

In presenting the dissertation as a partial fulfillment of the requirements for an advanced degree from the Georgia Institute of Technology, I agree that the Library of the Institute shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to copy from, or to publish from, this dissertation may be granted by the professor under whose direction it was written, or, in his absence, by the Dean of the Graduate Division when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. It is understood that any copying from, or publication of, this dissertation which involves potential financial gain will not be allowed without written permission.

[Handwritten signature]

7/25/68

ESTIMATION OF DISCRETE SIGNALS
CONTAINING A NONRANDOM COMPONENT

A THESIS

Presented to

The Faculty of the Graduate Division

by

Norman Dale Crump

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
in the School of Electrical Engineering

Georgia Institute of Technology

June, 1969

ESTIMATION OF DISCRETE SIGNALS
CONTAINING A NONRANDOM COMPONENT

Approved:

Chairman

J. G. Proakis

Date approved by Chairman: *June 2, 1969*

ACKNOWLEDGMENTS

Grateful acknowledgment is extended to my advisor, Dr. Roger P. Webb. His guidance and assistance in the preparation of this dissertation are deeply appreciated. I would also like to express my appreciation to Dr. J. L. Hammond and Dr. A. M. Bush for their time and service as members of my reading committee.

My thanks go to Dr. B. J. Dasher and the School of Electrical Engineering for providing me with financial assistance in the form of an NDEA fellowship and a DuPont fellowship. I am also grateful for the rewarding teaching experiences which I received as a graduate teaching assistant.

My deepest gratitude goes to my wife, Paula, for her faithful support, patience, and encouragement during the long effort of this study.

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS.	ii
LIST OF ILLUSTRATIONS.	iv
SUMMARY.	v
Chapter	
I. INTRODUCTION.	1
Statement of the Problem	
History of the Problem	
II. OPTIMUM DISCRETE SIGNAL ESTIMATION— THE KALMAN ESTIMATOR.	9
The Kalman Estimator Equations	
The Sensitivity Problem	
Modifications of the Kalman Estimator	
III. INTERPRETATIONS OF DISCRETE SIGNAL ESTIMATION	39
Least Squares Curve Fitting and the Kalman Estimator	
An Alternative Model Approach	
IV. THE AUGMENTED MEMORY APPROACH TO DISCRETE SIGNAL ESTIMATION.	60
The Augmented Memory Estimator	
Computation of Weighting Coefficients	
Interpretations of the Augmented Memory Estimator	
V. NUMERICAL COMPARISONS AND EXAMPLES.	85
VI. CONCLUSIONS AND RECOMMENDATIONS	101
APPENDIX	106
BIBLIOGRAPHY	123
VITA	126

LIST OF ILLUSTRATIONS

Figure	Page
1. The Kalman Estimator.	14
2. Kalman Estimator Error, Second Degree Polynomial Signal with First Degree Polynomial Model	19
3. Configuration for Re-entry Tracking Problem	20
4. Re-entry Signal Trajectories.	22
5. Kalman Estimator Error, Re-entry Trajectory with Polynomial Range Model	24
6. Kalman Estimator Error, Re-entry Trajectory with Inaccurate Drag Constant in Signal Model	27
7. Flow Diagram of the Augmented Memory Estimator.	71
8. Mean-Square Error Improvement in the Augmented Memory Estimator with Second Degree Polynomial Signal	87
9. Mean-Square Error Improvement in the Augmented Memory Estimator with First Degree Polynomial Signal.	87
10. Average Error of Augmented Memory Estimator, Second Degree Polynomial Signal with First Degree Polynomial Model.	90
11. Orbit Altitude Signal	93
12. Average Error of Augmented Memory Estimator, Orbit Altitude Signal with Second Degree Polynomial Model.	95
13. Average Error of Augmented Memory Estimator, Orbit Altitude Signal with First Degree Polynomial Model.	96
14. Average Error of the Augmented Memory Estimator, Re-entry Signal with Second Degree Polynomial Model	98
15. Average Error of the Augmented Memory Estimator, Re-entry Signal with First Degree Polynomial Model.	99

SUMMARY

This dissertation is concerned with the problem of estimating a signal contained in noisy discrete measurements. The signals considered are those which may be expressed as the sum of a random noise component and a nonrandom component. The nonrandom component is assumed to consist of the sum of known functions of time with unknown coefficients. The random component of the signal and the measurement noise are assumed to have autocorrelation and cross-correlation functions which are known at the sampling instants. This study is restricted to the consideration of linear realizable discrete estimators. Particular emphasis is on the sensitivity of linear estimators to errors in the assumed model of the nonrandom component of the signal. A new estimation technique is presented which is shown to have a combination of mean-square error and modeling error sensitivity properties which represent an improvement over currently available methods.

The discrete Kalman estimator is known to minimize mean-square estimation error for the class of signals considered in this dissertation. A detailed discussion of the Kalman estimator is given, with emphasis on the sensitivity of the estimates to errors in the model of the nonrandom characteristic of the signal. Various modifications of the Kalman estimator which attack the sensitivity problem are discussed. An interpretation of the Kalman estimator as a weighted least squares curve fitting procedure is used to relate the signal model of the Kalman estimator to an alternative modeling technique which is more

suitable for implementing a fixed memory estimator. The fixed memory estimator computes estimates based on data obtained over a fixed length of time immediately preceding the most recent measurement. It is shown that the fixed memory estimator has an advantage over the Kalman estimator in its relative insensitivity to errors in the signal dynamical model.

A new estimation technique, called the augmented memory estimator, is formulated with the objective of computing estimates which are relatively insensitive to errors in the signal dynamical model while making use of all available data. The augmented memory estimator uses a computational structure similar to that of the fixed memory estimator, but with a fixed number of past estimates retained rather than past measurements. The resulting estimate is linear in all past data but has sensitivity properties comparable to those of the fixed memory estimator. The mean-square error which results from the use of an accurate signal model lies between that of the Kalman and the fixed memory estimators.

The mean-square error and the sensitivity properties of the augmented memory estimator are illustrated by digital computer simulations. Linear models for nonlinear signals are included in the examples.

CHAPTER I

INTRODUCTION

The purpose of this research is to develop a new technique for estimating signals contained in discrete noisy measurements. The class of signals considered consists of signals containing both a nonrandom component and a random noise component. A suboptimal linear estimator structure is used so as to obtain estimates which are relatively insensitive to errors in the assumed model of the signal dynamical characteristics. The uniqueness of the technique to be developed is in the method used to combine all available measurement data in such a way that errors due to an inaccurate signal model do not accumulate. Subject to the constraints imposed by the structure of the estimator and the unbiased property imposed on the estimate, the estimator variables are chosen to minimize mean-square estimation error.

Signals and measurements having the assumed properties arise in controls and communications applications. For example, the trajectory of an airborne vehicle has a deterministic nature which may be altered by random noise forcing functions or disturbances. Discrete time radar measurements of such a trajectory contain an additional random noise component due to the limitations of the equipment. The problem then may consist of estimating the true position of the vehicle at the time of the most recent measurement.

While methods exist for optimally combining all past measurement

data so as to give estimates which minimize mean-square error for a large class of signals, the additional consideration of the sensitivity of the estimates to errors in the assumed model of the nonrandom component of the signal may be very important. In fact, in the presence of errors in the signal dynamical model, a suboptimal method of computing estimates can give considerably improved performance over the theoretically optimum method for the model used. The errors in the signal dynamical model may be of an unknown nature or may be intentional inaccuracies imposed to reduce the complexity of the model.

Statement of the Problem

The signal estimation problem consists of determining methods for combining measurements to obtain an estimate of the desired signal in a way that will extremize some criterion of effectiveness, or cost function. The estimate computed is an estimate of the signal at the time of the most recent measurement. In addition to cost function minimization, other constraints may be imposed on the estimator in order to satisfy certain performance criteria.

Various types of discrete estimators differ primarily with respect to four initial assumptions: the model used by the estimator for the structure of the signal generator, information assumed known about the statistics of noise sources in the signal and measurement models, the cost function to be minimized, and assumptions or constraints on the structure of the estimator. An estimator which is optimum with respect to one combination of these four initial assumptions is not in general optimum with respect to any other combination.

The class of signals considered in this dissertation consists of signals which contain the sum of a deterministic, or nonrandom, component and a random noise component. The nonrandom component is assumed to be expressible as an unknown linear combination of known functions of time, as follows:

$$x(t) = \underline{a}^T \underline{p}(t) + r(t) \quad (1.1)$$

where \underline{a} is an n -vector of unknown coefficients, $\underline{p}(t)$ is an n -vector of known functions of time, and $r(t)$ is the random noise component of the signal $x(t)$. The $\underline{p}(t)$ vector of time functions constitutes the model of the dynamic characteristics of the signal. Measurements are assumed to consist of discrete samples of the signal with additive measurement noise, as follows:

$$y(kT) = x(kT) + v(kT) \quad (1.2)$$

where T is the sample period, k is the sample number, and $v(kT)$ is random measurement noise. It is assumed that the autocorrelation functions and the cross-correlation function of $r(t)$ and $v(t)$ are known at the sampling instants. Stationarity is not a requirement for either noise process.

The cost function used in this study is mean-square estimation error taken at the time of the estimate, so the problem consists of minimizing

$$J(kT) = E\{[x(kT) - \hat{x}(kT)]^2\} \quad (1.3)$$

for each succeeding value of k , where $\hat{x}(kT)$ is the estimate of $x(kT)$.

Estimators considered in this dissertation are assumed to be linear and realizable, meaning that each estimate is a linear combination of past and present measurements only. It is desirable to include the requirement that the estimate is an unbiased estimate of $x(kT)$, so that

$$E[x(kT) - \hat{x}(kT)] = 0 \quad (1.4)$$

It will be shown that additional constraints on the linear estimator structure will lead to an estimation procedure which is less sensitive to errors in the model of the signal dynamic characteristics than the optimum, unbiased, unconstrained linear estimator.

History of the Problem

The estimation of a signal imbedded in measurement noise has been a subject of research for many years. The most well known of the early work is that of Wiener (1)^{*}, who derived minimum mean-square error continuous filters for smoothing, estimating, and predicting a signal in a noisy measurement. Both signal and measurement noise are assumed to be sample functions of stationary random processes with both auto-correlation functions and the cross-correlation function known.

^{*} Numbers in parentheses following a citation refer to items in the Bibliography.

The resulting estimator is the output of a linear system described by its transfer function or impulse response. Wiener's methods have also been extended to discrete signal estimation, as in Freeman's (2) work.

The signal estimation problem was solved by state variable techniques by Kalman (3) and Kalman and Bucy (4). In the state variable formulation the signal is a state vector which is describable as the output of a linear system forced by a white noise input. Measurements are described by the sum of a linear transformation on the signal state vector and a measurement noise vector. The signal noise and measurement noise processes are assumed to be independent zero mean white Gaussian noise processes with known variance. The state variable formulation can be used to describe signals containing both random and nonrandom components, so the Kalman estimator is applicable to a large class of signals. The Kalman estimator for discrete measurements and discrete time signals containing a nonrandom component is of primary importance in this dissertation, and is discussed in more detail in Chapter II, particularly with respect to the sensitivity of the estimates to errors in the signal dynamical model.

Both the Wiener and Kalman estimators are growing memory linear estimators, in the sense that as new estimates are generated they are expressible as a linear combination of an increasing quantity of measurement data. Much research has also been done with finite, fixed memory linear estimators which generate estimates expressible as a linear combination of measurement data taken over a fixed period of time up to and including the time at which the estimate is being computed. This type of estimator is useful for the estimation of

signals containing a nonrandom component because it tends to be less sensitive to errors in the model of the nonrandom component than is a growing memory estimator.

A fixed memory linear estimator for the estimation of a signal containing a nonrandom component was derived by Zadeh and Ragazzini (5). They derived equations for obtaining the transfer function of a fixed memory continuous filter which minimizes mean-square error. The filter input is continuous and consists of the signal and measurement noise. The signal is assumed to consist of the sum of a random noise component and a polynomial of known degree with unknown coefficients. Blum (6) extended Zadeh and Ragazzini's method to include a nonrandom signal component consisting of an unknown linear combination of any finite set of known functions of time. With the same signal model as that of Zadeh and Ragazzini, Lees (7) and Johnson (8) each extended the continuous filter to the discrete measurement case. Blum (9) and Hsieh (10) each extended the discrete fixed memory estimator to include signals with a nonrandom component consisting of an unknown linear combination of any finite set of known, linearly independent, functions of time. Each of these solutions to the fixed memory discrete estimation problem consists of determining weights for the most recent measurement and the measurements contained in the fixed length time interval immediately preceding the most recent measurement. The weights are chosen to minimize mean-square error in the estimate of the signal at the time of the most recent measurement. The weights also must satisfy constraints imposed by requiring the estimate to be unbiased.

Attempts at using fixed memory estimators to estimate signals having both random and nonrandom components have not been limited to the use of the mean-square error cost function. For example, Johanson (11) derived equations for computing the transfer function of a continuous fixed memory estimator which minimizes the maximum expected squared error. The signal is assumed to consist of the sum of a noise component and a polynomial of known degree, with a known bound on the highest order nonzero derivative of the polynomial. The estimator in this method is constrained to be distortionless, meaning that for noise-free measurements of a polynomial signal only, the estimate is without error. Zahl (12) derived equations for computing weights for a fixed memory linear discrete signal estimator which minimizes the maximum expected squared error. In this method, the signal model is the sum of a random component and a nonrandom component which is any known function of a vector of unknown constants. The vector of unknown constants is assumed to be contained in a known set of vectors, which could be either a finite or an infinite set. This approach results in a quadratic programming problem which generally must be solved by iterative methods.

The linear discrete signal estimators described above, and others in the literature, all result in an estimate which is expressible as a weighted sum of past and present measurements. The various methods mentioned are separated as to growing memory or fixed memory estimator structures. This separation is exploited in later chapters of this dissertation as a means of comparing sensitivity of various discrete estimation techniques to errors in the assumed model of the nonrandom

characteristic of the signal. Existing approaches to this sensitivity problem are discussed in detail in Chapter II.

CHAPTER II

OPTIMUM DISCRETE SIGNAL ESTIMATION—

THE KALMAN ESTIMATOR

In Chapter I it was stated that Kalman (3) presented a particularly compact solution to the signal estimation problem for a large class of signals. This chapter will examine the discrete Kalman estimator in more detail. Of particular importance in this dissertation is the behavior of the Kalman estimator in the presence of inaccuracies in the assumed model of signals having a nonrandom component. Various techniques for correcting the adverse effects of such errors will be presented and discussed.

The emphasis of this chapter on the Kalman estimator as opposed to other estimation techniques is warranted for several reasons. First, the Kalman estimator minimizes mean-square error for a larger class of signals than most methods. Also, state variable formulations of discrete signal systems are compact and especially suitable for digital signal processing. The assumptions regarding knowledge of noise statistics in the Kalman estimator correspond to those of most other linear estimators, including the new method presented in this dissertation. Finally, since there is only one linear estimate which minimizes mean-square error for a given class of signals, the estimation error behavior of the Kalman estimator will be the same as the error behavior of any other method which minimizes mean-square error subject to the same assumptions.

The Kalman Estimator Equations

In the Kalman estimator formulation, the signal to be estimated is described recursively by

$$\underline{x}(kT) = \Phi(kT, (k-1)T)\underline{x}((k-1)T) + G((k-1)T)\underline{u}((k-1)T) \quad (2.1)$$

where $\underline{x}(kT)$ is the signal state vector of dimension n , $\Phi(kT, (k-1)T)$ is the state transition matrix, $G((k-1)T)$ is a known multiplying matrix, $\underline{u}((k-1)T)$ is a zero mean white noise process with a known covariance matrix, and T is the sampling period. Measurements consist of the sum of measurement noise and a linear product of the state vector:

$$\underline{y}(kT) = H(kT)\underline{x}(kT) + \underline{v}(kT) \quad (2.2)$$

where $\underline{y}(kT)$ is the measurement vector of dimension m , $H(kT)$ is a known measurement matrix, and $\underline{v}(kT)$ is zero mean white noise with a known covariance matrix. In the remainder of this dissertation, it will be assumed that the sampling period is normalized to $T=1$ second, so that the signal and measurement models become

$$\underline{x}(k) = \Phi(k, k-1)\underline{x}(k-1) + G(k-1)\underline{u}(k-1) \quad (2.3)$$

$$\underline{y}(k) = H(k)\underline{x}(k) + \underline{v}(k) \quad (2.4)$$

The white noise covariance matrices are given by

$$E[\underline{u}(k)\underline{u}^T(j)] = \begin{cases} R(k), & k=j \\ [0], & k \neq j \end{cases} \quad (2.5)$$

$$E[\underline{v}(k)\underline{v}^T(j)] = \begin{cases} V(k), & k=j \\ [0], & k \neq j \end{cases} \quad (2.6)$$

The measurement noise and signal noise are assumed to be uncorrelated, so that

$$E[u_i(k)v_q(j)] = 0 \quad (2.7)$$

for all k , j , i , and q . Bryson and Johanson (13) have derived a method for handling colored noise processes with the Kalman estimator, but the white noise restriction is not important to the problem considered in this dissertation.

The problem solved by the Kalman estimator consists of determining the linear estimator which minimizes mean-square error $P(k)$ at time k , based on measurements up to and including $\underline{y}(k)$. Mean-square error is minimized for the entire state vector, so

$$P(k) = E\{[\underline{x}(k) - \hat{\underline{x}}(k)][\underline{x}(k) - \hat{\underline{x}}(k)]^T\} = \text{minimum} \quad (2.8)$$

where $\hat{\underline{x}}(k)$ is the estimate of $\underline{x}(k)$. Kalman showed that the estimate which minimizes mean-square error for signals representable in the form

of Equations (2.3, 2.4), when $\underline{u}(k)$ and $\underline{v}(k)$ are white Gaussian noise processes, can be written in the form

$$\begin{aligned} \hat{\underline{x}}(k) = & \Phi(k,k-1)\hat{\underline{x}}(k-1) \\ & + K(k)[\underline{y}(k) - H(k)\Phi(k,k-1)\hat{\underline{x}}(k-1)] \end{aligned} \quad (2.9)$$

The matrix $K(k)$, defined as the "gain" of the estimator, is computed by

$$K(k) = P'(k)H^T(k)[H(k)P'(k)H^T(k) + V(k)]^{-1} \quad (2.10)$$

and

$$P'(k) = \Phi(k,k-1)P(k-1)\Phi^T(k,k-1) \quad (2.11)$$

$$+ G(k-1)R(k-1)G^T(k-1)$$

$$P(k) = [I - K(k)H(k)]P'(k)[I - K(k)H(k)]^T \quad (2.12)$$

$$+ K(k)V(k)K^T(k)$$

where I is the unit matrix, $P'(k)$ is considered defined by Equation (2.11), and the other terms are as previously defined. Equations (2.9) through (2.12) constitute the Kalman estimator for the signal and measurement models of Equations (2.3) through (2.7). Inspection of Equation (2.9) shows that the Kalman estimator generates the optimum

estimate as the sum of two terms. The first term, $\Phi(k,k-1)\hat{\underline{x}}(k-1)$, is the best estimate of $\underline{x}(k)$ based on all measurements prior to $\underline{y}(k)$. The second term consists of the product of the gain $K(k)$ and the difference of the latest measurement $\underline{y}(k)$ and the best estimate of $\underline{y}(k)$ before it is obtained. Another way to interpret the estimator is as a weighted sum of the latest measurement $\underline{y}(k)$ and the previous estimate $\hat{\underline{x}}(k-1)$. These interpretations will be discussed in greater detail in Chapter III.

The Kalman estimator must be initialized by using a priori information to assign values to $\hat{\underline{x}}(0)$ and $P(0)$. These quantities should be chosen to be

$$\hat{\underline{x}}(0) = E[\underline{x}(0)] \quad (2.13)$$

$$P(0) = E[\underline{x}(0)\underline{x}^T(0)] \quad (2.14)$$

to insure optimum performance.

A block diagram representation of the signal and estimator equations is shown in Figure 1. This figure illustrates the feedback structure of the estimator, the presence of the signal model in the estimator, and the presence of the gain, $K(k)$, in the estimator structure.

The Sensitivity Problem

As shown in Figure 1, the Kalman estimator structure contains a model of the signal generating system. When the signal has a nonrandom

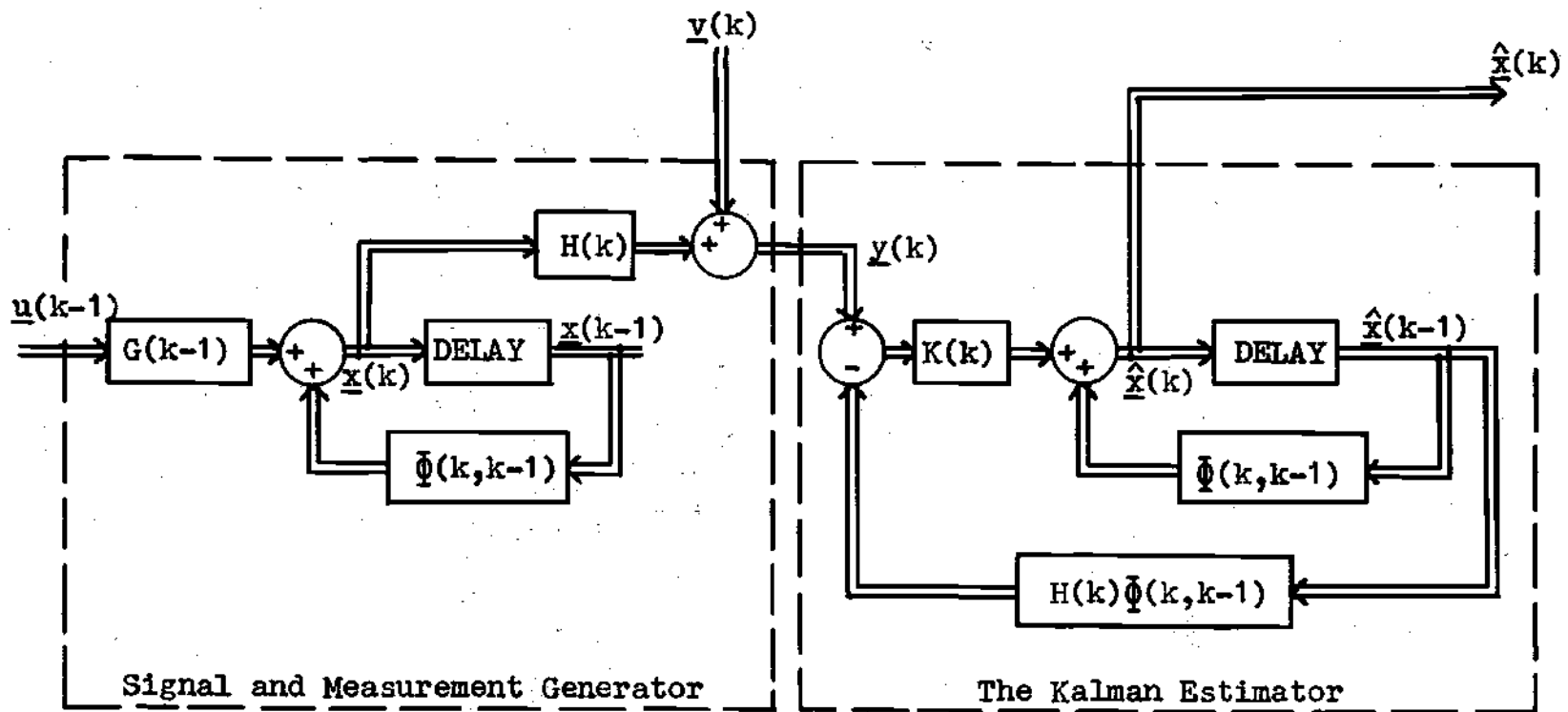


Figure 1. The Kalman Estimator

component, that component is described by the state transition matrix, $\Phi(k,k-1)$. Errors in the assumed nature of the signal dynamics are therefore characterized by errors in the state transition matrix. The sensitivity of the estimate to such errors is of primary concern in this dissertation.

Much of the recent literature has been concerned with the causes and nature of the Kalman estimator sensitivity problem. It is generally known that the actual estimation error, or the actual error variance, can diverge when an inaccurate state transition matrix is used in the Kalman estimator equations. Fagin (14) has determined expressions for the actual error variance resulting when an inaccurate state transition matrix is used in the Kalman estimator equations. Griffin and Sage (15) have derived algorithms for computing sensitivity coefficients associated with errors in the state transition matrix. Similar expressions were derived by Price (16), who included sufficiency conditions for uniform asymptotic stability in the large of the actual error covariance matrix. These contributions are useful for analyzing the behavior of the estimate when an approximation to a known state transition matrix property is used, but they do not give any alternative estimation techniques to use in the presence of unknown errors in the signal dynamical model.

Error divergence in the Kalman estimator can be related to the behavior of the computed error covariance $P(k)$ and the estimator gain $K(k)$. For signals which contain no random noise ($u(k) \equiv 0$ and $R(k) \equiv 0$), the gain and the computed error covariance both tend to vanish as k gets large. O'Donnell(17) has proven this for scalar signals and

Sorenson (18) has derived sufficiency conditions for the vanishing of the gain and computed error covariance matrix. The vanishing of the gain matrix means that new measurements are ignored in favor of the past estimates. After a large number of noisy measurements of a deterministic signal (no signal noise component), the estimator has effectively matched the data with the assumed model and therefore computes each new estimate by using only the preceding estimate and the assumed state transition matrix. This behavior corresponds to open loop operation in the block diagram of Figure 1, as the feedback loop has no effect when the gain vanishes. Since the gain and the computed error covariance are proportional, the vanishing of the gain corresponds to the fact that after a large number of noisy samples of a noiseless signal the error covariance computed using the assumed model becomes arbitrarily small. Therefore, the estimator can be computing estimates having a computed error covariance which decreases with time while the actual error covariance is diverging.

The literature contains many examples of the behavior of the Kalman estimator in the presence of an inaccurate signal dynamical model. Schlee, Standish, and Toda (19) showed that error divergence can occur in a Kalman estimator used for autonomous navigation using known landmark tracking in a low earth orbit. The error sources which they showed can cause error divergence are small errors in drag acceleration and computational errors due to finite word length computer computations. Brogan and LeMay (20) also illustrated error divergence in a similar orbit navigation system. Error divergence can also result when ground radar range measurements of an earth orbit trajectory are

processed with an inaccurate model. Jazwiński (21) showed that error divergence can occur in this problem due to small errors in the assumed gravity constant.

