MIMO SYSTEMS PARAMETERS IDENTIFICATION

by

Abdelhak Bennia

Thesis submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Master of Science

in

Electrical Engineering

APPROVED:

Hugh F. VanLandingham, Chairman

V. Vorperian

W. T. Baumann

June, 1986

Blacksburg, Virginia

MIMO SYSTEMS PARAMETERS IDENTIFICATION

by

Abdelhak Bennia

Hugh F. VanLandingham, Chairman

Electrical Engineering

(ABSTRACT)

In this thesis, a presentation of a new canonical representation of multi-input multi-output systems is given. The new characterization covers the full range of practical situations in linear systems according to the structural properties and model of the perturbations which are known. Its direct link to ARMA processes as well as to classical state space representation is also given.

The importance of the new representation lies in the fact that all unknown parameters and state variables appear linearly multiplied by either external variables (inputs and outputs) that appear in the data record, or by matrices that are only composed of zeroes and ones. This property enables us to perform a joint state and parameters estimation. Moreover, if the noises are gaussian and their statistics are known, an on-line algorithm that involves a standard discrete-time time-varying Kalman filter is proposed and used successfully in the estimation of unknown parameters for simulated examples.

Acknowledgements

The author wishes to express his sincere gratitude to his advisor Dr Hugh. F. VanLandingham, for his invaluale assistance and encouragement during the preparation of this thesis. It has been a real pleasure working with him and will continue to be so.

The author is also grateful to Dr V. Vorperian and Dr W. T. Baumann who agreed to spend their valuable time by being on his graduate commitee.

The author finally, would like to thank his parent without whom he would not be here for his graduate studies.

Table of Contents

Attention Patron:

Page vi omitted from numbering

INTRODUCTION

During the last two decades, the importance of model building and identification has increased greatly and has attracted considerable attention in many fields e.g chemical processes, biomedical engineering, economic systems, hydrology, aeronautics, etc. When faced with an unknown system, control engineers sharpen their ability to identify it, to obtain a more precise description of it.

Whether it is question of an automatic pilot of a ship, distillation unit, cement kiln, or even the economy of a country, we often seek to give a mathematical model of the phenomena, that is to say find a system of equations whose resolution gives results and predictions in compliance with the observations. The automatic systems will be elaborated according to those mathematical models.

To obtain the system equations, we rely on physical laws which will often be of certain utility provided that the phenomena remain relatively simple, but it is often impractical to put in equations all the detail of the phenomena governing a complicated process.

The control engineer is often only interested in the input-output relation of the process, the inputs being the set of the control values (e.g. flow of fuel in a thermal power station) and the outputs being the variables to control (e.g. steam pressure of a turbine).

In those conditions, instead of trying to establish equations for the process, the usual procedure of the control engineer is to identify it, that is to say obtain a model from the observation of the inputs and outputs of the system.

In the last several years, different algorithms for parameter identification have been developed. Most of these algorithms were developed for discrete-time systems. Some of them have already been successfully applied to engineering systems. For off-line identification , the maximum likelihood method [3] is a very successful approach to obtaining unbiased estimates even when the noise level is quite high. The generalized least squares method [9] also gives unbiased estimates, it is an essantially iterative procedure for solving a highly nonlinear minimization problem. Also, the instrumental variable method [33,38], is less complicated to use, however, less efficient than the others. For on-line identification, several methods were proposed. Some of them used the same algorithms described above; an on-line maximum likelihood algorithm [12,31] which is less convenient for on-line identification. an instrumental variable algorithm [3,38,31] a somewhat unattractive method because of its complexity. Among others, the stochastic approximation [3,31] requires less computation and gives good estimates. The crosscorelation algorithms [3,31], the extended Kalman filter [31], which suffers from a large computational programming, and numerous variations of those algorithms have been developed to improve the rate of convergence and other computational aspects.

1.0 PROBLEM STATEMENT AND IDENTIFICATION APPROACH

At the present time, many formulations of the problem of identification of multivariable systems exist, and numerous techniques have also been developed to estimate systems' parameters. One formulation is that of state space representation. The basic idea behind it, is to explore a class of state space canonical models, linked by particularly simple relations to input-output data , that can be directly identified from the input-output sequences. A fundamental contribution in this field can be found in the papers by Ho and Kalman [14], Budin [6], their methods present a direct procedure for minimal realization in a well defined structure. Guidorzi [13] has identified a system in state space form by first estimating the parameters of the equivalent input-output difference equation form. Then the system matrices (A,B,C) are recovered from these estimated parameters. Because of the nonuniqueness of the representation in the state space form, several canonical forms that reduce the number of parameters to be estimated, as well as make the problem of identification simpler, have been proposed [11]. Using canonical state space representation in identification requires the knowledge of some extra structural parameters [5,8,11]. Most of the exciting algorithms for identification of multivariable systems are based on one of the following methods: least squares, stochastic approximation, maximum likelihood and correlation. These algorithms are similar in many aspects and they concentrate mainly on the problem of removing the bias in the parameter estimates introduced due to the choice of the model. In the identification approach proposed in this work, an algorithm that involves an ordinary Kalman filter is presented.

The problem investigated in this study is as follows: consider an ARMA model described by

$$A(z)y(k) = B(z)u(k) \tag{1.1}$$

$$G(z) = A^{-1}(z)B(z) \tag{1.2}$$

where A(z) and B(z) are polynomial matrices and G(z) a rational transfer-function matrix.

An explicit canonical form is available see Salut [29,30], where all unknown parameters appear linearly when an input-output record is available, is developed. The scheme utilized here, is additionally suitable when state estimation is required along with parameter identification, it invoves augmenting the state variables of the system by adjoining to them the unknown parameters vectors and treating them as a part of the new state variable vector.

$$S(k) = [\ X^T(k) \mid A^T \mid B^T\ ]^T \tag{1.3}$$

The importance of this canonical form, lies in the fact that a pseudo-linear description of the system can be derived. Using the extended state vector, the system can be described by:

$$S(k+1) = F[y(k),u(k)]S(k) \tag{1.4}$$

$$y(k) = H[y(k),u(k)]S(k) \tag{1.5}$$

This structure will enable us to construct an optimal estimator for all unknown parameters and variables of the system. However, in this thesis we will focus on the identification of the parameters that describe the system's rational transfer-function matrix. The algorithm developed for the identification approach involves, as mentioned earlier, a standard discrete-time Kalman filter. However, besides its advantages in control theory, the approach is found to provide relatively fast and accurate results.

1.1 OUTLINE OF THE THESIS

In chapter 2 a study of the problem of identification for single-input-single-output (SISO) systems is described. A state variable modeling approach is used to combine the identification of parameters and the estimation of the states of an ARMA model. The choice of a parametric state space model implies the preliminary determination of the system order. A method for estimating the order from the data is discussed. The resulting identification algorithm involves a standard discrete-time Kalman filter. Finally, an example is included to illustrate the application of this procedure.

In chapter 3 a state space representation is formulated for the identification of multivariable systems in the transfer-function matrix representation. The proposed approach decomposes the system into subsystems (one for each output) such that the parameters of each subsystem are estimated independently. Canonical forms are presented for each of the following cases.

a) Single-input-multi-output (SIMO) systems

b) Multi-input-single-output (MISO) systems

In each case, an algorithm that extends the identification approach presented earlier is used. Also, some examples are included to illustrate the performance of this approach.

In chapter 4 we study the minimal realization of multivariable discrete-time systems from the input-output observations. The problem is to find an irreducible realization of rational transfer-function matrices, and to perform an estimation of the parameters. The possibility of identification of such systems is discussed. The identification requires the preliminary estimation of some structural parameters, a method for structural identification is also presented for that purpose. Finally, an illustration of the canonical representation and an example of identification are presented.

In chapter 5 concluding remarks are made and areas of further study and extensions are discussed.

SISO SYSTEMS IDENTIFICATION

2.0 Introduction

In many control problems, the main objective is to feedback the states of the system in order to modify its behavior. Hence, in addition to the estimation of the parameters of a model, the estimation of the states of the system from noise contaminated measurements of the inputs and outputs is also necessary.

The problem of combined state and parameter estimation was originally posed as a nonlinear state estimation problem by augmenting the state vector with the parameter vector [29,31].

In this chapter a presentation of a new canonical representation [29] of single-input single-output (SISO) discrete-time linear stochastic systems is given. Its equivalence with pulse transfer function form interpretation, directly linked to autoregressive-moving average (ARMA) processes as well as with classical state-space representation, is presented.

The importance of this canonical representation is mainly in the fact that a simultaneous joint state and parameter estimation can be performed. A pseudo-linear formulation of this canonical representation is derived, that is linear in the extended vector but time-varying. Therefore, a discrete Kalman filter can be implemented to yield the estimates of the states and unknown parameters. An estimation algorithm for optimal linear estimation of the parameters is proposed for this purpose.

It will be assumed that the order of the model is known a priori, and that equispaced samples of the input/output data are available. In practise, the order is not known a priori. Hence, in this chapter we shall also study a method [13] for estimating the order of single-input single-output systems.

2.1 Canonical Representation

Consider a single-input single-output ARMA model given by its n th order discrete-time transfer function

$$\frac{Y(z)}{U(z)} = \frac{b_n z^n + b_{n-1} z^{n-1} + \cdots + b_1 z + b_0}{z^n - a_{n-1} z^{n-1} - \cdots - a_1 z - a_0} \tag{2.1}$$

which can be written as

$$H(z) = \frac{b_n + b_{n-1} z^{-1} + \cdots + b_1 z^{-n+1} + b_0 z^{-n}}{1 - a_{n-1} z^{-1} - \cdots - a_1 z^{-n-1} - a_0 z^{-n}} \tag{2.2}$$

from equation (2.2) let us solve for Y(z)

$$Y = b_n U + z^{-1}(b_{n-1} U + a_{n-1} Y) + \cdots + z^{-n}(b_0 U + a_0 Y) \tag{2.3}$$

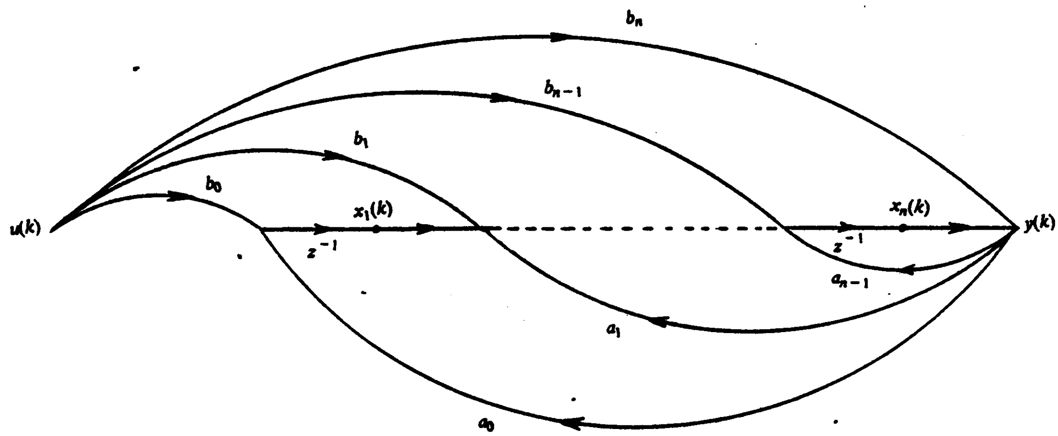

Figure 1. Signal-flow graph for equation (2.3)

The resulting state model, called the obsevable state model, is given by the following structure

$$X(k+1) = \Phi X(k) + \Gamma u(k) \tag{2.4}$$

$$y(k) = CX(k) + Dy(k) \tag{2.5}$$

where

$$\Phi(n \times n) = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & a_0 \\ 1 & 0 & 0 & \dots & 0 & a_1 \\ 0 & 1 & 0 & \dots & 0 & a_2 \\ : & : & : & \dots & : & : \\ : & : & : & \dots & : & : \\ 0 & 0 & 0 & \dots & 0 & a_{n-2} \\ 0 & 0 & 0 & \dots & 1 & a_{n-1} \end{bmatrix} \qquad \Gamma(n \times 1) = \begin{bmatrix} b_0 + a_0 b_n \\ b_1 + a_1 b_n \\ b_2 + a_2 b_n \\ :: \\ :: \\ b_{n-2} + a_{n-2} b_n \\ b_{n-1} + a_{n-1} b_n \end{bmatrix}$$

$$C(1 \times n) = [0\ 0\ 0\ \dots\ 0\ 1] \qquad D(1 \times 1) = [b_n]$$

Thus, our problem is the determination of the parameters a_0, a_1, ..., a_{n-1}, b_0, b_1, ..., b_n from the input/output data. The number of parameters to be estimated is $N = 2n + 1$.

Let the parameters vectors be

$$A^T = [\ a_0\ a_1\ a_2\ \dots\ a_{n-1}\ ]$$

and

$$B_p^T = [\ b_0\ b_1\ b_2\ \dots\ b_{n-1}\ ]$$

It is clear that the matrices Φ and Γ can be expressed as follows:

$$\Phi = T + A \times C \tag{2.6}$$

$$\Gamma = B_p + A \times D \tag{2.7}$$

where

$$T(n \times n) = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ : & : & : & \dots & : & : \\ : & : & : & \dots & : & : \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix}$$

subtituting the expressions of Φ and Γ into (2.4), it can be shown that

$$X(k+1) = (T + A \times C)X(k) + (B_p + A \times D)u(k)$$

ie

$$X(k+1) = TX(k) + A(CX(k) + Du(k)) + B_p u(k)$$

using equation (2.5) this leads to

$$X(k+1) = TX(k) + Ay(k) + B_p u(k)$$

Therefore equations (2.4) and (2.5) may be expressed as:

$$X(k+1) = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ : & : & : & \dots & : & : \\ : & : & : & \dots & : & : \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix} X(k) + \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ : \\ : \\ a_{n-2} \\ a_{n-1} \end{bmatrix} y(k) + \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ : \\ : \\ b_{n-2} \\ b_{n-1} \end{bmatrix} u(k) \quad (2.8)$$

$$y(k) = [0\ 0\ 0\ \dots\ 0\ 1]\ X(k) + [b_n]\ u(k) \tag{2.9}$$

let the total B parameter vector be

$$B^T = \left[B_p^T | b_n\right]$$

Now let us introduce the total parameter vector

$$P^T(k) = \left[A^T | B^T\right] \quad (2n + 1) \times 1 \quad matrix$$

It can be shown that equations (2.8) and (2.9) may also be written as

$$X(k+1) = TX(k) + G(k)P(k) \tag{2.10}$$

$$P(k+1) = \mathbf{I}_{2n+1}P(k) \tag{2.11}$$

$$y(k) = CX(k) + N(k)P(k) \tag{2.12}$$

where

$$G(k) = \left[y(k) \times \mathbf{I}_n \mid u(k) \times \mathbf{I}_n \mid \mathbf{0}\right] \quad n \times (2n+1) \quad matrix$$

$$\mathbf{I}_i = i \times i \quad identity \quad matrix$$

$$N(k) = [\ 0\ 0\ 0\ \dots\ 0\ u(k)\ ] \quad 1 \times (2n+1) \quad row \quad matrix$$

Using the extended vector that involves augmenting the state variables of the system by adjoining to them the unknown parameter vectors and treating them as part of the new state variable vector.

$$S^T(k) = \left[X^T(k) \mid A^T \mid B^T\right] = \left[X^T(k) \mid P^T(k)\right] \tag{2.13}$$

The equations (2.8) , (2.9) and (2.10) will have the following pseudo-linear form for the augmented state vector

$$S(k+1) = F(k)S(k) \tag{2.14}$$

$$y(k) = H(k)S(k) \tag{2.15}$$

Where

$$F(k) = \begin{bmatrix} T & G(k) \\ 0 & I_{2n+1} \end{bmatrix} \quad (3n+1) \times (3n+1) \quad \textit{matrix}$$

and

$$H(k) = \left[C \mid N(k) \right] \quad 1 \times (3n+1) \ \textit{row matrix}$$

The state representation given by equations (2.14) and (2.15) can be considered as a time-varying linear system. The F(k) and H(k) are time-varying matrices, which can be determined from the available record of input and output measurements at any time k.

If the form of eqns (2.14) and (2.15) is observable, it is possible to determine an optimal linear estimator for the system.

2.2 Observability of the representation

The question of observability of the system given in equations (2.14) and (2.15) can be reduced to whether or not S(0) can be calculated given the sequence of input and output data.

From equations (2.14) and (2.15)

$$S(k+1) = F(k)S(k)$$

$$y(k) = H(k)S(k)$$

this leads to

$$y(0) = H(0)S(0)$$

$$y(1) = H(1)S(1) = H(1)F(0)S(0)$$

$$y(2) = H(2)S(2) = H(2)F(1)F(0)S(0)$$

$$\ldots\ldots = \ldots\ldots$$

$$\ldots\ldots = \ldots\ldots$$

$$y(3n) = H(3n)S(3n) = H(3n)F(3n)F(3n-1)\ldots\ldots F(1)F(0)S(0)$$

Hence,

$$\begin{bmatrix} y(0) \\ y(1) \\ y(2) \\ : \\ : \\ : \\ y(3n) \end{bmatrix} = \begin{bmatrix} H(0) \\ H(1)F(0) \\ H(2)F(1)F(0) \\ :: \\ :: \\ :: \\ H(3n)F(3n-1)F(3n-2)\ldots\ldots F(1)F(0) \end{bmatrix} S(0) = \mathbf{O}S(0) \qquad (2.16)$$

If S(0) is to be determined uniquely, **O** must be nonsingular. The observability condition is that the matrix **O** have rank 3n+1

O is found to have the following structure

$$
\mathbf{O} = \left[\begin{array}{cccccccccccccc}
000\ldots01 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & u_0 \\
000\ldots10 & 0 & 0 & 0 & \cdots & 0 & y_0 & 0 & 0 & 0 & \cdots & 0 & u_0 & u_1 \\
000\ldots00 & 0 & 0 & 0 & \cdots & y_0 & y_1 & 0 & 0 & 0 & \cdots & u_0 & u_1 & u_1 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
010\ldots00 & 0 & 0 & y_0 & \cdots & y_{n-4} & y_{n-3} & 0 & 0 & u_0 & \cdots & u_{n-4} & u_{n-3} & u_{n-2} \\
100\ldots00 & 0 & y_0 & y_1 & \cdots & y_{n-3} & y_{n-2} & 0 & u_0 & u_1 & \cdots & u_{n-3} & u_{n-2} & u_{n-1} \\
000\ldots00 & y_0 & y_1 & y_2 & \cdots & y_{n-2} & y_{n-1} & u_0 & u_1 & u_2 & \cdots & u_{n-2} & u_{n-1} & u_n \\
000\ldots00 & y_1 & y_2 & y_3 & \cdots & y_{n-1} & y_n & u_1 & u_2 & u_3 & \cdots & u_{n-1} & u_n & u_{n+1} \\
000\ldots00 & y_2 & y_3 & y_4 & \cdots & y_n & y_{n+1} & u_2 & u_3 & u_4 & \cdots & u_n & u_{n+1} & u_{n+2} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
000\ldots00 & y_{2n-1} & y_{2n} & y_{2n+1} & \cdots & y_{3n-3} & y_{3n-2} & u_{2n-1} & u_{2n} & u_{2n+1} & \cdots & u_{3n-3} & u_{3n-2} & u_{3n-1} \\
000\ldots00 & y_{2n} & y_{2n+1} & y_{2n+2} & \cdots & y_{3n-2} & y_{3n-1} & u_{2n} & u_{2n+1} & u_{2n+2} & \cdots & u_{3n-2} & u_{3n-1} & u_{3n}
\end{array}\right]
$$

The structure of **O** shows that it is a function of the input and output data, hence, the observability of the system depends on the continued excitation of the system by the input. It is well known that a random excitation, such as a white noise sequence, satisfies this requirement.

2.3 Identification Algorithm

Consider the discrete-time dynamical system given by equations (2.14) and (2.15)

$$S(k+1) = F(k)S(k) \tag{2.14}$$

$$y(k) = F(k)S(k) \tag{2.15}$$

The system (2.14), (2.15) is completely observable in Kalman's sense, assuming O to be nonsingular, the state vector can therefore be reconstructed from at most 3n+1 independent measurements of the output signal. The state variables can, however, be reconstructed from a mathematical model of the system.

Consider for example the following model

$$\hat{S}(k+1) = F(k)\hat{S}(k) \tag{2.17}$$

If (2.17) is a perfect model, i.e, if the model parameters are identical to those of the system, and if the initial conditions of (2.14) and (2.17) are identical, then the state $\hat{S}$ of the model will be identical to the true state $\hat{S}$. If the initial conditions of (2.14) and (2.15) differ, the reconstruction $\hat{S}$ will converge to the true value only if the system (2.14) is asymptotically stable.

By exploiting the difference (y - HS) , we can adjust the estimate $\hat{S}$ given by (2.17), for example by using the reconstruction :

$$\hat{S}(k+1) = F(k)\hat{S}(k) + K(k)\left[y(k) - H(k)\hat{S}(k)\right] \tag{2.18}$$

Where K(k) is a suitably chosen matrix

To get some insight into the proper choice of K(k), we will consider the reconstruction error

$$\bar{S}(k) = S(k) - \hat{S}(k)$$

By substracting (2.18) from (2.14), we get

$$\bar{S}(k+1) = [F(k) - K(k)H(k)]\ \bar{S}(k) \tag{2.19}$$

Hence by introducing a feedback in the observable model, it is possible to reconstruct state variables even in the case the system itself is unstable. By a proper choice of K(k), the reconstruction error $\bar{S}(k)$ will always converge to zero for arbitrary initial states of (2.18).

It is, however, more realistic to model the states of the system as being subject to random perturbations and to model the measurements as being noisy.

The question now is there an optimal choice of K(k). To pose such problem, we must introduce more structure into the problem. To do so, we assume that the system is actually governed by stochastic difference equations.

Consider the random signals y(k), S(k) described by the following equations, which include the processes generating the noises and perturbations

$$S(k+1) = F(k)S(k) + V(k) \tag{2.20}$$

$$y(k) = H(k)S(k) + w(k) \tag{2.21}$$

Where:

S(k) is the extended state vector of the system

y(k) the measured output signal

V(k) and w(k) are white gaussian sequences with correlation functions

$$E\left\{\begin{bmatrix} V(k) \\ w(k) \end{bmatrix}\begin{bmatrix} V(l) \\ w(l) \end{bmatrix}^T\right\} = \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \delta_{kl} \tag{2.22}$$

and more precisely, V(k) and w(k) are independent of all the past of the system

$$E\left\{\begin{bmatrix} V(k) \\ w(k) \end{bmatrix}[S(0) \ \ldots \ S(k-1) \ y(0) \ \ldots \ y(k-1)]\right\} = \quad 0 \tag{2.23}$$

V(k) and w(k) are also, assumed to be independent of the state S(k)

$$E\left\{\begin{bmatrix} V(k) \\ w(k) \end{bmatrix}[\ S(k)\ ]^T\right\} = \quad 0 \tag{2.24}$$

The matrices F(k) and H(k) are available at each time k from the record of the input-output data. We assume also that the matrices Q and R are known, as are the statistical characteristics of the initial state S(0), that is to say its mean value and covariance matrix.

At the time k, we assume the measurement vector

$$\mathbf{Y}^T(k) = [y(0) \ y(1) \ \ldots \ y(k)] \tag{2.25}$$

and seek to estimate the unknown vector

$$\mathbf{S}^T(k) = \left[S^T(0) \ S^T(1) \ \ldots \ S^T(k)\right] \tag{2.26}$$

The assumptions of the problem allow us in theory to calculate the mean values, variances and covariances of S(k) and Y(k), hence, the possibility exists to apply the general formulas of the optimal linear estimation to obtain S(k) as a linear function of Y(k)

The estimation problem can be stated as the choice of the optimal gain K(k) such that the reconstruction error (2.18) is minimal.

The Kalman filtering can be applied to provide a technique for the identification of the parameters and the estimation of the states. From a standard discrete Kalman filter an estimation algorithm is developed. let

S(k/k) = optimal filtered estimate of S(k)

S(k+1/k) = optimal predicted estimate of S(k)

If the optimal filtered estimate S(k/k) and the covariance matrix P(k/k) of the corresponding filtering error S(k/k) are known for some k, then the single-stage optimal predicted estimate for all admissible loss functions is given by the expression

$$\hat{S}(k + 1/k) = F(k)\hat{S}(k/k) \tag{2.27}$$

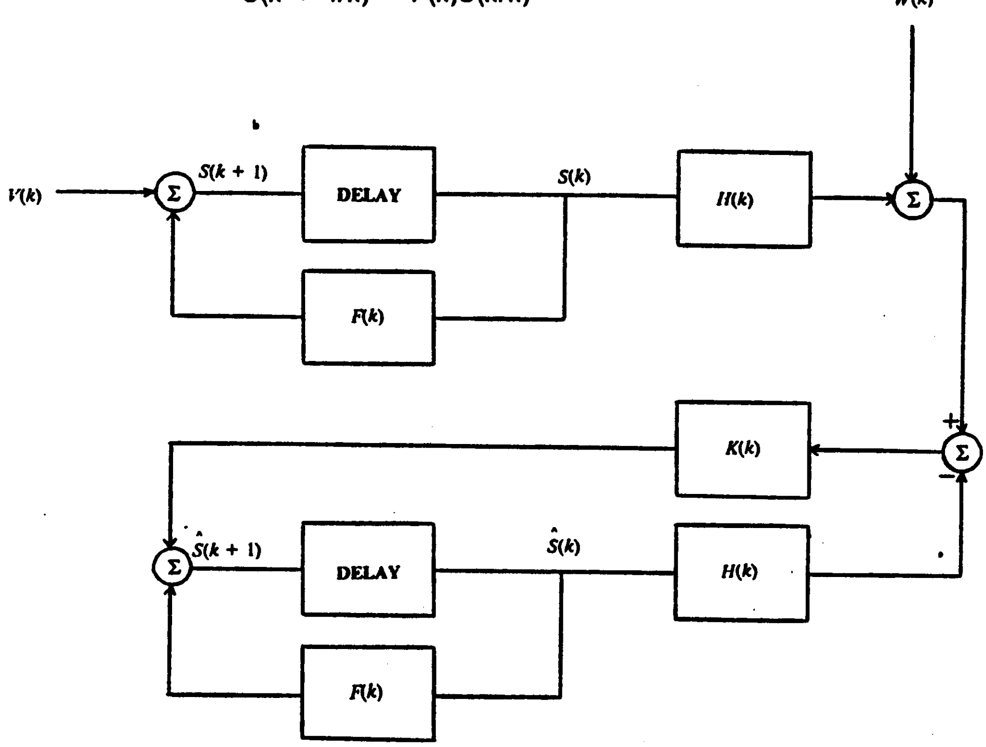

Figure 2. Discrete-time Kalman filter

The stochastic process $\{ S(k+1/k),\ k=0,1,..\}$ defined by the single-stage predicted error relation

$$\bar{S}(k + 1/k) = S(k + 1) - \hat{S}(k + 1/k)$$

is a zero mean gauss-markov process whose covariance matrix is given by the relation

$$P(k + 1/k) = F(k)P(k/k)F^{T}(k) + Q \tag{2.28}$$

The optimal filtered estimate $S(k+1/k+1)$ is given by the recursive relation

$$\hat{S}(k + 1/k + 1) = F(k)S(k/k) + K(k + 1)\left[y(k + 1) - H(k + 1)F(k)\hat{S}(k/k)\right] \tag{2.29}$$

for $k=0,1,....$ where $\hat{S}(0/0) = S(0)$

$K(k+1)$ is a vector of dimension $(3n+1)$ specified by the set of relations

$$K(k + 1) = P(k + 1/k)H^{T}(k + 1)\left[H(k + 1)P(k + 1/k)H^{T}(k + 1) + R\right]^{-1} \tag{2.30}$$

$$P(k + 1/k) = F(k)P(k/k)F^{T}(k) + Q \tag{2.31}$$

$$P(k + 1/k + 1) = [I - K(k + 1)H(k + 1)]P(k + 1/k) \tag{2.32}$$

For $k=0,1,....$ and $P(0/0)=P(0)$

2.4 Structural Identification

2.4.1 Deterministic case

When modeling a single-input single-output system by a difference equation a value for the model order must usually be assumed, and the system order is finally determined by comparing the goodness of fit of several orders of the model. There are some ways to test for the order of the system without first fitting coefficients to models. They give intuitive rather than statistical tests. However, once the order, or a range of orders, is determined an iterative modeling procedure such as maximum likelihood can be applied and stronger tests for the system order can be used.

