Mirecki, Anthony Vito

DECENTRALIZED IDENTIFICATION

City University of New York PH.D. 1986

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark __✓__.

1. Glossy photographs or pages _____
2. Colored illustrations, paper or print _____
3. Photographs with dark background _____
4. Illustrations are poor copy _____
5. Pages with black marks, not original copy __✓__
6. Print shows through as there is text on both sides of page _____
7. Indistinct, broken or small print on several pages __✓__
8. Print exceeds margin requirements _____
9. Tightly bound copy with print lost in spine _____
10. Computer printout pages with indistinct print _____
11. Page(s) _____ lacking when material received, and not available from school or author.
12. Page(s) _____ seem to be missing in numbering only as text follows.
13. Two pages numbered _____. Text follows.
14. Curling and wrinkled pages _____
15. Dissertation contains pages with print at a slant, filmed as received __✓__
16. Other _____

_____

_____

University
Microfilms
International

DECENTRALIZED IDENTIFICATION

by

ANTHONY MIRECKI

A dissertation submitted to the Graduate Faculty in Engineering in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

1966

This manuscript has been read and accepted for the Graduate Faculty in Engineering in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

1/3/86 — date

Frederick E. Thau — Chairman of Examining Committee

1/3/86 — date

Paul R. Kramer — Executive Officer

Professor Joseph Barba

Professor Shee-Ming Chen

Professor Vincent Del Toro

Professor George Kranc

Professor Frederick Thau

Supervisory Committee

The City University of New York

Abstract

DECENTRALIZED IDENTIFICATION

by

Anthony Mirecki

Adviser: Professor Frederick Tung

This thesis definitively solves the problem of system identification for linear time-invariant discrete systems when available measurements are "decentralized" as in the theory of decentralized control. In decentralized control a global, usually large scale, system has a total number of outputs. Instead of measuring and processing all outputs together in a "centralized" manner, nonoverlapping subsets of the outputs are measured and are available for feedback by subsystems. The decentralized pole-placement problem was formulated and solved in a 1973 paper by Wang and Davison. Since then other papers have appeared on this problem. In every paper it is given that the global state description is known to every subsystem. This thesis investigates the problem when subsystems do not know the global state-space description and use their decentralized measurements to try and identify it.

First, canonical forms for multi-input, multi-output (MIMO) systems are reviewed with greater emphasis on Luenberger canonical form. This canonical form is especially important because results have been published on

now to identify systems in this form from the outputs.

Next, from the decentralized set of outputs of a MIMO system a measurement equation, called the Decentralized Measurement Equation (DME), is derived for a system in any arbitraly state description with measurement noise and state noise multiplying unknown coefficients. The variation of Luenberger canonical form used for identification requires that the system be observable from a given set of outputs. This thesis solves the problem of system identification when the entire system is unobservable either from all outputs taken together or, more relevantly to the thesis title, when any decentralized subset of outputs is measurable.

Three second order systems are formulated in Chapter 4 and identified in accordance with the previously derived decentralized identification results. Chapter 5 again uses these results but now the system, a large flexible beam space structure, is very high order. The motion of the beam has been simulated under a NASA grant and is based on previous research on the Space Shuttle. Finally, Chapter 6 relates identification results to decentralized control results and also discusses further research.

ACKNOWLEDGEMENTS

First and foremost, I want to thank my mother for her moral and financial support over the years of this effort. I also want to thank the members of the Examining Committee and my adviser, Professor Frederick Thau, for reading of the manuscript and their contributions.

TABLE OF CONTENTS

ABSTRACT iv

Chapter Page

I. INTRODUCTION 1

II. LINEAR SYSTEM CANONICAL FORMS 12

2.1) SISO Canonical Forms 13

2.2) MIMO Canonical Forms 16

III. THE DECENTRALIZED MEASUREMENT EQUATION 49

3.1) The Identification Problem 51

3.2) Derivation of the DME 66

3.3) System Order Determination - Scalar Case 79

3.4) Derivation of the Vector DME 85

3.5) System Order and Structural Invariant Determination - Vector Case 94

IV. IDENTIFICATION OF THE DECENTRALIZED MEASUREMENT EQUATION 105

4.1) Least Squares System Identification . . . 105

4.2) Gradient Methods of System Identification 139

V. APPLICATION OF DECENTRALIZED IDENTIFICATION TO A FLEXIBLE BEAM STRUCTURE 166

VI. FUTURE RESEARCH 203

6.1) Combined Decentralized Identification and Control . . . . . 20..

6.2) Additional Future Research . . . . . 21[illegible]

APPENDIX A IDENTIFIABILITY RESULTS 220
A.1) N-identifiability 220
A.2) 1-identifiability 221
APPENDIX B MIMO SYSTEM IDENTIFICATION
IN ROW COMPANION FORM 226
BIBLIOGRAPHY 236

LIST OF FIGURES

Figure	Title	page
1.1	A two controller decentralized system	3
2.1	Inverted pendulum on moving cart	46
4.1-4.3	Recursive least squares convergences - observable example	126-128
4.4-4.6	Recursive least squares convergences - unobservable example	132-134
4.7-4.9	Recursive least squares convergences - stirred tank	136-140
4.10	Block diagram of deterministic gradient algorithm	145
4.11	Block diagram of stochastic gradient algorithm with measurement noise	154
4.12-4.14	Deterministic gradient convergences - observable example	166-170
4.15-4.17	Deterministic gradient convergences - unobservable example	174-176
4.18-4.20	Deterministic gradient convergences - stirred tank	178-180
4.21-4.23	Stochastic gradient convergences - observable example	182-184
4.24-4.26	Stochastic gradient convergences - unobservable example	185-187
4.27-4.29	Stochastic gradient convergences - stirred tank	188-190

Figure Title page

5.1 Block diagram of backup identification scheme for flexible space structure 197

5.2 Decentralized row companion canonical form for modal state-space description in eq. (5.9) 201

5.3 The least squares identification of the dependence coefficients contained in A^* and B^* in eq. (5.24) . 205

CHAPTER 1

INTRODUCTION

Over the past ten to fifteen years a number of papers and books (see Wang and Davison, Aoki, Corfmat and Morse, Jamshidi, and Sandell, etal.) have been published on the field of decentralized control. In modern control theory such problems as regulator design, pole-placement, optimal control, etc. often presuppose centrality, i.e., the entire state is assumed available for processing. In the decentralized case any given control problem is attempted to be solved under the assumption that there are one or more controllers each with only a part of the total system state available for measurement and feedback. This thesis is partly motivated by this area of research. Here the concept of decentralization is applied to the very well researched area of system identification thus the term "decentralized identification." Little work has been published on the concept of identifying linear time-invariant systems with such limited or decentralized information. Some work has been done on identifiability of subsystems in interconnected systems (see Brettnauer). This work is formulated in the frequency domain and no direct relation to the system state,

system controllability and observability, and other important system properties is made. Since system information is more generally given by arbitrary outputs than system states, decentralized outputs are generally considered. This thesis definitively solves the problem of identifying linear time-invariant discrete systems with only limited, decentralized information available. It is shown what can be identified about the system and how to do it using various identification algorithms, old and new. It is found that the full order system can always be identified if the global system is observable from the decentralized measurements with an input sequence compatible with the identification algorithm used. If the global system is unobservable from the decentralized measurements a system with order lower than the global system can always be identified completely describing the dynamics of the decentralized set of outputs except in the trivial case in which the state vector is in the null space of the output matrix. Again this assumes an appropriate identification algorithm is applied depending on the complexity of noise present and that sufficient excitation is present for all terms in the measurement equation the algorithm uses. In the centralized case several control problems, e.g., pole-placement, are found to be dual to observer problems (such as the placement of observer poles). The centralized identification problem is found to be significantly different from either controller or observer analysis and no

duality to their analysis exists. As in the centralized case of controller/observer analysis versus identification the problem here is in no way dual to the problem of decentralized control.

To consider the scope of the decentralized identification problem, consider a decentralized syster as shown in Fig. 1.1 taken from

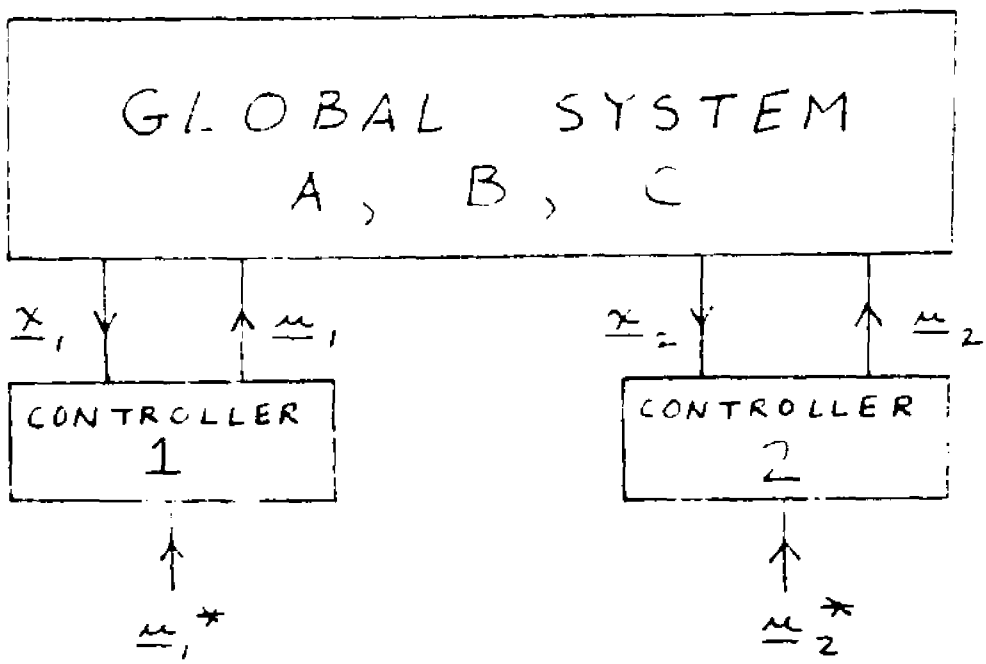

Figure 1.1. A two controller decentralized system.

Sandell, etal. The global system has the completely general linear time-invariant state description given by

$$\dot{x}(t) = Ax(t) + Bu(t) \qquad (1.1)$$

and in the figure two subsystems or controllers receive a part of x. Partition x and u as

$$\underline{x} = \begin{bmatrix} \underline{x}_1 \\ --- \\ \underline{x}_2 \end{bmatrix}, \quad \underline{u} = \begin{bmatrix} \underline{u}_1 \\ --- \\ \underline{u}_2 \end{bmatrix} \tag{1.2}$$

and as shown controller 1 only knows $\underline{x}_1$ and can apply a vector of inputs $\underline{u}_1$ that can be a function of $\underline{x}_1$ or arbitrary input $\underline{u}_1^*$ but not $\underline{x}_2$ and likewise for controller 2, $\underline{x}_2$ and $\underline{u}_2$. This is complete state and input decentralization. If any information is exchanged between controllers the system is termed partially decentralized.

An important paper on decentralized control by Wang and Davison appeared in 1973. They consider L subsystems each receiving a subvector of outputs of the total measurement vector

$$\underline{y}(t) = C\underline{x}(t) \tag{1.3}$$

so that $\underline{y}$ is partitioned

$$\underline{y} = \left[\underline{y}_1 \ \cdots \ \underline{y}_L\right]^T \tag{1.4}$$

and appears as in Fig. 1.1 except using $\underline{y}$ instead of $\underline{x}$. Inputs are still applied as shown. They consider each controller applies

$$\underline{u}_i = K_i \underline{y}_i, \ i=1, \ \ldots, \ L \tag{1.5}$$

so that with all L inputs eq. (1.1) results in

$$\dot{\underline{x}}(t) = A\underline{x}(t) + B \begin{bmatrix} K_1 & & \\ & \ddots & \\ & & K_L \end{bmatrix} \begin{bmatrix} \underline{y}_1 \\ \vdots \\ \underline{y}_L \end{bmatrix}$$

$$= A\underline{x}(t) + BKC\underline{x}(t) \qquad (1.6)$$

where K is block diagonal. In the centralized case it is well known the eigenvalues of A+bKC can be arbitrarily placed if (A,B) is a controllable pair provided complex poles occur in complex conjugate pairs. With a block diagonal K matrix this doesn't hold and stability depends on the fixed modes of A+bKC which are the same as the eigenvalues if K is full, i.e., all elements of the K matrix can be arbitrarily picked. The fixed modes must be stable modes for the system to be stabilizable with block diagonal K. Other system properties have been investigated for decentralized systems and found to hold under conditions that are quite different from those required in the centralized case. Aoki used the following example to show that controllability of unstable poles does not imply stabilizability as it does in the centralized case. The controllable continuous time system has the state description

$$\dot{\underline{x}}(t) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \underline{u}(t) \qquad (1.7)$$

and the matrix

$$K = \begin{bmatrix} k_1 & 0 & 0 \\ 0 & k_2 & 0 \end{bmatrix} \tag{1.8}$$

feeds back the first and second states in a decentralized manner. The characteristic equation is

$$|A + BK - sI| = (1 + k_1 - s)(s^2 - 3s + 2 - k_2) \tag{1.9}$$

with one eigenvalue at

$$s = 1 + k_1 \tag{1.10}$$

which can be placed arbitrarily on the real axis. Applying the Routh-Hurwitz array to the remaining quadratic yields

$$\begin{array}{lll} s^2 : & 1 & 2-k_2 \\ s : & -3 & \\ 1 : & 2-k_2 & \end{array} \tag{1.11}$$

and no matter what is chosen for k_2 there is at least one sign change in the first column and at least one eigenvalue in the right half plane.

The system identification problem invariably means finding an equivalent input-output description from which parameters are identified. The term dependence coefficients is used for these parameters of the input-output model which are then related back to the state description. So far in the literature (R.C.K. Lee, Shrikhande, etal., Irwin and Roberts) the state description is identified in phase-variable canonical form for single-input, single-output systems which is a special case of the Luenberger canonical form for multi-input, multi-output systems. Other state descriptions can always be obtained depending on how states are chosen from the input-output equation(s) and this fact

has not been utilized or even mentioned to date in the literature.

The variation of the Luenberger canonical form used here and in the literature is the row companion form. This form yields a canonical A matrix in eq. (1.1) with main diagonal sub-blocks

$$\begin{bmatrix} 0 & & & \\ \cdot & & & \\ \cdot & 1 & & \\ & & r_i - 1 & \\ \cdot & & & \\ 0 & & & \\ x\ x\ x\ x\ x\ x\ x\ x\ x\ x & & & \end{bmatrix}$$

and off diagonal blocks

$$\begin{bmatrix} 0 & & & 0 \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ 0 & & & 0 \\ x\ x\ x\ x\ x\ x\ x\ x\ x\ x & & & \end{bmatrix}$$

where x's are resultant variable elements for i=1, ..., m, the number of measurements. For an observable global system and observability maintained from the decentralized set of outputs the number of these sub-blocks changes from the number of outputs of the global system to the number of decentralized outputs. The sub-block sizes in general also change and how they change is non-unique in the MIMO case. Chapter 3 clarifies the block determination procedure and

explains the non-uniqueness when testing for the block sizes. In the literature to date no explanation of this inherent non-uniqueness is given. The case where the global system is unobservable from decentralized outputs has never been addressed and is completely resolved in this thesis. Part of the global system can always be identified except in the trivial case of the state vector being in the null space of the output matrix. This identifiable part is in theory and almost always (see section 4.2 for a full discussion of these conditions) in practice the observable part. What elements are identified is determined by use of the observability canonical form reviewed in Chapter 2 and Luenberger canonical form. Some motivation for considering decentralization comes from the following considerations:

1) Many large scale systems already have such a structure. Large urban traffic networks, some economic systems, and large rivers controlled for pollution or resource allocation have the property of only some measurements being available to a given location or processor and only that number is available for feedback.

2) Only partial system information is available and the total system state may be unobservable from this information. System control, identification, stabilization, etc. amounts to solving these problems in a decentralized framework. Two unobservable systems are identified in this thesis.

3) Lower order processing of large scale systems can be

implemented to achieve the same problem objectives. For example in decentralized observer design each subsystem reconstructs only part of the global state thus avoiding high order reconstruction of the whole global state.

4) Some outputs can be processed leaving the others to be processed redundantly as a backup in case of equipment failure. This is the idea behind the identification of the large flexible beam in Chapter 5.

Chapter 2 reviews SISO and MIMO canonical forms especially the Luenberger form and the observability canonical form that are critical to understanding later chapters. In Chapter 3 decentralized measurement equations - input/output equations using the decentralized outputs only - are derived for scalar and vector sets of measurements. The state model used is a completely general linear time-invariant model with state and measurement noise given by

$$\underline{x}(k+1) = A\underline{x}(k) + B\underline{u}(k) + T\underline{w}(k)$$

$$\underline{y}(k) = C\underline{x}(k) + \underline{v}(k) \; . \qquad (1.12)$$

System order determination is also discussed for both observable and unobservable cases.

Chapter 4 solves the decentralized measurement equations with existing identification algorithms. Deterministic and stochastic gradient, and batch and recursive least squares algorithms are discussed and computer simulation results given. Additional algorithms - partial correlation,

lattice, covariance - are also discussed (see Chaupe B, Schichor). Two abstract discrete systems and a discretized model of a continuous time model of a stirred tank are decentralized and dependence coefficients calculated and identified.

Chapter 5 (see Thau and Montgomery) describes a high order large scale system and applies the decentralized identification results to it. A large flexible beam experiment is being conducted at the NASA Langley research center to study the deployment of large objects from the Space Shuttle. The program takes a continuous time modal description of the beam and calculates a discrete time equivalent model. In Chapter 5 using four modes with four available outputs, two outputs are used to identify the entire beam leaving the remaining two to be used as a backup identifier in the event of sensor or processor failure. Computer simulation results of the resulting canonical form and its identification are included.

Chapter 6 discusses further research. The issues involved with relating decentralized identification to decentralized control are raised first. The 1973 paper by Wang and Davison b assumes the entire global system known and that all inputs from all subsystems can be collected for processing at one place to solve the decentralized control problem. If all subsystems identify the whole system then these results are directly applicable as is. If subsystems

measure only part of the system state or operate in a completely independent manner without regard to other subsystems the gain matrix in eq. (1.6) for say the first subsystem is further constrained to

$$K = \mathrm{diag}(K_1 \mid 0 \mid \ldots \mid 0) \qquad (1.13)$$

for which the Wang and Davison results can be applied with some explanation and modification. Problems with using full block diagonal K when the global system is unobservable from some or all subsystems should also be investigated in future research.

Further research needs to be done on identification of ARMA models with noise. So far more research is needed for identification of eq. (1.12) when T is unknown. Identification in other canonical forms should be studied. Finally, very little research has been published for system identification in a particular arbitraly state description, e.g., in the actual physical state description. This is needed since often some elements of a particular model are known and are not easily related to any canonical form. Because of the fewer number of unknown elements for systems in canonical form, identification is made possible using only available measurements. For purposes of having a state description for input/output behavior alone, a canonical form is sufficient and adequate. In this thesis only the K-identifiability result is mentioned and summarized in Appendix A and that applies to only SISO systems.

CHAPTER 2

LINEAR SYSTEM CANONICAL FORMS

In this chapter canonical forms for linear systems are reviewed first. Considerable attention is given to the Luenberger canonical form since identification results for systems in this form have been published (see Shrikhande, et al. and Irwin and Roberts) and are used widely in this thesis. For a MIMO system in this form output decentralization is studied here for the first time and examples are calculated to illustrate different ways of decentralizing the outputs for different sets of output measurements. Before system identification is undertaken in Chapters 3 to 5 the existence of canonical forms must be established for systems given different sets of outputs. This is illustrated with numerical examples. Systems in Luenberger canonical form have a sub-block structure in the output matrix. The sizes of these sub-blocks are called "structural invariants." For a given multi-output system, if fewer than all the outputs are available as in a decentralized system, then a Luenberger canonical form will exist in general from the decentralized outputs but the structural invariants will in general be different from

those describing the full output state description. The existence of these canonical state descriptions is significant in itself since the system input/output behavior is described by fewer arbitrary elements which is the advantage of a canonical form. But also, the existence of these canonical descriptions must be established before system identification can be attempted. Depending on available measurements, identification may be possible only in a specific canonical form.

2.1 SISO Canonical Form

The study of decentralized systems is related to and dependent on the study of large scale systems. Time-invariant, linear, discrete systems, small or large, are described in state-space form using the state and output equations

$$\underline{x}(k+1) = A\underline{x}(k) + B\underline{u}(k) \quad \text{and}$$
$$\underline{y}(k) = C\underline{x}(k). \qquad (2.1)$$

An nth order system has n components or states in the state vector, $\underline{x}(k)$, and the A matrix is nXn with n^2 elements or parameters. A general multivariable system can have r inputs and m outputs so that $\underline{u}(k)$ is an r-vector and $\underline{y}(k)$, an m-vector. Then the B (input) matrix is nXr and the C (measurement) matrix is mXn. These three constant matrices have a total of $(n^2+nr+mn)$ parameters to define the system.

The field of parameter identification is based on finding the system parameters given a history of inputs and outputs. If not enough measurements or the wrong measurements are made, the identification might not be possible. If measurements are restricted then we might expect only some parameters to be identifiable. What can be measured and what can be identified are inter-related. These issues will be further explored after the need for and use of canonical forms is examined.

A state-space description is unique only up to a linear nonsingular state transformation. A state transformation results in new A, B, and C matrices but the outputs are unchanged. Certain transformations exist so that the new matrices assume canonical form where it is known which matrix elements take on the values zero or one. For the identification problem if the system can be identified in canonical form there will be fewer unknown parameters than in noncanonical form and this may facilitate or make possible the identification depending on available measurements.

An important canonical form for identification is phase-variable canonical form. Given a single-input, single-output (SISO) system the system equations have the form

$$\underline{x}(k+1) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & 0 & 1 \\ -a_0 & \cdots & \cdots & \cdots & -a_{n-1} \end{bmatrix} \underline{x}(k) + \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{bmatrix} u(k) \quad (2.2)$$

$y(k) = C\underline{x}(k)$. No special structure is taken by the C matrix after the transformation. The characteristic polynomial for the system is

$$s^n + a_{n-1} s^{n-1} + \cdots + a_1 s + a_0$$

where $a_n = 1$. So the numbers in the last row of the transformed A matrix, $\tilde{A}$, are the coefficients of the characteristic polynomial.

The transformed state variable, $\underline{x}'$, is related to the old state by $\underline{x}(k) = T\underline{x}'(k)$ (see Kuo). The T matrix equals the product of P and M. P is the controllability matrix,

$$P = (b, Ab, A^2 b, \ldots, A^{n-1} b) \quad (2.3)$$

and M is an upper triangular matrix given by

$$M = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \\ a_2 & a_3 & \cdots & a_n & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n-1} & a_n & 0 & \cdots & 0 \\ a_n & 0 & \cdots & \cdots & 0 \end{bmatrix} . \quad (2.4)$$

For $a_n = 1$ the determinant of M = 1 and the matrix is non-singular. Substituting $\underline{x} = T\underline{x}'$ and $\underline{x}(k+1) = T\underline{x}'(k+1)$ into eqs. (2.1) we have

$$T\,\underline{x}'(k+1) = AT\underline{x}'(k) + bu(k) \text{ and}$$

$$y = CT\underline{x}'(k). \quad (2.5)$$

Premultiplying the first equation by T^{-1} gives

$$\underline{x}'(k+1) = T^{-1}AT\underline{x}'(k) + T^{-1}b\underline{u}(k) \qquad (2.6)$$

and so T must be nonsingular for the inverse to exist. Since M is nonsingular P must be nonsingular which also means the system must be controllable to obtain phase-variable canonical form.

A variation of the above phase-variable canonical form is to use the observability matrix for the transformation. This matrix has the form

$$Q = \begin{bmatrix} C \\ \hline CA \\ \hline \vdots \\ \hline CA^{n-1} \end{bmatrix} \qquad (2.7)$$

and is nonsingular if the matrix pair (A,C) is observable. The new state is related to the old state by $\underline{x}' = Q\underline{x}$. The resulting system equations are

$$\underline{x}'(k+1) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & \cdots & & 0 & 1 \\ -a_0 & \cdots & & & -a_{n-1} \end{bmatrix} \underline{x}'(k) + bu(k) \qquad (2.8)$$

$$y(k) = (1, 0, \ldots, 0)\underline{x}(k).$$

This time no structure is assumed by the new B matrix although the elements will be different.

In Appendix A it is shown how the system can be identified in the second phase-variable form given measurements of any state over 2n iterations. It is also shown how the actual A matrix can be identified if all states are measured for 2n iterations.

Another canonical form uses a state transformation called a similarity transformation consisting of the eigenvectors of the A matrix. The state is transformed through $\underline{x} = P\underline{x}'$ where the columns of P are the eigenvectors of A. The new matrix is diagonal with the eigenvalues of A on the main diagonal so that the new system equations are

$$\underline{x}'(k+1) = \begin{bmatrix} L_1 & 0 & 0 & \cdots & 0 \\ 0 & L_2 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & L_n \end{bmatrix} \underline{x}'(k) + \tilde{B}u(k) \qquad (2.9)$$

and

$$\underline{y}(k) = \tilde{C}\underline{x}(k)$$

where $L_1, \ldots, L_n$ are the eigenvalues of A. $\tilde{B}$ and $\tilde{C}$ don't have any special form. For the purpose of identification, however, little work has been done to identify systems in this form. Besides the lack of identification algorithms for this form, the eigenvalues must all be distinct unless A is symmetric and real. For a nonsymmetric A matrix with repeated eigenvalues a transformation exists to make it

almost diagonal. This form, called the Jordan canonical form, consists of Jordan blocks along the main diagonal. Typical blocks are

$$\begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{bmatrix} \qquad (2.10)$$

and

$$\begin{bmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{bmatrix}$$

for a third and second order repeated eigenvalue, respectively.

The variations of the phase-variable canonical form are widely used in the literature on system identification and are extensively used in this thesis. In the SISO case these canonical forms are unique for a given system. In the MIMO case this uniqueness doesn't hold and is discussed in the next section.

2.2 MIMO Canonical Form

In the single-input, single-output case the transformation matrix was found to be unique for phase-variable canonical form. With multi-input, multi-output (MIMO) canonical forms the transformation matrices are not unique in general. Different plans (see Luenberger) have been developed to narrow down the number of transformation

matrices but even within a plan there is still much room for variation.

Six MIMO canonical forms have been summarized in a 1976 paper by Sinha and Rosza. They are the input and output identifiable forms, the row and column companion forms, and the controllable and observable forms.

The output identifiable form was introduced to be used in an identification algorithm mentioned in the above paper. The transformation matrix uses the inverse of the first n rows of the observability matrix. The resulting system equations take the form

$$A = \begin{bmatrix} 0 & I_{p} \\ A_1 & A_2 \end{bmatrix} \tag{2.11}$$

$C = (I_m, 0)$ and B is arbitrary. With $p = n - m$, I_p is a pXp identity matrix, I_m is an mXm identity matrix, and A_1 and A_2 are arbitrary mXm and mXp matrices, respectively. These matrix dimensions have the same meaning as in eq. (2.1), i.e., n is the system order, m is the number of outputs, and r is the number of inputs. The zeros in A and C are null matrices of appropriate sizes. The input identifiable form is the dual of the previous form. Now C is arbitrary and A and B take the forms

$$A = \begin{bmatrix} 0 & A_1 \\ I_q & A_2 \end{bmatrix} \text{ and } B = \begin{bmatrix} I_r \\ 0 \end{bmatrix} \tag{2.12}$$

with q = n-r.

The column companion form, also called the first canonical form of Luenberger, was published in 1967 by Luenberger. The A matrix can be partitioned into sub-blocks

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdot & \cdot & \cdot & A_{1r} \\ A_{21} & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ A_{r1} & \cdot & \cdot & \cdot & \cdot & A_{rr} \end{bmatrix}. \qquad (2.13)$$

The next four canonical forms are characterized by these sub-blocks. The dimensions of the main diagonal sub-blocks are $p_i X p_i$ for i=1, ..., m and are the structural invariants mentioned at the beginning of this chapter. These dimensions are obtained from sequences of rows derived from the observability matrix and the process is reviewed in section 2.3. Sub-blocks on the main diagonal have the structure

$$A_{ii} = \begin{bmatrix} 0 & \cdot & \cdot & \cdot & \cdot & 0 & a_{ii}(1) \\ 1 & 0 & \cdot & \cdot & \cdot & 0 & a_{ii}(2) \\ 0 & 1 & 0 & \cdot & \cdot & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & 0 & 1 & a_{ii}(p_i) \end{bmatrix} \qquad (2.14)$$

with dimension p_i X p_i where $a_{ii}(1)$ through $a_{ii}(p_i)$ are

arbitrary numbers. Off diagonal blocks are

$$A_{ij} = \begin{bmatrix} 0 & \cdots & 0 & a_{ij}(1) \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & a_{ij}(p_j) \end{bmatrix} \qquad (2.15)$$

and have dimension $p_i \times p_j$. The B matrix takes the form

$$B = \begin{bmatrix} 1\ 0 & \cdots & \cdots & 0 \\ 0 \cdots 0\ 1\ 0 & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 1\ 0 \cdots & 0 \end{bmatrix} \qquad (2.16)$$

with a "1" appearing in each new column of an interval p_i elements wide. The C matrix is arbitrary.

The row companion form can be regarded as the dual to the column companion form. The state matrix is again composed of sub-blocks now appearing as

$$A_{ii} = \begin{bmatrix} 0 & 1\ 0 & \cdots & 0 \\ 0 & 0\ 1\ 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & 1 \\ a_{ii}(1) & \cdots & \cdots & a_{ii}(p_i) \end{bmatrix} \qquad (2.17)$$

and

$$A_{ij} = \begin{bmatrix} 0 & \cdots & 0 \\ \cdots & & \cdots \\ 0 & & 0 \\ a_{ij}(1) & \cdots & a_{ij}(p_j) \end{bmatrix} \qquad (2.18)$$

with the B matrix being arbitrary. The C matrix transforms to

$$C = \begin{bmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & 1 & 0 & \cdots & 0 \end{bmatrix} \qquad (2.19)$$

In view of these last two canonical forms one might question as to whether other variations of the above exist. There are two more possibilities. One is the controllable canonical form in which the A matrix takes on rows of parameters as in the row companion form but in which the B matrix takes on the B as in the column companion form. C is then arbitrary. The other and last case is the observable canonical form in which A has columns of parameters as in the column companion form while C is the same as in row companion form and B is arbitrary.

The last four canonical forms have a similarity to and appear to be a MIMO extension of phase-variable SISO form. Main diagonal sub-blocks in the state matrix of the four forms look like the phase-variable state matrix. Also either C or B is in canonical form while the other is arbitrary in both MIMO and SISO forms. In Appendix A following eq. (A.12) it is shown how identification of the system A matrix in the SISO phase-variable form is obtained. In Appendix B, following eq. (B.18), identification of all system matrices

is achieved for a MIMO, centralized system in row companion form. All the necessary measurements must be available and processed at a single location so parameters are identified under a centralized structure. The sub-blocks that appear in the Luenberger canonical A matrix have dimension $p_i \times p_i$ and operate on p_i states that are to the right of the sub-block in the global system state description. For a given row of sub-blocks in A the off-diagonal sub-blocks contain parameters that provide the coupling from states in other subsystems to the next iteration of the states in a given subsystem. Each subsystem's state description looks almost like a SISO phase-variable description except for additional coupling parameters in the A_{ij} and that more than one input can be present in its model. For the decentralized system defined by Wang and Davison (see Fig. 1.1) and for the case of one scalar measurement available to each subsystem, the number of subsystems is equal to the number of measurements in $y(k)$. In row companion form each subsystem measures the first state in its set of p_i states. Had controllable canonical form been used where C is arbitrary, then b would assign one input to each sub-block. In general a decentralized subsystem can apply a vector of inputs to the global system (see Fig. 1.1). Therefore, to describe a more general decentralized system row companion form must be used.

The lowest order, simplest system that also fits the model in Fig. 1.1 is a two state, two input model. At least two states are clearly needed to have decentralized state information with each of the two subsystems having only one state available to it. Each subsystem can apply an input that is either a function of the single state it measures or is arbitrary. Two separate inputs must be present to have input decentralization. If there is only one input from one subsystem the A matrix reduces to a single, phase-variable canonical block. All parameters can be identified from measuring any one state as shown in Appendix A. So it is seen that at least two independent inputs, each with limited state information, are present for this simple decentralized model.

4.3 State Transformation for Luenberger Canonical Form

It is established that the row companion canonical form is appropriate for a linear time-invariant, discrete, decentralized system. This form was first introduced in a 1967 paper by David Luenberger. In that paper the matrix pair (A,b) must be controllable for the transformation matrix to exist. Controllability and observability are "dual" system properties and the b and C matrices are dual matrix quantities. From duality canonical form for (A,C) requires that the pair be observable. The observability

matrix, $(C^T, (CA)^T, \ldots, (CA^{n-1})^T)^T$, must have full rank. The rank then equals the order of the system, n. Independent rows from the observability matrix are used as a basis for the transformed state-space. With

$$C = \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_m \end{bmatrix} \tag{2.20}$$

written in terms of its rows, independent rows are found using the matrix

$$\begin{bmatrix} c_1 \\ c_1A \\ \cdot \\ \cdot \\ \cdot \\ c_1A^{\mu_1-1} \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_2A^{\mu_2-1} \\ \cdot \\ \cdot \\ \cdot \\ c_mA^{\mu_m-1} \end{bmatrix} \tag{2.21}$$

The p_i that appear are called structural invariants and are the sizes of the main diagonal sub-blocks that appear in the Luenberger state transformed A matrix, $\tilde{A}$, as described in section 3.2. Each p_i depends on one c_i and the A matrix so that each measurement (depending on c_i and the state vector) has a p_i associated with it. Each p_i is found by multiplying a row c_i by A raised to a successively higher power until a row $c_i A^{p_i}$ is found. This row is linearly dependent on the previous rows of the ith interval and so is not included in the matrix of linearly independent rows.

The state transformation matrix is not unique since in general there are more than n linearly independent rows. One plan for choosing rows is to use the first n rows in the matrix above. This matrix will be called Luenberger's plan 1 matrix. The second plan is to choose rows in the order

$$c_1, \ldots, c_m, c_1 A, \ldots, c_m A, c_1 A^2, \ldots, c_m A^2, \ldots$$

until n rows are available. This transformation matrix is Luenberger's plan 2 matrix. The number of and size of main diagonal blocks will be different using different plans and parameters will also be different.

To illustrate the above procedures, apply state transformations under both plans to obtain row companion canonical form for a second order, two input model. The model is

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} u_1(k) \\ u_2(k) \end{bmatrix}$$

$$\begin{bmatrix} y_1(k) \\ y_2(k) \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} \qquad (2.22)$$

where elements c_{12} , and c_{21} are zero under the decentralized structure. The model directly fits Fig. 1.1 except that all vectors shown are scalars.

The Luenberger plan 1 matrix is

$$P = \begin{bmatrix} c_1 \\ --- \\ c_1 A \end{bmatrix} \qquad (2.23)$$

where $c_1 A = (c_{11} a_{11} , \ c_{11} a_{12})$. The new state $\underline{x}' = P\underline{x}$. The inverse of P is

$$P^{-1} = \begin{bmatrix} 1 & \\ --- & 0 \\ c_{11} & \\ & \\ -a_{11} & 1 \\ ------ & ------ \\ a_{12} c_{11} & a_{12} c_{11} \end{bmatrix} \qquad (2.24)$$

and the canonical system can then be found

$$\tilde{A} = PAP^{-1} = \begin{bmatrix} c_{11} & 0 \\ c_{11}\, a_{11} & c_{11}\, a_{12} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$\times \begin{bmatrix} \dfrac{1}{c_{11}} & 0 \\ -\dfrac{a_{11}}{a_{12}\, c_{11}} & \dfrac{1}{a_{12}\, c_{11}} \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 \\ a_{12}\, a_{21} - a_{11}\, a_{22} & a_{11} + a_{22} \end{bmatrix}$$

$$\tilde{B} = PB = \begin{bmatrix} c_{11} & 0 \\ c_{11}\, a_{11} & c_{11}\, a_{12} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

$$= \begin{bmatrix} b_{11}\, c_{11} & b_{12}\, c_{11} \\ c_{11}\, a_{11}\, b_{11} + c_{11}\, a_{12}\, b_{21} & c_{11}\, a_{12}\, b_{22} + c_{11}\, a_{11}\, b_{12} \end{bmatrix}$$

and

$$\tilde{C} = CP^{-1} = \begin{bmatrix} c_{11} & 0 \\ 0 & c_{22} \end{bmatrix} \begin{bmatrix} \dfrac{1}{c_{11}} & 0 \\ -\dfrac{a_{11}}{a_{12}c_{11}} & \dfrac{1}{a_{12}c_{11}} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ -\dfrac{a_{11}c_{22}}{a_{12}c_{11}} & \dfrac{c_{22}}{a_{12}c_{11}} \end{bmatrix}. \qquad (2.25)$$

Under the first plan it is seen that the decentralized structure is not preserved. In $\tilde{C}$, x_1 is measured by subsystem 1 and a linear combination of x_1 and x_2 is measured by subsystem 2. For two subsystems two phase-variable sub-blocks are expected in $\tilde{A}$ but only one 2 X 2 block appears.

In the second plan for choosing rows the rows are used in the order c_1, c_2, c_1A, c_2A. For a second order system only the first two rows are used so

$$P = \begin{bmatrix} c_1 \\ -- \\ c_2 \end{bmatrix}. \qquad (2.26)$$

The inverse of P is

$$P^{-1} = \begin{bmatrix} \frac{1}{c_{11}} & 0 \\ 0 & \frac{1}{c_{22}} \end{bmatrix} \qquad (2.27)$$

and the canonical system is

$$\tilde{A} = PAP^{-1} = \begin{bmatrix} c_{11} & 0 \\ 0 & c_{22} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$\times \begin{bmatrix} \frac{1}{c_{11}} & 0 \\ 0 & \frac{1}{c_{22}} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & \frac{c_{11} a_{12}}{c_{22}} \\ \frac{c_{22} a_{21}}{c_{11}} & a_{22} \end{bmatrix}$$

$$\tilde{B} = PB = \begin{bmatrix} c_{11} & 0 \\ 0 & c_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

$$= \begin{bmatrix} b_{11}\, c_{11} & c_{11}\, b_{12} \\ b_{21}\, c_{22} & c_{22}\, b_{22} \end{bmatrix}$$

$$\tilde{C} = CP^{-1} = CC^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \qquad (2.26)$$

Now it is seen that each of the two subsystems measures one of the other state. Also, there are two main diagonal sub-blocks in $\tilde{A}$ as required and both are scalar. Note that for this simplest possible example $\tilde{A}$ has as many parameters as A and the benefit of canonical form for A is not obtained. Larger order systems result in fewer parameters in $\tilde{A}$ as will be seen in the third order example to follow.

Although the Luenberger canonical form is non-unique, in the second example the state transformation narrows down to the second plan for picking independent rows. More generally the choice of rows in the plan leading to the sizes of the sub-blocks is non-unique and a variety of state-space descriptions is possible.

To see the different variations possible consider this third order system

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \end{bmatrix} = \begin{bmatrix} 7.7 & 1 & -11 \\ 1 & 1 & -1 \\ 6.16 & 1 & -6.7 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix} + B\, u(k) \qquad (2.29)$$

$y(k) = Cx(k)$ will be transformed four different ways. The A and C matrices must have different orders under different structures and so cannot be specified yet.

As a first case let two states be assigned to the sub-block of system 1 and the third state to subsystem 2. The state transformation to accomplish this is

$$F = \begin{bmatrix} c_1 \\ --- \\ c_1 A \\ --- \\ c_2 \end{bmatrix} \tag{2.30}$$

plan 1 is used within a sub-block but the number of rows is the number of states in the sub-block. Let

$$C = \begin{bmatrix} 10 & 0 & 0 \\ & & \\ 0 & 0 & 2 \end{bmatrix} \tag{2.31}$$

where a state decentralized structure is already imposed. Subsystem 1 receives 10x and no information about x_3. Subsystem 2 receives $2x_3$ and nothing from x_1 or x_2.

If the whole system is state decentralized (as in Fig. 1.1) then C will already have the form in the example except c_{12} may also be nonzero. If C has all arbitrary elements so that say subsystem 1 gets a linear combination of 3 states then the state transformation will still impose a state decentralized canonical form on the system.

Calculating the second row of P,

$$c_1 A = \begin{pmatrix} 77 & 10 & -110 \end{pmatrix} \tag{2.32}$$

the transformation matrix is

$$P = \begin{bmatrix} 10 & 0 & 0 \\ 77 & 10 & 110 \\ 0 & 0 & 2 \end{bmatrix} \tag{2.33}$$

with an inverse

$$P^{-1} = \begin{bmatrix} .1 & 0 & 0 \\ -.77 & .1 & 10.5 \\ 0 & 0 & .5 \end{bmatrix}. \tag{2.34}$$

Then

$$\tilde{A} = PAP^{-1} = \begin{bmatrix} 0 & 1 & 0 \\ 10.24 & -2.3 & -76.5 \\ -.308 & .2 & 2.3 \end{bmatrix} \tag{2.35}$$

and

$$\tilde{C} = CP^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.36}$$

So subsystem 1 is described by a second order main diagonal sub-block with a coupling element to x_3. Subsystem 2 has a first order or scalar sub-block with coupling elements $\tilde{a}_{31}$ and $\tilde{a}_{32}$ to x_1 and x_2.

An input or vector of inputs is applied separately to the global system by station 1 and station 2. The order of the B (input) matrix is $n \times r$ where r is the sum of all inputs from both stations. B can be some arbitrary matrix and after

the state transformation $\tilde{B} = Pb$ is also arbitrary, therefore, $\tilde{B}$ will not be calculated for some b in the four examples.

In the next case suppose states x_2 and x_3 are available to subsystem 1. Let the measurement matrix be

$$C = \begin{bmatrix} 10 & 0 & 0 \\ & & \\ 0 & 2 & 0 \end{bmatrix} \tag{2.37}$$

The transformation matrix is

$$P = \begin{bmatrix} c_1 \\ --- \\ c_2 \\ --- \\ c_2A \end{bmatrix} . \tag{2.38}$$

hows c_1 and c_2 are obtained from C and row

$$c_1A = (2, 2, -2) . \tag{2.39}$$

Then

$$P = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 2 & 0 \\ 2 & 2 & -2 \end{bmatrix} \tag{2.40}$$

and the inverse is

$$\begin{bmatrix} .1 & 0 & 0 \\ 0 & .5 & 0 \\ .1 & .5 & -.5 \end{bmatrix} \tag{2.41}$$

The canonical state and measurement matrices are

$$\tilde{A} = PAP^{-1} = \begin{bmatrix} -3.3 & -50 & 55 \\ 0 & 0 & 1 \\ -.152 & -2.3 & 3.3 \end{bmatrix} \tag{2.42}$$

and

$$\tilde{C} = CP^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ & & \\ 0 & 1 & 0 \end{bmatrix} \tag{2.43}$$

The resulting system is analogous to the previous case in which subsystem 1 received x_1 and (implicitly) x_2. Element $\tilde{a}_{31}$ is a coupling element from second order subsystem 2 to first order subsystem 1. Elements $\tilde{a}_{12}$ and $\tilde{a}_{13}$ couple the two states in subsystem 2 to subsystem 1.

A third case for a decentralized system where the overall system is third order is to have three subsystems each receiving one state. Let

$$C = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.44}$$

then

$$P = \begin{bmatrix} c_1 \\ -- \\ c_2 \\ -- \\ c_3 \end{bmatrix} \tag{2.45}$$

and

$$P^{-1} = \begin{bmatrix} .1 & 0 & 0 \\ 0 & .5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad (2.46)$$

Calculating the canonical state and measurement matrices gives

$$\tilde{A} = PAP^{-1} = \begin{bmatrix} 7.7 & 5 & -110 \\ .2 & 1 & -2 \\ .616 & .5 & -8.7 \end{bmatrix} \qquad (2.47)$$

and $\tilde{C} = CP^{-1} = CC^{-1} = I_3$ or the third order identity matrix. As in the second order model $\tilde{A}$ has the same number of parameters as A. Each $\tilde{A}_{ii}$ is scalar and each row of $\tilde{A}$ has 2 elements that couple the ith subsystem to the two others.

The last case is the centralized case where there is one system that can process all three states. Let

$C = (1/15.3, 0, 0)$

and

$$P = \begin{bmatrix} C_1 \\ C_1 A \\ C_1 A^2 \end{bmatrix}$$

$$= \frac{1}{15.3} \begin{bmatrix} 1 & 0 & 0 \\ 7.7 & 1 & -11 \\ -7.47 & -2.3 & 10 \end{bmatrix} \qquad (2.48)$$

then

$$\tilde{A} = PAP^{-1} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ .01 & .23 & 0 \end{bmatrix} \qquad (2.49)$$

and $\hat{C} = CP^{-1} = (1, 0, 0)$.

Two canonical forms that are important in the study of controllability and observability will now be reviewed (see Sivan and Kwakernaak). Both forms are canonical, i.e., have fewer arbitrary elements than total system elements when the system is uncontrollable or unobservable, but are full otherwise. The observability canonical form will be used in Chapter 3 to find parameters of the global syster when the global system is unobservable from the decentralized set of outputs for a given subsystem.

First, the controllability canonical form will be studied. The system in eq. (2.1) has the controllability matrix

$$P = (b, Ab, \ldots, A^{n-1} b) \tag{2.50}$$

and let the system be uncontrollable. Then P will be rank deficient. Let P have controllability index, i.e., rank, m. Then m linearly independent columns can be chosen from P. This or any other set of m vectors that spans the linearly independent subspace of P, the controllable subspace, forms a basis for this subspace. Define the matrix

$$T_1 = (p_1, p_2, \ldots, p_m) \tag{2.51}$$

containing these m vectors.

Any n vector, $g \neq 0$, that satisfies the equation

$$T_1^T g = 0 \tag{2.52}$$

is by definition in the null space of T_1. The number of linearly independent vectors that solve eq. (2.52) is n-m. Define

$$T_2 = (p_{m+1} , \ldots , p_n) \tag{2.53}$$

as a matrix containing a basis for the null space (uncontrollable subspace) of T_1 and P. Define the nonsingular transformation matrix

$$T = (T_1 , T_2) \tag{2.54}$$

and transformed state

$$\underline{x}'(k) = T^{-1}\underline{x}(k) \tag{2.55}$$

and substitute eq. (2.55) into eq. (2.1) yielding

$$\begin{aligned} T\underline{x}'(k+1) &= AT\underline{x}'(k) + B\underline{u}(k) \\ \underline{y}(k) &= CT\underline{x}'(k) \end{aligned} \tag{2.56}$$

or

$$\begin{aligned} \underline{x}'(k+1) &= T^{-1}AT\underline{x}'(k) + T^{-1}B\underline{u}(k) \\ \underline{y}(k) &= CT\underline{x}'(k) \ . \end{aligned} \tag{2.57}$$

The canonical structure eq. (2.57) takes on is given by

$$\underline{x}'(k+1) = \begin{bmatrix} A_{11}' & A_{12}' \\ 0 & A_{22}' \end{bmatrix} \underline{x}'(k) + \begin{bmatrix} b_1' \\ 0 \end{bmatrix} \underline{u}(k)$$

$$\underline{y}(k) = CT\underline{x}'(k) = (C_1' \quad 0)\underline{x}(k) \tag{2.58}$$

in which (A_{11}' , b_1') is a controllable pair and no particular structure is taken by CT. This follows by partitioning

$$T^{-1} = \begin{bmatrix} D_1 \\ \hline D_2 \end{bmatrix} \tag{2.59}$$

where D_1 has m rows and D_2 has n-m rows. Then the product

$$T^{-1}T = \begin{bmatrix} D_1T_1 & D_1T_2 \\ D_2T_1 & D_2T_2 \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ 0 & I_{n-m} \end{bmatrix} \tag{2.60}$$

and the products

$$T^{-1}AT = \begin{bmatrix} D_1AT_1 & D_1AT_2 \\ D_2AT_1 & D_2AT_2 \end{bmatrix} \tag{2.61}$$

and

$$T^{-1}B = \begin{bmatrix} D_1B \\ --- \\ D_2B \end{bmatrix} \tag{2.62}$$

follow from the partitioning. From eq. (2.60)

$$D_2T_1 = 0 \tag{2.63}$$

where T_1 spans the controllable subspace. Each column on the right hand side is a null $(n-m)\times m$ vector so that each column of T_1 is in the null space of D_2. Any vector in the controllable space is therefore also in the null space of D_2. The product AT_1 is also in the controllable subspace because the controllable subspace is invariant under A as shown in almost any basic text on control theory, e.g., see Kuo. In eq. (2.61) D_2AT_1 is then a null matrix and D_2B in eq. (2.62) is null since the columns of B are in the controllability matrix and so obviously in the controllable subspace.

