AD 741437

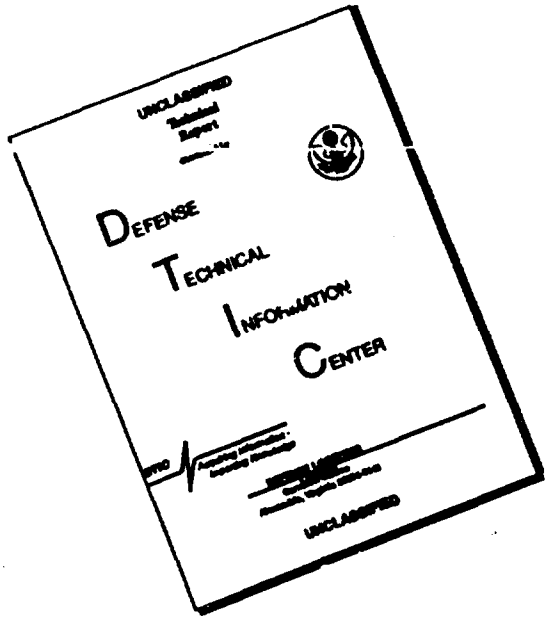
UNITED STATES AIR FORCE

AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

DISCLAIMER NOTICE



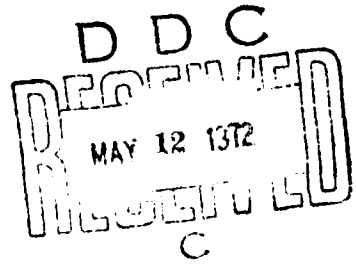
THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

THE IDENTIFICATION OF LINEAR STOCHASTIC SYSTEMS

THESIS

GSA/MA/72-5

Adolph Harris



GSA/MA/72-5

THE IDENTIFICATION OF LINEAR STOCHASTIC SYSTEMS

THESIS

Presented to the Faculty of the School of Engineering
of the Air Force Institute of Technology
Air University
in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

by

Adolph Harris, B.S.As.E., M.S.E.E.

March 1972

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*

Air Force Institute of Technology
Wright-Patterson Air Force Base, Ohio

2a. REPORT SECURITY CLASSIFICATION

Unclassified

2b. GROUP

3. REPORT TITLE

THE IDENTIFICATION OF LINEAR STOCHASTIC SYSTEMS

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

AFIT Thesis

5. AUTHOR(S) *(First name, middle initial, last name)*

Adolph Harris

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
March 1972	106	185

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO. N/A	GSA/MA/72-5
c.	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
d.	

10. DISTRIBUTION STATEMENT

Approved for public release, distribution unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY

13. ABSTRACT

This research considers the problem of the identfication of linear stochastic systems. A current state-of-the-art assessment of the general field of system identification is given, and criteria for the classification and selection of identification methods are presented and discussed. Several of the more popular identification methods from the literature are investigated and summarized. A bibliography containing 185 references, keyed by identification methods and other relevant headings, is included.

Using the state variable formulation for a discrete linear stochastic system, a detailed exposition of a few of the on-line identification methods currently appearing in the literature is presented. One such method, based on the autocorrelation function of the output measurements, is developed to identify the state transition matrix and the output noisc covariance (vector-input, scalar-output case). It is shown that a canonical or phase variable system representation can be used to reduce the number of unknown parameters requiring identification. Finally, an on-line identification method called Levy's proper canonical form, which is based on the Kalman filter representation of the system, is derived using the innovations sequence and certain results from optimal estimation theory. It is shown that this identification method results in still additional advantages over the identification methods previously developed.

DD FORM 1 NOV 65 1473

Unclassified
Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Identification						
Stochastic Estimation and Control						
Kalman Filtering						
Linear Stochastic Systems						
Estimation						
Canonical Forms						

Preface

The desire to undertake research in the field of system identification initially grew out of my experience as a practicing aerospace engineer working principally on the design and analysis of stability-and-control characteristics of vertical take-off and landing (VTOL) aircraft. The myriad problems encountered in this regard soon led me to graduate study in both aerospace and electrical engineering, chiefly in the areas of aeromechanics and automatic control theory. As a result, it soon became increasingly apparent that the VTOL parameter identification problem was conceptually similar to identification problems frequently occurring in various areas of electrical engineering. My interest was further nurtured when subsequent graduate work in economics and systems analysis revealed still other applications requiring identification of the parameters of linear stochastic systems.

The first three chapters of this thesis should be of particular interest to the reader who has a general curiosity about the subject area, but who, for one reason or another, is not particularly interested in the mathematical details. In these chapters I have endeavored to introduce basic concepts and present a non-technical assessment of the current state-of-the-art, including a descriptive exposition of some of the more popular methods of system identification. Where material has been condensed and summarized from the

literature, it has been my aim to do so in such a way as to maintain an objective, illuminating perspective on the subject.

My research into the subject area began with a fairly extensive literature search, the results of which are contained in the bibliography. The articles it lists are keyed not only by identification methods, but also by general headings directly related to system identification. The bibliography is, I think, relatively complete, and should serve as a ready source for the reader who wishes to investigate further certain areas of system identification.

I wish to express my sincere gratitude to my thesis advisor, Dr. David R. Barr, Department of Mathematics, for the many helpful suggestions and constructive comments he provided throughout this study. Finally, special thanks are offered to my wife, Veronica. Without her patience and understanding this investigation could not have been completed.

Contents

List of Figures

Notation

Symbol	Meaning	Page Where First Occurs
a_i	scalar coefficients of the characteristic polynomial of $\underset{\sim}{\Phi}$	45
α	the innovations sequence	63
$\underset{\sim}{B}_j$	n x n nonsingular matrix	43
$\underset{\sim}{C}_k$	autocorrelation function	43
δ_{jk}	Kronecker delta	34
$E[\cdot]$	mathematical expectation	32
$\underset{\sim}{F}$	p x p nonsingular matrix	58
$\underset{\sim}{G}$	n x 1 vector	69
$\underset{\sim}{\Gamma}$	n x p disturbance transition matrix	34
$\underset{\sim}{\Gamma}_e$	n x p equivalent disturbance transition matrix	58
$\underset{\sim}{\Gamma}^*$	canonical form of $\underset{\sim}{\Gamma}$	51
$\underset{\sim}{H}$	r x n measurement matrix	34
$\underset{\sim}{H}^*$	canonical form of $\underset{\sim}{H}$	51
$\hat{}$	conditional mean of a vector	38
i	integer, index parameter	40
I	index set	30
$\underset{\sim}{I}$	identity matrix	45
j	integer, index parameter	31
k	integer, index parameter	30
$\underset{\sim}{K}$	n x 1 Kalman filter gain	71
$\underset{\sim}{K}^*$	n x 1 optimal Kalman filter gain	73
λ	an eigenvalue of $\underset{\sim}{\Phi}$	45

Symbol	Meaning	Page Where First Occurs
n	integer, order of the system	31
p	integer	31
$\underset{\sim}{P}$	n x n covariance matrix of $\underset{\sim}{X}$	35
P_Y	scalar covariance of $\underset{\sim}{Y}$	67
$\underset{\sim}{\Phi}$	n x n state transition matrix	34
$\underset{\sim}{\Phi}^*$	canonical form of $\underset{\sim}{\Phi}$	51
$\underset{\sim}{Q}$	p x p positive definite input-noise covariance matrix	34
r	integer	34
$\underset{\sim}{R}$	r x r positive definite output-noise covariance matrix	35
s	integer, index parameter	37
σ	integer, index parameter	70
T	number of observations	48
$\underset{\sim}{T}$	n x n nonsingular transformation matrix	51
$\sim$	denotes a vector or matrix ; above denotes estimation error of a vector	30 38
$'$	denotes the transpose of a vector of matrix	32
$\underset{\sim}{U}$	p x 1 gaussian white-noise system disturbance vector	34
$\underset{\sim e}{U}$	p x 1 equivalent gaussian white-noise system disturbance vector	58
$\underset{\sim}{V}$	r x 1 gaussian white-noise measurement error vector	34
$\underset{\sim}{X}$	n x 1 state vector	34
$\underset{\sim}{X}^*$	canonical form of $\underset{\sim}{X}$	51

Symbol	Meaning	Page Where First Occurs
$\underset{\sim}{Y}$	r x 1 vector, the message	39
$\underset{\sim}{Z}$	r x 1 output measurement vector	34

Abstract

This research considers the problem of the identification of linear stochastic systems. A current state-of-the-art assessment of the general field of system identification is given, and criteria for the classification and selection of identification methods are presented and discussed. Several of the more popular identification methods from the literature are investigated and summarized. A bibliography containing 185 references, keyed by identification methods and other relevant headings, is included.

Using the state variable formulation for a discrete linear stochastic system, a detailed exposition of a few of the on-line identification methods currently appearing in the literature is presented. One such method, based on the autocorrelation function of the output measurements, is developed to identify the state transition matrix and the output noise covariance (vector-input, scalar-output case). It is shown that a canonical or phase variable system representation can be used to reduce the number of unknown parameters requiring identification. Finally, an on-line identification method called Levy's proper canonical form, which is based on the Kalman filter representation of the system, is derived using the innovations sequence and certain results from optimal estimation theory. It is shown that this identification method results in still additional advantages over the identification methods previously developed.

GSA/MA/72-5

THE IDENTIFICATION OF LINEAR STOCHASTIC SYSTEMS

I. Introduction

Purpose

The purpose of this thesis is to present:

1. A survey of the current state-of-the-art of the identification of linear stochastic systems.

2. A classification of many of the various identification methods currently appearing in the literature, in terms of the type of the model, input signal, and optimization criterion used.

3. The relevant properties of the different identification methods that differentiate one from another, and provide the systems analyst with valuable insight into which identification method to use in a given application.

4. The development of selective topics from the literature of system identification, including Mehra's on-line identification method and system identification via adaptive Kalman filtering.

The Identification Problem

Very basically, the identification problem is simply the determination of the mathematical model that describes a given physical process (i.e., the structure and parameters of the mathematical model). If the system designer is only interested in the parameter identification problem, he may

choose to model the process as shown in Fig. 1. In the model depicted in Fig. 1, the plant is that element which relates the system input to the system output. As such, the plant

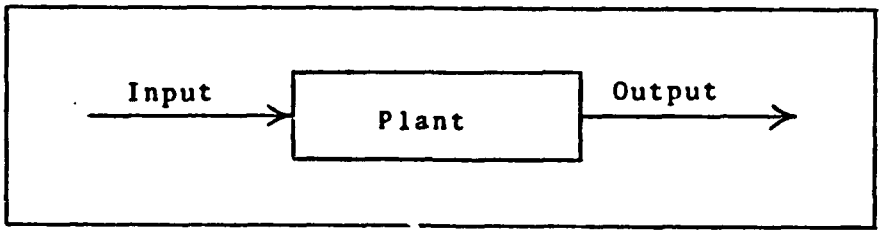


Fig. 1. Model for Parameter Identification Only.

can be thought of as consisting of the process parameters interconnected through some mathematical relationship such that, for a given system input, the output states of the process are completely determined. Hence, for a discrete-time process, the coefficients (constant or time-varying) of the difference equation relating system input to output may be regarded as functions of the process parameters. In many physical processes, some or all of the process parameters may be unknown, or may be known initially but change stochastically as the process unfolds. The parameter identification problem consists of the determination of these unknown parameters.

During the past few years an increasing number of papers have appeared in the literature on the general subject of system identification. These papers have appeared in such diverse fields as economics (Refs 105 and 106), industrial

processes (Refs 107 and 111), and biology (Ref 104). Generally, there are two primary motivations for the current interest in system identification. First, the system designer may only be interested in identifying the parameters of a physical process under consideration. Second, he may require knowledge of the process parameters for use in a subsequent control engineering application. Although obviously related, these two motivations can have distinctly different implications, which are discussed at length in Chapter II.

An example of a parameter identification problem that arises often in connection with economic analysis work concerns the Leontief input-output model (Ref 109). In this model, the national economy is broken down into N industries, each of which can be thought of as producing a certain output product. In general, the output product of each industry is used as an input, along with certain other inputs (e.g., raw materials), to the other industries in the economy. An input-output table is then constructed. The elements, a_{ij}, $i, j = 1, \ldots, N$, appearing in the table are called input-output coefficients, and refer to the amount of input i required by industry j to produce one unit of output product j. The identification problem consists of the determination of the input-output coefficients necessary to sustain a given level of demand for the various products in the economy.

On the other hand, if the system designer is interested in the identification of process parameters for subsequent

use in a control engineering application, he may choose to model the process as shown in Fig. 2. The model shown in Fig. 2 differs from that shown in Fig. 1 due to the addition of a feedback control loop. As shown, the feedback loop

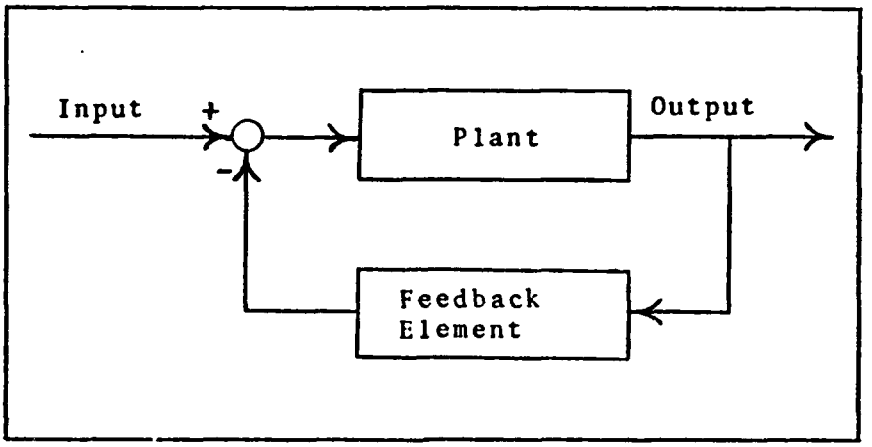


Fig. 2. Model for Parameter Identification Used for Control Application.

contains an element that operates on the output signal to yield a signal that is compared to the input signal. The difference between these two signals is then used as the effective input signal to the plant. If the design problem under consideration is one of control system synthesis, the system designer must determine the structure and parameters of the unknown feedback element that will enable the system output to follow, as closely as possible, some desired signal. The design of such a control strategy for a given process not only requires knowledge of the structure and parameters of the process to be controlled, but also knowledge of the various process states and how they change with time (the

estimation problem). However, in order to solve the estimation problem, one must first postulate a structure for the process model, and then identify the corresponding parameters of the model.

As an example of a situation in which parameter identification is required for use in a control engineering application, consider the problem of designing a stability augmentation system for a hovering vertical take-off and landing (VTOL) aircraft (Ref 108). If the process under consideration is the VTOL aircraft hovering in still air, it is fairly straightforward to calculate the process parameters (i.e., aircraft stability-and-control derivatives), and hence the state of the aircraft. However, as the VTOL aircraft transitions from hovering flight to aerodynamic wing-supported flight, the aerodynamic and propulsion forces acting on the aircraft interact in such a way as to make a reasonable prediction of the resulting aircraft stability-and-control derivatives extremely difficult. Thus, a parameter identification problem exists during this transition flight regime which must be solved before a satisfactory stability augmentation system can be designed.

In connection with the current interest in system identification for control engineering application, two points should be made. First, tremendous advances have been made in computer technology during the past few years. These advances now enable the system designer to develop computer-

based algorithms to solve complicated equations which in many instances were not computationally practical before. Second, many control engineering applications are such that if the process parameters are unknown, they can often be determined experimentally (Ref 135).

Outline

In Chapter II, an overview of system identification is presented. The general characteristics of the identification problem are discussed, including its assumptions and limitations. A convenient method for classifying various identification schemes is also provided, along with a discussion of those factors that might be considered in selecting an identification method for a particular application.

The identification of linear stochastic systems is discussed in Chapter III. Identification methods included are maximum likelihood, instrumental variables, adaptive estimation, model reference, stochastic approximation, linear least-squares, and correlation techniques. The approach taken in this chapter is to summarize, from the sytem identification literature, a few of the outstanding papers on each of the identification methods listed above. A comprehensive bibliography, keyed by identification method, is included to indicate other publications not specifically referenced in this paper.

Mehra's on-line identification scheme is the subject of Chapter IV. The development presented includes identification

of the state transition matrix, as well as the output noise covariance. A canonical form representation of the system is also developed for the case where the designer is concerned only with obtaining a system model that yields the desired output, and not interested in the particular structure of the model.

Chapter V approaches the identification problem through the use of adaptive Kalman filtering. The particular system representation that is discussed is Levy's proper canonical form, which is derived using the innovations sequence and certain results from optimal estimation theory. A method for determination of the optimal filtering gains is also discussed. Finally, the conclusions and recommendations for further study are presented in Chapter VI.

II. An Overview of System Identification

In this chapter, the general field of system identification is examined from an overview perspective. Without question, such an endeavor as this must of necessity remain somewhat incomplete. This is not surprising when one considers the rapid development presently occurring in system identification, as witnessed by the multitude of recent papers published on the subject. For a more comprehensive survey, the reader is referred to the literature (Refs 142 and 156).

An investigation into the literature of system identification reveals one immediate fact: there appears to be little, if any, unification of the field. In addition to the dozen-or-so recognized identification methods appearing in the literature, there are a sizeable number of publications within each method. The result is that the non-expert is often left somewhat bewildered and confused. Therefore, the major aim of this chapter is to provide an exposition of the underlying rationale necessary for improved unification and classification of the field of system identification.

General Characteristics of Identification Problems

A crucial first step in the formulation of any identification problem consists of answering the following two questions:

1. What is the system being considered?

2. What is the purpose of the desired identification? Is the system under consideration a portion of some larger process, the entire process itself, or the process in interaction with its environment? Obviously, answers to such questions are necessary in order to clearly define the system, and indicate possible limitations in the resulting identification analysis. Similarly, having the purpose for which the identification is desired clearly in mind can often suggest possible identification methods that might be appropriate for a particular application.

On the question of the purpose of the desired identification, it has been pointed out earlier in this paper that a primary motivation appears to be possible engineering control applications. However, many situations exist where the purpose of the desired identification is only to identify the parameters of the process. These two motivations can be distinctly different, possibly implying the application of different identification methods. For example, consider as the process an aircraft in steady unaccelerated flight. If the design task is to determine whether the aircraft has the required degree of static stability at this flight condition, the problem becomes one of parameter identification alone (i.e., determination of the static stability-and-control derivatives). In this case, rather precise parameter identification is usually required. However, if the design task is to determine whether the aircraft has the required degree of maneuver capability, the problem becomes one of parameter

identification for control application. In this case, the parameter identification part of the problem is generally not as critical as in the former case, because the aircraft's flight control system can usually be designed so as to be flexible enough to cover a variety of "close" models.

The preceding example illustrates an important relationship that often exists between the identification problem and the control problem. In many situations where parameter identification is only an intermediate step to a subsequent control application, the assumption is tacitly made that the problems of identification and control can be separated. Generally speaking, a control system designed using this assumption will not be optimum, primarily because the parameter identification obtained is seldom exact. For this reason, a significant number of papers have appeared recently in the literature dealing with adaptive estimation and model reference techniques. These methods attempt to deal with the problems of identification and control more or less simultaneously, and are taken up in Chapter III.