Now, we shall consider the problem where the order of the system is unknown. In other words, we know that the observations at our disposal came from a linear dynamic system but we know neither the parameters nor the order.

Although the order of the system is unknown, let us assume that it is less than n, where n is some integer consider the following pulse transfer function

$$\frac{y(z)}{u(z)} = \frac{b_m + b_{m-1}z^{-1} + ... + b_1z^{1-m} + b_0z^{-m}}{1 + a_{n-1}z^{-1} + ... + a_1z^{1-n} + a_0z^{-n}} \quad (m \leq n) \tag{2.33}$$

The system may also be expressed in the form of a difference equation as

$$y(k) = \sum_{i=1}^{n} a_{n-i}y(k-i) + \sum_{i=0}^{m} b_{m-i}u(k-i) \tag{2.34}$$

Consider now the matrix of input/output data given by

$$\begin{bmatrix} y(k) & y(k+1) & \ldots\ldots & u(k) & u(k+1) & \ldots\ldots \\ y(k+1) & y(k+2) & \ldots\ldots & u(k+1) & u(k+2) & \ldots\ldots \\ : & : & : & : & : & : \\ : & : & : & : & : & : \\ y(k+N) & \ldots. & \ldots\ldots & u(k+N) & \ldots. & \ldots. \end{bmatrix}$$

$$= [\ Y(k)\ Y(k+1) \ldots.. \mid U(k)\ U(k+1) \ldots..\ ] \tag{2.35}$$

Equation (2.34) shows that the dependence relations among the vectors of (2.35), is directly related to the number of parameters that can be uniquely determined. The vectors of (2.35) will therefore be selected in the following order.

$$Y(k),\ U(k),\ Y(k+1),\ U(k+1),\ \ldots.., \tag{2.36}$$

A vector is retained if and only if it is independent from previously selected ones; when a dependent vector is found, all the remaining vectors belonging to the same submatrix will also be dependent so that their test is unnecessary.

The integer N in (2.35) must be large enough in order to permit the selection of the necessary number of independent vectors.

It is certainly not advisable to carry out the structural identification directly on the vectors of (2.35) because of the large amount of storage necessary; moreover, the required storage would be a function of N. Since for every matrix , D, rankD=rank(D^T D) a more useful algorithm follows. let

$$L_i(Y) = [\ Y(k)\ \ Y(k+1)\ \ \ldots\ldots\ \ Y(k+i-1)\ ] \tag{2.37}$$

$$L_i(U) = [\ U(k)\ \ U(k+1)\ \ \ldots\ldots\ \ U(k+i-1)\ ] \tag{2.38}$$

Then the matrix (2.35) taking δ_1 vectors in the first submatrix and δ_2 vectors in the second can be written as

$$R(\delta_1, \delta_2) = \{ L_{\delta_1}(Y) \mid L_{\delta_2}(U) \} \tag{2.39}$$

Define the product R^TR as

$$S(\delta_1, \delta_2) = R^T(\delta_1, \delta_2)R(\delta_1, \delta_2) \tag{2.40}$$

$S(\delta_1, \delta_2)$ is therefore a square matrix whose dimension is given by $\delta_1 + \delta_2$

Construct then the sequence of increasing dimension matrices

$$S(2,1), S(2,2), \ldots\ldots, S(5,4), \ldots \tag{2.41}$$

and select from (2.41) nonsingular ones. When a singular matrix is found the procedure ends.

let $S(\mu_1, \mu_2)$ be a singular matrix in (2.41) and let μ_i be the index increased by one with respect to the previous nonsingular matrix in the sequence. Then the order is given by $n = \mu_i - 1$.

We may assume m=n with the knowledge that if $m < n$ the parameters b_i for i= 2m-n, ... , n will be found to be equal to zero.

2.4.2 In the presence of noise

It will be assumed here that the input and output sequences are corrupted by an additive uncorrelated noise with zero-mean, the noisy components of the input/output vectors will be denoted with

$$u_i^*(k) = u(k) + \varepsilon_u(k) \tag{2.42}$$

$$y_l^*(k) = y(k) + \varepsilon_y(k) \tag{2.43}$$

If the statistics of the noise are known, the origin of its effect can, however, be easily detected and eliminated. Because of the independence of the errors and ergodicity of all signals, the noisy covariance matrix can be decomposed, as the amount of data becomes large, into parts due to the signal and the noise.

$$\lim_{N \to \infty} \frac{1}{N} S^* = \lim_{N \to \infty} \frac{1}{N} S + R(\varepsilon) \tag{2.44}$$

Where R(ε) is the covariance matrix of the noise vector. When N is large enough a consistent estimation of S is therefore

$$S = S^* - N \times R(\varepsilon) \tag{2.45}$$

The identification of the system structure will be performed on the sequence of matrices $S(\mu_1, \mu_2)$.

When the same amount of uncorrelated zero-mean noise is added to the input-output sequences and nonoverlapping sets of data are used, then $R(\varepsilon) = \sigma^2 I$ where σ^2 the variance of the noise so that

$$\lim_{N \to \infty} \frac{1}{N} S^* = \lim_{N \to \infty} \frac{1}{N} S + \sigma^2 I \tag{2.46}$$

It is also important to note that, an estimate of σ^2 is given by the least eigenvalue of the symmetrical matrix $S(\mu, \mu)$ for $\mu > n$ so that a consistent estimate of the structure can be obtained. The same technique can be applied when different amounts of nois are present of the various inputs and outputs by performing a previous scaling on the data; the ratio of the different noises must however be known.

2.5 Illustrative example

To illustrate the identification approach developed earlier, an example has been considered. The proposed algorithm was applied to the identification of a simulated system. The order of the system was assumed a priori, thus, our only problem is the determination of the parameters from the input-output data.

Consider the following second order continous-time transfer function

$$G(s) = \frac{K}{s(s + p)} \tag{E.1}$$

Where

$$K = 6 \qquad and \qquad p = 3$$

A discrete-time signal { u(kT), k=0,1,...} as an input to excite the system

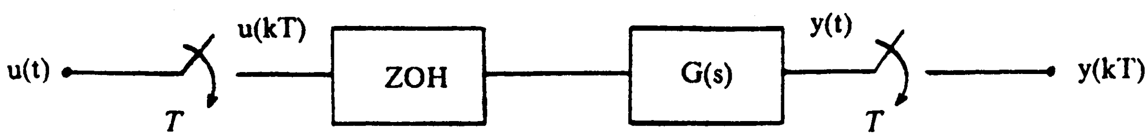

Figure 3. Sampled-data system for the example

The equivalent discrete-time transfer function of the clamped-input, sampled-ouput plan is given by the following expression

$$G(z) = \frac{b_1 z + b_0}{z^2 - a_1 z - a_0} \tag{E.2}$$

Where

$$a_0 = -e^{-pT} = -0.548812$$

$$a_1 = 1 + e^{-pT} = 1.548812$$

$$b_0 = \frac{K}{p^2}(1 - pTe^{-pT} - e^{-pT}) = 0.081268$$

$$b_1 = \frac{K}{p^2}(-1 + pT + e^{-pT}) = 0.099208$$

Using the canonical representation developed in equations (2.8) and (2.9) the system described by equation (E.2) can be expressed in the following state-space representation .

$$X(k+1) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} X(k) + \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} y(k) + \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} u(k) \qquad (E.3)$$

$$y(k) = \begin{bmatrix} 0 & 1 \end{bmatrix} X(k) \qquad (E.4)$$

Acoording to equations (2.14) and (2.15), the system (E.3) and (E.4) can be expressed as

$$S(k+1) = F(k)S(k) \qquad (E.5)$$

$$y(k) = H(k)S(k) \qquad (E.6)$$

Where

$$S(k) = \left[x_1(k)\ x_2(k)\ a_0\ a_1\ b_0\ b_1\right]$$

$$F(k) = \begin{bmatrix} T & G(k) \\ 0 & I_4 \end{bmatrix} \qquad G(k) = \left[\, y(k)I_2 \mid u(k)I_2 \,\right] \ \text{and} \qquad T = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

$$H(k) = [0\ 1\ 0\ 0\ 0\ 0\ ] \qquad I_i = i \times i \quad identity \quad matrix$$

The proposed algorithm was used to estimate the parameters of the system from the record of input-output data. The estimation method was performed on simulated data generated by the system (E.2). A unit step signal to which was added a zero-mean, white gaussian noise was used as an input to excite the system. We know from the linearity of the system that the signal is in part the response due to the noise-free input and in part the output noise.

From the simulated noise contaminated input-output data the following algorithm is used for the identification of the parameters of the system

$$M(k+1) = F(k)P(k)F^T(k) + Q$$

$$K(k+1) = M(k)H^T(k)\left[H(k+1)M(k+1)H^T(k) + R\right]^{-1}$$

$$P(k+1) = [I - K(k+1)H(k+1)]M(k+1)$$

And

$$\hat{S}(k+1) = F(k)\hat{S}(k) + K(k+1)\left[y(k+1) - H(k+1)F(k)\hat{S}(k)\right]$$

The initial state is assumed to be null and perfectly known, the input and output noises are independent normalized gaussian white noises, with known covariance matrices so,

$$S(0) = \underline{0} \quad and \quad P(0) = 100I$$

$$Q = diag(.001,.001,0.000,0.000,0.000,0.000)$$

$$R = 0.001$$

The computational results of the estimation algorithm are summarized in the figures shown at the end of the chapter. Only the estimated parameters of the characteristic polynomial are shown. Two different noise levels ($\sigma^2 = 0.10$, $\sigma^2 = 0.30$) were used in the experiment.

2.6 Discussion

The identification approach developed was applied to the identification of a simulated system and it was found that the method is easy to use and provides a valuable technique for parameter estimation.

The computed estimates are fairly good, and reasonable estimates are obtained after only 7 iterations as expected, which shows that this method converges to the true value very fast indeed. Hence, it requires less computing time than other methods. The great disadvantage of the method is the assumption that all random variables involved are gaussian. Also, the method is mostly suitable when state estimation is required.

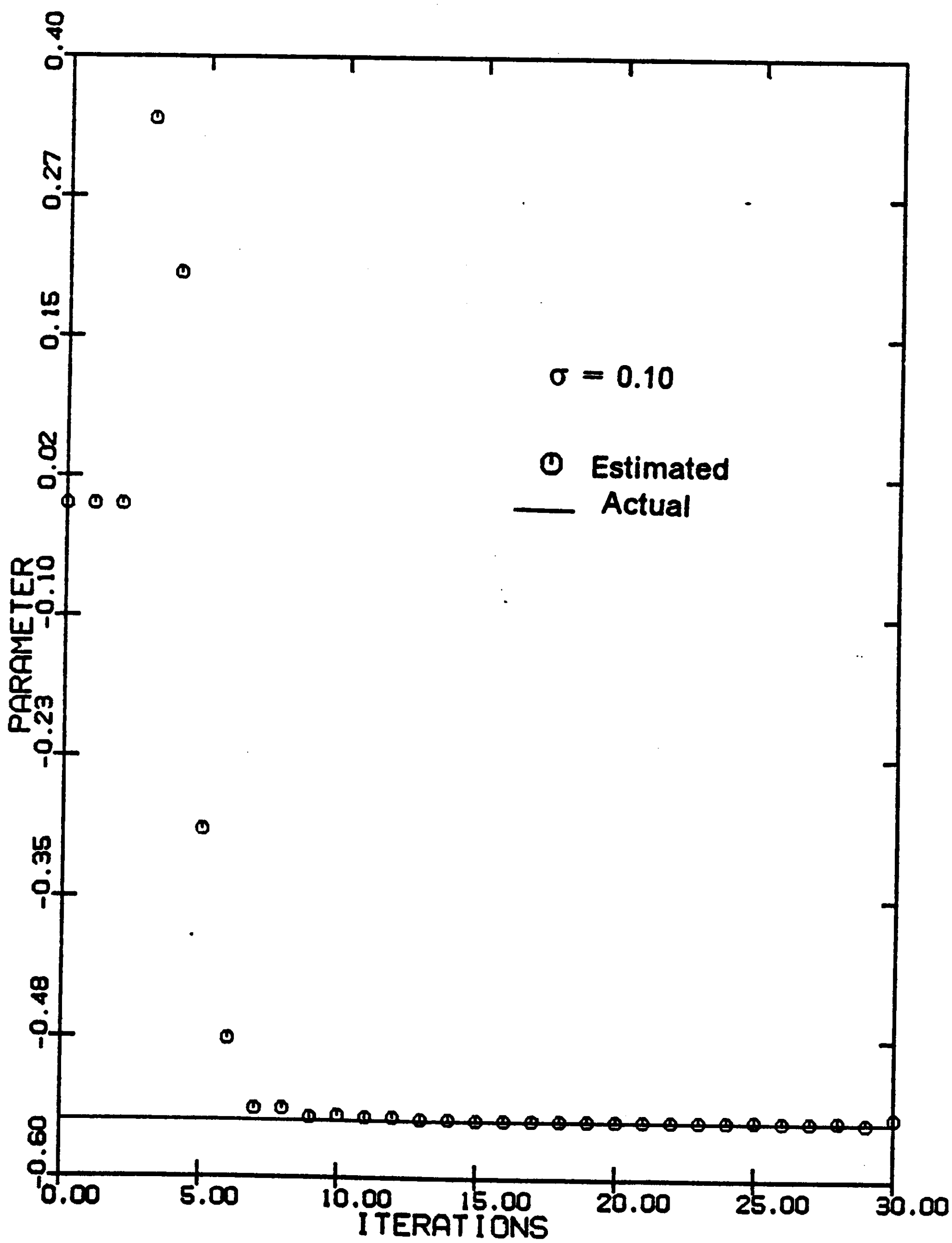

Figure 4. Estimated and actual values, in the case of low order noise level, of the parameter a(0), for the SISO case.

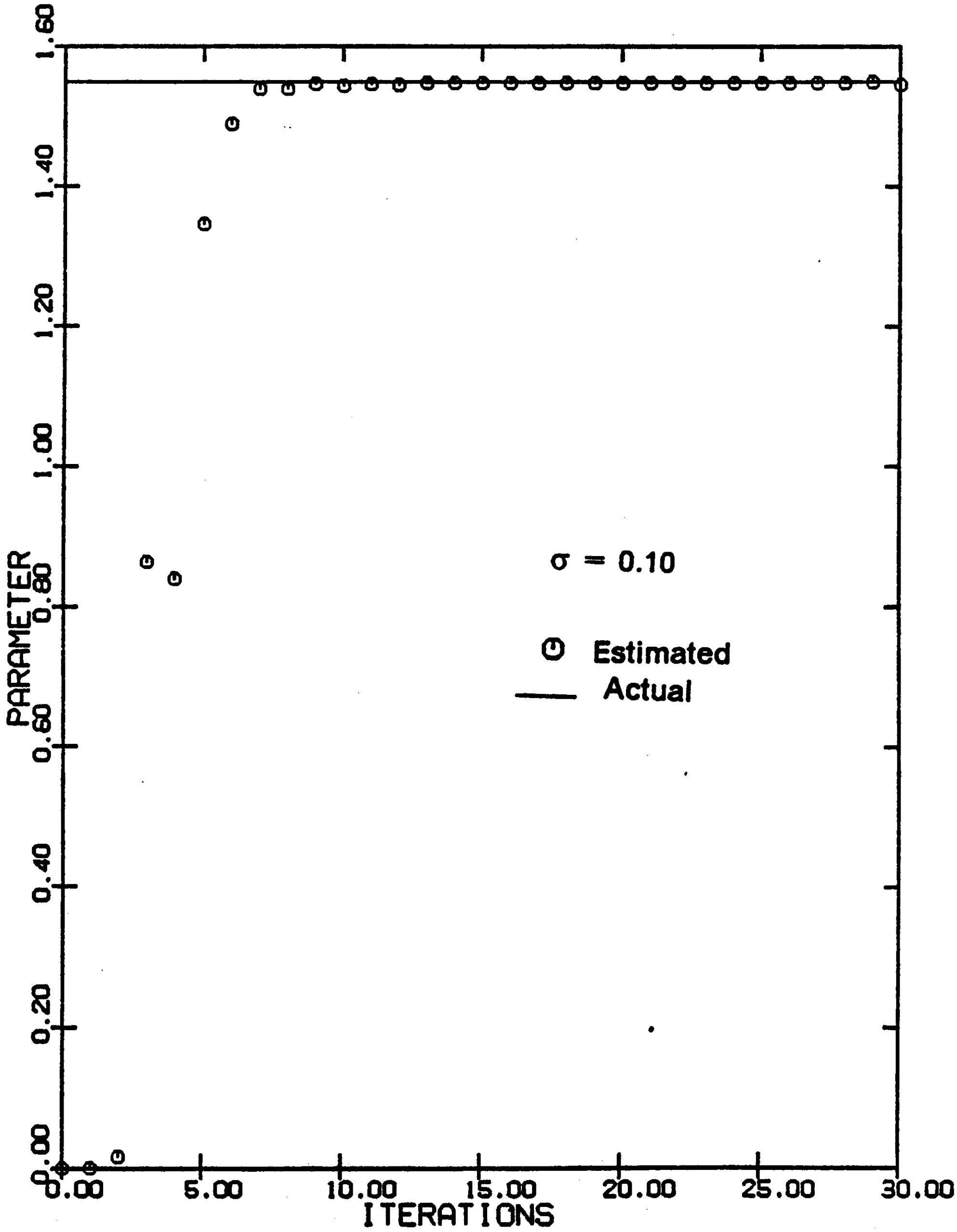

Figure 5. Estimated and actual values, in the case of low order noise level, of the parameter a(1), for the SISO case.

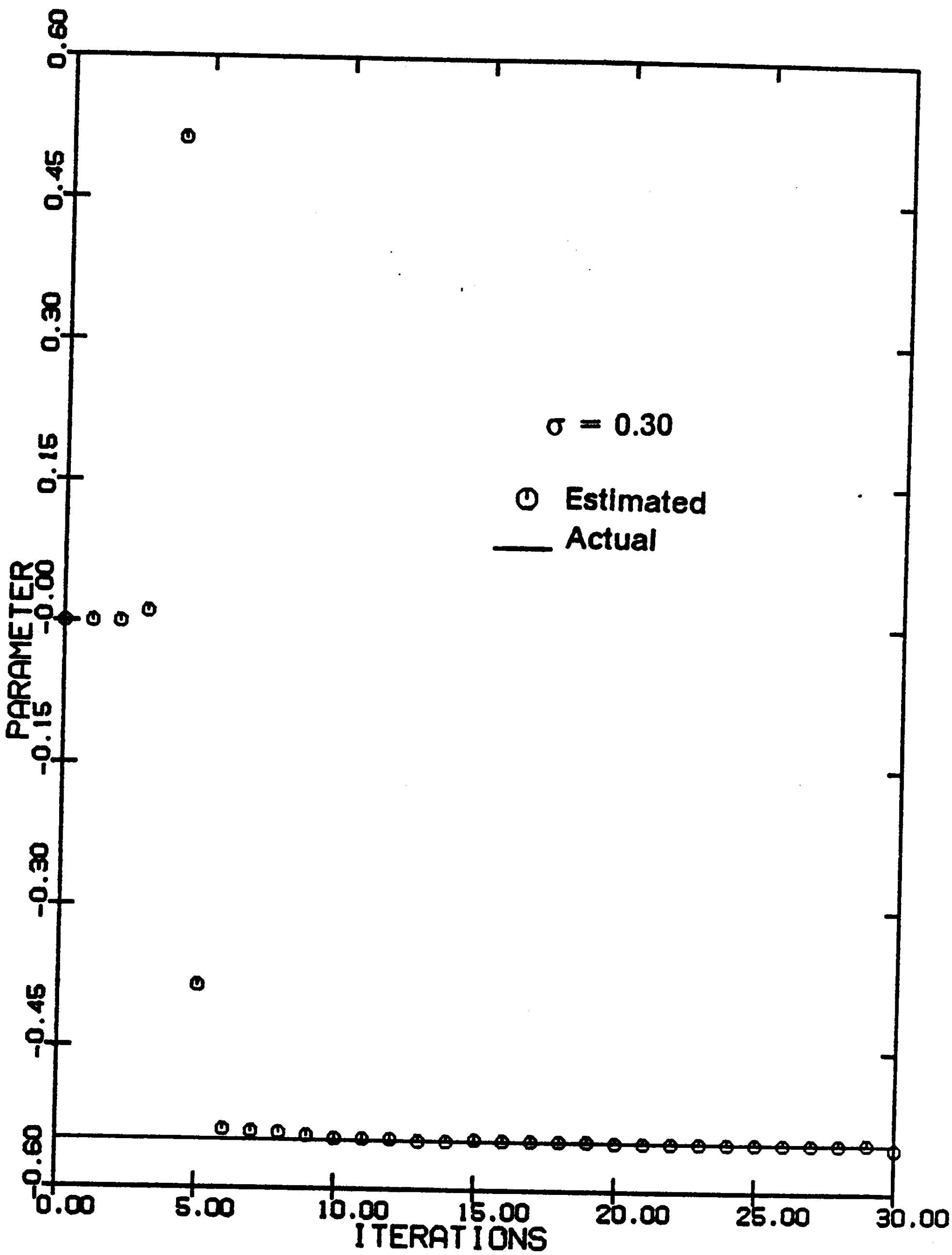

Figure 6. Estimated and actual values, in the case of high order noise level, of the parameter a(0), for the SISO case.

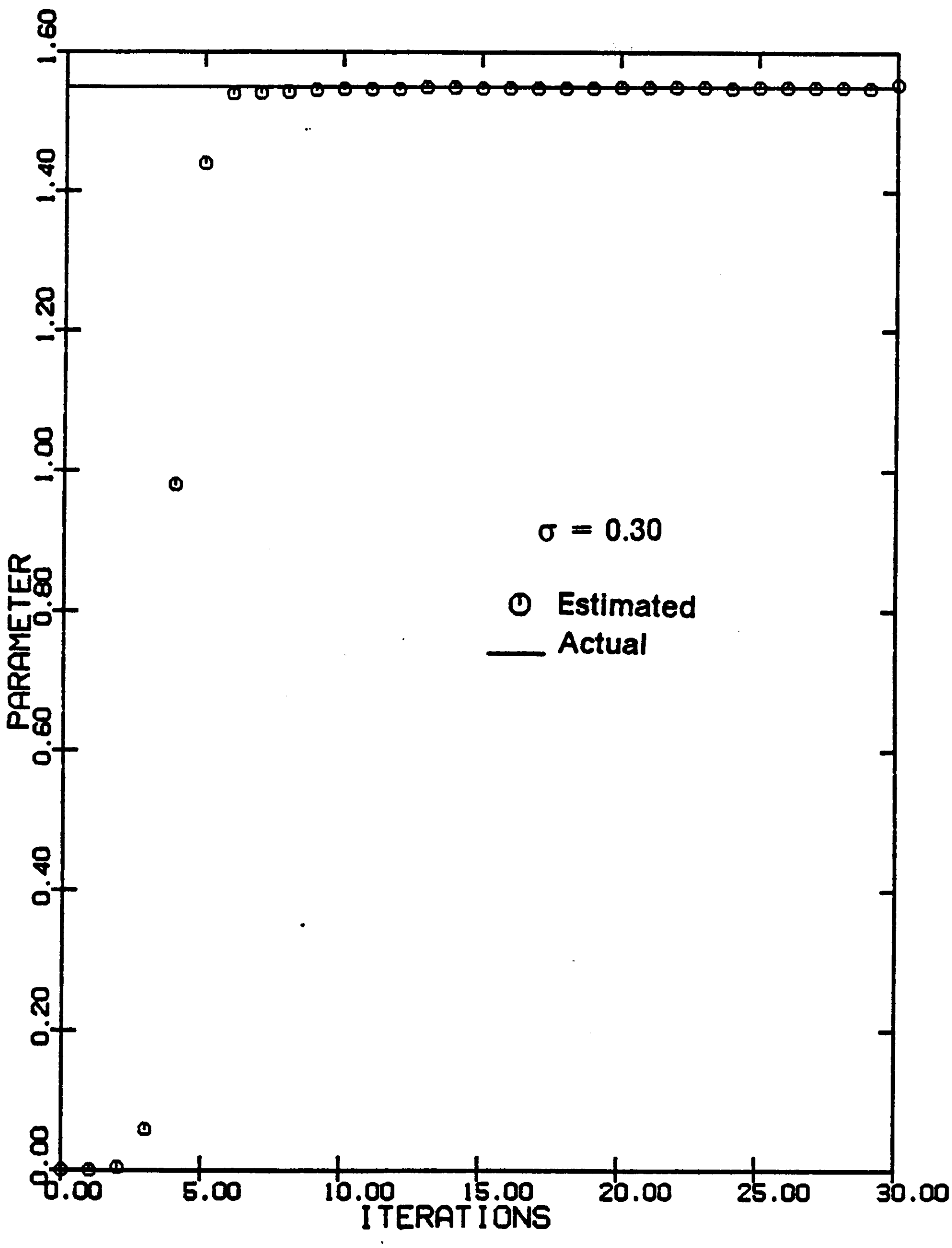

Figure 7. **Estimated and actual values, in the case of high order noise level, of the parameter a(1), for the SISO case.**

MISO AND SIMO SYSTEMS IDENTIFICATION

3.0 Introduction

The problem of identification of multivariable systems from input/output data has received much attention during the last decade. One of the important factors involved in this problem is the choice of model representation to be used for identification. Each model chosen affects the complexity and the biasedness of the identification algorithm. Numerous papers have been published on the problem, and many of these papers have used the state-space description for identification. This description has the advantage, as it will be shown in chapter 3, that certain canonical forms can be used to avoid overparametrization in the model. On the other hand it has the disavantage that in order to obtain canonical forms for the state-space model one must first determine the structural indices of the system, which may be quite involved. An alternative approach to the problem, which avoids this difficulty is to estimate the parameters of the system in the transfer-function matrix repersentation. In this case it is only required to know the order of the system. Many algorithms were developed to compute the order of a multivariable linear discrete-time system. Starting from a well defined transfer-function matrix with unknown parameters, we will attempt to combine state and parameter estimation for the considered model using the canonical representation introduced in chapter 1, which will be extended to the cases of multiple-input single-output (MISO) and single-input multiple-output (SIMO) linear systems. A discussion of the multiple-input multiple-output case is also given.