To illustrate error divergence in a Kalman estimator, consider the problem of estimating a nonrandom signal having the form

$$x(t) = a_1 + a_2 t + a_3 t^2 \quad (2.15)$$

where a_1 , a_2 , and a_3 are unknown constants. Discrete measurements have the form

$$y(k) = x(k) + v(k) \quad (2.16)$$

where $v(k)$ is a white Gaussian noise sequence with unit variance. The signal of Equation (2.15) may represent the range of an airborne vehicle having constant acceleration. If discrete time state variables are defined by $x_1(k) = x(k)$, $x_2(k) = \dot{x}_1(k)$, $x_3(k) = \ddot{x}_1(k)$, then the accurate model for the Kalman estimator would be

$$\underline{x}(k+1) = \begin{bmatrix} 1 & 1 & \frac{1}{2} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \underline{x}(k) \quad (2.17)$$

$$\underline{y}(k) = [1 \ 0 \ 0] \underline{x}(k) + v(k) \quad (2.18)$$

This model used in the Kalman estimator equations would result in the

minimum mean-square error estimate of $x(k)$. Now suppose that the signal of Equation (2.15) is modeled by assuming that $x(k)$ is a polynomial signal having constant velocity:

$$x_m(t) = c_1 + c_2 t \quad (2.19)$$

where $x_m(t)$ is the assumed model for $x(t)$, and c_1 and c_2 are unknown constants. This model has a state variable formulation given by

$$\underline{x}_m(k+1) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \underline{x}_m(k) \quad (2.20)$$

Figure 2 illustrates the error divergence resulting from the use of the model of Equations (2.20, 2.16) for the signal and measurement of Equations (2.15, 2.16). For clarity, only every tenth error value is shown on the curve. The constants used to generate the signal for Figure 2 are $a_1 = 10.0$, $a_2 = -0.16$, and $a_3 = 0.0008$, so that $x(k)$ is a parabolic arc of magnitude ten at $k = 0$ and $k = 200$, with a minimum value of two at $k = 100$.

As another example of the Kalman estimator behavior, consider the problem of estimating the altitude of an object re-entering the earth's atmosphere. Measurements consist of noisy slant range radar outputs as shown in Figure 3.

Estimation
Error
 $= x_1(k) - \hat{x}_1(k)$

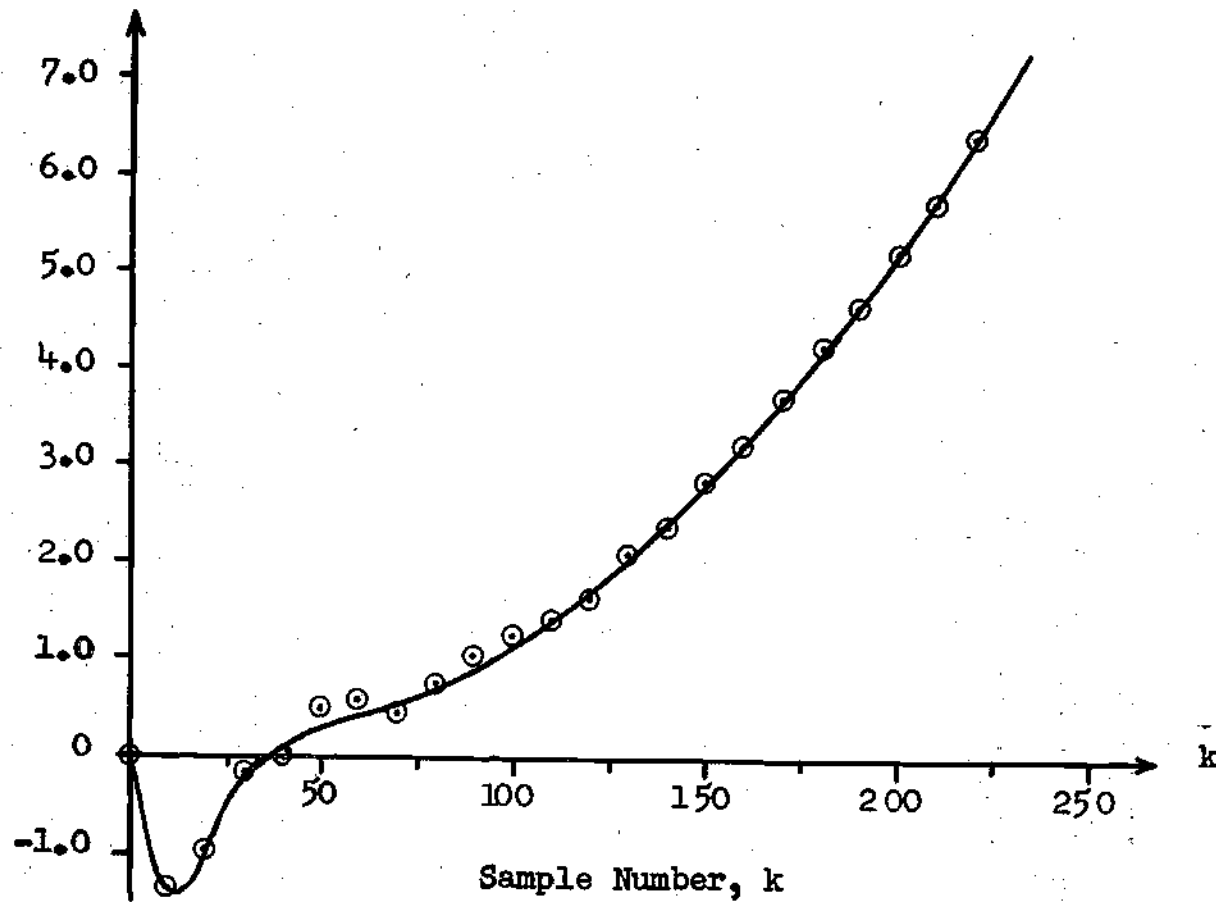


Figure 2. Kalman Estimator Error, Second Degree Polynomial Signal with First Degree Polynomial Model

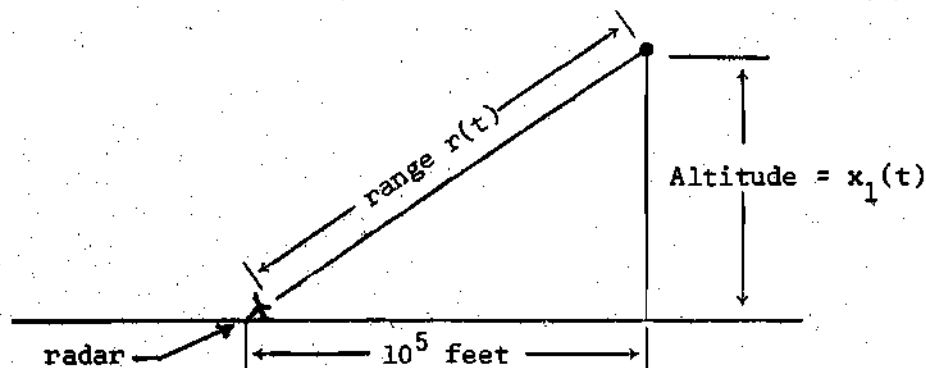


Figure 3. Configuration for Re-entry Tracking Problem.

The trajectory is described by a nonlinear state variable differential equation of the form

$$\dot{\underline{x}}(t) = \underline{f}(\underline{x}(t)) \quad (2.21)$$

where

$$\underline{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} = \begin{bmatrix} \text{altitude (feet)} \\ \text{velocity (feet/second)} \\ \text{a constant ballistic parameter} \end{bmatrix} \quad (2.22)$$

and

$$\underline{f}(\underline{x}(t)) = \begin{bmatrix} -x_2(t) \\ x_2^2(t)x_3(t)e^{-\gamma x_1(t)} \\ 0 \end{bmatrix} \quad (2.23)$$

For purposes of simulation, the drag constant γ is chosen to be 0.000050

and the initial state vector is given by

$$\underline{x}(0) = \begin{bmatrix} 300,000 \\ 20,000 \\ 0.001 \end{bmatrix} \quad (2.24)$$

Radar range measurements are taken each second, so

$$\underline{y}(k) = \underline{g}(\underline{x}(k)) + \underline{v}(k) \quad (2.25)$$

where in this example $\underline{g}(\underline{x}(k)) = r(k) = \sqrt{10^{10} + x_1^2(k)}$ and $v(k)$ is white noise with variance $V(k) = 10^4$ feet². The resulting altitude and range are shown in Figure 4. The above model and numbers coincide with those used by Athans, Wishner, and Bertolini (22) in their recent contribution to the application of the Kalman estimator to nonlinear signal estimation. They showed that the inclusion of second order as well as first order terms in a Taylor expansion linearization of a nonlinear estimation problem could result in estimates which are better than those obtained using only the first order terms. They were not concerned with the sensitivity problem being considered in this dissertation.

In the absence of accurate information about the model of Equation (2.23), one approach to the altitude estimation problem might be to apply a Kalman estimator by assuming that the range variable $r(k)$ is a second degree polynomial with unknown coefficients:

$$r_m(k) = a_1 + a_2 k + a_3 k^2 \quad (2.26)$$

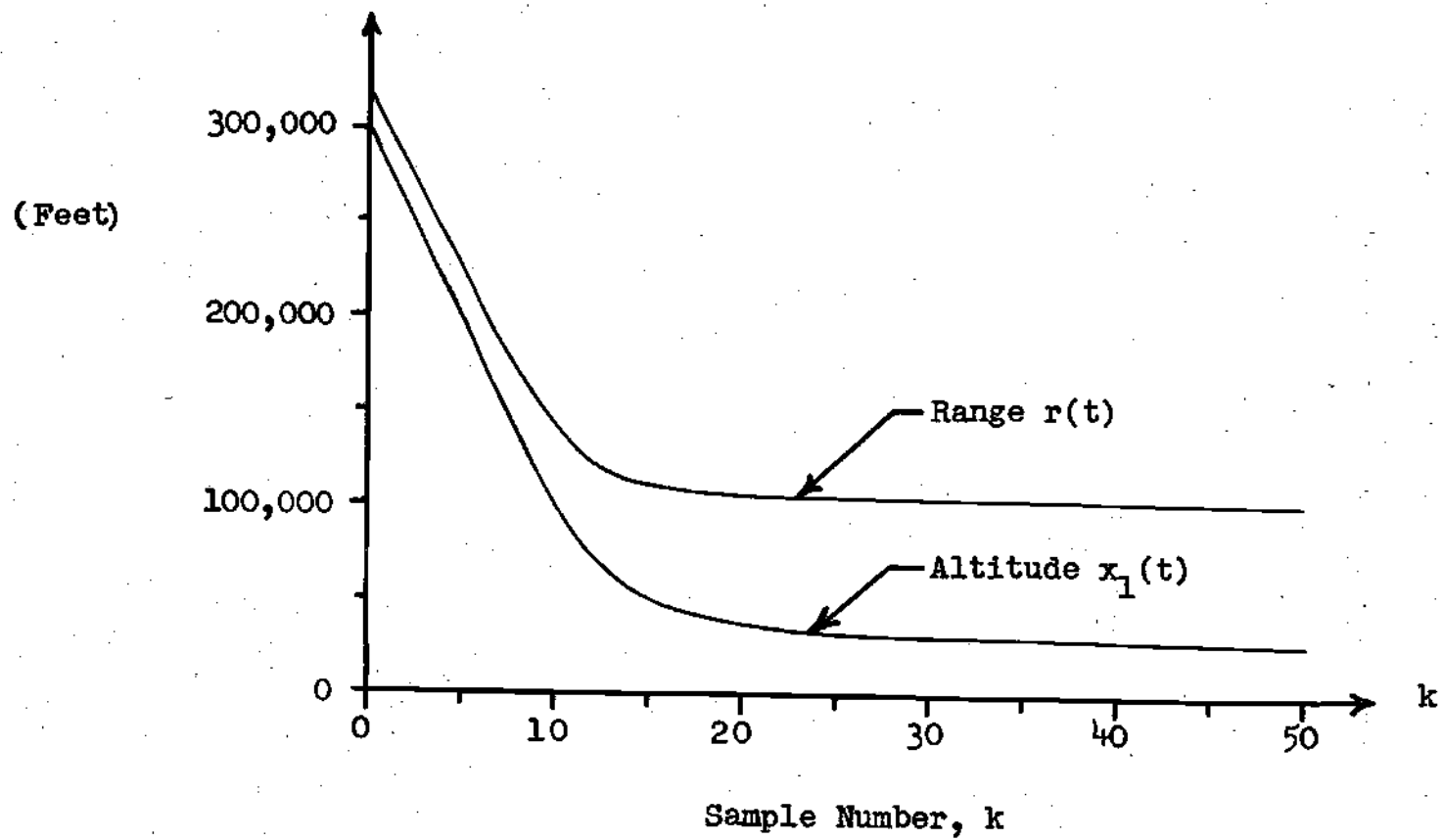


Figure 4. Re-entry Signal Trajectories

where the subscript m indicates that Equation (2.26) is the assumed model for $r(k)$. Then the Kalman estimator, using the transition matrix of Equation (2.17), with $r_m(k)$, $\dot{r}_m(k)$, and $\ddot{r}_m(k)$ as the state variables, would generate estimates $\hat{r}_m(k)$ of $r_m(k)$. From these estimates, the altitude could then be estimated by

$$x_1(k) = \sqrt{[\hat{r}_m(k)]^2 - 10^{10}} \quad (2.27)$$

The resulting error diverges, as shown in Figure 5. For clarity, the error values are shown for even sample numbers only. The initial estimate used to obtain the data for Figure 5 was accurate, so the divergence results because of the inaccurate model of the dynamics of the range variable.

If the accurate model of Equations (2.21 through 2.25) is known, the accurate model can be used in a variation of the Kalman estimator. This method, common in the literature, is discussed by Sorenson (25), Schmidt (24), Athans (22), and others. The signal and measurement are first linearized for small variations by taking only the first order terms in a Taylor series expansion. The variations are then described by

$$\underline{\delta \dot{x}}(t) = A \underline{\delta x}(t) \quad (2.28)$$

$$\underline{\delta y}(t) = H \underline{\delta x}(t) + \underline{v}(t) \quad (2.29)$$

with $\underline{\delta x}(t)$ and $\underline{\delta y}(t)$ defined by

Altitude
Estimation
Error
 $= x_1(k) - \hat{x}_1(k)$
(feet)

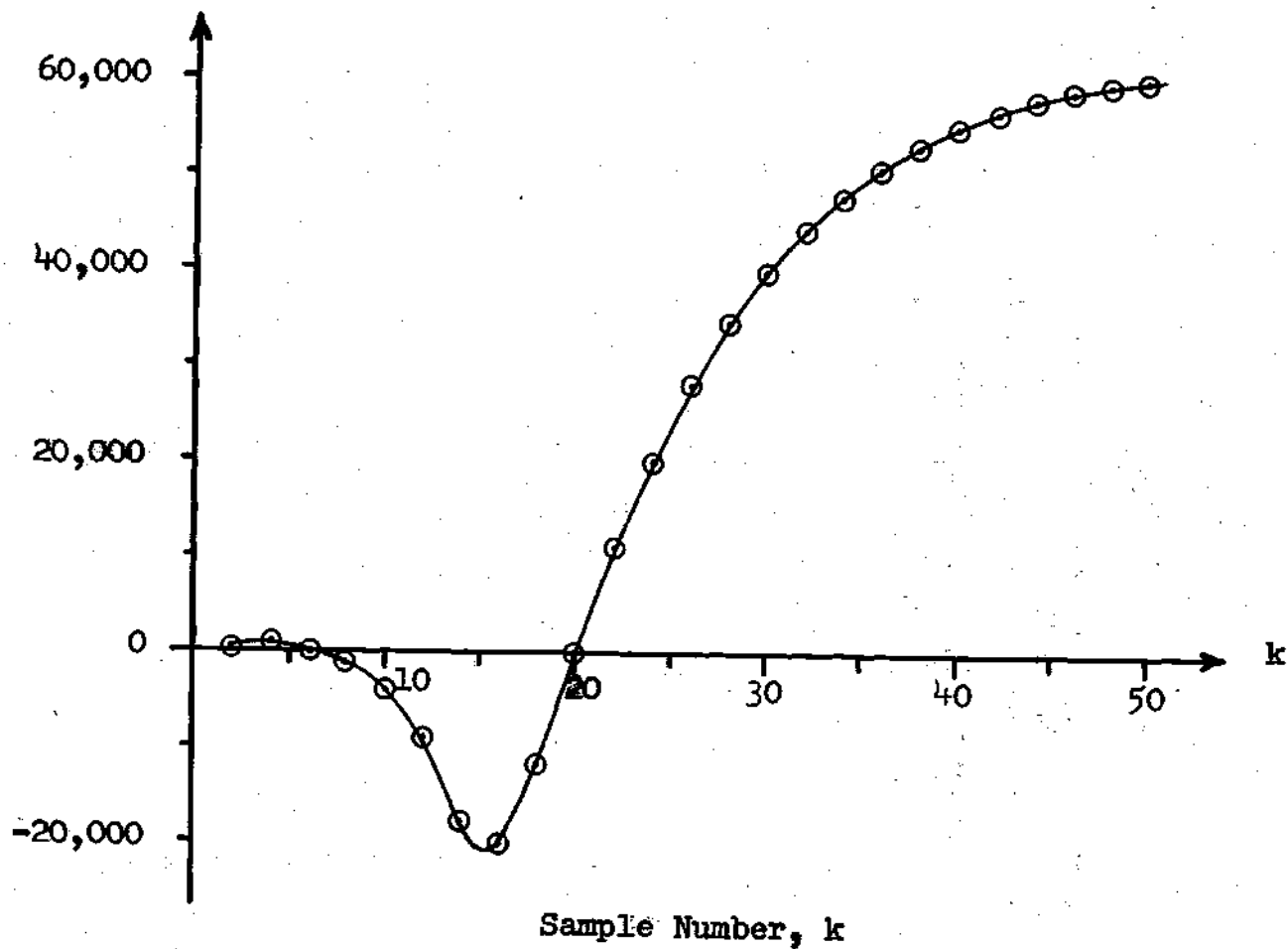


Figure 5. Kalman Estimator Error, Re-entry Trajectory
with Polynomial Range Model

$$\underline{\delta x}(t) \triangleq \underline{x}(t) - \underline{x}_0(t) \quad (2.30)$$

$$\underline{\delta y}(t) \triangleq \underline{y}(t) - \underline{g}(\underline{x}_0(t)) \quad (2.31)$$

A and H are found from Equations (2.21, 2.25) by

$$A = \left. \frac{\partial \underline{f}(\underline{x})}{\partial \underline{x}} \right|_{\underline{x}=\underline{x}_0} \quad (2.32)$$

$$H = \left. \frac{\partial \underline{g}(\underline{x})}{\partial \underline{x}} \right|_{\underline{x}=\underline{x}_0} \quad (2.33)$$

where \underline{x}_0 is the value of the reference trajectory. In the case of estimation, \underline{x}_0 is the most recent estimate of the state vector. To apply the discrete Kalman estimator technique, Equations (2.28, 2.29) must be written in terms of their equivalent discrete system:

$$\underline{\delta x}(k) = \Phi(k, k-1) \underline{\delta x}(k-1) \quad (2.34)$$

$$\underline{\delta y}(k) = H(k) \underline{\delta x}(k) + \underline{v}(k) \quad (2.35)$$

Once this is accomplished, a modification of the Kalman estimator is applied by computing the estimate of $\underline{\delta x}(k)$ by

$$\underline{\hat{\delta x}}(k) = \underline{\hat{\delta x}}'(k) + K(k) [\underline{\delta y}(k) - H(k) \underline{\hat{\delta x}}'(k)] \quad (2.36)$$

where $\underline{\hat{\delta x}}'(k)$ is the estimate of $\underline{\delta x}(k)$ obtained from $\underline{\hat{\delta x}}(k-1)$ by numerically integrating Equation (2.28). Comparison of Equation (2.36) with Equation (2.9) shows that $\underline{\hat{\delta x}}'(k)$ merely replaces $\Phi(k,k-1) \underline{\hat{\delta x}}(k-1)$ as the estimate of $\underline{\delta x}(k)$ obtained without the use of the measurement $\underline{y}(k)$. The gain $K(k)$ in Equation (2.36) is computed by Equations (2.10, 2.11, 2.12), where $R(k) \equiv 0$ and $\Phi(k,k-1)$ and $H(k)$ are the discrete versions of Equations (2.32, 2.33). The estimate of the signal $\underline{x}(k)$ is obtained from $\underline{\hat{\delta x}}(k)$ by

$$\underline{\hat{x}}(k) = \underline{\hat{\delta x}}(k) + \underline{x}_0(k) \quad (2.37)$$

where \underline{x}_0 is the reference trajectory value. In the estimation problem \underline{x}_0 is not known exactly but is replaced by the best estimate of $\underline{x}(k)$ prior to receiving $\underline{y}(k)$. This best estimate is obtained by integrating Equation (2.21) initialized by $\underline{\hat{x}}(k-1)$. This technique applied to nonlinear signals gives excellent results when an accurate model is known and the signal variations are adequately small between samples so that Equations (2.28, 2.29) are good approximations. Athans (22) illustrated the behavior of this method with the re-entry body tracking problem. The sensitivity of this method to errors in the nonlinear signal model, which corresponds to errors in the linearized state transition matrix, can be illustrated by examining the altitude error behavior resulting when the drag constant is inaccurate. Figure 6 illustrates the altitude error behavior when the drag constant used by the Kalman estimator is in error by 10 per cent of the actual value. For clarity, the error values are shown for even samples only. While the error does not

Altitude
Estimation
Error
 $= x_1(k) - \hat{x}_1(k)$
(feet)

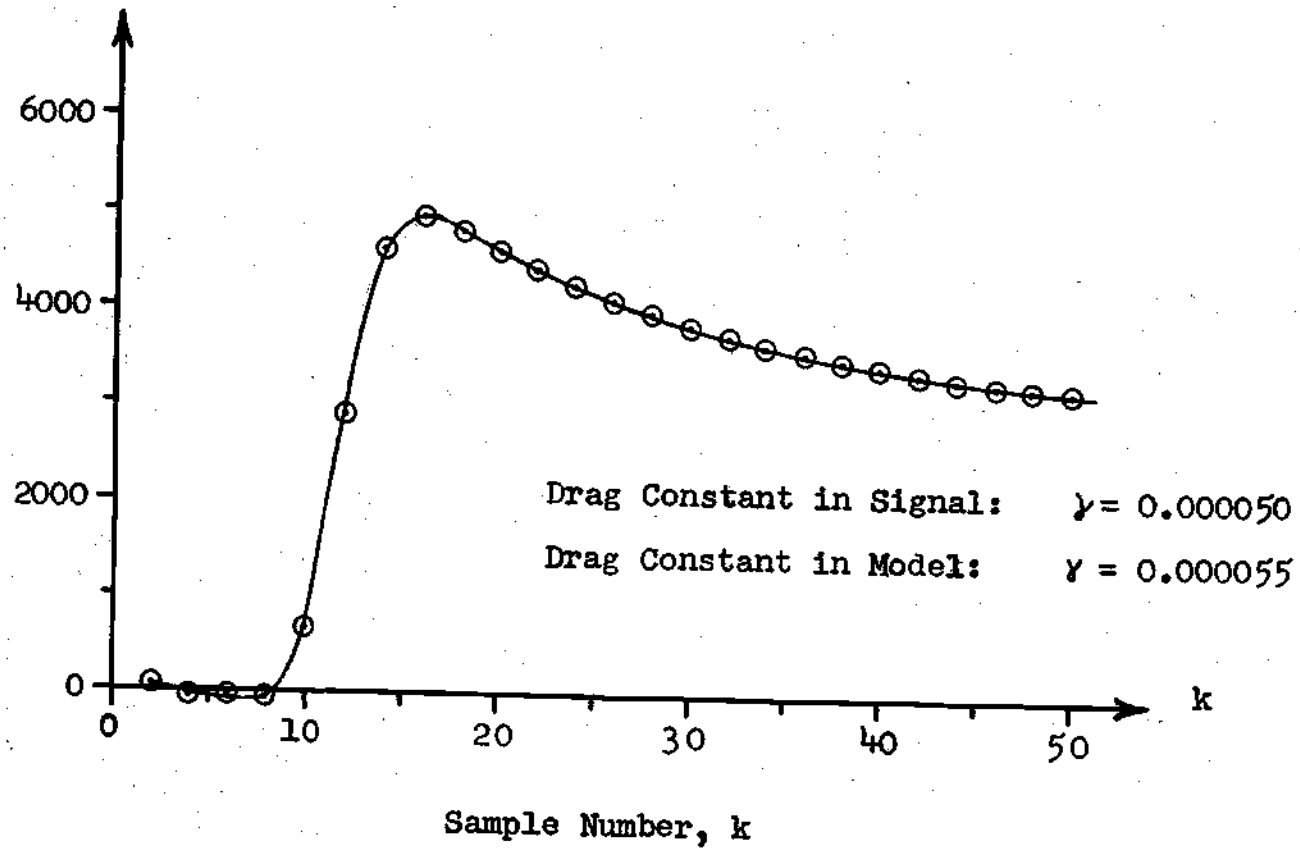


Figure 6. Kalman Estimator Error, Re-entry Trajectory with Inaccurate Drag Constant in Signal Model

diverge in this case, a considerable bias is present in the estimate and this bias is not detectable by inspection of the estimator equations.

Modifications of the Kalman Estimator

The examples discussed above serve to point out some results of the sensitivity problems encountered in the application of the Kalman estimator. Many of the recent publications on the sensitivity problem have suggested modifications of the Kalman estimator intended to make it less sensitive to errors in the signal dynamical model. One of the most common of these methods, as presented in Sorenson (23) and Schmidt (24) for example, is the method of state augmentation. This method consists of including unknown parameters in the state vector so that they are estimated along with the desired signal. The augmented system will have the same form as Equations (2.3, 2.4) but the state vector will have a larger number of terms. The Kalman estimation equations retain the form of Equations (2.9) through (2.12), but the augmented estimate includes estimates of the unknown parameters. In order to use this method, the state transition properties for the unknown parameters must be known.

