1972

Methods and applications of systems identification

Allen Glenn Behring

Department: Mechanical and Aerospace Engineering

METHODS AND APPLICATIONS OF SYSTEM IDENTIFICATION

by

ALLEN GLENN BEHRING, 1943-

A DISSERTATION

Presented to the Faculty of the Graduate School of the

UNIVERSITY OF MISSOURI - ROLLA

In Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

in

MECHANICAL ENGINEERING

1972

Virgil J. Flanigan
Advisor

R. T. Johnson

John S. Pazdera

C. Y. Ho

Lawrence K. Sieck

J. R. Fawcett

PUBLICATION THESIS OPTION

The papers presented within the body of this thesis have been prepared in the style utilized by the American Society of Mechanical Engineers. Pages 1-150 will be submitted to the A.S.M.E. Journal of Dynamic Systems, Measurement and Control for publication.

Because of journal requirements, matrices and vectors have been denoted by placing a single wavy line below their corresponding symbols. Symbols designated in this manner will appear in bold-face type within the journal copy.

An index and an appendix have been added for purposes normal to thesis writing.

ABSTRACT

A "Schmidt filter" is proposed to compute an optimal orthonormal basis for a set of "noisy" filter input functions. Procedures for determining the transfer function and inverse transfer function of the filter are given.

The Schmidt filter is applied to the problem of determining mathematical models of discrete, stationary, linear, dynamic systems for the case where measurements may be corrupted by noise of unknown statistics.

The identification problem is reconsidered for the case where noise and signal moments are specified. Procedures are given which insure unbiased, adaptive estimates of system order and parameters for this case.

These theoretical propositions are applied to the modeling of speculative prices. The stock market is formulated as a discrete, linear, dynamic system and the results of several simulation studies are presented. Evidence indicates that certain segments of the market can be approximated by high-order linear systems computed from small samples and tends to refute the random walk hypothesis.

Computer programs (written in PL/1) are presented which allow for efficient digital realization of the theoretical procedures discussed in the body of this work.

PREFACE

During the last decade, mathematical analysis and simulation have become common in nearly all areas of scientific inquiry. In fact, it may be observed that these tools have themselves become respected scientific disciplines.

State space (modern) control theory is a powerful mathematical tool that has enabled engineers to design and regulate complex mechanical/electrical devices. Modern control theory is based upon the premise that a process can be described by a system of differential equations in time, and that such a system of equations has an equivalent representation with respect to a single multi-dimensional vector space from which its "state" can be determined. Given the initial state of the system and a time-ordered set of "independent" (input) variables, it is possible to predict the dependent variables (outputs) of the system for times defined by the input set. To date, the greatest portion of work in modern control theory deals with processes which are describable by systems of linear differential equations.

Recently, the "identification" problem has received considerable attention in the literature on automatic control. Identification involves determination of the describing equations of a process directly from input/output data. Work in identification has been frequently directed toward systems which are linear, stationary, and possibly subject to random error in measurement. Occasionally, publications appear dealing with non-stationary systems, unknown inputs, and correlated noise.

Social scientists depend more and more upon mathematical methods for the detection and analysis of relationships within an increasingly complex and mobile societal structure. The science of econometrics, which deals with the quantization and analysis of economic phenomena, has long been a topic of considerable interest. Among the best-known tools employed in the studies of social and economic phenomena are regression analysis and factor analysis, which are commonly directed to the problem of linear approximation. The identification problem in econometrics also involves the establishment of a mathematical model of a process from observed variables.

The identification problems in engineering and social science are fundamentally similar. Both involve the abstraction of physical phenomena as a set of observable variables followed by a testing of hypotheses concerning relationships between these variables. Furthermore, a thorough examination of pertinent literature reveals that the differences in representation and methods of analyzing the identification problem in the two disciplines are largely superficial. Surprisingly, this fact seems to have been obscured even though significant contributions have been made in both areas.

Consider the problem of stock market investment. It has been theorized that the determinant of speculative prices is a set of expectations on the part of market participants concerning future conditions. These expectations are determined by past and current prices and other information which is presumed to affect future gains. Naturally, investment involves risk. The indicators of economic gain may change in the next instant -- fortunes may be gained or lost.

Actually, we are all investors in a sense, regardless of whether or not we choose to participate in the stock market. Our activities, and hence those of our society, are primarily based upon expectations concerning the future over which we have little control. For this reason, it is the Author's contention that scholarly research into the area of speculative prices will yield benefits in the analysis of social processes which will far exceed any prospect for financial reward.

In order for our governing bodies to cope democratically with the ever-increasing complexity of our society, it appears that we need to achieve a much greater quantitative understanding of the phenomena which motivate human behavior. While it is doubtless a great oversimplification to presume that the world's problems can be overcome simply by the study of speculative prices, this problem is a convenient one for investigative research. Generally, data related to speculative prices is easy to obtain. Also, there exists a "natural" interest in this topic which tends to reduce the barriers of communication between the sciences.

The problems which face our society today belong to us all. Hence, engineering research into sociological problems, while somewhat rare, is not inappropriate. This thesis responds to a need for increased communication between the engineering and social science communities. The identification problem is approached rigorously from the viewpoint of state space control theory. Three technical papers are first advanced to deal with the mechanics of the identification problem. Here, several new propositions are presented which allow for greater efficiency and increased generality of realization procedures for both the noise-free and noisy cases. A fourth paper deals with the application of the above propositions toward the understanding of speculative prices. The Appendix of this thesis contains a listing of computer programs which have been developed during the course of the Author's research. These programs have general applicability to the linear modeling of all processes that are ammenable to numerical quantification. It is hoped that this overall approach will motivate the interest of both the engineer and the social scientist.

The Author is indebted to the University of Missouri and the Department of Mechanical and Aerospace Engineering of the University of Missouri - Rolla for supporting this research and for provision of the Author's graduate assistantship during its conduct.

I am especially grateful to Dr. V. J. Flanigan, my major Advisor, for his guidance, encouragement, and technical assistance in the preparation of this work. Drs. R. T. Johnson, J. S. Pazdera, C. Y. Ho, and L. K. Sieck have also contributed significantly to this thesis and have my sincere thanks.

I reserve my deepest appreciation for my wife, Elizabeth, who exhibited great patience, provided continual encouragement, and typed the manuscript.

A.G.B.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

I. A Schmidt Orthonormal Filter for System Identification

Figures

IV. On Stock Market Dynamics

Figures

LIST OF TABLES

A SCHMIDT ORTHONORMAL FILTER FOR SYSTEM IDENTIFICATION

by

A. G. Behring

and

V. J. Flanigan*

ABSTRACT

A Schmidt filter is proposed to compute an optimal orthonormal basis for a set of "noisy" filter input functions. Procedures for determining the transfer function and inverse transfer function of the filter are given. Several interesting properties of the filter are noted and applied to the problem of system identification.

---

*The Authors are associated with the Department of Mechanical and Aerospace Engineering, University of Missouri - Rolla, Rolla, Missouri 65401, where Mr. Behring is a Graduate Student and Dr. Flanigan (member ASME) is an Associate Professor.

NOTATION

In this paper all bold-face capital letters denote matrices. Vectors are defined in column format and are denoted by lower case letters in bold face type. All scalars will be denoted by plain upper or lower case letters. Occasionally it will be necessary to display the format of a vector or matrix explicitly, e.g.,

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}.$$

Any exceptions to these general rules will be clearly specified in the text.

INTRODUCTION

In recent years several authors [1-17][1] have noted the simplifying features of orthogonal sets of functions in problems of identification and optimal control. The basic philosophy underlying most of these investigations consists of an expansion of the system input set as a series of

---

[1] Numbers in brackets denote references at the end of the paper.

orthogonal functions, the sum of whose additive contributions to the system output (system transfer function) is then determined by an optimal choice of independently-adjustable parameters.

Ho [18], Gopinath [19], and Budin [20] have developed algorithms for computing minimum-order mathematical models of discrete, time-invariant linear systems from input/output data. A central problem in the implementation of these algorithms (especially where the system order is unknown) is the determination of matrix rank (possibly in the presence of additive noise). Budin [20] proposes a solution to the problem of noise-corrupted observations using a modified Gaussian elimination algorithm.

The Gram-Schmidt procedure (sometimes called orthonormalization due to Erhard Schmidt), Epstein [21], Drygas [22], for computing an orthonormal basis of a set of vectors is well-known in the literature of linear algebra. Bingulac [23] gave an original method for computing an orthonormal basis from a set of linearly independent functions and demonstrated some apparent computational advantages of his procedure over the Gram-Schmidt process.

Several authors, including Penrose [24], Greville [25], Rao [26], and Mayne [27], have considered the problem of finding an inverse of singular matrices and have demonstrated the utility of such a "generalized inverse" or "pseudo inverse" in the solution of linear systems of equations. Mayne [27] used the Gram-Schmidt procedure to compute a pseudo inverse with allowances for computational round-off error that seem applicable to the noisy case.

PROBLEM STATEMENT

Given a "noisy" function $\underset{\sim}{y}(t)$, the problem is to define a linear filter with transfer function $\underset{\sim}{S}$ to compute an "optimal" orthonormal basis function $\underset{\sim}{x}(t)$. We also seek to define an "inverse filter" with transfer function $\underset{\sim}{S}^{+}$ which maps $\underset{\sim}{x}(t)$ into $\underset{\sim}{y}(t)$. We then consider $\underset{\sim}{y}(t)$ to be the output of a linear system and proceed to exploit the unique features of the orthonormal basis function $\underset{\sim}{x}(t)$ in the problem of system identification.

BASIC DEFINITIONS

Let $\underset{\sim}{X}$ be defined as the set of real $n\times 1$ vectors

$$\underset{\sim}{X} = [\underset{\sim}{x}(t_1),\underset{\sim}{x}(t_2),\underset{\sim}{x}(t_3),\ldots,\underset{\sim}{x}(t_r)], \tag{1}$$

where $t_i<t_{i+1}, \forall i$. Here $\underset{\sim}{x}(t)$ is a real-valued vector function of time t. Note that both the range and domain of $\underset{\sim}{x}(t)$ are defined by equation (1). Similarly, let $\underset{\sim}{Y}$ be defined as the set of real $p\times 1$ vectors

$$\underset{\sim}{Y} = [\underset{\sim}{y}(t_1),\underset{\sim}{y}(t_2),\underset{\sim}{y}(t_3),\ldots,\underset{\sim}{y}(t_r)]. \tag{2}$$

In the event that r approaches infinity, it will be assumed that both $\underset{\sim}{y}(t)$ and $\underset{\sim}{x}(t)$ are sectionally continuous. Further, it will be required that $\underset{\sim}{y}(t)$ and $\underset{\sim}{x}(t)$ exist, $\forall t_i$.

Orthogonality

Two functions $\underset{\sim}{x}(t)$ and $\underset{\sim}{y}(t)$ as in equations (1) and (2) will be called orthogonal if

$$\underset{\sim}{X}\underset{\sim}{Y}' = \underset{\sim}{0} \tag{3}$$

where $\underset{\sim}{0}$ is the n×p null matrix. Note that the condition of orthogonality given by equation (3) implies that

$$\underset{\sim}{Y}\underset{\sim}{X}' = \underset{\sim}{0}' \tag{4}$$

since

$$\underset{\sim}{Y}\underset{\sim}{X}' = [\underset{\sim}{X}\underset{\sim}{Y}']'. \tag{5}$$

Now, let $\underset{\sim}{A}$ be a constant m×p matrix where m and the elements of $\underset{\sim}{A}$ are arbitrary. We now consider the orthogonality of functions $\underset{\sim}{x}(t)$ and $\underset{\sim}{A}\underset{\sim}{y}(t)$, i.e., we write

$$\underset{\sim}{X}(\underset{\sim}{A}\underset{\sim}{Y})' = \underset{\sim}{X}\underset{\sim}{Y}'\underset{\sim}{A}'. \tag{6}$$

If equation (3) is satisfied, we conclude from equation (6) that

$$\underset{\sim}{X}(\underset{\sim}{A}\underset{\sim}{Y})' = \underset{\sim}{0} \tag{7}$$

where $\underset{\sim}{A}$ is arbitrary. This result significantly implies that $\underset{\sim}{x}(t)$ is orthogonal to the function space spanned by $\underset{\sim}{y}(t)$ and vice-versa.

If $\underset{\sim}{x}(t)$ and $\underset{\sim}{y}(t)$ are point functions, we compute $\underset{\sim}{X}\underset{\sim}{Y}'$ from

$$\underset{\sim}{X}\underset{\sim}{Y}' = \sum_{i=1}^{r} \underset{\sim}{x}(t_i)\underset{\sim}{y}'(t_i). \tag{8}$$

If $\underset{\sim}{x}(t)$ and $\underset{\sim}{y}(t)$ are sectionally continuous functions, we define $\underset{\sim}{X}\underset{\sim}{Y}'$ as

$$\underset{\sim}{X}\underset{\sim}{Y}' = \int_{t_1}^{t_r} \underset{\sim}{x}(t)\underset{\sim}{y}'(t)dt. \tag{9}$$

Orthonormality

We will say that the function $\underset{\sim}{x}(t)$ of equation (1) is orthonormal if

$$\underset{\sim}{X}\underset{\sim}{X}' = \underset{\sim}{I}_n \tag{10}$$

where $\underset{\sim}{I}_n$ is the n×n identity matrix. It is apparent that the condition of orthonormality expressed by equation (10) implies that $\underset{\sim}{X}$ is full rank.

LEAST SQUARES

Let $\underset{\sim}{X}$ and $\underset{\sim}{Y}$ be defined as in equations (1) and (2), rsspectively. Let $\hat{\underset{\sim}{A}}$ be a n×p constant matrix. It can be shown that, e.g., Sage and Melsa [28], that an optimal linear conditional estimate of $\underset{\sim}{y}(t)$ given $\underset{\sim}{x}(t)$ in the sense of minimum mean square[2] $\{\hat{\underset{\sim}{y}}(t)|\underset{\sim}{x}(t)\}$ is

$$\hat{\underset{\sim}{y}}(t)|\underset{\sim}{x}(t) = \hat{\underset{\sim}{A}}\underset{\sim}{x}(t) \tag{11}$$

where $\hat{\underset{\sim}{A}}$ is a constant matrix which satisfies

$$\hat{\underset{\sim}{A}}\underset{\sim}{X}\underset{\sim}{X}' = \underset{\sim}{Y}\underset{\sim}{X}'. \tag{12}$$

If $\underset{\sim}{X}\underset{\sim}{X}'$ is non-singular, equation (12) yields

$$\hat{\underset{\sim}{A}} = \underset{\sim}{Y}\underset{\sim}{X}'[\underset{\sim}{X}\underset{\sim}{X}']^{-1}. \tag{13}$$

Also, we can demonstrate the significant result that the p×r matrix $\underset{\sim}{E}$ defined by

$$\underset{\sim}{E} = \underset{\sim}{Y}-\hat{\underset{\sim}{A}}\underset{\sim}{X} \tag{14}$$

is orthogonal to $\underset{\sim}{X}$.

---

[2]Here, $\hat{\underset{\sim}{A}}$ minimizes $\{[\underset{\sim}{Y}-\underset{\sim}{A}\underset{\sim}{X}][\underset{\sim}{Y}-\underset{\sim}{A}\underset{\sim}{X}]^{-1}\}$

Using (13) and (14) we can write

$$\underset{\sim}{E}\underset{\sim}{X}' = \{\underset{\sim}{Y}-\underset{\sim}{Y}\underset{\sim}{X}'(\underset{\sim}{X}\underset{\sim}{X}')^{-1}\underset{\sim}{X}\}\underset{\sim}{X}' \tag{15}$$

or

$$\underset{\sim}{E}\underset{\sim}{X}' = \underset{\sim}{0}. \tag{16}$$

THE SCHMIDT FILTER

General Description

Given the vector function $\underset{\sim}{y}(t)$ of equation (2), we wish to produce a vector function $\underset{\sim}{x}(t)$ of equation (1) such that

$$\underset{\sim}{X}\underset{\sim}{X}' = \underset{\sim}{I}_n, \tag{17}$$

$$\underset{\sim}{X} = \underset{\sim}{S}\underset{\sim}{Y}, \tag{18}$$

and

$$\underset{\sim}{Y} = \underset{\sim}{S}^+\underset{\sim}{X}. \tag{19}$$

The linear transformation $\underset{\sim}{S}$ of equation (18) may be considered as the transfer function of a filter (here called the Schmidt filter). The linear transformation $\underset{\sim}{S}^+$ may be considered as the inverse transfer function of the Schmidt filter (or alternately as the transfer function of a restoring filter). The relationship between the variables can be conveniently displayed as in Figure 1.

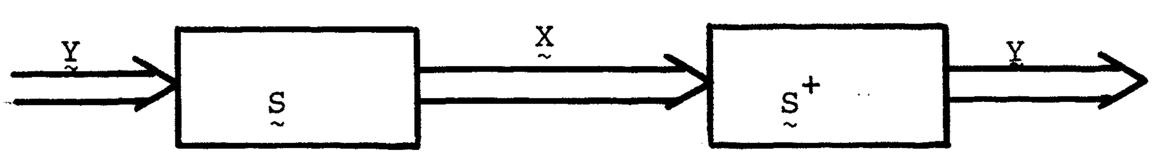

Fig. 1 Defining Relationships for the Schmidt Filter

We now consider some of the interesting properties of the filter which can be inferred directly from equations (17) through (19). First, if we post-multiply equation (19) by $\underset{\sim}{X}'$ and employ equation (17), it is easy to see that

$$\underset{\sim}{S}^{+} = \underset{\sim}{Y}\underset{\sim}{X}'. \tag{20}$$

From equation (20) we observe that

$$\underset{\sim}{S}\underset{\sim}{S}^{+} = \underset{\sim}{S}\underset{\sim}{Y}\underset{\sim}{X}'. \tag{21}$$

Using equations (18) and (17), equation (21) becomes

$$\underset{\sim}{S}\underset{\sim}{S}^{+} = \underset{\sim}{I}_n. \tag{22}$$

Using (22) we find that $\underset{\sim}{S}^{+}$ qualifies as a generalized inverse of $\underset{\sim}{S}$, Rao [26], i.e.,

$$\underset{\sim}{S}\underset{\sim}{S}^{+}\underset{\sim}{S} = \underset{\sim}{S}, \tag{23}$$

and

$$\underset{\sim}{S}^{+}\underset{\sim}{S}\underset{\sim}{S}^{+} = \underset{\sim}{S}^{+}. \tag{24}$$

The Filter Algorithm

For convenience, we define $\underset{\sim}{E}$ as the set of p×1 vectors

$$\underset{\sim}{E} = [\underset{\sim}{e}(t_1), \underset{\sim}{e}(t_2), \underset{\sim}{e}(t_3), \ldots, \underset{\sim}{e}(t_r)]. \tag{25}$$

We will denote the jth row of matrix $\underset{\sim}{E}$ as $\underset{\sim}{E}_j$, j=1,p. Let $\underset{\sim}{X}$ and $\underset{\sim}{Y}$ be defined as in equations (1) and (2), respectively, with meaning as shown in Figure 1. Also, we let $\underset{\sim}{X}_i$, i=1,n denote the ith row of $\underset{\sim}{X}$ and $\underset{\sim}{Y}_j$, j=1,p denote the jth row of $\underset{\sim}{Y}$. Finally, it will be convenient to let $\underset{\sim}{Z}_i$, i=1,n denote the <u>first i rows</u> of $\underset{\sim}{X}$, where it should be noted for later analysis that

$$\underset{\sim}{Z}_i\underset{\sim}{Z}_i' = \underset{\sim}{I}_i, \tag{26}$$

where $\underset{\sim}{I}_i$ is the i×i identity matrix. In order to initiate the algorithm, we will assume that $\underset{\sim}{Y}_1$ is non-trivial. In this event, we let

$$\underset{\sim}{E}_1 = \underset{\sim}{Y}_1 \tag{27}$$

and

$$\underset{\sim}{X}_1 = \underset{\sim}{E}_1/||\underset{\sim}{E}_1|| \tag{28}$$

where $||\underset{\sim}{E}_1||$ denotes the "norm" of $\underset{\sim}{E}_1$, given by

$$||\underset{\sim}{E}_1|| = [\underset{\sim}{E}_1\underset{\sim}{E}_1']^{\frac{1}{2}} \tag{29}$$

We generate successive rows of $\underset{\sim}{X}$ from the relations

$$\underset{\sim}{E}_j = \underset{\sim}{Y}_j - [\underset{\sim}{Y}_j\underset{\sim}{Z}_i']\underset{\sim}{Z}_i, \; j=2,p, \; i=1,n-1, \tag{30}$$

and

$$\underset{\sim}{X}_{i+1} = \underset{\sim}{E}_j/||\underset{\sim}{E}_j|| \quad \text{if } (\underset{\sim}{E}_j\underset{\sim}{E}_j')/(\underset{\sim}{Y}_j\underset{\sim}{Y}_j') > \underset{\sim}{\varepsilon}_j, \tag{31}$$

where the index i always indicates the number of "currently defined" rows of $\underset{\sim}{X}$ and $\underset{\sim}{\varepsilon}_j$ is prespecified (see sections entitled Computational Error and Noise). Equations (30) and (31) are executed p-1 times, beginning with j=2 and i=1. Each time these equations have been executed, the subscript j is incremented by one. If the condition indicated by equation (31) is true, then the (i+1)th row of $\underset{\sim}{X}$ is defined as shown and the subscript i is incremented by one. If this condition is not true the subscript i is not incremented.

Note that a least squares procedure, simplified by equation (26), has been used to compute $\underset{\sim}{E}_j$ in equation (31). We can therefore conclude that each $\underset{\sim}{E}_j$ is orthogonal to the function space spanned by the rows of $\underset{\sim}{Z}_i$. Provided that $\underset{\sim}{E}_j$ is non-trivial, this implies that $\underset{\sim}{E}_j/||\underset{\sim}{E}_j||$ should be included in the orthonormal matrix $\underset{\sim}{X}$, i.e., as $\underset{\sim}{X}_{i+1}$.

We can use equation (31) to re-write equation (30) as

$$\underset{\sim}{Y}_j = ||\underset{\sim}{E}_j||\underset{\sim}{X}_{i+1}+[\underset{\sim}{Y}_j\underset{\sim}{Z}_i']\underset{\sim}{Z}_i \tag{32}$$

from which it is apparent that $\underset{\sim}{Y}$ is expressible as a linear function of $\underset{\sim}{X}$. In fact, $\underset{\sim}{Y} = \underset{\sim}{S}^+\underset{\sim}{X}$, where $\underset{\sim}{S}^+$ is given by equation (20).

However, it is not necessary to actually compute $\underset{\sim}{S}^+$ from equation (20) because the elements of $\underset{\sim}{S}^+$ have been defined in the process of generating the orthonormal matrix $\underset{\sim}{X}$. In order to demonstrate this fact, we let $\underset{\sim}{S}_j^+$ denote the jth row of matrix $\underset{\sim}{S}^+$. From equations (27) and (28), we can write the first row of $\underset{\sim}{S}^+$ as

$$\underset{\sim}{S}_1^+ = [||\underset{\sim}{E}_1||,0,0,0,\ldots,0], \tag{33}$$

where $\underset{\sim}{S}_1^+$ is a 1×n matrix.

Let the 1×n matrix $\underset{\sim}{Q}_i(j)$ be defined as

$$\underset{\sim}{Q}_i(j) = [0,0,0,\ldots,0,\Delta_j,0,\ldots,0] \tag{34}$$

where Δ_j, defined by

$$\begin{aligned} \Delta_j &= ||\underset{\sim}{E}_j|| \quad \text{if } (\underset{\sim}{E}_j\underset{\sim}{E}_j')/(\underset{\sim}{Y}_j\underset{\sim}{Y}_j')>\varepsilon_j \\ \Delta_j &= 0, \quad \text{otherwise} \end{aligned} \tag{35}$$

is found in the ith column.

Now, using equations (32) and (34), we generate successive rows of $\underset{\sim}{S}^+$ from the relation

$$\underset{\sim}{S}_j^+ = [(\underset{\sim}{Y}_j\underset{\sim}{Z}_i'),0,0,0,\ldots,0]+\underset{\sim}{Q}_{i+1}(j). \tag{36}$$

The matrix $\underset{\sim}{S}^+$ is formed by executing equation (36) p-1 times, beginning with j=2 and i=1. After equation (36) has been executed, the subscript i is incremented by one, only if $\underset{\sim}{Q}_{i+1}(j)$ is non-trivial. The subscript j is <u>always</u> incremented by one after this decision.

Thus, given $\underset{\sim}{Y}$, we have computed the orthonormal matrix $\underset{\sim}{X}$ and the <u>inverse</u> filter $\underset{\sim}{S}^+$. We now show that $\underset{\sim}{S}$ can easily be determined.

Using equation (31), we can write equation (30) as

$$\underset{\sim}{X}_{i+1} = [\underset{\sim}{Y}_j - (\underset{\sim}{Y}_j \underset{\sim}{Z}_i')\underset{\sim}{Z}_i]/||\underset{\sim}{E}_j||, \quad \text{if } (\underset{\sim}{E}_j \underset{\sim}{E}_j')/(\underset{\sim}{Y}_j \underset{\sim}{Y}_j') > \varepsilon_j. \tag{37}$$

Let s_{ij}^+, i=1,p, j=1,n denote the elements of $\underset{\sim}{S}^+$, and $\underset{\sim}{S}_i$ denote the ith row of $\underset{\sim}{S}$. Using (20), we note that equation (37) can be expressed as

$$\underset{\sim}{X}_{i+1} = \left\{ \underset{\sim}{Y}_j - [s_{j1}^+, s_{j2}^+, s_{j3}^+, \ldots, s_{ji}^+] \begin{bmatrix} \underset{\sim}{X}_1 \\ \underset{\sim}{X}_2 \\ \underset{\sim}{X}_3 \\ \cdot \\ \cdot \\ \cdot \\ \underset{\sim}{X}_i \end{bmatrix} \right\} / ||E_j||. \tag{38}$$

Also, it is easy to see that

$$\underset{\sim}{X}_i = \underset{\sim}{S}_i \underset{\sim}{Y}, \quad i=1,n. \tag{39}$$

Let the $1 \times p$ matrix $\underset{\sim}{U}_j$ be defined as

$$\underset{\sim}{U}_j = [0,0,0,\ldots,0,1,0,\ldots,0], \tag{40}$$

where the unity element is found in the jth column.

From equations (27) and (28), it is clear that $\underset{\sim}{S}_1$ can be expressed as

$$\underset{\sim}{S}_1 = [1,0,0,0,\ldots,0]/||\underset{\sim}{E}_1|| = \underset{\sim}{U}_1/||\underset{\sim}{E}_1||. \tag{41}$$

Using equations (40) and (39), with equation (38), we have

$$\underset{\sim}{X}_{i+1} = [\underset{\sim}{U}_j - s_{j1}^+ \underset{\sim}{S}_1 - s_{j2}^+ \underset{\sim}{S}_2 - s_{j3}^+ \underset{\sim}{S}_3 - \ldots - s_{ji}^+ \underset{\sim}{S}_i]\underset{\sim}{Y}/||\underset{\sim}{E}_j||. \tag{42}$$

Again, using (39), it is clear that equation (42) becomes

$$\underset{\sim}{S}_{i+1} = [\underset{\sim}{U}_j - s^+_{j1}\underset{\sim}{S}_1 - s^+_{j2}\underset{\sim}{S}_2 - s^+_{j3}\underset{\sim}{S}_3 - \ldots - s^+_{ji}\underset{\sim}{S}_i]/||\underset{\sim}{E}_j||, \quad \text{if } (\underset{\sim}{E}_j\underset{\sim}{E}_j')/(\underset{\sim}{Y}_j\underset{\sim}{Y}_j') > \varepsilon_j \tag{43}$$

The rows of $\underset{\sim}{S}$ are generated by considering equation (43) p-1 times, beginning with j=2 and i=1. Equation (43) is only executed in the event that the indicated condition is true. Each time the equation has been considered, the subscript j is incremented by one. The subscript i is incremented by one only following actual execution of the equation.

We note from equation (43) that all columns of $\underset{\sim}{S}$ numbered k, such that $(\underset{\sim}{E}_k\underset{\sim}{E}_k')/(\underset{\sim}{Y}_k\underset{\sim}{Y}_k') \leq \varepsilon_k$ will be trivial. In fact, it can be shown that exactly p-n columns of $\underset{\sim}{S}$ will be trivial. This feature will be important in later analysis and deserves further consideration. We now ask the reader to re-examine equation (18) in view of the fact that certain of the columns of $\underset{\sim}{S}$ may be zero. If the kth column of $\underset{\sim}{S}$ is trivial, we may reason that the kth row of $\underset{\sim}{Y}$ is "blocked" by the filter, i.e., it has no projection into the vector space spanned by $\underset{\sim}{X}$. Let $\underset{\sim}{B}$ denote those rows of $\underset{\sim}{Y}$ which are blocked in this manner and $\underset{\sim}{P}$ those rows which are not blocked (passed). We observe that equations (30) and (31) determine whether a particular row of $\underset{\sim}{Y}$ is blocked or passed. In equation (30), $\underset{\sim}{E}_j$ represents that component of $\underset{\sim}{Y}_j$ which is orthogonal to $\underset{\sim}{Z}_i$. If $\underset{\sim}{E}_j$ is not significant it cannot contribute to the set of basis functions $\underset{\sim}{x}(t)$ and $\underset{\sim}{Y}_j$ is blocked. From this we conclude that the space spanned by $\underset{\sim}{B}$ is a subspace of the space spanned by $\underset{\sim}{X}$, and that $\underset{\sim}{X}$ and $\underset{\sim}{P}$ span the same vector space, i.e., $\underset{\sim}{X} = \underset{\sim}{D}^{-1}\underset{\sim}{P}$, where $\underset{\sim}{D}^{-1}$ is an appropriate n×n non-singular matrix. Further, we conclude that $\underset{\sim}{P}$ is a maximum set of linearly independent rows of $\underset{\sim}{Y}$.

For clarity, we restate the filter generating equations together as

$$\underset{\sim}{E}_j = \underset{\sim}{Y}_j - [\underset{\sim}{Y}_j \underset{\sim}{Z}_i'] \underset{\sim}{Z}_i, \tag{44}$$

$$\underset{\sim}{X}_{i+1} = \underset{\sim}{E}_j / ||\underset{\sim}{E}_j||, \quad \text{if } (\underset{\sim}{E}_j \underset{\sim}{E}_j')/(\underset{\sim}{Y}_j \underset{\sim}{Y}_j') > \varepsilon_j, \tag{45}$$

$$\underset{\sim}{S}_j^+ = [(\underset{\sim}{Y}_j \underset{\sim}{Z}_i'),0,0,0,\ldots,0] + \underset{\sim}{Q}_{i+1}(j), \tag{46}$$

and

$$\underset{\sim}{S}_{i+1} = [\underset{\sim}{U}_j - \sum_{k=1}^{i} s_{jk}^+ \underset{\sim}{S}_k] / ||\underset{\sim}{E}_j||, \qquad \text{if } (\underset{\sim}{E}_j \underset{\sim}{E}_j')/(\underset{\sim}{Y}_j \underset{\sim}{Y}_j') > \varepsilon_j, \tag{47}$$

where $\underset{\sim}{Q}_i(j)$ is given by equation (34) and $\underset{\sim}{U}_j$ is given by equation (40). The set of equations (44) through (47) is executed p-1 times, beginning with j=2 and i=1. Each time the set of equations is executed, the subscript j is incremented by one. Equations (45) and (47) are only defined if the indicated condition is true. Following consideration of equation (47), the subscript i is incremented by one, if the indicated condition was found to be true.

We note, contrary to Bingulac [23] that it is not actually necessary to generate the orthonormal matrix $\underset{\sim}{X}$ in order to determine the matrices $\underset{\sim}{S}$ and $\underset{\sim}{S}^+$. This fact is easy to demonstrate if we let the i×p matrix $\underset{\sim}{T}_i$, i=1,n denote the first i rows of $\underset{\sim}{S}$. In this event, we write

$$\underset{\sim}{Z}_i = \underset{\sim}{T}_i \underset{\sim}{Y}, \tag{48}$$

from which $\underset{\sim}{Y}_j \underset{\sim}{Z}_i'$ of equations (44) and (46) becomes

$$Y_j Z_i' = [Y_j Y'] T_i'. \tag{49}$$

Further, we can show that

$$E_j E_j' = Y_j Y_j' - [Y_j Y'] T_i' T_i [Y_j Y']'. \tag{50}$$

Clearly, the algorithm is able to proceed exactly as in equations (44) through (46), with appropriate substitions in lieu of definition of X as in equation (45).

It is evident from equations (49) and (50) that the filter is defined completely by transformations on the Gram matrix YY'. This conclusion allows extremely efficient digital realization of the algorithm (on the order of 25-30 executable statements). Also, this fact leads to a straightforward definition of the algorithm for the problem of adaptive filtering, since YY' is easily computed as observations are added to the set.

Finally, we state, without proof, the interesting fact that

$$S'S = [YY']^+, \tag{51}$$

i.e., the matrix $S'S$ is a generalized inverse of the given matrix YY'.

COMPUTATIONAL ERROR

In equation (31) we proposed a relation for determination of linear dependence, which we now restate as

$$b = (E_j E_j')/(Y_j Y_j'). \tag{52}$$

From equation (30) and the propositions of ordinary least squares, it appears that

$$0 \leq b \leq 1. \tag{53}$$

If b=0, we would certainly conclude that $\underset{\sim}{Y}_j$ is linearly dependent on $\underset{\sim}{Z}_i$ of equation (30) and should be "blocked" by the filter. Likewise, if b=1, we would conclude that $\underset{\sim}{Y}_j$ should be "passed." However, computational error is present in any algorithm and it is quite likely in practice that

$$0 < b < 1 \tag{54}$$

for any case.

We are, therefore, forced to choose some small number ε_j such that the condition that $\underset{\sim}{E}_j$ be blocked is

$$b \leqq \varepsilon_j. \tag{55}$$

Unfortunately, it is not possible to give any general rules for determination of ε_j, since any such rules are related to arithmetic precision and problem magnitude in a very complex way. However, we note emphatically that a choice of ε_j which is too small (e.g., zero) can lead to the result that a vector consisting of computational error is added to the orthonormal set.

The authors have been primarily concerned with the digital implementation of the algorithm for point functions using 16 digit arithmetic. For the type of problems at hand and for $\underset{\sim}{Y}$ of order 10×100 the authors have had no problems for $\varepsilon_j \approx 10^{-10}$.

NOISE

In this discussion, we will assume that $\underset{\sim}{F}$ is a noise-free "signal matrix" consisting of a set of p-vectors such that

$$\underset{\sim}{F} = [\underset{\sim}{f}(t_1),\underset{\sim}{f}(t_2),\underset{\sim}{f}(t_3),\ldots,\underset{\sim}{f}(t_r)]. \quad (56)$$

We also let $\underset{\sim}{V}$ be a "noise matrix" consisting of a set of p-vectors such that

$$\underset{\sim}{V} = [\underset{\sim}{v}(t_1),\underset{\sim}{v}(t_2),\underset{\sim}{v}(t_3),\ldots,\underset{\sim}{v}(t_r)]. \quad (57)$$

We allow that neither $\underset{\sim}{F}$ nor $\underset{\sim}{V}$ is known explicitly, but that we have observed the matrix $\underset{\sim}{Y}$, where the format of $\underset{\sim}{Y}$ is given by equation (2) such that

$$\underset{\sim}{Y} = \underset{\sim}{F}+\underset{\sim}{V}. \quad (58)$$

Now, we again let $\underset{\sim}{X}$ be the output of a Schmidt filter and let the relationship between $\underset{\sim}{Y}$ and $\underset{\sim}{X}$ be indicated by Figure 1.