3.1 MIMO systems

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

$$Y(k) = G(z)U(k) \tag{3.1}$$

Where Y(k) is the m-dimensional output vector sequence, U(k) is the p-dimensional input vector sequence and z the unit advance operator. The structural parameters required to characterize the transfer-function matrix of the system are the orders of the numerator and the denominator of each of the elements of G(z).

The transfer-function matrix G(z) can be written as

$$G(z) = \begin{bmatrix} g_{11}(z) & \ldots\ldots & g_{1p}(z) \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ g_{m1}(z) & \ldots\ldots & g_{?}(z) \end{bmatrix} \tag{3.2}$$

Where each element, $g_{ij}(z)$, i=1,..,m ; j=1,..,p is a rational function of z. The transfer-function matrix, G(z), can be expressed as

$$G(z) = \frac{1}{D(z)} \begin{bmatrix} g^*_{11}(z) & \ldots & g^*_{1p}(z) \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ g^*_{m1}(z) & \ldots & g^*_{mp}(z) \end{bmatrix} \tag{3.3}$$

Where D(z) is the characteristic polynomial of the system, defined as the least common monic denominator of all minors of G(z). The entries $g^*_{ij}(z)$ are polynomials in z. The structural parameters required to characterize G(z) in the form (3.3) are the order of the polynomial D(z), the order of the system, and the order of the polynomials $g^*_{ij}(z)$.

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

$$D(z) = z^n - a_{n-1}z^{n-1} - \ldots - a_1 z - a_0 \tag{3.4}$$

And

$$g^*_{ij}(z) = b^n_{ij}z^n + b^n_{ij}z^n + \ldots + b^1_{ij}z^1 + b^0_{ij} \tag{3.5}$$

[If the true order of any of the $g^*_{ij}(z)$ is less than n the corresponding coefficients in equation (3.5) will be found when estimated as zero].

From the definition of G(z) given in equation (3.3) the outputs can be written as

$$y_i(z) = \frac{1}{D(z)} \sum_{j=1}^{p} g^*_{ij}(z) u_j(z) \tag{3.6}$$

The given system can be decomposed into m subsystems (m = number of outputs) as shown in equation (3.6). Each of these subsystems corresponds to one row of the matrix G(z) and can be regarded as a multiple-input single-output (MISO) system. Hence the parameters of each subsystem can be estimated independently and the identification of the whole system is accomplished in m separate steps. 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 computation in the proposed algorithm, it is possible to avoid estimating the parameters of D(z) more than one time. This can be accomplished by taking them as constants after the identification of the first subsystem, or by introducing in parallel state space representation for each element $g_{ij}(z)$ of G(z) with D(z) as the common denominator.

3.2 Multiple-Input Single-Output Systems

Consider a two-input single-output linear discrete-time system given by its ARMA model transfer-function matrix. This system can be represented as follows

$$y(z) = \frac{b_{11}^{n} z^{n} + \dots + b_{11}^{1} z + b_{11}^{0}}{z^{n} - a_{n-1} z^{n-1} - \dots - a_{1} z - a_{0}} u_{1}(z) + \frac{b_{12}^{n} z^{n} + \dots + b_{12}^{1} z + b_{12}^{0}}{z^{n} - a_{n-1} z^{n-1} - \dots - a_{1} z - a_{0}} u_{2}(z) \quad (3.7)$$

Our problem can be stated as the determination of the parameter a_i ; i = 1,n-1 and b_{1j}^{k} ; j = 1,2 and k = 1,n based on the measured input/output data sequence.

Using a state variable representation, this systen can be written in the following observable canonical form

$$X(k+1) = \Phi X(k) + \Gamma U(k) \quad (3.8)$$

$$y(k) = CX(k) + DU(k) \tag{3.9}$$

where

$$U(k) = \begin{bmatrix} u_1(k) \\ u_2(k) \end{bmatrix}$$

$$\Phi(n \times n) = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & a_0 \\ 1 & 0 & 0 & \dots & 0 & a_1 \\ 0 & 1 & 0 & \dots & 0 & a_2 \\ : & : & : & \dots & : & : \\ : & : & : & \dots & : & : \\ 0 & 0 & 0 & \dots & 0 & a_{n-2} \\ 0 & 0 & 0 & \dots & 1 & a_{n-1} \end{bmatrix} \qquad \Gamma(n \times 1) = \begin{bmatrix} b_{11}^0 + a_0 b_{11}^n & b_{12}^0 + a_0 b_{12}^n \\ b_{11}^1 + a_1 b_{11}^n & b_{12}^1 + a_1 b_{12}^n \\ : \quad : & : \quad : \\ : \quad : & : \quad : \\ b_{11}^{n-1} + a_{n-1} b_{11}^n & b_{12}^{n-1} + a_{n-1} b_{12}^n \end{bmatrix}$$

$$C(1 \times n) = [0\ 0\ 0\ \dots\ 0\ 1] \qquad D(2 \times 1) = \begin{bmatrix} b_{11}^n & b_{12}^n \end{bmatrix}$$

The number of parameters to be estimated is N=3n+2.

Let the parameters vectors be

$$A^T = [a_0\ a_1\ a_2\ \dots\ a_{n-1}]$$

and

$$B_{1jp}^T = \begin{bmatrix} b_{1j}^1\ b_{1j}^0\ b_{1j}^2\ \dots\ b_{1j}^{n-1} \end{bmatrix} \quad \textit{for } j = 1,2$$

It is clear that the matrices Φ and Γ can be expressed as follows:

$$\Phi = T + A \times C \tag{3.10}$$

$$\Gamma = \begin{bmatrix} B_{11p}\ B_{12p} \end{bmatrix} + A \times D \tag{3.11}$$

where

$$T(n \times n) = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ : & : & : & \dots & : & : \\ : & : & : & \dots & : & : \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix}$$

subtituting the expressions of Φ and Γ into (3.8), it can be shown that

$$X(k + 1) = (T + A \times C)X(k) + ([B_{11p}\ B_{12p}] + A \times D)U(k)$$

ie

$$X(k + 1) = TX(k) + A(CX(k) + DU(k)) + [B_{11p}\ B_{12p}]\ U(k)$$

using equation (3.9) this leads to

$$X(k + 1) = TX(k) + Ay(k) + [B_{11p}\ B_{12p}]\ U(k)$$

Therefore equations (3.8) and (3.9) may be expressed as:

$$X(k+1) = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ : & : & : & \dots & : & : \\ : & : & : & \dots & : & : \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix} X(k) + \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ : \\ : \\ a_{n-2} \\ a_{n-1} \end{bmatrix} y(k) + \begin{bmatrix} b_{11}^0 & b_{12}^0 \\ b_{11}^0 & b_{12}^0 \\ b_{11}^2 & b_{12}^2 \\ : & : \\ : & : \\ b_{11}^{n-2} & b_{12}^{n-2} \\ b_{11}^{n-2} & b_{12}^{n-2} \end{bmatrix} U(k) \qquad (3.12)$$

$$y(k) = [0\ 0\ 0\ \dots\ 0\ 1]\ X(k) + \left[b_{11}^n\ b_{12}^n\right] U(k) \qquad (3.13)$$

let the total B parameters vector

$$B_{1j}^T = \left[B_{1jp}^T | b_{1j}^n\right]$$

Now let us introduce the total parameter vector

$$P^T(k) = \left[\ A^T \mid B_{11}^T \mid B_{12}^T\right] \quad (3n+2) \times 1 \quad \textit{matrix}$$

It can be shown that equations (3.12) and (3.13) may also be written as

$$X(k+1) = TX(k) + G(k)P(k) \qquad (3.14)$$

$$P(k+1) = I_{3n+2}P(k) \qquad (3.15)$$

$$y(k) = CX(k) + N(k)P(k) \qquad (3.16)$$

where

$G(k) = \left[\ y(k)I_n \mid M_1(k) \mid M_2(k)\ \right]$ $\quad n \times (3n+2)$ *matrix* $\quad M_i = \left[\ u_i(k)I_n \mid 0\ \right]$ *for* $i = 1,2$

$I_i = i \times i$ *identity matrix*

$$N(k) = \left[\, n_1(k) \mid n_2(k) \,\right] \quad where \quad n_i(k) = \left[\, 0\,0 \ldots\ldots 0\, u_i(k) \,\right] \quad for \quad i = 1,2$$

Using the extended vector that involves augmenting the state variables of the system by adjoining to them the unknown parameter vectors and treating them as part of the new state variable vector.

$$S^T(k) = \left[X^T(k) \mid A^T \mid B^T\right] = \left[X^T(k) \mid P^T(k)\right] \tag{3.17}$$

The equations (3.14), (3.15) and (3.16) will have the following form for the augmented state vector

$$S(k+1) = F(k)S(k) \tag{3.18}$$

$$y(k) = H(k)S(k) \tag{3.19}$$

Where

$$F(k) = \begin{bmatrix} T & G(k) \\ 0 & I_{3n+2} \end{bmatrix} \quad (4n+2) \times (4n+2) \quad matrix$$

and

$$H(k) = \left[C \mid N(k)\right] \quad 1 \times (4n+2) \ \ row\ matrix$$

The representation is considered to be pseudo-linear, linear and time-varying.

3.2.1 Generalization to the case of r-input and one-output

Now, we consider the case when it is question of a subsystem of a discrete-time multivariable system.

In general the system will be represented by

$$y(z) = \sum_{j=1}^{r} \frac{b_{1j}^{n} z^{n} + b_{1j}^{n-1} z^{n-1} + \dots + \dots + b_{1j}^{1} z + b_{1j}^{0}}{z^{n} - a_{n-1} z^{n-1} - \dots - \dots - a_1 z - a_0} u_j(k) \tag{3.20}$$

Let the parameters vectors be

$$A^T = [a_0 \; a_1 \; a_2 \dots \; a_{n-1}]$$

$$B_{1jp}^T = \left[b_{1j}^0 \; b_{1j}^1 \; b_{1j}^2 \; \dots \; b_{1j}^{n-1}\right] \quad \textit{for } j = 1,\dots,r$$

$$B_p^T = \left[B_{11p}^T \mid B_{12p}^T \mid \dots \; \dots \; B_{1rp}\right]$$

The number of parameters to be estimated is N = r(n + 1) + n

The system can be expressed as follows

$$X(k+1) = TX(k) + Ay(k) + B_p U(k) \tag{3.21}$$

$$y(k) = CX(k) + B^n U(k) \tag{3.22}$$

$$B^n = \left[b_{11}^n \, b_{12}^n \dots b_{1r}^n\right] \quad U^T(k) = \left[u_1(k) \; u_2(k) \; \dots \; u_r(k)\right]$$

$$B_{1j}^T = \left[B_{1jp}^T \mid b_{1j}^n\right]$$

Let us introduce the total parameter vector

$$P^T(k) = \left[\, A^T \mid B_{11}^T \mid B_{12}^T \mid \dots \mid B_{1r}^T \,\right]$$

It can be shown that equations (3.21) and (3.22) may also be written as

$$X(k+1) = TX(k) + G(k)P(k) \tag{3.23}$$

$$P(k+1) = IP(k) \tag{3.24}$$

$$y(k) = CX(k) + N(k)P(k) \tag{3.25}$$

Where

$$G(k) = \left[\, y(k) \times I_n \mid M_1(k) \mid M_2(k) \;\dots\; M_r(k) \right] \quad n \times (rn + n + r) \;\; matrix$$

$$M_j = \left[\, u_j(k) I_n \mid 0 \,\right] \;\; for \; j = 1,\dots,r$$

$$I_i = i \times i \;\; identity \;\; matrix$$

$$N(k) = \left[\, n_1(k) \mid n_2(k) \;\dots\; n_r(k) \right] \; where \quad n_j(k) = \left[\, 0\, 0 \dots 0\, u_j(k) \,\right]$$

using the extended vector that involves augmenting the state variables of the system by adjoining to them the unknown parameter vectors and treating them as part of the new state variable vector.

$$S^T(k) = \left[X^T(k) \mid A^T \mid B^T\right] = \left[X^T(k) \mid P^T(k)\right] \tag{3.26}$$

The equations (2.8) , (2.9) and (2.10) will have the following pseudo-linear form for the augmented state vector

$$S(k+1) = F(k)S(k) \tag{3.27}$$

$$y(k) = H(k)S(k) \tag{3.28}$$

Where

$$F(k) = \begin{bmatrix} T & G(k) \\ 0 & I \end{bmatrix} \quad ((n+1)r + 2n) \times ((n+1)r + 2n) \qquad matrix$$

and

$$H(k) = \left[C \mid N(k)\right] \quad 1 \times ((n+1)r + 2n) \ \textit{row matrix}$$

3.2.2 Identification algorithm

At this point, we have considered only the deterministic part of the system and assumed that the input-output data is free of noise. Since most practical systems have considerable measurement noise, a realistic algorithm should assume that the available data be corrupted by noise. In such practical situations one may model the system as

$$S(k+1) = F(k)S(k) + V(k) \tag{3.29}$$

$$y(k) = H(k)S(k) + w(k) \tag{3.30}$$

Where:

S(k) is the extended state vector of the system

y(k), the measured output signal

V(k) and w(k) are white gaussian sequences with correlation functions

$$E\left\{\begin{bmatrix} V(k) \\ w(k) \end{bmatrix}\begin{bmatrix} V(l) \\ w(l) \end{bmatrix}^T\right\} = \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \delta_{kl} \tag{3.31}$$

More precisely, V(k) and w(k) are independent of past and present states of the system Equations (3.29) and (3.30) are derived from stochastic difference equations, which include the processes generating the noises and perturbations

The matices F(k) and H(k) are available at each time k from the record of the input-output data. We assume also that the matrices Q and R are known, as well as the statistical characteristics of the initial state S(0), that is to say its mean value and covariance.

At time k, we assume the measurements vector

$$\mathbf{Y}^T(k) = [\ Y(0)\ Y(1)\ \dots\ Y(k)\ ] \tag{3.32}$$

and seek to estimate the unknown vector

$$\mathbf{S}^T(k) = \left[\ S^T(0)\ S^T(1)\ \dots\ S^T(k)\ \right] \tag{3.33}$$

The assumptions of the problem allow us in theory to calculate the mean values, variances and covariances of S(k) and Y(k), hence, the possibility to apply the general formulas of the optimal linear estimation to obtain S(k) as a linear function of Y(k)

The estimation problem can be stated as the choice of the optimal gain K(k) such that the reconstruction error (3.34) is minimal.

$$\hat{S}(k+1) = F(k)\hat{S}(k) + K(k)\left[\ y(k) - H(k)\hat{S}(k)\ \right] \tag{3.34}$$

The Kalman filtering algorithm can be applied to provide a technique for the identification of the parameters and estimation of the states.

From a standard discrete Kalman filter an estimation algorithm is developed. let

S(k/k) = optimal filtered estimate of S(k)

S(k + 1/k) = optimal predicted estimate of S(k)

If the optimal filtered estimate $\hat{S}(k/k)$ and the covariance matrix P(k/k) of the corresponding filtering error $\hat{S}(k/k)$ are known for some k, then the single-stage optimal predicted estimate for all admissible loss functions is given by the expression

$$\hat{S}(k+1/k) = F(k)\hat{S}(k/k) \tag{3.35}$$

The stochastic process $\{\hat{S}(k \quad 1/k),\ k=0,1,..\}$ defined by the single-stage predicted error relation

$$\bar{S}(k+1/k) = S(k+1) - \hat{S}(k+1/k)$$

is a zero mean gauss-markov process whose covariance matrix is given by the relation

$$P(k+1/k) = F(k)P(k/k)F^T(k) + Q \tag{3.36}$$

The optimal filtered estimate S(k+1/k+1) is given by the recursive relation

$$\hat{S}(k+1/k+1) = F(k)S(k/k) + K(k+1)\left[y(k+1)\_H(k+1)F(k)\hat{S}(k/k)\right] \tag{3.37}$$

for k=0,1,.... where $\hat{S}(0/0) = S(0)$

K(k+1) is the filter gain specified by the set of relations

$$K(k+1) = P(k+1/k)H^T(k+1)\left[H(k+1)P(k+1/k)H^T(k+1) + R\right]^{-1} \tag{3.38}$$

$$P(k+1/k) = F(k)P(k/k)F^T(k) + Q \tag{3.39}$$

$$P(k+1/k+1) = [I - K(k+1)H(k+1)]P(k+1/k) \tag{3.40}$$

For k=0,1,.... and P(0/0)=P(0)

3.2.3 Illustrative example

An example of identification is presented to illustrate the proposed approach. A strictly proper pulse transfer function matrix describes the two-input one-output system.

Consider the 1 × 2 pulse transfer function matrix

$$H(z) = \frac{1}{D(z)} [\ H_{11}(z)\ \ H_{12}(z)\ ] \tag{E.1}$$

For the given example it was assumed

$$D(z) = z^3 - a_2 z^2 - a_1 z - a_0 \tag{E.2}$$

$$H_{11}(z) = b_{11}^2 z^2 + b_{11}^1 z + b_{11}^0 \tag{E.3}$$

$$H_{12}(z) = b_{12}^2 z^2 + b_{12}^1 z + b_{12}^0 \tag{E.4}$$

With the following values of the parameters

$$a_0 = 0.375 \qquad a_1 = 0.625 \qquad a_2 = -0.75$$

$$b_{11}^0 = 3.375 \qquad b_{11}^1 = -4.0 \qquad b_{11}^2 = 1.0$$

$$b_{12}^0 = -1.125 \qquad b_{12}^1 = 3.0 \qquad b_{12}^2 = -2.0$$

[The fact that the discrete-time transfer function matrix is strictly proper leads to $b_{11}^3 = 0$ and $b_{12}^3 = 0$, which reduces the number of parameters to estimate]

Using a state variable representation (3.13), this system can be written in the following observable canonical form

$$X(k+1) = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} X(k) + \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} y(k) + \begin{bmatrix} b_{11}^0 & b_{12}^0 \\ b_{11}^1 & b_{12}^1 \\ b_{11}^2 & b_{12}^2 \end{bmatrix} U(k) \qquad (E.5)$$

$$y(k) = [\, 0 \;\; 0 \;\; 1 \,]\, X(k) \qquad (E.6)$$

where

$$A^T = [\, a_0 \; a_1 \; a_2 \,]$$

$$B_{i1}^T = \left[\, b_{i1}^0 \; b_{i1}^1 \; b_{i1}^2 \,\right] \quad for \quad i = 1,2$$

We define the total parameters vector as: $P^T(k) = \left[\, A^T \mid B_{11}^T \mid B_{21}^T \,\right]$

From equations (3.38) and (3.39) we may also write

$$X(k+1) = TX(k) + G(k)P(k) \qquad (E.7)$$

$$P(k+1) = I_9 P(k) \qquad (E.8)$$

$$Y(k) = CX(k) \qquad (E.9)$$

Where

$$G(k) = \left[\, y(k)I_3 \mid M_1(k) \mid M_2(k) \,\right] \quad (3 \times 9) \;\; matrix \quad M_i = \left[\, u_i(k)I_3 \mid 0 \,\right] \;\; for \; i = 1,2$$

$$T = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \qquad C = [\, 0 \;\; 0 \;\; 1 \,]$$

$$I_i = i \times i \quad identity \quad matrix$$

Using the extended vector $S^T(k) = \left[\, X^T(k) \mid P^T(k) \,\right]$, it leads to

$$S(k+1) = F(k)S(k) \tag{E.10}$$

$$Y(k+1) = H(k)S(k) \tag{E.11}$$

Where

$$F(k) = \begin{bmatrix} T & G(k) \\ 0 & I_9 \end{bmatrix} \ (12 \times 12) \quad H(k) = \begin{bmatrix} C \mid 0 \end{bmatrix} \ (1 \times 12)$$

It will be more realistic to consider the system as being subject to random perturbations and the measurements as being noisy, so, we reformulate our system as follows

$$S(k+1) = F(k)S(k) + V(k) \tag{E.12}$$

$$Y(k) = H(k)S(k) + w(k) \tag{E.13}$$

Where the noises V(k) and w(k) are white gaussian noises, assumed to be of zero-mean and uncorrelated.

A unit step corrupted by an additive zero-mean, white gaussian noise was used to excite the system. From the simulated noise contaminated input-output data, the proposed approach was used to estimate the parameters.

The following algorithm was used

$$M(k) = F(k-1)P(k-1)F^T(k) + Q \tag{E.14}$$

$$K(k) = M(k)H^T(k) \left[\ H(k)M(k)H^T(k) + R \ \right]^{-1} \tag{E.15}$$

$$K(k) = [\ I - K(k)H(k)\ ]M(k) \tag{E.16}$$

and

$$\hat{S}(k) = F(k-1)\hat{S}(k-1) + K(k)\left[\ Y(k) - H(k)F(k-1)\hat{S}(k-1)\ \right] \tag{E.17}$$

The initial state is assumed to be null, the input and output noises are independent normalized gaussian white noises, with known covariance matrices

$$S(0) = 0 \quad \textit{and} \quad P(0) = 100I$$

$$Q(12 \times 12) = \begin{bmatrix} Q_s & 0 \\ 0 & 0 \end{bmatrix} \qquad Q_s(3 \times 3) = diag(.001,.001,.001)$$

$$R = .001$$

The results of the first 30 iterations are shown in the figures at the end of the chapter. Only the estimates of the parameters of the characteristic polynomial are shown. As before, two different noise levels were used ($\sigma^2 = 0.10$, $\sigma^2 = 0.30$).

As shown in the figures, the estimates are very accurate and good; reasonable estimates are obtained after just 10 iterations, which proves again that this approach converges very fast.

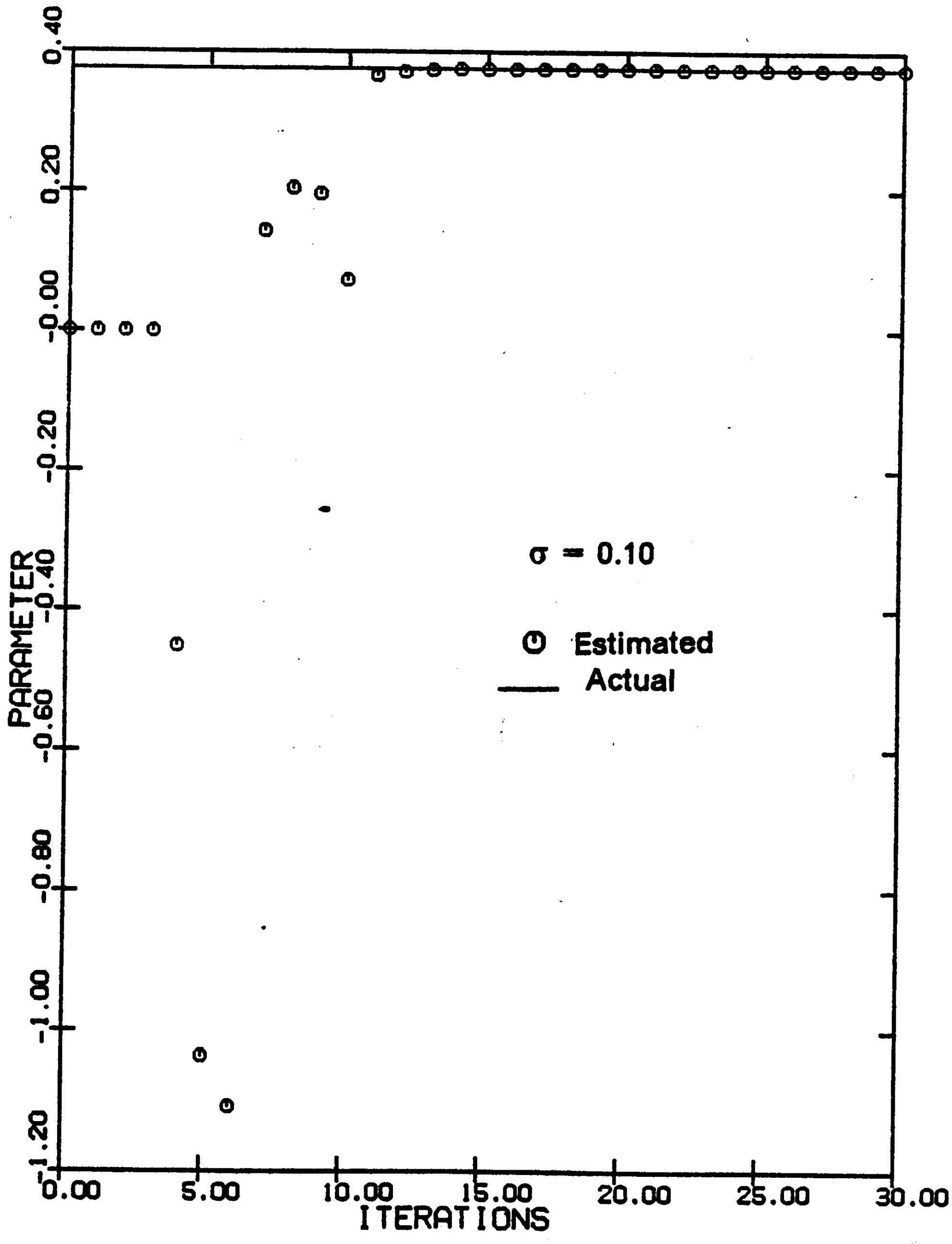

Figure 8. Estimated and actual values, in the case of low order noise level, of the parameter a(0), for the MISO case.