A method which is similar to the state augmentation method is given by Schmidt (24). In this method the state vector to be estimated and the measurements are described by

$$\underline{x}(k) = \Phi(k,k-1)\underline{x}(k-1) + U(k,k-1)\underline{u}(k-1) \quad (2.38)$$

$$\underline{y}(k) = H(k)\underline{x}(k) + F(k)\underline{u}(k) + \underline{v}(k) \quad (2.39)$$

where $\underline{u}(k)$ is a vector of unknown parameters described by

$$\underline{u}(k) = \Psi(k,k-1)\underline{u}(k-1) \quad (2.40)$$

$\underline{v}(k)$ is white measurement noise with known covariance matrix $V(k)$, and $\Phi(k,k-1)$, $U(k,k-1)$, $H(k)$, $F(k)$, and $\Psi(k,k-1)$ are known matrices. This method has an advantage over the state augmentation method in that the unknown parameters are not estimated. The state vector estimate is computed by

$$\hat{\underline{x}}(k) = \Phi(k,k-1)\hat{\underline{x}}(k-1) + K(k)[\underline{y}(k) - H(k)\Phi(k,k-1)\hat{\underline{x}}(k-1)] \quad (2.41)$$

which is the same as Equation (2.9). To compute the gain $K(k)$, first define the following matrices:

$$P(k) \triangleq E\{[\underline{x}(k) - \hat{\underline{x}}(k)][\underline{x}(k) - \hat{\underline{x}}(k)]^T\} \quad (2.42)$$

$$C(k) \triangleq E\{[\underline{x}(k) - \hat{\underline{x}}(k)]\underline{u}^T(k)\} \quad (2.43)$$

$$D(k) \triangleq E\{\underline{u}(k)\underline{u}^T(k)\} \quad (2.44)$$

The gain is then computed by

$$\begin{aligned}
K(k) = & [P'(k)H^T(k) + C'(k)F^T(k)][H(k)P'(k)H^T(k) & (2.45) \\
& + V(k) + F(k)D(k)F^T(k) + H(k)C'(k)F^T(k) \\
& + F(k)C'^T(k)H^T(k)]
\end{aligned}$$

where $P'(k)$ and $C'(k)$ are defined by

$$\begin{aligned}
C'(k) = & \phi(k,k-1)C(k-1)\Psi^T(k,k-1) & (2.46) \\
& + U(k,k-1)D(k-1)\Psi^T(k,k-1)
\end{aligned}$$

$$\begin{aligned}
P'(k) = & \phi(k,k-1)P(k-1)\phi^T(k,k-1) & (2.47) \\
& + \phi(k,k-1)C(k-1)U^T(k,k-1) \\
& + U(k,k-1)C^T(k-1)\phi^T(k,k-1) \\
& + U(k,k-1)D(k-1)U^T(k,k-1)
\end{aligned}$$

and $P(k)$, $C(k)$, and $D(k)$ are updated for the next sample by

$$P(k) = [I - K(k)H(k)] P'(k) - K(k)F(k)C'(k) \quad (2.48)$$

$$C(k) = [I - K(k)H(k)] C'(k) - K(k)F(k)D(k) \quad (2.49)$$

$$D(k) = \Psi(k,k-1)D(k-1)\Psi^T(k,k-1) \quad (2.50)$$

Another approach to the problem of estimating signals by using an inaccurate state transition matrix was presented by Neal (25). In his method, the actual signal and measurement have the form

$$\underline{x}(k) = [\Phi(k,k-1) + \delta\Phi(k,k-1)] \underline{x}(k-1) + \underline{u}(k-1) \quad (2.51)$$

$$\underline{y}(k) = H(k)\underline{x}(k) \quad (2.52)$$

where $\underline{u}(k)$ is a white noise process having known covariance matrix $R(k)$. The model $\underline{x}_m(k)$ assumed for the signal is

$$\underline{x}_m(k) = \Phi(k,k-1)\underline{x}_m(k-1) + \underline{u}(k-1) \quad (2.53)$$

and the estimate has the form

$$\hat{\underline{x}}(k) = \Phi(k,k-1)\hat{\underline{x}}(k-1) + K(k)[\underline{y}(k) - H(k)\Phi(k,k-1)\hat{\underline{x}}(k-1)] \quad (2.54)$$

This is the form of the Kalman estimator obtained by using $\Phi(k,k-1)$ as an approximation to the actual transition matrix $\Phi(k,k-1) + \delta\Phi(k,k-1)$. Neal presented equations for computing the gain $K(k)$ in Equation (2.54) in such a way that the resulting estimate is closer to the optimum estimate than that which would be obtained by using the model of Equation (2.53) in the Kalman estimator equations. This procedure is equivalent to making the signal noise covariance $R(k)$ larger than its

actual value, a technique to be described below. Neal's method and the other modifications discussed above all require that the sources of error in the signal dynamical model be known and have a known model. Therefore, these models do not represent useful modifications to the Kalman estimator applied to signals with an unknown model or unknown error source behavior.

A modification of the Kalman estimator which does not assume knowledge of all error sources was given by Schmidt (26). This method, referred to here as the ϵ -method, has the effect of preventing the estimator gain and error covariance matrices from losing numerical significance. The Kalman estimator for a noiseless signal can be written in the form

$$\underline{\hat{x}}(k) = \Phi(k, k-1)\underline{\hat{x}}(k-1) + \Delta\underline{\hat{x}}(k) \quad (2.55)$$

where

$$\Delta\underline{\hat{x}}(k) = P'(k)H^T(k)[H(k)P'(k)H^T(k) + V(k)]^{-1} \quad (2.56)$$

$$[y(k) - H(k)\Phi(k, k-1)\underline{\hat{x}}(k-1)]$$

The terms in the above equations are the same as those in Equations (2.9, 2.10). Rather than using Equation (2.56), the ϵ -method computes $\Delta\underline{\hat{x}}(k)$ as the sum of $\Delta\underline{\hat{x}}_1(k)$ and $\Delta\underline{\hat{x}}_2(k)$, where

$$\Delta\underline{\hat{x}}_1(k) = P'(k)H^T(k)[H(k)P'(k)H^T(k) + V(k)]^{-1} \quad (2.57)$$

$$[\underline{y}(k) - H(k)\phi(k,k-1)\hat{\underline{x}}(k-1)]$$

$$\underline{\Delta\hat{x}}_2(k) = \epsilon H^T(k)[H(k)P'(k)H^T(k) + V(k)]^{-1} \quad (2.58)$$

$$[\underline{y}(k) - H(k)\phi(k,k-1)\hat{\underline{x}}(k-1)]$$

The $\underline{\Delta\hat{x}}_1(k)$ component is the same as the $\underline{\Delta\hat{x}}(k)$ computed for the Kalman estimator and the $\underline{\Delta\hat{x}}_2(k)$ component is proportional to the gain for a Kalman estimator for which no a priori information is available. The estimator computed by the ϵ -method can then be written

$$\hat{\underline{x}}(k) = \phi(k,k-1)\hat{\underline{x}}(k-1) + [P'(k)H^T(k) \quad (2.59)$$

$$+ \epsilon H^T(k)][H(k)P'(k)H^T(k)$$

$$+ V(k)]^{-1} [\underline{y}(k) - H(k)\phi(k,k-1)\hat{\underline{x}}(k-1)]$$

The equations for updating the $P'(k)$ matrix become

$$P'(k) = \phi(k,k-1)P(k-1)\phi^T(k,k-1) \quad (2.60)$$

$$P(k) = P'(k) - P'(k)H^T(k)[H(k)P'(k)H^T(k) \quad (2.61)$$

$$+ V(k)]^{-1} H(k)P'(k)$$

$$+ \epsilon^2 [H(k)P'(k)H^T(k) + V(k)]^{-1} H^T(k)H(k)$$

where $P(k)$ is defined as in Equation (2.8). For the case of $\epsilon=0$, the above equations can be shown to be equivalent to the Kalman estimator of Equations (2.9) through (2.12) with $R(k) \equiv [0]$ (no signal noise). The ϵ^2 term in Equation (2.61) has the effect of increasing the eigenvalues of $P(k)$ over their optimum values. The value of ϵ would be a choice of the user of the ϵ -method. If ϵ were chosen to be $NV(k)$, then N represents the fraction of the variance of an individual measurement within which one can expect to estimate the signal after a very large number of measurements. The advantage of the ϵ -method is that it allows the use of a bound on acceptable accuracy computed by the estimator. The variance in the estimate of $[H(k)\underline{x}(k)]$, given by

$$\begin{aligned} E\{[H(k)\underline{x}(k) - H(k)\hat{\underline{x}}(k)][H(k)\underline{x}(k) - H(k)\hat{\underline{x}}(k)]^T\} & \quad (2.62) \\ & = H(k)P(k)H^T(k) \end{aligned}$$

has the property that in the limit of an infinite number of measurements the ϵ -method gives

$$\lim_{k \rightarrow \infty} [H(k)P(k)H^T(k)] = \epsilon \quad (2.63)$$

if the model is accurate. This limit would be zero for the Kalman estimator applied to a noiseless signal. The ϵ -method has the effect of keeping the gain, $[P'(k)H^T(k) + \epsilon H^T(k)][H(k)P'(k)H^T(k) + V(k)]^{-1}$, nonzero regardless of how long measurements are taken. Since this means that new measurements will never be ignored, the ϵ -method is less

sensitive to dynamical model errors than is the Kalman estimator. However, Schmidt has shown that an inaccurate transition matrix used with the ϵ -method can result in a steady-state bias when the Kalman estimator would diverge.

Any modification of the Kalman estimator which preserves the form of Equation (2.9) and produces a gain or computed error covariance larger than the optimum values results in an estimate which is less sensitive to modeling errors than is the Kalman estimator. As mentioned above, the ϵ -method computes a gain larger than that computed by the Kalman estimator equations. Other methods of computing larger than optimum gains have also been used. Schmidt (26), Schlee (19), and Brogan (20) each show that an effective modification of the Kalman estimator equations is to multiply the diagonal terms of the error covariance matrix by $(1 + \epsilon)$ after each computation of $P(k)$, where ϵ is chosen by the user of the method. This technique tends to keep $P(k)$ nonnegative definite while making the gain larger than the optimum. Another method is to impose predetermined lower bounds on the gain or the diagonal terms of the error covariance matrix.

A common method of keeping the estimator gain from vanishing is to include in the signal model a noise component which does not actually exist, or increase an existing signal noise covariance matrix. Inspection of Equations (2.11, 2.12) verifies that this procedure will increase the eigenvalues of the computed error covariance matrix and therefore increase the gain. The choice of the artificial signal noise covariance matrix can be made experimentally by simulating the

types of signals expected, as discussed by Brogan (20) and Neal (25). The technique of introducing an artificial signal noise component can be related to the magnitude of estimation error expected. To illustrate this relationship, first combine Equations (2.11) and (2.12) to obtain

$$\begin{aligned}
 P(k) = & [I-K(k)H(k)] \Phi(k,k-1)P(k-1) \Phi^T(k,k-1)[I-K(k)H(k)]^T \quad (2.64) \\
 & + [I-K(k)H(k)] G(k-1)R(k-1)G^T(k-1)[I-K(k)H(k)]^T \\
 & + K(k)V(k)K^T(k)
 \end{aligned}$$

Suppose that in computing the $\Phi(k,k-1)\hat{\underline{x}}(k-1)$ term in Equation (2.9), an error is introduced. This error is described by a zero mean process \underline{n}_x , so the actual $\Phi(k,k-1)\hat{\underline{x}}(k-1)$ term is replaced by $\Phi(k,k-1)\hat{\underline{x}}(k-1) + \underline{n}_x$. The estimate computed using the Kalman estimator equations then becomes

$$\begin{aligned}
 \hat{\underline{x}}(k) = & \Phi(k,k-1)\hat{\underline{x}}(k-1) \quad (2.65) \\
 & + K(k)[\underline{y}(k) - \Phi(k,k-1)H(k)\hat{\underline{x}}(k-1)] \\
 & + [I - K(k)H(k)]\underline{n}_x
 \end{aligned}$$

The error covariance resulting from this estimate is given by

$$P(k) = E\{[\underline{x}(k) - \hat{\underline{x}}(k)][\underline{x}(k) - \hat{\underline{x}}(k)]^T\} \quad (2.66)$$

Assuming that \underline{n}_x is independent of $\hat{x}(k-1)$ and $y(k)$, the error covariance resulting from the estimate of Equation (2.65) becomes

$$\begin{aligned}
 P(k) = & [I - K(k)H(k)] \Phi(k,k-1)P(k-1) \Phi^T(k,k-1)[I - K(k)H(k)]^T \quad (2.67) \\
 & + [I - K(k)H(k)] R_x [I - K(k)H(k)]^T \\
 & + K(k)V(k)K^T(k)
 \end{aligned}$$

where R_x is the covariance matrix of the \underline{n}_x noise process. Comparison of Equation (2.67) with Equation (2.64) shows that R_x contributes to the error covariance matrix in the same manner as the covariance of a signal noise component. The terms in R_x can be chosen by simulation or by a knowledge of the round-off error in the computing system.

An additional method for preventing error divergence in the Kalman estimator is to "reset" the estimator before the error becomes too large. Such a method was proposed by Jazwinski (21). His method consists of using the Kalman estimator for $2N$ samples, then recomputing new initial conditions based on the last N measurements, where N is a fixed integer. This technique results in a finite memory estimator whose memory varies from N to $2N$. The error resulting from the use of this method will remain relatively small if the assumed signal model is an adequate approximation to the actual signal over time periods of $2N$ samples. Data collected more than $2N$ samples in the past is ignored.

It is evident from the above discussion that the designer of a signal estimator has a wide selection of available modifications to the

Kalman estimator. Each of the techniques available will result in a suboptimal estimator which is less sensitive than the Kalman estimator to errors in the signal dynamical model. Chapter III will contain a useful interpretation of the Kalman estimator, which will imply an alternative signal modeling technique. Comparison of the structure and modeling techniques available for fixed memory estimators and the Kalman estimator will lay the framework for a new estimation technique presented in Chapter IV.

CHAPTER III

INTERPRETATIONS OF DISCRETE SIGNAL ESTIMATION

In Chapter II the Kalman estimator was discussed in terms of its sensitivity to errors in the model of the signal dynamics, and various techniques for modifying the estimator were described. In this chapter the Kalman estimator will be interpreted as a curve fitting algorithm which minimizes a deterministic least squares cost function. In addition to providing more insight into the sensitivity problem discussed in the previous chapter, the curve fitting interpretation verifies that the Kalman estimator equations represent a computational simplification of a growing memory estimator which computes estimates which are expressible as weighted sums of all available measurements. The computational simplification is a result of the recursive signal model used in the Kalman estimator equations. The curve fitting interpretation of the Kalman estimator also serves as motivation for the use of an alternative signal model presented in this chapter. The alternative signal model will be shown to result in computational simplicity when a fixed memory estimator is formulated. It will be shown that the fixed memory estimator using the alternative signal model is also interpretable as a least squares curve fitting procedure, but with the least squares cost function depending only on errors at sampling instants contained in a time interval of fixed length. The curve fitting interpretations of the Kalman estimator and the fixed memory estimator

will be shown to imply an advantage of the fixed memory estimator in its relative insensitivity to errors in the signal model.

The results of this chapter will be used in Chapter IV to postulate an estimator which will be shown to have properties which represent the desirable features of the Kalman and fixed memory estimators.

Least Squares Curve Fitting and the Kalman Estimator

As described in Chapter II, the Kalman estimator minimizes mean-square error when the signal and measurement are described by Equations (2.3) and (2.4) and the noise processes are white and Gaussian. When the signal is deterministic, so that $\underline{u}(k) \equiv 0$ in Equation (2.3), the Kalman estimator is also obtained by minimizing a deterministic least squares cost function rather than mean-square error. The Kalman estimator is therefore interpretable as a curve fitting algorithm. This relationship was proven by Sage and Masters (27), and a different proof is given below.

Consider a dynamic signal described by

$$\underline{x}(k) = \phi(k, k-1)\underline{x}(k-1) \quad (3.1)$$

with discrete measurements given by

$$\underline{y}(k) = H(k)\underline{x}(k) + \underline{v}(k), \quad k=0,1,2,\dots \quad (3.2)$$

where $\underline{v}(k)$ is white noise with covariance matrix $V(k)$. The problem is to compute a linear estimate $\hat{\underline{x}}(k)$ which minimizes the cost function given by

$$J(k) = [\underline{y}_T(k) - H_T(k)\hat{\underline{x}}(k)]^T v_T^{-1}(k)[\underline{y}_T(k) - H_T(k)\hat{\underline{x}}(k)] \quad (3.3)$$

This cost function is a weighted sum of the squares of deviations between all past measurements and the estimates of the measurements.

The terms in $J(k)$ are defined by

$$\underline{y}_T(k) \triangleq \begin{bmatrix} \underline{y}(0) \\ \underline{y}(1) \\ \underline{y}(2) \\ \vdots \\ \underline{y}(k) \end{bmatrix} \quad (3.4)$$

$$H_T(k) \triangleq \begin{bmatrix} H(1)\phi(1,k) \\ H(2)\phi(2,k) \\ H(3)\phi(3,k) \\ \vdots \\ H(k)\phi(k,k) \end{bmatrix} \quad (3.5)$$

where $\underline{x}(j)$ and $\underline{y}_T(k)$ can be written as

$$\underline{x}(j) = \phi(j,k)\underline{x}(k) \quad (3.6)$$

$$\underline{y}_T(k) = H_T(k)\underline{x}(k) + \underline{v}_T(k) \quad (3.7)$$

and $\underline{v}_T(k)$ is defined by

$$\underline{y}_T(k) \triangleq \begin{bmatrix} \underline{v}(0) \\ \underline{v}(1) \\ \underline{v}(2) \\ \vdots \\ \underline{v}(k) \end{bmatrix} \quad (3.8)$$

with covariance matrix $V_T(k)$ given by

$$V_T(k) = \begin{bmatrix} V(0) & 0 & 0 & \cdots & 0 \\ 0 & V(1) & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & V(k) \end{bmatrix} \quad (3.9)$$

To minimize $J(k)$, set $\frac{\partial J(k)}{\partial \underline{\hat{x}}(k)} = \underline{0}$, giving

$$\underline{0} = -2H_T^T(k)V_T^{-1}(k)[\underline{y}_T(k) - H_T(k)\underline{\hat{x}}(k)] \quad (3.10)$$

which gives as the estimate

$$\underline{\hat{x}}(k) = [H_T^T(k)V_T^{-1}(k)H_T(k)]^{-1} H_T^T(k)V_T^{-1}(k)\underline{y}_T(k) \quad (3.11)$$

In this form, the estimate is expressed as a weighted sum of all measurement data. To obtain a recursive form for the estimate, consider the addition of a new measurement, $\underline{y}(k+1)$. Equation (3.7) can then be replaced by

$$\begin{bmatrix} \underline{y}_T(k) \\ \underline{y}(k+1) \end{bmatrix} = \begin{bmatrix} \underline{H}_T(k)\phi(k,k+1) \\ H(k+1) \end{bmatrix} \underline{x}(k+1) + \begin{bmatrix} \underline{v}_T(k) \\ \underline{v}(k+1) \end{bmatrix} \quad (3.12)$$

Then the estimate of $\underline{x}(k+1)$, given by Equation (3.11) evaluated at (k+1) becomes

$$\hat{\underline{x}}(k+1) = \left(\begin{bmatrix} \underline{H}_T(k)\phi(k,k+1) \\ H(k+1) \end{bmatrix}^T \begin{bmatrix} \underline{V}_T(k) & 0 \\ 0 & V(k+1) \end{bmatrix}^{-1} \begin{bmatrix} \underline{H}_T(k)\phi(k,k+1) \\ H(k+1) \end{bmatrix} \right)^{-1} \quad (3.13)$$

$$\begin{bmatrix} \underline{H}_T(k)\phi(k,k+1) \\ H(k+1) \end{bmatrix}^T \begin{bmatrix} \underline{V}_T(k) & 0 \\ 0 & V(k+1) \end{bmatrix}^{-1} \begin{bmatrix} \underline{y}_T(k) \\ \underline{y}(k+1) \end{bmatrix}$$

Equation (3.13) can be reduced to

$$\begin{aligned} \hat{\underline{x}}(k+1) = & [\phi^T(k,k+1)\underline{H}_T^T(k)\underline{V}_T^{-1}(k)\underline{H}_T(k)\phi(k,k+1) \\ & + \underline{H}^T(k+1)\underline{V}^{-1}(k+1)\underline{H}(k+1)]^{-1} \\ & [\phi^T(k,k+1)\underline{H}_T^T(k)\underline{V}_T^{-1}(k)\underline{y}_T(k) \\ & + \underline{H}^T(k+1)\underline{V}^{-1}(k+1)\underline{y}(k+1)] \end{aligned} \quad (3.14)$$

A property of the state transition matrix is that $\phi(k,k+1) = \phi^{-1}(k+1,k)$, so that $\hat{\underline{x}}(k+1)$ can also be written

$$\begin{aligned} \hat{\underline{x}}(k+1) = & F(k) [\phi^{T^{-1}}(k+1,k) H_T^T(k) V_T^{-1}(k) \underline{y}_T(k) \\ & + H^T(k+1) V^{-1}(k+1) \underline{y}(k+1)] \end{aligned} \quad (3.15)$$

where $F(k)$ is defined by

$$\begin{aligned} F(k) \triangleq & [\phi^{T^{-1}}(k+1,k) H_T^T(k) V_T^{-1}(k) H_T(k) \phi^{-1}(k+1,k) \\ & + H^T(k+1) V^{-1}(k+1) H(k+1)]^{-1} \end{aligned} \quad (3.16)$$

It is now useful to apply the following Matrix Inversion Lemma (27):

$$(A^{-1} + B^T C^{-1} B)^{-1} = A - AB^T (BAB^T + C)^{-1} BA \quad (3.17)$$

After applying the Matrix Inversion Lemma to Equation (3.16), $F(k)$

becomes

$$F(k) = \phi(k+1,k) [H_T^T(k) V_T^{-1}(k) H_T(k)]^{-1} \phi^T(k+1,k) \quad (3.18)$$

$$- \phi(k+1,k) [H_T^T(k) V_T^{-1}(k) H_T(k)]^{-1} \phi^T(k+1,k)$$

$$H^T(k+1) \{H(k+1) \phi(k+1,k) [H_T^T(k) V_T^{-1}(k) H_T(k)]^{-1}$$

$$\phi^T(k+1,k) H^T(k+1) + V(k+1)\}^{-1} H(k+1) \phi(k+1,k)$$

$$[H_T^T(k) V_T^{-1}(k) H_T(k)]^{-1} \phi^T(k+1,k)$$

Now define two new matrices, $P(k)$ and $P'(k+1)$, by

$$P(k) \triangleq [H_T^T(k)V_T^{-1}(k)H_T(k)]^{-1} \quad (3.19)$$

$$P'(k+1) \triangleq \phi(k+1,k)P(k)\phi^T(k+1,k) \quad (3.20)$$

It will be verified below that $P(k)$ and $P'(k)$ are the same as defined for the Kalman estimator in Chapter II. Substituting Equations (3.18), (3.19), and (3.20) in Equation (3.15) and rearranging gives

$$\hat{\underline{x}}(k+1) = \phi(k+1,k)[H_T^T(k)V_T^{-1}(k)H_T(k)]^{-1}\phi^T(k+1,k) \quad (3.21)$$

$$\phi^{T^{-1}}(k+1,k)H_T^T(k)V_T^{-1}(k)\underline{y}_T(k) + \{P'(k+1)$$

$$- P'(k+1)H^T(k+1)[H(k+1)P'(k+1)H^T(k+1) + V(k+1)]^{-1}$$

$$H(k+1)P'(k+1)\}H^T(k+1)V^{-1}(k+1)\underline{y}(k+1)$$

$$- \{P'(k+1)H^T(k+1)[H(k+1)P'(k+1)H^T(k+1) + V(k+1)]^{-1}$$

$$H(k+1)\phi(k+1,k)[H_T^T(k)V_T^{-1}(k)H_T(k)]^{-1}\phi^T(k+1,k)$$

$$\phi^{T^{-1}}(k+1,k)H_T^T(k)V_T^{-1}(k)\}\underline{y}_T(k)$$

Using Equation (3.11) to substitute in Equation (3.21) gives

$$\hat{\underline{x}}(k+1) = \phi(k+1,k)\hat{\underline{x}}(k) + \{P'(k+1) - P'(k+1)H^T(k+1) \quad (3.22)$$

$$\begin{aligned}
& [H(k+1)P'(k+1)H^T(k+1) + V(k+1)]^{-1}H(k+1)P'(k+1) \\
& H^T(k+1)V^{-1}(k+1)\} \underline{y}(k+1) \\
& - \{P'(k+1)H^T(k+1)[H(k+1)P'(k+1)H^T(k+1) + V(k+1)]^{-1} \\
& H(k+1)\Phi(k+1,k)\} \hat{\underline{x}}(k)
\end{aligned}$$

The coefficient of $\underline{y}(k+1)$ in the above equation can be written

$$F(k)H^T(k+1)V^{-1}(k+1) = P'(k+1)H^T(k+1) \{V^{-1}(k+1) \quad (3.23)$$

$$- [H(k+1)P'(k+1)H^T(k+1) + V(k+1)]^{-1}$$

$$H(k+1)P'(k+1)H^T(k+1)V^{-1}(k+1)\}$$

$$= P'(k+1)H^T(k+1)[H(k+1)P'(k+1)H^T(k+1) + V(k+1)]^{-1}$$

Combining Equations (3.22, 3.23) gives

$$\hat{\underline{x}}(k+1) = \Phi(k+1,k)\hat{\underline{x}}(k) \quad (3.24)$$

$$+ \{P'(k+1)H^T(k+1)[H(k+1)P'(k+1)H^T(k+1) + V(k+1)]^{-1}\}$$

$$[\underline{y}(k+1) - H(k+1)\Phi(k+1,k)\hat{\underline{x}}(k)]$$

Defining the gain as the coefficient of $\underline{y}(k+1)$,

$$K(k+1) \triangleq P'(k+1)H^T(k+1)[H(k+1)P'(k+1)H^T(k+1) + V(k+1)]^{-1} \quad (3.25)$$

the estimate becomes

$$\hat{\underline{x}}(k+1) = \Phi(k+1,k)\hat{\underline{x}}(k) \quad (3.26)$$

$$+ K(k+1)[\underline{y}(k+1) - H(k+1)\Phi(k+1,k)\hat{\underline{x}}(k)]$$

Equations (3.25, 3.26) have the same form as the Kalman estimator, but it remains to derive equations for computing $P(k)$ and $P'(k+1)$ recursively. By Equation (3.19), and using the results of Equations (3.13 through 3.16), $P(k+1)$ has the form

$$P(k+1) = [\Phi^{T^{-1}}(k+1,k)P^{-1}(k)\Phi^{-1}(k+1,k) + H^T(k+1)V(k+1)H(k+1)]^{-1} \quad (3.27)$$

which by the Matrix Inversion Lemma of Equation (3.17) becomes

$$P(k+1) = \Phi(k+1,k)P(k)\Phi^T(k+1,k) \quad (3.28)$$

$$- \Phi(k+1,k)P(k)\Phi^T(k+1,k)H^T(k+1)$$

$$[H(k+1)\Phi(k+1,k)P(k)\Phi^T(k+1,k)H^T(k+1) + V(k+1)]^{-1}$$

$$H(k+1)\Phi(k+1,k)P(k)\Phi^T(k+1,k)$$

$$\begin{aligned}
&= P'(k+1) - P'(k+1)H^T(k+1)[H(k+1)P'(k+1)H^T(k+1) \\
&\quad + V(k+1)]^{-1}H(k+1)P'(k+1)
\end{aligned}$$

where $P'(k+1)$ is defined by Equation (3.20). The matrix $P(k+1)$ can be written in a different form by adding and subtracting an additional term, as follows:

$$P(k+1) = P'(k+1) - P'(k+1)H^T(k+1) \quad (3.29)$$

$$\begin{aligned}
&[H(k+1)P'(k+1)H^T(k+1) + V(k+1)]^{-1}H(k+1)P'(k+1) \\
&+ P'(k+1)H^T(k+1)[H(k+1)P'(k+1)H^T(k+1) + V(k+1)]^{-1} \\
&\{[H(k+1)P'(k+1)H^T(k+1) + V(k+1)][H(k+1)P'(k+1) \\
&\quad H^T(k+1) + V(k+1)]^{-1} - I\} H(k+1)P'(k+1) \\
&= P'(k+1) - K(k+1)H(k+1)P'(k+1) \\
&\quad + K(k+1)[H(k+1)P'(k+1)H^T(k+1)K^T(k+1) \\
&\quad + V(k+1)K^T(k+1) - H(k+1)P'(k+1)] \\
&= [I - K(k+1)H(k+1)]P'(k+1)[I - K(k+1)H(k+1)]^T \\
&\quad + K(k+1)V(k+1)K^T(k+1)
\end{aligned}$$

Equations (3.26, 3.25, 3.20, 3.29) constitute a complete recursive algorithm for computing the estimate which minimizes the $J(k)$ cost function given by Equation (3.3), and this algorithm is identical to that of the Kalman estimator of Equations (2.9 through 2.12) for a deterministic signal. The equivalence of the Kalman estimator equations and Equation (3.11) verifies that the recursive signal model used by the Kalman estimator results in recursive estimation equations which compute estimates which are expressible as weighted sums of all past measurements.