We would like the row dimension of the orthonormal function $\underset{\sim}{X}$ to be equal to $n=\text{rank}(\underset{\sim}{F})$. However, if $\underset{\sim}{V}$ is chosen completely at random, it is quite likely that

$$\text{rank}(\underset{\sim}{Y}) = p \quad (59)$$

where $p \geqq n$. (60)

From previous discussion, however, it can be seen that the Schmidt filter will produce an orthonormal matrix of rank=p for the case where ε_j, j=1,p are chosen to account only for round-off error.

In view of this difficulty, there is a certain temptation to choose the ε_j substantially larger, e.g., 10^{-2}, to allow for approximate dependence in the sense of least squares. In this way it is certainly possible to reduce the rank of the filter output $\underset{\sim}{X}$. Unfortunately, it is possible, using this procedure, to produce an $\underset{\sim}{X}$ such that $\text{rank}(\underset{\sim}{X}) < \text{rank}(\underset{\sim}{F})$. We now consider this method of analysis and some of the problems which can arise in its wake.

We now let

$$d = (1-b)\cdot 100\% \tag{61}$$

where b is given by equation (52).

In equation (61), d can be interpreted as the percent of the mean-square of $\underset{\sim}{Y}_j$ which will be realized by the restoring filter $\underset{\sim}{S}^+$ if $\underset{\sim}{E}_j$ is judged as insignificant, i.e., equation (55) is true. As an example, consider that we let $\varepsilon_j = .01, \forall j$ and performed the filtering process. In this event, we could be sure that at least 99% of the mean square of each row of $\underset{\sim}{Y}$ would be realized by the restoring filter, or that the mean square of the restoration error of $\underset{\sim}{Y}_j$ would not be greater than 1% of the mean square of $\underset{\sim}{Y}_j$. Choices of ε_j in this type of analysis are generally governed by what is "acceptable" to the investigator as far as restoration is concerned. Although this type of analysis is intuitively appealing, there are some disadvantages that must be made clear.

First of all, "percent mean square recovery" is not always as good an indicator as it might seem and if this method of analysis is used, it is generally advisable to compare the "recovered function" closely with the original. Secondly, since the choice of ε_j is arbitrary, we still have no way of knowing whether $\text{rank}(\underset{\sim}{X}) = \text{rank}(\underset{\sim}{F})$ and hence whether or not a complete set of basis functions is represented in $\underset{\sim}{X}$. Although the authors advise that great caution be exercised in the use of this method, we are impressed by its simplicity and recommend its use where signal and noise statistics are essentially unknown.

In the event that sample noise and signal statistics are available, i.e., $\underset{\sim}{V}\underset{\sim}{V}'$ and $\underset{\sim}{F}\underset{\sim}{V}'$ are specified, the above procedure certainly is not the "best" in a statistical sense. The treatment of this class of function is beyond the scope of the current presentation. However, this problem has been solved using a modified Schmidt filter and will be the topic of a forthcoming paper.

SYSTEM IDENTIFICATION

We now illustrate a novel application of the Schmidt filter to the problem of creating a minimum-order mathematical model of a linear, discrete-time system from <u>noise-free</u> operating data. The problems involved in applying this method to the case where observation noise of unknown statistics is present should be evident from previous discussion.

Consider the linear, discrete, autonomous[3] minimal realization $\underset{\sim}{\Sigma}$ described by

$$\underset{\sim}{x}(k+1) = \underset{\sim}{A}\underset{\sim}{x}(k) \tag{62}$$

$$\underset{\sim}{y}(k) = \underset{\sim}{C}\underset{\sim}{x}(k) \tag{63}$$

---

[3]Application of the Schmidt filter to systems with inputs is discussed extensively in ref. [29].

where $\underset{\sim}{x}(k)$ is a n×1 state vector, $\underset{\sim}{y}(k)$ is a p×1 output vector, and $\underset{\sim}{A}$ and $\underset{\sim}{C}$ are constant matrices of orders n×n and p×n, respectively. As usual, n indicates the "order" of the system and k indicates reference to the system at the beginning of the kth equal interval of time.

Equations (62) and (63) are also known as the "internal description" of $\underset{\sim}{\Sigma}$. We will assume, however, that our only knowledge of $\underset{\sim}{\Sigma}$ consists of an "external description" given by the sequence

$$\underset{\sim}{Y} = [\underset{\sim}{y}(1), \underset{\sim}{y}(2), \underset{\sim}{y}(3), \ldots, \underset{\sim}{y}(N+r-1)] \tag{64}$$

where the meaning of N and r will become evident later.

Given the external description of $\underset{\sim}{\Sigma}$ we wish to determine any equivalent internal description $\underset{\sim}{\Theta}$, defined by

$$\underset{\sim}{q}(k+1) = \underset{\sim}{F}\underset{\sim}{q}(k) \tag{65}$$

and

$$\underset{\sim}{y}(k) = \underset{\sim}{H}\underset{\sim}{q}(k) \tag{66}$$

where

$$\underset{\sim}{q}(k) = \underset{\sim}{P}^{-1}\underset{\sim}{x}(k), \tag{67}$$

$$\underset{\sim}{F} = \underset{\sim}{P}^{-1}\underset{\sim}{A}\underset{\sim}{P}, \tag{68}$$

and

$$\underset{\sim}{H} = \underset{\sim}{C}\underset{\sim}{P} \tag{69}$$

where $\underset{\sim}{P}^{-1}$ is any non-singular n×n constant matrix. Kalman [30] showed that all equivalent minimal realizations are completely observable. We now state the condition of observability as

$$\text{rank}(\underset{\sim}{K}) = n, \tag{70}$$

where

$$\underset{\sim}{K} = \begin{bmatrix} \underset{\sim}{C} \\ \underset{\sim}{C}\underset{\sim}{A} \\ \underset{\sim}{C}\underset{\sim}{A}^2 \\ \cdot \\ \cdot \\ \cdot \\ \underset{\sim}{C}\underset{\sim}{A}^{N-1} \end{bmatrix}, \forall N \geq n. \qquad (71)$$

Associated with the minimal realization $\underset{\sim}{\Sigma}$ we postulate the existence of the sequence of states $\underset{\sim}{X}$, defined by

$$\underset{\sim}{X} = [\underset{\sim}{x}(1),\underset{\sim}{x}(2),\underset{\sim}{x}(3),\ldots,\underset{\sim}{x}(r)], r>n, \qquad (72)$$

where

$$\text{rank}(\underset{\sim}{X}) = n. \qquad (73)$$

We can show that equations (67) and (71) imply that

$$\text{rank}(\underset{\sim}{W}) = n \qquad (74)$$

where

$$\underset{\sim}{W} = \underset{\sim}{K}\underset{\sim}{X}. \qquad (75)$$

The $Np \times 1$ matrix $\underset{\sim}{W}$ can be expressed as a set of $p \times 1$ elementary vectors where

$$(\underset{\sim}{W})_{ij} = \underset{\sim}{C}\underset{\sim}{A}^{i-1}\underset{\sim}{x}(j),\ i=1,N;\ j=1,r. \qquad (76)$$

Using equations (62), (63), (71), (72), and (76) with (75), we find that

$$\underset{\sim}{W} = \begin{bmatrix} \underset{\sim}{y}(1) & \underset{\sim}{y}(2) & \underset{\sim}{y}(3) & \cdots & \underset{\sim}{y}(r) \\ \underset{\sim}{y}(2) & \underset{\sim}{y}(3) & \underset{\sim}{y}(4) & \cdots & \underset{\sim}{y}(r+1) \\ \underset{\sim}{y}(3) & \underset{\sim}{y}(4) & \underset{\sim}{y}(5) & \cdots & \underset{\sim}{y}(r+2) \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \underset{\sim}{y}(N) & \underset{\sim}{y}(N+1) & \underset{\sim}{y}(N+2) & \cdots & \underset{\sim}{y}(r+N-1) \end{bmatrix} \qquad (77)$$

where it appears that $\underset{\sim}{W}$ is known from the external description. We can alternately express $\underset{\sim}{W}$ as a sequence of Np×l vectors, given by

$$\underset{\sim}{W} = [\underset{\sim}{w}(1), \underset{\sim}{w}(2), \underset{\sim}{w}(3), \ldots, \underset{\sim}{w}(r)]. \tag{78}$$

Let $\underset{\sim}{Q}$ be a sequence of n×l state vectors of the equivalent system $\underset{\sim}{\Theta}$, defined by

$$\underset{\sim}{Q} = [\underset{\sim}{q}(1), \underset{\sim}{q}(2), \underset{\sim}{q}(3), \ldots, \underset{\sim}{q}(r)]. \tag{79}$$

From equation (67), $\underset{\sim}{Q}$ and $\underset{\sim}{X}$ are related by

$$\underset{\sim}{Q} = \underset{\sim}{P}^{-1}\underset{\sim}{X} \tag{80}$$

so apparently

$$\text{rank}(\underset{\sim}{Q}) = n. \tag{81}$$

Now, since $\underset{\sim}{W}$ given by equation (75) is formed as a linear function of $\underset{\sim}{X}$, it follows that we can find a n×NP matrix $\underset{\sim}{S}$ such that

$$\underset{\sim}{Q} = \underset{\sim}{S}\underset{\sim}{W}. \tag{82}$$

We now employ the Schmidt filter to determine $\underset{\sim}{S}$ and $\underset{\sim}{Q}$ with the rationale that $\underset{\sim}{Q}$ is a basis of $\underset{\sim}{W}$.

For clarity in later analysis, we let

$$\underset{\sim}{W}_1 = [\underset{\sim}{w}(1), \underset{\sim}{w}(2), \underset{\sim}{w}(3), \ldots \underset{\sim}{w}(r-1)], \tag{83}$$

$$\underset{\sim}{W}_2 = [\underset{\sim}{w}(2), \underset{\sim}{w}(3), \underset{\sim}{w}(4), \ldots \underset{\sim}{w}(r)], \tag{84}$$

$$\underset{\sim}{Q}_1 = [\underset{\sim}{q}(1), \underset{\sim}{q}(2), \underset{\sim}{q}(3), \ldots \underset{\sim}{q}(r-1)], \tag{85}$$

and

$$\underset{\sim}{Q}_2 = [\underset{\sim}{q}(2), \underset{\sim}{q}(3), \underset{\sim}{q}(4), \ldots \underset{\sim}{q}(r)]. \tag{86}$$

The matrices $\underset{\sim}{S}$, $\underset{\sim}{S}^+$, and $\underset{\sim}{Q}_1$ are determined by filtering $\underset{\sim}{W}_1$, as shown in Figure 2.

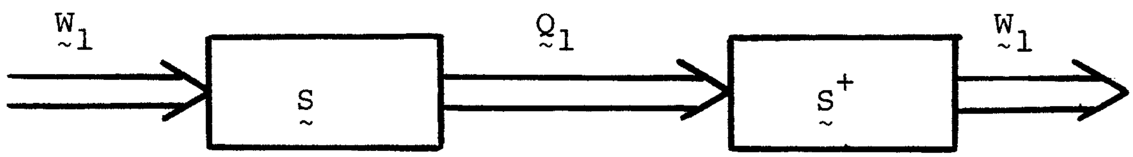

Fig. 2 Use of the Schmidt Filter for System Identification

We now form Q_2 from[4]

$$Q_2 = SW_2. \tag{87}$$

Using ordinary least squares, we now determine F of equation (65) from

$$F = Q_2 Q_1' [Q_1 Q_1']^{-1}. \tag{88}$$

However, since Q_1 is an orthonormal matrix, equation (88) becomes

$$F = Q_2 Q_1'. \tag{89}$$

The matrix H of equation (66) is simply equal to a sub-matrix of S^+. In fact, H is simply equal to the first p rows of S^+.

Besides being a very convenient method of obtaining a minimal realization of equation (64), the algorithm results in a sequence of states Q_1 which is orthonormal. This feature can be quite valuable with respect to visual inspection

---

[4] Since we already know $q(k)$, k=1,r-1, we only need to determine $Sq(r)$.

of results, especially in the case where additive noise is known to be present. Also, the coefficients of $\underset{\sim}{F}$ are all "independently determined" in the sense of least squares with the results that their associated contributions to regression sums of squares and coefficients of determination are easily computed.

As an interesting sidelight, Rowe [31] has shown that an equivalent "cannonical" description of the system $\underset{\sim}{\Sigma}$ is given by

$$\underset{\sim}{y}(k) = \underset{\sim}{B}_1\underset{\sim}{y}(k-1)+\underset{\sim}{B}_2\underset{\sim}{y}(k-2)+\underset{\sim}{B}_3\underset{\sim}{y}(k-3)+\ldots+\underset{\sim}{B}_L\underset{\sim}{y}(k-L), \tag{90}$$

where $L \leqq N$ and the $\underset{\sim}{B}_i, \forall_i$ are constant $p \times p$ matrices. We can re-write equation (90) as

$$\underset{\sim}{y}(k) = \underset{\sim}{B}\underset{\sim}{z}(k-1), \tag{91}$$

where

$$\underset{\sim}{B} = [\underset{\sim}{B}_1, \underset{\sim}{B}_2, \underset{\sim}{B}_3, \ldots, \underset{\sim}{B}_L], \tag{92}$$

and

$$\underset{\sim}{z}(k-1) = \begin{bmatrix} \underset{\sim}{y}(k-1) \\ \underset{\sim}{y}(k-2) \\ \underset{\sim}{y}(k-3) \\ \cdot \\ \cdot \\ \cdot \\ \underset{\sim}{y}(k-L) \end{bmatrix}. \tag{93}$$

We now partition the matrix $\underset{\sim}{S}$ into $n \times p$ block matrices $\underset{\sim}{S}_i$, $i=1,N$, such that

$$\underset{\sim}{S} = [\underset{\sim}{S}_1, \underset{\sim}{S}_2, \underset{\sim}{S}_3, \ldots, \underset{\sim}{S}_N]. \tag{94}$$

For the case where $L<N$, it is easily determined from the

properties of the Schmidt filter and equation (90) that exactly N-L of the $\underset{\sim}{S}_i$ in equation (94) will be trivial. In fact, it is obvious that the "first" trivial block matrix is $\underset{\sim}{S}_{L+1}$, or mathematically,

$$\underset{\sim}{S}_i \neq \underset{\sim}{0}, i<L+1. \tag{95}$$

Assume that L<N. Let the p×Lp matrix $\underset{\sim}{R}$ be defined by rows numbered Lp+1 through (L+1)p and columns numbered 1 through pL of the matrix $\underset{\sim}{S}^+\underset{\sim}{S}$. Let $\underset{\sim}{R}$ be partitioned into p×p matrices such that

$$\underset{\sim}{R} = [\underset{\sim}{R}_1, \underset{\sim}{R}_2, \underset{\sim}{R}_3, \ldots, \underset{\sim}{R}_L]. \tag{96}$$

We now state, without proof, that the matrices $\underset{\sim}{B}_i$, i=1,L of equation (90) are given in the sense of least squares by the relation

$$\underset{\sim}{B}_i = \underset{\sim}{R}_{L+1-i}, \; i=1,L. \tag{97}$$

Thus, we have indicated how information obtained from the Schmidt filter can be used to identify the system in state form, as in equations (62) and (63), and in the "cannonical" form given by equation (90).

CONCLUDING REMARKS

A "Schmidt orthonormal filter" has been proposed to compute a set of orthonormal basis functions of a set of noisy filter input functions for the case where noise statistics are essentially unknown. Well-defined procedures have been given to compute the transfer function and inverse transfer function of the filter. The utility of the filter has been demonstrated with respect to the problem of identifying discrete, linear systems of unknown order.

As a final note, the authors submit that the problem of order determination and unbiased estimation of system parameters for the case where noise statistics are specified has been solved using a modified Schmidt filter algorithm. This problem will be discussed thoroughly in a forthcoming paper.

ACKNOWLEDGMENTS

The Authors wish to express gratitude to Dr. J. S. Pazdera and Dr. R. T. Johnson for their assistance in preparation of this work.

REFERENCES

1. Kitamori, T., "Applications of Orthogonal Functions to the Determination of Process Dynamic Characteristics and to the Construction of Self-Optimizing Control Systems," Proc. I.F.A.C., Moscow, 1960, pp. 613-617.

2. Wolf, A. A., "A Statistical Theory for Parameter Identification in Physical Systems," J. Franklin Inst., Vol. 274, 1962, pp. 279-296.

3. Lubbock, J. K., "Self-Optimizing Filters Employing Orthogonal Systems of Functions," J. Electron. and Contr., Vol. 13, 1962, p. 365.

4. Clement, P. R., "On Completeness of Basis Functions Used for Signal Analysis," S.I.A.M. Rev., Vol. 5, No. 2, Apr. 1963, pp. 131-139.

5. Lubbock, J. K., and Barker, H. A., "A Solution of the Identification Problem," J.A.C.C. preprints, Minneapolis, 1963, pp. 191-199.

6. Puri, N. N., "Transfer Function Tracking of a Linear Time-Varying System by Means of Orthogonal Networks," I.E.E.E. Trans. Appl. and Ind., July 1963, pp. 234-238.

7. Elkind, J. I., "Application of Multiple Regression Analysis to Identification of Time-Varying Linear Dynamic Systems," I.E.E.E. Trans. Autom. Contr., AC-8, 1963, pp. 163-166.

8. Paul, R. J. A., "Orthogonal Functional Approximations to Linear Dynamic Systems," Proc. I.E.E., Vol. 3, No. 2, Nov. 1964, pp. 1913-1922.

9. Douce, J. L., and Roberts, P. D., "Orthogonal Functions," Control, Vol. 8, Sept. 1964, pp. 457-461.

10. Raevskii, S. Ya., "Orthogonal Representations of Random Processes in Problems of Statistical Dynamics of Automatic Control," Automat. and Rmt. Contr., Vol. 26, No. 6, June 1965, p. 1028.

11. Barker, H. A., and Hawley, D. A., "Orthogonalizing Techniques for System Identification and Optimization," Proc. I.F.A.C., London, 1966, pp. 36A.1-36A.10.

12. Roberts, P. D., "Self-Adjusting Orthogonal Digital Filter for Identification of Unknown Processes," Electron. Lett., Vol. 2, No. 5, May 1966, pp. 187-188.

13. Roberts, P. D., "Self-Adjustable Orthogonal Digital Filters for System Identification and Optimisation," Proc. I.E.E., Vol. 114, No. 5, May 1967, pp. 666-670.

14. Bruni, C., "The Identification of Linear Time-Invariant Systems by Means of Orthonormalized Exponential Series Expansions II. Case of Non-negligible Measurement Errors and Noise," Automaz. E Strum. (Italy), Vol. 16, No. 11, Nov. 1968, pp. 538-543.

15. Fuji, K., and Nakao, K., "Identification of Non-linear Dynamical Systems Through Expansion of Weiner Kernels Into Variable Orthonormal Functions," J. Japan Ass. Autom. Contr. Eng., Vol. 14, No. 10, Oct. 1970, pp. 593-601.

16. Higuchi, S., "Low Order Process Identification Using Orthonormal Functions," Trans. Soc. Instrum. Contr. Eng. (Japan), Vol. 16, No. 5, 1970, pp. 422-428.

17. Aksenov, G. A., and Baglai, R. D., "Identification of Nonlinear Relationships With the Aid of Orthogonal Converters," Autometriya (USSR), Sept.-Oct. 1970, pp. 36-40.

18. Ho, B. L., and Kalman, R. E., "Effective Construction of Linear State-Variable Models From Input/Output Functions," Regelungstechnik, Vol. 14, 1966, pp. 545-548.

19. Gopinath, B., "On the Identification of Linear Time-Invariant Systems From Input-Output Data," Bell System Tech. J., Vol. 48, No. 5, May 1969, pp. 1101-1113.

20. Budin, M. A., "Minimal Realization of Discrete Linear Systems From Input-Output Observations," I.E.E.E. Trans. Auto. Contr., AC-16, No. 5, Oct. 1971, pp. 395-401.

21. Epstein, B., Orthogonal Families of Analytic Functions, MacMillan, N.Y., 1965.

22. Drygas, H., The Coordinate-Free Approach to Gauss-Markov Estimation, Springer-Verlag, Berlin, 1970.

23. Bingulac, S., "The Orthonormalization of Functions," I.E.E.E. Trans. Auto. Contr., AC-10, July 1965, pp. 374-375.

24. Penrose, R., "A Generalized Inverse for Matrices," Proc. Cambridge Phil. Soc., Vol. 51, Pt. 3, 1955, pp. 406-413.

25. Greville, T. N. E., "The Pseudo Inverse of a Rectangular or Singular Matrix and Its Application to the Solution of Systems of Linear Equations," S.I.A.M. Newsletter, Vol. 5, 1957.

26. Rao, C. R., "A Note on the Generalized Inverse of a Matrix With Applications to Problems in Mathematical Sta-

tistics," J. Royal Statistical Soc., Ser. B., Vol. 24, 1962, pp. 152-158.

27. Mayne, D. Q., "An Algorithm for Computing the Pseudo-Inverse of a Singular Matrix," Computer J., Vol. 9, Nov. 1966, pp. 312-317.

28. Sage, A. P., and Melsa, J. L., Estimation Theory With Applications to Communications and Control, McGraw-Hill, N.Y., 1970.

29. Behring, A. G., "Methods and Applications of System Identification," Ph.D. thesis, University of Missouri-Rolla, Rolla, Mo., Dec. 1971.

30. Kalman, R. E., "Mathematical Description of Linear Dynamical Systems," S.I.A.M. J. Contr., Ser. A, Vol. 1, No. 2, 1963, pp. 152-192.

31. Rowe, I. H., "A Bootstrap Method for the Statistical Estimation of Model Parameters," Int. J. Contr., Vol. 12, No. 5, Nov. 1970, pp. 721-738.

MINIMUM-ORDER MATHEMATICAL MODELS OF DISCRETE LINEAR DYNAMIC SYSTEMS

by

A. G. Behring

and

V. J. Flanigan*

ABSTRACT

An orthonormal filter is used to obtain minimum-order, mathematical models of linear, multi-variable, discrete, time-invariant systems from input/output data. Some common problems associated with obtaining such descriptions are considered, including the problem of noise-corrupted observations.

---

*The Authors are associated with the Department of Mechanical and Aerospace Engineering, University of Missouri - Rolla, Rolla, Missouri 65401, where Mr. Behring is a Graduate Student and Dr. Flanigan (member ASME) is an Associate Professor.

NOTATION

In this paper all bold-face capital letters denote matrices. Vectors are defined in column format and are denoted by lower case letters in bold face type. All scalars will be denoted by plain upper or lower case letters. Occasionally it will be necessary to display the format of a vector or matrix explicitly, e.g.,

$$\underset{\sim}{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}$$

Any exceptions to these general rules will be clearly specified in the text.

Numbers in brackets designate references at the end of the paper.

INTRODUCTION

The foundation for much of the current activity in the analysis of linear, multiple input/output systems lies in the early work of Kalman [1].

Kalman [2] introduced an algorithm to compute a "minimal realization" of an impulse response matrix and showed that all such realizations are equivalent corresponding to that part of a system which is controllable and observable.

A new method for computing a minimal realization of an impulse response matrix (using Markov parameters) was later introduced by Ho [3]. Ho [4] extended his impulse response method using an indirect procedure to accommodate the presence of initial conditions and a selected class of inputs.

Apparently the first procedure for directly computing a minimal realization from input/output data for the case of discrete time systems was introduced by Gopinath [5,6] using least squares. Gopinath [5] also considers the realization problem for systems whose inputs and outputs are corrupted by zero mean noise with known statistics and shows the resulting parameter estimates to be consistent. However, the computational method suggested for arriving at numerical estimates of system parameters has proven to be undesirable in that it essentially depends upon the success of a trial and error procedure.

In a recent paper Budin [7] has reduced the original computational method of Gopinath to a deterministic algorithm and formulated certain other labor saving features.

During the last few years several authors have noted the simplifying features associated with the use of orthogonal functions in optimal control and identification problems. Among those who have considered the problem from this viewpoint, the works of Kitamori [8], Lubbock and Barker [9], Barker [10], and Roberts [11] are the most relevent to this paper.

PROBLEM STATEMENT

Let the dynamic equations $\underset{\sim}{\Sigma}$ of a process be given by

$$\underset{\sim}{x}(k+1) = \underset{\sim}{A}\underset{\sim}{x}(k)+\underset{\sim}{B}\underset{\sim}{u}(k) \tag{1}$$

$$\underset{\sim}{y}(k) = \underset{\sim}{C}\underset{\sim}{x}(k)+\underset{\sim}{D}\underset{\sim}{u}(k) \tag{2}$$

where $\underset{\sim}{x}$ is a n×1 state vector, $\underset{\sim}{u}$ is a m×1 input vector, and $\underset{\sim}{y}$ is a p×1 output vector. The constant matrices $\underset{\sim}{A}$, $\underset{\sim}{B}$, $\underset{\sim}{C}$, and $\underset{\sim}{D}$ are of order n×n, n×m, p×n, and p×m, respectively. As usual, the integer k indicates reference to the system at the beginning of the kth equal interval of time.

Equations (1) and (2) are also known as an "internal" description of the process. We will assume, however, that the only knowledge of the system consist of an "external" description given by a sequence of corresponding observations on the input $\underset{\sim}{u}$ and output $\underset{\sim}{y}$. For reference, see Kalman [2], Ho [4], Gopinath [6], and

Budin [7]. The problem is to determine the internal description given the external description.

Equivalence

Two constant, linear, discrete systems will be considered equivalent when

1. Their corresponding state vectors are related by constant non-singular transformations[1]
2. Their input/output descriptions are identical, $\forall k$.

Let $\underset{\sim}{\Sigma}$ and $\underset{\sim}{\Theta}$ be equivalent systems with $\underset{\sim}{x} \varepsilon \underset{\sim}{\Sigma}$, $\underset{\sim}{q} \varepsilon \underset{\sim}{\Theta}$. In particular, let

$$\underset{\sim}{q}(k) = \underset{\sim}{P}^{-1}\underset{\sim}{x}(k), \forall k \tag{3}$$

where $\underset{\sim}{P}^{-1}$ is a constant, non-singular matrix. From conditions 1 and 2 of equivalence, we can now use equations (1), (2), and (3) to formulate $\underset{\sim}{\Theta}$ as:

$$\underset{\sim}{q}(k+1) = \underset{\sim}{F}\underset{\sim}{q}(k)+\underset{\sim}{G}\underset{\sim}{u}(k) \tag{4}$$

$$\underset{\sim}{y}(k) = \underset{\sim}{H}\underset{\sim}{q}(k)+\underset{\sim}{D}\underset{\sim}{u}(k) \tag{5}$$

where

$$\underset{\sim}{F} = \underset{\sim}{P}^{-1}\underset{\sim}{A}\underset{\sim}{P}, \tag{6}$$

$$\underset{\sim}{G} = \underset{\sim}{P}^{-1}\underset{\sim}{B}, \tag{7}$$

and

$$\underset{\sim}{H} = \underset{\sim}{C}\underset{\sim}{P}. \tag{8}$$

---

[1]This condition corresponds to Kalman's [2] definition of "strict" equivalence.

If two systems are equivalent, we will say that their state vectors are equivalent.

Minimal Realizations

A minimal realization $\underset{\sim}{\Lambda}$ of $\underset{\sim}{\Sigma}$ is a system of minimal dimension which duplicates the input/output (external) description of $\underset{\sim}{\Sigma}$ where dim $(\underset{\sim}{\Lambda}) \leqq$ dim $(\underset{\sim}{\Sigma})$. Note that $\underset{\sim}{\Lambda}$ and $\underset{\sim}{\Sigma}$ are not necessarily equivalent. Kalman [2] gave the formal theorems dealing with minimal realizations of impulse response matrices. These theorems were later extended to deal with a more general class of inputs and initial conditions by Ho [4], Gopinath [6], and Budin [7]. Kalman's principle theorems on minimal realizations can be summarized as:

1. Any two minimal realizations of $\underset{\sim}{\Sigma}$ are equivalent.
2. A minimal realization of $\underset{\sim}{\Sigma}$ corresponds only to that portion of $\underset{\sim}{\Sigma}$ which is completely controllable and completely observable.
3. All minimal realizations of $\underset{\sim}{\Sigma}$ are completely controllable and completely observable.

Identification

Examination of pertinent literature appears to yield general agreement that if we were able to obtain the unique internal description of $\underset{\sim}{\Sigma}$ (not one which is merely equivalent) the system would be identified. The above situation might be called identification in a "parametric" sense. In the absence of constraints on parameter values, however, it is generally impossible to obtain such parametric identification, Fisher [12].

Identification, as used herein, will be considered accomplished if we are able to determine a $\underset{\sim}{\Lambda}$ which is equivalent to $\underset{\sim}{\Sigma}$. Thus, it appears that if $\underset{\sim}{\Lambda}$ is a minimal realization of $\underset{\sim}{\Sigma}$ then $\underset{\sim}{\Lambda}$ does not identify $\underset{\sim}{\Sigma}$ unless $\dim(\underset{\sim}{\Lambda}) = \dim(\underset{\sim}{\Sigma})$.

In view of the above statements, Kalman [2] implies that necessary conditions for system identification are those of complete controllability and complete observability. The condition for controllability of $\underset{\sim}{\Sigma}$ is that

$$\operatorname{rank}(\underset{\sim}{J}) = n, \tag{9}$$

where

$$\underset{\sim}{J} = [\underset{\sim}{B}, \underset{\sim}{A}\underset{\sim}{B}, \underset{\sim}{A}^2\underset{\sim}{B}, \ldots, \underset{\sim}{A}^{n-1}\underset{\sim}{B}]. \tag{10}$$

The condition that $\underset{\sim}{\Sigma}$ is observable is that

$$\operatorname{rank}(\underset{\sim}{K}) = n, \tag{11}$$

where

$$\underset{\sim}{K} = \begin{bmatrix} \underset{\sim}{C} \\ \underset{\sim}{C}\underset{\sim}{A} \\ \underset{\sim}{C}\underset{\sim}{A}^2 \\ \cdot \\ \cdot \\ \cdot \\ \underset{\sim}{C}\underset{\sim}{A}^{n-1} \end{bmatrix} \tag{12}$$

From (10) and (12) it is clear that these conditions do not depend upon any particular external description but are uniquely dependent upon system characteristics.

It happens that there are other conditions which, quite apart from the conditions of controllability and observability, determine the success of an attempt at identification. Lee [13] has shown that if initial conditions and/or inputs are insufficient to stimulate all system modes within a given input/output description, the system is not identifiable from that description.[2] It can also be shown that the system input functions must possess sufficient "generality" within a given external description for the system to be identified from that description, Ho [4], Gopinath [6], and Budin [7]. A somewhat obvious case of this malady may occur if one or more of the system inputs are linearly dependent within the external description. An insufficient number of observations of system inputs and outputs can also result in failure of an identification attempt. In this paper, however, it will be assumed that a sufficient number of observations is available for the purpose at hand.

Noise

Four classes of additive noise are frequently considered in the literature on linear systems.

The first class admits the possibility that errors in

---

[2]Lee refers to this condition as n-identifiability.

measurement of input and output variables are present where the statistics of such errors are completely unknown. It can be shown that the classical method of least squares produces asymtotically biased estimates of system parameters for this case. The seriousness of such bias is, however, largely dependent on the particular problem at hand, Rowe [14].

The second involves the assumption that inputs and outputs are subject to corruption by stationary noise processes with known statistics. In this case, Gopinath [5], it is possible to show that a modified least squares procedure results in consistent estimates of system parameters.

The third class of noise considers the possible presence of unknown zero mean, uncorrelated noise which forces the system. Under these conditions, it is possible to show that the least squares procedure produces biased parameter estimates. Rowe [14] investigates this problem using the method of instrumental variables.

The fourth class assumes the presence of colored (serially correlated) noise, Sage and Melsa [15]. This class of noise also introduces asymtotic bias in the least squares process. Classical treatment of this problem involves "extension" of the system state space to include the noise process.

The presence of any type of noise greatly increases the difficulty involved with the identification and realization problems. In fact, under this assumption it is quite likely impossible to construct <u>any</u> realization of $\underset{\sim}{\Sigma}$ much less identify $\underset{\sim}{\Sigma}$ in the sense described above. Where noise is present, the identification problem has been called the optimal identification problem, Lee [13]. Extension to the terminology of optimal realization seems natural.

The problem of noise will not be actively treated in this paper, however extension to the noisy case will be considered. Furthermore, when the presence of noise is admitted, it will be assumed that the noise is of the first type.

LINEAR LEAST SQUARES ESTIMATION

Consider that

$$\underset{\sim}{Y} = [\underset{\sim}{y}(1),\underset{\sim}{y}(2),\underset{\sim}{y}(3),\ldots,\underset{\sim}{y}(r)], \tag{13}$$

and

$$\underset{\sim}{X} = [\underset{\sim}{x}(1),\underset{\sim}{x}(2),\underset{\sim}{x}(3),\ldots,\underset{\sim}{x}(r)] \tag{14}$$

constitute sets of r corresponding discrete time observations of $\underset{\sim}{y}(k)$ and $\underset{\sim}{x}(k)$. Then, it can be shown, Sage and Melsa [15], that an optimal linear conditional estimate of $\underset{\sim}{y}(k)$ given $\underset{\sim}{x}(k)$ in the sense of minimum mean square $\{\underset{\sim}{\hat{y}}(k)|\underset{\sim}{x}(k)\}$ is

$$\underset{\sim}{\hat{y}}(k)|\underset{\sim}{x}(k) = \underset{\sim}{\hat{A}}\underset{\sim}{x}(k), \tag{15}$$

where $\underset{\sim}{\hat{A}}$ is a constant matrix which satisfies

$$\underset{\sim}{\hat{A}}\underset{\sim}{X}\underset{\sim}{X}' = \underset{\sim}{Y}\underset{\sim}{X}'. \tag{16}$$

Additionally, if it so happens that

$$E[\underset{\sim}{\hat{y}}(k)|\underset{\sim}{x}(k)] = \underset{\sim}{y}(k), \forall k, \tag{17}$$

where E{ } denotes the mathematical expectation, then $\hat{y}(k)|x(k)$ is called an unbiased estimate of $y(k)$ given $x(k)$.

THE SCHMIDT FILTER

We now assume the availability of a process, which we will call the Schmidt filter, described below.

Let the matrix Y be defined as the r-sequence of p-vectors

$$Y = [y(1), y(2), y(3), \ldots, y(r)]. \tag{18}$$

Assume that Y has arbitrary rank. If Y appears as the input to the Schmidt filter, the output of the filter will be an r-sequence of n-vectors x, where

$$X = [x(1), x(2), x(3), \ldots, x(r)], \tag{19}$$

such that X is an orthonormal matrix, defined by

$$XX' = I_n \tag{20}$$

where I_n is the n×n identity matrix and where Y and X are related by filter-defined constant linear transformations S and S^+ such that

$$X = SY \tag{21}$$

and

$$Y = S^+X. \tag{22}$$

It can be shown that S and S^+ possess the properties of the generalized inverse:[3]

---

[3]For reference, see Penrose [16] and Rao [17].

$$\tilde{S}\tilde{S}^{+}\tilde{S} = \tilde{S} \tag{23}$$

and

$$\tilde{S}^{+}\tilde{S}\tilde{S}^{+} = \tilde{S}^{+}, \tag{24}$$

where $\tilde{S}^{+}$ denotes the generalized inverse of $\tilde{S}$. Additionally, for the Schmidt filter, it can be shown that

$$\tilde{S}\tilde{S}^{+} = \tilde{I}_n. \tag{25}$$

In the literature, e.g., Drygas [18], Lubbock and Barker [9], and Barker [10], the algorithm which generates $\tilde{X}$ as in equation (19) is called "orthonormalization due to Erhard Schmidt," or the "Gram-Schmidt orthonormalization procedure." Mayne [19] used the Gram-Schmidt procedure to compute the pseudo-inverse of a matrix with allowances for computational round-off error that seem to be applicable to the "noisy" case.