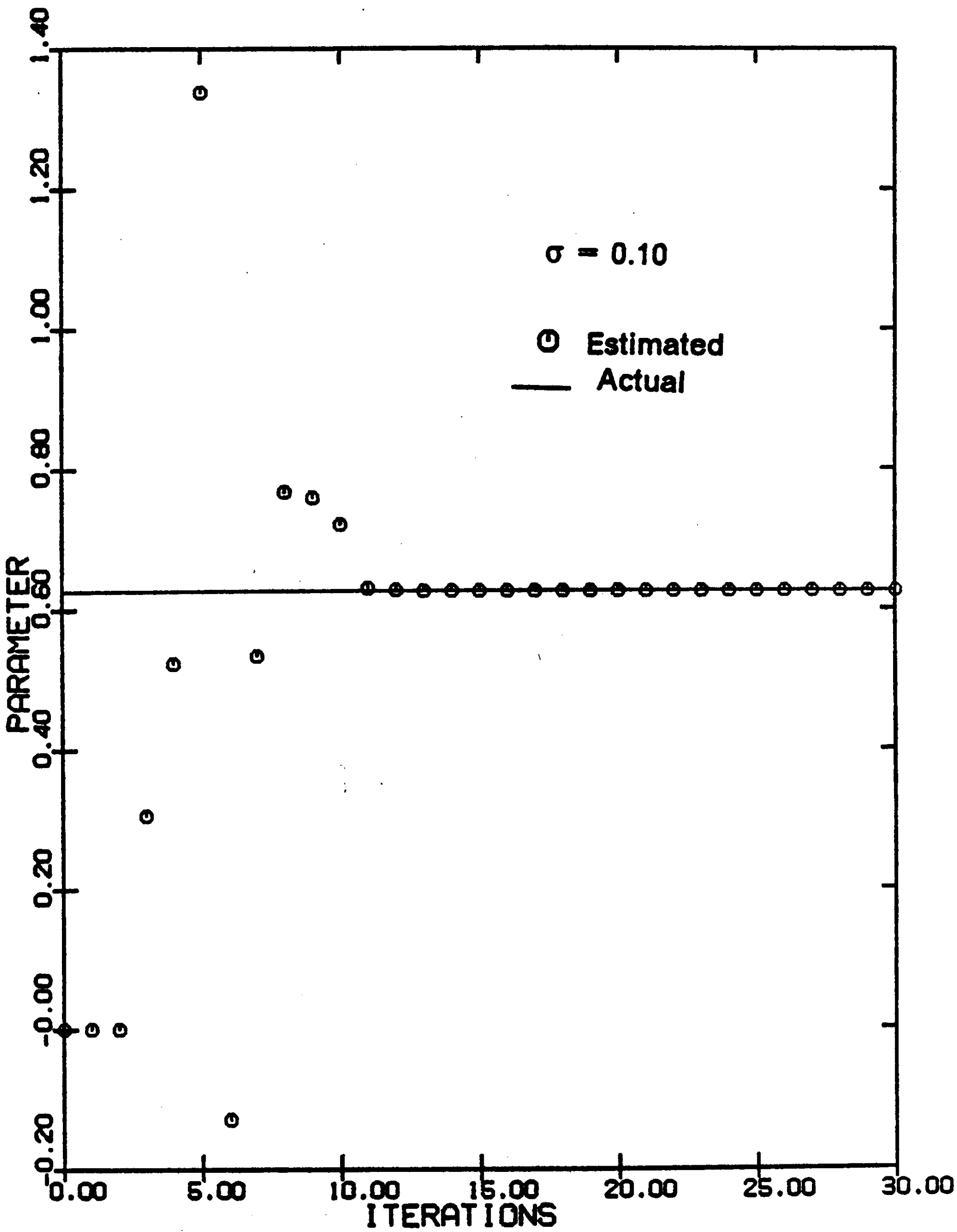

Figure 9. Estimated and actual values, in the case of low order noise level, of the parameter a(1), for the MISO case.

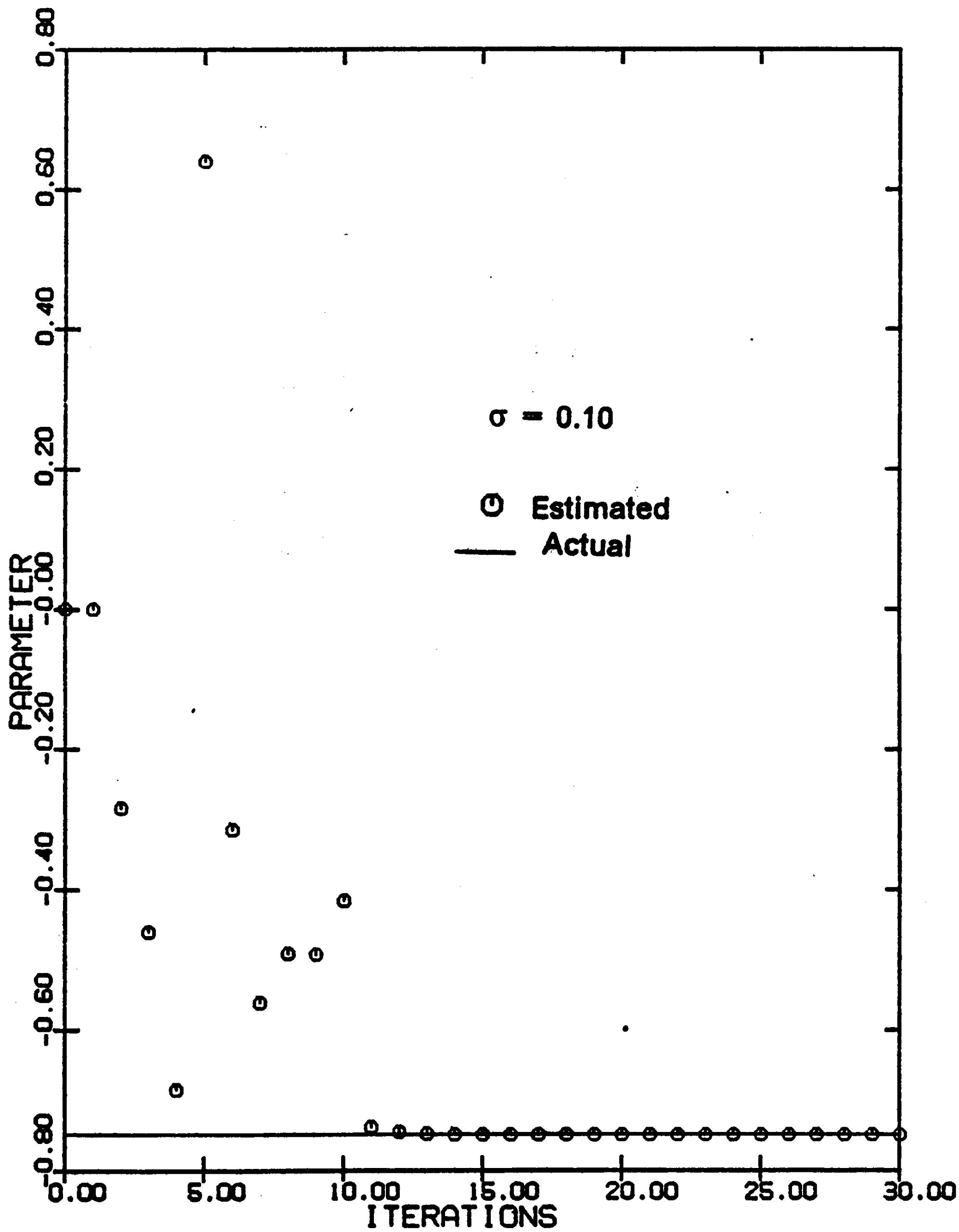

Figure 10. Estimated and actual values, in the case of low order noise level, of the parameter a(2), for the MISO case.

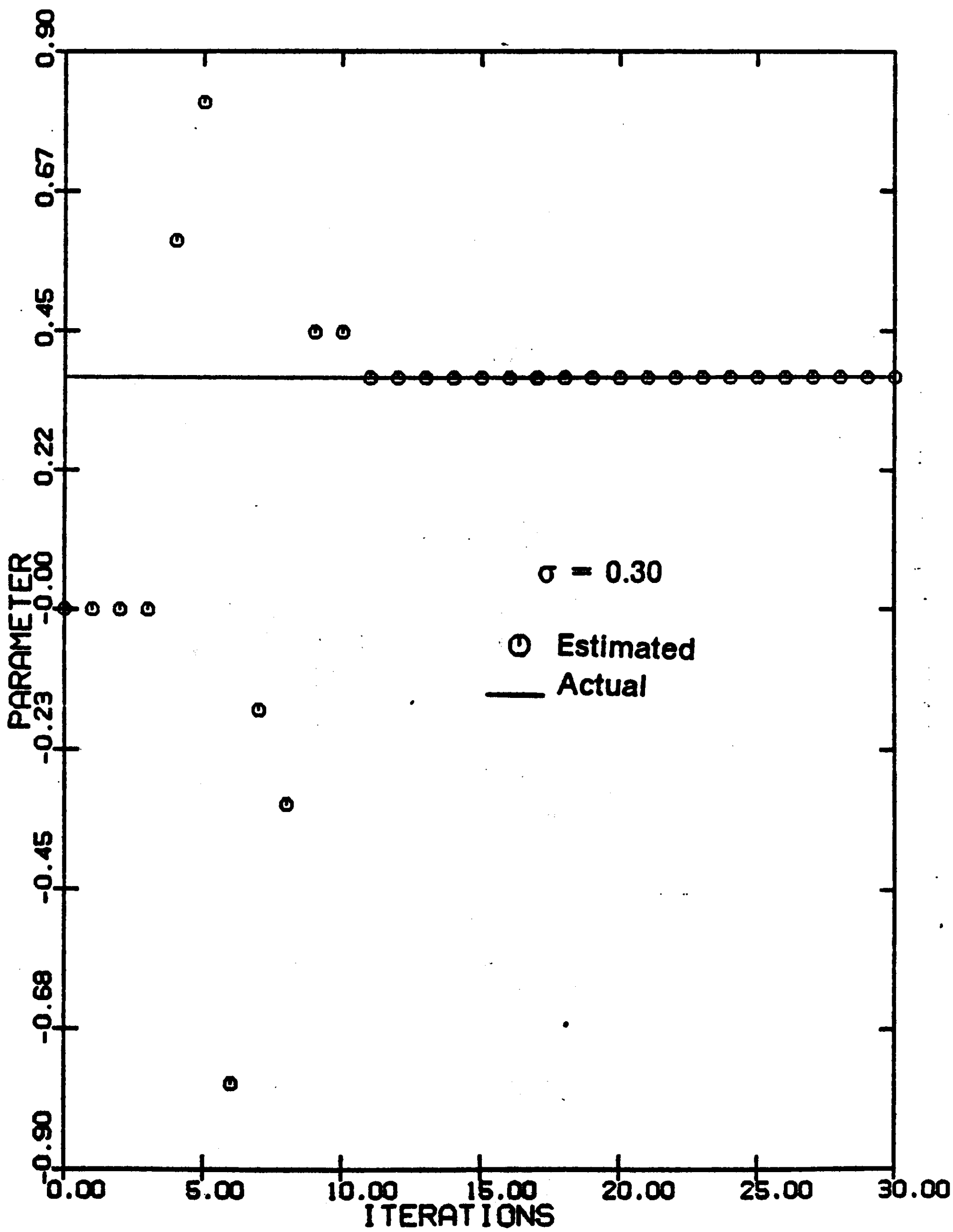

Figure 11. Estimated and actual values, in the case of high order noise level, of the parameter a(0), for the MISO case.

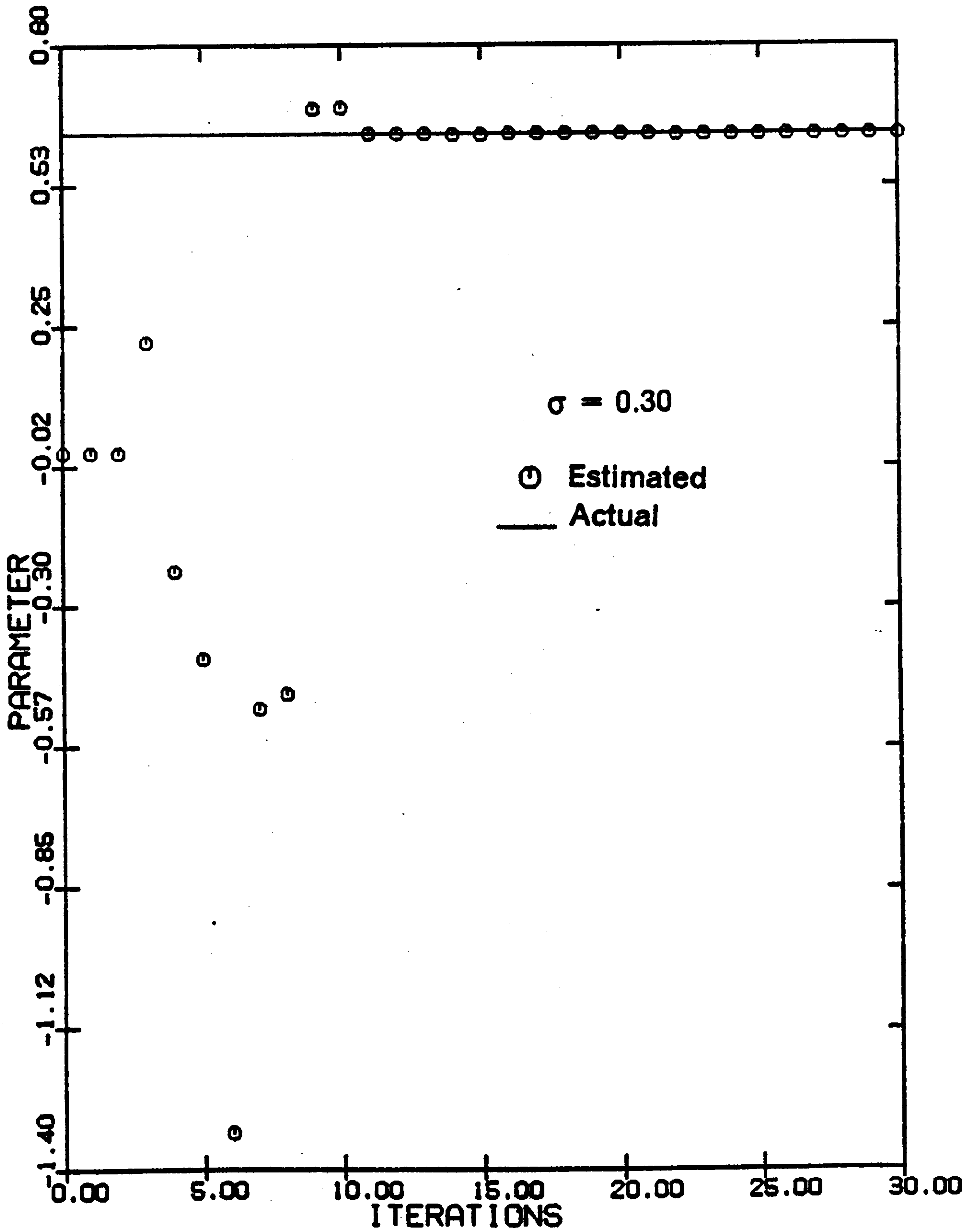

Figure 12. Estimated and actual values, in the case of high order noise level, of the parameter a(1), for the MISO case.

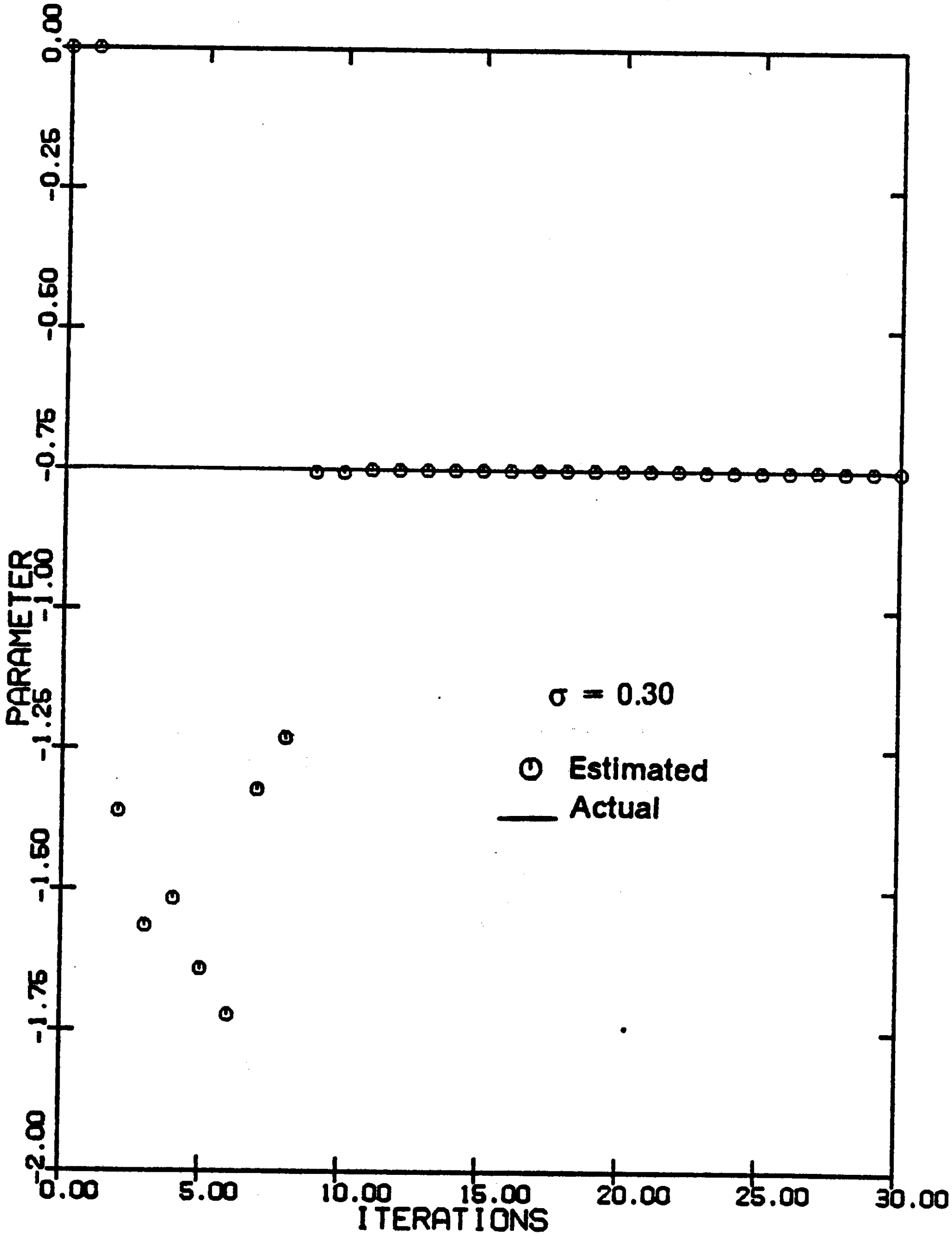

Figure 13. Estimated and actual values, in the case of high order noise level, of the parameter a(2), for the MISO case.

3.3 Single-input multiple-output systems

In considering SIMO systems, we attempt to show that there is another possibility of estimating MIMO systems by considering a state space representation of each element of G(z) in parallel with those of the other elements in the matrix transfer-function.

Consider a single-input two-output system given by

$$Y(z) = \frac{1}{z^n - a_{n-1}z^{n-1} - \dots - a_1 z - a_0} \begin{bmatrix} b_{11}^n z^n + b_{11}^{n-1} z^{n-1} + \dots + b_{11}^1 z + b_{11}^0 \\ b_{21}^n z^n + b_{21}^{n-1} z^{n-1} + \dots + b_{21}^1 z + b_{21}^0 \end{bmatrix} \quad (3.35)$$

This system can be written in the following state variable form

$$X(k+1) = \begin{bmatrix} \Phi & 0 \\ 0 & \Phi \end{bmatrix} X(k) + \begin{bmatrix} \Gamma_1 \\ \Gamma_2 \end{bmatrix} u(k) \quad (3.36)$$

$$Y(k) = \begin{bmatrix} C & 0 \\ 0 & C \end{bmatrix} X(k) + \begin{bmatrix} b_{11}^n \\ b_{21}^n \end{bmatrix} u(k) \quad (3.37)$$

Where

$$\Phi(n \times n) = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & a_0 \\ 1 & 0 & 0 & \dots & 0 & a_1 \\ 0 & 1 & 0 & \dots & 0 & a_2 \\ : & : & : & \dots & : & : \\ : & : & : & \dots & : & : \\ 0 & 0 & 0 & \dots & 0 & a_{n-2} \\ 0 & 0 & 0 & \dots & 1 & a_{n-1} \end{bmatrix} \qquad \Gamma_i(n \times 1) = \begin{bmatrix} b_{i1}^0 + a_0 b_{i1}^n \\ b_{i1}^1 + a_1 b_{i1}^n \\ b_{i1}^2 + a_2 b_{i1}^n \\ : \quad : \\ : \quad : \\ b_{i1}^{n-2} + a_{n-2} b_{i1}^n \\ b_{n-1} + a_{n-1} b_{i1}^n \end{bmatrix}$$

$$C(1 \times n) = [0\ 0\ 0\ \dots\ 0\ 1]$$

The number of parameters to be estimated is $N = 3n + 2$

Let the parameters vectors be

$$A^T = [a_0\ a_1\ a_2\ \dots\ a_{n-1}]$$

$$B_{i1p}^T = \left[b_{i1}^0\ b_{i1}^1\ b_{i1}^2\ \dots\ b_{i1}^{n-1}\right] \quad \text{for } i = 1,2$$

$$B_{i1}^T = \left[B_{i1p}^T \mid b_{i1}^n\right]$$

Let us introduce the total parameters vector

$$P^T(k) = \left[A^T \mid B_{11}^T \mid B_{21}^T \right] \quad (3n + 2) \times 1 \quad \textit{matrix}$$

Using the augmentation used in the previous paragraph, it can be shown that our system can be represented by the following set of equations

$$X(k+1) = \begin{bmatrix} T & 0 \\ 0 & T \end{bmatrix} X(k) + \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} Y(k) + \begin{bmatrix} B_{11p} \\ B_{21p} \end{bmatrix} \qquad (3.38)$$

$$Y(k) = \begin{bmatrix} C & 0 \\ 0 & C \end{bmatrix} X(k) + \begin{bmatrix} b_{11}^n \\ b_{21}^n \end{bmatrix} u(k) \tag{3.39}$$

It can be shown that equations (3.38) and (3.39) may also be written as

$$X(k+1) = \mathbf{T}X(k) + G(k)P(k) \tag{3.40}$$

$$P(k+1) = I_{3n+2}P(k) \tag{3.41}$$

$$y(k) = \mathbf{C}X(k) + N(k)P(k) \tag{3.42}$$

where

$$G(k) = \left[\begin{array}{c|c|c} y_1(k)I_n & u(k)I_n \,|\, 0 & 0 \\ y_2(k)I_n & 0 & u(k)I_n \,|\, 0 \end{array}\right] \quad (2n) \times (3n+2)$$

$$N(k) = \left[\begin{array}{c|c|c|c} 0 & u(k) & 0 & 0 \\ 0 & 0 & 0 & u(k) \end{array}\right] \quad 2 \times (3n+2)$$

$$\mathbf{T} = \begin{bmatrix} T & 0 \\ 0 & T \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} C & 0 \\ 0 & C \end{bmatrix}$$

The equations (3.40), (3.41) and (3.42) will have the following pseudo-linear form for the augmented state vector

$$S(k+1) = F(k)S(k) \tag{3.43}$$

$$y(k) = H(k)S(k) \tag{3.44}$$

$$F(k) = \begin{bmatrix} \mathbf{T} & G(k) \\ 0 & I_{3n+2} \end{bmatrix} \quad (5n+2) \times (5n+2) \qquad H(k) = [\mathbf{C} \;\; N(k)] \quad (2 \times (5n+2))$$

3.3.1 Generalization to the case of one-input and m-outputs

Now, we show that our representation contains all the parameters of the first column elements of G(z), and can be extended to the other elements for MIMO systems. However, we will end up with a huge system.

In general the system will be represented by

$$Y(z) = \frac{1}{z^n - a_{n-1}z^{n-1} - \dots - a_1 z - a_0} \begin{bmatrix} b_{11}^n z^n + b_{11}^{n-1} z^{n-1} + \dots + b_{11}^1 z + b_{11}^0 \\ : \; : \; : \; : \; : \; : \\ : \; : \; : \; : \; : \; : \\ b_{m1}^n z^n + b_{m1}^{n-1} z^{n-1} + \dots + b_{m1}^1 z + b_{m1}^0 \end{bmatrix} \quad (3.51)$$

The number of parameters to be estimated is $N = m(n+1) + n$

This system can be written in the following state variable form

$$X(k+1) = \begin{bmatrix} \Phi & 0 & \dots & 0 \\ 0 & \Phi & \dots & 0 \\ : & : & : & : \\ : & : & : & : \\ 0 & 0 & \dots & \Phi \end{bmatrix} X(k) + \begin{bmatrix} \Gamma_1 \\ \Gamma_2 \\ : \\ : \\ \Gamma_m \end{bmatrix} u(k) \quad (3.52)$$

$$Y(k) = \begin{bmatrix} C & 0 & \dots & 0 \\ 0 & C & \dots & 0 \\ : & : & : & : \\ : & : & : & : \\ 0 & 0 & \dots & C \end{bmatrix} X(k) + \begin{bmatrix} b_{11}^n \\ b_{21}^n \\ : \\ : \\ b_{m1}^n \end{bmatrix} u(k) \quad (3.53)$$

The parameters vectors are

$$A^T = [a_0\ a_1\ a_2\ \dots\ a_{n-1}]$$

$$B_{l1p}^T = \left[b_{l1}^0\ b_{l1}^1\ b_{l1}^2\ \dots\ b_{l1}^{n-1}\right] \quad for\ l = 1,\dots,m$$

$$B_{l1}^T = \left[B_{l1p}^T \mid b_{l1}^n\right]$$

Using the argumentation used in the previous paragraph, it can be shown that our system can be represented by the following set of equations

$$X(k+1) = \mathbf{T}X(k) + \mathbf{A}y(k) + \mathbf{B}u(k) \tag{3.54}$$

$$Y(k) = \mathbf{C}X(k) + B^n u(k) \tag{3.55}$$

$$B^n = \left[b_{11}^n\, b_{21}^n \dots b_{m1}^n\right]^T$$

$$\mathbf{B}^T = \left[\, B_{11p}^T \mid B_{21p}^T \mid \dots\dots \mid B_{m1}^T\,\right]$$

$$\mathbf{A} = \begin{bmatrix} A & 0 & \dots\dots & 0 \\ 0 & A & \dots\dots & 0 \\ : & : & : & : \\ : & : & : & : \\ 0 & 0 & \dots\dots & A \end{bmatrix} \quad \mathbf{T} = \begin{bmatrix} T & 0 & \dots\dots & 0 \\ 0 & T & \dots\dots & 0 \\ : & : & : & : \\ : & : & : & : \\ 0 & 0 & \dots\dots & T \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} C & 0 & \dots\dots & 0 \\ 0 & C & \dots\dots & 0 \\ : & : & : & : \\ : & : & : & : \\ 0 & 0 & \dots\dots & C \end{bmatrix}$$

Let us introduce the total parameters vector

$$P^T(k) = \left[\, A^T \mid B_{11}^T \mid B_{21}^T \mid \dots \mid B_{m1}^T\right] \times 1 \quad matrix$$

It can be shown that equations (3.54) and (3.55) may also be written as

$$X(k+1) = \mathbf{T}X(k) + G(k)P(k) \tag{3.56}$$

$$P(k+1) = \mathbf{I}P(k) \tag{3.57}$$

$$y(k) = \mathbf{C}X(k) + N(k)P(k) \tag{3.58}$$

Where

$$G(k) = \left[\begin{array}{c|c|c|c} y_1(k)I_n & u(k)I_n|0 & 0 & 0 \\ y_2(k)I_n & 0 & u(k)I_n|0 & 0 \\ :\,: & :\,: & :\,: & :\,: \\ y_m(k)I_m & 0 & 0 & u(k)I_n|0 \end{array}\right] \quad (mn + n + m) \times (mn + n + m)$$

$$N(k) = \begin{bmatrix} 0|u(k) & 0 & \ldots & 0 \\ 0 & 0|u(k) & \ldots & 0 \\ :\,: & :\,: & :\,: & :\,: \\ 0 & 0 & \ldots & 0|u(k) \end{bmatrix} \quad m \times (mn + n + m)$$

Finally, the pseudo-linear representation is given by the following equations

$$S(k + 1) = F(k)S(k) \tag{3.59}$$

$$y(k) = H(k)S(k) \tag{3.60}$$

Where

$$F(k) = \begin{bmatrix} \mathbf{T} & G(k) \\ 0 & \mathbf{I} \end{bmatrix}$$

and

$$H(k) = \left[\mathbf{C} \mid N(k)\right]$$

3.3.2 Identification algorithm

At this point, the results derived are of theoritical interest only, since the measurements are always contaminated with noise. In practical situations one may model the system as follows

$$S(k+1) = F(k)S(k) + V(k) \tag{3.29}$$

$$y(k) = H(k)S(k) + w(k) \tag{3.30}$$

where:

S(k) is the extended state vector of the system

y(k) the measured output signal

V(k) and w(k) are white gaussian sequences with correlation functions

$$E\{\begin{bmatrix} V(k) \\ w(k) \end{bmatrix}\begin{bmatrix} V(l) \\ w(l) \end{bmatrix}^T\} = \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \delta_{kl} \tag{3.31}$$

and more precisely, V(k) and w(k) are independent of the past and present states of the system. Equations (3.29) and (3.30) are derived from stochastic difference equations, which include the processes generating the noises and perturbations

The matices F(k) and H(k) are available at each time k from the record of the input-output data. We assume also that the matrices Q and R are known, so are the statistical characteristics of the initial state S(0), that is to say its mean value and covariance.

