

PRENTICE-HALL SIGNAL PROCESSING SERIES

Alan V. Oppenheim, Editor

ANDREWS and HUNT *Digital Image Restoration*

BRIGHAM *The Fast Fourier Transform*

BURDIC *Underwater Acoustic System Analysis*

CASTLEMAN *Digital Image Processing*

COWAN and GRANT *Adaptive Filters*

CROCHIERE and RABINER *Multirate Digital Signal Processing*

DUDGEON and MERSEREAU *Multidimensional Digital Signal Processing*

HAMMING *Digital Filters, 2nd ed.*

HAYKIN, ED. *Array Signal Processing*

JAYANT and NOLL *Digital Coding of Waveforms*

KINO *Acoustic Waves: Devices, Imaging, and Analog Signal Processing*

LEA, ED. *Trends in Speech Recognition*

LIM, ED. *Speech Enhancement*

MARPLE *Digital Spectral Analysis with Applications*

MCCELLAN and RADER *Number Theory in Digital Signal Processing*

MENDEL *Lessons in Digital Estimation Theory*

OPPENHEIM, ED. *Applications of Digital Signal Processing*

OPPENHEIM, WILLSKY, with YOUNG *Signals and Systems*

OPPENHEIM and SCHAFER *Digital Signal Processing*

RABINER and GOLD *Theory and Applications of Digital Signal Processing*

RABINER and SCHAFER *Digital Processing of Speech Signals*

ROBINSON and TRIETEL *Geophysical Signal Analysis*

STEARNS and DAVID *Signal Processing Algorithms*

TRIBOLET *Seismic Applications of Homomorphic Signal Processing*

WIDROW and STEARNS *Adaptive Signal Processing*

**Lessons
in Digital
Estimation Theory**

Jerry M. Mendel

Department of Electrical Engineering
University of Southern California
Los Angeles, California

Prentice-Hall, Inc., Englewood Cliffs, New Jersey 07632

***Lessons
in Digital
Estimation Theory***

Library of Congress Cataloging-in-Publication Data

MENDEL, JERRY M.. (date)
Lessons in digital estimation theory.

Bibliography: p.

Includes index.

1. Estimation theory. I. Title.

QA276.8.M46 1986 511'.4 86-9365

ISBN 0-13-530809-7

*To my parents, Eleanor and Alfred Mendel
and
my wife, Letty Mendel*

Editorial/production supervision: Gretchen K. Chenenko

Cover design: Lundgren Graphics

Manufacturing buyer: Gordon Osbourne

© 1987 by Prentice-Hall, Inc.
A Division of Simon & Schuster
Englewood Cliffs, New Jersey 07632

All rights reserved. No part of this book may be reproduced, in any form or by any means, without permission in writing from the publisher.

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

ISBN 0-13-530809-7 025

PRENTICE-HALL INTERNATIONAL (UK) LIMITED, London
PRENTICE-HALL OF AUSTRALIA PTY. LIMITED, Sydney
PRENTICE-HALL CANADA INC., Toronto
PRENTICE-HALL HISPANOAMERICANA, S.A., Mexico
PRENTICE-HALL OF INDIA PRIVATE LIMITED, New Delhi
PRENTICE-HALL OF JAPAN, INC., Tokyo
PRENTICE-HALL OF SOUTHEAST ASIA PTE. LTD., Singapore
EDITORIAL PRENTICE-HALL DO BRASIL, LTDA., Rio de Janeiro

Contents

PREFACE	xiii	
LESSON 1	INTRODUCTION, COVERAGE, AND PHILOSOPHY	1
	Introduction	1
	Coverage	2
	Philosophy	5
LESSON 2	THE LINEAR MODEL	7
	Introduction	7
	Examples	7
	Notational Preliminaries	14
	Problems	15
LESSON 3	LEAST-SQUARES ESTIMATION: BATCH PROCESSING	17
	Introduction	17
	Number of Measurements	18
	Objective Function and Problem Statement	18
	Derivation of Estimator	19
	Fixed and Expanding Memory Estimators	23
	Scale Changes	23
	Problems	25

LESSON 4	LEAST-SQUARES ESTIMATION: RECURSIVE PROCESSING	26
Introduction 26		
Recursive Least-Squares: Information Form 27		
Matrix Inversion Lemma 30		
Recursive Least-Squares: Covariance Form 30		
Which Form to Use 31		
Problems 32		
LESSON 5	LEAST-SQUARES ESTIMATION: RECURSIVE PROCESSING (continued)	35
Generalization to Vector Measurements 36		
Cross-Sectional Processing 37		
Multistage Least-Squares Estimators 38		
Problems 42		
LESSON 6	SMALL SAMPLE PROPERTIES OF ESTIMATORS	43
Introduction 43		
Unbiasedness 44		
Efficiency 46		
Problems 52		
LESSON 7	LARGE SAMPLE PROPERTIES OF ESTIMATORS	54
Introduction 54		
Asymptotic Distributions 54		
Asymptotic Unbiasedness 57		
Consistency 57		
Asymptotic Efficiency 60		
Problems 61		
LESSON 8	PROPERTIES OF LEAST-SQUARES ESTIMATORS	63
Introduction 63		
Small Sample Properties of Least-Squares Estimators 63		
Large Sample Properties of Least-Squares Estimators 68		
Problems 70		

LESSON 9	BEST LINEAR UNBIASED ESTIMATION	71
Introduction 71		
Problem Statement and Objective Function 72		
Derivation of Estimator 73		
Comparison of $\hat{\theta}_{BLU}(k)$ and $\hat{\theta}_{WLS}(k)$ 74		
Some Properties of $\hat{\theta}_{BLU}(k)$ 75		
Recursive BLUEs 78		
Problems 79		
LESSON 10	LIKELIHOOD	81
Introduction 81		
Likelihood Defined 81		
Likelihood Ratio 84		
Results Described by Continuous Distributions 85		
Multiple Hypotheses 85		
Problems 87		
LESSON 11	MAXIMUM-LIKELIHOOD ESTIMATION	88
Likelihood 88		
Maximum-Likelihood Method and Estimates 89		
Properties of Maximum-Likelihood Estimates 91		
The Linear Model ($\mathcal{H}(k)$ deterministic) 92		
A Log-Likelihood Function for an Important Dynamical System 94		
Problems 97		
LESSON 12	ELEMENTS OF MULTIVARIATE GAUSSIAN RANDOM VARIABLES	100
Introduction 100		
Univariate Gaussian Density Function 100		
Multivariate Gaussian Density Function 101		
Jointly Gaussian Random Vectors 101		
The Conditional Density Function 102		
Properties of Multivariate Gaussian Random Variables 104		
Properties of Conditional Mean 104		
Problems 106		

LESSON 13	ESTIMATION OF RANDOM PARAMETERS: GENERAL RESULTS	108
------------------	---	------------

- Introduction 108
- Mean-Squared Estimation 109
- Maximum a Posteriori Estimation 114
- Problems 116

LESSON 14	ESTIMATION OF RANDOM PARAMETERS: THE LINEAR AND GAUSSIAN MODEL	118
------------------	---	------------

- Introduction 118
- Mean-Squared Estimator 118
- Best Linear Unbiased Estimation,
Revisited 121
- Maximum a Posteriori Estimator 123
- Problems 126

LESSON 15	ELEMENTS OF DISCRETE-TIME GAUSS-MARKOV RANDOM PROCESSES	128
------------------	--	------------

- Introduction 128
- Definitions and Properties of Discrete-Time
Gauss-Markov Random Processes 128
- A Basic State-Variable Model 131
- Properties of the Basic State-Variable
Model 133
- Signal-to-Noise Ratio 137
- Problems 138

LESSON 16	STATE ESTIMATION: PREDICTION	140
------------------	-------------------------------------	------------

- Introduction 140
- Single-Stage Predictor 140
- A General State Predictor 142
- The Innovations Process 146
- Problems 147

LESSON 17	STATE ESTIMATION: FILTERING (THE KALMAN FILTER)	149
------------------	--	------------

- Introduction 149
- A Preliminary Result 150
- The Kalman Filter 151
- Observations About the Kalman Filter 153
- Problems 158

LESSON 18	STATE ESTIMATION: FILTERING EXAMPLES	160
------------------	---	------------

- Introduction 160
- Examples 160
- Problems 169

LESSON 19	STATE ESTIMATION: STEADY-STATE KALMAN FILTER AND ITS RELATIONSHIP TO A DIGITAL WIENER FILTER	170
------------------	---	------------

- Introduction 170
- Steady-State Kalman Filter 170
- Single-Channel Steady-State Kalman Filter 173
- Relationships Between the Steady-State
Kalman Filter and a Finite Impulse
Response Digital Wiener Filter 176
- Comparisons of Kalman and Wiener Filters 181
- Problems 182

LESSON 20	STATE ESTIMATION: SMOOTHING	183
------------------	------------------------------------	------------

- Three Types of Smoothers 183
- Approaches for Deriving Smoothers 184
- A Summary of Important Formulas 184
- Single-Stage Smoother 184
- Double-Stage Smoother 187
- Single- and Double-Stage Smoothers
as General Smoothers 189
- Problems 192

LESSON 21	STATE-ESTIMATION: SMOOTHING (GENERAL RESULTS)	193
Introduction 193		
Fixed-Interval Smoothers 193		
Fixed-Point Smoothing 199		
Fixed-Lag Smoothing 201		
Problems 202		
LESSON 22	STATE ESTIMATION: SMOOTHING APPLICATIONS	204
Introduction 204		
Minimum-Variance Deconvolution (MVD) 204		
Steady-State MVD Filter 207		
Relationship Between Steady-State MVD Filter and an Infinite Impulse Response Digital Wiener Deconvolution Filter 213		
Maximum-Likelihood Deconvolution 215		
Recursive Waveshaping 216		
Problems 222		
LESSON 23	STATE ESTIMATION FOR THE NOT-SO-BASIC STATE-VARIABLE MODEL	223
Introduction 223		
Biases 224		
Correlated Noises 225		
Colored Noises 227		
Perfect Measurements: Reduced-Order Estimators 230		
Final Remark 233		
Problems 233		
LESSON 24	LINEARIZATION AND DISCRETIZATION OF NONLINEAR SYSTEMS	236
Introduction 236		
A Dynamical Model 237		
Linear Perturbation Equations 239		

Discretization of a Linear Time-Varying State-Variable Model 242		
Discretized Perturbation State-Variable Model 245		
Problems 246		
LESSON 25	ITERATED LEAST SQUARES AND EXTENDED KALMAN FILTERING	248
Introduction 248		
Iterated Least Squares 248		
Extended Kalman Filter 249		
Application to Parameter Estimation 255		
Problems 256		
LESSON 26	MAXIMUM-LIKELIHOOD STATE AND PARAMETER ESTIMATION	258
Introduction 258		
A Log-Likelihood Function for the Basic State-Variable Model 259		
On Computing $\hat{\theta}_{ML}$ 261		
A Steady-State Approximation 264		
Problems 269		
LESSON 27	KALMAN-BUCY FILTERING	270
Introduction 270		
System Description 271		
Notation and Problem Statement 271		
The Kalman-Bucy Filter 272		
Derivation of KBF Using a Formal Limiting Procedure 273		
Derivation of KBF When Structure of the Filter Is Prespecified 275		
Steady-State KBF 278		
An Important Application for the KBF 280		
Problems 281		

LESSON A	SUFFICIENT STATISTICS AND STATISTICAL ESTIMATION OF PARAMETERS	282
Introduction	282	
Concept of Sufficient Statistics	282	
Exponential Families of Distributions	284	
Exponential Families and Maximum-Likelihood Estimation	287	
Sufficient Statistics and Uniformly Minimum-Variance Unbiased Estimation	290	
Problems	294	
APPENDIX A	GLOSSARY OF MAJOR RESULTS	295
REFERENCES		300
INDEX		305

Preface

Estimation theory is widely used in many branches of science and engineering. No doubt, one could trace its origin back to ancient times, but Karl Friederich Gauss is generally acknowledged to be the progenitor of what we now refer to as estimation theory. R. A. Fisher, Norbert Wiener, Rudolph E. Kalman, and scores of others have expanded upon Gauss' legacy, and have given us a rich collection of estimation methods and algorithms from which to choose. This book describes many of the important estimation methods and shows how they are interrelated.

Estimation theory is a product of need and technology. Gauss, for example, needed to predict the motions of planets and comets from telescopic measurements. This "need" led to the method of least squares. Digital computer technology has revolutionized our lives. It created the "need" for recursive estimation algorithms, one of the most important ones being the Kalman filter. Because of the importance of digital technology, this book presents estimation theory from a digital viewpoint. In fact, it is this author's viewpoint that estimation theory is a natural adjunct to classical digital signal processing. It produces time-varying digital filter designs that operate on random data in an optimal manner.

This book has been written as a collection of lessons. It is meant to be an *introduction* to the general field of estimation theory, and, as such, is not encyclopedic in content or in references. It can be used for self-study or in a one-semester course. At the University of Southern California, we have covered all of its contents in such a course, at the rate of two lessons a week. We have been doing this since 1978.

Approximately one half of the book is devoted to parameter estimation and the other half to state estimation. For many years there has been a tendency to treat state estimation as a stand-alone subject and even to treat parameter estimation as a special case of state estimation. Historically, this is incorrect. In the musical "Fiddler on the Roof," Tevye argues on behalf of tradition . . . "Tradition! . . ." Estimation theory also has its tradition, and it begins with Gauss and parameter estimation. In Lesson 2 we show that state estimation is a special case of parameter estimation, i.e., it is the problem of estimating random parameters, when these parameters change from one time to the next. Consequently, the subject of state estimation flows quite naturally from the subject of parameter estimation.

Most of the book's important results are summarized in theorems and corollaries. In order to guide the reader to these results, they have been summarized for easy reference in Appendix A.

Problems are included for most lessons, because this book is meant to be used as a textbook. The problems fall into two groups. The first group contains problems that ask the reader to fill in details, which have been "left to the reader as an exercise." The second group contains problems that are related to the material in the lesson. They range from theoretical to computational problems.

This book is an outgrowth of a one-semester course on estimation theory, taught at the University of Southern California. Since 1978 it has been taught by four different people, who have encouraged me to convert the lecture notes into a book. I wish to thank Mostafa Shiva, Alan Laub, George Papavassiliopoulos, and Rama Chellappa for their encouragement. Special thanks goes to Rama Chellappa, who provided supplementary Lesson A on the subject of sufficient statistics and statistical estimation of parameters. This lesson fits in very nicely just after Lesson 14.

While writing this text, the author had the benefit of comments and suggestions from many of his colleagues and students. I especially wish to acknowledge the help of Guan-Zhong Dai, Chong-Yung Chi, Phil Burns, Youngby Kim, Chung-Chin Lu, and Tom Hebert. Special thanks goes to Georgios Giannakis. The book would not be in its present form without their contributions.

Additionally, the author wishes to thank Marcel Dekker, Inc. for permitting him to include material from Mendel, J. M., 1973, *Discrete Techniques of Parameter Estimation : the Equation-Error Formulation*, in Lessons 1-9, 11, 18, and 24; and Academic Press, Inc. for permitting him to include material from Mendel, J. M., *Optimal Seismic Deconvolution: an Estimation-Based Approach*, copyright © 1983 by Academic Press, Inc., in Lessons 11-17, 19-22, and 26.

JERRY M. MENDEL
Los Angeles, California

Lesson 1

Introduction, Coverage, and Philosophy

INTRODUCTION

This book is all about estimation theory. It is useful, therefore, for us to understand the role of estimation in relation to the more global problem of modeling. Figure 1-1 decomposes modeling into four problems: representation, measurement, estimation, and validation. As Mendel (1973, pp. 2-4) states, "The *representation problem* deals with how something should be modeled. We shall be interested only in mathematical models. Within this class of models we need to know whether the model should be static or dynamic, linear or nonlinear, deterministic or random, continuous or discretized, fixed or varying, lumped or distributed . . . , in the time-domain or in the frequency-domain . . . , etc.

"In order to verify a model, physical quantities must be measured. We distinguish between two types of physical quantities, *signals*, and *parameters*. Parameters express a relation between signals. . . .

"Not all signals and parameters are measurable. The *measurement problem* deals with which physical quantities should be measured and how they should be measured.

"The *estimation problem* deals with the determination of those physical quantities that cannot be measured from those that can be measured. We shall distinguish between the estimation of signals (i.e., states) and the estimation of parameters. Because a subjective decision must sometimes be made to classify a physical quantity as a signal or a parameter, there is some overlap between signal estimation and parameter estimation. . . .

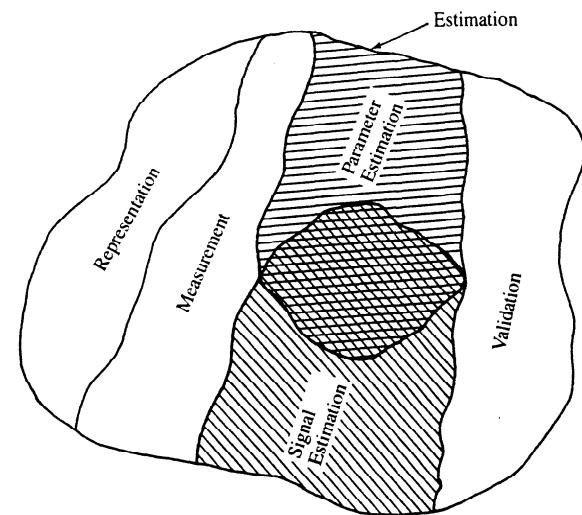


Figure 1-1 Modeling Problem (reprinted from Mendel, 1973, p. 4, by courtesy of Marcel Dekker, Inc).

"After a model has been completely specified, through choice of an appropriate mathematical representation, measurement of measurable signals, estimation of nonmeasurable signals, and estimation of its parameters, the model must be checked out. The *validation problem* deals with demonstrating confidence in the model. Often, statistical tests involving confidence limits are used to validate a model."

In this book we shall be interested in parameter estimation, state estimation, and combined state and parameter estimation. In Lesson 2 we provide six examples each of which can be categorized either as a parameter or state estimation problem. Here we just mention that: the problem of identifying the sampled values of a linear and time-invariant system's impulse response from input/output data is one of parameter estimation; the problem of reconstructing a state vector associated with a dynamical system, from noisy measurements, is one of state estimation (state estimates might be needed to implement a linear-quadratic-Gaussian optimal control law, or to perform postflight data analysis, or signal processing such as deconvolution).

COVERAGE

This book focuses on a wide range of estimation techniques that can be applied either to linear or nonlinear models. Both parameter and state estimation techniques are treated. Some parameter estimation techniques are for

deterministic parameters, whereas others are for random parameters, however, state estimation techniques are for random states. Table 1-1 summarizes the book's coverage.

Four lessons (Lessons 3, 4, 5, and 8) are devoted to least-squares estimation because it is a very basic and important technique, and, under certain often-occurring conditions, other techniques reduce to it. Consequently, once we understand the nuances of least squares and have established that a different technique has reduced to least squares, we do not have to restudy the nuances of that technique.

In order to fully study least-squares estimators we must establish their small and large sample properties. What we mean by such properties is the subject of Lessons 6 and 7.

Having spent four lessons on least-squares estimation, we cover best linear unbiased estimation (BLUE) in one lesson, Lesson 9. We are able to do this because BLUE is a special case of least squares.

In order to set the stage for maximum-likelihood estimation, which is covered in Lesson 11, we describe the concept of likelihood and its relationship to probability in Lesson 10.

Lesson 12 provides a transition from our study of estimation of deterministic parameters to our study of estimation of random parameters. It provides much useful information about elements of Gaussian random vari-

TABLE 1-1 Estimation Techniques

I. LINEAR MODELS
A. Parameter Estimation
1. Deterministic parameters
a. Least-squares (batch and recursive processing)
b. Best linear unbiased estimation (BLUE)
c. Maximum-likelihood
2. Random Parameters
a. Mean-squared
b. Maximum a posteriori
c. BLUE
d. Weighted least squares
B. State Estimation
1. Mean-squared prediction
2. Mean-squared filtering (Kalman filter/Kalman-Bucy filter)
3. Mean-squared smoothing
II. NONLINEAR MODELS
A. Parameter Estimation
Iterated least squares for deterministic parameters
B. State Estimation
Extended Kalman filter
C. Combined State and Parameter Estimation
1. Extended Kalman filter
2. Maximum-likelihood

ables. To some readers, this lesson may be a review of material already known to them.

General results for both mean-squared and maximum a posteriori estimation of random parameters are covered in Lesson 13. These results are specialized to the important case of the linear and Gaussian model in Lesson 14. Best linear unbiased and weighted least-squares estimation are also revisited in Lesson 14. Lesson 14 is quite important, because it gives conditions under which mean-squared, maximum a posteriori, best-linear unbiased, and weighted least-squares estimates of random parameters are identical. Lesson A, which is a supplemental one, is on the subject of sufficient statistics and statistical estimation of parameters. It fits in very nicely after Lesson 14.

Lesson 15 provides a transition from our study of parameter estimation to our study of state estimation. It provides much useful information about elements of discrete-time Gauss-Markov random processes, and also establishes the *basic state-variable model*, and its statistical properties, for which we derive a wide variety of state estimators. To some readers, this lesson may be a review of material already known to them.

Lessons 16 through 22 cover state estimation for the Lesson 15 basic state-variable model. Prediction is treated in Lesson 16. The important innovations process is also covered in that lesson. Filtering is the subject of Lessons 17, 18, and 19. The mean-squared state filter, commonly known as the Kalman filter, is developed in Lesson 17. Five examples which illustrate some interesting numerical and theoretical aspects of Kalman filtering are presented in Lesson 18. Lesson 19 establishes a bridge between mean-squared estimation and mean-squared digital signal processing. It shows how the steady-state Kalman filter is related to a digital Wiener filter. The latter is widely used in digital signal processing. Smoothing is the subject of Lessons 20, 21 and 22. Fixed-interval, fixed-point, and fixed-lag smoothers are developed in Lessons 20 and 21. Lesson 22 presents some applications which illustrate interesting numerical and theoretical aspects of fixed-interval smoothing. These applications are taken from the field of digital signal processing and include minimum-variance deconvolution, maximum-likelihood deconvolution, and recursive waveshaping.

Lesson 23 shows how to modify results given in Lessons 16, 17, 19, 20 and 21 from the basic state-variable model to a state-variable model that includes the following effects:

1. nonzero mean noise processes and/or known bias function in the measurement equation,
2. correlated noise processes,
3. colored noise processes, and
4. perfect measurements.

Lesson 24 provides a transition from our study of estimation for linear models to estimation for nonlinear models. Because many real-world systems are continuous-time in nature and nonlinear, this lesson explains how to linearize and discretize a nonlinear differential equation model.

Lesson 25 is devoted primarily to the extended Kalman filter (EKF), which is a form of the Kalman filter that has been “extended” to nonlinear dynamical systems of the type described in Lesson 24. The EKF is related to the method of iterated least squares (ILS), the major difference between the two being that the EKF is for dynamical systems whereas ILS is not. This lesson also shows how to apply the EKF to parameter estimation, in which case states and parameters can be estimated simultaneously, and in real time.

The problem of obtaining maximum-likelihood estimates of a collection of parameters that appears in the basic state-variable model is treated in Lesson 26. The solution involves state and parameter estimation, but calculations can only be performed off-line, after data from an experiment has been collected.

The Kalman-Bucy filter, which is the continuous-time counterpart to the Kalman filter, is derived from two different viewpoints in Lesson 27. We include this lesson because the Kalman-Bucy filter is widely used in linear stochastic optimal control theory.

PHILOSOPHY

The digital viewpoint is emphasized throughout this book. Our estimation algorithms are digital in nature; many are recursive. The reasons for the digital viewpoint are:

1. much real data is collected in a digitized manner, so it is in a form ready to be processed by digital estimation algorithms, and
2. the mathematics associated with digital estimation theory are simpler than those associated with continuous estimation theory.

Regarding (2), we mention that very little knowledge about random processes is needed to derive digital estimation algorithms, because digital (i.e., discrete-time) random processes can be treated as vectors of random variables. Much more knowledge about random processes is needed to design continuous-time estimation algorithms.

Suppose our underlying model is continuous-time in nature. We are faced with two choices: develop a continuous-time estimation theory and then implement the resulting estimators on a digital computer (i.e., discretize the continuous-time estimation algorithm), or discretize the model and develop a discrete-time (i.e., digital) estimation theory that leads to estimation algo-

rithms readily implemented on a digital computer. If both approaches lead to algorithms that are implemented digitally, then we advocate the *principle of simplicity* for their development, and this leads us to adopt the second-choice. *For estimation, our modeling philosophy is, therefore, discretize the model at the front end of the problem.*

Estimation theory has a long and glorious history (e.g., see Sorenson, 1970); however, it has been greatly influenced by technology, especially the computer. Although much of estimation theory was developed in the mathematics, statistical, and control theory literatures, we shall adopt the following viewpoint towards that theory: *estimation theory is the extension of classical signal processing to the design of digital filters that process uncertain data in an optimal manner.* In fact, estimation algorithms are just filters that transform input streams of numbers into output streams of numbers.

Most of classical digital filter design (e.g., Oppenheim and Schafer, 1975; Hamming, 1983; Peled and Liu, 1976) is concerned with designs associated with deterministic signals, e.g., low-pass and bandpass filters, and, over the years specific techniques have been developed for such designs. The resulting filters are usually “fixed” in the sense that their coefficients do not change as a function of time. Estimation theory, on the other hand, leads to filter structures that are time-varying. These filters are designed (i.e., derived) using time-domain performance specifications (e.g., smallest error variance), and, as mentioned above, process random data in an optimal manner. *Our philosophy about estimation theory is that it can be viewed as a natural adjunct to digital signal processing theory.*

Example 1-1

At one time or another we have all used the sample mean to compute an “average.” Suppose we are given a collection of k measured values of quantity X , namely $x(1), x(2), \dots, x(k)$. The sample mean of these measurements, $\bar{x}(k)$, is

$$\bar{x}(k) = \frac{1}{k} \sum_{j=1}^k x(j) \quad (1-1)$$

A recursive formula for the sample mean is obtained from (1-1), as follows:

$$\begin{aligned} \bar{x}(k+1) &= \frac{1}{k+1} \sum_{j=1}^{k+1} x(j) = \frac{1}{k+1} \left[\sum_{j=1}^k x(j) + x(k+1) \right] \\ \bar{x}(k+1) &= \frac{k}{k+1} \bar{x}(k) + \frac{1}{k+1} x(k+1) \end{aligned} \quad (1-2)$$

This recursive version of the sample mean is used for $k = 0, 1, \dots$ by setting $\bar{x}(0) = 0$.