Classification of Identification Methods

Generally speaking, there are two ways in which the system identification problem may be approached:

1. Theoretical (Mathematical-Physics) Analysis
2. Experimental Analysis.

In the theoretical approach, one typically begins by considering the physical laws that govern a given process, and

attempts to formulate a mathematical model of the process. The identification problem in this case becomes the determination of the parameters of the model in terms of the physical data of the process. On the other hand, the experimental approach attempts to identify the parameters of the process by measuring the dynamic behavior (input-output) of the process. The identification methods discussed in Chapter III are chiefly of the experimental approach.

A useful framework for considering the problems of system classification and identification is provided by Zadeh (Ref 185), who defines system classification as follows:

> "Given a black box B and a family of classes of systems C_1, C_2, ..., C_k, ..., C_n such that B belongs to one of these classes, say C_k, the (classification) problem is to determine C_k by observing the responses of B to various inputs."

The problem of system identification is then defined as a special case of the classification problem in which each of the classes C_1, C_2, ..., C_n has just one member:

> "Given a class of systems C with each member of C completely characterized, the (identification) problem is to determine a system in C which is equivalent to B."

The foregoing definitions imply that in order to formulate the identification problem, one must specify a class of systems, a class of input signals, and some criterion for defining "equivalence" between model and system. Each of these requirements is discussed briefly below.

A variety of representations may be used to characterize different classes of mathematical models, e.g., linear, non-linear, continuous, discrete, etc. The form of model selected can, of course, significantly affect the ensueing identification procedure. For the identification of linear stochastic systems, models are often classified as either parametric or non-parametric, depending upon whether or not the appropriate probability distributions are known. Most of the models discussed in Chapter III represent discrete linear systems with some form of additive "noise" present.

Although a wide range of input signals may conceivably be used in the identification of process parameters, major simplifications in the identification procedure can often be achieved by choosing input signals of a special type. This situation is analogous to the simplification that occurs in the identification of linear deterministic systems when an impulse function is used as the input signal. Aoki and Staley (Ref 83) discuss the input signal synthesis problem of parameter identification, and derive conditions necessary for obtaining a class of asymptotically unbiased and efficient parameter estimates. In Chapter III, the input signal used most frequently in the various identification methods is a gaussian "white-noise" sequence.

Once the form of the mathematical model and input signal have been selected, the remaining requirement in formulating the identification problem is the specification of some criterion to be used to measure the equivalence between the

model and system. The selection of such a criterion essentially transforms the identification problem into an optimization problem. For many of the identification methods appearing in the literature, this optimization usually involves the minimization of some function of an "error signal". The error signals most frequently used are input errors, output errors, and (where necessary probability distributions are available) errors in the parameter values.

Selection of an Identification Method

Before moving on to a discussion of various identification methods, it might be appropriate at this point to make a few comments concerning some factors that might be considered in selecting an identification method for a particular application. Certainly, a primary consideration concerns the purpose for which the identification is desired, as discussed earlier in this chapter. In most practical engineering control applications, seldom does the designer have available a priori complete knowledge about the process to be controlled and its environment. However, any reasonable steps that the designer can take to increase his knowledge concerning the system--e.g., rough calculations, comparisons with similar systems, use of engineering experience and judgement--can often provide valuable insight into possible identification methods that might be appropriate.

Another factor that can sometimes influence the choice of a particular identification procedure concerns the required

accuracy of the identification analysis. Regardless of whether the ultimate purpose of the identification analysis is parameter identification or control application, the degree of accuracy required is an important factor that can often dictate the type of identification method to employ. However, a major problem exists in this regard in determining the degree of accuracy required of the identification analysis, and formulating a criterion for measuring that accuracy. Unfortunately, the literature of system identfication is far from complete in the area of the accuracy of identification analyses.

III. The Identification of Linear Stochastic Systems

The overview of system identification presented in the previous chapter is used as the basis for a general discussion of various identification methods in this chapter. The approach selected to accomplish this objective is to summarize, from the system identification literature, one or more recent papers felt to be illustrative of each identification method discussed. In order to keep the discussion in this chapter as generalized as possible, most of the notational details peculiar to the various identification methods have been omitted. However, many of the definitions are included in the first two sections of Chapter IV; the others may be found by consulting the indicated references. This approach is not meant to be totally comprehensive, but rather merely to expose the reader to some of the underlying concepts attendant to the different identification methods.

An endeavor such as this is almost certain to be somewhat incomplete, as no attempt has been made to include every identification method or variation appearing in the literature. Furthermore, it is conceivable that some applicable papers may have been inadvertently overlooked. It is hoped, however, that the more popular methods have been included for discussion in this chapter. A list of references for each identification method is contained in the bibliography.

Maximum Likelihood

The maximum likelihood identification of the parameters of a discrete, stationary linear stochastic system from noisy input-output measurements is discussed by Kashyap (Ref 4). In particular, he considers the problem of fitting linear models to the observed output data of a physical process. The mathematical model used to describe the process consists of a vector input-output, linear nth order difference equation having constant but unknown matrix coefficients. The system is assumed to be driven by a zero-mean input noise disturbance which is completely specified by a set of n unknown correlation matrices. The output measurements are assumed to be contaminated by a zero-mean, uncorrelated measurement noise sequence having a constant but unknown correlation matrix. The input and output noise sequences are assumed to be mutually independent, and the identification problem consists of identifying, from the input-output record of the process, optimal estimates for the various unknown matrices. The approach selected by Kashyap is to denote the parameters of the unknown matrices by a parameter vector, and assume that the conditional probability density function of the output measurement (given all previous output measurements and the unknown parameter vector) has a multivariate normal distribution. This allows the likelihood function to be derived, and then maximized with respect to the unknown parameter vector. The optimal estimates are then obtained by solving a resulting parameter optimization

problem. Conditions are also derived under which the maximum likelihood estimates (MLE) are unique, asymptotically unbiased, and consistent.

The problem of determining the unknown parameters of a dynamic system from noisy input-output observations is also considered by Aoki and Yue (Ref 1). Their paper examines in detail the estimation errors of two algorithms that approximately compute the maximum likelihood estimates of the system parameters. A scalar input-output, linear nth order difference equation with constant but unknown coefficients is used as the mathematical model of the system. The input and output observations are assumed to be corrupted by mutually independent, additive white-noise sequences having a standard normal distribution. The approach taken by Aoki and Yue is to combine the unknown coefficients into a parameter vector, and show that the approximate MLE of the unknown parameter vector is equivalent to the computation of an eigenvector of a real symmetric matrix. It is shown that the algorithms for the approximate MLE yield a global optimum for the unknown parameter vector, whereas the algorithms for the true MLE yield only local optimum and hence require iteration to obtain the global optimum. Necessary and sufficient conditions are derived for the approximate MLE to converge with probability one to its true value as the sample size tends to infinity. For finite sample sizes, explicit bounds are derived for the mean-square error of the approximate MLE.

The maximum likelihood identification of stochastic linear dynamic systems using the Kalman filtering representation is discussed by Mehra (Ref 5). The mathematical model used is a state variable representation of a discrete, single-input, single-output linear dynamic system with constant but unknown system matrices. The input and output observations are assumed to be corrupted by zero-mean, mutually independent, gaussian white-noise sequences having constant but unknown covariance matrices. The approach followed by Mehra is to transform the given model to an equivalent state space representation called "Levy's proper canonical form" (this representation is discussed in detail in Chapter V). In effect, the output measurements are "whitened" through use of a causal invertible linear transformation (Kalman filter). The MLE of the unknown system parameters is then obtained by maximizing the conditional probability density function of the output measurement (given all previous output measurements and the unknown system parameters), subject to the equations defined by Levy's proper canonical form. Conditions are indicated under which the MLE are unbiased, consistent, and efficient.

Instrumental Variables

Wong and Polak (Ref 15) consider the use of the instrumental variable method to estimate the parameters of discrete, linear time-invariant systems. In general, the method is restricted to applications in which the controlled-input

(if any) is noise-free, while the observed output may be corrupted by white-noise.

Essentially, the instrumental variable method estimates a set of unknown parameters from an array of linear algebraic equations involving these parameters, a set of controlled-input observations, and a set of noise-corrupted output observations. The output measurement-noise sequence is assumed to be a sample from a zero-mean stationary noise process whose covariance function, $r(t)$, tends to zero at a rate faster than $1/t$ as $t \to \infty$. The controlled-input sequence is assumed to be either deterministic, or else a sample from a stationary random process statistically independent of the output noise process. A further restriction in the method is that the number of equations in the array must be greater than the number of unknown parameters to be estimated. Wong and Polak show that by first premultiplying the given array by a suitable rectangular matrix, called the instrumental matrix, a square invertible array is obtained which may then be solved for the required parameter estimates. The elements of the instrumental matrix are called instrumental variables.

Since the instrumental variable method cannot be used with certain inputs (e.g., controlled-inputs corrupted by noise), Wong and Polak derive a necessary and sufficient condition for the instrumental matrix to exist when the input is deterministic and bounded. They also show that under suitable conditions on the system input and the output

measurement-noise, optimal instrumental variables exist corresponding to two criteria of optimality. The method, when applicable, is always shown to yield consistent estimates of the parameters.

Adaptive Estimation

In most physical processes, the mathematical model representation of the process can usually be specified only up to an unknown set of parameters. If $\underset{\sim}{X}(t)$ denotes the state vector of a process and $\underset{\sim}{\theta}$ the time-invariant parameter vector, then the problem of obtaining the optimal (in some sense) estimate of the state vector under the condition of the uncertainty of the parameter vector is often referred to as adaptive estimation. In essence, it is the simultaneous estimation of the state vector and identification of the process parameters.

The problem of the optimal estimation of a sampled, Gauss-Markov stochastic process when certain parameters of the process are initially unknown is discussed by Magill (Ref 26). His approach is to assume that the unknown parameters belong to a set that contains a finite number of possibilities which are known a priori. The given stochastic process may then be represented by a set of "elemental stochastic processes" (one for each possible combination of parameters), a switch that is permanently but randomly connected to one of the elemental stochastic processes, and a set of a priori probabilities for the set of switch

positions. The elemental stochastic processes are represented as the outputs of linear dynamic systems excited by white gaussian noise.

Magill shows that the optimal estimate is one that is obtained by minimizing a generalized mean-square-error criterion. In particular, he shows that the optimal state estimate is obtained by taking the complete set of state estimates conditioned on all available output observations, weighting each with the conditional probability that the appropriate parameter vector is true, and summing over the space of all possible parameter values. The conditional weighting coefficients are determined by an application of Bayes' rule, where the elemental stochastic processes are assumed to have a multivariate normal density function.

Hillborn and Lainiotis (Ref 22) have extended the work of Magill to cover non-Gaussian, Markov processes with unknown parameters. The processes are characterized as having probability distributions of known functional form, but containing a set of unknown parameters. It is assumed that all initial knowledge of the unknown parameters can be expressed by appropriate probability distributions (Bayesian estimation), with the result that the optimal state estimate may be determined without going through the intermediate step of parameter identification per se. For sampled stochastic processes having finite-state unknown parameters and a generalized Markov property, Hillborn and Lainiotis show that the optimal state estimates can be formed from a

set of optimal estimates based on known parameters, and a set of "learning" statistics which are updated recursively. Necessary and sufficient conditions are also established for the convergence or "learning" of the constant unknown parameters.

The adaptive estimation techniques discussed in the previous two papers have been extended by Lainiotis (Ref 24) to cover both structural and parameter adaptation, where structural adaptation refers to the unknown dimensionality of the state vector. The Bayesian approach to the adaptive estimation problem is utilized by Lainiotis in assuming that the system generating the random processes involved is chosen at random from a finite collection of possible systems. These systems are characterized as having state vector dimensionality σ, parameter vector value θ_σ, and known or assumed a priori probability $Pr[\sigma,\theta_\sigma]$. It is further assumed that the model structure σ is less than some fixed number n, and the defining parameter vector θ is time-invariant. The assumption of an upper bound n to the system dimensionality permits structural adaptation to be imbedded into parameter adaptation. This is accomplished by arbitrarily choosing the dimensionality of the model as n, and determining those system parameters that are zero if the correct model structure is less than n.

Using the approach of augmenting the state vector with the parameter vector, Lainiotis shows that the optimal state

estimator can be decomposed into two parts: a linear non-adaptive part consisting of a bank of ordinary Kalman-Bucy filters matched to each admissible value of the unknown parameter vector, and a nonlinear part consisting of likelihood ratios that incorporate the adaptive learning nature of the estimator. In addition, the conditional-error covariance matrix is also derived for on-line performance evaluation.

Model Reference

Model reference or model tracking techniques are ideally suited to applications where it is desired to obtain the parameter identification result recursively as the process develops (called on-line or real-time identification). Such techniques also enjoy wide applicability in the identification of processes with time-varying parameters. A typical model reference diagram is shown in Fig. 3. Use of the model

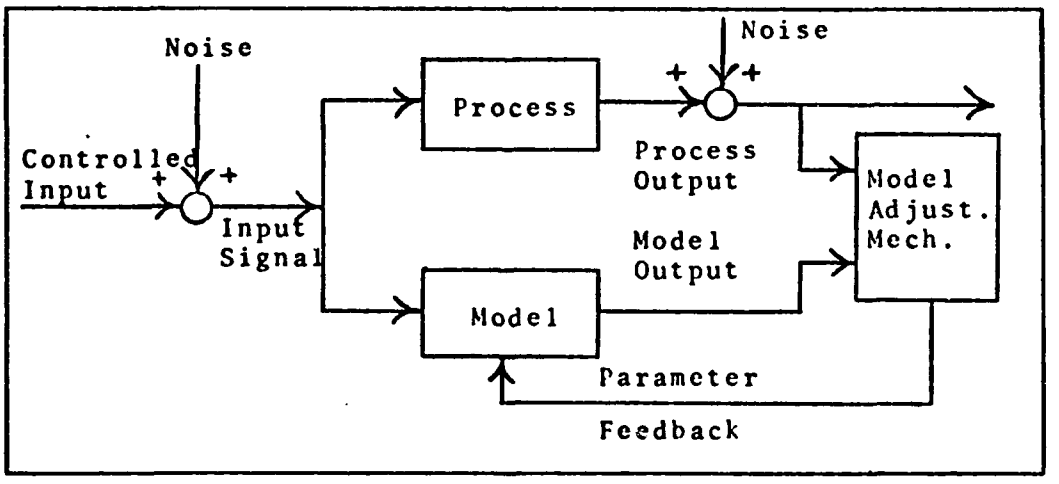


Fig. 3. Identification by Model Reference.

reference technique begins by formulating, for the physical process under consideration, a mathematical model having adjustable parameters. The system input signal is then simultaneously fed into the actual process and its mathematical model, after which the output of each is fed into a model adjustment mechanism. This mechanism operates on the adjustable model parameters in such a way as to make the model output follow, as closely as possible, the output of the actual process.

A model reference technique for identifying the parameters of dynamic systems modeled by differential equations is discussed by Hsia and Vimolvanich (Ref 37). An identification algorithm, based on the learning model concept, is derived using the state-variable formulation. A multiple-input, single-output noise-free system is first considered, with the results then extended to the case where the system has multiple-outputs and noise present. The system under consideration is characterized by an nth order vector differential equation, and a corresponding model with unknown but adjustable parameters is assumed. Arbitrary initial values of the model parameters are assumed, and the identification objective becomes to adjust the model parameters so that they converge toward the corresponding parameter values of the actual system. The parameter adjustment procedure adopted by Hsia and Vimolvanich is to minimize the identification error, i.e., the difference between the

system's output and the model's output. To accomplish this, equations are derived for the adjustment of the model parameters, and the identification error is shown to converge in the mean.

Stochastic Approximation

Stochastic approximation is an identification method that can be used when the system input is assumed to be a stationary stochastic process. Although some variations of the method may be used without prior knowledge of the statistics of the process, most require knowledge of the noise covariances to yield consistent estimates. The parameter estimates obtained using stochastic approximation methods generally have larger variances than those obtained using other methods (e.g., least-squares estimates).

Ho and Lee (Ref 51) consider the problem of deriving a real-time-convergence identification scheme for linear dynamical systems using stochastic approximation. A linear discrete model in state variable form is used to derive an algorithm for determining the elements of the state transition matrix (in canonical or phase variable form). The method is based upon the following assumptions: the system is excited by a zero-mean, white-noise process having known variance; the system output is measured exactly, i.e., the output measurement-noise is zero; and the transfer function between the input noise sequence and the system output has no numerator dynamics. The resulting algorithm is shown to

yield estimates that, although not optimal in a stochastic sense, nevertheless represent the least-square fit to the measurement data. The parameter estimates obtained using the algorithm are shown to converge to their true values in the mean-square sense.

The work of Ho and Lee has been generalized by Saridis and Stein (Ref 58) to cover the on-line identification of forced, discrete linear systems from a sequence of white-noise-corrupted output measurements. The system is modeled in canonical or phase variable form using the state variable representation. The Robbins-Monro stochastic approximation procedure is used to derive an identification algorithm that uses only measurements of the system input and output, and does not require knowledge of the output measurement noise statistics. However, if the input measurements are also corrupted with white-noise, the algorithm requires knowledge of the variance of this noise sequence. The algorithm is also shown to converge to the true value of the parameters in the mean-square sense.

A slightly different approach to the parameter identification problem using stochastic approximation is discussed by Sakrison (Ref 57). He considers the on-line estimation of the poles and zeros of a rational transfer function whose order is known not to exceed some fixed number n. A linear, time-invariant system model is assumed, with both the system input and output observable in the presence of mutually uncorrelated, stationary random noise processes having

zero-mean and known correlation functions. The approach taken by Sakrison is to use stochastic approximation methods to compute on-line an optimum filter, from which the desired coefficients in the system transfer function may be obtained. The parameter estimates generated using the derived algorithm are shown to converge, in the mean-square sense, to their true values.

Linear Least Squares

Steiglitz and McBride (Ref 69) discuss an iterative technique to identify a linear system from samples of its input and output in the presence of noise. The model assumed is a linear sampled-data system, with input and output related by a rational Z-transform. The approach selected by Steiglitz and McBride is to minimize the mean-square error between the model output and the observed output of the plant. A technique is derived for carrying out this minimization by iteratively performing a sequence of Kalman least-squares linear regressions on the system's input-output data. Each iteration is shown to be computationally equivalent to an ordinary Kalman linear regression, except for prefiltering of the input and output data. Experimental results are presented to show that the iterative identification method converges more slowly than the ordinary Kalman linear regression method, but results in improved parameter estimates.

The problem of identifying, from input-output measurements, the unknown parameters in linear dynamical systems with transport lags is considered by Hsia (Ref 63). An nth order, linear differential-difference equation model is assumed, with noise-free input and output. The identification technique derived by Hsia essentially involves two steps: (1) the use of a finite difference technique to reduce the differential-difference equation to an ordinary difference equation, and (2) estimation of the system parameters through identification of the resulting discrete model via Kalman's least-square method.

Correlation

Anderson, et. al. (Ref 72) discuss the problem of determining consistent estimates of the parameters of a linear dynamic system. A discrete linear model in the state variable formulation is assumed, with uncorrelated, additive noise present at both the input and output. The state transition matrix and input-output noise covariance matrices are assumed to be unknown, and thus require identification. Estimates of these matrices are derived and shown to be strongly consistent when the linear system is stable, i.e., when the eigenvalues of the state transition matrix lie within the unit circle. In addition, the asymptotic properties of the model are investigated, and shown to remain unchanged when the unknown system parameters are replaced by their strongly consistent estimates.

