s} \end{bmatrix}$$

The Markov parameters in the above equation can be calculated using either correlation techniques (Sinha et al. [36]) or stochastic approximation (El-Sherief and Sinha [29]).

6. The algorithm for combined parameter and state estimation proposed in Chapter 6 can be extended to solve the problem of combined identification and control of linear multivariable systems. A bootstrap identification estimation and control algorithm can be obtained by adding a small independent dither signal to the input during normal

operation. This algorithm will be more efficient than that of Lobbia and Saridis [45] as it will not need any transformation of the estimated system parameters to obtain the state space matrices and hence no matrix inversion will be needed. Also this algorithm will avoid the bias introduced into the parameter estimates [45] due to the truncation of the impulse response sequence.

APPENDIX I

THE RESIDUAL ERROR TECHNIQUE

A1.1 The Noise-free Case

Suppose the set of vectors $\{y, x_1, x_2, \ldots, x_n\}$ is given and we want to study the problem whether or not the vector y is a linear combination of the set of vectors $\{x_1, x_2, \ldots, x_n\}$. This problem is equivalent to (Suen and Liu [54])

$$y = X\,\theta \quad \text{for some vector } \theta \tag{A1.1}$$

where $X \triangleq [x_1 \quad x_2 \quad \ldots \quad x_n]$

For any vector $\hat{\theta}$, let

$$\hat{y} = X\,\hat{\theta} \tag{A1.2}$$

be estimated by $\hat{\theta}$ and let

$$e(\hat{\theta}) = ||y-\hat{y}||^2 \tag{A1.3}$$

be the residual error for $\hat{\theta}$. Then y is a linear combination of $\{x_1, x_2, \ldots, x_n\}$ if and only if $e(\hat{\theta}) = 0$ for some $\hat{\theta}$. However, such $\hat{\theta}$ may be difficult to find. Alternatively, consider

$$\theta^o = X^+\,y \tag{A1.4}$$

where X^+ is the pseudo-inverse of X. Then from (A1.1) we get

$$y^o = X\,X^+\,y \tag{A1.5}$$

and from (A1.3) we get

$$e(\theta^o) = (y - X X^+ y)^T (y - X X^+ y)$$

$$= y^T y - y^T X X^+ y - y^T (X X^+)^T y + y^T (X X^+)^T (X X^+) y$$

$$= y^T (I - X X^+) y \qquad \text{(A1.6)}$$

It can be shown Wiberg [101] that $e(\theta^o) \leq e(\theta)$ for all θ (Suen and Liu [54]).

Lemma 1

The vector y is a linear combination of $\{x_1, x_2, \ldots, x_n\}$ if and only if $e^o \triangleq e(\theta^o) = 0$.

It can be noticed that the above lemma does not require the linear independence of the set of vectors $\{x_1, x_2, \ldots, x_n\}$.

A1.2 The Noisy Case

Let a vector y be corrupted by a zero-mean noise vector v and let y^* be the noisy observation

$$y^* = y + v \qquad \text{(A1.7)}$$

From a given set of vectors $\{y^*, x_1, x_2, \ldots, x_n\}$ we want to study the problem whether or not y is a linear combination of the set of vectors $\{x_1, x_2, \ldots, x_n\}$ (Suen and Liu [54])

$$\text{i.e.} \quad y = X\,\theta \quad \text{for some vector } \theta \qquad \text{(A1.8)}$$

From equations (A1.7) and (A1.8) we get

$$y^* = X\,\theta + v \qquad \text{(A1.9)}$$

Hence substituting in equation (A1.6) for y by y^* we get

$$e^o \triangleq e(\theta^o) = y^{*T} (I - X X^+) y^* \tag{A1.10}$$
$$= y^T (I - X X^+) y + 2y^T (I - X X^+) v + v^T (I - X X^+) v$$

<u>Lemma 2</u>

Assuming v is a zero-mean vector uncorrelated with y, then if y is a linear combination of $\{x_1, x_2, \ldots, x_n\}$ we have

$$E\{e^o\} = E\{v^T (I - X X^+) v\} \tag{A1.11}$$

and otherwise

$$E\{e^o\} = E\{v^T (I - X X^+) v\} + y^T (I - X X^+) y \tag{A1.12}$$

It can be observed that the R.H.S. of equation (A1.12) is strictly greater than that of equation (A1.11).