Observe that *the sample mean, as expressed in (1-2), is a time-varying recursive digital filter whose input is measurement $x(k)$.* In later lessons we show that the sample mean is also an optimal estimation algorithm; thus, although the reader may not have been aware of it, the sample mean, which he or she has been using since early school-days, is an estimation algorithm. \square

Lesson 2

The Linear Model

INTRODUCTION

In order to estimate unknown quantities (i.e., parameters or signals) from measurements and other a priori information, we must begin with model representations and express them in such a way that attention is focused on the explicit relationship between the unknown quantities and the measurements. Many familiar models are linear in the unknown quantities (denoted θ), and can be expressed as

$$\mathbf{z}(k) = \mathcal{H}(k)\boldsymbol{\theta} + \mathbf{v}(k) \quad (2-1)$$

In this model, $\mathbf{z}(k)$, which is $N \times 1$, is called the *measurement vector*; $\boldsymbol{\theta}$, which is $n \times 1$, is called the *parameter vector*; $\mathcal{H}(k)$, which is $N \times n$ is called the *observation matrix*; and, $\mathbf{v}(k)$, which is $N \times 1$, is called the *measurement noise* vector. Usually, $\mathbf{v}(k)$ is random. By convention, the argument “ k ” of $\mathbf{z}(k)$, $\mathcal{H}(k)$, and $\mathbf{v}(k)$ denotes the fact that the last measurement used to construct (2-1) is the k th. All other measurements occur “before” the k th.

Strictly speaking, (2-1) represents an “affine” transformation of parameter vector $\boldsymbol{\theta}$ rather than a linear transformation. We shall, however, adhere to traditional estimation-theory literature, by calling (2-1) a “linear model.”

EXAMPLES

Some examples that illustrate the formation of (2-1) are given in this section. What distinguishes these examples from one another are the nature of and interrelationships between $\boldsymbol{\theta}$, $\mathcal{H}(k)$ and $\mathbf{v}(k)$. The following situations can occur.

A. θ is deterministic

1. $\mathcal{H}(k)$ is deterministic.
2. $\mathcal{H}(k)$ is random.
 - a. $\mathcal{H}(k)$ and $\mathcal{V}(k)$ are statistically independent.
 - b. $\mathcal{H}(k)$ and $\mathcal{V}(k)$ are statistically dependent.

B. θ is random

1. $\mathcal{H}(k)$ is deterministic.
2. $\mathcal{H}(k)$ is random.
 - a. $\mathcal{H}(k)$ and $\mathcal{V}(k)$ are statistically independent.
 - b. $\mathcal{H}(k)$ and $\mathcal{V}(k)$ are statistically dependent.

Example 2-1 Impulse Response Identification

It is well known that the output of a single-input single-output, linear, time-invariant, discrete-time system is given by the following convolution-sum relationship

$$y(k) = \sum_{i=-\infty}^{\infty} h(i)u(k-i) \quad (2-2)$$

where $k = 1, 2, \dots, N$, $h(i)$ is the system's impulse response (IR), $u(k)$ is its input and $y(k)$ its output. If $u(k) = 0$ for $k < 0$, and the system is causal, so that $h(i) = 0$ for $i \leq 0$, and, $h(i) = 0$ for $i > n$, then

$$y(k) = \sum_{i=1}^n h(i)u(k-i) \quad (2-3)$$

Signal $y(k)$ is measured by a sensor which is corrupted by additive measurement noise, $v(k)$, i.e., we only have access to measurement $z(k)$, where

$$z(k) = y(k) + v(k) \quad (2-4)$$

and $k = 1, 2, \dots, N$. We now collect these N measurements as follows:

$$\underbrace{\begin{pmatrix} z(N) \\ z(N-1) \\ \vdots \\ z(n) \\ \vdots \\ z(2) \\ z(1) \end{pmatrix}}_{\mathcal{Z}(N)} = \underbrace{\begin{pmatrix} u(N-1) & u(N-2) & u(N-3) & \cdots & u(N-n) \\ u(N-2) & u(N-3) & u(N-4) & \cdots & u(N-n-1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ u(n-1) & u(n-2) & u(n-3) & \cdots & u(0) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ u(1) & u(0) & 0 & \cdots & 0 \\ u(0) & 0 & 0 & \cdots & 0 \end{pmatrix}}_{\mathcal{H}(N-1)}$$

$$\times \underbrace{\begin{pmatrix} h(1) \\ h(2) \\ \vdots \\ h(n) \end{pmatrix}}_{\theta} + \underbrace{\begin{pmatrix} v(N) \\ v(N-1) \\ \vdots \\ v(n) \\ \vdots \\ v(2) \\ v(1) \end{pmatrix}}_{\mathcal{V}(N)} \quad (2-5)$$

Examples

Clearly, (2-5) is in the form of (2-1).

In this application the n sampled values of the IR, $h(i)$, play the role of unknown parameters, i.e., $\theta_i = h(1), \theta_2 = h(2), \dots, \theta_n = h(n)$, and these parameters are deterministic. If input $u(k)$ is deterministic and is known ahead of time (or can be measured) without error, then $\mathcal{H}(N-1)$ is deterministic so that we are in case A.1. Often, however, $u(k)$ is random so that $\mathcal{H}(N-1)$ is random; but $u(k)$ is in no way related to measurement noise $v(k)$, so we are in case A.2.a. \square

Example 2-2 Identification of the Coefficients of a Finite-Difference Equation

Suppose a linear, time-invariant, discrete-time system is described by the following n th-order finite-difference equation

$$y(k) + \alpha_1 y(k-1) + \cdots + \alpha_n y(k-n) = u(k-1) \quad (2-6)$$

This model is often referred to as an all-pole or autoregressive (AR) model. It occurs in many branches of engineering and science, including speech modeling, geophysical modeling, etc. Suppose, also, that N perfect measurements of signal $y(k)$ are available. Parameters $\alpha_1, \alpha_2, \dots, \alpha_n$ are unknown and are to be estimated from the data. To do this, we can rewrite (2-6) as

$$y(k) = -\alpha_1 y(k-1) - \cdots - \alpha_n y(k-n) + u(k-1) \quad (2-7)$$

and collect $y(1), y(2), \dots, y(N)$ as we did in Example 2-1. Doing this, we obtain

$$\underbrace{\begin{pmatrix} y(N) \\ y(N-1) \\ \vdots \\ y(n) \\ \vdots \\ y(2) \\ y(1) \end{pmatrix}}_{\mathcal{Z}(N)} = \underbrace{\begin{pmatrix} y(N-1) & y(N-2) & y(N-3) & \cdots & y(N-n) \\ y(N-2) & y(N-3) & y(N-4) & \cdots & y(N-n-1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y(n-1) & y(n-2) & y(n-1) & \cdots & y(0) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y(1) & y(0) & 0 & \cdots & 0 \\ y(0) & 0 & 0 & \cdots & 0 \end{pmatrix}}_{\mathcal{H}(N-1)} \times \underbrace{\begin{pmatrix} -\alpha_1 \\ -\alpha_2 \\ \vdots \\ -\alpha_n \end{pmatrix}}_{\theta} + \underbrace{\begin{pmatrix} u(N-1) \\ u(N-2) \\ \vdots \\ u(n) \\ \vdots \\ u(1) \\ u(0) \end{pmatrix}}_{\mathcal{V}(N-1)} \quad (2-8)$$

which, again, is in the form of (2-1).

In this example $\theta = \text{col}(-\alpha_1, -\alpha_2, \dots, -\alpha_n)$, and these parameters are deterministic. If input $u(k-1)$ is deterministic, then the system's output $y(k)$ will also be deterministic, so that both $\mathcal{H}(N-1)$ and $\mathcal{V}(N-1)$ are deterministic. This is a very special case of case A.1, because usually \mathcal{V} is random. If, however, $u(k-1)$ is random then $y(k)$ will also be random; but, the elements of $\mathcal{H}(N-1)$ will now depend on those in $\mathcal{V}(N-1)$, because $y(k)$ depends upon $u(0), u(1), \dots, u(k-1)$. In this situation we are in case A.2.b. \square

Example 2-3 Identification of the Initial Condition Vector in an Unforced State Equation Model

Consider the problem of identifying the $n \times 1$ initial condition vector $\mathbf{x}(0)$ of the linear, time-invariant, discrete-time system

$$\mathbf{x}(k+1) = \Phi \mathbf{x}(k) \quad (2-9)$$

from the N measurements $z(1), z(2), \dots, z(N)$, where

$$z(k) = \mathbf{h}' \mathbf{x}(k) + v(k) \quad (2-10)$$

The solution to (2-9) is

$$\mathbf{x}(k) = \Phi^k \mathbf{x}(0) \quad (2-11)$$

so that

$$z(k) = \mathbf{h}' \Phi^k \mathbf{x}(0) + v(k) \quad (2-12)$$

Collecting the N measurements, as before, we obtain

$$\underbrace{\begin{pmatrix} z(N) \\ z(N-1) \\ \vdots \\ z(1) \end{pmatrix}}_{\mathcal{Z}(N)} = \underbrace{\begin{pmatrix} \mathbf{h}' \Phi^N \\ \mathbf{h}' \Phi^{N-1} \\ \vdots \\ \mathbf{h}' \Phi \end{pmatrix}}_{\mathcal{H}(N)} \underbrace{\mathbf{x}(0)}_{\theta} + \underbrace{\begin{pmatrix} v(N) \\ v(N-1) \\ \vdots \\ v(1) \end{pmatrix}}_{\mathcal{V}(N)} \quad (2-13)$$

Once again, we have been led to (2-1), and we are in case A.1. \square

Example 2-4 State Estimation

State-variable models are widely used in control and communication theory, and in signal processing. Often, we need the entire state vector of a dynamical system in order to implement an optimal control law for it, or, to implement a digital signal processor. Usually, we cannot measure the entire state vector, and our measurements are corrupted by noise. In state estimation, our objective is to estimate the entire state vector from a limited collection of noisy measurements.

Here we consider the problem of estimating $n \times 1$ state vector $\mathbf{x}(k)$, at $k = 1, 2, \dots, N$ from a scalar measurement $z(k)$, where $k = 1, 2, \dots, N$. The model for this example is

$$\mathbf{x}(k+1) = \Phi \mathbf{x}(k) + \gamma u(k) \quad (2-14)$$

$$z(k) = \mathbf{h}' \mathbf{x}(k) + v(k) \quad (2-15)$$

We are keeping this example simple by assuming that the system is time-invariant and has only one input and one output; however, the results obtained in this example are easily generalized to time-varying and multichannel systems or multichannel systems alone.

If we try to collect our N measurements as before, we obtain

$$\left. \begin{aligned} z(N) &= \mathbf{h}' \mathbf{x}(N) + v(N) \\ z(N-1) &= \mathbf{h}' \mathbf{x}(N-1) + v(N-1) \\ &\dots \\ z(1) &= \mathbf{h}' \mathbf{x}(1) + v(1) \end{aligned} \right\} \quad (2-16)$$

Examples

Observe that a different (unknown) state vector appears in each of the N measurement equations; thus, there does not appear to be a common “ θ ” for the collected measurements. Appearances can sometimes be deceiving.

So far, we have not made use of the state equation. Its solution can be expressed as

$$\mathbf{x}(k) = \Phi^{k-j} \mathbf{x}(j) + \sum_{i=j+1}^k \Phi^{k-i} \gamma u(i-1) \quad (2-17)$$

where $k \geq j+1$. We now focus our attention on the value of $\mathbf{x}(k)$ at $k = k_1$, where $1 \leq k_1 \leq N$. Using (2-17), we can express $\mathbf{x}(N), \mathbf{x}(N-1), \dots, \mathbf{x}(k_1+1)$ as an explicit function of $\mathbf{x}(k_1)$, i.e.,

$$\mathbf{x}(k) = \Phi^{k-k_1} \mathbf{x}(k_1) + \sum_{i=k_1+1}^k \Phi^{k-i} \gamma u(i-1) \quad (2-18)$$

where $k = k_1+1, k_1+2, \dots, N$. In order to do the same for $\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(k_1-1)$, we solve (2-17) for $\mathbf{x}(j)$ and set $k = k_1$,

$$\mathbf{x}(j) = \Phi^{j-k_1} \mathbf{x}(k_1) - \sum_{i=j+1}^{k_1} \Phi^{j-i} \gamma u(i-1) \quad (2-19)$$

where $j = k_1-1, k_1-2, \dots, 2, 1$. Using (2-18) and (2-19), we can reexpress (2-16) as

$$\left. \begin{aligned} z(k) &= \mathbf{h}' \Phi^{k-k_1} \mathbf{x}(k_1) + \mathbf{h}' \sum_{i=k_1+1}^k \Phi^{k-i} \gamma u(i-1) + v(k) \\ k &= N, N-1, \dots, k_1+1 \\ z(k_1) &= \mathbf{h}' \mathbf{x}(k_1) + v(k_1) \\ z(l) &= \mathbf{h}' \Phi^{l-k_1} \mathbf{x}(k_1) - \mathbf{h}' \sum_{i=l+1}^{k_1} \Phi^{l-i} \gamma u(i-1) + v(l) \\ l &= k_1-1, k_1-2, \dots, 1 \end{aligned} \right\} \quad (2-20)$$

These N equations can now be collected together, to give

$$\left. \begin{aligned} \underbrace{\begin{pmatrix} z(N) \\ z(N-1) \\ \vdots \\ z(1) \end{pmatrix}}_{\mathcal{Z}(N)} &= \underbrace{\begin{pmatrix} \mathbf{h}' \Phi^{N-k_1} \\ \mathbf{h}' \Phi^{N-1-k_1} \\ \vdots \\ \mathbf{h}' \Phi^{1-k_1} \end{pmatrix}}_{\mathcal{H}(N, k_1)} \underbrace{\mathbf{x}(k_1)}_{\theta} \\ &+ \underbrace{\mathbf{M}(N, k_1) \begin{pmatrix} u(N-1) \\ u(N-2) \\ \vdots \\ u(0) \end{pmatrix}}_{\mathcal{U}(N, k_1)} + \underbrace{\begin{pmatrix} v(N) \\ v(N-1) \\ \vdots \\ v(1) \end{pmatrix}}_{\mathcal{V}(N, k_1) \quad (2-21)} \end{aligned} \right.$$

where the exact structure of matrix $\mathbf{M}(N, k_1)$ is not important to us at this point. Observe that the state at $k = k_1$ plays the role of parameter vector θ and that both \mathcal{H} and \mathcal{V} are different for different values of k_1 .

If $\mathbf{x}(0)$ and the system input $u(k)$ are deterministic, then $\mathbf{x}(k)$ is deterministic for

all k . In this case θ is deterministic, but $\mathcal{V}(N, k_1)$ is a superposition of deterministic and random components. On the other hand, if either $x(0)$ or $u(k)$ are random then θ is a vector of *random parameters*. This latter situation is the more usual one in state estimation. It corresponds to case B.1. \square

Example 2-5 A Nonlinear Model

Many of the estimation techniques that are described in this book in the context of linear model (2-1) can also be applied to the estimation of unknown signals or parameters in nonlinear models, when such models are suitably linearized. Suppose, for example, that

$$z(k) = f(\theta, k) + v(k) \quad (2-22)$$

where, $k = 1, 2, \dots, N$, and the structure of nonlinear function $f(\theta, k)$ is known explicitly. To see the forest from the trees in this example we assume θ is a scalar parameter.

Let θ^* denote a *nominal value* of θ , $\delta\theta = \theta - \theta^*$, and $\delta z = z - z^*$, where

$$z^*(k) = f(\theta^*, k) \quad (2-23)$$

Observe that the *nominal measurements*, $z^*(k)$, can be computed once θ^* is specified, because $f(\cdot, k)$ is assumed to be known.

Using a first-order Taylor series expansion of $f(\theta, k)$ about $\theta = \theta^*$, it is easy to show that

$$\delta z(k) = \frac{\partial f(\theta, k)}{\partial \theta} \Big|_{\theta=\theta^*} \delta\theta + v(k) \quad (2-24)$$

where $k = 1, 2, \dots, N$. It is easy to see how to collect these N equations, to give

$$\underbrace{\begin{pmatrix} \delta z(N) \\ \vdots \\ \delta z(1) \end{pmatrix}}_{\mathcal{Z}(N)} = \underbrace{\begin{pmatrix} \partial f(\theta^*, N)/\partial\theta^* \\ \vdots \\ \partial f(\theta^*, 1)/\partial\theta^* \end{pmatrix}}_{\mathcal{H}(N, \theta^*)} \delta\theta + \underbrace{\begin{pmatrix} v(N) \\ \vdots \\ v(1) \end{pmatrix}}_{\mathcal{V}(N)} \quad (2-25)$$

in which $\partial f(\theta^*, k)/\partial\theta^*$ is short for “ $\partial f(\theta, k)/\partial\theta$ evaluated at $\theta = \theta^*$.”

Observe that \mathcal{H} depends on θ^* . We will discuss different ways for specifying θ^* in Lesson 25. \square

Example 2-6 Deconvolution (Mendel, 1983b)

In Example 2-1 we showed how a convolutional model could be expressed as the linear model $\mathcal{Z} = \mathcal{H}\theta + \mathcal{V}$. In that example we assumed that both input and output measurements were available, and, we wanted to estimate the sampled values of the system's impulse response. Here we begin with the same convolutional model, written as

$$z(k) = \sum_{i=1}^k \mu(i)h(k-i) + v(k) \quad (2-26)$$

Examples

where $k = 1, 2, \dots, N$. Noisy measurements $z(1), z(2), \dots, z(N)$ are available to us, and we assume that we know the system's impulse response $h(j)$, $\forall j$. What is not known is the input to the system $\mu(1), \mu(2), \dots, \mu(N)$. *Deconvolution is the signal processing procedure for removing the effects of $h(j)$ and $v(j)$ from the measurements so that one is left with an estimate of $\mu(j)$.*

In deconvolution we often assume that input $\mu(j)$ is white noise, but is not necessarily Gaussian. This type of deconvolution problem occurs in reflection seismology. We assume further that

$$\mu(k) = r(k)q(k) \quad (2-27)$$

where $r(k)$ is white Gaussian noise with variance σ_r^2 , and $q(k)$ is a random event location sequence of zeros and ones (a Bernoulli sequence). Sequences $r(k)$ and $q(k)$ are assumed to be statistically independent.

We now collect the N measurements, but in such a way that $\mu(1), \mu(2), \dots, \mu(N)$ are treated as the unknown parameters. Doing this we obtain the following linear deconvolution model:

$$\underbrace{\begin{pmatrix} z(N) \\ z(N-1) \\ \vdots \\ z(2) \\ z(1) \end{pmatrix}}_{\mathcal{Z}(N)} = \underbrace{\begin{pmatrix} h(N-1) & h(N-2) & \dots & h(1) & h(0) \\ h(N-2) & h(N-3) & \dots & h(0) & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ h(1) & h(0) & \dots & 0 & 0 \\ h(0) & 0 & \dots & 0 & 0 \end{pmatrix}}_{\mathcal{H}(N-1)}$$

$$\times \underbrace{\begin{pmatrix} \mu(1) \\ \mu(2) \\ \vdots \\ \mu(N-1) \\ \mu(N) \end{pmatrix}}_{\theta} + \underbrace{\begin{pmatrix} v(N) \\ v(N-1) \\ \vdots \\ v(2) \\ v(1) \end{pmatrix}}_{\mathcal{V}(N)} \quad (2-28)$$

We shall often refer to θ as μ .

Using (2-27), we can also express $\theta = \mu$ as

$$\mu = Q_q r \quad (2-29)$$

where

$$r = \text{col}(r(1), r(2), \dots, r(N)) \quad (2-30)$$

and

$$Q_q = \text{diag}(q(1), q(2), \dots, q(N)) \quad (2-31)$$

In this case (2-28) can be expressed as

$$\mathcal{Z}(N) = \mathcal{H}(N-1)Q_q r + \mathcal{V}(N) \quad (2-32)$$

When event locations $q(1), q(2), \dots, q(N)$ are known, then we can view (2-32) as a linear model for determining \mathbf{r} .

Regardless of which linear deconvolution model we use as our starting point for determining μ , we see that deconvolution corresponds to case B.1. Put another way, we have shown that *the design of a deconvolution signal processing filter is isomorphic to the problem of estimating random parameters in a linear model*. Note, however, that the dimension of θ , which is $N \times 1$, increases as the number of measurements increase. In all other examples θ was $n \times 1$ where n is a fixed integer. We return to this point in Lesson 14 where we discuss convergence of estimates of μ to their true values.

In Lesson 14, we shall develop minimum-variance and maximum-likelihood deconvolution filters. Equation (2-28) is the starting point for derivation of the former filter, whereas Equation (2-32) is the starting point for derivation of the latter filter. \square

NOTATIONAL PRELIMINARIES

Equation (2-1) can be interpreted as a *data generating model*; it is a mathematical representation that is associated with the data. Parameter vector θ is assumed to be unknown and is to be estimated using $\mathcal{E}(k)$, $\mathcal{H}(k)$ and possibly other a priori information. We use $\hat{\theta}(k)$ to denote the estimate of constant parameter vector θ . Argument k in $\hat{\theta}(k)$ denotes the fact that the estimate is based on measurements up to and including the k th. In our preceding examples, we would use the following notation for $\hat{\theta}(k)$:

- Example 2-1 [see (2-5)]: $\hat{\theta}(N)$ with components $\hat{h}(i | N)$
- Example 2-2 [see (2-8)]: $\hat{\theta}(N)$
- Example 2-3 [see (2-13)]: $\hat{x}(0 | N)$
- Example 2-4 [see (2-21)]: $\hat{x}(k_1 | N)$
- Example 2-5 [see (2-25)]: $\hat{\theta}\hat{\theta}(N)$
- Example 2-6 [see (2-28)]: $\hat{\theta}(N)$ with components $\hat{\mu}(i | N)$

The notation used in Examples 1, 3, 4, and 6 is a bit more complicated than that used in the other examples, because we must indicate the time point at which we are estimating the quantity of interest (e.g., k_1 or i) as well as the last data point used to obtain this estimate (e.g., N). We often read $\hat{x}(k_1 | N)$ as “the estimate of $x(k_1)$ conditioned on N ” or as “ x hat at k_1 conditioned on N .”

In state estimation (or deconvolution) three situations are possible depending upon the relative relationship of N to k_1 . For example, when $N < k_1$ we are estimating a future value of $x(k_1)$, and we refer to this as a *predicted estimate*. When $N = k_1$ we are using all past measurements and the most recent measurement to estimate $x(k_1)$. The result is referred to as a *filtered estimate*. Finally, when $N > k_1$ we are estimating an earlier value of $x(k_1)$ using past, present and future measurements. Such an estimate is referred to as a

smoothed or *interpolated* estimate. Prediction and filtering can be done in real time whereas smoothing can never be done in real time. We will see that the impulse responses of predictors and filters are causal, whereas the impulse response of a smoother is noncausal.

We use $\tilde{\theta}(k)$ to denote *estimation error*, i.e.,

$$\tilde{\theta}(k) = \theta - \hat{\theta}(k) \quad (2-33)$$

In state estimation, $\bar{x}(k_1 | N)$ denotes state estimation error, and, $\bar{x}(k_1 | N) = x(k_1) - \hat{x}(k_1 | N)$. In deconvolution $\hat{\mu}(i | N)$ is defined in a similar manner.

Very often we use the following *estimation model* for $\mathcal{E}(k)$,

$$\hat{\mathcal{E}}(k) = \mathcal{H}(k)\hat{\theta}(k) \quad (2-34)$$

To obtain (2-34) from (2-1), we assume that $\mathcal{V}(k)$ is zero-mean random noise that cannot be measured. In some applications (e.g., Example 2-2) $\hat{\mathcal{E}}(k)$ represents a predicted value of $\mathcal{E}(k)$. Associated with $\hat{\mathcal{E}}(k)$ is the error $\tilde{\mathcal{E}}(k)$, where

$$\tilde{\mathcal{E}}(k) = \mathcal{E}(k) - \hat{\mathcal{E}}(k) \quad (2-35)$$

satisfies the equation

$$\tilde{\mathcal{E}}(k) = \mathcal{H}(k)\tilde{\theta}(k) + \mathcal{V}(k) \quad (2-36)$$

In those applications where $\tilde{\mathcal{E}}(k)$ is a predicted value of $\mathcal{E}(k)$, $\tilde{\mathcal{E}}(k)$ is known as a *prediction error*. Other names for $\tilde{\mathcal{E}}(k)$ are *equation error* and *measurement residual*.

In the rest of this book we develop specific structures for $\hat{\theta}(k)$. These structures are referred to as *estimators*. *Estimates* are obtained whenever data is processed by an estimator. Estimator structures are associated with specific estimation techniques, and these techniques can be classified according to the natures of θ and $\mathcal{H}(k)$, and what a priori information is assumed known about noise vector $\mathcal{V}(k)$. See Lesson 1 for an overview of all the different estimation techniques that are covered in this book.

PROBLEMS

1. Suppose $z(k) = \theta_1 + \theta_2 k + v(k)$, where $z(1) = 0.2$, $z(2) = 1.4$, $z(3) = 3.6$, $z(4) = 7.5$, and $z(5) = 10.2$. What are the explicit structures of $\mathcal{E}(5)$ and $\mathcal{H}(5)$?
- 2-2. According to thermodynamic principles, pressure P and volume V of a given mass of gas are related by $PV^\gamma = C$, where γ and C are constants. Assume that N measurements of P and V are available. Explain how to obtain a linear model for estimation of parameters γ and $\ln C$.
- 2-3. (Mendel, 1973, Exercise 1-16(a), pg. 46). Suppose we know that a relationship exists between y and x_1, x_2, \dots, x_n of the form

$$y = \exp(a_1x_1 + a_2x_2 + \dots + a_nx_n)$$

We desire to estimate a_1, a_2, \dots, a_n from measurements of y and $x = \text{col}(x_1, x_2, \dots, x_n)$. Explain how to do this.

- 2-4. (Mendel, 1973, Exercise 1-17, pp. 46–47). The efficiency of a jet engine may be viewed as a linear combination of functions of inlet pressure $p(t)$ and operating temperature $T(t)$; that is to say,

$$E(t) = C_1 + C_2 f_1[p(t)] + C_3 f_2[T(t)] + C_4 f_3[p(t), T(t)] + \nu(t)$$

where the structures of f_1, f_2 , and f_3 are known a priori and $\nu(t)$ represents modeling error of known mean and variance. From tests on the engine a table of values of $E(t)$, $p(t)$, and $T(t)$ are given at discrete values of t . Explain how C_1, C_2, C_3 , and C_4 are estimated from these data.

Lesson 3

Least-Squares

Estimation:

Batch Processing

INTRODUCTION

The method of least squares dates back to Karl Gauss around 1795, and is the cornerstone for most estimation theory, both classical and modern. It was invented by Gauss at a time when he was interested in predicting the motion of planets and comets using telescopic measurements. The motions of these bodies can be completely characterized by six parameters. The estimation problem that Gauss considered was one of inferring the values of these parameters from the measurement data.