The problem of estimating the autoregressive parameters of a mixed moving-average time series of known order using output data alone is considered by Gersch (Ref 75). The problem is shown to be equivalent to the estimation of the denominator terms of a scalar transfer function of a stationary, linear discrete-time system. Three formulations of the system are discussed: state variable, time series, and Z-transform representations. It is assumed that the system is excited by a zero-mean, uncorrelated input noise sequence of unknown variance, and the output observations are exact (i.e., the output measurement-noise is zero). The approach followed by Gersch is to derive a modified set of Yule-Walker equations, which are then used to solve for an asymptotically unbiased estimator of the unknown autoregressive parameters. The estimator is also shown to be unbiased in the presence of additive white output measurement-noise of arbitrary finite correlation time.

In the next chapter, the on-line identification of discrete linear stochastic systems is considered. The identification methods discussed are based largely upon the statistical correlation properties of the output measurements. As such, the methods are similar to those employed in time series analysis (Refs 75, 80, and 82).

IV. On-Line Identification of Discrete Linear Stochastic Systems

In this chapter, attention is turned to an in-depth examination of an identification method recently proposed by Mehra (Refs. 78 and 79). The method is more general than some others currently appearing in the literature (e.g., instrumental variables and stochastic approximation), in that complete knowledge of the input-output noise covariance matrices is not required. Additionally, the method is capable of providing on-line identification when implemented using a digital computer.

In the interest of notational simplicity and continuity, only the discrete case will be discussed in the remainder of this paper. This should cause no serious limitations, since the method can be extended to the continuous case with little difficulty.

Preliminaries

For the convenience of the reader, a few preliminary definitions are introduced in this section. The terminology and notation appearing in the remainder of this paper are fairly standard, being used rather extensively in the literature of system identification and estimation (Ref 180:106-121).

Definition 1: A stochastic process is a family of random vectors $\{\underset{\sim}{X}(k), k\in I\}$ indexed by a parameter k all of whose values lie in some appropriate index set I.

Definition 2: Let $\underset{\sim}{X}$ be an n vector. Then the stochastic process $\{\underset{\sim}{X}(k), k\epsilon I\}$ is said to be independent if, for any m time points $k_1, \ldots, k_m$ in I, where m is any integer, the joint probability distribution function of the m random n vectors $\underset{\sim}{X}(k_1), \ldots, \underset{\sim}{X}(k_m)$ is equal to the product of the probability distribution functions of the m individual n vectors. That is

$$P[\underset{\sim}{X}(k_1) \leq \underset{\sim}{x}_1, \ldots, \underset{\sim}{X}(k_m) \leq \underset{\sim}{x}_m] = \prod_{i=1}^{m} P_i[\underset{\sim}{X}(k_i) \leq \underset{\sim}{x}_i]$$

for all n vectors $\underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_m$.

Definition 3: Let $\{\underset{\sim}{X}(k), k\epsilon I\}$ and $\{\underset{\sim}{Y}(k), k\epsilon I\}$ be two stochastic processes, where $\underset{\sim}{X}$ is an n vector and $\underset{\sim}{Y}$ is a p vector. The two stochastic processes are said to be independent of each other if, for any m time points $k_1, \ldots, k_m$ in I, where m is any integer, the joint probability distribution function of the 2m random vectors $\underset{\sim}{X}(k_1), \ldots. \underset{\sim}{X}(k_m), \underset{\sim}{Y}(k_1), \ldots, \underset{\sim}{Y}(k_m)$ is equal to the product of joint probability distribution functions of the m random n vectors and m random p vectors. That is,

$$P[\underset{\sim}{X}(k_1) \leq \underset{\sim}{x}_1, \ldots, \underset{\sim}{X}(k_m) \leq \underset{\sim}{x}_m, \underset{\sim}{Y}(k_1) \leq \underset{\sim}{y}_1, \ldots, \underset{\sim}{Y}(k_m) \leq \underset{\sim}{y}_m]$$

$$= P_1[\underset{\sim}{X}(k_1) \leq \underset{\sim}{x}_1, \ldots, \underset{\sim}{X}(k_m) \leq \underset{\sim}{x}_m]P_2[\underset{\sim}{Y}(k_1) \leq \underset{\sim}{y}_1, \ldots, \underset{\sim}{Y}(k_m) \leq \underset{\sim}{y}_m]$$

for all n vectors $\underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_m$ and all p vectors $\underset{\sim}{y}_1, \ldots, \underset{\sim}{y}_m$.

Definition 4: A stochastic process $\{\underset{\sim}{X}(k), k\epsilon I\}$ is said to be uncorrelated if, for all $j, k\epsilon I$, $j \neq k$

$$E[\underset{\sim}{X}(j)\underset{\sim}{X}'(k)] = E[\underset{\sim}{X}(j)]E[\underset{\sim}{X}'(k)]$$

Definition 5: Two stochastic processes $\{\underset{\sim}{X}(k), k\epsilon I\}$ and $\{\underset{\sim}{Y}(k), k\epsilon I\}$ are said to be uncorrelated if, for all $j, k\epsilon I$

$$E[\underset{\sim}{X}(j)\underset{\sim}{Y}'(k)] = E[\underset{\sim}{X}(j)]E[\underset{\sim}{Y}'(k)]$$

Definition 6: A stochastic process is said to be stationary if the probability laws governing the mechanism producing the process remain time-invariant as the process evolves in time.

Definition 7: A stochastic process $\{\underset{\sim}{X}(k), k\epsilon I\}$ is said to be gaussian or normal if, for any m time points $k_1, \ldots, k_m$ in I, where m is any integer, the set of m random n vectors $\underset{\sim}{X}(k_1), \ldots, \underset{\sim}{X}(k_m)$ is jointly gaussian distributed.

Definition 8: A stochastic process $\{\underset{\sim}{X}(k), k\epsilon I\}$ is said to be a gaussian white process if, for any m time points $k_1, \ldots, k_m$ in I, where m is any integer, the m random n vectors $\underset{\sim}{X}(k_1), \ldots, \underset{\sim}{X}(k_m)$ are independent gaussian random vectors.

Definition 9: A stochastic process $\{\underset{\sim}{X}(k), k\epsilon I\}$ is said to be a Markov process if, for any m time points $k_1 < k_2 < \ldots < k_m$ in I, where m is any integer, the condition probability distribution function of $\underset{\sim}{X}(k_m)$ for given values of $\underset{\sim}{X}(k_1), \ldots, \underset{\sim}{X}(k_{m-1})$ has the property that

$$P[\underset{\sim}{X}(k_m) \leq \underset{\sim}{x}_m | \underset{\sim}{X}(k_1) = \underset{\sim}{x}_1, \ldots, \underset{\sim}{X}(k_{m-1}) = \underset{\sim}{x}_{m-1}]$$

$$= P[\underset{\sim}{X}(k_m) \leq \underset{\sim}{x}_m | \underset{\sim}{X}(k_{m-1}) = \underset{\sim}{x}_{m-1}]$$

for all n vectors $\underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_m$.

Definition 10: A stochastic process is said to be Gauss-Markov if and only if it is both Gaussian and Markov.

The Model

The system model proposed by Mehra has been discussed many times in the system estimation and identification literature, e.g., Kalman (Ref 100), Ho and Lee (Ref 51), and Meditch (Ref 180). Essentially, the model represents a discrete linear dynamic system operating in a stochastic environment, and is formulated using the state variable approach. A block diagram representation of the model is shown in Fig. 4.

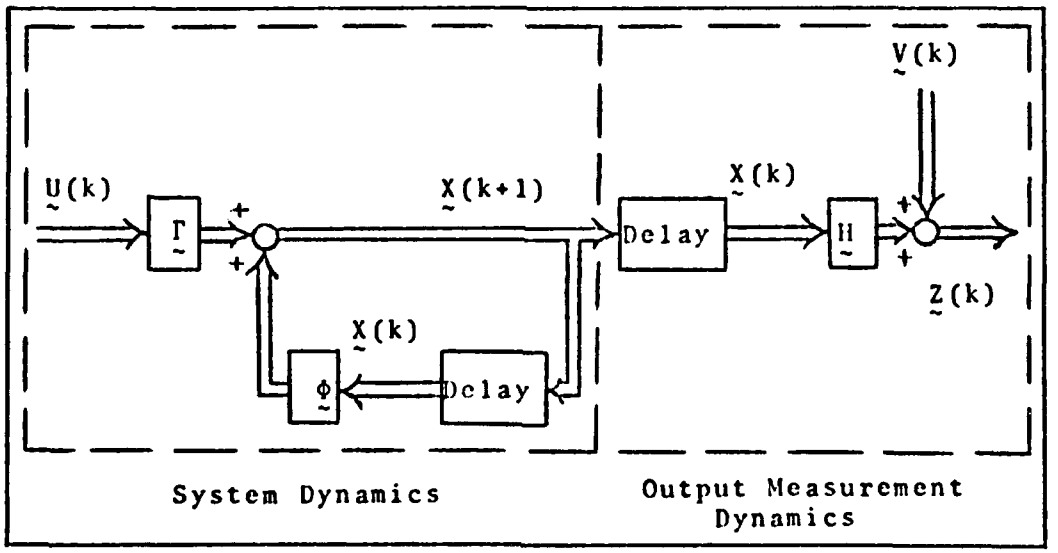


Fig. 4. Discrete Linear Stochastic Model.

The system dynamics and output measurement dynamics are governed, respectively, by the following two linear difference equations:

$$\underset{\sim}{X}(k+1) = \underset{\sim\sim}{\Phi}\underset{\sim}{X}(k) + \underset{\sim\sim}{\Gamma}\underset{\sim}{U}(k) \tag{1}$$

$$\underset{\sim}{Z}(k) = \underset{\sim\sim}{H}\underset{\sim}{X}(k) + \underset{\sim}{V}(k) \tag{2}$$

for k = 0, 1, ..., where

$\underset{\sim}{X}$ is the n x 1 state vector

$\underset{\sim}{\Phi}$ is the n x n state transition matrix (constant)

$\underset{\sim}{U}$ is a p x 1 vector of gaussian white-noise called the system disturbance vector

$\underset{\sim}{\Gamma}$ is an n x p matrix called the disturbance transition matrix (constant)

$\underset{\sim}{Z}$ is a r x 1 vector called the output measurement vector

$\underset{\sim}{H}$ is a r x n matrix called the measurement matrix (constant)

$\underset{\sim}{V}$ is a r x 1 vector of gaussian white-noise called the measurement error vector

Equation (1) is often called the state equation, and Eq (2) the measurement equation.

Assumptions

1. $$E[\underset{\sim}{U}(k)] = \underset{\sim}{0} \tag{3}$$

$$E[\underset{\sim}{U}(j)\underset{\sim}{U}'(k)] = \underset{\sim}{Q}\delta_{jk} \tag{4}$$

$$E[\underset{\sim}{V}(k)] = \underset{\sim}{0} \tag{5}$$

$$E[\underset{\sim}{V}(j)\underset{\sim}{V}'(k)] = \underset{\sim}{R}\delta_{jk} \tag{6}$$

$$E[\underset{\sim}{U}(j)\underset{\sim}{V}'(k)] = \underset{\sim}{0} \tag{7}$$

$$E[\underset{\sim}{X}(0)] = \underset{\sim}{0} \tag{8}$$

$$E[\underset{\sim}{X}(0)\underset{\sim}{X}'(0)] = \underset{\sim}{P}(0) \tag{9}$$

for j, k = 0, 1, ..., where

$\underset{\sim}{Q}$ is a p x p positive definite matrix called the input-noise covariance matrix (constant)

$\underset{\sim}{R}$ is a r x r positive definite matrix called the output-noise covariance matrix (constant)

$\underset{\sim}{X}(0)$ is a zero-mean guassian n vector called the initial state vector

$\underset{\sim}{P}(0)$ is an n x n matrix called the initial state covariance matrix

δ_{jk} is the Kronecker delta

2. The system is completely observable and controllable. Kalman (Ref 100) shows that this is equivalent to the following two conditions:

$$\text{rank } [\underset{\sim}{\Gamma}, \underset{\sim}{\Phi}\underset{\sim}{\Gamma}, \ldots, (\underset{\sim}{\Phi}^{n-1})\underset{\sim}{\Gamma}] = n \tag{10}$$

$$\text{rank } [\underset{\sim}{H}', \underset{\sim}{\Phi}'\underset{\sim}{H}', \ldots, (\underset{\sim}{\Phi}^{n-1})'\underset{\sim}{H}'] = n \tag{11}$$

This condition, along with the positive definiteness of $\underset{\sim}{Q}$ and $\underset{\sim}{R}$, is necessary to ensure the asymptotic global

stability of the Kalman filter, as shown by Deyst and Price (Ref 95).

3. The state transition matrix, $\underset{\sim}{\Phi}$, is nonsingular, with all eigenvalues located inside the unit circle. DeRusso, et. al. (Ref 177:447) show that this condition is necessary to ensure stable dynamics for a time-invariant linear system.

4. The system is minimum phase, i.e., the system transfer function has no zeros located in the right-half plane. Truxal (Ref 182:426-427) shows that this condition is necessary to ensure the physical realizability of a linear system.

5. The system is time-invariant, and the identification procedure begins after steady state conditions have been reached. This assumption is necessary to allow the system matrices $\underset{\sim}{\Phi}$, $\underset{\sim}{\Gamma}$, and $\underset{\sim}{H}$, the noise covariance matrices $\underset{\sim}{Q}$ and $\underset{\sim}{R}$, and the covariance matrix of the state vector $\underset{\sim}{X}(\cdot)$ to be treated as constants.

6. The problem of interest is to identify the unknown matrices $\underset{\sim}{\Phi}$, $\underset{\sim}{\Gamma}$, $\underset{\sim}{H}$, $\underset{\sim}{R}$, and $\underset{\sim}{Q}$ from a record of the output measurements $\underset{\sim}{Z}(0)$, $\underset{\sim}{Z}(1)$, ...

Properties

1. Meditch (Ref 180:168-169) shows that the model described by Eqs (1) through (9) has the following properties:

a. The stochastic processes $\{\underset{\sim}{X}(k), k = 0, 1, \ldots\}$ and $\{\underset{\sim}{Z}(i), i = 0, 1, \ldots, j\}$ are Gauss-Markov sequences with identically zero means.

b. $$E[\underset{\sim}{X}(j)\underset{\sim}{U}'(k)] = \underset{\sim}{0} \quad \forall\ k \geq j,\ j = 0, 1, \ldots \tag{12}$$

c. $$E[\underset{\sim}{Z}(j)\underset{\sim}{U}'(k)] = \underset{\sim}{0} \quad \forall\ k \geq j,\ j = 0, 1, \ldots \tag{13}$$

d. $$E[\underset{\sim}{X}(j)\underset{\sim}{V}'(k)] = \underset{\sim}{0} \quad \forall\ j \text{ and } k \tag{14}$$

e. $$E[\underset{\sim}{Z}(j)\underset{\sim}{V}'(k)] = \underset{\sim}{0} \quad \forall\ k > j,\ j = 0, 1, \ldots \tag{15}$$

These properties follow from the linearity inherent in the model, and the fact that $\underset{\sim}{X}(k)$ depends only upon $\underset{\sim}{X}(0)$, $\underset{\sim}{U}(0), \ldots, \underset{\sim}{U}(k-1)$, where $\{\underset{\sim}{U}(j), j = 0, 1, \ldots\}$ and $\{\underset{\sim}{V}(j), j = 0, 1, \ldots\}$ are uncorrelated zero-mean white gaussian sequences independent of the gaussian random vector $\underset{\sim}{X}(0)$.

2. At this point, it is convenient to state some general properties from optimal estimation theory (Refs 100 and 180).

a. The optimal estimate of the state $\underset{\sim}{X}(k)$ given the output measurements $\underset{\sim}{Z}(s)$, $0 \leq s \leq j$, is given by the conditional mean of $\underset{\sim}{X}(k)$, denoted as follows:

$$\hat{\underset{\sim}{X}}(k|j) = E[\underset{\sim}{X}(k)|\underset{\sim}{Z}(s),\ 0 \leq s \leq j] \qquad (16)$$

If $k > j$, the estimation problem is one of prediction; if $k = j$, the estimation problem is one of filtering; and if $k < j$, the problem is one of smoothing. The filtered and predicted state estimates are used in the adaptive Kalman filtering approach to the identification problem discussed in Chapter V.

b. Let the estimation error of the state $\underset{\sim}{X}(k)$ be defined by

$$\tilde{\underset{\sim}{X}}(k|j) \equiv \underset{\sim}{X}(k) - \hat{\underset{\sim}{X}}(k|j) \qquad (17)$$

Then the single-stage optimal predicted estimate of $\underset{\sim}{X}(k+1)$ is given by the following relation:

$$\hat{\underset{\sim}{X}}(k+1|k) = \underset{\sim\sim}{\Phi}\hat{\underset{\sim}{X}}(k|k) \qquad (18)$$

where $\hat{\underset{\sim}{X}}(k|k)$ is the optimal filtered estimate of $\underset{\sim}{X}(k)$.

c. $\hat{\underset{\sim}{X}}(k|j)$ is a linear estimate, i.e., a linear combination of the available output measurements $\underset{\sim}{Z}(s)$, $0 \leq s \leq j$.

d. $\hat{\underset{\sim}{X}}(k|j)$ is unique.

e. $\hat{\underset{\sim}{X}}(k|j)$ and $\tilde{\underset{\sim}{X}}(k|j)$ are gaussian random n vectors.

f. The stochastic process $\{\tilde{\underset{\sim}{X}}(k+1|k),\ k = 0, 1, \ldots\}$ is a zero-mean Gauss-Markov sequence.

g. $\tilde{\underset{\sim}{X}}(k|j)$ is independent of <u>any</u> linear combination of the available output measurements. In particular, $\tilde{\underset{\sim}{X}}(k|j)$ is independent of $\hat{\underset{\sim}{X}}(s|t)$, which implies that

$$E[\tilde{\underset{\sim}{X}}(k|j)\hat{\underset{\sim}{X}}'(s|t)] = 0 \qquad (19)$$

for all k, j, s, and t.

h. If the r x 1 vector $\underset{\sim}{Y}(k)$ is defined by

$$\underset{\sim}{Y}(k) \equiv \underset{\sim\sim}{H}\underset{\sim}{X}(k)$$

so that $\tilde{\underset{\sim}{Y}}(k|k-1) = \underset{\sim\sim}{H}\tilde{\underset{\sim}{X}}(k|k-1)$, then $\underset{\sim}{Z}(j)$ and $\tilde{\underset{\sim}{Y}}(k|k-1)$ are independent random vectors, implying that

$$E[\underset{\sim}{Z}(j)\tilde{\underset{\sim}{Y}}'(k|k-1)] = \underset{\sim}{0} \qquad (20)$$

for all k and j.