For the sake of economy, the detailed internal structure of the Schmidt filter will not be presented here. A comprehensive discussion of the filter can be found in reference [20]. Some structural properties of the filter, however, are of central importance to the development that follows. These are:

1. The filter is essentially an inversion-free linear least squares process.
2. As the algorithm proceeds, each row of Y is considered for dependence[4] in serial order

---

[4]Specifically, the test for dependence consists of a

with respect to any currently-defined rows of $\underset{\sim}{X}$.

3. If a row of $\underset{\sim}{Y}$ is judged as linearly independent as in 2, the associated vector of residuals is divided by its norm and enters as the next row of $\underset{\sim}{X}$.
4. If a row of $\underset{\sim}{Y}$ is judged as linearly dependent, no new rows of $\underset{\sim}{X}$ are formed and consideration moves to the next row of $\underset{\sim}{Y}$. In this event, the column of $\underset{\sim}{S}$ corresponding to this row is set uniformly equal to zero, since this row contributes no "new" information to $\underset{\sim}{X}$.
5. If the presence of "noise" is assumed <u>and accounted for</u> (see footnote 4) then equation (22) may only be approximately true, otherwise no approximation is involved.

THE SELECTOR MATRIX

Consider that it is desired to form a matrix $\underset{\sim}{N}$ from certain rows of a matrix $\underset{\sim}{M}$. This activity can be accomplished conveniently in a mathematical sense by premultiplying $\underset{\sim}{M}$ by a so-called selector matrix, say $\underset{\sim}{T}$, and setting $\underset{\sim}{N}$ equal to the result, i.e.,

---

comparison of the ratio (residual mean square/total mean square) to a prespecified "noise/signal ratio."

$$\underset{\sim}{N} = \underset{\sim}{T}\underset{\sim}{M}.$$

The structure of the selector matrix has been thoroughly discussed by Gopinath [6] and Budin [7]. Basically, each row of $\underset{\sim}{T}$ can have but one non-zero element per row, equal to 1, in its columns which correspond to the selected rows of $\underset{\sim}{M}$. To be useful in the following development we must also require that the selector matrix have a maximum of one non-zero element per column (i.e., no two rows of $\underset{\sim}{N}$ can be equal to a single row of $\underset{\sim}{M}$).

STRUCTURAL RELATIONSHIPS

The structural theorems of Gopinath [6] are now considered from a unified point of view and extended to account for the possible presense of passive system elements.

We begin by assuming that $\underset{\sim}{\Sigma}$ described by equations (1) and (2) is a minimal realization of an external description of a linear system $\underset{\sim}{\Theta}$. The extent to which $\underset{\sim}{\Sigma}$ identifies $\underset{\sim}{\Theta}$ will be discussed later.

In association with the minimal realization $\underset{\sim}{\Sigma}$ we now postulate the existence of an r-sequence of states

$$\underset{\sim}{X} = [\underset{\sim}{x}(1),\underset{\sim}{x}(2),\underset{\sim}{x}(3),\ldots,\underset{\sim}{x}(r)] \qquad (26)$$

where

$$r>n+m \qquad (27)$$

and

$$\text{rank }(\underset{\sim}{X}) = n. \qquad (28)$$

Since the minimal realization $\underset{\sim}{\Sigma}$ is observable, we can now assume that

$$\text{rank } (\underset{\sim}{K}) = n, \tag{29}$$

where

$$\underset{\sim}{K} = \begin{bmatrix} \underset{\sim}{C} \\ \underset{\sim}{C}\underset{\sim}{A} \\ \underset{\sim}{C}\underset{\sim}{A}^2 \\ \cdot \\ \cdot \\ \cdot \\ \underset{\sim}{C}\underset{\sim}{A}^{N-1} \end{bmatrix} \tag{30}$$

for all N such that

$$N \geqq n. \tag{31}$$

Using (28) and (29) we can show that the realization must satisfy the condition

$$\text{rank } (\underset{\sim}{K}\underset{\sim}{X}) = n, \tag{32}$$

where in terms of elementary p×1 vectors

$$[\underset{\sim}{K}\underset{\sim}{X}]_{ij} = \underset{\sim}{C}\underset{\sim}{A}^{i-1}\underset{\sim}{x}(j),\ i=1,N;\ j=1,r. \tag{33}$$

Equation (33) can be used to show that

$$\underset{\sim}{K}\underset{\sim}{X} = \underset{\sim}{Z}-\underset{\sim}{E}\underset{\sim}{V}, \tag{34}$$

where

$$\underset{\sim}{Z} = \begin{bmatrix} \underset{\sim}{y}(1) & \underset{\sim}{y}(2) & \underset{\sim}{y}(3) & \cdots & \underset{\sim}{y}(r) \\ \underset{\sim}{y}(2) & \underset{\sim}{y}(3) & \underset{\sim}{y}(4) & \cdots & \underset{\sim}{y}(r+1) \\ \underset{\sim}{y}(3) & \underset{\sim}{y}(4) & \underset{\sim}{y}(5) & \cdots & \underset{\sim}{y}(r+2) \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \underset{\sim}{y}(N) & \underset{\sim}{y}(N+1) & \underset{\sim}{y}(N+2) & \cdots & \underset{\sim}{y}(N+r-1) \end{bmatrix}, \tag{35}$$

$$\underset{\sim}{V} = \begin{bmatrix} \underset{\sim}{u}(1) & \underset{\sim}{u}(2) & \underset{\sim}{u}(3) & \cdots & \underset{\sim}{u}(r) \\ \underset{\sim}{u}(2) & \underset{\sim}{u}(3) & \underset{\sim}{u}(4) & \cdots & \underset{\sim}{u}(r+1) \\ \underset{\sim}{u}(3) & \underset{\sim}{u}(4) & \underset{\sim}{u}(5) & \cdots & \underset{\sim}{u}(r+2) \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \underset{\sim}{u}(N) & \underset{\sim}{u}(N+1) & \underset{\sim}{u}(N+2) & \cdots & \underset{\sim}{u}(N+r-1) \end{bmatrix}, \tag{36}$$

and

$$\underset{\sim}{E} = \begin{bmatrix} \underset{\sim}{D} & \underset{\sim}{0} & \underset{\sim}{0} & \cdots & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{C}\underset{\sim}{B} & \underset{\sim}{D} & \underset{\sim}{0} & \cdots & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{C}\underset{\sim}{A}\underset{\sim}{B} & \underset{\sim}{C}\underset{\sim}{B} & \underset{\sim}{D} & \cdots & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{C}\underset{\sim}{A}^2\underset{\sim}{B} & \underset{\sim}{C}\underset{\sim}{A}\underset{\sim}{B} & \underset{\sim}{C}\underset{\sim}{B} & \cdots & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \underset{\sim}{D} & \underset{\sim}{0} & \underset{\sim}{0} \\ \cdot & \cdot & \cdot & \cdots & \underset{\sim}{C}\underset{\sim}{B} & \underset{\sim}{D} & \underset{\sim}{0} \\ \underset{\sim}{C}\underset{\sim}{A}^{N-2}\underset{\sim}{B} & \underset{\sim}{C}\underset{\sim}{A}^{N-3}\underset{\sim}{B} & \underset{\sim}{C}\underset{\sim}{A}^{N-4}\underset{\sim}{B} & \cdots & \underset{\sim}{C}\underset{\sim}{A}\underset{\sim}{B} & \underset{\sim}{C}\underset{\sim}{B} & \underset{\sim}{D} \end{bmatrix}. \tag{37}$$

Now, from (32) it is evident that we can always select n linearly independent rows from $\underset{\sim}{H}\underset{\sim}{X}$, say $\underset{\sim}{Q}$, where $\underset{\sim}{Q}$ is an $n \times r$ matrix of the selected rows. Using the concept of a selector matrix, we can symbolically represent this action as

$$\underset{\sim}{Q} = \underset{\sim}{T}\underset{\sim}{K}\underset{\sim}{X}, \tag{38}$$

where $\underset{\sim}{T}$ is an appropriate $n \times Np$ selector matrix. Clearly, however, the operation described in (38) is equivalent to

$$\underset{\sim}{Q} = \underset{\sim}{P}^{-1}\underset{\sim}{X}, \tag{39}$$

where $\underset{\sim}{P}^{-1}$ is a $n \times n$ non-singular matrix given by

$$\underset{\sim}{P}^{-1} = \underset{\sim}{T}\underset{\sim}{K}. \tag{40}$$

The matrix $\underset{\sim}{Q}$ can be written as

$$\underset{\sim}{Q} = [\underset{\sim}{q}(1),\underset{\sim}{q}(2),\underset{\sim}{q}(3),\ldots,\underset{\sim}{q}(r)], \tag{41}$$

where the $\underset{\sim}{q}(k)$, $k=1,r$ are $n\times 1$ vectors.

Using the fact that $\underset{\sim}{P}^{-1}$ is non-singular, we can write

$$\underset{\sim}{q}(k) = \underset{\sim}{P}^{-1}\underset{\sim}{x}(k), \forall k. \tag{42}$$

It should be clear now from the previous discussion of equivalence that the $\underset{\sim}{q}(k),\forall k$ are equivalent state vectors.

Using (42) with equations (1) and (2) we can now proceed to construct an equivalent minimal realization as in equations (4) through (8). We now write $\underset{\sim}{Z}$ as the sequence of $Np\times 1$ vectors

$$\underset{\sim}{Z} = [\underset{\sim}{z}(1),\underset{\sim}{z}(2),\underset{\sim}{z}(3),\ldots,\underset{\sim}{z}(r)], \tag{43}$$

and $\underset{\sim}{V}$ as the sequence of $Nm\times 1$ vectors

$$\underset{\sim}{V} = [\underset{\sim}{v}(1),\underset{\sim}{v}(2),\underset{\sim}{v}(3),\ldots,\underset{\sim}{v}(r)]. \tag{44}$$

Using equations (26), (34), (38), (43), and (44), we can write

$$\underset{\sim}{q}(k) = \underset{\sim}{T}[\underset{\sim}{z}(k)-\underset{\sim}{E}\underset{\sim}{v}(k)],\ k=1,r. \tag{45}$$

Substitution of equation (45) into equations (4) and (5) gives

$$\underset{\sim}{T}[\underset{\sim}{z}(k+1)-\underset{\sim}{E}\underset{\sim}{v}(k+1)] = \underset{\sim}{F}\underset{\sim}{T}[\underset{\sim}{z}(k)-\underset{\sim}{E}\underset{\sim}{v}(k)]+\underset{\sim}{G}\underset{\sim}{u}(k) \tag{46}$$

and

$$\underset{\sim}{y}(k) = \underset{\sim}{H}\underset{\sim}{T}[\underset{\sim}{z}(k)-\underset{\sim}{E}\underset{\sim}{v}(k)]+\underset{\sim}{D}\underset{\sim}{u}(k), \tag{47}$$

where $\underset{\sim}{F}$, $\underset{\sim}{G}$, and $\underset{\sim}{H}$ are defined by equations (6), (7), and (8), respectively.

Equation (46) can be rearranged to give

$$\underset{\sim}{T}\underset{\sim}{z}(k+1) = \underset{\sim}{F}\underset{\sim}{T}\underset{\sim}{z}(k)-\underset{\sim}{F}\underset{\sim}{T}\underset{\sim}{E}\underset{\sim}{v}(k)+\underset{\sim}{G}\underset{\sim}{u}(k)+\underset{\sim}{T}\underset{\sim}{E}\underset{\sim}{v}(k+1). \tag{48}$$

Using equations (40) and (30) with equation (6), we have

$$\mathbf{G} = \mathbf{T}\begin{bmatrix} \mathbf{CB} \\ \mathbf{CAB} \\ \mathbf{CA}^2\mathbf{B} \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{CA}^{N-1}\mathbf{B} \end{bmatrix}. \tag{49}$$

Using (49), (37), and (36) with (48) we have

$$\mathbf{Gu}(k)+\mathbf{TEv}(k+1) = \mathbf{T}\begin{bmatrix} \mathbf{CB} \\ \mathbf{CAB} \\ \mathbf{CA}^2\mathbf{B} \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{CA}^{N-1}\mathbf{B} \end{bmatrix}\mathbf{u}(k)+$$

$$\mathbf{T}\begin{bmatrix} \mathbf{D} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{CB} & \mathbf{D} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{CA}^2\mathbf{B} & \mathbf{CB} & \mathbf{D} & \cdots & \mathbf{0} & \mathbf{0} \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \mathbf{D} & \cdot \\ \mathbf{CA}^{N-2}\mathbf{B} & \mathbf{CA}^{N-3}\mathbf{B} & \mathbf{CA}^{N-4}\mathbf{B} & \cdots & \mathbf{CB} & \mathbf{D} \end{bmatrix}\begin{bmatrix} \mathbf{u}(k+1) \\ \mathbf{u}(k+2) \\ \mathbf{u}(k+3) \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{u}(k+N) \end{bmatrix}. \tag{50}$$

Expansion and recondensation of (50) reveals that

$$\mathbf{Gu}(k)+\mathbf{TEv}(k+1) = \mathbf{TMv}^*(k), \tag{51}$$

where

$$\underset{\sim}{T}\underset{\sim}{M} = \underset{\sim}{T}\begin{bmatrix} \underset{\sim}{C}\underset{\sim}{B} & \underset{\sim}{D} & \underset{\sim}{0} & \cdots & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{C}\underset{\sim}{A}\underset{\sim}{B} & \underset{\sim}{C}\underset{\sim}{B} & \underset{\sim}{D} & \cdots & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{C}\underset{\sim}{A}^2\underset{\sim}{B} & \underset{\sim}{C}\underset{\sim}{A}\underset{\sim}{B} & \underset{\sim}{C}\underset{\sim}{B} & \cdots & \underset{\sim}{0} & \underset{\sim}{0} \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \underset{\sim}{D} & \underset{\sim}{0} \\ \underset{\sim}{C}\underset{\sim}{A}^{N-1}\underset{\sim}{B} & \underset{\sim}{C}\underset{\sim}{A}^{N-2}\underset{\sim}{B} & \underset{\sim}{C}\underset{\sim}{A}^{N-3}\underset{\sim}{B} & \cdots & \underset{\sim}{C}\underset{\sim}{B} & \underset{\sim}{D} \end{bmatrix}, \quad (52)$$

and

$$\underset{\sim}{v}^*(k) = \begin{bmatrix} \underset{\sim}{u}(k) \\ \underset{\sim}{u}(k+1) \\ \underset{\sim}{u}(k+2) \\ \cdot \\ \cdot \\ \cdot \\ \underset{\sim}{u}(k+N-1) \\ \underset{\sim}{u}(k+N) \end{bmatrix} = \begin{bmatrix} \underset{\sim}{v}(k) \\ \hline \underset{\sim}{u}(k+N) \end{bmatrix} . \quad (53)$$

The expression $-\underset{\sim}{F}\underset{\sim}{T}\underset{\sim}{E}\underset{\sim}{v}(k)$ from equation (48) can be written as

$$-\underset{\sim}{F}\underset{\sim}{T}\underset{\sim}{E}\underset{\sim}{v}(k) = -\underset{\sim}{F}\underset{\sim}{T}\underset{\sim}{L}\underset{\sim}{v}^*(k), \quad (54)$$

where

$$\underset{\sim}{L} = \begin{bmatrix} \underset{\sim}{D} & \underset{\sim}{0} & \underset{\sim}{0} & \cdots & \cdot & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{C}\underset{\sim}{B} & \underset{\sim}{D} & \underset{\sim}{0} & \cdots & \cdot & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{C}\underset{\sim}{A}\underset{\sim}{B} & \underset{\sim}{C}\underset{\sim}{B} & \underset{\sim}{D} & \cdots & \cdot & \underset{\sim}{0} & \underset{\sim}{0} \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \underset{\sim}{D} & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{C}\underset{\sim}{A}^{N-2}\underset{\sim}{B} & \underset{\sim}{C}\underset{\sim}{A}^{N-3}\underset{\sim}{B} & \underset{\sim}{C}\underset{\sim}{A}^{N-4}\underset{\sim}{B} & \cdots & \underset{\sim}{C}\underset{\sim}{B} & \underset{\sim}{D} & \underset{\sim}{0} \end{bmatrix}, \quad (55)$$

and $\underset{\sim}{v}^*(k)$ is given by (53). We can now use equations (51) and (54) to write

$$-\underset{\sim}{F}\underset{\sim}{T}\underset{\sim}{E}\underset{\sim}{v}(k)+\underset{\sim}{G}\underset{\sim}{u}(k)+\underset{\sim}{T}\underset{\sim}{E}\underset{\sim}{v}(k+1) = \underset{\sim}{R}\underset{\sim}{v}^*(k), \tag{56}$$

where

$$\underset{\sim}{R} = -\underset{\sim}{F}\underset{\sim}{T}\underset{\sim}{L}+\underset{\sim}{T}\underset{\sim}{M}. \tag{57}$$

Substitution of equation (56) into (48) reveals that $\underset{\sim}{T}\underset{\sim}{z}(k+1)$ can be written as

$$\underset{\sim}{T}\underset{\sim}{z}(k+1) = [\underset{\sim}{R},\underset{\sim}{F}]\begin{bmatrix} \underset{\sim}{v}^*(k) \\ \hline \underset{\sim}{T}\underset{\sim}{z}(k)\end{bmatrix}. \tag{58}$$

Gopinath [5] was the first to determine that, given $\underset{\sim}{T}$, knowledge of $\underset{\sim}{R}$ and $\underset{\sim}{F}$ suffices to identify the system. Let

$$\underset{\sim}{W} = \left[\begin{array}{c|c|c|c|c} \underset{\sim}{v}^*(1) & \underset{\sim}{v}^*(2) & \underset{\sim}{v}^*(3) & \cdots & \underset{\sim}{v}^*(r-1) \\ \hline \underset{\sim}{T}\underset{\sim}{z}(1) & \underset{\sim}{T}\underset{\sim}{z}(2) & \underset{\sim}{T}\underset{\sim}{z}(3) & \cdots & \underset{\sim}{T}\underset{\sim}{z}(r-1) \end{array}\right]. \tag{59}$$

Then $[\underset{\sim}{R},\underset{\sim}{F}]$ can be determined uniquely in the sense of least squares provided that

$$\det[\underset{\sim}{W}\underset{\sim}{W}'] > 0. \tag{60}$$

We will assume, for the moment, that $\underset{\sim}{T}$ is known and that (60) is true in order to expedite further analysis (i.e., $[\underset{\sim}{R},\underset{\sim}{F}]$ is known). The task is now to show that given $[\underset{\sim}{R},\underset{\sim}{F}]$ the realization $\underset{\sim}{\Sigma}$ is identified in the sense described earlier.

We now let

$$\underset{\sim}{R} = [\underset{\sim}{R}_0,\underset{\sim}{R}_1,\underset{\sim}{R}_2,\ldots,\underset{\sim}{R}_N] \tag{61}$$

where the $\underset{\sim}{R}_i$, i=0,1,2,...N are n×m constant matrices. After Gopinath [6], equation (57) is now evaluated in terms of the $\mathbf{R}_i$ (beginning with $\mathbf{R}_N$ and working back to $\mathbf{R}_0$) and the results used to determine the set of equations:

$$\mathbf{T}\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{D} \end{bmatrix} = \mathbf{R}_N \tag{62.1}$$

$$\mathbf{T}\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{0} \\ \mathbf{D} \\ \mathbf{CB} \end{bmatrix} = \mathbf{R}_{N-1}+\mathbf{F}\mathbf{R}_N \tag{62.2}$$

$$\mathbf{T}\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{D} \\ \mathbf{CB} \\ \mathbf{CAB} \end{bmatrix} = \mathbf{R}_{N-2}+\mathbf{F}\mathbf{R}_{N-1}+\mathbf{F}^2\mathbf{R}_N \tag{62.3}$$

$$\underset{\sim}{T}\begin{bmatrix} \underset{\sim}{D} \\ \underset{\sim}{C}\underset{\sim}{B} \\ \underset{\sim}{C}\underset{\sim}{A}\underset{\sim}{B} \\ \cdot \\ \cdot \\ \cdot \\ \underset{\sim}{C}\underset{\sim}{A}^{N-2}\underset{\sim}{B} \end{bmatrix} = \underset{\sim}{R}_1+\underset{\sim}{F}\underset{\sim}{R}_2+\underset{\sim}{F}^2\underset{\sim}{R}_3+\ldots+\underset{\sim}{F}^{N-1}\underset{\sim}{R}_N \tag{62.N}$$

$$\underset{\sim}{T}\begin{bmatrix} \underset{\sim}{C}\underset{\sim}{B} \\ \underset{\sim}{C}\underset{\sim}{A}\underset{\sim}{B} \\ \underset{\sim}{C}\underset{\sim}{A}^2\underset{\sim}{B} \\ \cdot \\ \cdot \\ \cdot \\ \underset{\sim}{C}\underset{\sim}{A}^{N-1}\underset{\sim}{B} \end{bmatrix} = \underset{\sim}{R}_0+\underset{\sim}{F}\underset{\sim}{R}_1+\underset{\sim}{F}^2\underset{\sim}{R}_2+\ldots+\underset{\sim}{F}^{N}\underset{\sim}{R}_N \tag{62.N+1}$$

Now, from equations (49) and (62.N+1) it appears that

$$\underset{\sim}{G} = \underset{\sim}{R}_0+\underset{\sim}{F}\underset{\sim}{R}_1+\underset{\sim}{F}^2\underset{\sim}{R}_2+\ldots+\underset{\sim}{F}^{N}\underset{\sim}{R}_N \tag{63}$$

Thus, $\underset{\sim}{F}$ and $\underset{\sim}{G}$ of equation (4) are now determined.

The procedure for determining the matrices $\underset{\sim}{H}$ and $\underset{\sim}{D}$ of equation (5) is greatly simplified if it can be assumed that the matrix $\underset{\sim}{C}$ is full rank. Assuming that this is true, we can conveniently let the first p columns of $\underset{\sim}{T}$ have non-zero selection elements. In this event, the choice

$$\underset{\sim}{H} = [\underset{\sim}{I}_p,\underset{\sim}{0}], \tag{64}$$

where $\underset{\sim}{I}_p$ is the p×p identity matrix, satisfies equation (47) since

$$[\underset{\sim}{I}_p,\underset{\sim}{0}]\underset{\sim}{T}[\underset{\sim}{z}(k)-\underset{\sim}{E}\underset{\sim}{v}(k)]+\underset{\sim}{D}\underset{\sim}{u}(k) = \underset{\sim}{y}(k)-\underset{\sim}{D}\underset{\sim}{u}(k)+\underset{\sim}{D}\underset{\sim}{u}(k) = \underset{\sim}{y}(k).$$

Also, under the assumption that the first p columns of $\underset{\sim}{T}$

are non-trivial, the matrix $\underset{\sim}{D}$ is equal to the first p rows of the matrix given by equation (62.N).

Further, it is now a simple matter to compute the equivalent sequence of states $\underset{\sim}{Q}$ using equation (45). The vectors $\underset{\sim}{z}(k)$ and $\underset{\sim}{v}(k)$ are known and the matrix $\underset{\sim}{T}\underset{\sim}{E}$ can be directly formed from expressions (62.1) through (62.N) taken in reverse order, i.e.,

$$\underset{\sim}{T}\underset{\sim}{E} = \left(\underset{\sim}{T}\begin{bmatrix} \underset{\sim}{D} \\ \underset{\sim}{C}\underset{\sim}{B} \\ \underset{\sim}{C}\underset{\sim}{A}\underset{\sim}{B} \\ \cdot \\ \cdot \\ \cdot \\ \underset{\sim}{C}\underset{\sim}{A}^{N-2}\underset{\sim}{B} \end{bmatrix}, \underset{\sim}{T}\begin{bmatrix} \underset{\sim}{0} \\ \underset{\sim}{D} \\ \underset{\sim}{C}\underset{\sim}{B} \\ \cdot \\ \cdot \\ \cdot \\ \underset{\sim}{C}\underset{\sim}{A}^{N-3}\underset{\sim}{B} \end{bmatrix}, \ldots, \underset{\sim}{T}\begin{bmatrix} \underset{\sim}{0} \\ \underset{\sim}{0} \\ \underset{\sim}{0} \\ \cdot \\ \cdot \\ \cdot \\ \underset{\sim}{D} \end{bmatrix} \right) \tag{65}$$

Thus, a minimal realization equivalent to $\underset{\sim}{\Sigma}$ has been obtained.

We now return our consideration to determination of the triplet $(\underset{\sim}{T},\underset{\sim}{R},\underset{\sim}{F})$ in equation (58). Gopinath [5,6] determined that a choice of $\underset{\sim}{T}$ with row dimension greater than n leads to certain singularity of $\underset{\sim}{W}\underset{\sim}{W}'$ in expression (60). Also, it can be observed from (59) that any linear dependence among the inputs (i.e., the first (N+1)m rows of $\underset{\sim}{W}$) leads to the violation of (60). In order to determine the triplet $(\underset{\sim}{T},\underset{\sim}{R},\underset{\sim}{F})$ Gopinath assumes that the system inputs are sufficiently "general" such that singularity of $\underset{\sim}{W}\underset{\sim}{W}'$ depends only upon the choice of $\underset{\sim}{T}$ as in (59). His method begins by assuming a system dimension greater than n (with an associated $\underset{\sim}{T}$) which leads to singularity of $\underset{\sim}{W}\underset{\sim}{W}'$. The row dimensions of $\underset{\sim}{T}$ is then successively reduced until $\underset{\sim}{W}\underset{\sim}{W}'$ becomes non-singular, i.e., the row dimension of $\underset{\sim}{T}$ is equal to the order of a minimal realization. Determination of matrices $\underset{\sim}{R}$ and $\underset{\sim}{F}$ is then straightforward via least squares.

For convenience, we now define $\underset{\sim}{\Delta}$ as

$$\underset{\sim}{\Delta} = \left[\begin{array}{c|c|c|c|c} \underset{\sim}{v}^*(1) & \underset{\sim}{v}^*(2) & \underset{\sim}{v}^*(3) & & \underset{\sim}{v}^*(r-1) \\ \hline \underset{\sim}{z}(1) & \underset{\sim}{z}(2) & \underset{\sim}{z}(3) & \cdots & \underset{\sim}{z}(r-1) \end{array}\right] = \left[\begin{array}{c} \underset{\sim}{\Delta}_{(N+1)m} \\ \hline \underset{\sim}{\Delta}_{Np} \end{array}\right] \tag{66}$$

where $\underset{\sim}{\Delta}_{(N+1)m}$ is a $(N+1)m\times(r-1)$ matrix of input observations and $\underset{\sim}{\Delta}_{Np}$ is a $Np\times(r-1)$ matrix of output observations. Büdin [7] observed that the problem of determining the matrix $\underset{\sim}{T}$ can be reduced to selecting a maximum number of linearly independent rows from $\underset{\sim}{\Delta}_{Np}$ which are themselves linearly independent of the rows of $\underset{\sim}{\Delta}_{(N+1)m}$. In order to accomplish this feat, Budin employed the process of Gaussian elimination along with the assumption of input generality.

OBTAINING REALIZATIONS

We now present a new algorithm for computing a minimal realization based on structural relationships due to Gopinath. The algorithm is extremely simple, makes no assumption as to input generality, and can be applied to the noisy case in its present form.

We begin by passing the matrix $\underset{\sim}{\Delta}$, given by equation (66), through a Schmidt filter to produce the orthonormal matrix $\underset{\sim}{\Delta}$ and matrices $\underset{\sim}{S}$ and $\underset{\sim}{S}^+$ such that

$$\underset{\sim}{\Theta} = \underset{\sim}{S}\underset{\sim}{\Delta}, \tag{67}$$

$$\underset{\sim}{\Delta} = \underset{\sim}{S}^+\underset{\sim}{\Theta}, \tag{68}$$

$$\underset{\sim}{\Theta}\underset{\sim}{\Theta}' = \underset{\sim}{I}, \tag{69}$$

and

$$\underset{\sim}{S}\underset{\sim}{S}^+ = \underset{\sim}{I}, \tag{70}$$

in view of the earlier discussion of the Schmidt filter. Further, considering (66) we can write

$$\underset{\sim}{\Theta} = [\underset{\sim}{S}_{(N+1)m}, \underset{\sim}{S}_{Np}] \begin{bmatrix} \underset{\sim}{\Delta}_{(N+1)m} \\ \hline \underset{\sim}{\Delta}_{Np} \end{bmatrix} \tag{71}$$

where

$$[\underset{\sim}{S}_{(N+1)m}, \underset{\sim}{S}_{Np}] = \underset{\sim}{S}. \tag{72}$$

In equation (72), the column dimensions of the partitions of $\underset{\sim}{S}$ are given by their subscripts.

As a result of the filtering process, certain of the columns of $\underset{\sim}{S}$ may be zero. Recall that the appearance of zero columns in $\underset{\sim}{S}$ corresponds to rejection (by the filter) of those corresponding rows of $\underset{\sim}{\Delta}$ which do not contribute linearly independent information to $\underset{\sim}{\Theta}$. Thus, rows of selected (not rejected) in this manner constitute a linearly independent set. Also, since the filter considers the rows in serial order, we can be sure that the rows selected (not rejected) by the filter from $\underset{\sim}{\Delta}_{Np}$, i.e., the output elements, constitute a maximum number of such linearly independent rows which are themselves linearly independent of the rows of $\underset{\sim}{\Delta}_{(N+1)m}$, i.e., the input elements.

It should now be clear that $\underset{\sim}{S}_{Np}$ contains sufficient information to determine $\underset{\sim}{T}$ uniquely. In fact, the row dimension of the selector matrix $\underset{\sim}{T}$ is given by the number of non-trivial columns of $\underset{\sim}{S}_{Np}$. The column position of the unity element in each succeeding row of $\underset{\sim}{T}$ is given by the column number of each succeeding non-trivial column of $\underset{\sim}{S}_{Np}$ (numbering the columns of $\underset{\sim}{S}_{Np}$ as 1, 2, 3, ..., Np).

Now, since $\underset{\sim}{T}$ is determined, equation (58) can, in principle, be used to solve for $[\underset{\sim}{R},\underset{\sim}{F}]$ in the sense of least squares. However, we must still be concerned that condition (60) is satisfied, as no assumptions concerning the class of $\underset{\sim}{u}(k)$ have been made. In order to circumvent this difficulty, and to avoid the possibility of having to invert the matrix $\underset{\sim}{W}\underset{\sim}{W}'$, we now make maximal use of information provided by the Schmidt filter.

Let

$$\underset{\sim}{\Lambda} = \underset{\sim}{T}[\underset{\sim}{z}(2),\underset{\sim}{z}(3),\underset{\sim}{z}(4),\ldots,\underset{\sim}{z}(r)]. \tag{73}$$

Using equations (66) and (73) with equation (58), it is clear that there exists a $n\times\{(N+1)m+Np\}$ matrix $\underset{\sim}{\Gamma}$ (not necessarily unique) such that

$$\underset{\sim}{\Lambda} = \underset{\sim}{\Gamma}\underset{\sim}{\Delta}. \tag{74}$$

Using the basic relationship of linear least squares (16), we can write[5]

---

[5]For the noisy case, $\underset{\sim}{\Gamma}$ becomes $\hat{\underset{\sim}{\Gamma}}$

$$\Gamma\Delta\Delta' = \Lambda\Delta'. \quad (75)$$

Substituting equation (68) into equation (75), the result is

$$\Gamma S^{+}\Theta\Theta' S^{+\prime} = \Lambda\Theta' S^{+\prime}. \quad (76)$$

But, in view of equation (69), (76) becomes

$$\Gamma S^{+} S^{+\prime} = \Lambda\Theta' S^{+\prime}. \quad (77)$$

Now, a solution for Γ in the sense of least squares is

$$\Gamma = \Lambda\Theta' S \quad (78)$$

which can be verified by direct substitution and noting the property (70).

The matrix Γ can be written as

$$\Gamma = [\Gamma_{(N+1)m}, \Gamma_{Np}]. \quad (79)$$

Now, from equations (72) and (78), it is clear that zero columns in S_{Np} will result in corresponding zero columns in Γ_{Np}, so that

$$F = (T\Gamma'_{Np})' = \Gamma_{Np} T'. \quad (80)$$

Also, it is easy to see that

$$R = \Gamma_{(N+1)m}. \quad (81)$$

Summary of Procedure

Since the development has been somewhat lengthy, the necessary activity involved in obtaining a minimal realization is summarized below.

1. Form Δ as in equation (66).
2. Pass Δ through a Schmidt filter and obtain Θ and S as in equations (67) through (70).
3. Determine T from S_{Np} as in equation (71).
4. Form Λ using equation (73).

5. Determine $\underset{\sim}{\Gamma} = \underset{\sim}{\Lambda}\underset{\sim}{\Theta}'\underset{\sim}{S}$.
6. Select $[\underset{\sim}{R},\underset{\sim}{F}]$ using equations (80) and (81).
7. Implement the realization procedure as in equations (61) through (65).

Steps 1 through 7 are very easily implemented on a digital computer. No matrix inversion whatsoever is required. The only subprograms necessary are a Schmidt filter and a matrix multiplication routine.

CONCLUDING REMARKS

An algorithm for computing a minimal realization of the external description of a linear discrete time system has been presented. We now consider some of the difficulties which have been known to arise in a practical modeling problem including the noisy case.

Order Determination

It has been uniformly assumed throughout this presentation that an integer N is known as in equation (35). It is easy to see that N directly governs the computational labor involved in the realization algorithm.

In the interest of economy it has been observed, Budin [7], that equation (31) is not a necessary condition for the construction of a minimal realization. Without loss of generality, we can clearly define N to be the minimum integer such that equation (29) is satisfied. In this case, analysis proceeds exactly as before, but with the possibility of greater efficiency.

Nonetheless, the problem of an appropriate selection of N remains. Unfortunately, there is no easy answer to this problem unless definite prior knowledge of system structure is available. Noise, and the possibility of "distributed lags" increase the difficulty of the problem.

One thing is certain, namely, a choice of N that is too small can lead to results which are quite incorrect. However, this observation leads to some salvation for the noise-free case. If an N is chosen which is too small, it will be impossible to exactly realize the given input/output description of a linear system.

Adequate Observations

The number of corresponding observations of the system input and output vectors is determined by the choice of N (see "Order Determination" above) and the number of system inputs and outputs. Reference to equations (66) and (73) indicate that we should select r such that

$$r > (N+1)m+Np.$$

For the case where noise is present, it is generally advisable to choose r considerably larger than indicated above.

Identification

Identification of a system simply cannot be accomplished (in the sense defined earlier) unless the system to be identified is a minimal realization (i.e., it is completely controllable and completely observable). It is worthwhile to note here that a minimal realization can only represent those modes of the system which are excited during the input/output sequence. For purposes of analysis, modes not excited may be considered as resulting in uncontrollable and unobservable states. Also, as will be discussed in the next section, it is impossible to even identify a minimal realization unless the system inputs are sufficiently "general."

Input Generality

In previous discussions, Ho [4], Gopinath [6], and Budin [7], input generality was taken to be a necessary condition. As previously stated, our algorithm makes no such assumption in the computation of a minimal realization. However, some difficulty can be experienced under the latter specification and this will now be discussed.