We shall study least-squares estimation from two points of view: the classical batch-processing approach, in which all the measurements are processed together at one time, and the more modern recursive processing approach, in which measurements are processed only a few (or even one) at a time. The recursive approach has been motivated by today's high-speed digital computers; however, as we shall see, the recursive algorithms are outgrowths of the batch algorithms. In fact, as we enter the era of very large scale integration (VLSI) technology, it may well be that VLSI implementations of the batch algorithms are faster than digital computer implementations of the recursive algorithms.

The starting point for the method of least squares is the linear model

$$\mathfrak{Z}(k) = \mathcal{H}(k)\theta + \mathcal{V}(k) \quad (3-1)$$

where $\mathfrak{Z}(k) = \text{col}(z(k), z(k-1), \dots, z(k-N+1))$, $z(k) = \mathbf{h}'(k)\theta + \nu(k)$, and the estimation model for $\mathfrak{Z}(k)$ is

$$\hat{\mathfrak{Z}}(k) = \mathcal{H}(k)\hat{\theta}(k) \quad (3-2)$$

We denote the (weighted) least-squares estimator of θ as $[\hat{\theta}_{WLS}(k)]\hat{\theta}_{LS}(k)$. In this lesson and the next two we shall determine explicit structures for this estimator.

NUMBER OF MEASUREMENTS

Suppose that θ contains n parameters and $\mathcal{Z}(k)$ contains N measurements. If $N < n$ we have fewer measurements than unknowns and (3-1) is an underdetermined system of equations that does not lead to unique or very meaningful values for $\theta_1, \theta_2, \dots, \theta_n$. If $N = n$, we have exactly as many measurements as unknowns, and as long as the n measurements are linearly independent, so that $\mathcal{H}^{-1}(k)$ exists, we can solve (3-1) for θ , as

$$\theta = \mathcal{H}^{-1}(k)\mathcal{Z}(k) - \mathcal{H}^{-1}(k)\mathcal{V}(k) \quad (3-3)$$

Because we cannot measure $\mathcal{V}(k)$, it is usually neglected in the calculation of (3-3). For small amounts of noise this may not be a bad thing to do but for even moderate amounts of noise this will be quite bad. Finally, if $N > n$ we have more measurements than unknowns, so that (3-1) is an overdetermined system of equations. The extra measurements can be used to offset the effects of the noise; i.e., they let us “filter” the data. Only this last case is of real interest to us.

OBJECTIVE FUNCTION AND PROBLEM STATEMENT

Our method for obtaining $\hat{\theta}(k)$ is based on minimizing the objective function

$$J[\hat{\theta}(k)] = \tilde{\mathcal{Z}}'(k)\mathbf{W}(k)\tilde{\mathcal{Z}}(k) \quad (3-4)$$

where

$$\tilde{\mathcal{Z}}(k) = \text{col}[\tilde{z}(k), \tilde{z}(k-1), \dots, \tilde{z}(k-N+1)] \quad (3-5)$$

and *weighting matrix* $\mathbf{W}(k)$ must be symmetric and positive definite, for reasons explained below.

No general rules exist for how to choose $\mathbf{W}(k)$. The most common choice is a diagonal matrix such as

$$\mathbf{W}(k) = \text{diag}[\mu^{k-N+1}, \mu^{k-N+2}, \dots, \mu^k]$$

When $|\mu| < 1$ recent measurements are weighted more heavily than past ones. Such a choice for $\mathbf{W}(k)$ provides the weighted least-squares estimator with an “aging” or “forgetting” factor. When $|\mu| > 1$ recent measurements are weighted less heavily than past ones. Finally, if $\mu = 1$, so that $\mathbf{W}(k) = \mathbf{I}$, then all measurements are weighted by the same amount. When $\mathbf{W}(k) = \mathbf{I}$, $\hat{\theta}(k) = \hat{\theta}_{LS}(k)$, whereas for all other $\mathbf{W}(k)$, $\hat{\theta}(k) = \hat{\theta}_{WLS}(k)$.

Our objective is to determine the $\hat{\theta}_{WLS}(k)$ that minimizes $J[\hat{\theta}(k)]$.

DERIVATION OF ESTIMATOR

To begin, we express (3-4) as an explicit function of $\hat{\theta}(k)$, using (3-2):

$$\begin{aligned} J[\hat{\theta}(k)] &= \tilde{\mathcal{Z}}'(k)\mathbf{W}(k)\tilde{\mathcal{Z}}(k) = [\mathcal{Z}(k) - \mathcal{H}(k)]'\mathbf{W}(k)[\mathcal{Z}(k) - \mathcal{H}(k)] \\ &= [\mathcal{Z}(k) - \mathcal{H}(k)\hat{\theta}(k)]'\mathbf{W}(k)[\mathcal{Z}(k) - \mathcal{H}(k)\hat{\theta}(k)] \\ &= \mathcal{Z}'(k)\mathbf{W}(k)\mathcal{Z}(k) - 2\mathcal{Z}'(k)\mathbf{W}(k)\mathcal{H}(k)\hat{\theta}(k) \\ &\quad + \hat{\theta}'(k)\mathcal{H}'(k)\mathbf{W}(k)\mathcal{H}(k)\hat{\theta}(k) \end{aligned} \quad (3-6)$$

Next, we take the vector derivative of $J[\hat{\theta}(k)]$ with respect to $\hat{\theta}(k)$, but before doing this recall from vector calculus, that:

If \mathbf{m} and \mathbf{b} are two $n \times 1$ nonzero vectors, and \mathbf{A} is an $n \times n$ symmetric matrix, then

$$\frac{d}{d\mathbf{m}}(\mathbf{b}'\mathbf{m}) = \mathbf{b} \quad (3-7)$$

and

$$\frac{d}{d\mathbf{m}}(\mathbf{m}'\mathbf{A}\mathbf{m}) = 2\mathbf{A}\mathbf{m} \quad (3-8)$$

Using these formulas, we find that

$$\frac{dJ[\hat{\theta}(k)]}{d\hat{\theta}(k)} = -2[\mathcal{Z}'(k)\mathbf{W}(k)\mathcal{H}(k)]' + 2\mathcal{H}'(k)\mathbf{W}(k)\mathcal{H}(k)\hat{\theta}(k) \quad (3-9)$$

Setting $dJ[\hat{\theta}(k)]/d\hat{\theta}(k) = \mathbf{0}$ we obtain the following formula for $\hat{\theta}_{WLS}(k)$,

$$\hat{\theta}_{WLS}(k) = [\mathcal{H}'(k)\mathbf{W}(k)\mathcal{H}(k)]^{-1}\mathcal{H}'(k)\mathbf{W}(k)\mathcal{Z}(k) \quad (3-10)$$

Note, also, that

$$\hat{\theta}_{LS}(k) = [\mathcal{H}'(k)\mathcal{H}(k)]^{-1}\mathcal{H}'(k)\mathcal{Z}(k) \quad (3-11)$$

Comments

- For (3-10) to be valid, $\mathcal{H}'(k)\mathbf{W}(k)\mathcal{H}(k)$ must be invertible. This is always true when $\mathbf{W}(k)$ is positive definite, as assumed, and $\mathcal{H}(k)$ is of maximum rank.
- How do we know that $\hat{\theta}_{WLS}(k)$ minimizes $J[\hat{\theta}(k)]$? We compute $d^2J[\hat{\theta}(k)]/d\hat{\theta}^2(k)$ and see if it is positive definite [which is the vector calculus analog of the scalar calculus requirement that $\hat{\theta}$ minimizes $J(\hat{\theta})$ if $dJ(\hat{\theta})/d\hat{\theta} = 0$ and $d^2J(\hat{\theta})/d\hat{\theta}^2$ is positive]. Doing this, we see that

$$\frac{d^2J[\hat{\theta}(k)]}{d\hat{\theta}^2(k)} = 2\mathcal{H}'(k)\mathbf{W}(k)\mathcal{H}(k) > 0 \quad (3-12)$$

because $\mathcal{H}'(k)\mathbf{W}(k)\mathcal{H}(k)$ is invertible.

3. Estimator $\hat{\theta}_{WLS}(k)$ processes the measurements $\tilde{z}(k)$ linearly; thus, it is referred to as a *linear estimator*. It processes the data contained in $\mathcal{H}(k)$ in a very complicated and nonlinear manner.
4. When (3-9) is set equal to zero we obtain the following system of *normal equations*

$$[\mathcal{H}'(k)\mathbf{W}(k)\mathcal{H}(k)]\hat{\theta}_{WLS}(k) = \mathcal{H}'(k)\mathbf{W}(k)\tilde{z}(k) \quad (13-13)$$

This is a system of n linear equations in the n components of $\hat{\theta}_{WLS}(k)$.

In practice, one does not compute $\hat{\theta}_{WLS}(k)$ using (3-10), because computing the inverse of $\mathcal{H}'(k)\mathbf{W}(k)\mathcal{H}(k)$ is fraught with numerical difficulties. Instead, the normal equations are solved using stable algorithms from numerical linear algebra that involve orthogonal transformations (see, e.g., Stewart, 1973; Bierman, 1977; and Dongarra, et al., 1979). Because it is not the purpose of this book to go into details of numerical linear algebra, we leave it to the reader to pursue this important subject.

Based on this discussion, we must view (3-10) as a useful “*theoretical*” formula and not as a useful computational formula. Remember that this is a book on estimation *theory*, so for our purposes, theoretical formulas are just fine.

5. Equation (3-13) can also be reexpressed as

$$\tilde{z}'(k)\mathbf{W}(k)\mathcal{H}(k) = 0' \quad (13-14)$$

which can be viewed as an *orthogonality condition* between $\tilde{z}(k)$ and $\mathbf{W}(k)\mathcal{H}(k)$. Orthogonality conditions play an important role in estimation theory. We shall see many more examples of such conditions throughout this book.

6. Estimates obtained from (3-10) will be random! This is because $\tilde{z}(k)$ is random, and, in some applications even $\mathcal{H}(k)$ is random. It is therefore instructive to view (3-10) as a complicated transformation of vectors or matrices of random variables into the vector of random variables $\hat{\theta}_{WLS}(k)$. In later lessons, when we examine the properties of $\hat{\theta}_{WLS}(k)$, these will be statistical properties, because of the random nature of $\hat{\theta}_{WLS}(k)$.

Example 3-1 (Mendel, 1973, pp. 86–87)

Suppose we wish to calibrate an instrument by making a series of uncorrelated measurements on a constant quantity. Denoting the constant quantity as θ , our measurement equation becomes

$$z(k) = \theta + v(k) \quad (3-15)$$

where $k = 1, 2, \dots, N$. Collecting these N measurements, we have

$$\begin{pmatrix} z(N) \\ z(N-1) \\ \vdots \\ z(1) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}\theta + \begin{pmatrix} v(N) \\ v(N-1) \\ \vdots \\ v(1) \end{pmatrix} \quad (3-16)$$

Derivation of Estimator

Clearly, $\mathcal{H} = \text{col}(1, 1, \dots, 1)$; hence;

$$\hat{\theta}_{LS}(N) = \frac{1}{N} \sum_{i=1}^N z(i) \quad (3-17)$$

which is the sample mean of the N measurements. We see, therefore, that *the sample mean is a least-squares estimator*. \square

Example 3-2 (Mendel, 1973)

Figure 3-1 depicts simplified third-order pitch-plane dynamics for a typical, high-performance, aerodynamically controlled aerospace vehicle. Cross-coupling and body-bending effects are neglected. Normal acceleration control is considered with feedback on normal acceleration and angle-of-attack rate. Stefani (1967) shows that if the system gains are chosen as

$$K_{Ni} = \frac{C_2}{100 M_\delta Z_\alpha} \quad (3-18)$$

$$K_{\dot{\alpha}} = \frac{C_1 - 100 \left(\frac{Z_\alpha 1845}{\mu} \right) + M_\alpha}{100 M_\delta} \quad (3-19)$$

and

$$K_{Na} = \frac{C_2 + 100 M_\alpha}{100 M_\delta Z_\alpha} \quad (3-20)$$

then

$$\frac{N_\alpha}{N_i}(s) = \frac{C_2}{s^3 + 100s^2 + C_1 s + C_2} \quad (3-21)$$

Stefani assumes $Z_\alpha 1845/\mu$ is relatively small, and chooses $C_1 = 1400$ and $C_2 = 14,000$. The closed-loop response resembles that of a second-order system with a bandwidth of 2 Hz and a damping ratio of 0.6 that responds to a step command of input acceleration with zero steady-state error.

In general, M_α , M_δ , and Z_α are dynamic parameters and all vary through a large range of values. Also, M_α may be positive (unstable vehicle) or negative (stable vehicle). System response must remain the same for all values of M_α , M_δ , and Z_α ; thus, it is necessary to estimate these parameters so that K_{Ni} , $K_{\dot{\alpha}}$, and K_{Na} can be adapted to keep C_1 and C_2 invariant at their designed values. For present purposes we shall assume that M_α , M_δ , and Z_α are frozen at specific values.

From Fig. 3-1,

$$\ddot{\theta}(t) = M_\alpha \alpha(t) + M_\delta \delta(t) \quad (3-22)$$

and

$$N_\alpha(t) = Z_\alpha \alpha(t) \quad (3-23)$$

Our attention is directed at the estimation of M_α and M_δ in (3-22). We leave it as an exercise for the reader to explore the estimation of Z_α in (3-23).

Our approach will be to estimate M_α and M_δ from the equation

$$\ddot{\theta}_m(k) = M_\alpha \alpha(k) + M_\delta \delta(k) + v_\alpha(k) \quad (3-24)$$

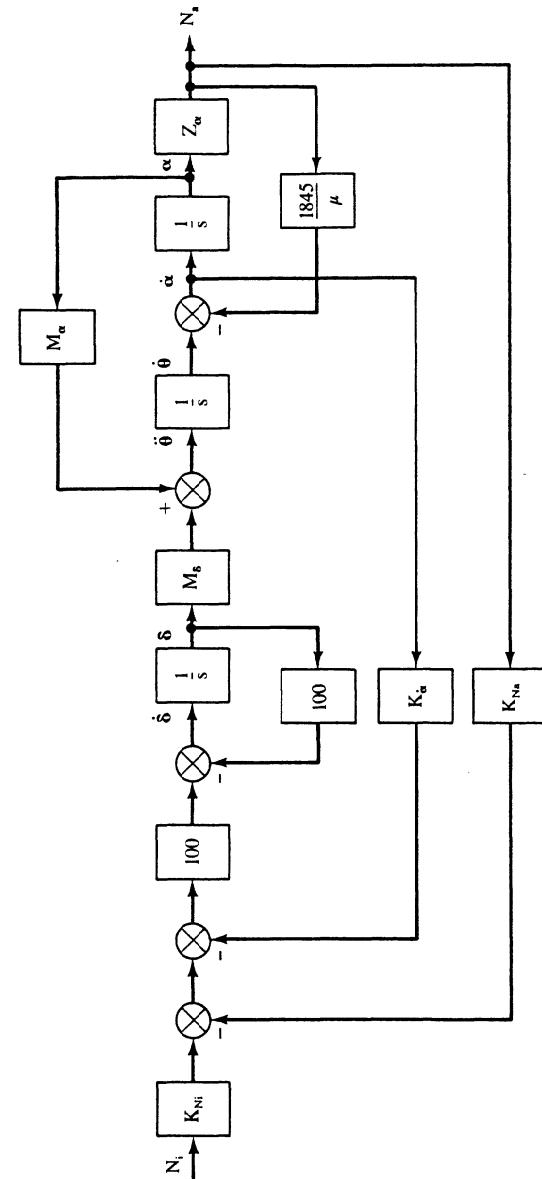


Figure 3-1 Pitch-plane dynamics and nomenclature: N_i , input normal acceleration along the negative Z_a axis; $K_{N\dot{a}}$, gain on N_i ; δ , control-surface deflection; M_s , control-surface effectiveness; θ , rigid-body acceleration; α , angle-of-attack; M_a , aerodynamic moment effectiveness; K_a , control gain on $\dot{\alpha}$; Z_a , normal acceleration along the negative Z axis; μ , axial velocity; N_a , system-achieved normal acceleration along the negative Z_a axis; K_{Na} , control gain on N_a (reprinted from Mendel, 1973, p. 33, by courtesy of Marcel Dekker, Inc.).

Scale Changes

where $\ddot{\theta}_m(k)$ denotes the measured value of $\ddot{\theta}(k)$, that is corrupted by measurement noise $v_{\ddot{\theta}}(k)$. We shall assume (somewhat unrealistically) that $\alpha(k)$ and $\delta(k)$ can both be measured perfectly. The concatenated measurement equation for N measurements is

$$\begin{pmatrix} \ddot{\theta}_m(k) \\ \ddot{\theta}_m(k-1) \\ \vdots \\ \ddot{\theta}_m(k-N+1) \end{pmatrix} = \begin{pmatrix} \alpha(k) & \delta(k) \\ \alpha(k-1) & \delta(k-1) \\ \vdots & \vdots \\ \alpha(k-N+1) & \delta(k-N+1) \end{pmatrix} \times \begin{pmatrix} M_a \\ M_s \end{pmatrix} + \begin{pmatrix} v_{\ddot{\theta}}(k) \\ v_{\ddot{\theta}}(k-1) \\ \vdots \\ v_{\ddot{\theta}}(k-N+1) \end{pmatrix} \quad (3-25)$$

Hence, the least-squares estimates of M_a and M_s are

$$\begin{pmatrix} \hat{M}_a(k) \\ \hat{M}_s(k) \end{pmatrix} = \begin{pmatrix} \sum_{j=0}^{N-1} \alpha^2(k-j) & \sum_{j=0}^{N-1} \alpha(k-j)\delta(k-j) \\ \sum_{j=0}^{N-1} \alpha(k-j)\delta(k-j) & \sum_{j=0}^{N-1} \delta^2(k-j) \end{pmatrix}^{-1} \times \begin{pmatrix} \sum_{j=0}^{N-1} \alpha(k-j)\ddot{\theta}_m(k-j) \\ \sum_{j=0}^{N-1} \delta(k-j)\ddot{\theta}_m(k-j) \end{pmatrix} \quad \square \quad (3-26)$$

FIXED AND EXPANDING MEMORY ESTIMATORS

Estimator $\hat{\theta}_{WLS}(k)$ uses the measurements $z(k-N+1), z(k-N+2), \dots, z(k)$. When N is fixed ahead of time, $\hat{\theta}_{WLS}(k)$ uses a fixed window of measurements, a window of length N , and, $\hat{\theta}_{WLS}(k)$ is then referred to as a *fixed-memory estimator*. A second approach for choosing N is to set it equal to k ; then, $\hat{\theta}_{WLS}(k)$ uses the measurements $z(1), z(2), \dots, z(k)$. In this case $\hat{\theta}_{WLS}(k)$ uses an expanding window of measurements, a window of length k , and, $\hat{\theta}_{WLS}(k)$ is then referred to as an *expanding-memory estimator*.

SCALE CHANGES

Least-squares (LS) estimates may not be invariant under changes of scale. One way to circumvent this difficulty is to use normalized data.

For example, assume that observers A and B are observing a process; but observer A reads the measurements in one set of units and B in another.

Let \mathbf{M} be a symmetric matrix of scale factors relating A to B ; $\mathcal{Z}_A(k)$ and $\mathcal{Z}_B(k)$ denote the total measurement vectors of A and B , respectively. Then

$$\mathcal{Z}_B(k) = \mathcal{H}_B(k)\theta + \mathcal{V}_B(k) = \mathbf{M}\mathcal{Z}_A(k) = \mathbf{M}\mathcal{H}_A(k)\theta + \mathbf{M}\mathcal{V}_A(k) \quad (3-27)$$

which means that

$$\mathcal{H}_B(k) = \mathbf{M}\mathcal{H}_A(k) \quad (3-28)$$

Let $\hat{\theta}_{A,WLS}(k)$ and $\hat{\theta}_{B,WLS}(k)$ denote the WLSE's associated with observers A and B , respectively; then $\hat{\theta}_{B,WLS}(k) = \hat{\theta}_{A,WLS}(k)$ if

$$\mathbf{W}_B(k) = \mathbf{M}^{-1}\mathbf{W}_A(k)\mathbf{M}^{-1} \quad (3-29)$$

It seems a bit peculiar though to have different weighting matrices for the two WLSE's. In fact, if we begin with $\hat{\theta}_{A,LS}(k)$ then it is impossible to obtain $\hat{\theta}_{B,LS}(k)$ such that $\hat{\theta}_{B,LS}(k) = \hat{\theta}_{A,LS}(k)$. The reason for this is simple. To obtain $\hat{\theta}_{A,LS}(k)$, we set $\mathbf{W}_A(k) = \mathbf{I}$, in which case (3-29) reduces to $\mathbf{W}_B(k) = (\mathbf{M}^{-1})^2 \neq \mathbf{I}$.

Next, let \mathbf{N}_A and \mathbf{N}_B denote symmetric normalization matrices for $\mathcal{Z}_A(k)$ and $\mathcal{Z}_B(k)$, respectively. We shall assume that our data is always normalized to the same set of numbers, i.e., that

$$\mathbf{N}_A\mathcal{Z}_A(k) = \mathbf{N}_B\mathcal{Z}_B(k) \quad (3-30)$$

Observe that

$$\mathbf{N}_A\mathcal{Z}_A(k) = \mathbf{N}_A\mathcal{H}_A(k)\theta + \mathbf{N}_A\mathcal{V}_A(k) \quad (3-31)$$

and

$$\mathbf{N}_B\mathcal{Z}_B(k) = \mathbf{N}_B\mathbf{M}\mathcal{Z}_A(k) = \mathbf{N}_B\mathbf{M}\mathcal{H}_A(k)\theta + \mathbf{N}_B\mathbf{M}\mathcal{V}_A(k) \quad (3-32)$$

From (3-30), (3-31), and (3-32), we see that

$$\mathbf{N}_A = \mathbf{N}_B\mathbf{M} \quad (3-33)$$

We now find that

$$\hat{\theta}_{A,WLS}(k) = (\mathcal{H}_A'\mathbf{N}_A\mathbf{W}_A\mathbf{N}_A\mathcal{H}_A)^{-1}\mathcal{H}_A'\mathbf{N}_A\mathbf{W}_A\mathbf{N}_A\mathcal{Z}_A(k) \quad (3-34)$$

and

$$\hat{\theta}_{B,WLS}(k) = (\mathcal{H}_A'\mathbf{M}\mathbf{N}_B\mathbf{W}_B\mathbf{N}_B\mathcal{H}_A)^{-1}\mathcal{H}_A'\mathbf{M}\mathbf{N}_B\mathbf{W}_B\mathbf{N}_B\mathcal{Z}_A(k) \quad (3-35)$$

Substituting (3-33) into (3-35), we then find

$$\hat{\theta}_{B,WLS}(k) = (\mathcal{H}_A'\mathbf{N}_A\mathbf{W}_B\mathbf{N}_A\mathcal{H}_A)^{-1}\mathcal{H}_A'\mathbf{N}_A\mathbf{W}_B\mathbf{N}_A\mathcal{Z}_A(k) \quad (3-36)$$

Comparing (3-36) and (3-34), we conclude that $\hat{\theta}_{B,WLS}(k) = \hat{\theta}_{A,WLS}(k)$ if $\mathbf{W}_B(k) = \mathbf{W}_A(k)$. This is precisely the result we were looking for. It means that, under proper normalization, $\hat{\theta}_{B,WLS}(k) = \hat{\theta}_{A,WLS}(k)$, and, as a special case $\hat{\theta}_{B,LS}(k) = \hat{\theta}_{A,LS}(k)$.

PROBLEMS

- 3-1. Derive the formula for $\hat{\theta}_{WLS}(k)$ by completing the square on the right-hand side of the expression for $J[\hat{\theta}(k)]$ in (3-6).
- 3-2. Here we explore the estimation of Z_a in (3-23). Assume that N noisy measurements of $N_a(k)$ are available, i.e., $N_{am}(k) = Z_a\alpha(k) + v_{Na}(k)$. What is the formula for the least-squares estimator of Z_a ?
- 3-3. Here we explore the simultaneous estimation of M_a , M_δ , and Z_a in (3-22) and (3-23). Assume that N noisy measurements of $\hat{\theta}(k)$ and $N_a(k)$ are available, i.e., $\hat{\theta}_m(k) = M_a\alpha(k) + M_\delta\delta(k) + v_\theta(k)$ and $N_{am}(k) = Z_a\alpha(k) + v_{Na}(k)$. Determine the least-squares estimator of M_a , M_δ , and Z_a . Is this estimator different from $\hat{M}_{a,LS}$ and $\hat{M}_{\delta,LS}$ obtained just from $\hat{\theta}_m(k)$ measurements, and $\hat{Z}_{a,LS}$ obtained just from $N_{am}(k)$ measurements?
- 3-4. In a *curve fitting problem* we wish to fit a given set of data $z(1), z(2), \dots, z(N)$ by the approximating function

$$\hat{z}(k) = \sum_{j=1}^n \hat{\theta}_j \phi_j(k)$$

where $\phi_j(k)$ ($j = 1, 2, \dots, n$) are a set of prespecified basis functions.

- (a) Obtain a formula for $\hat{\theta}_{LS}(N)$ that is valid for any set of basis functions.
- (b) The simplest approximating function to a set of data is the straight line. In this case $\hat{z}(k) = \hat{\theta}_1 + \hat{\theta}_2 k$, which is known as the least-squares or regression line. Obtain closed-form formulas for $\hat{\theta}_{1,LS}(N)$ and $\hat{\theta}_{2,LS}(N)$.
- 3-5. Suppose $z(k) = \theta_1 + \theta_2 k$, where $z(1) = 3$ miles per hour and $z(2) = 7$ miles per hour. Determine $\hat{\theta}_{1,LS}$ and $\hat{\theta}_{2,LS}$ based on these two measurements. Next, redo these calculations by scaling $z(1)$ and $z(2)$ to the units of feet per second. Are the least-squares estimates obtained from these two calculations the same? Use the results developed in the section entitled "Scale Changes" to explain what has happened here.
- 3-6. (a) Under what conditions on scaling matrix \mathbf{M} is scale invariance preserved for a least-squares estimator?
 (b) If our original model is nonlinear in the measurements [e.g., $z(k) = \theta z^2(k-1) + v(k)$] can anything be done to obtain invariant WLSE's under scaling?