APPENDIX II

PROOF OF CONVERGENCE OF STOCHASTIC APPROXIMATION ALGORITHM

A2.1 Introduction

In this appendix a proof [102] of the unbiasedness and convergence in the mean-square sense of the parameter estimates, obtained by the normalized stochastic approximation algorithms (section 4.2 and [29]), is derived in this appendix.

The algorithms used in section 4.2 and [29] estimate the parameters θ_i's of the following model

$$z_i(k+L) = U_{(k)}^{T(k+L-1)} \theta_i + e_i(k+L) \quad i = 1, 2, \ldots, m \qquad (A2.1)$$

where

$$U_{(k)}^{T(k+L-1)} := [u^T(k+L-1) \;\; u^T(k+L-2) \;\ldots\; u^T(k)]$$

$$e_i(k+L) = \varepsilon_i(k-1) + v_i(k+L) \qquad \text{for the algorithm of [29]}$$

$$= v_i(k+L) \qquad \text{for the algorithm of section 4.2}$$

$u(k)$ is the p-dimensional input vector, $z_i(k)$ is the ith output, $v_i(k)$ is the noise at the ith output, $\varepsilon_i(k-1)$ is a truncation term, m is the number of outputs,, L is the identification interval and θ_i is the parameter vector characterizing the ith output. $e_i(k+L)$ is a zero-mean noise sequence independent from the vector sequence $U_{(k)}^{T(k+L-1)}$ i.e.

$$E\{e_i(k+L)\} = 0, \; E\{U_{(k)}^{T(k+L-1)}\} = 0$$

and $$E\{U^{T(k+L-1)}_{(k)} e_i(k+L)\} = 0 \qquad i = 1, 2, \ldots, m \qquad (A2.2)$$

The estimation algorithm is as follows

$$\hat{\theta}_i(k+L) = \hat{\theta}_i(k-1) + \nu(k) \frac{U^{T(k+L-1)}_{(k)}}{||U^{T(k+L-1)}_{(k)}||^2} [z_i(k+L) - U^{T(k+L-1)}_{(k)} \hat{\theta}_i(k-1)]$$

$$i = 1, 2, \ldots, m; \; k = 1, L+1, 2L+3 \qquad (A2.3)$$

where $\hat{\theta}_i(k)$ is the estimate of θ_i at the kth iteration and the sequence $\nu(k)$ satisfies the following conditions

$$\lim_{k\to\infty} \nu(k) = 0, \quad \sum_{k=1}^{\infty} \nu(k) = \infty \quad \text{and} \quad \sum_{k=1}^{\infty} \nu^2(k) < \infty \qquad (A2.4)$$

In the next sections a proof of the unbiasedness and convergence in the mean-square sense of the parameter estimates will be obtained. Proof of a similar algorithms has been obtained in different ways by Albert and Gardner [103], Kwatny and Shen [104] and has been exploited by Nagumo and Noda [105] and Kwatny [82].

A2.2 Proof of Unbiasedness

The algorithm (A2.3) minimizes the following criterion

$$J_i \triangleq f(\theta_i) = \frac{1}{2} E \left\{ \frac{e_i^2(k+L)}{||U(k+L-1)||^2} \right\} \qquad i = 1,2,\ldots,m \qquad (A2.5)$$

where $e_i(k+L) = z_i(k+L) - U^T(k+L-1)\,\theta_i$

and $U(k+L-1) \triangleq U^{(k+L-1)}_{(k)}$

Actually $e_i(k)$ cannot be obtained because θ_i is not known, hence an estimate of θ_i can be used and the criterion to be minimized will be

$$\hat{J}_i = \frac{1}{2} E \left\{ \frac{[z_i(k+L)-U^T(k+L-1)\hat{\theta}_i(k-1)]^2}{||U(k+L-1)||^2} \mid \hat{\theta}_i(k-1) \right\} \quad i=1,2,\ldots,m \tag{A2.6}$$

where $\hat{\theta}_i(k)$ is the estimate of θ_i at the kth iteration.