Input "generality" will be taken to mean only that

$$\text{rank } [\underset{\sim}{\Delta}_{(N+1)m}] = (N+1)m, \tag{83}$$

where $\underset{\sim}{\Delta}_{(N+1)m}$ is defined by equation (66). Note that the idea of input generality does not necessarily bear any relationship to the consideration of modal excitation as discussed earlier. If the elements of $\underset{\sim}{u}(k)$ are drawn at random, then no problem exists. There is, however, an important class of inputs for which equation (83) is almost certainly not true.

As an example, consider a single input/single output, second order system where the input is equal to a constant in the external description. This might correspond to an identification attempt using a simple step function imposed at time t=0. For this case, it is easy to determine that (83) is not satisfied. In fact, this same difficulty would also occur if this same system were forced by a discrete-valued sine function.

Under the above circumstances there is no algorithm which can successfully <u>identify</u> the system (even though all modes may be excited and the system is controllable and observable). However, this does not preclude the possibility of obtaining a minimal realization of the input/output description. It can easily be seen from equation (58) that this is true.

Given an external description of a linear system, we can always find an $\underset{\sim}{R}$ such that equation (58) is satisfied. The problem is that if equation (83) is not true, $\underset{\sim}{R}$ is not <u>uniquely</u> specified in the sense of least squares (even though the algorithm described in this paper yields a suitable $\underset{\sim}{R}$). Given $\underset{\sim}{T}$, each $\underset{\sim}{R}$ which satisfies equation (58) results in a different minimal realization. Simple examples can be constructed to show that these realizations do not generally satisfy condition 1 of equivalence (and hence are not equivalent). Thus, the implication of previous discussion is that if equation (83) is not true, it is impossible to identify a system, even though the system is a minimal realization.

However, it can be observed from equation (58) that $\underset{\sim}{F}$, and hence the order of the system, are still uniquely determined. Also, recall that a minimal realization of the external description is obtained regardless of the validity of (83). Such realizations are not usually sufficient for purposes of optimal control, but are obviously sufficient for prediction if the input for future times is "consistent." In any case, the realization obtained for this case is quite likely the best available under the circumstances.

Noise

For the case where input and output observations are corrupted by additive noise of essentially unknown statistics, the Schmidt filter can be used to empirically determine system order by judicious adjustment of the specified "noise to signal ratios" discussed earlier. This procedure produces good results when the contribution of noise processes is "small." The particular advantage to using the Schmidt filter is that linear dependence is always determined in the sense of mean square, as opposed to the more arbitrary specification required in Gaussian elimination. If noise is known to be present, we additionally recommend a visual inspection of the orthonormal filter output with the observation that the last rows of the output tend to have a proportionately larger noise content.

We note that if noise is present, the procedure described in this paper may produce estimates of system parameters which are statistically biased. However, for the class of noise described above, there is no procedure which can guarantee unbiased estimates of system parameters. Finally, we note that the terms "biased" and "useless" are not necessarily synonomous, depending upon the problem at hand, and let the utility of our algorithm speak for itself.

FURTHER RESEARCH

The Authors realize that there may be difficulty involved with "knowing" both the input and output of the Schmidt filter simultaneously in some identification problems due to possible constraints on computer storage. Where this difficulty can be circumvented, we believe the approach presented here to be extremely valuable. We do not recommend, however, that the procedure presented here be applied to the adaptive filtering problem since added input/output observations normally require complete regeneration of the orthonormal sequence.

The above difficulties have led the authors to develop a modified algorithm, based upon the analysis presented here, in which it is never necessary to know the output <u>or</u> input of the Schmidt filter explicitly in order to accomplish the purpose at hand. Besides providing an exact solution for the noise free case, the modified algorithm yields unbiased estimates of system order and parameters for the case where noise statistics are known. This extremely efficient algorithm, which will be considered in a forthcoming publication, can also be conveniently applied to the adaptive filtering problem.

ACKNOWLEDGEMENTS

The Authors wish to express their gratitude to Dr. J. S. Pazdera and Dr. R. T. Johnson for their assistance in preparation of this work.

REFERENCES

1. Kalman, R. E., "On the General Theory of Control Systems," Proc. I.F.A.C., Moscow, Vol. 1, 1960, pp. 481-492.

2. Kalman, R. E., "Mathematical Description of Linear Dynamical Systems," S.I.A.M. J. Control, Ser. A, Vol. 1, No. 2, 1963, pp. 152-192.

3. Ho, B. L., and Kalman, R. E., "Effective Construction of Linear State-Variable Models From Input/Output Functions," Regelungstechnik, Vol. 14, 1966, pp. 545-548.

4. Ho, B. L., "An Effective Construction of Realization From Input/Output Descriptions," Ph.D. thesis, Stanford University, 1966.

5. Gopinath, B., and Lange, B. O., "On the Identification and Control of Linear Systems," Stanford University Center for Systems Research, SUDAAR No. 351, June, 1968.

6. Gopinath, B., "On the Identification of Linear Time-Invariant Systems From Input-Output Data," Bell System Tech. J., Vol. 48, No. 5, May 1969, pp. 1101-1113.

7. Budin, M. A., "Minimal Realization of Discrete Linear Systems From Input-Output Observations," IEEE Trans. Auto. Contr., AC-16, No. 5, Oct. 1971, pp. 395-401.

8. Kitamori, T., "Applications of Orthogonal Functions to the Determination of Process Dynamic Characteristics and to Construction of Self Optimizing Control Systems," Proc. I.F.A.C., Moscow, Vol. 2, 1960, pp. 613-618.

9. Lubbock, J. K., and Barker, H. A., "A Solution to

the Identification Problem," Preprints J.A.C.C., Minneapolis, 1963, pp. 191-199.

10. Barker, H. A., "Orthogonalizing Techniques for System Identification and Optimization," Proc. I.F.A.C., London, 1966, p. 36A.1.

11. Roberts, P. D., "Self Adjustable Orthogonal Digital Filters for System Identification and Optimization," Proc. IEE, Vol. 114, No. 5, May 1967, pp. 666-670.

12. Fisher, F. M., The Identification Problem in Econometrics, McGraw-Hill, New York, 1966.

13. Lee, R. C. K., Optimal Estimation, Identification and Control, MIT Press, Cambridge, Mass., 1964.

14. Rowe, I. H., "A Bootstrap Method for the Statistical Estimation of Model Parameters," Int. J. Contr., Vol. 12, No. 5, Nov. 1970, pp. 721-738.

15. Sage, A. P., and Melsa, J. L., Estimation Theory With Applications to Communications and Control, McGraw-Hill, New York, 1970.

16. Penrose, R., "A Generalized Inverse for Matrices," Proc. Cambridge Phil. Soc., Vol. 51, Pt. 3, 1955, pp. 406-413.

17. Rao, C. R., "A Note on the Generalized Inverse of a Matrix With Applications to Problems in Mathematical Statistics," J. Royal Statistical Soc., Ser. B, Vol. 24, 1962, pp. 152-158.

18. Drygas, H., The Coordinate-Free Approach to Gauss-Markov Estimation, Springer-Verlag, Berlin, 1970.

19. Mayne, D. Q., "An Algorithm for Computing the Pseudo-Inverse of a Singular Matrix," Computer J., Vol. 9, Nov. 1966, pp. 312-317.

20. Behring, A. G., "Methods and Applications of System Identification," Ph.D. thesis, University of Missouri-Rolla, Rolla, Mo., Dec. 1971.

UNBIASED LINEAR SYSTEM IDENTIFICATION IN A NOISY ENVIRONMENT

by

A. G. Behring

and

V. J. Flanigan*

ABSTRACT

An optimal filter is proposed to compute a basis of a set of "noisy" filter input functions for the case where signal and noise statistics are specified. Procedures for determining the filter transfer function and the transfer function of a restoring filter are given. The filter is then applied to the problem of minimum-order mathematical modeling of discrete, multi-variable, dynamic systems in noisy environments. It is shown that the resulting estimates of system order and system parameters are unbiased.

---

*The Authors are associated with the Department of Mechanical and Aerospace Engineering, University of Missouri - Rolla, Rolla, Missouri 65401, where Mr. Behring is a Graduate Student and Dr. Flanigan (member ASME) is an Associate Professor.

NOTATION

In this paper all bold-face capital letters denote matrices. Vectors are defined in column format and are denoted by lower case letters in bold-face type. All scalars will be denoted by plain upper or lower case letters. Occasionally it will be necessary to display the format of a vector or matrix explicitly, e.g.,

$$\underset{\sim}{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}.$$

Any exceptions to these general rules will be clearly specified in the text.

INTRODUCTION

In an earlier paper, the Authors [1][1] introduced a "Schmidt filter" to compute a set of orthonormal basis functions of a set of "noisy" input functions for the case where noise statistics were essentially unspecified. The novel features of the filter were then shown to be applica-

---

[1]Numbers in brackets designate references at end of paper.

ble to the problem of system identification. The method demonstrated in our previous paper was considered to be the "best available" for unspecified noise, but was described as non-optimal in the event that noise and signal statistics are known. In this paper we develop a modified Schmidt filter for the latter case and show how the filter can be used to obtain unbiased estimates of system parameters.

The Schmidt filter algorithm is based on the well-known "Gram-Schmidt orthonormalization process," Epstein [2], also known as "orthonormalization due to Erhard Schmidt," Drygas [3]. The Gram-Schmidt procedure has been used by others working in the area of system identification, notably by Kitamori [4], Lubbock [5], Clement [6], Lubbock and Barker [7], Douce and Roberts [8], Barker and Hawley [9], and Roberts [10]. The most popular approach to the problem proposes an expansion of system inputs as a set of orthogonal functions. A system transfer function is then formed as a set of optimal, independently-determined parameters which most nearly approximates known operating data. The problem of noise has not been thoroughly considered in these analyses.

The first direct method for computing a minimal realization of a linear, multi-variable, discrete system from input/output data was derived by Gopinath [11]. Budin [12] gives a much improved algorithm based on the earlier work of Gopinath. A central problem in the extension of either algorithm to the noisy case is the determination of matrix rank in the presence of additive noise - a problem that has not been satisfactorily solved to date.

PROBLEM STATEMENT

We begin by assuming the existence of real $p \times 1$ vector functions of time $\underset{\sim}{f}(t)$, $\underset{\sim}{v}(t)$, and $\underset{\sim}{y}(t)$. Further, let the $p \times r$ matrices $\underset{\sim}{F}$, $\underset{\sim}{V}$, and $\underset{\sim}{Y}$ be ordered sets which define the ranges and domains of $\underset{\sim}{f}(t)$, $\underset{\sim}{v}(t)$, and $\underset{\sim}{y}(t)$, respectively, such that

$$\underset{\sim}{F} = [\underset{\sim}{f}(t_1), \underset{\sim}{f}(t_2), \underset{\sim}{f}(t_3), \ldots, \underset{\sim}{f}(t_r)], \tag{1}$$

$$\underset{\sim}{V} = [\underset{\sim}{v}(t_1), \underset{\sim}{v}(t_2), \underset{\sim}{v}(t_3), \ldots, \underset{\sim}{v}(t_r)], \tag{2}$$

and

$$\underset{\sim}{Y} = [\underset{\sim}{y}(t_1), \underset{\sim}{y}(t_2), \underset{\sim}{y}(t_3), \ldots, \underset{\sim}{y}(t_r)], \tag{3}$$

where $t_i < t_{i+1}, \forall i$. Similarly, let the real $n \times 1$ vector function $\underset{\sim}{x}(t)$ exist and let the $n \times r$ matrix $\underset{\sim}{X}$ be defined as

$$\underset{\sim}{X} = [\underset{\sim}{x}(t_1), \underset{\sim}{x}(t_2), \underset{\sim}{x}(t_3), \ldots, \underset{\sim}{x}(t_r)]. \tag{4}$$

In the event that r in equations (1) through (4) approaches infinity, we require that $\underset{\sim}{f}(t)$, $\underset{\sim}{v}(t)$, $\underset{\sim}{y}(t)$, and $\underset{\sim}{x}(t)$ approach sectionally continuous functions. In any case, we require that $\underset{\sim}{f}(t)$, $\underset{\sim}{v}(t)$, $\underset{\sim}{y}(t)$, and $\underset{\sim}{x}(t)$ exist, $\forall t_i$.

We will say that two functions, say $\underset{\sim}{x}(t)$ and $\underset{\sim}{y}(t)$, are <u>orthogonal</u> if it happens that

$$\underset{\sim}{X}\underset{\sim}{Y}' = \underset{\sim}{0}, \tag{5}$$

which clearly implies that

$$\underset{\sim}{Y}\underset{\sim}{X}' = \underset{\sim}{0}', \tag{6}$$

where $\underset{\sim}{Y}$ and $\underset{\sim}{X}$ are given by equations (3) and (4), respectively, and $\underset{\sim}{0}$ is the $n \times p$ null matrix. Given (5) or (6), it is easy to show that $\underset{\sim}{Y}$ is orthogonal to the function space spanned by $\underset{\sim}{X}$, and conversely. If $\underset{\sim}{x}(t)$ and $\underset{\sim}{y}(t)$ are discrete functions, we compute $\underset{\sim}{X}\underset{\sim}{Y}'$ from

$$\underset{\sim}{X}\underset{\sim}{Y}' = \sum_{i=1}^{r} \underset{\sim}{x}(t_i)\underset{\sim}{y}'(t_i), \tag{7}$$

whereas, if $\underset{\sim}{x}(t)$ and $\underset{\sim}{y}(t)$ are sectionally continuous, we define $\underset{\sim}{X}\underset{\sim}{Y}'$ as

$$\underset{\sim}{X}\underset{\sim}{Y}' = \int_{t_1}^{t_r} \underset{\sim}{x}(t)\underset{\sim}{y}'(t)dt. \tag{8}$$

A function, say $\underset{\sim}{x}(t)$, will be called <u>orthonormal</u> if

$$\underset{\sim}{X}\underset{\sim}{X}' = \underset{\sim}{I}_n, \tag{9}$$

where $\underset{\sim}{I}_n$ is the $n \times n$ identity matrix.

In order to illustrate the problem at hand, we let $\underset{\sim}{f}(t)$ be a noise-free "signal" function, $\underset{\sim}{v}(t)$ be an "observation noise" function, and $\underset{\sim}{y}(t)$ be an "observed" function, where the relation between $\underset{\sim}{f}(t)$, $\underset{\sim}{v}(t)$, and $\underset{\sim}{y}(t)$ is given by

$$\underset{\sim}{Y} = \underset{\sim}{F}+\underset{\sim}{V}. \tag{10}$$

We assume that $\underset{\sim}{Y}$ is known, but that we have no explicit knowledge of $\underset{\sim}{F}$ or $\underset{\sim}{V}$ except for the statistics $\underset{\sim}{F}\underset{\sim}{V}'$ and $\underset{\sim}{V}\underset{\sim}{V}'$. The fundamental problem is to optimally compute a set of basis functions of expression (10). We then show that the solution of this problem leads to unbiased identification of system parameters for the case where noise statistics are known.

LINEAR LEAST SQUARES

Let $\underset{\sim}{Y}$ and $\underset{\sim}{X}$ be defined as in equations (3) and (4), respectively. Let $\underset{\sim}{A}$ be a p×n constant matrix. It can be shown, e.g., Sage and Melsa [13], that an optimal, linear conditional estimate of $\underset{\sim}{y}(t)$ given $\underset{\sim}{x}(t)$ in the sense of minimum mean square[2] $\{\hat{\underset{\sim}{y}}(t)|\underset{\sim}{x}(t)\}$ is

$$\underset{\sim}{y}(t)|\underset{\sim}{x}(t) = \hat{\underset{\sim}{A}}\underset{\sim}{x}(t), \tag{11}$$

where $\hat{\underset{\sim}{A}}$ is a constant matrix which satisfies

$$\hat{\underset{\sim}{A}}\underset{\sim}{X}\underset{\sim}{X}' = \underset{\sim}{Y}\underset{\sim}{X}'. \tag{12}$$

If $\underset{\sim}{X}\underset{\sim}{X}'$ is non-singular, $\hat{\underset{\sim}{A}}$ is uniquely determined as

$$\hat{\underset{\sim}{A}} = \underset{\sim}{Y}\underset{\sim}{X}'[\underset{\sim}{X}\underset{\sim}{X}']^{-1}. \tag{13}$$

Using (13), it is easy to demonstrate the significant result that the matrix $\underset{\sim}{E}$ defined by

$$\underset{\sim}{E} = \underset{\sim}{Y}-\underset{\sim}{P}, \tag{14}$$

where

$$\underset{\sim}{P} = \hat{\underset{\sim}{A}}\underset{\sim}{X}, \tag{15}$$

is orthogonal to $\underset{\sim}{X}$, i.e.,

$$\underset{\sim}{E}\underset{\sim}{X}' = \underset{\sim}{0}. \tag{16}$$

Also, using (13), (14), and (15), we can show that

$$\underset{\sim}{E}\underset{\sim}{P}' = \underset{\sim}{0}. \tag{17}$$

From equation (14) we conclude that the function $\underset{\sim}{Y}$ can be written as

$$\underset{\sim}{Y} = \underset{\sim}{E}+\underset{\sim}{P}, \tag{18}$$

---

[2] Here, $\hat{\underset{\sim}{A}}$ minimizes $\{[\underset{\sim}{Y}-\underset{\sim}{A}\underset{\sim}{X}][\underset{\sim}{Y}-\underset{\sim}{A}\underset{\sim}{X}]'\}$.

where $\underset{\sim}{E}$ may be considered as the component of $\underset{\sim}{Y}$ which is orthogonal to the space spanned by $\underset{\sim}{X}$ and $\underset{\sim}{P}$ as the orthogonal projection of $\underset{\sim}{Y}$ on the space spanned by $\underset{\sim}{X}$.

THE SCHMIDT FILTER

The defining relationships for the Schmidt filter, derived in [1], are now repeated for convenience.

Given $\underset{\sim}{Y}$ of equation (3), the Schmidt filter produces a function $\underset{\sim}{X}$ of equation (4), where $n \leq p$, and linear transformations $\underset{\sim}{S}$ and $\underset{\sim}{S}^+$ such that

$$\underset{\sim}{X}\underset{\sim}{X}' = \underset{\sim}{I}_n, \tag{19}$$

$$\underset{\sim}{X} = \underset{\sim}{S}\underset{\sim}{Y}, \tag{20}$$

and

$$\underset{\sim}{Y} = \underset{\sim}{S}^+\underset{\sim}{X}, \tag{21}$$

where $\underset{\sim}{S}$ is the transfer function of the Schmidt filter, and $\underset{\sim}{S}^+$ can be considered as the transfer function of a restoring filter. Further, we can determine directly from (19), (20), and (21) that

$$\underset{\sim}{S}^+ = \underset{\sim}{Y}\underset{\sim}{X}', \tag{22}$$

and

$$\underset{\sim}{S}\underset{\sim}{S}^+ = \underset{\sim}{I}_n. \tag{23}$$

It is obvious from (23) that $\underset{\sim}{S}^+$ qualifies as a "generalized inverse" of $\underset{\sim}{S}$, Rao [14], namely,

$$\underset{\sim}{S}^+\underset{\sim}{S}\underset{\sim}{S}^+ = \underset{\sim}{S}^+ \tag{24}$$

and

$$\underset{\sim}{S}\underset{\sim}{S}^+\underset{\sim}{S} = \underset{\sim}{S}. \tag{25}$$

Further, it is straightforward to show that the matrix $\underset{\sim}{S}'\underset{\sim}{S}$

is a generalized inverse of the Gram matrix $\underset{\sim}{Y}\underset{\sim}{Y}'$, i.e.,

$$\underset{\sim}{S}'\underset{\sim}{S} = [\underset{\sim}{Y}\underset{\sim}{Y}']^{+}, \tag{26}$$

where $[\underset{\sim}{Y}\underset{\sim}{Y}']^{+}$ has extensive application in the area of linear estimation, Drygas [3].

We let $\underset{\sim}{E}$ be defined as the set of p×1 vectors

$$\underset{\sim}{E} = [\underset{\sim}{e}(t_1), \underset{\sim}{e}(t_2), \underset{\sim}{e}(t_3), \ldots, \underset{\sim}{e}(t_r)]. \tag{27}$$

The notation $\underset{\sim}{E}_j$, j=1,p will denote the jth row of $\underset{\sim}{E}$. Also, we let $\underset{\sim}{Y}_j$, j=1,p denote the jth row of $\underset{\sim}{Y}$; $\underset{\sim}{X}_i$, i=1,n denote the ith row of $\underset{\sim}{X}$; and $\underset{\sim}{Z}_i$, i=1,n denote the <u>first i</u> rows of $\underset{\sim}{X}$.

Let the notation $||\underset{\sim}{E}_j||$ denote the "norm" of $\underset{\sim}{E}_j$, defined by

$$||\underset{\sim}{E}_j|| = [\underset{\sim}{E}_j\underset{\sim}{E}_j']^{\frac{1}{2}} \tag{28}$$

Also, we let $\underset{\sim}{S}_i$, i=1,n denote the ith row of $\underset{\sim}{S}$; $\underset{\sim}{S}_j^{+}$, j=1,p denote the jth row of $\underset{\sim}{S}^{+}$; and s_{ji}^{+}, j=1,p, i=1,n denote the elements of matrix $\underset{\sim}{S}^{+}$.

The Schmidt algorithm begins by assuming that the first row of $\underset{\sim}{Y}$ is non-trivial. In this event, we let

$$\underset{\sim}{X}_1 = \underset{\sim}{Y}_1/||\underset{\sim}{Y}_1||, \tag{29}$$

$$\underset{\sim}{S}_1^{+} = [||\underset{\sim}{Y}_1||, 0, 0, 0, \ldots, 0], \tag{30}$$

and

$$\underset{\sim}{S}_1 = [1, 0, 0, 0, \ldots, 0]/||\underset{\sim}{Y}_1||. \tag{31}$$

Successive rows of $\underset{\sim}{X}$, $\underset{\sim}{S}^{+}$, and $\underset{\sim}{S}$ are generated by

$$\underset{\sim}{E}_j = \underset{\sim}{Y}_j - [\underset{\sim}{Y}_j\underset{\sim}{Z}_i']\underset{\sim}{Z}_i, \tag{32}$$

$$\underset{\sim}{X}_{i+1} = \underset{\sim}{E}_j/||\underset{\sim}{E}_j||, \quad \text{if } (\underset{\sim}{E}_j\underset{\sim}{E}_j')/(\underset{\sim}{Y}_j\underset{\sim}{Y}_j') > \varepsilon_j, \tag{33}$$

$$\underset{\sim}{S}_j^{+} = [\underset{\sim}{Y}_j\underset{\sim}{Z}_i'), 0, 0, 0, \ldots, 0] + \underset{\sim}{Q}_{i+1}(j), \tag{34}$$

and

$$\underset{\sim}{S}_{i+1} = [\underset{\sim}{U}_j - \sum_{k=1}^{i} s_{jk}^{+}\underset{\sim}{S}_k]/||\underset{\sim}{E}_j||, \text{ if } (\underset{\sim}{E}_j\underset{\sim}{E}_j')/(\underset{\sim}{Y}_j\underset{\sim}{Y}_j') > \varepsilon_j. \quad (35)$$

The 1×n matrix $\underset{\sim}{Q}_i(j)$ is defined as

$$\underset{\sim}{Q}_i(j) = [0,0,0,\ldots,0,\Delta_j,0,\ldots,0], \quad (36)$$

where

Δ_j, defined by

$$\Delta_j = ||\underset{\sim}{E}_j|| \quad \text{if } (\underset{\sim}{E}_j\underset{\sim}{E}_j')/(\underset{\sim}{Y}_j\underset{\sim}{Y}_j') > \varepsilon_j, \quad (37)$$

$$\Delta_j = 0, \quad \text{otherwise},$$

is found in the ith column.

The i×p matrix $\underset{\sim}{U}_j$ is defined as

$$\underset{\sim}{U}_j = [0,0,0,\ldots,0,1,0,\ldots,0], \quad (38)$$

where unity element is found in the jth column.

The set of equations (32) through (35) is executed p-1 times, beginning with j=2 and i=1. Each time these equations are executed, the subscript j is incremented by one. Equations (33) and (35) are only defined in the event that the indicated condition is true, where ε_j, j=1,p is ordinarily specified to account for small computational errors. Following execution of equation (34), the subscript i is incremented by one only if the indicated condition is true.

From equation (32) and the discussion of least squares, it is clear that $\underset{\sim}{E}_j$ represents the component of $\underset{\sim}{Y}_j$ which is orthogonal to the space spanned by $\underset{\sim}{Z}_i$.

THE MODIFIED SCHMIDT FILTER

We now consider an alternate version of the Schmidt filter algorithm which is extremely valuable in adaptive filtering and the analysis of the noisy case. In this development, we show that it is never necessary to know $\underset{\sim}{Y}$ or $\underset{\sim}{X}$ to determine the filter pair $(\underset{\sim}{S},\underset{\sim}{S}^+)$, but only the Gram matrix of $\underset{\sim}{Y}$, defined by

$$\underset{\sim}{G} = \underset{\sim}{Y}\underset{\sim}{Y}', \tag{39}$$

where $\underset{\sim}{G}$ is a p×p matrix. Let $\underset{\sim}{G}_j$, j=1,p denote the jth row of matrix $\underset{\sim}{G}$ and g_{ji}, j=1,p, i=1,p denote the elements of matrix $\underset{\sim}{G}$. From (39) we notice that

$$g_{ji} = \underset{\sim}{Y}_j\underset{\sim}{Y}_i' \tag{40}$$

$$\underset{\sim}{G}_j = \underset{\sim}{Y}_j\underset{\sim}{Y}'. \tag{41}$$

Let $\underset{\sim}{T}_i$, i=1,n denote the <u>first i</u> rows of the matrix $\underset{\sim}{S}$. From previous discussion, it is clear that

$$\underset{\sim}{Z}_i = \underset{\sim}{T}_i\underset{\sim}{Y}. \tag{42}$$

Using equation (42) with equation (32), we can show that

$$\underset{\sim}{E}_j\underset{\sim}{E}_j' = \underset{\sim}{Y}_j\underset{\sim}{Y}_j' - [\underset{\sim}{Y}_j\underset{\sim}{Y}'\underset{\sim}{T}_i'][\underset{\sim}{Y}_j\underset{\sim}{Y}'\underset{\sim}{T}_i']'. \tag{43}$$

Using equations (40) and (41) with (43) and (32), we have

$$\underset{\sim}{E}_j\underset{\sim}{E}_j' = g_{jj} - [\underset{\sim}{G}_j\underset{\sim}{T}_i'][\underset{\sim}{G}_j\underset{\sim}{T}_i']' \tag{44}$$

$$\underset{\sim}{S}_j^+ = [\underset{\sim}{G}_j\underset{\sim}{T}_i',0,0,0,\ldots,0] + \underset{\sim}{Q}_{i+1}(j). \tag{45}$$

Using equations (30), (31), (35), (44), and (45), we now give the complete set of filter generating equations as

$$\underset{\sim}{S}_1^+ = [g_{11}^{1/2},0,0,0,\ldots,0] \tag{46}$$

$$\underset{\sim}{S}_1 = [1,0,0,0,\ldots,0]/g_{11}^{1/2} \tag{47}$$

$$\underset{\sim}{E}_j\underset{\sim}{E}_j' = g_{jj} - [\underset{\sim}{G}_j\underset{\sim}{T}_i'][\underset{\sim}{G}_j\underset{\sim}{T}_j']' \tag{48}$$

$$\underset{\sim}{S}_j^+ = [\underset{\sim}{G}_j\underset{\sim}{T}_i',0,0,0,\ldots,0] + \underset{\sim}{Q}_{i+1}(j) \tag{49}$$

$$\underset{\sim}{S}_{i+1} = [\underset{\sim}{U}_j - \sum_{k=1}^{i} s_{jk}^+\underset{\sim}{S}_k]/||\underset{\sim}{E}_j\underset{\sim}{E}_j'||, \text{ if } [\underset{\sim}{E}_j\underset{\sim}{E}_j']/g_{jj} > \varepsilon_j, \tag{50}$$

where $\underset{\sim}{Q}_i(j)$ is defined by equation (36) and $\underset{\sim}{U}_j$ is defined by equation (38).

The algorithm described by equations (48) through (50) proceeds exactly as before except that no knowledge of $\underset{\sim}{Y}$ or $\underset{\sim}{X}$ is ever required. Of course, given $\underset{\sim}{Y}$, $\underset{\sim}{X}$ can be readily computed using the filter transform $\underset{\sim}{S}$.

The modified filter algorithm is more conveniently applied with a digital computer because of simplified logic and greatly reduced storage requirements. However, we note that no effort in computation is actually saved by using this form, since the Gram matrix of $\underset{\sim}{Y}$ was not required in the original algorithm. In fact, the computation effort required is approximately equal for both versions of the algorithm.

One of the most outstanding advantages of the modified algorithm is that it is very easily applied to problems of adaptive filtering since the Gram matrix $\underset{\sim}{Y}\underset{\sim}{Y}'$ is easily updated for added observations. Once the Gram matrix is defined anew, the computational effort required to recompute $\underset{\sim}{S}$ and $\underset{\sim}{S}^+$ is practically nil using equations (46) through (50). Finally, we note the interesting fact that, in general, there are an infinite number of functions $\underset{\sim}{Y}$ which produce the same Gram matrix $\underset{\sim}{G}$. Obviously, this fact implies that the pair $(\underset{\sim}{S},\underset{\sim}{S}^+)$ does not ordinarily determine the pair $(\underset{\sim}{Y},\underset{\sim}{X})$.

The applicability of the modified Schmidt algorithm to the noisy case will be discussed in the next section.

NOISE

Now, if $\underset{\sim}{Y}$ is the sum of a signal function and a noise function, as in (10), it is clear that $\underset{\sim}{X}$ will also be a "noise-corrupted" function, since $\underset{\sim}{X}$ is merely a linear transformation of $\underset{\sim}{Y}$. The implications of this observation are that the function $\underset{\sim}{X}$ is a "noisy basis" for the noisy function $\underset{\sim}{Y}$ and that the row dimension of $\underset{\sim}{X}$ does not necessarily reflect the rank of $\underset{\sim}{F}$. There is some hope that we could conceivably solve the rank determination problem by adjusting the ε_j upward to allow for "approximate dependence," but there certainly is no guarantee that this empirical procedure will produce the desired results, especially where the contribution of the noise process $\underset{\sim}{V}$ is large. Also, using the empirical method, it can easily develop that rank $(\underset{\sim}{X})$ < rank $(\underset{\sim}{F})$ in which case the set $\underset{\sim}{X}$ does not contain a complete set of basis functions of $\underset{\sim}{F}$. It is thus clear that any results obtained in this fashion are purely heuristic.

We now consider how the modified Schmidt algorithm, defined by equations (46) through (50), can be conveniently applied to the case where the noise and signal statistics are known. We show that, given $\underset{\sim}{Y}$ and the statistics $\underset{\sim}{F}\underset{\sim}{V}'$ and $\underset{\sim}{V}\underset{\sim}{V}'$, it is possible to determine the filters $\underset{\sim}{S}$ and $\underset{\sim}{S}^+$ associated with the function $\underset{\sim}{F}$.

Using equation (10), we write

$$\underset{\sim}{Y}\underset{\sim}{Y}' = \underset{\sim}{F}\underset{\sim}{F}'+\underset{\sim}{F}\underset{\sim}{V}'+[\underset{\sim}{F}\underset{\sim}{V}']'+\underset{\sim}{V}\underset{\sim}{V}' \qquad (51)$$

Thus, given $\underset{\sim}{Y}\underset{\sim}{Y}'$, $\underset{\sim}{F}\underset{\sim}{V}'$, and $\underset{\sim}{V}\underset{\sim}{V}'$, it is trivial to compute the statistic $\underset{\sim}{F}\underset{\sim}{F}'$ using equation (51). Hence, we assume the statistic $\underset{\sim}{F}\underset{\sim}{F}'$ to be known. In order to apply the Schmidt filter to the noisy case, we now use the modified form to determine $\underset{\sim}{S}$ and $\underset{\sim}{S}^+$ based on $\underset{\sim}{F}\underset{\sim}{F}'$. This allows exact determination of rank $(\underset{\sim}{F})$. Further, if $\underset{\sim}{F}$ were the input to such a filter, then the filter output would obviously be a complete set of orthonormal basis functions of $\underset{\sim}{F}$. We note, however, that in this case the filter $\underset{\sim}{S}$ applied to $\underset{\sim}{Y}$, which is known, does not ordinarily produce an orthonormal matrix $\underset{\sim}{X}$, although it is easy to see that $\underset{\sim}{X}$ is complete, i.e., it contains a complete set of linearly independent components of $\underset{\sim}{F}$. Further, if we consider that $\underset{\sim}{X}$ is the sum of a "signal function" $\underset{\sim}{X}_S$ and a "noise function" $\underset{\sim}{X}_N$, i.e.,

$$\underset{\sim}{X} = \underset{\sim}{X}_S + \underset{\sim}{X}_N, \tag{52}$$

we can compute the statistics $\underset{\sim}{X}_S\underset{\sim}{X}_N'$ and $\underset{\sim}{X}_N\underset{\sim}{X}_N'$ from $\underset{\sim}{F}\underset{\sim}{V}'$ and $\underset{\sim}{V}\underset{\sim}{V}'$ using equation (10) as

$$\underset{\sim}{X}_S\underset{\sim}{X}_N' = \underset{\sim}{S}\underset{\sim}{F}\underset{\sim}{V}'\underset{\sim}{S}', \tag{53}$$

and

$$\underset{\sim}{X}_N\underset{\sim}{X}_N' = \underset{\sim}{S}\underset{\sim}{V}\underset{\sim}{V}'\underset{\sim}{S}'. \tag{54}$$

We now consider the application of the above results to the identification of systems in noisy environments.

THE IDENTIFICATION PROBLEM

We direct the reader's attention to a problem considered earlier by the Authors [1]. The distinction between this and the previous analysis is that we now seek to obtain an optimal solution of the identification problem for the case where noise statistics are known. Here we show that unbiased estimates of system order and system parameters are easily obtained using the modified Schmidt filter algorithm.

It is important to note that the meaning of symbols used in this section does not necessarily correspond with definitions in previous sections.

We begin by assuming that we have an "external description" of a minimal realization $\underset{\sim}{\Sigma}$, where $\underset{\sim}{\Sigma}$ is described by

$$\underset{\sim}{x}(k+1) = \underset{\sim}{A}\underset{\sim}{x}(k)+\underset{\sim}{B}\underset{\sim}{u}(k) \tag{55}$$

and

$$\underset{\sim}{y}(k) = \underset{\sim}{C}\underset{\sim}{x}(k)+\underset{\sim}{D}\underset{\sim}{u}(k), \tag{56}$$

where $\underset{\sim}{x}$ is a $n\times 1$ state vector, $\underset{\sim}{u}$ is a $m\times 1$ input vector, and $\underset{\sim}{y}$ is a $p\times 1$ output vector. The constant matrices $\underset{\sim}{A}$, $\underset{\sim}{B}$, $\underset{\sim}{C}$, and $\underset{\sim}{D}$ are of orders $n\times n$, $n\times m$, $p\times n$, and $p\times m$, respectively. The integer k indicates observation of the system at the beginning of the kth equal interval of time. We assume that the external description is given by the sequences

$$\underset{\sim}{Y} = [\underset{\sim}{y}(1),\underset{\sim}{y}(2),\underset{\sim}{y}(3),\ldots,\underset{\sim}{y}(N+r)] \tag{57}$$

and

$$\underset{\sim}{U} = [\underset{\sim}{u}(1),\underset{\sim}{u}(2),\underset{\sim}{u}(3),\ldots,\underset{\sim}{u}(N+r)] \tag{58}$$

where the significance of N and r is shown later. The functions $\underset{\sim}{Y}$ and $\underset{\sim}{U}$ are assumed to be noise-free.