Lesson 4

Least-Squares

Estimation:

Recursive Processing

INTRODUCTION

In Lesson 3 we assumed that $\mathbf{z}(k)$ contained N elements, where $N > \dim \boldsymbol{\theta} = n$. Suppose we decide to add more measurements, increasing the total number of them from N to N' . Formula (3-10) in Lesson 3 would not make use of the previously calculated value of $\hat{\boldsymbol{\theta}}$ that is based on N measurements during the calculation of $\hat{\boldsymbol{\theta}}$ that is based on N' measurements. This seems quite wasteful. We intuitively feel that it should be possible to compute the estimate based on N' measurements from the estimate based on N measurements, and a modification of this earlier estimate to account for the $N' - N$ new measurements. In this lesson we shall justify our intuition.

In Lesson 3 we also assumed that $\hat{\boldsymbol{\theta}}$ is determined for a fixed value of n . In many system modeling problems one is interested in a preliminary model in which dimension n is a variable. This is becoming increasingly more important as we begin to model large scale societal, energy, economic, etc. systems in which it may not be clear at the onset what effects are most important. One approach is to recompute $\hat{\boldsymbol{\theta}}$ by means of Formula (3-10) in Lesson 3 for different values of n . This may be very costly, especially for large scale systems, since the number of flops to compute $\hat{\boldsymbol{\theta}}$ is on the order of n^3 . A second approach is to obtain $\hat{\boldsymbol{\theta}}$ for $n = n_1$, and to use that estimate in a computationally effective manner to obtain $\hat{\boldsymbol{\theta}}$ for $n = n_2$, where $n_2 > n_1$. These estimators are recursive in the dimension of $\boldsymbol{\theta}$. We shall also examine these estimators.

RECURSIVE LEAST-SQUARES: INFORMATION FORM

To begin, we consider the case when one additional measurement $z(k+1)$, made at t_{k+1} , becomes available:

$$z(k+1) = \mathbf{h}'(k+1)\boldsymbol{\theta} + v(k+1) \quad (4-1)$$

When this equation is combined with our earlier linear model we obtain a new linear model,

$$\mathbf{z}(k+1) = \mathcal{H}(k+1)\boldsymbol{\theta} + \mathcal{V}(k+1) \quad (4-2)$$

where

$$\mathbf{z}(k+1) = \text{col}(z(k+1)|\mathbf{z}(k)) \quad (4-3)$$

$$\mathcal{H}(k+1) = \left(\frac{\mathbf{h}'(k+1)}{\mathcal{H}(k)} \right) \quad (4-4)$$

and

$$\mathcal{V}(k+1) = \text{col}(v(k+1)|\mathcal{V}(k)) \quad (4-5)$$

Using (3-10) from Lesson 3 and (4-2), it is clear that

$$\begin{aligned} \hat{\boldsymbol{\theta}}_{WLS}(k+1) \\ = [\mathcal{H}'(k+1)\mathbf{W}(k+1)\mathcal{H}(k+1)]^{-1}\mathcal{H}'(k+1)\mathbf{W}(k+1)\mathbf{z}(k+1) \end{aligned} \quad (4-6)$$

To proceed further we must assume that \mathbf{W} is diagonal, i.e.,

$$\mathbf{W}(k+1) = \text{diag}(w(k+1)|\mathbf{W}(k)) \quad (4-7)$$

We shall now show that it is possible to determine $\hat{\boldsymbol{\theta}}(k+1)$ from $\hat{\boldsymbol{\theta}}(k)$ and $z(k+1)$.

Theorem 4-1. (Information Form of Recursive LSE). *A recursive structure for $\hat{\boldsymbol{\theta}}_{WLS}(k)$ is*

$$\hat{\boldsymbol{\theta}}_{WLS}(k+1) = \hat{\boldsymbol{\theta}}_{WLS}(k) + \mathbf{K}_w(k+1)[z(k+1) - \mathbf{h}'(k+1)\hat{\boldsymbol{\theta}}_{WLS}(k)] \quad (4-8)$$

where

$$\mathbf{K}_w(k+1) = \mathbf{P}(k+1)\mathbf{h}(k+1)w(k+1) \quad (4-9)$$

and

$$\mathbf{P}^{-1}(k+1) = \mathbf{P}^{-1}(k) + \mathbf{h}(k+1)w(k+1)\mathbf{h}'(k+1) \quad (4-10)$$

These equations are initialized by $\hat{\boldsymbol{\theta}}_{WLS}(n)$ and $\mathbf{P}^{-1}(n)$, where $\mathbf{P}(k)$ is defined below in (4-13), and are used for $k = n, n+1, \dots, N-1$.

Proof. Substitute (4-3), (4-4) and (4-7) into (4-6) (sometimes dropping the dependence upon k and $k + 1$, for notational simplicity) to see that

$$\begin{aligned}\hat{\theta}_{\text{WLS}}(k+1) &= [\mathcal{H}'(k+1)\mathbf{W}(k+1)\mathcal{H}(k+1)]^{-1}[\mathbf{h}wz + \mathcal{H}'\mathbf{W}\mathcal{Z}] \quad (4-11)\end{aligned}$$

Express $\hat{\theta}_{\text{WLS}}(k)$ as

$$\hat{\theta}_{\text{WLS}}(k) = \mathbf{P}(k)\mathcal{H}'(k)\mathbf{W}(k)\mathcal{Z}(k) \quad (4-12)$$

where

$$\mathbf{P}(k) = [\mathcal{H}'(k)\mathbf{W}(k)\mathcal{H}(k)]^{-1} \quad (4-13)$$

From (4-12) and (4-13) it is straightforward to show that

$$\mathcal{H}'(k)\mathbf{W}(k)\mathcal{Z}(k) = \mathbf{P}^{-1}(k)\hat{\theta}_{\text{WLS}}(k) \quad (4-14)$$

and

$$\mathbf{P}^{-1}(k+1) = \mathbf{P}^{-1}(k) + \mathbf{h}(k+1)w(k+1)\mathbf{h}'(k+1) \quad (4-15)$$

It now follows that

$$\begin{aligned}\hat{\theta}_{\text{WLS}}(k+1) &= \mathbf{P}(k+1)[\mathbf{h}wz + \mathbf{P}^{-1}(k)\hat{\theta}_{\text{WLS}}(k)] \\ &= \mathbf{P}(k+1)\{\mathbf{h}wz + [\mathbf{P}^{-1}(k+1) - \mathbf{h}\mathbf{h}']\hat{\theta}_{\text{WLS}}(k)\} \\ &= \hat{\theta}_{\text{WLS}}(k) + \mathbf{P}(k+1)\mathbf{h}w[z - \mathbf{h}'\hat{\theta}_{\text{WLS}}(k)] \\ &= \hat{\theta}_{\text{WLS}}(k) + \mathbf{K}_w(k+1)[z - \mathbf{h}'\hat{\theta}_{\text{WLS}}(k)]\end{aligned} \quad (4-16)$$

which is (4-8) when gain matrix \mathbf{K}_w is defined as in (4-9).

Based on preceding discussions about $\dim \theta = n$ and $\dim \mathcal{Z}(k) = N$, we know that the first value of N for which (3-10) in Lesson 3 can be used is $N = n$; thus, (4-8) must be initialized by $\hat{\theta}_{\text{WLS}}(n)$, which is computed using (3-10) in Lesson 3. Equation (4-10) is also a recursive equation for $\mathbf{P}^{-1}(k+1)$, which is initialized by $\mathbf{P}^{-1}(n) = \mathcal{H}'(n)\mathbf{W}(n)\mathcal{H}(n)$. \square

Comments

1. Equation (4-8) can also be expressed as

$$\begin{aligned}\hat{\theta}_{\text{WLS}}(k+1) &= [\mathbf{I} - \mathbf{K}_w(k+1)\mathbf{h}'(k+1)]\hat{\theta}_{\text{WLS}}(k) \\ &\quad + \mathbf{K}_w(k+1)z(k+1)\end{aligned} \quad (4-17)$$

which demonstrates that the recursive least-squares estimate (LSE) is a time-varying digital filter that is excited by random inputs (i.e., the measurements), one whose plant matrix may itself be random, because \mathbf{K}_w and $\mathbf{h}(k+1)$ may be random. The random natures of \mathbf{K}_w and $(\mathbf{I} - \mathbf{K}_w\mathbf{h}')$ make the analysis of this filter exceedingly difficult. If \mathbf{K}_w and \mathbf{h} are deterministic, then stability of this filter can be studied using Lyapunov stability theory.

2. In (4-8), the term $\mathbf{h}'(k+1)\hat{\theta}_{\text{WLS}}(k)$ is a prediction of the actual measurement $z(k+1)$. Because $\hat{\theta}_{\text{WLS}}(k)$ is based on $\mathcal{Z}(k)$, we express this predicted value as $\hat{z}(k+1|k)$, i.e.,

$$\hat{z}(k+1|k) = \mathbf{h}'(k+1)\hat{\theta}_{\text{WLS}}(k), \quad (4-18)$$

so that

$$\hat{\theta}_{\text{WLS}}(k+1) = \hat{\theta}_{\text{WLS}}(k) + \mathbf{K}_w(k+1)[z(k+1) - \hat{z}(k+1|k)]. \quad (4-19)$$

3. Two recursions are present in our recursive LSE. The first is the *vector recursion* for $\hat{\theta}_{\text{WLS}}$ given by (4-8). Clearly $\hat{\theta}_{\text{WLS}}(k+1)$ cannot be computed from this expression until measurement $z(k+1)$ is available. The second is the *matrix recursion* for \mathbf{P}^{-1} given by (4-10). Observe that values for \mathbf{P}^{-1} (and subsequently \mathbf{K}_w) can be precomputed before measurements are made.

4. A digital computer implementation of (4-8)–(4-10) proceeds as follows:

$$\mathbf{P}^{-1}(k+1) \rightarrow \mathbf{P}(k+1) \rightarrow \mathbf{K}_w(k+1) \rightarrow \hat{\theta}_{\text{WLS}}(k+1).$$

5. Equations (4-8)–(4-10) can also be used for $k = 0, 1, \dots, N-1$ using the following values for $\mathbf{P}^{-1}(0)$ and $\hat{\theta}_{\text{WLS}}(0)$:

$$\mathbf{P}^{-1}(0) = \frac{1}{a^2} \mathbf{I}_n + \mathbf{h}(0)w(0)\mathbf{h}'(0) \quad (4-20)$$

and

$$\hat{\theta}_{\text{WLS}}(0) = \mathbf{P}(0) \left[\frac{1}{a} \boldsymbol{\epsilon} + \mathbf{h}(0)w(0)z(0) \right] \quad (4-21)$$

In these equations (which are derived in Mendel, 1973, pp. 101–106; see, also, Problem 4-1) a is a very large number, $\boldsymbol{\epsilon}$ is a very small number, $\boldsymbol{\epsilon}$ is $n \times 1$, and $\boldsymbol{\epsilon} = \text{col}(\epsilon, \epsilon, \dots, \epsilon)$. When these initial values are used in (4-8)–(4-10) for $k = 0, 1, \dots, n-1$, then the resulting values obtained for $\hat{\theta}_{\text{WLS}}(n)$ and $\mathbf{P}^{-1}(n)$ are the very same ones that are obtained from the batch formulas for $\hat{\theta}_{\text{WLS}}(n)$ and $\mathbf{P}^{-1}(n)$.

Often $z(0) = 0$, or there is no measurement made at $k = 0$, so that we can set $z(0) = 0$. In this case we can set $w(0) = 0$ so that $\mathbf{P}^{-1}(0) = \mathbf{I}_n/a^2$ and $\hat{\theta}(0) = a\boldsymbol{\epsilon}$. By choosing $\boldsymbol{\epsilon}$ on the order of $1/a^2$, we see that (4-8)–(4-10) can be initialized by setting $\hat{\theta}(0) = \mathbf{0}$ and $\mathbf{P}(0)$ equal to a diagonal matrix of very large numbers.

6. The reason why the results in Theorem 4-1 are referred to as the “information form” of the recursive LSE is deferred until Lesson 11 where connections are made between least-squares and maximum-likelihood estimators (see the section entitled The Linear Model ($\mathcal{H}(k)$ deterministic), Lesson 11).

MATRIX INVERSION LEMMA

Equations (4-10) and (4-9) require the inversion of $n \times n$ matrix \mathbf{P} . If n is large than this will be a costly computation. Fortunately, an alternative is available, one that is based on the following *matrix inversion lemma*.

Lemma 4-1. *If the matrices \mathbf{A} , \mathbf{B} , \mathbf{C} , and \mathbf{D} satisfy the equation*

$$\mathbf{B}^{-1} = \mathbf{A}^{-1} + \mathbf{C}'\mathbf{D}^{-1}\mathbf{C} \quad (4-22)$$

where all matrix inverses are assumed to exist, then

$$\mathbf{B} = \mathbf{A} - \mathbf{AC}'(\mathbf{CAC}' + \mathbf{D})^{-1}\mathbf{CA}. \quad (4-23)$$

Proof. Multiply \mathbf{B} by \mathbf{B}^{-1} using (4-23) and (4-22) to show that $\mathbf{BB}^{-1} = \mathbf{I}$. For a constructive proof of this lemma see Mendel (1973), pp. 96–97. \square

Observe that if \mathbf{A} and \mathbf{B} are $n \times n$ matrices, \mathbf{C} is $m \times n$, and \mathbf{D} is $m \times m$, then to compute \mathbf{B} from (4-23) requires the inversion of one $m \times m$ matrix. On the other hand, to compute \mathbf{B} from (4-22) requires the inversion of one $m \times m$ matrix and two $n \times n$ matrices [\mathbf{A}^{-1} and $(\mathbf{B}^{-1})^{-1}$]. When $m < n$ it is definitely advantageous to compute \mathbf{B} using (4-23) instead of (4-22). Observe, also, that in the special case when $m = 1$, matrix inversion in (4-23) is replaced by division.

RECURSIVE LEAST-SQUARES: COVARIANCE FORM

Theorem 4-2. (Covariance Form of Recursive LSE). *Another recursive structure for $\hat{\theta}_{WLS}(k)$ is:*

$$\begin{aligned} \hat{\theta}_{WLS}(k+1) &= \hat{\theta}_{WLS}(k) + \mathbf{K}_w(k+1)[z(k+1) - \mathbf{h}'(k+1)\hat{\theta}_{WLS}(k)] \end{aligned} \quad (4-24)$$

where

$$\begin{aligned} \mathbf{K}_w(k+1) &= \mathbf{P}(k)\mathbf{h}(k+1)\left[\mathbf{h}'(k+1)\mathbf{P}(k)\mathbf{h}(k+1) + \frac{1}{w(k+1)}\right]^{-1} \end{aligned} \quad (4-25)$$

and

$$\mathbf{P}(k+1) = [\mathbf{I} - \mathbf{K}_w(k+1)\mathbf{h}'(k+1)]\mathbf{P}(k) \quad (4-26)$$

These equations are initialized by $\hat{\theta}_{WLS}(n)$ and $\mathbf{P}(n)$ and are used for $k = n, n+1, \dots, N-1$.

Proof. We obtain the results in (4-25) and (4-26) by applying the matrix inversion lemma to (4-10), after which our new formula for $\mathbf{P}(k+1)$ is substituted into (4-9). In order to accomplish the first part of this, let $\mathbf{A} = \mathbf{P}(k)$,

Which Form To Use

$\mathbf{B} = \mathbf{P}(k+1)$, $\mathbf{C} = \mathbf{h}'(k+1)$ and $\mathbf{D} = 1/w(k+1)$. Then (4-10) looks like (4-22), so, using (4-23) we see that

$$\begin{aligned} \mathbf{P}(k+1) &= \mathbf{P}(k) - \mathbf{P}(k)\mathbf{h}(k+1)[\mathbf{h}'(k+1)\mathbf{P}(k)\mathbf{h}(k+1) \\ &\quad + w^{-1}(k+1)]^{-1}\mathbf{h}'(k+1)\mathbf{P}(k) \end{aligned} \quad (4-27)$$

Consequently,

$$\begin{aligned} \mathbf{K}_w(k+1) &= \mathbf{P}(k+1)\mathbf{h}(k+1)w(k+1) \\ &= [\mathbf{P} - \mathbf{Ph}(\mathbf{h}'\mathbf{Ph} + w^{-1})^{-1}\mathbf{h}'\mathbf{P}]\mathbf{h}w \\ &= \mathbf{Ph}[\mathbf{I} - (\mathbf{h}'\mathbf{Ph} + w^{-1})^{-1}\mathbf{h}'\mathbf{Ph}]\mathbf{h}w \\ &= \mathbf{Ph}(\mathbf{h}'\mathbf{Ph} + w^{-1})^{-1}(\mathbf{h}'\mathbf{Ph} + w^{-1} - \mathbf{h}'\mathbf{Ph})\mathbf{h}w \\ &= \mathbf{Ph}(\mathbf{h}'\mathbf{Ph} + w^{-1})^{-1} \end{aligned}$$

which is (4-25). In order to obtain (4-26), express (4-27) as

$$\begin{aligned} \mathbf{P}(k+1) &= \mathbf{P}(k) - \mathbf{K}_w(k+1)\mathbf{h}'(k+1)\mathbf{P}(k) \\ &= [\mathbf{I} - \mathbf{K}_w(k+1)\mathbf{h}'(k+1)]\mathbf{P}(k) \end{aligned} \quad \square$$

Comments

1. The recursive formula for $\hat{\theta}_{WLS}$, (4-24), is unchanged from (4-8). Only the matrix recursion for \mathbf{P} , leading to gain matrix \mathbf{K}_w has changed. A digital computer implementation of (4-24)–(4-26) proceeds as follows: $\mathbf{P}(k) \rightarrow \mathbf{K}_w(k+1) \rightarrow \hat{\theta}_{WLS}(k+1) \rightarrow \mathbf{P}(k+1)$. This order of computations differs from the preceding one.
2. When $z(k)$ is a scalar then the covariance form of the recursive LSE requires no matrix inversions, and only one division.
3. Equations (4-24)–(4-26) can also be used for $k = 0, 1, \dots, N-1$ using the values for $\mathbf{P}^{-1}(0)$ and $\hat{\theta}_{WLS}(0)$ given in (4-20) and (4-21).
4. The reason why the results in Theorem 4-2 are referred to as the “covariance form” of the recursive LSE is deferred to Lesson 9 where connections are made between least-squares and best linear unbiased minimum-variance estimators (see p. 79).

WHICH FORM TO USE

We have derived two formulations for a recursive least-squares estimator, the information and covariance forms. In on-line applications, where speed of computation is often the most important consideration, the covariance form is preferable to the information form. This is because a smaller matrix needs to be inverted in the covariance form, namely an $m \times m$ matrix rather than an $m \times n$ matrix (m is often much smaller than n).

The information form is often more useful than the covariance form in analytical studies. For example, it is used to derive the initial conditions for $\mathbf{P}^{-1}(0)$ and $\hat{\theta}_{\text{WLS}}(0)$, which are given in (4-20) and (4-21) (see Mendel, 1973, pp. 101–106). The information form is also to be preferred over the covariance form during the startup of recursive least squares. We demonstrate why this is so next.

We consider the case when

$$\mathbf{P}(0) = a^2 \mathbf{I}_n \quad (4-28)$$

where a^2 is a very, very large number. Using the information form, we find that, for $k = 0$, $\mathbf{P}^{-1}(1) = \mathbf{h}(1)\mathbf{w}(1)\mathbf{h}'(1) + 1/a^2 \mathbf{I}_n$, and, therefore, $\mathbf{K}_w(1) = [\mathbf{h}(1)\mathbf{w}(1)\mathbf{h}'(1) + 1/a^2 \mathbf{I}_n]^{-1} \mathbf{h}(1)\mathbf{w}(1)$. No difficulties are encountered when we compute $\mathbf{K}_w(1)$ using the information form.

Using the covariance form we find, first, that $\mathbf{K}_w(1) \approx a^2 \mathbf{h}(1)[\mathbf{h}'(1)a^2 \mathbf{h}(1)]^{-1} = \mathbf{h}(1)[\mathbf{h}'(1)\mathbf{h}(1)]^{-1}$, and then that

$$\mathbf{P}(1) = \{\mathbf{I} - \mathbf{h}(1)[\mathbf{h}'(1)\mathbf{h}(1)]^{-1} \mathbf{h}'(1)\}a^2 \quad (4-29)$$

however, this matrix is singular. To see this, postmultiply both sides of (4-29) by $\mathbf{h}(1)$, to obtain

$$\mathbf{P}(1)\mathbf{h}(1) = \{\mathbf{h}(1) - \mathbf{h}(1)[\mathbf{h}'(1)\mathbf{h}(1)]^{-1} \mathbf{h}'(1)\mathbf{h}(1)\}a^2 = 0 \quad (4-30)$$

Neither $\mathbf{P}(1)$ nor $\mathbf{h}(1)$ equal zero; hence, $\mathbf{P}(1)$ must be a singular matrix for $\mathbf{P}(1)\mathbf{h}(1)$ to equal zero. In fact, once $\mathbf{P}(1)$ becomes singular, all other $\mathbf{P}(j)$, $j \geq 2$, will be singular.

In Lesson 9 we shall show that when $\mathbf{W}^{-1}(k) = \mathbf{E}\{\mathbf{V}(k)\mathbf{V}'(k)\} = \mathcal{R}(k)$, then $\mathbf{P}(k)$ is the covariance matrix of the estimation error, $\hat{\theta}(k)$. This matrix must be positive definite, and it will be quite difficult to maintain this property if $\mathbf{P}(k)$ is singular; hence, it is advisable to initialize the recursive least-squares estimator using the information form. However, it is also advisable to switch to the covariance formulation as soon after initialization as possible, in order to reduce computing time.

PROBLEMS

- 4-1.** In order to derive the formulas for $\mathbf{P}^{-1}(0)$ and $\hat{\theta}_{\text{WLS}}(0)$, given in (4-20) and (4-21), respectively, one proceeds as follows. Introduce n artificial measurements $z^a(-1), z^a(-2), \dots, z^a(-n)$, where $z^a(-j) \triangleq \epsilon$ in which ϵ is a very small number. Then, assume that the model for $z^a(-j)$ is $z^a(-j) = \frac{1}{a} \theta_j$ ($j = 1, 2, \dots, n$) where a is a very large number.

(a) Show that $\hat{\theta}_{\text{LS}}(-1) = a\epsilon$, where $n \times 1$ vector $\epsilon = \text{col}(\epsilon, \epsilon, \dots, \epsilon)$.

Additionally, show that $\mathbf{P}^a(-1) = a^2 \mathbf{I}_n$.

(b) Show that $\hat{\theta}_{\text{WLS}}(0) = \mathbf{P}^a(0)[\epsilon/a + \mathbf{h}(0)\mathbf{w}(0)z(0)]$ and $\mathbf{P}^a(0) = [\mathbf{I}_n/a^2 + \mathbf{h}(0)\mathbf{w}(0)\mathbf{h}'(0)]^{-1}$.

Lesson 4 Problems

- (c) Show that when the measurements $z^a(-n), \dots, z^a(-1), z(0), z(1), \dots, z(l+1)$ are used, then

$$\hat{\theta}_{\text{WLS}}^a(l+1) = \mathbf{P}^a(l+1) \left[\epsilon/a + \sum_{j=0}^{l+1} \mathbf{h}(j)\mathbf{w}(j)z(j) \right]$$

and

$$\mathbf{P}^a(l+1) = \left[\mathbf{I}_n/a^2 + \sum_{j=0}^{l+1} \mathbf{h}(j)\mathbf{w}(j)\mathbf{h}'(j) \right]^{-1}$$

- (d) Show that when the measurements $z(0), z(1), \dots, z(l+1)$ are used (i.e., the artificial measurements are not used), then

$$\hat{\theta}_{\text{WLS}}(l+1) = \mathbf{P}(l+1) \left[\sum_{j=0}^{l+1} \mathbf{h}(j)\mathbf{w}(j)z(j) \right]$$

and

$$\mathbf{P}(l+1) = \left[\sum_{j=0}^{l+1} \mathbf{h}(j)\mathbf{w}(j)\mathbf{h}'(j) \right]^{-1}$$

- (e) Finally, show that for $a \ggg$ and $\epsilon \lll$

$$\mathbf{P}^a(l+1) \rightarrow \mathbf{P}(l+1)$$

and

$$\hat{\theta}_{\text{WLS}}^a(l+1) \rightarrow \hat{\theta}_{\text{WLS}}(l+1)$$

- 4-2.** Prove that once $\mathbf{P}(1)$ becomes singular, all other $\mathbf{P}(j)$, $j \geq 2$, will be singular.

- 4-3.** (Mendel, 1973, Exercise 2-12, pg. 138). The following weighting matrix weights past measurements less heavily than the most recent measurements,

$$\mathbf{W}(k+1) = \text{diag}(w(k+1), w(k), \dots, w(1)) \triangleq \begin{pmatrix} \bar{w}(k+1) & \mathbf{0}' \\ 0 & \beta^{k+1-k} \mathbf{W}(k) \end{pmatrix}$$

where

$$0 < \beta < 1$$

- (a) Show that $w(j) = \bar{w}(j)\beta^{k+1-j}$ for $j = 1, \dots, k+1$.

- (b) How must the equations for the recursive weighted least-squares estimator be modified for this weighting matrix? The estimator thus obtained is known as a *fading memory estimator* (Morrison, 1969).

- 4-4.** We showed, in Lesson 3, that it doesn't matter how one chooses the weights, $w(j)$, in the method of least squares, because the weights cancel out in the batch formula for $\hat{\theta}_{\text{LS}}(k)$. On the other hand, $w(k+1)$ appears explicitly in the recursive WLSE, but only in the formula for $\mathbf{K}_w(k+1)$, i.e.,

$$\mathbf{K}_w(k+1) = \mathbf{P}(k)\mathbf{h}(k+1)[\mathbf{h}'(k+1)\mathbf{P}(k)\mathbf{h}(k+1) + w^{-1}(k+1)]^{-1}$$

It would appear that if we set $w(k+1) = w_1$, for all k , we would obtain a $\hat{\theta}_{\text{LS}}(k+1)$ value that would be different from that obtained by setting $w(k+1) = w_2$, for all k . Of course, this cannot be true. Show, using the

formulas for the recursive WLSE, how they can be made independent of $w(k+1)$, when $w(k+1) = w$, for all k .