Identification of $\underset{\sim}{\Phi}$

The identification method proposed by Mehra (Ref 78) to estimate the elements of the state transition matrix $\underset{\sim}{\Phi}$ is based on the autocorrelation function of the output measurements $\underset{\sim}{Z}(i)$. In order to formally derive the method,

however, it is first necessary to establish that under steady state conditions the stochastic process $\{\underset{\sim}{Z}(i),\ i = 0, 1, \ldots\}$ is a stationary gaussian sequence. Using Eqs (2), (6), and (14), the covariance matrix of the output measurements may be written as follows for $k > 0$:

$$\begin{aligned} E[\underset{\sim}{Z}(i)\underset{\sim}{Z}'(i-k)] &= E\{[\underset{\sim}{\underset{\sim}{H}}\underset{\sim}{X}(i)+\underset{\sim}{V}(i)][\underset{\sim}{\underset{\sim}{H}}\underset{\sim}{X}(i-k)+\underset{\sim}{V}(i-k)]'\} \\ &= \underset{\sim}{\underset{\sim}{H}}E[\underset{\sim}{X}(i)\underset{\sim}{X}'(i-k)]\underset{\sim}{\underset{\sim}{H}}'+\underset{\sim}{\underset{\sim}{H}}E[\underset{\sim}{X}(i)\underset{\sim}{V}'(i-k)] \\ &+ E[\underset{\sim}{V}(i)\underset{\sim}{X}'(i-k)]\underset{\sim}{\underset{\sim}{H}}'+E[\underset{\sim}{V}(i)\underset{\sim}{V}'(i-k)] \\ &= \underset{\sim}{\underset{\sim}{H}}E[\underset{\sim}{X}(i)\underset{\sim}{X}'(i-k)]\underset{\sim}{\underset{\sim}{H}}' \end{aligned} \tag{21}$$

Rewriting Eq (1) in terms of the initial state vector $\underset{\sim}{X}(0)$ results in

$$\underset{\sim}{X}(i) = \underset{\sim}{\Phi}^{i}\underset{\sim}{X}(0) + \sum_{s=0}^{i-1} \underset{\sim}{\Phi}^{i-1-s}\ \underset{\sim}{\underset{\sim}{\Gamma}}\underset{\sim}{U}(s) \tag{22}$$

$$\underset{\sim}{X}(i-k) = \underset{\sim}{\Phi}^{i-k}\ \underset{\sim}{X}(0) + \sum_{s=0}^{i-k-1} \underset{\sim}{\Phi}^{i-k-1-s}\ \underset{\sim}{\underset{\sim}{\Gamma}}\underset{\sim}{U}(s) \tag{23}$$

Solving Eq (23) for $\underset{\sim}{X}(0)$ and substituting the result into (22) yields

$$\begin{aligned} \underset{\sim}{X}(i) &= \underset{\sim}{\Phi}^{i}[\underset{\sim}{\Phi}^{k-i}\underset{\sim}{X}(i-k) - \sum_{s=0}^{i-k-1} \underset{\sim}{\Phi}^{-1-s}\ \underset{\sim}{\underset{\sim}{\Gamma}}\underset{\sim}{U}(s)] \\ &\quad + \sum_{s=0}^{i-1} \underset{\sim}{\Phi}^{i-1-s}\ \underset{\sim}{\underset{\sim}{\Gamma}}\underset{\sim}{U}(s) \\ &= \underset{\sim}{\Phi}^{k}\underset{\sim}{X}(i-k) + \sum_{s=i-k}^{i-1} \underset{\sim}{\Phi}^{i-1-s}\ \underset{\sim}{\underset{\sim}{\Gamma}}\underset{\sim}{U}(s) \end{aligned} \tag{24}$$

Substituting Eq (24) into (21), and using Eq (12) yields

$$E[\underset{\sim}{Z}(i)\underset{\sim}{Z}'(i-k)] = \underset{\sim}{H}E\{[\underset{\sim}{\Phi}^{k}\underset{\sim}{X}(i-k) + \sum_{s=i-k}^{i-1} \underset{\sim}{\Phi}^{i-1-s}\,\underset{\sim}{\Gamma}\underset{\sim}{U}(s)] \cdot$$

$$\cdot\,[\underset{\sim}{X}'(i-k)]\}\underset{\sim}{H}'$$

$$= \underset{\sim}{H}\underset{\sim}{\Phi}^{k}E[\underset{\sim}{X}(i-k)\underset{\sim}{X}'(i-k)]\underset{\sim}{H}'$$

$$+ \underset{\sim}{H}\sum_{s=i-k}^{i-1} \underset{\sim}{\Phi}^{i-1-s}\,\underset{\sim}{\Gamma}E[\underset{\sim}{U}(s)\underset{\sim}{X}'(i-k)]\underset{\sim}{H}'$$

$$= \underset{\sim}{H}\underset{\sim}{\Phi}^{k}E[\underset{\sim}{X}(i-k)\underset{\sim}{X}'(i-k)]\underset{\sim}{H}' \qquad (25)$$

If the n x n covariance matrix of the state vector $\underset{\sim}{X}(i)$ is defined by

$$\underset{\sim}{P}(i) \equiv E[\underset{\sim}{X}(i)\underset{\sim}{X}'(i)] \qquad (26)$$

then the following expression may be obtained for Eq (25):

$$E[\underset{\sim}{Z}(i)\underset{\sim}{Z}'(i-k)] = \underset{\sim}{H}\underset{\sim}{\Phi}^{k}\underset{\sim}{P}(i-k)\underset{\sim}{H}' \qquad (27)$$

By assumption five of the model, however, steady state conditions have been reached. This implies that the covariance matrix of the state vector $\underset{\sim}{X}(\cdot)$ is constant and may be written as

$$\underset{\sim}{P} = \underset{\sim}{P}(\cdot) \qquad (28)$$

Substituting Eq (28) into (27) results in the following relation for the covariance matrix of the output measurements:

$$E[\underset{\sim}{Z}(i)\underset{\sim}{Z}'(i-k)] = \underset{\sim\sim}{H}\underset{\sim\sim}{\Phi}^{k}\underset{\sim\sim}{P}\underset{\sim\sim}{H}', \quad k > 0 \tag{29}$$

A side condition on $\underset{\sim}{P}$ may be obtained by substituting Eq (1) into (26), and using Eqs (4), (12), and (28):

$$\begin{aligned}\underset{\sim}{P}(i) &= E\{[\underset{\sim\sim}{\Phi}\underset{\sim}{X}(i-1)+\underset{\sim\sim}{\Gamma}\underset{\sim}{U}(i-1)][\underset{\sim\sim}{\Phi}\underset{\sim}{X}(i-1)+\underset{\sim\sim}{\Gamma}\underset{\sim}{U}(i-1)]'\} \\ &= \underset{\sim\sim}{\Phi}E[\underset{\sim}{X}(i-1)\underset{\sim}{X}'(i-1)]\underset{\sim\sim}{\Phi}'+\underset{\sim\sim\sim}{\Gamma Q \Gamma}' \\ &= \underset{\sim\sim}{\Phi}\underset{\sim\sim}{P}(i-1)\underset{\sim\sim}{\Phi}'+\underset{\sim\sim\sim}{\Gamma Q \Gamma}'\end{aligned}$$

or

$$\underset{\sim}{P} = \underset{\sim\sim\sim}{\Phi P \Phi}'+\underset{\sim\sim\sim}{\Gamma Q \Gamma}' \tag{30}$$

Similarly, using Eqs (2), (6), (14), and (28), the autocorrelation function of $\underset{\sim}{Z}(i)$ may be written as follows for the case of k = 0:

$$\begin{aligned}E[\underset{\sim}{Z}(i)\underset{\sim}{Z}'(i)] &= E\{[\underset{\sim\sim}{H}\underset{\sim}{X}(i)+\underset{\sim}{V}(i)][\underset{\sim\sim}{H}\underset{\sim}{X}(i)+\underset{\sim}{V}(i)]'\} \\ &= \underset{\sim}{H}E[\underset{\sim}{X}(i)\underset{\sim}{X}'(i)]\underset{\sim}{H}'+\underset{\sim}{R} \\ &= \underset{\sim\sim\sim}{HPH}'+\underset{\sim}{R}, \quad k = 0\end{aligned} \tag{31}$$

Therefore, it can be seen from Eqs (29) and (31) that under steady state conditions, $E[\underset{\sim}{Z}(i)\underset{\sim}{Z}'(i-k)]$ is independent of i for all k. This implies that $\{\underset{\sim}{Z}(i), i = 0, 1, \ldots\}$ is a stationary sequence, which is also gaussian by property one of the model.

It should be noted that in general, the autocorrelation function of the output measurements is not a scalar but

rather a r x r matrix. However, if only scalar outputs are considered (i.e., r = 1), the corresponding autocorrelation function is also a scalar. In the remainder of this paper, therefore, only the scalar-output case will be considered, and the scalar autocorrelation function of the output measurements is given by

$$C_k = \begin{cases} \underset{\sim}{H}\underset{\sim}{P}\underset{\sim}{H}' + R, & k = 0 \\ \underset{\sim}{H}\underset{\sim}{\Phi}^k\underset{\sim}{P}\underset{\sim}{H}', & k > 0 \end{cases} \tag{32}$$

Having established the stationarity of the above vector-input scalar-output system, attention is returned to a consideration of Mehra's proposed on-line identification method (Ref 78). The autocorrelation function of the output measurements, C_k, is used to derive a set of equations analogous to the Yule-Walker equations in the statistical analysis of purely autoregressive time series (Ref 183). The derived set of equations are then shown to form the basis for the proposed on-line identification method. Using Eq (32) for k = j, j + 1, ..., j + n - 1, where n is the order of the system and $j \geq 1$, results in

$$\begin{bmatrix} C_j \\ C_{j+1} \\ \vdots \\ C_{j+n-1} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{H}\underset{\sim}{\Phi}^j\underset{\sim}{P}\underset{\sim}{H}' \\ \underset{\sim}{H}\underset{\sim}{\Phi}^{j+1}\underset{\sim}{P}\underset{\sim}{H}' \\ \vdots \\ \underset{\sim}{H}\underset{\sim}{\Phi}^{j+n-1}\underset{\sim}{P}\underset{\sim}{H}' \end{bmatrix}$$

$$= \underset{\sim}{B}_j\underset{\sim}{P}\underset{\sim}{H}' \tag{33}$$

where

$$\underset{\sim}{B}_j = \begin{bmatrix} \underset{\sim}{H}\underset{\sim}{\Phi}^j \\ \underset{\sim}{H}\underset{\sim}{\Phi}^{j+1} \\ \vdots \\ \underset{\sim}{H}\underset{\sim}{\Phi}^{j+n-1} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{H} \\ \underset{\sim}{H}\underset{\sim}{\Phi} \\ \vdots \\ \underset{\sim}{H}\underset{\sim}{\Phi}^{n-1} \end{bmatrix} \underset{\sim}{\Phi}^j \qquad (34)$$

It should be observed that the n x n matrix $\underset{\sim}{B}_j$ is nonsingular This follows from assumptions three and two of the model which state, respectively, that the state transition matrix $\underset{\sim}{\Phi}$ is nonsingular and the n x n observability matrix is of full rank, i.e.,

$$\text{rank} \begin{bmatrix} \underset{\sim}{H} \\ \underset{\sim}{H}\underset{\sim}{\Phi} \\ \vdots \\ \underset{\sim}{H}\underset{\sim}{\Phi}^{n-1} \end{bmatrix} = n$$

Therefore, since $\underset{\sim}{B}_j$ from Eq (34) is seen to be the product of the observability matrix having rank n and the nonsingular matrix $\underset{\sim}{\Phi}^j$, it follows that the rank of $\underset{\sim}{B}_j$ is also n. Thus the rows of $\underset{\sim}{B}_j$ are linearly independent, and the nonsingularity of $\underset{\sim}{B}_j$ is established. Solving for $\underset{\sim}{P}\underset{\sim}{H}'$ from Eq (33) yields

$$\underset{\sim}{PH}' = \underset{\sim}{B}_j^{-1} \begin{bmatrix} C_j \\ C_{j+1} \\ \cdot \\ \cdot \\ \cdot \\ C_{j+n-1} \end{bmatrix} \tag{35}$$

Substituting Eq (35) into (32) for k = j + n results in

$$C_{j+n} = \underset{\sim}{H}\underset{\sim}{\Phi}^{j+n}\underset{\sim}{B}_j^{-1} \begin{bmatrix} C_j \\ C_{j+1} \\ \cdot \\ \cdot \\ \cdot \\ C_{j+n-1} \end{bmatrix} \tag{36}$$

Equation (36) represents a recursive relationship for the autocorrelation function C_k. In its present form, Eq (36) cannot be used in the identification of $\underset{\sim}{\Phi}$ because of the presence of the unknown matrix $\underset{\sim}{H}$. However, $\underset{\sim}{H}$ can be eliminated by use of the Caley-Hamilton Theorem (Ref 177), which states that the square matrix $\underset{\sim}{\Phi}$ satisfies its own characteristic equation. If the characteristic polynomial of $\underset{\sim}{\Phi}$ is denoted

$$f(\lambda) = |\lambda\underset{\sim}{I} - \underset{\sim}{\Phi}| = \lambda^n + a_n\lambda^{n-1} + \ldots + a_2\lambda + a_1$$

where λ is a root of the characteristic equation of $\underset{\sim}{\Phi}$ and the a_i, i=1, ..., n are unknown scalars, then this implies that

$$\underset{\sim}{\Phi}^n + a_n\underset{\sim}{\Phi}^{n-1} + \ldots + a_2\underset{\sim}{\Phi} + a_1\underset{\sim}{\Phi} = \underset{\sim}{0} \tag{37}$$

Premultiplying Eq (37) by $\underset{\sim}{H\Phi}^{j}$ and solving the resulting equation for $\underset{\sim}{H\Phi}^{j+n}$ yields

$$\underset{\sim}{H\Phi}^{j+n} = -a_n \underset{\sim}{H\Phi}^{j+n-1} - \ldots - a_2 \underset{\sim}{H\Phi}^{j+1} - a_1 \underset{\sim}{H\Phi}^{j}$$

$$= -[a_1, a_2, \ldots, a_n] \begin{bmatrix} \underset{\sim}{H\Phi}^{j} \\ \underset{\sim}{H\Phi}^{j+1} \\ \vdots \\ \underset{\sim}{H\Phi}^{j+n-1} \end{bmatrix}$$

$$= -[a_1, a_2, \ldots, a_n] \underset{\sim}{B}_j \tag{38}$$

Postmultiplying Eq (38) by $\underset{\sim}{B}_j^{-1}$ results in

$$\underset{\sim}{H\Phi}^{j+n} \underset{\sim}{B}_j^{-1} = -[a_1, a_2, \ldots, a_n] \tag{39}$$

Substituting Eq (39) into (36), the following result is obtained:

$$C_{j+n} = -[a_1, a_2, \ldots, a_n] \begin{bmatrix} C_j \\ C_{j+1} \\ \vdots \\ C_{j+n-1} \end{bmatrix}$$

$$C_{j+n} = -\sum_{i=1}^{n} a_i C_{i+j-1}, \quad j \geq 1 \tag{40}$$

Equation (40) represents a modified set of Yule-Walker equations (Ref 183). These equations may be written in

matrix notation by allowing j to run from 1 to n in Eq (40):

$$\begin{bmatrix} C_{n+1} \\ C_{n+2} \\ \vdots \\ C_{2n} \end{bmatrix} = - \begin{bmatrix} C_1 & C_2 & \cdots & C_n \\ C_2 & C_3 & \cdots & C_{n+1} \\ & & \vdots & \\ C_n & C_{n+1} & \cdots & C_{2n-1} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \tag{41}$$

The development leading to the result in Eq (41) is similar to the approach used by Gersch in the analysis of a mixed autoregressive moving-average time series (Ref 75). In his work, Gersch derives an unbiased estimator for the autoregressive time series parameters. He refers to Eq (41) as the "normal equation" of the system, and shows that for j = 1, ..., n, the n x n matrix of autocorrelations appearing therein is nonsingular. Hence, Eq (41) may be rewritten as

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = - \begin{bmatrix} C_1 & C_2 & \cdots & C_n \\ C_2 & C_3 & \cdots & C_{n+1} \\ & & \vdots & \\ C_n & C_{n+1} & \cdots & C_{2n-1} \end{bmatrix}^{-1} \begin{bmatrix} C_{n+1} \\ C_{n+2} \\ \vdots \\ C_{2n} \end{bmatrix} \tag{42}$$

In order to use Eq (42) to estimate the a_i, and hence $\underset{\sim}{\Phi}$, it is necessary to first estimate the autocorrelation functions C_k. If $\hat{C}_k$ is used to denote an estimate of C_k, one possibility is to make use of the fact that the Z(i) are a

stationary gaussian sequence, and substitute the empirical autocorrelation functions of the output measurements for $\hat{C}_k$:

$$\hat{C}_k = 1/T \sum_{i=k}^{T} Z(i)Z(i-k) \tag{43}$$

where T is the number of sample points of the output. For finite sample sizes, the estimates of $\hat{C}_k$ obtained using Eq (43) are biased, requiring that T in the denominator be replaced by T-k for unbiased finite-sample estimates. However, Heffes (Ref 97) shows that the estimates obtained using Eq (43) result in less mean-square error than the corresponding unbiased estimates, and hence are preferable. Therefore, substituting the values of $\hat{C}_k$ given by Eq (43) into Eq (42), and letting $\hat{a}_i$ denote an estimate of a_i, the following result is obtained:

$$\begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \cdot \\ \cdot \\ \cdot \\ \hat{a}_n \end{bmatrix} = - \begin{bmatrix} \hat{C}_1 & \hat{C}_2 & \cdots & \hat{C}_n \\ \hat{C}_2 & \hat{C}_3 & \cdots & \hat{C}_{n+1} \\ & & \cdot & \\ & & \cdot & \\ & & \cdot & \\ \hat{C}_n & \hat{C}_{n+1} & \cdots & \hat{C}_{2n-1} \end{bmatrix}^{-1} \begin{bmatrix} \hat{C}_{n+1} \\ \hat{C}_{n+2} \\ \cdot \\ \cdot \\ \cdot \\ \hat{C}_{2n} \end{bmatrix} \tag{44}$$

Mehra shows that the identification method represented by Eq (44) has the following properties:

1. The $\hat{a}_i$ are asymptotically unbiased, normal, and consistent estimates of the a_i. This follows from the fact that for large T, the errors in estimating the a_i are approximately linearly related to the errors in estimating the C_k.

Parzen (Ref 80) shows that $\hat{C}_k$ is an asymptotically unbiased, normal, and consistent estimator of C_k. Since the errors in estimating the C_k are asymptotically normal with zero mean, the stated property is established.