Partitioning the state as

$$\underline{x}'(k) = \begin{bmatrix} \underline{x}'_1(k) \\ - - - \\ \underline{x}'_2(k) \end{bmatrix} \tag{2.64}$$

where $\underline{x}'(k)$ is an m vector then

$$\underline{x}'(k+1) = A'_{22}\ \underline{x}'_2(k) \tag{2.65}$$

from eq. (2.58) describes a completely decoupled subsystem in which the input vector has no effect. This uncontrollable subsystem does contribute coupling to the first m states which describe a controllable subsystem.

Since the rows of the T matrix are not unique then the primed submatrices of eq. (2.58) are not unique. It can be shown using the canonical A matrix as written in eq. (2.61) and

$$A = M \Lambda M^{-1} \tag{2.66}$$

where M is the modal matrix of A and Λ is a diagonal matrix containing the eigenvalues of A that the poles of A'_{11} and A'_{22} are the same no matter how the transformation matrix is chosen. The poles of A'_{11} are called the controllable poles of the system and the eigenvectors corresponding to these poles span the controllable subspace of the system. The poles of A'_{22} are called uncontrollable poles and their eigenvectors span the uncontrollable subspace of the system.

Another system property arises if these results are related to system stability. The system is stabilizable if the uncontrollable poles are all stable poles - inside the

unit circle. If controllable poles are unstable, feedback can be used to make the system stable, or stabilize, it.

The observability canonical form is dual to the controllability canonical form. The observability matrix is

$$Q = \begin{bmatrix} C \\ CA \\ \cdot \\ \cdot \\ \cdot \\ CA^{n-1} \end{bmatrix} \qquad (2.67)$$

for the system in eq. (2.1). If the system is unobservable a state transformation exists to transform the system to a canonical form that brings some insight into the system structure. If observable, the state transformation results in a system with as many arbitrary elements as eq. (2.1) as in the controllability canonical form and so cannot technically be said to be canonical.

Let Q be rank deficient and have observability index, or rank, m. Then m linearly independent rows can be found in eq. (2.67). Define the matrix

$$T = \begin{bmatrix} t_1 \\ t_2 \\ \cdot \\ \cdot \\ \cdot \\ t_m \end{bmatrix} \qquad (2.68)$$

containing these linearly independent vectors. Then n-m linearly independent vectors can be found using

$$T_1 \underline{g} = \underline{0} \tag{2.69}$$

where $\underline{g}$ is an n vector to span the null space of ζ. Since ζ is mnXn then n vectors will exist to span the observable and unobservable (null) subspaces. Define the rows in the matrix

$$T_2 = \begin{bmatrix} f_{m+1} \\ f_{m+2} \\ \cdot \\ \cdot \\ \cdot \\ f_n \end{bmatrix} \tag{2.70}$$

using the rows spanning the null space. Define the nonsingular transformation matrix

$$T = \begin{bmatrix} T_1 \\ --- \\ T_2 \end{bmatrix} \tag{2.71}$$

and transformed state

$$\underline{x}'(k) = T\underline{x}(k) \tag{2.72}$$

Substituting into eq. (2.1) gives

$$T^{-1}\underline{x}'(k+1) = AT^{-1}\underline{x}'(k) + b\underline{u}(k)$$

$$\underline{y}(k) = CT^{-1}\underline{x}'(k) \tag{2.73}$$

and then

$$\underline{x}'(k+1) = TAT^{-1}\underline{x}'(k) + Tb\underline{u}(k)$$

$$y(k) = CT^{-1}\underline{x}'(k) \ . \tag{2.74}$$

The canonical form eq. (2.74) takes is given by

$$\underline{x}^*(k+1) = \begin{bmatrix} A^*_{11} & 0 \\ A^*_{21} & A^*_{22} \end{bmatrix} \underline{x}^*(k) + \begin{bmatrix} B^*_1 \\ B^*_2 \end{bmatrix} \underline{u}(k)$$

$$y(k) = (C^*_1 \ , \ 0) \underline{x}^*(k) \tag{2.75}$$

where $(A^*_{11} \ , \ C^*_1)$ is an observable pair. To see this define

$$T^{-1} = (D_1 \ , \ D_2) \tag{2.76}$$

where D_1 is mXn and D_2 is (n-m)Xn. Form the products

$$TT^{-1} = \begin{bmatrix} T_1 D_1 & T_1 D_2 \\ T_2 D_1 & T_2 D_2 \end{bmatrix}$$

$$= \begin{bmatrix} I_m & 0 \\ 0 & I_{n-m} \end{bmatrix} \tag{2.77}$$

$$T^{-1}AT = \begin{bmatrix} T_1 A D_1 & T_1 A D_2 \\ T_2 A D_1 & T_2 A D_2 \end{bmatrix} \tag{2.78}$$

and

$$CT^{-1} = (CD_1 \ , \ CD_2) \tag{2.79}$$

using the partitioned matrices. From eq. (2.77)

$$T_1 D_2 = 0 \tag{2.80}$$

so that all column vectors of D_2 are in the unobservable subspace which is the null space of T_1, the vectors spanning the observable subspace. Any vector $\underline{y}$ that satisfies $T_1 \underline{y} = \underline{0}$ also satisfies $C\underline{y} = \underline{0}$, so $\underline{y}$ is in the unobservable subspace.

The product AD_2 is in the unobservable subspace because the subspace is invariant under A (see Kuo). Then in eq. (2.78)

$$T_1 AD_2 = 0 \tag{2.81}$$

also since the rows of C are rows of Q, and the columns of D_2 are in the null space of Q then

$$CD_2 = 0 \tag{2.82}$$

in eq. (2.79).

Partitioning the state as

$$\underline{x}'(k) = \begin{bmatrix} \underline{x}'_1(k) \\ ---- \\ \underline{x}'_2(k) \end{bmatrix} \tag{2.83}$$

with $\underline{x}'_1(k)$ an n vector then

$$\underline{x}'(k+1) = A'_{11}\,\underline{x}'(k) + B'_1\underline{u}(k)$$

$$\underline{y}(k) = C'_1\underline{x}(k) \tag{2.84}$$

describes a completely observable subsystem of eq. (2.1). The remaining system description is represented by

$$\underline{x}'_2(k+1) = A'_{21}\,\underline{x}_1(k) + A'_{22}\,\underline{x}_2(k) + B'_2\underline{u}(k) \tag{2.85}$$

with zero measurements so that this part is completely unobservable.

This canonical form is important in decentralized identification and is used in Chapter 3. An example from Sivan and Kwakernaak in continuous time will be transformed to observability canonical form to illustrate the method.

An inverted pendulum consisting of a pendulum attached to a pivot on a moving cart (see Fig. 2.1) has the state description

$$\dot{\underline{x}}(t) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 1 & 0 \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} u(t)$$

$$y(t) = (-1, 0, 1, 0)\underline{x}(t) \qquad (2.66)$$

where the system parameters - effective pendulum mass, friction coefficient, and force of gravity - are all set to one for simplicity. The observed variable or output is the angle the pendulum makes with the vertical. The observability matrix is

$$\mathcal{O} = \begin{bmatrix} -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ -1 & 1 & 1 & 0 \\ 0 & -2 & 0 & 1 \end{bmatrix} \qquad (2.67)$$

which has rank 3.

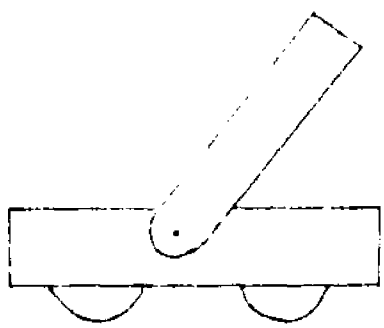

Fig. 2.1 Inverted pendulum on moving cart.

The first three row vectors span the rows of $\mathcal{O}$:

$$(-1,\ 0,\ 1,\ 0),\ (0,\ -1,\ 0,\ 1) \text{ and } (0,\ 1,\ 0,\ 0), \tag{2.88}$$

that is, any row of Q can be written as a linear combination of these three vectors. Any vector in the null space must satisfy

$$\begin{bmatrix} -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = 0 \tag{2.89}$$

or

$$\begin{aligned} -x_1 + x_3 &= 0 \\ -x_2 + x_4 &= 0 \\ x_2 &= 0 \end{aligned} \tag{2.90}$$

so that a vector that spans the unobservable space is

$$(1,\ 0,\ 1,\ 0) . \tag{2.91}$$

This vector is the sum of the two vectors

$$(1,\ 0,\ 0,\ 0) \text{ and } (0,\ 0,\ 1,\ 0) \tag{2.92}$$

and either of these can be used to form a nonsingular transformation matrix. Suppose the first vector in eq. (2.92) is used. Then the transformation matrix is

$$T = \begin{bmatrix} -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \tag{2.93}$$

and its inverse is

$$T^{-1} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} . \qquad (2.94)$$

Calculating all matrix products in eq. (2.74) results in the canonical representation

$$\underline{x}'(t) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \underline{x}'(t) + \begin{bmatrix} 0 \\ -1 \\ -1 \\ 0 \end{bmatrix} u(t)$$

$$y(t) = (1, 0, 0, 0)\underline{x}'(t) \qquad (2.95)$$

where the partitioning clearly shows the corresponding submatrices with eq. (2.75). The submatrix A'_{11} has the observable poles -1, -1, and +1. The single unobservable pole is 0.

Another system property arises if the previous results are related to the system stability. If any state in the unobservable subspace is also in the stable subspace, the system is said to be detectable. Whatever initial estimate is guessed for the unobservable component of the state, the error between the actual state and the observable states plus initial guess for unobservable states will not grow indefinitely as the observable states are reconstructed by an observer.

In summary this section has reviewed SISO and MIMO canonical forms with emphasis on the Luenberger canonical form and the observability canonical form. Systems in row companion form have been discussed and for systems with decentralized output(s) the existence of a new canonical description using those output(s) was discussed and illustrated with examples. Chapter 3 derives a measurement equation first using only one scalar output and then for a vector of outputs from which parameters can be identified and used to obtain a state description in row companion form. Chapter 3 also gives procedures for determining the structural invariants from measurable outputs and measurements of all inputs. Chapter 4 then investigates ways of identifying the parameters with and without noise and presents three examples to illustrate parameter identification from decentralized outputs and the calculation procedure leading to the corresponding state description. Chapter 5 identifies a high order system using decentralized outputs and presents calculations for obtaining the state description of the system.

CHAPTER 3

THE DECENTRALIZED MEASUREMENT EQUATION

This chapter begins with a discussion of the simplest type of identification problem. A memoryless equation with only two parameters is cast into measurement equation form and one way of solving for the parameters is given. Next recursive equations are considered and cast into measurement equation form. Inputs are included yielding ARMA (autoregressive, moving-average) model equations and then issues are raised such as bias and solutions discussed when noise is included. Such equations result when a system's state description is modeled by an equivalent input/output equation(s). Section 3.2 considers a MIMO system when only one output is measurable. Only a scalar measurement is thus given in this decentralized scheme. The derivation for the scalar measurement is achieved here yielding a scalar "decentralized measurement equation" from which parameters relating to the system state description can be identified (the actual identification is studied in Chapter 4). If the global system is observable from the given output then it is shown that the global system can be identified, for example,

in row companion form, using Appendix B. The section also solves the problem when the global system is unobservable from the given output. It is shown that since an unobservable system can be transformed to observability canonical form in which a decoupled observable portion appears, this portion can be identified using the order of this observable part in place of the global system order. Appendix B can then be used to get a state description of this observable portion. Section 3.3 modifies a technique used for SISO systems used in R.C.K. Lee that finds the system order. More than one input must be included in the test and further insight is given in section 3.3 B, showing that columns used in the test are precisely those that appear in the nonsingular least squares solution to the identification problem.

Section 3.4 considers a given vector of measurements and derives a vector decentralized measurement equation. Again as in the scalar case parameters can be identified that can be related to a state-space description of the system, e.g., in row companion form, using Appendix B. Also, the problem of the global system being unobservable from the decentralized set of outputs is again addressed and solved. As in the scalar case, the observability canonical form is used to show that the observable portion of the system using Appendix B can be identified provided an appropriate identification algorithm is correctly applied.

Section 3.5 solves the problem of finding the system order but the situation is more complicated in the multi-output case. The structural invariants which are not unique in general must be found with a test procedure partly similar to the scalar case in section 3.3. Some papers (Bonivento and Guidorzi and Shrikhande, etal.) have given specific sequences to use in the test but this section definitively solves the problem to take into account any test sequence so that the system can be identified in many different forms depending on the sequence used. Again further insight is given by relating the columns tested to the least squares solution to the parameters in the vector DME.

3.1 The Identification Problem

A very simple identification problem is the identification of two constant parameters. Given two time-varying signals, a(k) and b(k), the constant parameters, A and B, are to be found. The output, z(k), is linearly related to two other signals by

$$z(k) = A\,a(k) + B\,b(k) \ .$$

The constants A and B are unknown and are to be identified or solved for. The time-varying signals z(k), a(k), and b(k) are available for all k and are known, i.e., can be measured.

One way to solve this simple problem is to collect two measurements into a system of two equations

$$z(k) = A\,a(k) + B\,b(k)$$

$$z(k+1) = A\,a(k+1) + B\,b(k)$$

or in matrix form

$$\begin{bmatrix} z(k) \\ z(k+1) \end{bmatrix} = \begin{bmatrix} a(k) & b(k) \\ a(k+1) & b(k+1) \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix}$$

by pre-multiplying both sides of the equation by the inverse of the 2 X 2 matrix, often called a measurement matrix, the parameters are found

$$\begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} a(k) & b(k) \\ a(k+1) & b(k+1) \end{bmatrix}^{-1} \begin{bmatrix} z(k) \\ z(k+1) \end{bmatrix}.$$

This approach fails if the determinant of the measurement matrix is zero because the two equations are not linearly independent in that case. It also fails if even a small measurement noise perturbs the deterministic output z(k).

Another negative aspect of this approach is that for large scale, high order systems the measurement matrix will have high order. The calculation of the inverse requires a large number of arithmetic operations that require a large amount of computer time and als , for large order, may not be numerically stable. A whole class of algorithms have been developed which solve the problem recursively (See Saridis and Stein, Mendel, Graupe a, Graupe b, and Astrom).

An estimate of the solution is generated at each iteration step and if conditions for algorithm convergence hold, such as the avoidance of orthogonality in deterministic gradient and stochastic gradient algorithms to be explained later, then the estimates converge to the true parameters. A test for terminating the algorithm such as finding the change in parameters from one iteration to the next could then be applied. If the change is below a set limit, the algorithm ends. The previous nonrecursive approach requiring inversion finds the parameters all at once or all in one "batch" hence is called a batch solution. One important advantage of some of the recursive techniques is they do not require the calculation of any inverses (see Saridis and Stein).

In the two parameter example, at least two measurements had to be concatenated and the measurement matrix to be inverted was second order. The problem can be generalized to n parameters with an output z(k) depending on n time varying signals that are measurable

$$z(k) = P1\ x_1(k) + P2\ x_2(k) + \ldots + PN\ x_n(k) .$$

This equation in vector notation is

$$z(k) = \left[x_1(k),\ x_2(k),\ \ldots,\ x_n(k) \right] \begin{bmatrix} P_1 \\ -- \\ P2 \\ -- \\ \cdot \\ \cdot \\ \cdot \\ -- \\ PN \end{bmatrix}$$

$$= \underline{x}^T(k)\ \underline{P}\ .$$

In the batch approach measurements at successive times are concatenated or stacked together until there is an overdetermined set of equations

$$\begin{bmatrix} z(k) \\ \cdot \\ \cdot \\ \cdot \\ z(k+N-1) \end{bmatrix} = \begin{bmatrix} x^T(k) \\ \cdot \\ \cdot \\ \cdot \\ x^T(k+N-1) \end{bmatrix} \underline{P}\ .$$

The solution is then obtained by

$$\underline{P} = \begin{bmatrix} \underline{x}^T(k) \\ \cdot \\ \cdot \\ \cdot \\ x^T(k+N-1) \end{bmatrix}^{+} \begin{bmatrix} z(k) \\ \cdot \\ \cdot \\ \cdot \\ z(k+N-1) \end{bmatrix}$$

where "+" is the pseudo-inverse.

An important application of this formulation is in the identification of the coefficients in a finite difference equation. Suppose there is an nth order finite difference equation

$$y(k+n) + a_1 y(k+n-1) + \ldots + a_n y(k) = bu(k) \qquad (3.2)$$

with unknown coefficients $a_1, \ldots, a_n$, b and $y(k+n), \ldots, y(k), u(k)$ are perfectly measurable. Solving for $y(k+n)$ and using vector notation

$$y(k+n) = \left[y(k+n-1), \ldots, y(k), u(k)\right] \begin{bmatrix} -a_1 \\ \cdot \\ \cdot \\ \cdot \\ -a_n \\ b \end{bmatrix}.$$

This example is cast into exactly the same measurement equation formulation as the previous identification problem (see Mendel for a detailed description of the equation error formulation). From here measurements at successive times can be concatenated. After n+1 measurements are collected the inverse of the n+1 X n+1 measurement matrix could be pre-multiplied on both sides of the concatenated equation to obtain the n+1 parameters in the difference equation.

The identification of linear time-invariant systems in state-space canonical form can now be achieved since an nth order state description can be represented by an nth order finite difference equation. The following state-space description

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ \cdot \\ \cdot \\ \cdot \\ x_n(k+1) \end{bmatrix} = \begin{bmatrix} 0 & & \\ \cdot & & \\ \cdot & I_{n-1} & \\ \cdot & & \\ 0 & & \\ -a_n & \cdots & -a_1 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ \cdot \\ \cdot \\ \cdot \\ x_n(k) \end{bmatrix} + \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ b \end{bmatrix} \underline{u}(k) \qquad (3.3)$$

is equivalent to the finite difference equation in (3.2) If the following states are defined

$$x_1(k) = y(k)$$

$$x_2(k) = y(k+1)$$

$$x_3(k) = y(k+2)$$

$$\cdot$$

$$\cdot$$

$$\cdot$$

$$x_n(k) = y(k+n-1) \; . \qquad (3.4)$$

To see the equivalence first step up $x_n(k)$ to $x_n(k+1) = y(k+n)$ then from the last row in eq. (3.3)

$$x_n(k+1) = y(k+n) = -a_n x_1(k) - a_{n-1} x_2(k) - \cdots - a_1 x_n(k) + bu(k) \; . \qquad (3.5)$$

Next substituting for the n states from eq. (3.4) into (3.5)

$$y(k+n) = -a_n y(k) - a_{n-1} y(k+1) - \cdots - a_1 y(k+n-1) + bu(k)$$

or

$$y(k+n) + a_n y(k) + a_{n-1} y(k+1) + \cdots + a_1 y(k+n-1) = bu(k) \; .$$

With the finite difference equation/state-space description equivalence established the identification of parameters in a state-space model is cast into the

measurement equation formulation of the memoryless problem at the beginning of this chapter. That example as well as the others were deterministic, involving no random disturbances or measurement noise. An important issue in identification is to identify system parameters in the presence of noise corrupting or being added onto any or all signals that are measured. Returning to the nth order memoryless example consider that a noise source is present being added to a linear combination of n time-varying signals $g_1(k)$ to $g_n(k)$ so that

$$z(k) = P1\ g_1(k) + \ldots + PN\ g_n(k) + v(k)$$
$$= \underline{g}\ (k)\ \underline{P} + v(k)$$

in vector notation. Assume the values that the noise source $v(k)$ takes are unknown and values at different times are statistically independent of each other. Again as before concatenate n measurements

$$z(k) = \underline{g}^T(k)\underline{P} + v(k)$$
$$z(k+1) = \underline{g}^T(k+1)\underline{P} + v(k+1)$$
$$\vdots$$
$$z(k+n-1) = \underline{g}^T(k+n-1)\underline{P} + v(k+n-1)\ . \qquad (3.6)$$

The parameter vector $\underline{P}$ is common to each equation and can be factored out. Let

$$\underline{Z}(k) = (z(k),\ z(k+1),\ \ldots,\ z(k+n-1))^T,$$
$$\underline{G}(k) = (g(k),\ g(k+1),\ \ldots,\ g(k+n-1))^T$$

and $\underline{V}(k) = (v(k),\ v(k+1),\ \ldots,\ v(k+n-1))^T$.

Then the concatenated measurement equation (3.6) can be written as

$$\underline{z}(k) = G(k)\underline{x} + \underline{V}(k) \quad . \tag{3.7}$$

In the deterministic case $\underline{V}(k)$ is zero and (3.7) would be solved by pre-multiplying by $G^{-1}(k)$. If this is repeated here

$$\underline{x} = G^{-1}(k)\underline{z}(k) - G^{-1}(k)\underline{V}(k)$$

and the true solution is altered by the term $G^{-1}(k)\underline{V}(k)$.

If L>n measurement equations are concatenated the system of equations is then overdetermined (with regard to signals and parameters only) and this fact can be used to offset the effect of measurement errors. Based on the available measurements an estimate, $\hat{\underline{x}}(k)$, of the true parameters will be generated. The parameter error is $\tilde{\underline{x}}(k) = \underline{x} - \hat{\underline{x}}(k)$. A linear estimate of $\underline{z}(k)$ is obtained through $\hat{\underline{z}}(k) = G(k)\hat{\underline{x}}(k)$ and a measurement error is defined by $\tilde{\underline{z}}(k) = \underline{z}(k) - \hat{\underline{z}}(k)$. (Here $G(k)$ is an LXn matrix, L>n.)

An estimate $\hat{\underline{x}}(k)$ along the lines of the batch solution with no noise is obtained through the method of generalized least squares: by minimizing the weighted sum of the squares of the components of the measurement error vector, $\tilde{\underline{z}}(k)$. The weighted sum is

$$S(\hat{\underline{x}}(k)) = w(k)\tilde{z}^2(k) + w(k+1)\tilde{z}^2(k+1) + \ldots$$
$$+ w(k+L-1)\tilde{z}^2(k+L-1)$$

which can be written as

$$S = \tilde{\underline{z}}^T(k)\,W(k)\,\tilde{\underline{z}}(k)$$

with

$$W(K) = \begin{bmatrix} w(k) & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & w(k+1) & 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 & w(k+L-1) \end{bmatrix} \quad .$$

To find the $\hat{\underline{F}}(k)$ that minimizes S substitute for $\tilde{\underline{Z}}(k)$ with

$$\tilde{\underline{Z}}(K) = \underline{Z}(k) - \hat{\underline{Z}}(K) = \underline{Z}(k) - G(K)\hat{\underline{F}}(K)$$

and then set the derivative of the resulting expression with respect to $\hat{\underline{F}}(k)$ to the null vector:

$$\frac{dS}{d\hat{\underline{F}}} = -2(G^T(k)W(k)\underline{Z}(k)) + 2G^T(k)W(k)G(k)\hat{\underline{F}}(k) = \underline{0} \quad .$$

This equation is called the normal equation. Solving the equation yields

$$\hat{\underline{F}}(K) = (G^T(k)W(k)G(K))^{-1}(G^T(K)W(k)\underline{Z}(k)) \quad . \qquad (3.7b)$$

The second derivative of S must be positive definite to minimize S. Taking the derivative of $dS/d\hat{\underline{F}}$,

$$\frac{d^2S}{d\hat{\underline{F}}^2} = 2G^T(K)W(K)G(K).$$

A requirement for a generalized least squares solution is that the inverse $(G^TWG)^{-1}$ exist. If all rows of G are linearly independent so that G has maximum rank, the inverse exists. If $W(K) = wI$ where w is a scalar and I the nth order identity matrix then $\hat{\underline{F}}(k)$ is the least squares estimate of $\underline{F}$ given by

$$\hat{\underline{F}}(K) = (G^T(K)G(K))^{-1}G^T(K)\underline{Z}(K) \quad .$$

In eq. (3.7) the solution with measurement noise was incorrect because of the term $G^{-1}(k)\underline{V}(k)$. The solution is then clearly incorrect by the amount $G^{-1}(k)\underline{V}(k)$. Without noise the generalized least squares estimate is

$$\hat{\underline{P}}(k) = (G^T(k)W(k)G(k))^{-1}G^T(k)W(k)\underline{Z}(k)$$
$$= (G^T(k)W(k)G(k))^{-1}G^T(k)W(k)G(k)\underline{P} = \underline{P}$$

which is clearly correct. With noise the estimate is

$$\hat{\underline{P}}(k) = (G^T(k)W(k)G(k))^{-1}G^T(k)W(k)(G(k)\underline{P} + \underline{V}(k))$$
$$= \underline{P} + (G^T(k)W(k)G(k))^{-1}G^T(k)W(k)\underline{V}(k) \quad . \qquad (3.8)$$

This does not solve the problem but the estimate, $\hat{\underline{P}}(k)$, approaches $\underline{P}$ as L, the number of measurements, increases if the second term on the right hand side of (3.8) has zero mean. Taking the expectation on both sides of (3.8)

$$E(\hat{\underline{P}}(k)) = \underline{P} + E((G^T(k)W(k)G(k))^{-1}G^T(k)W(k)\underline{V}(k))$$

The second term is the bias and if it has zero mean then

$$E(\hat{\underline{P}}(k)) = \underline{P} \quad .$$

One way for that second term to have zero mean is if $G(k)$ and $\underline{V}(k)$ are statistically independent. In that case

$$E((G^TWG)^{-1}G^TW\underline{V}) = E((G^TWG)^{-1}G^TW)E(\underline{V}) = \underline{0}$$

because $E(\underline{V}(k)) = \underline{0}$.

Now consider the identification of parameters in a system with memory and noise. A system with memory can be modeled by the finite difference equation previously studied. Now suppose a noise source, $w(k)$, is present so that

$$y(k+n) + a_1y(k+n-1) + \ldots + a_ny(k) = bu(k) + w(k). \qquad (3.9)$$

This closely resembles the decentralized measurement equation to be derived in the next section.

Writing (3.9) in vector form

$$y(k+n) = y(k+n-1), \ldots, y(k), u(k) \begin{bmatrix} -a_1 \\ -a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \\ b \end{bmatrix} + w(k)$$

L measurements are then concatenated with L>n so that

$$\begin{bmatrix} y(k+n) \\ y(k+n+1) \\ \cdot \\ \cdot \\ \cdot \\ y(k+n+L-1) \end{bmatrix} = \begin{bmatrix} y(k+n-1) & \cdots & y(k) & u(k) \\ y(k+n) & \cdots & y(k+1) & u(k+1) \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ y(k+n+L-2) & \cdots & y(k+L-1) & u(k+L-1) \end{bmatrix} \begin{bmatrix} -a_1 \\ -a_2 \\ \cdot \\ \cdot \\ \cdot \\ -a_n \\ b \end{bmatrix}$$

$$+ \begin{bmatrix} w(k) \\ w(k+1) \\ \cdot \\ \cdot \\ \cdot \\ w(k+L-1) \end{bmatrix}$$

or

$$y(k) = G^{T}(k)\underline{\theta} + \underline{w}(k) \ .$$

An unbiased least squares solution exists if $G(k)$ and $\underline{w}(k)$ are independent. For recursive measurement equations it is found that they are not independent. This will be shown by finding what elements of $G(k)$ are dependent on what elements of the noise. If the elements of $G(k)$ are dependent on noise terms that also appear in the noise vector $\underline{w}(k)$ then $G(k)$ and $\underline{w}(k)$ are correlated and the least squares solution could be biased.

To see what noise terms $G(k)$ depends on, each element $y(\ )$ must be analyzed for dependence on noise terms. This dependence can be seen by obtaining the equivalent state space description of the finite difference equation in which $x_1(k) = y(k)$ and solving the state equation for $x_1(k)$. The required form results if

$$x_1(k) = y(k)$$

$$x_2(k) = y(k+1)$$

$$\vdots$$

$$x_n(k) = y(k+n-1)$$

then eq. (3.9) is equivalent to

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ \cdot \\ \cdot \\ \cdot \\ x_n(k+1) \end{bmatrix} = \begin{bmatrix} 0 & & \\ \cdot & & \\ \cdot & I_{n-1} & \\ \cdot & & \\ 0 & & \\ -a_n \; \cdot \; \cdot \; \cdot \; -a_1 & & \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ \cdot \\ \cdot \\ \cdot \\ x_n(k) \end{bmatrix} + \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ b \end{bmatrix} u(k)$$

$$+ \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 1 \end{bmatrix} w(k) \quad .$$

Re-writing the equation as

$$\underline{x}(k+1) = A\underline{x}(k) + \underline{b}u(k) + \underline{t}w(k)$$

the solution to this equation is

$$\underline{x}(k) = A^{k}\underline{x}(0) + \sum_{i=1}^{k} A^{k-i}\underline{b}u(i-1) + \sum_{i=1}^{k} A^{k-i}\underline{t}w(i-1).$$

Since $y(k) = x_1(k)$ the noise terms on which $y(k)$ depends can be determined from the second summation in the solution. The second summation contributes a sum of terms beginning with $w(0)$ for $i=1$ up to $w(k-1)$ for $i=k$ multiplying whatever elements from A^{k-i} and $\underline{t}$. The earliest and latest outputs appearing in $G(k)$, $y(k)$ and $y(k+n+L-2)$, respectively, are found to be dependent on $w(0)$ to $w(k)$ to $w(k+L-1)$ up to $w(k+n+L+3)$. Comparing $G(k)$ to

$$\underline{W}^{T}(k) = (w(k), w(k+1), \ldots, w(k+L-1))$$

it is seen that $G(k)$ and $\underline{w}(k)$ are functions of identical noise terms which means they are not independent. This doesn't mean they might not be orthogonal and the estimate could possibly be unbiased. The condition of orthogonality, however, depends on the specific sequence of values in $\underline{w}(k)$ which is not measurable, and so, not available. Orthogonality cannot be proved before using the generalized least squares estimate which may or may not be biased making this approach ineffective for the current problem.

A batch approach applied to the parameter estimation problem for a system modeled by a difference equation with a noise source is seen to be ineffective and this problem is of a simpler form than the equation to be identified in the next section. In that section a vector of noise sources, $\underline{w}(k)$, is present in the state model and a finite difference equation derived from the model has components of the noise vector multiplying unknown coefficients. As yet this development hasn't taken into account measurement noise as in the memoryless problem in eq. (3.7). These more complicated problems as well as problems solvable by the batch approach can be dealt with using recursive identification methods.

In general most recursive methods or algorithms have the form

$$\hat{\underline{p}}(k+1) = \hat{\underline{p}}(k) + \underline{G}(k)(e_1) \qquad (3.11)$$

in which $\hat{\underline{P}}(k)$ is the kth estimate of a vector of unknown parameters, $\underline{P}$. The gain K multiplies an error, er, which is the difference between an available measurement that may be corrupted by noise and an estimate of this measurement. For example, in recursive gradient type algorithms a measurement estimate is generated through

$$z(k) = \underline{g}^{T}(k)\,\underline{\hat{P}}(k)$$

where $\underline{g}^{T}(k)$ is a vector of past measurements. In the absence of noise the true measurement is generated by the process according to

$$z(k) = \underline{g}^{T}(k)\,\underline{P}.$$

Then the error is

$$er(k) = z(k) - \hat{z}(k) = \underline{g}^{T}(k)\,(\underline{P} - \underline{\hat{P}}(k)).$$

As the error decreases due to new estimates closer to the true parameters the correction term, Ker, in the algorithm contributes less to new estimates. The gain may depend on the measurements and arbitrary weights, constrained by conditions for convergence to the true parameters.

Enough background discussion and examination of identification problems has been considered to lead into the next section. There a measurement equation using limited, i.e., decentralized, state information is derived from which parameters can be identified. In Chapter 4 deterministic and stochastic algorithms will be studied in greater detail and will be applied along with batch techniques to the decentralized state measurement equation.

3.2 Derivation of the DME

This section considers the problem of identifying elements of the system matrices for linear time-invariant MIMO systems when only part of the total system state is available. More generally the decentralized structure allows for measurements that are linear combinations of that available part of the total state. For the linear time-invariant system

$$\underline{x}(k+1) = A\underline{x}(k) + B\underline{u}(k) + T\underline{w}(k)$$

$$\underline{ym}(k) = C\underline{x}(k) + \underline{v}(k) \qquad (3.12)$$

the structure under consideration allows one measurement $y_i(k)$ of the mth order output vector to be measured by a subsystem.

Given a history of the single measurement $ym_i(k)$ an equation, called the Decentralized Measurement Equation, (DME), will be derived in terms of past and present measurements of the form

$$ym_i(k+n) = pNym_i(k+n-1) + \ldots + p1ym_i(k)$$
$$+ \underline{pu1}^T\underline{u}(k+n-1) + \ldots + \underline{pun}^T\underline{u}(k)$$
$$+ \underline{pw1}^T\underline{w}(k+n-1) + \ldots + \underline{pwn}^T\underline{w}(k) \quad .$$

From this equation the coefficients are identified in Chapter 4 using various identification algorithms and for varying complexity of the noise terms appearing in (3.12). From the identified parameters a state description in row companion form is realizable using the results of Appendix

b. Since the DME is a function of only $ym_i(k)$ and no other outputs, the equation is "output decentralized." Note that the equation involves a linear combination of past and present inputs from all subsystems.

All identification algorithms considered in this thesis require that all inputs, $\underline{u}(k)$, be known. All input information must be available to a subsystem so that input information is centralized.

The derivation of the DME will proceed for an arbitrary ith subsystem of an overall nth order global system. The ith subsystem "sees"

$$ym_i(k) = c_i \underline{x}(k) + v_i(k) \tag{3.13}$$

where c_i is the ith of m rows in C in eq. (3.12). A concatenated vector equation will be obtained in terms of the total state, $\underline{x}(k)$, by linking n successive measurements leading to the DME. The next measurement is

$$ym_i(k+1) = c_i \underline{x}(k+1) + v_i(k+1). \tag{3.14}$$

Substituting for $\underline{x}(k+1)$ from eq. (3.12) then

$$ym_i(k+1) = c_i A\underline{x}(k) + c_i B\underline{u}(k) + c_i \Gamma\underline{w}(k) + v_i(k+1). \tag{3.15}$$

Stepping up (3.15) and again substituting out $\underline{x}(k+1)$

$$ym_i(k+2) = c_i A\underline{x}(k+1) + c_i B\underline{u}(k+1)$$
$$+ c_i T\underline{w}(k+1) + v_i(k+2)$$
$$= c_i A(A\underline{x}(k) + b\underline{u}(k) + T\underline{w}(k)) + c_i B\underline{u}(k+1)$$
$$+ c_i T\underline{w}(k+1) + v_i(k+2)$$
$$= c_i A^2\underline{x}(k) + c_i AB\underline{u}(k) + c_i b\underline{u}(k+1)$$
$$+ c_i AT\underline{w}(k) + c_i T\underline{w}(k+1) + v_i(k+2). \qquad (3.16)$$

Similarly, the next measurement is found

$$ym_i(k+3) = c_i A^2\underline{x}(k+1) + c_i AB\underline{u}(k+1)$$
$$+ c_i b\underline{u}(k+2) + c_i AT\underline{w}(k+1) + c_i T\underline{w}(k+2)$$
$$+ v_i(k+3)$$
$$= c_i A^3\underline{x}(k) + c_i A^2B\underline{u}(k) + c_i AB\underline{u}(k+1)$$
$$+ c_i B\underline{u}(k+2) + c_i A^2T\underline{w}(k) + c_i AT\underline{w}(k+1)$$
$$+ c_i T\underline{w}(k+2) + v_i(k+3). \qquad (3.17)$$

The final nth measurement is

$$ym_i(k+n-1) = c_i A^{n-1}\underline{x}(k) + c_i A^{n-2} b\underline{u}(k)$$
$$+ c_i A^{n-3} b\underline{u}(k+1) + \ldots + c_i Ab\underline{u}(k+n-3) + c_i B\underline{u}(k+n-2)$$
$$+ c_i A^{n-2} T\underline{w}(k) + c_i A^{n-3} T\underline{w}(k+1) + \ldots + c_i AT\underline{w}(k+n-3)$$
$$+ c_i T\underline{w}(k+n-2) + v_i(k+n-1). \qquad (3.18)$$

The n measurements are now collected to form the concatenated decentralized measurement equation (CDME)

$$\begin{bmatrix} ym_i(k) \\ ym_i(k+1) \\ \vdots \\ ym_i(k+n-2) \\ ym_i(k+n-1) \end{bmatrix} = \begin{bmatrix} c_i \\ c_i A \\ \vdots \\ c_i A^{n-2} \\ c_i A^{n-1} \end{bmatrix} \underline{x}(k) + \begin{bmatrix} 0 \\ c_i B \\ \vdots \\ c_i A^{n-3} B \\ c_i A^{n-2} B \end{bmatrix} \underline{u}(k)$$

$$+\begin{bmatrix} 0 \\ 0 \\ c_i B \\ \cdot \\ \cdot \\ \cdot \\ c_i A^{n-3} B \end{bmatrix} \underline{u}(k+1) + \begin{bmatrix} 0 \\ 0 \\ 0 \\ c_i B \\ \cdot \\ \cdot \\ c_i A^{n-4} B \end{bmatrix} \underline{u}(k+2) + \cdots$$

$$+\begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ c_i B \end{bmatrix} \underline{u}(k+n-2) + \begin{bmatrix} 0 \\ c_i T \\ \cdot \\ \cdot \\ \cdot \\ c_i A^{n-3} T \\ c_i A^{n-2} T \end{bmatrix} \underline{w}(k)$$

$$+\begin{bmatrix} 0 \\ 0 \\ c_i T \\ \cdot \\ \cdot \\ \cdot \\ c_i A^{n-3} T \end{bmatrix} \underline{w}(k+1) + \begin{bmatrix} 0 \\ 0 \\ 0 \\ c_i T \\ \cdot \\ \cdot \\ c_i A^{n-4} T \end{bmatrix} \underline{w}(k+2) + \cdots$$

$$+ \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ c_i I \end{bmatrix} \underline{w}(k+n-2) + \begin{bmatrix} v_i(k) \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ v_i(k+n-1) \end{bmatrix} . \qquad (3.19)$$

More compactly the CDME is rewritten as

$$\overline{ym}_i(k) = M\underline{x}(k) + M\underline{u}(k) + M1\underline{u}(k+1) + M2\underline{u}(k+2) +$$
$$+ \ldots + (Mn-2)\underline{u}(k+n-2) + N\underline{w}(k) + N1\underline{w}(k+1) +$$
$$+ N2\underline{w}(k+2) + \ldots + (Nn-2)\underline{w}(k+n-2) + \overline{v}_i(k) . \qquad (3.20)$$

Next step up the CDME one iteration so that

$$\overline{ym}_i(k+1) = M\underline{x}(k+1) + M\underline{u}(k+1) + M1\underline{u}(k+2)$$
$$+ M2\underline{u}(k+3) + \ldots + (Mn-2)\underline{u}(k+n-1) + N\underline{w}(k+1)$$
$$+ N1\underline{w}(k+2) + N2\underline{w}(k+3) + \ldots + (Nn-2)\underline{w}(k+n-1)$$
$$+ \overline{v}_i(k+1) . \qquad (3.21)$$

Substituting $\underline{x}(k+1)$ in terms of $\underline{x}(k)$ from eq. (3.12) results in

$$\overline{ym}_i(k+1) = hA\underline{x}(k) + hB\underline{u}(k) + M\underline{u}(k+1)$$
$$+ M1\underline{u}(k+2) + M2\underline{u}(k+3) + \ldots + (Mn-2)\underline{u}(k+n-1) + H1\underline{w}(k)$$
$$+ N\underline{w}(k+1) + N1\underline{w}(k+2) + N2\underline{w}(k+3) + \ldots$$
$$+ (Nn-2)\underline{w}(k+n-1) + \overline{v}_i(k+1) . \qquad (3.22)$$

The full state, $\underline{x}(k)$, which is unknown and in general could be unobservable is eliminated using the CDME, eq. (3.20), where

$$\underline{x}(k) = H^{-1}\overline{y\underline{m}}_i(k) - H^{-1}M\underline{u}(k)$$
$$- H^{-1}M1\underline{u}(k+1) - H^{-1}M2\underline{u}(k+2) - \ldots$$
$$- H^{-1}(Mn-2)\underline{u}(k+n-2) - H^{-1}N\underline{w}(k) - H^{-1}N1\underline{w}(k+1)$$
$$- H^{-1}N2\underline{w}(k+2) - \ldots - H^{-1}(Nn-2)\underline{w}(k+n-2) -$$
$$- H^{-1}\overline{v}_i(k) \quad . \tag{3.23}$$

Here it is assumed that H^{-1} exists. The possibility of singular H is treated in the next section.

Substituting $\underline{x}(k)$ into (3.22) results in

$$\overline{y\underline{m}}_i(k+1) = HAH^{-1}\overline{y\underline{m}}_i(k) - HAH^{-1}M\underline{u}(k) - HAH^{-1}M1\underline{u}(k+1)$$
$$- HAH^{-1}M2\underline{u}(k+2) - \ldots - HAH^{-1}Mn-2\underline{u}(k+n-2)$$
$$- HAH^{-1}N\underline{w}(k) - HAH^{-1}N1\underline{w}(k+1) - HAH^{-1}N2\underline{w}(k+2)$$
$$- \ldots - HAH^{-1}Nn-2\underline{w}(k+n-2) - HAH^{-1}\overline{v}_i(k) + HB\underline{u}(k)$$
$$+ M\underline{u}(k+1) + M1\underline{u}(k+2) + M2\underline{u}(k+3) + \ldots + Mn-2\underline{u}(k+n-1)$$
$$+ HI\underline{w}(k) + N\underline{w}(k+1) + N1\underline{w}(k+2) + N2\underline{w}(k+3) + \ldots$$
$$+ Nn-2\underline{w}(k+n-1) + \overline{v}_i(k+1) \quad . \tag{3.24}$$

Defining $\overline{H}=HAH^{-1}$ and adding like terms in (3.24) gives

$$\overline{y\underline{m}}_i(k+1) = \overline{H}\overline{y\underline{m}}_i(k) + (HB-\overline{H}M)\underline{u}(k) + (M - \overline{H}M1)\underline{u}(k+1)$$
$$+ (M1 - \overline{H}M2)\underline{u}(k+2) + \ldots + (Mn-3 - \overline{H}Mn-2)\underline{u}(k+n-2)$$
$$+ Mn-2\underline{u}(k+n-1) + (HI-\overline{H}N)\underline{w}(k) + (N-\overline{H}N1)\underline{w}(k+1)$$
$$+ (N1-\overline{H}N2)\underline{w}(k+2) + \ldots + (Nn-3-\overline{H}Nn-2)\underline{w}(k+n-2)$$
$$+ Nn-2\underline{w}(k+n-1) - \overline{H}\overline{v}_i(k) + \overline{v}_i(k+1) \quad . \tag{3.25}$$

The next step in the derivation of the DME is to evaluate dependence coefficient matrices in (3.25) which multiply time varying functions of k. The first such matrix in eq. (3.25) is $\overline{H}=HAH^{-1}$. To evaluate this first return to eq. (3.19) where H is seen to be composed of rows beginning with

C_i and multiplying powers of A up to $C_i A^{n-1}$. This matrix is the "local" observability matrix of the ith subsystem. Alternatively, it can be looked at as the observability matrix of a single output system. As is well known from the study of canonical system state transformations and reviewed in chapter 2, this matrix product is

$$\bar{H} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ \cdot & 0 & 1 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdot & \cdot & 0 & 1 \\ -a_0 & -a_1 & - & \cdots & & -a_{n-1} \end{bmatrix}. \qquad (3.26)$$

The coefficients in the last row are the negatives of the coefficients in the characterisitic polynomial of A:

$$s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0$$

If H^{-1} is singular and the total state is unobservable from the subsystem in question, less than the total global system in row companion can be identified. What can be identified is explained in the next section but the DMT still applies with a simple modification.

Next the two matrix products whose difference is the coefficient dependence matrix of $\underline{y}(k)$ are evaluated

$$HB = \begin{bmatrix} c_i \\ c_i A \\ \cdot \\ \cdot \\ \cdot \\ c_i A^{n-1} \end{bmatrix} \quad b = \begin{bmatrix} c_i b \\ c_i Ab \\ \cdot \\ \cdot \\ \cdot \\ c_i A^{n-1} b \end{bmatrix} \tag{3.27}$$

$$\bar{H}\bar{M} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdots & \cdots & 0 & 1 \\ P1 & \cdots & \cdots & PN \end{bmatrix} \begin{bmatrix} 0 \\ c_i b \\ \cdot \\ \cdot \\ c_i A^{n-2} B \end{bmatrix} = \begin{bmatrix} c_i B \\ c_i AB \\ \cdot \\ \cdot \\ c_i A^{n-2} B \\ \underline{P}^T M \end{bmatrix} \tag{3.28}$$

where

$$\underline{P}^T = (P1, \ldots, PN) \tag{3.29}$$

is the last row of $\bar{H}$. Hence,

$$HB - \bar{H}\bar{M} = \begin{bmatrix} 0 & \cdots & 0 \\ \cdot & \cdot & \cdot \\ 0 & \cdots & 0 \\ c_i A^{n-1} B - \underline{P}^T M \end{bmatrix}. \tag{3.30}$$

For the coefficient of $u(k+1)$ the needed matrix product is

$$\bar{H}M1 = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & \cdot & \cdot & \cdots & 0\ 1 \\ P1 & \cdot & \cdot & \cdots & PN \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ c_i B \\ \cdot \\ \cdot \\ c_i A^{n-3} B \end{bmatrix} = \begin{bmatrix} 0 \\ c_i B \\ \cdot \\ \cdot \\ \cdot \\ c_i A^{n-3} B \\ \underline{P}^T M1 \end{bmatrix}. \quad (3.31)$$

The dependence coefficient matrix is

$$M - \bar{H}M1 = \begin{bmatrix} 0 & \cdots & 0 \\ \cdot & \cdots & \cdot \\ 0 & \cdots & 0 \\ c_i A^{n-2} B - \underline{P}^T M1 \end{bmatrix}. \quad (3.32)$$

Similarly, $\underline{u}(k+2)$ uses the matrix product

$$\bar{H}M2 = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & \cdot & \cdot & \cdots & 0\ 1 \\ P1 & \cdot & \cdot & \cdots & PN \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ c_i B \\ \cdot \\ \cdot \\ c_i A^{n-4} B \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ c_i B \\ \cdot \\ \cdot \\ c_i A^{n-4} B \\ \underline{P}^T M2 \end{bmatrix}. \quad (3.33)$$

so that the dependence coefficient matrix for $\underline{u}(k+2)$ is

$$M1 - \bar{H}M2 = \begin{bmatrix} 0 & \cdots & 0 \\ \cdot & \cdots & \cdot \\ 0 & \cdots & 0 \\ c_i A^{n-3} B - \underline{P}^T M2 \end{bmatrix}. \quad (3.34)$$

The coefficient matrix of the input $\underline{u}(k+n-2)$ uses the product

$$\bar{H}M_{n-2} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 & 1 \\ P1 & \cdot & \cdot & \cdot & \cdot & \cdot & PN \end{bmatrix} \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ c_A b \end{bmatrix} = \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ c_A b \\ \underline{P}^T M_{n-2} \end{bmatrix} \qquad (3.35)$$

with the resulting dependence coefficient matrix

$$M_{n-3}\bar{H}M_{n-2} = \begin{bmatrix} 0 & \cdots & 0 \\ \cdot & \cdot & \cdot \\ 0 & \cdots & 0 \\ c_A Ab \; \underline{P}^T M_{n-2} \end{bmatrix}.$$

The latest measurement of the input $\underline{u}(k+n-1)$ requires no calculation in its coefficient matrix.

The next series of dependence coefficient matrices multiply iterations of the state noise input. These terms are completely analogous to those for the input and the corresponding matrices can be obtained from the previous results if M is replaced with N and b with I. This completes the evaluation of dependence coefficient matrices in eq. (3.25).