In association with the minimal realization $\underset{\sim}{\Sigma}$ we require the existence of the sequence of states $\underset{\sim}{X}$ where

$$\underset{\sim}{X} = [\underset{\sim}{x}(1),\underset{\sim}{x}(2),\underset{\sim}{x}(3),\ldots,\underset{\sim}{x}(r)] \tag{59}$$

such that $r > n+p$ (60)

and rank $(\underset{\sim}{X}) = n.$ (61)

Since all minimal realizations are observable, we require that

$$\text{rank } (\underset{\sim}{K}) = n, \tag{62}$$

where[3]

$$\underset{\sim}{K} = \begin{bmatrix} \underset{\sim}{C} \\ \underset{\sim}{C}\underset{\sim}{A} \\ \underset{\sim}{C}\underset{\sim}{A}^2 \\ \cdot \\ \cdot \\ \cdot \\ \underset{\sim}{C}\underset{\sim}{A}^N \end{bmatrix}, \forall_{N \geq n}. \tag{63}$$

Let the $(N+1)p \times r$ matrix $\underset{\sim}{W}$ be defined as

$$\underset{\sim}{W} = \begin{bmatrix} \underset{\sim}{y}(1) & \underset{\sim}{y}(2) & \underset{\sim}{y}(3) & \cdots & \underset{\sim}{y}(r) \\ \underset{\sim}{y}(2) & \underset{\sim}{y}(3) & \underset{\sim}{y}(4) & \cdots & \underset{\sim}{y}(r+1) \\ \underset{\sim}{y}(3) & \underset{\sim}{y}(4) & \underset{\sim}{y}(5) & \cdots & \underset{\sim}{y}(r+2) \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \underset{\sim}{y}(N+1) & \underset{\sim}{y}(N+2) & \underset{\sim}{y}(N+3) & \cdots & \underset{\sim}{y}(N+r) \end{bmatrix}, \tag{64}$$

and the $(N+1)p \times r$ matrix $\underset{\sim}{V}$ be defined as

---

[3]Note that the last p rows of $\underset{\sim}{K}$ are always linearly dependent on the first Np rows in the definition employed here.

$$\underset{\sim}{V} = \begin{bmatrix} \underset{\sim}{u}(1) & \underset{\sim}{u}(2) & \underset{\sim}{u}(3) & \cdots & \underset{\sim}{u}(r) \\ \underset{\sim}{u}(2) & \underset{\sim}{u}(3) & \underset{\sim}{u}(4) & \cdots & \underset{\sim}{u}(r+1) \\ \underset{\sim}{u}(3) & \underset{\sim}{u}(4) & \underset{\sim}{u}(5) & \cdots & \underset{\sim}{u}(r+2) \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \underset{\sim}{u}(N+1) & \underset{\sim}{u}(N+2) & \underset{\sim}{u}(N+3) & \cdots & \underset{\sim}{u}(N+r) \end{bmatrix} . \tag{65}$$

Also, let $\underset{\sim}{\alpha}_W$ denote the first column of $\underset{\sim}{W}$ and $\underset{\sim}{\omega}_W$ denote the last column of $\underset{\sim}{W}$. Let $\underset{\sim}{\omega}_V$ denote the last column of $\underset{\sim}{V}$. Further, let the matrices $\underset{\sim}{W}_1$, $\underset{\sim}{W}_2$, and $\underset{\sim}{V}_1$ be defined by

$$\underset{\sim}{W} = [\underset{\sim}{W}_1, \underset{\sim}{\omega}_W] = [\underset{\sim}{\alpha}_W, \underset{\sim}{W}_2] \tag{66}$$

$$\underset{\sim}{V} = [\underset{\sim}{V}_1, \underset{\sim}{\omega}_V] \tag{67}$$

Using equations (57) through (63) and (66) and (67), we can show that[4]

$$\underset{\sim}{T}\underset{\sim}{W}_2 = [\underset{\sim}{R}, \underset{\sim}{F}] \begin{bmatrix} \underset{\sim}{V}_1 \\ \hline \underset{\sim}{T}\underset{\sim}{W}_1 \end{bmatrix} \tag{68}$$

where $\underset{\sim}{F}$ is a n×n matrix, $\underset{\sim}{R}$ is a n×(N+1)m matrix and $\underset{\sim}{T}$ is a suitable n×(N+1)p "selector matrix." Basically, the selector matrix is allowed to have but one non-zero (unity) element per row and one non-zero (unity) element per column. The column position of each unity element of the selector matrix corresponds the column position of a row selected

---

[4]The reader is directed to reference [1] for proof of this statement.

from its operand. Further, it can be shown[5] that knowledge of the matrix $(\underset{\sim}{R},\underset{\sim}{F})$ is sufficient to yield a minimal realization of the external descriptions of $\underset{\sim}{\Sigma}$. Here, $\underset{\sim}{T}$ is assumed known, but it in fact must be determined indirectly from the external description. The problem reduces to the selection of a maximum number of linearly independent rows of $\underset{\sim}{W}$ which are themselves linearly independent of the rows of $\underset{\sim}{V}$. This problem was solved in [1] essentially by passing the matrix

$$\begin{bmatrix} \underset{\sim}{V}_1 \\ \hline \underset{\sim}{W}_1 \end{bmatrix} \tag{69}$$

through a Schmidt filter to determine which rows of $\underset{\sim}{W}_1$ are "passed" by the filter as linearly independent, leading to direct determination of $\underset{\sim}{T}$. Following determination of $\underset{\sim}{T}$, the filter transfer function and orthonormal output were used in an efficient manner to determine $[\underset{\sim}{R},\underset{\sim}{F}]$.

This procedure is certainly applicable to noise-free linear systems and is perhaps one of the best available where noise statistics are unknown, even though it is acknowledged that parameter estimates may be biased for the noisy case.

---

[5] Ibid. [1]

A Direct Method

The procedure developed here is based on our original analysis except that the procedure discussed here allows for much more convenient digital implementation, especially for the noisy case.

We begin by requiring that the last p columns of the selector matrix $\underset{\sim}{T}$ be trivial. This specification does not in any way restrict the analysis, since it is easily determined from equation (63) and footnote 3 that the last p rows of $\underset{\sim}{T}$ of equation (69) will be linearly dependent on the first (N+1)m+Np rows.

In order to implement the new method, we begin by passing the matrix

$$\begin{bmatrix} \underset{\sim}{V}_1 \\ \hline \underset{\sim}{W}_1 \end{bmatrix}$$

through a Schmidt filter to produce matrices $\underset{\sim}{S}$, $\underset{\sim}{S}^+$ and orthonormal output $\underset{\sim}{Q}$, where

$$\underset{\sim}{Q} = \underset{\sim}{S} \begin{bmatrix} \underset{\sim}{V}_1 \\ \hline \underset{\sim}{W}_1 \end{bmatrix} \tag{70}$$

$$\begin{bmatrix} \underset{\sim}{V}_1 \\ \hline \underset{\sim}{W}_1 \end{bmatrix} = \underset{\sim}{S}^+ \underset{\sim}{Q} \tag{71}$$

$$\underset{\sim}{Q}\underset{\sim}{Q}' = \underset{\sim}{I}. \tag{72}$$

We now partition $\underset{\sim}{S}$ as

$$\underset{\sim}{S} = [\underset{\sim}{S}_{(N+1)m}, \underset{\sim}{S}_{(N+1)p}] \tag{73}$$

where the column dimensions of the partitions of $\underset{\sim}{S}$ are given by their respective subscripts. We note that certain of the columns of $\underset{\sim}{S}_{(N+1)p}$ may be trivial (those corresponding to rows of $\underset{\sim}{W}$ which are "blocked" by the filter because of linear dependence). As in our previous analysis [1], we formulate $\underset{\sim}{T}$ based upon the occurence of non-trivial columns of $\underset{\sim}{S}_{(N+1)p}$.

We let the non-zero selection elements of $\underset{\sim}{T}$ be placed in the columns of $\underset{\sim}{T}$ which correspond to the non-trivial columns of $\underset{\sim}{S}_{(N+1)p}$. We note from previous discussion that the last p columns of $\underset{\sim}{T}$ are trivial after this method. As a result of this procedure, it is easy to show that

$$\underset{\sim}{S}_{(N+1)p}\underset{\sim}{T}'\underset{\sim}{T} = \underset{\sim}{S}_{(N+1)p} \tag{74}$$

We now form a second $n\times(N+1)p$ selector matrix $\underset{\sim}{T}^*$ from the relation

$$\underset{\sim}{T}^*\underset{\sim}{W}_1 = \underset{\sim}{T}\underset{\sim}{W}_2. \tag{75}$$

In particular, we form $\underset{\sim}{T}^*$ by shifting the selection elements of $\underset{\sim}{T}$ to the right by p positions. We note that the last p columns of $\underset{\sim}{T}^*$ are not necessarily trivial. Using equations (70) and (71), we can write

$$\underset{\sim}{T}^*\underset{\sim}{W}_1 = [\underset{\sim}{0},\underset{\sim}{T}^*][\underset{\sim}{S}^+\underset{\sim}{S}]\begin{bmatrix}\underset{\sim}{V}_1\\ \hline \underset{\sim}{W}_1\end{bmatrix} \tag{76}$$

where $\underset{\sim}{0}$ is a $n\times(N+1)m$ null matrix.

The matrix $\underset{\sim}{S}^+$ is now partitioned as

$$\underset{\sim}{S}^+ = \begin{bmatrix}\underset{\sim}{S}^+_{(N+1)m}\\ \hline \underset{\sim}{S}^+_{(N+1)p}\end{bmatrix} \tag{77}$$

where $\underset{\sim}{S}^+_{(N+1)m}$ represents the first $(N+1)m$ rows of $\underset{\sim}{S}^+$ and

$S^+_{(N+1)p}$ represents the last (N+1)p rows of S^+.

Using (73), (74), (75), and (77) with (76), we have

$$TW_2 = [T^*S^+_{(N+1)p}S_{(N+1)m}, T^*S^+_{(N+1)p}S_{(N+1)p}T'] \begin{bmatrix} V_1 \\ --- \\ W_1 \end{bmatrix} \tag{78}$$

where $T^*S^+_{(N+1)p}S_{(N+1)m}$ is a $n\times(N+1)m$ matrix and $T^*S^+_{(N+1)p}S_{(N+1)p}T'$ is a $n\times n$ matrix. Comparison of (78) with (68) reveals that

$$R = T^*S^+_{(N+1)p}S_{(N+1)m} \tag{79}$$

$$F = T^*S^+_{(N+1)p}S_{(N+1)p}T' \tag{80}$$

Further, the matrix $S^+_{(N+1)p}S_{(N+1)m}$ is simply the matrix defined by the last (N+1)p rows and the first (N+1)m columns of S^+S, and that the matrix $S^+_{(N+1)p}S_{(N+1)p}$ is the matrix defined by the last (N+1)p rows and last (N+1)p columns of S^+S. Also, it is clear that we can substitute W for W_1 and V for V_1 in equation (76) and still obtain the results indicated by equations (79) and (80).

We now summarize the complete realization procedure as

1. Form the matrix ZZ', where

$$Z = \begin{bmatrix} V \\ --- \\ W \end{bmatrix}. \tag{81}$$

2. Use the modified Schmidt algorithm with ZZ' to obtain S and S^+.
3. Form T from S and, hence T^*.
4. Use T and T^* to select $[R,F]$ from S^+S as in equations (79) and (80).

5. Use $[\underset{\sim}{R},\underset{\sim}{F}]$ to obtain a minimal realization of $\underset{\sim}{\Sigma}$ via the method suggested in reference [1].

We believe this to be the most general and most efficient realization procedure published to date. Also, we note that Steps 1 through 5 are easily implemented in an adaptive algorithm.

Finally, we note that all of the coefficients in the "external" form derived by Rowe [15] are easily selected from the matrix $\underset{\sim}{S}^{+}\underset{\sim}{S}$. Such an external form was determined from $\underset{\sim}{S}^{+}\underset{\sim}{S}$ by the Authors in reference [1] for the case of autonomous systems. Extension to the more general class of systems discussed in this paper is straightforward.

Additive Noise

We now show that the problem of additive noise does not affect the accuracy of or appreciably increase the computational labor involved with application of the previously discussed identification procedure for the case where noise statistics are known.

We let $\underset{\sim}{Z}$ be defined exactly as in equation (81) where $\underset{\sim}{W}$ and $\underset{\sim}{V}$ are given by equations (64) and (65), respectively. We let the unknown $(N+1)(m+p)\times r$ matrix $\underset{\sim}{\eta}$ be the noise matrix associated with attempted observation of $\underset{\sim}{Z}$, such that

$$\underset{\sim}{H} = \underset{\sim}{Z}+\underset{\sim}{\eta} \tag{82}$$

where $\underset{\sim}{H}$ is the $(N+1)(m+p)\times r$ matrix of noisy observations. For further analysis, we require that the statistics $\underset{\sim}{\eta}\underset{\sim}{\eta}'$ and $\underset{\sim}{Z}\underset{\sim}{\eta}'$ be known. From equations (64) and (65), it is clear that this requirement implies knowledge of an extensive set of autocorrelations and crosscorrelations of signal and noise processes.

Now, if $\underset{\sim}{\eta}\underset{\sim}{\eta}'$ and $\underset{\sim}{Z}\underset{\sim}{\eta}'$ are known, it is trivial to compute $\underset{\sim}{Z}\underset{\sim}{Z}'$ from equation (82). It should be clear that we can now use $\underset{\sim}{Z}\underset{\sim}{Z}'$ as in the direct identification method to obtain an unbiased estimate of $[\underset{\sim}{R},\underset{\sim}{F}]$ and hence an unbiased minimal realization for the noisy data.

CONCLUDING REMARKS

The Authors have shown that unbiased estimates of linear, multi-variable system parameters can be obtained for the case where signal and noise statistics are known. In some respects our approach to the problem has been from a "deterministic" point of view in that such terms as "expected value," "stationary," "ergodic," etc., have not been introduced into the analysis. Instead, we have preferred to approach the problem from what might be called a "small sample" point of view. We note for this case that any approximations to signal and noise statistics may seriously bias results. In fact, it certainly can happen that results obtained in this manner can be worse than those obtained by assuming that no prior knowledge of noise statistics is available, for which case the original Schmidt algorithm might be employed.

The Authors realize only too well that in a practical problem the noise and signal statistics may only be approximately specified. In some cases such approximations may be quite reasonable. For instance, it may be concluded that the noise and signal statistics are specified in the limit as the sample size becomes large.[6] For these cases it is easily deduced from previous discussion that the process described in this paper yields results which are "asymtotically unbiased."

Finally, we note that a modular collection of PL/1 subroutines which allow efficient structuring of a variety of digital realizations (including the adaptive case) of the procedures described in this paper is given in [1].

ACKNOWLEDGMENTS

The Authors wish to express their gratitude to Dr. J. S. Pazdera and Dr. R. T. Johnson for their assistance in the preparation of this work.

---

[6] A common assumption is that the signal and noise processes are uncorrelated.

REFERENCES

1. Behring, A. G., "Methods and Applications of System Identification," Ph.D. thesis, University of Missouri - Rolla, Rolla, Mo., Jan., 1971.

2. Epstein, B., Orthogonal Families of Analytic Functions, MacMillan, N.Y., 1965.

3. Drygas, H., The Coordinate-Free Approach to Gauss-Markov Estimation, Springer-Verlag, Berlin, 1970.

4. Kitamori, T., "Applications of Orthogonal Functions to the Determination of Process Dynamic Characteristics and to the Construction of Self-Optimizing Control Systems," Proc. I.F.A.C., Moscow, 1960, pp. 613-617.

5. Lubbock, J. K., "Self-Optimizing Filters Employing Orthogonal Systems of Functions," J. Electron. and Contr., Vol. 13, 1962, p. 365.

6. Clement, P. R., "On Completeness of Basis Functions Used for Signal Analysis," S.I.A.M. Rev., Vol. 5, No. 2, Apr., 1963, pp. 131-139.

7. Lubbock, J. K., and Barker, H. A., "A Solution of the Identification Problem," J.A.C.C. preprints, Minneapolis, 1963, pp. 191-199.

8. Douce, J. L., and Roberts, P. D., "Orthogonal Functions," Control, Vol. 8, Sept. 1964, pp. 457-461.

9. Barker, H. A., and Hawley, D. A., "Orthogonalizing Techniques for System Identification and Optimization," Proc. I.F.A.C., London, 1966, pp. 36A.1-36A.10.

10. Roberts, P. D., "Self-Adjustable Orthogonal Digital Filters for System Identification and Optimization," Proc. I.E.E., Vol. 114, No. 5, May 1967, pp. 666-670.

11. Gopinath, B., "On the Identification of Linear Time-Invariant Systems From Input-Output Data," Bell System Tech. J., Vol. 48, No. 5, May 1969, pp. 1101-1113.

12. Budin, M. A., "Minimal Realization of Discrete Linear Systems From Input-Output Observations," I.E.E.E. Trans. Auto. Contr., AC-16, No. 5, Oct. 1971, pp. 395-401.

13. Sage, A. P., and Melsa, J. L., Estimation Theory With Applications to Communications and Control, McGraw-Hill, N.Y., 1970.

14. Rao, C. R., "A Note on the Generalized Inverse of a Matrix With Applications to Problems in Mathematical Statistics," J. Royal Statistical Soc., Ser. B, Vol. 24, 1962, pp. 152-158.

15. Rowe, I. H., "A Bootstrap Method for the Statistical Estimation of Model Parameters," Int. J. Contr., Vol. 12, No. 5, Nov. 1970, pp. 721-738.

ON STOCK MARKET DYNAMICS

by

A. G. Behring

and

V. J. Flanigan*

ABSTRACT

The mathematical analysis of security prices is traced from its early beginnings through the current state-of-the art. The concept of a multiple input/output dynamic system is presented from the viewpoint of modern control theory and examined for relevance to the simulation of speculative prices. An efficient procedure for obtaining adaptive realizations from input/output data is briefly described. The stock market is formulated as a linear, discrete-time, multiple input/output system and the results of several simulation studies are presented. Evidence indicates that at least some segments of the market can be approximated by high-order linear systems computed from small samples and tends to refute the random walk hypothesis.

---

*The Authors are associated with the Department of Mechanical and Aerospace Engineering, University of Missouri - Rolla, Rolla, Missouri 65401, where Mr. Behring is a Graduate Student and Dr. Flanigan (member ASME) is an Associate Professor.

NOTATION

In this paper all bold-face capital letters denote matrices. Vectors are defined in column format and are denoted by lower case letters in bold-face type. All scalars will be denoted by plain upper or lower case letters. Occasionally it will be necessary to display the format of a vector or matrix explicitly, e.g.,

$$\underset{\sim}{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}.$$

Any exceptions to these general rules will be clearly specified in the text.

INTRODUCTION

Speculative Prices

It is generally recognized that the first mathematical analysis of speculative prices was undertaken by Bachelier [1].[1] Because it reflects significantly upon the problem

---

[1]Numbers in brackets refer to references at end of paper.

at hand and has particular relevance to later analysis, we quote a selected portion of the introduction to Bachelier's (1900) doctoral dissertation in which he gives his interpretation of the French bond market:

> The influences which determine fluctuations on the Exchange are innumerable; past, present and even discounted future events are reflected in market price, but often show no apparent relation to price changes. Besides the somewhat natural causes of price changes, artificial causes also intervene: the Exchange reacts on itself, and the current fluctuation is a function not only of previous fluctuations, but also of the orientation of the current state....

Bachelier's mathematical treatise is considered to be the foundation of what is referred to as the "random walk" hypothesis of stock price behavior. Bachelier proposed a model of speculative prices in which successive changes in price were considered as independent events, drawn from a Gaussian population with zero mean. In a description of what was later to be called a "perfect" market[2] by Working [2], Bachelier clearly inferred that the best estimate of future price is always the current price.

---

[2]The "perfect" market is one in which expectations concerning future economic conditions are reflected by current price. Expectations are in turn influenced by incoming information which is rapidly disseminated to market participants.

Contrary to the propositions of Bachelier, there appears to have consistently been a cult of professional traders who believe that existing knowledge in fact sheds light on future market movements. This group may be further divided into "technicians" and "fundamentalists," where the technicians believe that future prices are determined by the existing history of prices and the fundamentalists tend to examine associated information as well, such as profits, dividends, capital investment, etc., in hopes of predicting future price movements. It appears that a large number of the methods employed by the technicians and fundamentalists are highly empirical in nature (some being based substantially on intuition), which is not to conclude that they are without value. Cowles [3], however, investigated the historical predictive performance of eleven financial publications and concluded that these professionals were powerless (on the basis of realized profit) to predict market movements. Unfortunately, it is likely that many of the methods employed by the professional forecasters have gone unpublished, exceptions being [4,5,6].

It appears that vigorous interest in mathematical analysis of security prices was not revived until the 1950's. Kendall [7] examined a wide variety of economic series, drawn from the British Stock Exchange, and found no substantial indications of non-randomness using methods of multiple regression and autocorrelation. The random walk hypothesis was revived by a number of authors during this period, including Osborne [8]. Osborne [8], in the tradition of Bachelier, concluded from observed data that common stock prices tend to follow a random walk. However, Osborne chose not to examine the distribution of simple price differences, but the distribution of differences in the logarithms of successive prices (logged price relatives). The adjustment in the data tends to "normalize" results drawn from extended intervals where the magnitude of price changes appears to vary with price. Also, the possibility of "non-positive" prices is conveniently excluded from consideration. Roberts [9] gave haunting graphical examples showing how familiar series of stock price data might be generated by a random walk model. Additional support for the random walk model is given by Larson [10], employing the "perfect market" hypothesis of Working [2] with the assumption that "new" information arrives at the marketplace in a random fashion. It is worthy of note that Larson uses a statistical test of continuity as a basis for his results rather than the frequently-used techniques of autocorrelation and regression. The findings of Larson were later largely substantiated by Samuelson [11].

It appears that the first substantive challenge to random walk proponents was issued by Alexander [12], who concluded that "self-supporting" trends existed in stock prices and proposed what he called an x percent filter to profitably exploit such trends. Alexander [13] described the essential character of his filter as:

> Choose a percentage x, say 5%. If the average rises by 5%, buy, and if it declines by 5%, sell.

Alexander [12] showed that his simple strategy yielded eye-popping yearly profits (before commission) when applied to Dow Jones and Standard and Poor's Industrial closing price data from 1897 through 1959, from which it might be inferred that the random walk theory left something to be desired. Alexander's original work [12] drew considerable criticism from random walk proponents, much of which is reflected in works by Fama and Blume [14], and Dryden [15]. Accepting some of his antagonists criticisms, dealing primarily with the application of this filter rule, Alexander [13] was still able to show that his filter produced noticeable profits, although resultant yields were considerably reduced. Following Alexander's original work, several papers appeared which tended to reject the idea that a simple random walk model could completely explain observed phenomena. Osborne [16] concluded that significant time-periodic frequencies are present in the moments of stock price changes, implying non-stationarity. Cootner [17] proposed a modified random walk model of stock price differences which consisted of a random walk with "reflecting barriers" superimposed on a large number of short-term trends. Steiger [18] noted significant non-randomness in stock price behavior with statistical results which tended to support the hypothesis of Cootner [17]. Granger [19,20] applied the technique of spectral analysis to stock prices and concluded that low frequencies (periods of 6, 12, 40 months) are noticeably significant, while higher frequencies (periods of days and weeks) contribute little spectral power.

A tendency has been found by various investigators in random walk theory for the distribution of price differences and logged price relatives to be leptokurtic (similar to a Gaussian distribution, but having a significantly "higher peak" and "longer tails") rather than normal. It appears, however, that the first thorough analysis of this phenomenon was due to Mandelbrot [21,22]. Mandelbrot proposed that the observed changes in stock prices were actually drawn from a distribution called "stable Paretian," which explained the characteristics of observed frequency functions. Mandelbrot's hypothesis was extremely controversial because of the properties of the "stable Paretian" distribution, which has finite mean but infinite variance, a feature which directly challenged the applicability of most statistical techniques previously used. The validity of Mandelbrot's hypothesis and the properties of the stable Paretian distribution were later considered in detail by Fama [23,24]. It is noteworthy that Press [25], in the tradition of random walk, concluded later that the distribution of logged price relatives could be described by a Poisson mixture of normal distributions, which is not stable.

A recent movement in the literature concerns the distribution of individual stock price differences over intervals other than time, an idea actually first considered by Alexander [12]. Brada [26] considered the distribution of stock price differences taken over successive transactions and found notable dependence. Brada also concluded that such methods result in distributions which are more nearly normal than the distribution of time price differences. About the same time, Niederhoffer and Osborne [27] also noted such dependency over transactions and observed a tendency for stock price reversals to cluster at integer prices. Elsewhere, Niederhoffer [28] also examined the clustering phenomenon and noted that the mechanics of stock trading appear to result in a tendency for stock prices to be concentrated at integers. Granger [29] and Simmons [30] tended to reconfirm the dependence-over-transactions theory and defend the random walk hypothesis for longer intervals. Simmons, however, concluded that price dependence over transactions is generated by the actual mechanism of trading, and that the underlying price change sequence is in fact a random walk.

It appears to the Authors that a paper by Osborne [31] marks the beginning of another trend in the literature which we believe has yet to fully emerge. Osborne [31] viewed the stock market as an automatic control system and derived a discrete difference equation to describe its dynamics. Inputs to the model consisted of orders to buy and sell and outputs of the model consisted of a sequence of stock prices. While the system (black box) represented the mechanics of the market, a feedback loop is included to represent the activity of the market participants who view the system output and generate new orders to buy and sell. Although a specific model is never constructed, Osborne [31] evaluated the results qualitatively from the properties of the difference equations and found them to be "in qualitative agreement with the folklore of stock trading." Later, Osborne [32] compared the results "expected" from the random walk model, the discrete model and a continuous differential equation model with actual price and volume series data using a statistical theory of "coincident events." "Positive agreement" with historical Dow Jones Industrials and New York Stock Exchange total volume data was found for both the random walk model and the discrete model for differencing intervals on the order of days and weeks.

Osborne's papers are significant because they make the first attempt at a description of the mechanics of the stock market as a mathematical entity. Just as important, however, is the fact that they tend to be descriptive of the effects of new types of information on the outputs (e.g., price) of the trading mechanism. It has generally been recognized [21,24,25,30] that the stock market adjusts rapidly to new information and manifests such adjustments as stock price. Recently, several papers have appeared which attempt to assess the significance of certain types of new information on price adjustments. Although these publications tend to be more statistical and less mathematically descriptive than the recent works of Osborne, they contribute an important body of knowledge to the understanding of the market mechanism.

Ying [33] considered the relationship between stock market price and volume. Using daily Standard and Poor 500 closing price data, and NYSE total volume data, he concluded that not only are volume and price related, but that volume is an important indicator of future price changes. Ying [33], describing his findings, states:

> The relationships between stock prices and volumes of sales are examined with the view that they are joint products of a single market mechanism. The results found here tend to support the notion that any model of the stock market which separates prices from volumes, or vice versa, will inevitably yield incomplete if not erroneous results.

Fama [34] investigated the behavior or rates of return on individual stocks before and after the announcement date of a stock split by examining regression residuals. Summarizing some striking results, Fama concludes that rates of return (including price and dividends) tend to be high during the months preceding a stock split announcement, an effect which he attributed to "anticipation" of increased dividends. Using a similar statistical technique, Waud [35] analyzed the so-called "announcement effect" of Federal Reserve discount rate changes. Waud concluded from data on discount rate changes and the Standard and Poor's 500 stock price index that an announcement effect not only exists, but that a "public consensus" exists as to what economic future is indicated by a discount rate increase as compared to a discount rate decrease. Interestingly, Waud infers from his results that discount rate decreases tend to be anticipated several days prior to the earliest announcement date. Unfortunately, it is not possible to determine from Waud's results if an increase in the discount rate has a positive effect on expectations and vice versa, or whether the magnitude of a change is significant. In an interesting paper, Niederhoffer [36] has analyzed the effect of world news items (derived from headlines in the New York Times) upon the Standard and Poor's Composite Index. Niederhoffer and his fellow workers classified each "important" news item into one of twenty categories and rated each item on a seven point bad-good scale with respect to its apparent economic news content. Niederhoffer concluded that some types of world news items do have significant effect on stock prices in the several days following their appearance.

After a thorough examination of the mathematical literature on security prices, the Authors conclude that the random walk hypothesis continues to be dominant theory of stock price behavior (at least among academians). In fact it forms a basis for some of the more recent analyses of price adjustment [34,35] and factor analysis of economic series [37]. However, the Authors also note the continued publications of papers [38,39,40] concerning schemes of predicting price movements.

While the Authors have a great deal of respect for the hypotheses presented in the literature, we cannot conclude that a satisfactory explanation of stock price behavior has yet been presented. We close this section with a quote from Bliss [41] which we find appropriate in view of our findings:

> ...Furthermore, it is sometimes inferred that nature behaves in precisely the was which the mathematics indicates. As a matter of fact, nature never does behave in this way, and there are always more mathematical theories than one whose results depart from a given set of data by less than the errors of observation....

Dynamic Systems

We now examine a form for the representation of dynamic processes which has become popular during the last decade in the literature on automatic control. For basic reference, the Authors cite texts by Ogata [42] and Takahashi, Rabins and Auslander [43]. An advanced exposition of system theory is given by Kalman, Falb and Arbib [44].

Consider the representation of a dynamic system given by Figure 1. In this description, the u_i, i=1,m are known as system inputs, the y_i, i=1,p are known as system outputs, and the x_i, i=1,n are called state variables. The number n is said to be the "order" of the system. The u_i may be considered as independent (exogeneous) variables with respect to the system. The y_i represent dependent (endogenous) variables with respect to the system. Both the u_i and y_i may be classed as "observables" in that they may be measured directly (except possibly for observation error). In general, the x_i, which offer a complete description of the system at a given instant with respect to an n-dimensional Euclidean space, cannot be measured directly. However, under certain conditions (discussed later), the complete set can be inferred from input/output data.

Denoting t as a reference to time, the describing equations of a system are frequently written as

$$\dot{\underset{\sim}{x}}(t) = \underset{\sim}{f}[\underset{\sim}{x}(t),\underset{\sim}{u}(t),t], \tag{1}$$

and

$$\underset{\sim}{y}(t) = \underset{\sim}{g}[\underset{\sim}{x}(t),\underset{\sim}{u}(t),t], \tag{2}$$

where $\underset{\sim}{u}(t)$ is the $m \times 1$ input vector, $\underset{\sim}{y}(t)$ is the $p \times 1$ output vector, $\underset{\sim}{x}(t)$ is the $n \times 1$ state vector, and $\dot{\underset{\sim}{x}}(t)$ is the time-derivative of $\underset{\sim}{x}(t)$. Equation (1) is called the state equation and equation (2) is called the output equation. Equations (1) and (2), considered as a set, are called the dynamic equations of the system.

If a system is linear, equations (1) and (2) can be expressed as

$$\underset{\sim}{x}(t) = \underset{\sim}{F}(t)\underset{\sim}{x}(t) + \underset{\sim}{G}(t)\underset{\sim}{u}(t), \tag{3}$$

and

$$\underset{\sim}{y}(t) = \underset{\sim}{H}(t)\underset{\sim}{x}(t) + \underset{\sim}{I}(t)\underset{\sim}{u}(t), \tag{4}$$

where $\underset{\sim}{F}$, $\underset{\sim}{G}$, $\underset{\sim}{H}$, and $\underset{\sim}{I}$ are generally time-varying matrices of orders $n \times n$, $n \times m$, $p \times n$, and $p \times m$, respectively. A system described by equations (3) and (4) is said to be time-varying (non-stationary). If $\underset{\sim}{F}$, $\underset{\sim}{G}$, $\underset{\sim}{H}$, and $\underset{\sim}{I}$ are constant matrices, the system is said to be time-invariant (stationary).

In order to motivate the following discussion, let us assume that the word STOCK MARKET replaces SYSTEM in Figure 1. Further, assume that the input vector $\tilde{u}(t)$ represents the complete set of information which influences the stock market. Let the output vector $\tilde{y}(t)$ represent the complete set of observable information relative to market performance. An example of an input might be, e.g., the Federal Reserve discount rate (N.Y.), while an example of an output might be the price of Texaco common stock. The reader has doubtless already reasoned that the stock market is a rather abstract entity. Certainly many definitions could be formulated. For instance, one might theorize that it is appropriate to consider only the New York Exchange, or the American Exchange, or both taken together. In our analysis the choice of system inputs and outputs determines to a certain extent the system that is being considered.

The class of systems defined by equations (3) and (4) is quite extensive, though certainly not all-inclusive. Kalman [45] gives the basic theorems of controllability, observability and equivalence for this class. The condition that a system is completely observable requires that the complete state vector be determinable from input/output data. The concept of complete controllability requires, given the current system state, that the system can be driven to an arbitrarily-selected state in finite time by an appropriate choice of input values. Kalman [45] also demonstrated that any system of the form of (3) and (4) can be cannonically factored into four mutually exclusive parts which are

1. Completely controllable, but unobservable
2. Completely controllable and completely observable
3. Uncontrollable and unobservable
4. Uncontrollable, but completely observable.

Further, Kalman [45] determined the significant fact that the <u>only</u> part of a system that can be determined from input/output data is that part of the system which is completely controllable and completely observable. The concepts of controllability and observability seem to portend some interesting discussions by stock market theorists. Kalman's [45] definition of equivalence addresses the fact that the quartet (F,G,H,I) may not be uniquely defined by input/output data. In other words, there may be a number (possibly infinite) of choices of (F,G,H,I) which define a suitable (equivalent) input/output map. Generally, equivalent systems are related by a transformation of coordinates in the state space.

Frequently, a record of inputs and outputs associated with a system is called an "external description" of that system. A system of minimum order which is able to reproduce a specified input/output description is said to be a minimal realization of that external description. Minimal realizations are an expedient for obtaining mathematical models from input/output data that are suitable for purposes of prediction and optimal control. As such, minimal realizations may not be equivalent (in the sense described earlier) to the system which originated an input/output description. For most purposes, however, the set of all minimal realization may be considered as equivalent.

The Authors emphasize that a great volume of engineering literature has been devoted to the analysis of equations (3) and (4). This form is an extremely convenient framework for manipulations leading to determination of natural frequencies, controllability, observability, stability and policies of optimal control. Recent years have also seen an intensive investigation into the problem of identification (as in economics), which is a central issue of our current paper. In automatic control, the identification problem amounts to specification of F(t), G(t), H(t), I(t), and x(t) in equations (3) and (4). For summaries of the status of identification theory in engineering, the reader is directed to Cuenod and Sage [46], Eykhoff [47], and Nieman, Fisher and Seborg [48].