2. The identification procedure is capable of being implemented on-line, i.e., the $\hat{a}_i$ can be calculated resursively. If $\hat{C}_k^{T+1}$ is used to denote the estimate of C_k based on T + 1 sample points, it follows that

$$\hat{C}_k^{T+1} = \frac{1}{T+1} \sum_{i=k}^{T+1} Z(i)Z(i-k)$$

$$= \frac{1}{T+1} [Z(T+1)Z(T+1-k)] + \frac{T}{T+1} \hat{C}_k^T \qquad (45)$$

The $\hat{a}_i$ can then be calculated recursively using Eq (44) for T + 1 observations:

$$\begin{bmatrix} \hat{a}_1^{T+1} \\ \hat{a}_2^{T+1} \\ \vdots \\ \hat{a}_n^{T+1} \end{bmatrix} = - \begin{bmatrix} \hat{C}_1^{T+1} & \hat{C}_2^{T+1} & \cdots & \hat{C}_n^{T+1} \\ \hat{C}_2^{T+1} & \hat{C}_3^{T+1} & \cdots & \hat{C}_{n+1}^{T+1} \\ \vdots & \vdots & & \\ \hat{C}_n^{T+1} & \hat{C}_{n+1}^{T+1} & \cdots & \hat{C}_{2n-1}^{T+1} \end{bmatrix}^{-1} \begin{bmatrix} \hat{C}_{n+1}^{T+1} \\ \hat{C}_{n+2}^{T+1} \\ \vdots \\ \hat{C}_{2n}^{T+1} \end{bmatrix} \qquad (46)$$

3. The identification method, while identifying the coefficients a_i in the characteristic polynomial of $\underset{\sim}{\Phi}$, does not actually identify the elements of the $\underset{\sim}{\Phi}$ matrix itself. This can be a serious disadvantage if the exact structure

of $\underset{\sim}{\Phi}$ is required. The reason for this is that once the a_i are known, the Cayley-Hamilton theorem, viz., Eq (37), can theoretically be used to solve for the actual elements of the $\underset{\sim}{\Phi}$ matrix. However, the resulting system of n algebraic equations in n unknowns is nonlinear, with the result that non-unique solutions will exist. In fact, Birkhoff and MacLane (Ref 176) show that corresponding to each nth degree monic polynomial

$$g(\lambda) = \lambda^n + a_n\lambda^{n-1} + \ldots + a_2\lambda + a_1$$

an n x n matrix having characteristic polynomial $g(\lambda)$ can be constructed. This matrix is called the companion matrix of $g(\lambda)$, and has the following form:

$$\begin{bmatrix} 0 & 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 1 & \ldots & 0 \\ & & & & \vdots & \\ 0 & 0 & 0 & 0 & \ldots & 1 \\ -a_1 & -a_2 & -a_3 & -a_4 & \ldots & -a_n \end{bmatrix}$$

However, matrices other than the companion matrix can be found which also satisfy the characteristic equation of $\underset{\sim}{\Phi}$.

Identification of $\underset{\sim}{\Phi}$, $\underset{\sim}{\Gamma}$, and $\underset{\sim}{H}$ Using a Canonical Form System Representation

The possible disadvantages associated with use of the above method for identifying the state transition matrix $\underset{\sim}{\Phi}$

may be eliminated if one is only interested in modeling the output observations of the system. That is, if the particular structure of $\underset{\sim}{\Phi}$ is unimportant, then the system represented by Eqs (1) and (2) can be transformed into a canonical or phase variable form (Ref 90) containing a fewer number of parameters to be identified. Consider the following linear transformation:

$$\underset{\sim}{X}^*(k) = \underset{\sim}{T}\underset{\sim}{X}(k) \tag{47}$$

where $\underset{\sim}{T}$ is a constant n x n matrix that maps the n x 1 state vector X(k) into the n x 1 vector $\underset{\sim}{X}^*(\kappa)$. Mehra suggests the following definition for $\underset{\sim}{T}$:

$$\underset{\sim}{T} \equiv \begin{bmatrix} \underset{\sim}{H} \\ \underset{\sim}{H\Phi} \\ \vdots \\ \underset{\sim}{H\Phi}^{n-1} \end{bmatrix} \tag{48}$$

Using the assumption of complete observability, viz., Eq (11), it follows that $\underset{\sim}{T}$ is also nonsingular. Premultiplying Eq (1) by $\underset{\sim}{T}$ and using (47), Eqs (1) and (2) become, respectively

$$\underset{\sim}{X}^*(k+1) = \underset{\sim}{\Phi}^*\underset{\sim}{X}^*(k) + \underset{\sim}{\Gamma}^*\underset{\sim}{U}(k) \tag{49}$$

$$Z(k) = \underset{\sim}{H}^*\underset{\sim}{X}^*(k) + V(k) \tag{50}$$

where

$$\underset{\sim}{\Phi}^* = \underset{\sim\sim}{T}\underset{\sim}{\Phi}\underset{\sim\sim}{T}^{-1} \tag{51}$$

$$\underset{\sim}{\Gamma}^* = \underset{\sim\sim}{T}\underset{\sim}{\Gamma} \tag{52}$$

$$\underset{\sim}{H}^* = \underset{\sim\sim}{H}\underset{\sim\sim}{T}^{-1} \tag{53}$$

Equations (49) and (50) represent the original system in canonical or phase variable form, with $\underset{\sim}{\Phi}^*$, $\underset{\sim}{\Gamma}^*$, and $\underset{\sim}{H}^*$ derived as follows. Postmultiplying Eq (48) by $\underset{\sim}{\Phi}$, and using Eq (34) with j = 1 results in

$$\underset{\sim\sim}{T}\underset{\sim}{\Phi} = \begin{bmatrix} \underset{\sim\sim}{H}\underset{\sim}{\Phi} \\ \underset{\sim\sim}{H}\underset{\sim}{\Phi}^2 \\ \vdots \\ \underset{\sim\sim}{H}\underset{\sim}{\Phi}^n \end{bmatrix} = \underset{\sim}{B}_1 \tag{54}$$

from which

$$\underset{\sim}{T}^{-1} = \underset{\sim\sim}{\Phi}\underset{\sim}{B}_1^{-1} \tag{55}$$

Substituting Eq (55) into (51) yields

$$\underset{\sim}{\Phi}^* = \underset{\sim\sim}{T}\underset{\sim}{\Phi}(\underset{\sim\sim}{\Phi}\underset{\sim}{B}_1^{-1})$$

$$= \underset{\sim\sim}{T}\underset{\sim}{\Phi}^2\underset{\sim}{B}_1^{-1} \tag{56}$$

An expression for $\underset{\sim}{\Phi}^*$ that is independent of $\underset{\sim}{T}$ may be obtained by substituting Eq (54) into (56), and using Eq (34) with $j = 1$:

$$\underset{\sim}{\Phi}^* = (\underset{\sim\sim}{T\Phi})\underset{\sim\sim}{\Phi B}_1^{-1}$$

$$= \underset{\sim}{B}_1\underset{\sim\sim}{\Phi B}_1^{-1}$$

$$= \begin{bmatrix} \underset{\sim\sim}{H\Phi} \\ \vdots \\ \underset{\sim\sim}{H\Phi}^{n-1} \\ \underset{\sim\sim}{H\Phi}^{n} \end{bmatrix} \underset{\sim\sim}{\Phi B}_1^{-1}$$

$$= \begin{bmatrix} \underset{\sim\sim}{H\Phi}^2\underset{\sim}{B}_1^{-1} \\ \vdots \\ \underset{\sim\sim}{H\Phi}^{n}\underset{\sim}{B}_1^{-1} \\ \underset{\sim\sim}{H\Phi}^{n+1}\underset{\sim}{B}_1^{-1} \end{bmatrix} \tag{57}$$

But, from Eq (54) it is clear that

$$\underset{\sim\sim\sim}{T\Phi B}_1^{-1} = \begin{bmatrix} \underset{\sim\sim\sim}{H\Phi B}_1^{-1} \\ \underset{\sim\sim}{H\Phi}^2\underset{\sim}{B}_1^{-1} \\ \vdots \\ \underset{\sim\sim}{H\Phi}^{n}\underset{\sim}{B}_1^{-1} \end{bmatrix} = \underset{\sim}{I} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ & & & \vdots & \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \tag{58}$$

Therefore, using Eq (58) to solve for the first n - 1 rows of $\underset{\sim}{\Phi}^*$ in the representation given by Eq (57), and Eq (39) with j = 1 for the last row of $\underset{\sim}{\Phi}^*$, it follows that Eq (57) can be rewritten as

$$\underset{\sim}{\Phi}^* = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ & & & \vdots & \\ 0 & 0 & 0 & \cdots & 1 \\ -a_1 & -a_2 & -a_3 & \cdots & -a_n \end{bmatrix} \tag{59}$$

It should be noted in passing that $\underset{\sim}{\Phi}^*$, the canonical or phase variable representation of the state transition matrix $\underset{\sim}{\Phi}$, is precisely the companion matrix of the characteristic polynomial of $\underset{\sim}{\Phi}$ that was discussed in the previous section. Furthermore, other representations of $\underset{\sim}{\Phi}$ exist, since the transformation given by Eq (51) identifies $\underset{\sim}{\Phi}$ only to within a similarity transformation (Ref 176).

Continuing, an expression for $\underset{\sim}{H}^*$ can be obtained by substituting Eq (55) into (53), yielding

$$\underset{\sim}{H}^* = \underset{\sim\sim\sim}{H\Phi B_1^{-1}}$$

From Eq (58), however, $\underset{\sim\sim\sim}{H\Phi B_1^{-1}}$ is seen to be precisely the first row of the n x n identity matrix. Thus, $\underset{\sim}{H}^*$ may be written as

$$H^* = [1, 0, 0, \ldots, 0] \tag{60}$$

Finally, it remains to determine $\underset{\sim}{\Gamma}^*$. Using the canonical form system representation given by Eqs (49) and (50), and denoting the covariance matrix of $\underset{\sim}{X}^*(\cdot)$ by $\underset{\sim}{P}^*$, it is easily shown that the following equation, analogous to Eq (30), can be derived:

$$\underset{\sim}{P}^* = \underset{\sim}{\Phi}^* \underset{\sim}{P}^* \underset{\sim}{\Phi}^{*\prime} + \underset{\sim}{\Gamma}^* \underset{\sim}{Q} \underset{\sim}{\Gamma}^{*\prime} \tag{61}$$

Assuming for the moment that the input noise covariance matrix $\underset{\sim}{Q}$ is known, Eq (61) still cannot be used to solve for $\underset{\sim}{\Gamma}^*$ because of the unknown matrix $\underset{\sim}{P}^*$. However, an auxilary equation for $\underset{\sim}{P}^*$ may be obtained as follows. Once again starting with Eqs (49) and (50), it is straightforward to show that the following equation, analogous to Eq (34) with j = 1, can be obtained:

$$\underset{\sim}{B}_1^* = \begin{bmatrix} \underset{\sim}{H}^* \underset{\sim}{\Phi}^* \\ \underset{\sim}{H}^* (\underset{\sim}{\Phi}^*)^2 \\ \vdots \\ \underset{\sim}{H}^* (\underset{\sim}{\Phi}^*)^n \end{bmatrix} \tag{62}$$

Equations (54) and (56), however, imply that

$$(\underset{\sim}{\Phi}^*)^2 = \underset{\sim\sim}{T} \underset{\sim\sim}{\Phi}^2 \underset{\sim}{B}_1^{-1} (\underset{\sim\sim}{T} \underset{\sim\sim}{\Phi}) \underset{\sim}{\Phi} \underset{\sim}{B}_1^{-1}$$

$$= \underset{\sim\sim}{T} \underset{\sim\sim}{\Phi}^3 \underset{\sim}{B}_1^{-1}$$

and by induction

$$(\underset{\sim}{\Phi}^*)^i = \underset{\sim\sim}{T}\underset{\sim\sim}{\Phi}^{i+1}\underset{\sim}{B}_1^{-1} \tag{63}$$

Substituting Eqs (53) and (63) with i = 1, ..., n into Eq (62), and using Eq (57) results in

$$\underset{\sim}{B}_1^* = \begin{bmatrix} \underset{\sim\sim}{H}\underset{\sim\sim}{\Phi}^2\underset{\sim}{B}_1^{-1} \\ \underset{\sim\sim}{H}\underset{\sim\sim}{\Phi}^3\underset{\sim}{B}_1^{-1} \\ \vdots \\ \underset{\sim\sim}{H}\underset{\sim\sim}{\Phi}^{n+1}\underset{\sim}{B}_1^{-1} \end{bmatrix} = \underset{\sim}{\Phi}^* \tag{64}$$

Now, using Eqs (49) and (50) to derive an equation analogous to Eq (35) with j = 1, i.e.,

$$\underset{\sim}{P}^*\underset{\sim}{H}^{*\prime} = (\underset{\sim}{B}_1^*)^{-1} \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_n \end{bmatrix}$$

and substituting for $\underset{\sim}{B}_1^*$ from Eq (64), the following auxiliary equation for $\underset{\sim}{P}^*$ is obtained:

$$\underset{\sim}{P}^*\underset{\sim}{H}^{*\prime} = (\underset{\sim}{\Phi}^*)^{-1} \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_n \end{bmatrix} \tag{65}$$

Thus when $\underset{\sim}{Q}$ is known, Eqs (61) and (65) may be used to solve for $\underset{\sim}{\Gamma}^*$ and $\underset{\sim}{P}^*$. Mehra (Ref 181) presents an iterative procedure for solving equations of the form of Eqs (61) and (65) for the case of $\underset{\sim}{Q}$ equal to the identity matrix. Identification of the covariance matrices $\underset{\sim}{Q}$ and R is discussed in the next section.

One of the advantages of using the canonical form system representation discussed above is that the number of unknown parameters required to identify the various system matrices has been reduced. In the original system representation given by Eqs (1) and (2), there are a total of $n^2 + np + n$ unknown parameters required to identify $\underset{\sim}{\Phi}$, $\underset{\sim}{\Gamma}$, and $\underset{\sim}{H}$. On the other hand, in the canonical form representation given by Eqs (49) and (50), there are only a total of $n + np$ unknown parameters required to identify $\underset{\sim}{\Phi}^*$, $\underset{\sim}{\Gamma}^*$, and $\underset{\sim}{H}^*$. Hence, use of the canonical or phase variable system representation has significantly reduced the total number of unknown parameters requiring identification.

Identification of $\underset{\sim}{Q}$ and R

In the system representation given by Eqs (1) and (2), the covariance matrix of the input gaussian white-noise sequence $\underset{\sim}{U}(k)$, having disturbance transition matrix $\underset{\sim}{\Gamma}$, is assumed to be a symmetric positive definite matrix $\underset{\sim}{Q}$. In this section it is shown that no generality is lost in the system model if the covariance of $\underset{\sim}{U}(k)$ is assumed to be the identity matrix, provided that $\underset{\sim}{\Gamma}$ is adjusted accordingly.

The development begins by considering the following well known result from matrix algebra (Ref 176): if a p x p symmetric matrix $\underset{\sim}{Q}$ is positive definite, then there exists a nonsingular p x p matrix $\underset{\sim}{F}$ such that

$$\underset{\sim}{Q} = \underset{\sim}{F}'\underset{\sim}{F} \tag{66}$$

Substituting Eq (66) into (30), the following result is obtained:

$$\underset{\sim}{P} = \underset{\sim\sim}{\Phi P \Phi}' + \underset{\sim\sim}{\Gamma F' F \Gamma}'$$

$$= \underset{\sim\sim}{\Phi P \Phi}' + (\underset{\sim\sim}{\Gamma F'})(\underset{\sim\sim}{\Gamma F'})' \tag{67}$$

Letting the n x p matrix $\underset{\sim}{\Gamma}_e \equiv \underset{\sim\sim}{\Gamma F}'$ be an equivalent disturbance transition matrix, then Eq (67) implies that Eq (1) of the original system model representation may be rewritten as

$$\underset{\sim}{X}(k+1) = \underset{\sim\sim}{\Phi X}(k) + \underset{\sim e}{\Gamma}_e \underset{\sim e}{U}(k) \tag{68}$$

In Eq (68), $\underset{\sim e}{U}(k)$ is a p x 1 vector of equivalent gaussian white-noise having the following properties:

$$E[\underset{\sim e}{U}(k)] = \underset{\sim}{0}$$

$$E[\underset{\sim e}{U}(j)\underset{\sim e}{U}'(k)] = \underset{\sim}{I}\delta_{jk}$$

and $\underset{\sim}{I}$ is the p x p identity matrix.

Conceptually, in arriving at Eq (68), the input disturbance matrix $\underset{\sim}{\Gamma}$ is adjusted in such a way (viz., $\underset{\sim\sim}{\Gamma F}$) that the gaussian white-noise input sequence $\{\underset{\sim}{U}(k), k = 0, 1, \ldots\}$

appears to have its covariance equal to the identity matrix. That is, the original input model $\underset{\sim}{\Gamma}\underset{\sim}{U}(k)$ can be thought of as being replaced by an equivalent input model $\underset{\sim}{\Gamma}_e\underset{\sim}{U}_e(k)$. Statistically, however, the two input models are equivalent with respect to the state vector $\underset{\sim}{X}(k)$. Therefore, no loss in generality occurs in the system model given by Eqs (1) and (2) if $\underset{\sim}{Q}$ is replaced by $\underset{\sim}{I}$ in Eq (4):

$$E[\underset{\sim}{U}(j)\underset{\sim}{U}'(k)] = \underset{\sim}{I}\delta_{ij} \tag{69}$$

The only parameter remaining to be identified in the system representation given by Eqs (1) and (2) is R, the covariance of the gaussian white-noise measurement error. A method that yields asymptotically unbiased, normal, and consistent estimates of R is suggested by Mehra (Ref 78). The method is based on the autocorrelation function of the output measurements when k = 0. Substituting Eq (35) with j = 1 into Eq (32) with k = 0 yields

$$C_0 = \underset{\sim}{H}\underset{\sim}{B}_1^{-1} \begin{bmatrix} C_1 \\ C_2 \\ . \\ . \\ . \\ C_n \end{bmatrix} + R \tag{70}$$

Premultiplying Eq (37) by $\underset{\sim}{H}$ and solving the resulting equation for $a_1\underset{\sim}{H}$ results in

$$a_1 \underset{\sim}{H} = -a_{n+1} \underset{\sim}{H\Phi}^n - \ldots - a_3 \underset{\sim}{H\Phi}^2 - a_2 \underset{\sim}{H\Phi}$$

$$= -[a_2, a_3, \ldots, a_{n+1}] \begin{bmatrix} \underset{\sim}{H\Phi} \\ \underset{\sim}{H\Phi}^2 \\ \cdot \\ \cdot \\ \cdot \\ \underset{\sim}{H\Phi}^n \end{bmatrix} \qquad (71)$$

where $a_{n+1} \equiv 1$. Substituting Eq (34) with j = 1 into Eq (71), and solving the resulting equation for $\underset{\sim}{HB}_1^{-1}$ yields

$$\underset{\sim}{HB}_1^{-1} = -\frac{1}{a_1} [a_2, a_3, \ldots, a_{n+1}] \qquad (72)$$

Substituting Eq (72) into (70), the following result is obtained:

$$a_1 C_0 = -\sum_{j=1}^{n} a_{j+1} C_j + a_1 R$$

from which the indicated estimate of R is found to be

$$\hat{R} = \frac{1}{\hat{a}_1} \sum_{j=0}^{n} \hat{a}_{j+1} \hat{C}_j \qquad (73)$$

The $\hat{C}_j$ in Eq (73) can be calculated using Eq (43), and the $\hat{a}_j$ can be calculated using Eq (44).