The first observation to be made is that all matrices except H have zero in the first n-1 rows. The first n-1 rows of eq. (3.25) are

$$ym_i(k+1) = ym_i(k+1) - v_i(k+1) + v_i(k+1)$$

$$\vdots$$

$$ym_i(k+n-1) = ym_i(k+n-1) - v_i(k+n-1) + v_i(k+n-1) \qquad (3.36)$$

due to the structure of H. These rows are trivial but the last row is an equation from which parameters can be identified and is the decentralized measurement equation or DME. In Table 3.1 a summary of the dependence coefficients and their corresponding signals is made along with vector variables defining the parameters. The resulting DME can then be written

$$\begin{aligned} ym_i(k+n) &= \underline{p}^T \overline{ym}\ (k) + \underline{pbn}^T \underline{u}(k+n-1) + \underline{pbn-1}^T \underline{u}(k+n-2) \\ &+ \ldots + \underline{pb3}^T \underline{u}(k+2) + \underline{pb2}^T \underline{u}(k+1) + \underline{pb1}^T \underline{u}(k) \\ &+ \underline{pwn}^T \underline{w}(k+n-1) + \underline{pwn-1}^T \underline{w}(k+n-2) + \\ &+ \ldots + \underline{pw3}^T \underline{w}(k+2) + \underline{pw2}^T \underline{w}(k+1) + \underline{pw1}^T \underline{w}(k) \\ &- \underline{p}^T \overline{v}_i(k) + v_i(k+n) \ . \end{aligned} \qquad (3.37)$$

To see what is identified as coefficients related to the elements of the state-space description requires knowing the order of the system, n. Then the unknown combinations of elements of A, B, c_i are algebraically determined using the vector/matrix relations in Table 3.1. This procedure, summarized at the end of this chapter, is illustrated by detailed examples in Chapter 4. The system in Luenberger row companion canonical form is realizable using the procedure given by eqs. (B.19) - (B.22) in Appendix B. There

the DMc with all noise set equal to zero is directly given by eq. (b.17). Decentralized system identification is thus achieved.

TABLE 3.1

DEPENDENCE COEFFICIENTS OF INPUT TERMS IN DAE

INPUT	DEPENDENCE COEFFICIENT VECTOR
$\underline{u}(k)$	$\underline{pb1} = (c_\lambda A^{n-1} B - \underline{p}^T M)^T$
$\underline{u}(k+1)$	$\underline{pb2} = (c_\lambda A^{n-2} B - \underline{p}^T M1)^T$
$\underline{u}(k+2)$	$\underline{pb3} = (c_\lambda A^{n-3} B - \underline{p}^T M2)^T$
•	
•	
•	
$\underline{u}(k+n-2)$	$\underline{pbn-1} = (c_\lambda AB - \underline{p}^T Mn-2)^T$
$\underline{u}(k+n-1)$	$\underline{pbn} = (c_\lambda b)^T$
$\underline{w}(k)$	$\underline{pw1} = (c_\lambda A^{n-1} I - \underline{p}^T N)^T$
$\underline{w}(k+1)$	$\underline{pw2} = (c_\lambda A^{n-2} I - \underline{p}^T N1)^T$
$\underline{w}(k+2)$	$\underline{pw3} = (c_\lambda A^{n-3} I - \underline{p}^T N2)^T$
•	
•	
•	
$\underline{w}(k+n-2)$	$\underline{pwn-1} = (c_\lambda AI - \underline{p}^T Nn-2)^T$
$\underline{w}(k+n-1)$	$\underline{pwn} = (c_\lambda I)^T$

3.3 System Order Determination - Scalar Case

In section 3.2 a measurement equation was developed from which parameters (dependence coefficients) related to the elements of the A, b, and C matrices in the linear, state-space description may be determined. The section concluded in the statement that the system order must be known to see what elements and how they are related to the dependence coefficients are identified. The problem now is to find the order of the global system. Given the linear time-invariant discrete system:

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

$$y_i(k) = c_i x(k)$$

in which $y_i(k)$ and $u(k)$ are known this section will find the order of the system if (A, c_i) is an observable pair or the observability index if (A, c_i) is an unobservable pair.

For the scalar measurement case one output of a MIMO system is measurable and as explained in section 3.2 all inputs from all subsystems entering the global system must be known. This output and these inputs are presumed to be available for as many iterations as is necessary.

How the procedure works can be seen from the batch or least squares approach to identifying the coefficients in eq. (3.37). The approach is more fully studied in Chapter 4. Neglecting noise terms eq. (3.37) reduces to

$$y_i(k+n) = \underline{p}^T \underline{y}_i(k) + \underline{pb_n}^T \underline{u}(k+n-1) + \underline{pb_{n-1}}^T \underline{u}(k+n-2) + \ldots + \underline{pb_1}^T \underline{u}(k) \qquad (3.38)$$

with all signals available after N=n(r+1) iterations of eq. (3.38) the measurements can be concatenated into the fully determined matrix equation

$$\begin{bmatrix} y_i(k+n) \\ y_i(k+n+1) \\ \vdots \\ y_i(k-1) \end{bmatrix} = \begin{bmatrix} y_i(k+n-1) & \cdots & y_i(k) & \underline{u}^T(k+n-1) \\ y_i(k+n) & & y_i(k+1) & \underline{u}^T(k+n) \\ & & & \\ & & & \\ y_i(K-2) & & y_i(k+N-1) & \underline{u}^T(k-2) \end{bmatrix}$$

$$\begin{bmatrix} \cdots & \underline{u}^T(k) \\ & \underline{u}^T(k+1) \\ & \vdots \\ & \underline{u}^T(k+N-1) \end{bmatrix} \begin{bmatrix} \underline{p} \\ \underline{pbn} \\ \underline{pbn-1} \\ \vdots \\ \underline{pb1} \end{bmatrix} \qquad (3.39)$$

where K=k+n+N. More compactly

$$\underline{Y}_i(k+n) = S(k)\underline{P}$$

so that the parameters can be solved for by

$$\underline{P} = S^{-1}(K)\underline{Y}_i(K+n) \ . \qquad (3.40)$$

The input-output sequence must be identifiable (see R.C.K. Lee and Appendix A) and S must be nonsingular because if it is not then S^{-1} doesn't exist for the existence of a unique solution. Identifiability requires that the global system be observable from y_i. The case in which it is unobservable will be analyzed shortly.

If $N>n(r+1)$ iterations are used the additional measurements will give a better estimate if data is noisy. Without noise the estimate is the same as for $N=n(r+1)$. Either way the measurement matrix, S, is rectangular, $N\times n(r+1)$, and

$$S^T \underline{Y}_i(k+n) = S^T S\underline{P}$$

$$(S^T S)^{-1} S^T \underline{Y}_i(k+n) = \underline{P}$$

or

$$\underline{P} = S^+ \underline{Y}_i(k+n) \quad . \tag{3.41}$$

The matrix

$$S^+ = (S^T S)^{-1} S^T$$

is called the pseudo-inverse of S and is the inverse of S if $N=n(r+1)$. The justification for the following approach for structural invariant and system order determination can be readily seen in light of the least squares solution.

Sequences of inputs and outputs are as follows

$$\begin{matrix} y(k) & \underline{u}^T(k) & y(k+1) & \underline{u}^T(k+1) & \cdots \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \\ y(k+N-1) & \underline{u}^T(k+N-1) & y(k+N) & \underline{u}^T(k+N) & \cdots \end{matrix} \tag{3.42}$$

The columns can be tested beginning with the first two then adding one more column at a time and testing for linear dependence. These are precisely the columns used in the least squares estimate and if a linearly dependent column is found the inverse (or pseudo-inverse) would not exist in eq. (3.40) (or (3.41)) if retained.

The order of the system is then the number of output sequences retained. This assumes the system is observable from the output (see Appendix A and R.C.K. Lee). The unobservable case is examined later in this section.

The retained columns can then be used in either eq. (3.40) or eq. (3.41) depending on N as long as the condition holds that $N>n(r+1)$. Practically, N must be assumed large enough. Returning to the problem with inputs present the columns are examined in the order

$$\hat{y}(0)\ \hat{u}_1(0)\ \ldots\ \hat{u}_r(0)\ \hat{y}(1)\ \hat{u}_1(1)\ \ldots\ \hat{u}_r(1)$$
$$\hat{y}(2)\ \hat{u}_1(2)\ \ldots \qquad (3.42)$$

The dependence test need not be performed for new columns of inputs since they do not relate to the system order. They are necessary as Shrikhande, etal. point out to identify the nonaynamic part of the state model. This is because the wrong inputs would make the S matrix singular in eq. (3.39), for example, a sequence of all zero inputs would cause the dependence test and hence the identification test to fail.

Suppose that the output is such that the global system is not observable from the single output, y_i. Then from the conditions for identifiability (see R.C.K. Lee, Liu and Suen, and Ise and Anton) the global system is not identifiable in phase-variable canonical form. This problem, unaddressed in the literature on identification until now, is resolved here.

Less than the entire A, b, and C matrix triple can be identified and exactly how much can be determined is given by using the observability canonical form. Recalling from Chapter 2 (also see Sivan and Kwakernaak) a nonsingular transformation, T, exists to transform the unobservable system to the following form

$$\underline{x}'(k+1) = \begin{bmatrix} A'_{11} & 0 \\ A'_{21} & A'_{22} \end{bmatrix} \underline{x}'(k) + \begin{bmatrix} b'_1 \\ b'_2 \end{bmatrix} \underline{u}(k)$$

$$y(k) = \begin{bmatrix} C'_1 & 0 \end{bmatrix} \underline{x}'(k) \qquad (3.45)$$

in which the pair (A'_{11}, C'_1) is completely observable.

The observability matrix from y_i is

$$\zeta = \begin{bmatrix} c_i \\ c_i A \\ \cdot \\ \cdot \\ \cdot \\ c_i A^{N-1} \end{bmatrix} . \qquad (3.46)$$

With regard to finding dependent vectors the number of columns at any point retained in eq. (3.42) can be tested by forming the inner product of that set of columns with itself, the determinant of the inner product will be zero if a dependency occurs. If it happens that more columns are retained than the length of a column, N, then N is too short and a longer length with more measurements needed. Another technique (which is used in this thesis) is to find

the eigenvalues of the matrix formed by taking the inner product mentioned above. If one or more zero singular values appear then a dependency has occurred. An IMSL (International Mathematical Societies Library) library subroutine, LSVDF, used in this thesis orders the singular values from largest to smallest in a vector.

The rank of Q, $m<n$, is the observability index. There are m linearly independent rows in Q which span the observable subspace of (A, c_i). This or any other m rows spanning this space are used in the transformation matrix, T. The remaining n-m rows span the unobservable space or null space of Q.

When the input/output sequence dependence test is performed properly, i.e., with $N>m(r+1)$ inclusion of the $(m(r+1)+1)$th output sequence results in a dependency. This is based on the same considerations as the observable case discussed previously. The global system order cannot be found.

An mth order system is then identifiable. Using m instead of n in Table 3.1 the parameters of any realization that can be identified are found. From Chapter 2 it was seen that the observability canonical form is not unique. The observable mth order system can be identified in phase-variable canonical form. The parameters obtained in an identification scheme can be used in Appendix B to realize an mth order system.

Some issues to note are: first, although the global system is not identifiable in any form the dynamics that determine the available output or, in other words, the m observable poles, can be found. The stability of the available response for the given subsystem can then be ascertained because the relevant system state-space description is known. Second, the controllability of the system producing the available response can be ascertained since the input matrix is identifiable under requirement of this thesis that all inputs be known. Finally, under the structure of decentralized systems studied in this thesis in which only a subset of all inputs entering the global system can be controlled by a given subsystem then the issue of decentralized control enters the picture. This is further explored in Chapter 6 where decentralized control results are reviewed and applied to the problem.

3.4 Derivation of the Vector Decentralized Measurement Equation

This section parallels section 3.2 where each subsystem of a MIMO system receives one output. Here each subsystem is assumed to receive a subvector of outputs of the total output vector from the global system. Due to the added complexity additional notation is introduced. The global system generates m outputs and using L for the total number

of subsystems. m_ℓ will be used for the number of outputs entering the ℓth subsystem. For the sake of allowing the first output of each subvector to have the subscript "1" a super-script "1" will be assigned to each output to indicate which subsyster it is assigned to. The global system can be described in row companion canonical form with structural invariants p_1 to p_m. Each subsystem has m_ℓ structural invariants p_1^ℓ to $p_{m_\ell}^\ell$ for m_ℓ outputs. In this section we assume the structural invariants are known. Section 3.5 gives a procedure for finding the structural invariants.

The development to find a vector measurement equation now proceeds for an ith measurement of an ℓth subsystem. The ith output propagates as

$$ym_i(k) = c_i \underline{x}(k) + v_i(k)$$

$$ym_i(k+1) = c_i \underline{x}(k+1) + v_i(k+1)$$

$$= c_i A \underline{x}(k) + c_i B \underline{u}(k) + c_i \Gamma \underline{w}(k) + v_i(k+1)$$

$$\vdots$$

$$ym_i(k+p_i^\ell - 1) = c_i A^{p_i^\ell - 1} \underline{x}(k) + c_i A^{p_i^\ell - 2} B \underline{u}(k) + \ldots$$

$$+ c_i B \underline{u}(k+p_i^\ell - 2) + c_i A^{p_i^\ell - 2} \Gamma \underline{w}(k) + \ldots$$

$$+ c_i \Gamma \underline{w}(k+p_i^\ell - 2) + v_i(k+p_i^\ell - 1) \quad . \qquad (3.47)$$

These sets of measurements will be concatenated and the resulting equation stepped up. The state is eliminated using the state equation and a vector measurement equation results with dependence coefficients that can be identified. The m_ℓ sets of measurements are now concatenated

$$
\begin{bmatrix}
y_{m_1}^{\ell}(k) \\
\vdots \\
y_{m_1}^{\ell}(k+p_{m_1}^{\ell}-1) \\
\vdots \\
y_{m_M}^{\ell}(k) \\
\vdots \\
y_{m_M}^{\ell}(k+p_{m_M}^{\ell}-1) \\
\vdots \\
y_{m_{m_\ell}}^{\ell}(k) \\
\vdots \\
y_{m_{m_\ell}}^{\ell}(k+p_{m_\ell}^{\ell}-1)
\end{bmatrix}
=
\begin{bmatrix}
c_{m_1}^{\ell} \\
\vdots \\
c_{m_1}^{\ell} A^{p_{m_1}^{\ell}-1} \\
\vdots \\
c_{m_M}^{\ell} \\
\vdots \\
c_{m_M}^{\ell} A^{p_{m_M}^{\ell}-1} \\
\vdots \\
c_{m_{m_\ell}}^{\ell} \\
\vdots \\
c_{m_{m_\ell}}^{\ell} A^{p_{m_\ell}^{\ell}-1}
\end{bmatrix}
\underline{x}(k)
$$

$$
+
\begin{bmatrix}
0 \\
c_1^{\ell} b \\
\vdots \\
c_1^{\ell} A^{\nu_1^{\ell}-2} B \;\cdots\; c_1^{\ell} B \\
\vdots \\
c_M^{\ell} b \\
\vdots \\
c_M^{\ell} A^{\nu_M^{\ell}-2} B \;\cdots\; c_M^{\ell} B \\
\vdots \\
c_{m_\ell}^{\ell} B \\
\vdots \\
c_{m_\ell}^{\ell} A^{\nu_{m_\ell}^{\ell}-2} B \;\cdots\; c_{m_\ell} B
\end{bmatrix}
\begin{bmatrix}
\underline{u}(K) \\
\vdots \\
\underline{u}(K+\nu_M^{\ell}-2)
\end{bmatrix}
$$

$$
+
\begin{bmatrix}
0 & & & \\
c_1^{\ell} I & & & \\
\vdots & & & \\
c_1^{\ell} A^{p_1^{\ell}-2} & I & \cdots & c_1^{\ell} I \\
\vdots & & & \\
c_M^{\ell} I & & & \\
\vdots & & & \\
c_M^{\ell} A^{p_M^{\ell}-2} & I & \cdots & c_M^{\ell} I \\
\vdots & & & \\
c_{m_\ell}^{\ell} I & & & \\
\vdots & & & \\
c_{m_\ell}^{\ell} A^{p_{m_\ell}^{\ell}-2} & I & \cdots & c_{m_\ell}^{\ell} I
\end{bmatrix}
\begin{bmatrix}
\underline{w}(k) \\
\vdots \\
\underline{w}(k+p_M^{\ell}-2)
\end{bmatrix}
$$

$$
+
\begin{bmatrix}
v_1^{\ell}(k) \\
\vdots \\
v_{m_\ell}^{\ell}(k+p_{m_\ell}^{\ell}-1)
\end{bmatrix}
.
$$

More compactly the measurement equation is written

$$\underline{Y}^{\ell}(k) = H\underline{x}(k) + M\underline{U}(k) + N\underline{W}(k) + \underline{V}^{\ell}(k) \tag{3.49}$$

from which the state can be solved for:

$$\underline{x}(k) = H^{-1}\underline{Y}^{\ell}(k) - H^{-1}M\underline{U}(k) - H^{-1}N\underline{W}(k) - H^{-1}\underline{V}^{\ell}(k) \ . \tag{3.50}$$

Stepping up eq. (3.49) gives

$$\underline{Y}^{\ell}(k+1) = H\underline{x}(k+1) + M\underline{U}(k+1) + N\underline{W}(k+1) + \underline{V}^{\ell}(k+1) \ . \tag{3.51}$$

Next substitute for $\underline{x}(k+1)$ from the state description in eq. (2.1) so that

$$\underline{Y}^{\ell}(k+1) = HA\underline{x}(k) + HB\underline{U}(k) + HI\underline{W}(k) + M\underline{U}(k+1) + N\underline{W}(k+1) + \underline{V}^{\ell}(k+1) \ . \tag{3.52}$$

Finally, substitute for $\underline{x}(k)$ from eq. (3.50) to give

$$\underline{Y}^{\ell}(k+1) = HAH^{-1}\underline{Y}^{\ell}(k) - HAH^{-1}M\underline{U}(k) - HAH^{-1}N\underline{W}(k) - HAH^{-1}\underline{V}^{\ell}(k) + HB\underline{U}(k) + HI\underline{W}(k) + M\underline{U}(k+1) + N\underline{W}(k+1) + \underline{V}^{\ell}(k+1) \ . \tag{3.53}$$

If the input sequence $\bar{U}(k)$ is defined

$$\bar{U}(k) = \begin{bmatrix} \underline{U}(k) \\ \cdot \\ \cdot \\ \cdot \\ \underline{U}(k+p_M^{\ell}-1) \end{bmatrix} \tag{3.54}$$

eq. (3.53) can be rewritten in the form

$$\underline{XM}^{\ell}(k+1) = HAH^{-1}\underline{XM}^{\ell}(k)$$
$$+ \left(-HAH^{-1}M \mid 0 \right)\bar{U}(k) + \left(HB \mid M \right)\bar{U}(k)$$
$$+ \left(-HAH^{-1}N \mid 0 \right)\bar{W}(k) + \left(HI \mid N \right)\bar{W}(k)$$
$$- HAH^{-1}\underline{V}^{\ell}(k) + \underline{V}^{\ell}(k+1) \qquad (3.55)$$

and like terms combined.

Equation (3.55) will be simplified after taking note of the structure of HAH^{-1}. The H matrix is the same matrix used to transform the A matrix to row companion form so that

HAH^{-1} =

$$\begin{bmatrix}
0 & & & & & & & & & & & & & & & & \\
\cdot & & & & & & & & & & & & & & & & \\
\cdot & & I_{p_1^{\ell}-1} & & & & & & & & & & & & & & \\
\cdot & & & & & & & & & & & & & & & & \\
0 & & & & & & & & & & & & & & & & \\
x & x & x & x & x & x & x & x & x & \cdot & \cdot & \cdot & x & x & x & x & x & x & x & x \\
\cdot & & & & & & & & & & & & & & & & \\
\cdot & & & & & & & & & & & & & & & & \\
\cdot & & & & & & & & & & & & & & & & \\
0 & & & & & & & & 0 & & & 0 & & & & & \\
\cdot & & & & & & & & & & & & & & & & \\
\cdot & & & & & & & & & & I_{p_{m_{\ell}}^{\ell}-1} & & & & & & \\
\cdot & & & & & & & & & & & & & & & & \\
0 & & & & & & & & 0 & & & 0 & & & & & \\
x & x & x & x & x & x & x & x & x & \cdot & \cdot & \cdot & x & x & x & x & x & x & x & x
\end{bmatrix} \qquad (3.56)$$

where the x's are arbitrary distinct numbers. Next the following matrices are partitioned

$$M = (M_1 \quad \cdots \quad M_{p_m^{\ell}-1})$$

$$N = (N_1 \quad \cdots \quad N_{p_m^{\ell}-1}) \; . \tag{3.57}$$

The n rowed matrix (HB ¦ M) has the following p_i^{ℓ}th through $p_{m_\ell}^{\ell}$th or in other words significant rows

$$c_i^{\ell} A^{p_i^{\ell}-1} b \mid c_i^{\ell} A^{p_i^{\ell}-2} B \mid \cdots \mid c_i^{\ell} b$$

$$\vdots$$

$$c_{m_\ell}^{\ell} A^{p_{m_\ell}^{\ell}-1} b \mid c_{m_\ell}^{\ell} A^{p_{m_\ell}^{\ell}-2} b \mid \cdots \mid c_{m_\ell}^{\ell} B \; . \tag{3.58}$$

and the n rowed matrix (HT ¦ N) has significant rows

$$c_i^{\ell} A^{p_i^{\ell}-1} T \mid c_i^{\ell} A^{p_i^{\ell}-2} T \mid \cdots \mid c_i^{\ell} T$$

$$\vdots$$

$$c_{m_\ell}^{\ell} A^{p_{m_\ell}^{\ell}-1} T \mid c_{m_\ell}^{\ell} A^{p_{m_\ell}^{\ell}-2} T \mid \cdots \mid c_{m_\ell}^{\ell} T \; . \tag{3.59}$$

Also, let the rows of x's in eq. (3.56), the significant rows of HAH^{-1}, be

$$\underline{a}_1$$

$$\vdots$$

$$\underline{a}_{m_\ell}$$

for $i=1, \ldots, p_\ell$.

All rows of eq. (3.55) are trivial except the p_1^{ℓ}th, p_2^{ℓ}th, etc. or significant rows. For example, the first $p_1^{\ell}-1$ rows of eq. (3.55) have the form

$$ym_i^{\ell}(k+1) = ym_i^{\ell}(k+1) + (c_i^{\ell} B \mid 0 \ldots 0)\bar{U}(k)$$
$$- (c_i^{\ell} B \mid 0 \ldots 0)\bar{U}(k) + (c_i^{\ell} T \mid 0 \ldots 0)\bar{W}(k)$$
$$- (c_i^{\ell} T \mid 0 \ldots 0)\bar{W}(k) + v_i^{\ell}(k+1) - v_i^{\ell}(k+1)$$

$$\vdots$$

$$ym_i^{\ell}(k+p_i^{\ell}-1) = ym_i^{\ell}(k+p_i^{\ell}-1)$$
$$+ (c_i^{\ell} A^{p_{m_\ell}^{\ell}-2} B \mid 0 \ldots 0)\bar{U}(k)$$
$$- (c_i^{\ell} A^{p_{m_\ell}^{\ell}-2} B \mid 0 \ldots 0)\bar{U}(k)$$
$$+ (c_i^{\ell} A^{p_{m_\ell}^{\ell}-2} T \mid 0 \ldots 0)\bar{W}(k)$$
$$- (c_i^{\ell} A^{p_{m_\ell}^{\ell}-2} T \mid 0 \ldots 0)\bar{W}(k)$$
$$+ v_i^{\ell}(k+p_i^{\ell}-1) - v_i^{\ell}(k+p_i^{\ell}-1) .$$

For $i=p_i^{\ell}$ th to $p_{m_\ell}^{\ell}$th, the significant rows of eq. (3.55), are decentralized measurement equations from which parameters can be identified:

$$ym_{\lambda}^{\ell}(k+p_{\lambda}^{\ell}) = \underline{a}_{\lambda}\underline{YM}^{\ell}(k)$$
$$- (\underline{a}_{\lambda}M_1 \mid \ldots \mid \underline{a}_{\lambda}M_{p_{\lambda}^{\ell}-1} \mid 0)\bar{U}(k)$$
$$+ (c_{\lambda}^{\ell} A^{p_{\lambda}^{\ell}-1} B \mid \ldots \mid c_{\lambda}^{\ell} B)\bar{U}(k)$$
$$- (\underline{a}_{\lambda}N_1 \mid \ldots \mid \underline{a}_{\lambda}N_{p_{\lambda}^{\ell}-1} \mid 0)\bar{W}(k)$$
$$+ (c_{\lambda}^{\ell} A^{p_{\lambda}^{\ell}-1} T \mid \ldots \mid c_{\lambda}^{\ell} T)\bar{W}(k)$$
$$- \underline{a}_{\lambda}\underline{V}^{\ell}(k) + v_{\lambda}(k+p_{\lambda}^{\ell}) . \qquad (3.61)$$

Note that no "1" superscript goes with $\bar{U}(k)$ or $\bar{W}(k)$ since these sequences represent centralized information.

Chapters 4 and 5 explore means of identifying the vector coefficients of eq. (3.61). The significant rows of the A matrix in row companion canonical form are identifiable at

sufficient excitation is present on all signals and that an appropriate identification algorithm is used. As in the scalar case the order must be known to apply the results of this section. In addition since there is more than one structural invariant in the vector case this must also be known in order to apply the correct order decentralized measurement vector equation. How to obtain this information is given in the next section.

3.5 System Order and Structural Invariant Determination - Vector Case

This section has some similarity to section 3.3 where sequences of a single output and all the inputs are available. The presence of more than one output, however, complicates the procedure and introduces some non-uniqueness to the vector output case. First, eq. (3.61) will be explicitly written without the noise terms which only obfuscate the theory behind the procedure. Then a batch equation will written analogous to eq. (3.39) that will clearly show why the procedure works.

For the ith of m_l measurements that the lth subsystem receives, dropping the l superscript and noise terms, eq. (3.61) becomes

$$y_i(k+p_i) = \left[y_1(k) \;\cdots\; y_1(k+p_1-1) \;\cdots\; y_{m_i}(k) \;\cdots\; y_{m_i}(k+m_i-1) \;\; \underline{u}^T(k) \;\cdots\; \underline{u}^T(k+p_M-1) \right] \times \begin{bmatrix} \underline{a}_i \\ --- \\ \underline{b}^* \end{bmatrix} \tag{3.62}$$

For m_i outputs after $N>n+1p_{M_i}$ iterations the following batch equation is written

$$\begin{bmatrix} y_1(k+p_1) & \cdots & y_{m_i}(k+p_{m_i}) \\ \vdots & & \vdots \\ y_1(k+p_1+N-1) & \cdots & y_{m_i}(k+p_{m_i}+n-1) \end{bmatrix}$$

$$= \begin{bmatrix} y_1(k) & \cdots & y_1(k+p_1-1) & \cdots & y_{m_i}(k) \\ \vdots & & & & \\ y_1(k+N-1) & \cdots & y_1(k+p_1+N-2) & \cdots & y_{m_i}(k+N-1) \end{bmatrix}$$

$$\begin{bmatrix} \cdots & y_{m_i}(k+p_{m_i}) & \underline{u}^T(k) & \cdots & \underline{u}^T(k+p_M-1) \\ & \vdots & & & \\ \cdots & y_{m_i}(k+p_{m_i}+N-1) & \underline{u}^T(k+N-1) & \cdots & \underline{u}^T(k+p_M+N-2) \end{bmatrix} \tag{3.63}$$

$$\begin{bmatrix} \underline{a}_1 & \cdots & \underline{a}_{m_i} \\ \underline{b}^* & \cdots & \underline{b}^*_{m_i} \end{bmatrix}$$

More compactly (3.63) is written

$$Y_{m_i} = \Im P_{m_i} \tag{3.64}$$

which is an Nth order concatenated matrix equation. For now the global system will be assumed to be observable from the m_f outputs and have an input sequence such that S is nonsingular in eq. (3.64). The system is then identifiable.

As in the scalar case the columns that appear in eq. (3.63) are precisely those used in the following dependence test. Before discussing the non-uniqueness of the dependence test first the technique suggested by Bonivento and Guidorzi is reviewed. In that paper columns would be tested in this order

$$\hat{y}_1(0) \ \ldots \ \hat{y}_{m_f}(0) \ \hat{u}^T(0) \ \hat{y}_1(1) \ \ldots \ \hat{y}_{m_f}(1) \ \hat{u}^T(1) \ \ldots \quad (3.65)$$

where input sequences are not tested but included to ensure that the inverse in eq. (3.64) exists. Each output has a structural invariant associated with it. When an output sequence is tested that is found to be dependent on previous columns it is discarded and the structural invariant for that output equals the number of columns of that output already retained. This testing must be performed for all outputs and the sum of all structural invariants equals the global system order.

The following example from Irwin and Roberts

$$\underline{x}(K+1) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & .7 & .2 \\ .1 & -.7 & .3 \end{bmatrix} \underline{x}(k) + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u(k)$$

$$\underline{y}(K) = \begin{bmatrix} 1 & 0 & 0 \\ & & \\ 0 & 0 & 1 \end{bmatrix} \underline{x}(k) \qquad (3.66)$$

has the following inputs applied and has the corresponding outputs

k	u(k)	$y_1(k)$	$y_2(k)$
0	1	1	1
1	-1	1	.5
2	0	.5	.37
3	-1	.37	.015
4	-1	.015	-.1045
5	1	-.1045	-.3024
6	0	-.3024	-.4655
7	-1	-.4655	-.1618
8	0	-.1618	-.0049

(3.67)

The sequence length N must be greater than 5. Suppose a sequence length of 6 is guessed. Using eq. (3.65) columns are tested in the order

$$\hat{y}_1(0)\ \hat{y}_2(0)\ \hat{u}(0)\ \hat{y}_1(1)\ \hat{y}_2(1)\ \hat{u}(1)\ \hat{y}_1(2)\ \ldots \qquad (3.68)$$

Then $\hat{y}_2(1)$ and $\hat{y}_1(2)$ are found to be dependent so that $\mu_1=2$ and $\mu_2=1$ as in the A matrix of eq. (3.66).

Now suppose the following sequence had been tested:

$$\hat{y}_2(0)\ \hat{y}_1(0)\ \hat{u}(0)\ \hat{y}_2(1)\ \hat{y}_1(1)\ \hat{u}(1)\ \hat{y}_2(2)\ \ldots \qquad (3.69)$$

Columns $\hat{y}_1(1)$ and $\hat{y}_2(2)$ are found to be dependent with $\rho_1=1$ and $\rho_2=2$ whose sum must equal the order of the system or 3. The system can then be identified in the form

$$\underline{x}(k+1) = \begin{bmatrix} .143 & .429 & -1.43 \\ 0 & 0 & 1 \\ -.056 & -.307 & .857 \end{bmatrix} \underline{x}(k) + \begin{bmatrix} 0 \\ 1 \\ .3 \end{bmatrix} u(k)$$

$$\underline{y}(k) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \underline{x}(k) \qquad (3.70)$$

The same conclusion would be reached if this sequence were used:

$$\hat{y}_1(0)\ \hat{y}_2(0)\ \hat{y}_2(1)\ \hat{u}(0)\ \hat{u}(1)\ \hat{y}_1(1)\ \hat{y}_2(2)$$
$$\hat{y}_2(3)\ \hat{u}(2)\ \hat{u}(3)\ \ldots \qquad (3.71)$$

Just as the canonical form is non-unique in general so too is the dependence test plan. How many output sequences one can test at one interval between input sequences is non-unique. Using more sequences of a particular output could depend on such factors as convenience, desirability of a certain set of structural invariants, or other application oriented reasons. Different numbers will give different canonical forms. In general the output sequence can take the form

$$\hat{y}_1(0)\ \ldots\ \hat{y}_1(q_1)\ \hat{y}_2(0)\ \ldots\ \hat{y}_2(q_2)\ \ldots\ \hat{y}_{m_\ell}(0)\ \ldots$$
$$\hat{y}_{m_\ell}(q_{m_\ell})\ \hat{\underline{u}}^T(0)\ \ldots\ \hat{\underline{u}}^T(q_M)\ \hat{y}_1(q_1+1)\ \ldots$$

in which the q_i are guessed by the user with the constraint that $\hat{y}_i(q_i)$ is an independent column and where $q_M=\max(q_i)$, $i=1,\ldots,m_\ell$.

Once the p_i are known the system order is known and the system matrix elements that can be identified are given in eq. (3.61).

Next suppose the global system is not observable from the given m_ℓ outputs. By definition of observability n linearly independent rows cannot be found for H in eq. (3.49) and the global system is not identifiable. What is identifiable is found by using the observability canonical form for multi-output systems.

The multi-output system addressed is the global system along with m_ℓ outputs decentralized or separated from the others described by

$$\underline{x}(k+1) = A\underline{x}(k) + B\underline{u}(k)$$
$$\underline{y}^\ell = C^\ell \underline{x}(k) \qquad (3.73)$$

where C^ℓ is an $m_\ell Xn$ matrix. The observability matrix from this set of outputs is

$$\begin{bmatrix} C^\ell \\ C^\ell A \\ \cdot \\ \cdot \\ \cdot \\ C^\ell A^{n-1} \end{bmatrix} \qquad (3.74)$$

which is rank deficient by assumption. Transformation to observability canonical form is obtained by using a set of n linearly independent rows that span the observability

subspace of eq. (3.74) (see Sivan and Kwakernaak). Any n linearly independent rows in eq. (3.74) form such a set. The rest of the rows in the transformation matrix must span the unobservable subspace or null space of the matrix in eq. (3.74). Let the n independent rows be called T_1. Then n-n rows can be found by solving for n-m independent rows that solve the equation

$$T_1 \underline{q} = \underline{0} \tag{3.75}$$

where $\underline{q}$ is an n vector. Let these n-m rows be called matrix T_2. The state transformation

$$\underline{x}'(k) = \begin{bmatrix} T_1 \\ \hline T_2 \end{bmatrix} \underline{x}(k) = T\underline{x}(k) \tag{3.76}$$

transforms eq. (3.73) to

$$\underline{x}'(k+1) = \begin{bmatrix} A'_{11} & 0 \\ A'_{21} & A'_{22} \end{bmatrix} \underline{x}'(k) + \begin{bmatrix} B'_1 \\ B'_2 \end{bmatrix} \underline{u}(k)$$

$$y^{\ell}(k) = (C'_{\ell} \mid 0) \underline{x}'(k) \tag{3.77}$$

in which (A'_{11}, C'_{ℓ}) is an observable pair. The system the lth subsystem "sees" or can identify is

$$x'_{\ell}(k+1) = A'_{11} x'_{\ell}(k) + \begin{bmatrix} B'_1 \\ \hline B'_2 \end{bmatrix} u(k)$$

$$y^{\ell}(k) = C'_{\ell} \underline{x}'_{\ell}(k) \tag{3.78}$$

in which $\underline{x}'_{\ell}(k)$ are the first n states of $\underline{x}'(k)$.

Applying the dependence test to an input/output sequence of columns generated by eq. (3.77) or eq. (3.78) will yield

n_i structural invariants whose values will depend on A_i^*, C_i^*, and the sequence in which columns are tested as previously discussed. The global system cannot be identified in any form and its order cannot be determined. The order of the observable subspace of the global system is found. The nth order realization can be in the row companion form (see Appendix b) but the identified parameters can also realize other state descriptions.

In summary the linear time-invariant system given by

$$\underline{x}(K+1) = A\underline{x}(K) + B\underline{u}(K)$$
$$\underline{y}(K) = C\underline{x}(k) \qquad (3.79)$$

is identified in whole or in part by the following procedure when measurements are decentralized. First, the scalar case procedure is summarized and then the vector case is presented.

Given one measurement

$$y_i(K) = c_i\underline{x}(K) \qquad (3.80)$$

and information of the entire input vector, $\underline{u}(k)$, the first step is to find the system order or, in the case of unobservability, the observability index of the global system. Columns of the single output and of the inputs are ordered as follows

$$\hat{y}(0)\ \hat{u}_1(0)\ \ldots\ \hat{u}_r(0)\ \hat{y}(1)\ \hat{u}_1(1)\ \ldots\ \hat{u}_r(1)$$
$$\hat{y}(2)\ \hat{u}_1(2)\ \ldots \qquad (3.81)$$

and tested as follows:

1) Start with the first two columns of outputs and all input columns up to the time of the latest output column and test for dependence of the second output column. Input columns are not tested for dependence on previous columns but included to make sure the input matrix can be identified. The length of the columns is guessed and a length that is too short ($N<n(r+1)$) is readily discovered as will be shown later.

2) Continue testing new columns of outputs for linear dependence. If a dependency is found at say iteration n+1 then the system order is n if the global system is observable from $y_i(k)$. If unobservable the order obtained is the observability index of the global system from $y_i(k)$. If system observability from $y_i(k)$ is not known, the dependence test does not supply this information.

3) If a dependency occurs for $N=n(r+1)$ a longer column length is needed and the test repeated. The case of $N=n(r+1)$ will always result in a dependency because the space of the output columns is completely spanned at that point.

Second, n obtained from the dependence test is used in Table 3.1 and eq. (3.37) as the order of the decentralized measurement equation. From this equation the dependence coefficients are identified using an applicable identification algorithm.

Third and last the identified parameters are used in the procedure given in Appendix B starting with equation (B.17) on to model the system in row companion form.

In the vector case the vector of measurements

$$\underline{Y}^{\ell}(K) = C^{\ell}\,\underline{X}(K) \tag{3.62}$$

and measurements of $\underline{U}(K)$ are given. The first step is to find the structural invariants of the system and system order (observability index if $(A,\ C^{\ell})$ is an unobservable pair). Columns of outputs and inputs are ordered as follows for K=0:

$$\hat{Y}_1(0)\ \ldots\ \hat{Y}_1(q_1)\ \hat{Y}_2(0)\ \ldots\ \hat{Y}_2(q_2)\ \ldots\ \hat{Y}_{m_\ell}(0)\ \ldots$$
$$\hat{Y}_{m_\ell}(q_{m_\ell})\ \hat{\underline{U}}^{T}(0)\ \ldots\ \hat{\underline{U}}^{T}(q_M)\ \hat{Y}_1(q_1+1)\ \ldots \tag{3.63}$$

where $q_M=\max(q_i)$, $i=1,\ \ldots,\ m_\ell$. The test for structural invariants is as follows:

1) Start with q_1 through q_{m_ℓ} columns of outputs and test for linear independence. Include q_M input columns to find the input matrix but do not test them for linear dependence. The column length must be $N>n+rp$ which must be guessed. If the length is too short this is readily detected.

2) Continue adding columns until a dependency occurs then discard that column. The number of columns retained of that output is the structural invariant for that output. After a dependency is found for all outputs the sum of the structural invariants obtained is the order of the global system if $(A,\ C^{\ell})$ is an observable pair. If $(A,\ C^{\ell})$ is an unobservable pair this sum is the order of the subspace of the observability matrix for $(A,\ C^{\ell})$.

3) If it is found that $n+rp_M = N$ a longer column length is needed and the test repeated. This is because the space spanned by the columns in eq. (3.83) has been spanned before the order is found.

Second, the $p_i^{\hat{\lambda}}$ obtained are used in eq. (3.63) to determine the order of that equation. The coefficients in these m_ℓ equations are then identified using an applicable identification algorithm.

Third, and last, the identified parameters make up the matrices A^* and B^* in eq. (b.17). Then eqs. (b.18) - (b.22) are used to obtain a state description of the system in row companion form.

CHAPTER 4

IDENTIFICATION OF THE DECENTRALIZED MEASUREMENT EQUATION

4.1 Least Squares System Identification

Under the decentralized framework of this thesis scalar and vector measurement equations were derived in sections 3.2 and 3.4. This section applies the least squares method and its variations that have been developed over the past few decades (see Graupe b, Johnstone, Panuska, and Mendel) to the identification of the parameters related to the system matrices in the decentralized measurement equations. Algorithms will be applied to deterministic parameter identification and for varying degrees of complexity of the noise terms in eqs. (3.37) and (3.61). Examples will be given to find dependence coefficients in sections 4.1 and 4.2 and the state descriptions derived from those dependence coefficients will be given in section 4.3.

In the deterministic, no noise, case the global system state model is

$$\underline{x}(k+1) = A\underline{x}(k) + b\underline{u}(k)$$

$$y(k) = C\underline{x}(k) \qquad (4.1)$$

where the m is dropped to indicate no noise. The scalar decentralized measurement equation is

$$y_i(k+n) = \underline{p}^T \underline{\bar{y}}(k) + \underline{pbn}^T \underline{u}(k+n-1) + \underline{pbn-1}^T \underline{u}(k+n-2) + \ldots + \underline{pb1}^T \underline{u}(k). \quad (4.2)$$

Define the nondynamic coefficients in one vector

$$\underline{pb*} = (\underline{pbn} \mid \ldots \mid \underline{pb1})^T \quad (4.3)$$

so that eq. (4.2) is written in vector form as

$$y_i(k+n) = \left[\bar{y}^T(k) \quad \underline{u}^T(k+n-1) \mid \ldots \mid \underline{u}^T(k) \right] \begin{bmatrix} \underline{p} \\ --- \\ \underline{pb*} \end{bmatrix} \quad (4.4)$$

The corresponding equation in the vector case is similar with the understanding that $\bar{y}(k)$ would involve the whole vector of available measurements. This equation is in the form to be used in deterministic and some stochastic identification algorithms.

The batch solution to the parameter identification of unknown coefficients in a measurement equation was partly discussed in Chapter 3 to explain structural invariant and system order determination. The solution is obtained by concatenating iterations of eq. (4.4) into

$$\begin{bmatrix} y_i(k+n) \\ \cdot \\ \cdot \\ \cdot \\ y_i(k+n+N-1) \end{bmatrix} = \begin{bmatrix} \bar{y}^T(k) & \underline{u}^T(k+n-1) & \ldots & \underline{u}^T(k) \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ \bar{y}^T(k+N-1) & \underline{u}^T(k+n+N-2) & \ldots & \underline{u}^T(k+N-1) \end{bmatrix}$$

$$\lambda \begin{bmatrix} \underline{\theta} \\ --- \\ \underline{\theta}\underline{b}^{*} \end{bmatrix} \tag{4.5}$$

or

$$\bar{y}_{\lambda}(K) = S\underline{\theta} \quad . \tag{4.6}$$

If $k=n(r+1)$ then S is square and the parameters are found by

$$\underline{\theta} = S^{-1}\bar{y}_{\lambda}(K) \tag{4.7}$$

If $k>n(r+1)$ then S has more rows than columns the parameters can be solved for by

$$S^{T}\bar{y}_{\lambda}(K) = S^{T}S\underline{\theta}$$

$$\underline{\theta} = (S^{T}S)^{-1}S^{T}\bar{y}_{\lambda}(K)$$

$$\underline{\theta} = S^{+}\bar{y}_{\lambda}(K)$$

where S^{+} is the pseudo-inverse. Without noise the solution is the same as eq. (4.7). With noise and a memryless measurement equation the additional measurements give a better least squares estimate (see Mendel). This benefit applies to measurement equations with memory but the estimate will still have a bias. The amount of bias depends on the statistics of the noise as discussed in Chapter 3 also see Mendel.

Aside from the bias other disadvantages are the need to store large amounts of data. The inversion is not suitable for on-line identification especially for large order systems (see Sinha). In computer simulations an inverse calculation subroutine is needed. These subroutines for example in the IMsl library of subroutines use Gaussian

elimination which requires finding a pivot element. For high order a pivot element is often not found and the inverse not calculated. The data in the interval over which the estimate is made must be stored using up large amounts of computer memory.

The problem with the batch solution can be removed with the technique of recursive least squares (see Mendel, Graupe a). A new estimate is generated at each iteration (sampling period) using the old estimate based on old data plus the new available data. Two formulations are applied here to identification of decentralized measurement equations.

First, for ease of notation write the measurement equation (4.4) in the form

$$z(K) = \underline{h}^{T}(K)\underline{P} \qquad (4.6)$$

where the correspondence is obvious but some explanation is in order as to agreement in indices for example with

$$z(K) = y_{i}(k+n) \qquad (4.9)$$

why isn't $z(K+1)$ used? The answer is that $z(K)$ is a generic symbol and shorthand for describing any arbitrary measurement equation such as in the discussion of measurement equations in section 3.1. The nongeneric quantities are used in the algorithm and it is these indices that are used in the simulation.

Including measurement noise eq. (4.8) is written

$$z(K) = \underline{h}^{T}(K)\underline{P} + v(K)$$

and suppose L+1 measurements are accumulated from z(k-L) to z(k) and then one new measurement z(k+1) becomes available. Let the first L+1 measurements, measurement vectors, noise terms and weights be defined as

$$\underline{z}(k) = (z(k), \ldots, z(k-L))^T$$

$$h(k) = (\underline{h}(k), \ldots, \underline{h}(k-L))^T$$

$$\underline{V}(k) = (v(k), \ldots, v(k-L))^T$$

$$W(k) = \mathrm{diag}(w(k), \ldots, w(k-L)) \qquad (4.10)$$

so that the concatenated measurement equation is

$$\underline{Z}(k) = H(k)\underline{x} + \underline{V}(k) .$$

Also define

$$\underline{z}(k+1) = \begin{bmatrix} z(k+1) \\ \hline \underline{z}(k) \end{bmatrix}$$

$$\underline{V}(k+1) = \begin{bmatrix} v(k+1) \\ \hline \underline{V}(k) \end{bmatrix}$$

$$H(k+1) = \begin{bmatrix} \underline{h}(k+1) \\ \hline H(k) \end{bmatrix}$$

$$W(k+1) = \mathrm{diag}(w(k+1), \ldots, w(k-L)) . \qquad (4.12)$$

The least squares solution for L+2 measurements from eq. (3.7b) is

$$\hat{\underline{x}}(k+1) = (H^T(k+1)W(k+1)H(k+1))^{-1} X H^T(k+1)W(k+1)\underline{Z}(k+1) \qquad (4.13)$$

and using eq.(4.12) can be written

$$\hat{\underline{x}}(k+1) = (\underline{h}(k+1)w(k+1)\underline{h}^{T}(k+1) + H^{T}(k)W(k)H(k))^{-1}$$
$$(\underline{h}(k+1)w(k+1)z(k+1) + H^{T}(k)W(k)\underline{z}(k)) \qquad (4.14)$$

In eq. (4.14) let

$$P(k) = (H^{T}(k)W(k)H(k))^{-1} \qquad (4.15)$$

then inside eq. (4.13)

$$P(k+1) = (H^{T}(k+1)W(k+1)H(k+1))^{-1}$$
$$= (\underline{h}(k+1)w(k+1)\underline{h}^{T}(k+1) + Q^{-1}(k))^{-1} \qquad (4.16)$$

so that

$$Q^{-1}(k+1) = Q^{-1}(k) + \underline{h}(k+1)w(k+1)\underline{h}^{T}(k+1) . \qquad (4.17)$$

which is a recursive equation for $Q^{-1}(k+1)$. Using eq. (3.7b) again and eq. (4.15) yields

$$\hat{\underline{x}}(k) = Q(k)H^{T}(k)W(k)\underline{z}(k) \qquad (4.18)$$

or

$$H^{T}(k)W(k)\underline{z}(k) = Q^{-1}(k)\hat{\underline{x}}(k) \qquad (4.19)$$

which is substituted into eq. (4.14) to get

$$\hat{\underline{x}}(k+1) = Q(k+1)(\underline{h}(k+1)w(k+1)z(k+1)$$
$$+ Q^{-1}(k)\hat{\underline{x}}(k)) . \qquad (4.20)$$

Using eq. (4.17) in eq. (4.20) yields

$$\hat{\underline{x}}(k+1) = Q(k+1)(\underline{h}(k+1)w(k+1)z(k+1)$$
$$+ (Q^{-1}(k+1) - \underline{h}(k+1)w(k+1)\underline{h}^{T}(k+1))\hat{\underline{x}}(k))$$
$$= \hat{\underline{x}}(k) + Q(k+1)\underline{h}(k+1)w(k+1)$$
$$\times (z(k+1) - \underline{h}^{T}(k+1)\hat{\underline{x}}(k)) \qquad (4.21)$$

and this equation together with eq. (4.17) constitute one version of the recursive least squares algorithm. Since the effect of all past data is in the last estimate, the past data is used without having to be stored and processed for e

new estimate. Only the most recent measurement $z(k+1)$ is needed. In contrast a new measurement in the batch method requires using it and all past data for a new estimate if all the data is to be taken into account. Eq. (4.17) still requires matrix inversion of an nXn matrix and at each iteration of the algorithm. Inversion can be completely avoided using the matrix inversion lemma.

The lemma applies to equations of the form

$$A^{-1} = B^{-1} + C^T D^{-1} C \tag{4.22}$$

where A, B and D are nonsingular matrices. Premultiplying by A gives

$$I = AB^{-1} + AC^T D^{-1} C \tag{4.23}$$

and postmultiplying by BC^T gives

$$BC^T = AB^{-1}BC^T + AC^T D^{-1} CBC^T$$

$$= AC^T D^{-1}(D + CBC^T) . \tag{4.24}$$

Postmultiplying by $(D+CBC^T)^{-1}CB$ gives

$$BC^T(D + CBC^T)^{-1}CB = AC^T D^{-1} CB \tag{4.25}$$

and subtracting (4.25) from A gives

$$A - BC^T(D + CBC^T)^{-1}CB$$

$$= A - AC^T D^{-1} CB . \tag{4.26}$$

Postmultiplying eq. (4.23) by B

$$B = A + AC^T D^{-1} CB \tag{4.27}$$

and substitute into eq. (4.26) to get

$$A = B - BC^T(D + CBC^T)^{-1}CB . \tag{4.28}$$

Then applying this lemma to eq. (4.17) gives

$$\underline{x}(k+1) = \underline{x}(k) - \underline{w}(k)\underline{h}(k+1)$$

$$\lambda (\underline{h}^T(k+1)\underline{w}(k)\underline{h}(k+1) + 1/\lambda(k+1))^{-1} \tag{4.29}$$

and since D and CBC^T correspond to scalars no matrix inversion is necessary.