At time k, we assume the measurements vector

$$Y^T(k) = [\ Y(0)\ Y(1)\ \dots\ Y(k)\ ] \tag{3.32}$$

and seek to estimate the unknown vector

$$\mathbf{S}^T(k) = [\ S^T(0)\ S^T(1)\ \ldots\ S^T(k)\ ] \tag{3.33}$$

The assumptions of the problem allow us in theory to calculate the mean values, variances and covariances of S(k) and Y(k), hence, the possibility to apply the general formulas of the optimal linear estimation to obtain S(k) as a linear function of Y(k)

The estimation problem can be stated as the choice of the optimal gain K(k) such that the reconstruction error (3.34) is minimal.

$$\hat{S}(k+1) = F(k)\hat{S}(k) + K(k)\left[\ y(k) - H(k)\hat{S}(k)\ \right] \tag{3.34}$$

Kalman theory can be applied to yield an estimation technique for the identification of the parameters and the estimation of the states.

From a standard discrete Kalman filter an estimation algorithm is developed. let

S(k/k) = optimal filtered estimate of S(k)

S(k+1/k) = optimal predicted estimate of S(k)

If the optimal filtered estimate $\hat{S}(k/k)$ and the covariance matrix P(k/k) of the corresponding filtering error $\hat{S}(k/k)$ are known for some k, then the single-stage optimal predicted estimate for all admissible loss functions is given by the expression

$$\hat{S}(k+1/k) = F(k)\hat{S}(k/k) \tag{3.35}$$

The stochastic process $\{\hat{S}(k\ \ 1/k),\ k=0,1,..\}$ defined by the single-stage predicted error relation

$$\bar{S}(k+1/k) = S(k+1) - \hat{S}(k+1/k)$$

is a zero mean gauss-markov process whose covariance matrix is given by the relation

$$P(k+1/k) = F(k)P(k/k)F^T(k) + Q \qquad (3.36)$$

The optimal filtered estimate $\hat{S}(k \quad 1/k \quad 1)$ is given by the recursive relation

$$\hat{S}(k+1/k+1) = F(k)S(k/k) + K(k+1)\left[y(k+1) \_ H(k+1)F(k)\hat{S}(k/k)\right] \qquad (3.37)$$

for k=0,1,.... where $\hat{S}(0/0)$ =S(0)

K(k+1) is the filter gain specified by the set of relations

$$K(k+1) = P(k+1/k)H^T(k+1)\left[H(k+1)P(k+1/k)H^T(k+1) + R\right]^{-1} \qquad (3.38)$$

$$P(k+1/k) = F(k)P(k/k)F^T(k) + Q \qquad (3.39)$$

$$P(k+1/k+1) = [I - K(k+1)H(k+1)]P(k+1/k) \qquad (3.40)$$

For k=0,1,.... and P(0/0)=P(0)

3.3.3 Illustrative example

To illustrate the proposed approach, a system with one-input and two-outputs was considered for identification. The discrete-time transfer function matrix that describes the system is strictly proper, which reduces the number of parameters to estimate.

Consider the 2 × 1 pulse transfer function matrix

$$H(z) = \frac{1}{D(z)} \begin{bmatrix} H_{11}(z) \\ H_{21}(z) \end{bmatrix} \tag{E.1}$$

For the given example it was assumed

$$D(z) = z^3 - a_2 z^2 - a_1 z - a_0 \tag{E.2}$$

$$H_{11}(z) = b_{11}^2 z^2 + b_{11}^1 z + b_{11}^0 \tag{E.3}$$

$$H_{21}(z) = b_{21}^2 z^2 + b_{21}^1 z + b_{21}^0 \tag{E.4}$$

With the following values of the parameters

$$a_0 = 0.32 \qquad a_1 = 0.88 \qquad a_2 = 0.2$$

$$b_{11}^0 = -2.28 \qquad b_{11}^1 = 3.2 \qquad b_{11}^2 = 1.0$$

$$b_{21}^0 = -4.08 \qquad b_{21}^1 = 2.4 \qquad b_{21}^2 = 2.0$$

[The assumption that the pulse transfer function matrix is strictly proper leads to $b_{21}^3 = 0$ and $b_{11}^3 = 0$]

Using the augmentation used in the previous paragraph, it can be shown that our system can be represented by the following set of equations

$$X(k+1) = \begin{bmatrix} T & 0 \\ 0 & T \end{bmatrix} X(k) + \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} Y(k) + \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix} u(k) \tag{E.5}$$

$$Y(k) = \begin{bmatrix} C & 0 \\ 0 & C \end{bmatrix} X(k) \tag{E.6}$$

Where

$$A^T = [\ a_0\ a_1\ a_2\ ]$$

$$B_{i1}^T = \left[\ b_{i1}^0\ b_{i1}^1\ b_{i1}^2\ \right] \quad for \quad i = 1,2$$

$$T = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad C = [\ 0\ \ 0\ \ 1\ ]$$

Let us define the characteristic vector

$P^T(k) = \left[\ A^T \mid B_{11}^T \mid B_{21}^T\ \right]$

From equations (3.38) and (3.39) we may also write

$$X(k+1) = \mathbf{T}X(k) + G(k)P(k) \tag{E.7}$$

$$P(k+1) = IP(k) \tag{E.8}$$

$$Y(k) = \mathbf{C}\,X(k) \tag{E.9}$$

Where

$$\mathbf{T} = \begin{bmatrix} T & \mathbf{0} \\ \mathbf{0} & T \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} C & \mathbf{0} \\ \mathbf{0} & T \end{bmatrix}$$

$$G(k) = \begin{bmatrix} y_1(k)I_3 & | & u(k)I_3 & | & \mathbf{0} \\ y_2(k)I_3 & & \mathbf{0} & & u(k)I_n \end{bmatrix}$$

Using the extended vector $S^T(k)\ \left[\ X^T(k)\ \mid\ P^T(k)\ \right]$, it leads to

$$S(k+1) = F(k)S(k) \tag{E.10}$$

$$Y(k+1) = H(k)S(k) \tag{E.11}$$

Taking into account the presence of noise, and assuming that the measurements are contaminated with a well defined noise, we reformulate our system as follows

$$S(k+1) = F(k)S(k) + V(k) \tag{E.12}$$

$$Y(k) = H(k)S(k) + W(k) \tag{E.13}$$

Where the noises V(k) and W(k) are white gaussian noises, assumed to be of zero-mean and uncorrelated.

The proposed algorithm was used to perform the estimation of the parameters from a simulated input-output data. A unit step to which was added a zero-mean, white gaussian noise was used to excite the system.

The following algorithm was used

$$M(k) = F(k-1)P(k-1)F^T(k) + Q \tag{E.14}$$

$$K(k) = M(k)H^T(k)\left[\, H(k)M(k)H^T(k) + R \,\right]^{-1} \tag{E.15}$$

$$K(k) = [\, I - K(k)H(k)\,]M(k) \tag{E.16}$$

and

$$\hat{S}(k) = F(k-1)\hat{S}(k-1) + K(k)\left[\, Y(k) - H(k)F(k-1)\hat{S}(k-1)\,\right] \tag{E.17}$$

The initial state estimate is assumed to be null, the input and output noises are independent normalized gaussian white noises, with known covariance matrices

$$S(0) = \mathbf{0} \quad and \quad P(0) = 100I$$

$$Q(15 \times 15) = \begin{bmatrix} Q_s & 0 \\ 0 & 0 \end{bmatrix} \qquad Q_s(6 \times 6) = diag(.001,.001,.001,.001,.001,.001)$$

$$R(2 \times 2) = diag(.001,.001)$$

The results of the first 30 iterations are shown in the figures at the end of the chapter. Only the estimates of the parameters of the characteristic polynomial are shown. As before, two different noise levels were used ($\sigma^2 = 0.10$, $\sigma^2 = 0.30$).

As shown in the figures, the estimates are very good; reasonable estimates are obtained after just 10 iterations, which proves that this approach converges very fast.

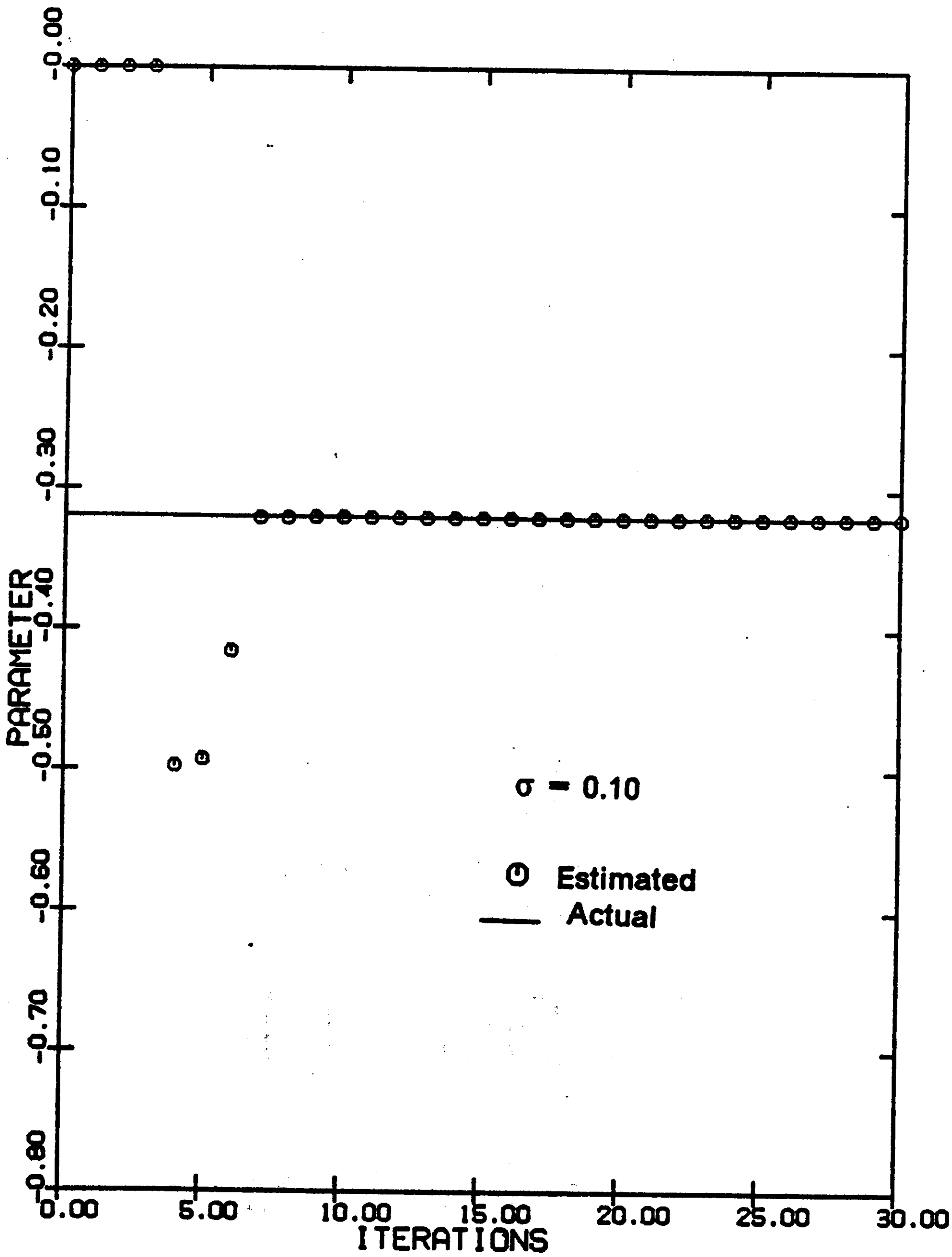

Figure 14. Estimated and actual values, in the case of low order noise level, of the parameter a(0), in the SIMO case.

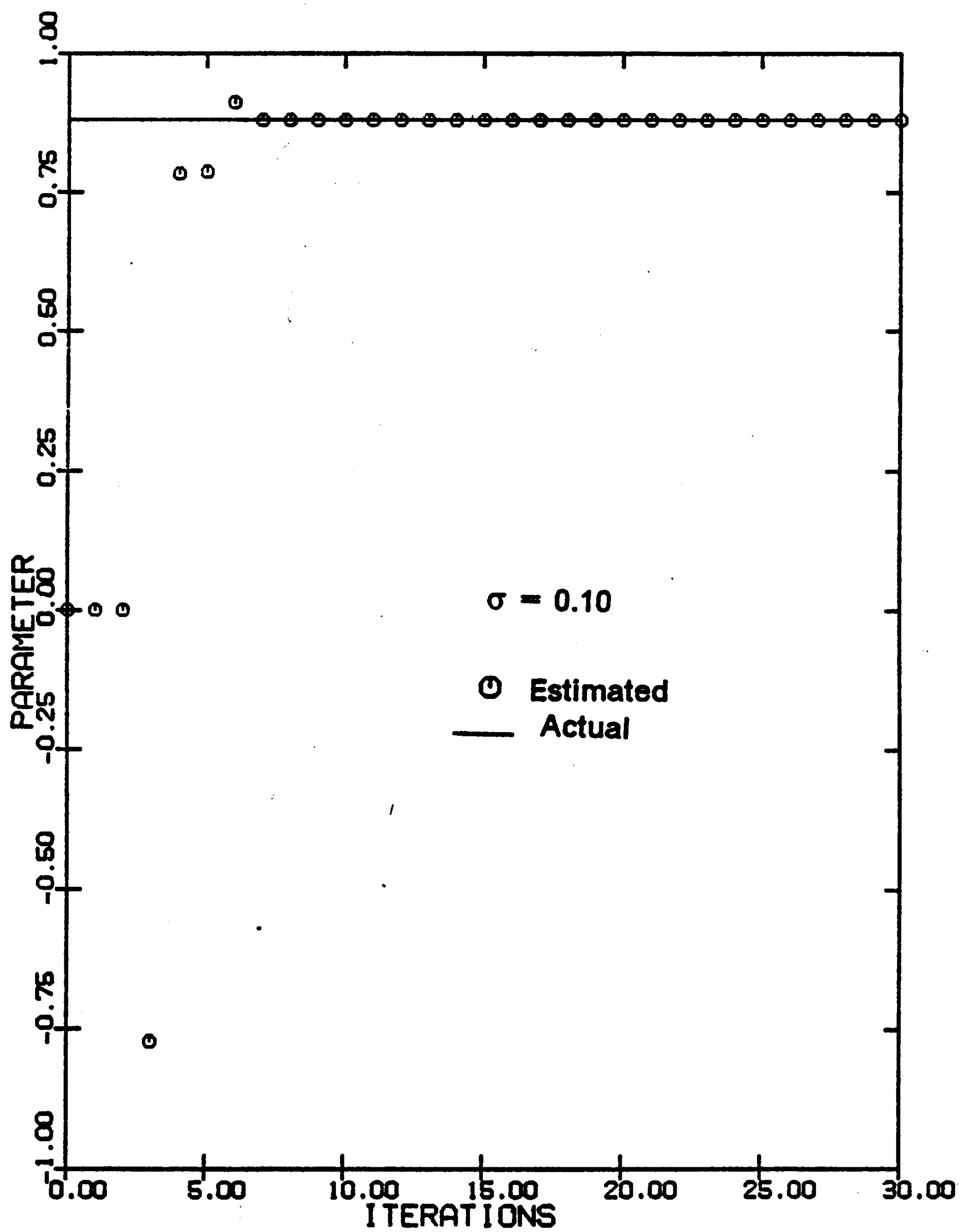

Figure 15. Estimated and actual values, in the case of low order noise level, of the parameter a(1), for the SIMO case.

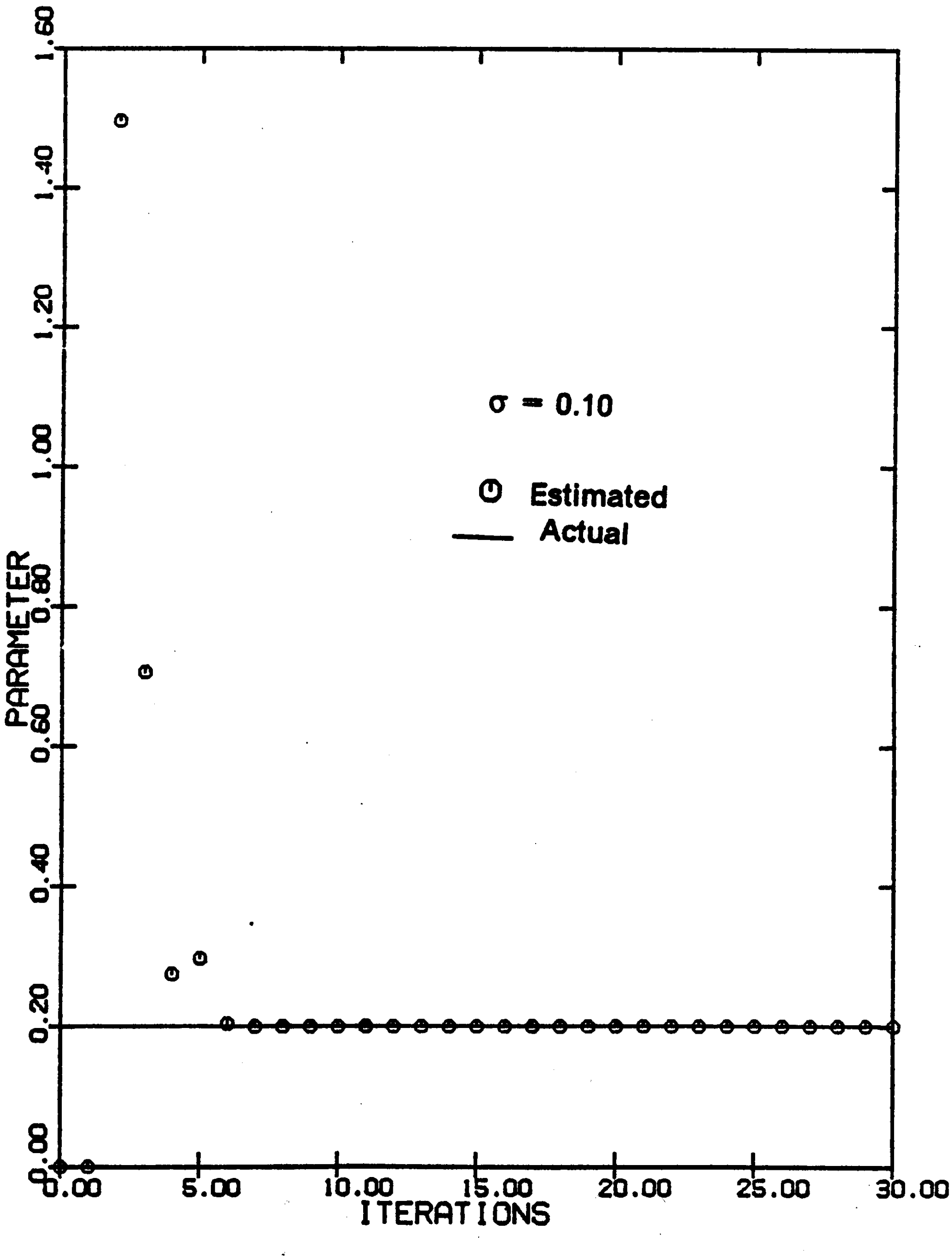

Figure 16. Estimated and actual values, in the case of low order noise level, of the parameter a(2), for the SIMO case.

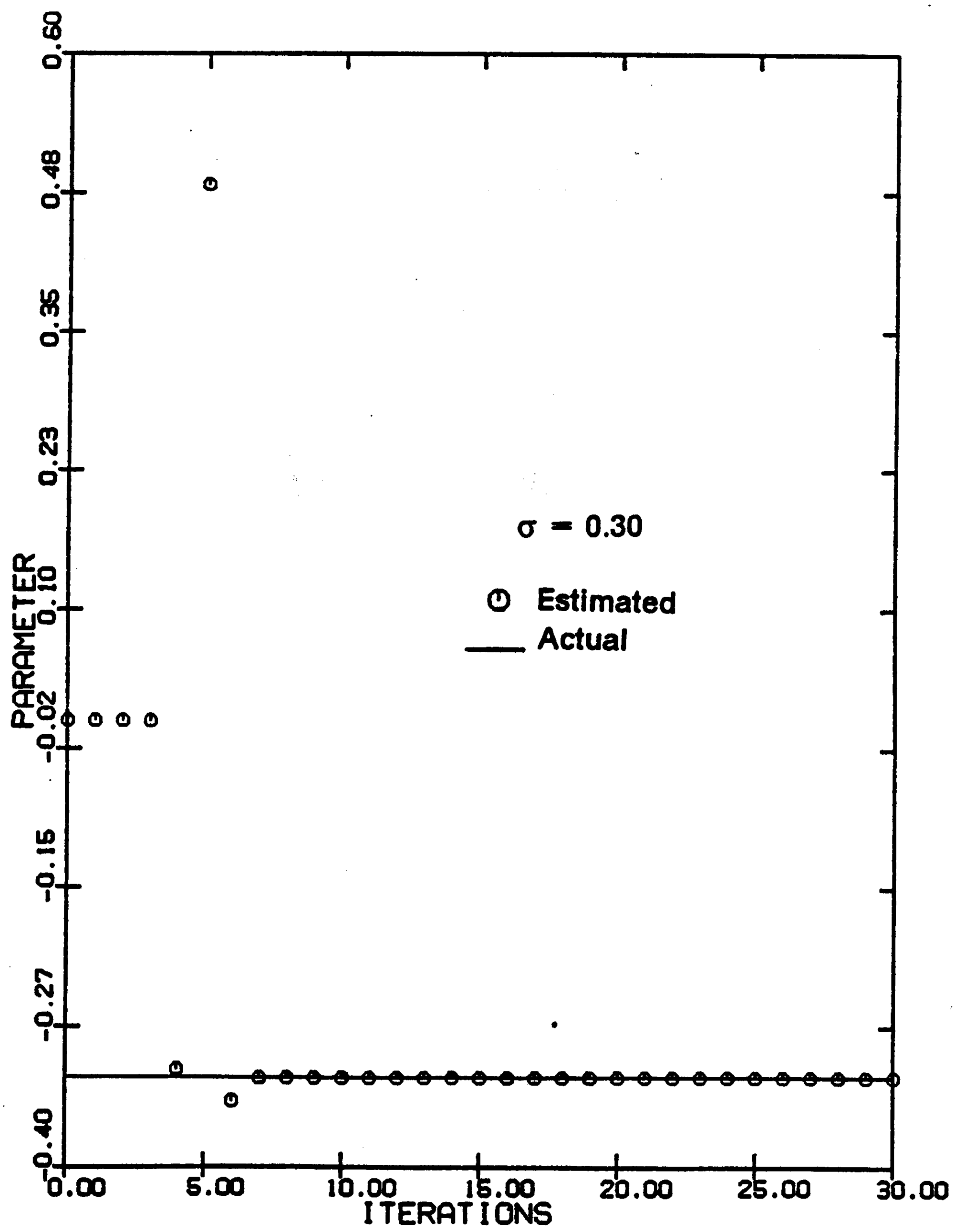

Figure 17. Estimated and actual values, in the case of high order noise level, of the parameter a(0), for the SIMO case.

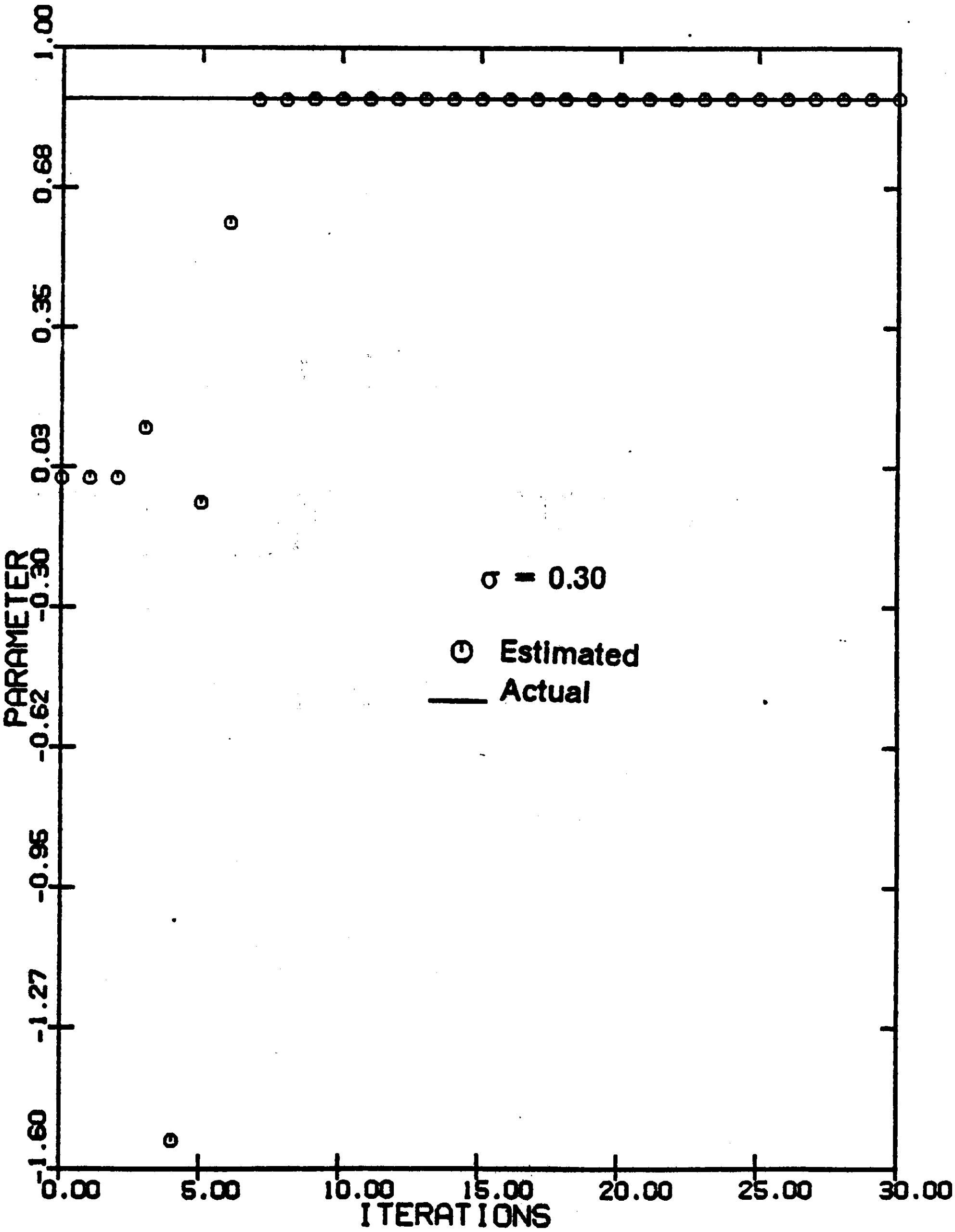

Figure 18. Estimated and actual values, in the case of high order noise level, of the parameter a(1), for the SIMO case.