With respect to the statistical properties of the state vector $\underset{\sim}{X}(k)$, it was shown earlier in this chapter that no loss in generality occurs in the system model if the covariance of the input noise sequence is assumed to be the

identity matrix. However, a major problem exists if $\underset{\sim}{Q}$ must be identified explicitly. The crux of the problem is that even though the existence of a matrix $\underset{\sim}{F}$ that satisfies Eq (66) is guaranteed when $\underset{\sim}{Q}$ is symmetric and positive definite, $\underset{\sim}{F}$ must still be identified in order to calculate $\underset{\sim}{Q}$. One possible way in which the problem of identifying $\underset{\sim}{Q}$ and $\underset{\sim}{R}$ may be approached is through adaptive Kalman filtering (Ref 79). Although identification of the noise covariances will not be pursued further in this paper, Kalman filtering will be used in the next chapter to develop another identification method.

V. System Identification Using Adaptive Kalman Filtering

Parameter identification using two system representations was discussed in the preceding chapter. It was shown that distinct advantages could be obtained by using a canonical form system representation rather than the original system representation. In this chapter another model of the system, called "Levy's proper canonical form" (Ref 74), is developed. This model is shown to offer yet additional advantages in solving the identification problem.

Levy's proper canonical form is based on the Kalman filter representation of the system (Ref 101), and is derived using the "innovations sequence" of the filter (Ref 76). Hence, the chapter begins with a definition of the innovations sequence, and proceeds to a development of some of its properties. Using these results, the optimal filtered state estimate is derived, and combined with certain results from optimal estimation theory to yield the desired canonical form.

The Innovations Sequence

The innovations approach to linear least-squares estimation is discussed by Kailath (Ref 76). Basically, the innovations approach is to use a causal and causally invertible linear transformation to first "whiten" the observed output data, i.e., convert the observed output to a white-noise sequence. The reason for doing this is that

with white-noise observations, the estimation problem is greatly simplified. Then, once the solution to this simplified problem is obtained, the inverse of the original "whitening" filter can be used to express the solution in terms of the original output observations.

For the convenience of the reader, at this point the basic system model for the vector-input scalar-output case is restated, and certain additional terminology is introduced (Ref 100):

$$\underset{\sim}{X}(k+1) = \underset{\sim\sim}{\Phi}\underset{\sim}{X}(k) + \underset{\sim\sim}{\Gamma}\underset{\sim}{U}(k) \tag{74}$$

$$Z(k) = Y(k) + V(k) \tag{75}$$

$$Y(k) = \underset{\sim\sim}{H}\underset{\sim}{X}(k) \tag{76}$$

where $Y(k)$ is the output message, $V(k)$ is the output measurement noise, and $Z(k)$ is the output signal (message plus noise). Using Eqs (16) and (17), the conditional mean and estimation error of the message sequence are, respectively

$$\hat{Y}(k|k-1) = E[Y(k)|Z(j),\ 0 \leq j \leq k-1] \tag{77}$$

$$\tilde{Y}(k|k-1) = Y(k) - \hat{Y}(k|k-1) \tag{78}$$

The innovations sequence of $Z(\cdot)$, denoted by $\alpha(\cdot)$, is defined by Kailath (Ref 76) to be the difference between the output signal and the conditional mean of the message sequence. Substituting Eqs (75) and (78) into this difference

yields

$$\alpha(k) \equiv Z(k) - \hat{Y}(k|k-1) = Y(k) + V(k) - \hat{Y}(k|k-1)$$

$$= \tilde{Y}(k|k-1) + V(k) \qquad (79)$$

Another interpretation of the innovations sequence $\alpha(\cdot)$ may be deduced as follows. Taking the conditional expectation of Eq (75) given $Z(j)$, $0 \leq j \leq k - 1$, results in

$$\hat{Z}(k|k-1) = \hat{Y}(k|k-1) + \hat{V}(k|k-1) \qquad (80)$$

From Eq (15), however, it is seen that $V(k)$ is independent of $Z(j)$, $0 \leq j \leq k - 1$. That is, future measurement noise is independent of past output signals, and hence

$$\hat{V}(k|k-1) = E[V(k)|Z(j),\ 0 \leq j \leq k - 1]$$

$$= E[V(k)]$$

$$= 0 \qquad (81)$$

Therefore substituting Eq (81) into (80) yields

$$\hat{Z}(k|k-1) = \hat{Y}(k|k-1) \qquad (82)$$

after which Eq (79) becomes

$$\alpha(k) = Z(k) - \hat{Z}(k|k-1) = \tilde{Z}(k|k-1)$$

The previous equation indicates that the innovations sequence $\alpha(k)$ may be regarded as the estimation error of the output signal, i.e., the difference between the output at time k and

the optimal estimate of the output at time k given all previous observations of the output. This difference is precisely the "new information" brought by the latest output observation Z(k).

Properties of the Innovations Sequence

The following properties of the innovations sequence $\alpha(\cdot)$ are discussed by Kailath (Ref 76) and Mehra (Ref 79):

1. The stochastic process defined by the innovations sequence $\{\alpha(k), k = 0, 1, \ldots\}$ is a zero-mean Gauss-Markov sequence. This property follows from the fact that by using property 2c of the system model and Eq (76), $\hat{Y}(k|k-1)$ is seen to be a linear combination of the available Z(s), $0 \leq s \leq k - 1$. Thus, Eq (79) indicates that $\alpha(k)$ is also a linear combination of the available Z(s), $0 \leq s \leq k$. It follows, therefore, that since the $Z(\cdot)$ form a zero-mean Gauss-Markov sequence (property 1a of the system model), the $\alpha(\cdot)$ also form a zero-mean Gauss-Markov sequence.

2. The innovations sequence $\alpha(\cdot)$ is a white-noise sequence. This property may be established by using Eq (79) and considering the covariance of the innovations sequence:

$$E[\alpha(j)\alpha(k)] = E\{[\tilde{Y}(j|j-1)+V(j)][\tilde{Y}(k|k-1)+V(k)]\}$$

$$= E[\tilde{Y}(j|j-1)\tilde{Y}(k|k-1)] + E[\tilde{Y}(j|j-1)V(k)]$$

$$+ E[V(j)\tilde{Y}(k|k-1)] + E[V(j)V(k)] \qquad (83)$$

However, using Eqs (14), (17), and (76), it follows that

$$E[\tilde{Y}(j|j-1)V(k)] = \underset{\sim}{H}E[\tilde{\underset{\sim}{X}}(j|j-1)V(k)]$$

$$= \underset{\sim}{H}E[\underset{\sim}{X}(j)V(k)] - \underset{\sim}{H}E[\hat{\underset{\sim}{X}}(j|j-1)V(k)]$$

$$= -\underset{\sim}{H}E[\hat{X}(j|j-1)V(k)], \quad k > j$$

Using property 2c of the system model and Eq (15), the last equation becomes

$$E[\tilde{Y}(j|j-1)V(k)] = 0, \quad k > j \tag{84}$$

Also, using Eqs (19), (76), and (78), it follows that

$$E[\tilde{Y}(j|j-1)\tilde{Y}(k|k-1)] = E[Y(j)\tilde{Y}(k|k-1)] - E[\hat{Y}(j|j-1)\tilde{Y}(k|k-1)]$$

$$= E[Y(j)\tilde{Y}(k|k-1)], \quad k > j \tag{85}$$

Therefore, substituting Eqs (84) and (85) into (83), and using Eqs (20) and (75), the following result is obtained:

$$E[\alpha(j)\alpha(k)] = E[Y(j)\tilde{Y}(k|k-1)] + E[V(j)\tilde{Y}(k|k-1)]$$

$$+ E[V(j)V(k)], \quad k > j$$

$$= E[Z(j)\tilde{Y}(k|k-1)] + E[V(j)V(k)]$$

$$= E[V(j)V(k)], \quad k > j \tag{86}$$

By a similar development it can be shown that Eq (86) also holds for $k < j$. Therefore, substituting Eq (6) into (86) yields

$$E[\alpha(j)\alpha(k)] = 0, \quad j \neq k \tag{87}$$

The variance of the innovations sequence may be obtained by substituting j = k into Eq (83), and using Eqs (76) and (78):

$$\begin{aligned} E[\alpha(k)\alpha(k)] &= E[V(k)V(k)] + 2E[\tilde{Y}(k|k-1)V(k)] \\ &\quad + E[\tilde{Y}(k|k-1)\tilde{Y}(k|k-1)] \\ &= E[V(k)V(k)] + 2\underset{\sim}{H}E[\underset{\sim}{X}(k)V(k)] \\ &\quad - 2\underset{\sim}{H}E[\hat{\underset{\sim}{X}}(k|k-1)V(k)] + E[\tilde{Y}(k|k-1)\tilde{Y}(k|k-1)] \end{aligned} \tag{88}$$

Using Eqs (14) and (15), and property 2c of the system model, Eq (88) becomes

$$E[\alpha(k)\alpha(k)] = E[V(k)V(k)] + E[\tilde{Y}(k|k-1)\tilde{Y}(k|k-1)] \tag{89}$$

Defining the steady state variance of the estimation error of the message sequence as

$$P_Y \equiv E[\tilde{Y}(k|k-1)\tilde{Y}(k|k-1)], \ \forall\, k \tag{90}$$

and using Eq (6), the following result is obtained from Eq (89) for the variance of the innovations sequence:

$$E[\alpha(k)\alpha(k)] = R + P_Y, \ j = k \tag{91}$$

Hence, combining Eqs (87) and (91), the variance-covariance matrix of the innovations sequence may be written as

$$E[\alpha(j)\alpha(k)] = (R + P_Y)\delta_{jk} \qquad (92)$$

A comparison of Eqs (6) and (92) reveals that the innovations $\alpha(\cdot)$ is a white-noise sequence like the measurement noise $V(\cdot)$, but with a different variance.

3. For causal linear operations (i.e., linear operations that do not require the future value of one variable to determine the current value of another variable), $\alpha(\cdot)$ and $Z(\cdot)$ are "statistically equivalent" and may be obtained from one another. For the case where $Z(j)$, $0 \leq j \leq k$, is known, then $\hat{Y}(k|k-1)$ can be calculated using Eq (82). Hence, from Eq (79), it is seen that $\alpha(k)$ is completely determined by $Z(j)$, $0 \leq j \leq k$. For the case where $\alpha(\cdot)$ is known, $Z(\cdot)$ may be obtained by using the celebrated Kalman-Bucy formula, and the proof is given in Ref 76.

Levy's Proper Canonical Form

In this section, Levy's proper canonical form representation of the system is derived. This representation is based on the Kalman filter for the system, and is discussed by Geesey and Kailath (Ref 74) and Mehra (Refs 78 and 79). Preparatory to the formal derivation that follows, however, a preliminary result is stated in the form of the following theorem (Ref 76):

Theorem 2: The Projection Theorem

In the original system representation given by Eqs (74) through (76), the best estimate of the message error $\tilde{Y}(k|k-1)$

is unique and satisfies the following condition:

$$\tilde{Y}(k|k-1) = Y(k) - \hat{Y}(k|k-1) \perp Z(s), \quad 0 \leq s \leq k-1 \tag{93}$$

where

$$Y \perp Z \quad \text{means} \quad E[YZ] = 0 \tag{94}$$

In words, the projection theorem states that the instantaneous message error is uncorrelated with the output signal.

The first step in the derivation of Levy's proper canonical form involves the development of a useful result from optimal estimation theory. Meditch (Ref 180) shows that any linear estimate of the state $\underset{\sim}{X}(k)$ given the set of output measurements $\{Z(s) = Y(s) + V(s), \ 0 \leq s \leq k\}$ can be written as

$$\hat{\underset{\sim}{X}}(k|k) = \sum_{s=0}^{k} \underset{\sim}{G}(k,s)Z(s) \tag{95}$$

where the $\underset{\sim}{G}(k,s)$ are n x 1 vectors. By property 3 of the innovations sequence, $\alpha(\cdot)$ and $Z(\cdot)$ are statistically equivalent for causal linear operations, and therefore Eq (95) implies that

$$\hat{\underset{\sim}{X}}(k|k) = \sum_{s=0}^{k} \underset{\sim}{G}(k,s)\alpha(s) \tag{96}$$

In Eq (96), the vector $\underset{\sim}{G}(k,\cdot)$ acts as a linear filter that is selected such that the instantaneous error in measuring the state vector is independent of the output signal, and

hence is also independent of the innovations. That is,

$$\tilde{\underset{\sim}{X}}(k|k) = \underset{\sim}{X}(k) - \hat{\underset{\sim}{X}}(k|k) \perp \alpha(s),\ 0 \leq s \leq k \tag{97}$$

Postmultiplying Eq (97) by $\alpha(s)$, taking the expected value of the result, and using the projection theorem results in

$$E[\tilde{\underset{\sim}{X}}(k|k)\alpha(s)] = \underset{\sim}{0} = E[\underset{\sim}{X}(k)\alpha(s)] - E[\hat{\underset{\sim}{X}}(k|k)\alpha(s)]$$

or

$$E[\underset{\sim}{X}(k)\alpha(s)] = E[\hat{\underset{\sim}{X}}(k|k)\alpha(s)] \tag{98}$$

Substituting Eq (96) into (98), and using Eq (92) yields

$$E[\underset{\sim}{X}(k)\alpha(s)] = E\left[\sum_{\sigma=0}^{k} \underset{\sim}{G}(k,\sigma)\alpha(\sigma)\alpha(s)\right] \tag{99}$$

$$= \sum_{\sigma=0}^{k} \underset{\sim}{G}(k,\sigma)E[\alpha(\sigma)\alpha(s)] \tag{100}$$

$$= \underset{\sim}{G}(k,s)(R + P_Y),\ 0 \leq s \leq k \tag{101}$$

Equation (101) says that in the steady state, the correlation function between the state vector and the output signal is proportional to the linear filter $\underset{\sim}{G}(k,\cdot)$. Since in the scalar-output case, R and P_Y are both constant scalars, the existence or physical realizability of $\underset{\sim}{G}(k,\cdot)$ is assured. Hence, from Eq (101) it follows that

$$\underset{\sim}{G}(k,s) = E[\underset{\sim}{X}(k)\alpha(s)](R + P_Y)^{-1} \tag{102}$$

Postmultiplying Eq (102) by $\alpha(s)$, summing the result from 0 to k, and using Eq (96) yields

$$\hat{\underset{\sim}{X}}(k|k) = \sum_{s=0}^{k} \{E[\underset{\sim}{X}(k)\alpha(s)](R + P_Y)^{-1}\alpha(s)\} \qquad (103)$$

which implies that

$$\hat{\underset{\sim}{X}}(k+1|k+1) = \sum_{s=0}^{k+1} \{E[\underset{\sim}{X}(k+1)\alpha(s)](R + P_Y)^{-1}\alpha(s)\}$$

$$= \sum_{s=0}^{k} \{E[\underset{\sim}{X}(k+1)\alpha(s)](R + P_Y)^{-1}\alpha(s)\}$$

$$+ E[\underset{\sim}{X}(k+1)\alpha(k+1)](R + P_Y)^{-1}\alpha(k+1) \qquad (104)$$

If the steady state Kalman filter gain is now defined as the n x 1 vector $\underset{\sim}{K}$, where

$$\underset{\sim}{K} \equiv E[\underset{\sim}{X}(\cdot)\alpha(\cdot)](R + P_Y)^{-1}$$

then Eq (104) becomes

$$\hat{\underset{\sim}{X}}(k+1|k+1) = \sum_{s=0}^{k} \{E[\underset{\sim}{X}(k+1)\alpha(s)](R + P_Y)^{-1}\alpha(s)\} + \underset{\sim}{K}\alpha(k+1)$$

Substituting Eq (74) into the last equation yields

$$\hat{\underset{\sim}{X}}(k+1|k+1) = \sum_{s=0}^{k} \{E[\underset{\sim}{\Phi}\underset{\sim}{X}(k)\alpha(s)+\underset{\sim}{\Gamma}\underset{\sim}{U}(k)\alpha(s)](R + P_Y)^{-1}\alpha(s)\}$$

$$+ \underset{\sim}{K}\alpha(k+1)$$

$$= \underset{\sim}{\Phi} \sum_{s=0}^{k} \{E[\underset{\sim}{X}(k)\alpha(s)](R + P_Y)^{-1}\alpha(s)\}$$

$$+ \underset{\sim}{\Gamma} \sum_{s=0}^{k} \{E[\underset{\sim}{U}(k)\alpha(s)](R + P_Y)^{-1}\alpha(s)\}+\underset{\sim}{K}\alpha(k+1) \quad (105)$$

However, using property 2c of the system model, and Eqs (13), (76), and (79), it follows that

$$\sum_{s=0}^{k} \{E[\underset{\sim}{U}(k)\alpha(s)]\} = \sum_{s=0}^{k} \{E[\underset{\sim}{U}(k)Z(s)] - E[\underset{\sim}{U}(k)\hat{Y}(s|s-1)]\}$$

$$= \underset{\sim}{0} \quad (106)$$

Substituting Eq (106) into (105), and using Eq (103) yields the following result:

$$\hat{\underset{\sim}{X}}(k+1|k+1) = \underset{\sim}{\Phi}\hat{\underset{\sim}{X}}(k|k) + \underset{\sim}{K}\alpha(k+1) \quad (107)$$

For convenience, Eq (18) is rewritten below:

$$\hat{\underset{\sim}{X}}(k+1|k) = \underset{\sim}{\Phi}\hat{\underset{\sim}{X}}(k|k) \quad (108)$$

The optimal filtered estimate of the state $\underset{\sim}{X}(k+1)$ is given by Eq (107), while the optimal predicted estimate is given by Eq (108). Together, Eqs (107) and (108) constitute a partial description of the steady state Kalman filter for

the system given by Eqs (74) through (76). In recent years the Kalman filter has enjoyed wide applicability in the field of optimal control, principally because its recursive form is ideally suited for implementation using digital computers. For a complete mathematical description of the Kalman filter, it is also necessary to determine the optimal filtering gain and the optimal covariances of the filtered and predicted estimates. The interested reader is referred to the literature for details (Refs 101 and 180).

The derivation of Levy's proper canonical form continues by rewriting Eq (107) as

$$\hat{\underset{\sim}{X}}(k+1|k+1) = \hat{\underset{\sim}{X}}(k+1|k) + \underset{\sim}{K}\alpha(k+1)$$

which implies that

$$\hat{\underset{\sim}{X}}(k|k) = \hat{\underset{\sim}{X}}(k|k-1) + \underset{\sim}{K}\alpha(k) \tag{109}$$

Substituting Eq (109) into (108) yields

$$\hat{\underset{\sim}{X}}(k+1|k) = \underset{\sim}{\Phi}[\hat{\underset{\sim}{X}}(k|k-1)+\underset{\sim}{K}\alpha(k)] \tag{110}$$

Premultiplying Eq (110) by $\underset{\sim}{T}$, where $\underset{\sim}{T}$ is defined in Eq (48), and using the linear transformation given by Eq (47), the following form is obtained for Eq (110):

$$\hat{\underset{\sim}{X}}^*(k+1|k) = \underset{\sim}{\Phi}^*[\hat{\underset{\sim}{X}}^*(k|k-1)+\underset{\sim}{K}^*\alpha(k)] \tag{111}$$

where $\underset{\sim}{\Phi}^*$ is defined in Eq (51), and

$$\underset{\sim}{K}^* = \underset{\sim}{T}\underset{\sim}{K} \tag{112}$$

Solving for Z(k) from Eq (79) and using Eq (76) results in

$$Z(k) = \hat{Y}(k|k-1) + \alpha(k)$$

$$= \underset{\sim}{H}\hat{\underset{\sim}{X}}(k|k-1) + \alpha(k)$$

Using the transformation given by Eqs (47) and (48), the last equation becomes

$$Z(k) = \underset{\sim}{H}^*\hat{\underset{\sim}{X}}^*(k|k-1) + \alpha(k) \qquad (113)$$

where $\underset{\sim}{H}^*$ is defined in Eq (53).