The $P(k)$ and $P'(k+1)$ matrices defined by Equations (3.19, 3.20) can be shown to have the same meaning as in the Kalman estimator. The mean-square error of the estimator in the form of Equation (3.11) is given by

$$E\{[\underline{x}(k) - \hat{\underline{x}}(k)][\underline{x}(k) - \hat{\underline{x}}(k)]^T\} \quad (3.30)$$

$$\begin{aligned} &= E[\underline{x}(k)\underline{x}^T(k)] - \{[H_T^T(k)V_T^{-1}(k)H_T(k)]^{-1} \\ &\quad H_T^T(k)V_T^{-1}(k)\} E[\underline{y}_T(k)\underline{x}^T(k)] \\ &- E[\underline{x}(k)\underline{y}_T^T(k)] \{V_T^{-1}(k)H_T(k)[H_T^T(k)V_T^{-1}(k)H_T(k)]^{-1}\} \\ &+ \{[H_T^T(k)V_T^{-1}(k)H_T(k)]^{-1}H_T^T(k)V_T^{-1}(k)\} \\ &\quad E[\underline{y}_T(k)\underline{y}_T^T(k)] \{V_T^{-1}(k)H_T(k)[H_T^T(k)V_T^{-1}(k)H_T(k)]^{-1}\} \end{aligned}$$

Using Equation (3.7), $E[\underline{y}_T(k)\underline{y}_T^T(k)]$, $E[\underline{y}_T(k)\underline{x}^T(k)]$, and $E[\underline{x}(k)\underline{y}_T^T(k)]$

can be written as

$$E[\underline{y}_T(k)\underline{y}_T^T(k)] = H_T(k)E[\underline{x}(k)\underline{x}^T(k)]H_T^T(k) + V_T(k) \quad (3.31)$$

$$E[\underline{y}_T(k)\underline{x}^T(k)] = H_T(k)E[\underline{x}(k)\underline{x}^T(k)] \quad (3.32)$$

$$E[\underline{x}(k)\underline{y}_T^T(k)] = E[\underline{x}(k)\underline{x}^T(k)]H_T^T(k) \quad (3.33)$$

Substituting the three above equations in Equation (3.30) and combining the terms involving $E[\underline{x}(k)\underline{x}^T(k)]$ gives

$$E\{[\underline{x}(k) - \hat{\underline{x}}(k)][\underline{x}(k) - \hat{\underline{x}}(k)]^T\} = [H_T^T(k)V_T^{-1}(k)H_T(k)]^{-1} \quad (3.34)$$

which is the expression defined as $P(k)$ in Equation (3.19).

The significance of the $P'(k+1)$ matrix defined in Equation (3.20) can be shown by computing the mean-square error of $\phi(k+1,k)\hat{\underline{x}}(k)$ as an estimate of $\underline{x}(k+1)$. This mean-square error is given by

$$\begin{aligned} E\{[\underline{x}(k+1) - \phi(k+1,k)\hat{\underline{x}}(k)][\underline{x}(k+1) - \phi(k+1,k)\hat{\underline{x}}(k)]^T\} \\ = \phi(k+1,k)P(k)\phi^T(k+1,k) \end{aligned} \quad (3.35)$$

where the result is obtained by replacing $\underline{x}(k+1)$ by $\phi(k+1,k)\underline{x}(k)$.

Comparison of Equations (3.20, 3.35) shows that $P'(k+1)$ represents the mean-square error in the estimate of $\underline{x}(k+1)$ obtained by using measurements through $\underline{y}(k)$ only. In Equation (3.35), if $\underline{x}(k+1)$ is replaced by

$\Phi(k+1,k)\underline{x}(k) + G(k)\underline{u}(k)$, representing a signal with a white noise term as treated in Chapter II, then $P'(k+1)$ would have the same meaning if it were defined by

$$P'(k+1) = \Phi(k+1,k)P(k)\Phi^T(k+1,k) + G(k)R(k)G^T(k) \quad (3.36)$$

with $P(k)$ defined as the mean-square error of $\hat{\underline{x}}(k)$. Comparing Equation (3.36) with Equation (2.11) shows that the least squares curve fitting interpretation of the Kalman estimator is also valid when a signal noise term is present.

The above interpretation of the Kalman estimator also provides another interpretation of the sensitivity problem discussed in Chapter II. The Kalman estimator computes an estimate which represents a weighted least squares curve fit of the assumed model to all past data. If the model is inaccurate, error divergence can result as the curve fit becomes progressively worse as more data is taken. One way to avoid error divergence is to use a fixed memory estimator, which computes each estimate as a weighted sum of data taken over a fixed period of time immediately preceding the most recent measurement. This particular approach is easily implemented by using a signal model which has a different form than the model used in the Kalman estimator equations. This alternative model is discussed below, and the resulting estimator is discussed in terms of a curve fitting interpretation.

An Alternative Model Approach

The state variable representation of a signal is not a unique model. The transition matrix and choice of state variables for a signal with a nonrandom component requires some knowledge of the form of the signal. This knowledge can also be used to express the signal model in other forms. In the particular case of scalar measurements being linearly combined to estimate a scalar signal, the following model is useful:

$$x(k) = \underline{a}^T p(k) + r(k) \quad (3.37)$$

$$y(k) = x(k) + v(k) \quad (3.38)$$

In these equations, $x(k)$ represents the signal to be estimated, $y(k)$ is the measurement, $r(k)$ and $v(k)$ are signal and measurement noise processes, \underline{a} is an n -vector of unknown constants,

$$\underline{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \quad (3.39)$$

and

$$\underline{p}(k) = \begin{bmatrix} p_1(k) \\ p_2(k) \\ \vdots \\ p_n(k) \end{bmatrix} \quad (3.40)$$

is a vector of known functions of time, representing the model of the signal dynamics. In an example in Chapter II, the Kalman estimator was applied to a signal of the above form. The signal, represented by $x(k) = a_1 + a_2k + a_3k^2$ required the three-state Kalman estimator model given by Equations (2.17, 2.18).

With the model of Equations (3.37, 3.38), the estimate of $x(k)$ can be written

$$\hat{x}(k) = \underline{c}^T(k)\underline{m}(k) \quad (3.41)$$

where

$$\underline{m}(k) \triangleq \begin{bmatrix} y(k) \\ y(k-1) \\ \vdots \\ y(k-L) \end{bmatrix} \quad (3.42)$$

is a vector of past measurements and

$$\underline{c}(k) = \begin{bmatrix} c_0(k) \\ c_1(k) \\ \vdots \\ c_L(k) \end{bmatrix} \quad (3.43)$$

is a vector of coefficients to be computed by the estimator equations. Then $\hat{x}(k)$ is a weighted sum of past measurements, as is the Kalman estimator in the form of Equation (3.11). The parameter L , called the "memory" of the estimator, can be either fixed or increasing with time. If L increases with time, say $L = k$, then the estimator is a growing memory estimator, computing each estimate as a weighted sum of all past data. If L is a fixed integer, then the estimator is a fixed memory estimator which computes each estimate as a weighted sum of data taken over a fixed time interval of length L immediately preceding the latest measurement $y(k)$. The fixed memory estimator is often referred to as a "moving-window" or "sliding-arc" estimator. If the cross-correlation and autocorrelation functions of $r(k)$ and $v(k)$ are known, then the minimum mean-square error growing memory estimator must be equivalent to the Kalman estimator (with zero a priori information), since both would minimize mean-square error under the same assumptions. The fixed memory estimator is a suboptimal structure, so that its minimum mean-square error estimate would have a larger mean-square error than the Kalman estimator.

The equations for computing $\underline{c}(k)$ so as to minimize mean-square error were derived by Blum (9). The estimate is computed by Equation (3.41), with

$$\underline{c}(k) = Q^{-1}(k) \{ \underline{Y}(k) + D(k) [D^T(k) Q^{-1}(k) D(k)]^{-1} \quad (3.44)$$

$$[\underline{p}(k) - D^T(k) Q^{-1}(k) \underline{Y}(k)] \}$$

where

$$D(k) \triangleq \begin{bmatrix} \underline{p}^T(k) \\ \underline{p}^T(k-1) \\ \vdots \\ \underline{p}^T(k-L) \end{bmatrix} \quad (3.45)$$

$$\underline{Y}(k) \triangleq E[r(k) \underline{m}(k)] = E \begin{bmatrix} r^2(k) + r(k)v(k) \\ r(k)r(k-1) + r(k)v(k-1) \\ \vdots \\ r(k)r(k-L) + r(k)v(k-L) \end{bmatrix} \quad (3.46)$$

and

$$Q(k) \triangleq E \left[\begin{bmatrix} r(k) + v(k) \\ r(k-1) + v(k-1) \\ \vdots \\ r(k-L) + v(k-L) \end{bmatrix} \begin{bmatrix} r(k) + v(k) \\ r(k-1) + v(k-1) \\ \vdots \\ r(k-L) + v(k-L) \end{bmatrix}^T \right] \quad (3.47)^*$$

* Equation (3.47) represents a slight extension to Blum's method. Blum assumed that the signal and measurement noise sequences were independent and stationary, which are not necessary assumptions in the above equation.

A useful interpretation of the estimator of Equations (3.41, 3.44) can be obtained by considering the estimation of a deterministic signal. This requires $r(k) \equiv 0$, so Equation (3.44) becomes

$$\underline{c}(k) = \underline{V}^{-1}(k) \underline{D}(k) [\underline{D}^T(k) \underline{V}^{-1}(k) \underline{D}(k)]^{-1} \underline{p}(k) \quad (3.48)$$

where $\underline{V}(k)$ is the covariance matrix given by

$$\underline{V}(k) \triangleq \underline{E} \begin{pmatrix} \underline{v}(k) & \underline{v}(k) \\ \underline{v}(k-1) & \underline{v}(k-1) \\ \vdots & \vdots \\ \underline{v}(k-L) & \underline{v}(k-L) \end{pmatrix}^T \quad (3.49)$$

The estimate then becomes

$$\hat{\underline{x}}(k) = \underline{p}^T(k) [\underline{D}^T(k) \underline{V}^{-1}(k) \underline{D}(k)]^{-1} \underline{D}^T(k) \underline{V}^{-1}(k) \underline{m}(k) \quad (3.50)$$

This estimate can now be related to a curve fitting algorithm for the same signal by determining the estimate $\hat{\underline{a}}(k)$ of \underline{a} which minimizes the weighted least squares cost function

$$J(k) = \sum_{i=0}^L W_i^{\frac{1}{2}} [\hat{\underline{a}}^T(k) \underline{p}(k-i) - \underline{y}(k-i)]^2 \quad (3.51)$$

where the $W_i^{\frac{1}{2}}$ coefficients are weights given to the squares of deviations between measurements $\underline{y}(k-i)$ and estimates $\hat{\underline{a}}^T(k) \underline{p}(k-i)$. Equation (3.51) can also be written as

$$J(k) = [D(k)\hat{\underline{a}}(k) - \underline{m}(k)]^T W(k)[D(k)\hat{\underline{a}}(k) - \underline{m}(k)] \quad (3.52)$$

where $D(k)$ and $\underline{m}(k)$ are as previously defined and $W(k)$ is defined by

$$W(k) \triangleq \begin{bmatrix} W_k & & & & 0 \\ & W_{k-1} & & & \\ & & \ddots & & \\ & & & \ddots & \\ 0 & & & & W_{k-L} \end{bmatrix} \quad (3.53)$$

Choosing the $\hat{\underline{a}}(k)$ which makes $\frac{\partial J(k)}{\partial \hat{\underline{a}}(k)} = \underline{0}$ gives

$$\hat{\underline{a}}(k) = [D^T(k)W(k)D(k)]^{-1} D^T(k)W(k)\underline{m}(k) \quad (3.54)$$

The resulting estimate of $x(k) = \underline{a}^T \underline{p}(k)$ becomes

$$\hat{x}(k) = \hat{\underline{a}}^T(k)\underline{p}(k) = \underline{p}^T(k)\hat{\underline{a}}(k) \quad (3.55)$$

$$= \underline{p}^T(k)[D^T(k)W(k)D(k)]^{-1} D^T(k)W(k)\underline{m}(k)$$

Comparison of Equations (3.50, 3.55) shows that the minimum mean-square error estimate and the weighted least squares curve fit estimate are identical if the least squares cost function is weighted by $W(k) = V^{-1}(k)$ and $v(k)$ is a white noise sequence. In this case, $W_i = 1/\sqrt{E[v^2(k)]}$. This result corresponds to that proven earlier in this

chapter for the Kalman estimator.

The curve fitting interpretation discussed above for the estimator of Equation (3.41) is valid for both the growing memory estimator ($L=k$) and the fixed memory estimator ($L=\text{constant}$). As discussed above, the growing memory estimator is equivalent to the Kalman estimator with no a priori information, and each is interpretable as a curve fitting procedure which fits the assumed signal model to all past data in a least squares sense. The fixed memory estimator, interpreted as a curve fitting technique, fits the assumed signal model to only those measurements contained in a fixed length time interval immediately preceding the most recent measurement. In this case the signal model is used only over the latest time interval of length L sample periods, and information obtained by measurements more than L samples in the past is ignored. The bias resulting in an estimate due to an inaccurate signal model is only a function of the accuracy of the model over the last L sample periods. Therefore, the fixed memory estimator is not as sensitive to errors in the signal model as is the Kalman or growing memory estimator. Also, since the fixed memory estimator is computed by using fewer measurements than the Kalman or growing memory estimator, its resulting mean-square error will be larger than the optimum when an accurate signal model is used.

The curve fitting interpretation of the fixed memory estimator illustrates the dependence of the estimator error behavior on the choice of the memory length L . The value used for L determines the mean-square error of the estimates when an accurate signal model is used and determines the amount of bias in the estimates when an inaccurate signal

model is used. In order to make the estimates relatively insensitive to errors in the signal model, the user of the fixed memory estimator must choose a value for L so that the assumed model is an adequate approximation of the signal over time periods of length L sample periods. Small values of L imply less sensitivity to signal modeling errors, at a cost in mean-square estimation error over periods of time when the model is accurate.

The conclusions obtained in this chapter will be used in Chapter IV to postulate an estimator which will be shown to have mean-square error and modeling error sensitivity properties which represent a compromise between those of the Kalman and fixed memory estimators. That is, the estimator will be less sensitive to modeling errors than the Kalman estimator, while resulting in smaller mean-square error than the fixed memory estimator when an accurate signal model is used.

CHAPTER IV

THE AUGMENTED MEMORY APPROACH TO
DISCRETE SIGNAL ESTIMATION

Previous chapters have contained discussions of the growing memory or Kalman estimator and the fixed memory estimator with respect to mean-square error and various interpretations of the sensitivity problem. It was pointed out in Chapter III that while the Kalman estimator results in minimum mean-square error when an accurate model is used, the fixed memory estimator has an advantage in its relative insensitivity to signal modeling errors.

This chapter will present a development of equations for a different linear discrete estimator, which will be shown to have mean-square error and modeling error sensitivity properties which represent a compromise between those of the Kalman and fixed memory estimators. Interpretations of the new estimation technique developed in this chapter and the numerical results presented in Chapter V will be shown to imply that the mean-square error resulting when an accurate model is used will be smaller than that of the comparable fixed memory estimator, and the estimation error will be relatively insensitive to modeling errors when compared with the Kalman estimator.

The procedure used in this chapter to develop the new estimation technique is to first postulate an estimator structure based on intuitive arguments supported by the results of the previous chapter. After

the estimator equations are developed, the resulting estimator will then be interpreted in terms of an equivalent constrained growing memory estimator and in terms of the sensitivity of the estimates to errors in the signal model. The sensitivity properties of the estimator will be interpreted by considering the bias which results in the estimates when an inaccurate signal model is used.

The Augmented Memory Estimator

The linear discrete estimator developed below, designated the "augmented memory estimator," assumes a signal model and estimator computational structure similar to that of the fixed memory estimator discussed in the previous chapter. The signal and measurement models have the form

$$x(k) = \underline{a}^T \underline{p}(k) + r(k) \quad (4.1)$$

$$y(k) = x(k) + v(k) \quad (4.2)$$

where, as for the fixed memory estimator, \underline{a} is an unknown vector of n constants and $\underline{p}(k)$ is a vector of known, linearly independent, functions of time:

$$\underline{p}(k) = \begin{bmatrix} p_1(k) \\ p_2(k) \\ \vdots \\ p_n(k) \end{bmatrix} \quad (4.3)$$

The noise sequences $r(k)$ and $v(k)$ may be zero mean colored noise processes with a nonzero cross-correlation function. The noise processes are not required to be stationary. It is assumed that the correlation functions defined below are known at all sampling instants:

$$\rho(k,j) \triangleq E[r(k)r(j)] \quad (4.4)$$

$$\sigma(k,j) \triangleq E[v(k)v(j)] \quad (4.5)$$

$$\phi(k,j) \triangleq E[r(k)v(j)] \quad (4.6)$$

One additional assumption must be made on the nature of the correlation functions in order to assure that the augmented memory estimator will not require an increasing number of computations as k increases. This assumption is that the correlation functions defined by Equations (4.4, 4.5, 4.6) are zero when the noise product terms are separated in time by more than M . That is, $\rho(k,j)$, $\sigma(k,j)$, and $\phi(k,j)$ are all zero for all (k,j) such that $|k-j| > M$. This restriction becomes necessary in the derivation of equations given in the Appendix.

The augmented memory estimator computes an estimate of $x(k)$ by

$$\hat{x}(k) = \underline{w}^T(k)\underline{m}(k) \quad (4.7)$$

where $\underline{w}(k)$ is a vector of $L + 1$ weighting coefficients which must be computed,

$$\underline{w}(k) = \begin{bmatrix} w_0(k) \\ w_1(k) \\ \vdots \\ w_L(k) \end{bmatrix} \quad (4.8)$$

and $\underline{m}(k)$ is a vector containing the latest measurement and the last L estimates:

$$\underline{m}(k) \triangleq \begin{bmatrix} y(k) \\ \hat{x}(k-1) \\ \hat{x}(k-2) \\ \vdots \\ \hat{x}(k-L) \end{bmatrix} \quad (4.9)$$

Equations (4.1, 4.2, 4.7, 4.9) show that with L constant the augmented memory estimator uses the signal model and computational structure of the fixed memory estimator discussed in Chapter III. The important difference is that the $\underline{m}(k)$ vector of Equation (4.9) contains a fixed number of past estimates where the fixed memory estimator used past measurements. If past estimates are better statistical approximations to past values of the signal than are past measurements, then the augmented memory estimator with an accurate signal model should produce estimates having smaller mean-square error than those computed by the fixed memory estimator. Also, the computational structure of the augmented estimator implies error sensitivity properties comparable to those of the fixed memory estimator. The use of past estimates improves

the information in the time interval over which the signal model is to be used, but the time interval remains fixed in length. At this point in the development, these conclusions are based on intuitive arguments. Later in this chapter the augmented memory estimator will be interpreted in greater detail in terms of its structure and sensitivity to signal modeling errors.

The constant parameter L in Equation (4.9) is analogous to the memory length in the fixed memory estimator, but for the augmented memory estimator L will be defined as the "interval of confidence." Later in this chapter, it will be shown that L can be interpreted as a parameter which is chosen to represent the number of sample periods over which the assumed model can be used with confidence.

Subject to the assumed estimator structure of Equation (4.7), the weighting coefficients $w(k)$ are chosen so as to minimize mean-square estimation error defined by

$$P(k) \triangleq E\{[x(k) - \hat{x}(k)]^2\} \quad (4.10)$$

An additional constraint is imposed on the choice of $w(k)$ by requiring the resulting estimate to be unbiased, so that

$$E[x(k) - \hat{x}(k)] = 0 \quad (4.11)$$

This requirement insures that the estimator give perfect estimates of the signal when the $r(k)$ and $v(k)$ noise processes are identically zero.

Equations for choosing the weighting coefficients which minimize $P(k)$ subject to the constraint of Equation (4.11) are derived below. Interpretations of the augmented memory estimator will then be presented to illustrate its properties and compare its structure to that of the growing memory and fixed memory estimators.

Computation of Weighting Coefficients

Consider first the conditions imposed on the weighting coefficients by the requirement that the estimate be unbiased. Combining Equations (4.1, 4.7, 4.11) results in a constraint of the form

$$\begin{aligned} E[\underline{w}^T(k)\underline{m}(k)] &= \underline{w}^T(k)E[\underline{m}(k)] & (4.12) \\ &= \underline{w}^T(k)D(k)\underline{a} \\ &= \underline{a}^T D^T(k)\underline{w}(k) \end{aligned}$$

where

$$D(k) \triangleq \begin{bmatrix} P^T(k) \\ P^T(k-1) \\ \vdots \\ P^T(k-L) \end{bmatrix} \quad (4.13)$$

Combining Equations (4.11, 4.12) gives

$$\underline{a}^T D^T(k)\underline{w}(k) = \underline{a}^T \underline{p}(k) \quad (4.14)$$

Equation (4.14) must be true for any vector \underline{a} , so the resulting constraint relation is

$$\underline{D}^T(k)\underline{w}(k) = \underline{p}(k) \quad (4.15)$$

The weighting coefficients $\underline{w}(k)$ must be computed so as to minimize the mean-square error of Equation (4.10) while satisfying the constraints of Equation (4.15). Expanding Equation (4.10) with $\underline{x}(k)$ and $\hat{\underline{x}}(k)$ in the forms given by Equations (4.1, 4.7) results in the mean-square error having the form

$$\begin{aligned} P(k) = & \underline{w}^T(k) E[\underline{m}(k)\underline{m}^T(k)]\underline{w}(k) & (4.16) \\ & - 2\underline{a}^T \underline{p}(k) E[\underline{m}^T(k)\underline{w}(k)] \\ & - 2E[r(k)\underline{m}^T(k)] \underline{w}(k) + [\underline{a}^T \underline{p}(k)]^2 \\ & + 2E[r(k)]\underline{a}^T \underline{p}(k) + E[r^2(k)] \end{aligned}$$

Using the assumed properties of the noise process $r(k)$ and the constraints of Equation (4.15), the terms in Equation (4.16) can be written

$$\begin{aligned} \underline{w}^T(k)E[\underline{m}(k)\underline{m}^T(k)]\underline{w}(k) = & \underline{w}^T(k)D(k)\underline{a}\underline{a}^T D^T(k)\underline{w}(k) & (4.17) \\ & + \underline{w}^T(k)S(k)\underline{w}(k) \end{aligned}$$

$$= [\underline{a}_p^T(k)]^2 + \underline{w}^T(k)S(k)\underline{w}(k)$$

$$\underline{a}_p^T(k) E[\underline{m}^T(k)\underline{w}(k)] = [\underline{a}_p^T(k)]^2 \quad (4.18)$$

$$E[r(k)\underline{m}^T(k)]\underline{w}(k) = \underline{\gamma}^T(k)\underline{w}(k) \quad (4.19)$$

$$E[r(k)]\underline{a}_p^T(k) = 0 \quad (4.20)$$

where $S(k)$ and $\underline{\gamma}(k)$ are defined by

$$S(k) \triangleq E \begin{bmatrix} r(k) + v(k) & r(k) + v(k) \\ r(k-1) + n(k-1) & r(k-1) + n(k-1) \\ r(k-2) + n(k-2) & r(k-2) + n(k-1) \\ \vdots & \vdots \\ r(k-L) + n(k-L) & r(k-L) + n(k-L) \end{bmatrix}^T \quad (4.21)$$

$$\underline{\gamma}(k) \triangleq E[r(k)\underline{m}(k)] \quad (4.22)$$

and $n(j)$ is defined as the estimation error by

$$n(j) \triangleq \hat{x}(j) - x(j) \quad (4.23)$$

Substituting Equations (4.17) through (4.20) in Equation (4.16) gives

$$P(k) = \underline{w}^T(k)S(k)\underline{w}(k) - 2\underline{\gamma}^T(k)\underline{w}(k) + \rho(k,k) \quad (4.24)$$

In order to minimize $P(k)$ subject to the constraints of Equation (4.15),

the constraints are adjoined to $P(k)$ by a Lagrange multiplier vector $2\underline{\lambda}(k)$, producing a function $P'(k)$ defined by

$$P'(k) = \underline{w}^T(k)S(k)\underline{w}(k) - 2\underline{\gamma}^T(k)\underline{w}(k) + \rho(k,k) \quad (4.25)$$

$$+ [D^T(k)\underline{w}(k) - \underline{p}(k)]^T[2\underline{\lambda}(k)]$$

The function $P'(k)$ is minimized with respect to $\underline{w}(k)$ by computing the $\underline{w}(k)$ and $\underline{\lambda}(k)$ which give

$$\frac{\partial P'(k)}{\partial \underline{w}(k)} = \underline{0} \quad (4.26)$$

and

$$\frac{\partial P'(k)}{\partial \underline{\lambda}(k)} = \underline{0} \quad (4.27)$$

Substituting Equation (4.25) in Equations (4.26, 4.27) gives

$$S(k)\underline{w}(k) - [\underline{\gamma}(k) + D(k)\underline{\lambda}(k)] = \underline{0} \quad (4.28)$$

$$\underline{p}(k) - D^T(k)\underline{w}(k) = \underline{0} \quad (4.29)$$

Combining Equations (4.28, 4.29) to eliminate $\underline{w}(k)$ gives

$$\underline{\lambda}(k) = [D^T(k)S^{-1}(k)D(k)]^{-1} [\underline{p}(k) - D^T(k)S^{-1}(k)\underline{\gamma}(k)] \quad (4.30)$$

Substituting Equation (4.30) for $\underline{\lambda}(k)$ in Equation (4.28) gives the desired equation for $\underline{w}(k)$:

$$\begin{aligned} \underline{w}(k) = S^{-1}(k) \{ & \underline{\gamma}(k) \\ & + D(k) [D^T(k)S^{-1}(k)D(k)]^{-1} [\underline{p}(k) \\ & - D^T(k)S^{-1}(k)\underline{\gamma}(k)] \} \end{aligned} \quad (4.31)$$

This value of $\underline{w}(k)$ is used in Equation (4.7) to produce the augmented memory estimator.