MATHEMATICAL MODELING

Basic Models

The model of particular concern here is the special case of equations (3) and (4) given by[3]

---

[3]We realize that difference equations per se are nothing new to the economist. We use the form shown here because it is convenient for interpretation.

$$\underset{\sim}{x}(k+1) = \underset{\sim}{A}\underset{\sim}{x}(k)+\underset{\sim}{B}\underset{\sim}{u}(k), \tag{5}$$

and

$$\underset{\sim}{y}(k) = \underset{\sim}{C}\underset{\sim}{x}(k)+\underset{\sim}{D}\underset{\sim}{u}(k), \tag{6}$$

where it is assumed that $\underset{\sim}{A}$, $\underset{\sim}{B}$, $\underset{\sim}{C}$, and $\underset{\sim}{D}$ are constant matrices of orders $n \times n$, $n \times m$, $p \times n$, and $p \times m$, respectively. It can be shown that equations (3) and (4) reduce to (5) and (6) under the assumption that $\underset{\sim}{F}$, $\underset{\sim}{G}$, $\underset{\sim}{H}$, and $\underset{\sim}{I}$ are constant matrices and that the system inputs are able to vary only in an instantaneous manner at the beginnings of fixed equal length intervals of time. The integer index k denotes reference to the system at the beginning of the kth equal interval of time. It is important to realize that the system itself can be continuous and still be consistent with (5) and (6) if the input restriction is satisfied. If the system is discrete-time, but non-stationary, its dynamic equations are ordinarily written as

$$\underset{\sim}{x}(k+1) = \underset{\sim}{A}(k)\underset{\sim}{x}(k)+\underset{\sim}{B}(k)\underset{\sim}{u}(k), \tag{7}$$

and

$$\underset{\sim}{y}(k) = \underset{\sim}{C}(k)\underset{\sim}{x}(k)+\underset{\sim}{D}(k)\underset{\sim}{u}(k), \tag{8}$$

where the restrictions on time-variation now apply to the coefficient matrices as well as inputs.

Gopinath [49] gave the first direct procedure for obtaining a minimal realization of (5) and (6) from input/output data by the method of least squares. Increased efficiency, generality and measurement noise were topics to be later considered by Budin [50] and Behring [51].

Behring gave a set of PL/1 subroutines for the realization procedure, which have the following features:

1. Procedures are modular and can be structured for a number of "off-line" and "on-line" applications.
2. Realization order need not be known a priori.
3. Updated estimates of system parameters are easily obtained as new input/output observations become available, with optional exponential weighting of "old" data.
4. Procedures allow for sequential changes in system structure, such as system order and numbers of inputs and outputs.
5. Highly correlated input and output variables can be handled conveniently.
6. Procedures are not affected by addative measurement noise if noise statistics are known.

Relation to Previous Work

We now give our interpretation of how the simple model defined by equations (5) and (6) might be used to formulate a realization of the stock market and examine the implications of such a model with respect to previous theory and results.

Assume that the stock market is realizable as a stationary, linear dynamic system of unknown order. Let $\underset{\sim}{u}$ represent the complete set of information which influences market performance, and let $\underset{\sim}{y}$ represent the complete set of observables indicating market performance. Further, assume that we can find an interval which is suitably small such that the discrete representation of equations (5) and (6) is valid. Finally, for convenience let the system given by equations (5) and (6) be denoted as $\underset{\sim}{\Sigma}$.

Under the preceding assumptions our model is not consistent with the random walk model proposed by Bachelier except for trivial cases and the assumption of random inputs. Examination of $\underset{\sim}{\Sigma}$ reveals that current information does influence the expected model output at the next sample time. However, $\underset{\sim}{\Sigma}$ is consistent with the theories of those who would interpret the statement that current information is not "sufficient" to predict future outputs, since the model output at the next sampling interval (k+1) is not determined completely by current information (k). Incidentally, as noted in our introduction, Bachelier did in fact offer a qualitative explanation of market behavior (other than random walk) which has some basic similarities with the model submitted here.

We also believe that experimental findings to date do not tend to either confirm or deny that the market can be represented, or at least approximated, by $\underset{\sim}{\Sigma}$. This opinion is based upon a supposition of high system order and many unknown (possibly random) inputs. Under these conditions, even the most powerful of current statistical tests would fail to determine the viability of $\underset{\sim}{\Sigma}$.

Output Specification

Assume that the components of $\underset{\sim}{y}$ represent individual stock prices. There may be some inclination, however, to use indices such as the Dow Jones 30 Industrials to form an "artificial" set of outputs, say $\underset{\sim}{y}^*$. Generally, there is no theoretical difficulty involved with this assumption as long as such indices are constructed as a constant linear weighting of individual stock prices over the interval used for model estimation. It can be observed from equation (6) that $\underset{\sim}{y}$ is simply a linear function of system state $\underset{\sim}{x}$ and the input $\underset{\sim}{u}$. Any linear combination $\underset{\sim}{y}^*$ of $\underset{\sim}{y}$ is still a linear combination of $\underset{\sim}{x}$ and $\underset{\sim}{u}$. However, there is some loss of generality in the minimal realization obtained in this manner, since if an index is used in lieu of individual stock prices, then those individual prices cannot be recouped from the model output. Also, if a restrictive index is used, such as the 30 Industrials, there is a possibility that the order of a minimal realization computed from the index $\underset{\sim}{y}^*$ and one computed using $\underset{\sim}{y}$ will not be of the same order. This case would arise if the linear vector space spanned by observations on $\underset{\sim}{u}$ and $\underset{\sim}{y}^*$ does not contain the space spanned by observations on $\underset{\sim}{u}$ and $\underset{\sim}{y}$. However, this is no problem if one is simply content to obtain a realization of $\underset{\sim}{y}^*$ rather than, say, the entire stock market. The same theoretical propositions can be advanced for trading volumes as long as volumes are linear functions of $\underset{\sim}{x}$ and $\underset{\sim}{u}$. In fact, in this case it is possible to combine price and volume data in a single index and still obtain a minimal realization of that index. In view of the above discussion, we reject the proposition of Ying [33] which effectively states that any valid model of price must include volume simply because volume is a joint product of the market mechanism. Failure to consider all outputs or aggregation of outputs does not result in an inability to realize those outputs which are considered.

Input Specification

The problem of input specification is crucial. Failure to recognize significant inputs and/or linear aggregation of inputs can lead to disastrous results. Observations of selected inputs must span the space generated by all relevant inputs, or estimates of system order, parameters and state will be biased. In fact, input specification error generally precludes the possibility of obtaining any realization. The only exception is where the set of all relevant inputs is a product of an autonomous linear system. Even though a realization can be obtained for this case, its order will not be minimal.

Other Considerations

Another problem in obtaining useful minimal realizations is that the state and/or input must be "sufficient" to excite all natural modes of the actual system at some time during the interval used to estimate the realization. If this is not the case, a realization can still be obtained, but can prove to be quite inappropriate on another interval. Finally, a non-arbitrary distinction between the contributions of two or more inputs cannot be made if it happens that those inputs are linearly dependent over the interval of estimation. Again, a minimal realization can be obtained for this case but can fail over another interval. These last items of discussion strongly suggest an adaptive model, i.e., one whose structure is updated for each "new" pair of input/output observations as they become available.

SIMULATION EXPERIMENTS

Data

Numeric data was derived from BARRON'S [52] and the Federal Reserve BULLETIN [53]. Information on world events was selected from FACTS ON FILE [54].

To the Authors, the year 1968 constitutes an extremely interesting period in stock price history. A brief analysis of price data reveals that 1968 began as a bear market with the Dow Jones Industrials falling about 9% in the first 30 trading days. After than substantial slide, followed by what appeared to be a 30-day period of uncertainty, the market became a powerful bull with the Dow Jones Industrials rising about 9% in the next 15 trading days. Following this abrupt surge, the Industrials began a slow oscillatory (and sometimes uncertain) rise to a near all-time high close of 985.21 on December 3rd.

The year of 1968 was also a period of great unrest at home and abroad. Economic news was mixed, with concern for the value of the dollar, 10% surtax, three changes in the Federal Reserve discount rate, one increase in stock margin requirements, and the usual concern over inflation. Politically, the year 1968 was volatile. Two assasinations of political leaders, an announced halt in bombing, the seizure of a spy ship and a presidential election grasped national attention.

Initial Results

The results of our first simulation attempts discussed here are not displayed graphically because of space limitations, but are discussed because they appear to be of some tutorial value.

Our first model of the New York Stock Exchange consisted of two inputs which were the Federal Reserve discount rate (N.Y.) and the percent stock margin requirement and four outputs which were the daily Dow Jones closing price averages: 30 Industrials, 20 Rails (now Transportation), 15 Utilities, and 40 Bonds. From this choice the Authors naively hoped to obtain a linear model in the form of equations (5) and (6) which would closely approximate the specified indices over the interval of estimation and would allow reasonable prediction for at least a short time beyond the estimation interval. One-hundred days from the beginning of 1968 were used to estimate model coefficients with the remaining data from 1968 used to check predicted results. Several model orders were tried, the lowest being 4th order and the highest 12th.

It was expected that the predictions yielded by the model from day to day over the interval of estimation would be fairly good, since the coefficients would probably at least to some extent be biased to the interval chosen. Our philosophy for initial testing was instead to start the model on the first day of the estimation interval and make recursive estimates of the above indices for each trading day assuming only that input quantities (discount rate and margin) were known. The results were striking! Even the fourth order model followed the oscillations of the market indices (which happened to be fairly distinct for this interval) and the Authors were plainly enthusiastic. We then started the model at several intermediate points on the interval with the same result -- the model seemed to know exactly where it was supposed to go. The margin did not change over this particular interval, but the discount rate (which changed twice) seemed to be contributing significantly to the quality of the model. The discount rate changes were both increases, the first from 4 1/2 to 5% and the second from 5 to 5 1/2%. In both cases the model predicted significant declines in all of the price indices for a period of about five days, with the modes excited by these changes quickly damping out, allowing the market to proceed on its destined path. Furthermore, comparison of predicted results with actual data seemed to verify predicted dynamics following discount rate changes. Our only misgivings were that the predicted industrials did not seem to be able to quite keep up with the great bullish rise which began about April 1.

Finally we started the model on the last day of the estimation interval and made recursive predictions, again assuming knowledge of the Federal Reserve discount rate (N.Y.). This time the graceful dynamics which the model displayed over the estimation interval was gone. The model seemed to quickly seek a level of equilibrium where it resided until excited later by another discount rate change (5 1/2 to 5 1/4%), the effects of which were dissipated rapidly. A quick check against actual data revealed that while the model generated a rather mundane performance, the actual stock market data continued to be oscillatory, a fact which led to the ultimate demise of our first modeling effort. Several increases in model order (from 4 through 12) improved the ability of the model to follow data on the estimation interval, but did not seem to significantly add to its predictive quality during the later interval.

We have stated this embarrassing result in order to illustrate some basic problems associated with time series analysis using regression methods. First, we cannot conclude from the above results that the model failed to explain (significantly) the dynamics of the market on the estimation interval. This is because we cannot guarantee that the dynamics of the market remain constant into the next 100 days. From experience, however, the Authors believe that the very act of parameter estimation "defines" a stable trajectory in space which the model tends to approach regardless of where it is started on the estimation interval. This could at least in part explain some of the "successes" noted earlier. Most importantly, however, we cannot conclude that a linear model of the market is inappropriate. It was found later that a 10th order model was required to reasonably describe the dynamics of the bond market alone. It is possible that our decision to "scrap" our original model based upon a 12th order approximation was a bit premature. However, if a linear model is indeed appropriate, the problem more likely is input specification. In retrospect, the two items of information chosen seem incapable of describing market adjustment phenomena completely.

Perhaps the most important conclusion to be drawn from the foregoing results is the fact that a model appears to "explain" the data on the interval of estimation does not necessarily indicate its successful use for predictions at a later time. We reason that the only conclusive evidence which can be obtained for this case is that derived from its performance on an independent interval.

Further Results

The Authors chose to be somewhat more conservative in future simulation studies of the stock market and concentrate on more restrictive models, hoping to glean some significant information as to the character of speculative prices. We chose to examine the dynamics of the market segment defined by the Dow Jones 40 Bonds, again using a constant-coefficient linear model. As before, data from 1968 was used in the analysis. For the results presented here we used the 15 Utilities closing price as the input to the system (N.Y.S.E. Bond Market?) that is implied (in the cannonical sense) by the choice of input and output. A complete set of closing price data for the 40 Bonds and 15 Utilities is included in the Appendix.

The schematic representation of the system is shown in Figure 2. In Figure 2 the unity input has the effect of introducing a constant into the state equations and allows for any appropriate translation in scale among the variables.

From their earlier investigations, the Authors noted, at least for the year 1968, that there was a general (although sometimes rather obscure) tendency for the 15 Utilities to "lead" the 40 Bonds. We ask the reader to examine the relationship between Figures 3 and 4. Figure 3 is a plot of the 15 Utilities (close) versus trading day for the first 40 trading days in our sample. Figure 4 is a plot of the 40 Bonds (close) for that same interval. We note that the time scale for this and all succeeding plots, where comparison is appropriate, is identical. Observe that both series rise rapidly to their maximum values within the first 15 days. The 15 Utilities reach their maximum value on day 10, while the 40 Bonds do not peak until day 14. We also observe that the relationship between the series up until day 15 can be nearly approximated by a "pure delay." However, following day 15 the pure delay hypothesis is evidently destroyed since the utilities drop quite rapidly while the bonds continue to "float" at nearly their peak value. There appears to be little noticeable correlation in the daily movements of the two series.

Since the indices seemed to have loosely similar properties, the Authors decided to investigate their mechanical relationship in greater detail. From [55] the Authors observed that during 1968 the 40 Bond averages were as usual computed as a linear weighting of 10 industrial bonds, 10 higher grade rails, 10 second grade rails, and 10 utilities. A more thorough analysis revealed that of the 40 Bonds, 15 were issued by parent firms that were also represented in the Dow Stock price indices (30 Industrials, 20 Rails, and 15 Utilities). Of the 15 firms whose index representation overlapped into the 40 Bonds, 3 were Industrials, 9 were Rails, and 3 were Utilities. Thus, we were able to determine that only 7.5% of the firms represented by the 40 Bonds had mutual representation in the 15 Utilities. Thus, the Authors chose to investigate the model shown in

Figure 2 with some uncertainty as to what the subsequent analysis would yield.

In view of earlier investigations, it was decided that no initial model testing would be done for the estimation interval itself, but that analysis would be done on the data drawn from a following "independent" interval. Several different model orders were postulated for the system shown in Figure 2, the lowest order model being 1 and the highest 12. However, all subsequent graphical results are for the 10th order case.

The first model was estimated from data given by Figures 1 and 2. This model was then started on the last day of the estimation interval and used to predict "into the future" (trading days past number 40) assuming that the 15 Utilities index was known. Figure 5 shows the activity of the utilities (model input) for the next 30 trading days. Figure 6 shows both the predicted model output and the actual data (not used to condition the model in any way) for that same interval. We note primarily that the model tends to follow the actual data quite well, in fact much better than indicated by chance. However, we also observe that there is a noticeable amount of co-movement between the input, output, and predicted output for this interval, warranting a closer inspection of the results. Actually, the Authors conclude that the model does not tend to "follow" the Utilities to a greater extent than it tends to follow the actual data. This is especially evident beyond day number 62 when the Utilities begin a substantial rise but the model "remains in the doldroms" for an additional 5 days and predicts (quite well, we might add) an upward swing in the Bonds index.

Several orders other than 10 were used on this interval to check the significance of model order on results obtained for the independent interval. In fact, model order was increased from 1 through 12 by increments of one, with the results that extremely poor results were initially demonstrated, with each succeeding increase in model order leading (almost uniformly) to better results. Now, if model predictive quality were examined only on the interval of estimation, this result would be expected. However, results were examined on the <u>independent</u> interval, where if there were no basis for using a linear model, increased order should have little or <u>no</u> significance. This fact was made abundantly clear to the Authors in their earlier investigation.

Feeling that some measure of significance had been demonstrated by the above results, we decided to continue to test the modeling procedure on subsequent data. The manner in which we chose to do this was to expand both the interval of estimation and the interval of prediction, while continuing to use a 10th order model. The reader is now asked to examine Figures 7 and 8. In this example, the first 60 days were used to estimate model coefficients and the following 50 days used to make predictions, again assuming the utilities to be known. In this case, the predictive quality of the model is noticeably worse. We do note, however, that the model does tend to compromise the distinction between the utilities and bonds, which over this interval seem to be almost completely out-of-phase.

Continuing in a similar manner, we used the first 80 days to estimate the model. The following 70 days were used to examine simulation results, which are expressed by Figures 9 and 10. Examination of Figures 9 and 10 reveals some of the character of our discrete model. The "high frequency cut-off" for the model seems to be somewhat lower than the frequencies present in the utilities average. This can be clearly seen by examining corresponding data for days following 115, when the utilities surge upward rapidly.

Finally, we ask the reader to examine Figures 11 and 12. In this case the first 100 days of our sample were used to estimate the model and the next 90 days to make predictions. In this interval we see a phenomenon which is almost directly analagous to the use of a step input to determine system characteristics. It is from this interval that one can graphically determine the validity of a linear approximation to the dynamics of the bond market. We leave detailed analysis of these figures to the reader.

CONCLUDING REMARKS

Why does the model appear to offer some explanation of the dynamics of the bond market? The Authors are not sure of the answer to this question. We need not conclude that the utilities control the bond market in any sense, although there is a certain temptation to do so. It has been previously noted in several quarters that speculative prices (including indices) adjust rapidly to new information of various types. In a sense then, these indices act as numeric "observers" of information which otherwise would have little quantifiable meaning. It is simply a possibility that the 15 Utilities is such an observer and reflects information which is important to future bond prices. Also, the Authors do not conclude that the utilities are in fact the only significant input to the bond market. Others, such as the Industrials, and Rails, were not employed in simulation experiments.

The Authors are not naive enough to believe that the bond market is describable exactly by a linear model. However, we believe it to be of a nature which can be approximated by a high-order difference equation. In a discussion advocating the random walk hypothesis, Roberts [9] states

> ...there should be great interest in the possibility that, to a first approximation, stock-market behavior may be statistically the simplest, by far, of all economic time series.

We believe that there should be great interest on the part of random walk theorists in the fact that results presented here infer that a simple linear model is able to condition the expected value of speculative price series.

Why haven't the Authors conducted statistical tests of residuals, listed correlation coefficients, determined natural frequencies, and system stability? Experience has led us to believe these tests are significant in a predictive model only after a suitable model has been determined from independent interval analysis, otherwise they appear to be nearly meaningless. While we believe that the results presented here are significant, we do not believe that they are good enough to imply that the bond market has certain natural frequencies or is stable or unstable. Besides, over what interval does one attempt to make such determinations?

What do the results imply for further analysis? The Authors have held the intuitive belief from the outset that the stock market is a system, which has time-varying dynamics. This might imply, for instance, that a model of the form described by equations (7) and (8) is appropriate. Our thesis rests upon the basic understanding that, after all, the stock market is itself a result of the activities of people. It is no secret, for instance, that the normative attitudes and behavior of our society have undergone vast changes in just the last several years. No doubt this fact is reflected in stock market activity. As a simple example, we can cite the introduction of new technology (e.g., TV, digital computer, nuclear power, space travel) which has changed and continues to change our entire lifestyle, much less the structure of the stock market. This time-varying aspect is one which is apparently important, but one which is generally neglected by mathematical economists. The tendency is to use several years of daily data for instance to estimate the coefficients of a simple, constant-coefficient, linear model, the objective being to obtain a "large sample" parameter estimates. This activity, while intuitively appealing, can yield simply erroneous results if the underlying process is time-varying. However, a model with constant coefficients, as given in equations (5) and (6), can be a fairly good approximation over short intervals if the underlying process is "slowly" time-varying.

In view of the above reasoning, the Authors unfortunately cannot guarantee that the model (or the modeling procedure) used here will yield successful results when used with data drawn from another sample (not attempted by the Authors). The work presented here can only be cited as "evidence" in a growing body of knowledge concerning the behavior of speculative prices.

We do believe that the method used here is a powerful tool for time series analysis and that the results presented can form the basis for several interesting future investigations, even with the Authors' original model, as future data on significant market-affecting information becomes available. Unfortunately, even with the simple linear, time-invariant model of the bond market presented here, there are many possibilities reflected by the choices of interval, order and inputs.

Lastly, we are convinced that ultimate conclusions regarding stock market dynamics will have significant import with respect to the understanding of other socio-economic systems, which seem to display "expectations-adjustment" phenomena, but for which data is much harder to obtain. For this reason alone, further research into the understanding of speculative prices seems justified.

ACKNOWLEDGMENTS

The Authors wish to express their gratitude to Dr. J. S. Pazdera and Dr. R. T. Johnson for their assistance in the preparation of this work.

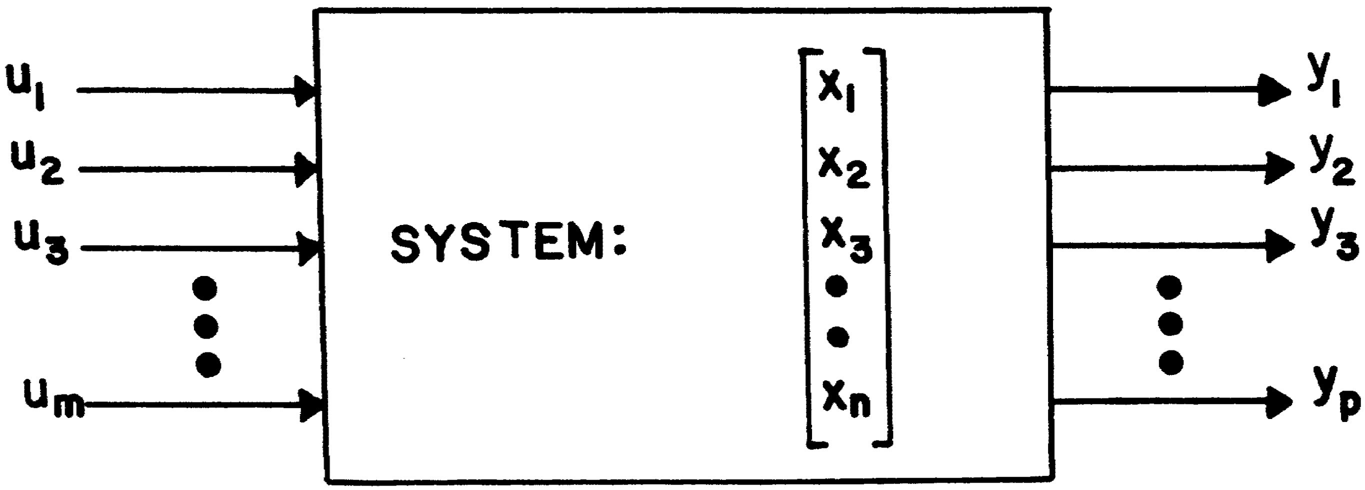

Fig. 1 Dynamic System

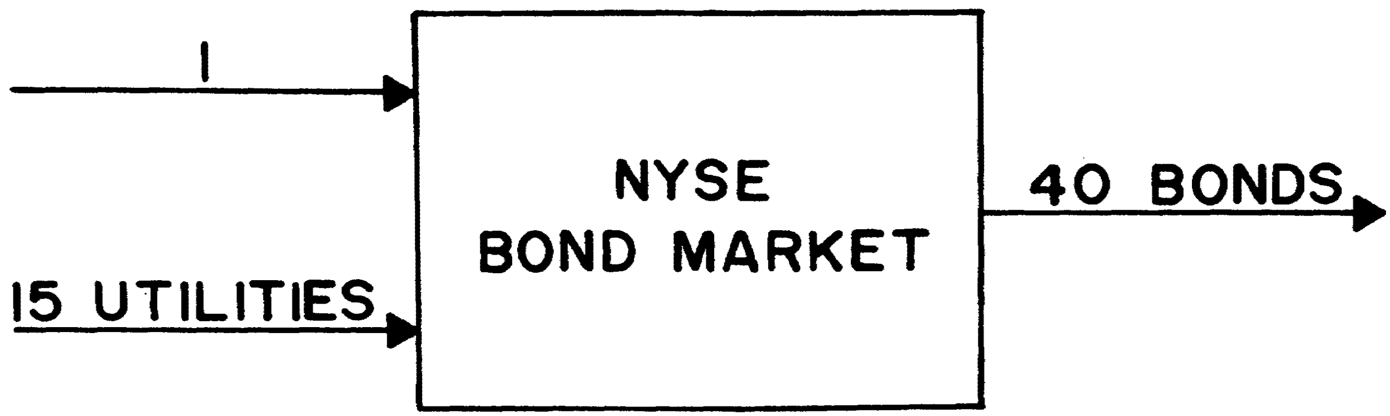

Fig. 2 The NYSE Bond Market as a Dynamic System

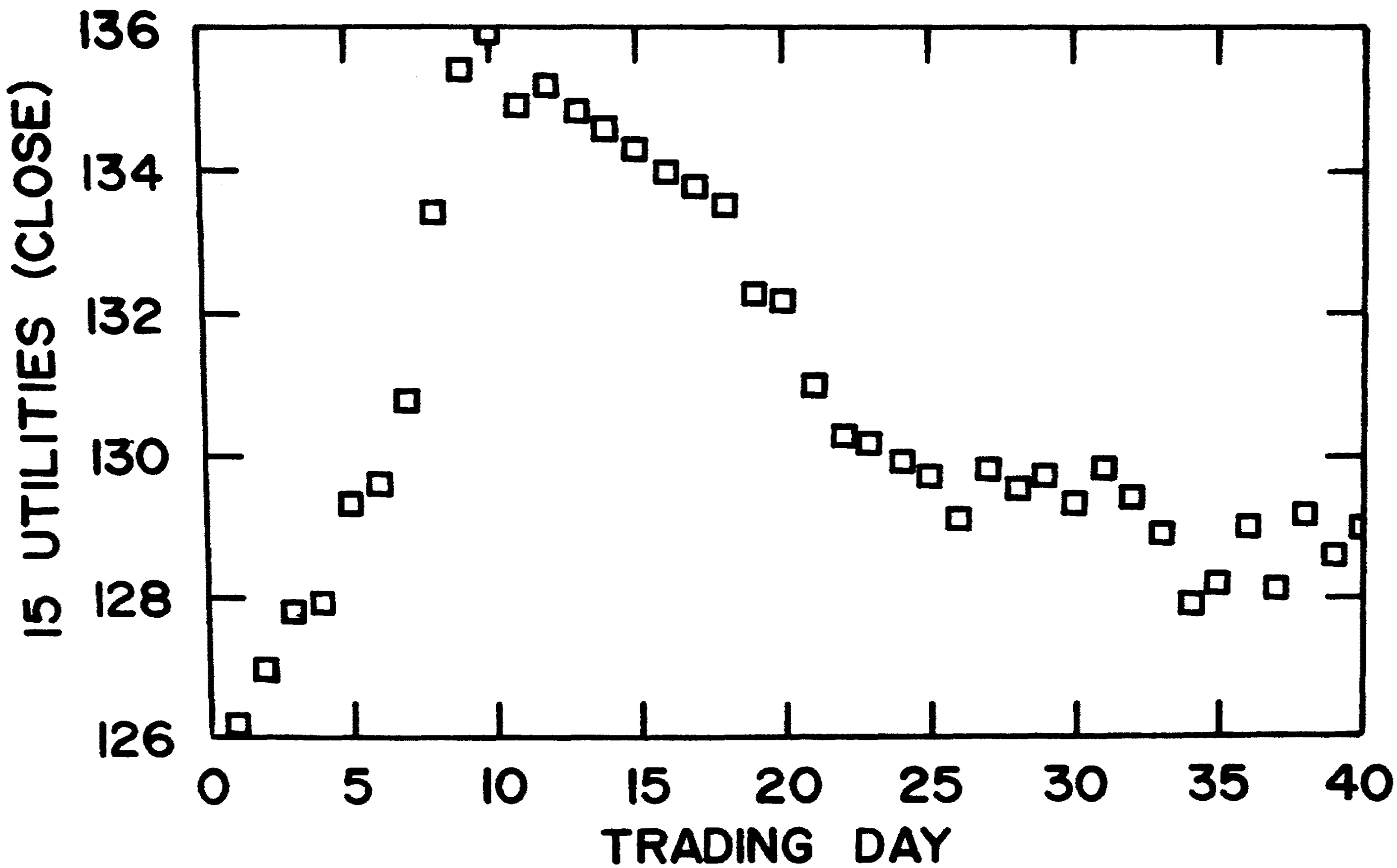

Fig. 3   Observed Utilities

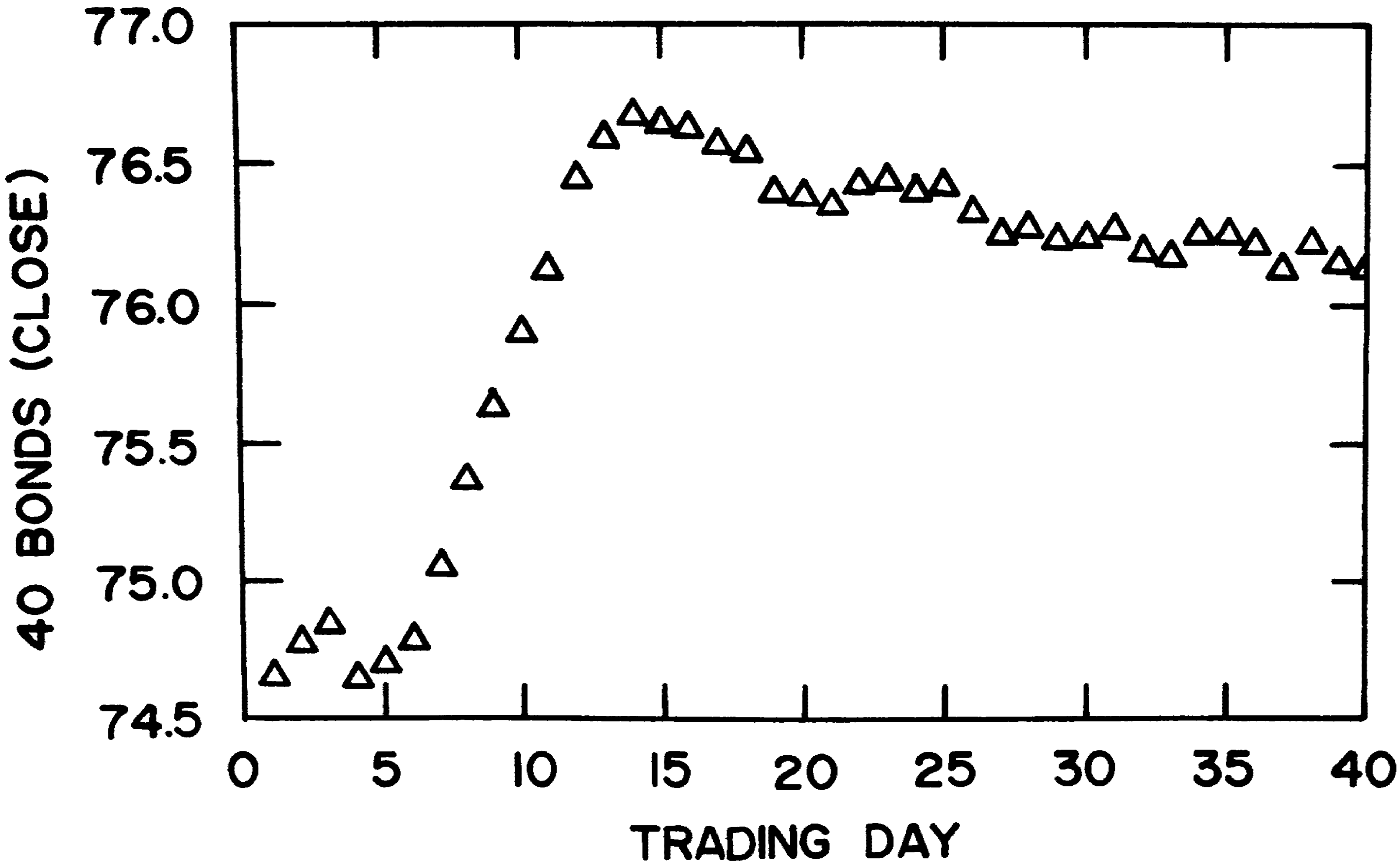

Fig. 4 Observed Bonds Average

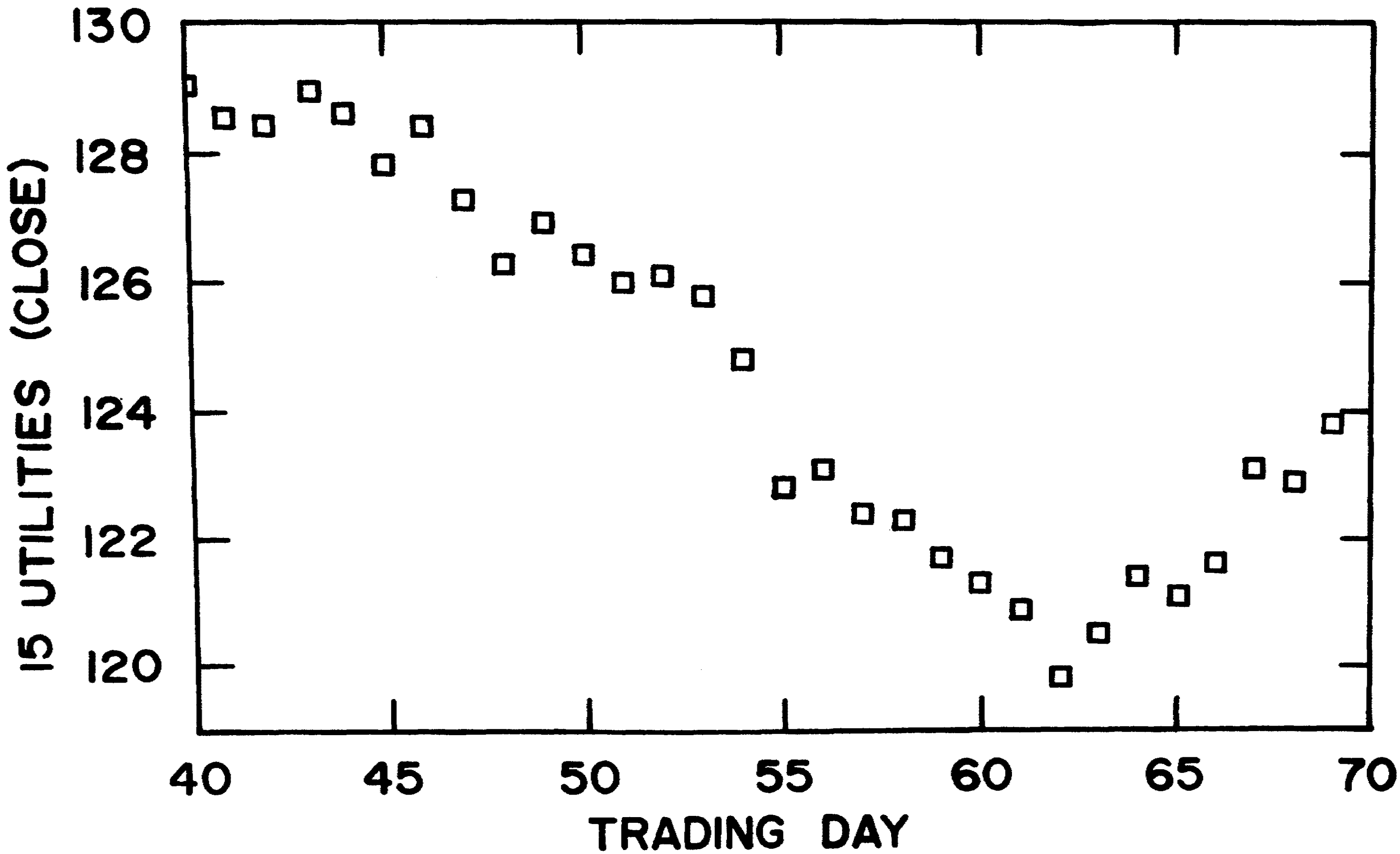

Fig. 5   Model Input   (Days 40-69)