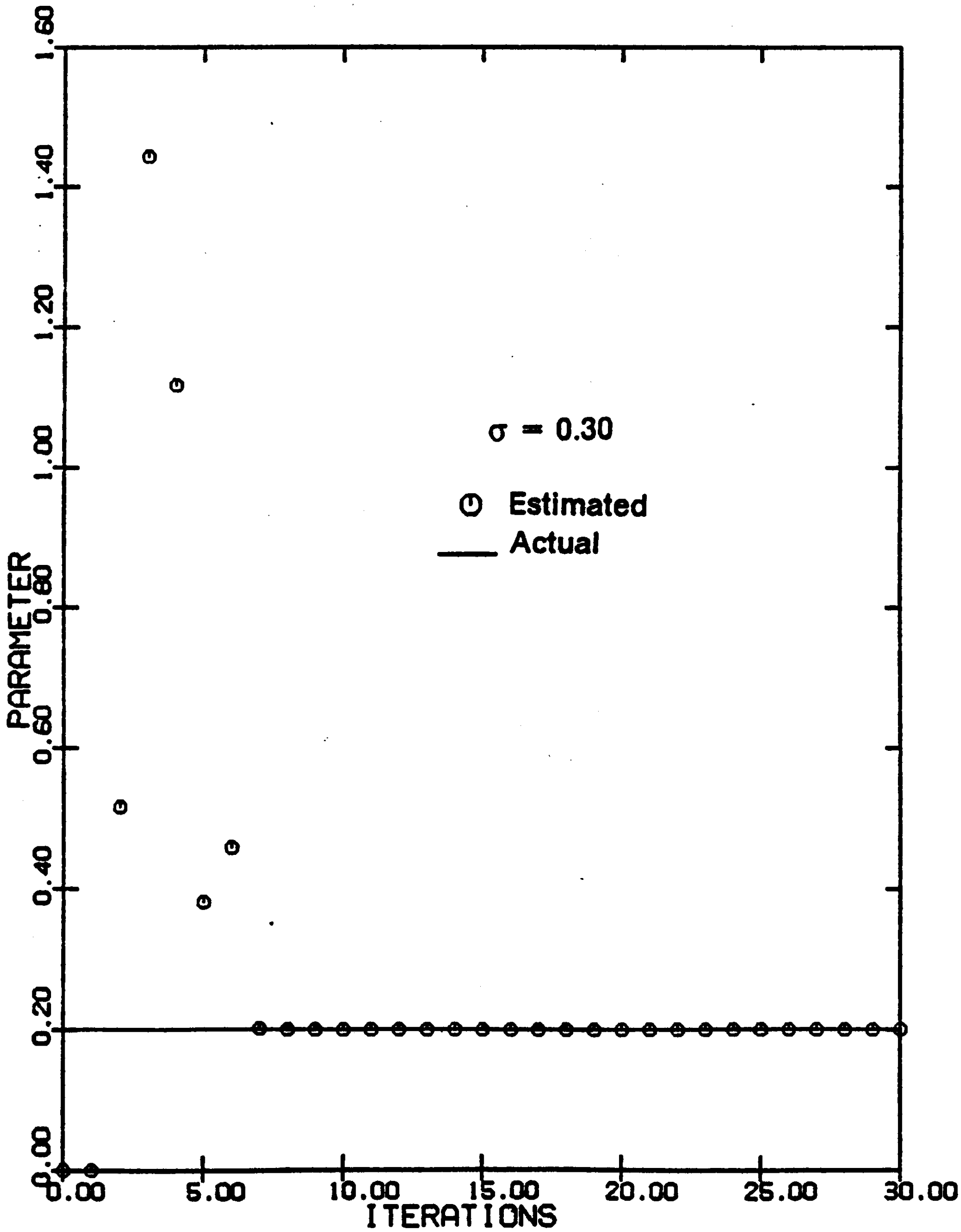

Figure 19. Estimated and actual values, in the case of high order noise level, of the parameter a(2), for the SIMO case.

MINIMAL REALIZATION AND IDENTIFICATION OF MIMO SYSTEMS

4.0 Introduction

The problem of deriving a minimal realization of a linear dynamical system, first introduced by Kalman (1963), can be generally defined as the problem of determining a triple of matrices (A,B,C), which describes the system behavior in the usual state space representation, from the knowledge of any other given characterization of the system. This problem is still an interesting area of research both for theoretical implications and for the role that the state space representation plays in the development of unitary and efficient algorithms for analysis and synthesis purposes.

Many authors [24,37] have been concerned with the matrices of rational proper functions, partial fraction expansions, weighting patterns as initial system data. Moreover, the direct realization from input/output observations has been emphasized by some authors, mainly extending the original work by Ho and Kalman (1965). Some authors, starting from an input/output canonical representation give a canonical realization from a sequence of Markov parameters [1,5]. However, these methods intend to estimate more parameters than systems actually need, therefore, the efficiency of identification can be greatly degraded in case of the presence of noises. Among the models used for identification, the canonical state space model are primarily used 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. For a given input/output description, the matrices (A,B,C) are not unique, and any non-singular linear transformation of the state will give another set of these matrices. This has led to the development of several canonical forms of the state space formulation that reduce the number of parameters to be estimated, as well as make the problem of identification simpler [11]. However, each canonical form introduces some additional structural parameters which must also be determined before the model parameters are estimated.

The problem of system identification consists of two main steps: structural determination and parameter estimation. The first step is more difficult, especially for the case of noisy data, and successful identification depends heavily on the model used; every model has its own structural parameters. After estimating or assuming the structural parameters, the parameters of the model can be readily identified. Another factor in the problem of identification from noisy data is the assumption made about the noise. The contribution presented here regards a new realization procedure from external data observations. The key to the method consists of an input/output representation which unequivocally and in a very simple way is connected with the irreducible matrix-fraction description (MFD) of ARMA models. This property allows us to perform data selection leading to a minimal realization. In this chapter, the canonical representation introduced in chapter 1 is extended to the minimal realization case of MIMO linear discrete-time systems. A structural identification algorithm [13] is given, and a discrete Kalman filter is used to combine state and parameter estimation using the approach developed for the SISO linear systems.

4.1 Matrix-Fraction Description of ARMA Models

We shall study the problem of specifying unique models for input-output relations of the auto-regressive moving-average (ARMA) type model.

$$A_0Y(k) + A_1Y(k-1) + ... + A_pY(k-p) = B_0U(k) + B_1U(k-1) + ... + B_pU(k-p) \quad (4.1)$$

We shall assume that Y(.) is m-dimensional and U(.) r-dimensional vector sequences. Taking the z-transform of equation (4.1) gives

$$A(z)Y(z) = B(z)U(z) \quad (4.2)$$

Where

$$A(z) = \sum_{i=0}^{p} A_iz^{-i} \qquad B(z) = \sum_{i=0}^{p} B_iz^{-i} \quad (4.3)$$

To assure that the Y(.) sequence is uniquely determined by U(.), we shall require that:

$$Det A(z) \neq 0 \quad (4.4)$$

at all but a finite number of points

Finally, we assume that

$$T(z) = A^{-1}(z)B(z) \tag{4.5}$$

is a proper rational matrix, namely that

$$\lim T(z) = E < \infty \quad for \quad z \rightarrow \infty \tag{4.6}$$

This guarantees that T(z) is the transfer function of some linear dynamical system. (A(z),B(z)) is known as a left matrix-fraction description (MFD) of T(z).

Among all MFD's of T(z) are some in which the numerator and denominator are left coprime, that is they have no common left factors except unimodular matrices (polynomial matrices with polynomial inverses). Such MFD's will be called irreducible, it follows that the determinantal degree of the denominator matrix of every irreducible MFD of T(z) is the same and is minimal among the class of all MFD's of T(z). In this sense, the irreducible MFD's are the simplest of all possible MFD's of a given T(z).

It should be clear that the problem of determining a set of canonical forms for MFD's, is precisely the classical problem of determining canonical forms for one-sided equivalence of full rank polynomial matrices. Since an irreducible MFD can always be found from an arbitrary MFD by removing the greatest common left divisor of numerator and denominator, a set of canonical forms for all MFD's can be found in this way.

A subset of the irreducible MFD's has the additional property that the determinantal degree of A(z), is equal to the sum of the degrees of the rows of A(z). Such an A(z) will be said to be row proper, following Wolovich, and the corresponding MFD will be called standard. By row permutations, the row degrees may be arranged in decreasing order. Furthermore, it can be shown that the irreducibility of the MFD (A(z),B(z)) ensures the observability and controllability of the realization (A,B,C), the determinantal degree of A(z), in an irreducible MFD of T(z), is the dimension of a minimal realization of T(z).

The classical canonical form for left equivalence of polynomial matricies provides the first canonical form for an irreducible MFD. The Hermite canonical MFD ($A_H(z)$, $B_H(z)$), is obtained by taking any irreducible MFD(A(z),B(z)) and applying the elementary row operations that bring A(z) to its Hermite form $A_H(z)$. The same operations applied to B(z) then give $B_H(z)$. In this form, the matrix $A_H(z)$ is lower triangular, with monic polynomials as diagonal elements and with each diagonal element having the greatest degree in its column. Although $A_H(z)$ is not necessarily row proper, it is column proper. Rosenbrock [28] used the Hermite MFD to specify a unique MFD form .

A second canonical MFD is is the echelon MFD introduced by Popov[27], denoting this form by ($A_E(z)$, $B_E(z)$), it is specified by the following requirements. There is a set of pivot indices $\{s_i ; i \in m\}$ defined so that if $A(z) = (A_{ij}(z))$, then:

- 0/ $n_1 \geq n_2 \geq \ldots \geq n_m$
- 1/ degree of $A_{is_i}(z) = n_i$
- 2/ $A_{is_i}(z)$ is monic polynomial
- 3/ degree of $A_{is_j}(z) < n_j$ for $i \neq j$
- 4/ degree of $A_{ij}(z) < n_i$ for $j > s_i$
- 5/ If $n_i = n_j$ and $i > j$ then

$s_i > s_j$

In other words, $A_\varepsilon(z)$ is a row proper matrix with row degrees $\{n_i\}$. The polynomial $A_{is_i}(z)$ is monic and has a degree n_i that is larger than that of any other polynomial in column s_i and of any polynomial to its right in row i. Finally if two rows have the same degree, the pivot index of the first is smaller than that of the second. Since ($A_\varepsilon(z)$, $B_\varepsilon(z)$) is a standard, irreducible MFD, an observer form realization may be associated with it.

In this chapter, we shall show that a more complete canonical MFD is closely related to observable canonical state space models.

Consider the equation(4.2)

$$A(z)Y(z) = B(z)U(z) \tag{4.2}$$

Where

$$A(z) = \begin{bmatrix} A_{11}(z) & .... & A_{1m}(z) \\ : & : & : \\ : & : & : \\ A_{m1}(z) & .... & A_{mm}(z) \end{bmatrix} \qquad B(z) = \begin{bmatrix} B_{11}(z) & .... & B_{1r}(z) \\ : & : & : \\ : & : & : \\ B_{m1}(z) & .... & B_{mr}(z) \end{bmatrix}$$

The polynomials of A(z) are given by:

$$A_{ii}(z) = z^{n_i} + a_{ii}^{n_i-1} z^{n_i-1} + ... + a_{ii}^1 z + a_{ii}^0 \tag{4.7}$$

$$A_{ij}(z) = a_{ij}^{n_{ij}} z^{n_{ij}} + a_{ij}^{n_{ij}-1} z^{n_{ij}-1} + ... + a_{ij}^1 z + a_{ij}^0 \tag{4.8}$$

Where $n_{ij} \le n_i$ for all $j \neq i$

i.e. A(z) is a row reduced polynomial matrix, with the diagonal polynomials as monic polynomials with the highest degree in their rows. Then the set of pivot indices is defined such that $s_i = i$. We will also assume that the rational matrix T(z) is proper. i.e.

deg of i^{th} column of B(z) $\leq$ deg of i^{th} column of A(z)

Also

$$B_{ij}(z) = b_{ij}^{n_{ij}} z^{n_{ij}} + b_{ij}^{n_{ij}-1} z^{n_{ij}-1} + \ldots\ldots + b_{ij}^{1} z + b_{ij}^{0} \tag{4.9}$$

we assume that: $n_m \leq \ldots \leq n_2 \leq n_1$ Where $\{n_i\ ,\ i=1,..,m\}$ are the structural indices, or The observability indices since they, in the dual sense, correspond to the controllability indices [28]. It is shown elsewhere [23,24] that only this number of parameters is uniquely identifiable from the given input/output data. Therefore, the number of these parameters represents the minimal number of parameters by which a given linear multivariable system may be described. The observability indices are obtained by constructing the vector sequences

$$\begin{array}{c} C_1,\ A^T C_1,\ A^{T^2} C_1, \ldots\ldots \quad , \\ \ldots,\ \ldots,\ \ldots,\ \ldots \\ C_m,\ A^T C_m,\ A^{T^2} C_m, \ldots\ldots \end{array} \tag{4.10}$$

and selecting them in the following order

$$C_1,\ C_2,\ \ldots\ldots,\ C_m,\ A^T C_1,\ \ldots\ldots,\ A^T C_m,\ \ldots\ldots \tag{4.11}$$

Using the formula of retaining a vector $A^{T^k} C_i$ if and only if it is independent from previously selected ones. Let $n_1, n_2, \ldots\ldots, n_m$ be the numbers of vectors selected from the first, second,...., m^{th} sequence in (4.10) Because of the complete observability of the system it follows

$$n = \sum_{1}^{m} n_l \tag{4.12}$$

4.2 Structural Identification

It is well known that the basic feature of multivariable systems is not specified by a single order n, but by a set of integers $(n_1, n_2, ..., n_m)$ which will be refered to as "structural indices". Therefore, the problem of structure identification is to determine those unknown indices from input/output observations.

A method developed by Guidorzi [4,5], and that has been successfully used, mainly in the deterministic case, is presented to give some insight into the problem of structural identification.

Consider the s^{th} equation in the set (4.2),i.e.

$$\sum_{l=1}^{m} A_{sl}(z) y_l(k) = \sum_{l=1}^{r} B_{sl}(z) u_l(k) \tag{4.13}$$

That, on the basis of the relations (4.7) , (4.8) and (4.9), can also be written in the following form

$$y_s(k + n_s) = - \sum_{l=1}^{m} \sum_{j=1}^{n_{sl}} a_{sl}^{j} y_l(k + j - 1) + \sum_{l=1}^{r} \sum_{j=1}^{n_{sl}} b_{sl}^{j} u_l(k + j - 1) \qquad (in\ n_{ll} = n_l) \tag{4.14}$$

Consider now the matrix of input/output data given by

$$
\begin{bmatrix}
y_1(k) & y_1(k+1) & \dots & y_m(k) & \dots & u_1(k) & u_1(k+1) & \dots & u_r(k) & \dots \\
y_1(k+1) & y_1(k+2) & \dots & y_m(k+1) & \dots & u_1(k+1) & u_1(k+2) & \dots & u_r(k+1) & \dots \\
: & : & : & : & : & : & : & : & : & : \\
: & : & : & : & : & : & : & : & : & : \\
y_1(k+N) & \dots & \dots & y_m(k+N) & \dots & u_1(k+N) & \dots & \dots & u_r(k+N) & \dots
\end{bmatrix}
$$

$$
= \left[\; Y_1(k)\; Y_1(k+1) \dots \mid \dots \mid Y_m(k) \dots \mid U_1(k) \dots \mid \dots \mid U_r(k) \dots \;\right] \tag{4.15}
$$

Equation (4.14) shows that the dependence relations among the vectors of (4.15) are the same, also taking into account the input, as those among the vectors (4.10). This property allows the determination of the structural indices by selecting the vectors (4.15) according to the same selection plan in (4.11). The vectors of (4.15) will therefore be selected in the following order.

$$
Y_1(k),\ \dots,\ Y_m(k),\ U_1(k),\ \dots,\ U_r(k),\ Y_1(k+1),\ \dots,\ Y_m(k+1),\ \dots \tag{4.16}
$$

A vector is retained if and only if it is independent from previously selected ones; when a dependent vector $Y_s(k + v_s)$ is found, all the remaining vectors belonging to the same submatrix will also be dependent so that their test is unnecessary. The selection ends when a dependent vector has been found in every output submatrix; the numbers of vectors selected from these submatrices will be $n_1,\ n_2,\ \dots,\ n_m$

The integer N in (4.15) must be large enough ($N > n \;\; rv_M$ *where* v_M max(v_i)) in order to permit the selection of the necessary number of independent vectors. In other words, the input-output sequence must be of sufficient length to permit the complete structural identification of the system. Usually the number of available data is many times the system order so that the previous condition is largely fulfilled.

For the structural identification, the input sequence must satisfy well known conditions, i.e. must excite all modes of the system. Also the requirement can easily be met when the length of the data be long enough with respect to the system order. It is certainly not advisable to carry out the structural identification directly on the vectors of (4.15) because of the large amount of storage necessary; morever, the required storage would be a function of N. Since for every matrix , D, rankD = rank(D^TD) a more useful algorithm follows.

4.2.0 Algorithm for the structural identification

Let

$$L_i(Y_j) = \left[\ Y_j(k)\, Y_j(k+1) \ldots\ldots Y_j(k+i-1)\ \right] \tag{4.17}$$

$$L_i(U_j) = \left[\ U_j(k)\, U_j(k+1) \ldots\ldots U_j(k+i-1)\ \right] \tag{4.18}$$

Then the matrix (4.15) taking δ_1 vectors in the firstsubmatrix, δ_2 vectors in the second, etc... can bewritten as

$$R(\delta_1,\ \delta_2,\ \ldots,\ \delta_{m+r}) = \{\ L_{\delta_1}(Y_1) \ldots\ldots L_{\delta_m}(Y_m)\ |\ L_{\delta_{m+1}}(U_1) \ldots\ldots L_{\delta_{m+r}}(U_r)\ \} \tag{4.19}$$

Define the product R^TR as

$$S(\delta_1,\ \delta_2,\ \ldots,\ \delta_{m+r}) = R^T(\delta_1,\ \delta_2,\ \ldots,\ \delta_{m+r})R(\delta_1,\ \delta_2,\ \ldots,\ \delta_{m+r}) \tag{4.20}$$

$S(\delta_1, \delta_2, \ldots, \delta_{m+r})$ is therefore a square matrix whose dimension is given by $\delta_1 + \delta_2 + \ldots. + \delta_{m+r}$

Construct then the sequence of increasing dimension matrices

$$S(2,1,...,1), S(2,2,...,1), ..., S(2,2,...,2), ... \tag{4.21}$$

and select from (4.21) nonsingular ones. When a singular matrix is found, one of the indices is determined; the procedure ends, all remaining matrices are singular, when all m indices are determined. If two adjacent matrices in sequence (4.21) are considered, all the elements of the first are present in the subsequent one that can thus be obtained by computing only a limited number of terms.

let

$S(\mu_1, \mu_2, ..., \mu_{m+r})$ be a singular matrix in (4.21) and let μ_i be the index increased by one with respect to the previous nonsingular matrix in the sequence. Then $n_i = \mu_i - 1$ while the indices n_{ij} are given by $n_{ij} = \mu_i$ (j = 1,...,m) ($i \neq j$).

4.2.1 Numerical example

The input/output sequences of a system with one input and two outputs are given by

k	u(k)	$y_1(k)$	$y_2(k)$
0	1.	0.	0.
1	2.	0.	0.
2	4.	1.	0.
3	5.	0.	2.
4	-5.	2.	-1.
5	-12.	5.	4.
6	15.	-3.	-9.

7	50.	5.	10.
8	10.	-25.	8.
9	-60.	56.	-7.
10	30.	-3.	10.
11	0.	-27.	-4.
12	0.	40.	5.

Taking N=9, the sequences of matrices S(2,1,1), S(2,2,1), ..., S(3,2,2), ..., S(3,3,2), ... is constructed.

As an example let us show how the matrix S(2,1,1) was obtained:

Consider $S(\delta_1, \delta_2, \delta_3)$ with $\delta_1 = 2$, $\delta_2 = 1$ and $\delta_3 = 1$.

We know that

$$S(\delta_1, \delta_2, \delta_3) = R^T(\delta_1, \delta_2, \delta_3)R(\delta_1, \delta_2, \delta_3)$$

where

$$R(\delta_1, \delta_2, \delta_3) = \{\ L_{\delta_1}(Y_1)\, L_{\delta_2}(Y_2) \mid L_{\delta_3}(U)\ \}$$

and

$$L_{\delta_1}(Y_1) = L_2(Y_1) = [\ Y_1(k)\, Y_1(k+1)\ ]$$

$$L_{\delta_2}(Y_2) = L_1(Y_2) = [\ Y_2(k)\ ]$$

$$L_{\delta_3}(U) = L_1(U) = [\ U(k)\ ]$$

so

$$R(2,1,1) = [\ Y_1(k)\quad Y_1(k+1)\quad Y_2(k)\quad U(k)\ ]$$

We choose k=0 and N=9, R(2,1,1) is a 4 × 9 matrix given by

$$R(2,1,1) = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 1 & 0 & 0 & 4 \\ 0 & 2 & 2 & 5 \\ 2 & 5 & -1 & -5 \\ 5 & -3 & 4 & -12 \\ -3 & 5 & -9 & 15 \\ 5 & -25 & 10 & 50 \\ -25 & 56 & 8 & 10 \end{bmatrix}$$

since

$$S(2,1,1) = R^T(2,1,1)R(2,1,1)$$

we find that

$$S(2,1,1) = \begin{bmatrix} 689 & -1545 & -105 & -111 \\ -1545 & 3684 & 140 & -742 \\ -105 & 140 & 266 & 412 \\ -111 & -742 & 412 & 3040 \end{bmatrix}$$

is a nonsingular matrix. The first singular matrix found is S(3,3,2) so that $n_2 = 3 - 1 = 2$ this matrix is deleted from the sequence that continues as S(3,2,3), S(4,2,3), ... The second singular matrix found is S(4,2,3) so that $n_1 = n_m = 3$ since m=2 the structural identification is terminated.

4.2.2 In the presence of noise

In the presence of noise, the procedure fails since in practice any vector is linearly independent of th previous ones. Moreover, the introduction in the test, for instance, of an acceptability threshold is an easy source of erroneous evaluations.

An equivalent way of testing, as shown in (4.20), is as known the singularity test of matrices. This approach is particularly useful in the case of noise, since it allows one to perform an averaging of the data before the singularity test is made.

It will be assumed here that the input and output sequences are corrupted by an additive noise with zero-mean, the noisy components of the input and output vectors will be denoted:

$$y_i^*(k) = y_i(k) + d(y_i(k)) \tag{4.21}$$

$$u_i^*(k) = u_i(k) + d(u_i(k)) \tag{4.22}$$

The origin of the bias in the structural identification can, however, be easily detected and eliminated, if the statistics of the noise are known. The elements of the matrix $S = R^T R$ are essentially correlations or cross-correlations. Because of the assumption that the covariance matrix of the additive noise, which is independent of the input and output sequences, is diagonal, the following statements stand

$$\lim_{N \to \infty} \frac{1}{N} S^* = \lim_{N \to \infty} \frac{1}{N} S + N(d_s) \tag{4.23}$$

Where N(d_s) is the covariance matrix of the noise vector

$$d_s = \{\ d(y_1(k)) \ \ldots\ldots\ d(y_1(k + n_{s1}))\ \mid\ \ldots\ \mid\ d(u_1(k))\ \ldots\ldots\ d(u_r(k + n_{s-1}))\ \} \tag{4.24}$$

If the statistics of the noise are known it is therefore possible, for N large enough, to obtain a consistent estimation of the quantities S by means of the expression

$$S = S^{*} - N \times N(d_s) \tag{4.25}$$

The identification of the system structure will be performed on the sequence of matrices $S(\mu_1, \mu_2, \ldots, \mu_{m+r})$

When the same amount of uncorrelated zero-mean noise is added to the input-output sequences and nonoverlapping sets of data are used, then $N(d_s) = \sigma^2 I$ where σ^2 is the variance of the noise so that

$$\lim_{N \to \infty} \frac{1}{N} S_s^{*} = \lim_{N \to \infty} \frac{1}{N} S_s + \sigma^2 I \tag{4.26}$$

It is also important to note that, an estimate of σ^2 is given by the least eigenvalue of the symmetrical matrix $S(\mu, \mu, \ldots, \mu)$ for $\mu > n_M$ so that a consistent estimate of the structure can be obtained. The same technique can be applied when different amounts of noise are present of the various inputs and outputs by performing a previous scaling on the data; the ratio of the different noises must however be known.

It is much easier to determine the structure of the system if the input output observations are free from noise corruption. In the case of noisy observations, the algorithm given here is, still far from being satisfactory.