Eq. (4.29) in conjunction with eq. (4.21) constitute a recursive least squares algorithm with no matrix inversion.

Because of correlation between the measurement matrix and the measurement noise, recursive or batch least squares estimates are biased with measurement noise present. If the noise is very large the estimates could be unacceptable. An algorithm introduced by Panuska is a modified version of recursive least squares that can be shown to converge in mean square and gives unbiased estimates with measurement noise present. The algorithm applies to equations of the form

$$y(k+n) = a_1 y(k+n-1) + \ldots + a_n y(k) + b_1 u(k+n-1) + \ldots + b_n u(k) + c_1 v(k+n-1) + \ldots + c_n v(k) + v(k+n) \quad (4.30)$$

which in vector form is written

$$z(k) = \underline{h}^T(k)\underline{P} + e(k) \quad (4.31)$$

with

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

$$\underline{h}^T(k) = (y(k+n-1), \ldots, y(k), u(k+n-1), u(k), v(k+n-1), \ldots, v(k))$$

$$\underline{P}^T = (a_1, \ldots, a_n, b_1, \ldots, b_n, c_1, \ldots, c_n)$$

$$e(k) = v(k+n) \quad (4.32)$$

and $e(k)$ is an unmeasurable random disturbance with a rational spectral density. It is assumed to be a sequence of random variables with zero mean and variance σ^2.

The coefficients $c_{i,\lambda}$, $i=1, \ldots, n$ are disturbance parameters and $\underline{h}^T(k)$ is an enlarged measurement vector in contrast to the situation examined thusfar. Since the $e(k)$'s are not available then computed errors are used for eq. (4.31)

$$\hat{e}(k) = z(k) - \hat{\underline{h}}^T(k)\underline{p} \tag{4.33}$$

where

$$\hat{\underline{h}}^T(k) = (y(k+n-1), \ldots, y(k), u(k+n-1), \ldots, u(k), \hat{v}(k+n-1), \ldots, \hat{v}(k)).$$

which is an enlarged measurement vector using the computed error estimates. The algorithm is given by these recursive equations

$$\hat{\underline{p}}(k+1) = \hat{\underline{p}}(k) + \underline{\gamma}(k)(y(k) - \hat{\underline{p}}^T(k)\underline{h}(k))\hat{h}(k)$$

$$\underline{\gamma}(k+1) = \underline{\gamma}(k) - \underline{\gamma}(k)\hat{h}(k+1)$$
$$(\hat{\underline{h}}^T(k+1)\underline{\gamma}(k)\hat{\underline{h}}(k+1) + 1)^{-1}\hat{\underline{h}}^T(k+1)\underline{\gamma}(k) .$$

The scalar decentralized measurement equation is cast into the form of eq. (4.30) if no state noise is present so that eq. (4.30) reduces to

$$y_{\lambda}(k+n) = \underline{1}^T\overline{y}_{\lambda}(k) + \underline{c}_{1\lambda}^T\underline{v}(k+n-1) + \ldots + \underline{c}_{n\lambda}^T\underline{v}(k) - \underline{\alpha}^T\overline{y}_{\lambda}(k) + v_{\lambda}(k+n) \tag{4.34}$$

where the notation is the same as in eq. (4.32) except that more than one input is present. Since the order of the

elements in $\underline{L}(k)$ can be reversed (revised the order in $\hat{\underline{L}}(k)$ is also, the inner product in eq. (4.31) is unchanged and

$$\underline{L}^T(k) = (\overline{\underline{y}}_{\underline{L}}^T(k), \underline{u}^T(k+n-1), \ldots, \underline{u}^T(k), \underline{v}_{\underline{L}}(k))$$

$$\underline{L} = (\underline{L}^T, \underline{p}_{\underline{L}}\underline{L}^T, \ldots, \underline{p}_{\underline{L}}\underline{L}^T, \underline{L}^T)^T \qquad (4.36)$$

where the peculiar result that the dynamic dependence coefficients and disturbance parameters are the same, in effect, making this a special case of a more general problem.

The relevant global system state description is then

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

$$\underline{y}_m(k) = C\underline{x}(k) + \underline{v}(k) \qquad (4.37)$$

To prove convergence these assumptions must be satisfied

1) The system described by eq. (4.37) or eq. (4.36) and the predictor in eq. (4.33) must be stable. Also, the system must be controllable from the input $\underline{u}(k)$ or disturbance $e(k)$.

2) Both input $\underline{u}(k)$ and disturbance $e(k)$ have finite second and fourth order moments and $\underline{u}$ and e are mutually independent.

3) The covariance matrix

$$L\left[\hat{\underline{n}}(k)\hat{\underline{n}}^T(k) \mid \hat{\underline{L}}(1), \ldots, \hat{\underline{P}}(k)\right]$$

is positive definite as $k \to \infty$.

4) The condition

$$\lim_{k\to\infty} K_{\underline{v}}(k) = \text{positive definite constant matrix}$$

holds.

Mahoska's algorithm was seen to be applicable to identification of the scalar DMB with measurement noise but not state noise present in the system. With a vector of measurements available and measurement noise only eq. (3.61) for the ith measurement is

$$ym_i(k+p_i) = \underline{a}_i ym(k) - \underline{b}^* U(k) + \underline{a}_i V(k) + v_i(k+p_i) \qquad (4.38)$$

where

$$\underline{b}^* = (c_i A^{p_i-1} B \mid \cdots \mid c_i B) - (\underline{a}_i M_1 \mid \cdots \mid \underline{a}_i M_{p_i-1} \mid 0) \qquad (4.39)$$

and the superscript, 1, has been dropped for notational simplicity.

As in the scalar measurement case computed errors can be calculated since v(k) is not measurable by

$$\hat{\epsilon}_i(k) = ym_i(k+p_i) - \hat{\underline{p}}^T(k)\hat{\underline{h}}(k+1) \qquad (4.40)$$

where

$$\hat{\underline{h}}(k+1) = (\overline{ym}^T(k) \mid \bar{U}^T(k) \mid \hat{\bar{V}}^T(k))^T$$

$$\underline{k} = (\underline{a}_i \mid \underline{k}^{*T} \mid \underline{a}_i)$$

$$\hat{\epsilon}_i(k) = \hat{v}_i(k+p_i)$$

for $i=1, \ldots, n_\ell$.

Other variations of the least squares approach (based on minimizing a weighted sum of squares of errors) have been developed. These algorithms differ in the numerical solution of the problem and have different convergence rates, sensitivities to round-off errors, and computational speeds.

These algorithms hold for the general autoregressive model

$$y(k+n) + a_1 y(k+n-1) + \ldots + a_n y(k) = w(k+n) \qquad (4.41)$$

where the white noise source $w(k)$ is zero mean and second moment ergodic and the covariance of the noise is finite and constant. The first methods to be discussed are the covariance, partial correlation, and autocorrelation methods which are nonrecursive (see Mehra, Box and Jenkins, Gibson and Melsa).

As discussed in the beginning of the chapter the batch least squares estimate is given by

$$\hat{\underline{a}} = (S^T S)^{-1} S^T \bar{y}(N)$$

where

$$\bar{y}(N) = (\bar{y}(n), \ldots, \bar{y}(N+n-1))^T \qquad (4.42)$$

and

$$S = (\bar{y}(N-1), \bar{y}(N-2), \ldots, \bar{y}(N-n))^T \qquad (4.43)$$

Next, the sample covariance is defined by

$$R(i,l) = 1/N \sum_{k=n}^{N+n-1} y(k-i)y(k-l); \quad i,l=1,\ldots,N \qquad (4.44)$$

so that

$$S^T \bar{y}(N) = (R(1,1), \ldots, R(1,n))^T \qquad (4.45)$$

and

$$S^T S = \begin{bmatrix} R(1,1) & \cdots & R(1,N) \\ \vdots & & \vdots \\ R(N,1) & \cdots & R(N,N) \end{bmatrix} \qquad (4.46)$$

let

$$M = S^T S$$

and

$$I(i) = S^T y(N) \tag{4.47}$$

so that

$$M = (I(1), \ldots, I(n))$$

and

$$\hat{P} = M^{-1} I(0) \ . \tag{4.48}$$

Therefore, this algorithm is the batch estimate except for a difference in numerical execution. Also, since R(i,1) is the autocorrelation R(i-1) of y(k) with delay i-1, eq. (4.48) is also called an autocorrelation estimate.

The batch partial correlation method is a nonrecursive method that reduces the n dimensional problem in eq. (4.46) to n one dimensional problems by successively increasing the order of eq. (4.46) (see Box and Jenkins and Gibson and Melsa). First, eq. (4.48) is written in Yule-Walker form

$$\sum_{i=1}^{n} \hat{P}_i I(i-1) = -I(1)$$

$$I(i-1) = 1/N \sum_{k=n}^{N+n-1} y(k-i)\, y(k-1) \ . \tag{4.49}$$

Let the ith coefficient of a jth order autoregressive process be P_i^j. For j=1 from eq. (4.48)

$$P_1^1 = -I(1)/I(0) \tag{4.50}$$

for j=2

$$\begin{bmatrix} P_1^2 \\ P_2^2 \end{bmatrix} = \begin{bmatrix} I(0) & I(1) \\ I(1) & I(0) \end{bmatrix}^{-1} \begin{bmatrix} I(1) \\ I(2) \end{bmatrix}$$

and so on which can be reduced to the form

$$P_{\lambda}^{j+1} = P_{\lambda}^{j} + k_{j+1} P_{j+1-\lambda}^{j} \quad \text{for } \lambda = 1, \ldots, j$$

$$k_j = P_j^j \; ; \; P_0^j = 1 \text{ for } j = 0, 1, \ldots, n-1 \tag{4.51}$$

and

$$P_{j+1}^{j+1} = \sum_{\lambda=0}^{j} P_{\lambda}^{j} T(j+1-\lambda) \Big/ \sum_{\lambda=0}^{j} P_{\lambda}^{j} T(\lambda) \tag{4.52}$$

where the P_{j+1}^{j+1} are the partial correlation coefficients.

A sequential covariance and partial correlation algorithm can be implemented using the recursive equation

$$T^{j+1}(l) = T^{j}(l) + 1/(j+1)\left(y(j+1)y(j+1-l) - R^{j}(l) \right) \tag{4.53}$$

from Merhav and Gabay where T^j is the correlation in eq. (4.47) based on j data points. The correlation is then used in eq. (4.47). Using this correlation in conjunction with eq. (4.52) gives sequential partial correlation estimates.

The next variation of least squares to be discussed is the lattice method. The method involves computing forward and backward error residuals. A function of the residuals is minimized with respect to coefficients that converge to the desired ones in eq. (4.41).

Lattice algorithms have been used in other fields of engineering, e.g., in digital signal processing and several papers have appeared in the control theory literature (see Glaupe B, Shichor, Burg, and Makhoul). The first step is to compute error residuals, some papers find forward residuals given by

$$e_j(i,k) = y(k) - \hat{y}(k) = \sum_{\lambda=0}^{j} a_{\lambda}^{j} y_{k-\lambda} \; ; \; a_0^j = 1;$$

$$e_{+,k}^{0} = y(k) \tag{4.54}$$

which depend on past outputs y(k-1), y(k-2), etc. while others find backward residuals given by

$$e_b(j,k) = y(k-j)y(k-j) = \sum_{i=0}^{j} a_i^j y(k+i-j); \; a_0^j = 1$$

$$e_b(0,k-1) = y(k-1) \tag{4.55}$$

which depend on future outputs y(k-j+1), y(k-j+2), etc.

The functions that have been used in the minimization are, where E is the expectation operator, the mean square forward error

$$E_F^j = E(e_f^2(j,k)) \tag{4.56}$$

the mean square backward error

$$E_B^j = E(e_b^2(j,k)) \tag{4.57}$$

and at least one paper (Srinath and Viswanathan) uses the average sum of the mean square forward and backward errors

$$E_A^j = \frac{E_F^j + E_B^j}{2} \; . \tag{4.58}$$

Also define

$$E_C^j = E(e_f(j,k)e_b(j,k-1)) \tag{4.59}$$

and after minimizing eqs. (4.56) to (4.56) with respect to $a_1^1, a_2^2, \ldots, a_{j+1}^{j+1}$ (see Graupe b) yields

$$a_{j+1}^{j+1}\,(\text{forward}) = E_C^j / E_B^j$$

$$a_{j+1}^{j+1}\,(\text{backward}) = E_C^j / E_F^j$$

$$a_{j+1}^{j+1}\,(\text{average}) = 2E_C^j / (E_F^j + E_B^j) \tag{4.60}$$

which are the lattice algorithms.

With various least algorithms reviewed three examples will be formulated and results from computer simulations

given. The recursive least squares formulation described by eqs. (4.29) and (4.21) which avoids the calculation of inverses has been programmed and applied to three MIMO examples. Outputs are then decentralized and all possible parameters (as much of the system as possible) identified in accordance with the identifiability results of Chapter 3.

The first example is given by the following state second order abstract discrete system

$$\begin{bmatrix} x_1(K+1) \\ \\ x_2(K+1) \end{bmatrix} = \begin{bmatrix} 1 & .3 \\ \\ -1 & -.7 \end{bmatrix} \begin{bmatrix} x_1(K) \\ \\ x_2(K) \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ \\ 0 & 2 \end{bmatrix} \begin{bmatrix} u_1(K) \\ \\ u_2(K) \end{bmatrix}$$

$$\underline{y}(K) = \underline{x}(K) \qquad (4.61)$$

with eigenvalues at .8 and -.5.

Let $y_1(K)$, but not $y_2(K)$, be measurable by one identifier. The other output, $y_2(K)$, may or may not be available to some other identifier but will not concern the processing by the first identifier.

The observability matrix from the first output is

$$V = \begin{bmatrix} 1 & 0 \\ \\ 1 & .3 \end{bmatrix} \qquad (4.62)$$

therefore the single structural invariant is 2 and the identified model will be second order which is the global system order. With n=2 the dependence coefficients identified are found from Table 3.1

$$\underline{c}(K) = \underline{cb}\,1^{T} \quad \underline{c}_1\underline{AB} - \underline{\alpha}^{T}\underline{M}$$

$$\underline{y}(k+1) : \underline{pb2}^T = c_1 B \tag{4.63}$$

where

$$\underline{p}^T = (-|A|, \ tr(A)) = (.4 \quad .3) \tag{4.64}$$

contains the negatives of the coefficients of the characteristic equation of A in reverse. From Table 3.1

$$M = \begin{bmatrix} 0 \\ c_1 B \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \tag{4.65}$$

and then

$$c_1 AB = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & .3 \\ -1 & -.7 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & .6 \end{bmatrix} \tag{4.66}$$

$$\underline{p}^T M = (.3 \quad 0) \tag{4.67}$$

$$\underline{pb1}^T = (.7 \quad .6) \ . \tag{4.68}$$

The characteristic equation of A is

$$z^2 - .3z - .4 = 0 \ . \tag{4.69}$$

For the inputs

$$c_1 AB - \underline{p}^T M = (.7 \quad .6) \ . \tag{4.70}$$

and the DME is

$$y_1(k+2) = .3y_1(k+1) + .4y_1(k) + u_1(k+1) + .7u_1(k) + .6u_2(k) \tag{4.71}$$

and is identified by the recursive least squares identifier described previously in eqs. (4.14)-(4.16). Also,

$$\underline{pb2}^T = (1 \quad 0) \tag{4.72}$$

as the dependence coefficient vector of $\underline{y}(k+1)$. The total number of coefficients is given by

$$n + np_k = 2 + (2)(2) = 6 \tag{4.73}$$

and the total coefficient vector is

$$\underline{\theta} = (.4, .3, 1, 0, .7, .6)^T \quad . \tag{4.74}$$

Then the decentralized measurement equation has the form

$$z(k) = y(k+2) = \underline{g}^T(k)\underline{\theta} \tag{4.75}$$

where the measurement vector

$$\underline{g}^T(k) = (y_1(k), y_1(k+1), u_1(k+1), u_2(k+1), u_1(k), u_2(k)) \quad . \tag{4.76}$$

To start the algorithm $Q(0)$ and an initial parameter guess, $\underline{\hat{\theta}}(0)$, are needed. As suggested in Mendel they are calculated using the two equations

$$Q(0) = (1/a^2 I + \underline{g}(0)w(0)\underline{g}(0))^{-1}$$
$$\underline{\hat{\theta}}(0) = Q(0)(1/a\underline{m} + \underline{g}(0)z(0)) \tag{4.77}$$

where "a" is a very large number and $\underline{m}$ is a vector of very small numbers. If the initial weight $w(0)$ and initial measurement $z(0)$ are both zero eq. (4.77) reduces to

$$Q(0) = a^2 I$$
$$\underline{\hat{\theta}}(0) = 1/a^3 \underline{m} \approx \underline{0} \quad . \tag{4.78}$$

If initial conditions and inputs are set to zero such that

$$y_1(0) = y_1(1) = u_1(1) = u_1(0) = u_2(1) = u_2(0) = 0$$

are all zero in eq. (4.76) then $z(0)$ is zero. This was done in the simulation and a^2 was set to 10^6 in eq. (4.78). Arbitrary inputs were used as follows

$$u_1(k) = 5\sin(3k)$$
$$u_2(k) = -4\sin(7k) + 1 \quad .$$

The 6 parameters converge to their true values very fast and are shown in Figs. (4.1)-(4.3).

Another example will now be identified using the recursive least squares algorithm in which the global system is unobservable from the available measurement. The second order abstract discrete system is given by

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = \begin{bmatrix} .5 & .1 \\ .2 & .85 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1(k) \\ u_2(k) \end{bmatrix}$$

$$\begin{bmatrix} y_1(k) \\ y_2(k) \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} \qquad (4.80)$$

and suppose only $y_1(k)$ is available. The system is stable with eigenvalues at .9 and .45. The observability matrix from this output is

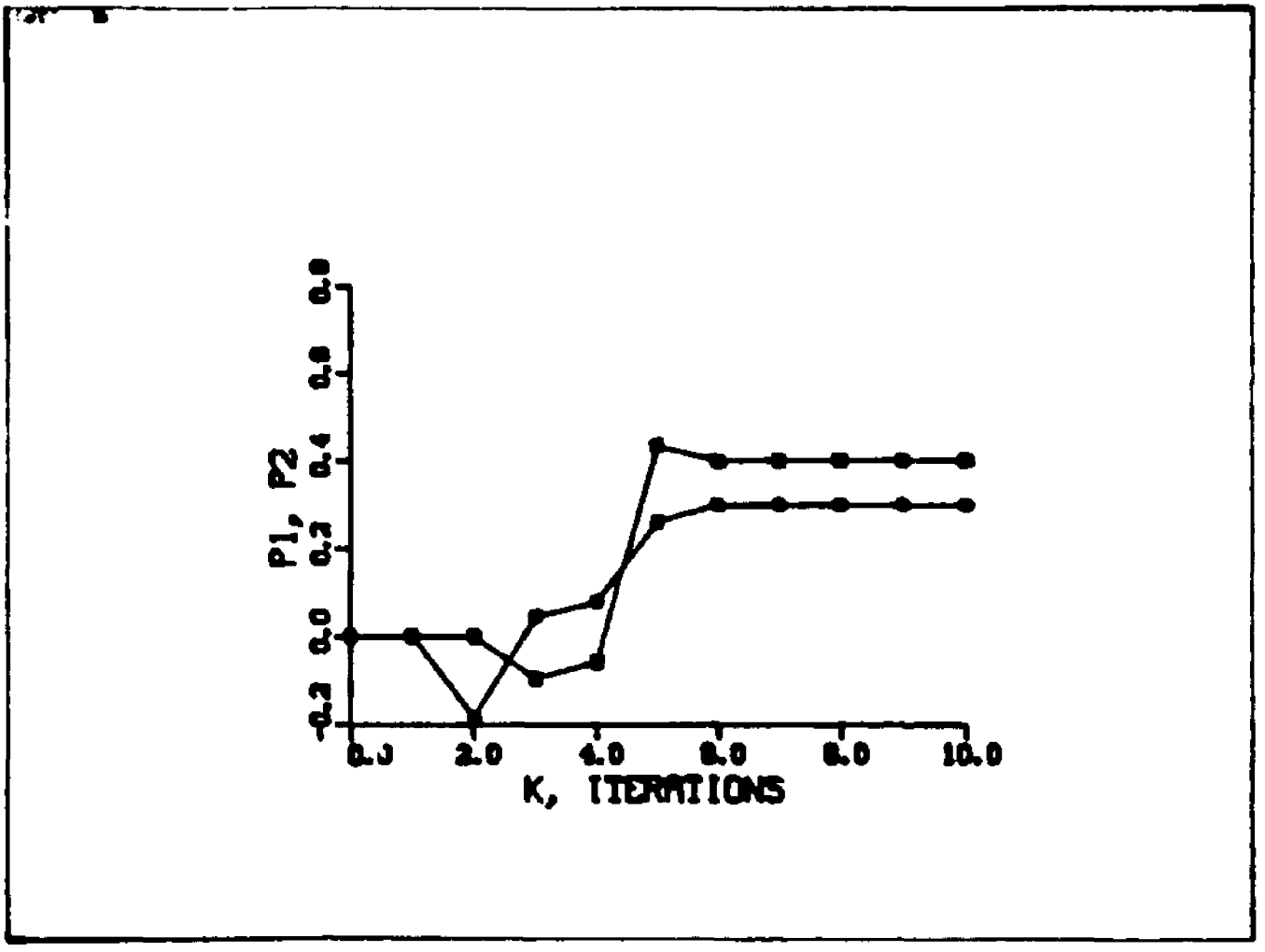

Fig. 4.1 Parameter P1 and P2 convergence using recursive least squares in observable example.

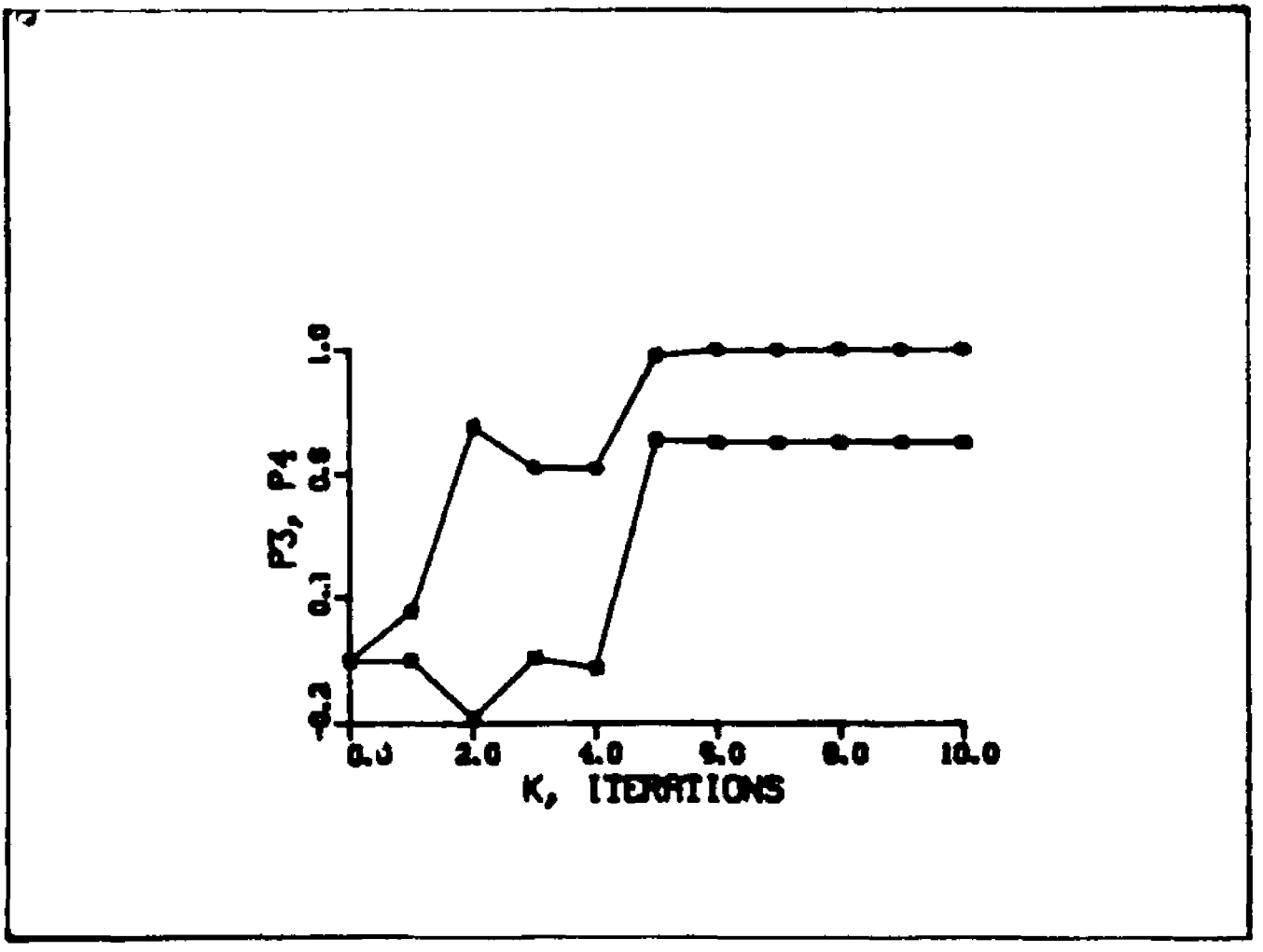

Fig. 4.2 Parameter P3 and P4 convergence using recursive least squares in observable example.

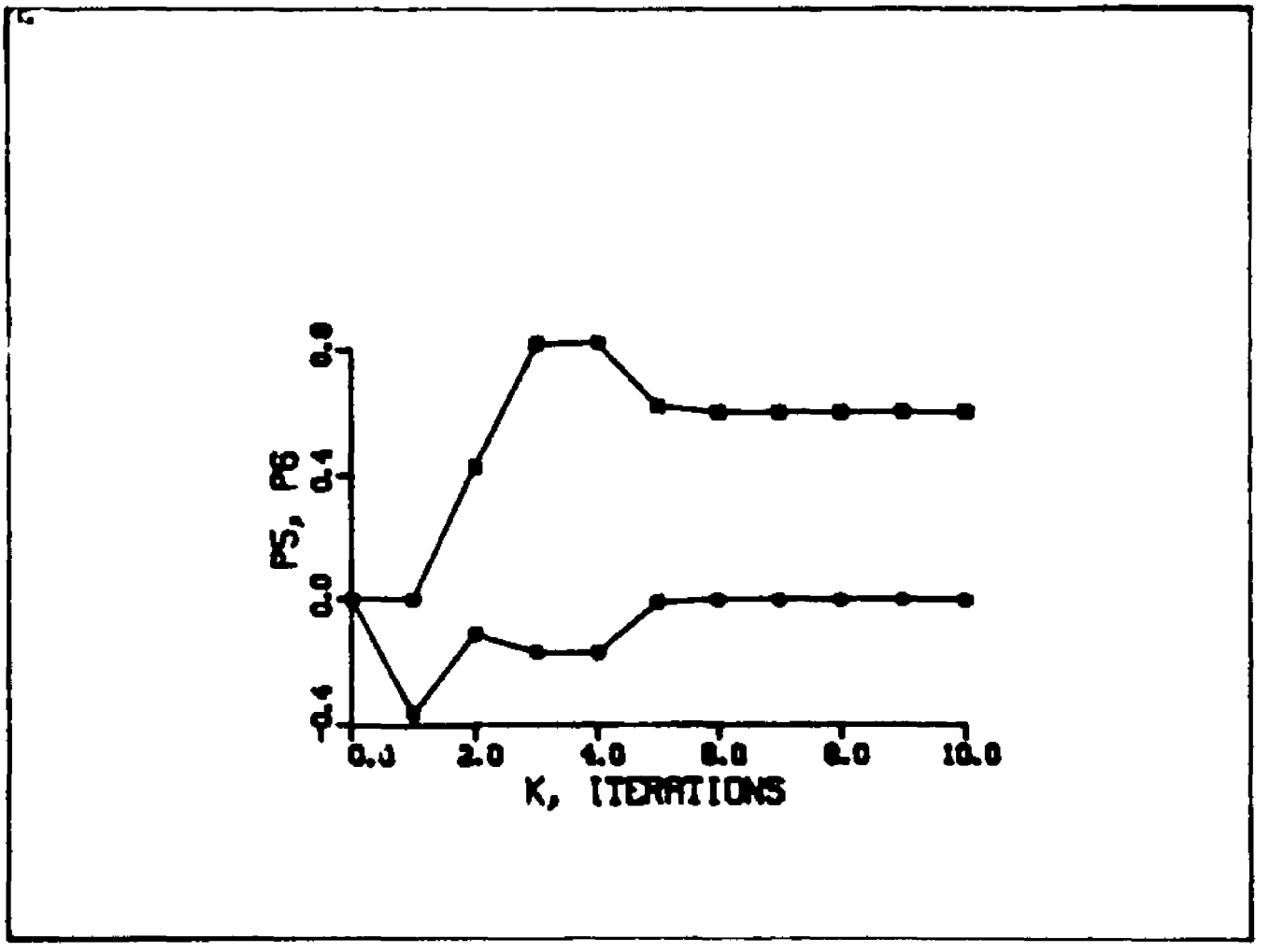

Fig. 4.3 Parameter P5 and P6 convergence using recursive least squares in observable example.

$$C = \begin{bmatrix} 1 & 2 \\ .9 & 1.8 \end{bmatrix} \qquad (4.81)$$

which is singular. With reference to the results in sections 3.2 and 3.3 only the decoupled part of the system in observability canonical form can be identified. To show that the identifier output to follow is correct the system will be transformed to this canonical form.

As discussed in Chapter 2 the first m rows of the transformation matrix span the observable subspace of the subsystem and such a set is obtained using the largest possible number of linearly independent rows in eq. (4.81). This amounts to the first row in eq. (4.81). A set of linearly independent rows spanning the null space of eq. (4.81) is obtained from

$$c_1 e = 0$$

$$(1 \quad 2)e = 0 \qquad (4.82)$$

or

$$e_1 + 2e_2 = 0 .$$

Letting e_2=1 the second row of the transformation matrix is then

$$(-2 \quad 1)$$

and so

$$T = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \qquad (4.83)$$

with an inverse

$$T^{-1} = \begin{bmatrix} .2 & -.4 \\ .4 & .2 \end{bmatrix} . \qquad (4.84)$$

The transformed system matrices are then calculated

$$TAT^{-1} = \begin{bmatrix} .9 & 0 \\ .1 & .45 \end{bmatrix}$$

$$CT^{-1} = (1 \quad 0)$$

$$Tb = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \qquad (4.85)$$

where

$$(A_{11}' , C_1') = (.9 \quad 1) \qquad (4.86)$$

is an observable pair. The single eigenvalue of A_{11}' is .9 which is an observable pole while .45 is an unobservable pole.

From the column dependence test discussed in section 3.3 the order of the observable part of the system is found to be one. This value is then used in Table 3.1 to find what dependence coefficients will be identified. The characteristic equation of A_{11}' is simply

$$z - .9 = 0 \qquad (4.87)$$

so that

$$p^T = p = .9 \qquad (4.88)$$

and the dependence coefficients for the inputs are

$$\underline{\psi}(k): \underline{\psi\psi}^T = c_1' \tilde{B} = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$$

$$= \begin{pmatrix} 1 & 2 \end{pmatrix} . \qquad (4.89)$$

The DME is then

$$y_1(k+1) = .9y_1(k) + u_1(k) + 2u_2(k) \qquad (4.90)$$

where the parameter vector

$$\underline{\theta} = \begin{pmatrix} .9 & 1 & 2 \end{pmatrix}^T \qquad (4.91)$$

is identified in Figs. (4.4)-(4.6) using the same input functions and startup procedure as in the observable example.

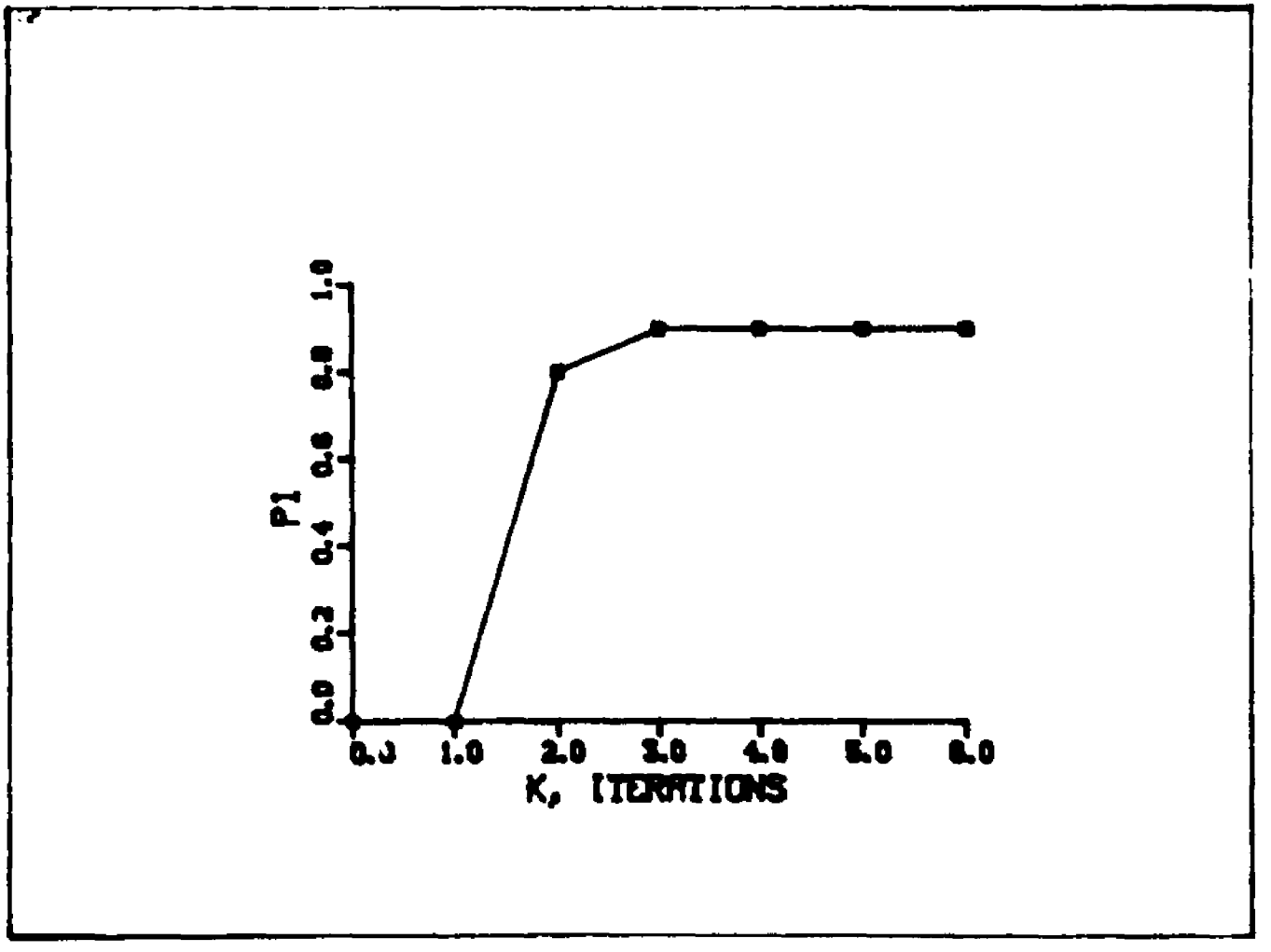

Fig. 4.4 Recursive least squares convergence to parameter P1 of unobservable example.

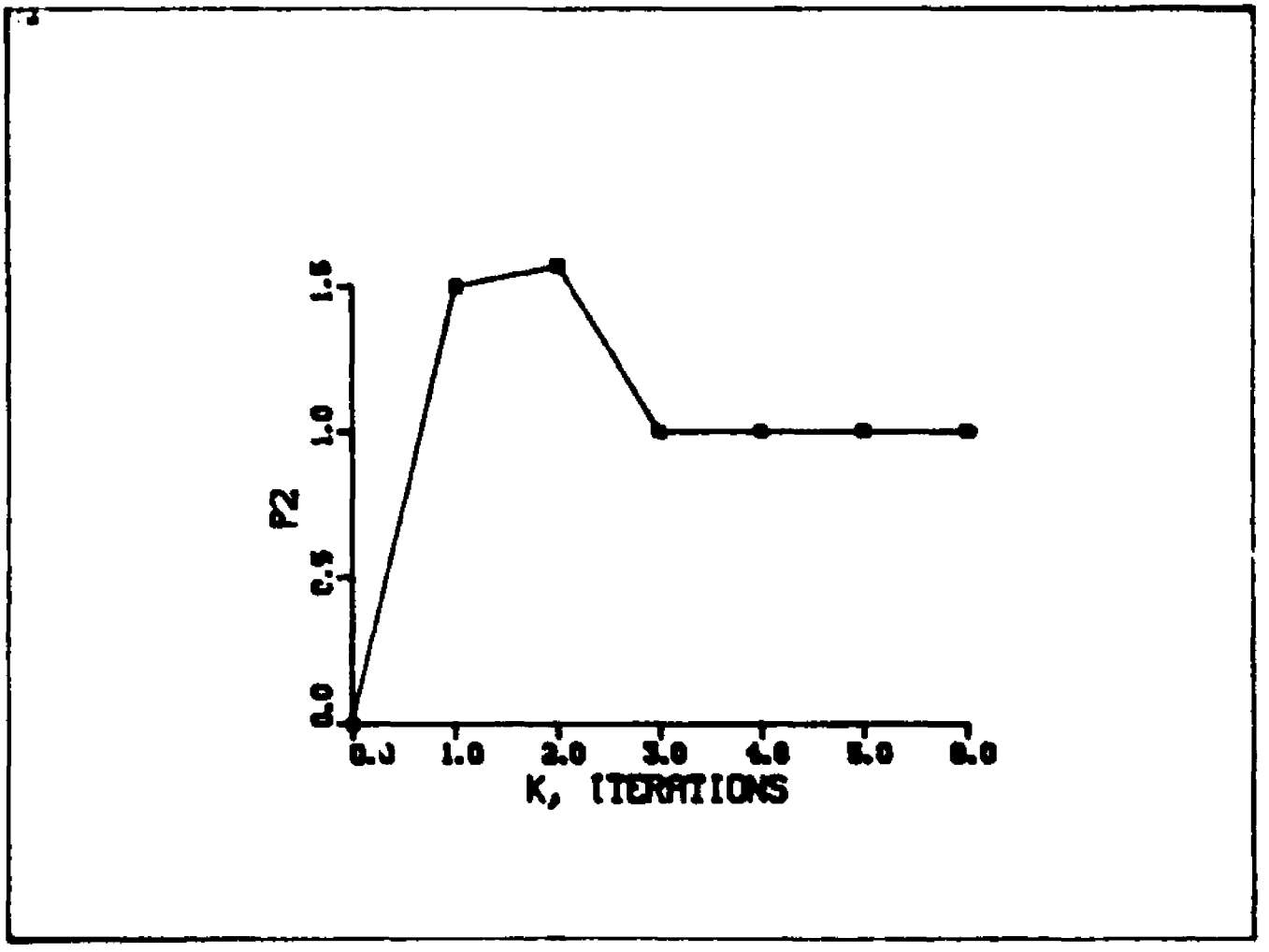

Fig. 4.5 Recursive least squares convergence to Parameter P2 of unobservable example.

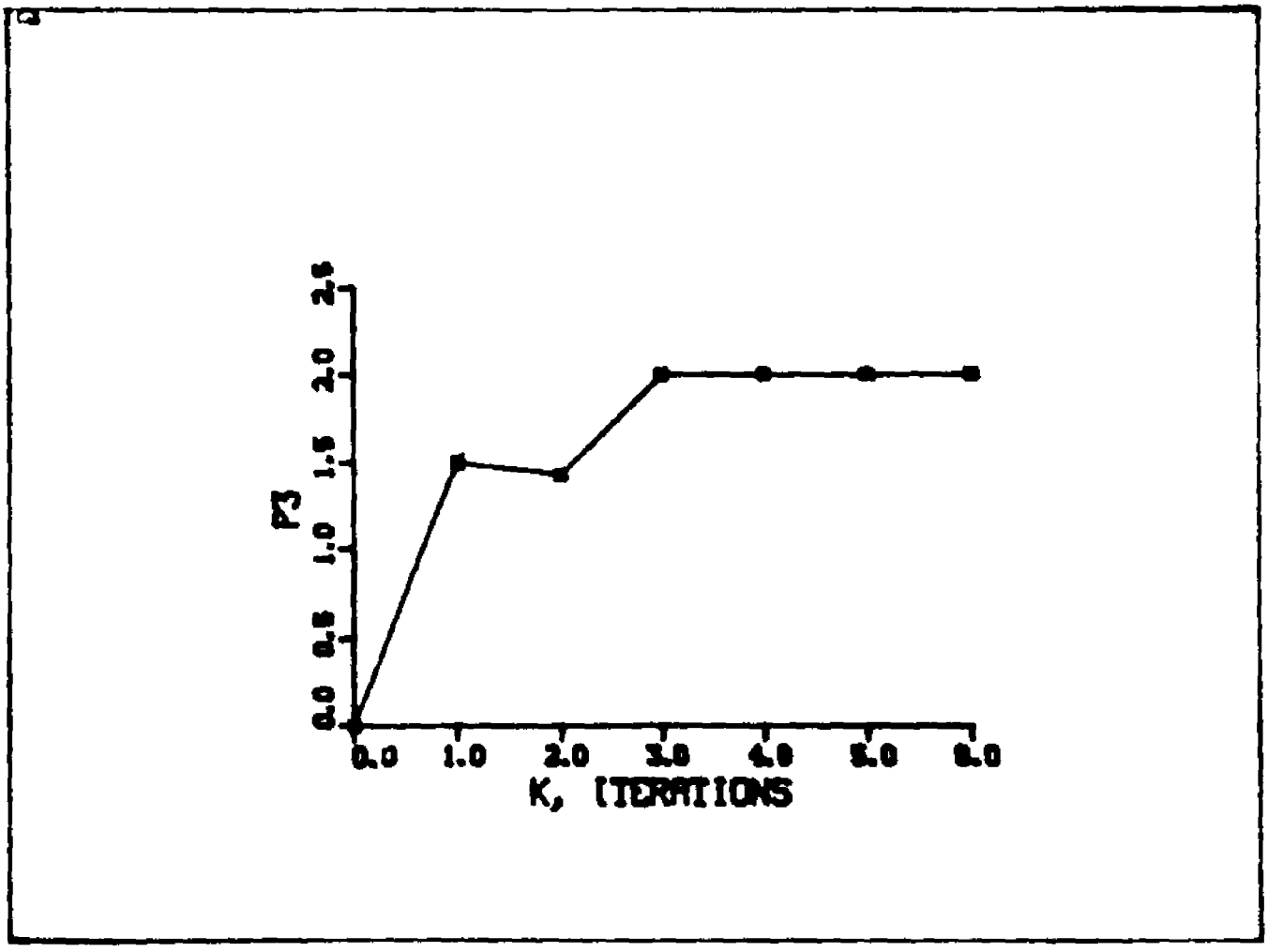

Fig. 4.6 Recursive least squares convergence to Parameter P3 of unobservable example.

A third example is a sampled data version of a linearized continuous time model of a stirred tank. This example is typical of some process control systems. Two feeds are the inputs to the stirred tank carrying time-varying flows. The flows contain dissolved material with constant concentrations c_1 and c_2. The propellor in the tank maintains a constant concentration in the tank which is the concentration of the outgoing flow at the bottom.

The mass balance and output flow rate equations describing the tank have been linearized in Sivan and Kwakernaak and lead to the state-space description

$$\dot{\underline{x}}(t) = \begin{bmatrix} -1/2h & 0 \\ 0 & -1/h \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 1 & 1 \\ (c_1-c_o)/V_o & (c_2-c_o)/V_o \end{bmatrix} \underline{u}(t) \quad (4.92)$$

where h is the holdup time and

$$h = V_o / F_o \quad (4.93)$$

and

V_o = nominal steady state tank volume = 1 m^3

F_o = nominal output flow rate = .02 m^3/s

h = 50 s

c_1 = feed 1 concentration = 1 kmol/m^3

c_2 = feed 2 concentration = 2 kmol/m^3 .

Then eq. (4.92) is given by

$$\dot{\underline{x}}(t) = \begin{bmatrix} -.01 & 0 \\ 0 & -.02 \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 1 & 1 \\ -.25 & .75 \end{bmatrix} \underline{u}(t) \quad (4.94)$$

where the components of the state are the incremental volume in the tank and the incremental concentration.

Next, consider the tank forms part of a process commanded by a process control computer. As a result, the value settings change at discrete time instants only and are constant in between. Let the interval between time instants be D. The discrete time matrices are then given by

$$A_d = e^{AD}$$
$$b_d = \left(\int_0^D e^{A\tau} d\tau \right) B \qquad (4.95)$$

and let D=5 seconds. The discrete time system is then computed to be

$$\underline{x}(K+1) = \begin{bmatrix} .9512 & 0 \\ 0 & .9048 \end{bmatrix} \underline{x}(K) + \begin{bmatrix} 4.877 & 4.877 \\ -1.1695 & 3.569 \end{bmatrix} \underline{u}(K) \qquad (4.96)$$

and also if the whole sampling interval is used to process the data and defining the states as outputs the output equation takes the form

$$\underline{y}(K) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \underline{x}(K) \quad . \qquad (4.97)$$

Now suppose the outputs are decentralized so that only the first state or incremental volume is measurable by the identifier. The observability matrix from this output is

$$V = \begin{bmatrix} 1 & 0 \\ .9512 & 0 \end{bmatrix} \qquad (4.98)$$

which is singular. Less than the entire global system is identifiable. Since the system is already in observable canonical form the dependence coefficients can be calculated from Table 3.1 using n'=1 in place of n. The characteristic equation of A'_{11} is

$$z - .9512 = 0$$

and so

$$\underline{k}^T = k = .9512 \quad .$$

For the inputs

$$\underline{u}(k): \underline{pb1}^T = c_1 B = (4.877 \quad 4.877)$$

and the DME is

$$y_1(k+1) = .9512y_1(k) + 4.877u_1(k) + 4.877u_2(k)$$

where the parameter vector

$$\underline{k} = (.9512 \quad 4.877 \quad 4.877)^T$$

is identified, by the recursive least squares identifier described previously, in Figs. (4.7)-(4.9).