The above derivation does not include equations for computing either $S(k)$ or $\underline{\gamma}(k)$. Inspection of Equations (4.21, 4.22) shows that $S(k)$ and $\underline{\gamma}(k)$ require correlation terms involving the error $n(k-j)$ for $j=1,2,\dots,L$. The evaluation of $S(k)$ and $\underline{\gamma}(k)$ is accomplished by a recursive algorithm which computes the necessary error correlation terms. This algorithm is derived and related to digital computer implementation in the Appendix. The recursive algorithm is entered after L measurements have been taken, with the first L measurements being considered to be the first L estimates. If measurements start with $y(0)$, this means that the recursive algorithm is first used to compute $\hat{x}(L)$, with $\underline{m}(L)$, $\underline{\gamma}(L)$, and $S(L)$ initialized by

$$\underline{m}(L) = \begin{bmatrix} \hat{y}(L) \\ y(L-1) \\ \vdots \\ y(0) \end{bmatrix} \quad (4.32)$$

$$\underline{\gamma}(L) = E[r(L)\underline{m}(L)] = \begin{bmatrix} \rho(L,L) + \phi(L,L) \\ \rho(L,L-1) + \phi(L,L-1) \\ \vdots \\ \rho(L,0) + \phi(L,0) \end{bmatrix} \quad (4.33)$$

$$S(L) = E \left[\begin{bmatrix} r(L) + v(L) \\ r(L-1) + v(L-1) \\ \vdots \\ r(0) + v(0) \end{bmatrix} \begin{bmatrix} r(L) + v(L) \\ r(L-1) + v(L-1) \\ \vdots \\ r(0) + v(0) \end{bmatrix}^T \right] \quad (4.34)$$

For $k \geq L$, the vector $\underline{m}(L)$ stores the estimates $\hat{x}(k)$ rather than measurements $y(k)$, and the recursive algorithm is used to compute $S(k)$ and $\underline{\gamma}(k)$.

The procedure for computing estimates by the augmented memory estimator method is summarized by the flow diagram of Figure 7. This diagram is especially suitable for implementation of the augmented memory estimator by a digital computer program. The necessary equations and procedures are referenced on the diagram.

Interpretations of the Augmented Memory Estimator

The augmented memory estimator expressed by Equation (4.7) and computed by the above procedure insures a fixed number of computations after each sample. However, the above representation of the augmented

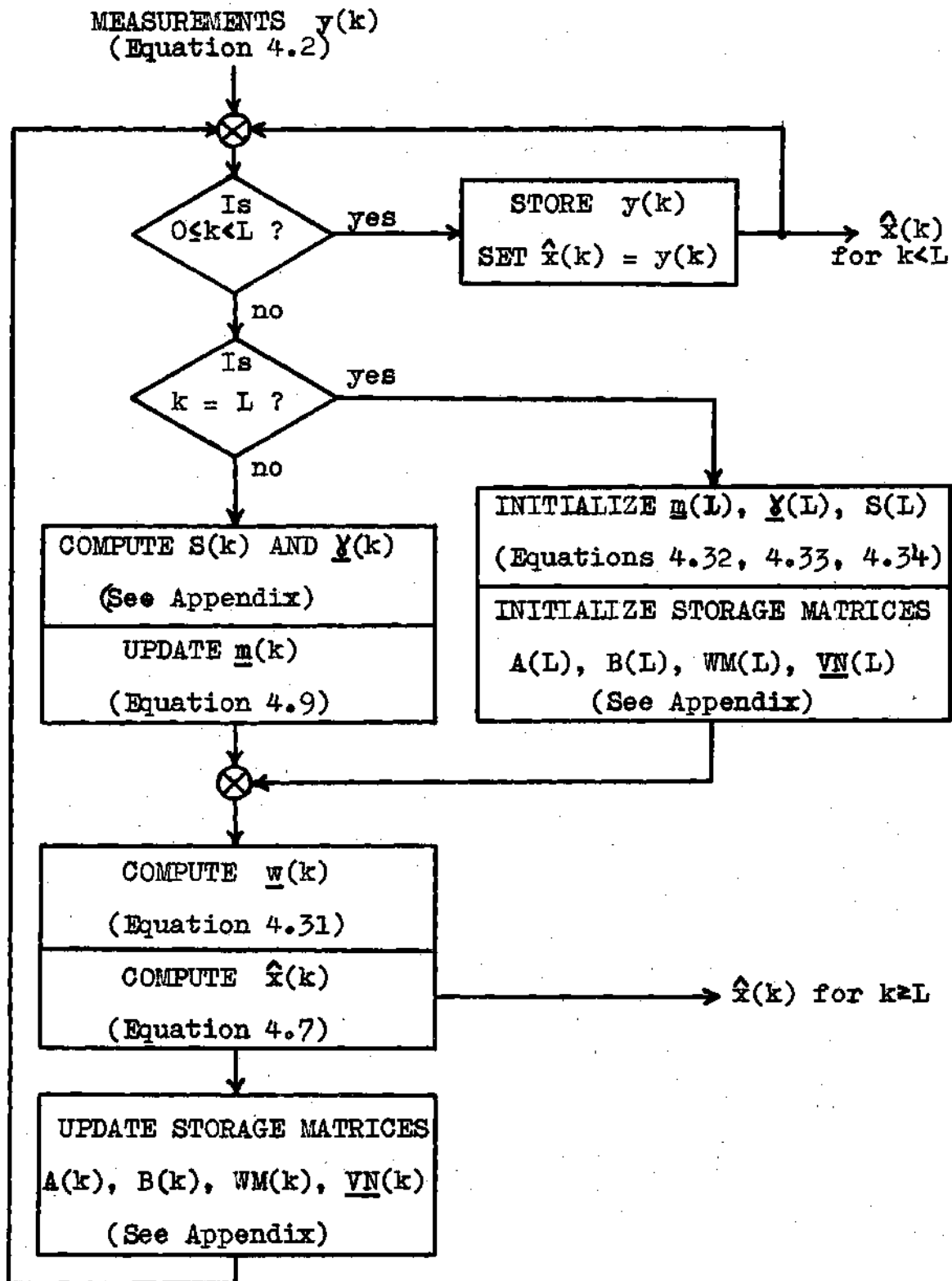


Figure 7. Flow Diagram of the Augmented Memory Estimator

memory estimator only partially explains its advantages or justifies its use as opposed to the fixed memory or Kalman estimator. It remains to interpret the augmented memory estimator in ways which will permit an evaluation of its performance with respect to mean-square error and sensitivity to errors in the dynamical model of the signal.

Because of its dependence on past estimates contained in $\underline{m}(k)$, the augmented memory estimator may be interpreted in terms of an equivalent growing memory estimator. The nature of the equivalent growing memory estimator can be illustrated by considering the augmented memory estimator with an interval of confidence of $L = 3$. Then the augmented memory estimator computes $\hat{x}(k)$ by

$$\begin{aligned} \hat{x}(k) = & \bar{w}_0(k)y(k) + \bar{w}_1(k)\hat{x}(k-1) & (4.35)^* \\ & + \bar{w}_2(k)\hat{x}(k-2) + \bar{w}_3(k)\hat{x}(k-3) \end{aligned}$$

where $\bar{w}_0(k)$, $\bar{w}_1(k)$, $\bar{w}_2(k)$, and $\bar{w}_3(k)$ are coefficients computed after the measurement $y(k)$ is obtained. The estimator is initialized by using the first L measurements as the first L estimates, so

$$\hat{x}(0) = y(0) \quad (4.36)$$

$$\hat{x}(1) = y(1) \quad (4.37)$$

* The bar over a coefficient in this and later equations represents a coefficient which is computable. All other terms which appear are predetermined.

$$\hat{x}(2) = y(2) \quad (4.38)$$

Then $\hat{x}(3)$ has the form

$$\hat{x}(3) = \bar{w}_0(3)y(3) + \bar{w}_1(3)\hat{x}(2) + \bar{w}_2(3)\hat{x}(1) \quad (4.39)$$

$$+ \bar{w}_3(3)\hat{x}(0) = \bar{w}_0(3)y(3) + \bar{w}_1(3)y(2)$$

$$+ \bar{w}_2(3)y(1) + \bar{w}_3(3)y(0)$$

This estimate $\hat{x}(3)$ is identical to the estimate which would be computed by the fixed memory, unbiased, minimum mean-square error estimator with a memory of three samples. However, the fixed memory estimator and the augmented memory estimator cease to be equivalent for $k > L$. For the case being considered, $\hat{x}(4)$ has the form

$$\hat{x}(4) = \bar{w}_0(4)y(4) + \bar{w}_1(4)\hat{x}(3) \quad (4.40)$$

$$+ \bar{w}_2(4)\hat{x}(2) + \bar{w}_3(4)\hat{x}(1)$$

Substituting Equations (4.39, 4.38, 4.37) for $\hat{x}(3)$, $\hat{x}(2)$, and $\hat{x}(1)$ in Equation (4.40) gives

$$\hat{x}(4) = \bar{w}_0(4)y(4) + [\bar{w}_1(4)\bar{w}_0(3)]y(3) \quad (4.41)$$

$$+ [\bar{w}_1(4)w_1(3) + \bar{w}_2(4)]y(2)$$

$$+ [\bar{w}_1(4)w_2(3) + \bar{w}_3(4)]y(1)$$

$$+ [\bar{w}_1(4)w_3(3)]y(0)$$

Continuing the process of substituting for past estimates in Equation (4.35) to produce estimates as weighted sums of past measurements results in

$$\hat{x}(5) = \bar{w}_0(5)y(5) + \bar{w}_1(5)\hat{x}(4) + \bar{w}_2(5)\hat{x}(3) + \bar{w}_3(5)\hat{x}(2) \quad (4.42)$$

$$= \bar{w}_0(5)y(5) + [\bar{w}_1(5)w_0(4)]y(4)$$

$$+ [\bar{w}_1(5)w_1(4)w_0(3) + \bar{w}_2(5)w_0(3)]y(3)$$

$$+ [\bar{w}_1(5)w_1(4)w_1(3) + \bar{w}_1(5)w_2(4) + \bar{w}_2(5)w_1(3)$$

$$+ \bar{w}_3(5)]y(2) + [\bar{w}_1(5)w_1(4)w_2(3) + \bar{w}_1(5)w_3(4)$$

$$+ \bar{w}_2(5)w_2(3)]y(1) + [\bar{w}_1(5)w_1(4)w_3(3) + \bar{w}_2(5)w_3(3)]y(0)$$

$$\hat{x}(6) = \bar{w}_0(6)y(6) + \bar{w}_1(6)\hat{x}(5) + \bar{w}_2(6)\hat{x}(4) + \bar{w}_3(6)\hat{x}(3) \quad (4.43)$$

$$= \bar{w}_0(6)y(6) + [\bar{w}_1(6)w_0(5)]y(5)$$

$$\begin{aligned}
& + [\bar{w}_1(6)w_1(5)w_0(4) + \bar{w}_2(6)w_0(4)]y(4) \\
& + [\bar{w}_1(6)w_1(5)w_1(4)w_0(3) + \bar{w}_1(6)w_2(5)w_0(3) \\
& + \bar{w}_2(6)w_1(4)w_0(3) + \bar{w}_3(6)w_0(3)]y(3) \\
& + [\bar{w}_1(6)w_1(5)w_1(4)w_1(3) + \bar{w}_1(6)w_1(5)w_2(4) \\
& + \bar{w}_1(6)w_2(5)w_1(3) + \bar{w}_1(6)w_3(5) + \bar{w}_2(6)w_1(4)w_1(3) \\
& + \bar{w}_2(6)w_2(4) + \bar{w}_3(6)w_1(3)]y(2) \\
& + [\bar{w}_1(6)w_1(5)w_1(4)w_2(3) + \bar{w}_1(6)w_1(5)w_3(4) \\
& + \bar{w}_1(6)w_2(5)w_2(3) + \bar{w}_2(6)w_1(4)w_2(3) \\
& + \bar{w}_2(6)w_3(4) + \bar{w}_3(6)w_2(3)]y(1) \\
& + [\bar{w}_1(6)w_1(5)w_1(4)w_3(3) + \bar{w}_1(6)w_2(5)w_3(3) \\
& + \bar{w}_2(6)w_1(4)w_3(3) + \bar{w}_3(6)w_3(3)]y(0)
\end{aligned}$$

Using the above procedure, the augmented memory estimator for $L=3$ can be expressed as a growing memory estimator which computes estimates having the form

$$\begin{aligned}
\hat{x}(k) = & \bar{w}_0(k)y(k) + [\bar{w}_1(k)w_0(k-1)]y(k-1) & (4.44) \\
& + [\bar{w}_1(k)w_1(k-1)w_0(k-2) + \bar{w}_2(k)w_0(k-2)]y(k-2) \\
& + [\bar{w}_1(k)d_{13}(k) + \bar{w}_2(k)d_{23}(k) + \bar{w}_3(k)d_{33}(k)]y(k-3) \\
& + [\bar{w}_1(k)d_{14}(k) + \bar{w}_2(k)d_{24}(k) + \bar{w}_3(k)d_{34}(k)]y(k-4) \\
& \vdots \\
& + [\bar{w}_1(k)d_{1k}(k) + \bar{w}_2(k)d_{2k}(k) + \bar{w}_3(k)d_{3k}(k)]y(0)
\end{aligned}$$

The $d_{ij}(k)$ coefficients in the above equation represent numbers which are computed prior to the k th sample. Therefore, as far as the k th estimate is concerned, the $d_{ij}(k)$ terms are predetermined constants.

Inspection of Equation (4.44) leads to an interpretation of the augmented memory estimator in terms of an equivalent growing memory estimator. The augmented memory estimator is expressed as a weighted sum of all past data in the same form as the optimum growing memory estimator described earlier, namely

$$\hat{x}(k) = \sum_{i=0}^k c_i(k)y(k-i) \quad (4.45)$$

With both types of estimators required to be unbiased, the only difference is in the values of the $c_i(k)$ coefficients used in Equation (4.45). For either estimator, the $c_i(k)$ coefficients must satisfy the constraint relation represented by an extension of Equation (4.15)

obtained by replacing $\underline{w}(k)$ in that equation by $\underline{c}(k)$, the vector of the $c_i(k)$ coefficients, and allowing the $D^T(k)$ matrix to include the model over the entire past. The constraint relation becomes

$$D_T^T(k) \underline{c}(k) = \underline{p}(k) \quad (4.46)$$

with $D_T(k)$ defined by

$$D_T(k) \triangleq \begin{bmatrix} \underline{p}^T(k) \\ \underline{p}^T(k-1) \\ \underline{p}^T(k-2) \\ \vdots \\ \underline{p}^T(0) \end{bmatrix} \quad (4.47)$$

The difference between the augmented memory estimator and the optimum growing memory estimator is evident in Equations (4.44, 4.45). Only the four coefficients in the $\underline{w}(k)$ vector are computed by the estimator equations, while there are $k+1$ measurements to be linearly weighted. The estimator has four degrees of freedom corresponding to the four coefficients to be computed. These four degrees of freedom can be interpreted as the freedom to choose the coefficients of the four most recent measurements, and once these four coefficients are chosen all the other $k-3$ coefficients are fixed. In general, for any interval of confidence L , the augmented memory estimator has $L+1$ degrees of freedom, corresponding to the choices of values for the coefficients of the most recent $L+1$ measurements. The optimum growing

memory estimator has $k+1$ degrees of freedom, corresponding to the freedom to choose the coefficients of all $k+1$ measurements. Therefore, the augmented memory estimator is suboptimal with respect to mean-square error minimization. A later interpretation and examples will be used to show that it has advantages over the optimal growing memory estimator in its sensitivity to signal modeling errors.

The augmented memory estimator in the form of Equations (4.44, 4.45) is equivalent to an unbiased growing memory estimator which minimizes mean-square error subject to an increasing number of constraints on the coefficients of past measurements. These constraints are the result of the decrease in the number of degrees of freedom as described above. The additional constraints can be written in the form

$$Z(k)\underline{c}(k) = \underline{0} \quad (4.48)$$

where $\underline{c}(k)$ is the vector of coefficients in Equation (4.45), $Z(k)$ is defined by

$$Z(k) \triangleq \begin{bmatrix} 0 & z_{11} & z_{12} & z_{13} & 1 & 0 & 0 & \dots & 0 \\ 0 & z_{21} & z_{22} & z_{23} & 0 & 1 & 0 & \dots & 0 \\ \vdots & & & & & & & & \\ \vdots & & & & & & & & \\ 0 & z_{k-3,1} & z_{k-3,2} & z_{k-3,3} & 0 & 0 & 0 & \dots & 1 \end{bmatrix} \quad (4.49)$$

and the z_{ij} terms can be computed from the relationships given by Equation (4.44). The above $Z(k)$ matrix represents the $L=3$ case. In general, $Z(k)$ is a $(k-L)$ by $(k+1)$ matrix and Equation (4.48) represents a set of $k-L$ equations which the $\underline{c}(k)$ vector of coefficients must satisfy in addition to the unbiased constraints imposed by Equation (4.46).

The equivalence of the augmented memory estimator and a constrained growing memory estimator only serves to show that the former is different from the optimum growing memory estimator. The original form for the augmented memory estimator, given by Equation (4.7), is the simplest way to implement the estimator because its computational complexity does not depend on k .

An additional interpretation of the augmented memory estimator can be used to illustrate the sensitivity of the error to inaccuracies in the signal dynamical model. This interpretation is obtained by computing equations for the error bias resulting from the use of an inaccurate signal dynamical model. Consider a signal having the form of Equation (4.1). The augmented memory estimator assumes a signal model $x_m(k)$ of the form

$$x_m(k) = a_{-m}^T p_m(k) + r(k) \quad (4.50)$$

where the m subscript emphasizes that this is a model of the actual signal. Suppose also that the noise and measurement models are accurate. The augmented memory estimator computes an estimate of

$x(k)$ by Equation (4.7), repeated here for convenience:

$$\hat{x}(k) = \underline{w}^T(k) \underline{m}(k) \quad (4.51)$$

The error bias resulting when $\underline{p}_m(k)$ is not the same as $\underline{p}(k)$ is given by

$$E[n(k)] = \underline{w}^T(k) E[\underline{m}(k)] - \underline{a}^T \underline{p}(k) \quad (4.52)$$

with $n(k)$ defined by Equation (4.23). But $E[\underline{m}(k)]$ can be expressed as

$$E[\underline{m}(k)] = D(k) \underline{a} + E[\underline{N}(k)] \quad (4.53)$$

Where $D(k)$ is defined by Equation (4.13) and $\underline{N}(k)$ is defined by

$$\underline{N}(k) \triangleq \begin{bmatrix} r(k) + v(k) \\ r(k-1) + n(k-1) \\ r(k-2) + n(k-2) \\ \vdots \\ r(k-L) + n(k-L) \end{bmatrix} \quad (4.54)$$

If a fixed memory estimator is used instead of the augmented memory estimator, $\underline{m}(k)$ contains only measurements, all of the $n(k-i)$ terms in Equation (4.54) are replaced by $v(k-i)$, and the expected value of $\underline{N}(k)$ becomes zero. In this case, the resulting bias, defined as $b(k)$, would be given by

$$b(k) = \underline{w}^T(k)D(k)\underline{a} - \underline{a}^T\underline{p}(k) \quad (4.55)$$

Therefore, $b(k)$ can be considered a measure of the accuracy of the model over the time interval $(k-L, k)$. The bias resulting in the augmented memory estimator can now be written in terms of present and past values of $b(k)$.

With $\hat{x}(k) = y(k)$ for $k < L$, the bias is zero for $k < L$. For $k = L$ the augmented memory estimator is equivalent to the fixed memory estimator, so

$$E[n(L)] = b(L) \quad (4.56)$$

Combining Equations (4.52, 4.53, 4.55) gives the bias in the augmented memory estimator for $k > L$ in the form

$$E[n(k)] = b(k) + \underline{w}^T(k)E[\underline{N}(k)] \quad (4.57)$$

$$= b(k) + \underline{w}^T(k) \begin{bmatrix} 0 \\ E[n(k-1)] \\ E[n(k-2)] \\ \vdots \\ E[n(k-L)] \end{bmatrix}$$

Therefore, the error bias in the augmented memory estimator for $k > L$ is the sum of $b(k)$ and a weighted sum of past error biases. To illustrate the nature of the bias, it is helpful to carry out some of the error

bias expressions, using Equations (4.56, 4.57) as follows:

$$E[n(k)] = 0 \text{ for } k=0,1,2,\dots,L-1 \quad (4.58)$$

$$E[n(L)] = b(L) \quad (4.59)$$

$$E[n(L+1)] = b(L+1) + w_1(L+1) E[n(L)] \quad (4.60)$$

$$= b(L+1) + w_1(L+1)b(L)$$

$$E[n(L+2)] = b(L+2) + w_1(L+2)E[n(L+1)] \quad (4.61)$$

$$+ w_2(L+1)E[n(L)] = b(L+2)$$

$$+ w_1(L+2)b(L+1) + [w_1(L+2)w_1(L+1)$$

$$+ w_2(L+1)]b(L)$$

$$E[n(L+3)] = b(L+3) + w_1(L+3)E[n(L+2)] \quad (4.62)$$

$$+ w_2(L+3)E[n(L+1)] + w_3(L+3)E[n(L)]$$

$$= b(L+3) + w_1(L+3)b(L+2)$$

$$+ [w_1(L+3)w_1(L+2) + w_2(L+3)]b(L+1)$$

$$\begin{aligned}
& + [w_1(L+3)w_1(L+2)w_1(L+1) + w_1(L+3)w_2(L+1) \\
& + w_2(L+3)w_1(L+1) + w_3(L+3)]b(L)
\end{aligned}$$

Continuing this procedure results in a bias of the form

$$E[n(k)] = b(k) + \sum_{i=1}^{k-L} q_i(k)b(k-i) \quad (4.63)$$

where the $q_i(k)$ are coefficients which are sums of products of elements in $w(k)$, $w(k-1)$, ..., $w(L+1)$. Equations (4.55, 4.63) can be used to determine whether an unacceptable error bias results when an approximation to a correct signal model is used. The correct model must be known to obtain numerical bias results from Equations (4.55, 4.63). However, Equation (4.63) does show that the error bias in the augmented memory estimator is a weighted sum of $b(k-i)$ bias terms, each of which is a measure of the inaccuracy of the signal model only over a time period of L sample periods.

The interpretations of the augmented memory estimator discussed above provide some insight into the structure of the estimator. The interpretation of the augmented memory estimator as a constrained growing memory estimator shows that the augmented memory estimator is suboptimal with respect to mean-square error minimization. However, the mean-square error of the augmented memory estimator is smaller than that obtained by the comparable fixed memory estimator when the signal model is accurate, as will be illustrated in the following chapter for the case of polynomial signals. Equations developed for the bias

resulting from the use of an inaccurate signal model in the augmented memory estimator equations show that the bias is a weighted sum of terms which depend on the accuracy of the signal model only over past time periods of length L sample periods. Since the results of Chapter III show that the Kalman estimator fits the assumed model to the data over the entire past, these bias equations imply sensitivity advantages over the Kalman estimator when the assumed model is an adequate approximation of the signal over all past time periods of length L sample periods. This conclusion must be verified by numerical results, which could be obtained by using the bias equations of this chapter or by simulating various combinations of signals and inaccurate models. The following chapter uses the latter approach, presenting the results of digital computer simulations of various signals with inaccurate polynomial models.

CHAPTER V

NUMERICAL COMPARISONS AND EXAMPLES

In Chapter IV the augmented memory estimator was proposed and its equations were derived. The estimator structure was interpreted in terms of an equivalent growing memory estimator and in terms of the bias resulting from the use of an inaccurate dynamical model. This chapter will consist of numerical results of digital computer simulations which illustrate the behavior of the augmented memory estimator with respect to mean-square error when an accurate model is used and the sensitivity of the estimator to errors in the signal dynamical model.

As stated in Chapter IV, the purpose for using the augmented memory estimator is to obtain an estimate which has two desirable properties. First, the mean-square error of the augmented memory estimator should be smaller than that of the fixed memory estimator with a memory of L samples. Also, the estimator error should be less sensitive to dynamical modeling errors than is the estimation error of the growing memory or Kalman estimator. The numerical results presented below will verify that the augmented memory estimator has these properties for examples considered in this chapter. Each of these examples consists of a deterministic signal corrupted by white Gaussian measurement noise of constant variance. In each case the augmented memory estimator is used with a model which assumes a noiseless polynomial signal.