The gradient of $\hat{J}_i$ with respect to $\hat{\theta}_i(k-1)$ is as follows

$$\frac{\partial \hat{J}_i}{\partial \hat{\theta}_i(k-1)} = - E \left\{ U(k+L-1) \frac{[z_i(k+L)-U^T(k+L-1)\hat{\theta}_i(k-1)]}{||U(k+L-1)||^2} \mid \hat{\theta}_i(k-1) \right\}$$

$$= E \left\{ U(k+L-1) \frac{[U^T(k+L-1)\theta_i - U^T(k+L-1)\hat{\theta}_i(k-1) + e_i(k+L)]}{||U(k+L-1)||^2} \mid \hat{\theta}_i(k-1) \right\} \quad i = 1,2,\ldots,m \tag{A2.7}$$

Defining $\tilde{\theta}_i(k) = \hat{\theta}_i(k) - \theta_i$ we get

$$\frac{\partial \hat{J}_i}{\partial \hat{\theta}_i(k-1)} = - E \left\{ U(k+L-1) \frac{[e_i(k+L)-U^T(k+L-1)\tilde{\theta}_i(k-1)]}{||U(k+L-1)||^2} \mid \hat{\theta}_i(k-1) \right\} \quad i = 1,2,\ldots,m \tag{A2.8}$$

Now, from conditions (A2.2) on the sequences $e_i(k)$ and $U(k)$ and writing $\hat{\theta}_i(k-1) = \theta_i$ equation (A2.8) gives

$$E \left\{ \frac{\partial \hat{J}_i}{\partial \hat{\theta}_i(k-1)} \mid \hat{\theta}_i(k-1) = \theta_i \right\} = 0 \qquad i = 1,2,\ldots,m \tag{A2.9}$$

Hence the estimate of the parameters θ_i's is unbiased (Wasan [106]).

A2.3 Proof of Convergence

Subtracting θ_i from both sides of equation (A2.3) and substituting for $z_i(k+L)$ from equation (A2.1) we get

$$\tilde{\theta}_i(k+L) = \tilde{\theta}_i(k-1) + \nu(k)\,\frac{U(K+L-1)}{||U(k+L-1)||^2}\,[e_i(k+L) - U^T(k+L-1)\tilde{\theta}_i(k-1)]$$

$$i = 1,2,\ldots,m \quad \text{(A2.10)}$$

where $\tilde{\theta}_i(k)$ is as defined before. Taking the norm of both sides of equation (A2.10) we get

$$||\tilde{\theta}_i(k+L)||^2 = \tilde{\theta}_i^T(k-1)\,[I - 2\,\frac{\nu(k)}{||U(k+L-1)||^2}\,U(k+L-1)\,U^T(k+L-1)]\,\tilde{\theta}_i(k-1)$$

$$+ \frac{\nu^2(k)e_i(k+L)}{(||U(k+L-1)||^2)^2}\quad \tilde{\theta}_i^T(k-1)U(k+L-1)U^T(k+L-1)U(k+L-1)U^T(k+L-1)$$

$$+ \frac{2\nu(k)e_i(k+L)}{(||U(k+L-1)||^2)^2}\,\tilde{\theta}_i^T(k-1)U(k+L-1)$$

$$- \frac{2\nu^2(k)e_i(k+L)}{(||U(k+L-1)||^2)^2}\,\tilde{\theta}_i^T(k-1)U(k+L-1)U^T(k+L-1)U(k+L-1)$$

$$+ \frac{\nu^2(k)e_i^2(k+L)}{(||U(k+L-1)||^2)^2}\,U^T(k+L-1)U(k+L-1)$$

$$= \tilde{\theta}_i^T(k-1)\,[I - \frac{2\nu(k)}{||U(k+L-1)||^2}\,U(k+L-1)U^T(k+L-1)]\,\tilde{\theta}_i(k-1)$$

$$+ \frac{\nu^2(k)}{||U(k+L-1)||^2}\quad \tilde{\theta}_i^T(k-1)U(k+L-1)U^T(k+L-1)\tilde{\theta}_i(k-1)$$

$$+ \frac{2\nu(k)e_i(k+L)}{(||U(k+L-1)||^2)^2} \tilde{\theta}_i^T(k-1)U(k+L-1)$$

$$- \frac{2\nu^2(k)e_i(k+L)}{||U(k+L-1)||^2} \tilde{\theta}_i^T(k-1)U(k+L-1) + \frac{\nu^2(k)e_i^2(k+L)}{||U(k+L-1)||^2}$$