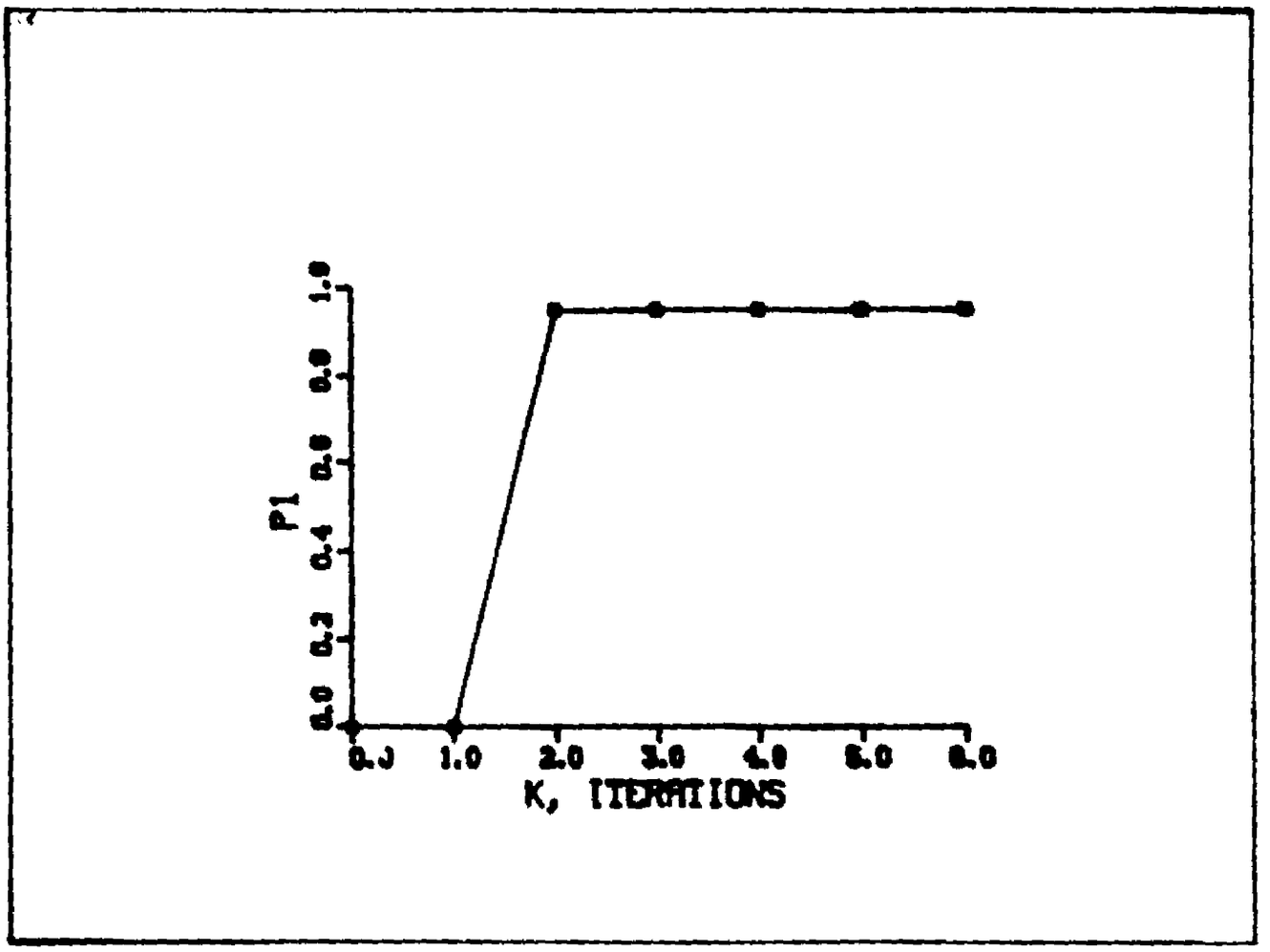

Fig. 4.7 Parameter P1 convergence using recursive least squares in stirred tank example.

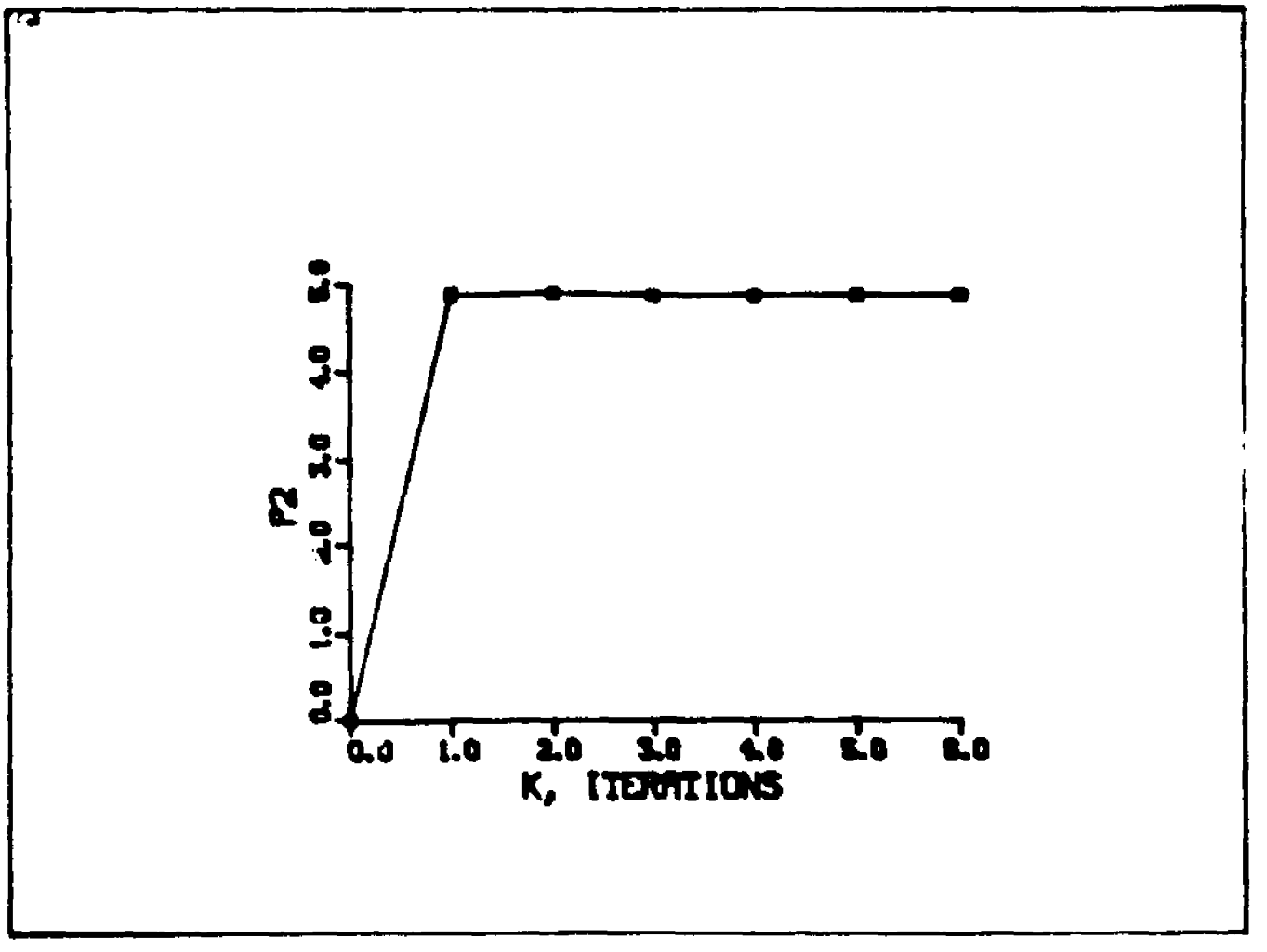

Fig. 4.8 Parameter P2 convergence using recursive least squares in stirred tank example.

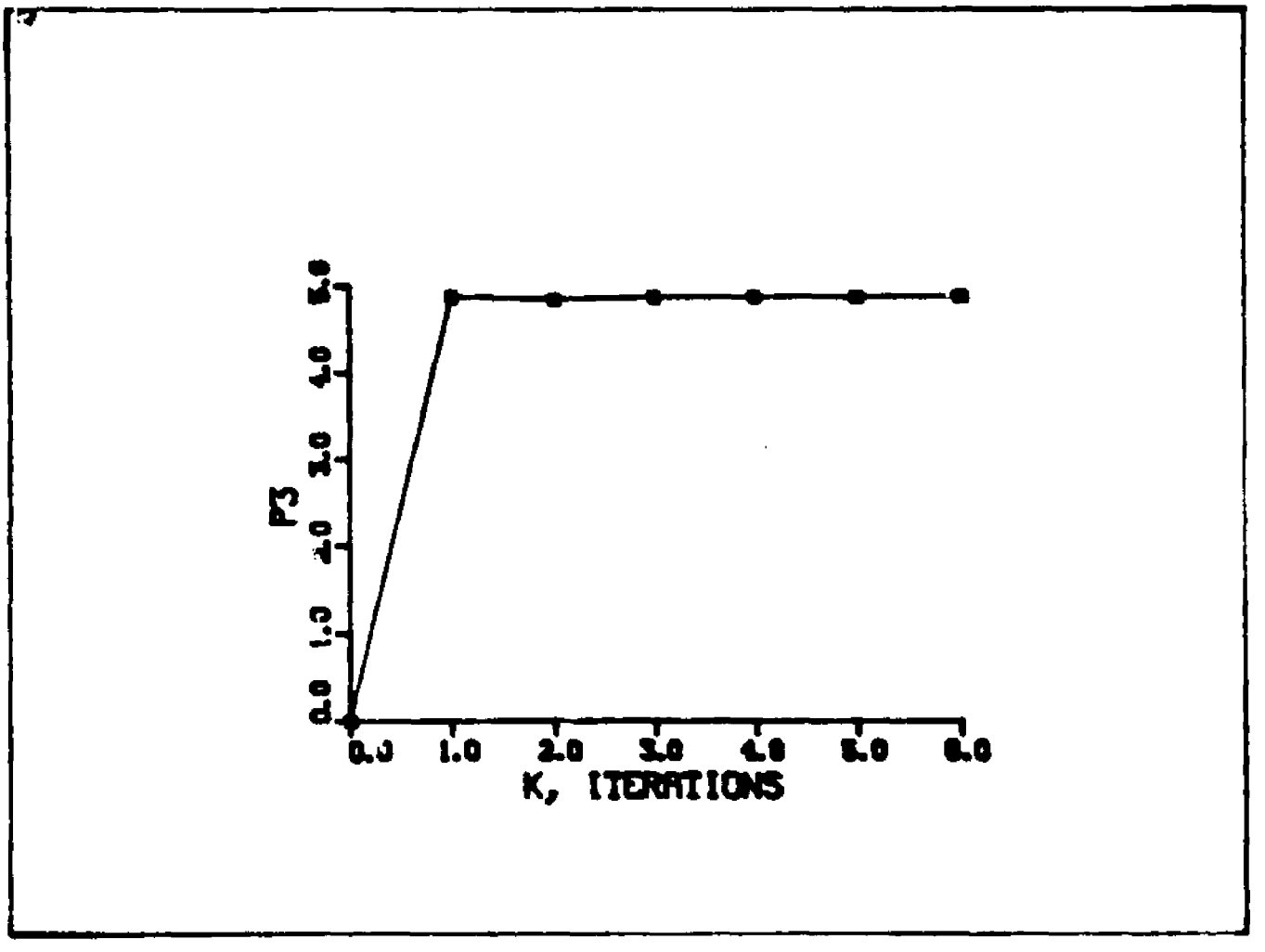

Fig. 4.9 Parameter P3 convergence using recursive least squares in stirred tank example.

4.2 Gradient Methods of System Identification

In this section as in the last identification algorithms will be reviewed and applied to varying levels of complexity of the scalar and vector decentialized measurement equations. The algorithms in this section are all based on the principle of gradient descent whereas the least squares algorithms are based on minimizing a sum of squares of errors. In gradient descent a measure of the square of the measurement error is minimized with respect to the parameter estimates. The minimized function is the gradient which from linear algebra gives the direction of maximum increase of the function. Here since the criterion is a measure of error it is to be minimized so that the negative of the gradient is used in a correction term to update new estimates.

The use of the least squares approach began with the work of Karl Gauss around 1795. His estimation problem was to find parameters that described the motions of planets, comets and other bodies from telescopic measurements (see Mendel). The gradient approach also known as stochastic approximation was first applied to parametric identification in the 1950's in papers by Blum, Dvoretsky, and others. These early papers identified parameters of memoryless equations. In the 1960's and 70's papers appeared on identifying linear, time-invariant, single-input, single-

output system in phase-variable canonical form. The identification would invariably consist of obtaining an nth order input-output description of the nth order state model which could be used as a measurement equation from which parameters could be identified. Initially ways of identifying systems with measurement noise only were studied and proved to converge. Later, state noise multiplying unknown coefficients was included and conditions for convergence and other requirements established (see Sakrison, Saridis and Stein, and Saridis). Some books on parameter identification have appeared that discuss and apply the more basic algorithms to SISO system identification (see Graupe a and Mendel).

The beginnings of identification of MIMO systems were made possible once row companion canonical form was published in a paper by Luenberger in 1967. A set of input/output equations for systems in that canonical form in papers by Bonivento and Guidorzi, and Shrikhance, etal. without noise and with state and measurement noise in Irwin and Roberts are derived. Whether the dependence coefficients can be identified in the noisy cases depends on the algorithm used and if the noise and equation complexity are constrained to fit the requirements of the algorithm. The same considerations will apply to the scalar and vector DME's derived in Chapter 3 when the existing stochastic approximation algorithms are adapted and applied.

In general gradient algorithms are not as fast to converge as least squares algorithms. This is because the latter have least squares convergence at each estimation step past the first n' estimates where n' is the dimension of the parameter vector and is greater than or equal to the order of the measurement equation.

The deterministic gradient approach (see Mendel) is a recursive method of identifying parameters in equations of the form

$$z(k) = P_1 g_1(k) + P_2 g_2(k) + \ldots + P_m g_m(k) \qquad (4.99)$$

in which P_1 through P_m are unknown constant parameters. The equation in vector form is

$$z(k) = \underline{g}^T(k)\underline{P} \qquad (4.100)$$

and an estimate to $z(k)$ is generated by

$$z(k) = \underline{g}^T(k)\hat{\underline{P}}(k) \qquad (4.101)$$

with $\hat{\underline{P}}(k)$ the kth estimate of $\underline{P}$ generated by a parameter identification algorithm. The difference, $z(k)-\hat{z}(k)$, is called the measurement error, $\tilde{z}(k)$.

Like other recursive identification algorithms the gradient method has the form

New estimate = old estimate + gain matrix X measurement error (4.102)

where the gain matrix is obtained by minimizing the performance criterion

$$J = 1/2\ \tilde{z}^2(k) \qquad (4.103)$$

with respect to $\hat{\underline{p}}(k)$. As previously discussed the negative of the gradient will be used in the gain matrix. Using the negative of this result in eq. (4.102)

$$\hat{\underline{p}}(k+1) = \hat{\underline{p}}(k) + K(k)\underline{g}(k)\tilde{z}(k) \qquad (4.106)$$

in which $K(k)$ is a weighting matrix to be determined later. Fig. 4.10 shows all relationships in a block diagram.

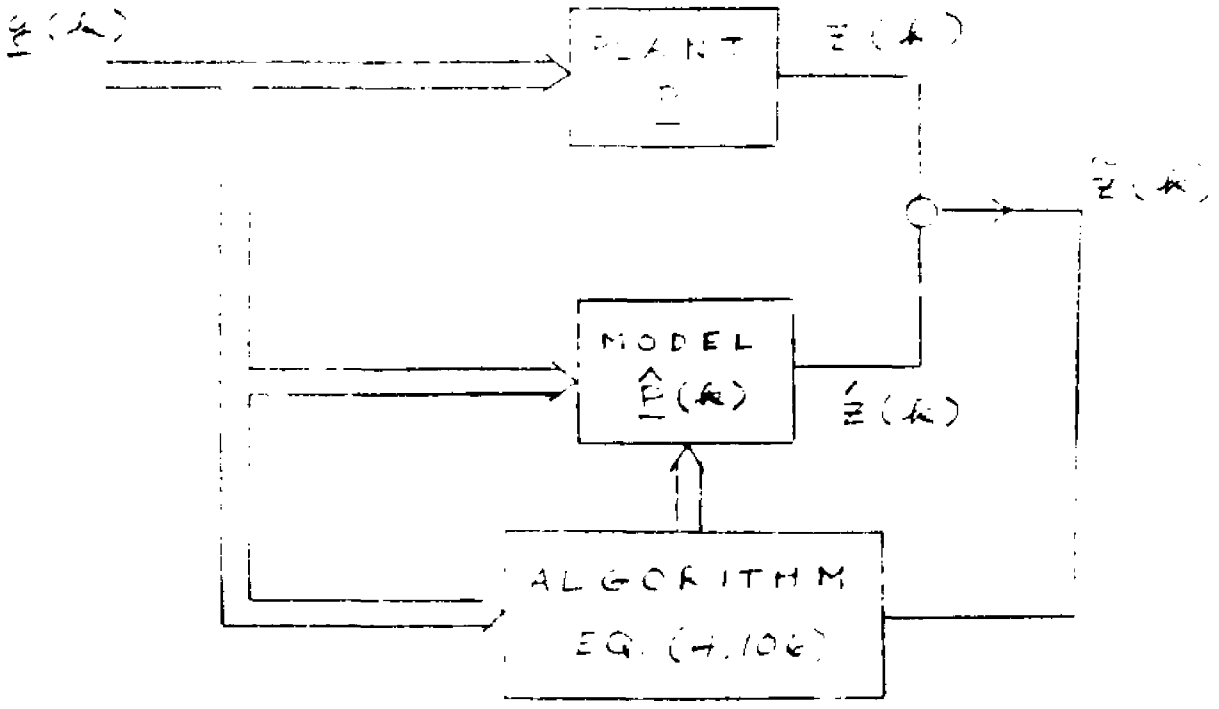

Fig. 4.10 Block diagram of deterministic gradient algorithm.

The weighting matrix, $R(k)$, is related to weights, $h_i(k)$, that are determined as follows. At least one of n weights, $h_i(k)$, satisfies

$$\frac{h_m(k) - h_m(k+1)}{h_m(k)} > \frac{h_i(k) - h_i(k+1)}{h_i(k)} \tag{4.110}$$

which can also be written as

$$\frac{h_m(k+1)}{h_m(k)} < \frac{h_i(k+1)}{h_i(k)} \tag{4.111}$$

for $i=1, \ldots, n$. The vector equation in eq. (4.106) reduces to n decoupled scalar equations if the weighting matrix is diagonal so that

$$R(k) = c(k)\,\mathrm{diag}(\, h_1(k), \ldots, h_n(k)\,) \tag{4.112}$$

where $c(k)$ is a scalar and is defined later. Then eq. (4.106) reduces to

$$\hat{P}_i(k+1) = \hat{P}_i(k) + c(k)h_i(k)g_i(k)\tilde{z}(k) \tag{4.113}$$

which greatly simplifies its being programmed removing the need for matrix arithmetic subroutines.

First order gradient algorithms have $R(k)=c(k)I$ so $h_i(k)=1$ for all i. Second order algorithms have more general weighting matrices such as in eq. (4.112) or derived by more complicated means which will be seen later. These second order algorithms usually have much faster rates of convergence.

Another rather obvious constraint on the weights is that they be finite and nonzero. Mendel proves that the estimates converge for any initial estimate if

$$0 < c(k) < 2/\sum_{i=1}^{m} h_i(k) g_i^2(k) \qquad (4.114)$$

provided the parameter error, $\tilde{\underline{p}}(k)$, and measurement, $\underline{g}(k)$ are never orthogonal, i.e.,

$$\tilde{\underline{p}}^T(k)\underline{g}(k) \neq 0 \qquad (4.115)$$

as explained in section 3.1. This result is obtained from the condition that the rate of change of the Lyapunov function be negative. The function

$$V(\tilde{\underline{p}},k) = V(\tilde{\underline{p}},k+1) - V(\tilde{\underline{p}},k) < 0 \qquad (4.116)$$

is a measure of the rate of change of the Lyapunov function as well as a measure of the rate of convergence of the algorithm in eq. (4.106). Minimization of eq. (4.116) gives the optimum choice for $c(k)$

$$c^*(k) = 1/\sum_{i=1}^{m} h_i(k) g_i^2(k) \qquad (4.117)$$

and when used in eq. (4.112) for $R(k)$ gives the Lyapunov optimum weighting matrix.

SISO systems in the form

$$\underline{x}(k+1) = A\underline{x}(k) + bu(k)$$

$$y(k) = c\underline{x}(k) \qquad (4.118)$$

with no state or measurement noise can be identified in phase-variable canonical form using the deterministic gradient algorithm. First, one notes that complete observability is implicit in the proposed canonical form and hence will be assumed in the following discussion. Second, complete controllability of the pair (A,b) is a sufficient but not necessary condition for identifiability. A necessary and sufficient condition for identification is

that all terms in the input/output equation for (4.118), eq. (4.119), have sufficient excitation (provided any algorithm dependent conditions are also met). Two cases will now be considered in which the pair is uncontrollable. First, suppose the A matrix is unstable. With adequate input excitation on nondynamic terms the coefficients can be identified since output terms receive excitation because the output is increasing due to instability regardless of inputs or initial conditions. Second, suppose the A matrix is stable. The only source of excitation left for dynamic terms is from the initial condition of the system. This may or may not be sufficient excitation depending on how long the identification algorithm takes, what parameter precision is desired, and whatever other numerical requirements are present.

Eq. (4.118) is equivalent to this input/output equation

$$y(k+n) + a_n y(k+n-1) + \ldots + a_1 y(k) = b_n^* u(k+n-1) + \ldots + b_1^* u(k) \qquad (4.119)$$

in which a_n through a_1 are the coefficients of the characteristic equation of A and

$$\underline{b}^* = (b_1^*, \ldots, b_n^*) = \underline{1}\tilde{b} \qquad (4.120)$$

with

$$
1 = \begin{bmatrix}
a_1 & a_2 & \cdots & a_n & 1 \\
a_3 & a_4 & & 1 & \\
\cdot & & & & \\
\cdot & & & & \\
\cdot & & & & \\
a_n & & & 0 & \\
1 & & & &
\end{bmatrix} \qquad (4.121)
$$

and $\tilde{b}$ is the canonical although still full b matrix discussed in Chapter 2. With

$$\underline{a}^T = (-a_1, \ldots, -a_n) \qquad (4.122)$$

and

$$\bar{y}(k) = (y(k), \ldots, y(k+n-1))^T$$

$$\bar{u}(k) = (u(k), \ldots, u(k+n-1))^T$$

eq. (4.119) can be written as

$$y(k+n) = \begin{bmatrix} \bar{y}^T(k), \bar{u}^T(k) \end{bmatrix} \begin{bmatrix} \underline{a} \\ --- \\ \underline{b}^* \end{bmatrix} \qquad (4.123)$$

which is in the form of eq. (4.100). In eq. (4.106) these associations are made:

$$z(k) = y(k+1)$$

$$\underline{k}(k) = (\underline{a}^T, \underline{b}^{*T})^T$$

$$\underline{g}(k) = (\bar{y}^T(k), \bar{u}^T(k))^T \qquad (4.124)$$

to use the gradient algorithm.

In the MIMO no noise case the state description is

$$\underline{x}(k+1) = A\underline{x}(k) + b\underline{u}(k)$$

$$\underline{y}(k) = C\underline{x}(k) \qquad (4.125)$$

and let it be controllable for the reasons given in the SISO case. If (A,C) is an observable pair it can be transformed to row companion canonical form as seen in Chapter 2. If (A,C) is unobservable part of the system in eq. (4.125) can be identified in row companion form using the vector DME to be discussed shortly. The details for obtaining the input/output equations that describe eq. (4.125) are summarized in Appendix B. There are m of these equations with the form (see eq. (B.16))

$$y_i(k+p_i) = \underline{a}_i^{*T} y(k) + \underline{b}_i^{*T} u'(k) \qquad (4.126)$$

for $i=1, \ldots, m$. Rewriting eq. (4.126)

$$y_i(k+p_i) = \left[y^T(k) \;\vdots\; u'^T(k) \right] \begin{bmatrix} \underline{a}_i^* \\ --- \\ \underline{b}_i^* \end{bmatrix} \qquad (4.127)$$

which is in the form of eq. (4.100) and these associations are made

$$\begin{aligned} z(k) &= y_i^*(k+1) \\ \underline{f}(k) &= \left(\underline{a}_i^{*T} , \underline{b}^{*T} \right)^T \\ g(k) &= \left(\bar{y}^T(k) \;\vdots\; \bar{u}'^T(k) \right)^T \end{aligned} \qquad (4.128)$$

to use eq. (4.106).

Deterministic gradient identification will now be applied to identification of the scalar and vector DME's. From eq. (4.4) the no noise scalar DME can be written as

$$y_i(k+n) = \left[y_i^T(k) \;\vdots\; \underline{u}^T(k+n-1) \;\vdots\; \ldots \;\vdots\; \underline{u}^T(k) \right] \begin{bmatrix} \underline{b} \\ --- \\ \underline{pt} \end{bmatrix} \qquad (4.129)$$

which is in the form of eq. (4.100). The order of the equation is determined using the method in section 3.3. If the system in eq. (4.125) is observable from $y_i(k)$ the order of eq. (4.129) equals the order of eq. (4.125). If unobservable the order of eq. (4.129) will be less than that of eq. (4.125) with n replaced by this order and a lower order state description can be identified that yields the same input/output behavior as eq. (4.125). Eq. (4.106) can be used with the associations

$$z(k) = y_i(k+n)$$
$$\underline{\xi} = (\underline{p}^T, \underline{\xi}\underline{\xi}^{*T})^T$$
$$\underline{\zeta}(k) = (\bar{y}^T(k) \mid \underline{u}^T(k+n-1) \mid \ldots \mid \underline{u}^T(k))^T \qquad (4.130)$$

From eq. (4.30) a deterministic vector DME can be written in the form

$$y_i^{\ell}(k+p_i^{\ell}) = \left[\bar{y}_i^{\ell\,T}(k) \mid \bar{U}^T(k) \right] \begin{bmatrix} \underline{a}_i \\ --- \\ \underline{b}_i^* \end{bmatrix} \qquad (4.131)$$

for i=1, ..., m which casts it in the form of eq. (4.100). In section 3.4 a method of finding the p_i^{ℓ}, the structural invariants of the decentralized system, was discussed. If the entire system in eq. (4.125) is unobservable from the m_0 outputs available then the sum of the p_i^{ℓ} will be less than the order of eq. (4.125). A state description in row companion or other canonical form can be identified with order less than eq. (4.125). It will give the same input/output behavior, however, as the decentralized version

of eq. (4.125). If eq. (4.125) is observable from the n_i outputs, the order of the system identified from eq. (4.131) is the order of eq. (4.125). Eq. (4.106) can be used with these associations

$$\begin{aligned} z(k) &= y_{\lambda}^{i}(k+p_{\lambda}^{i}) \\ \underline{I}(k) &= (\underline{a}_{\lambda}^{T} \quad \underline{b}^{*T}_{\lambda})^{T} \\ \underline{g}(k) &= (\underline{\bar{y}}_{\lambda}^{i\,T}(k) \quad \underline{U}^{T}(k))^{T} \end{aligned} \tag{4.132}$$

in m_{λ} separate deterministic gradient identifiers.

Now that the use of the deterministic gradient algorithm for system identification has been considered, the measurement equation complexity will be complicated by considering that noise is added to the output measurement(s). Since the measurement vector contains delayed outputs, noise is then added to the measurement vector as well. Past inputs are also present in the measurement vector and the case of noise on the inputs is also included in the algorithm to be discussed. If inputs are deterministic the algorithms to be discussed are entirely applicable as a special case. The noisy measurement vector and output are defined as

$$\begin{aligned} \underline{gm}(k) &= \underline{g}(k) + \underline{n}(k) \\ ym(k) &= y(k) + v(k) \end{aligned} \tag{4.133}$$

where $\underline{n}(k)$ includes noise added to past output and/or input measurements.

Repeating the calculation of the gradient of the performance criterion in eq. (4.103) except using $\underline{g}_m(k)$ instead of $\underline{g}(k)$ gives

$$\text{grad } J = -\underline{g}_m(k)\tilde{z}(k) \quad . \tag{4.134}$$

Then eq. (4.101) is modified to

$$\hat{\underline{p}}(k+1) = \hat{\underline{p}}(k) + R(k)\underline{g}_m(k)\tilde{z}(k) \tag{4.135}$$

and Mendel shows that the algorithm converges in mean square and with probability 1 if eq. (4.135) is updated every s iterations where s is a spacing parameter that is dependent on the order of the measurement equation whose coefficients are being identified. Eq. (4.135) then can be written

$$\hat{\underline{p}}(k+s) = \hat{\underline{p}}(k) + R(k)\underline{g}_m(k)\tilde{z}(k) \tag{4.136}$$

for k=s, 2s, 3s, ... For mean square convergence the following assumptions are required:

1) The vectors, $\underline{g}(k)$ and $\hat{\underline{p}}(k)$, are independent and the covariance matrix of $\underline{g}(k)$, conditional on $\hat{\underline{p}}(k)$, is positive definite and constant.
2) The vectors, $\underline{g}(k)$ and $\underline{n}(k)$ are independent.
3) The noise terms $v(k)$ and $\underline{n}(k)$ are independent.
4) All noise terms have zero mean and the covariance of $\underline{n}(k)$ is constant and known.

With these conditions eq. (4.135) is found to give estimates biased by the amount

$$-R(k)\,\Sigma_n\, E(\,\underline{n}(k)\underline{n}^T(k)\,) \quad . \tag{4.137}$$

Therefore, if Σ_n is known and since $R(k)$ is known the bias can be added to eq. (4.135) to give

$$\hat{\underline{p}}(k+s) = (\,I + h(k)\,\Sigma_n\,)\hat{\underline{p}}(k) + h(k)\underline{g}_m(k)\tilde{z}(k) \tag{4.138}$$

for k=s, 2s, ... Fig. 4.11 illustrates the algorithm using the generic symbols.

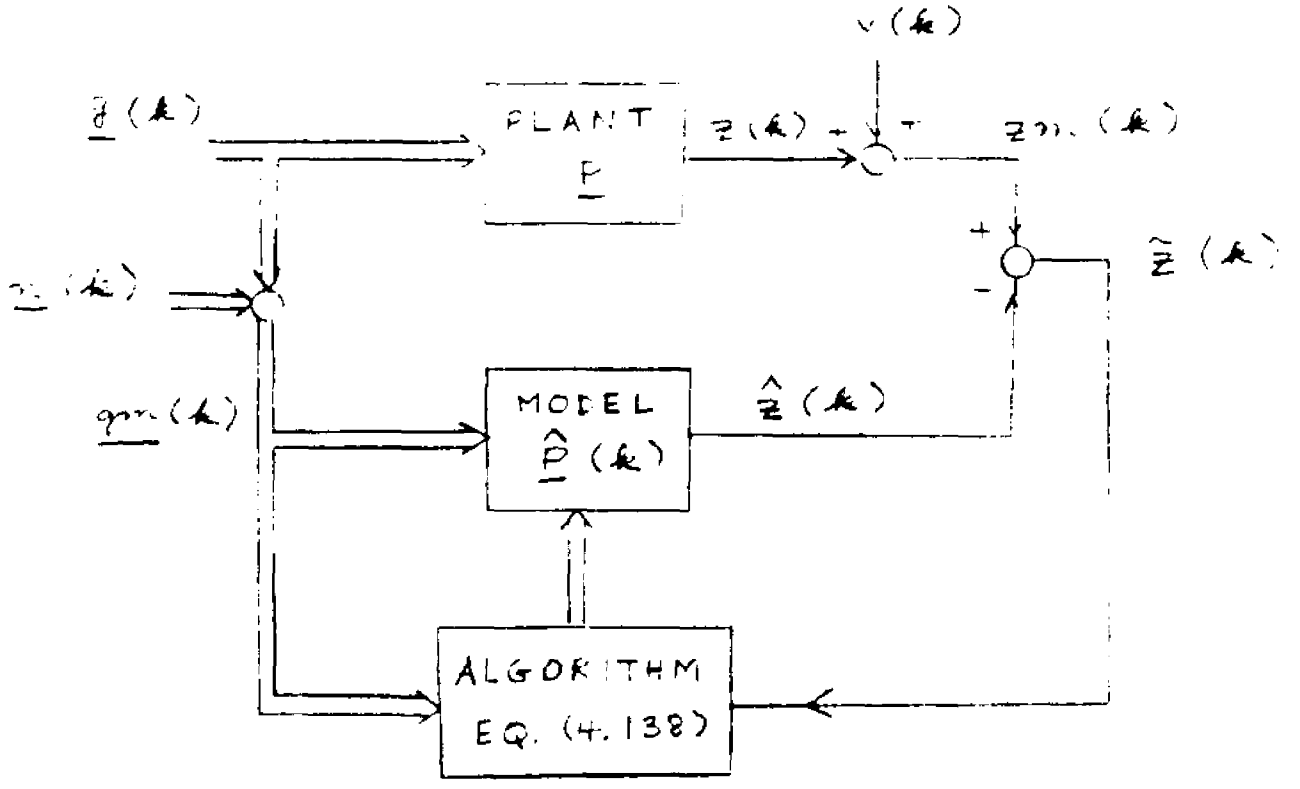

Fig. 4.11 Block diagram of stochastic gradient algorithm with measurement noise only.

A condition to avoid correlation between successive measurement vectors that would result in biased estimates is

$$s > n-1 \ . \tag{4.149}$$

The smallest s one can use in both cases is n which is desirable to update estimates as fast as possible.

The MIMO case and vector LMS have $\underline{g}(k)$ of the general form

$$\underline{g}(k) = (y_1(k), \ldots, y_1(k+p_1-1) , \ldots, y_m(k), \ldots, y_m(k+p_m-1), \underline{u}^T(k), \ldots, \underline{u}^T(k+n-1)) \tag{4.152}$$

dropping the decentralized notation in the second case. Since all p_i must be less than n then the earliest term in $\underline{g}(k-s)$ is $\underline{u}(k+n-1-s)$ and the latest signal in eq. (4.152) has index k then eq. (4.149) again applies and the smallest s is $s=n$.

To apply this algorithm to the four measurement equations (from SISO, MIMO input/output equations, vector and scalar LMS's) for system identification the $\underline{g}(k)$ and $\underline{p}(k)$ are the same as in eqs. (4.124), (4.128), (4.130), and (4.132). Additional associations will be made but first allowing for noise on input measurements define

$$u_m(k) = u(k) + n_u(k) \ . \tag{4.153}$$

For the SISO case define the additional quantities

$$z_m(k) = y_m(k+n)$$

$$\underline{g}_m(k) = \left(\overline{\underline{y}_m}^T(k) \;\vdots\; \overline{\underline{u}_m}^T(k) \right)^T$$

$$\underline{n}(k) = \left(v(k), \ldots, v(k+n-1), n_{\mu}(k), \ldots, n_{\mu}(k+n-1) \right)^T \tag{4.154}$$

to use eq. (4.138).

Next, for MIMO systems and centralized measurements make the associations

$$z(k) = y_{\lambda}(k+p_{\lambda})$$

$$\underline{g}_m(k) = \left(y_{m_1}(k), \ldots, y_{m_1}(k+p_1), \ldots, y_{m_m}(k), \ldots, y_{m_m}(k+p_m-1), \underline{u}_m^T(k), \ldots, \underline{u}_m^T(k+n-1) \right)^T$$

$$\underline{n}(k) = \left(v_1(k), \ldots, v_1(k+p_1-1), \ldots, v_m(k), \ldots, v_m(k+p_m-1), \underline{n}_{\mu}^T(k), \ldots, \underline{n}_{\mu}^T(k+n-1) \right)^T \tag{4.155}$$

to use eq. (4.138).

For the scalar DME the additional associations are

$$z(k) = y_{m_{\lambda}}(k+n)$$

$$\underline{g}_m(k) = \left(y_{m_{\lambda}}(k), \ldots, y_{m_{\lambda}}(k+n-1), \underline{u}_m^T(k+n-1), \ldots, \underline{u}_m^T(k) \right)^T$$

$$\underline{n}(k) = \left(v_{\lambda}(k), \ldots, v_{\lambda}(k+n-1), \underline{n}_{\mu}^T(k+n-1), \ldots, \underline{n}_{\mu}^T(k) \right)^T \tag{4.156}$$

for eq. (4.136).

For the vector DME the associations for a subsystem are

$$z(k) = y_{\lambda}^{\ell}(k+p_{\lambda}^{\ell})$$

$$\underline{g}_m(k) = \left(y_{m_1}^{\ell}(k), \ldots, y_{m_1}^{\ell}(k+p_1^{\ell}-1), \ldots, y_{m_{\ell}}^{\ell}(k), \ldots, y_{m_{\ell}}^{\ell}(k+p_{m_{\ell}}^{\ell}-1), \underline{u}_m^T(k), \ldots, \underline{u}_m^T(k+n-1) \right)^T$$

$$D(k) = (v_1(k), \ldots, v_1(k+p_1^{\ell}-1), \ldots, v_{m_\ell}(k), \ldots, v_{m_\ell}(k+p_{m_\ell}^{\ell}-1), D_{\mu}^{T}(k+n-1), \ldots, D_{\mu}^{T}(k))^{T}$$

to use eq. (4.138). In eqs. (4.156) and (4.157) it is understood from sections 3.4 and 3.5 that $n^{*}<n$ is used instead of n if the entire system is unobservable with the given outputs.

In the stochastic convergence proof of Mendel with $R(k)$ taking the specific form

$$R(k) = p(k) I_n \tag{4.158}$$

the conditions on $p(k)$ are

$$\begin{aligned} &1)\ p(k)>0 \\ &2)\ \sum_{k=0}^{\infty} p(k) \\ &3)\ \sum_{k=0}^{\infty} p^2(k) < \infty \end{aligned} \tag{4.159}$$

for mean square convergence. One possibility for $R(k)$ meeting these criteria is

$$R(k) = \frac{n}{k^m} I_n \tag{4.160}$$

where n is a constant and $1/2<m<1$. In practice it is found that convergence with this $R(k)$ is very slow because $p(k)$ decreases as k increases providing too much attenuation of the correction term in eq. (4.138). Mendel shows that beginning the algorithm with the Lyapunov optimum weighting matrix and at some arbitrary time switching to eq. (4.160) gives good convergence. This technique is used in the simulations of this chapter successfully.

An algorithm for system identification with measurement noise only was applied (in eqs. (4.154) to (4.157)) to SISO and MIMO systems plus the scalar and vector DML's. The input/output description will now be complicated by the addition of state noise. The SISO linear discrete time nth order system is now

$$\underline{x}(k+1) = A\underline{x}(k) + bu(k) + tw(k)$$

$$ym(k) = y(k) + v(k) \qquad (4.161)$$

where $w(k)$, $n_{ml}(k)$ and $v(k)$ are mutually independent sequences of random variables with zero means and constant known variances σ_w^2, σ_{ml}^2 and σ_v^2. As before assume the system is completely observable and controllable so that it can be transformed to phase-variable canonical form.

As in eq. (4.119) eq. (4.161) can be written as a finite-difference equation but with extra terms due to $w(k)$

$$y(k+n) + a_n y(k+n-1) + \ldots + a_1 y(k) = b_n^* u(k+n-1) + \ldots + b_1^* u(k) + t_n^* w(k+n-1) + \ldots + t_1^* w(k) \quad . \qquad (4.162)$$

The coefficients of the inputs form the vector $\underline{b}^*$ as in eq. (4.120). Define

$$\underline{t}^* = (t_1^* , \ldots , t_n^*)^T = T\tilde{t} \qquad (4.163)$$

where $\tilde{t}$ is the canonical, but still arbitrary, t matrix after transformation to phase-variable canonical form and must be known. The T matrix is defined as in eq. (4.121). Define these quantities

$$\underline{g}(k) = (y(k), \ldots, y(k+n-1), u(k), \ldots, u(k+n-1))^T$$

$$\bar{u}(k) = (u(k), \ldots, u(k+n-1))^T$$

$$\bar{w}(k) = (w(k), \ldots, w(k+n-1))^T$$

$$\underline{n}(k) = (v(k), \ldots, v(k+n-1), n(k), \ldots, n(k+n-1))^T$$

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

$$\underline{g_m}(k) = (y_m(k), \ldots, y_m(k+n-1), u_m(k), \ldots, u_m(k+n-1))^T$$

$$\underline{g}(k) + \underline{n}(k)$$

and

$$\underline{\theta} = (\underline{a}^T, \underline{b}^{*T})^T. \tag{4.164}$$

The input/output equation can be rewritten as

$$z(k) = y(k+n) = \underline{g}^T(k)\underline{\theta} + \bar{w}^T(k)\underline{b}^* \tag{4.165}$$

and a measured input/output equation is given by

$$z_m(k) = y_m(k+n) = \underline{g_m}^T(k)\underline{\theta} - \underline{n}^T(k)\underline{\theta}$$
$$\bar{w}^T(k)\underline{b}^* + v(k+n) \tag{4.166}$$

where the last three terms on the right can be regarded as a composite noise term and easily shown to have zero mean. Saridis and Stein used the following algorithm

$$\hat{\underline{\theta}}(k+n) = \hat{\underline{\theta}}(k-1) + \rho(k-1)\underline{g_m}(k)$$
$$X \left[z_m(k) - \overline{g_m}^T(k)\hat{\underline{\theta}}(k-1) \right]$$
$$+ \begin{bmatrix} \sigma_v^2 I & 0 \\ 0 & \sigma_w^2 I \end{bmatrix} \hat{\underline{\theta}}(k-1)$$
$$+ \begin{bmatrix} \sigma_{ww}^2 DD^T & 0 \\ 0 & 0 \end{bmatrix} \hat{\underline{\theta}}(k-1) - \begin{bmatrix} \sigma_{ww}^2 d^* \\ 0 \end{bmatrix} \tag{4.167}$$

for K=1, n+2, 2n+3, ... such that the spacing parameter previously discussed is n+1. The matrix D and vector d* are

$$L = \begin{bmatrix} 0 & & & 0 \\ \tilde{t}_1 & 0 & & \\ \tilde{t}_2 & \tilde{t}_1 & & \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ \tilde{t}_{n-1} & \tilde{t}_{n-2} \cdots & \tilde{t}_1 & 0 \end{bmatrix} \qquad d^* = D \begin{bmatrix} \tilde{t}_n \\ \tilde{t}_{n-1} \\ \cdot \\ \cdot \\ \cdot \\ \tilde{t}_1 \end{bmatrix} \tag{4.168}$$

derived from the canonical $\underline{t}$ vector, $\tilde{\underline{t}}$. The negatives of the last three terms in eq. (4.167) are precisely the amounts by which the estimates are biased by the state and measurement noise and therefore the added terms cancel the these biases. The conditions on the weighting sequence are the same as for the algorithm in eq. (4.106) and are given in eq. (4.111). Proof of mean square convergence is given in the paper.

Now consider the MIMO centralized case which has the state description

$$\underline{x}(k+1) = A\underline{x}(k) + b\underline{u}(k) + \Gamma\underline{w}(k)$$

$$\underline{y}_m(k) = C\underline{x}(k) + \underline{v}(k) \tag{4.169}$$

and assuming complete observability the corresponding input/output equation from Appendix B

$$y_i(k+p_i) = \underline{e}_i^{*T}\, \underline{y}(k) + \underline{p}_i^{*T}\, \bar{u}'(k) + \underline{t}_i^{*T}\, \bar{w}'(k) \tag{4.170}$$

for i=1, ..., m. Define the vectors

$$\underline{g}(k) = (y_1(k), \ldots, y_1(k+p_1-1), \ldots, y_m(k), \ldots, y_m(k+p_m-1), \underline{u}^T(k), \ldots, \underline{u}^T(k+n-1))^T$$

$$\bar{u}(k) = (\underline{u}^T(k), \ldots, \underline{u}^T(k+n-1))^T$$

$$\bar{w}(k) = (\underline{w}^T(k), \ldots, \underline{w}^T(k+n-1))^T$$

$$\underline{h}(k) = (v_1(k), \ldots, v_1(k+p_1-1), \ldots, v_m(k), \ldots, v_m(k+p_m-1), n_1(k), \ldots, n_1(k+n-1))^T$$

$$z(k) = y_i(k+p_i)$$

$$z_m(k) = y_{m_i}(k+p_i)$$

$$\underline{g}_m(k) = (y_{m_1}(k), \ldots, y_{m_1}(k+p_1-1), \ldots, y_{m_m}(k), \ldots, y_{m_m}(k+p_m-1), \underline{u}_m^T(k), \ldots, \underline{u}_m^T(k+n-1))^T$$

and

$$\underline{P} = (\underline{a}^{*T}_i , \underline{b}^{*T}_i)^T . \qquad (4.171)$$

Rewriting the input/output equation in vector form yields

$$z(k) = y_i(k+p_i) = \underline{g}^T(k)\underline{P} + \bar{w}^T(k)\underline{t}^*_i \qquad (4.172)$$

and the measured input/output equation is

$$z_m(k) = y_{m_i}(k+p_i) = \underline{g}_m^T(k)\underline{P} - \underline{h}^T(k)\underline{P} - \bar{w}(k)\underline{t}^* + v_i(k+p_i) \qquad (4.173)$$

where again the last three terms on the right constitute a composite noise term. As long as every noise source, all elements in $\underline{w}(k)$ and $\underline{v}(k)$, has zero mean and constant variance, and all sources are mutually independent then all terms in the composite noise sum are independent random variables and that sum has zero mean. The algorithm is then applicable to the MIMO case provided variances and $\underline{t}^*_i$ are known and applicable changes are made in the bias terms of eq. (4.167).

For the scalar DME the state description is

$$\underline{x}(k+1) = A\underline{x}(k) + B\underline{u}(k) + I\underline{w}(k)$$

$$ym_i(k) = c_i\underline{x}(k) + v_i(k) \qquad (4.174)$$

with this equivalent input/output description

$$ym_i(k+n) = \underline{\alpha}^T \overline{ym}(k) + \underline{\beta}\underline{b}^{*T} u(k)$$

$$+ \underline{\beta}\underline{w}^{*T} \bar{w}(k) - \underline{\beta}^T \bar{v}_i(k) + v_i(k+n) \qquad (4.175)$$

where n is replaced by $n' \le n$ if (A, c_i) is an unobservable pair. Define the quantities

$$\underline{g}(k) = (y_i(k), \ldots, y_i(k+n-1), \underline{u}^T(k), \ldots, \underline{u}^T(k+n-1))^T$$

$$\bar{u}(k) = (\underline{u}^T(k), \ldots, \underline{u}^T(k+n-1))^T$$

$$\bar{w}(k) = (\underline{w}^T(k), \ldots, \underline{w}^T(k+n-1))^T$$

$$\underline{n}(k) = (v_i(k), \ldots, v_i(k+n-1))^T$$

$$z(k) = y_i(k+n)$$

$$zm(k) = ym_i(k+n)$$

$$\underline{gm}(k) = (ym_i(k), \ldots, ym_i(k+n-1), \underline{um}^T(k), \ldots, \underline{um}^T(k+n-1))^T$$

and

$$\underline{\beta} = (\underline{\alpha}^T, \underline{\beta}\underline{b}^{*T})^T . \qquad (4.176)$$

the input/output equation in vector form is

$$z(k) = y_i(k+n) = \underline{g}^T(k)\underline{\beta} + \bar{w}^T(k)\underline{\beta}\underline{w}^* \qquad (4.177)$$

and the measured input/output equation is

$$zm(k) = ym_i(k+n) = \underline{gm}^T(k)\underline{\beta} - \underline{n}^T(k)\underline{\beta}$$

$$- \bar{w}^T(k)\underline{\beta}\underline{w}^* + v_i(k+n) \qquad (4.178)$$

where again if all noise sources are independent and their variances known and $\underline{\beta}\underline{w}^*$ is known then the algorithm in eq.

(4.167) is applicable with the corresponding changes. For the vector DME the state description is

$$\underline{x}(k+1) = A\underline{x}(k) + B\underline{u}(k) + T\underline{w}(k)$$
$$\underline{ym}^{\ell}(k) = C^{\ell}\underline{x}(k) + \underline{v}^{\ell}(k) \qquad (4.179)$$

and the corresponding DME is

$$ym_i^{\ell}(k+p_i^{\ell}) = \underline{a}_i^{\top}\overline{\underline{ym}}^{\ell}(k) + \underline{b}^{*\top}\overline{u}(k) + \underline{t}^{*\top}_i\overline{w}(k) - \underline{e}_i^{\top}\overline{v}^{\ell}(k) + v_i(k+p_i^{\ell}) \qquad (4.180)$$

for $i=1, \ldots, m_{\ell}$. If (A,C^{ℓ}) is an unobservable pair then

$$\sum_{i=1}^{m_{\ell}} p_i^{\ell} = n^* < n \qquad (4.181)$$

and a lower order system can be identified in row companion form. Define the vectors

$$\underline{u}(k) = (y_1^{\ell}(k), \ldots, y_1^{\ell}(k+p_1^{\ell}-1), \ldots, y_{m_{\ell}}^{\ell}(k), \ldots, y_{m_{\ell}}^{\ell}(k+p_{m_{\ell}}-1), \underline{u}^{\top}(k), \ldots, \underline{u}^{\top}(k+n-1))^{\top}$$
$$\overline{u}(k) = (\underline{u}^{\top}(k), \ldots, \underline{u}^{\top}(k+n-1))^{\top}$$
$$\overline{w}(k) = (\underline{w}^{\top}(k), \ldots, \underline{w}^{\top}(k+n-1))^{\top}$$
$$\underline{v}(k) = (v_1^{\ell}(k), \ldots, v_1^{\ell}(k+p_1^{\ell}-1), \ldots, v_{m_{\ell}}^{\ell}(k), \ldots, v_{m_{\ell}}^{\ell}(k+p_{m_{\ell}}-1), \underline{n}_{\mu}^{\top}(k), \ldots, \underline{n}_{\mu}^{\top}(k+n-1))^{\top}$$
$$z(k) = y_i^{\ell}(k+p_i^{\ell})$$
$$zm(k) = ym_i^{\ell}(k+p_i^{\ell})$$
$$\underline{zm}(k) = (ym_1^{\ell}(k), \ldots, ym_1^{\ell}(k+p_1^{\ell}-1), \ldots, ym_{m_{\ell}}^{\ell}(k), \ldots, ym_{m_{\ell}}^{\ell}(k+p_{m_{\ell}}-1), \underline{um}^{\top}(k), \ldots, \underline{um}^{\top}(k+n-1))^{\top}$$

and

$$\underline{l} = (\underline{e}_i^{\top}, \underline{t}^{*\top}_i)^{\top}. \qquad (4.182)$$

The input/output equation is then

$$z(k) = y_{\lambda}^{\ell}(k+p_{\lambda}^{\ell}) = \underline{g}^{T}(k)\underline{\ell} - \bar{w}^{T}(k)\underline{t}_{\lambda}^{*} \qquad (4.163)$$

and the measured input/output equation is

$$z_m(k) = y_{m\lambda}^{\ell}(k+p_{\lambda}^{\ell}) = \underline{g}_m^{T}(k)\underline{\ell} - \underline{n}^{T}(k)\underline{l} - \bar{w}^{T}(k)\underline{t}_{\lambda}^{*} + v_{\lambda}(k+p_{\lambda}^{\ell}) \qquad (4.164)$$

and once again the algorithm in eq. (4.167) is applicable with the proper modifications.