To illustrate mean-square error behavior when an accurate model is used, consider a polynomial signal of the form $x(k) = \underline{a}^T \underline{p}(k)$ with $\underline{p}(k)$ given by

$$\underline{p}(k) = \begin{bmatrix} 1 \\ k \\ k^2 \end{bmatrix} \quad (5.1)$$

Measurements are defined by Equation (4.2), with $v(k)$ a white Gaussian noise process with zero mean and constant variance. For this signal, the fixed memory estimator has a mean-square error which is constant for all $k > L$. The Kalman estimator has a mean-square error which is asymptotic to zero. The mean-square error of the augmented memory estimator is the same as that of the fixed memory estimator for $k = L$, but decreases for $k > L$ until a steady-state value is obtained. Figure 8 illustrates the mean-square error improvement obtained by using the augmented memory estimator with an interval of confidence L as opposed to the fixed memory estimator with a memory length L . The mean-square error improvement is computed by subtracting the mean-square error in the augmented memory estimator at $k = 10L$ from the mean-square error of the fixed memory estimator and expressing the difference as a percentage of the mean-square error of the fixed memory estimator. Figure 9 shows the mean-square error improvement computed in the same manner for a polynomial signal of the form $x(k) = \underline{a}^T \underline{p}(k)$ with $\underline{p}(k)$ given by

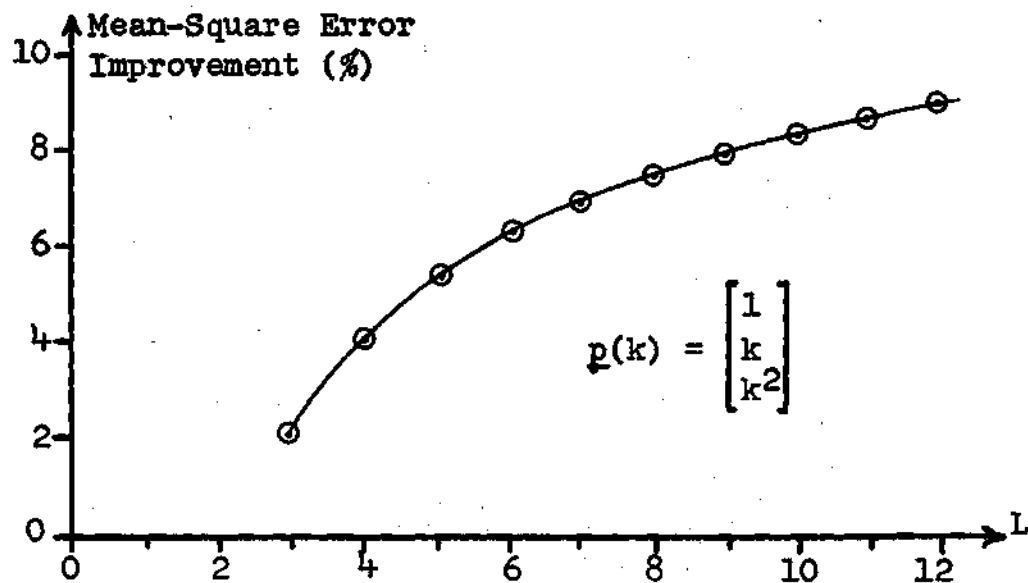


Figure 8. Mean-Square Error Improvement in the Augmented Memory Estimator with Second Degree Polynomial Signal

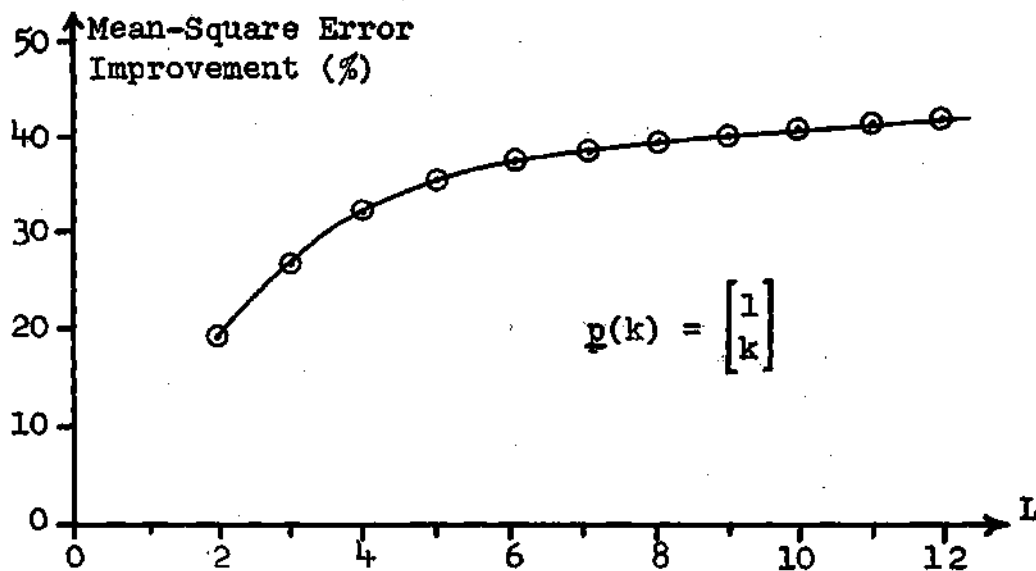


Figure 9. Mean-Square Error Improvement in the Augmented Memory Estimator with First Degree Polynomial Signal

$$\underline{p}(k) = \begin{bmatrix} 1 \\ k \end{bmatrix} \quad (5.2)$$

Figures 8 and 9 show that the improvement in mean-square error increases as the interval of confidence L increases.

In order to illustrate the sensitivity of the augmented memory estimator to dynamical modeling errors, three examples are presented below. In each example the augmented memory estimator assumes a noiseless polynomial signal model which is not a correct representation of the true signal. The first example uses a signal which is a polynomial of higher degree than that of the assumed model, and the last two examples consist of signals which are generated as solutions to nonlinear differential equations. For each example the error at each sampling instant is averaged over 30 computer runs, each run with a different measurement noise sequence. By averaging over 30 computer runs, any bias in the error will be reinforced and the random component of the error will be reduced. The error averages are plotted versus time to show the error bias behavior. In addition, the average error and sample error variance are computed for each computer run, and the 30 run averages of these quantities are computed. The averaged error is defined as μ and averaged sample error variance by $\hat{\sigma}^2$. Two values for the interval of confidence L are used for each signal model in order to illustrate the dependence of the error bias on L .

Consider first a noiseless polynomial signal described by

$$x(k) = 10.0 - 0.4k + 0.005k^2 \quad (5.3)$$

This signal has the form $\underline{a}^T \underline{p}(k)$ with

$$\underline{p}(k) = \begin{bmatrix} 1 \\ k \\ k^2 \end{bmatrix} \quad (5.4)$$

This signal is identical to that used to illustrate Kalman estimator error divergence in Chapter II, except for a difference in sampling rate. The signal is a parabola with value ten at $k = 0$ and $k = 80$ and with a minimum value of two at $k = 40$. Assume the measurement $y(k)$ is given by

$$y(k) = x(k) + v(k) \quad (5.5)$$

with $v(k)$ white Gaussian noise having a constant variance of one.

Figure 10 shows the result of applying the augmented memory estimator to this signal with an inaccurate signal dynamical model. The estimator assumes the signal has the form $x(k) = \underline{a}^T \underline{p}_m(k)$ with $\underline{p}_m(k)$ given by

$$\underline{p}_m(k) = \begin{bmatrix} 1 \\ k \end{bmatrix} \quad (5.6)$$

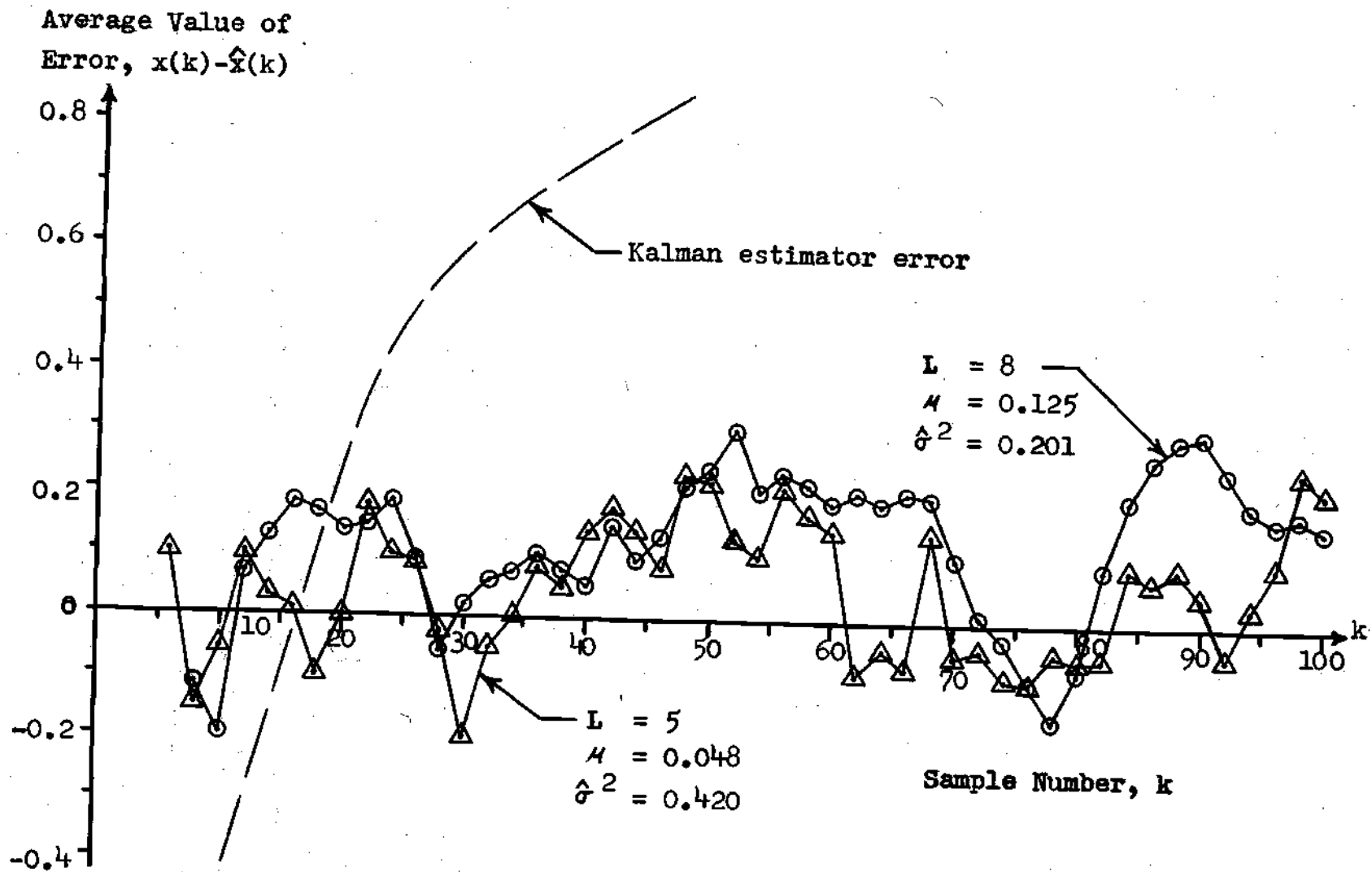


Figure 10. Average Error of Augmented Memory Estimator, Second Degree Polynomial Signal with First Degree Polynomial Model

The curves of Figure 10 represent the errors averaged over 30 runs as previously discussed, with curves for $L = 5$ and $L = 8$. For clarity, only the errors for even values of k are shown. Random effects are not completely eliminated by the averaging process, but a bias is visible over certain ranges of time, such as over the range $40 \leq k \leq 60$. The magnitude of the average error is confined to less than one-third the value of the measurement noise variance for both values of L . The average error for $L = 8$ has a larger range of values than for $L = 5$ because the inaccurate model results in larger biases as L increases. The error which results when the Kalman estimator is applied to the same signal with the same inaccurate model, as shown in Figure 2, is also shown in Figure 10. The Kalman estimator error diverges in this case.

As a second illustration of error behavior in the augmented memory estimator, consider the problem of estimating the altitude of an object in an elliptical earth orbit. The position vector $\underline{x}(t)$ of such an object is deterministic, satisfying the nonlinear differential equation

$$\ddot{\underline{x}}(t) + \left[\frac{\beta}{x^3(t)} \right] \underline{x}(t) = \underline{0} \quad (5.7)$$

where $\underline{x}(t)$ can be expressed as

$$\underline{x}(t) = x(t)e^{j\theta(t)} \quad (5.8)$$

with $x(t)$ and $\theta(t)$ the magnitude and angle of $\underline{x}(t)$ in polar coordinates. Substituting Equation (5.8) into Equation (5.7) gives the following simultaneous nonlinear differential equations:

$$\ddot{x}(t) - x(t)[\dot{\theta}(t)]^2 + \frac{\beta}{x^2(t)} = 0 \quad (5.9)$$

$$x(t)\ddot{\theta}(t) + 2\dot{x}(t)\dot{\theta}(t) = 0 \quad (5.10)$$

In the digital computer simulation, the units of $x(t)$ are earth radii, abbreviated er, and $\theta(t)$ is in radians. The gravity constant β used in the simulation is $\beta = 19.9094165$. Measurements are assumed to consist of discrete samples of $x(t)$ taken once every 15 minutes, with additive Gaussian measurement noise having a constant variance of 0.01 er^2 .

Then if k represents the sample number, with $k = 0$ for $t = 0$, measurements are described by

$$y(k) = x(k) + v(k) \quad (5.11)$$

For simulation purposes $x(t)$, $\dot{x}(t)$, and $\dot{\theta}(t)$ are initialized by $x(0) = 2.0 \text{ er}$, $\dot{x}(0) = 0.0 \text{ er/second}$, and $\dot{\theta}(0) = 1.8 \text{ radians/second}$. The resulting altitude $x(t)$ is shown in Figure 11. In Figure 11, time in minutes is obtained by multiplying the value of k by 15.

In order to use the augmented memory estimator to estimate orbit altitude, a linear model must be assumed for $x(k)$. One possibility is to assume $x(k)$ has the form of a noiseless second-degree polynomial,

Orbit
Altitude = $x(k)$
(earth radii)

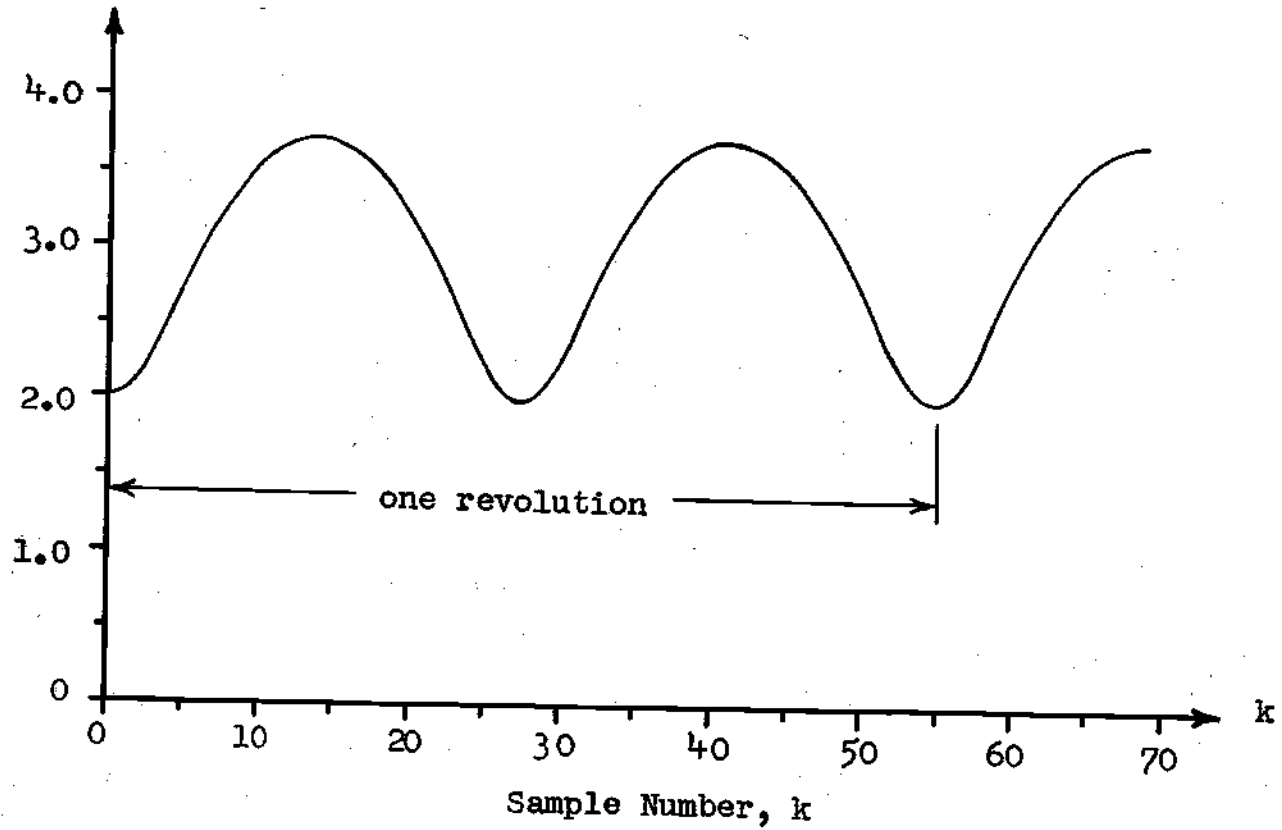


Figure 11. Orbit Altitude Signal

so that $x(k) = \underline{a}^T \underline{p}_m(k)$ with $\underline{p}_m(k)$ given by

$$\underline{p}_m(k) = \begin{bmatrix} 1 \\ k \\ k^2 \end{bmatrix} \quad (5.12)$$

The interval of confidence L is then chosen so that the assumed model of $x(k)$ is a reasonably accurate approximation of the true signal over time periods of length L sample intervals. Figure 12 shows the results of computer simulations using the model of Equation (5.12) with $L = 6$ and $L = 10$. As in the previous example, the average error is the result of averaging the error for each sample over 30 computer runs and average error is shown for even samples only. Again, larger biases result when the larger value of L is used because of the inaccuracy of the model. Computer simulation has also shown that the Kalman estimator which uses the second-degree polynomial model for $x(t)$ has an error larger than 1.5 earth radii for all $k > 20$.

Figure 13 shows the average error in the augmented memory estimator when the altitude model is a noiseless first-degree polynomial of the form $x(t) = \underline{a}^T \underline{p}_m(k)$ with $\underline{p}_m(k)$ given by Equation (5.6). Simulation results are given for $L = 3$ and $L = 6$. In the $L = 6$ case the model is a worse approximation to the signal over the interval of confidence than is the second-degree polynomial approximation, as evidence by the larger error biases.

For the last example, consider the problem of estimating the altitude of an object re-entering the earth's atmosphere. This problem

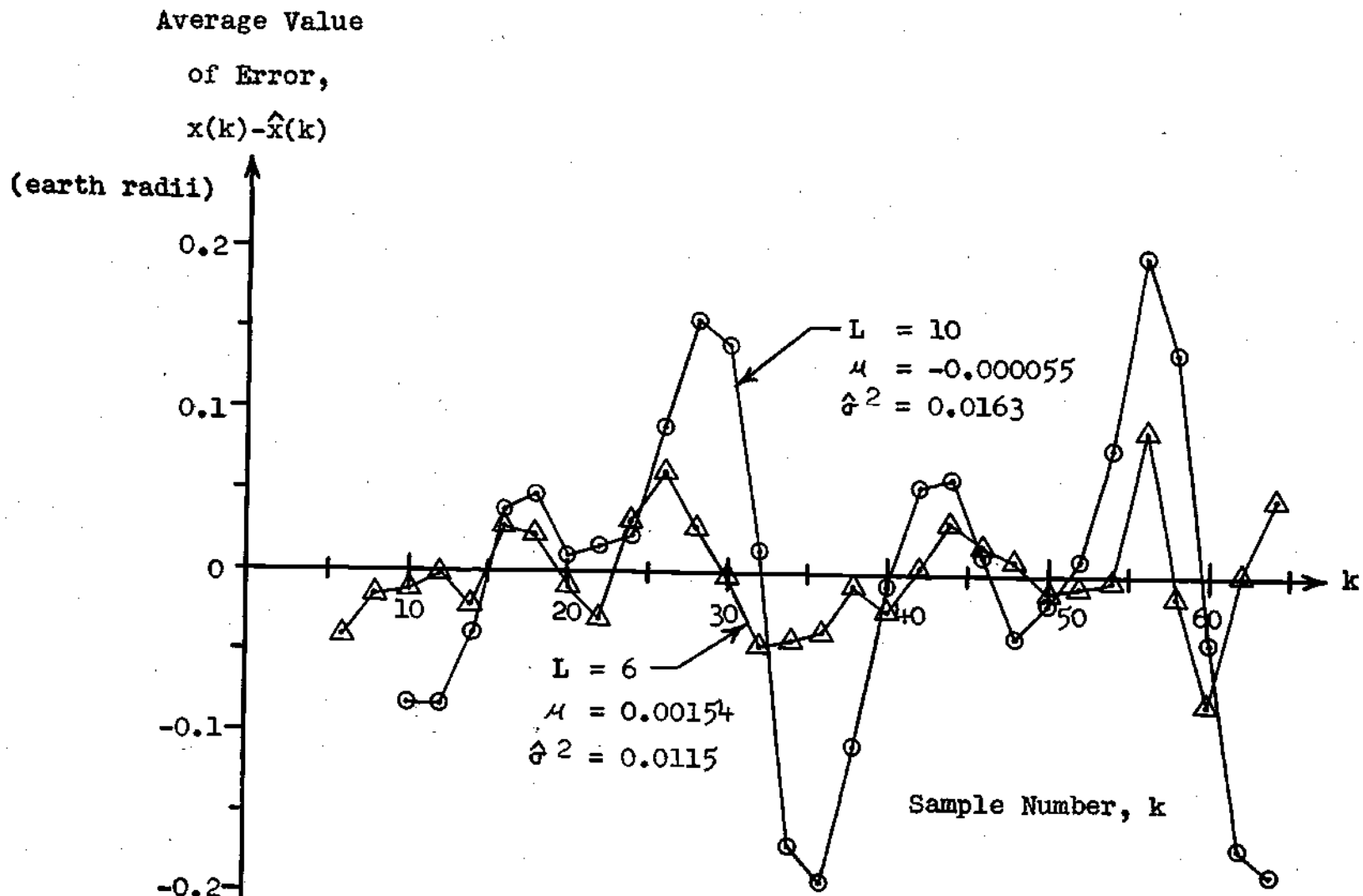


Figure 12. Average Error of Augmented Memory Estimator, Orbit Altitude Signal with Second Degree Polynomial Model

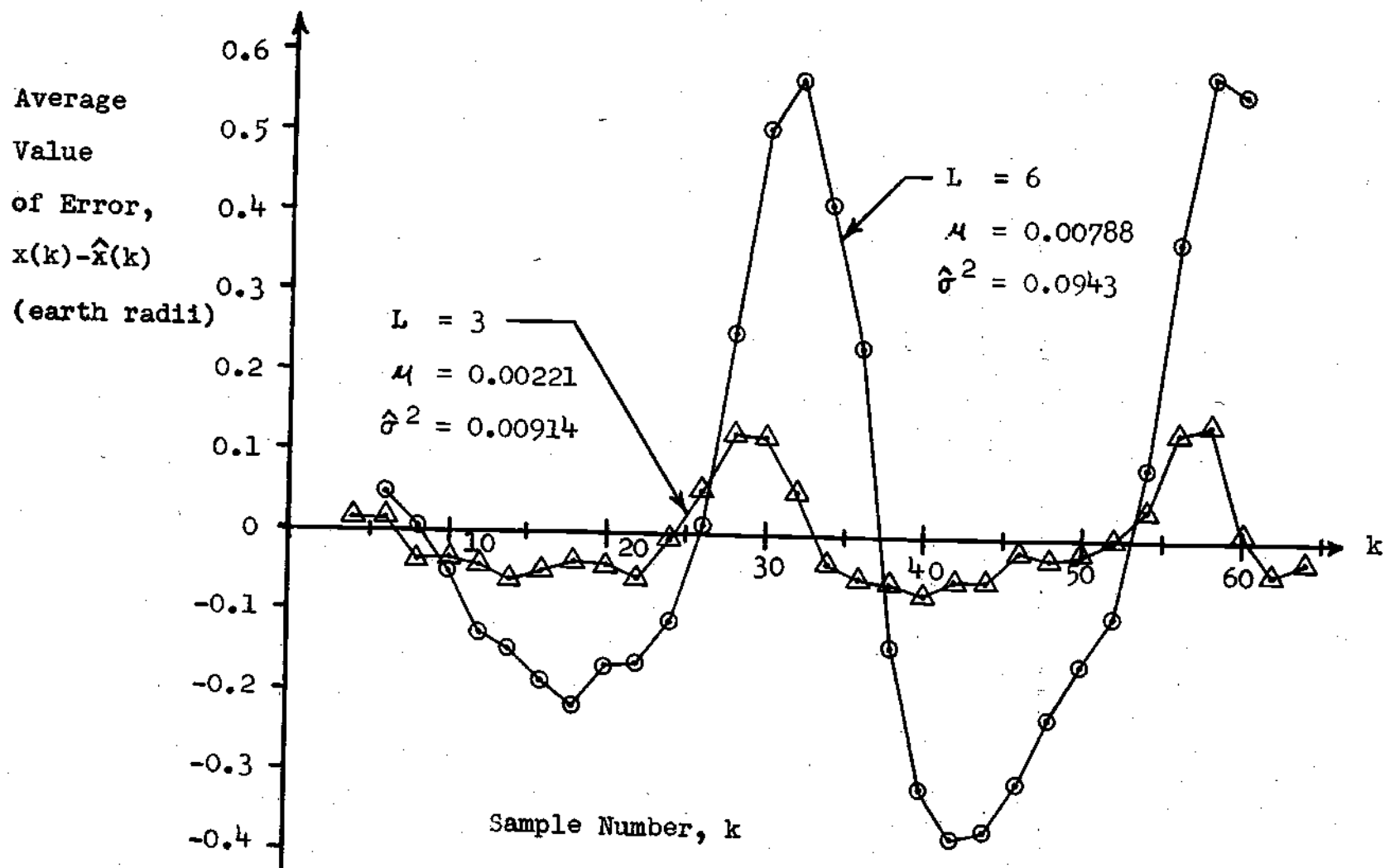


Figure 13. Average Error of Augmented Memory Estimator, Orbit Altitude Signal With First Degree Polynomial Model

was discussed in detail in Chapter II. The deterministic signal and discrete noisy measurements are described in Equations (2.21) through (2.25) and Figure 3, with the resulting signal shown in Figure 4. The linear range model of Equations (2.26, 5.12) can be used with the augmented memory estimator to compute $\hat{r}(k)$, with the altitude estimate $\hat{x}_1(k)$ then being computed by Equation (2.27). This model corresponds to the assumption that the range variable can be approximated by a noiseless second-degree polynomial over time intervals of L sample periods. Figure 14 shows the average altitude error obtained by using the second-degree polynomial range model. Results are shown for $L = 4$ and $L = 6$ with the average error for only the even values of k shown on the figure. For both values of L , a bias is evident between $k = 10$ and $k = 20$. Inspection of Figure 4 shows that these values of k are in the time interval in which the signal is most nonlinear, so that the model is less accurate. For $k > 30$ the range variable appears to have a form which the model can fit well, as verified by the lack of error bias for $k > 30$. The most significant result obtained from this example is that the relatively large biases present in the altitude estimates for $k < 20$ do not prevent the estimator error from being relatively unbiased in the $k > 30$ range.