- 4-5. For the data in the accompanying table, do the following:

- (a) Obtain the least-squares line $\hat{y}(t) = a + bt$ by means of the batch processing least-squares algorithm;
- (b) Obtain the least-squares line by means of the recursive least-squares algorithm, using the recursive startup technique (let $a = 10^8$ and $\epsilon = 10^{-16}$).

t	$y(t)$
0	1
1	5
2	9
3	11

Lesson 5

Least-Squares Estimation: Recursive Processing (continued)

Example 5-1

In order to illustrate some of the Lesson 4 results we shall obtain a recursive algorithm for the least-squares estimator of the scalar θ in the instrument calibration example (Lesson 3). Gain $K_w(k+1)$ is computed using (4-9) of Lesson 4, and $P(k+1)$ is computed using (4-13) of Lesson 4. Generally, we do not compute $P(k+1)$ using (4-13); but, the simplicity of our example allows us to use this formula to obtain a closed-form expression for $P(k+1)$ in the most direct way. Recall that $\mathcal{H} = \text{col}(1, 1, \dots, 1)$, which is a $k \times 1$ vector, and $h(k+1) = 1$; thus, setting $W(k) = I$ and $w(k+1) = 1$ (in order to obtain the recursive LSE of θ) in the preceding formulas, we find that

$$P(k+1) = [\mathcal{H}'(k+1)\mathcal{H}(k+1)]^{-1} = \frac{1}{k+1} \quad (5-1)$$

and

$$K_w(k+1) = P(k+1) = \frac{1}{k+1} \quad (5-2)$$

Substituting these results into (4-8) of Lesson 4, we then find

$$\hat{\theta}_{LS}(k+1) = \hat{\theta}_{LS}(k) + \frac{1}{k+1}[z(k+1) - \hat{\theta}_{LS}(k)]$$

or,

$$\hat{\theta}_{LS}(k+1) = \left(\frac{k}{k+1}\right)\hat{\theta}_{LS}(k) + \frac{1}{k+1}z(k+1) \quad (5-3)$$

Formula (5-3), which can be used for $k = 0, 1, \dots, N - 1$ by setting $\hat{\theta}_{LS}(0) = 0$, lets us reinterpret the well-known sample mean estimator as a time-varying digital filter [see, also, (1-2) of Lesson 1]. We leave it to the reader to study the stability properties of this first-order filter.

Usually it is in only the simplest of cases that we can obtain closed-form expressions for \mathbf{K}_w and \mathbf{P} and subsequently $\hat{\theta}_{LS}$ [or $\hat{\theta}_{WLS}$]; however, we can always obtain values for $\mathbf{K}_w(k + 1)$ and $\mathbf{P}(k + 1)$ at successive time points using the results in Theorems 4-1 or 4-2. \square

GENERALIZATION TO VECTOR MEASUREMENTS

A vector of measurements can occur in any application where it is possible to use more than one sensor; however, it is also possible to obtain a vector of measurements from certain types of individual sensors. In spacecraft applications, it is not unusual to be able to measure attitude, rate, and acceleration. In electrical systems applications, it is not uncommon to be able to measure voltages, currents, and power. Radar measurements often provide information about range, azimuth, and elevation. Radar is an example of a single sensor that provides a vector of measurements.

In the vector measurement case, (4-1) of Lesson 4 is changed from $\mathbf{z}(k + 1) = \mathbf{h}'(k + 1)\boldsymbol{\theta} + \mathbf{v}(k + 1)$ to

$$\mathbf{z}(k + 1) = \mathbf{H}(k + 1)\boldsymbol{\theta} + \mathbf{v}(k + 1) \quad (5-4)$$

where \mathbf{z} is now an $m \times 1$ vector, \mathbf{H} is $m \times n$ and \mathbf{v} is $m \times 1$.

We leave it to the reader to show that all of the results in Lessons 3 and 4 are unchanged in the vector measurement case; but some notation must be altered (see Table 5-1).

TABLE 5-1 Transformations from Scalar to Vector Measurement Situations, and Vice-Versa

Scalar Measurement	Vector of Measurements
$z(k + 1)$	$\mathbf{z}(k + 1)$, an $m \times 1$ vector
$v(k + 1)$	$\mathbf{v}(k + 1)$, an $m \times 1$ vector
$w(k + 1)$	$\mathbf{w}(k + 1)$, an $m \times m$ matrix
$\mathbf{h}'(k + 1)$, a $1 \times n$ matrix	$\mathbf{H}(k + 1)$, an $m \times n$ matrix
$\mathbf{Z}(k)$, an $N \times 1$ vector	$\mathbf{Z}(k)$, an $Nm \times 1$ vector
$\mathbf{V}(k)$, an $N \times 1$ vector	$\mathbf{V}(k)$, an $Nm \times 1$ vector
$\mathbf{W}(k)$, an $N \times N$ matrix	$\mathbf{W}(k)$, an $Nm \times Nm$ matrix
$\mathbf{H}(k)$, an $N \times n$ matrix	$\mathbf{H}(k)$, an $Nm \times n$ matrix

Source: Reprinted from Mendel, 1973, p. 110. Courtesy of Marcel Dekker, Inc., NY.

CROSS-SECTIONAL PROCESSING

Suppose that at each sampling time t_{k+1} there are q sensors or groups of sensors that provide our vector measurement data. These sensors are corrupted by noise that is uncorrelated from one sensor group to another. The m -dimensional vector $\mathbf{z}(k + 1)$ can be represented as

$$\mathbf{z}(k + 1) = \text{col}(\mathbf{z}_1(k + 1), \mathbf{z}_2(k + 1), \dots, \mathbf{z}_q(k + 1)) \quad (5-5)$$

where

$$\mathbf{z}_i(k + 1) = \mathbf{H}_i(k + 1)\boldsymbol{\theta} + \mathbf{v}_i(k + 1) \quad (5-6)$$

$$\dim \mathbf{z}_i(k + 1) = m_i \times 1,$$

$$\sum_{i=1}^q m_i = m \quad (5-7)$$

$$\mathbf{E}\{\mathbf{v}_i(k + 1)\} = \mathbf{0}, \text{ and}$$

$$\mathbf{E}\{\mathbf{v}_i(k + 1)\mathbf{v}_j'(k + 1)\} = \mathbf{R}_i(k + 1)\delta_{ij} \quad (5-8)$$

An alternative to processing all m measurements in one batch (i.e., simultaneously) is available, and is one in which we freeze time at t_{k+1} and recursively process the q batches of measurements one batch at a time. Data $\mathbf{z}_i(k + 1)$ are used to obtain an estimate (for notational simplicity, in this section we omit the subscript WLS or LS on $\boldsymbol{\theta}$) $\hat{\theta}_i(k + 1)$ with $\hat{\theta}_1(k) \triangleq \hat{\theta}(k)$ and $\mathbf{z}(k + 1) \triangleq \mathbf{z}_1(k + 1)$. When these calculations are completed $\mathbf{z}_2(k + 1)$ is processed to obtain the estimate $\hat{\theta}_2(k + 1)$. Estimate $\hat{\theta}_1(k + 1)$ is used to initialize $\hat{\theta}_2(k + 1)$. Each set of data is processed in this manner until the final set $\mathbf{z}_q(k + 1)$ has been included. Then time is advanced to t_{k+2} and the cycle is repeated. This type of processing is known as cross-sectional or sequential processing. It is summarized in Figure 5-1 and is contrasted with more usual simultaneous recursive processing in Figure 5-2.

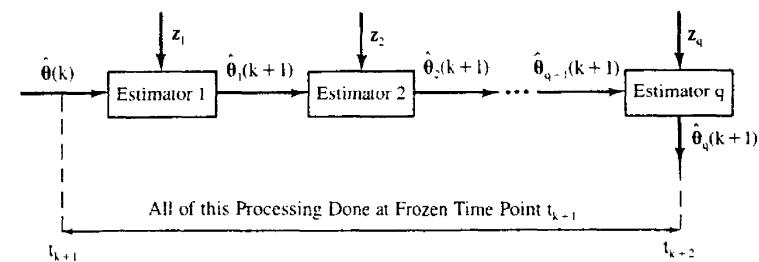


Figure 5-1 Cross-sectional processing of m measurements.

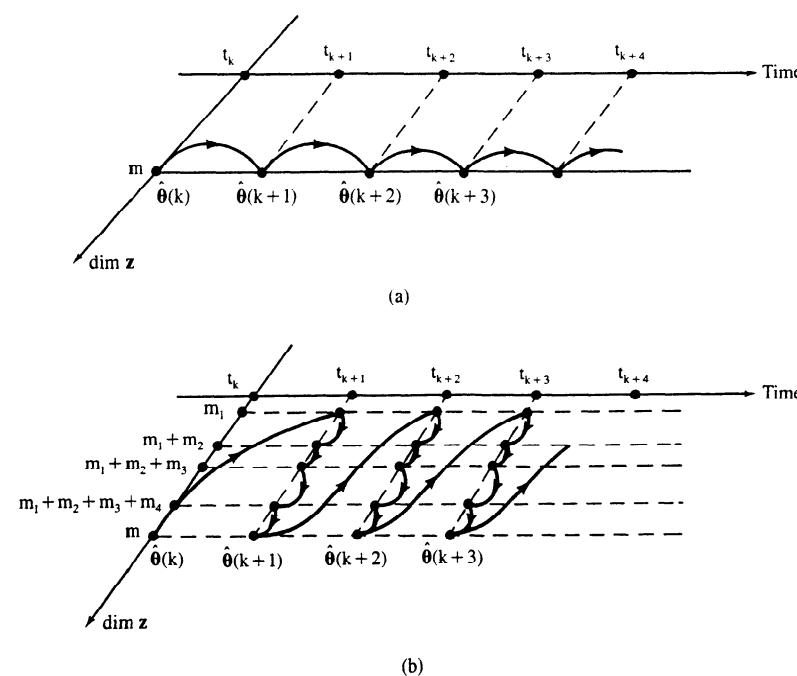


Figure 5-2 Two ways to reach $\hat{\theta}(k+1), \hat{\theta}(k+2), \dots$ (a) Simultaneous recursive processing performed along the line $\text{dim } z = m$, and (b) cross-sectional recursive processing where, for example, at t_{k+1} the processing is performed along the line $\text{TIME} = t_{k+1}$ and stops when that line intersects the line $\text{dim } z = m$.

The remarkable property about cross-sectional processing is that

$$\hat{\theta}_q(k+1) = \hat{\theta}(k+1) \quad (5-9)$$

where $\hat{\theta}(k+1)$ is obtained from simultaneous recursive processing.

A very large computational advantage exists for cross-sectional processing if $m_i = 1$. In this case the matrix inverse $[\mathbf{H}(k+1)\mathbf{P}(k)\mathbf{H}'(k+1) + \mathbf{w}^{-1}(k+1)]^{-1}$ needed in (4-25) of Lesson 4 is replaced by the division $[\mathbf{h}_i'(k+1)\mathbf{P}_i(k)\mathbf{h}_i(k+1) + 1/w_i(k+1)]^{-1}$. See Mendel, 1973, pp. 113–118 for a proof of (5-9).

MULTISTAGE LEAST-SQUARES ESTIMATORS

Suppose we are given a linear model $\mathcal{Z}(k) = \mathcal{H}_1(k)\theta_1 + \mathcal{V}(k)$ with n unknown parameters θ_1 , datum $\{\mathcal{Z}(k), \mathcal{H}_1(k)\}$, and the LSE of θ_1 , $\hat{\theta}_{1,LS}^*(k)$, where

$$\hat{\theta}_{1,LS}^*(k) = [\mathcal{H}_1'(k)\mathcal{H}_1(k)]^{-1}\mathcal{H}_1(k)\mathcal{Z}(k) \quad (5-10)$$

We extend this model to include l additional parameters, θ_2 , so that our model is given by

$$\mathcal{Z}(k) = \mathcal{H}_1(k)\theta_1 + \mathcal{H}_2(k)\theta_2 + \mathcal{V}(k) \quad (5-11)$$

For this model, datum $\{\mathcal{Z}(k), \mathcal{H}_1(k), \mathcal{H}_2(k)\}$ is available. We wish to compute the least-squares estimates of θ_1 and θ_2 for the $n + l$ parameter model using the previously computed $\hat{\theta}_{1,LS}^*(k)$.

Theorem 5-1. Given the linear model in (5-11), where θ_1 is $n \times 1$ and θ_2 is $l \times 1$. The LSEs of θ_1 and θ_2 based on datum $\{\mathcal{Z}(k), \mathcal{H}_1(k), \mathcal{H}_2(k)\}$ are found from the following equations:

$$\hat{\theta}_{1,LS}(k) = \hat{\theta}_{1,LS}^*(k) - \mathbf{G}(k)\mathbf{C}(k)\mathcal{H}_2'(k)[\mathcal{Z}(k) - \mathcal{H}_1(k)\hat{\theta}_{1,LS}^*(k)] \quad (5-12)$$

and

$$\hat{\theta}_{2,LS}(k) = \mathbf{C}(k)\mathcal{H}_2'(k)[\mathcal{Z}(k) - \mathcal{H}_1(k)\hat{\theta}_{1,LS}^*(k)] \quad (5-13)$$

where

$$\mathbf{G}(k) = [\mathcal{H}_1'(k)\mathcal{H}_1(k)]^{-1}\mathcal{H}_1'(k)\mathcal{H}_2(k) \quad (5-14)$$

and

$$\mathbf{C}(k) = [\mathcal{H}_2'(k)\mathcal{H}_2(k) - \mathcal{H}_2'(k)\mathcal{H}_1(k)\mathbf{G}(k)]^{-1} \quad (5-15)$$

The results in this theorem were worked out by Åström (1968) and emphasize operations which are performed on the vector of residuals for the n parameter model, namely $\mathcal{Z}(k) - \mathcal{H}_1(k)\hat{\theta}_{1,LS}^*(k)$. Other forms for $\hat{\theta}_{1,LS}(k)$ and $\hat{\theta}_{2,LS}(k)$ appear in Mendel (1975).

Proof. The derivation of (5-12) and (5-13) is based primarily on the block decompositon method for inverting $\mathcal{H}'\mathcal{H}$, where $\mathcal{H} = (\mathcal{H}_1|\mathcal{H}_2)$. See Mendel (1975) for the details. \square

Similar results to those in Theorem 5-1 can be developed for the removal of l parameters from a model, for adding or removing parameters one at a time, and for recursive-in-time versions of all these results (see Mendel, 1975). All of these results are referred to as “multistage LSEs.”

We conclude this section and lesson with some examples which illustrate problems for which multistage algorithms can be quite useful.

Example 5-2 Identification of a Sampled Impulse Response: Zoom-In Algorithm

A test signal, $u(t)$ is applied at $t = t_0$ to a linear, time-invariant, causal, but unknown system whose output, $y(t)$, and input are measured. The unknown impulse response, $w(t)$, is to be identified using sampled values of $u(t)$ and $y(t)$. For such a system

$$y(t_k) = \int_{t_0}^{t_k} w(\tau)u(t_k - \tau)d\tau \quad (5-16)$$

One approach to identifying $w(t)$ is to discretize (5-16) and to only identify $w(t)$ at discrete values of time. If we assume that (1) $w(t) \approx 0$ for all $t \geq t_w$, (2) $[t_0, t_w]$ is divided into n equal intervals, each of width T , so that $n = (t_w - t_0)/T$ and (3) for $\tau \in [t_{i-1}, t_i]$, $w(\tau) \equiv w(t_{i-1})$ and $u(t-\tau) \equiv u(t-t_{i-1})$, then

$$y(t_k) \approx \sum_{i=1}^n w_1(t_{i-1})u(t_k - t_{i-1}) \quad (5-17)$$

where

$$w_1(t_{i-1}) = Tw(t_{i-1}) \quad (5-18)$$

It is straightforward to identify the n unknown parameters $w_1(t_0)$, $w_1(t_1)$, ..., $w_1(t_{n-1})$ via least squares (see Example 2-1 of Lesson 2); however, for n to be known, t_w must be accurately known and T must be chosen most judiciously. In actual practice t_w is not known that accurately so that n may have to be varied. Multistage LSEs can be used to handle this situation. Sometimes T can be too coarse for certain regions of time, in which case significant features of $w(t)$ such as a ripple may be obscured. In this situation, we would like to "zoom-in" on those intervals of time and rediscretize $y(t)$ just over those intervals, thereby adding more terms to (5-17). Multistage LSEs can also be used to handle this situation, as we demonstrate next.

For illustrative purposes, we present this procedure for the case when the interval of interest equals T , i.e., when $t \in [t_x, t_{x+1}] = T$ which is further divided into q equal intervals, each of width ΔT_x , so that $q = (t_{x+1} - t_x)/\Delta T_x = T/\Delta T_x$. Observe that

$$\begin{aligned} \int_{t_x}^{t_{x+1}} w(\tau)u(t_k - \tau)d\tau &= \sum_{j=0}^{q-1} \int_{t_x + j\Delta T_x}^{t_x + (j+1)\Delta T_x} w(\tau)u(t_k - \tau)d\tau \\ &= \sum_{j=0}^{q-1} w(t_x + j\Delta T_x)u(t_k - t_x - j\Delta T_x)\Delta T_x \end{aligned} \quad (5-19)$$

hence,

$$\begin{aligned} y(t_k) &= \sum_{\substack{i=1 \\ i \neq x}}^n w_1(t_{i-1})u(t_k - t_{i-1}) + \frac{\Delta T_x}{T} w_1(t_x)u(t_k - t_x) \\ &\quad + \sum_{j=1}^{q-1} w_2(t_x + j\Delta T_x)u(t_k - t_x - j\Delta T_x) \end{aligned} \quad (5-20)$$

where

$$w_2(t_p) = w(t_p)\Delta T_x \quad (5-21)$$

Equation (5-20) contains $n + q - 1$ parameters.

Let

$$\theta_1 = \text{col}(w_1(t_0), w_1(t_1), \dots, w_1(t_x), \dots, w_1(t_{n-1})) \quad (5-22)$$

and assume that a least-squares estimate of θ_1 , $\hat{\theta}_{1,LS}^*$, which is based on (5-17), is available. Let

$$\theta_2 = \text{col}[w_2(t_x + \Delta T_x), w_2(t_x + 2\Delta T_x), \dots, w_2(t_x + (q-1)\Delta T_x)] \quad (5-23)$$

To obtain $\hat{\theta}_{LS} = \text{col}(\hat{\theta}_{1,LS}, \hat{\theta}_{2,LS})$ proceed as follows: 1) modify $\hat{\theta}_{1,LS}^*$ by scaling $\hat{w}_{1,LS}(t_x)$ to $\hat{w}_{1,LS}(t_x)\Delta T_x/T$ and call the modified result $\hat{\theta}_{1,LS}^*$; and 2) apply Theorem 5-1 to obtain

TABLE 5-2 Impulse Response Estimates

Number of Parameters in the Model <i>n</i>	Average Estimates of Parameters ^{a,b}								Average Performance
	θ_1 (0.39)	θ_2 (0.76)	θ_3 (0.38)	θ_4 (-0.23)	θ_5 (-0.48)	θ_6 (-0.25)	θ_7 (0.14)	θ_8 (0.30)	
1	0.34								0.96
2	0.41	0.75							0.75
3	0.39	0.78	0.40						0.63
4	0.39	0.79	0.40	-0.12					0.62
5	0.40	0.79	0.43	-0.17	0.41				0.44
6	0.39	0.79	0.43	-0.15	-0.44	-0.29			0.40
7	0.39	0.80	0.44	-0.16	-0.44	-0.27	0.06		0.40
8	0.41	0.81	0.43	-0.15	-0.47	-0.30	0.12	0.30	0.29
9	0.43	0.83	0.43	-0.14	-0.47	-0.33	0.10	0.35	0.18
10	0.43	0.81	0.42	-0.15	-0.46	-0.32	0.10	0.35	0.13

Source: Reprinted from Mendel, 1975, pg. 782, © 1975 IEEE.

^a Values for parameters shown in parentheses are true values.

^b Blank spaces in the table denote "no value" for parameters.

$\hat{\theta}_{LS}$. Note that the scaling of $\hat{w}_{1,LS}(t_x)$ in Step 1 is due to the appearance of $w_1(t_x)\Delta T_x / T$ instead of $w_1(t_x)$ in (5-20).

It is straightforward to extend the approach of this example to regions that include more than one sampling interval, T . \square

Example 5-3 Impulse Response Identification

An n -stage batch LSE was used to estimate impulse response models having from one to ten parameters [i.e., n in (5-17) ranged from 1–10] for a second-order system (natural frequency of 10.47 rad/sec, damping ratio of 0.12, and, unity gain). The system was forced with an input of randomly chosen ± 1 's each of which was equally likely to occur. The system's output was corrupted by zero-mean pseudo-random Gaussian white noise of unity variance. Fifty consecutive samples of the noisy output and noise-free input were then processed by the n -stage batch LSE, and this procedure was repeated ten times from which the average values for the parameter estimates given in Table 5-2 (page 41) were obtained.

The ten models were tested using ten input sequences, each containing 50 samples of randomly chosen ± 1 's. The last column in Table 5-2 gives values for the normalized average performance

$$\left[\sum_{i=1}^{50} [\bar{y}_{\text{model}}(t_i) - \bar{y}_{\text{actual}}(t_i)]^2 / \sum_{i=1}^{50} \bar{y}_{\text{actual}}^2(t_i) \right]^{1/2}$$

in which $\bar{y}(t_i)$ denotes an average over the ten runs. Not too surprisingly, we see that average predicted performance improves as n increases.

All the θ_i results were obtained in one pass through the n -stage LSE using approximately 3150 flops. The same results could have been obtained using 10 LSEs (Lesson 3); but, that would have required approximately 5220 flops. \square

PROBLEMS

- 5-1. Show that, in the vector measurement case, the results given in Lessons 3 and 4 only need to be modified using the transformations listed in Table 5-1.
- 5-2. Prove that, using cross-sectional processing, $\hat{\theta}_q(k+1) = \hat{\theta}(k+1)$.
- 5-3. Prove the multistage least-squares estimation Theorem 5-1.
- 5-4. Extend Example 5-2 to regions of interest equal to mT , where m is a positive integer and T is the original data sampling time.

Lesson 6

Small Sample Properties of Estimators

INTRODUCTION

How do we know whether or not the results obtained from the LSE, or for that matter any estimator, are good? To answer this question, we make use of the fact that all estimators represent transformations of random data. For example, our LSE, $[\mathcal{H}'(k)\mathcal{W}(k)\mathcal{H}(k)]^{-1}\mathcal{H}'(k)\mathcal{W}(k)\mathcal{E}(k)$, represents a linear transformation on $\mathcal{E}(k)$. Other estimators may represent nonlinear transformations of $\mathcal{E}(k)$. The consequence of this is that $\hat{\theta}(k)$ is itself random. Its properties must therefore be studied from a statistical viewpoint.

In the estimation literature, it is common to distinguish between small-sample and large-sample properties of estimators. The term "sample" refers to the number of measurements used to obtain $\hat{\theta}$, i.e., the dimension of \mathcal{E} . The phrase "small-sample" means *any number of measurements* (e.g., 1, 2, 100, 10^4 , or even an infinite number), whereas the phrase "large-sample" means *an infinite number of measurements*. Large-sample properties are also referred to as asymptotic properties. It should be obvious that *if an estimator possesses a small-sample property it also possesses the associated large-sample property; but, the converse is not always true*.

Why bother studying large-sample properties of estimators if these properties are included in their small-sample properties? Put another way, why not just study small-sample properties of estimators? For many estimators it is relatively easy to study their large-sample properties and virtually impossible to learn about their small-sample properties. An analogous situation occurs in stability theory, where most effort is directed at infinite-time stability behavior rather than at finite-time behavior.

Although “large-sample” means an infinite number of measurements, estimators begin to enjoy their large-sample properties for much fewer than an infinite number of measurements. How few, usually depends on the dimension of θ , n .

A thorough study into $\hat{\theta}$ would mean determining its probability density function $p(\hat{\theta})$. Usually, it is too difficult to obtain $p(\hat{\theta})$ for most estimators (unless $\hat{\theta}$ is multivariate Gaussian); thus, it is customary to emphasize the first- and second-order statistics of $\hat{\theta}$ (or, its associated error $\bar{\theta} = \theta - \hat{\theta}$), namely the mean and covariance.

We shall examine the following small- and large-sample properties of estimators: unbiasedness and efficiency (small-sample), and asymptotic unbiasedness, consistency, and asymptotic efficiency (large-sample). Small-sample properties are the subject of this lesson, whereas large-sample properties are studied in Lesson 7.

UNBIASEDNESS

Definition 6-1. Estimator $\hat{\theta}(k)$ is an unbiased estimator of deterministic θ , if

$$E\{\hat{\theta}(k)\} = \theta \quad \text{for all } k \quad (6-1)$$

or of random θ , if

$$E\{\hat{\theta}(k)\} = E\{\theta\} \quad \text{for all } k \quad \square \quad (6-2)$$

In terms of estimation error, $\bar{\theta}(k)$, unbiasedness means, that

$$E\{\bar{\theta}(k)\} = 0 \quad \text{for all } k \quad (6-3)$$

Example 6-1

In the instrument calibration example of Lesson 3, we determined the following LSE of θ :

$$\hat{\theta}_{LS}(N) = \frac{1}{N} \sum_{i=1}^N z(i) \quad (6-4)$$

where

$$z(i) = \theta + v(i) \quad (6-5)$$

Suppose $E\{v(i)\} = 0$ for $i = 1, 2, \dots, N$; then,

$$E\{\hat{\theta}_{LS}(N)\} = \frac{1}{N} \sum_{i=1}^N E\{z(i)\} = \frac{1}{N} \sum_{i=1}^N \theta = \theta$$

which means that $\hat{\theta}_{LS}(N)$ is an unbiased estimator of θ . \square

Unbiasedness

Many estimators are linear transformations of the measurements, i.e.,

$$\hat{\theta}(k) = F(k)\mathcal{Z}(k) \quad (6-6)$$

In least-squares, we obtained this linear structure for $\hat{\theta}(k)$ by solving an optimization problem. Sometimes, we begin by assuming that (6-6) is the desired structure for $\hat{\theta}(k)$. We now address the question “when is $F(k)\mathcal{Z}(k)$ an unbiased estimator of deterministic θ ?”

Theorem 6-1. When $\mathcal{Z}(k) = \mathcal{H}(k)\theta + \mathcal{V}(k)$, $E\{\mathcal{V}(k)\} = 0$, and $\mathcal{H}(k)$ is deterministic, then $\hat{\theta}(k) = F(k)\mathcal{Z}(k)$ [where $F(k)$ is deterministic] is an unbiased estimator of θ if and only if

$$F(k)\mathcal{H}(k) = I \quad \text{for all } k \quad (6-7)$$

Note that this is the first place where we have had to assume any a priori knowledge about noise $\mathcal{V}(k)$.