$$i = 1,2,\ldots,m \qquad (A2.11).$$

Taking the following conditional expectation and from the properties of the sequences $e_i(k)$ and $U(k)$, described in equation (A2.2), we get

$$E\{||\tilde{\theta}_i(k+L)||^2|\tilde{\theta}_i(k-1) = \Omega_i\} = \Omega_i^T [I - 2\nu(k) E \{ \frac{U(k+L-1)U^T(k+L-1)}{||U(k+L-1)||^2} \}] \Omega_i$$

$$+ \nu^2(k) \Omega_i^T E \{ \frac{U(k+L-1)U^T(k+L-1)}{||U(k+L-1)||^2} \} \Omega_i + \nu^2(k) E \{ \frac{e_i^2(k+L)}{||U(k+L-1)||^2} \}$$

$$i = 1,2,\ldots,m \qquad (A2.12)$$

which can be reduced to

$$E\{||\tilde{\theta}_i(k+L)||^2|\tilde{\theta}_i(k-1) = \Omega_i\} \leq \Omega_i^T [I-2\nu(k)E\{ \frac{U(k+L-1)U^T(k+L-1)}{||U(k+L-1)||^2} \}] \Omega_i$$

$$+ c_1\nu^2(k)||\Omega_i||^2 + c_2\nu^2(k) \qquad i = 1,2,\ldots,m \qquad (A2.13)$$

where c_1 and c_2 are positive constants.

By assumption (A2.2), the matrix $E\{ \frac{U(k+L-1)U^T(k+L-1)}{||U(k+L-1)||^2} \}$ is positive definite, hence it has a minimum eigenvalue $\lambda_{min} > 0$ (Graybill [107]). Therefore equation (A2.13) can be written as

$$E\{||\tilde{\theta}_i(k+L)||^2|\tilde{\theta}_i(k-1) = \Omega_i\} \leq ||\Omega_i||^2 - 2\lambda_{min} \nu(k)||\Omega_i||^2$$

$$+ c_1 \nu^2(k) ||\Omega_i||^2 + c_2 \nu^2(k) \qquad i = 1,2,\ldots,m \qquad (A2.14)$$

and hence

$$E\{||\tilde{\theta}_i(k+L)||^2 | \tilde{\theta}_i(k-1) = \Omega_i\} \leq [1-2\,\lambda_{min}\,\nu(k)$$

$$+ c_1 \nu^2(k)] E\{||\tilde{\theta}_i(k-1)||^2\} + c_2 \nu^2(k) \qquad i = 1,2,\ldots,m \qquad (A2.15)$$

Under conditions (A2.4) equation (A2.15) satisfies the requirements of a convergence proof due to Dovretzky [91] from which it follows that

$$\lim_{k\to\infty} E\{||\hat{\theta}_i(k+L)||^2\} = 0$$

and $$\text{prob}\,[\lim_{k\to\infty} \hat{\theta}_i(k+L) = \theta_i] = 1 \qquad i = 1,2,\ldots,m$$

REFERENCES

[1] R.C. Desai and C.S. Lalwani, "Identification Techniques", Tata McGraw-Hill, Bombay, 1972.

[2] D. Graupe, "Identification of Systems", Van Nostrand Reinhold, New York, 1972.

[3] A.P. Sage and J.L. Melsa, "System Identification", Academic Press, New York, 1972.

[4] J.M. Mendel, "Discrete Techniques of Parameter Estimation: the equation error approach", Marcell Dekker Inc., New York, 1973.

[5] P. Eykhoff, "System Identification", John Wiley and Sons, Ltd., London, 1974.

[6] R.K. Mehra and D.G. Lainiotis, "System Identification: Advances and Case Studies", Academic Press, Inc., 1976.

[7] M. Cuenod and A.P. Sage, "Comparison of some methods used for process identification", Automatica, vol. 4, pp. 235-269, 1968.

[8] P. Eykhoff, "Process parameter and state estimation", Automatica, vol. 4, pp. 205-233, 1968.

[9] K.J. Astrom and P. Eykhoff, "System identification - a survey", Automatica, vol. 7, pp. 123-162, 1971.

[10] R.E. Nieman, D.G. Fischer and D.E. Seborg, "A review of process identification and parameter estimation techniques", Int. J. of Control, pp. 209-264, 1971.

[11] K.J. Astrom, "System identification", in "Lecture notes in mathematics", vol. 294, pp. 35-55, Springer, 1972.