4.3 State Descriptions for Three Examples

The three examples at the end of section 4.1 have been identified in accordance with the results of Chapter 3 using the deterministic gradient algorithm of eq. (4.106). In each example the initial parameter estimates were set equal to zero arbitrarily.

The state description of the first example is given in eq. (4.61). The measurement equation for the scalar DME in eq. (4.129) for n=2 for this observable example is

$$y_1(k+2) = \left[\bar{y}_1^{T}(k) \;\vdots\; \underline{u}^{T}(k+1) \;\vdots\; \underline{u}^{T}(k)\right] \begin{bmatrix} \underline{p} \\ \cdots \\ \underline{pb}^{*} \end{bmatrix} \qquad (4.165)$$

where the order is assumed either known or to have been determined using the procedure in section 3.3. Then in eq. (4.130) these associations are made

$$z(k) = y_1(k+2)$$

$$\underline{\ell} = (\,.4,\ .5,\ 1,\ 0,\ .7,\ .6\,)^{T}$$

$$\underline{u}(k) = (\,y_1(k),\ y_1(k+1),\ u_1(k+1),\ u_2(k+1),$$

$$u_1(k),\ u_2(k)\ )^T \qquad (4.166)$$

and the Lyapunov optimum weighting matrix in eq. (4.117) was used. With the order in $g(k)$ shown above the weights used were

$$\underline{h}(k) = (\ 1,\ 2,\ 2,\ 2,\ 1,\ 1\ )^T \qquad (4.167)$$

which is a simplified fading memory approach where later signals are weighted twice as much as the earlier signals.

The inputs used were derived from an IMSL subroutine that gives uniform numbers from a pseudo-random number generator. Inputs for $u_1(k)$ were uniformly obtained on the interval (-10, 10) and $u_2(k)$ from the interval (-20, 20) both with zero mean. Figs. (4.12)-(4.14) show the convergence of the 6 parameters. Convergence is much slower for the same precision than the recursive least squares algorithm.

Using Appendix B and the identified parameters the state model will now be found. First, assemble all pertinent variables in Appendix B

$$F = \begin{bmatrix} 0 & 0 \\ & \\ b_1(1) & \end{bmatrix}$$

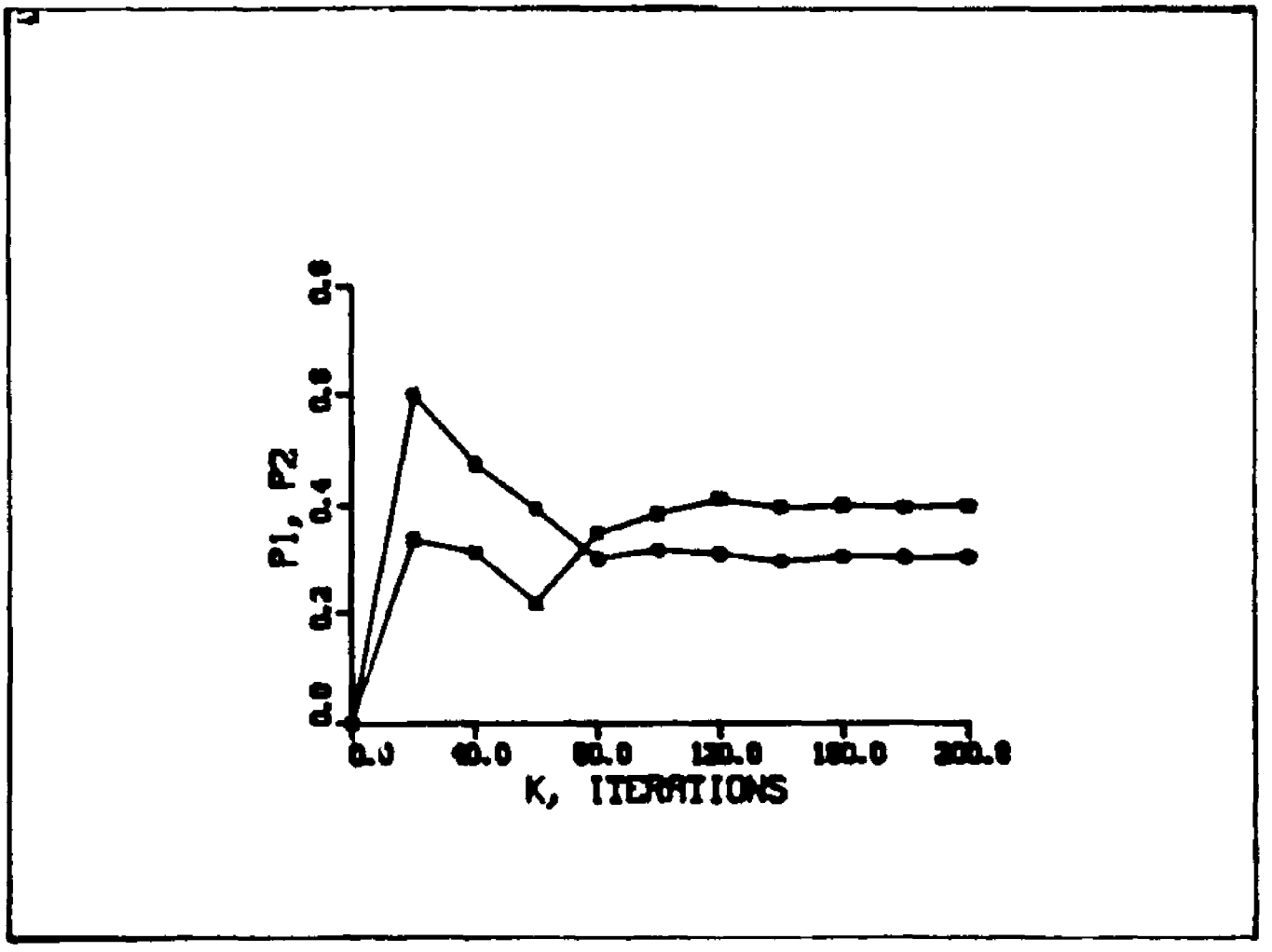

Fig. 4.12 Parameter P1 and P2 convergence using deterministic gradient algorithm in observable example.

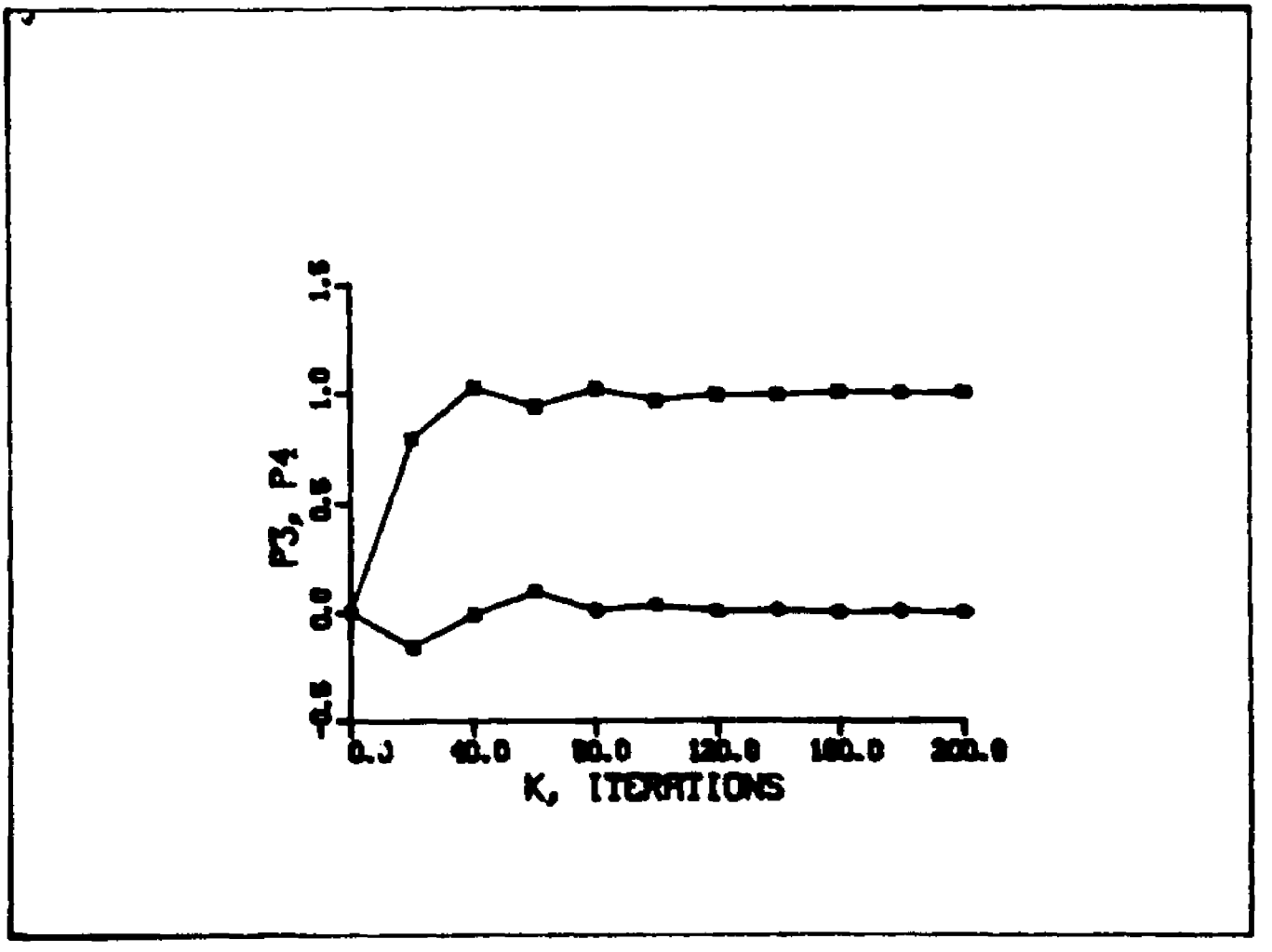

Fig. 4.13 Parameter P3 and P4 convergence using deterministic gradient algorithm in observable example.

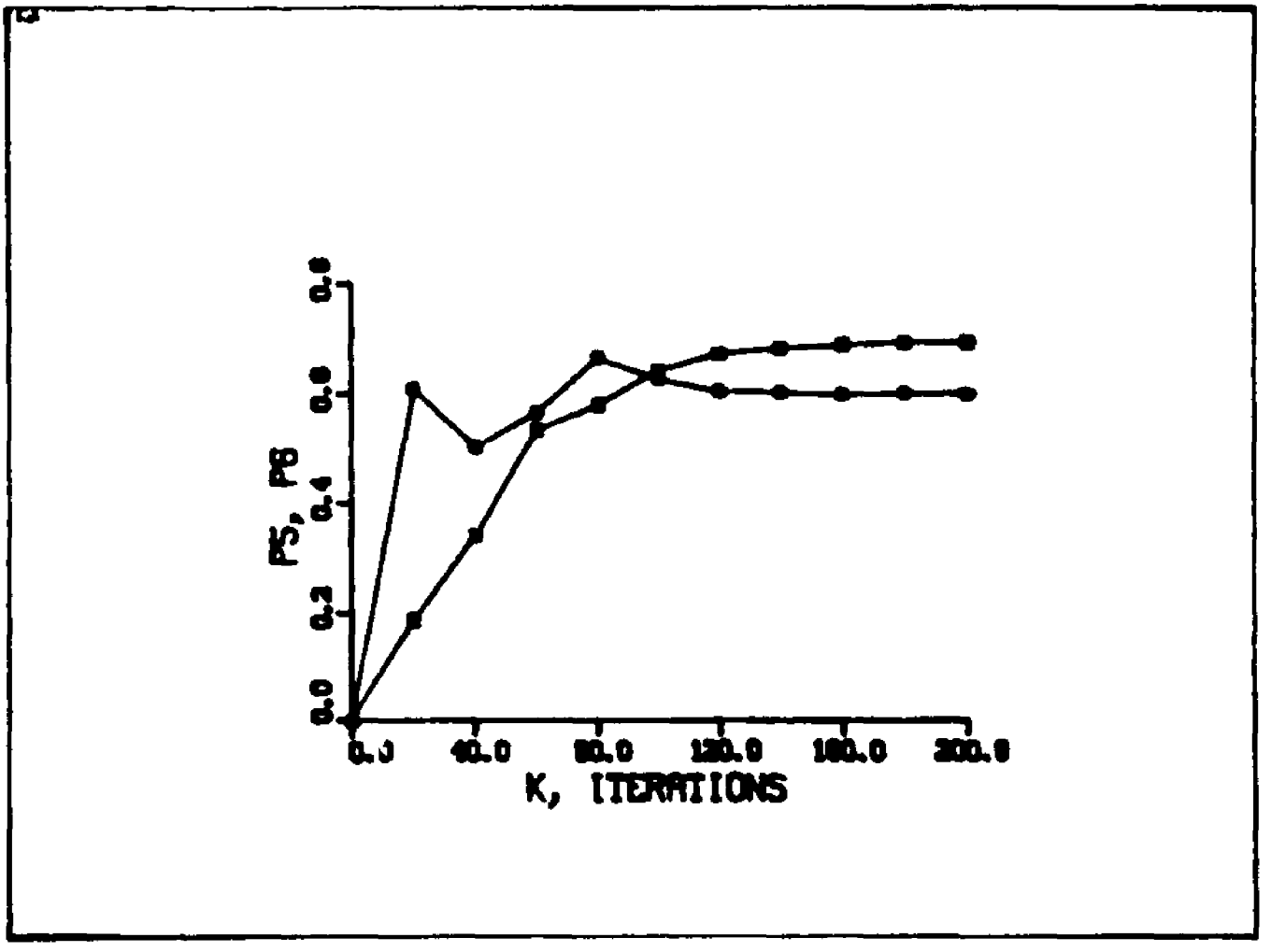

Fig. 4.14 Parameter P5 and P6 convergence using deterministic gradient algorithm in observable example.

$$B = \begin{bmatrix} b_1(1) \\ \\ b_1(2) \end{bmatrix}$$

$$B^* = \left[B \mid M \right] = \begin{bmatrix} b_1(1) & 0 & 0 \\ \\ b_1(2) & I_1(1) \end{bmatrix}$$

$$B^{*C} = \left[BC \mid MC(1) \right] = \left[b_1(2) \quad b_1(1) \right]$$

$$M0 = \left[M \mid 0 \right] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ \\ b_1(1) & 0 & 0 \end{bmatrix}$$

$$BC = b_1(2)$$

$$MC(1) = b_1(1) \qquad (4.188)$$

where B is the identified B matrix. From eq. (b.19)

$$\left[b_1(2) \quad b_1(1) \right] - \underline{a}^* \begin{bmatrix} 0 & 0 & 0 & 0 \\ \\ b_1(1) & 0 & 0 \end{bmatrix}$$

$$= \left[BC^*(1) \mid BC^*(2) \right] \qquad (4.189)$$

where $\underline{a}^*$ is the single significant row of the A matrix. From the parameter vector we have

$$\underline{a}^* = (.4 \quad .3)$$

$$BC^*(1) = (.7 \quad .6)$$

$$BC^*(2) = (1 \quad 0) \qquad (4.190)$$

and using the Shrikhande, etal. procedure described in Appendix B yields

$$b_1(1) - \underline{a}^* \lambda 0 = (1 \quad 0) \qquad (4.191)$$

starting from the right of eq. (4.189) so that

$$I_1(1) = (1 \quad 0) .$$

Next, on the left of eq. (4.189)

$$b_1(2) - \underline{a}^* \begin{bmatrix} 0 & 0 \\ & \\ b_1(1) & \end{bmatrix} = \begin{bmatrix} .7 & .6 \end{bmatrix} \tag{4.192}$$

or

$$b_1(2) = (.7 \quad .6) - (.3 \quad 0)$$

$$= (1 \quad .6) \ . \tag{4.193}$$

The identified state model is

$$\underline{x}(k+1) = \begin{bmatrix} 0 & 1 \\ & \\ .4 & .3 \end{bmatrix} \underline{x}(k) + \begin{bmatrix} 1 & 0 \\ & \\ 1 & .6 \end{bmatrix} \underline{u}(k) \tag{4.194}$$

$$y(k) = (1 \quad 0) \ \underline{x}(k) \ . \tag{4.195}$$

Eq. (4.129) for the second example with the state-space description in eq. (4.80) is for $n^*=1$

$$y_1(k+1) = \left[y_1(k) \ \vdots \ \underline{u}(k) \right] \begin{bmatrix} p \\ \cdots \\ \underline{p}\underline{b}^* \end{bmatrix} \tag{4.196}$$

where n^* is known or determined using section 3.3. In eq. (4.130) the associations are made

$$z(k) = y_1(k+1)$$

$$\underline{p} = (.9 \quad 1 \quad 2)^T$$

$$\underline{g}(k) = (y_1(k), \ u_1(k), \ u_2(k))^T \tag{4.197}$$

with the Lyapunov optimum weighting and the weights were simply

$$\underline{p}^T(k) = (1 \quad 1 \quad 1) \tag{4.198}$$

and inputs were derived uniformly on the interval (-10, 10). Figs. (4.15)-(4.17) show the convergence of the 3 parameters.

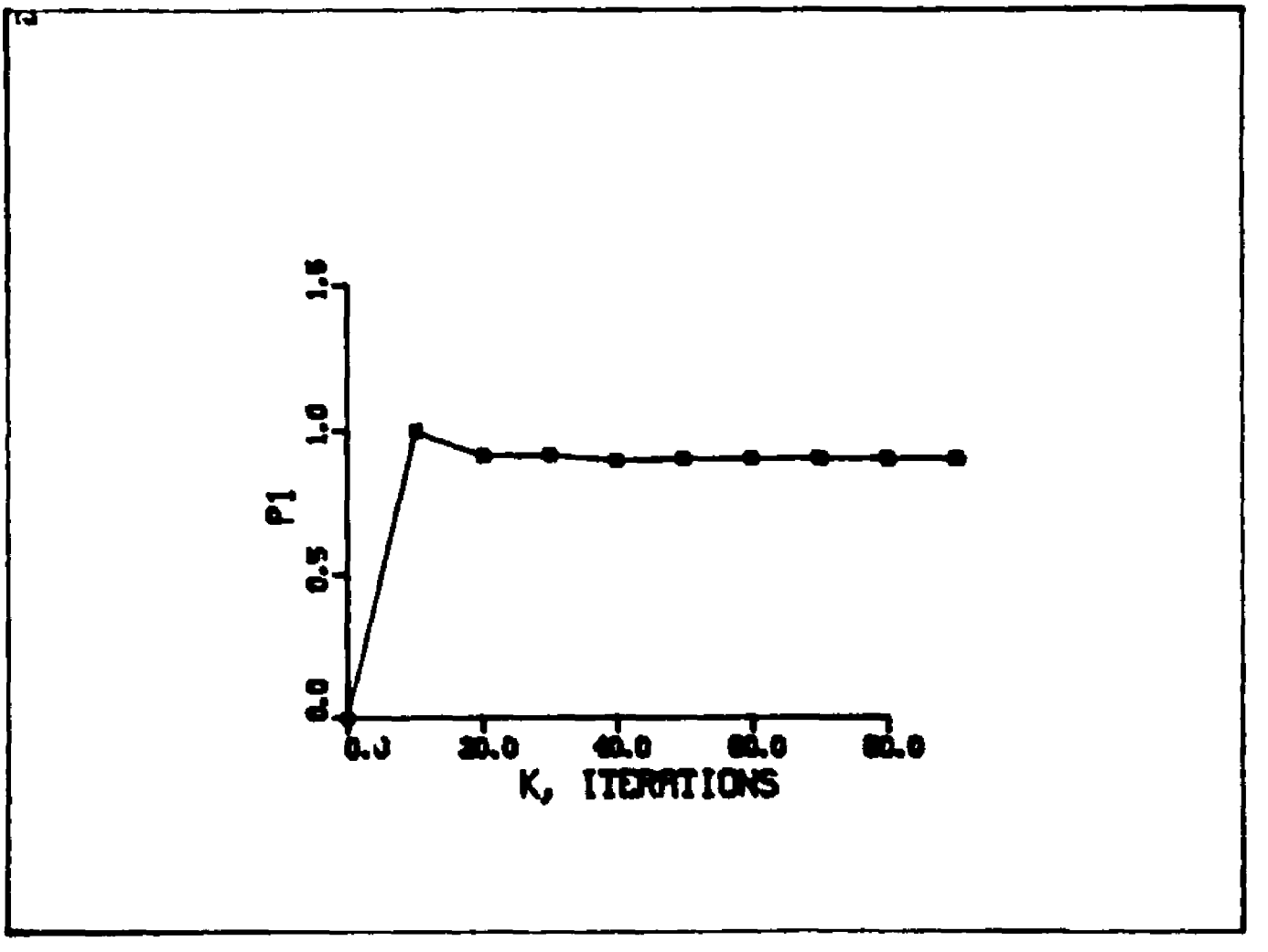

Fig. 4.15 Parameter P1 convergence using deterministic gradient algorithm in unobservable example.

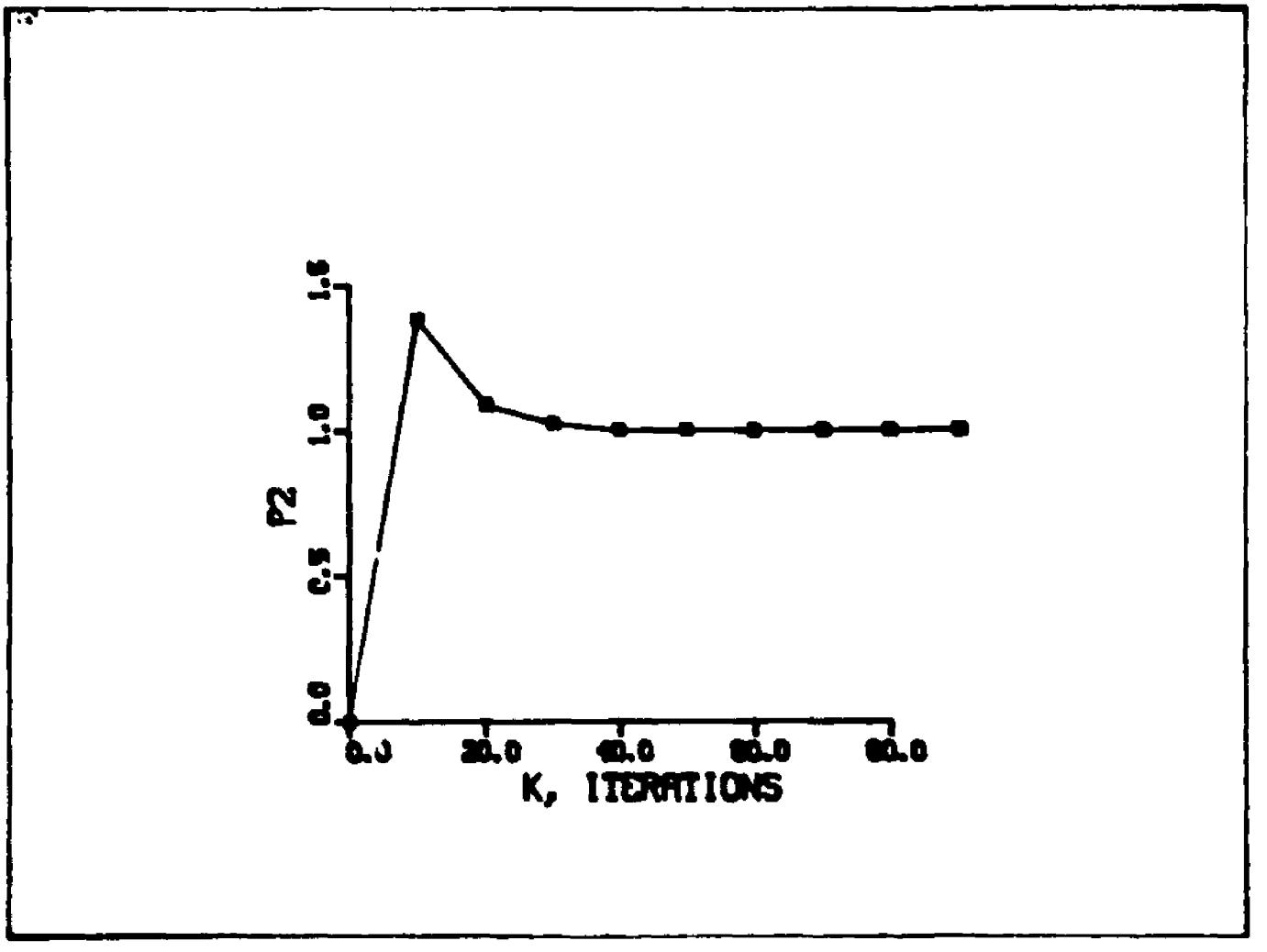

Fig. 4.16 Parameter P2 convergence using deterministic gradient algorithm in unobservable example.

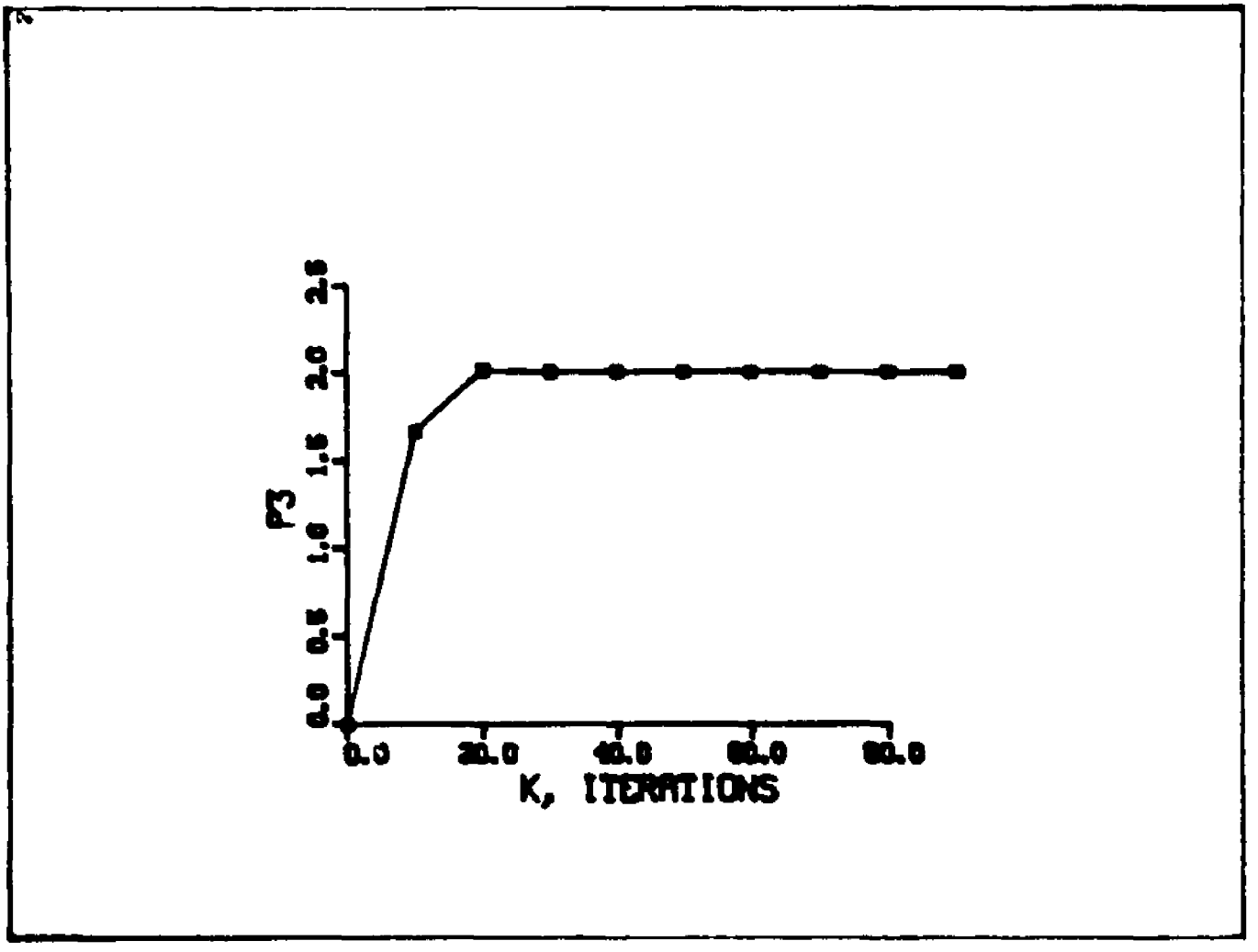

Fig. 4.17 Parameter P3 convergence using deterministic gradient algorithm in unobservable example.

The state description is simply

$$x(k+1) = .9x(k) + \begin{pmatrix} 1 & 2 \end{pmatrix} \underline{u}(k)$$

$$y(k) = x(k) = x_1(k)$$

where this single state is the first state of the system in eq. (4.125) after being transformed to observability canonical form.

The third example or stirred tank has the same eqs. (4.196)-(4.198) except

$$\underline{b} = \begin{pmatrix} .9512 & 4.877 & 4.877 \end{pmatrix}^T . \qquad (4.199)$$

Figs. (4.18)-(4.20) show the deterministic gradient convergence. The state model is

$$x(k+1) = .9512x(k) + \begin{pmatrix} 4.877 & 4.877 \end{pmatrix} \underline{u}(k)$$

$$y(k) = x(k) = x_1(k) . \qquad (4.200)$$

The three examples using the stochastic gradient algorithm in eq. (4.136) will now be identified with measurement noise on outputs. Inputs were used as in the deterministic gradient cases and without noise. The weights also were the same as before. All measurement noise used had a variance $\sigma^2 = .1$ and zero mean. An IMSL subroutine generated the Gaussian noise deviates.

The first example in eq. (4.61) now has the output equation

$$ym_1(k) = x_1(k) + v_1(k)$$

and additional associations to those in eq. (4.186) are made from eq. (4.15c)

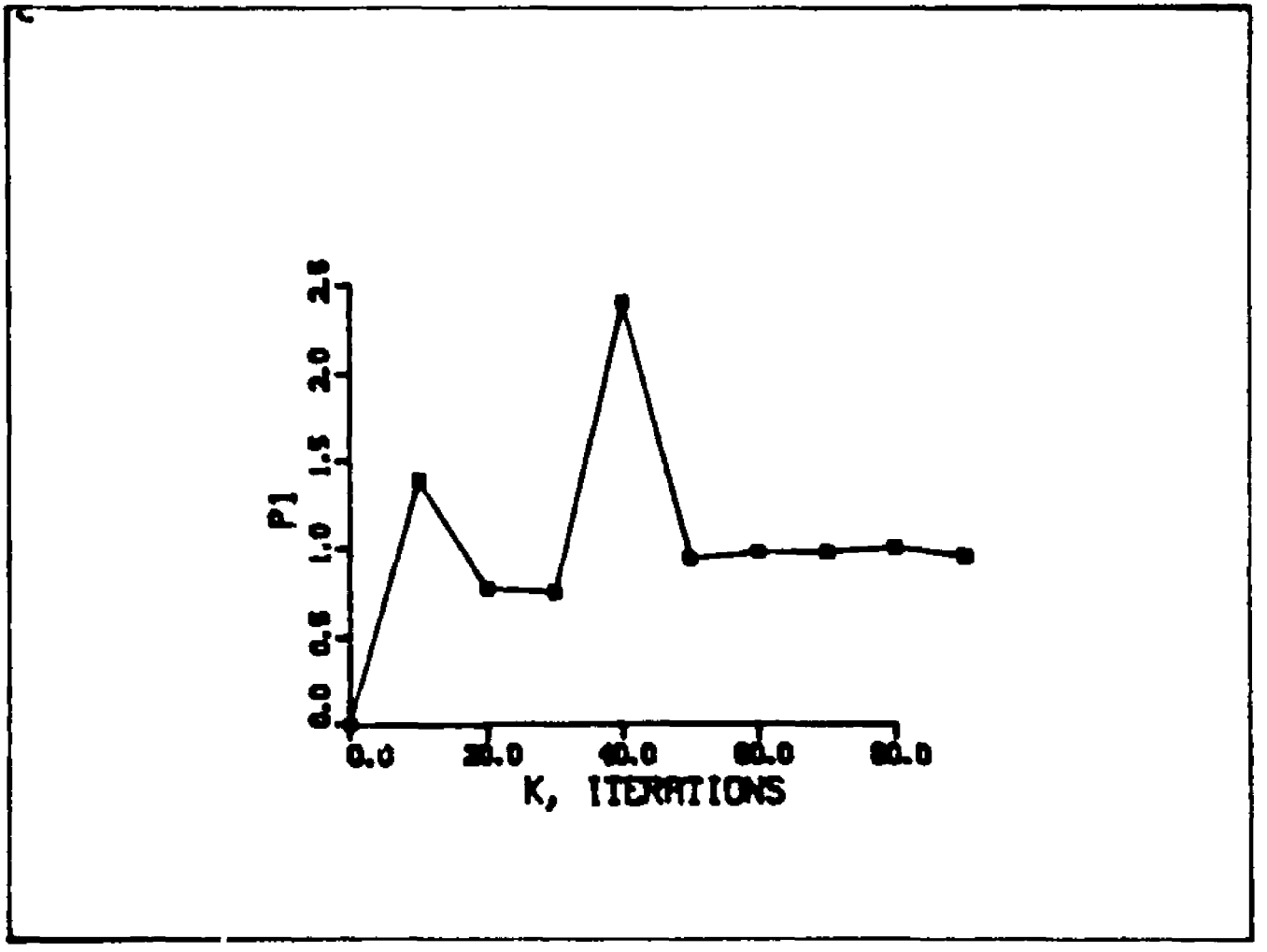

Fig. 4.16 Parameter P1 convergence using deterministic gradient algorithm in stirred tank example.

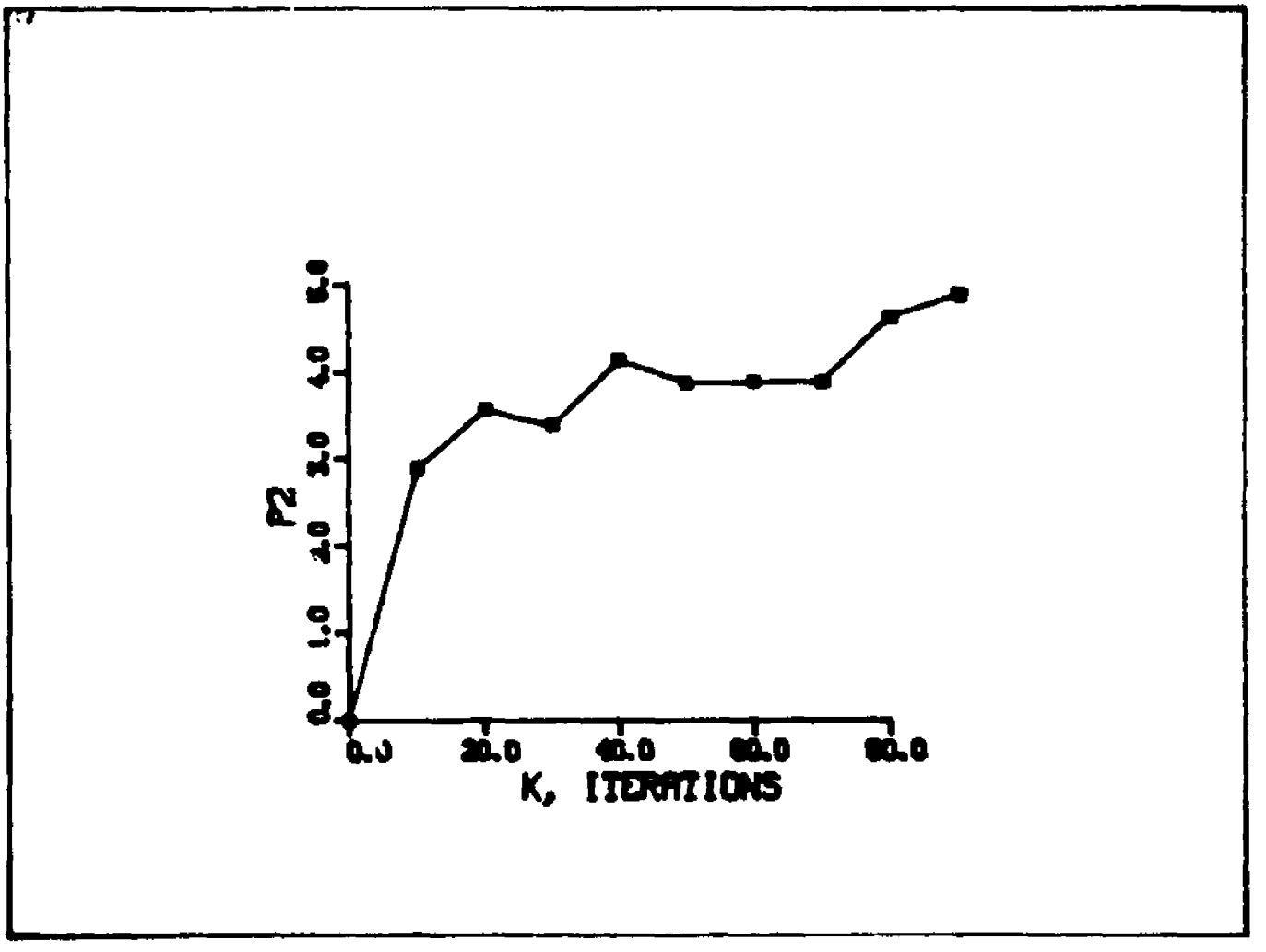

Fig. 4.19 Parameter P2 convergence using deterministic gradient algorithm in stirred tank example.

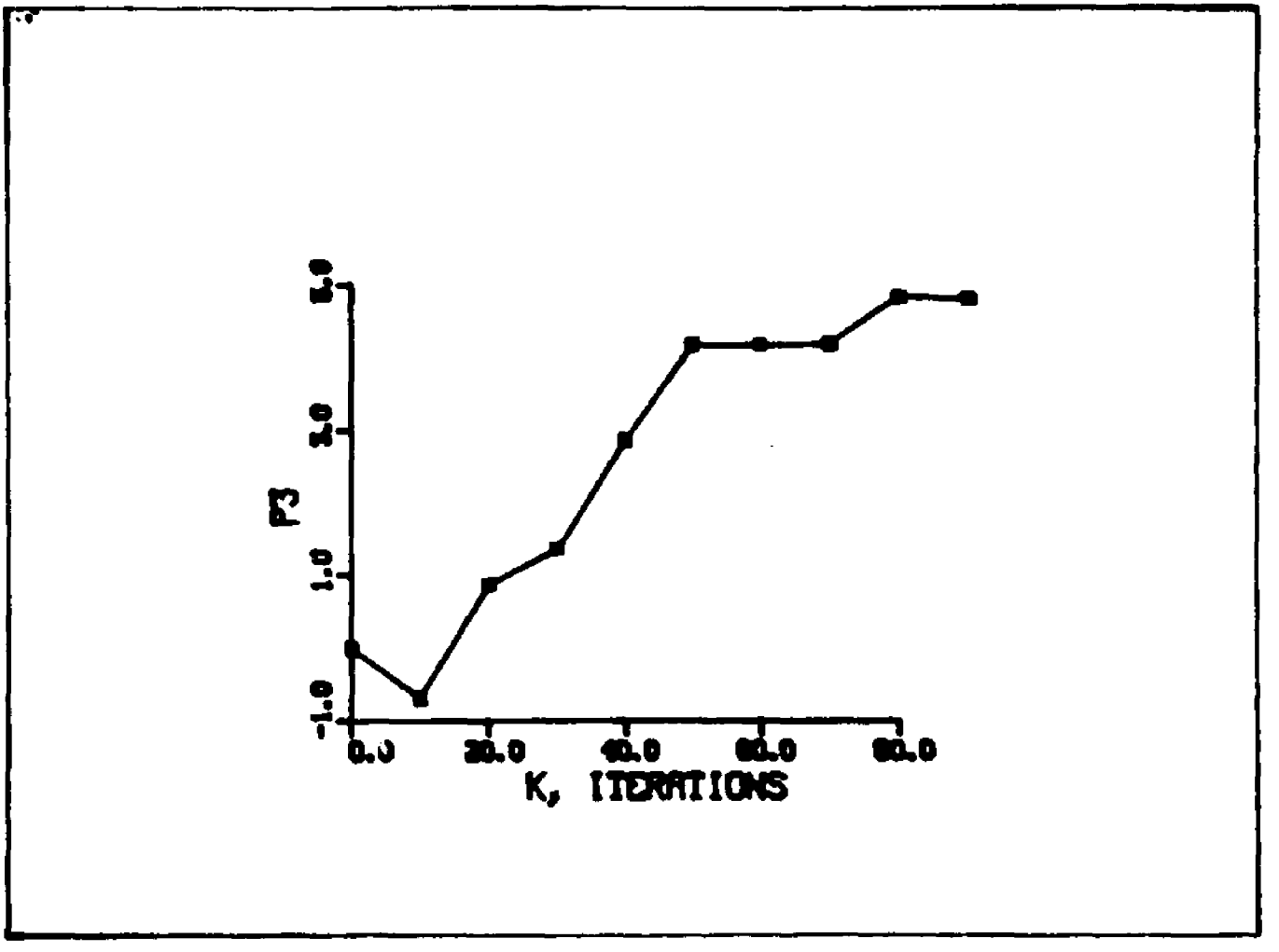

Fig. 4.20 Parameter P3 convergence using deterministic gradient algorithm in stirred tank example.

$$z_m(k) = y_{m_1}(k+2)$$

$$\underline{g}_m(k) = (y_{m_1}(k),\ y_{m_1}(k+1),\ u_1(k+1),\ u_2(k+1),\ u_1(k),\ u_2(k))$$

$$\underline{n}(k) = (v_1(k),\ v_1(k+1),\ 0,\ 0,\ 0,\ 0) \quad . \qquad (4.201)$$

The noise covariance matrix needed in the algorithm is given by

$$\Sigma_n = \mathrm{diag}(.1,\ .1,\ 0,\ 0,\ 0,\ 0) \quad . \qquad (4.202)$$

Figs. (4.21)-(4.23) show the convergences to the parameters in eq. (4.186). The spacing parameter used was s=2.

Both of the next two examples have the same additional associations which are

$$z_m(k) = y_{m_1}(k+1)$$

$$\underline{g}_m(k) = (y_{m_1}(k),\ u_1(k),\ u_2(k))^T$$

$$\underline{n}(k) = (v_1(k),\ 0,\ 0)^T \qquad (4.203)$$

and also

$$\Sigma_n = \mathrm{diag}(.1 \quad 0 \quad 0) \ .$$

Figs. (4.24)-(4.26) identify the parameters in eq. (4.197) and Figs. (4.27)-(4.29) identify the stirred tank parameters in eq. (4.199).

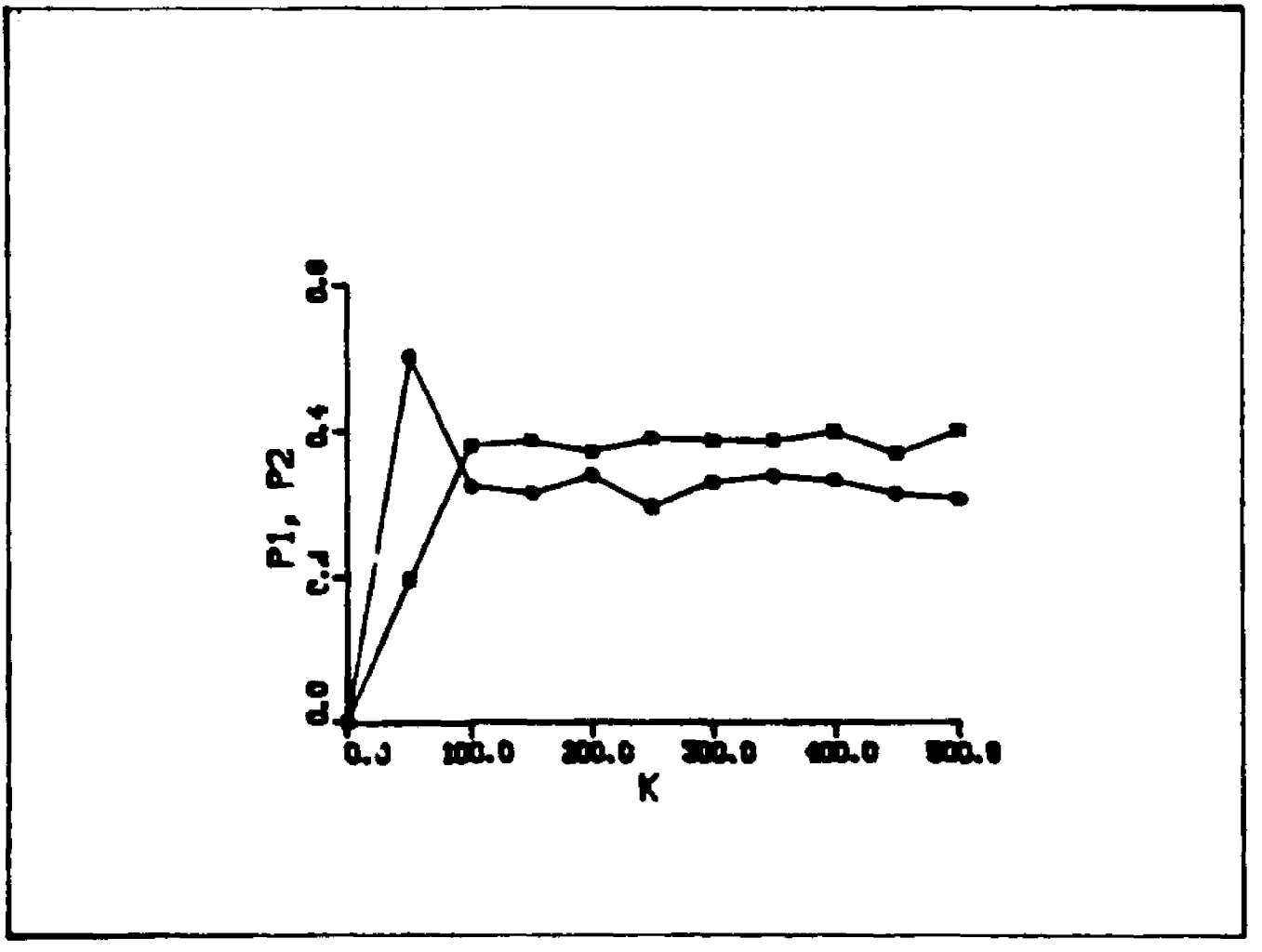

Fig. 4.21 Parameter P1 and P2 convergence using stochastic gradient algorithm in observable example.