Proof

a. (Necessity). From the model for $\mathcal{Z}(k)$ and the assumed structure for $\hat{\theta}(k)$, we see that

$$\hat{\theta}(k) = F(k)\mathcal{H}(k)\theta + F(k)\mathcal{V}(k) \quad (6-8)$$

If $\hat{\theta}(k)$ is an unbiased estimator of θ , $F(k)$ and $\mathcal{H}(k)$ are deterministic, and $E\{\mathcal{V}(k)\} = 0$, then

$$E\{\hat{\theta}(k)\} = \theta = F(k)\mathcal{H}(k)\theta$$

or

$$[I - F(k)\mathcal{H}(k)]\theta = 0 \quad (6-9)$$

Obviously, for $\theta \neq 0$, (6-7) is the solution to this equation.

b. (Sufficiency). From (6-8) and the nonrandomness of $F(k)$ and $\mathcal{H}(k)$, we have

$$E\{\hat{\theta}(k)\} = F(k)\mathcal{H}(k)\theta \quad (6-10)$$

Assuming the truth of (6-7), it must be that

$$E\{\hat{\theta}(k)\} = \theta$$

which, of course, means that $\hat{\theta}(k)$ is an unbiased estimator of θ . \square

Example 6-2

Matrix $F(k)$ for the WLSE of θ is $[\mathcal{H}'(k)W(k)\mathcal{H}(k)]^{-1}\mathcal{H}'(k)W(k)$. Observe that this $F(k)$ matrix satisfies (6-7); thus, when $\mathcal{H}(k)$ is deterministic the WLSE of θ is unbiased. Unfortunately, in many interesting applications $\mathcal{H}(k)$ is random, and we cannot apply Theorem 6-1 to study the unbiasedness of the WLSE. We return to this issue in Lesson 8. \square

Suppose that we begin by assuming a linear recursive structure for $\hat{\theta}$, namely

$$\hat{\theta}(k+1) = \mathbf{A}(k+1)\hat{\theta}(k) + \mathbf{b}(k+1)z(k+1) \quad (6-11)$$

We then have the following counterpart to Theorem 6-1.

Theorem 6-2. When $z(k+1) = \mathbf{h}'(k+1)\theta + v(k+1)$, $E\{v(k+1)\} = 0$, and $\mathbf{h}(k+1)$ is deterministic, then $\hat{\theta}(k+1)$ given by (6-11) is an unbiased estimator of θ if

$$\mathbf{A}(k+1) = \mathbf{I} - \mathbf{b}(k+1)\mathbf{h}'(k+1) \quad (6-12)$$

where $\mathbf{A}(k+1)$ and $\mathbf{b}(k+1)$ are deterministic. \square

We leave the proof of this result to the reader. Unbiasedness means that our recursive estimator does not have two independent design matrices (degrees of freedom), $\mathbf{A}(k+1)$ and $\mathbf{b}(k+1)$. Unbiasedness constraints $\mathbf{A}(k+1)$ to be a function of $\mathbf{b}(k+1)$. When (6-12) is substituted into (6-11), we obtain the following important structure for an unbiased linear recursive estimator of θ ,

$$\hat{\theta}(k+1) = \hat{\theta}(k) + \mathbf{b}(k+1)[z(k+1) - \mathbf{h}'(k+1)\hat{\theta}(k)] \quad (6-13)$$

Our recursive WLSE of θ has this structure; thus, as long as $\mathbf{h}(k+1)$ is deterministic, it produces unbiased estimates of θ . Many other estimators that we shall study will also have this structure.

EFFICIENCY

Did you hear the story about the conventioning statisticians who all drowned in a lake that was on the average 6 in. deep? The point of this rhetorical question is that unbiasedness by itself is not terribly meaningful. We must also study the dispersion about the mean, namely the variance. If the statisticians had known that the variance about the 6 in. average depth was 120 ft, they might not have drowned!

Ideally, we would like our estimator to be unbiased and to have the smallest possible error variance. We consider the case of a scalar parameter first.

Definition 6-2. An unbiased estimator, $\hat{\theta}(k)$ of θ is said to be more efficient than any other unbiased estimator, $\hat{\theta}(k)$, of θ , if

$$\text{Var}(\hat{\theta}(k)) \leq \text{Var}(\hat{\theta}(k)) \quad \text{for all } k \quad \square \quad (6-14)$$

Very often, it is of interest to know if $\hat{\theta}(k)$ satisfies (6-14) for all other unbiased estimators, $\hat{\theta}(k)$. This can be verified by comparing the variance of

Efficiency

$\hat{\theta}(k)$ with the smallest error variance that can ever be attained by any unbiased estimator. The following theorem provides a lower bound for $E\{\hat{\theta}^2(k)\}$ when θ is a scalar deterministic parameter. Theorem 6-4 generalizes these results to the case of a vector of deterministic parameters.

Theorem 6-3 (Cramer-Rao Inequality). Let \mathbf{z} denote a set of data [i.e., $\mathbf{z} = \text{col}(z_1, z_2, \dots, z_k)$; \mathbf{z} is also short for $\mathbf{z}(k)$]. If $\hat{\theta}(k)$ is an unbiased estimator of deterministic θ , then

$$E\{\hat{\theta}^2(k)\} \geq \frac{1}{E\left\{\left[\frac{\partial}{\partial \theta} \ln p(\mathbf{z})\right]^2\right\}} \quad \text{for all } k \quad (6-15)$$

Two other ways for expressing (6-15) are

$$E\{\hat{\theta}^2(k)\} \geq \frac{1}{\int_{-\infty}^{\infty} \left[\frac{\partial p(z)}{\partial \theta}\right]^2 \frac{1}{p(z)} dz} \quad \text{for all } k \quad (6-16)$$

and

$$E\{\hat{\theta}^2(k)\} \geq \frac{1}{-E\left\{\frac{\partial^2 \ln p(\mathbf{z})}{\partial \theta^2}\right\}} \quad \text{for all } k \quad (6-17)$$

where dz is short for dz_1, dz_2, \dots, dz_k . \square

Inequalities (6-15), (6-16) and (6-17) are named after Cramer and Rao, who discovered them. They are functions of k because \mathbf{z} is.

Before proving this theorem it is instructive to illustrate its use by means of an example.

Example 6-3

We are given M statistically independent observations of a random variable z that is known to have a Cauchy distribution, i.e.,

$$p(z_i) = \frac{1}{\pi[1 + (z_i - \theta)^2]} \quad (6-18)$$

Parameter θ is unknown and will be estimated using z_1, z_2, \dots, z_M . We shall determine the lower bound for the error-variance of any unbiased estimator of θ using (6-15). Observe that we are able to do this without having to specify an estimator structure for $\hat{\theta}$. Without further explanation, we calculate:

$$p(\mathbf{z}) = \prod_{i=1}^M p(z_i) = 1 / \left(\pi^M \prod_{i=1}^M [1 + (z_i - \theta)^2] \right) \quad (6-19)$$

$$\ln p(\mathbf{z}) = -M \ln \pi - \sum_{i=1}^M \ln [1 + (z_i - \theta)^2] \quad (6-20)$$

and

$$\frac{\partial \ln p(\mathbf{z})}{\partial \theta} = \sum_{i=1}^M 2(z_i - \theta) / [1 + (z_i - \theta)^2] \quad (6-21)$$

so that

$$E\left\{\left[\frac{\partial}{\partial \theta} \ln p(\mathbf{z})\right]^2\right\} = E\left\{\left[\sum_{i=1}^M \frac{2(z_i - \theta)}{1 + (z_i - \theta)^2}\right]\left[\sum_{j=1}^M \frac{2(z_j - \theta)}{1 + (z_j - \theta)^2}\right]\right\} \quad (6-22)$$

Next, we must evaluate the right-hand side of (6-22). This is tedious to do, but can be accomplished, as follows. Note that

$$\sum_{i \neq j} \sum_{i,j} + \sum_{i=j} \sum_{i,j} = TA + TB \quad (6-23)$$

Consider **TA** first, i.e.,

$$TA = E\left\{\sum_{i=1}^M 2(z_i - \theta) / [1 + (z_i - \theta)^2]\right\} E\left\{\sum_{j=1}^M 2(z_j - \theta) / [1 + (z_j - \theta)^2]\right\} \quad (6-24)$$

where we have made use of statistical independence of the measurements. Observe that

$$E\left\{\frac{z_i - \theta}{1 + (z_i - \theta)^2}\right\} = \int_{-\infty}^{\infty} \frac{y}{1 + y^2} \cdot \frac{1}{\pi} \frac{1}{1 + y^2} dy = 0 \quad (6-25)$$

where $y = z_i - \theta$. The integral is zero because the integrand is an odd function of y . Consequently,

$$TA = 0 \quad (6-26)$$

Next, consider **TB**, i.e.,

$$TB = E\left\{\sum_{i=1}^M 4(z_i - \theta)^2 / [1 + (z_i - \theta)^2]^2\right\} \quad (6-27)$$

which can also be written as

$$TB = 4 \sum_{i=1}^M TC \quad (6-28)$$

where

$$TC = E\{(z_i - \theta)^2 / [1 + (z_i - \theta)^2]^2\} \quad (6-29)$$

or

$$TC = \int_{-\infty}^{\infty} \frac{y^2}{(1 + y^2)^2} \frac{1}{\pi} \frac{1}{1 + y^2} dy \quad (6-30)$$

Integrating (6-30) by parts twice, we find that

$$TC = \frac{1}{8\pi} \int_{-\infty}^{\infty} \frac{dy}{1 + y^2} = \frac{1}{8} \quad (6-31)$$

because the integral in (6-31) is the area under the Cauchy probability density function, which equals π . Substituting (6-31) into (6-28), we determine that

$$TB = M/2 \quad (6-32)$$

thus, when (6-23) is substituted into (6-22), and that result is substituted into (6-15), we find that

$$E\{\hat{\theta}^2(M)\} \geq \frac{2}{M} \quad \text{for all } M \quad (6-33)$$

Observe that the Cramer-Rao bound depends on the number of measurements used to estimate θ . For large numbers of measurements, this bound equals zero. \square

Proof of Theorem 6-3. Because $\hat{\theta}(k)$ is an unbiased estimator of θ ,

$$E\{\hat{\theta}(k)\} = \int_{-\infty}^{\infty} [\hat{\theta}(k) - \theta] p(\mathbf{z}) d\mathbf{z} = 0 \quad (6-34)$$

Differentiating (6-34) with respect to θ , we find that

$$\int_{-\infty}^{\infty} [\hat{\theta}(k) - \theta] \frac{\partial p(\mathbf{z})}{\partial \theta} d\mathbf{z} - \int_{-\infty}^{\infty} p(\mathbf{z}) d\mathbf{z} = 0$$

which can be rewritten as

$$1 = \int_{-\infty}^{\infty} [\hat{\theta}(k) - \theta] \frac{\partial p(\mathbf{z})}{\partial \theta} d\mathbf{z} \quad (6-35)$$

As an aside, we note that

$$\frac{\partial}{\partial \theta} \ln p(\mathbf{z}) = \frac{\partial p(\mathbf{z})}{\partial \theta} \frac{1}{p(\mathbf{z})} \quad (6-36)$$

so that

$$\frac{\partial p(\mathbf{z})}{\partial \theta} = p(\mathbf{z}) \frac{\partial}{\partial \theta} \ln p(\mathbf{z}) \quad (6-37)$$

Substitute (6-37) into (6-35) to obtain

$$1 = \int_{-\infty}^{\infty} [(\hat{\theta}(k) - \theta) \sqrt{p(\mathbf{z})}] \left[\sqrt{p(\mathbf{z})} \frac{\partial}{\partial \theta} \ln p(\mathbf{z}) \right] d\mathbf{z} \quad (6-38)$$

Recall the Schwarz inequality

$$\left[\int_{-\infty}^{\infty} a(\mathbf{z}) b(\mathbf{z}) d\mathbf{z} \right]^2 \leq \left[\int_{-\infty}^{\infty} a^2(\mathbf{z}) d\mathbf{z} \right] \left[\int_{-\infty}^{\infty} b^2(\mathbf{z}) d\mathbf{z} \right] \quad (6-39)$$

where equality is achieved when $b(\mathbf{z}) = ca(\mathbf{z})$ in which c is an arbitrary constant. Next, square both sides of (6-38) and apply (6-39) to the new right-hand side, to see that

$$1 \leq \left[\int_{-\infty}^{\infty} [\hat{\theta}(k) - \theta]^2 p(\mathbf{z}) d\mathbf{z} \right] \left[\int_{-\infty}^{\infty} \left[\frac{\partial}{\partial \theta} \ln p(\mathbf{z}) \right]^2 p(\mathbf{z}) d\mathbf{z} \right]$$

or

$$1 \leq E\{\hat{\theta}^2(k)\} E\left\{\left[\frac{\partial}{\partial \theta} \ln p(\mathbf{z}) \right]^2\right\} \quad (6-40)$$

Finally, to obtain (6-15), solve (6-40) for $E\{\hat{\theta}^2(k)\}$.

In order to obtain (6-16) from (6-15), observe that

$$\begin{aligned} E\left\{\left[\frac{\partial}{\partial \theta} \ln p(\mathbf{z})\right]^2\right\} &= \int_{-\infty}^{\infty} \left[\frac{\partial p(\mathbf{z})}{\partial \theta} \frac{1}{p(\mathbf{z})}\right]^2 p(\mathbf{z}) d\mathbf{z} \\ &= \int_{-\infty}^{\infty} \left[\frac{\partial p(\mathbf{z})}{\partial \theta}\right]^2 \frac{1}{p(\mathbf{z})} d\mathbf{z} \end{aligned} \quad (6-41)$$

To obtain (6-41) we have also used (6-36).

In order to obtain (6-17), we begin with the identity

$$\int_{-\infty}^{\infty} p(\mathbf{z}) d\mathbf{z} = 1$$

and differentiate it twice with respect to θ , using (6-37) after each differentiation, to show that

$$\int_{-\infty}^{\infty} p(\mathbf{z}) \frac{\partial^2 \ln p(\mathbf{z})}{\partial \theta^2} d\mathbf{z} = - \int_{-\infty}^{\infty} \left[\frac{\partial \ln p(\mathbf{z})}{\partial \theta}\right]^2 p(\mathbf{z}) d\mathbf{z}$$

which can also be expressed as

$$E\left\{\left[\frac{\partial}{\partial \theta} \ln p(\mathbf{z})\right]^2\right\} = -E\left\{\frac{\partial^2 \ln p(\mathbf{z})}{\partial \theta^2}\right\} \quad (6-42)$$

Substitute (6-42) into (6-15) to obtain (6-17). \square

It is sometimes easier to compute the Cramer-Rao bound using one form [i.e., (6-15) or (6-16) or (6-17)] than another. The logarithmic forms are usually used when $p(\mathbf{z})$ is exponential (e.g., Gaussian).

Corollary 6-1. If the lower bound is achieved in Theorem 6-3, then

$$\hat{\theta}(k) = \frac{1}{c} \frac{\partial \ln p(\mathbf{z})}{\partial \theta} \quad (6-43)$$

where c is an arbitrary constant.

Proof. In deriving the Cramer-Rao bound we used the Schwarz inequality (6-39) for which equality is achieved when $b(\mathbf{z}) = c a(\mathbf{z})$. In our case $a(\mathbf{z}) = [\hat{\theta}(k) - \theta] \sqrt{p(\mathbf{z})}$ and $b(\mathbf{z}) = \sqrt{p(\mathbf{z})} (\partial/\partial \theta) \ln p(\mathbf{z})$. Setting $b(\mathbf{z}) = c a(\mathbf{z})$, we obtain (6-43). \square

Equation (6-43) links the structure of an estimator to the property of efficiency, because the left-hand side of (6-43) depends explicitly on $\hat{\theta}(k)$.

We turn next to the general case of a vector of parameters.

Definition 6-3. An unbiased estimator, $\hat{\theta}(k)$, of vector θ is said to be more efficient than any other unbiased estimator, $\hat{\theta}(k)$, of θ , if

$$E\{[\theta - \hat{\theta}(k)][\theta - \hat{\theta}(k)]'\} \leq E\{[\theta - \hat{\theta}(k)][\theta - \hat{\theta}(k)]'\} \quad \square \quad (6-44)$$

For a vector of parameters, we see that a more efficient estimator has the smallest error covariance among all unbiased estimators of θ , "smallest" in the sense that $E\{[\theta - \hat{\theta}(k)][\theta - \hat{\theta}(k)]'\} - E\{[\theta - \hat{\theta}(k)][\theta - \hat{\theta}(k)]'\}$ is negative semi-definite.

The generalization of the Cramer-Rao inequality to a vector of parameters is given next.

Theorem 6-4 (Cramer-Rao inequality for a vector of parameters). Let \mathbf{z} denote a set of data as in Theorem 6-3, and $\hat{\theta}(k)$ be any unbiased estimator of deterministic θ based on \mathbf{z} . Then

$$E\{\hat{\theta}(k)\hat{\theta}'(k)\} \geq \mathbf{J}^{-1} \quad \text{for all } k \quad (6-45)$$

where \mathbf{J} is the "Fisher information matrix,"

$$\mathbf{J} = E\left\{\left[\frac{\partial}{\partial \theta} \ln p(\mathbf{z})\right]\left[\frac{\partial}{\partial \theta} \ln p(\mathbf{z})\right]'\right\} \quad (6-46)$$

which can also be expressed as

$$\mathbf{J} = -E\left\{\frac{\partial^2}{\partial \theta^2} \ln p(\mathbf{z})\right\} \quad (6-47)$$

Equality holds in (6-45), if and only if

$$\left[\frac{\partial}{\partial \theta} \ln p(\mathbf{z})\right]' = c(\theta)\hat{\theta}(k) \quad \square \quad (6-48)$$

A complete proof of this result is given by Sorenson (1980, pp. 94–96). Although the proof is similar to our proof of Theorem 6-3, it is a bit more intricate because of the vector nature of θ .

Inequality (6-45) demonstrates that any unbiased estimator can have a covariance no smaller than \mathbf{J}^{-1} . Unfortunately, \mathbf{J}^{-1} is not a greatest lower bound for the error covariance. Other bounds exist which are tighter than (6-45) [e.g., the Bhattacharyya bound (Van Trees, 1968)], but they are even more difficult to compute than \mathbf{J}^{-1} .

Corollary 6-2. Let \mathbf{z} denote a set of data as in Theorem 6-3, and $\hat{\theta}_i(k)$ be any unbiased estimator of deterministic θ_i based on \mathbf{z} . Then

$$E\{\hat{\theta}_i^2(k)\} \geq (\mathbf{J}^{-1})_{ii} \quad i = 1, 2, \dots, n \text{ and all } k \quad (6-49)$$

where $(\mathbf{J}^{-1})_{ii}$ is the i -ith element in matrix \mathbf{J}^{-1} .

Proof. Inequality (6-45) means that $E\{\hat{\theta}(k)\hat{\theta}'(k)\} - \mathbf{J}^{-1}$ is a positive semi-definite matrix, i.e.,

$$\mathbf{a}'[E\{\hat{\theta}(k)\hat{\theta}'(k)\} - \mathbf{J}^{-1}]\mathbf{a} \geq 0 \quad (6-50)$$

where \mathbf{a} is an arbitrary nonzero vector. Choosing $\mathbf{a} = \mathbf{e}_i$ (the i th unit vector) we obtain (6-49). \square

Results similar to those in Theorem 6-4 and Corollary 6-2 are also available for a vector of random parameters (e.g., Sorenson, 1980, pp. 99–100). Let $p(\mathbf{z}, \theta)$ denote the joint probability density function between \mathbf{z} and θ . The Cramer-Rao inequality for random parameters is obtained from Theorems 6-3 and 6-4 by replacing $p(\mathbf{z})$ by $p(\mathbf{z}, \theta)$. Of course, the expectation is now with respect to \mathbf{z} and θ .

- 6-9. Suppose $\hat{\theta}(k)$ is a biased estimator of deterministic θ , with bias $B(\theta)$. Show that

$$\mathbb{E}\{\hat{\theta}^2(k)\} \geq \frac{\left[1 + \frac{\partial B(\theta)}{\partial \theta}\right]^2}{\mathbb{E}\left\{\left[\frac{\partial}{\partial \theta} \ln p(\mathbf{z})\right]^2\right\}} \quad \text{for all } k$$

PROBLEMS

- 6-1. Prove Theorem 6-2, which provides an unbiasedness constraint for the two design matrices that appear in a linear *recursive* estimator.
- 6-2. Prove Theorem 6-4, which provides the Cramer-Rao bound for a vector of parameters.
- 6-3. Random variable $X \sim N(x; \mu, \sigma^2)$, and we are given a random sample $\{x_1, x_2, \dots, x_N\}$. Consider the following estimator for μ ,

$$\hat{\mu}(N) = \frac{1}{N+a} \sum_{i=1}^N x_i$$

where $a \geq 0$. For what value(s) of a is $\hat{\mu}(N)$ an unbiased estimator of μ ?

- 6-4. Suppose z_1, z_2, \dots, z_N are random samples from a Gaussian distribution with unknown mean, μ , and variance, σ^2 . Reasonable estimators of μ and σ^2 are the sample mean and sample variance,

$$\bar{z} = \frac{1}{N} \sum_{i=1}^N z_i$$

and

$$s^2 = \frac{1}{N} \sum_{i=1}^N (z_i - \bar{z})^2$$

Is s^2 an unbiased estimator of σ^2 ? [Hint: Show that $\mathbb{E}[s^2] = (N-1)\sigma^2/N$]

- 6-5. (Mendel, 1973, first part of Exercise 2-9, pg. 137). Show that if $\hat{\theta}$ is an unbiased estimate of θ , $a\hat{\theta} + b$ is an unbiased estimate of $a\theta + b$.
- 6-6. Suppose that N independent observations (x_1, x_2, \dots, x_N) are made of a random variable X that is Gaussian, i.e.,

$$p(x_i | \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp[-(x_i - \mu)^2/2\sigma^2]$$

In this problem only μ is unknown. Derive the Cramer-Rao lower bound of $\mathbb{E}\{\hat{\mu}^2(N)\}$ for an unbiased estimator of μ .

- 6-7. Repeat Problem 6-6, but in this case assume that only σ^2 is unknown, i.e., derive the Cramer-Rao lower bound of $\mathbb{E}\{\hat{\sigma}^2(N)\}$ for an unbiased estimator of σ^2 .
- 6-8. Repeat Problem 6-6, but in this case assume both μ and σ^2 are unknown, i.e., compute \mathbf{J}^{-1} when $\theta = \text{col}(\mu, \sigma^2)$.

Large Sample Properties of Estimators

INTRODUCTION

To begin, we reiterate the fact that “if an estimator possesses a small-sample property it also possesses the associated large-sample property; but, the converse is not always true.” In this lesson we shall examine the following large-sample properties of estimators: asymptotic unbiasedness, consistency, and asymptotic efficiency. The first and third properties are natural extensions of the small-sample properties of unbiasedness and efficiency in the limiting situation of an infinite number of measurements. The second property is about convergence of $\hat{\theta}(k)$ to θ .

Before embarking on a discussion of these three large-sample properties, we digress a bit to introduce the concept of asymptotic distribution and its associated asymptotic mean and variance (or covariance, in the vector situation). Doing this will help us better understand these large-sample properties.

ASYMPTOTIC DISTRIBUTIONS

According to Kmenta (1971, pg. 163), “... if the distribution of an estimator tends to become more and more similar in form to some specific distribution as the sample size increases, then such a specific distribution is called the *asymptotic distribution* of the estimator in question.... What is meant by the asymptotic distribution is not the ultimate form of the distribution, which may

be degenerate, but the form that the distribution tends to put on in the last part of its journey to the final collapse (if this occurs).” Consider the situation depicted in Figure 7-1, where $p_i(\hat{\theta})$ denotes the probability density function associated with estimator $\hat{\theta}$ of the scalar parameter θ , based on i measurements. As the number of measurements increases, $p_i(\hat{\theta})$ changes its shape (although, in this example, each one of the density functions is Gaussian). The density function eventually centers itself about the true parameter value θ , and the variance associated with $p_i(\hat{\theta})$ tends to get smaller as i increases. Ultimately, the variance will become so small that in all probability $\hat{\theta} = \theta$. The asymptotic distribution refers to $p_i(\hat{\theta})$ as it evolves from $i = 1, 2, \dots$, etc., especially for large values of i .

The preceding example illustrates one of the three possible cases that can occur for an asymptotic distribution, namely the case when an estimator has a distribution of the same form regardless of the sample size, and this form is known (e.g., Gaussian). Some estimators have a distribution that, although not necessarily always of the same form, is also known for every sample size. For example, $p_5(\hat{\theta})$ may be uniform, $p_{20}(\hat{\theta})$ may be Rayleigh, and $p_{200}(\hat{\theta})$ may be Gaussian. Finally, for some estimators the distribution is not necessarily known for every sample size, but is known only for $k \rightarrow \infty$.

Asymptotic distributions, like other distributions, are characterized by their moments. We are especially interested in their first two moments, namely, the asymptotic mean and variance.

Definition 7-1. *The asymptotic mean is equal to the asymptotic expectation, namely $\lim_{k \rightarrow \infty} E\{\hat{\theta}(k)\}$.* □

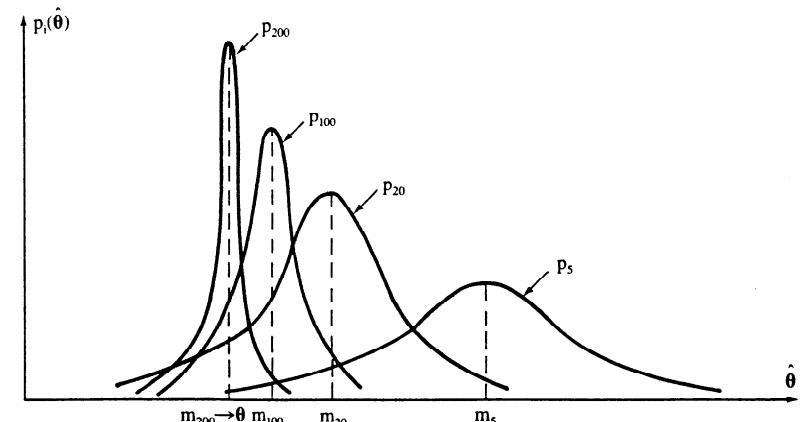


Figure 7-1 Probability density function for estimate of scalar θ , as a function of number of measurements, e.g., p_{20} is the p.d.f. for 20 measurements.

As noted in Goldberger (1964, pg. 116), if $E\{\hat{\theta}(k)\} = m$ for all k , then $\lim_{k \rightarrow \infty} E\{\hat{\theta}(k)\} = \lim_{k \rightarrow \infty} m = m$. Alternatively, suppose that

$$E\{\hat{\theta}(k)\} = m + k^{-1}c_1 + k^{-2}c_2 + \dots \quad (7-1)$$

where the c 's are finite constants; then,

$$\lim_{k \rightarrow \infty} E\{\hat{\theta}(k)\} = \lim_{k \rightarrow \infty} \{m + k^{-1}c_1 + k^{-2}c_2 + \dots\} = m \quad (7-2)$$

Thus if $E\{\hat{\theta}(k)\}$ is expressible as a power series in $k^0, k^{-1}, k^{-2}, \dots$, the asymptotic mean of $\hat{\theta}(k)$ is the leading term of this power series; as $k \rightarrow \infty$ the terms of "higher order of smallness" in k vanish.