4.2 Characterization of multivariable unknown linear systems

Consider a linear multivariable system with m-dimensional control input { U(k)} and r-dimensional output { Y(k) } described by the following irreducible matrix-fraction description.

$$A(z)Y(z) = B(z)U(z)$$

Where A(z) and B(z) are $(m \times m)$ and $(m \times r)$ left coprime polynomial matrices in z. Given the set of structural indices $n_1, n_2, \ldots, n_m$, these representations allow us to describe the whole class of completely observable systems. The matrix A(z) must be row reduced, so that

$$\sum_{l=1}^{m} n_l = deg\, det\, A(z) \tag{4.27}$$

The pulse transfer matrix given by

$$G(z) = A^{-1}(z)B(z) \tag{4.28}$$

is assumed to be proper, therefore

$$deg\ of\ i^{th}\ column\ of\ B(z) \leq deg\ of\ i^{th}\ column\ of\ A(z) \quad for\ i = 1, \ldots, m$$

Define $n = \sum_{l=1}^{m} n_l$ and $n_m \leq \ldots \leq n_2 \leq n_1$

We shall show how to construct an observable realization of order equal to degdetA(z). To display the row degrees more explicitly, we can write

$$A(z) = H(z)A_h + L(z)A_l \tag{4.29}$$

$$B(z) = H(z)B_h + L(z)B_l \tag{4.30}$$

where

$$H(z) = \begin{bmatrix} z^{n_1} & 0 & \cdots\cdots & 0 \\ 0 & z^{n_2} & \cdots\cdots & 0 \\ : & : & : & : \\ : & : & : & : \\ 0 & 0 & \cdots\cdots & z^{n_m} \end{bmatrix} \quad L(z) = \begin{bmatrix} 1 \; z \; \cdots \; z^{n_1 - 1} & 0 & 0 \\ 0 & 1 \; z \; \cdots \; z^{n_2 - 1} & : \\ : & : & : \\ : & : & : \\ 0 & 0 & 1 \; z \; \cdots \; z^{n_m - 1} \end{bmatrix}$$

The n_i are the row degrees of A(z), and A_h is the highest row degree coefficient matrix of A(z), the term $L(z)A_l$ accounts for the reaining terms of lower row degree terms of A(z), with A_l a matrix of coefficients. The matrices A_h, A_l, B_h and B_l are given by:

$$A_l \; (n \times m) = \begin{bmatrix} A_1 \\ A_2 \\ : \\ : \\ A_m \end{bmatrix} \quad where \quad A_i \; (n_i \times m) = \begin{bmatrix} a_{i1}^0 & a_{i2}^0 & \cdots\cdots & a_{im}^0 \\ a_{i1}^1 & a_{i2}^1 & \cdots\cdots & a_{im}^1 \\ : & : & : & : \\ : & : & : & : \\ a_{i1}^{n_i - 1} & a_{i2}^{n_i - 1} & \cdots\cdots & a_{im}^{n_i - 1} \end{bmatrix}$$

$$A_h \; (m \times m) = \begin{bmatrix} 1 & a_{12}^{n_1} & \cdots\cdots & a_{1m}^{n_1} \\ a_{21}^{n_2} & 1 & \cdots\cdots & a_{2m}^{n_2} \\ : & : & : & : \\ : & : & : & : \\ a_{m1}^{n_m} & a_{m2}^{n_m} & \cdots\cdots & 1 \end{bmatrix}$$

$$B_l\ (n \times r) = \begin{bmatrix} B_1 \\ B_2 \\ : \\ : \\ B_m \end{bmatrix} \quad where \quad B_i\ (n_i \times m) = \begin{bmatrix} b_{i1}^{0} & b_{i2}^{0} & \ldots\ldots & b_{im}^{0} \\ b_{i1}^{1} & b_{i2}^{1} & \ldots\ldots & b_{im}^{1} \\ : & : & : & : \\ : & : & : & : \\ b_{i1}^{n_i-1} & b_{i2}^{n_i-1} & \ldots\ldots & b_{im}^{n_i-1} \end{bmatrix}$$

$$B_h\ (m \times m) = \begin{bmatrix} b_{11}^{n_1} & b_{12}^{n_1} & \ldots\ldots & b_{1r}^{n_1} \\ b_{21}^{n_2} & b_{22}^{n_2} & \ldots\ldots & b_{2r}^{n_2} \\ : & : & : & : \\ : & : & : & : \\ b_{m1}^{n_m} & b_{m2}^{n_m} & \ldots\ldots & b_{mr}^{n_m} \end{bmatrix}$$

Now consider the matrices E_h *and* E_l

$$E_l\ (n \times n) = \begin{bmatrix} E_1 & 0 & \ldots\ldots & 0 \\ 0 & E_2 & \ldots\ldots & 0 \\ : & : & : & : \\ : & : & : & : \\ 0 & 0 & \ldots\ldots & E_m \end{bmatrix} \quad Where \quad E_i\ (i \times i) = \begin{bmatrix} 0 & 0 & \ldots\ldots & 0 & 0 \\ 1 & 0 & \ldots\ldots & 0 & 0 \\ 0 & 1 & \ldots\ldots & 0 & 0 \\ : & : & : & : & : \\ : & : & : & : & : \\ 0 & 0 & \ldots\ldots & 1 & 0 \end{bmatrix}$$

$$E_h\ (m \times n) = \begin{bmatrix} 0 & 0 & \ldots & 1 & \ldots & 0 & \ldots & \ldots & 0 \\ 0 & 0 & \ldots & 0 & \ldots & 1 & \ldots & \ldots & 0 \\ : & : & : & : & : & : & : & : & : \\ : & : & : & : & : & : & : & : & : \\ 0 & 0 & \ldots & 0 & \ldots & 0 & ..1 & \ldots & 0 \\ : & : & : & : & : & : & : & : & : \\ 0 & 0 & \ldots & 0 & \ldots & 0 & \ldots & \ldots & 1 \end{bmatrix}$$

According to the hypotheses mentioned, We can state that the system under study can be written in the following observable canonical form:

$$X(k+1) = E_l X(k) - A_l Y(k) + B_l U(k) \tag{4.31}$$

$$0 = E_h X(k) - A_h Y(k) + B_h U(k) \tag{4.32}$$

Where X(k) is $(n \times 1)$ state vector

Proof: Consider the equation (4.26)

$$A(z)Y(z) = B(z)U(z)$$

As shown in equations (4.29), (4.30) we have

$$A(z) = H(z)A_h + L(z)A_l$$

$$B(z) = H(z)B_h + L(z)B_l$$

Substituting into equation (4.26), we get

$$H(z)A_h Y(z) + L(z)A_l Y(z) = H(z)B_h U(z) + L(z)B_l U(z) \tag{4.33}$$

Now starting from the set of equations given in (4.31) and (4.32) , let us apply the z-transform theorem, we get

$$X(z) = -(zI - E_l)^{-1}A_l Y(z) + (zI - E_l)^{-1}B_l U(z) \tag{4.34}$$

$$0 = E_h X(z) - A_h Y(z) + B_h U(z) \tag{4.35}$$

Subtituting equation (4.34) into equation (4.35), it leads to

$$E_h(zI - E_l)^{-1}A_l Y(z) + A_h Y(z) = E_h(zI - E_l)^{-1}B_l Y(z) + B_h Y(z) \tag{4.36}$$

It can be checked by direct calculation that

$$E_h(zI - E_l)^{-1} = H^{-1}(z)L(z) \tag{4.37}$$

By substituting , it leads to

$$H^{-1}(z)L(z)A_lY(z) + A_hY(z) = H^{-1}(z)L(z)B_lU(z) + B_hU(z)$$

Multiplying both members of this equation by H(z), we get

$$L(z)A_lY(z) + H(z)A_hY(z) = L(z)A_lU(z) + H(z)B_hU(z) \tag{4.38}$$

Hence, the system can be written in the state space form given by equations (4.31), (4.32) The main property of the form given by equations (4.31), (4.32) is in the fact that the matrices E_l and E_h that multiply the state convey only structural information about the system, as they contain only 0 and 1's. All the parametric information is rejected to the matrices A_l, A_h, B_l and B_h

Now let us consider the canonical representation given by

$$X(k + 1) = E_lX(k) - A_lY(k) + B_lU(k)$$

$$0 = E_hX(k) - A_hY(k) + B_hU(k)$$

According to the structure of A_h, we can write

$$A_h = I_m + A_{ph} \tag{4.39}$$

Where $I_m = (m \times m)$ *identity matrix*

$$X(k + 1) = E_lX(k) - A_lY(k) + B_lU(k) \tag{4.40}$$

$$Y(k) = E_hX(k) - A_{ph}Y(k) + B_hU(k) \tag{4.41}$$

It is immediately noted, that in our case all unknown parameters or state variables appear linearly multiplied by either the external variables that are available in the record of inputs and outputs, or by matrices that are only composed of zeroes and ones. This very important property enables us to construct a linear estimator for all unknown variables and parameters of the system.

Let us construct the total parameters vector $P = [A^T \mid B^T]^T$ Where A stands for the vector formed by all non null parameters Of A_h *and* A_l and B stands for all parameters of B_h *and* B_l

Remark: If the rational matrix G(z) is strictly proper, then $B_h = 0$

As all parameters in P are supposed unknown but constant , we shall formulate that by

$$P(k+1) = P(k) \tag{4.42}$$

with those considerations, and as shown for the SISO case, the equations can be rewritten in the following form

$$X(k+1) = E_l X(k) + G[Y(k),U(k)]P(k) \tag{4.43}$$

$$Y(k) = E_h X(k) + N[Y(k),U(k)]P(k) \tag{4.44}$$

Where the matrices G and N depend only on the values of the observed external variables Y(k) and U(k) at time k and their dependance and structure are fixed after the ordering of the unknown parameters in P.

Using the extended vector $S^T(k) = \left[X^T(k) \mid P^T \right]$, we may finally represent the system in the following way

$$S(k+1) = F[Y(k),U(k)]S(k) \tag{4.45}$$

$$Y(k) = H[Y(k),U(k)]S(k) \tag{4.46}$$

Where

$$F[Y(k),U(k)] = \begin{bmatrix} E_l & G(Y,U) \\ 0 & I \end{bmatrix} \qquad H[Y(k),U(k)] = \begin{bmatrix} E_h & \vert & N(Y,U) \end{bmatrix}$$

The final augmented system is a time-varying linear system. We may state our problem as follows: given a record of the external variables, find an optimal estimator for the unknown parameters of the model. It is known that such problem must be imbedded in a nonlinear filtering one.

Illustration of this canonical form

Consider a 5-th order two-input two-output linear discrete-time system, the system is represented by the form given in (4.2)

$$A(z)Y(z) = B(z)U(z)$$

the structural parameters of the system are $n_1 = 3$ and $n_2 = 2$, hence, the polynomial matrices A(z) and B(z) are given by:

$$A(z) = \begin{bmatrix} z^3 + a_{11}^2 z^2 + a_{11}^1 z + a_{11}^0 & a_{12}^3 z^3 + a_{12}^2 z^2 + a_{12}^1 z + a_{12}^0 \\ a_{21}^2 z^2 + a_{21}^1 z + a_{21}^0 & z^2 + a_{22}^1 z + a_{22}^0 \end{bmatrix}$$

$$B(z) = \begin{bmatrix} b_{11}^3 z^3 + b_{11}^2 z^2 + b_{11}^1 z + b_{11}^0 & b_{12}^3 z^3 + b_{12}^2 z^2 + b_{12}^1 z + b_{12}^0 \\ b_{21}^2 z^2 + b_{21}^1 z + b_{21}^0 & b_{22}^2 z^2 + b_{22}^1 z + b_{22}^0 \end{bmatrix}$$

As shown in equations (4.29), (4.30) we have

$$A(z) = H(z)A_h + L(z)A_l$$

$$B(z) = H(z)B_h + L(z)B_l$$

Where

$$H(z) = \begin{bmatrix} z^3 & 0 \\ 0 & z^2 \end{bmatrix} \quad L(z) = \begin{bmatrix} 1 & z & z^2 & 0 & 0 \\ 0 & 0 & 0 & 1 & z \end{bmatrix}$$

Therefore,

$$A_h(2 \times 2) = \begin{bmatrix} 1 & a_{21}^3 \\ a_{21}^2 & 1 \end{bmatrix} \quad B_h(2 \times 2) = \begin{bmatrix} b_{11}^3 & b_{12}^3 \\ b_{21}^2 & b_{22}^2 \end{bmatrix}$$

$$A_l(5 \times 2) = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \text{ where } A_1(3 \times 2) = \begin{bmatrix} a_{11}^0 & a_{12}^0 \\ a_{11}^1 & a_{12}^1 \\ a_{11}^2 & a_{12}^2 \end{bmatrix} \quad A_2(2 \times 2) = \begin{bmatrix} a_{21}^0 & a_{22}^0 \\ a_{21}^1 & a_{22}^1 \end{bmatrix}$$

$$B_l(5 \times 2) = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \text{ where } A_1(3 \times 2) = \begin{bmatrix} b_{11}^0 & b_{12}^0 \\ b_{11}^1 & b_{12}^1 \\ b_{11}^2 & b_{12}^2 \end{bmatrix} \quad A_2(2 \times 2) = \begin{bmatrix} b_{21}^0 & b_{22}^0 \\ b_{21}^1 & b_{22}^1 \end{bmatrix}$$

Now consider the matrices E_l and E_h

$$E_l(5 \times 5) = \begin{bmatrix} E_1 & 0 \\ 0 & E_2 \end{bmatrix} \text{ where } E_1(3 \times 3) = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad E_2(2 \times 2) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

$$E_h(2 \times 5) = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Therefore, our system can be written in the following form

$$X(k+1) = E_l X(k) - A_l Y(k) + B_l U(k)$$

$$0 = E_h X(k) - A_h Y(k) + B_h U(k)$$

As mentioned in equation (4.37), it can be checked by direct calculation that:

$$E_h(zI - E_l)^{-1} = H^{-1}(z)L(z)$$

4.4 Identification algorithm

Measurements of the system input { U(k) } and output { Y(k) } become available in real time, and the aim is to estimate the coefficients a_{ij}^k and b_{ij}^k using the measurements. If (4.1) is taken as the equation describing the system and the measurements are supposed to be noise-free, with sufficient measurements, the coefficients can be found by solving a set of linear equations. It is, however, more realistic to model the states of the system as being subject to random perturbations and to model the measurements as being noisy. So, we must introduce more structure into the problem, and derive as in chapter 1 an identification algorithm that involves a standard discrete-time Kalman filter.

Consider the random signals Y(k), S(k) described by the following equations, which include the processes generating the noises and perturbations

$$S(k+1) = F(k)S(k) + V(k) \tag{4.47}$$

$$Y(k) = H(k)S(k) + W(k) \tag{4.48}$$

Where:

S(k) is the extended state vector of the system

Y(k) the measured output signal

If the noise terms {V(k),W(k)}, as well as the initial condition S(0), are gaussian, then $\hat{S}(0)$ is determined by the Kalman filter.

We assume that V(k) and W(k) are white gaussian sequences with correlation functions

$$E\{\begin{bmatrix} V(k) \\ W(k) \end{bmatrix}\begin{bmatrix} V(l) \\ W(l) \end{bmatrix}^T\} = \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \delta_{kl} \tag{4.49}$$

and more precisely, V(k) and W(k) are independent of all the past of the system V(k) and W(k) are also, assumed to be independent of the state S(k).

Kalman filtering can be applied to provide a technique for the identification of the coefficients in a scalar ARMA equation of the form (4.1).

The matrices F(k) and H(k) are available at each time k from the record of the input-output data. We assume also that the matrices Q and R are known, so are the statistical characteristics of the initial state S(0), that is to say its mean value and covariance.

At the time k, we dispose of the measurements vector

$$\mathbf{Y}^T(k) = [\ Y(0)\ Y(1)\ \dots\ Y(k)\ ] \tag{4.50}$$

and we seek to estimate the unknown vector

$$\mathbf{s}^T(k) = \left[S^T(0)\ S^T(1)\ \dots\ S^T(k)\ S^T(k+1)\right] \tag{4.51}$$

The assumptions of the problem allow us in theory to calculate the mean values, variances and covariances of S(k) and Y(k), hence, the possibility to apply the general formulas of the optimal linear estimation to obtain S(k) as a linear function of Y(k)

$$S(k+1) = F(k)S(k) + K(k)\left[\,Y(k) - H(k)S(k)\,\right] \tag{4.52}$$

The estimation problem can be stated as the choice of the optimal gain K(k) such that the reconstruction error is minimal. From a standard discrete Kalman filter an estimation algorithm is developed. let

S(k/k) = optimal filtered estimate of S(k)

S(k+1/k) = optimal predicted estimate of S(k)

If the optimal filtered estimate S(k/k) and the covariance matrix P(k/k) of the corresponding filtering error S(k/k) are known for some k, then the single-stage optimal predicted estimate for all admissible loss functions is given by the expression

$$\hat{S}(k+1/k) = F(k)\hat{S}(k/k) \tag{4.53}$$

The stochastic process { S(k+1/k), k=0,1,..} defined by the single-stage predicted error relation

$$\bar{S}(k+1/k) = S(k+1) - \hat{S}(k+1/k)$$

is a zero mean gauss-markov process whose covariance matrix is given by the relation

$$P(k+1/k) = F(k)P(k/k)F^{T}(k) + Q \tag{4.54}$$

The optimal filtered estimate S(k+1/k+1) is given by the recursive relation

$$\hat{S}(k+1/k+1) = F(k)S(k/k) + K(k+1)\left[y(k+1)\_H(k+1)F(k)\hat{S}(k/k)\right] \tag{4.55}$$

for k=0,1,.... where $\hat{S}(0/0)$ =S(0)

The "Kalman gain" is given by the following set of relations

$$K(k+1) = P(k+1/k)H^T(k+1)\left[H(k+1)P(k+1/k)H^T(k+1) + R\right]^{-1} \tag{4.56}$$

$$P(k+1/k) = F(k)P(k/k)F^T(k) + Q \tag{4.57}$$

$$P(k+1/k+1) = [I - K(k+1)H(k+1)]P(k+1/k) \tag{4.58}$$

For k=0,1,.... and P(0/0)=P(0)

4.5 Example of identifiction

Consider a third order discrete-time linear system, whose structural parameters are supposed to be known a priori. From a simulated input/output data we attempt to estimate the parameters of the matrix fraction description of the system. To reduce the number of parameters we will assume that the matrix transfer function is strictly proper, we also assume that the monic polynomials have the highest degree, i.e. the diagonal polynomials of A(z) have a higher than any other polynomial in their row.

$$\begin{bmatrix} z^2 + a_{11}^1 z + a_{11}^0 & a_{12}^1 z + a_{12}^0 \\ a_{21}^0 & z + a_{22}^0 \end{bmatrix} Y(z) = \begin{bmatrix} b_{11}^1 z + b_{11}^0 & b_{12}^1 z + b_{12}^0 \\ b_{21}^0 & b_{22}^0 \end{bmatrix} U(z) \tag{E.1}$$

Our system can be written in form

$$A(z)Y(z) = B(z)U(z) \tag{E.2}$$

with the following values of the parameters

$$a_{11}^0 = 0.27 \qquad a_{11}^1 = 1.20 \qquad a_{12}^0 = 1.0$$

$$a_{12}^1 = 2.0 \qquad a_{22}^0 = 0.70 \qquad a_{21}^0 = 0.0$$

$$b_{11}^0 = -1.0 \qquad b_{11}^1 = 3.0 \qquad b_{12}^0 = -1.0$$

$$b_{12}^1 = 1.0 \qquad b_{21}^0 = 2.0 \qquad b_{22}^0 = -1.0$$

Using the observable-form realization we write

$$A(z) = H(z)A_h + L(z)A_l \tag{E.3}$$

$$B(z) = L(z)B_l \tag{E.4}$$

Where

$$H(z) = \begin{bmatrix} z^2 & 0 \\ 0 & z \end{bmatrix} \quad L(z) = \begin{bmatrix} 1 & z & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

with

$$A_l = \begin{bmatrix} a_{11}^0 & a_{12}^0 \\ a_{11}^1 & a_{12}^1 \\ a_{21}^0 & a_{22}^0 \end{bmatrix} \quad B_l = \begin{bmatrix} b_{11}^0 & b_{12}^0 \\ b_{11}^1 & b_{12}^1 \\ b_{21}^0 & b_{22}^0 \end{bmatrix} A_h = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Using the new approach developed earlier, the system can be expressed in the following form

$$X(k+1) = E_l X(k) - A_l Y(k) + B_l U(k) \tag{E.5}$$

$$0 = E_l X(k) - A_h Y(k) \tag{E.6}$$

Where

$$E_l = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ and } E_h = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

According to the structure of A_h our system may be rewritten as

$$X(k+1) = E_l X(k) - A_l Y(k) + B_l U(k) \tag{E.7}$$

$$Y(k) = E_h X(k) \tag{E.8}$$

let the parameters vectors be

$$A^T = \left[\ a_{11}^0\ a_{11}^1\ a_{21}^0\ a_{12}^0\ a_{12}^1\ a_{22}^0\ \right] \quad B^T = \left[\ b_{11}^0\ b_{11}^1\ b_{21}^0\ b_{12}^0\ b_{12}^1\ b_{22}^0\ \right]$$

Using the total parameters vector $P^T = \left[\ A^T \mid B^T\ \right]$ we can write

$$X(k+1) = E_l X(k) + G(k)P(k) \tag{E.9}$$

$$P(k+1) = P(k) \tag{E.10}$$

$$Y(k) = E_h X(k) \tag{E.11}$$

Where

$$G(k) = \left[\ -y_1 I_3 \mid -y_2 I_3 \mid u_1 I_3 \mid u_2 I_3\ \right]$$

Using the extended vector $S^T(k) = \left[\ X^T(k) \mid P^T(k)\ \right]$ we may express the system

$$S(k+1) = F(k)S(k) \tag{E.12}$$

$$Y(k) = H(k)S(k) \tag{E.13}$$

where

$$F(k) = \begin{bmatrix} E_l & G(k) \\ 0 & I_{12} \end{bmatrix} \quad and \quad H(k) = [\ E_h \ \ 0\ ]$$

It will be more realistic to consider the system as being subject to random perturbations and the measurements as being noisy, so, we reformulate our system as follows

$$S(k+1) = F(k)S(k) + V(k) \tag{E.14}$$

$$Y(k) = H(k)S(k) + w(k) \tag{E.15}$$

Where the noises V(k) and w(k) are white gaussian noises, assumed to be of zero-mean and uncorrelated.

A unit step corrupted by an additive zero-mean, white gaussian noise was used to excite the system. From the simulated noise contaminated input-output data, the proposed approach was used to estimate the parameters.

The following algorithm was used

$$M(k) = F(k-1)P(k-1)F^T(k) + Q \tag{E.16}$$

$$K(k) = M(k)H^T(k)\left[\ H(k)M(k)H^T(k) + R\ \right]^{-1} \tag{E.17}$$

$$K(k) = [\ I - K(k)H(k)\ ]M(k) \tag{E.18}$$

and

$$\hat{S}(k) = F(k-1)\hat{S}(k-1) + K(k)\Big[\; Y(k) - H(k)F(k-1)\hat{S}(k-1) \;\Big] \qquad (E.19)$$

The initial state is assumed to be null and known, the input and output noises are independent normalized gaussian white noises, with known covariance matrices

$$S(0) = 0 \quad \text{and} \quad P(0) = 100I$$

$$Q(15 \times 15) = \begin{bmatrix} Q_s & 0 \\ 0 & 0 \end{bmatrix} \qquad Q_s(3 \times 3) = diag(.001,.001,.001)$$

$$R(2 \times 2) = diag(.001,.001)$$

The same noise levels ($\sigma^2 = 0.10$, $\sigma^2 = 0.30$) were used in the simulated example given here. The first 30 parameters estimates are given in the figures at the end of the chapter. Only the estimates of the parameters of the characteristic polynomial are shown.

4.6 Discussion

The application of the identification algorithm to a simulated 3rd order 2×2 discrete-time system led to the same results obtained in the preceding chapters. The structural indices were known a priori, hence, only the parameters estimation step was performed. The computed estimates were found to be fairly good and accurate. Only 12 iterations were needed to obtain reasonable estimates, which sustains the claim that the algorithm converges very fast, and requires less computing time than other methods to obtain good estimates. However, in a practical situation there would be no preliminary knowledge of the system's structure, and a structural identification algorithm would have to be applied, which might get quite involved mainly in the noisy case.

In chapter 2 an alternative to this problem was presented to the structural identification step. Only the order of the system is required, but more parameters are estimated. This method attempts to identify a system in its reduced state space representation where the order is minimal.

It should be noticed that an overestimation of the structural parameters does not deteriorate the performance of the algorithm. such overestimation can be easily corrected as it leads usually to zero pole cancellation. An underestimation of these parameters may lead to an unsatisfactory identification.

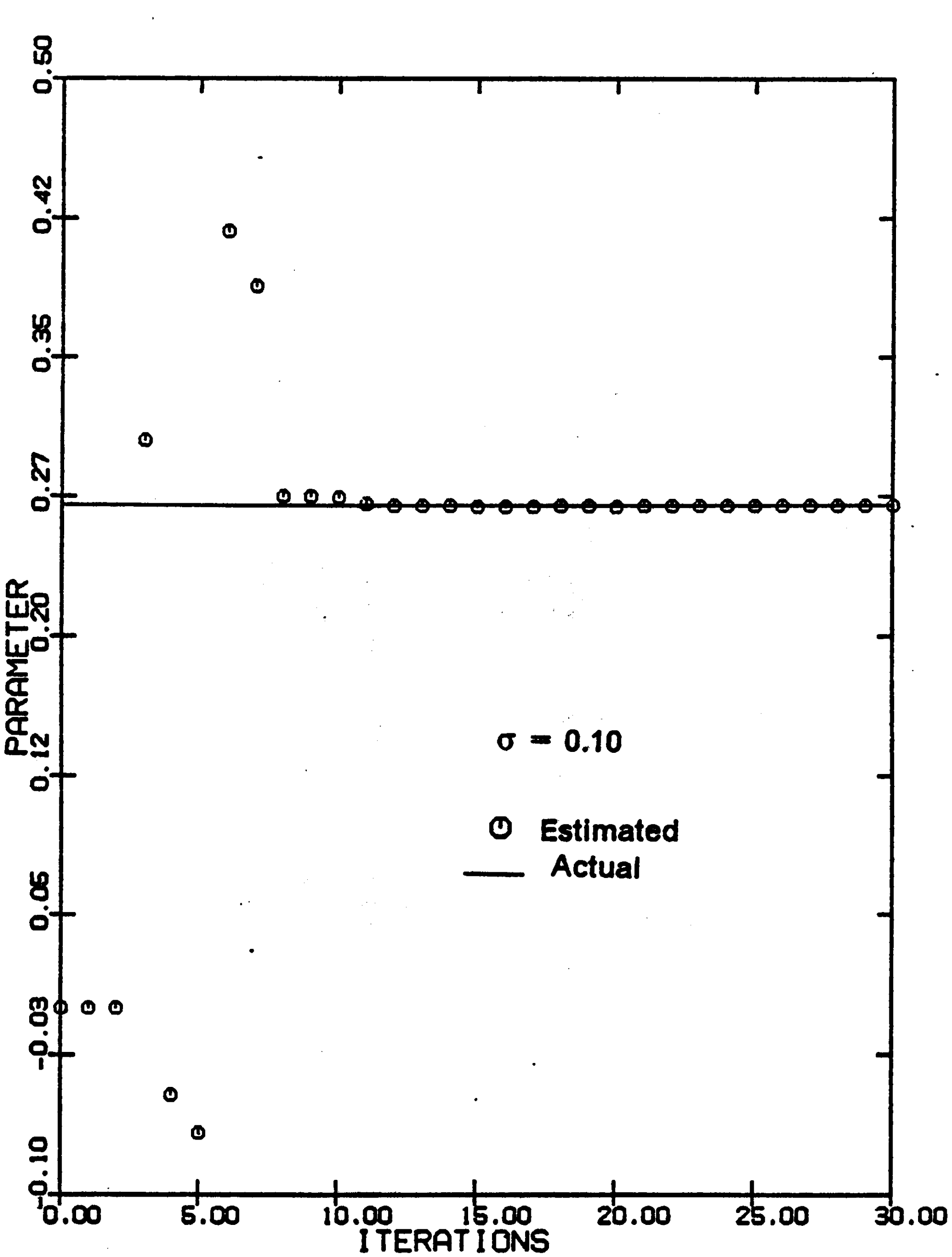

Figure 20. Estimated and actual values, in the case of low order noise level, of the parameter a11(0), for the MIMO case.