Taken together, Eqs (111) and (113) comprise Levy's proper canonical form, which is shown in Block diagram representation in Fig. 5. Levy's proper canonical form representation has the following advantages over the canonical form representation given by Eqs (49) and (50):

1. The optimal filtered and predicted estimates of the state vector are obtained directly.

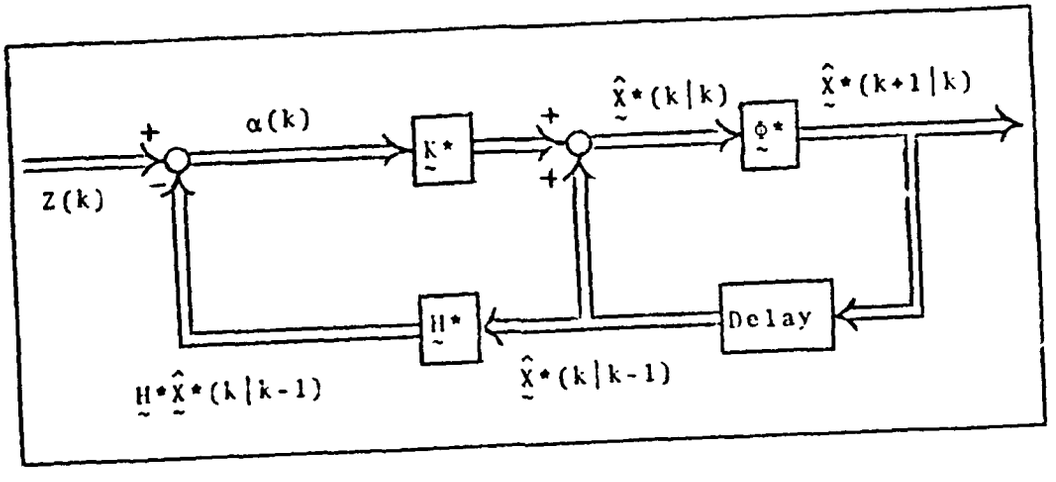


Fig. 5. Model of Levy's Proper Canonical Form.

2. The innovations sequence $\alpha(k)$ is obtained directly from the output measurements $Z(k)$ and the optimal predicted estimate of the state vector $\hat{\underset{\sim}{X}}^*(k|k-1)$, where $\hat{\underset{\sim}{X}}^*(k|k-1)$ can be realized using the feedback design shown in Fig. 5. Availability of the innovations sequence enables a check to be performed concerning the optimality of a Kalman filter constructed using parameter estimates, as well as providing information pertaining to the statistical quality of the parameter estimates. These tests are discussed by Mehra (Ref 79).

3. For vector-input scalar-output systems, Levy's proper canonical form representation effects a net reduction in the number of unknown parameters to be estimated, i.e., $\underset{\sim}{K}^*$ contains n unknown elements, whereas $\underset{\sim}{\Gamma}^*$ contains np unknown elements.

Identification of $\underset{\sim}{K}^*$

Although distinct advantages may be gained by using Levy's proper canonical form over other system representations, a problem nevertheless exists in that the optimal gain $\underset{\sim}{K}^*$ remains to be estimated. In this section, an algorithm for estimating K* is discussed (Ref 78). The algorithm is based on the innovations sequence $\alpha(\cdot)$, and is capable of being implemented on-line.

As pointed out in the preceding section, a complete mathematical description of the Kalman filter, and hence of Levy's proper canonical form, requires knowledge of the

optimal filter gain and the optimal covariance of the error in estimating the filtered and predicted states. When all of the parameters in the original system formulation (viz., $\underset{\sim}{\Phi}$, $\underset{\sim}{\Gamma}$, $\underset{\sim}{H}$, $\underset{\sim}{Q}$, and $\underset{\sim}{R}$) are known, expressions for determining the optimal gain and covariances in the Kalman filter are well documented in the literature (Refs 101 and 180). However, when the system parameters are initially unknown and require estimation, a Kalman filter constructed using these parameter estimates may not be optimal. In cases where the parameter identification is particularly poor, the Kalman filter may become unstable and even diverge (Refs 95 and 97). On the other hand, when the parameter estimates are good, the Kalman filter will be stable and the state vector will converge to its true value.

Mehra (Ref 79) shows that a necessary and sufficient condition for the optimality of a Kalman filter is that the innovations sequence $\alpha(\cdot)$ be white. This condition forms the basis for the algorithm that estimates the optimal filter gain $\underset{\sim}{K}^*$, for when the innovations sequence is white, the corresponding value of the filter gain is optimal. Very basically, the algorithm begins by arbitrarily assuming an initial value of the filter gain $\underset{\sim}{K}_0$ and filtering the output observations {Z(0), ... Z(n)} using Eqs (111) and (113). Next, the innovations sequence, which Mehra shows to be stationary under steady state conditions, is autocorrelated and tested for whiteness. If the innovations sequence is

not white, the filter gain is changed to $\underset{\sim}{K}_1$, where $\underset{\sim}{K}_1$ is determined by a recursive relationship involving $\underset{\sim}{K}_0$, $\underset{\sim}{\Phi}^*$, $\underset{\sim}{H}^*$, and the normalized autocorrelations of the innovations sequence. The above procedure is repeated until the sequence $\alpha(\cdot)$ is white. The sequence of filter gains, $\underset{\sim}{K}_0$, $\underset{\sim}{K}_1$, ..., is shown to converge to the optimal value $\underset{\sim}{K}^*$ in the mean-square sense.

VI. Conclusions and Recommendations for Further Study

Conclusions

This thesis considered the problem of the identification of linear stochastic systems. A basically non-technical state-of-the-art assessment of the subject area was made, and criteria for the classification and selection of identification methods were presented and discussed. Several of the more popular identification methods from the literature were investigated and summarized. The methods were shown to have application to such diverse fields as economics, industrial processes, and aircraft design.

Using the state variable formulation for a discrete linear stochastic system, a detailed exposition of a few of the on-line identification methods currently appearing in the literature was presented. One such method, based on the autocorrelation function of the output measurements, was developed to identify the state transition matrix and the output noise covariance. It was shown that a canonical or phase variable system representation could be used to reduce the number of unknown parameters requiring identification. Finally, an on-line identification method called Levy's proper canonical form, based on the Kalman filter representation of the system, was derived using the innovations sequence and certain results from optimal estimation theory. It was shown that this identification method resulted in still

additional advantages over the identification methods previously developed.

Recommendations for Further Study

Obviously, this thesis has not answered all interesting questions concerning the identification of linear stochastic systems. A fruitful area for further research is discussed below.

Regardless of whether the ultimate purpose of an identification analysis is parameter identification alone, or for control application, the degree of accuracy required is an important factor that can often dictate the type of identification method to employ. However, a major problem exists in this regard in determining the degree of accuracy attainable using a particular identification method, and formulating criteria for measuring that accuracy. Unfortunately, the literature of system identification is far from complete in this area. Therefore, it is recommended that a criteria be developed for measuring the accuracy and sensitivity of different identification methods, and that such criteria be validated by comparing several identification methods using simulation techniques.

Bibliography

Maximum Likelihood

1. Aoki, M. and Yue, P. C. "On a Priori Error Estimates of Some Identification Methods." IEEE Trans. Aut. Control, Vol. AC-15, 5:541-548 (October 1970).

2. Astrom, K. J. and Bohlin, T. "Numerical Identification of Linear Dynamic Systems From Normal Operating Records." Proc. Second IFAC Symp. Theory of Self-Adaptive Systems (1965).

3. Balakrishnan, A. B. "Identification of Linear Dynamic Systems." Third Southeastern Symposium on System Theory, Georgia Institute of Technology (April 1971).

4. Kashyap, R. L. "Maximum Likelihood Identification of Stochastic Linear Systems." IEEE Trans. Aut. Control, Vol. AC-15, 5:25-34 (February 1970).

5. Mehra, R. K. "Identification of Linear Dynamic Systems Using Kalman Filter Representation." AIAA Journal, Vol 9, 1:28-31 (January 1971).

6. Mehra, R. K. "Maximum Likelihood Identification of Aircraft Parameters." Joint Aut. Contr. Conf., Atlanta, Georgia (June 1970).

7. Panuska, V. "On the Maximum Likelihood Estimation of Rational Transfer Function Parameters." IEEE Trans. Aut. Control, Vol. AC-13:304 (1968).

8. Rauch, H. E., et. al. "Maximum Likelihood Estimates of Linear Dynamic Systems." AIAA Journal, Vol. 3, 8:1445-1450 (August 1965).

9. Rogers, A. E. and Steiglitze, K. "Maximum Likelihood Estimation of Rational Transfer Function Parameters." IEEE Trans. Aut. Control, Vol. AC-12:594-597 (1967).

10. Sage, A. P. "Maximum Likelihood Estimation of Bias in Sequential Estimation." Proceedings M. J. Kelly Communications Conference, University of Missouri at Rolla (October 1970).

11. Schweppe, F. C. "Evaluation of Likelihood Functions for Gaussian Signals." IEEE Trans. Inf. Theory, Vol. IT-11: 76-111 (January 1965).

23. Kopp, R. E. and Orford, R. J. "Linear Regression Applied to System Identification for Adaptive Control Systems." AIAA Journal, Vol. 1, 10:2300-2306 (October 1963).

24. Lainiotis, D. G. "Optimal Adaptive Estimation: Structure and Parameter Adaptation." IEEE Trans. Aut. Control, Vol. AC-16, 2:160-170 (April 1971).

25. Levy, L. J. "Adaptive Estimation Algorithms." Engineering Research Institute, Technical Report 93400, Iowa State University (December 1970).

26. Magill, D. T. "Optimal Adaptive Estimation of Sampled Stochastic Processes." IEEE Trans. Aut. Control, Vol AC-10, 4:434-439 (October 1965).

27. Oza, K. G. and Jury, E. I. "Adaptive Algorithm for Identification Problem." Automatica, Vol. 6:795-799 (1970).

28. Roy, R. J. and Sherman, J. ""A Learing Technique for Volterra Series Representation.""IEEE Trans. Aut. Control, Vol. AC-12:761-764 (1967).

29. Sage, A. P. and Husa, G. W. "Adaptive Filtering with Unknown Prior Statistics." Technical Report, Information and Control Sciences Center, SMU Institute of Technology (1970).

30. Sengbush, R. L. and Lainiotis, D. G. "Simplified Parameter Quantization Procedure for Adaptive Estimation." IEEE Trans. Aut. Control, Vol. AC-14:424-425 (August 1969(.

31. Shellenbarger, J. C. "Estimation of Covariance Parameters for an Adaptive Kalman Filter." Proc. Nat. Electronics Conf. pp. 698-702 (1966).

32. Sims, F. L., et. al. "Recursive Algorithm for the Calculation of the Adaptive Kalman Filter Weighting Coefficients." IEEE Trans. Aut. Control, Vol. AC-14: 215-218 (April 1969).

33. Smith, G. L. "Sequential Estimation of Observation Error Variances in a Trajectory Estimation Problem." AIAA Journal, Vol. 5, 11:1964-1970 (November 1967).

34. Tsypkin, Y. Z. "Generalized Algorithms of Learning." Aut. Remote Control, Vol. 31:97-104 (1970).

Model Reference

35. Belanger, P. R. "Comments on 'A Learning Method for System Identification'." IEEE Trans. Aut. Control, Vol. AC-13:207-208 (1968).

36. Dressler, R. M. "An Approach to Model-Referenced Adaptive Control Systems." IEEE Trans. Aut. Control, Vol. AC-12: 75-80 (February 1967).

37. Hsia, T. C. and Vimolvanich, V. "An On-Line Technique for System Identification." IEEE Trans. Aut. Control, Vol. AC-14:92-96 (1969).

38. Lion, P. M. "Rapid Identification of Linear and Nonlinear Systems." Joint Aut. Contr. Conf., pp. 605-615 (1966).

39. Meissinger, H. F. and Bekey, G. A. "An Analysis of Continuous Parameter Identification Method." Simulation, Vol. 6:95-102 (1966)

40. Nagumo, J. and Noda, A. "A Learning Method for System Identification." IEEE Trans. Aut. Control, Vol. AC-12: 282-287 (June 1967).

41. Parks, P. C. "Liapunov Redesign of Model Reference Adaptive Control Systems." IEEE Trans. Aut. Control, Vol. AC-11:362-367 (July 1966).

42. Schley, C. and Roy, R."Process Identification Using a Mode Learnirg Machine." Proc. Seventh Joint Aut. Contr. Conf. pp. 639-648, University of Washington (1966).

43. Whitaker, H. P., et. al. "Design of a Model Reference Adaptive Control System for Aircraft." Report R-164, Instrumentation Lab., MIT (1958).

Stochastic Approximation

44. Albert, A. E. and Gardner, L. A. Stochastic Approximation and Non-Linear Regression. MIT Press, Cambridge, Mass. (1967).

45. Bekey, G. A. and Neal, C. B. "Estimation of the Parameters of a Sampled Data System by Stochastic Approximation." Technical Report, Electronic Sciences Lab., Univ. of South. Calif. (1969).

46. Bojanov, E. S. "Application of Stochastic Approximation Method for Object Characteristic Regeneration." Autom remote Control, pp. 95-103 (1967)

47. Dvoretsky, A. "On Stochastic Approximation." Third Berkeley Symp. Math. Statist. and Prob., University of California Press (1956).

48. Elliot, D. F. and Sworder, D. D. "Applications of a Simplified Multidimensional Stochastic Approximation Algorithm." Proc. Joint Aut. Contr. Conf. pp. 148-154 (1969)

49. Fu, K. S., et. al. "Stochastic Approximation Algorithms for System Identification, Estimation, and Decomposition of Mixtures." Proc. Fifth Allerton Conf. on Circuit and System Theory (1967).

50. Ho, Y. C. "On the Stochastic Approximation Method and the Optimal Filtering Theory." Math. Analysis and Applications, Vol. 6:152-155 (February 1963).

51. Ho, Y. C. and Lee, R. C. K. "Identification of Linear Dynamic Systems." Information and Control, Vol. 8:93-110 (1965).

52. Ho, Y. C. and Whalen, B. H. "An Approach to the Identification and Control of Linear Dynamic Systems with Unknown Parameters." IEEE Trans. Aut. Control, Vol. AC-8:255-256 (July 1963).

53. Holmes, J. K. "Two Stochastic Approximation Procedures for Identifying Linear Systems." IEEE Trans. Aut. Control, Vol. AC-14:3 (June 1969).

54. Koginov, N. V. "Stochastic Approximation Methods." Automn remote Control, 4:185 (1966).

55. Neal, C. B. "Estimation of the Parameters of Sampled Data Systems by Stochastic Approximation." Tech. Report, Electronic Sciences Lab., Univ. of South. Calif. (1969).

56. Panuska, V. "A Stochastic Approximation Method for Identification of Linear System Using Adaptive Filtering." Proc. Ninth Joint Aut. Contr. Conf., University of Michigan (June 1968).

57. Sakrison, D. J. "The Use of Stochastic Approximation to Solve the System Identification Problem." IEEE Trans. Aut. Control, Vol. AC-12, 5:563-567 (October 1967).

58. Saridis, G. N. and Stein, G. "A New Algorithm for Linear System Identification." IEEE Trans. Aut. Control, Vol. AC-13, 5:592-594 (October 1968).

59. Saridis, G. N. and Stein, G. "Stochastic Approximation Algorithms for Linear Discrete-Time System Identification." IEEE Trans. Aut. Control, Vol. AC-13, 5:515-523 (October 1968).

60. Zhirogluadov, V. P. and Kaipov, V. K. "On Application of Stochastic Approximation Methods in the Identification Problem." Automn remote Control, 10:54 (1966).

Linear Least-Squares

61. Graupe, D., et. al. "Reduction and Identification of Multivariable Processes Using Regression Analysis." IEEE Trans. Aut. Control, Vol. AC-13:564-567 (1968).

62. Ho, Y. C. "The Method of Least Squares and Optimal Filtering Theory." The RAND Corp. Tech. Report RM-3329-PR, Santa Monica, Calif. (October 1962).

63. Hsia, T. C. "A Discrete Method for Parameter Identification in Linear Systems with Transport Lags." IEEE Trans. on Aerospace and Electronic Systems, Vol. AES-5, 2:236-239 (March 1969).

64. Kalman, R. E. "Design of a Self-Optimizing Control System, ASME, J. Bas. Eng., Vol. 80:468-478 (February 1958).

65. Levin, M. J. "Estimation of a System Pulse Transfer Function in the Presence of Noise." IEEE Trans. Aut. Control, Vol. AC-9:229-235 (July 1964).

66. Marquardt, D. W. "An Algorithm for Least-Squares Estimation of Non-Linear Parameters." Journal Society Industrial and Applied Mathematics, Vol. 11:431-441 (1963).

67. Panuska, V. "An Adaptive Recursive Least-Squares Identification Algorithm." Proc. IEEE Symp. on Adaptive Processes, Penn. State Univ. (1969).

68. Rubin, A. I., et. al. "Regression Analysis and Parameter Identification." Simulation, Vol. 9:39-47 (1967).

69. Steiglitz, K. and McBride, L. E. "A Technique for the Identification of Linear Systems." IEEE Trans. Aut. Control, Vol. AC-10:461-464 (October 1965).

70. Young, P. C. "On the Use of A Priori Parameter Variation Information to Enhance the Performance of a Recursive Least Squares Estimator." Naval Weapons Center Tech. Note 404-90, China Lake, Calif. (May 1969).

71. Young, P. C. "Regression Analysis and Process Parameter Estimation--A Cautionary Message." Simulation, Vol. 10:125-128 (1968)

Correlation

72. Anderson, W. N., et. al. "Consistent Estimates of the Parameters of a Linear System." Ann. of Math. Statis., Vol. 40, 6:2064-2075 (1969).

73. Farison, J. B., et. al. "Identification and Control of Linear Discrete Systems." IEEE Trans. Aut. Control, Vol. AC-12:109 (1967).

74. Geesey, R. and Kailath, T. "Comments on 'The Relationship of Alternate State-Space Representations in Linear Filtering Problems'." IEEE Trans. Aut. Control, Vol. AC-14 113-114 (February 1969).

75. Gersch, W. "Estimation of the Autoregressive Parameters of a Mixed Autoregressive Moving-Average Time Series." IEEE Trans. Aut. Control, Vol. AC-15:583-588 (October 1970).

76. Kailath, T. "An Innovations Approach to Least-Squares Estimation Part I: Filtering in Additive White Noise." IEEE Trans. Aut. Control, Vol. AC-13, 6:646-655 (December 1968).

77. Lichtenberger, W. W. "A Technique of Linear System Identification Using Correlating Filters." IRE Trans. Aut. Control, Vol. AC-6:183-199 (1961).

78. Mehra, R. K. "On-Line Identification of Linear Dynamic Systems with Applications to Kalman Filtering." IEEE Trans. Aut. Control, Vol. AC-16, 1:12-21 (February 1971).