Definition 7-2. *The asymptotic variance, which is short for "variance of the asymptotic distribution" is not equal to $\lim_{k \rightarrow \infty} \text{var}[\hat{\theta}(k)]$. It is defined as*

$$\text{asymptotic var}[\hat{\theta}(k)] = \frac{1}{k} \lim_{k \rightarrow \infty} E\left\{k[\hat{\theta}(k) - \lim_{k \rightarrow \infty} E\{\hat{\theta}(k)\}]^2\right\} \quad \square \quad (7-3)$$

Kmenta (1971, pg. 164) states "The asymptotic variance . . . is *not* equal to $\lim_{k \rightarrow \infty} \text{var}(\hat{\theta})$. The reason is that in the case of estimators whose variance decreases with an increase in k , the variance will approach zero as $k \rightarrow \infty$. This will happen when the distribution collapses on a point. But, as we explained, the asymptotic distribution is not the same as the collapsed (degenerate) distribution, and its variance is *not* zero."

Goldberger (1964, pg. 116) notes that if $E\{[\hat{\theta}(k) - \lim_{k \rightarrow \infty} E\{\hat{\theta}(k)\}]^2\} = v/k$ for all values of k , then $\text{asymptotic var}[\hat{\theta}(k)] = v/k$. Alternatively, suppose that

$$E\{[\hat{\theta}(k) - \lim_{k \rightarrow \infty} E\{\hat{\theta}(k)\}]^2\} = k^{-1}v + k^{-2}c_2 + k^{-3}c_3 + \dots \quad (7-4)$$

where the the c 's are finite constants; then,

$$\begin{aligned} \text{asymptotic var}[\hat{\theta}(k)] &= \frac{1}{k} \lim_{k \rightarrow \infty} (v + k^{-1}c_2 + k^{-2}c_3 + \dots) \\ &= \frac{v}{k} \end{aligned} \quad (7-5)$$

Thus if the variance of each $\hat{\theta}(k)$ is expressible as a power series in k^{-1}, k^{-2}, \dots , the asymptotic variance of $\hat{\theta}(k)$ is the leading term of this power series; as k goes to infinity the terms of "higher order of smallness" in k vanish. Observe that if (7-5) is true then the asymptotic variance of $\hat{\theta}(k)$ decreases as $k \rightarrow \infty \dots$ which corresponds to the situation depicted in Figure 7-1.

Extensions of our definitions of asymptotic mean and variance to sequences of random vectors [e.g., $\hat{\theta}(k), k = 1, 2, \dots$] are straightforward and can be found in Goldberger (1964, pg. 117).

ASYMPTOTIC UNBIASEDNESS

Definition 7-3. *Estimator $\hat{\theta}(k)$ is an asymptotically unbiased estimator of deterministic θ , if*

$$\lim_{k \rightarrow \infty} E\{\hat{\theta}(k)\} = \theta \quad (7-6)$$

or of random θ , if

$$\lim_{k \rightarrow \infty} E\{\hat{\theta}(k)\} = E\{\theta\} \quad \square \quad (7-7)$$

Figure 7-1 depicts an example in which the asymptotic mean of $\hat{\theta}$ has converged to θ (note that p_{200} is centered about $m_{200} = \theta$).

Note that for the calculation on the left-hand side of either (7-6) or (7-7) to be performed, the asymptotic distribution of $\hat{\theta}(k)$ must exist, because

$$E\{\hat{\theta}(k)\} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \hat{\theta}(k)p(\hat{\theta}(k))d\hat{\theta}(k) \quad (7-8)$$

Example 7-1

Recall our linear model $\mathbf{Z}(k) = \mathbf{H}(k)\theta + \mathbf{V}(k)$ in which $E\{\mathbf{V}(k)\} = \mathbf{0}$. Let us assume that each component of $\mathbf{V}(k)$ is uncorrelated and has the same variance σ_v^2 . In Lesson 8, we determine an unbiased estimator for σ_v^2 . Here, on the other hand, we just assume that

$$\hat{\sigma}_v^2(k) \triangleq \frac{\tilde{\mathbf{Z}}'(k)\tilde{\mathbf{Z}}(k)}{k} \quad (7-9)$$

where $\tilde{\mathbf{Z}}(k) = \mathbf{Z}(k) - \mathbf{H}(k)\hat{\theta}_{LS}(k)$. We leave it to the reader to show that (see Lesson 8)

$$E\{\hat{\sigma}_v^2(k)\} = \left(\frac{k-n}{k}\right)\sigma_v^2 \quad (7-10)$$

Observe that $\hat{\sigma}_v^2(k)$ is not an unbiased estimator of σ_v^2 ; but, it is an asymptotically unbiased estimator, because $\lim_{k \rightarrow \infty} [(k-n)/k]\hat{\sigma}_v^2 = \sigma_v^2$. \square

CONSISTENCY

We now direct our attention to the issue of stochastic convergence. The reader should review the different modes of stochastic convergence, especially convergence in probability and mean-squared convergence (see, for example, Papoulis, 1965).

Definition 7-4. *The probability limit of $\hat{\theta}(k)$ is the point θ^* on which the distribution of our estimator collapses. We abbreviate "probability limit of $\hat{\theta}(k)$ " by $\text{plim } \hat{\theta}(k)$. Mathematically speaking,*

$$\text{plim } \hat{\theta}(k) = \theta^* \Leftrightarrow \lim_{k \rightarrow \infty} \Pr[|\hat{\theta}(k) - \theta^*| \geq \epsilon] \rightarrow 0 \quad (7-11)$$

where ϵ is a small positive number. \square

Definition 7-5. $\hat{\theta}(k)$ is a consistent estimator of θ if

$$\text{plim } \hat{\theta}(k) = \theta \quad \square \quad (7-12)$$

Note that “consistency” means the same thing as “convergence in probability.” For an estimator to be consistent, its probability limit θ^* must equal its true value θ . Note, also, that a consistent estimator need not be unbiased or asymptotically unbiased.

Why is convergence in probability so popular and widely used in the estimation field? One reason is that $\text{plim}(\cdot)$ can be treated as an operator. For example, suppose X_k and Y_k are two random sequences, for which $\text{plim } X_k = X$ and $\text{plim } Y_k = Y$; then (see Tucker, 1962, for simple proofs of these and other facts),

$$\text{plim } X_k Y_k = (\text{plim } X_k)(\text{plim } Y_k) = XY \quad (7-13)$$

and

$$\text{plim} \left(\frac{X_k}{Y_k} \right) = \frac{\text{plim } X_k}{\text{plim } Y_k} = \frac{X}{Y} \quad (7-14)$$

Additionally, suppose A_k and B_k are two commensurate matrix sequences, for which $\text{plim } A_k = A$ and $\text{plim } B_k = B$ [note that $\text{plim } A_k$, for example, means $(\text{plim } a_{ij}(k))_{ij}$]; then

$$\text{plim } A_k B_k = (\text{plim } A_k)(\text{plim } B_k) = AB \quad (7-15)$$

$$\text{plim } A_k^{-1} = (\text{plim } A_k)^{-1} = A^{-1} \quad (7-16)$$

and

$$\text{plim } A_k^{-1} B_k = A^{-1} B \quad (7-17)$$

The treatment of $\text{plim}(\cdot)$ as an operator often makes the study of consistency quite easy. We shall demonstrate the truth of this in Lesson 8 when we examine the consistency of the least-squares estimator.

A second reason for the importance of consistency is the property that “consistency carries over”; i.e., any continuous function of a consistent estimator is itself a consistent estimator [see Tucker, 1967, for a proof of this property, which relies heavily on the preceding treatment of $\text{plim}(\cdot)$ as an operator].

Example 7-2

Suppose $\hat{\theta}$ is a consistent estimator of θ . Then $1/\hat{\theta}$ is a consistent estimator of $1/\theta$, $(\hat{\theta})^2$ is a consistent estimator of θ^2 , and $\ln \hat{\theta}$ is a consistent estimator of $\ln \theta$. These facts are all due to the consistency carry-over property. \square

The reader may be scratching his or her head at this point and wondering about the emphasis placed on these illustrative examples. Isn’t, for example, using $\hat{\theta}$ to estimate θ^2 by $(\hat{\theta})^2$ the “natural” thing to do? The answer is “Yes, but only if we know ahead of time that $\hat{\theta}$ is a consistent estimator of θ .” If you

do not know this to be true, then there is no guarantee that $\widehat{\theta^2} = (\hat{\theta})^2$. In Lesson 11 we show that maximum-likelihood estimators are consistent; thus $\widehat{(\theta^2)_{ML}} = (\hat{\theta}_{ML})^2$. We mention this property about maximum-likelihood estimators here, because one must know whether or not an estimator is consistent before applying the consistency carry-over property. Not all estimators are consistent!

Finally, this carry-over property for consistency does not necessarily apply to other properties. For example, if $\hat{\theta}(k)$ is an unbiased estimator of θ , then $A\hat{\theta}(k) + b$ will be an unbiased estimator of $A\theta + b$; but, $\hat{\theta}^2(k)$ will not be an unbiased estimator of θ^2 .

How do you determine whether or not an estimator is consistent? Often, the direct approach, which makes heavy use of $\text{plim}(\cdot)$ operator algebra, is possible. Sometimes, an indirect approach is used, one which examines whether both the bias in $\hat{\theta}(k)$ and variance of $\hat{\theta}(k)$ approach zero as $k \rightarrow \infty$. In order to understand the validity of this indirect approach, we digress to discuss mean-squared convergence and its relationship to convergence in probability.

Definition 7-6. Estimator $\hat{\theta}(k)$ converges to θ in a mean-squared sense, if

$$\lim_{k \rightarrow \infty} E\{[\hat{\theta}(k) - \theta]^2\} \rightarrow 0 \quad \square \quad (7-18)$$

Theorem 7-1. If $\hat{\theta}(k)$ converges to θ in mean-square, then it converges to θ in probability.

Proof (Papoulis, 1965, pg. 151). Recall the Inequality of Bienaymé

$$\Pr[|x - a| \geq \epsilon] \leq E\{|x - a|^2\}/\epsilon^2 \quad (7-19)$$

Let $a = 0$ and $x = \hat{\theta}(k) - \theta$ in (7-19), and take the limit as $k \rightarrow \infty$ on both sides of (7-19), to see that

$$\lim_{k \rightarrow \infty} \Pr[|\hat{\theta}(k) - \theta| \geq \epsilon] \leq \lim_{k \rightarrow \infty} E\{[\hat{\theta}(k) - \theta]^2\}/\epsilon^2 \quad (7-20)$$

Using the fact that $\hat{\theta}(k)$ converges to θ in mean-square, we see that

$$\lim_{k \rightarrow \infty} \Pr[|\hat{\theta}(k) - \theta| \geq \epsilon] \rightarrow 0 \quad (7-21)$$

thus, $\hat{\theta}(k)$ converges to θ in probability. \square

Recall, from probability theory, that although mean-squared convergence implies convergence in probability, the converse is not true.

Example 7-3 (Kmenta, 1971, pg. 166)

Let $\hat{\theta}(k)$ be an estimator of θ , and let the probability density function of $\hat{\theta}(k)$ be

$$\begin{array}{c} \hat{\theta}(k) & p(\hat{\theta}(k)) \\ \hline \theta & 1 - \frac{1}{k} \\ k & \frac{1}{k} \end{array}$$

In this example $\hat{\theta}(k)$ can only assume two different values, θ and k . Obviously, $\hat{\theta}(k)$ is consistent, because as $k \rightarrow \infty$ the probability that $\hat{\theta}(k)$ equals θ approaches unity, i.e., $\text{plim } \hat{\theta}(k) = \theta$. Observe, also, that $E\{\hat{\theta}(k)\} = 1 + \theta(1 - 1/k)$, which means that $\hat{\theta}(k)$ is biased.

Now let us investigate the mean-squared error between $\hat{\theta}$ and θ ; i.e.,

$$\begin{aligned}\lim_{k \rightarrow \infty} E\{[\hat{\theta}(k) - \theta]^2\} &= \lim_{k \rightarrow \infty} \left[(\theta - \theta)^2 \left(1 - \frac{1}{k}\right) + (k - \theta)^2 \frac{1}{k} \right] \\ &= \lim_{k \rightarrow \infty} k \rightarrow \infty\end{aligned}$$

In this pathological example, the mean-squared error is diverging to infinity; but, $\hat{\theta}(k)$ converges to θ in probability. \square

Theorem 7-2. Let $\hat{\theta}(k)$ denote an estimator of θ . If bias $\hat{\theta}(k)$ and variance $\hat{\theta}(k)$ both approach zero as $k \rightarrow \infty$, then the mean-squared error between $\hat{\theta}(k)$ and θ approaches zero, and, therefore, $\hat{\theta}(k)$ is a consistent estimator of θ .

Proof. From elementary probability theory, we know that

$$E\{[\hat{\theta}(k) - \theta]^2\} = [\text{bias } \hat{\theta}(k)]^2 + \text{variance } \hat{\theta}(k) \quad (7-22)$$

If, as assumed, bias $\hat{\theta}(k)$ and variance $\hat{\theta}(k)$ both approach zero as $k \rightarrow \infty$, then

$$\lim_{k \rightarrow \infty} E\{[\hat{\theta}(k) - \theta]^2\} \rightarrow 0 \quad (7-23)$$

which means that $\hat{\theta}(k)$ converges to θ in mean-square. Thus, by Theorem 7-1 $\hat{\theta}(k)$ also converges to θ in probability. \square

The importance of Theorem 7-2 is that it provides a constructive way to test for consistency.

ASYMPTOTIC EFFICIENCY

Definition 7-7. $\hat{\theta}(k)$ is an asymptotically efficient estimator of scalar parameter θ , if: 1. $\hat{\theta}(k)$ has an asymptotic distribution with finite mean and variance, 2. $\hat{\theta}(k)$ is consistent, and 3. the variance of the asymptotic distribution equals

$$1/E\left\{\left[\frac{\partial}{\partial \theta} \ln p(z)\right]^2\right\} \quad \square$$

For the case of a vector of parameters conditions 1. and 2. in this definition are unchanged; however, condition 3. is changed to read "if the covariance matrix of the asymptotic distribution equals J^{-1} (see Theorem 6-4)." □

PROBLEMS

7-1. Random variable $X \sim N(x; \mu, \sigma^2)$, and we are given a random sample $\{x_1, x_2, \dots, x_N\}$. Consider the following estimator for μ ,

$$\hat{\mu}(N) = \frac{1}{N+a} \sum_{i=1}^N x_i$$

where $a \geq 0$.

- (a) For what value(s) of a is $\hat{\mu}(N)$ an asymptotically unbiased estimator of μ ?
- (b) Prove that $\hat{\mu}(N)$ is a consistent estimator of μ for all $a \geq 0$.
- (c) Compare the results obtained in (a) with those obtained in Problem 6-3.

7-2. Suppose z_1, z_2, \dots, z_N are random samples from a Gaussian distribution with unknown mean, μ , and variance, σ^2 . Reasonable estimators of μ and σ^2 are the sample mean and sample variance

$$\bar{z} = \frac{1}{N} \sum_{i=1}^N z_i$$

and

$$s^2 = \frac{1}{N} \sum_{i=1}^N (z_i - \bar{z})^2$$

- (a) Is s^2 an asymptotically unbiased estimator of σ^2 ? [Hint: Show that $E\{s^2\} = (N-1)\sigma^2/N$.]
- (b) Compare the result in (a) with that from Problem 6-4.
- (c) One can show that the variance of s^2 is

$$\text{var}(s^2) = \frac{\mu_4 - \sigma^4}{N} - \frac{2(\mu_4 - 2\sigma^4)}{N^2} + \frac{\mu_4 - 3\sigma^4}{N^3}$$

Explain whether or not s^2 is a consistent estimator of σ^2 .

7-3. Random variable $X \sim N(x; \mu, \sigma^2)$. Consider the following estimator of the population mean obtained from a random sample of N observations of X ,

$$\hat{\mu}(N) = \bar{x} + \frac{a}{N}$$

where a is a finite constant, and \bar{x} is the sample mean.

- (a) What are the asymptotic mean and variance of $\hat{\mu}(N)$?
- (b) Is $\hat{\mu}(N)$ a consistent estimator of μ ?
- (c) Is $\hat{\mu}(N)$ asymptotically efficient?

7-4. Let X be a Gaussian variable with mean μ and variance σ^2 . Consider the problem of estimating μ from a random sample of observations x_1, x_2, \dots, x_N . Three estimators are proposed:

$$\hat{\mu}_1(N) = \frac{1}{N} \sum_{i=1}^N x_i$$

$$\hat{\mu}_2(N) = \frac{1}{N+1} \sum_{i=1}^N x_i$$

$$\hat{\mu}_3(N) = \frac{1}{2} x_1 + \frac{1}{2N} \sum_{i=2}^N x_i$$

You are to study unbiasedness, efficiency, asymptotic unbiasedness, consistency, and asymptotic efficiency for these estimators. Show the analysis that allows you to complete the following table.

Properties	Estimator		
	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\mu}_3$
Small sample			
Unbiasedness			
Efficiency			
Large sample			
Unbiasedness			
Consistency			
Efficiency			

Entries to this table are *yes* or *no*.

- 7-5. If $\text{plim } X_k = X$ and $\text{plim } Y_k = Y$ where X and Y are constants, then $\text{plim } (X_k + Y_k) = X + Y$ and $\text{plim } c X_k = cX$, where c is a constant. Prove that $\text{plim } X_k Y_k = XY$ {Hint: $X_k Y_k = \frac{1}{4} [(X_k + Y_k)^2 - (X_k - Y_k)^2]$ }.

Lesson 8

Properties of Least-Squares Estimators

INTRODUCTION

In this lesson we study some small- and large-sample properties of least-squares estimators. Recall that in least-squares we estimate the $n \times 1$ parameter vector θ of the linear model $\mathbf{Z}(k) = \mathcal{H}(k)\theta + \mathcal{V}(k)$. We will see that most of the results in this lesson require $\mathcal{H}(k)$ to be deterministic, or, $\mathcal{H}(k)$ and $\mathcal{V}(k)$ to be statistically independent. In some applications one or the other of these requirements is met; however, there are many important applications where neither is met.

SMALL SAMPLE PROPERTIES OF LEAST-SQUARES ESTIMATORS

In this section (parts of which are taken from Mendel, 1973, pp. 75–86) we examine the bias and variance of weighted least-squares and least-squares estimators.

To begin, we recall Example 6-2, in which we showed that, when $\mathcal{H}(k)$ is deterministic, the WLSE of θ is unbiased. We also showed [after the statement of Theorem 6-2 and Equation (6-13)] that our recursive WLSE of θ has the requisite structure of an unbiased estimator; but, that unbiasedness of the recursive WLSE of θ also requires $\mathbf{h}(k+1)$ to be deterministic.

When $\mathcal{H}(k)$ is random, we have the following important result:

Theorem 8-1. *The WLSE of θ .*

$$\hat{\theta}_{\text{WLS}}(k) = [\mathcal{H}'(k)\mathbf{W}(k)\mathcal{H}(k)]^{-1}\mathcal{H}'(k)\mathbf{W}(k)\mathbf{Z}(k) \quad (8-1)$$

is unbiased if $\mathcal{V}(k)$ is zero mean and if $\mathcal{H}(k)$ and $\mathcal{V}(k)$ are statistically independent.

Note that this is the first place where, in connection with least squares, we have had to assume any a priori knowledge about noise $\mathcal{V}(k)$.

Proof. From (8-1) and $\mathcal{Z}(k) = \mathcal{H}(k)\theta + \mathcal{V}(k)$, we find that

$$\begin{aligned}\hat{\theta}_{\text{WLS}}(k) &= (\mathcal{H}'\mathcal{W}\mathcal{H})^{-1}\mathcal{H}'\mathcal{W}(\mathcal{H}\theta + \mathcal{V}) \\ &= \theta + (\mathcal{H}'\mathcal{W}\mathcal{H})^{-1}\mathcal{H}'\mathcal{W}\mathcal{V} \quad \text{for all } k\end{aligned}\quad (8-2)$$

where, for notational simplification, we have omitted the functional dependences of \mathcal{H} , \mathcal{W} , and \mathcal{V} on k . Taking the expectation on both sides of (8-2), it follows that

$$\mathbb{E}\{\hat{\theta}_{\text{WLS}}(k)\} = \theta + \mathbb{E}\{\mathcal{H}'\mathcal{W}\mathcal{H}\}^{-1}\mathcal{H}'\mathcal{W}\mathbb{E}\{\mathcal{V}\} \quad \text{for all } k \quad (8-3)$$

In deriving (8-3) we have used the fact that $\mathcal{H}(k)$ and $\mathcal{V}(k)$ are statistically independent [recall that if two random variables, a and b , are statistically independent $p(a, b) = p(a)p(b)$; thus, $\mathbb{E}\{ab\} = \mathbb{E}\{a\}\mathbb{E}\{b\}$ and $\mathbb{E}\{g(a)h(b)\} = \mathbb{E}\{g(a)\}\mathbb{E}\{h(b)\}$]. The second term in (8-3) is zero, because $\mathbb{E}\{\mathcal{V}\} = \mathbf{0}$, and therefore

$$\mathbb{E}\{\hat{\theta}_{\text{WLS}}(k)\} = \theta \quad \text{for all } k \quad \square \quad (8-4)$$

This theorem only states *sufficient conditions* for unbiasedness of $\hat{\theta}_{\text{WLS}}(k)$, which means that, if we do not satisfy these conditions, we *cannot* conclude anything about whether $\hat{\theta}_{\text{WLS}}(k)$ is unbiased or biased. In order to obtain *necessary conditions* for unbiasedness, assume that $\mathbb{E}\{\hat{\theta}_{\text{WLS}}(k)\} = \theta$ and take the expectation on both sides of (8-2). Doing this, we see that

$$\mathbb{E}\{(\mathcal{H}'\mathcal{W}\mathcal{H})^{-1}\mathcal{H}'\mathcal{W}\mathcal{V}\} = \mathbf{0} \quad (8-5)$$

Letting $\mathbf{M} = (\mathcal{H}'\mathcal{W}\mathcal{H})^{-1}\mathcal{H}'\mathcal{W}$ and \mathbf{m}_i' denote the i th row of matrix \mathbf{M} , (8-5) can be expressed as the following collection of *orthogonality conditions*,

$$\mathbb{E}\{\mathbf{m}_i'\mathcal{V}\} = 0 \quad \text{for } i = 1, 2, \dots, N \quad (8-6)$$

Orthogonality [recall that two random variables a and b are orthogonal if $\mathbb{E}\{ab\} = 0$] is a weaker condition than statistical independence, but is often more difficult to verify ahead of time than independence, especially since \mathbf{m}_i' is a very nonlinear transformation of the random elements of \mathcal{H} .

Example 8-1

Recall the impulse response identification Example 2-1, in which $\theta = \text{col}[h(1), h(2), \dots, h(n)]$, where $h(i)$ is the value of the sampled impulse response at time t_i . System input $u(k)$ may be deterministic or random.

If $\{u(k), k = 0, 1, \dots, N - 1\}$ is deterministic, then $\mathcal{H}(N - 1)$ [see Equation (2-5)] is deterministic, so that $\hat{\theta}_{\text{WLS}}(N)$ is an unbiased estimator of θ . Often, one uses a random input sequence for $\{u(k), k = 0, 1, \dots, N - 1\}$, such as from a random

number generator. This random sequence is in no way related to the measurement noise process, which means that $\mathcal{H}(N - 1)$ and $\mathcal{V}(N)$ are statistically independent, and again $\hat{\theta}_{\text{WLS}}(N)$ will be an unbiased estimate of the impulse response coefficients. We conclude, therefore, that *WLSEs of impulse response coefficients are unbiased*. \square

Example 8-2

As a further illustration of an application of Theorem 8-1, let us take a look at the weighted least-squares estimates of the n α -coefficients in the Example 2-2 AR model. We shall now demonstrate that $\mathcal{H}(N - 1)$ and $\mathcal{V}(N - 1)$, which are defined in Equation (2-8), are dependent, which means of course that we cannot apply Theorem 8-1 to study the unbiasedness of the WLSEs of the α -coefficients.

We represent the explicit dependences of \mathcal{H} and \mathcal{V} on their elements in the following manner:

$$\mathcal{H} = \mathcal{H}[y(N - 1), y(N - 2), \dots, y(0)] \quad (8-7)$$

and

$$\mathcal{V} = \mathcal{V}[u(N - 1), u(N - 2), \dots, u(0)] \quad (8-8)$$

Direct iteration of difference equation (7) in Lesson 2 for $k = 1, 2, \dots, N - 1$, reveals that $y(1)$ depends on $u(0)$, $y(2)$ depends on $u(1)$ and $u(0)$, and finally, that $y(N - 1)$ depends on $u(N - 2), \dots, u(0)$; thus,

$$\mathcal{H}[y(N - 1), y(N - 2), \dots, y(0)] = \mathcal{H}[u(N - 2), u(N - 3), \dots, u(0), y(0)] \quad (8-9)$$

Comparing (8-8) and (8-9), we see that \mathcal{H} and \mathcal{V} depend on similar values of random input u ; hence, they are statistically dependent. \square

Example 8-3

We are interested in estimating the parameter a in the following first-order system:

$$y(k + 1) = -ay(k) + u(k) \quad (8-10)$$

where $u(k)$ is a zero-mean white noise sequence. One approach to doing this is to collect $y(k + 1), y(k), \dots, y(1)$ as follows,

$$\begin{bmatrix} y(k + 1) \\ y(k) \\ y(k - 1) \\ \vdots \\ y(1) \end{bmatrix} = \begin{bmatrix} -y(k) \\ -y(k - 1) \\ -y(k - 2) \\ \vdots \\ -y(0) \end{bmatrix} a + \underbrace{\begin{bmatrix} u(k) \\ u(k - 1) \\ u(k - 2) \\ \vdots \\ u(0) \end{bmatrix}}_{\mathcal{V}(k)} \quad (8-11)$$

and, to obtain \hat{a}_{LS} . In order to study the bias of \hat{a}_{LS} , we use (8-2) in which $\mathcal{W}(k)$ is set equal to \mathbf{I} , and $\mathcal{H}(k)$ and $\mathcal{V}(k)$ are defined in (8-11). We also set $\hat{\theta}_{\text{LS}}(k) = \hat{a}_{\text{LS}}(k + 1)$. The argument of \hat{a}_{LS} is $k + 1$ instead of k , because the argument of \mathcal{Z} in (8-11) is $k + 1$. Doing this, we find that

$$\hat{a}_{\text{LS}}(k + 1) = a - \frac{\sum_{i=0}^k u(i)y(i)}{\sum_{j=0}^k y^2(j)} \quad (8-12)$$

thus,

$$E\{\hat{a}_{LS}(k+1)\} = a - \sum_{i=0}^k E\left\{\sum_{j=0}^k y^2(j)\right\} \quad (8-13)$$

Note, from (8-10), that $y(j)$ depends at most on $u(j-1)$; therefore,

$$E\left\{\sum_{j=0}^k y^2(j)\right\} = E\{u(k)\}E\left\{\sum_{j=0}^k y^2(j)\right\} = 0 \quad (8-14)$$

because $E\{u(k)\} = 0$. Unfortunately, all of the remaining terms in (8-13), i.e., for $i = 0, 1, \dots, k-1$, will not be equal to zero; consequently,

$$E\{\hat{a}_{LS}(k+1)\} = a + 0 + k \text{ nonzero terms} \quad (8-15)$$

Unless we are very lucky so that the k nonzero terms sum identically to zero, $E\{\hat{a}_{LS}(k+1)\} \neq a$, which means, of course, that \hat{a}_{LS} is biased.