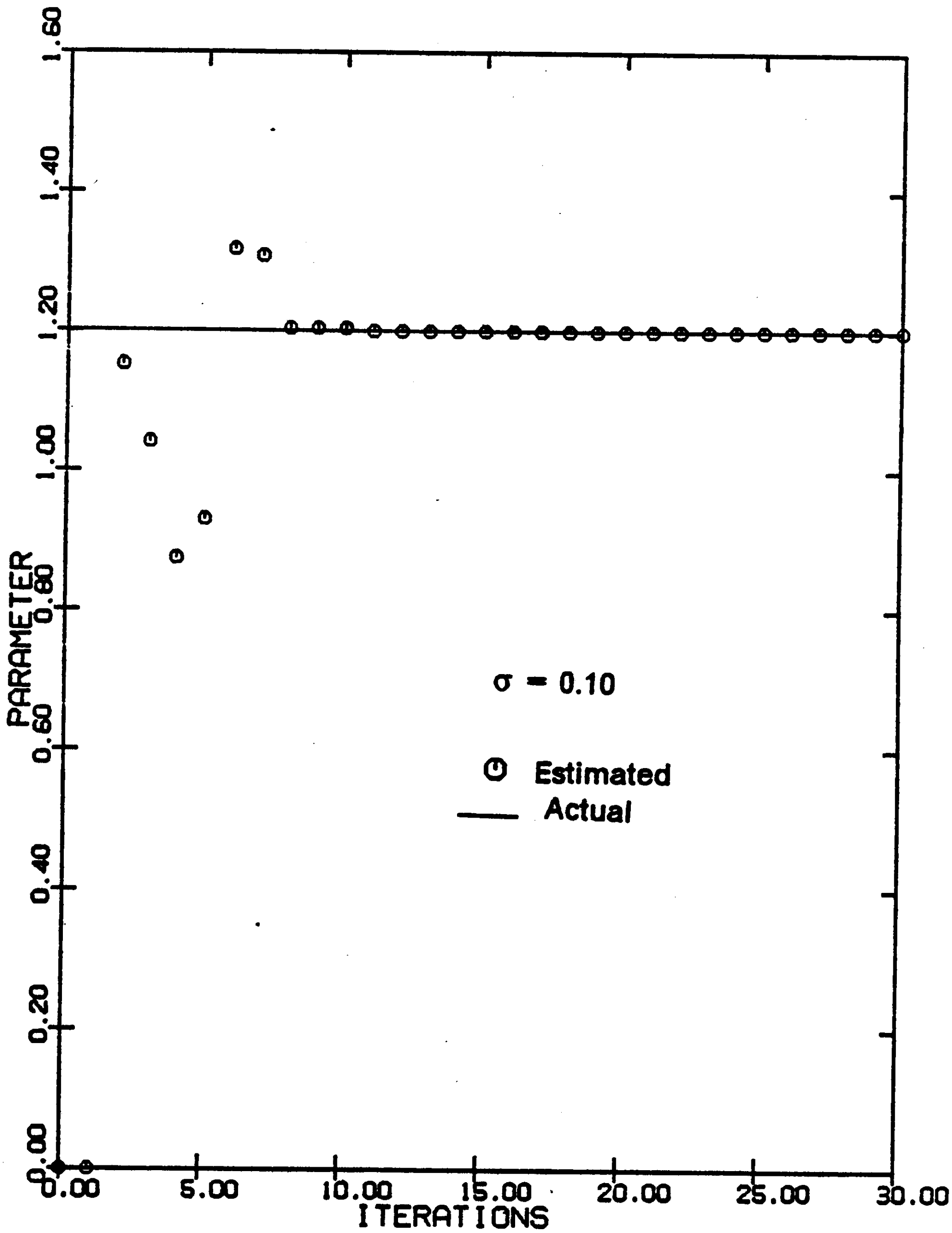

Figure 21. Estimated and actual values, in the case of low order noise level, of the parameter a11(1), for the MIMO case.

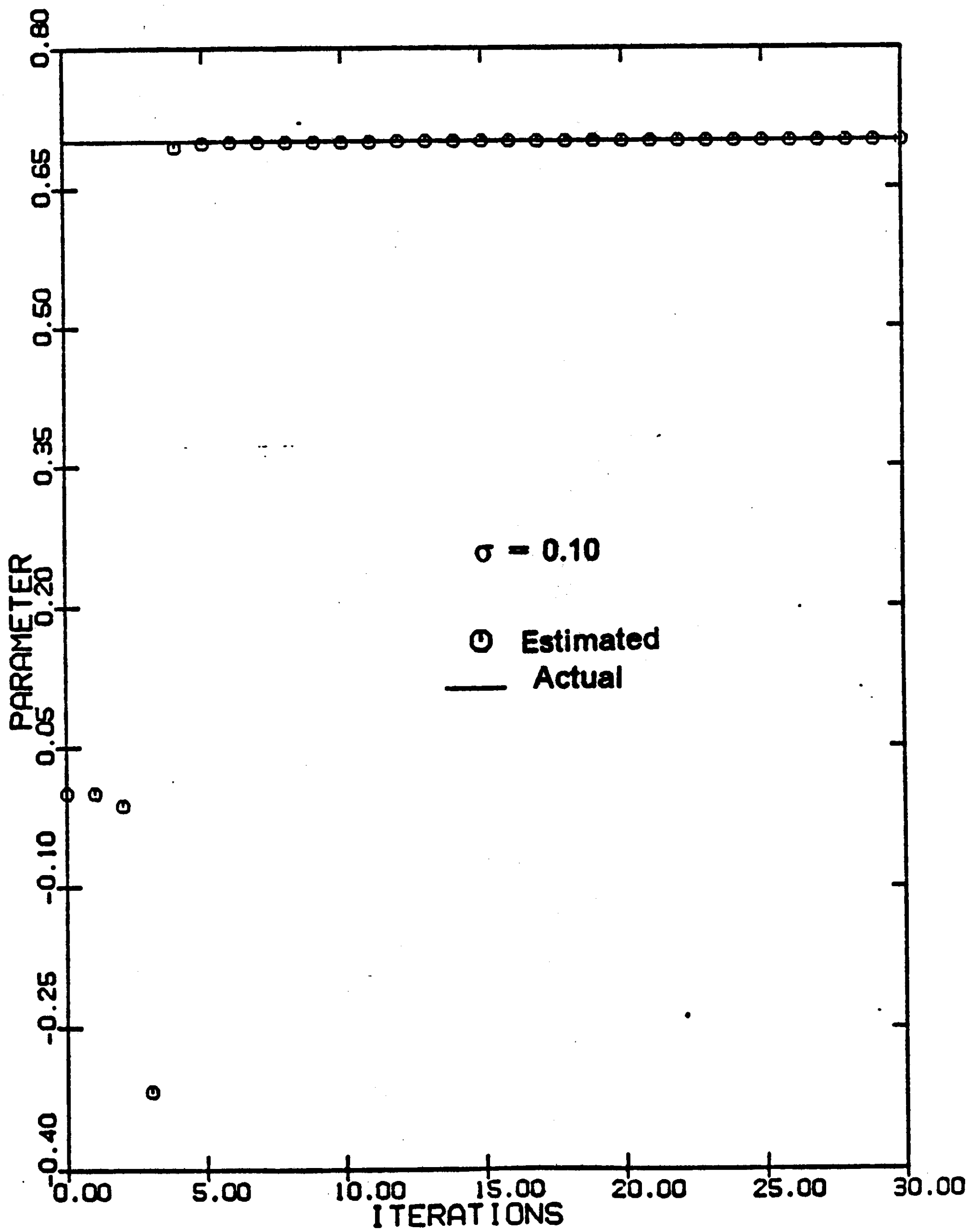

Figure 22. Estimated and actual values, in the case of low order noise level, of the parameter a22(0), for the MIMO case.

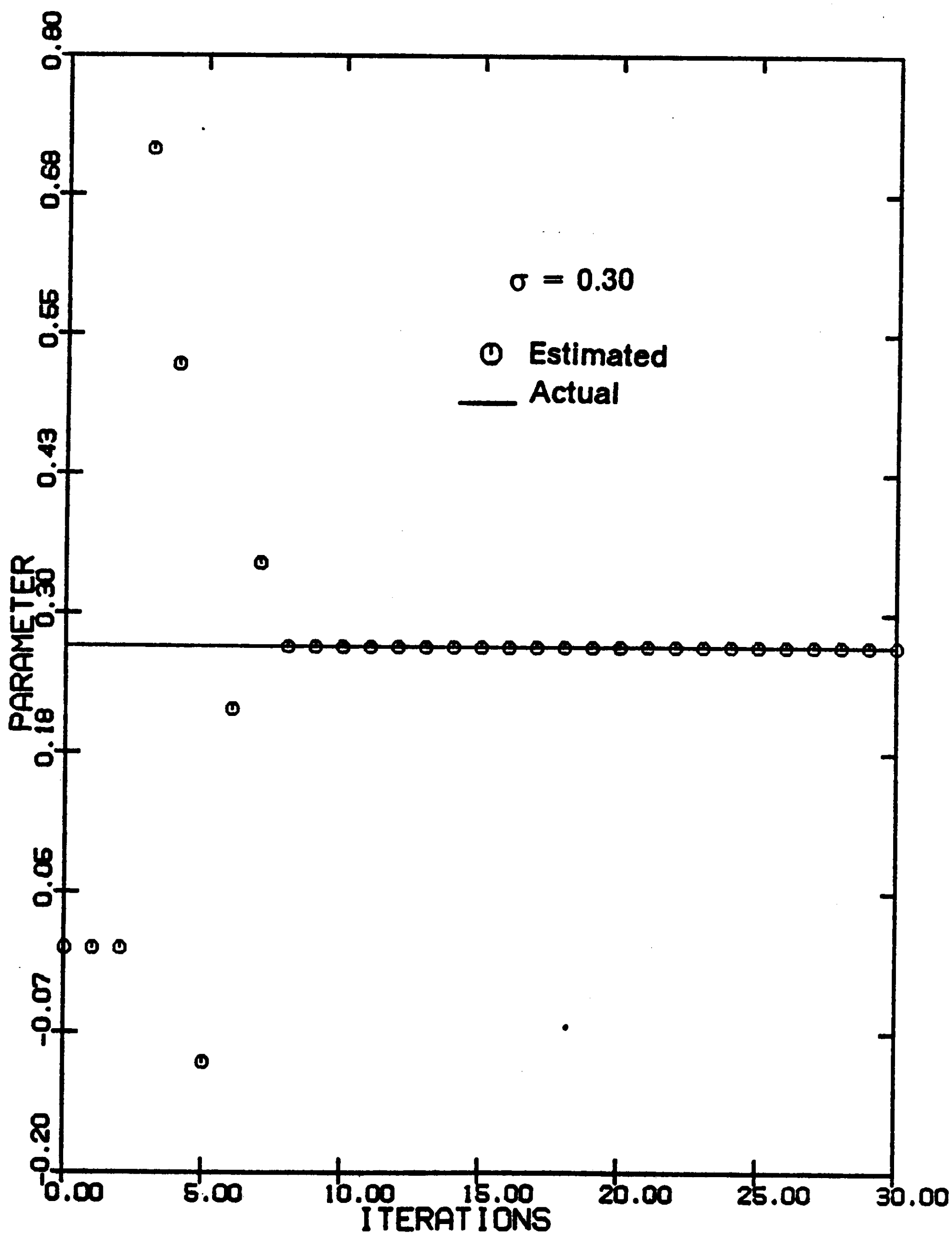

Figure 23. Estimated and actual values, in the case of high order noise level, of the parameter a11(0), for the MIMO case.

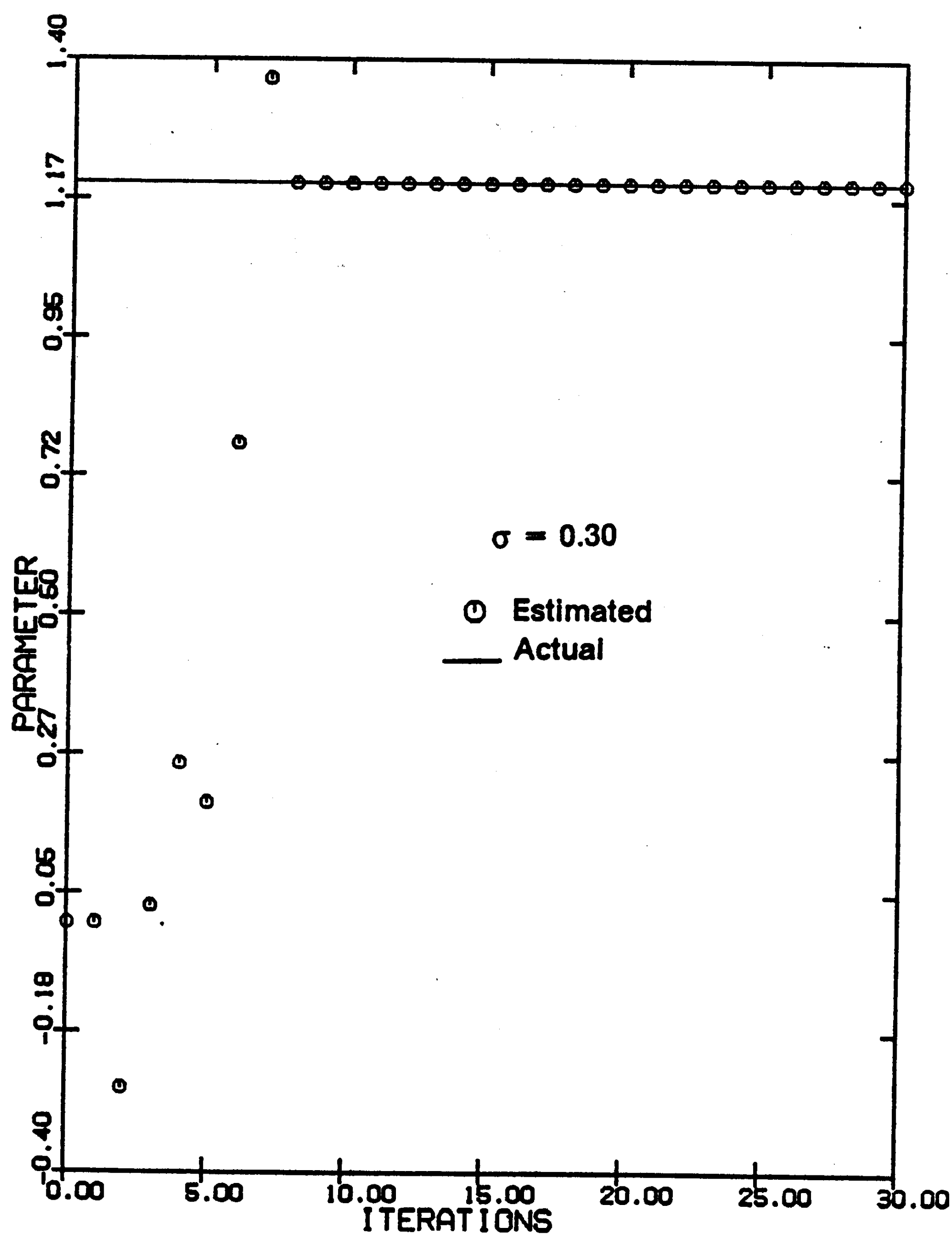

Figure 24. Estimated and actual values, in the case of high order noise level, of the parameter a11(1), for the MIMO case.

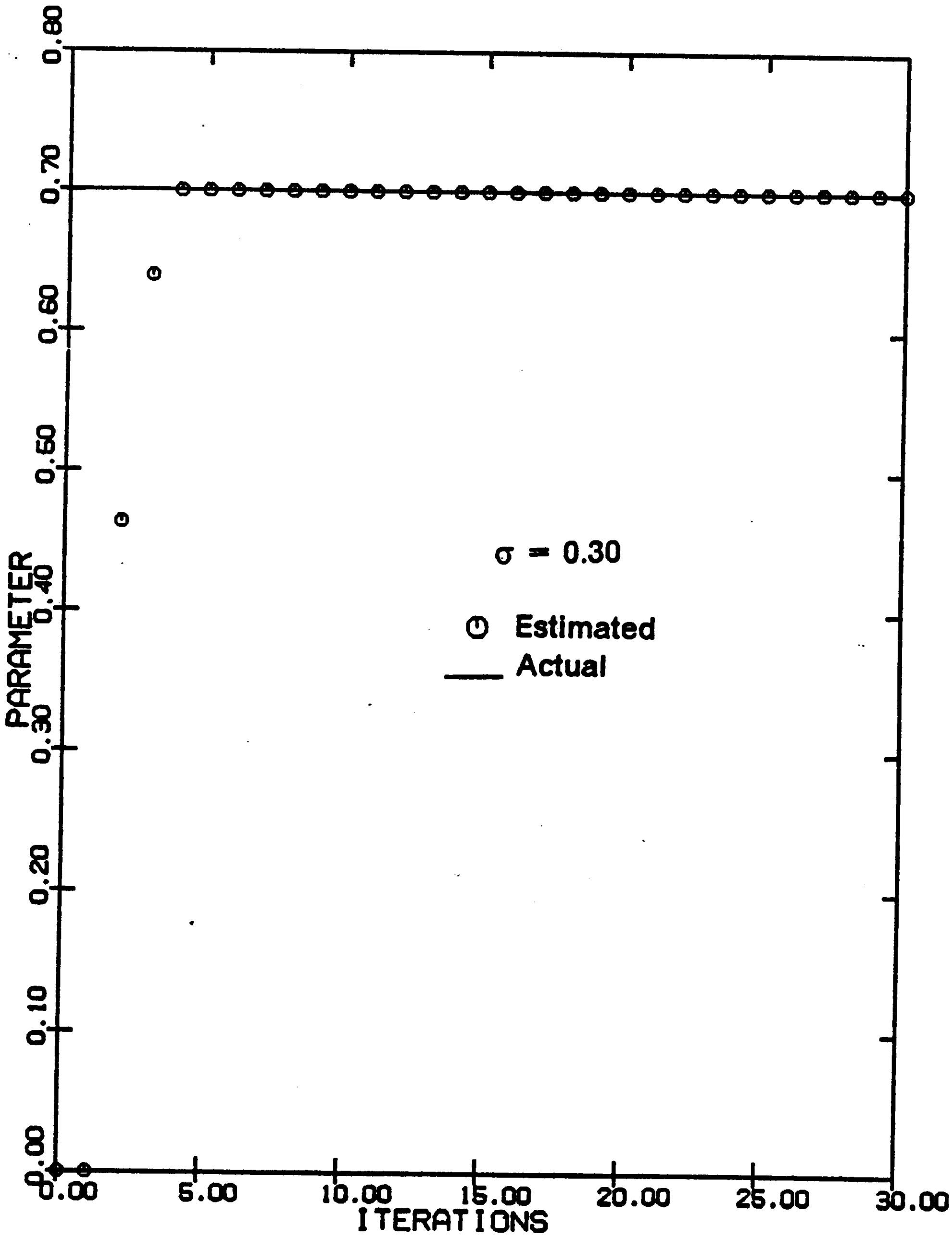

Figure 25. Estimated and actual values, in the case of high order noise level, of the parameter a22(0), for the MIMO case.

Conclusion

An algorithm has been proposed for on-line identification of the states and parameters of linear discrete-time multivariable systems. It is a one-stage algorithm, in which a special canonical form of the state equations provides a pseudo-linear representation where all unknown parameters and state variables appear linearly multiplied by either external variables, or by matrices that are only composed of zeroes and ones.

Deriving an augmented system, that involves augmenting the state variables of the system by adjoining to them the unknown parameter vectors and treating them as a part of the new state variable vector, unbiased estimates of the parameters are obtained using a discrete-time Kalman filter. The algorithm requires gaussian noise sequences with known statistics.

A nice feature of the given algorithm is that, it works directly on input/output data and leads to a state representation.

On the basis of the number of simulated examples treated, the identification approach appears to be working quite well. The convergence of the algorithm is very fast, the number of iterations needed to obtain conclusive convergence is very small, yet the estimation accuracy compares favorably with the performance of other more elaborate identification techniques. Actually, the minimal number of iterations necessary for convergence is equal to the total number of states and parameters to be estimated. By reducing the number of iterations, the method requires less computer time. It gives unbiased and consistent estimates even for high noise-to-signal ratios.

As to be expected, the examples considered illustrate the importance of using an input sufficiently rich in spectral components. Too, it is extremely important to use initial conditions on the augmented state and on the covariance matrices which are not in contradiction with the actual values of the parameters to be identified.

The identification method described here constitutes an effective on-line estimation technique which may be of value in practical situations. Though the preliminary results are encouraging, further work is needed to make the algorithm operational in a real time fashion and speed up the computational routines in the program as well as compare its performance in relation with other identification schemes.

In order to implement this estimation algorithm it is necessary to know the spectral components, which represent our knowledge about the perturbations and noises. To estimate the noise characteristics can be advantageously used in algorithms of the type [ ], where an on-line algorithm for the identification of the model is proposed, first the autocovariance matrices are estimated and then a minimal realization is derived. An investigation can be performed to check whether the Kalman filter is working optimally or not. If the Kalman filter is suboptimal, a technique as mentioned above can be developed to obtain asymptotically normal, unbiased and consistent estimates of Q and R.

Also, 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. From practical considerations, it is important to determine the adequacy of the model under the types of inputs the process is likely to encounter.

Bibliography

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

2. Astrom, K. J.:Introduction to Stochastic Control theory. Academic Press,1970.

3. Astrom, K. J. and Eykhoff, P. E.(1971),"System Identification - A survey," Automatica,vol.7, pp.123-162.

4. Beghelli, S. and Guidorzi, R. (1976),"A new input-output canonical form for multivariable systems," IEEE Trans. Automatic Control,AC-21, pp.692-696.

5. Bonivento, C. and Guidorzi, R.,"Parametric identification of linear multivariable systems," Preprints of the JACC, St Louis (1971).

6. Budin, M. A.(1971),"Minimal realization of discrete linear systems from input output observations,"IEEE Trans Automatic Control AC-16: PP 395-401

7. Chen, C. T.: Linear System Theory and Design. Holt, Rinehart and Winston,1970.

8. Chow, J. C.,"On the Structural Identification and Parameter Estimation in Linear Multivariable Systems," Preprint of the JACC, Colombus (1973)

9. Clarke, D. W. (1967), "Generalized least squares estimation of the parameters of a dynamical model,"Preprints of the First IFAC Symposium on Identification (Prague, Czechoslovakia).

10. El-Sherif, H., Sinha, N. K., 1977. A nonparametric stochastic approximation algorithm for on-line identification of multivariable systems. Proceedings 15th annual Allerton conference on circuit and system theory, pp. 231-236.

11. El-Sherif, H., Sinha, N. K.,"Identification and Modeling for Linear Multivariable Discrete-Time Systems: A Survey," Journal of Cybernitics, vol.9:43-71,1979.

12. Gertler, J. and Banyasz, C. (1974)," A recursive (on-line) maximum likelihood identification method," IEEE Trans. Aut Cont, vol.AC-19, pp.816-820

13. Guidorzi,R.(1975),"Canonical structures in the identification of multivariable systems," Automatica,Vol. 11, pp 361-374

14. Ho, B. L. and Kalman, R. E. (1966),"Effective constructionof linear state variable models from input-output functions," Regelungstechnik, vol. 14,pp.545-558

15. Hsia, T. C.(1976),"On least squares algorithm for system parameter identifiction," IEEE Trans. Aut Cont, vol.AC-21, pp.104-108.

16. Kailath, T.: Linear Systems. Prentice-hall, 1980.

17. Kalman, R. E." Mathematical description of linear systems,"SIAM. J. Cont., Ser.A.,vol.1,pp 152-192, 1963.

18. Kalman, R. E." irreducible realizations and the degree of rational matrix," SIAM. J. Cont.,vol.13,pp 520-544, 1965.

19. Kalman, R. E.,Falb P.L. and Arbib, M.A.: Tepics in Mathematical System Theory. New York: MacGraw-Hill,1966.

20. Kalman, R. E.,"On partial realization, transfer functions, and canonical forms," ACTA, polytechnica scandinavica, Ma 31,pp.9-32,1979.

21. Lee, R. C. K. : Optimal Estimation Identification and Control. MIT Press, Mass (1965)

22. Luenberger, D. G.(1969) "Canonical Forms for Linear Multivariable Systems," IEEE Trans Automatic Control.AC-12,pp.290-293.

23. Mayne, D. Q.,"Computational procedure for the minimal realization of transfer function matrices," Proc.IEE.,15.,pp1363-1368 (1968).

24. Mayne, D. Q.,"An elementary derivation of Rosenbrock's minimal realization algorithm," IEEE Trans. Aut Cont,vol.AC-18, pp.306-307,1973.

25. Mehra, R. K. (1971),"On-line identification of linear dynamic systems with applications to Kalman filtering," IEEE Trans. Aut Cont, vol.AC-16, pp.12-21.

26. Patel, R. V. and Munro, N.:Multivariable System Theory and Design. Pergamon Press,1982.

27. Popov, V. M.,"Invariant Description of Linear, Time-Invariant Controllable Systems," S.I.A.M.Journal of Control,vol.10,No.2, May1972.

28. Rosenbrock, H. H.: State Space and Multivariable Systems" New York, Wiley,1970.

29. Salut,G.,"Identifiabiblite' d'un systeme dynamique lineaire invariant," CR Acad.Sc Paris t.278 serie A,181,1974

30. Salut,G.,Aguilar-Martin, J.,Lefebvre,S.,"New Results on Optimal Joint Parameter and State Estimation of Linear Stochastic Systems," Journal of Dynamics Systems,Measurtement, and Control,Vol 102,pp 28-34,March 1980

31. Saridis, G. N.,"Comparison of Six On-Line Identification Algorithms," Automatica,vol.10,pp 69-79.

32. Sinha, N. K., and Kuszta, B.,"Modeling and Identification of Dynamic Systems", Van Nostrand Reinhold electrical computer science and engineering series, 1983.

33. Young, P. C. (1970), "An instrumental variable method for real-time identification of a noisy process," Automatica, vol.6,pp. 271-287.

34. VanLandingham, H. G.:Introduction to Digital Control Systems. Macmillan,1985.

35. Wolovich, W. A.: Linear Multivariable Systems. New York: Springer, 1974.

36. Wolovich, W.A.,"The determination of state-space representation for linear multivariable systems," Automatica, vol. 9, pp.97-106,1973.

37. Wolovich, W. A. and Falb, P. L.," On the structure of multivariable systems," SIAM. J. Control, vol.7.No3.1963.

38. Wong, K. Y. and Polak, E. (1967)," Identification of linear discrete time systems using the instrumental variable approach," IEEE Trans. Aut Cont, vol.AC-12, pp.707-718.

39. Woodside, C. M.,"Estimation of the order of linear systems," proc IFAC symp on identification and process parameter estimation, Prague (1970).

The vita has been removed from the scanned document