Figure 14 also shows the result of applying the Kalman estimator to the re-entry problem with the noiseless second-degree polynomial range model. The result, also shown in Figure 5, is a diverging altitude estimation error.

Figure 15 shows the average altitude estimation error obtained when the first-degree polynomial model of Equation (5.6) is assumed for

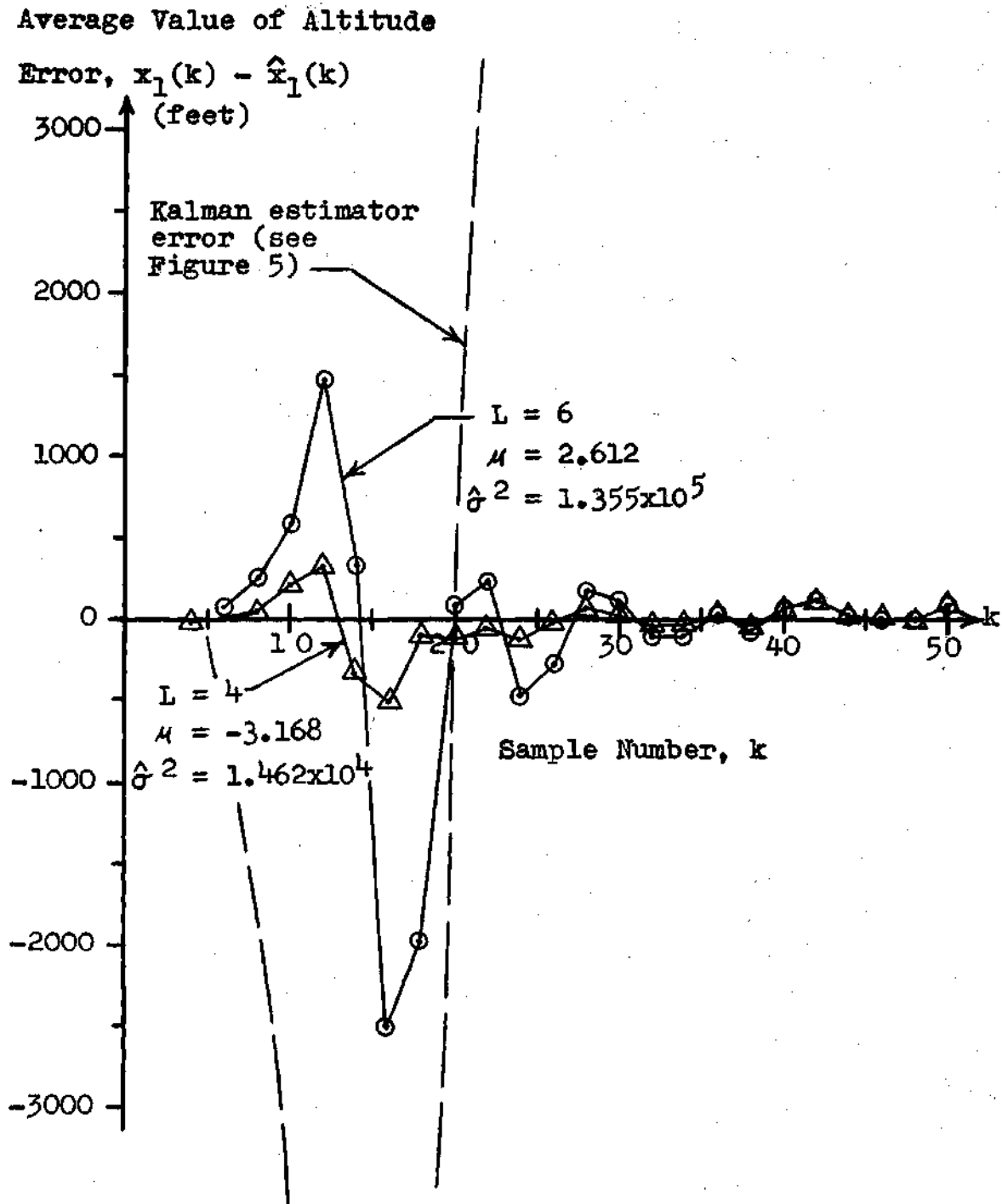


Figure 14. Average Error of the Augmented Memory Estimator, Re-entry Signal with Second Degree Polynomial Model

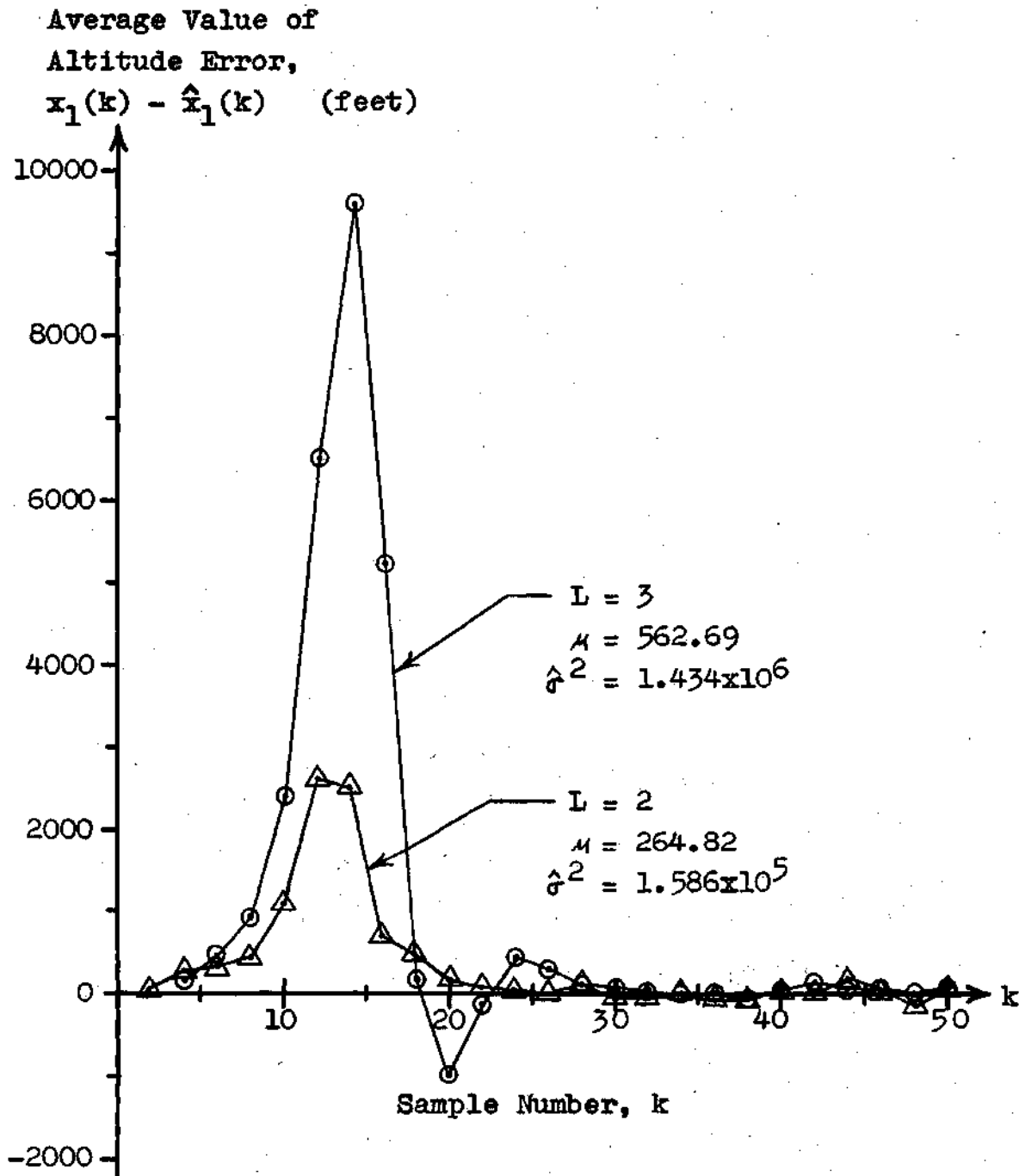


Figure 15. Average Error of the Augmented Memory Estimator, Re-entry Signal with First Degree Polynomial Model

the range variable. This model is even less accurate than the second-degree polynomial model for $k < 20$, so the resulting biases are larger than those of Figure 14. Simulation results are shown in Figure 15 for $L = 2$ and $L = 3$. Again, it is significant to note that relatively large biases for $k < 20$ do not prevent acceptable error behavior for $k > 30$, where the model is more accurate.

The examples described above show that the augmented memory estimator with a polynomial signal model demonstrates the desired properties outlined at the beginning of this chapter. Figures 7 and 8 verify that the augmented memory estimator with accurate polynomial models of noiseless polynomial signals results in a smaller mean-square error than the fixed memory estimator with the same models and memory length L . The other examples presented in this chapter illustrate that the augmented memory estimator with a polynomial signal model is less sensitive to signal modeling errors than is the Kalman estimator. They also show that the choice of the interval of confidence L can be of primary importance. Too large a value of L can result in considerable bias when the assumed model is very inaccurate over time periods of length L sample intervals. A significant result of the examples is that a large bias in estimates computed over one period of time do not necessarily cause biases in estimates in later time intervals.

CHAPTER VI

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

This dissertation contains a detailed discussion of the problem of estimating signals which contain a nonrandom component. Attention is restricted to discrete measurement systems and linear estimators. Particular emphasis in this dissertation is on the sensitivity of linear estimators to errors in the assumed model of the nonrandom component of the signal. After existing discrete estimation techniques are discussed and compared in terms of mean-square error and sensitivity to signal modeling errors, a new discrete estimator is derived. The new estimator satisfies the objective of computing estimates which have mean-square error and sensitivity properties representing an improvement over the existing methods.

In Chapter II the Kalman estimator equations are discussed. It is shown there that errors in the signal dynamical model, or state transition matrix, can result in considerable error bias or even error divergence when used in the Kalman estimator equations. Error divergence is related to the loss of significance of the gain and computed error covariance matrix, resulting in decreased dependence of estimates on the most recent measurements.

Various modifications of the Kalman estimator are also discussed in Chapter II. These modified Kalman estimators are each intended to

reduce the sensitivity of the estimates to signal dynamical modeling errors. Common among these estimators is the requirement of a known model for the error sources in the signal model, or the use of assumed bounds on the gain or error covariance. Only the limited memory method (21) allows direct use of a model which is only intended to represent an approximation to the signal over limited periods of time chosen by the user.

In Chapter III it is shown that the Kalman estimator can be interpreted as a weighted least-squares curve fitting procedure which fits all past data with the assumed model of the signal. This interpretation further explains the error sensitivity problem in the Kalman estimator. It also justifies the use of the alternative signal model of Equations (3.37, 3.38) when the measurement vector is first order. The alternative signal model leads directly to the fixed memory estimator which uses only the measurements obtained during the most recent fixed period of time. The fixed memory estimator has an advantage over the Kalman estimator in its relative insensitivity to modeling errors, but it achieves this advantage by disregarding all data taken earlier than a fixed amount of time in the past.

The augmented memory estimator equations are derived in Chapter IV and in the Appendix. The augmented memory estimator structure has a form similar to the fixed memory estimator, but with past measurements replaced by past estimates. This difference means that the augmented memory estimator depends on all past data, and in fact it is shown to be equivalent to a growing memory estimator which minimizes mean-square error subject to an increasing number of constraints. An interpretation

of the sensitivity of the error in the augmented memory estimator to errors in the signal model is also presented in Chapter IV. This interpretation proves that the bias in an estimate is expressible as a weighted sum of separate bias terms, each of which is a measure of the accuracy of the signal only over a period of time of length L sample periods.

Illustrations of the error behavior in the augmented memory estimator which uses a polynomial signal model are presented in Chapter V. It is shown that the use of an accurate polynomial signal model results in a mean-square error improvement over the comparable fixed memory estimator. The percentage of mean-square error improvement increases as the interval of confidence L increases. The results of simulation studies of particular examples are also presented in Chapter V in order to illustrate the error bias resulting from the use of inaccurate polynomial signal models. These examples indicate that the augmented memory estimator error is less sensitive to modeling errors than is the Kalman estimator. The amount of bias resulting from the use of an inaccurate signal model in the augmented memory estimator increases with the value chosen for the interval of confidence L . Therefore, L must be chosen to satisfy a trade-off between the criteria of mean-square error and sensitivity of estimator error to signal modeling errors. The examples also indicate that relatively large biases in estimates over a period of time do not necessarily prevent unbiased estimation at later times when the signal model is more accurate.

In addition to its sensitivity advantages, the augmented memory estimator has other advantages over the Kalman estimator. The presence of colored noise sequences in the signal and measurement, with a non-zero cross-correlation function, requires no changes in the estimator equations or the signal model. Also, the augmented memory estimator is easily reset when it is desirable to eliminate any effect of biases in past estimates. All that is required is to store L measurements and re-enter the estimator algorithm with new initialized storage matrices.

Recommendations for Further Research

There are several possible extensions of the augmented memory estimator presented in this dissertation. These include extensions of the class of signals to be estimated and modifications of the estimator to incorporate additional information about the signal to be estimated or to allow it to change itself adaptively. Some of the possible extensions are the following:

1. The augmented memory estimator structure, or some variation of it, might be applied to the estimation of arbitrary functions of the signal in addition to the signal itself. This extension might include the smoothing and prediction problems, in which estimates are computed for the signal or functions of the signal at time instants prior to and later than the most recent measurement.

2. Although this dissertation is concerned with signals containing a nonrandom component, the augmented memory estimator structure might be applied to the problem of estimating random signals with known power spectrum or autocorrelation function. Decreased sensitivity of

estimates to errors in the autocorrelation function would be an objective of this approach. In this case, the signal model of Equation (4.1) may be used with $\underline{a}^T \underline{p}(k) = 0$, or the random signal may be expanded in a series with some of its terms described as the $\underline{a}^T \underline{p}(k)$ term in Equation (4.1) and the remainder described by the $r(k)$ term.

3. A logical extension of the augmented memory estimator is to determine a method of using vector measurements $\underline{y}(k)$ rather than scalar measurements to estimate a signal. For example, in an orbit determination problem like the one presented in Chapter V, noisy measurements of both range and range rate may be available at each sampling instant.

4. The form of the signal model used by the augmented memory estimator permits simple mathematical descriptions of known bounds on the signal or functions of the signal. Such known bounds might be incorporated in the equations of the augmented memory estimator. This procedure should have the effect of limiting the biases which occur when an inaccurate signal model is used in the estimator equations without the consideration of known bounds.

5. The augmented memory estimator might be made adaptive by causing it to automatically adjust its own structure. Adjustments could take the form of changes in the interval of confidence or in the degree of a polynomial signal model. To determine when and how to make such adjustments, an error sensor consisting of observations of the differences between measurements and estimates might be used.

APPENDIX
DEVELOPMENT OF COMPUTATIONAL ALGORITHM FOR
ERROR CORRELATION TERMS

In Chapter IV the equations for the augmented memory estimator are derived. The resulting equation for the weighting coefficients involves two matrices, $S(k)$ and $\underline{\gamma}(k)$, which contain correlation terms involving the estimation error $n(k)$ and the noise sequences in the signal and measurement. In this appendix the derivation of a computational algorithm for recursively computing the error correlation terms needed in $S(k)$ and $\underline{\gamma}(k)$ will be presented. The algorithm will be described in a form which can be directly implemented by a digital computer program. Methods for initializing and updating required storage matrices will also be presented.

The matrix $S(k)$ is defined by Equation (4.21). Inspection of Equation (4.21) shows that $S(k)$ is symmetric, with the first row and column containing the terms

$$E \{ [r(k) + v(k)]^2 \} = \rho(k, k) + 2\phi(k, k) + \sigma(k, k) \quad (A.1)$$

$$\begin{aligned} E \{ [r(k) + v(k)] [r(k-i) + n(k-i)] \} \\ = \rho(k, k-i) + \alpha(k, k-i) + s(k, k-i) \\ + \phi(k-i, k) \text{ for } i = 1, 2, \dots, L \end{aligned} \quad (A.2)$$

where ρ , ϕ , and σ are correlation functions defined by Equations (4.4, 4.5, 4.6) and α and s are error correlation functions defined by

$$\alpha(k, j) \triangleq E [r(k)n(j)] \quad (A.3)$$

$$s(k, j) \triangleq E [v(k)n(j)] \quad (A.4)$$

with $n(j)$ defined by Equation (4.23). The second row, columns two through $(L+1)$, of $S(k)$ contains terms of the form

$$\begin{aligned} & E \{ [r(k-1) + n(k-1)] [r(k-i) + n(k-i)] \} \quad (A.5) \\ & = \rho(k-1, k-i) + \alpha(k-1, k-i) + \alpha(k-i, k-1) + \mu(k-1, k-i) \\ & \quad \text{for } i = 1, 2, \dots, L \end{aligned}$$

with $\mu(k, j)$ the error autocorrelation function defined by

$$\mu(k, j) \triangleq E [n(k)n(j)] \quad (A.6)$$

In addition to the terms of the form of Equations (A.2, A.5), the last $(L-1)$ rows of $S(k)$ contain terms having the forms

$$\begin{aligned} & E \{ [r(k-i) + n(k-i)] [r(k-j) + n(k-j)] \} \quad (A.7) \\ & = \rho(k-i, k-j) + \alpha(k-i, k-j) + \alpha(k-j, k-i) + \mu(k-i, k-j) \\ & \quad \text{for } i = 2, 3, \dots, L \text{ and } j = 2, 3, \dots, L \end{aligned}$$

Replacing k in Equations (A.5, A.7) by $k-1$ reveals that the terms in $S(k)$ having the form of Equation (A.7) can be obtained directly from terms in the $S(k-1)$ matrix. The matrix remaining after deleting the first two rows and columns of $S(k)$ is identical to the matrix remaining after deleting the first and last rows and columns of $S(k-1)$. This leaves only the first two rows and columns of $S(k)$ to be determined. All terms involving the correlation functions ρ , ϕ , and σ are assumed known, leaving only the terms involving the error $n(k-i)$ in the first two rows

and columns to be computed. These terms consist of five types of expressions, so it is convenient to define the following five forms:

$$\text{Form 1: } E[r(k)n(k-i)] = \alpha(k, k-i) \text{ for } i = 1, 2, \dots, L \quad (\text{A.8})$$

$$\text{Form 2: } E[r(k-1)n(k-i)] = \alpha(k-1, k-i) \text{ for } i = 1, 2, \dots, L \quad (\text{A.9})$$

$$\text{Form 3: } E[r(k-1)n(k-1)] = \alpha(k-1, k-1) \text{ for } i = 2, 3, \dots, L \quad (\text{A.10})$$

$$\text{Form 4: } E[v(k)n(k-i)] = s(k, k-i) \text{ for } i = 1, 2, \dots, L \quad (\text{A.11})$$

$$\text{Form 5: } E[n(k-1)n(k-i)] = \mu(k-1, k-i) \text{ for } i = 1, 2, \dots, L \quad (\text{A.12})$$

Algorithms for computing these terms will be derived below. After these terms are computed, the first two rows and columns of the $S(k)$ matrix are completely determined by using Equations (A.1, A.2, A.5).

The $\underline{\gamma}(k)$ vector defined by Equation (4.22) can be written

$$\underline{\gamma}(k) = E[r(k)\underline{m}(k)] = \begin{bmatrix} \rho(k, k) + \phi(k, k) \\ \rho(k, k-1) + \alpha(k, k-1) \\ \rho(k, k-2) + \alpha(k, k-2) \\ \vdots \\ \rho(k, k-L) + \alpha(k, k-L) \end{bmatrix} \quad (\text{A.13})$$

This result is obtained by using Equations (4.1, 4.23) to replace the estimates $\hat{x}(k-i)$ in $\underline{m}(k)$ by

$$\hat{x}(k-i) = \underline{a}^T \underline{p}(k-i) + r(k-i) + n(k-i) \quad (\text{A.14})$$

Equation (A.13) shows that $\underline{\gamma}(k)$ can be determined after the Form 1 terms of Equation (A.8) are determined.

In order to compute the error correlation terms of Equations (A.8-A.12) by a recursive algorithm it is necessary to store a fixed

amount of past information. It is convenient to summarize this information storage requirement by defining the storage matrices below:

$$A(k) \triangleq \begin{bmatrix} \alpha(k-1, k-2) & \alpha(k-2, k-2) & \dots & \alpha(k-L-1, k-2) \\ \alpha(k-1, k-3) & \alpha(k-2, k-3) & \dots & \alpha(k-L-1, k-3) \\ \vdots & \vdots & \ddots & \vdots \\ \alpha(k-1, k-L-1) & \alpha(k-2, k-L-1) & \dots & \alpha(k-L-1, k-L-1) \end{bmatrix} \quad (A.15)$$

$$B(k) \triangleq \begin{bmatrix} \mu(k-2, k-2) & 0 & 0 & 0 & \dots & 0 \\ \mu(k-2, k-3) & \mu(k-3, k-3) & 0 & 0 & \dots & 0 \\ \mu(k-2, k-4) & \mu(k-3, k-4) & \mu(k-4, k-4) & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu(k-2, k-L) & \mu(k-3, k-L) & \mu(k-4, k-L) & \dots & \mu(k-L, k-L) \\ \mu(k-2, k-L-1) & \mu(k-3, k-L-1) & \mu(k-4, k-L-1) & \dots & \mu(k-L, k-L-1) \end{bmatrix} \quad (A.16)$$

$$WM(k) \triangleq \begin{bmatrix} \underline{w}^T(k-1) \\ \underline{w}^T(k-2) \\ \vdots \\ \underline{w}^T(k-L) \end{bmatrix} \quad (A.17)$$

$$\underline{VN}(k) \triangleq \begin{bmatrix} s(k-1, k-2) \\ s(k-1, k-3) \\ \vdots \\ s(k-1, k-L) \end{bmatrix} \quad (A.18)$$

These storage matrices are considered to be available for the computation of the error correlation terms of Equations (A.8-A.12). After the algorithms for computing the error correlation terms are derived below, the procedure for initializing and updating the storage matrices will be derived.

Computation of Form 1 Terms

Define the vector $\underline{F1}(k)$ containing the Form 1 terms by

$$\underline{F1}(k) \triangleq \begin{bmatrix} \alpha(k, k-1) \\ \alpha(k, k-2) \\ \vdots \\ \alpha(k, k-L) \end{bmatrix} \quad (\text{A.19})$$

Equation (A.14) solved for $n(k-i)$ gives

$$n(k-i) = -\underline{a}^T \underline{p}(k-i) - r(k-i) + \hat{x}(k-i) \quad (\text{A.20})$$

Then by Equations (A.3, A.20, 4.7) the last term of $\underline{F1}(k)$ can be written as

$$\begin{aligned} \alpha(k, k-L) &= E \{ r(k) [-\underline{a}^T \underline{p}(k-L) - r(k-L) \\ &\quad + \underline{w}^T(k-L) \underline{m}(k-L)] \} \\ &= -\rho(k, k-L) + \underline{w}^T(k-L) E [r(k) \underline{m}(k-L)] \end{aligned} \quad (\text{A.21})$$

Substituting Equations (4.9, A.14) for $\underline{m}(k-L)$ in Equation (A.21) gives

$$\alpha(k, k-L) = -\rho(k, k-L) \quad (\text{A.22})$$

$$+ \underline{w}^T(k-L) \begin{bmatrix} \rho(k, k-L) + \phi(k, k-L) \\ \rho(k, k-L-1) + \alpha(k, k-L-1) \\ \rho(k, k-L-2) + \alpha(k, k-L-2) \\ \vdots \\ \rho(k, k-2L) + \alpha(k, k-2L) \end{bmatrix}$$

Each of the $\alpha(k, k-L-1)$ terms in the above equation could now be expanded in the same manner as $\alpha(k, k-L)$. In order to keep the number of computations constant for any value of k , it is necessary to make the

assumption that nonzero correlation time is bounded for the noise sequences in the signal and measurement. A reasonable assumption is that no correlation exists between samples from either noise process which are separated in time by more than L sample periods, where L is the interval of confidence. This assumption means that $\rho(k,j)$, $\sigma(k,j)$, and $\phi(k,j)$ are zero for all (k,j) such that $|k-j| > L$. Since $n(k-j)$ for $j > 0$ depends only on terms obtained no later than the k^{th} sample, it follows that $\alpha(k,k-j)$ and $s(k,k-j)$ are also zero for all $j > L$.