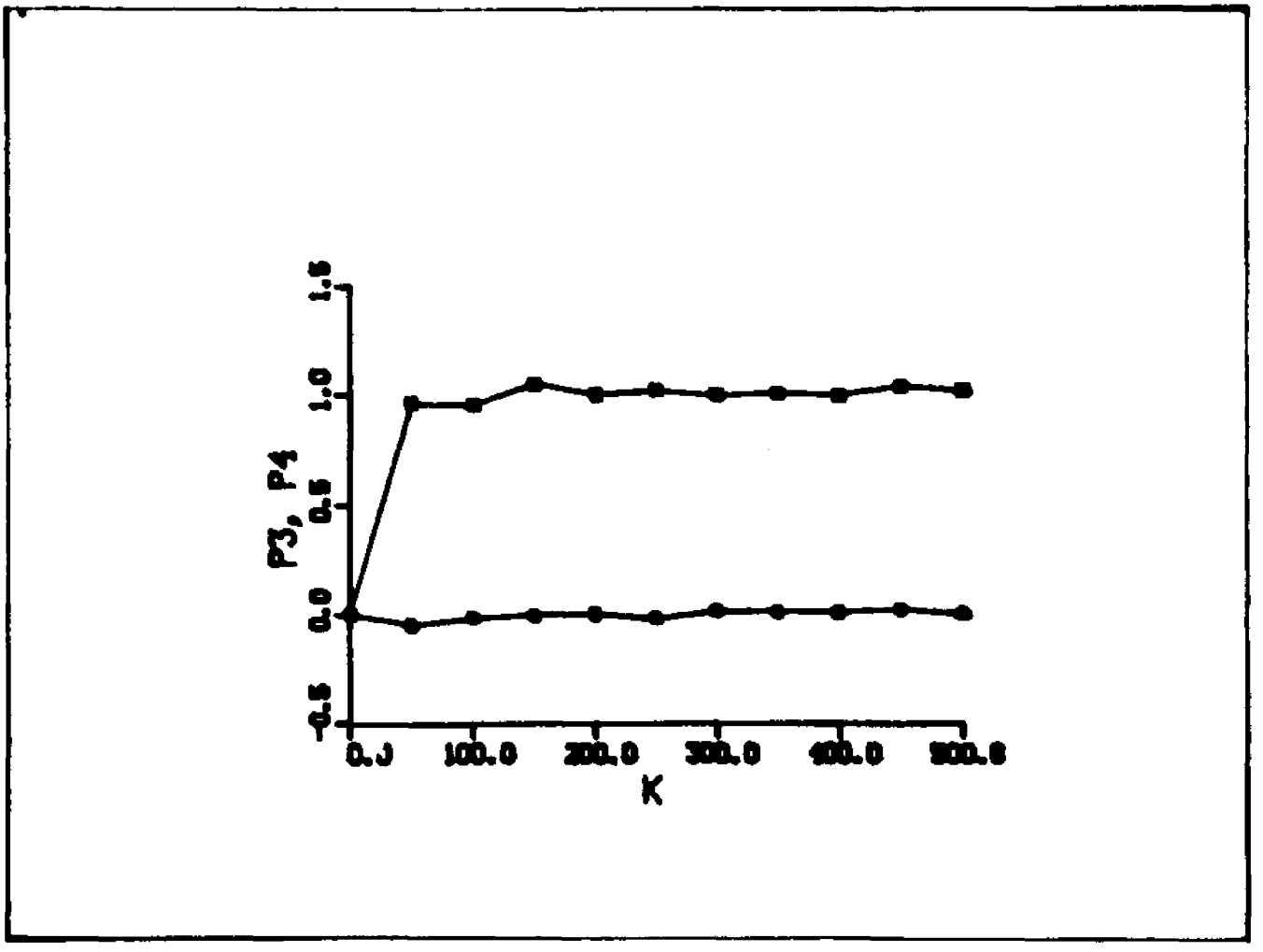

Fig. 4.22 Parameter P3 and P4 convergence using stochastic gradient algorithm in observable example.

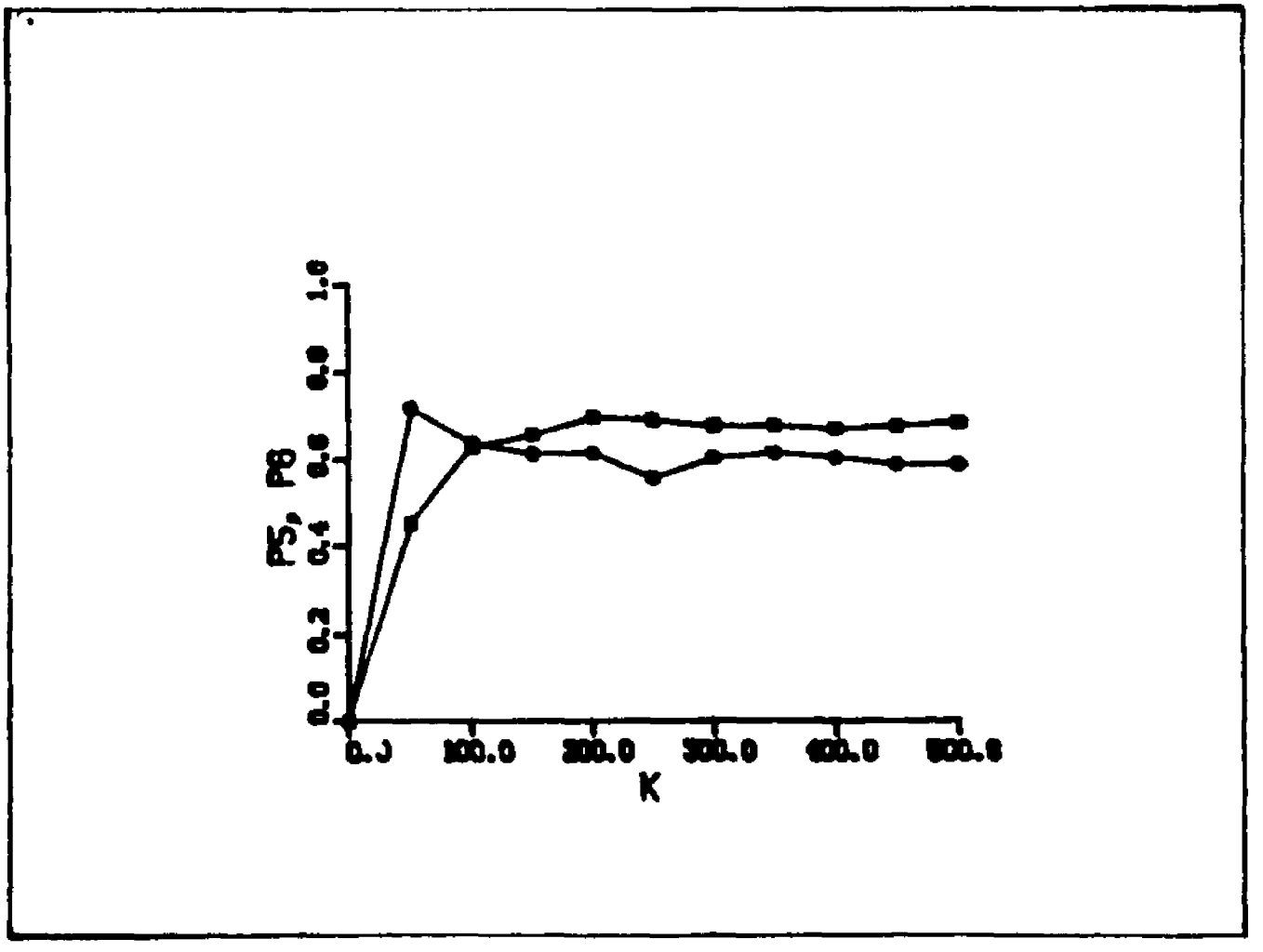

Fig. 4.23 Parameter P5 and P6 convergence using stochastic gradient algorithm in observable example.

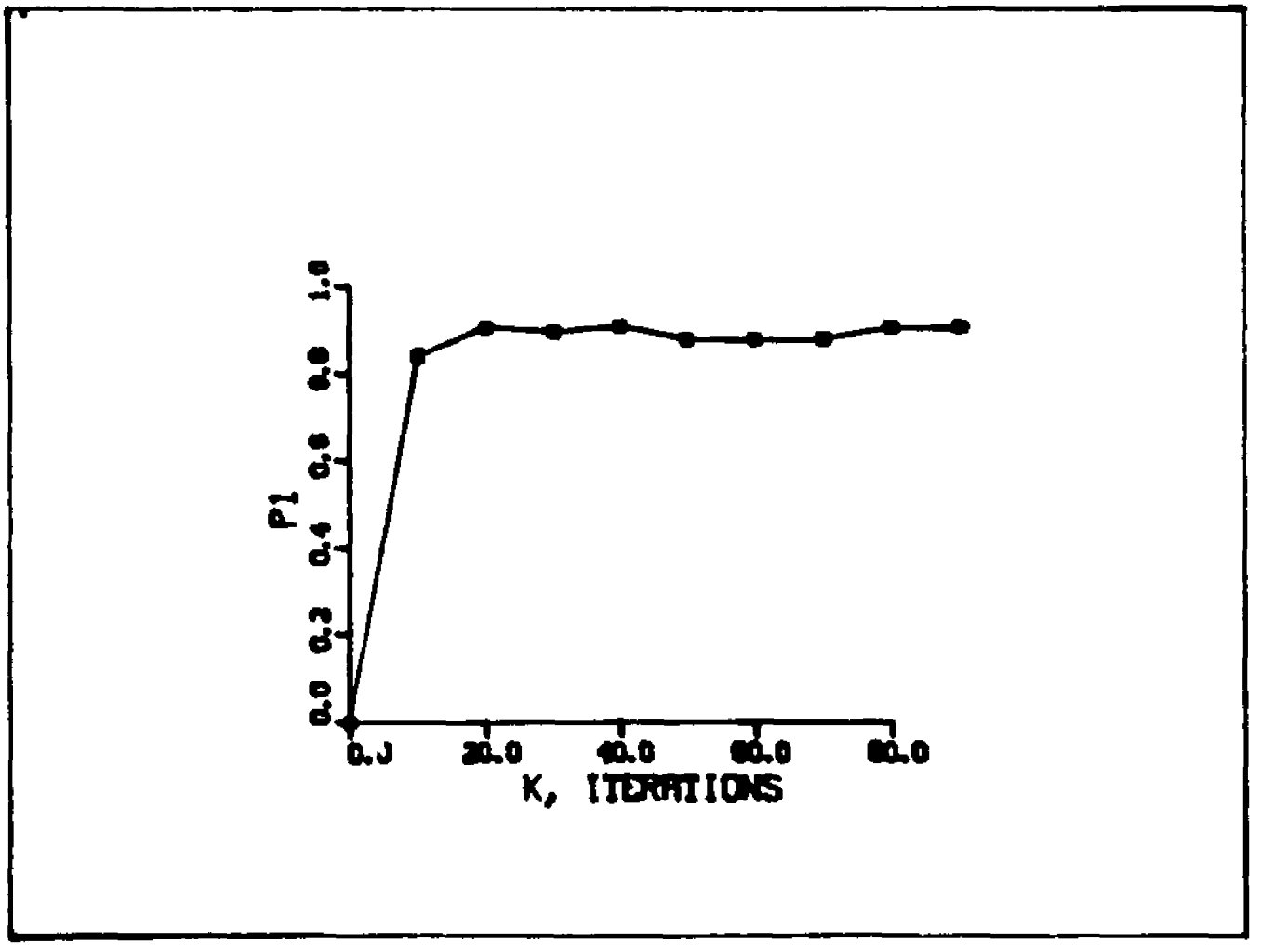

Fig. 4.24 Parameter P1 convergence using stochastic gradient algorithm in unobservable example.

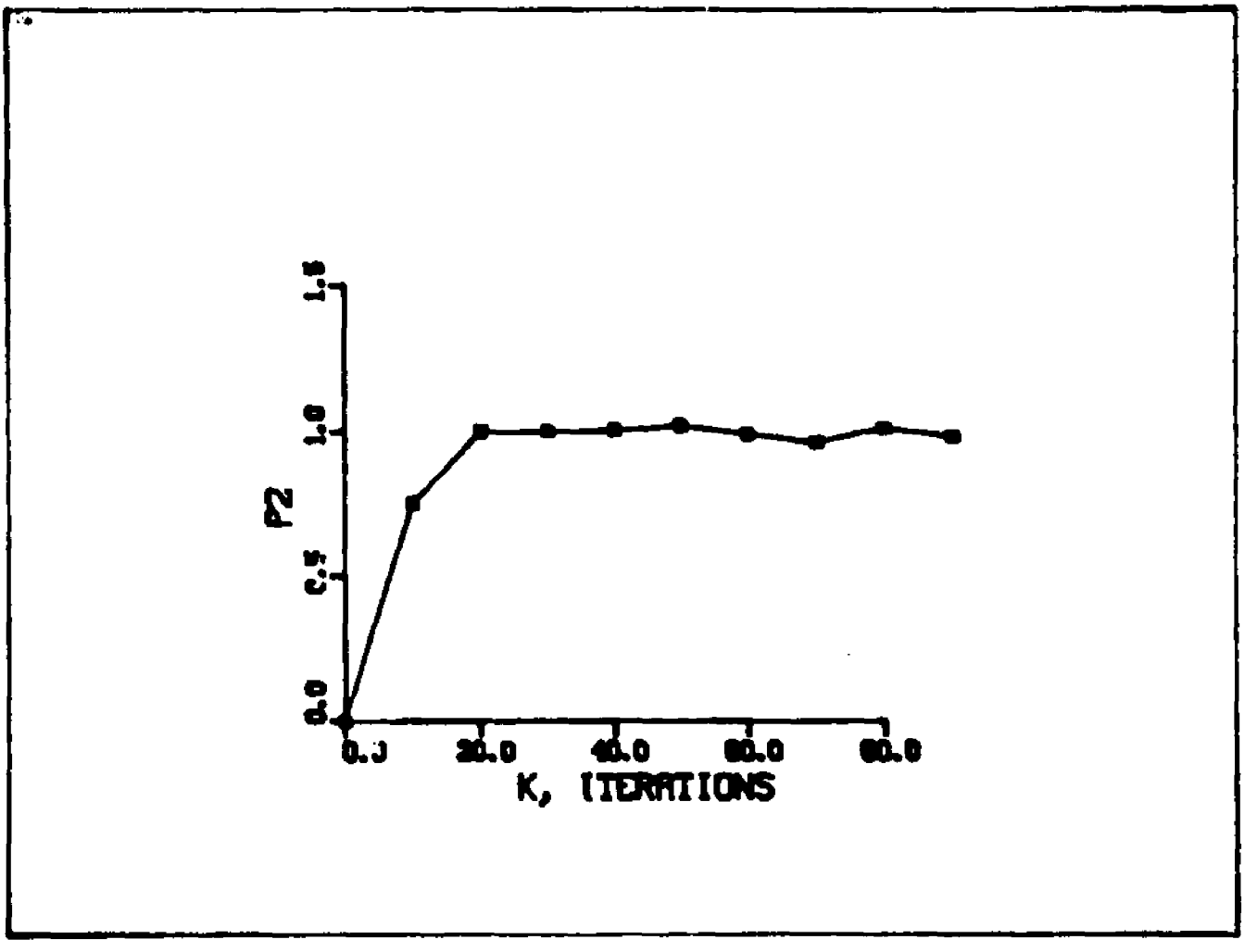

Fig. 4.25 Parameter P2 convergence using stochastic gradient algorithm in unobservable example.

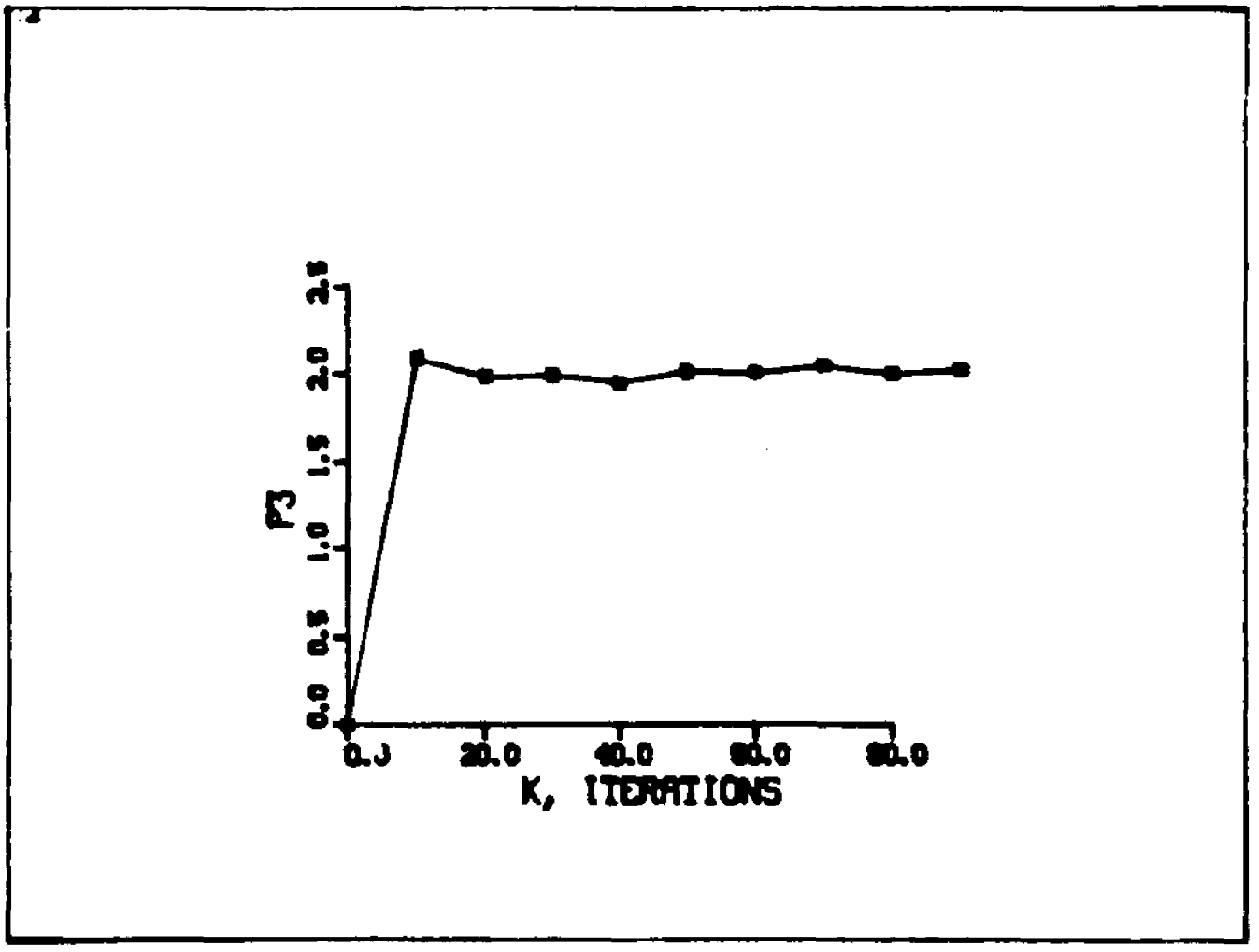

Fig. 4.26 Parameter P3 convergence using stochastic gradient algorithm in unobservable example.

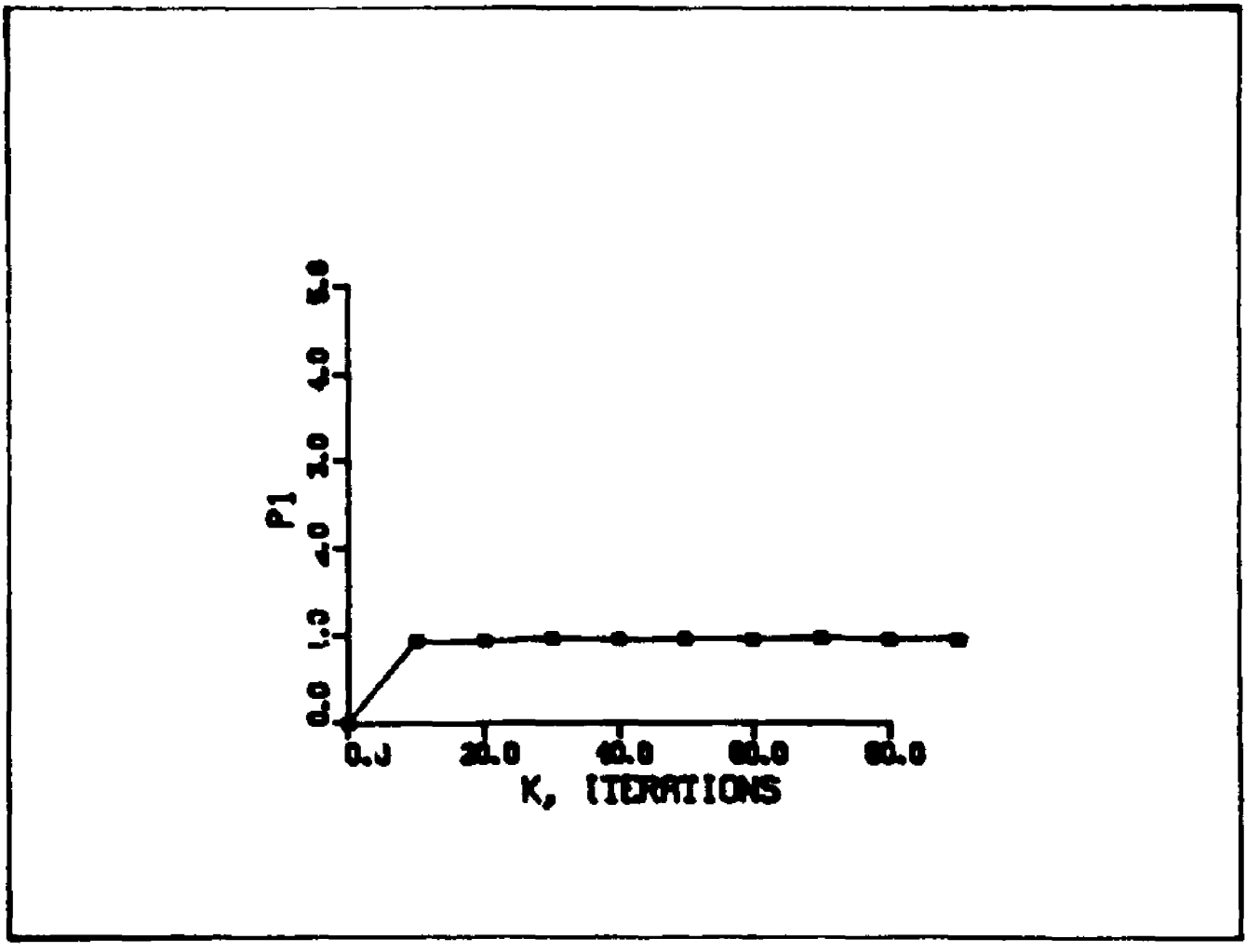

Fig. 4.27 Parameter P1 convergence using stochastic gradient algorithm in stirred tank example.

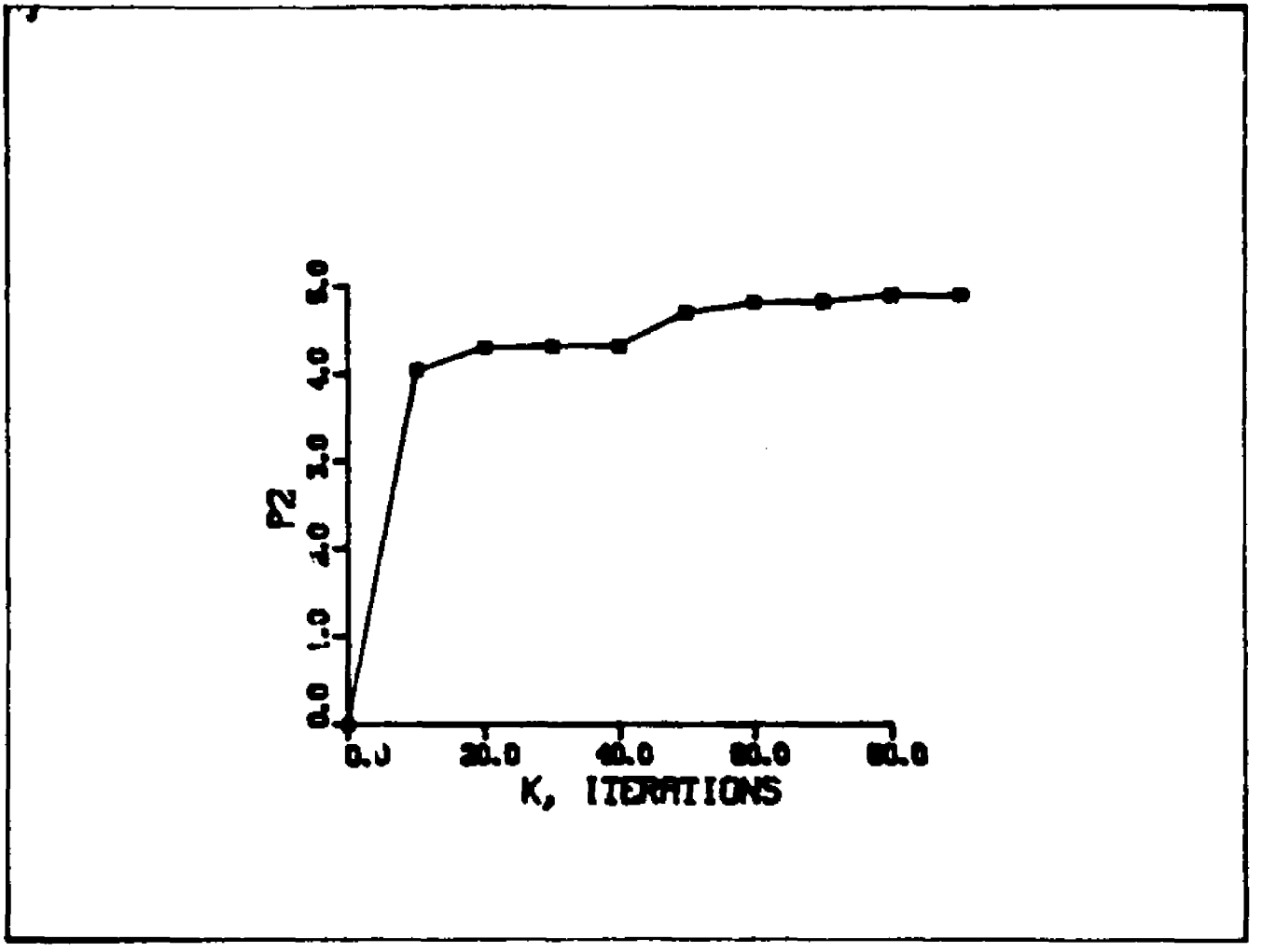

Fig. 4.28 Parameter P2 convergence using stochastic gradient algorithm in stirred tank example.

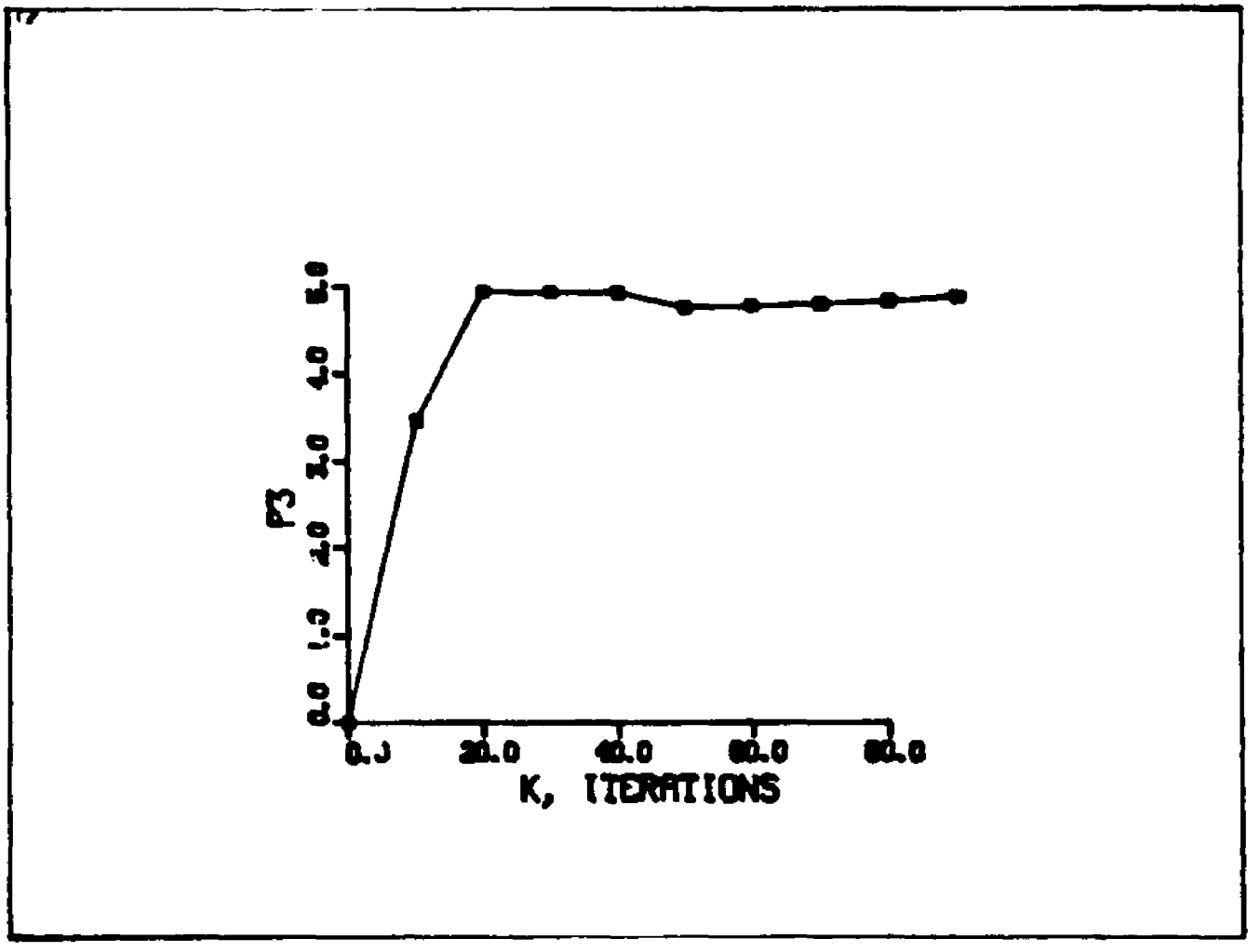

Fig. 4.29 Parameter P3 convergence using stochastic gradient algorithm in stirred tank example.

CHAPTER 5

APPLICATION OF DECENTRALIZED IDENTIFICATION TO A FLEXIBLE BEAM STRUCTURE

This chapter is devoted to the application of the results in decentralized identification theory to a high order system. The system - a flexible beam used in experiments at the NASA Langley research center - will be discussed and the state-space description derived. Whereas the examples identified in Chapter 4 were low order, this example is eighth order after some simplification from its original form. Once the state-space form is established, a set of decentralized measurements will be defined and the system identified with this reduced number. Motivation for this application will be discussed after a summary of the system description.

A simulation program for generating modal amplitudes that describe the dynamics of a flexible beam has been developed under a NASA grant at the City College of New York (see Thau and Montgomery). The modal amplitudes of the beam will later be shown to be directly related to a state-space description which is needed to directly apply the results of

Chapter 3 and 4. Estimation of the modal amplitudes is then equivalent to state estimation. Linear, non-linear, and repeated least squares algorithms have been applied to this estimation problem. Parameter identification using a version of the stochastic gradient algorithm in Chapter 4 in the presence of measurement noise was applied to the beam simulation.

Study of the flexible beam is motivated by the need to control large flexible space systems. The study of such systems has become important in recent years and promises to expand in the future. The structural design of a large space system aims toward low weight per surface area for efficiency. This results in low effective stiffness and very flexible structures. Precise knowledge of the system's structure dynamics is required when towing and assembling to accomplish adequate attitude and shape control. Attitude control involves maintaining the orientation of the spacecraft with respect to the sun or earth. Shape control involves maintaining the configuration of the large flexible space structure.

Structural knowledge at present cannot be determined analytically. Also, ground testing is not possible because of the structure size and because these structures are not designed to be self-supporting under gravity. An adaptive/learning approach used in the NASA simulation

assumes that structural testing is conducted during deployment or assembly. A problem is that control excitation levels usually required by adaptive systems are unacceptable. Therefore, the adaptation process is scheduled when it is convenient and when not then extrapolate and monitor the system performance.

So far estimation and parameter identification algorithms together with extrapolation of the system model are necessary. Input design must also be considered. For parameter identification proper design yields a large signal to noise environment. Speed of testing is important for parameter identification of large space structures because large amounts of time are needed due to the low frequencies of the vibrational modes of the structure.

The continuous time equations of the beam motion will now be discussed, followed by the corresponding discretized model. The following matrix differential equation

$$M\ddot{\underline{x}} + K\underline{x} = b\underline{u} \tag{5.1}$$

describes the large flexible space structure motion. The matrices M and K are the inertia and stiffness matrices, $\underline{x}$ is an n-vector of generalized coordinates, and $\underline{u}$ is an r-vector. Physically the inputs are the forces and moments applied to beam. The forces are applied through point-force actuators and the moments by torquers. Although there is small damping due to joint freeplay, material damping, etc.,

it is small enough to be neglected and does not appear in eq. (5.1).

Using the point transformation $\underline{x}=T\underline{s}$ where T is a modal matrix in eq. (5.1) yields

$$\ddot{\underline{s}} + \Lambda^2 \underline{s} = Tbu \tag{5.2}$$

where Λ^2 is a diagonal matrix of system natural frequencies. Thus, there are n decoupled second order differential equations. The first equation of (5.2) is

$$\ddot{s}_1 + \Lambda_1^2 s_1 = b_1^* u \tag{5.3}$$

where s_1 is the first element of $\underline{s}$, Λ_1 is the first frequency in Λ and b_1^* is the first row of Tb. Defining states

$$\begin{aligned} x_1 &= s_1 \\ x_2 &= \dot{s}_1 \end{aligned} \tag{5.3}$$

the state description of eq. (5.3) is

$$\dot{\underline{x}} = \begin{bmatrix} 0 & 1 \\ -\Lambda_1^2 & 0 \end{bmatrix} \underline{x} + \begin{bmatrix} 0 \\ b_1^* \end{bmatrix} u \tag{5.5}$$

which applies for one natural frequency or mode. All other modes can be similarly described and any number of modes can be concatenated.

The NASA simulation has been programmed to work with up to ten modes in discrete time. Table 5.1 lists the first six modal frequencies in radians

TABLE 5.1

MODAL RADIAN FREQUENCIES	ACTUATOR MODE SHAPES
11.4164	-5.9485
31.36	3.3258
61.2585	.8592
100.9	1.4178
150.185	3.3812
209.	4.9101

and the actuator mode shapes which correspond to b^* in eq. (5.5). One scalar input is present in the simulation. On a modal basis the discretization of eq. (5.5) is calculated in the simulation. As discussed in Chapter 4 for the discretization of the linearized stirred tank state-space description, eqs. (4.95) are used to find the discrete A, b, and C matrices.

The NASA simulation has only one scalar input so the discrete state-space model per mode is given by

$$\underline{x}(k+1) = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \underline{x}(k) + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} u(k) \qquad (5.6)$$

and

$$A_{11} = \cos \Omega_i \sqrt{1 - \Omega_i^2}\ T / \sqrt{1 - \Omega_i^2}$$

$$A_{12} = \sin \Omega_i \sqrt{1 - \Omega_i^2}\ T / \Omega_i \sqrt{1 - \Omega_i^2}$$

$$A_{21} = -\Omega_i \sin(\Omega_i \sqrt{1 - \Omega_i^2}\ T) / \sqrt{1 - \Omega_i^2}$$

$$A_{22} = \cos \Omega_i \sqrt{1 - \Omega_i^2}\ T$$

$$b_1 = ((1 - \cos \Omega_i T) / \Omega_i^2)\, b_i^*$$

$$b_2 = \sin \Omega_i T / \Omega_i\, b_i^* \qquad (5.7)$$

and in the program

$$\underline{y}(k) = H\underline{s}(k) \qquad (5.6)$$

in which $\underline{y}(k)$ is a 9X1 measurement vector, H is a 9X8 measurement matrix, and $\underline{s}(k)$ is an 8X1 vector with the first state of each mode in eq. (5.6).

In this thesis the results of decentralized identification are applied to the beam model. A four mode model is used. Four outputs are then available and the system is eighth order. An application of the results of this thesis to the beam is to decentralize the available outputs to two separate identifiers and thereby have a backup identifier in the event of a failure by one identifier. Fig. 5.1 shows a block diagram of the identification scheme.

In the NASA simulation all outputs are processed together in a centralized manner. This approach has the advantage of being very fast. A state-space description is identified that describes the input-output system behavior. With decentralized measurements two (or even more) state-space

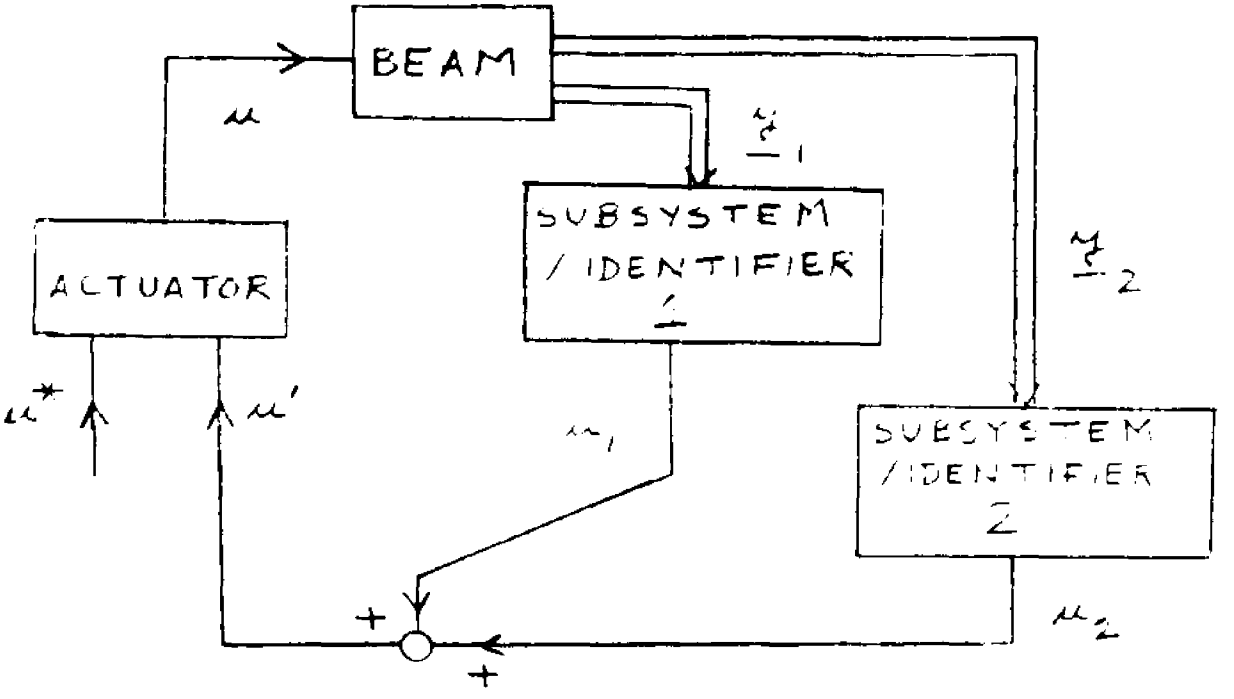

Fig. 5.1 block diagram of backup identification scheme for flexible space structure.

descriptions can be identified all describing the exact same input-output behavior (taken together) as the centralized case. Each identifier separately gives a full order system description since the system is observable from any output or any combination of outputs. Besides a backup for identifier failure the decentralized identifiers provide a backup for sensor failure.

Another point of view can be taken as to sensor economy. Since only an input-output description is desired, this can be achieved with fewer than the full 9 sensors currently

slated to be used. Once sensor and identifier reliability is proved from actual beam deployment, considerations of fewer than one output per mode can be undertaken in future designs. In every case of fewer outputs longer identification times result. With increasingly faster computers being introduced this becomes less and less of a problem and on-line identification schemes described in Chapter 4 are unimpaired.

The numerical state model and the decentralized version will now be calculated. Computer results will follow successfully identifying the beam using the decentralized outputs. For four modes the state model has the form

$$\underline{x}(k+1) = \begin{bmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & A_3 & \\ 0 & & & A_4 \end{bmatrix} \underline{x}(k) + b\underline{u}(k) \qquad (5.9)$$

where

$$A_1 = \begin{bmatrix} .8414042 & .04732768 \\ -6.170576 & .8414042 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} .002796269 & .03186763 \\ -31.35988 & .002796269 \end{bmatrix}$$

$$A_3 = \begin{bmatrix} -.9969073 & .001282867 \\ -4.814093 & -.9969073 \end{bmatrix}$$

$$A_4 = \begin{bmatrix} .32651218 & -.009367623 \\ 95.36998 & .32651218 \end{bmatrix} \qquad (5.10)$$

and

$$\underline{b}^T = (.00121641, .0473277, .00101359,$$
$$.0318876, .000532139, .00128287,$$
$$.661527X10^{-4}, -.0093676) . \qquad (5.11)$$

Next, the C matrix for 2 outputs is formed from the H matrix in eq. (5.6) using appropriately placed columns of zeros

$$C^T = \begin{bmatrix} -9.7111 & -5.9485 \\ 0 & 0 \\ 9.6354 & 3.3258 \\ 0 & 0 \\ 9.5304 & .8592 \\ 0 & 0 \\ -9.3913 & 1.4178 \\ 0 & 0 \end{bmatrix}$$

so that

$$\underline{y}(k) = C\underline{x}(k) \qquad (5.12)$$

where $\underline{y}(k)$ is a 2X1 vector. The dependence coefficients that will be identified directly are found using eq. (5.19)

$$b^{*} = B^{**} - A^{*}(M0)$$

$$= \begin{bmatrix} b_1(4) & b_1(3) & b_1(2) & b_1(1) & 0 \\ b_2(4) & b_2(3) & b_2(2) & b_2(1) & 0 \end{bmatrix}$$

$$- \begin{bmatrix} \underline{a}_1^T \\ \\ \underline{a}_2^T \end{bmatrix} \begin{bmatrix} 0 & & \\ b_1(1) & & 0 \\ b_1(2) & b_1(1) & \\ b_1(3) & b_1(2) & b_1(1) \\ 0 & & \\ b_2(1) & & 0 \\ b_2(2) & b_2(1) & \\ b_2(3) & b_2(2) & b_2(1) \end{bmatrix}$$

where the input matrix in eq. (5.9) is row partitioned as

$$\underline{b} = \begin{bmatrix} b_1(1) \\ b_1(2) \\ b_1(3) \\ b_1(4) \\ b_2(1) \\ b_2(2) \\ b_2(3) \\ b_2(4) \end{bmatrix} .$$

These equations are based on structural invariants

$$p_1 = p_2 = 4$$

which would emerge if equal numbers of columns of both outputs are used in the dependence test (see Chapter 3). The transformation matrix is

$$T = \begin{bmatrix} c_1 \\ c_1A \\ c_1A^2 \\ c_1A^3 \\ c_2 \\ c_2A \\ c_2A^2 \\ c_2A^3 \end{bmatrix} \qquad (5.15)$$

and with

$$x'(k) = Tx(k) \qquad (5.16)$$

eqs.(5.9) and (5.12) are transformed to

$$\underline{x}'(\kappa+1) = TAT^{-1}\underline{x}'(k) + T\underline{b}u(k)$$

$$\underline{y}(\kappa) = CT^{-1}\underline{x}(\kappa) \; . \qquad (5.17)$$

The resulting canonical system as well as B* are shown in Fig. (5.2).

```
CANONICAL A
 0.[illegible]E-06 [illegible] 01 0.1251-05 0.5961-07-[illegible]232E-05-[illegible]191E-05-0.352L-05-0.6E9E-[illegible]
 0.[illegible]E-05-[illegible]351E-06 0.1001 01 [illegible].120L-05 [illegible].190E-05-[illegible]244E-05 0.202E-05-0.849L-[illegible]
-[illegible]E-05-[illegible]611E-05-0.7751-06 0.1001 01-0.339E-06 0.197E-05 0.238E-05-0.241L-[illegible]
-0.1[illegible]E 01-[illegible].[illegible]E 01-0.2851 00-0.2021 01 [illegible].739E-05 0.739E 01-0.498E 01 0.734[illegible] 01
 0.9[illegible]L-07-0.[illegible]2L-06 0.279[illegible]-06-0.820F-07-0.771E-06 0.190E 01-0.137E-05 0.627L-[illegible]
-0.132[illegible]-06-[illegible].071E-07 0.2911-06 0.2[illegible]9[illegible]-06 0.767E-06-0.119E-05 0.100E 01-0.64EF-[illegible]
-0.[illegible]7E-06 0.[illegible]04E-06-0.156[illegible]-06 0.1711-06 0.[illegible]64E-06-[illegible].232E-05 0.213E-05 0.1[illegible]L [illegible]
-0.124F-05-[illegible].[illegible]0E 00 0.253[illegible] 00-0.300F 00-0.100E 01 [illegible].237E 01-0.274E 01 0.237E [illegible]

CANONICAL C
 [illegible].1[illegible]E 01 0.[illegible]77E-06 0.2391-06-[illegible].137F-05-[illegible].359F-06-0.[illegible]54F-06-0.286E-05 0.29EE-05
 0.279[illegible]-06-[illegible].14[illegible]-06 0.2961-0[illegible]-[illegible].287E-06 0.100E 01 0.[illegible]96[illegible]-07-0.107E-05 0.124E-05

B 15
 [illegible].240[illegible]7E-02-0.27[illegible]93L-01-0.[illegible]C30[illegible]-01-0.[illegible]194[illegible]-01
-[illegible].331[illegible]E-02-0.163[illegible]-01-0.[illegible]35[illegible]-01-0.[illegible]73EE-01

D* TRANSPOSE
 0.240[illegible]7E-02-0.331[illegible]-02
 0.[illegible]411E-02-0.[illegible]7571[illegible]-02
 0.[illegible]41EE-02-0.[illegible]756[illegible]-02
 0.240[illegible]E-02-0.331[illegible]-02
 [illegible].000[illegible] 00 0.000[illegible] 00
```

Fig. 5.2 Decentralized row companion canonical form for modal state-space description in eq. (5.9).

The unknown parameters to be identified are now extracted from Fig. 5.2. From the canonical A matrix

$$A^{*T} = \begin{bmatrix} -1. & .0 \\ -2.02 & -.3 \\ -.283 & .253 \\ -2.02 & -.3 \\ .0 & -1. \\ 7.34 & 2.37 \\ -4.98 & -2.74 \\ 7.34 & 2.37 \end{bmatrix} \tag{5.18}$$

b^* which appears in eqs. (5.13) and (B.19) is calculated in the program (neglecting the last two zeros to reduce the order of the measurement matrix to be inverted shortly) to be

$$B^{*T} = \begin{bmatrix} .00241 & -.00331 \\ .00124 & -.00776 \\ .00124 & -.00776 \\ .00241 & -.00331 \end{bmatrix} . \tag{5.19}$$

The vector DME is given by

$$y^*(k+1) = \left[A^* \mid B^* \right] \begin{bmatrix} \bar{y}(k) \\ \hline \bar{u}'(k) \end{bmatrix} \tag{5.20}$$

where

$$\begin{aligned} y^*(k+1) &= \left(y_1(k+4),\ y_2(k+4) \right)^T \\ \bar{y}(k) &= \left(y_1(k),\ \ldots,\ y_1(k+3),\ y_2(k),\ \ldots,\ y_2(k+3) \right)^T \\ \bar{u}'(k) &= \left(u(k),\ \ldots,\ u(k+4) \right)^T . \end{aligned} \tag{5.21}$$

A fully determined system is then formed by

$$\left[y^*(1), \ldots, y^*(12) \right] = \left[A^* \mid B^* \right] \begin{bmatrix} \bar{y}(0), & \ldots, & \bar{y}(11) \\ \bar{u}'(0), & \ldots, & \bar{u}'(11) \end{bmatrix} \tag{5.22}$$

or

$$Y^* = \left[A^* \mid B^* \right] G \tag{5.23}$$

then

$$\left[A^* \mid B^* \right] = Y^* G^{-1} \ . \tag{5.24}$$

The inputs and outputs used, U, Y1, and Y2, measurement matrix, G, and identified parameters, X, are shown in Fig. 5.3. Note that the elements of X in Fig. 5.3 are very close to those of A* and B* in eqs. (5.18) and (5.19).

Finally, the state description in Fig. 5.2 is obtained through using A* and B* in eq. (5.13) and using the procedure in Appendix B to reconstruct the B matrix.

Fig. 5.3 The least squares identification of the dependence coefficients contained in A* and B* in eq. (5.24).

CHAPTER 6

FUTURE RESEARCH

6.1 Combined Decentralized Identification and Control

In Chapters 2 through 5 the problem of identifying systems with a decentralized set of outputs was studied and decentralized measurement equations with and without noise derived. Important related topics of MIMO canonical forms and system identifiability, controllability, and observability were reviewed and/or discussed. The study of decentralized identification in this thesis was partially motivated by the relatively recent ongoing active research area of the study of decentralized control systems (see Sandell, etal. for a survey and lengthy bibliography of the subject; other papers relevant to the present discussion will be cited later). A more or less obvious possibility for further research arises from considering how this thesis relates to the decentralized control results. After briefly reviewing major results of decentralized control, the relation between these results and those of decentralized identification are discussed in the rest of this section.