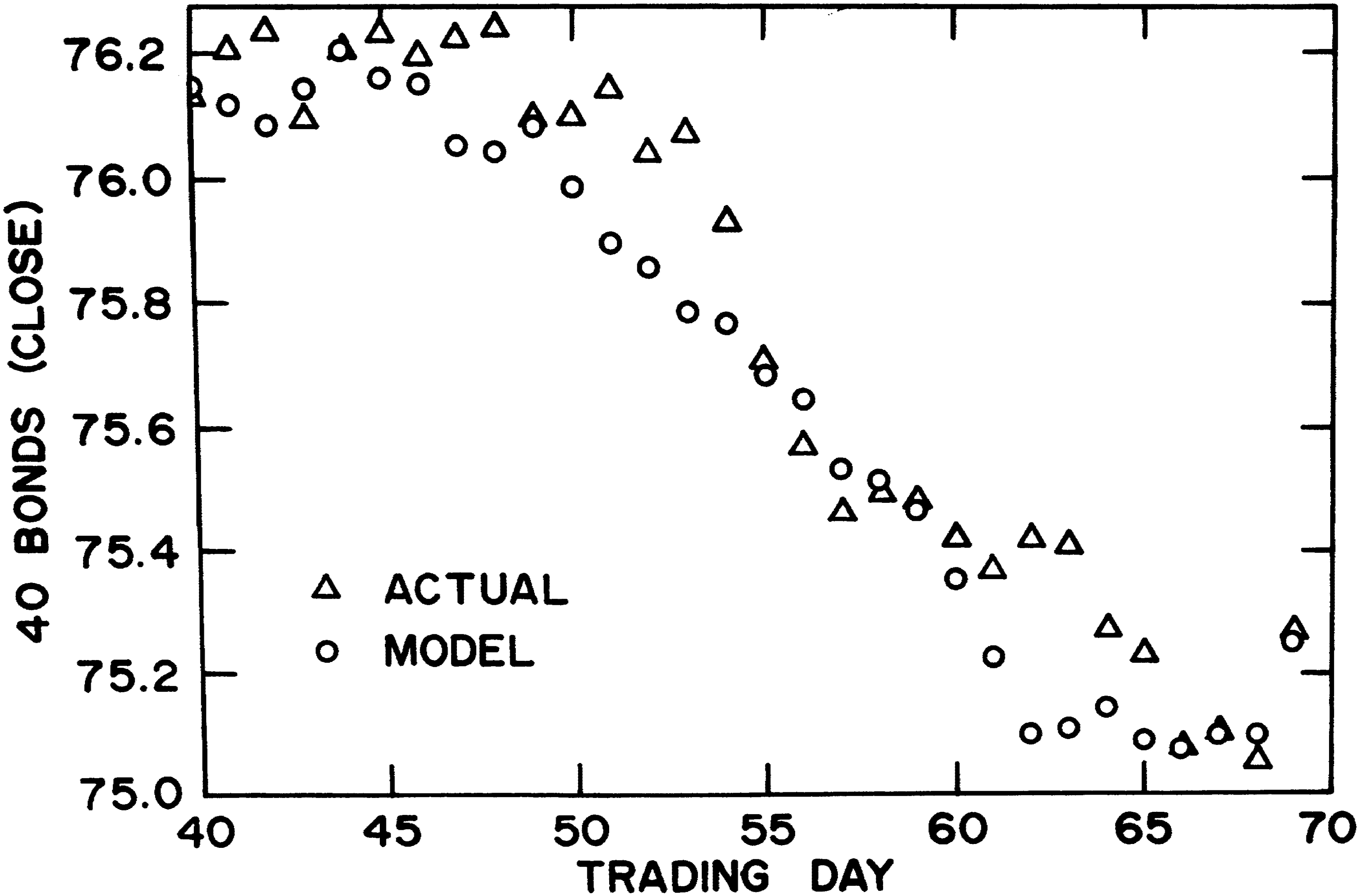

Fig. 6 Model Performance (Days 40-69)

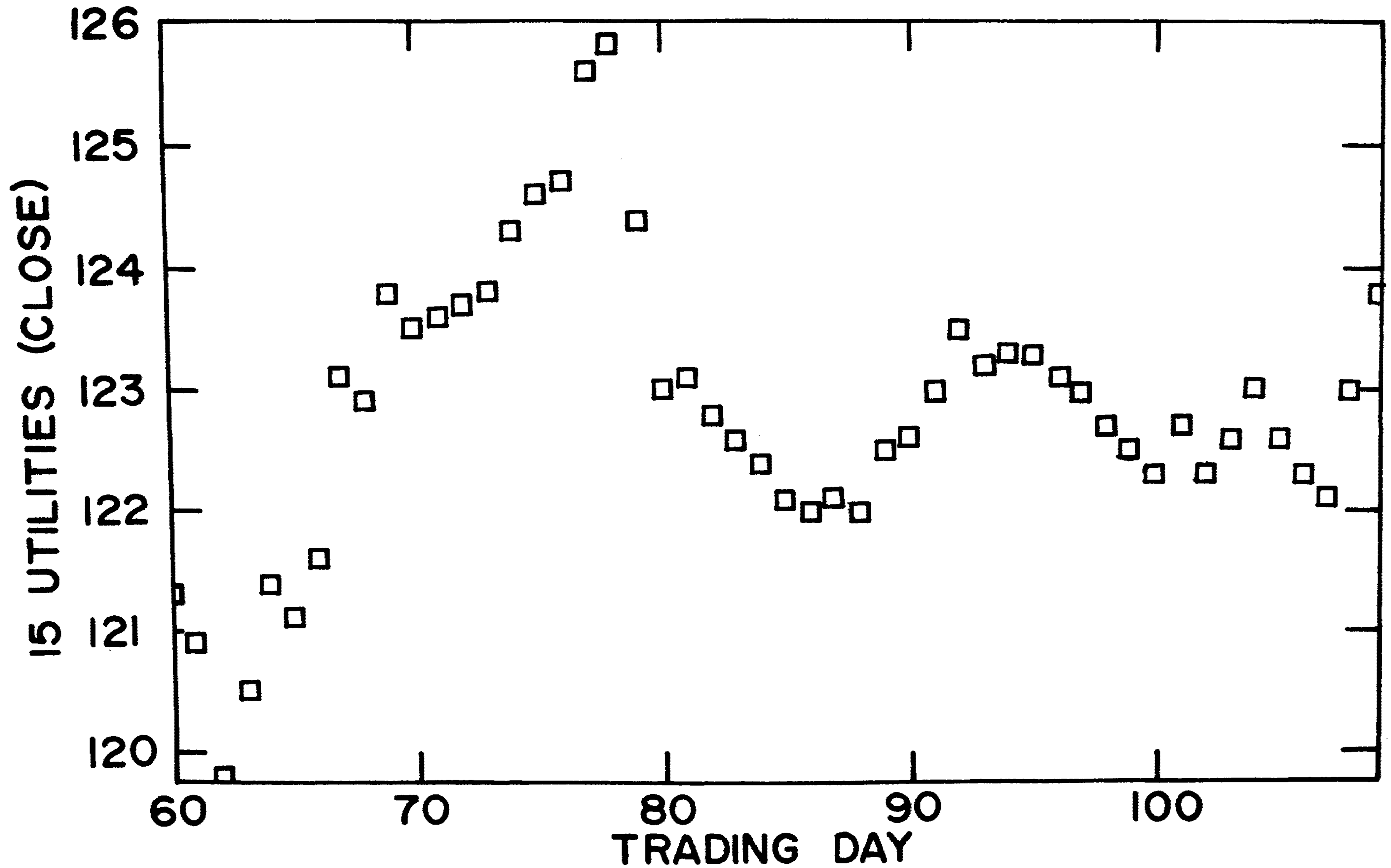

Fig. 7 Model Input (Days 60-109)

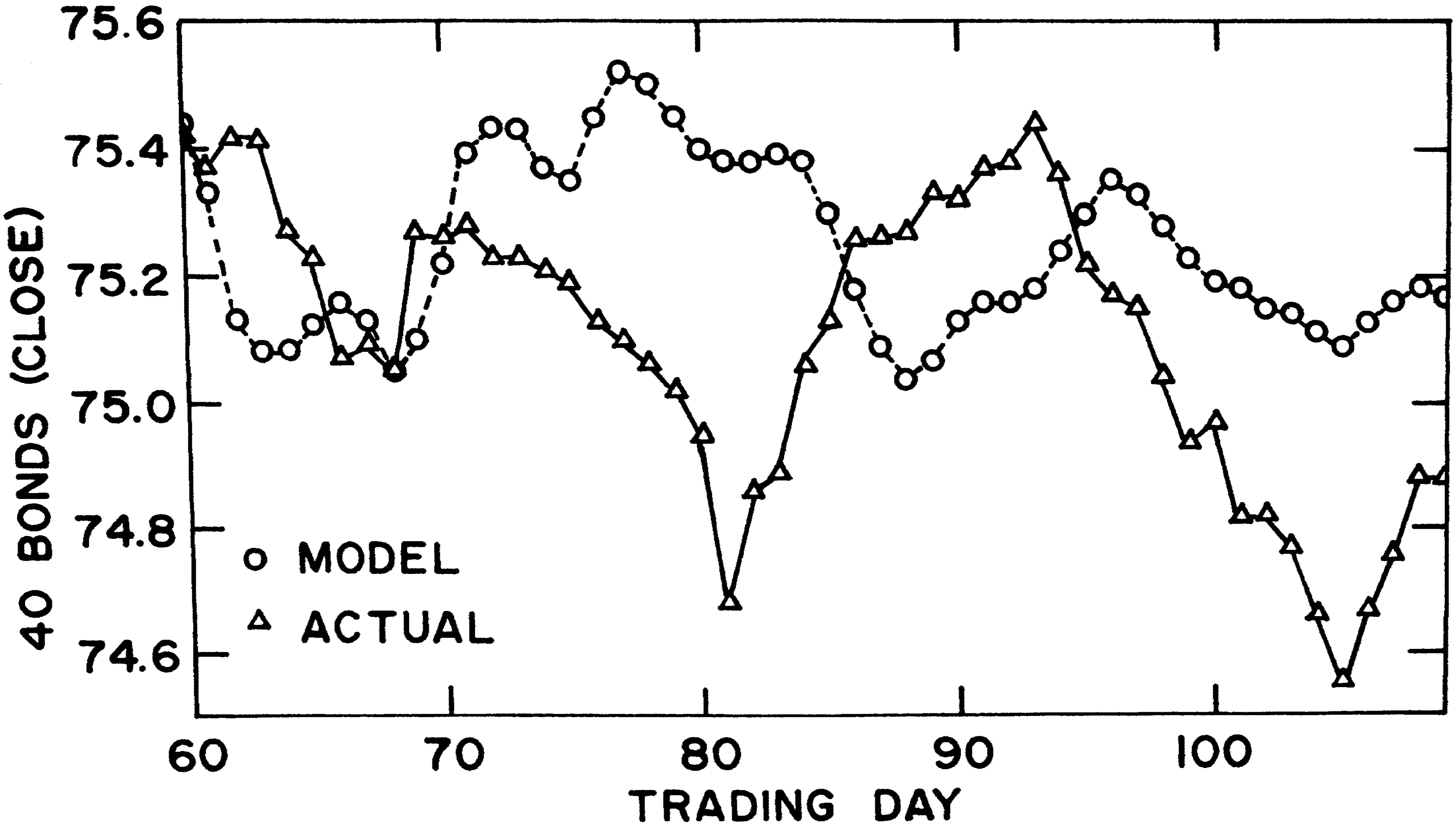

Fig. 8 Model Performance (Days 60-109)

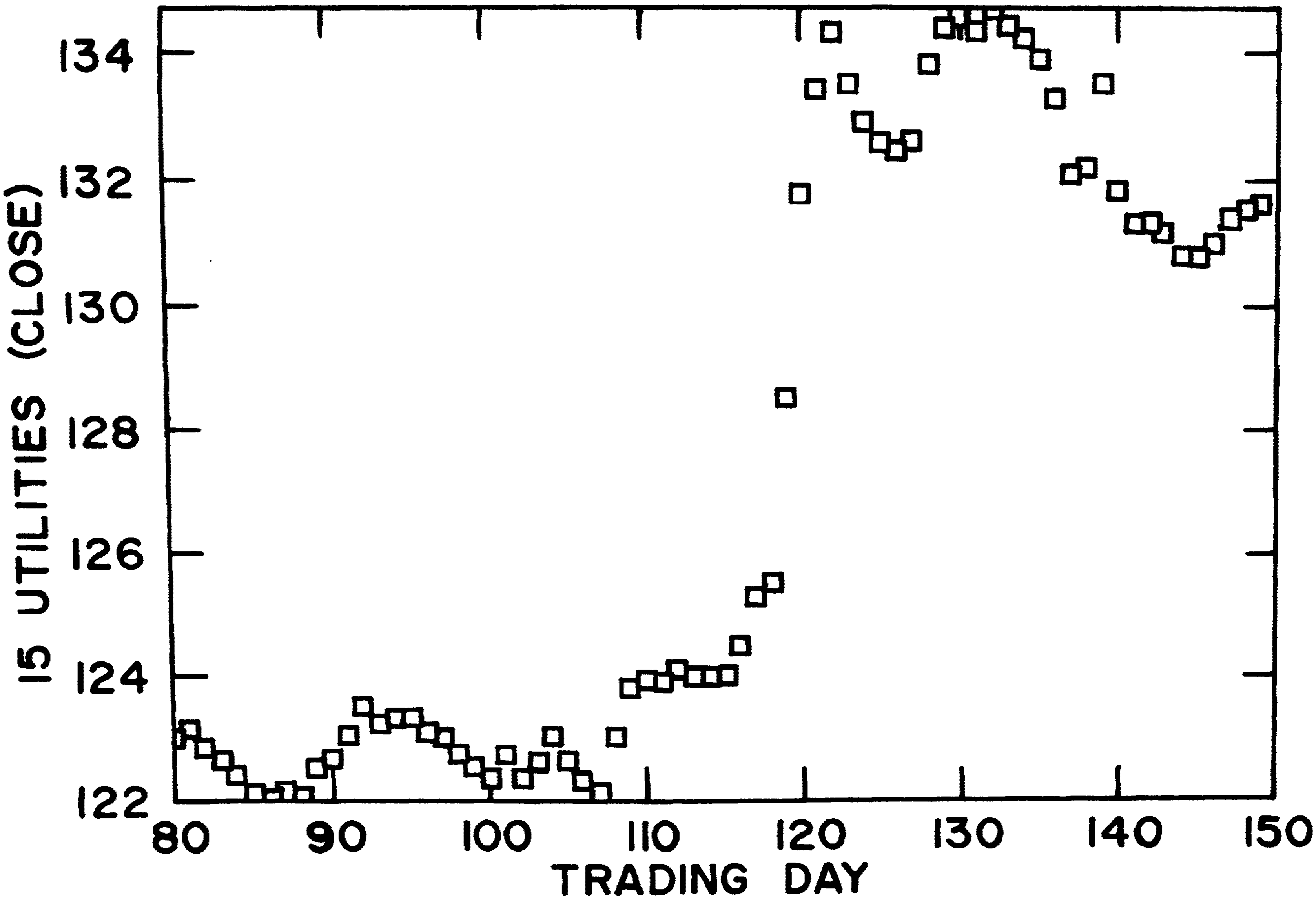

Fig. 9 Model Input (Days 80-149)

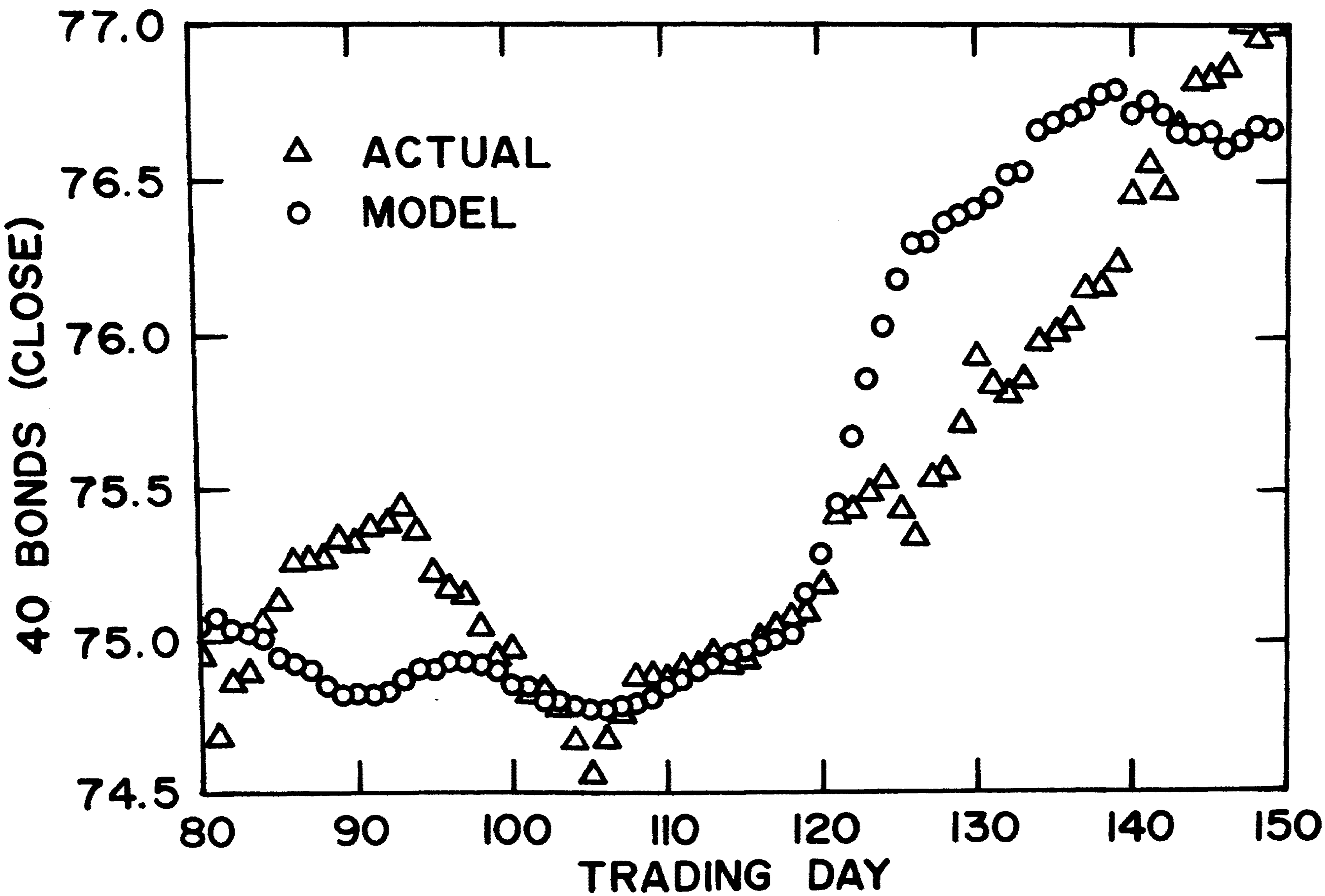

Fig. 10 Model Performance (Days 80-149)

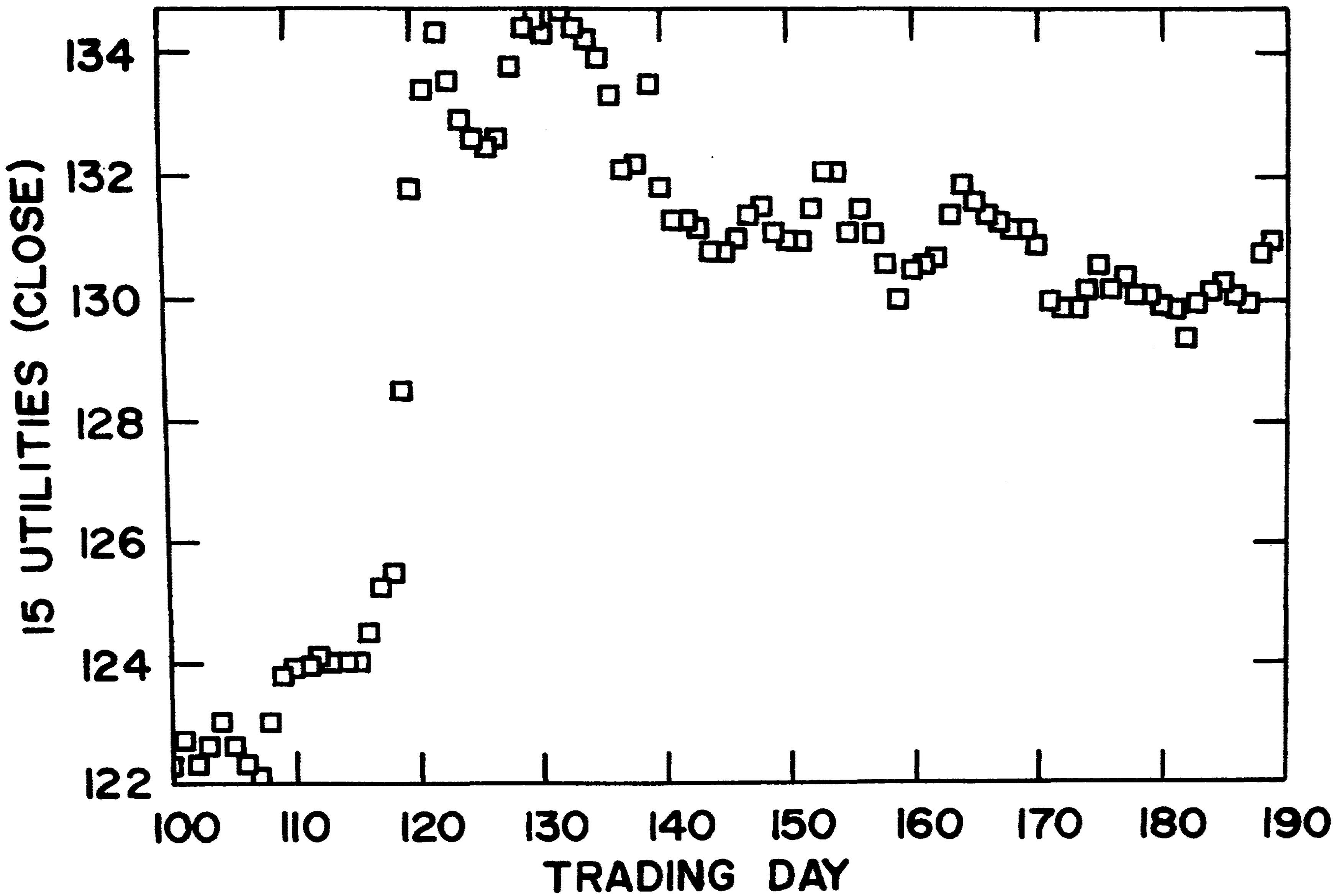

Fig. 11 Model Input (Days 100-189)

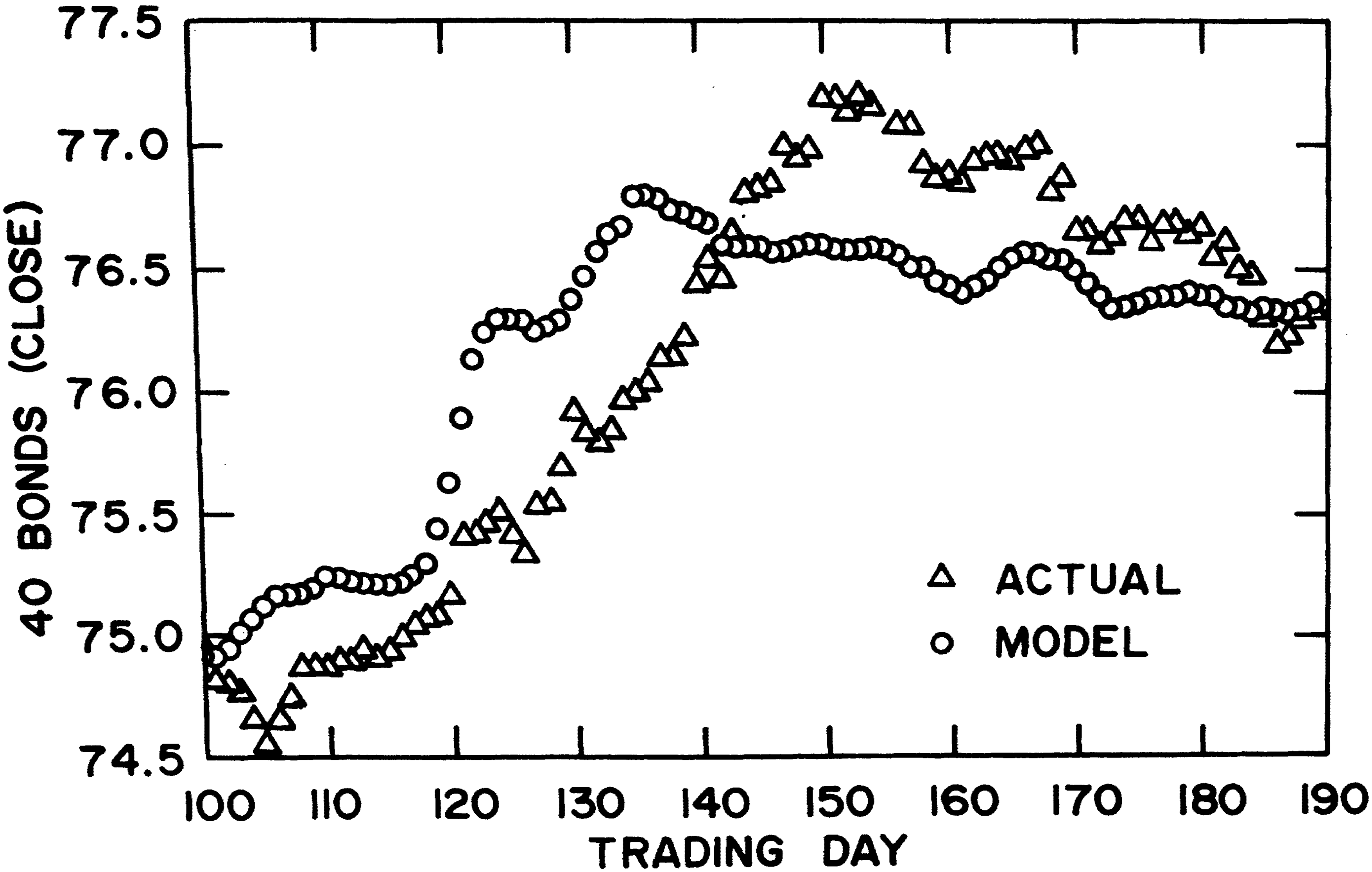

Fig. 12 Model Performance (Days 100-189)

REFERENCES

1. Bachelier, L., "Theorie de la Speculation," Gauthier-Villars, Paris, 1900, In Cootner, P. H. (ed.), The Random Character of Stock Prices, M.I.T. Press, Cambridge, Mass., 1964, pp. 17-78.

2. Working, H., "A Theory of Anticipatory Prices," Amer. Econ. Rev., May 1958, pp. 188-189.

3. Cowles, A., "Stock Market Forecasting," Econometrics, Vol. 7, Nos. 3-4, July-Oct. 1964.

4. Eiteman, W. J., Dice, C. A., and Eiteman, D. K., The Stock Market, 4th ed., McGraw-Hill, N.Y., 1966.

5. Edwards, R. D., and Magee, J., Technical Analysis of Stock Trends, 5th ed., John Magee, Springfield, Mass., 1967.

6. Rosenberg, C. N., Jr. (ed.), Stock Market Primer, World Pub. Co., N.Y., 1970

7. Kendall, M. G., "The Analysis of Economic Time Series -- Part I: Prices," J. Royal Statist. Soc., Vol. 96, Pt. 1, 1953, pp. 11-25, in Cootner, P. H. (ed.), The Random Character of Stock Market Prices, M.I.T. Press, Cambridge, Mass., 1964, pp. 85-99.

8. Osborne, M. F. M., "Brownian Motion in the Stock Market," Operations Research, Vol. 7, Mar.-Apr. 1959, pp. 145-173, in Cootner, P. H. (ed.), The Random Character of Stock Market Prices, M.I.T. Press, Cambridge, Mass., 1964, pp. 100-128.

9. Roberts, H. V., "Stock Market Patterns and Financial Analysis: Methodological Suggestions," J. Finance, Vol. 14, 1959, pp. 1-10, in Cootner, P. H. (ed.), The Random Character of Stock Market Prices, M.I.T. Press, Cambridge, Mass., 1964, pp. 7-16.

10. Larson, A. B., "Measurement of Random Process in Future Prices," Food Research Inst. Studies, Vol. 1, No. 3, Nov. 1960, pp. 313-324, in Cootner, P. H. (ed.), The Random Character of Stock Market Prices, M.I.T. Press, Cambridge, Mass., 1964, pp. 219-230.

11. Samuelson, P. A., "Proof That Properly Anticipated Prices Fluctuate Randomly," Ind. Mgmt. Rev., Vol. 6, 1965, pp. 41-49.

12. Alexander, S. S., "Price Movements in Speculative Markets: Trends or Random Walks," Ind. Mgmt. Rev., Vol. 2, No. 2, May 1961, pp. 7-26, in Cootner, P. H. (ed.), The Random Character of Stock Market Prices, M.I.T. Press, Cambridge, Mass., 1964, pp. 199-218.

13. Alexander, S. S., "Price Movements in Speculative Markets: Trends or Random Walks, No. 2," Ind. Mgmt. Rev., Vol. 5, 1964, pp. 25-46, in Cootner, P. H. (ed.), The Random Character of Stock Market Prices, M.I.T. Press, Cambridge, Mass., 1964, pp. 338-372.

14. Fama, E. F., and Blume, M. E., "Filter Rules and Stock Trading," J. Business, Vol. 39, No. 1, Pt. 2, Jan. 1966, pp. 226-241.

15. Dryden, M. M., "A Source of Bias in Filter Tests of Share Prices," J. Business, Vol. 42, No. 3, July 1969, pp. 321-325.

16. Osborne, M. F. M., "Periodic Structure in the Brownian Motion of Stock Prices," Operations Research, Vol. 10, May-June 1962, pp. 345-379, in Cootner, P. H. (ed.), The Random Character of Stock Market Prices, M.I.T. Press, Cambridge, Mass., 1964, pp. 262-296.

17. Cootner, P. H., "Stock Prices: Random vs. Systematic Changes," Ind. Mgmt. Rev., Vol. 3, No. 2, 1962, in Cootner, P. H. (ed.), The Random Character of Stock Market Prices, M.I.T. Press, Cambridge, Mass., 1964, pp. 231-252.

18. Steiger, W., "A Test of Nonrandomness in Stock Price Changes," in Cootner, P. H. (ed.), The Random Character of Stock Market Prices, M.I.T. Press, Cambridge, Mass., 1964, pp. 253-261.

19. Granger, C. W. J., and Morgenstern, O., "Spectral Analysis of New York Stock Market Prices," Kyklos, Vol. 16, 1963, pp. 1-27, in Cootner, P. H. (ed.), The Random Character of Stock Market Prices, M.I.T. Press, Cambridge, Mass., 1964, pp. 162-188.

20. Granger, C. W. J., "The Typical Spectral Shape of an Economic Variable," Econometrica, Vol. 34, No. 1, Jan. 1966.

21. Mandelbrot, B., "The Variation of Certain Speculative Prices," J. Business, Vol. 36, No. 4, Oct. 1963,

pp. 394-419, in Cootner, P. H. (ed.), <u>The Random Character of Stock Market Prices</u>, M.I.T. Press, Cambridge, Mass., 1964, pp. 307-332.

22. Mandelbrot, B., "Forecasts of Future Prices, Unbiased Markets, and Martingale Models," J. Business, Vol. 39, No. 1, Pt. 2, Jan. 1966, pp. 242-255.

23. Fama, E. F., "Mandelbrot and the Stable Paretian Hypothesis," J. Business, Vol. 36, No. 4, Oct. 1963, pp. 420-429, in Cootner, P. H. (ed.), <u>The Random Character of Stock Market Prices</u>, M.I.T. Press, Cambridge, Mass., 1964, pp. 297-306.

24. Fama, E. F., "The Behavior of Stock Market Prices," J. Business, Vol. 38, No. 1, 1965, pp. 34-105.

25. Press, S. J., "A Compound Events Model for Security Prices," J. Business, Vol. 40, No. 3, July 1967, pp. 317-335.

26. Brada, J., Ernst, H., and Van Tassel, J., "The Distribution of Stock Price Differences: Gaussian After All?" Operations Research, Vol. 14, 1966, pp. 334-340.

27. Niederhoffer, V., and Osborne, M. F. M., "Market Making and Reversal on the New York Stock Exchange," J. Am. Statistical Assn., Vol. 61, No. 316, Dec. 1966, pp. 897-916.

28. Niederhoffer, V., "A New Look at the Clustering of Stock Prices," J. Business, Vol. 39, No. 2, Apr. 1966, pp. 309-313.

29. Granger, C. W. J., "Some Aspects of the Random Walk Model of Stock Market Prices," Int. Econ. Rev., Vol. 9, No. 2, June 1968, pp. 253-257.

30. Simmons, D. M., "Common Stock Transaction Sequences and the Random Walk Model," Operations Research, Vol. 19, No. 4, July-Aug. 1971, pp. 845-861.

31. Osborne, M. F. M., "The Dyanmics of Stock Trading," Econometrica, Vol. 33, No. 1, Jan. 1965, pp. 88-113.

32. Osborne, M. F. M., "Some Quantitative Tests for Stock Price Generating Models and Trading Folklore," J. Am. Statist. Assn., Vol. 62, 1967, pp. 321-340.

33. Ying, C. C., "Stock Market Prices and Volumes of Sales," Econometrica, Vol. 34, No. 3, July 1966, pp. 676-685.

34. Fama, E. F., et. al., "The Adjustment of Stock Prices to New Information," Int. Econ. Rev., Vol. 10, No. 1, Feb. 1969, pp. 1-21.

35. Waud, R. N., "Public Interpretation of Federal Reserve Discount Rate Changes: Evidence on the Announcement Effect," Econometrica, Vol. 38, No. 2, 1970, p. 231.

36. Niederhoffer, V., "The Analysis of World Events and Stock Prices," J. Business, Vol. 44, No. 2, Apr. 1971, pp. 193-219.

37. King, B. F., "Market and Industry Factors in Stock Price Behavior," J. Business, Vol. 39, No. 1, Pt. 2, Jan. 1966, pp. 139-170.

38. Shelton, J. P., "The Value Line Contest: A Test of the Predictability of Stock-Price Changes," J. Business,

Vol. 40, No. 3, July 1967, pp. 251-269.

39. Barnes, R. M., "A Statistical Method for Setting Stops in Stock Trading," Operations Research, Vol. 18, No. 4, July-Aug. 1970, pp. 665-688.

40. Ho, C. Y., and Gaitros, D. L., "Computerized Evaluation of Daily Market Strength," Presented at the 38th National Meeting of the Operations Research Society of America, Detroit, Oct. 28-30, 1970.

41. Bliss, G. A., "Mathematical Interpretations of Geometrical and Physical Phenomena," Amer. Math. Monthly, Vol. 40, 1933, pp. 472-480.