With the above assumption on the noise processes, Equation (A.22) becomes

$$\alpha(k, k-L) = -\rho(k, k-L) \quad (\text{A.23})$$

$$+ \underline{w}^T(k-L) \begin{bmatrix} \rho(k, k-L) + \phi(k, k-L) \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Using the same procedure for expanding $\alpha(k, k-L)$ to expand $\alpha(k, k-L+1)$ gives

$$\alpha(k, k-L+1) = -\rho(k, k-L+1) \quad (\text{A.24})$$

$$+ \underline{w}^T(k-L+1) \begin{bmatrix} \rho(k, k-L+1) + \phi(k, k-L+1) \\ \rho(k, k-L) + \alpha(k, k-L) \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

The $\alpha(k, k-L)$ term in the above equation is known from the immediately preceding computation of Equation (A.23). Expanding $\alpha(k, k-L+2)$ in the same manner gives

$$\alpha(k, k-L+2) = -\rho(k, k-L+2) \quad (\text{A.25})$$

$$+ \underline{w}^T(k-L+2) \begin{bmatrix} \rho(k, k-L+2) + \phi(k, k-L+2) \\ \rho(k, k-L+1) + \alpha(k, k-L+1) \\ \rho(k, k-L) + \alpha(k, k-L) \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

The $\alpha(k, k-L+1)$ and $\alpha(k, k-L)$ terms in the above equation are known from Equations (A.23, A.24). Continuing in this manner, all terms in $\underline{F1}(k)$ are computed in sequence from $\alpha(k, k-L)$ to $\alpha(k, k-1)$, which has the form:

$$\alpha(k, k-1) = -\rho(k, k-1) \quad (\text{A.26})$$

$$+ \underline{w}^T(k-1) \begin{bmatrix} \rho(k, k-1) + \phi(k, k-1) \\ \rho(k, k-2) + \alpha(k, k-2) \\ \rho(k, k-3) + \alpha(k, k-3) \\ \vdots \\ \rho(k, k-L) + \alpha(k, k-L) \\ 0 \end{bmatrix}$$

Summarizing the procedure described above, $\underline{F1}(k)$ is computed term by term by the following equations, which must be used in the order given:

$$\underline{F1}_L(k) = \alpha(k, k-L) = -\rho(k, k-L) \quad (\text{A.27})$$

$$+ w_L(k-L) [\rho(k, k-L) + \phi(k, k-L)]$$

$$\underline{F1}_i(k) = \alpha(k, k-i) = -\rho(k, k-i)$$

$$+ w_i(k-i) [\rho(k, k-i) + \phi(k, k-i)]$$

$$+ \sum_{j=2}^{L+1-i} \{w_j(k-i) [\rho(k, k+1-i-j) + \phi(k, k+1-i-j)]\}$$

$$+ \alpha(k, k+1-i-j)] \} \text{ for } i=L-1, L-2, \dots, 1$$

In terms of the storage matrix $WM(k)$ defined by Equation (A.17), Equation (A.27) becomes

$$\begin{aligned} F1_L(k) &= \alpha(k, k-L) = -\rho(k, k-L) & (A.28) \\ &+ WM_{L,1}(k)[\rho(k, k-L) + \phi(k, k-L)] \\ F1_i(k) &= \alpha(k, k-i) = -\rho(k, k-i) \\ &+ WM_{i,1}(k)[\rho(k, k-i) + \phi(k, k-i)] \\ &+ \sum_{j=2}^{L+1-i} \{ WM_{i,j}(k)[\rho(k, k+1-i-j) \\ &+ F1_{i+j-1}(k)] \} \text{ for } i=L-1, L-2, \dots, 1 \end{aligned}$$

Computation of Form 2 Terms

Define the vector $\underline{F2}(k)$ containing the Form 2 terms by

$$\underline{F2}(k) \triangleq \begin{bmatrix} \alpha(k-1, k-1) \\ \alpha(k-1, k-2) \\ \vdots \\ \alpha(k-1, k-L) \end{bmatrix} \quad (A.29)$$

All but the first term of $\underline{F2}(k)$ are available as values in the first column of the $A(k)$ storage matrix defined by Equation (A.15). The $\alpha(k-1, k-1)$ term of $\underline{F2}(k)$ can be expanded by the same procedure used for the Form 1 terms in the preceding discussion. The result is

$$F2_1(k) = \alpha(k-1, k-1) = -\rho(k-1, k-1) \quad (A.30)$$

$$+ \underline{w}^T(k-1) \begin{bmatrix} \rho(k-1, k-1) + \phi(k-1, k-1) \\ \rho(k-1, k-2) + \alpha(k-1, k-2) \\ \rho(k-1, k-3) + \alpha(k-1, k-3) \\ \vdots \\ \rho(k-1, k-L) + \alpha(k-1, k-L) \\ \rho(k-1, k-L-1) + \alpha(k-1, k-L-1) \end{bmatrix}$$

The $\alpha(k-1, k-i)$ terms of the above equation are all contained in the first column of $A(k)$. The $\underline{w}^T(k-1)$ vector is the first row of the $WM(k)$ matrix defined by Equation (A.17). The procedure for computing $F2(k)$ is summarized by the following equations:

$$\begin{aligned} F2_1(k) &= \alpha(k-1, k-1) = -\rho(k-1, k-1) & (A.31) \\ &+ WM_{1,1}(k) [\rho(k-1, k-1) + \phi(k-1, k-1)] \\ &+ \sum_{j=2}^{L+1} \{ WM_{1,j}(k) [\rho(k-1, k-1) + A_{j-1,1}(k)] \} \end{aligned}$$

$$F2_i(k) = \alpha(k-1, k-i) = A_{i-1,1}(k) \quad \text{for } i=2, 3, \dots, L$$

Computation of Form 3 Terms

Define the vector $\underline{F3}(k)$ containing the Form 3 terms by

$$\underline{F3}(k) = \begin{bmatrix} \alpha(k-2, k-1) \\ \alpha(k-3, k-1) \\ \vdots \\ \alpha(k-L, k-1) \end{bmatrix} \quad (A.32)$$

Using Equations (A.3, A.20, 4.7), the first term in $\underline{F3}(k)$ can be written

$$\alpha(k-2, k-1) = -\rho(k-2, k-1) \quad (\text{A.33})$$

$$+ \underline{w}^T(k-1) \begin{bmatrix} \rho(k-2, k-1) + \phi(k-2, k-1) \\ \rho(k-2, k-2) + \alpha(k-2, k-2) \\ \rho(k-2, k-3) + \alpha(k-2, k-3) \\ \vdots \\ \rho(k-2, k-L) + \alpha(k-2, k-L) \\ \rho(k-2, k-L-1) + \alpha(k-2, k-L-1) \end{bmatrix}$$

The $\alpha(k-2, k-1)$ terms in the above equation are identical to the terms in the second column of $A(k)$. By the same procedure used to obtain Equation (A.33), the second term in $\underline{F3}(k)$ can be written as

$$\alpha(k-3, k-1) = -\rho(k-3, k-1) \quad (\text{A.34})$$

$$+ \underline{w}^T(k-1) \begin{bmatrix} \rho(k-3, k-1) + \phi(k-3, k-1) \\ \rho(k-3, k-2) + \alpha(k-3, k-2) \\ \rho(k-3, k-3) + \alpha(k-3, k-3) \\ \vdots \\ \rho(k-3, k-L) + \alpha(k-3, k-L) \\ \rho(k-3, k-L-1) + \alpha(k-3, k-L-1) \end{bmatrix}$$

The $\alpha(k-3, k-i)$ terms in the above equation are identical to the terms in the third column of $A(k)$. Continuing the above procedure, each term in $\underline{F3}(k)$ can be written in terms of values in the first row of $WM(k)$ and columns of $A(k)$. The algorithm for computing $\underline{F3}(k)$ is summarized by the following equations:

$$\begin{aligned} F3_{i-1}(k) &= \alpha(k-i, k-1) \quad (\text{A.35}) \\ &= -\rho(k-i, k-1) + WM_{1,1}(k) [\rho(k-i, k-1)] \end{aligned}$$

$$\begin{aligned}
& + \phi(k-i, k-1)] + \sum_{j=2}^{L+1} \{WM_{1,j}(k)[\rho(k-i, k-j) \\
& + A_{j-1,i}(k)]\} \quad \text{for } i=2, 3, \dots, L
\end{aligned}$$

Computation of Form 4 Terms

Define the vector $\underline{F4}(k)$ containing the Form 4 terms by

$$\underline{F4}(k) \triangleq \begin{bmatrix} s(k, k-1) \\ s(k, k-2) \\ \vdots \\ s(k, k-L) \end{bmatrix} \quad (\text{A.36})$$

Comparison of Equations (A.19, A.36) and the definitions of Equations (A.3, A.4) leads to the conclusion that $\underline{F4}(k)$ can be computed in the same manner as $\underline{F1}(k)$. The procedure is to start with the last term in $\underline{F4}(k)$ and work up to the $s(k, k-1)$ term. Again, the assumptions regarding the correlation times described in the discussion of the Form 1 terms are necessary to insure a fixed computational procedure. The resulting algorithm is summarized by the following equations:

$$F4_L(k) = s(k, k-L) = -\phi(k-L, k) \quad (\text{A.37})$$

$$+ WM_{L,1}(k)[\phi(k-L, k) + \sigma(k, k-L)]$$

$$F4_{i+1}(k) = s(k, k-i) = -\phi(k-i, k)$$

$$+ WM_{i+1,1}(k)[\phi(k-i, k) + \sigma(k, k-i)]$$

$$+ \sum_{j=2}^{L+1-i} \{WM_{i+1,j}(k)[\phi(k+1-i-j, k) + F4_{i+j-1}(k)]\}$$

for $i = L-1, L-2, \dots, 1$

Computation of Form 5 Terms

Define the vector $\underline{F5}(k)$ containing Form 5 terms by

$$\underline{F5}(k) \triangleq \begin{bmatrix} \mu(k-1, k-1) \\ \mu(k-1, k-2) \\ \vdots \\ \mu(k-1, k-L) \end{bmatrix} \quad (\text{A.38})$$

Since the first term of $\underline{F5}(k)$ is just the mean-square error of the estimate $\hat{x}(k-1)$, it can be computed by Equation (4.24) expressed at $k-1$:

$$\begin{aligned} F5_1(k) &= \mu(k-1, k-1) & (\text{A.39}) \\ &= \underline{w}^T(k-1)S(k-1)\underline{w}(k-1) \\ &\quad - 2 \underline{\gamma}^T(k-1)\underline{w}(k-1) + \rho(k-1, k-1) \end{aligned}$$

This term is computed after $\hat{x}(k-1)$ is obtained.

The second term of $\underline{F5}(k)$ can be expanded by Equations (A.6, A.20, 4.7) to obtain

$$\begin{aligned} F5_2(k) &= \mu(k-1, k-2) & (\text{A.40}) \\ &= E\{n(k-2)[-a^T p(k-1) - r(k-1) + \underline{w}^T(k-1)\underline{m}(k-1)]\} \\ &= -\alpha(k-1, k-2) + \underline{w}^T(k-1) \begin{bmatrix} \alpha(k-1, k-2) + s(k-1, k-2) \\ \alpha(k-2, k-2) + \mu(k-2, k-2) \\ \alpha(k-3, k-2) + \mu(k-2, k-3) \\ \vdots \\ \alpha(k-L-1, k-2) + \mu(k-2, k-L-1) \end{bmatrix} \end{aligned}$$

In the above equation all of the $\alpha(k-i, k-2)$ terms are available from the first row of the $A(k)$ storage matrix, the $s(k-1, k-2)$ term is

the first term in the $\underline{VN}(k)$ storage vector, and all of the $\mu(k-2, k-1)$ terms are available from the first column of the $B(k)$ storage matrix. Expanding the third term of $\underline{F5}(k)$ in a similar manner gives

$$\begin{aligned}
 F5_3(k) &= \mu(k-1, k-3) & (A.41) \\
 &= E\{n(k-3)[-a^T p(k-1) - r(k-1) + w^T(k-1)m(k-1)]\} \\
 &= -\alpha(k-1, k-3) + w^T(k-1) \begin{bmatrix} \alpha(k-1, k-3) + s(k-1, k-3) \\ \alpha(k-2, k-3) + \mu(k-2, k-3) \\ \alpha(k-3, k-3) + \mu(k-3, k-3) \\ \alpha(k-4, k-3) + \mu(k-3, k-4) \\ \vdots \\ \alpha(k-L-1, k-3) + \mu(k-3, k-L-1) \end{bmatrix}
 \end{aligned}$$

The $\alpha(k-i, k-3)$ terms in the above equation are available in the second row of $A(k)$, and $s(k-1, k-3)$ is the second term in $\underline{VN}(k)$. The $\mu(k-i, k-3)$ terms are in the second row of $B(k)$ and the $\mu(k-3, k-i)$ terms are in the second column of $B(k)$. Continuing the same procedure of expanding terms in $\underline{F5}(k)$ results in a general expression of the form

$$\begin{aligned}
 F5_1(k) &= \mu(k-1, k-i) & (A.42) \\
 &= E\{n(k-1)[-a^T p(k-1) - r(k-1) + w^T(k-1)m(k-1)]\} \\
 &= -\alpha(k-1, k-i) + w^T(k-1) \begin{bmatrix} \alpha(k-1, k-i) + s(k-1, k-i) \\ \alpha(k-2, k-i) + \mu(k-2, k-i) \\ \alpha(k-3, k-i) + \mu(k-3, k-i) \\ \vdots \\ \alpha(k-i, k-i) + \mu(k-1, k-i) \\ \alpha(k-i-1, k-i) + \mu(k-i, k-i-1) \\ \vdots \\ \alpha(k-L-1, k-i) + \mu(k-i, k-L-1) \end{bmatrix}
 \end{aligned}$$

The $\alpha(k-j, k-i)$ terms in the above equation are available from row $i-1$ of $A(k)$, and the $s(k-1, k-i)$ term is available as the value of $VN_{i-1}(k)$. The $\mu(k-j, k-i)$ terms are available from row $i-1$ of $B(k)$ and the $\mu(k-i, k-j)$ terms are available from column $i-1$ of $B(k)$.

The above procedure for computing $F5(k)$ is implemented by the following equations:

$$\begin{aligned}
 F5_1(k) &= P(k-1) & (A.43) \\
 &= \underline{w}^T(k-1)S(k-1)\underline{w}(k-1) \\
 &\quad - 2\underline{\gamma}^T(k-1)\underline{w}(k-1) + \rho(k-1, k-1) \\
 F5_i(k) &= -A_{i-1,1}(k) + WM_{1,1}(k)[A_{i-1,1}(k) + VN_{i-1}(k)] \\
 &\quad + \sum_{j=2}^i \{WM_{1,j}(k)[A_{i-1,j}(k) + B_{i-1,j-1}(k)]\} \\
 &\quad + \sum_{q=i+1}^{L+1} \{WM_{1,q}(k)[A_{i-1,q}(k) + B_{q-1,i-1}(k)]\} \\
 &\quad \text{for } i = 2, 3, \dots, L
 \end{aligned}$$

Procedure for Updating and Initializing the
Storage Matrices

After the above procedures are used to compute the terms needed for determining $S(k)$ and $\underline{\gamma}(k)$, the storage matrices defined by Equations (A.15-A.18) must be updated to $k+1$ so that they contain terms needed to compute $S(k+1)$ and $\underline{\gamma}(k+1)$. Substituting $k+1$ for k in the definitions of the storage matrices gives

$$A(k+1) = \begin{bmatrix} \alpha(k, k-1) & | & \alpha(k-1, k-1) & | & \alpha(k-2, k-1) & \dots & \alpha(k-L, k-1) \\ \alpha(k, k-2) & | & \alpha(k-1, k-2) & | & \alpha(k-2, k-2) & \dots & \alpha(k-L, k-2) \\ \vdots & & & & & & \\ \alpha(k, k-L) & | & \alpha(k-1, k-L) & | & \alpha(k-2, k-L) & \dots & \alpha(k-L, k-L) \end{bmatrix} \quad (A.44)$$

(A.45)

$$B(k+1) = \begin{bmatrix} \mu(k-1, k-1) & | & 0 & 0 & 0 & \dots & 0 \\ \mu(k-1, k-2) & | & \mu(k-2, k-2) & 0 & 0 & \dots & 0 \\ \mu(k-1, k-3) & | & \mu(k-2, k-3) & \mu(k-3, k-3) & 0 & \dots & 0 \\ \vdots & & & & & & \\ \mu(k-1, k-L+1) & | & \mu(k-2, k-L+1) & \mu(k-3, k-L+1) & \dots & \mu(k-L+1, k-L+1) \\ \mu(k-1, k-L) & | & \mu(k-2, k-L) & \mu(k-3, k-L) & \dots & \mu(k-L+1, k-L) \end{bmatrix}$$

$$\underline{WM}(k+1) = \begin{bmatrix} \underline{w}^T(k) \\ \underline{w}^T(k-1) \\ \vdots \\ \underline{w}^T(k-L+1) \end{bmatrix} \quad (A.46)$$

$$\underline{VN}(k+1) = \begin{bmatrix} s(k, k-1) \\ s(k, k-2) \\ \vdots \\ s(k, k-L+1) \end{bmatrix} \quad (A.47)$$

The partitioning of $A(k+1)$ and $B(k+1)$ shown above is to clarify the following updating procedure. The first column of $A(k+1)$ is equal to $\underline{F1}(k)$ defined by Equation (A.19). The second column of $A(k+1)$ is equal to $\underline{F2}(k)$ defined by Equation (A.29). The first row, columns three through $L+1$ of $A(k+1)$ is equal to the transpose of $\underline{F3}(k)$ defined by Equation (A.32). Only the submatrix obtained by deleting the first row and first two columns of $A(k+1)$ remains to be updated. This submatrix is equal to the submatrix obtained by deleting the last row and the first

and last columns of the $A(k)$ matrix defined by Equation (A.15).

The first column of $B(k+1)$ is equal to $\underline{F5}(k)$ defined by Equation (A.38). The submatrix obtained by deleting the first row and first column of $B(k+1)$ is equal to the submatrix obtained by deleting the last row and the last column of the $B(k)$ matrix defined by Equation (A.16).

The updated matrix $WM(k+1)$ of Equation (A.46) is obtained by shifting each of the first $L-1$ rows of $WM(k)$ down by one row and setting the first row of $WM(k+1)$ equal to the transpose of $\underline{w}(k)$ computed by Equation (4.31).

The updated $\underline{VN}(k+1)$ of Equation (A.47) is equal to the first $L-1$ terms of $\underline{F4}(k)$ defined by Equation (A.36).

The above procedure completes the updating of the storage matrices. As mentioned in Chapter IV, the augmented memory estimator is initialized by storing the first L measurements and starting the recursive procedure at $k=L$. The initial values of $\underline{m}(L)$, $\underline{\gamma}(L)$, and $S(L)$ are given in Equations (4.32-4.34). The storage matrices of Equations (A.15-A.18) must also be initialized for $k=L$. Since $\hat{x}(k) = y(k) = x(k) + v(k)$ for $k < L$, the error $n(k)$ for $k < L$ becomes $n(k) = v(k)$. Then the correlation functions involving the error become

$$\alpha(k, k-i) = E [r(k)v(k-i)] = \phi(k, k-i) \quad (A.48)$$

$$s(k, k-i) = E [v(k)v(k-i)] = \sigma(k, k-i)$$

$$\mu(k, k-i) = E [v(k)v(k-1)] = \sigma(k, k-1)$$

for $k < L$ and $i > 0$. Then for $k=L$ the initialized storage matrices are given by

$$A(L) = \begin{bmatrix} \phi(L-1, L-2) & \phi(L-2, L-2) & \dots & \phi(-1, L-2) \\ \phi(L-1, L-3) & \phi(L-2, L-3) & \dots & \phi(-1, L-3) \\ \vdots & \vdots & \ddots & \vdots \\ \phi(L-1, -1) & \phi(L-2, -1) & \dots & \phi(-1, -1) \end{bmatrix} \quad (A.49)$$

$$B(L) = \begin{bmatrix} \sigma(L-2, L-2) & 0 & 0 & 0 & \dots & 0 \\ \sigma(L-2, L-3) & \sigma(L-3, L-3) & 0 & 0 & \dots & 0 \\ \sigma(L-2, L-4) & \sigma(L-3, L-4) & \sigma(L-4, L-4) & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma(L-2, 0) & \sigma(L-3, 0) & \sigma(L-4, 0) & \dots & \sigma(0, 0) \\ \sigma(L-2, -1) & \sigma(L-3, -1) & \sigma(L-4, -1) & \dots & \sigma(0, -1) \end{bmatrix} \quad (A.50)$$

$$WM(L) = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \dots & 0 \end{bmatrix} \quad (A.51)$$

$$\underline{VN}(L) = \begin{bmatrix} \sigma(L-1, L-2) \\ \sigma(L-1, L-3) \\ \vdots \\ \sigma(L-1, 0) \end{bmatrix} \quad (A.52)$$

BIBLIOGRAPHY

1. Norbert Wiener, Extrapolation, Interpolation, and Smoothing of Stationary Time Series, John Wiley and Sons, New York, 1949.
2. Herbert Freeman, Discrete-Time Systems, An Introduction To The Theory, John Wiley and Sons, New York, 1965.
3. R. E. Kalman, "A New Approach to Linear Filtering and Control Problems," American Society of Mechanical Engineers, Transactions, Series D, Journal of Basic Engineering, Vol. 82, March 1960.
4. R. E. Kalman, R. S. Bucy, "New Results in Linear Filtering and Prediction Theory," ASME Transactions, Series D, Journal of Basic Engineering, Vol. 83, March 1961.
5. L. A. Zadeh, J. R. Ragazzini, "An Extension of Wiener's Theory of Prediction," Journal of Applied Physics, Vol. 21, July 1950.
6. Marvin Blum, "Generalization of the Class of Nonrandom Inputs of the Zadeh-Ragazzini Prediction Model," Institute of Radio Engineers, Transactions on Information Theory, Vol. IT-2, No. 2, June 1956.
7. A. B. Lees, "Interpolation and Extrapolation of Sampled Data," IRE Transactions on Information Theory, Vol. IT-2, No. 1, March 1956.
8. K. R. Johnson, "Optimum, Linear, Discrete Filtering of Signals Containing a Nonrandom Component," IRE Transactions on Information Theory, Vol. IT-2, No. 2, June 1956.
9. Marvin Blum, "An Extension of the Minimum Mean Square Prediction Theory for Sampled Input Signals," IRE Transactions on Information Theory, Vol. IT-2, No. 3, September 1956.
10. H. C. Hsieh, "On the Optimum Synthesis of Discrete Multipole Filters with Random and Nonrandom Inputs," IRE Transactions on Automatic Control, Vol. AC-5, No. 3, August 1960.
11. D. E. Johanson, "Solution of a Linear Mean Square Estimation Problem When Process Statistics are Undefined," Institute of Electrical and Electronics Engineers, Transactions on Automatic Control, Vol. AC-11 No. 1, January 1966.
12. Samuel Zehl, "The Linear Prediction of Deterministic Signals," IEEE Transactions on Information Theory, Vol. IT-10, No. 3, July 1964.

13. A. E. Bryson, Jr., D. E. Johanson, "Linear Filtering for Time-Varying Systems Using Measurements Containing Colored Noise," IEEE Transactions on Automatic Control, Vol. AC-10, No. 1, January 1965.
14. Samuel L. Fagin, "Recursive Linear Regression Theory, Optimal Filter Theory, and Error Analysis of Optimal Systems," IEEE International Convention Record (Part 1), 1964.
15. R. E. Griffin, A. P. Sage, "Large and Small Scale Sensitivity Analysis of Optimum Estimation Algorithms," IEEE Transactions on Automatic Control, Vol. AC-13, No. 4, August 1968.
16. C. F. Price, "An Analysis of the Divergence Problem in the Kalman Filter," IEEE Transactions on Automatic Control, Vol. AC-13, No. 6, December 1968.
17. J. J. O'Donnell, "Finite-Time Filtering: One-Dimensional Case," IEEE Transactions on Automatic Control, Vol. AC-9, No. 1, January 1964.
18. H. W. Sorenson, "On the Error Behavior in Linear Minimum Variance Estimation Problems," IEEE Transactions on Automatic Control, Vol. AC-12, No. 5, October 1967.
19. F. H. Schlee, C. J. Standish, N. F. Toda, "Divergence in the Kalman Filter," American Institute of Aeronautics and Astronautics Journal, Vol. 5, No. 6, June 1967.
20. W. L. Brogan, J. L. LeMay, "Orbit Navigation with Known Landmark Tracking," Joint Automatic Control Conference, Preprints of Technical Papers, 1968, Sponsored by American Automatic Control Council.
21. Andrew H. Jazwinski, "Limited Memory Optimal Filtering," IEEE Transactions on Automatic Control, Vol. AC-13, No. 5, October 1968.
22. Michael Athans, R. P. Wishner, Anthony Bertolini, "Suboptimal State Estimation for Continuous-Time Nonlinear Systems from Discrete Noisy Measurements," IEEE Transactions on Automatic Control, Vol. AC-13, No. 5, October 1968.
23. H. W. Sorenson, "Kalman Filtering Techniques," Advances in Control Systems, Vol. 3 (1966), Ed. by C. T. Leondes, Academic Press, New York.
24. Stanley F. Schmidt, "Application of State-Space Methods to Navigation," Advances in Control Systems, Vol. 3 (1966), Ed. by C. T. Leondes, Academic Press, New York.

25. S. R. Neal, "Linear Estimation in the Presence of Errors in Assumed Plant Dynamics," IEEE Transactions on Automatic Control, Vol. AC-12, No. 5, October 1967.
26. Stanley F. Schmidt, "Estimation of State with Acceptable Accuracy Constraints," Report No. 67-4, Contract NAS 5-9085, Analytical Mechanics Associates, Inc., January 1967.
27. A. P. Sage, Jr., G. W. Masters, Jr., "Least-Squares Curve Fitting and Discrete Optimum Filtering," IEEE Transactions on Education, Vol. E-10, No. 1, March 1967.

VITA

Norman Dale Crump was born in Dallas, Texas on February 9, 1942. He is the son of Clifton G. and Jewel Crump. He was married to Paula Sue Jenkins on December 18, 1964. They now have one child, Kelly Lynn Crump, born on October 5, 1968.

Mr. Crump attended public schools in Dallas, Texas, where he graduated from high school in 1960. He entered the Georgia Institute of Technology in September 1960, and received the B.E.E. degree in December 1964. During his sophomore and junior years he was a co-op student employed by Ling-Temco-Vought Aeronautics Division in Dallas, Texas. From December 1964 to September 1965 he was employed by the same company as an associate engineer in the operational analysis section.

In September 1965 he began his graduate studies in the School of Electrical Engineering at the Georgia Institute of Technology. During the time of his graduate studies he received financial support in the form of an NDEA fellowship and a grant provided by the E. I. DuPont DeNemours and Company. From March 1968 to June 1969 he held the position of graduate teaching assistant in the School of Electrical Engineering. He received the M.S.E.E. degree in 1967 and the Ph.D. degree in 1969.