79. Mehra, R. K. "On the Identification of Variances and Adaptive Kalman Filtering." IEEE Trans. Aut. Control, Vol. AC-15, 2:175-184 (April 1970).

80. Parzen, E. Multiple Time Series Modeling." Stanford University Tech. Report 12 (1968).

81. Rao, M. M. "Consistency and Limit Distributions of Estimators of Parameters in Explosive Stochastic Difference Equations." Ann. Math. Statis. Vol. 32:195-218 (1961).

82. Walker, A. M. "Large-Sample Estimation of Parameters for Autoregressive Processes with Moving Average Residuals, Parts I, II." Biometrika, Vol. 49:117-131 (1962).

General

Input Signal Synthesis

83. Aoki, M. and Staley, R. M. "On Input Signal Synthesis in Parameter Identification." Automatica, Vol. 6:431-440 (1970).

84. Gagliardi, R. M. "Input Selections for Parameter Identification in Discrete Systems." IEEE Trans. Aut. Control, Vol. AC-12, No. 5 (October 1967).

85. Levadi, V. S. "Design of Input Signals for Parameter Estimation." IEEE Trans. Aut. Control, Vol. AC-11, 2:205-211 (April 1966).

86. Litman, S. and Huggins, W. H. "Growing Exponentials as a Probing Signal for System Identification." Proc. IEEE (January 1963).

87. Nahi, N. E. "Optimal Inputs for Parameter Estimation in Dynamic Systems with White Observation Noise." Proc. Joint Aut. Contr. Conf., pp. 506-516 (1969).

88. Turin, G. L. "On the Estimation in the Presence of Noise of the Impulse Response of a Random Linear Filter." I.R.E. Trans. Inf. Theory, Vol. IT-3:5-10 (March 1967).

Canonical Forms

89. Bucy, R. S. "Canonical Forms for Multivariable Systems." IEEE Trans. Aut. Control, Vol. AC-13: 567-569 (1968).

90. Johnson, C. D. and Wonham, W. M. "A Note on the Transformation to Canonical (Phase-Variable) Form." IEEE Trans. Aut. Control, Vol. AC-9 :312-313 (July 1964).

91. Kalman, R. E. "Mathematic Description of Linear Systems." Journal SIAM Control, Ser. A, Vol. 1, 2:152-192 (1963).

92. Luenberger, D. G. "Canonical Forms for Linear Multivariable Systems." IEEE Trans. Aut. Control, Vol. AC-12:290-293 (1967).

93. Ramaswami, B. and Ramar, K. "Transformation to the Phase Variable Canonical Form." IEEE Trans. Aut. Control, Vol. AC-13:746-747 (December 1968).

94. Tuel, W. G. "Canonical Forms for Linear Systems I and II." IBM Research Report R. J. 375 and 390 (1966).

Kalman Filtering

95. Deyst, J. J. and Price, C. F. "Conditions for Asymptotic Stability of the Discrete Minimum-Variance Linear Estimator." IEEE Trans. Aut. Control, Vol. AC-13:702-705 (December 1968).

96. Friedland, B. "Treatment of Bias in Recursive Filtering." IEEE Trans. Aut. Control, Vol. AC-14, 4:359-367 (August 1969).

97. Heffes, H. "The Effect of Erroneous Models on the Kalman Filter Response." IEEE Trans. Aut. Control, Vol. AC-11: 541-543 (July 1966).

98. Huddle, J. R. and Wismer, D. A. "Degradation of Linear Filter Performance Due to Modeling Error." IEEE Trans. Aut. Control, Vol. AC-13:421-423 (August 1968).

99. Kalman, R. E. "A New Approach to Linear Filtering and Prediction Problems." Trans. ASME Journal Basic Engineering, Vol. 82:34-45 (March 1960).

100. Kalman, R. E. "New Methods and Results in Linear Prediction and Filtering Theory." Proc. Symp. on Eng. Applications of Random Function Theory and Probability at Purdue University (November 1960).

101. Kalman, R. E. and Bucy, R. S. "New Results in Linear Filtering and Prediction Theory." Trans. ASME Journal Basic Engineering, Ser. D, Vol. 83:95-107 (December 1961).

102. Neal, S. R. "Linear Estimation in the Presence of Errors in Assumed Plant Dynamics." IEEE Trans. Aut. Control, Vol. AC-12:592-594 (October 1967).

103. Price, C. F. "An Analysis of the Divergence Problem in the Kalman Filter." IEEE Trans. Aut. Control, Vol. AC-13:699-701 (December 1968).

Applications

104. Buel, J. D. and Kalaba, R. E. "Quasilinearization and the Fitting of Non-Linear Models of Drug Metabolism to Experimental Kinetic Data." Math. Biosci., Vol. 5:121-132 (1969).

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

106. Griliches, Z. "Distributed Lags: A Survey." Econometrica, Vol. 35:16-49 (1967).

107. Habegger, L. J. and Bailey, R. E. "Minimum Variance Estimation of Parameters and States in Nuclear Power Systems." Proc. Fourth IFAC Congress, Warsaw, Poland (1969).

108. Harris, A. The Effect of Various Pilot-Model Approximations on the Design of a Pitch-Attitude Stability Augmentation System for a Hovering V/STOL Aircraft, M.S. Thesis, Dept. of Elect. Eng., Ohio State University (1969).

109. Henderson, J. M. and Quandt, R. E. Microeconomic Theory, McGraw-Hill Book Co., New York (1958).

110. Sage, A. P. and Masters, G. W. "Identification and Modeling of States and Parameters of Nuclear Reactor Systems." IEEE Trans. Nuclear Sci. (February 1967).

111. Sastry, V. A. and Vetter, W. J. "The Application of Identification Methods to an Industrial Process." Preprints Joint Aut Contr. Conf., pp. 787-794 (1969).

Use of A Priori Information

112. Cox, H. "Sequential Minimax Estimation." IEEE Trans. Aut. Control, Vol. AC-11:323-324 (1966).

113. Denery, D. G. "An Identification Algorithm Which is Insensitive to Initial Parameter Estimates." AIAA Paper 70-34, New York (1970).

114. Ho, Y. C. and Lee, R. C. K. "A Bayesian Approach to Problems in Stochastic Estimation and Control." IEEE Trans. Aut. Control, Vol. AC-9:333-339 (October 1964).

115. Klinger, A. "Prior Estimation and Bias in Sequential Estimation." IEEE Trans. Aut. Control, IEEE Trans. Aut. Control, Vol. AC-13:102-103 (1968).

116. Lin, T. T. and Yau, S. S. "Bayesian Approach to the Optimization of Adaptive Systems." IEEE Trans. Sys. Sci. and Cyb., Vol. 3:77 (1967).

117. Nishimura, T. "On the A Priori Information in Sequential Estimation Problems." IEEE Trans. Aut. Control, Vol. AC-11, 2:197-204 (April 1966).

118. Smith, G. L. "Sequential Estimation of Observation Error Variances in a Trajectory Estimation Problem." AIAA Jour., Vol. 5:11 (1967).

119. Soong, T. T. "On A Priori Statistics in Minimum Variance Estimation Problems." Trans. ASME Journal Basic Eng., Ser. D, Vol. 87:109-112 (March 1965).

120. Tarn, T. J. "Stochastic Optimal Control with Noisy Observation Having Imperfectly Known Plant Disturbances." Joint Aut. Contr. Conf., Boulder, Colorado (1969).

Other Identification Methods

121. Balakrishnan, A. V. "A New Computing Technique in System Identification." Journal Comp. Syst. Sci., Vol. 2, No. 1 (1968).

122. Balakrishnan, A. V. "Stochastic System Identification Techniques." Chapter in Stochastic Optimization and Central., John Wiley, New York (1968).

123. Clement, P. R. and Smith, G. W. "Identification of Linear System Dynamics Using Exponential Spectral Coefficients." Journal Franklin Inst., Vol. 281:122-135 (1966).

124. Diskind, T. "Parameter Evaluation Using Complex Domain Analysis." Preprint Joint Aut. Contr. Conf., pp. 770-781 (1969).

125. Hilton, W. B. and Smith, F. W. "Monte Carlo Evaluation of Methods for Pulse Transfer Function Estimation." IEEE Trans. Aut. Control, Vol. AC-12:568-576 (1967).

126. Jacob, J. P. and Zadeh, L. A. "A Computer-Oriented Identification Technique." Preprint Joint Aut. Contr. Conf., pp. 979-981 (1969).

127. Kagiwada, H., et. al. Identification of Linear Systems via Numerical Inversion of Laplace Transforms, RAND Corp., RM-4262-PR (August 1964).

128. Kushner, H. J. "A Simple Iterative Procedure for the Identification of the Unknown Parameters of a Linear Time-Varying Discrete System." Trans. ASME Journal Basic Eng., pp. 227-235 (June 1963).

129. McGhee, R. B. and Walford, R. B. "A Monte Carlo Approach to the Evaluation of Conditional Expectation Parameter Estimates for Non-Linear Dynamic Systems." IEEE Trans. Aut. Control, Vol. AC-13:29-37 (1968).

130. Mintz, M. L. and Thorp, J. S. "A Method for Discrete System Identification Approximation." Journal Franklin Institute, Vol. 283:389-403 (1966).

131. Oza, K. G. and Jury, E. I. "System Identification and the Principle of Contraction Mapping." Journal SIAM Control, Vol. 6:244-257 (1968).

132. Pazdera, J. S. and Pottinger, H. J. "Linear System Identification via Liapunov Design Techniques." Preprint Joint Aut. Contr. Conf., pp. 795-801 (1969).

133. Rose, R. E. and Lance, G. M. "An Approximate Steepest Descent Method for Parameter Identification." Proc. Joint Aut. Contr. Conf., pp. 483-487 (1969).

134. Rowe, I. H. "A Bootstrap Method for the Statistical Estimation of Model Parameters." Preprint Joint Aut. Contr. Conf., pp. 989-987 (1969).

135. Shultz, E. R. "Estimation of Pulse Transfer Function Parameters by Quasilinearization." IEEE Trans. Aut. Control, Vol. AC-13:424-426 (1968).

136. Silverman, H. Identification of Linear Systems Using Fast Fourier Transform Technique. Ph.D. Dissertation, Brown Univ. (1970).

137. Smith, F. W. "System Laplace Transform Estimation from Sampled Data." IEEE Trans. Aut. Control, Vol. AC-13: 37-44 (February 1968).

138. Young, P. C. "Identification Problems Associated with the Equation Error Approach to Process Parameter Estimation." Proc. Second Asilomar Conf. on Circuits and Systems, pp. 416-422 (October 1968).

139. Young, P. C. "Parameter Estimation and the Method of Conjugate Gradients." Proc. IEEE, Vol. 54 (1966).

140. Young, P. C. "Recursive Methods of Parameter Estimation and Their Application to Process Identification: Parts I and II." Control Engineering, Vol. 16, Nos. 10 and 11 (1969).

General System Identifiability

141. Astrom, K. J. "Computational Aspects of a Class of Identification Problems." Report, Dept. of EE, University of Southern Calif. (1969).

142. Balakrishnan, A. V. and Peterka, V. "Identification in Automatic Control Systems." Automatica, Vol. 5:817-829 (1969).

143. Bellman, R. and Astrom, K. J. "On Structural Identifiability." Math. Biosci., Vol. 1:329-339 (1969).

144. Bellman, R., et. al. "Quasilinearization, System Identification, and Prediction." Memorandum RM-3812-PR, RAND Corp., Santa Monica, Calif. (1963).

145. Borisova, R. W. and Ulanov, G. M. "Evaluation of Parameters of Stationary Controlled Plants in the Identification Problem." Autom Remote Control, No. 11 (1966).

146. Cheruy, J. and Menedez, A. "A Comparison of Three Dynamic Identification Methods." Preprint Joint Aut. Contr. Conf., pp. 982-983 (1969).

147. Choate, W. C. and Sage, A. P. "Minimum Time Identification of Non-Stationary Dynamic Processes." Proc. Nat. Electronics Conf. (1965).

148. Cox, H. "On the Estimation of State Variables and Parameters for Noisy Dynamic Systems." IEEE Trans. Aut. Control, Vol. AC-9:5-12 (January 1964).

149. Davisson, L. D. "The Prediction of Stationary Gaussian Time-Series of Unknown Covariances." IEEE Trans. Inf. Theory, Vol. IT-11:527-532 (October 1965).

150. Eisenberg, B. R. and Sage, A. P. "Experiments in Non-Linear and Non-Stationary System Identification via Quasilinearization and Differential Approximation." Proc. Joint Aut. Contr. Conf. (August 1966).

151. Eykhoff, P. "Process Parameter and State Estimation." Automatica, Vol. 4:205-233 (1968).

152. Eykhoff, P. "Some Fundamental Aspects of Process Parameter Estimation." IEEE Trans. Aut. Control, Vol. AC-8:347-357 (1963).

153. Galtieri, C. A. The Problem of Identification of Discrete Time Processes, Report, Electronics Research Lab., University of Calif., Berkeley (1963).

154. Gluss, B., et. al. Identification of Differential Systems with Time-Varying Coefficients, RAND Corp., RM-4288-PR (November 1964).

155. Goodman, T. P. and Reswick, J. B. "Determination of System Characteristics from Normal Operating Records." Trans. ASME, Vol. 78:259-271 (1956).

156. Isermann, R. "New Results on the Identification of Processes." Automatica, Vol. 7:191-197 (1971).

157. Kushner, H. J. "On the Convergence of Lion's Identification Method with Random Inputs." IEEE Trans. Aut. Control, Vol. AC-15:652-654 (December 1970).

158. Lance, G. M. and Rose, R. E. "Analysis of the Behavior of Continuous Parameter Estimation Methods." Preprint Joint Aut. Contr. Conf., Univ. of Michigan (1968).

159. Lavi, A. and Strauss, J. C. "Parameter Identification in Continuous Dynamic Systems." IEEE International Convention Record, Symposium on Automatic Control; Systems Science; Cybernetics; Human Factors, Part IV, pp. 49-61 (1965).

160. Lee, R. C. K. Optimal Estimation, Identification, and Control. The MIT Press, Cambridge, Mass. (1964).

161. Lendaris, G. G. "The Identification of Linear Systems." Trans. AIEE (Applications and Industry), Vol. 81:231-240 (September 1962).

162. Lindenlaub, J. C. and Cooper, G. R. "Noise Limitations of System Identification Techniques." IEEE Trans. Aut. Control, Vol. AC-8:43-48 (1963).

163. Maslov, E. P. "On One Estimate of Identification Accuracy." Automn remote Control, Vol. 10:58-61 (1966).

164. Masters, G. W. and Sage, A. P. "On-Line Estimation-States and Parameters for Discrete Non-Linear Dynamic Systems." Proc. National Electronics Conf. No. 22 (1966).

165. Mayne, D. Q. "Optimal Non-Stationary Estimation of the Parameters of a Linear System with Gaussian Inputs." Journal Electronics Control, Vol. 14:101-112 (1963).

166. Mayne, D. Q. "Parameter Estimation." Automatica, Vol. 3:245-255 (January 1966).

167. Pearson, J. D. "Deterministic Aspects of the Identification of Linear Discrete Dynamic Systems." IEEE Trans. Aut. Control, Vol. AC-12:764-766 (1967).

168. Rauch, H. E. "Linear Estimation of Sampled Stochastic Processes with Unknown Parameters." Stanford Electronics Lab., TR 2108-1, Stanford University (April 1962).

169. Schwartz, S. C. and Steiglitz, K. "The Identification and Control of Unknown Linear Discrete Systems." Report 24, Information Sciences and Systems Laboratory, Dept. EE, Princeton Univ. (1968).

170. Shidhar R. and Kumar, K. S. "On the Identification of Control Systems by the Quasilinearization Method." IEEE Trans. Aut. Control, Vol. AC-9, No. 2 (1964).

171. Smith, O. J. M. and Eykhoff, P. "Optimizing Control with Process Dynamic Identification." IRE Trans. Aut. Control, Vol. AC-7:140-155 (1962).

172. Staley, R. M. and Yue, P. C. "On System Parameter Identifiability." Information Science, Vol. 2, No. 1 (1970).

173. Surber, W. H. and Kerr, R. B. "Precision and Impulse-Response Identification Based on Short, Normal Operating Records." IEEE Trans. Aut. Control, Vol. AC-6:173-182 (1961).

Other

174. Anderson, T. W. "The Choice of the Degree of a Polynomial Regression as a Multiple Decision Problem." Ann. Math. Statis., Vol. 33:255-265 (1962).

175. Bartlett, M. S. An Introduction to Stochastic Processes. Cambridge Univ. Press, London (1962).

176. Birkhoff, G. and MacLane, S. A Survey of Modern Algebra, Rev. ed., Macmillan Book Company, New York (1953).

177. DeRusso, P. M., et. al. State Variables for Engineers, Wiley Book Co., New York (1966).

178. Egorov, C. V. "The Way of Determination of Complex Plants Dynamic Characteristics." Automn remote Control, Vol. 12:37-46 (1966).

179. Kailath, T. "A General Likelihood-Ratio Formula for Random Signals in Gaussian Noise." IEEE Trans. Inf. Theory, Vol. IT-15, 3:350-361 (May 1969).

180. Meditch, J. S. Stochastic Optimal Estimation and Control. New York: McGraw-Hill Book Co. (1969).

181. Mehra, R. K. "An Algorithm to Solve Matrix Equations $PH' = G$ and $P = \Phi P \Phi' + \Gamma\Gamma'$." IEEE Trans. Aut. Control, Vol. AC-15, p. 600 (1970).

182. Truxal, J. G. Automatic Feedback Control System Synthesis, McGraw-Hill Book Co., New York (1956).

183. Wald, A. and Mann, H. B. "On the Statistical Treatment of Linear Stochastic Difference Equations." Econometrica, Vol. 11, Nos. 3 and 4, pp. 173-220 (July-October 1943).

184. Wilkie, D. F. and Perkins, W. R. "Essential Parameters in Sensitivity Analysis." Automatica, Vol. 5, No. 2 (March 1969).

185. Zadeh, L. A. "From Circuit Theory to System Theory." Proc. IRE, Vol. 50:856-865 (1962).

Vita

Adolph Harris was born on 22 November 1941 in Washington, Georgia. He was graduated from high school in Cincinnati, Ohio, in 1959. He attended the University of Cincinnati, where he was elected to Tau Beta Pi and participated in the co-operative work-study program at Aeronautical Systems Division (ASD), Wright-Patterson AFB, Ohio. After graduating from the University of Cincinnati in 1964 with the Bachelor of Science degree in Aerospace Engineering, he returned to ASD where he worked as an aerospace engineer in the Advanced Systems Design Division of the Deputy for Development Planning. While so employed, he attended the Ohio State University part-time, receiving the Master of Science degree in Electrical Engineering in 1969. He is a registered Professional Engineer, and a member of IEEE and AIAA. He was selected to attend the Graduate Systems Analysis (GSA-72) program starting in June 1970 at the Air Force Institute of Technology, Wright-Patterson AFB, Ohio, from where he received the degree of Master of Science in Systems Analysis in March 1972.

Permanent Address: 1124 Wilson Drive
Dayton, Ohio 45407

This thesis was typed by Jane Manemann.