[12] I. Gustavsson, "Comparison of different methods for identification of industrial processes", Automatica, vol. 8, pp. 127-142, 1972.

[13] N.K. Sinha and A. Sen, "On-line system identification: a critical survey", Internal Report, SOC-14, Faculty of Eng., McMaster Univ., Hamilton, Ont., Canada, 1973.

[14] R. Iserman, U. Baur, W. Bamberger, P. Knepo and H. Siebert, "Comparison of six on-line identification and parameter estimation methods", Automatica, vol. 10, pp. 81-103, 1974.

[15] P.N. James, P. Souter and D.C. Dixon, "A comparison of parameter estimation algorithms for discrete systems", Ch. Eng. Science, pp. 539-547, 1974.

[16] G.N. Saridis, "A comparison of six on-line identification algorithms", Automatica, vol. 10, pp. 69-79, 1974.

[17] G.N. Saridis, "Stochastic approximation methods for identification and control - a survey", IEEE Trans. Aut. Control, vol. AC-19, pp. 798-809, 1974.

[18] T. Soderstrom, L. Ljung and I. Gustavsson, "A comparative study of recursive identification methods", Report 7427, Lund Inst. of Tech., Dept. of Automatic Control, 1974.

[19] N.K. Sinha and A. Sen, "Comparison of six different methods for on-line estimation of the parameters of a dual-input heat exchanger", Automatic Control Theory and Application, vol. 3, pp. 62-66, 1975.

[20] N.S. Rajbman, "The application of identification methods in U.S.S.R. - a survey", Automatica, vol. 12, pp. 73-95, 1976.

[21] N.S. Rajbman and N.K. Sinha, "Identification of multivariable systems - a critical survey", Proc. of 14th Annual Allerton Conf. on Circuits and Systems, 1976.

[22] N.K. Sinha, "On-line identification of linear multivariable systems - a survey and new results", Proc. Int. Conf. on Information Sciences and Systems, Patras, Greece, pp. 301-305, 1977.

[23] N.K. Sinha and H. El-Sherief, "Stochastic approximation algorithms for system identification", Proc. 8th Annual Pittsburgh Conf. on Modeling and Simulation, pp. 371-375, 1977.

[24] B. Gopinath, "On the identification of linear time-invariant systems from input-output data", Bell System Technical Journal, vol. 48, pp. 1101-1113, 1969.

[25] M.A. Budin, "Minimal realization of discrete linear systems from input output observations", IEEE Trans. Aut. Control, vol. AC-16, pp. 395-401, 1971.

[26] L.A. Zadeh, "From circuit theory to system theory", Proc. IRE 50, 856-865, 1962.

[27] H. El-Sherief and N.K. Sinha, "Identification and modeling for linear multivariable discrete-time systems - a survey", Journal of Cybernetics, vol. 9, pp. 43-71, 1979.

[28] A. Sen and N.K. Sinha, "On-line estimation of the parameters of a multivariable system using matrix pseudo-inverse", Int. J. of Syst. Science, vol. 7, pp. 461-471, 1976.

[29] H. El-Sherief and N.K. Sinha, "Stochastic approximation for the identification of linear multivariable systems", IEEE Trans. Aut. Control, vol. AC-24, 1979.

[30] H. El-Sherief and N.K. Sinha, "Models for the identification of linear multivariable discrete-time systems", Proc. 16th Annual Allerton Conference on Communication, Control and Computing, October 1978.

[31] B.A. Abaza, "Multivariable parameter estimation using a 2-step least squares estimator", Elect. Lett., No. 13, pp. 331-333, 1976.

[32] S. Sinha and P.E. Caines, "On the use of register sequences as instrumental variables for the recursive identification of multivariable linear systems", Int. J. of Syst. Science, vol. 8, pp. 1041-1055, 1977.

[33] D.P. Mital and C.T. Chen, "Identification of multivariable systems from input-output data", Proc. of Int. Conf. on System, Man and Cyb., pp. 422-426, 1974.

[34] R.K. Mehra, "On-line identification of linear dynamic systems with applications to Kalman filtering", IEEE Trans. Aut. Control, vol. AC-16, pp. 12-21, 1971.

[35] H. El-Sherief and N.K. Sinha, "A nonparametric normalized stochastic approximation algorithm for on-line identification of multivariable systems", Int. J. of Syst. Science, accepted for publication.