42. Ogata, K., State Space Analysis of Control Systems, Prentice-Hall, Englewood Cliffs, N.J., 1967.

43. Takahashi, Y., Rabins, M. J., and Auslander, D. M., Control, Addison-Wesley, Reading, Mass., 1970.

44. Kalman, R. E., Falb, P. L., and Arbib, M. A., Topics in Mathematical System Theory, McGraw-Hill, N.Y., 1969.

45. Kalman, R. E., "Mathematical Description of Linear Dynamical Systems," S.I.A.M. J. Contr., Ser. A, Vol. 1, No. 2, 1963, pp. 152-192.

46. Cuenod, M., and Sage, A. P., "Comparison of Some Methods Used for Process Identification," Automatica, Vol. 4, No. 4, 1968, p. 235.

47. Eykhoff, P., "Process Parameter and State Estimation," Automatica, Vol. 4, No. 4, 1968, p. 205.

48. Nieman, R. E., Fisher, D. G., and Seborg, D. E., "A Review of Process Identification and Parameter Estimation Techniques," Int. J. Contr., Vol. 13, No. 2, 1971, pp. 209-264.

49. Gopinath, B., "On the Identification of Linear Time-Invariant Systems From Input-Output Data," Bell System Tech. J., Vol. 48, No. 5, May 1969, pp. 1101-1113.

50. Budin, M. A., "Minimal Realization of Discrete Linear Systems From Input-Output Observations," I.E.E.E. Trans. Auto. Contr., AC-16, No. 5, Oct. 1971, pp. 395-401.

51. Behring, A. G., "Methods and Applications of System Identification," Ph.D. thesis, University of Missouri-Rolla, Rolla, Mo., Dec. 1971.

52. BARRON'S National Business and Financial Weekly, Dow Jones and Co., Inc., N.Y.

53. Federal Reserve BULLETIN, published monthly by the Federal Reserve System, Washington, D.C..

54. Facts on File Yearbook 1968, Facts on File, Inc., Vol. XXVIII, N.Y., 1968.

55. The Dow Jones Investor's Handbook 1969, Dow Jones and Co., Inc., N.Y., 1969.

APPENDIX

TABLE 1

STOCK MARKET DATA[4]

Date	Day Number	15 Utilities (close)	40 Bonds (close)
12-26-67	1	126.18	74.66
12-27-67	2	127.04	74.78
12-28-67	3	127.84	74.84
12-29-67	4	127.91	74.64
1-02-68	5	129.31	74.70
1-03-68	6	129.63	74.78
1-04-68	7	130.75	75.05
1-05-68	8	135.37	75.37
1-08-68	9	135.42	75.63
1-09-68	10	135.93	75.91
1-10-68	11	134.87	76.13
1-11-68	12	135.22	76.45
1-12-68	13	134.84	76.59
1-15-68	14	134.65	76.67
1-16-68	15	134.33	76.64
1-17-68	16	134.04	76.63
1-18-68	17	133.82	76.58
1-19-68	18	133.53	76.56
1-22-68	19	132.35	76.40
1-23-68	20	132.19	76.39
1-24-68	21	131.04	76.37
1-25-68	22	130.34	76.44
1-26-68	23	130.24	76.44
1-29-68	24	129.92	76.42
1-30-68	25	129.73	76.42
1-31-68	26	129.06	76.34
2-01-68	27	129.76	76.26
2-02-68	28	129.54	76.28
2-05-68	29	129.70	76.23
2-06-68	30	129.25	76.24

[4]Data extracted from BARRON'S National Business and Financial Weekly [52]. Data are for successive trading days.

Date	Day Number	15 Utilities (close)	40 Bonds (close)
2-07-68	31	129.76	76.27
2-08-68	32	129.41	76.19
2-09-68	33	128.90	76.18
2-13-68	34	127.88	76.26
2-14-68	35	128.23	76.26
2-15-68	36	128.96	76.22
2-16-68	37	128.10	76.13
2-19-68	38	129.22	76.22
2-20-68	39	128.61	76.16
2-21-68	40	128.96	76.13
2-23-68	41	128.48	76.20
2-26-68	42	128.45	76.23
2-27-68	43	128.87	76.09
2-28-68	44	128.58	76.20
2-29-68	45	127.84	76.23
3-01-68	46	128.36	76.19
3-04-68	47	127.33	76.22
3-05-68	48	126.31	76.24
3-06-68	49	126.89	76.09
3-07-68	50	126.44	76.10
3-08-68	51	126.02	76.14
3-11-68	52	125.12	76.04
3-12-68	53	125.77	76.07
3-13-68	54	124.81	75.93
3-14-68	55	122.80	75.71
3-15-68	56	123.11	75.57
3-18-68	57	122.41	75.40
3-19-68	58	122.32	75.49
3-20-68	59	121.68	75.48
3-21-68	60	121.26	75.42
3-22-68	61	120.91	75.37
3-25-68	62	119.79	75.42
3-26-68	63	120.46	75.41
3-27-68	64	121.39	75.27
3-28-68	65	121.13	75.23
3-29-68	66	121.58	75.07
4-01-68	67	123.15	75.09
4-02-68	68	122.92	75.05
4-03-68	69	123.75	75.27
4-04-68	70	123.53	75.26
4-05-68	71	123.56	75.28
4-08-68	72	123.72	75.23
4-10-68	73	123.75	75.23
4-11-68	74	124.27	75.21
4-15-68	75	124.62	75.19
4-16-68	76	124.68	75.13
4-17-68	77	125.58	75.10
4-18-68	78	125.83	75.06
4-19-68	79	124.36	75.02

Date	Day Number	15 Utilities (close)	40 Bonds (close)
4-22-68	80	123.02	74.95
4-23-68	81	123.15	74.68
4-24-68	82	122.80	74.86
4-25-68	83	122.60	74.89
4-26-68	84	122.41	75.06
4-29-68	85	122.09	75.13
4-30-68	86	121.96	75.26
5-01-68	87	122.12	75.26
5-02-68	88	122.03	75.27
5-03-68	89	122.48	75.33
5-06-68	90	123.53	75.37
5-07-68	91	123.53	75.37
5-08-68	92	123.21	75.38
5-09-68	93	123.31	75.44
5-10-68	94	123.27	75.63
5-13-68	95	123.27	75.22
5-14-68	96	123.15	75.17
5-15-68	97	123.05	75.15
5-16-68	98	122.70	75.04
5-17-68	99	122.57	74.94
5-20-68	100	122.32	74.97
5-21-68	101	122.70	74.82
5-22-68	102	122.28	74.82
5-23-68	103	122.57	74.77
5-24-68	104	123.02	74.66
5-27-68	105	122.64	74.55
5-28-68	106	122.28	74.67
5-29-68	107	122.09	74.76
5-31-68	108	123.98	74.88
6-03-68	109	123.79	74.88
6-04-68	110	123.95	74.87
6-05-68	111	123.91	74.91
6-06-68	112	124.14	74.91
6-07-68	113	124.05	74.95
6-10-68	114	123.98	74.92
6-11-68	115	123.98	74.94
6-13-68	116	124.49	75.00
6-14-68	117	125.35	75.05
6-17-68	118	125.54	75.08
6-18-68	119	128.51	75.09
6-20-68	120	131.77	75.18
6-21-68	121	133.44	75.42
6-24-68	122	134.27	75.43
6-25-68	123	133.50	75.48
6-27-68	124	132.89	75.53
6-28-68	125	132.60	75.43
7-01-68	126	132.54	75.34
7-02-68	127	132.60	75.54
7-03-68	128	133.82	75.56
7-08-68	129	134.39	75.71

Date	Day Number	15 Utilities (close)	40 Bonds (close)
7-09-68	130	134.49	75.93
7-11-68	131	134.27	75.84
7-12-68	132	134.71	75.81
7-15-68	133	134.43	75.86
7-16-68	134	134.17	75.98
7-18-68	135	133.95	76.01
7-19-68	136	133.28	76.05
7-22-68	137	132.06	76.15
7-23-68	138	132.19	76.16
7-25-68	139	131.55	76.24
7-26-68	140	131.81	76.46
7-29-68	141	131.29	76.55
7-30-68	142	131.29	76.47
8-01-68	143	131.23	76.67
8-02-68	144	130.85	76.82
8-05-68	145	130.78	76.83
8-06-68	146	131.04	76.86
8-08-68	147	131.45	77.00
8-09-68	148	131.52	76.95
8-12-68	149	131.13	76.99
8-13-68	150	131.04	77.20
8-15-68	151	131.01	77.19
8-16-68	152	131.52	77.13
8-19-68	153	132.09	77.20
8-20-68	154	132.13	77.17
8-22-68	155	131.10	77.13
8-23-68	156	131.55	77.07
8-26-68	157	131.07	77.07
8-27-68	158	130.62	76.93
8-29-68	159	130.02	76.87
8-30-68	160	130.53	76.89
9-03-68	161	130.56	76.86
9-04-68	162	130.66	76.94
9-05-68	163	131.45	76.97
9-06-68	164	131.93	76.97
9-09-68	165	131.65	76.95
9-10-68	166	131.42	76.99
9-12-68	167	131.26	77.01
9-13-68	168	131.23	76.82
9-16-68	169	131.23	76.88
9-17-68	170	130.94	76.66
9-19-68	171	129.98	76.76
9-20-68	172	129.95	76.59
9-23-68	173	129.89	76.64
9-24-68	174	130.21	76.71
9-26-68	175	130.56	76.62
9-27-68	176	130.24	76.71
9-30-68	177	130.37	76.69
10-01-68	178	130.14	76.69

Date	Day Number	15 Utilities (close)	40 Bonds (close)
10-03-68	179	130.08	76.64
10-04-68	180	129.86	76.68
10-07-68	181	129.89	76.56
10-08-68	182	129.38	76.62
10-10-68	183	130.02	76.50
10-11-68	184	130.18	76.48
10-14-68	185	130.30	76.32
10-15-68	186	130.14	76.19
10-17-68	187	130.02	76.23
10-18-68	188	130.85	76.30
10-21-68	189	131.04	76.33
10-22-68	190	130.75	76.21
10-24-68	191	130.46	76.13
10-25-68	192	130.62	76.09
10-28-68	193	131.39	76.01
10-29-68	194	130.82	76.07
10-31-68	195	131.26	76.13
11-01-68	196	131.33	76.17
11-04-68	197	131.71	76.21
11-06-68	198	131.84	76.11
11-07-68	199	132.51	76.23
11-08-68	200	133.56	76.13

INDEX

VITA

Allen Glenn Behring was born on August 17, 1943, in St. Louis, Missouri. He received his primary and secondary education in St. Louis County, Missouri. He first enrolled at the University of Missouri - Rolla in September 1961 and was graduated with a Bachelor of Science in Mechanical Engineering in May 1966. He enrolled in the Graduate School of the University of Missouri - Rolla in September 1966. He was married to the former Miss Elizabeth Ann Wehrenbrecht in June 1967. He received a Master of Science in Mechanical Engineering from the University in August 1968.

He began a program leading to a Doctor of Philosophy in Mechanical Engineering at the University of Missouri - Rolla in September 1968.

He is a member of Phi Kappa Phi, Sigma Xi, Pi Tau Sigma, Beta Sigma Psi, and Blue Key.

Mr. Behring is a citizen of the United States of America.

APPENDIX

COMPUTER SOFTWARE

A collection of subroutines is now presented which can be used to implement the mathematical procedures outlined in the body of this thesis. These procedures (written in PL/1) are sufficient to generate minimal realizations of linear, discrete, multi-variable systems for both off-line and on-line applications given a suitable master program. The reader should note that all computations are carried out in "double precision," having an equivalent PL/1 mode/precision attribute of FLOAT BINARY (53).

In the following discussion, it will be useful to define

$$U = [u(1), u(2), u(3), \ldots, u(K)], \tag{1}$$

and

$$Y = [y(1), y(2), y(3), \ldots, y(K)], \tag{2}$$

where U represents a sequence of K observed $M \times 1$ system input vectors and Y represents a corresponding sequence of K observed $P \times 1$ system output vectors. Further, let

$$W = \left[\begin{array}{ccccc} u(1) & u(2) & u(3) & \cdots & u(r) \\ u(2) & u(3) & u(4) & \cdots & u(r+1) \\ u(3) & u(4) & u(5) & \cdots & u(r+2) \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ u(NN+1) & u(NN+2) & u(NN+3) & \cdots & u(NN+r) \\ \hdashline y(1) & y(2) & y(3) & \cdots & y(r) \\ y(2) & y(3) & y(4) & \cdots & y(r+1) \\ y(3) & y(4) & y(5) & \cdots & y(r+2) \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ y(NN+1) & y(NN+2) & y(NN+3) & \cdots & y(NN+r) \end{array}\right] \quad (3)$$

The relationship between K in equations (1) and (2) and r in equation (3) is

$$r = K\text{-}NN. \quad (4)$$

Lastly, a sequence of r assumed N×1 state vectors X is defined as

$$X = [x(1), x(2), x(3), \ldots, x(r)]. \quad (5)$$

"Time-wise" correspondence exists between the input, output, and state for observations 1 through r. In this discussion, the symbols M, P, and N will always refer to the number of inputs, outputs, and states, respectively.

In order to implement the algorithms which follow, it will be necessary to select a suitable value for the scalar NN as in equation (3), where a particular NN reflects the user's estimate of system order. A good rule of thumb is that NN should be chosen such that the product of NN and P exceeds the estimated system order. For further discussion, see page 57 of this thesis.

Subroutine GRAM

```
GRAM: PROC (NN,U,Y,GM);
      DCL NN FIXED BIN;
      DCL (U(*,*),Y(*,*),GM(*,*)) FLOAT BIN (53);
      DCL (I,J,K,L,M,P,II,JJ,KK,LL,MM,PP,R) FIXED BIN;
      M=HBOUND(U,1);
      P=HBOUND(Y,1);
      R=HBOUND(Y,2)-NN+1;
      IF R<1 THEN GO TO ER;
      GM=0;
      II=0;
      DO I=1 TO NN;
        DO J=1 TO M;
          II=II+1;
          JJ=II-1;
          LL=J;
          DO K=I TO NN;
            DO L=LL TO M;
              JJ=JJ+1;
              DO PP=1 TO R;
                GM(II,JJ)=GM(II,JJ)+U(J,I-1+PP)*U(L,K-1+PP);
              END;
            END;
            LL=1;
          END;
          DO K=1 TO NN;
            DO L=1 TO P;
              JJ=JJ+1;
              DO PP=1 TO R;
                GM(II,JJ)=GM(II,JJ)+U(J,I-1+PP)*Y(L,K-1+PP);
              END;
            END;
          END;
        END;
      END;
      II=NN*M;
      DO I=1 TO NN;
        DO J=1 TO P;
          II=II+1;
          JJ=II-1;
          LL=J;
```

```
          DO K=I TO NN;
            DO L=LL TO P;
              JJ=JJ+1;
              DO PP=1 TO R;
                GM(II,JJ)=GM(II,JJ)+Y(J,I-1+PP)*Y(L,K-1+PP);
              END;
            END;
            LL=1;
          END;
        END;
      END;
      DO I=1 TO NN*(M+P);
        DO J=1 TO NN*(M+P);
          GM(J,I)=GM(I,J);
        END;
      END;
      RETURN;
   ER: PUT LIST ('INSUFFICIENT OBSERVATIONS TO COMPUTE GRAM');
 END GRAM;
```

GRAM is an efficient procedure which may be used to compute the Gram matrix WW', where W is defined by equation (3). In order to accomplish this, a sample mainline program might read:

```
MAIN: PROC OPTIONS (MAIN);
     DCL (NN,M,P,K,MM) FIXED BIN;
     DCL ((U,Y,GM)(1,1)) CONTROLLED BIN (53);
     DCL GRAM ENTRY (FIXED BIN,(*,*) FLOAT BIN (53),
                          (*,*) FLOAT BIN (53),
                          (*,*) FLOAT BIN (53));
     GET LIST (M,P,K,NN);
     MM=(NN+1)*(M+P);
     ALLOCATE U(M,K),Y(P,K),GM(MM,MM);
     GET LIST (U,Y);
     CALL GRAM (NN+1,U,Y,GM);
END MAIN;
```

In this example, U and Y correspond to the U and Y in equations (1) and (2), respectively. The scalar, NN, is defined as in equation (3). The arguments M,P and K correspond to the earlier definitions. Here, the call to GRAM results in computation of GM, where GM is the Gram matrix WW' associated with (3).

It is important to note that storage for the matrix W itself is never required by GRAM. Also, the logic of GRAM has been designed to take computational advantage of the fact that WW' is a symmetric matrix.

Subroutine SCHMIDT

```
SCHMIDT: PROC (GM,F,R,PS,NS);
     DCL (GM(*,*),F(*,*),R(*,*),NS(*,*)) CONTROLLED BIN (53);
     DCL PS(*,*) BIT(1);
     DCL ((FF,RR)(HBOUND(GM,1),HBOUND(GM,1))) BIN (53);
     DCL (A,B,C,D,E) BIN (53);
     DCL (I,J,K,L,M,N) FIXED BIN;
     PS='0'B;
     FF,RR=0;
     I=0;
     DO J=1 TO HBOUND(GM,1);
       IF GM(J,J) < 1.0E-14 THEN GO TO CYC;
       DO K=1 TO I;
         DO L=1 TO J;
           RR(J,K)=RR(J,K)+GM(J,L)*FF(K,L);
         END;
       END;
       E=GM(J,J);
       DO K=1 TO I;
         E=E-RR(J,K)**2;
       END;
       B=E/GM(J,J);
       IF B > NS(J,1) THEN
         DO;
           C=SQRT (E);
           RR(J,I+1)=C;
           DO K=1 TO I;
             FF(I+1,*)=FF(I+1,*)-RR(J,K)*FF(K,*);
           END;
           FF(I+1,J)=1;
           FF(I+1,*)=FF(I+1,*)/C;
           I=I+1;
           PS(J,1)='1'B;
         END;
       NS(J,1)=B;
CYC: END;
     IF I < 1 THEN PUT LIST
                         ('ERROR IN SCHMIDT: GM IS TRIVIAL');
     ALLOCATE F(I,HBOUND(GM,1)),R(HBOUND(GM,1),I);
     DO K=1 TO I;
       DO L=1 TO HBOUND(GM,1);
         F(K,L)=FF(K,L);
```

```
        R(L,K)=RR(L,K);
      END;
    END;
END SCHMIDT;
```

Procedure Schmidt is a digital realization of the modified Schmidt filter discussed beginning on page 75 of this thesis. A sample mainline program to implement SCHMIDT might read in part:

```
MAIN: PROC OPTIONS (MAIN);
     .
     .
     .
     CALL GRAM (NN+1,U,Y,GM);
     NS(*,1)=1.0E-10;
     CALL SCHMIDT (GM,F,R,PS,NS);
END MAIN;
```

Note that values for the Gram matrix GM and NS are specified prior to the call to SCHMIDT. Let the row dimension of GM be MM = (M+P)(NN+1) as before. Then, NS must be dimensioned MM×1 in the calling program MAIN. The elements of NS correspond row-wise to the ε_j, j=1,MM in equation 50, page 76, i.e., NS(1,1) = ε_1, NS(2,1) = ε_2, etc. Following the test indicated in equation (50), page 76, the scalar

$$\underset{\sim}{E}_j \underset{\sim}{E}_j'/g_{jj}$$

replaces the specified ε_j in NS. The vector PS must be declared MM×1 with the attribute BIT(1) in the calling program MAIN. SCHMIDT will return either a one-bit or zero bit in each row of PS corresponding to those rows of W which are "passed" or "blocked," respectively, by the Schmidt filter. Matrices F and R are generated by SCHMIDT. The matrix F is the resultant transfer function of the Schmidt filter (S), while R is the transfer function of a restoring filter (S^+). Both F and R must be given the attribute CONTROLLED in the calling program MAIN. However, these matrices are actually allocated with appropriate dimensions and assigned values by SCHMIDT.

Subroutine MULT1

```
MULT1: PROC (X,Y,XY);
     DCL (X(*,*),Y(*,*),XY(*,*) FLOAT BIN (53);
     DCL (I,J,K) FIXED BIN;
     XY=0;
     DO I=1 TO HBOUND (X,1);
       DO J=1 TO HBOUND (Y,2);
         DO K=1 TO HBOUND (X,2);
           XY(I,J)=XY(I,J)+X(I,K)*Y(K,J);
         END;
       END;
     END;
     RETURN;
END MULT1;
```

Subroutine MULT1 is simply a procedure which is designed to pre-multiply a matrix Y by a matrix X and return the product matrix XY to the calling program. Note that matrices X, Y, and XY must be allocated by the calling program.

Subroutine SYSTEM

```
SYSTEM: PROC (M,P,NN,PJ,PS,A,B,C,D,T);
     DCL (M,P,NN) FIXED BIN;
     DCL PJ(*,*) FLOAT BIN (53);
     DCL PS(*,*) BIT(1);
     DCL (A(*,*),B(*,*),C(*,*),D(*,*),T(*,*))
                         CONTROLLED BIN (53);
     DCL MULT1 ENTRY ((*,*)FLOAT BIN (53),(*,*) FLOAT
                         BIN (53),(*,*) FLOAT BIN (53));
     DCL MATOUT ENTRY (CHAR(120)VAR,(*,*) FLOAT BIN (53),
                         CHAR(1),CHAR(1),FIXED BIN, FIXED BIN);
     DCL (I,J,K,L,N,II,JJ,KK,LL,MM,R) FIXED BIN;
```

```
DCL Q(1,1) CONTROLLED BIN (53);
MM=(NN+1)*(M+P);
N=0;
DO I=(NN+1)*M+1 TO MM-P;
  IF PS(I,1) THEN N=N+1;
END;
IF N<1 THEN PUT LIST ('ERROR IN SYSTEM: ORDER = 0');
ALLOCATE A(N,N),Q(N,(NN+1)*M);
K=0;
DO I=(NN+1)*M+1 TO MM-P;
  IF PS(I,1) THEN
    DO;
      K=K+1;
      DO J=1 TO (NN+1)*M;
        Q(K,J)=PJ(I+P,J);
      END;
      L=0;
      DO J=(NN+1)*M+1 TO MM-P;
        IF PS(J,1) THEN
          DO;
            L=L+1;
            A(K,L)=PJ(I+P,J);
          END;
      END;
    END;
END;
ALLOCATE B(N,M),C(N,M);
DO I=1 TO NN;
  KK=(NN+1)*M-I*M;
  DO J=1 TO N;
    DO L=1 TO M;
      B(J,L)=Q(J,KK+L);
    END;
  END;
  CALL MULT1 (A,B,C);
  DO J=1 TO N;
    DO L=1 TO M;
      Q(J,KK-M+L)=Q(J,KK-M+L)+C(J,L);
    END;
  END;
END;
FREE C;
ALLOCATE T(N,NN*(P+M));
T=0;
KK=0;
DO I=1 TO NN*P;
  IF PS((NN+1)*M+I,1) THEN
    DO;
      KK=KK+1;
      T(KK,I)=1;
    END;
END;
```

```
        DO I=1 TO N;
          DO J=1 TO NN*M;
            T(I,NN*P+J)=-Q(I,M+J);
          END;
        END;
        DO I=1 TO N;
          DO J=1 TO M;
            B(I,J)=Q(I,J);
          END;
        END;
        ALLOCATE C(P,N),D(P,M);
        C=0;
        DO I=1 TO P;
          C(I,I)=1;
          DO J=1 TO M;
            D(I,J)=Q(I,M+J);
          END;
        END;
        FREE Q;
        RETURN;
END SYSTEM;
```

Subroutine SYSTEM is a procedure designed to compute the coefficient matrices A, B, C, and D of equations (1) and (2), page 33, and a suitable transformation matrix T which allows determination of the system state given an appropriate set of input/output observations. The parameters M, P, NN, PJ, and PS are always specified in the calling program. Parameter M is the number of system inputs, P is the number of system outputs, and NN is given by equation (3). Again, let MM = (NN+1)(M+P). As before, PS is a MM×1 BIT(1) vector whose one-bits row-wise indicate the corresponding rows of W of equation (3) which are passed by a suitable Schmidt filter. The matrix PJ is defined as the product of matrices F and R (corresponding to $S^{+}S$ of the Schmidt filter). Matrices A, B, C, D, and T must be declared CONTROLLED in the calling program, but are actually allocated with appropriate row and column dimensions, and given values by SYSTEM. The matrix T can be used to determine system states as follows:

Let

$$W^* = \left[\begin{array}{ccccc} u(1) & u(2) & u(3) & \ldots & u(r) \\ u(2) & u(3) & u(4) & \ldots & u(r+1) \\ u(3) & u(4) & u(5) & \ldots & u(r+2) \\ \cdot & \cdot & \cdot & \ldots & \cdot \\ \cdot & \cdot & \cdot & \ldots & \cdot \\ \cdot & \cdot & \cdot & \ldots & \cdot \\ u(NN) & u(NN+1) & u(NN+2) & \ldots & u(NN+r-1) \\ \hdashline y(1) & y(2) & y(3) & \ldots & y(r) \\ y(2) & y(3) & y(4) & \ldots & y(r+1) \\ y(3) & y(4) & y(5) & \ldots & y(r+2) \\ \cdot & \cdot & \cdot & \ldots & \cdot \\ \cdot & \cdot & \cdot & \ldots & \cdot \\ \cdot & \cdot & \cdot & \ldots & \cdot \\ y(NN) & y(NN+1) & y(NN+2) & \ldots & y(NN+r-1) \end{array}\right] \tag{6}$$

Then the matrix X of equation (5) is given by

$$X = TW^*. \tag{7}$$

Subroutine SYSTEM might be implemented as shown in the following mainline program:

```
MAIN: PROC OPTIONS (MAIN);
     .
     .
     .
     CALL GRAM (NN+1,U,Y,GM);
     CALL SCHMIDT (GM,F,R,PS,NS);
     CALL MULT1 (R,F,PJ);
     CALL SYSTEM (M,P,NN,PJ,PS,A,B,C,D,T);
     .
     .
     .
END MAIN;
```

The subroutines described thus far are independent and, as such, can be used as modules in a number of application programs, including the on-line and off-line identification problems. These modules, of course, must be linked by suitable driving routines. The following two subroutines are examples of such driving routines.

Subroutine ID#1

```
ID#1: PROC (U,Y,X,A,B,C,D,T,NN,NS);
     DCL (U(*,*),Y(*,*),NS) FLOAT BIN (53);
     DCL (X(*,*),A(*,*),B(*,*),C(*,*),D(*,*),T(*,*))
                          CONTROLLED BIN (53);
     DCL NN FIXED BIN;
     DCL (I,J,K,L,M,N,II,JJ,KK,LL,MM,P,R) FIXED BIN;
     DCL PS(1,1) CONTROLLED BIT(1);
     DCL ((NSS,TX,TY,GM)(1,1)) CONTROLLED FLOAT BIN (53);
     DCL SCHMIDT ENTRY ((*,*) FLOAT BIN (53),(*,*) FLOAT
                          BIN (53),(*,*) FLOAT BIN (53),
                          (*,*)BIT(*),(*,*) FLOAT BIN (53));
     DCL GRAM ENTRY (FIXED BIN,(*,*) FLOAT BIN (53),
                          (*,*) FLOAT BIN (53),(*,*) FLOAT
                          BIN (53));
     DCL MATOUT ENTRY (CHAR(120)VAR,(*,*) FLOAT BIN (53),
                          CHAR(1),CHAR(1), FIXED BIN, FIXED
                          BIN);
     DCL SYSTEM ENTRY (FIXED BIN, FIXED BIN, FIXED BIN,
                          (*,*) FLOAT BIN (53),(*,*)BIT(*),
                          (*,*) FLOAT BIN (53), (*,*) FLOAT
                          BIN (53),(*,*) FLOAT BIN (53),
                          (*,*) FLOAT BIN (53),(*,*) FLOAT
                          BIN (53));
     DCL MULT1 ENTRY ((*,*) FLOAT BIN (53),(*,*) FLOAT
                          BIN (53),(*,*) FLOAT BIN (53));
     M=HBOUND (U,1);
     P=HBOUND (Y,1);
     R=HBOUND (Y,2)-NN;
     MM=(NN+1)*(M+P);
     ALLOCATE GM(MM,MM);
     CALL GRAM (NN+1,U,Y,GM);
     ALLOCATE NSS(MM,1),PS(MM,1);
     NSS=NS;
     DO I=MM-P+1 TO MM;
       NSS(I,1)=10;
     END;
     CALL SCHMIDT (GM,TX,TY,PS,NSS);
```

```
      CALL MULT1 (TY,TX,GM);
      FREE TY,TX;
      CALL SYSTEM (M,P,NN,GM,PS,A,B,C,D,T);
      FREE GM,PS,NSS;
      N=HBOUND (A,1);
      ALLOCATE X(N,R);
      X=0;
      DO I=1 TO N;
        DO J=1 TO R;
          KK=0;
          DO K=1 TO NN;
            DO L=1 TO P;
              KK=KK+1;
              X(I,J)=X(I,J)+T(I,KK)*Y(L,K-1+J);
            END;
          END;
          DO K=1 TO NN;
            DO L=1 TO M;
              KK=KK+1;
              X(I,J)=X(I,J)+T(I,KK)*U(L,K-1+J);
            END;
          END;
        END;
      END;
      RETURN;
END ID#1;
```

Procedure ID#1 is a subroutine which is suitable for the off-line identification of linear discrete dynamic systems. The parameters U, Y, NN, and NS are specified in the calling program. Here, U, Y, and NN are defined as before. The scalar NS, however, does not directly correspond to the definition given previously. Effectively, Procedure ID#1 sets all of the ε_j of equation (50), page 76, equal to the scalar NS specified as the last argument of ID#1 in the calling program. In otherwords, the results are that all of the ε_j are set uniformly equal. This is not a necessary feature, but only a simplifying one. Alterations to suit individual need can easily be made. Matrices X, A, B, C, D, and T must be declared CONTROLLED in the main program, but are actually allocated with appropriate dimensions and given values by ID#1 or its subroutines. It is important to note that calls to subroutines GRAM, SCHMIDT, and SYSTEM are issued by ID#1.

No special provision is made for the case of autonomous systems (systems with no inputs). This case can, however, be easily handled with little inconvenience by dimensioning the matrix U as 1×K, where K is defined by equation (1), and setting all of its elements equal to zero. The result will be that the elements of B and D returned by ID#1 will be uniformly zero.

Subroutine ID#2

```
ID#2: PROC (GM,U,Y,X,A,B,C,D,T,NS);
     DCL (GM(*,*),U(*,*),Y(*,*),NS) BIN (53);
     DCL (X(*,*),A(*,*),B(*,*),C(*,*),D(*,*),T(*,*))
                      CONTROLLED BIN (53);
     DCL (I,J,K,L,M,N,II,JJ,KK,LL,MM,P,R,NN) FIXED BIN;
     DCL PS(1,1) CONTROLLED BIT(1);
     DCL ((NSS,TX,TY,GM2)(1,1)) CONTROLLED BIN (53);
     DCL SCHMIDT ENTRY ((*,*) FLOAT BIN (53), (*,*) FLOAT
                      BIN (53), (*,*) FLOAT BIN (53),
                      (*,*) BIT(*),(*,*) FLOAT BIN (53));
     DCL GRAM ENTRY (FIXED BIN,(*,*) FLOAT BIN (53),
                      (*,*) FLOAT BIN (53), (*,*) FLOAT
                      BIN (53));
     DCL MATOUT ENTRY (CHAR(120)VAR,(*,*) FLOAT BIN (53),
                      CHAR(1), CHAR(1), FIXED BIN, FIXED
                      BIN);
     DCL SYSTEM ENTRY (FIXED BIN, FIXED BIN, FIXED BIN,
                      (*,*) FLOAT BIN (53), (*,*) BIT(*),
                      (*,*) FLOAT BIN (53), (*,*) FLOAT
                      BIN (53), (*,*) FLOAT BIN (53),
                      (*,*) FLOAT BIN (53), (*,*) FLOAT
                      BIN (53));
     DCL MULT1 ENTRY ((*,*) FLOAT BIN (53), (*,*) FLOAT
                      BIN (53), (*,*) FLOAT BIN (53));
     M=HBOUND (U,1);
     P=HBOUND (Y,1);
     MM=HBOUND (GM,1);
```

```
        NN=MM/(M+P)-1;
        R=HBOUND (Y,2)-NN;
        ALLOCATE GM2(MM,MM);
        CALL GRAM (NN+1,U,Y,GM2);
        ALLOCATE NSS(MM,1),PS(MM,1);
        NSS=NS;
        DO I=MM-P+1 TO MM;
          NSS(I,1)=10;
        END;
        GM=GM+GM2;
        CALL SCHMIDT (GM,TX,TY,PS,NSS);
        CALL MULT1 (TY,TX,GM2);
        FREE TY,TX;
        CALL SYSTEM (M,P,NN,GM2,PS,A,B,C,D,T);
        FREE PS, NSS,GM2;
        N=HBOUND(A,1);
        ALLOCATE X(N,R);
        X=0;
        DO I=1 TO N;
          DO J=1 TO R;
            KK=0;
            DO K=1 TO NN;
              DO L=1 TO P;
                KK=KK+1;
                X(I,J)=X(I,J)+T(I,KK)*Y(L,K-1+J);
              END;
            END;
            DO K=1 TO NN;
              DO L=1 TO M;
                KK=KK+1;
                X(I,J)=X(I,J)+T(I,KK)*U(L,K-1+J);
              END;
            END;
          END;
        END;
        RETURN;
END ID#2;
```

Procedure ID#2 differs from ID#1 essentially in that it is suited for on-line, or adaptive, identification problems. The same mathematical principles are used by both ID#1 and ID#2. The parameters GM, U, Y, and NS must be specified in the calling program. The scalar NS corresponds to its definition in ID#1. The matrix GM is the Gram matrix WW', where W is given by equation (3). The matrices U and Y are defined by equations (1) and (2), respectively.

Matrices X, A, B, C, D, and T must be declared CONTROLLED in the program which calls ID#2, however, these matrices are allocated and assigned values by ID#2. In order to use ID#2 for adaptive or on-line identification, the algorithm is called recursively as new input/output observations are added to the existing set. When ID#2 is called, it is convenient to consider that the matrix GM is always associated with the complete set of "old" observations, whereas U and Y represent a "new" set of K corresponding observations on the system input and output. Note that K must be sufficiently large to compute a Gram matrix consistent with the choice of NN, i.e., $K \geq NN+1$. When ID#2 is called, the matrix GM is immediately up-dated to include the "new" observations U and Y. Computation of the matrices A, B, C, D, and T is then based on the up-dated Gram matrix GM. The matrix X which is returned to the calling program is consistent with the updated T and is computed as a transformation on U and Y.

In order to initiate a recursive identification scheme, the matrix GM can be set uniformly equal to zero. Then the coefficient matrices computed by the first call to ID#2 will be based only on the first set of K observations on U and Y. It is important to note that the system order will always be estimated as zero if K for the first set of observations is less than or equal to (NN+1)M.

Considerable versatility is realized by the fact that successive calls to ID#2 need not necessarily reflect equal numbers of added observations. Also, it is easy to see that an arbitrary amount of time can be allowed to elapse before successive updating operations consistent with specific modeling requirements.

Since A, B, C, D, and T are computed anew based on the updated matrix GM, it is conceivable that successive computation may yield varying estimates of system order and structure commensurate with input "generality" and sample size. A simple, but non-trivial example is where the first set of observations is of insufficient number to yield a Gram matrix of rank consistent with system order.

Finally, it should be noted that since A, B, C, D, T, and X are allocated each time ID#2 is called, these values should be freed before the next call to ID#2 unless it is desired to "stack" successive allocations as is possible in PL/1.