Section 6.2 then discusses other areas and issues for further research.

Consider the linear time-invariant system

$$\underline{x}_i(k+1) = A_{ii}\underline{x}_i(k) + \sum_{\substack{j=1 \\ j\neq i}}^{L} A_{ij}\underline{x}_j(k) + \sum_{i=1}^{L} B_i u_i(k)$$

$$\underline{y}_i = C_i \underline{x}(k)$$

$i=1, \ldots, L$ (6.1)

where A_{ii} is the main diagonal block for the ith subsystem, A_{ij} are interconnection matrices from all other subsystems, and B_i are appropriately partitioned parts of the whole B matrix. All L subsystems together describe eq. (2.1). The ith subsystem receives the m_i vector $\underline{y}_i(k)$. The decentralized stabilization problem is to find L local output feedback control laws with dynamic compensation given by

$$\underline{u}_i(k) = H_i\underline{z}_i(k) + K_i\underline{y}_i(k) + L_i\underline{u}_i^*(k)$$

$$\underline{z}_i(k+1) = F_i\underline{z}_i(k) + S_i\underline{y}_i(k) + G_i\underline{u}_i^*(k) \quad (6.2)$$

so that acting together the resultant system has no eigenvalues outside the unit circle. The variable $\underline{z}_i(k)$ is the compensator state. This problem studied by Wang and Davison a, Corfmat and Morse a b, and others considers all matrices (A, B_i, C_i) of eq. (6.1) to be known.

When all subsystems of eq. (6.2) are augmented into the global compensator

$$\underline{u}(k) = H\underline{z}(k) + K\underline{y}(k) + L\underline{u}^*(k)$$

$$\underline{z}(k+1) = F\underline{z}(k) + S\underline{y}(k) + G\underline{u}^*(k) \quad (6.3)$$

the six coefficient matrices take on block diagonal form. With $\underline{u}_i$, an r_iX1 vector, and $\underline{y}_i$, an m_i X1 vector, the K matrix, for example, has the form

$$K = \begin{bmatrix} K_1 & & & 0 \\ & K_2 & & \\ & & \ddots & \\ & & & \\ 0 & & & K_L \end{bmatrix} \tag{6.4}$$

with $\sum_{i=1}^{L} m_i = m$ and $\sum_{i=1}^{L} r_i = r$ where $\underline{u}$ is the total rX1 input vector, $\underline{y}$ is the total mX1 measurement vector, and $\underline{u}^*$ is an external (non-feedback) input, and $\underline{z}(k)$ is the total compensator state. First, decentralized stabilization results will be summarized and then the rest of this section will apply or modify these results to state models obtained by identifying unknown systems using only decentralized measurements.

The augmented system from applying eq. (6.2) into eq. (6.1) is

$$\begin{bmatrix} \underline{x}(k+1) \\ \underline{z}(k+1) \end{bmatrix} = \begin{bmatrix} A+BKC & BH \\ SC & F \end{bmatrix} \begin{bmatrix} \underline{x}(k) \\ \underline{z}(k) \end{bmatrix} + \begin{bmatrix} BL \\ G \end{bmatrix} \underline{u}^*(k) \tag{6.5}$$

with

$$B = (B_1 \mid \ldots \mid B_s)$$

and

$$C = \begin{bmatrix} C_1 \\ \cdot \\ \cdot \\ \cdot \\ C_s \end{bmatrix} \quad . \tag{6.6}$$

From Wang and Davison [8] the dynamic compensator is found by defining the augmented triple

$$A_\eta = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} \begin{matrix} \}n \\ \}\eta \end{matrix} \qquad B_\eta = \begin{bmatrix} B & 0 \\ 0 & I \end{bmatrix} \begin{matrix} \}n \\ \}\eta \end{matrix}$$

$$C_\eta = \begin{bmatrix} C & 0 \\ 0 & I \end{bmatrix} \begin{matrix} \}m \\ \}\eta \end{matrix}$$

and

$$A_\eta = \begin{bmatrix} K_1 & & & H_1 & & \\ & \ddots & 0 & & \ddots & 0 \\ 0 & & & 0 & & \\ & & K_s & & & H_s \\ S_1 & & & F_1 & & \\ & \ddots & 0 & & \ddots & 0 \\ 0 & & & 0 & & \\ & & S_s & & & F_s \end{bmatrix} \tag{6.7}$$

with $\eta = \sum_{i=1}^{s} \eta_i$.

The major result of that paper, a generalization of Theorem 1 quoted below, is based on the following two definitions.

Def. 1. For a linear time-invariant system with matrices (C,A,B) and block diagonal K as in eq. (6.4) the set of fixed modes with respect to (C, A, b, K) is the intersection of the sets of eigenvalues of A+BKC for all possible K or

$$\Lambda(C, A, b, K) = \bigcap_{K \in \mathcal{K}} \lambda(A + BKC) \quad . \tag{6.8}$$

This includes the null matrix for K so that Λ is a subset of $\lambda(A)$. It is also to be noted that Λ is a subset of that $\lambda(A)$ that is not a function of K.

Def. 2. The fixed polynomial of (C, A, b, K) is the greatest common divisior of all possible polynomials

$$|\lambda I - A - BKC| = 0 \quad \text{for } K \in \mathcal{K} . \tag{6.9}$$

The fixed polynomial is therefore independent of K and with reference to Def. 1 the fixed modes are the roots of the fixed polynomial.

Thm 1. To stabilize eq. (6.1) with

$$u_i = K_i y_i(k) + L_i u_i^*(k)$$

requires that the fixed modes of (A, B, C, K) be inside the unit circle. This clearly is because the fixed modes are independent of K and no K can change them. If they are unstable modes the system remains unstable with the feedback.

A result due to Brasch and Pearson and applied in Wang and Davison [6] using the Kalman canonical structure theorem states that the fixed polynomial of

$$|\lambda I - A_\eta - B_\eta K_\eta C_\eta|$$

is the same as the fixed polynomial of

$$|\lambda I - A - BKC| .$$

The main result of Wang and Davison can then be stated in Thm. 2.

Thm. 2. For the system in eq. (6.5) and for the K in eq. (6.4) a necessary and sufficient condition for asymptotic stability is that

$$\Lambda(A, B, C, K) \subset \odot \qquad (6.10)$$

i.e., the fixed modes of (A, B, C, K) are inside the unit circle.

Wang and Davison a have shown that the fixed modes are the transmission zeros of the system in eq. (6.1) and K in eq. (6.4). With

$$B = (b_1, b_2, \ldots, b_n)^T$$

and

$$C = (c_1, c_2, \ldots, c_m)^T$$

and $\det(M_0)=0$ a lemma by Anderson and Wong states

$$\det\left(M_0 + \sum_{i=1}^{s} \mu_i D_i C_i^T \right) = 0 \quad \forall \mu_i \in R \qquad i = 1, \ldots, s \qquad (6.11)$$

where μ_i are real scalars if the following holds

$$\det \begin{bmatrix} M_0 & b_{i_1} & b_{i_2} & \cdots & b_{i_t} \\ c_{i_1}^T & 0 & \cdots & & 0 \\ c_{i_2}^T & & & & \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & 0 & \\ c_{i_t}^T & & & & \end{bmatrix} = 0$$

$$i_1, \ldots, i_t \in 1, 2, \ldots, L \qquad (6.12)$$

for all disjoint values of $i_1, \ldots, i_t$;t=1, ..., L.

Applied to eq. (6.1)

$$\det\left(A - \lambda I + \sum_{k=1}^{L} \mu_k b_{i_k} c_{j_k}^T \right) = 0$$

$$\forall \mu_k \in R;\ k = 1, \ldots, L \qquad (6.13)$$

if λ is a transmission zero of all subsystems

$$\begin{bmatrix} c_{j_{k_1}}^T \\ \cdot \\ \cdot \\ \cdot \\ c_{j_{k_2}}^T \end{bmatrix}, A, \left(b_{i_{k_1}}, \ldots, b_{i_{k_t}} \right)$$

$k_1 = 1, \ldots, L+1-t$

$k_2 = k_1+1, k_1+2, \ldots, L+2-t$

$\cdot$

$\cdot$

$\cdot$

$k_t = k_{t-1}+1, k_{t-2}+2, \ldots, L$

for $t = 1, 2, \ldots, \min(r, m)$. (6.14)

the matrix $A - \lambda I$ equates with M_o and the $c_{(\ )}^T$'s and $b_{(\ )}$'s in eq. (6.14) are the same as those in the augmented determinant above.

With decentralized pole placement results established, these results will now be related to an arbitrary subsystem that has used decentralized identification results to identify the global system.

The decentralized identification equation has been derived for a scalar y_i in section 3.2. The general nth order equation is (without noise)

$$y_i(k+n) = P_N y_i(k+n-1) + \ldots + P_1 y_i(k) + \underline{P_{BN}}^T \underline{u}(k+n-1) + \ldots + \underline{P_{B1}}^T \underline{u}(k) \quad . \tag{6.15}$$

Decentralized control theory will now be applied for

$$A = \begin{bmatrix} .1 & 0 \\ -.2 & 1.1 \end{bmatrix} \qquad B = \begin{bmatrix} -1 & 3 \\ 5 & 6 \end{bmatrix}$$

$$\underline{c}_1 = \begin{pmatrix} 1 & 0 \end{pmatrix}$$

$$\underline{c}_2 = \begin{pmatrix} 0 & 1 \end{pmatrix} \quad . \tag{6.16}$$

Dependence coefficients from Table 3.1 are computed

$$u_1(k): PB1(1) = b_{21}a_{12} - b_{11}a_{22}$$
$$= 0 - (-1)(1.1) = 1.1$$

$$u_2(k): PB2(2) = b_{22}a_{12} - b_{12}a_{22}$$
$$= 0 - 3(1.1) = -3.3$$

$$u_1(k+1): PB2(1) = b_{11} = -1$$

$$u_2(k+1): PB2(2) = b_{12} = 3$$

and $\lambda I - A = (\lambda - .1)(\lambda - 1.1) = \lambda^2 - 1.2 + .11$

$$y(k+1) : P2 = 1.2$$

$$y(k) : P1 = -.11 \qquad (6.17)$$

Eq. (6.15) for the model in eq. (6.16) is

$$y_1(k+2) = 1.2y_1(k+1) - .11y_1(k) - u_1(k+1)$$
$$+ 1.1u_1(k) + 3u_2(k+1) - 3.3u_2(k) \quad . \qquad (6.18)$$

All coefficients in eq. (6.18) are identified by any applicable identification algorithm. This second order ARMA equation can be put into state-space form by defining states as follows let

$$x_2^d(k+1) = y_1(k+2) - 3u_2(k+1) - 1.2y_1(k+1)$$
$$+ u_1(k+1) = -.11y_1(k) + 1.1u_1(k) - 3.3u_2(k) \qquad (6.19)$$

then

$$x_2^d(k) = y_1(k+1) - 3u_2(k) - 1.2y_1(k) + u_1(k) \qquad (6.20)$$

and let

$$x_1^d(k+1) = y_1(k+1)$$

then

$$x_1^d(k) = y_1(k) \quad .$$

From eq. (6.20)

$$x_2^d(k) = x_1^d(k+1) - 3u_2(k) - 1.2x_1^d(k) + u_1(k)$$

and

$$x_1^d(k+1) = 1.2x_1^d(k) + x_2^d(k) - u_1(k) + 3u_2(k) \quad . \tag{6.21}$$

From eq. (6.19)

$$x_2^d(k+1) = -.11x_1^d(k) + 1.1u_1(k) - 3.3u_2(k)$$

then

$$\underline{x}^d(k+1) = \begin{bmatrix} 1.2 & 1 \\ -.11 & 0 \end{bmatrix} \underline{x}^d(k) + \begin{bmatrix} -1 & 3 \\ 1.1 & -3.3 \end{bmatrix} \underline{u}(k) \quad . \tag{6.23}$$

The first observation to make is that

$$\lambda(A_d) = .1,\ 1.1 = \lambda(A) \quad . \tag{6.24}$$

Suppose the single available measurement is fed back so that

$$u_1(k) = K_1 y_1(k) = K_1 x_1^d(k) = K_1 x_1(k) \quad . \tag{6.25}$$

Then eq. (6.23) becomes

$$\underline{x}^d(k+1) = \begin{bmatrix} 1.2-K_1 & 1 \\ -.11+1.1K_1 & 0 \end{bmatrix} \underline{x}^d(k) + \begin{bmatrix} 3 \\ -3.3 \end{bmatrix} u_2(k) \tag{6.26}$$

where

$$A+BKC = \begin{bmatrix} 1.2 & 1 \\ -.11 & 0 \end{bmatrix} + \begin{bmatrix} -1 & 3 \\ 1.1 & -3.3 \end{bmatrix} \begin{bmatrix} K_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{6.27}$$

Notice that K in eq. (6.27) is not block diagonal. As mentioned at the end of the 1973 paper by Wang and Davison a the K matrix can take on other structures and the results of that paper still hold. If K is a full matrix of adjustable elements the fixed modes are the uncontrollable and unobservable poles of the system. Some authors have already studied other structures for K, for example, Aoki uses

blocks off the main diagonal to model communication between subsystems.

Calculating the new characteristic equation

$$|\lambda I - A - bKC| = \begin{vmatrix} \lambda - 1.2 + K_1 & -1 \\ .11 - 1.1K_1 & \lambda \end{vmatrix}$$

$$= \lambda^2 + (K_1 - 1.2) + .11 - 1.1K_1$$

$$= (\lambda - 1.1)(\lambda + K_1 - .1) \qquad (6.28)$$

the first factor is recognized as a fixed polynomial, independent of any possible K, and 1.1 is a fixed mode. Since it is outside the unit circle this system cannot be stabilized with respect to this K in accordance with Thm. 1.

Another approach in this simple case is a root locus analysis. Applying

$$u_1(k) = K_1 y_1(k)$$

to eq. (6.18) gives

$$y_1(k+2) = 1.2y_1(k+1) - .11y_1(k) - K_1 y_1(k+1) + 1.1K_1 y_1(k) + 3u_2(k+1) - 3.3u_2(k) \qquad (6.29)$$

with characteristic equation

$$z^2 + Kz - 1.2z + .11 - 1.1K_1 = 0 \quad . \qquad (6.30)$$

Obtaining root locus form

$$1 + \frac{K_1(z - 1.1)}{z^2 - 1.2z + .11} = 1 + \frac{K_1(z - 1.1)}{(z - 1.1)(z - .1)}$$

$$= 1 + \frac{K_1}{z - .1} = 0 \quad . \qquad (6.31)$$

reveals that 1.1 is a transmission zero. Since previously this mode was shown to be a fixed mode, in accordance with the results of the 1985 paper by Wang and Davison [8], the two are identical.

A second possibility for K arises by feeding back x_1^d and x_2^d in eqs. (6.20). All signals are measurable so the decentralized states are computable. In order to obtain the decentralized identification, knowledge of the value of $u_2(k)$ was necessary. This doesn't mean $u_2(k)$ is available to system 1 but if it is $x_2^d(k)$ is available and

$$u_1(k) = K_1 x_1^d(k) + K_2 x_2^d(k) \tag{6.32}$$

is available for feedback. If u_2 is not available then K takes the form shown in eq. (6.27) with the new K eq. (6.27) is

$$A^d + B^d K C^d = \begin{bmatrix} 1.2 & 1 \\ -.11 & 0 \end{bmatrix} + \begin{bmatrix} -1 & 3 \\ 1.1 & -3.3 \end{bmatrix} \begin{bmatrix} K_1 & K_2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1.2 & 1 \\ -.11 & 0 \end{bmatrix} + \begin{bmatrix} K_1+1 & -K_2 \\ 1.1K_1 & 1.1K_2 \end{bmatrix}$$

$$= \begin{bmatrix} 2.2+K_1 & 1-K_2 \\ 1.1K_1-.11 & 1.1K_2 \end{bmatrix} . \tag{6.33}$$

The characteristic equation is

$$|\lambda I - A^{d} - B^{d} K C^{d}| = \begin{vmatrix} \lambda - 2.2 - K_1 & K_2 - 1 \\ .11 - 1.1K_1 & \lambda - 1.1K_2 \end{vmatrix}$$

$$= \lambda^2 + (-2.2 - K_1 - 1.1K_2)\lambda + 2.42K_2 + 1.1K_1K_2 - K_2.11 + 1.1K_1K_2 + .11 - 1.1K_1$$

$$= \lambda^2 + (-2.2 - K_1 - 1.1K_2)\lambda - 1.1K_1 + 2.31K_2 + 2.2K_1K_2 + .11 \qquad (6.34)$$

or desired CE can be equated to eq. (6.34) resulting in two (one nonlinear) equations in two unknowns. For example, suppose two stable eigenvalues at .5 are desired with desired CE

$$(\lambda - .5)^2 = \lambda^2 - \lambda + .25 \quad . \qquad (6.35)$$

Two solutions are found

$$K_1 = -2.030089, \quad K_2 = .754626$$

and

$$K_1 = -.60627, \quad K_2 = -.539751 \quad . \qquad (6.36)$$

This clearly follows from Thm. 1 since there are no fixed modes that satisfy eq. (6.30). This can be seen by solving for the roots of eq. (6.34) or using the fixed mode finding algorithm given in Jamshidi.

Continuing the line of reasoning of raising the number of gains in K suppose a full K is attempted, in other words, can centralized state feedback be used on the decentralized model since both "d" states are available? The elements K_{21} and K_{22} are present in K if

$$u_2(k) = K_{21} x_1^{d}(k) + K_{22} x_2^{d}(k) \quad . \qquad (6.37)$$

So far measurement of $u_2(k)$ was necessary for the identification and subsystem 1 must supply $u_2(k)$ to generate $x_2^{\alpha}(k)$. But $u_2(k)$ is controlled and determined by subsystem 2 and under the decentralized model structure operates independent of any actions by subsystem 1. Since not all blocks in K will be present this means subsystem 1 cannot use centralized feedback based on the decentralized identification model.

6.2 Additional Future Research

The areas of future research can be summarized into two parts. First is the identification of the decentralized identification problem. The input/output description of the most general state model in eq. (3.12) has unknown coefficients multiplying unmeasurable components of state noise. This makes the problem nonlinear. Second, the solutions to the decentralized problems of pole placement, stabilization, optimal control, the servomechanism problem, and others have all been published with complete knowledge of the system presupposed. Now let the system be unknown and apply the results of decentralized identification. Do these results still apply or are modifications necessary? In section 6.1 some of these issues were discussed for the pole placement problem. Lastly, the relation of decentralized identification to state estimation needs further research.

In section 4.2 the stochastic gradient algorithm of Saridis and Stein solved the identification of eq. (4.174) wnen the T matrix and noise variances are known. With T unknown but equal state and measurement noise an approach given in section 4.1 identified the model using a recursive least squares variation from Panuska. Using Panuska's predictor approach in section 4.1 and a linear estimator (such as the stochastic gradient algorithm) Irwin and Roberts showed how to identify the model if it can be represented in row companion form. With unequal state and measurement noise and unknown T matrix nonlinear terms appear in the input/output equation making previously discussed approaches inadequate. Further research in the literature or new research is needed for the solution of the most general case.

In the 1973 paper by Wang and Davison b all inputs from all subsystems are brought together to be applied to the global system and solve the decentralized pole placement and stabilizability problems. This approach can directly apply to a decentralized system where the global system is unknown but observable from measurements by each subsystem and subsequently identifiable with an appropriate algorithm and if enough excitation on uncontrollable modes exists to identify parameters before their responses die out (or, even better, the global system is controllable). In general some subsystems may have measurements such that the global system

cannot be identified. Some subsystems may be able to identify the global system, some may not, or the global system may even be unidentifiable from any one subsystem by itself. It is not clear how the Wang and Davison problem is solved if subsystems do not know the whole global system and do not know what is the correct feedback to solve the decentralized pole placement problem globally. Further examination of this predicament is suggested. This also applies to any other decentralized control problem area in which the system model is not identifiable from some or all subsystems.

State estimation with parameter identification is another immediate area of further research. Irwin and Roberts show that this is a nonlinear problem in the centralized case and deterministically, or with equal state and measurement noise, can be reduced to two linear problems. Specifically, for the model

$$\underline{x}(k+1) = A\underline{x}(k) + B\underline{u}(k) + I\underline{e}(k)$$

$$\underline{y}(k) = C\underline{x}(k) + \underline{e}(k)$$

apply a Kalman filter

$$\hat{\underline{\theta}}_i^a(k+1) = \hat{\underline{\theta}}_i^a(k) + K_i(k+1)\left(z_i(k+1) - \hat{\underline{g}}^T(k+1)\,\hat{\underline{\theta}}_i^a(k) \right)$$

$$K_i(k+1) = Q_i(k)\hat{\underline{g}}(k+1)\left(\hat{\underline{g}}(k+1)Q_i(k)\hat{\underline{g}}(k+1) + V_{ii}(k) \right)^{-1}$$

$$Q_i(k+1) = \left(I - K_i(k+1)\hat{\underline{g}}^T(k+1) \right)Q_i(k)$$

$$\hat{e}_i(k) = z_i(k+1) - \hat{\underline{g}}^T(k+1)\,\hat{\underline{\theta}}_i^a(k+1)$$

$$\hat{\underline{g}}(k) = \left(\underline{y}^T(k), \ldots, \underline{y}^T(k-q)\ \underline{u}^T(k), \ldots, \underline{u}^T(k-q)\right.$$
$$\left.\hat{\underline{e}}^T(k), \ldots, \hat{\underline{e}}^T(k-q) \right)$$

$$E\left(e^2(k) \right) = V_{ii}(k)$$

where the augmented parameter and measurement vectors $\underline{\theta}$ and $\underline{g}(k)$ are given in eqs. (4.33)-(4.36). Then simultaneous state estimates can be obtained from

$$\hat{\underline{x}}(k+1) = \hat{A}(k)\hat{\underline{x}}(k) + \hat{B}(k)\underline{u}(k) + \hat{K}(k)\left(z(k) - C\hat{\underline{x}}(k) \right) \quad (6.38)$$

The parameters of the input/output model are estimated first which are related to the state model system matrices. Then these estimated matrices are used in the state estimator. Application to decentralized measurements of this approach should be made and a scheme for the more general state model in eq. (3.12) obtained.

Lastly, in practice some elements of the physical state model may be known. Systems identified in canonical form are not easily related to the physical model in that form. More research of this problem needs to be investigated.

APPENDIX A

IDENTIFIABILITY RESULTS

This Appendix summarizes some of the work of R.C.K. Lee in regard to identifiability. The problem is how to identify the constant matrix A in the linear time-invariant equation

$$\underline{x}(k+1) = A\underline{x}(k) \quad . \tag{A.1}$$

Section A.1 discusses how to obtain A when all states are measurable. This leads to the property of n-identifiability in which n refers to the number of states. Section A.2 allows for only a single measurement of a linear combination of the states. It is found that the A matrix transformed to phase-variable form can be identified and leads to the property of 1-identifiability.

A.1 N-identifiability

beginning with k=0 the new state is generated according to eq. (A.1) as follows:

$$\underline{x}(1) = A\underline{x}(0)$$

$$\underline{x}(2) = A\underline{x}(1)$$

$$\vdots$$

$$\underline{x}(n) = A\underline{x}(n-1) \quad . \tag{A.2}$$

Take the n measurements of the state and concatenate them to form

$$\left[\underline{x}(1) \mid \underline{x}(2) \mid \cdots \mid \underline{x}(n)\right]$$

$$= \left[A\underline{x}(0) \mid A\underline{x}(1) \mid \cdots \mid A\underline{x}(n-1)\right]$$

$$= A\left[\underline{x}(0) \mid \underline{x}(1) \mid \cdots \mid \underline{x}(n-1)\right] \quad .$$

The A matrix is then solved for by

$$A = \left[\underline{x}(1) \mid \underline{x}(2) \mid \cdots \mid \underline{x}(n)\right] X \left[\underline{x}(0) \mid \underline{x}(1) \mid \cdots \mid \underline{x}(n-1)\right]^{-1} \quad . \tag{A.4}$$

Therefore, if all n states are measurable and the inverse of the matrix in eq. (A.4) is nonsingular then the system is n-identifiable.

A.2 1-identifiability

Suppose only one measurement is available so that the system is described by

$$\underline{x}(k+1) = A\underline{x}(k)$$

$$y(k) = \underline{c}\ \underline{x}(k)$$

where row vector $\underline{c}$ is arbitrary. This single output system will be shown to be identifiable in phase-variable canonical form.

For k=1 to n the following outputs are generated

$$y(1) = \underline{c}\,\underline{x}(1) = \underline{c}\,A\underline{x}(0)$$

$$y(2) = \underline{c}\,\underline{x}(2) = \underline{c}\,A\underline{x}(1) = \underline{c}\,A^2\underline{x}(0)$$

$$\vdots$$

$$y(n) = \underline{c}\,A^n\underline{x}(0) \qquad \text{(A.6)}$$

and concatenated into the measurement vector

$$Y(n) = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(n) \end{bmatrix} = \begin{bmatrix} cA \\ cA^2 \\ \vdots \\ cA^n \end{bmatrix} \underline{x}(0) = \begin{bmatrix} c \\ cA \\ \vdots \\ cA^{n-1} \end{bmatrix} A\underline{x}(0) \qquad \text{(A.7)}$$

$$= HA\underline{x}(0)$$

in which H is the observability matrix. Step up $\underline{Y}(n)$ to get

$$Y(n+1) = \begin{bmatrix} y(2) \\ y(3) \\ \cdot \\ \cdot \\ \cdot \\ y(n+1) \end{bmatrix} = \begin{bmatrix} cAx(1) \\ cA^2x(1) \\ \cdot \\ \cdot \\ \cdot \\ cA^nx(1) \end{bmatrix}$$

$$= \begin{bmatrix} c \\ cA \\ \cdot \\ \cdot \\ \cdot \\ cA^{n-1} \end{bmatrix} Ax(1) = HAx(1) = HA^2x(0) \ . \tag{A.8}$$

by induction

$$\underline{Y}(2n-1) = HA^n\underline{x}(0) \ . \tag{A.9}$$

Define a measurement matrix by

$$\begin{aligned} S(2n-1) &= (\underline{Y}(n) \mid \underline{Y}(n+1) \mid \ldots \mid \underline{Y}(2n-1) \\ &= (HA\underline{x}(0) \mid HA^2\underline{x}(0) \mid \ldots \mid HA^n\underline{x}(0)) \\ &= HA(\underline{x}(0) \mid A\underline{x}(0) \mid \ldots \mid A^{n-1}\underline{x}(0)) \\ &= HAB \end{aligned} \tag{A.10}$$

where B is the n-identifiability matrix required to be nonsingular in eq. (A.4) for n-identifiability. Returning to eq. (A.8), it can be written as

$$\underline{Y}(n+1) = HA^2\underline{x}(0) = HAH^{-1}HA\underline{x}(0)$$

or with

$$A^* = HAH^{-1}$$

then

$$\underline{Y}(n+1) = A^*\underline{Y}(n) \ . \tag{A.11}$$

The product HAH^{-1} as shown in Chapter 2 is the A matrix in phase-variable canonical form. For n+2

$$\underline{Y}(n+2) = HA^3\underline{X}(0) = HAH^{-1}HA^2\underline{X}(0)$$

$$= A^*\underline{Y}(n+1) \qquad (A.12)$$

and so on until

$$\underline{Y}(2n) = A^*\underline{Y}(2n-1) \ . \qquad (A.13)$$

Stepping up S(2n-1)

$$S(2n) = (\underline{Y}(n+1) \mid \underline{Y}(n+2) \mid \ldots \mid \underline{Y}(2n))$$

$$= (A^*\underline{Y}(n) \mid A^*\underline{Y}(n+1) \mid \ldots \mid A^*\underline{Y}(2n-1))$$

$$= A^*S(2n-1) \ . \qquad (A.14)$$

Solving for A^*

$$A^* = S(2n)S^{-1}(2n-1) \ . \qquad (A.15)$$

From eq. (A.10)

$$S^{-1}(2n-1) = (HAB)^{-1} = B^{-1}A^{-1}H^{-1} \qquad (A.16)$$

so that if the system is n-identifiable and observable and A is nonsingular all three inverses on the right of eq. (A.16) exist and the system is 1-identifiable.

APPENDIX B

MIMO SYSTEM IDENTIFICATION IN ROW COMPANION FORM

This Appendix summarizes some of the work that has been published on the identification of linear time-invariant systems that can be represented in row companion canonical form. Different parts of papers by Shrikhande, etal., Irwin and Roberts, and Bonivento and Guidorzi, are used here to give a straighforward procedure for the system identification. Row companion form is a particular variation of Luenberger canonical form and is reviewed in Chapter 2.

Given the general linear time-invariant state-space description in row companion canonical form

$$\underline{x}(k+1) = A\underline{x}(k) + B\underline{u}(k) + T\underline{w}(k)$$
$$\underline{y}(k) = C\underline{x}(k) + \underline{v}(k) \qquad (B.1)$$

a measurement equation, from which parameters related to elements of the system matrices are identified, is derived given that all measurements, $\underline{y}(k)$, are available. Since all of $\underline{y}(k)$ must be known this is a case of centralized

parameter identification. In the input-output measurement equation derived below centralized state and input information is required.

First, a series of measurements is concatenated and formed into a matrix equation from which the state vector can be solved for. Then other measurements are chosen from which the state is eliminated resulting in an input-output equation. Dependence coefficients appear in this equation that are linearly related to the matrix elements in eq. (6.1). With no noise all elements in A, B, and C are then found. With some or all noise present different identification schemes have been proposed and are still being researched.

with

$$C = \begin{bmatrix} c_1 \\ \cdot \\ \cdot \\ \cdot \\ c_m \end{bmatrix}$$

and

$$B = \begin{bmatrix} b_1(1) \\ \cdot \\ \cdot \\ \cdot \\ b_1(p_1) \\ \cdot \\ \cdot \\ \cdot \\ b_m(1) \\ \cdot \\ \cdot \\ \cdot \\ b_m(p_m) \end{bmatrix}$$

derive these sets of equations

$$y_i(k) = c_i \underline{x}(k) + v_i(k)$$

$$y_i(k+1) = c_i A\underline{x}(k) + c_i B\underline{u}(k) + c_i T\underline{w}(k) + v_i(k+1)$$

$$\cdot$$

$$\cdot$$

$$\cdot$$

$$y_i(k+p_i-1) = c_i A^{p_i-1} \underline{x}(k) + c_i A^{p_i-2} B\underline{u}(k)$$
$$+ \ldots + c_i B\underline{u}(k+p_i-2) + c_i A^{p_i-2} T\underline{w}(k)$$
$$+ \ldots + c_i T\underline{w}(k+p_i-2) + v_i(k+p_i-1) \qquad (B.2)$$

for $i=1,\ldots,m$ and p_i are the structural invariants which are the sizes of the main diagonal sub-blocks of the A matrix as discussed in Chapter 2. Concatenating the sets of equations gives

$$
\begin{bmatrix}
y_1(k) \\
y_1(k+1) \\
\vdots \\
y_1(k+p_1-1) \\
\vdots \\
y_M(k) \\
\vdots \\
y_M(k+p_M-1) \\
\vdots \\
y_m(k) \\
\vdots \\
y_m(k+p_m-1)
\end{bmatrix}
=
\begin{bmatrix}
c_1 \\
c_1 A \\
\vdots \\
c_1 A^{p_1-1} \\
\vdots \\
c_M \\
\vdots \\
c_M A^{p_M-1} \\
\vdots \\
c_m \\
\vdots \\
c_m A^{p_m-1}
\end{bmatrix}
\underline{x}(k)
$$

$$
+\begin{bmatrix}
0 & & & & \\
b_1(1) & & 0 & & \\
\vdots & & & & \\
b_1(p_1-1) & \cdots & b_1(1) & \cdots & 0 \\
\vdots & & & & \\
0 & & & & \\
\vdots & & & & \\
b_M(p_M-1) & \cdots & b_M(1) & & 0 \\
\vdots & & & & \\
0 & & & & \\
\vdots & & & & \\
b_m(p_m-1) & \cdots & b_m(1) & \cdots & 0
\end{bmatrix}
\begin{bmatrix}
\underline{u}(k) \\
\underline{u}(k+1) \\
\\
\\
\\
\vdots \\
\\
\\
\\
\underline{u}(k+p_m-1)
\end{bmatrix}
$$

$$
\begin{bmatrix}
0 & & & & \\
t_1(1) & & 0 & & \\
\vdots & & & & \\
t_1(p_1-1) & \cdots & t_1(1) & \cdots & 0 \\
\vdots & & & & \\
0 & & & & \\
\vdots & & & & \\
t_M(p_M-1) & \cdots & t_M(1) & & 0 \\
\vdots & & & & \\
0 & & & & \\
\vdots & & & & \\
t_m(p_m-1) & \cdots & t_m(1) & \cdots & 0
\end{bmatrix}
\begin{bmatrix}
\underline{w}(k) \\
\underline{w}(k+1) \\
\\
\\
\\
\vdots \\
\\
\\
\\
\\
\\
\\
\\
\\
\underline{w}(k+p_m-1)
\end{bmatrix}
+
\begin{bmatrix}
v_1(k) \\
\vdots \\
v_m(k+p_m-1)
\end{bmatrix}
\qquad (b.3)
$$

where $p_M = \max(\mu_i)$, i=1, ..., m. Each c_i is a unit vector with a single "1" and due to the structure of A it multiplies a "1" in a row of A giving a unit vector with a "1" moved one place to the right. Each multiplication of c_i A by another A results in moving the "1" one more place to the right. This holds for powers of A less than or equal to $\mu_i - 1$. As a result the matrix multiplying $\underline{x}(k)$ is the identity matrix and (B.3) is now compactly written as

$$\bar{y}(k) = \underline{x}(k) + M\bar{u}(k) + N\bar{w}(k) + \bar{v}(k) \quad . \tag{B.4}$$

Stepping up eq. (B.4)

$$\bar{y}(k+1) = \underline{x}(k+1) + M\bar{u}(k+1) + N\bar{w}(k+1) + \bar{v}(k+1) \tag{B.5}$$

and using $\underline{x}(k+1)$ from eq. (B.1)

$$\bar{y}(k+1) = A\underline{x}(k) + B\underline{u}(k) + I\underline{w}(k) + M\bar{u}(k+1) + N\bar{w}(k+1) + \bar{v}(k+1) \quad . \tag{B.6}$$

The state vector can be solved for from eq. (B.4)

$$\underline{x}(k) = \bar{y}(k) - M\bar{u}(k) - N\bar{w}(k) - \bar{v}(k)$$

and substituting into eq. (B.6)

$$\bar{y}(k+1) = A\bar{y}(k) - AM\bar{u}(k) - AN\bar{w}(k) + B\underline{u}(k) + I\underline{w}(k) + M\bar{u}(k+1) + N\bar{w}(k+1) - A\bar{v}(k) + \bar{v}(k+1) \quad . \tag{B.7}$$

Defining

$$\bar{u}'(k) = \begin{bmatrix} \underline{u}(k) \\ \hline \bar{u}(k+1) \end{bmatrix} = \begin{bmatrix} \underline{u}(k) \\ \underline{u}(k+1) \\ \cdot \\ \cdot \\ \cdot \\ \underline{u}(k+p_m) \end{bmatrix} \tag{b.8a}$$

and

$$\bar{w}'(k) = \begin{bmatrix} \underline{w}(k) \\ \hline \bar{w}(k+1) \end{bmatrix} = \begin{bmatrix} \underline{w}(k) \\ \underline{w}(k+1) \\ \cdot \\ \cdot \\ \cdot \\ \underline{w}(k+p_m) \end{bmatrix} \tag{B.8b}$$

terms in eq. (E.7) are combined as follows

$$B\underline{u}(k) + M\bar{u}(k+1) = \left[B \mid M\right] \begin{bmatrix} \underline{u}(k) \\ \hline \bar{u}(k+1) \end{bmatrix} = B'\bar{u}'(k) \tag{B.9}$$

and

$$I\underline{w}(k) + N\bar{w}(k+1) = \left[I \mid N\right] \begin{bmatrix} \underline{w}(k) \\ \hline \bar{w}(k+1) \end{bmatrix} = I'\bar{w}'(k) \tag{B.9}$$

Augmenting the terms $-AM\bar{u}(k)$ and $-AN\bar{w}(k)$ with null matrices to

$$-AM\bar{u}(k) = -A\left[M \mid 0\right] \begin{bmatrix} \bar{u}(k) \\ \hline \underline{u}(k+p_m) \end{bmatrix} = -A(M0)\bar{u}'(k) \tag{b.11}$$

and

$$-AN\bar{w}(k) = -A\left[N \mid 0\right]\begin{bmatrix}\bar{w}(k)\\ \hline \underline{w}(k+p_m)\end{bmatrix} = -A(N0)\bar{w}'(k) \qquad (B.11)$$

they can then be added to (B.9) and (B.10) to give

$$(B' - A(M0))\bar{u}'(k) = B''\bar{u}'(k)$$

$$(T' - A(N0))\bar{w}'(k) = T''\bar{w}'(k) \quad . \qquad (B.13)$$

Eq. (B.7) is now written as

$$\bar{y}(k+1) = A\bar{y}(k) + B''\bar{u}'(k) + T''\bar{w}'(k)$$

$$- A\bar{v}(k) + \bar{v}(k+1) \quad . \qquad (B.14)$$

Next define A*, B*, and T* as matrices with only the p_1 th, p_2 th, ..., p_m th rows of A, B'', and T'' and define

$$y^*(k+1) = \begin{bmatrix} y_1(k+p_1)\\ y_2(k+p_2)\\ \cdot\\ \cdot\\ \cdot\\ y_m(k+p_m)\end{bmatrix} \qquad (B.15)$$

and

$$v^*(k+1) = \begin{bmatrix} v_1(k+p_1)\\ v_2(k+p_2)\\ \cdot\\ \cdot\\ \cdot\\ v_m(k+p_m)\end{bmatrix} . \qquad (B.16)$$

Eq. (B.14) is "compressed" to

$$y^*(k+1) = A^*\bar{y}(k) + B^*\bar{u}'(k) + T^*\bar{w}'(k)$$

$$- A^*\bar{v}(k) + v^*(k+1) \quad . \qquad (B.16)$$

Eq. (B.16) is an input-output measurement equation that can be used in identification algorithms.

The rest of this appendix will look at the deterministic case so that the last three terms in eq. (B.16) are zero. Eq. (B.16) is now written as

$$y^*(k+1) = \left[A^* \mid B^*\right] \begin{bmatrix} \bar{y}(k) \\ \hline \bar{u}'(k) \end{bmatrix} . \qquad (B.17)$$

Collecting N measurements

$$y^*(k+1), \ldots, y^*(k+N)$$

$$= \left[A^* \mid B^*\right] \begin{bmatrix} \bar{y}(k), & \ldots, & \bar{y}(k+N-1) \\ \bar{u}'(k), & \ldots, & \bar{u}'(k+N-1) \end{bmatrix} \qquad (B.18)$$

with $N=n+r(p_m+1)$, the number of unknown parameters, the parameters are obtained by post-multiplying eq. (B.18) by the inverse of the NXN measurement matrix. The matrix A^*, the significant rows of A, identifies A in row companion form but B^* is a matrix of dependence coefficients linearly related to B. Bonivento and Guidorzi relate B to B^* through the inverse of a complicated linear transformation matrix. In the SISO case this matrix reduces to eq. (2.4). The approach taken here from Shrikhande, etal. avoids inverses and is easier to implement.

Returning to eq. (B.13) and retaining p_1th, p_2th, to p_mth rows gives

$$L'^* - A^*(M0) = B^* \qquad (B.19)$$

where starred matrices are compressed. The matrices B'*, M0, and B* are partitioned into submatrices r columns wide

$$M0 = (M \mid 0) = (M(1) \mid \ldots \mid M(p_M) \mid 0)$$

$$B'^* = (BC \mid MC) = (BC \mid MC(1) \mid \ldots \mid MC(p_M))$$

$$B^* = (BC^*(1) \mid \ldots \mid BC^*(p_M+1)) \quad . \qquad (B.20)$$

Eq. (B.19) is explicitly written as

$$(BC \mid MC(1) \mid \ldots \mid MC(p_M)) - A^*(M(1) \mid \ldots \mid M(p_M) \mid 0)$$

$$= (BC^*(1) \mid \ldots \mid BC^*(p_M+1)). \qquad (B.21)$$

Beginning at the right gives

$$MC(p_M) - A^* X 0 = BC^*(p_M+1) \text{ or}$$

$$MC(p_M) = BC^*(p_M+1) .$$

The matrix on the right is known and $M(p_M)$ can be determined from $MC(p_M)$ (see M in eq. B.4). Second from the right in eq. (B.21)

$$MC(p_M-1) - A^*M(p_M) = BC^*(p_M) \quad .$$

This procedure is repeated until BC is found which has the p_1 th through p_m th rows of B and all of B is known.

BIBLIOGRAPHY

Alam, A. and Sage, A., "Error and Sensitivity Analysis of Stochastic Approximation Algorithms for Linear System Identification," IEEE Trans. Aut. Contr., vol. 19, pp. 810-815, Dec. 1974.

Aoki, M. and Li, M., "Controllability and Stabilizability of Decentralized Dynamic Systems," Proc. of 14th JACC, Ohio State Univ., Columbus, pp. 278-288, July 1973.

Aoki, M., "On Feedback Stabilizability of Decentralized Dynamic Systems," Automatica, vol. 8, pp. 163-173, Mar. 1973.

Anderson, B.D.O. and Hong, H.M., "Structured Controllability and Matrix Nets," Int. J. Contr., vol. 35, no. 3, pp. 397-416, 1982.

Balas, M., "Large Space Structure Control Theory; Current Trends," IEEE Trans. Aut. Contr., vol. 27, pp. 522-535, June 1982.

Bonivento, C. and Guidorzi, R., "Canonical Input-Output Description of Linear Multivariable Systems," Int. J. Control, vol. 23, pp. 290-293, June 1967.

Brasch, Jr., F. and Pearson, J., "Pole Placement Using Dynamic Compensators," IEEE Trans. Aut. Contr., vol. 15, pp. 34-43, Feb. 1970.

Bretthauer, G., "Identifiability of Subsystems in Interconnected Systems," Proc. of IFAC/IFORS Symposium, Warsaw, Poland, pp. 417-422, July 1983.

Budin, M., "Minimal Realization of Discrete Linear Systems from Input-Output Observations," IEEE Trans. Aut. Contr., vol. 16, pp. 395-401, Oct. 1971.

Corfmat, J. and Morse, A.a, "Decentralized Control of Linear Multivariable Systems," Automatica, vol. 12, pp. 479-495, Sept. 1976.

Corfmat, J. and Morse, A.b, "Stabilization with Decentralized Feedback Control," IEEE Trans. Aut. Contr., vol. 18, pp. 679-682, Dec. 1973.

Chen, H., and P. Caines, "The Strong Consistency of the Stochastic Gradient Algorithm of Adaptive Control," IEEE Trans. Aut. Contr., vol. 30, Feb. 1985.

Davison, E., "The Robust Decentralized Control of a General Servomechanism Problem," IEEE Trans. Aut. Contr., vol. 21, pp. 14-24, Feb. 1976.

Davison, E. and Wang, S.a, "Properties and Calculation of Transmission Zeros of Linear Multivariable Systems," Automatica, vol. 10, 1974.

Davison, E. and Wang, S.b, "Properties of Linear Time-Invariant Multivariable Systems Subject to Arbitrary Output and State Feedback," IEEE Trans. Aut. Contr., vol. 18, pp. 24-32, Feb. 1973.

Denham, M., "Canonical Forms for the Identification of Multivariable Linear Systems," IEEE Trans. Aut. Contr., vol. 19, pp.646-656, Dec. 1974.

Dickinson, b., "Canonical Matrix Fraction and State-Space Descriptions for Deterministic and Stochastic Linear Systems," IEEE Trans. Aut. Contr., vol. 19, pp. 656- 667, Dec. 1974.

Graupe, D.a, Identification of Systems, New York: Van Nostrand Reinhold Company, 1972.

Graupe, D.b, "On Convergence of Least-Squares Identifiers of Autoregressive Models Having Stable and Unstable Roots,"

IEEE Trans. Aut. Contr., vol. 25, pp. 999-1002, Oct. 1980.

Graupe, D., Jain, V., and Salahi, J., "A Comparative Analysis of Various Least-Squares Identification Algorithm," Automatica, vol. 16, pp. 663-681, Nov. 1980.

Holmes, J., "Two Stochastic Approximation Procedures for Identifying Linear Systems," IEEE Trans. Aut. Contr., vol. 14, pp. 292-295, 1969.

Irwin, G. and Roberts, A., "The Luenberger Canonical Form in the State/Parameter Estimation of Linear Systems," Int. J. Control, vol. 23, pp. 851-864, 1976.

Ikeda, M. and Siljak, D.D., "Decentralized Stabilization of Linear Time-Varying Systems," IEEE Trans. Aut. Contr., vol. 25, pp. 106-107, Feb. 1960.

Jamshidi, M., Large-Scale Systems Modeling and Control, New York: Elsevier Science Publishing Co., 1983.

Kalman, R. and Bertram, J., "Control System Analysis and Design Via the Second Method of Lyapunov: I. Continuous Time Systems," ASME J. Basic Eng., Series D, pp. 371-393, 1960.

Kobayashi, H., Hanafusa, H., and Yoshikawa, T., "Controllability Under Decentralized Information Structure," IEEE Trans. Aut. Contr., vol. 23, pp. 182-188, Apr. 1978.

Krtolica, R. and Siljak, D., "Suboptimality of Decentralized Stochastic Control and Estimation," IEEE Trans. Aut. Contr., vol. 25, pp. 76-83, Feb. 1980.

Kuo, B., Automatic Control Systems, Englewood Cliffs: Prentice Hall, Inc., 1982.

Ljung, L., "Identification of Linear, Multivariable Systems Operating Under Linear Feedback Control," IEEE Trans. Aut. Contr., vol. 19, pp. 836-840, Dec. 1974.

Luenberger, D., "Canonical Forms for Linear Multivariable Systems," IEEE Trans. Aut. Contr., vol. 12, pp. 290-293, June 1967.

Mehra, R., "On-Line Identification of Linear Dynamic Systems with Applications to Kalman Filtering," IEEE Trans. Aut. Contr., vol. 16, Feb. 1971.

Mendel, J., Discrete Techniques of Parameter Estimation - The Equation Error Formulation, N.Y.: Marcel Dekker, Inc., 1973.

Morse, A., "Structural Invariants of Linear Multivariable Systems," SIAM J. Control, vol. 11, pp. 446-465, Aug. 1973.

Morse, A. and Wonham, W., "Decoupling and Pole Assignment By Dynamic Compensation," SIAM J. Control, vol. 8, pp. 317-337, Aug. 1970.

Padulo, L. and Arbib, M., System Theory, Washington: W.B. Saunders Co., 1974.

Panuska, V., "An Adaptive Recursive Least Squares Identification Algorithm," Proc. 8th IEEE Symp. on Adaptive Processes, Paper 6-e-1, 1969.

Sandell, N., Jr., Varaiya, P., Athans, M., and Safonov, M., "Survey of Decentralized Control Methods for Large Scale Systems," IEEE Trans. Aut. Contr., vol. 23, pp. 108-128, Apr. 1978.

Seal, C., "Canonical Forms for Multiple Input Time Systems," IEEE Trans. Aut. Contr., vol. 14, pp. 704-707, Dec. 1969.

Shrikhande, V., Mital, D., and Ray, L., "On Minimal Canonical Realization from Input-Output Data Sequences," IEEE Trans. Aut. Contr., vol. 25, pp. 309-317, Apr. 1980.

Sivan, R., and Kwakernaak, H., Linear Optimal Control Systems, New York: John Wiley, 1972.

Sinha, N., and Rozsa, P., "Some Canonical Forms for Linear Multivariable Systems," Int. J. Control, vol. 23, pp. 865-883, 1976.

Srinath, M. and Viswanathan, M., "Sequential Algorithm for Identification of Parameters of an Autoregressive Process," IEEE vol. 20, pp. 542-546, Aug. 1976.

Thau, F. and Montgomery, R., "Adaptive Learning Control of Large Space Structures: System Identification Techniques," Proc. of JACC, Paper TP1-D, 1980.

Wang, S. and Davison, E.a, "A Characterization of Decentralized Fixed Modes in Terms of Transmission Zeros," IEEE Trans. Aut. Contr., vol. 30, pp. 81-82, Jan. 1985.

Wang, S. and Davison, E.b, "On the Stabilization of Decentralized Control Systems," IEEE Trans. Aut. Contr., vol. 18, pp. 473-478, Oct. 1973.