[36] N.K. Sinha, A.K. Mahalanabis and H. El-Sherief, "A nonparametric approach to the identification of linear multivariable systems", Int. J. of Syst. Science, vol. 9, pp. 420-430, 1978.

[37] R.L. Kashyap and R.E. Nasburg, "Parameter estimation in multivariable stochastic difference equations", IEEE Trans. Aut. Control, vol. AC-19, pp. 784-797, 1974.

[38] E.J. Hannan, "The identification program for multiple equation systems with moving average errors", Econometric, vol. 39, No. 5, pp. 751-765, 1971.

[39] H. Akaike, "Maximum likelihood identification of Gaussian autoregressive moving average errors", Biometrica, vol. 60, pp. 255-265, 1973.

[40] B.W. Dickinson, T. Kailath and M. Morf, "Canonical matrix fraction and state-space descriptions for deterministic and stochastic linear systems", IEEE Trans. Aut. Control, vol. AC-19, pp. 656-667, 1974.

[41] R. Guidorzi, "Canonical structures in the identification of multivariable systems", Automatica, vol. 11, pp. 361-374, 1975.

[42] N.K. Sinha and Y.H. Kwong, "Recursive identification of the parameters of multivariable systems", Proc. of IFAC Conf. on Multivariable Tech. Syst., pp. 323-327, Canada, 1977.

[43] H. El-Sherief and N.K. Sinha, "Algorithm for identification of multivariable systems combining stochastic approximation and pseudo-inverse", Automatic Control Theory and Applications, vol. 6, pp. 37-40, 1978.

[44] K. Furuta, "An application of realization theory to identification of multivariable process", Proc. of 3rd IFAC Symp. on Identification and System Parameter Estimation, pp. 939-941, (The Hague), 1973.

[45] R.N. Lobbia and G.N. Saridis, "Identification and control of multivariable stochastic discrete systems", Journal of Cybernetics, pp. 40-59, 1973.

[46] H. Weinert and J. Anton, "Canonical forms for multivariable system identification", Proc. of IEEE Conf. on Decision and Control, pp. 37-39, 1972.

[47] E. Tse and H. Weinert, "Structure determination and parameter identification for multivariable stochastic linear systems", IEEE Trans. Aut. Control, vol. AC-20, pp. 603-613, 1975.

[48] S.P. Bingulac and M.A.C. Farias, "Identification and minimal realization of multivariable systems", Proc. of IFAC Conf. on Multivariable Tech. Syst., pp. 373-377, Canada, 1977.

[49] H.J. Denham, "Canonical forms for the identification of multivariable linear systems", IEEE Trans. Aut. Control, vol. AC-19, pp 646-656, 1974.

[50] J.V. Beck and K.J. Arnold, "Parameter Estimation in Engineering and Science", John Wiley and Sons, New York, 1977.

[51] W.A. Wolovich, "Linear Multivariable Systems", Springer-Verlag, New York, 1974.

[52] H. El-Sherief and N.K. Sinha, "Identification of multivariable systems in the transfer-function matrix form", Journal of Cybernetics, vol. 9, pp. 113-125, 1979.

[53] N.K. Sinha and W. Pille, "On-line parameter estimation using the matrix pseudo-inverse", Proc. of IEE, vol. 118, pp. 1041-1046, 1971.

[54] L.C. Suen and R. Liu, "Determination of the structure of multivariable stochastic linear systems", IEEE Trans. Aut. Control, vol. AC-23, pp. 458-464, 1978.

[55] V. Panuska, "Adaptive recursive least squares identification algorithm", Proc. of 8th IEEE Symp. Adaptive Processes, 1969.

[56] A. Albert, "Regression and the Moore-Penrose Pseudo-Inverse", New York, Academic Press, 1972.

[57] B.L. Ho and R.E. Kalman, "Effective construction of linear state-variable models from input-output functions", Regelungstechnik, vol. 14, pp. 545-548, 1966.

[58] P. Rozsa and N.K. Sinha, "Minimal realization of a transfer function matrix in canonical forms", Int. J. of Control, vol. 21, pp. 273-284, 1975.

[59] Y. Bar-Shalom and S.C. Schwartz, "A nonparametric recursive identification algorithm for finite memory linear systems", Proc. 7th Annual Allerton Conf. on Circuit and System Theory, pp. 523-532, 1969.

[60] Y.C. Ho and R.C. Lee, "Identification of linear dynamic systems", J. Inform. Control, vol. 8, pp. 93-110, 1965.

[61] H.B. Mann and A. Wald, "On the statistical treatment of linear stochastic difference equations", Econometrica, vol. 11, pp. 173-220, 1943.

[62] N.K. Sinha and A. Sen, "On-line identification of multivariable systems", Proc. of 4th IFAC Symposium on System Identification (USSR), 1976.

[63] P. Rozsa and N.K. Sinha, "Efficient algorithm for irreducible realization of a rational matrix", Int. J. of Control, vol. 20, pp. 739-751, 1974.

[64] D.Q. Mayne, "A canonical model for identification of multivariable linear systems", IEEE Trans. Aut. Control, vol. AC-17, pp. 728-729, 1972.

[65] G.W. Irwin and A.P. Roberts, "The Luenberger canonical form in the state/parameter estimation of linear systems", Int. J. of Control, vol. 23, pp. 851-864, 1976.

[66] R. Liu and L.C. Suen, "Minimal dimension and identifiability of input-output sequences", IEEE Trans. Aut. Control, vol. AC-22, pp. 227-232, 1977.

[67] B.L. Ho. and R.E. Kalman, "Markov parameters, the moment problem and minimal realizations", SIAM J. Control, vol. 5, 1967.

[68] J. Valis, "On-line identification of multivariable linear systems of unknown structure from input-output data", Proc. of IFAC Conf. on Identification and Proc. Par. Estimation, Prague, 1970.

[69] B. Carew and P.R. Belanger, "Identification of optimum filter steady-state gain for system with unknown noise covariances", IEEE Trans. Aut. Control, vol. AC-18, pp. 582-587, 1973.

[70] L.W. Nelson and E. Stear, "The simultaneous on-line estimation of parameters and states in linear systems", IEEE Trans. Aut. Control, vol. AC-21, pp. 94-98, 1976.

[71] R.E. Kalman, "A new approach to linear filtering and prediction problems", Trans. ASME (J. Basic Eng.), Series D, vol. 82, pp. 35-45, 1960.

[72] W.C. Martin and A.R. Stubberud, "An uncoupling method for linear system identification", IEEE Trans. Aut. Control, vol. AC-21, pp. 506-509, 1976.

[73] E.V. Bohn and M.K. DeBeer, "Consistent parameter estimation in multi-input multi-output discrete systems", Automatica, vol. 13, pp. 301-305, 1977.

[74] P. Blessing, "Parameter estimation of state-space models for multivariable systems with correlation analysis and method of least squares", Proc. of IFAC Symp. on Multivariable Tech. Systems, pp. 385-394, Canada, 1977.

[75] Ph. De Larminat and C. Doncarli, "A real-time generalized least squares estimation method for identification of linear systems", Proc. of IFAC Symp. on Multivariable Tech. Systems, pp. 379-384, Canada, 1977.

[76] G. Salut and G. Gavier, "Identification et poursuite adaptive optimale des systemes lineaires stochastic", Report final DRME LAAS-GESSY no. 73/564, Nov. 1975.

[77] C. Doncarli, "Identification en temps reel des systemes lineaires stochastiques", Universite de Nanter, 1977.

[78] H. El-Sherief and N.K. Sinha, "Determination of the structure of a canonical model for the identification of linear multivariable systems", <u>5th IFAC Symp. on Identification and System Parameter Estimation</u>, Darmstadt, F.R. of Germany, Sept. 1979.

[79] S. Beghelli and R. Guidorzi, "A new input-output canonical form for multivariable systems", <u>IEEE Trans. Aut. Control</u>, vol. AC-21, pp. 692-696, 1976.

[80] H. El-Sherief and N.K. Sinha, "Bootstrap estimation of parameters and states of linear multivariable systems", <u>IEEE Trans. Aut. Control</u>, vol. AC-24, 1979.

[81] D.G. Luenberger, "Canonical forms for linear multivariable systems", <u>IEEE Trans. Aut. Control</u>, vol. AC-12, pp. 290-293, 1967.

[82] H.G. Kwatny, "A note on stochastic approximation algorithm in system identification", <u>IEEE Trans. Aut. Control</u>, vol. AC-17, pp. 571-572, 1972.

[83] A. Sen and N.K. Sinha, "On-line system identification algorithm combining stochastic approximation and pseudoinverse", <u>Automatica</u>, vol. 11, pp. 425-429, 1975.

[84] D.W. Clarke, "Generalized least squares estimation of the parameters of dynamic model", <u>Proc. of 1st IFAC Symposium on Identification</u> (Prague), 1967.

[85] J.B. Farison, R.E. Graham and R.C. Shelton, "Identification and control of linear discrete systems", <u>IEEE Trans. Aut. Control</u>, vol. AC-12, pp. 438-442, 1967.

[86] A.H. Jazwinski, "<u>Stochastic Processes and Filtering Theory</u>", New York, Academic, 1970.

[87] R.M. Prasad and A.K. Sinha, "On Bootstrap identification using stochastic approximation", <u>IEEE Trans. Aut. Control</u>, vol. AC-22, pp. 671-672, 1977.

[88] R.A. Padilla, C.S. Padilla and S.P. Bingulac, "Comments on the simultaneous on-line estimation of parameters and states in linear systems", IEEE Trans. Aut. Control, vol. AC-23, pp. 96-97, 1978.

[89] T.C. Hsia, "On least squares algorithms for system parameters identification", IEEE Trans. Aut. Control, vol. AC-23, pp. 104-108, 1976.

[90] N.K. Sinha and P. Rozsa, "Some canonical forms for linear multivariable systems", Int. J. of Control, vol. 23, pp. 365-383, 1976.

[91] A. Dvoretzky, "On stochastic approximation", in Proc. 3rd Berkeley Symp. on Mathematical Statistics and Probability, vol. 1, Los Angeles, Univ. of Calif. Press, pp. 39-56, 1956.

[92] J. Hickin and N.K. Sinha, "Canonical forms for aggregated models", Int. J. of Control, pp. 473-485, 1978.

[93] D. Graupe, "On identification of closed-loop systems", IEEE Trans. Aut. Control, vol. AC-20, pp. 553-555, 1975.

[94] G. Saridis and R. Lobbia, "Parameter identification and control of linear discrete-time systems", IEEE Trans. Aut. Control, AC-17, pp. 52-60, 1972.

[95] G. Box and J. MacGregor, "The analysis of closed-loop dynamic stochastic systems", Techometrics, vol. 16, pp. 391-398, 1974.

[96] L. Ljung, I. Gustavsson and T. Soderstrom, "Identification of linear multivariable systems operating under linear feedback control", IEEE Trans. Aut. Control, vol. AC-19, pp. 836-840, 1974.

[97] L. Ljung, "Analysis of recursive stochastic algorithms", IEEE Trans. Aut. Control, vol. AC-22, pp. 551-575, 1977.

[98] L. Ljung, "Convergence of an adaptive filter algorithm", Int. J. of Control, vol. 27, pp. 673-693, 1978.

[99] T. Soderstrom, L. Ljung and I. Gustavsson, "A theoretical analysis of recursive identification methods", Automatica, vol. 14, pp. 231-244, 1978.

[100] N.K. Sinha and M.Y. Tang, "On-line system identification in real time using a minicomputer", Trans. IMACS, vol. XIX, pp. 113-120, 1977.

[101] D.M. Wiberg, "State Space and Linear Systems", New York, McGraw-Hill, 1971.

[102] H. El-Sherief and N.K. Sinha, "On the Convergence and unbiasedness of stochastic approximation algorithm for the identification of linear multivariable systems", IEEE Trans. Aut. Control, vol. AC-24, 1979.

[103] A.E. Albert and L.A. Gardner, "Stochastic Approximation and Linear Regression", MIT Press, 1967.

[104] H.G. Kwatny and D.W.C. Shen, "Identification of nonlinear systems using a method of stochastic approximation", Proc. of 8th JACC, pp. 814-826, 1967.

[105] J. Nagumo and A. Noda, "A learning method for system identification", IEEE Trans. Aut. Control, vol. AC-12, pp. 282-287, 1967.

[106] M.T. Wasan, "Parametric Estimation", McGraw-Hill, New York, 1970.

[107] F.A. Graybill, "Introduction to Matrices with Application in Statistics", Wadsworth Pub. Co., 1969.