The results in this example generalize to higher-order difference equations, so that we can conclude that *least-squares estimates of coefficients in an AR model are biased*. \square

In the *method of instrumental variables* $\mathcal{H}(k)$ is replaced by $\mathcal{H}^*(k)$ where $\mathcal{H}^*(k)$ is chosen so that it is statistically independent of $\mathcal{V}(k)$. There can be, and in general there will be, many choices of $\mathcal{H}^*(k)$ that may qualify as instrumental variables. It is often difficult to check that $\mathcal{H}^*(k)$ is statistically independent of $\mathcal{V}(k)$.

Next we proceed to compute the covariance matrix of $\tilde{\theta}_{WLS}(k)$, where

$$\tilde{\theta}_{WLS}(k) = \theta - \hat{\theta}_{LS}(k) \quad (8-16)$$

Theorem 8-2. *If $E\{\mathcal{V}(k)\} = \mathbf{0}$, $\mathcal{V}(k)$ and $\mathcal{H}(k)$ are statistically independent, and*

$$E\{\mathcal{V}(k)\mathcal{V}'(k)\} = \mathcal{R}(k) \quad (8-17)$$

then

$$\text{cov}[\tilde{\theta}_{WLS}(k)] = E_{\mathcal{H}}\{(\mathcal{H}'\mathbf{W}\mathcal{H})^{-1}\mathcal{H}'\mathbf{W}\mathcal{R}\mathbf{W}\mathcal{H}(\mathcal{H}'\mathbf{W}\mathcal{H})^{-1}\} \quad (8-18)$$

Proof. Because $E\{\mathcal{V}(k)\} = \mathbf{0}$ and $\mathcal{V}(k)$ and $\mathcal{H}(k)$ are statistically independent, $E\{\tilde{\theta}_{WLS}(k)\} = \mathbf{0}$, so that

$$\text{cov}[\tilde{\theta}_{WLS}(k)] = E\{\tilde{\theta}_{WLS}(k)\tilde{\theta}_{WLS}'(k)\} \quad (8-19)$$

Using (8-2) in (8-16), we see that

$$\tilde{\theta}_{WLS}(k) = -(\mathcal{H}'\mathbf{W}\mathcal{H})^{-1}\mathcal{H}'\mathbf{W}\mathcal{V} \quad (8-20)$$

hence

$$\text{cov}[\tilde{\theta}_{WLS}(k)] = E_{\mathcal{H}, \mathcal{V}}\{(\mathcal{H}'\mathbf{W}\mathcal{H})^{-1}\mathcal{H}'\mathbf{W}\mathcal{V}\mathcal{V}'\mathbf{W}\mathcal{H}(\mathcal{H}'\mathbf{W}\mathcal{H})^{-1}\} \quad (8-21)$$

where we have made use of the fact that \mathbf{W} is a symmetric matrix and the transpose and inverse symbols may be permuted. From probability theory (e.g., Papoulis, 1965), recall that

$$E_{\mathcal{H}, \mathcal{V}}\{\cdot\} = E_{\mathcal{H}}\{E_{\mathcal{V}|\mathcal{H}}\{\cdot | \mathcal{H}\}\} \quad (8-22)$$

Applying (8-22) to (8-21), we obtain (8-18). \square

As it stands, Theorem 8-2 is not too useful because it is virtually impossible to compute the expectation in (8-18), due to the highly nonlinear dependence of $(\mathcal{H}'\mathbf{W}\mathcal{H})^{-1}\mathcal{H}'\mathbf{W}\mathcal{R}\mathbf{W}\mathcal{H}(\mathcal{H}'\mathbf{W}\mathcal{H})^{-1}$ on \mathcal{H} . The following special case of Theorem 8-2 is important in practical applications in which $\mathcal{H}(k)$ is deterministic and $\mathcal{R}(k) = \sigma_v^2 \mathbf{I}$.

Corollary 8-1. *Given the conditions in Theorem 8-2, and that $\mathcal{H}(k)$ is deterministic, and, the components of $\mathcal{V}(k)$ are independent and identically distributed with zero-mean and constant variance σ_v^2 , then*

$$\text{cov}[\tilde{\theta}_{LS}(k)] = \sigma_v^2 [\mathcal{H}'(k)\mathcal{H}(k)]^{-1} \quad (8-23)$$

Proof. When \mathcal{H} is deterministic, $\text{cov}[\tilde{\theta}_{WLS}(k)]$ is obtained from (8-18) by deleting the expectation on its right-hand side. To obtain $\text{cov}[\tilde{\theta}_{LS}(k)]$ when $\text{cov}[\mathcal{V}(k)] = \sigma_v^2 \mathbf{I}$, set $\mathbf{W} = \mathbf{I}$ and $\mathcal{R}(k) = \sigma_v^2 \mathbf{I}$ in (8-18). The result is (8-23). \square

Usually, when we use a least-squares estimation algorithm we do not know the numerical value of σ_v^2 . If σ_v^2 is known ahead of time, it can be used directly in the estimate of θ . We show how to do this in Lesson 9. Where do we obtain σ_v^2 in order to compute (8-23)? We can estimate it!

Theorem 8-3. *An unbiased estimator of σ_v^2 is*

$$\hat{\sigma}_v^2(k) = \tilde{\mathcal{L}}'(k)\tilde{\mathcal{L}}(k)/(k-n) \quad (8-24)$$

where

$$\tilde{\mathcal{L}}(k) = \mathcal{L}(k) - \mathcal{H}(k)\hat{\theta}_{LS}(k) \quad (8-25)$$

Proof. We shall proceed by computing $E\{\tilde{\mathcal{L}}'(k)\tilde{\mathcal{L}}(k)\}$ and then approximating it as $\tilde{\mathcal{L}}'(k)\tilde{\mathcal{L}}(k)$, because the latter quantity can be computed from $\mathcal{L}(k)$ and $\hat{\theta}_{LS}(k)$, as in (8-25).

First, we compute an expression for $\tilde{\mathcal{L}}(k)$. Substituting both the linear model for $\mathcal{L}(k)$ and the least-squares formula for $\hat{\theta}_{LS}(k)$ into (8-25), we find that

$$\begin{aligned} \tilde{\mathcal{L}} &= \mathcal{H}\theta + \mathcal{V} - \mathcal{H}(\mathcal{H}'\mathcal{H})^{-1}\mathcal{H}'(\mathcal{H}\theta + \mathcal{V}) \\ &= \mathcal{V} - \mathcal{H}(\mathcal{H}'\mathcal{H})^{-1}\mathcal{H}'\mathcal{V} \\ &= [\mathbf{I}_k - \mathcal{H}(\mathcal{H}'\mathcal{H})^{-1}\mathcal{H}']\mathcal{V} \end{aligned} \quad (8-26)$$

where \mathbf{I}_k is the $k \times k$ identity matrix. Let

$$\mathbf{M} = \mathbf{I}_k - \mathcal{H}(\mathcal{H}'\mathcal{H})^{-1}\mathcal{H}' \quad (8-27)$$

Matrix \mathbf{M} is idempotent, i.e., $\mathbf{M}' = \mathbf{M}$ and $\mathbf{M}^2 = \mathbf{M}$; therefore,

$$\begin{aligned} E\{\tilde{\mathcal{L}}'\tilde{\mathcal{L}}\} &= E\{\mathcal{V}'\mathbf{M}'\mathbf{M}\mathcal{V}\} = E\{\mathcal{V}'\mathbf{M}\mathcal{V}\} \\ &= E\{\text{tr } \mathbf{M}\mathcal{V}\mathcal{V}'\} \end{aligned} \quad (8-28)$$

Recall the following well-known facts about the trace of a matrix:

1. $E\{\text{tr } \mathbf{A}\} = \text{tr } E\{\mathbf{A}\}$
2. $\text{tr } c\mathbf{A} = c \text{ tr } \mathbf{A}$, where c is a scalar
3. $\text{tr } (\mathbf{A} + \mathbf{B}) = \text{tr } \mathbf{A} + \text{tr } \mathbf{B}$
4. $\text{tr } \mathbf{I}_N = N$
5. $\text{tr } \mathbf{AB} = \text{tr } \mathbf{BA}$

Using these facts, we now continue the development of (8-28), as follows:

$$\begin{aligned} E\{\tilde{\mathcal{L}}'\tilde{\mathcal{L}}\} &= \text{tr} [\mathbf{M}\mathcal{E}\{\mathcal{V}\mathcal{V}'\}] = \text{tr } \mathbf{M}\mathcal{R} = \text{tr } \mathbf{M}\sigma_v^2 \\ &= \sigma_v^2 \text{tr } \mathbf{M} = \sigma_v^2 \text{tr } [\mathbf{I}_k - \mathcal{H}(\mathcal{H}'\mathcal{H})^{-1}\mathcal{H}'] \\ &= \sigma_v^2 k - \sigma_v^2 \text{tr } \mathcal{H}(\mathcal{H}'\mathcal{H})^{-1}\mathcal{H}' \\ &= \sigma_v^2 k - \sigma_v^2 \text{tr } (\mathcal{H}'\mathcal{H})(\mathcal{H}'\mathcal{H})^{-1} \\ &= \sigma_v^2 k - \sigma_v^2 \text{tr } \mathbf{I}_n = \sigma_v^2(k - n) \end{aligned} \quad (8-29)$$

Solving this equation for σ_v^2 , we find that

$$\sigma_v^2 = \frac{E\{\tilde{\mathcal{L}}'\tilde{\mathcal{L}}\}}{(k - n)} \quad (8-30)$$

Although this is an exact result for σ_v^2 , it is not one that can be evaluated, because we cannot compute $E\{\tilde{\mathcal{L}}'\tilde{\mathcal{L}}\}$.

Using the structure of σ_v^2 as a starting point, we estimate σ_v^2 by the simple formula

$$\hat{\sigma}_v^2(k) = \tilde{\mathcal{L}}'(k)\tilde{\mathcal{L}}(k)/(k - n) \quad (8-31)$$

To show that $\hat{\sigma}_v^2(k)$ is an unbiased estimator of σ_v^2 , we observe that

$$\begin{aligned} E\{\hat{\sigma}_v^2(k)\} &= E\{\tilde{\mathcal{L}}'\tilde{\mathcal{L}}\}/(k - n) \\ &= \sigma_v^2 \end{aligned} \quad (8-32)$$

where we have used (8-29) for $E\{\tilde{\mathcal{L}}'\tilde{\mathcal{L}}\}$. \square

LARGE SAMPLE PROPERTIES OF LEAST-SQUARES ESTIMATORS

Many large sample properties of LSE's are determined by establishing that the LSE is equivalent to another estimator for which it is known that the large sample property holds true. In Lesson 11, for example, we will provide condi-

tions under which the LSE of θ , $\hat{\theta}_{LS}(k)$, is the same as the maximum-likelihood estimator of θ , $\hat{\theta}_{ML}(k)$. Because $\hat{\theta}_{ML}(k)$ is consistent, asymptotically efficient, and asymptotically Gaussian, $\hat{\theta}_{LS}(k)$ inherits all these properties.

Theorem 8-4. If

$$\text{plim } [\mathcal{H}'(k)\mathcal{H}(k)/k] = \Sigma_{\mathcal{X}} \quad (8-33)$$

$\Sigma_{\mathcal{X}}^{-1}$ exists, and

$$\text{plim } [\mathcal{H}'(k)\mathcal{V}(k)/k] = \mathbf{0} \quad (8-34)$$

then

$$\text{plim } \hat{\theta}_{LS}(k) = \theta \quad (8-35)$$

Note that the probability limit of a matrix equals a matrix each of whose elements is the probability limit of the respective matrix element. Assumption (8-33) postulates the existence of a probability limit for the second-order moments of the variables in $\mathcal{H}(k)$, as given by $\Sigma_{\mathcal{X}}$. Assumption (8-34) postulates a zero probability limit for the correlation between $\mathcal{H}(k)$ and $\mathcal{V}(k)$. $\mathcal{H}'(k)\mathcal{V}(k)$ can be thought of as a "filtered" version of noise vector $\mathcal{V}(k)$. For (8-34) to be true "filter $\mathcal{H}(k)$ " must be stable. If, for example, $\mathcal{H}(k)$ is deterministic and $\sigma_{\mathcal{V}(k)}^2 < \infty$, then (8-34) will be true.

Proof. Beginning with (8-2), but for $\hat{\theta}_{LS}(k)$ instead of $\hat{\theta}_{WLS}(k)$, we see that

$$\hat{\theta}_{LS}(k) = \theta + (\mathcal{H}'\mathcal{H})^{-1}\mathcal{H}'\mathcal{V} \quad (8-36)$$

Operating on both sides of this equation with plim, and using properties (7-15), (7-16), and (7-17), we find that

$$\begin{aligned} \text{plim } \hat{\theta}_{LS}(k) &= \theta + \text{plim } [(\mathcal{H}'\mathcal{H}/k)^{-1}(\mathcal{H}'\mathcal{V}/k)] \\ &= \theta + \text{plim } (\mathcal{H}'\mathcal{H}/k)^{-1} \text{plim } (\mathcal{H}'\mathcal{V}/k) \\ &= \theta + \Sigma_{\mathcal{X}}^{-1} \cdot \mathbf{0} \\ &= \theta \end{aligned}$$

which demonstrates that, under the given conditions, $\hat{\theta}_{LS}(k)$ is a consistent estimator of θ . \square

In some important applications Eq. (8-34) does not apply, e.g., Example 8-2. Theorem 8-4 then does not apply, and, the study of consistency is often quite complicated in these cases.

Theorem 8-5. If (8-33) and (8-34) are true, $\Sigma_{\mathcal{X}}^{-1}$ exists, and

$$\text{plim } [\mathcal{V}'(k)\mathcal{V}(k)/k] = \sigma_v^2 \quad (8-37)$$

then

$$\text{plim } \hat{\sigma}_v^2(k) = \sigma_v^2 \quad (8-38)$$

where $\hat{\sigma}_v^2(k)$ is given by (8-24).

Proof. From (8-26), we find that

$$\tilde{\mathcal{Z}}'\tilde{\mathcal{Z}} = \mathcal{V}'\mathcal{V} - \mathcal{V}'\mathcal{H}(\mathcal{H}'\mathcal{H})^{-1}\mathcal{H}'\mathcal{V} \quad (8-39)$$

Consequently,

$$\begin{aligned} \text{plim } \hat{\sigma}_v^2(k) &= \text{plim } \tilde{\mathcal{Z}}'\tilde{\mathcal{Z}}/(k-n) \\ &= \text{plim } \mathcal{V}'\mathcal{V}/(k-n) - \text{plim } \mathcal{V}'\mathcal{H}(\mathcal{H}'\mathcal{H})^{-1}\mathcal{H}'\mathcal{V}/(k-n) \\ &= \sigma_v^2 - \text{plim } \mathcal{V}'\mathcal{H}/(k-n) \cdot \text{plim } [\mathcal{H}'\mathcal{H}/(k-n)]^{-1} \\ &\quad \cdot \text{plim } \mathcal{H}'\mathcal{V}/(k-n) \\ &= \sigma_v^2 - \mathbf{0}' \cdot \Sigma_{\mathcal{H}}^{-1} \cdot \mathbf{0} \\ &= \sigma_v^2 \quad \square \end{aligned}$$

PROBLEMS

- 8-1. Suppose that $\hat{\theta}_{LS}$ is an unbiased estimator of θ . Is $\hat{\theta}_{LS}^2$ an unbiased estimator of θ^2 ? (Hint: Use the least-squares batch algorithm to study this question.)
- 8-2. For $\hat{\theta}_{WLS}(k)$ to be an unbiased estimator of θ we required $E\{\mathcal{V}(k)\} = \mathbf{0}$. This problem considers the case when $E\{\mathcal{V}(k)\} \neq \mathbf{0}$.
 - (a) Assume that $E\{\mathcal{V}(k)\} = \mathcal{V}_0$ where \mathcal{V}_0 is known to us. How is the concatenated measurement equation $\mathcal{Z}(k) = \mathcal{H}(k)\theta + \mathcal{V}(k)$ modified in this case so we can use the results derived in this lesson to obtain $\hat{\theta}_{WLS}(k)$ or $\hat{\theta}_{LS}(k)$?
 - (b) Assume that $E\{\mathcal{V}(k)\} = m_r$ where m_r is constant but is unknown. How is the concatenated measurement equation $\mathcal{Z}(k) = \mathcal{H}(k)\theta + \mathcal{V}(k)$ modified in this case so that we can obtain least-squares estimates of both θ and m_r ?
- 8-3. Consider the stable autoregressive model $y(k) = \theta_1 y(k-1) + \dots + \theta_K y(k-K) + \epsilon(k)$ in which the $\epsilon(k)$ are identically distributed random variables with mean zero and finite variance σ^2 . Prove that the least-squares estimates of $\theta_1, \dots, \theta_K$ are consistent (see also, Ljung, 1976).
- 8-4. In this lesson we have assumed that the $\mathcal{H}(k)$ variables have been measured without error. Here we examine the situation when $\mathcal{H}_m(k) = \mathcal{H}(k) + \mathcal{N}(k)$ in which $\mathcal{H}(k)$ denotes a matrix of true values and $\mathcal{N}(k)$ a matrix of measurement errors. The basic linear model is now

$$\mathcal{Z}(k) = \mathcal{H}(k)\theta + \mathcal{V}(k) = \mathcal{H}_m(k)\theta + [\mathcal{V}(k) - \mathcal{N}(k)\theta].$$

Prove that $\hat{\theta}_{LS}(k)$ is not a consistent estimator of θ .

Lesson 9

Best Linear Unbiased Estimation

INTRODUCTION

Least-squares estimation, as described in Lessons 3, 4 and 5, is for the linear model

$$\mathcal{Z}(k) = \mathcal{H}(k)\theta + \mathcal{V}(k) \quad (9-1)$$

where θ is a deterministic, but unknown vector of parameters, $\mathcal{H}(k)$ can be deterministic or random, and we do not know anything about $\mathcal{V}(k)$ ahead of time. By minimizing $\tilde{\mathcal{Z}}'(k)\tilde{\mathcal{Z}}(k)$, where $\tilde{\mathcal{Z}}(k) = \mathcal{Z}(k) - \mathcal{H}(k)\hat{\theta}_{WLS}(k)$, we determined that $\hat{\theta}_{WLS}(k)$ is a linear transformation of $\mathcal{Z}(k)$, i.e., $\hat{\theta}_{WLS}(k) = \mathbf{F}_{WLS}(k)\mathcal{Z}(k)$. After establishing the structure of $\hat{\theta}_{WLS}(k)$, we studied its small- and large-sample properties. Unfortunately, $\hat{\theta}_{WLS}(k)$ is not always unbiased or efficient. These properties were not built into $\hat{\theta}_{WLS}(k)$ during its design.

In this lesson we develop our second estimator. It will be both unbiased and efficient, *by design*. In addition, we want the estimator to be a linear function of the measurements $\mathcal{Z}(k)$. This estimator is called a best linear unbiased estimator (BLUE) or an unbiased minimum-variance estimator (UMVE). To keep notation relatively simple, we will use $\hat{\theta}_{BLU}(k)$ to denote the BLUE of θ .

As in least-squares, we begin with the linear model in (9-1), where θ is deterministic. Now, however, $\mathcal{H}(k)$ must be deterministic and $\mathcal{V}(k)$ is assumed to be zero mean with positive definite known covariance matrix $\mathcal{R}(k)$. An example of such a covariance matrix occurs for white noise. In the case of scalar measurements, $z(k)$, this means that scalar noise $v(k)$ is white, i.e.,

$$E\{v(k)v(j)\} = \sigma_v^2(k)\delta_{kj} \quad (9-2)$$

where δ_{kj} is the Kronecker δ (i.e., $\delta_{kj} = 0$ for $k \neq j$ and $\delta_{kj} = 1$ for $k = j$) thus,

$$\begin{aligned}\mathcal{R}(k) &= \mathbf{E}\{\mathbf{V}(k)\mathbf{V}'(k)\} \\ &= \text{diag}[\sigma_v^2(k), \sigma_v^2(k-1), \dots, \sigma_v^2(k-N+1)]\end{aligned}\quad (9-3)$$

In the case of vector measurements, $\mathbf{z}(k)$, this means that vector noise, $\mathbf{v}(k)$, is white, i.e.,

$$\mathbf{E}\{\mathbf{v}(k)\mathbf{v}'(j)\} = \mathbf{R}(k)\delta_{kj} \quad (9-4)$$

thus,

$$\mathcal{R}(k) = \text{diag}[\mathbf{R}(k), \mathbf{R}(k-1), \dots, \mathbf{R}(k-N+1)] \quad (9-5)$$

PROBLEM STATEMENT AND OBJECTIVE FUNCTION

We begin by assuming the following linear structure for $\hat{\theta}_{BLU}(k)$,

$$\hat{\theta}_{BLU}(k) = \mathbf{F}(k)\mathcal{Z}(k) \quad (9-6)$$

where, for notational simplicity, we have omitted subscripting $\mathbf{F}(k)$ as $\mathbf{F}_{BLU}(k)$. We shall design $\mathbf{F}(k)$ such that

- a. $\hat{\theta}_{BLU}(k)$ is an unbiased estimator of θ , and
- b. the error variance for each one of the n parameters is minimized. In this way, $\hat{\theta}_{BLU}(k)$ will be unbiased and efficient, by design.

Recall, from Theorem 6-1, that *unbiasedness constrains design matrix $\mathbf{F}(k)$, such that*

$$\mathbf{F}(k)\mathcal{H}(k) = \mathbf{I} \quad \text{for all } k \quad (9-7)$$

Our objective now is to choose the elements of $\mathbf{F}(k)$, subject to the constraint of (9-7), in such a way that the error variance for each one of the n parameters is minimized.

In solving for $\mathbf{F}_{BLU}(k)$, it will be convenient to partition matrix $\mathbf{F}(k)$, as

$$\mathbf{F}(k) = \begin{pmatrix} \mathbf{f}_1'(k) \\ \mathbf{f}_2'(k) \\ \vdots \\ \mathbf{f}_n'(k) \end{pmatrix} \quad (9-8)$$

Equation (9-7) can now be expressed in terms of the vector components of $\mathbf{F}(k)$. For our purposes, it is easier to work with the transpose of (9-7), $\mathcal{H}'\mathbf{F}' = \mathbf{I}$, which can be expressed as

$$\mathcal{H}'[\mathbf{f}_1 | \mathbf{f}_2 | \dots | \mathbf{f}_n] = [\mathbf{e}_1 | \mathbf{e}_2 | \dots | \mathbf{e}_n] \quad (9-9)$$

Derivation of Estimator

where \mathbf{e}_i is the i th unit vector,

$$\mathbf{e}_i = \text{col}(0, 0, \dots, 0, 1, 0, \dots, 0) \quad (9-10)$$

in which the nonzero element occurs in the i th position. Equating respective elements on both sides of (9-9), we find that

$$\mathcal{H}'(k)\mathbf{f}_i(k) = \mathbf{e}_i \quad i = 1, 2, \dots, n \quad (9-11)$$

Our single unbiasedness constraint on matrix $\mathbf{F}(k)$ is now a set of n constraints on the rows of $\mathbf{F}(k)$.

Next, we express $\mathbf{E}\{[\theta_i - \hat{\theta}_{i,BLU}(k)]^2\}$ in terms of \mathbf{f}_i ($i = 1, 2, \dots, N$). We shall make use of (9-11), (9-1), and the following equivalent representation of (9-6)

$$\hat{\theta}_{i,BLU}(k) = \mathbf{f}_i'(k)\mathcal{Z}(k) \quad i = 1, 2, \dots, n \quad (9-12)$$

Proceeding, we find that

$$\begin{aligned}\mathbf{E}\{[\theta_i - \hat{\theta}_{i,BLU}(k)]^2\} &= \mathbf{E}\{(\theta_i - \mathbf{f}_i'\mathcal{Z})^2\} = \mathbf{E}\{(\theta_i - \mathcal{Z}'\mathbf{f}_i)^2\} \\ &= \mathbf{E}\{\theta_i^2 - 2\theta_i\mathcal{Z}'\mathbf{f}_i + (\mathcal{Z}'\mathbf{f}_i)^2\} \\ &= \mathbf{E}\{\theta_i^2 - 2\theta_i(\mathcal{H}\theta + \mathcal{V}')\mathbf{f}_i + [(\mathcal{H}\theta + \mathcal{V}')\mathbf{f}_i]^2\} \\ &= \mathbf{E}\{\theta_i^2 - 2\theta_i\mathcal{H}'\mathbf{f}_i - 2\theta_i\mathcal{V}'\mathbf{f}_i + [\theta_i'\mathcal{H}'\mathbf{f}_i + \mathcal{V}'\mathbf{f}_i]^2\} \\ &= \mathbf{E}\{\theta_i^2 - 2\theta_i\theta'\mathbf{e}_i - 2\theta_i\mathcal{V}'\mathbf{f}_i + [\theta_i'\mathbf{e}_i + \mathcal{V}'\mathbf{f}_i]^2\} \\ &= \mathbf{E}\{\mathbf{f}_i'\mathcal{V}\mathcal{V}'\mathbf{f}_i\} = \mathbf{f}_i'\mathcal{R}\mathbf{f}_i\end{aligned}\quad (9-13)$$

Observe that the error-variance for the i th parameter depends only on the i th row of design matrix $\mathbf{F}(k)$. We, therefore, establish the following objective function:

$$\begin{aligned}J_i(\mathbf{f}_i, \lambda_i) &= \mathbf{E}\{[\theta_i - \hat{\theta}_{i,BLU}(k)]^2\} + \lambda_i'(\mathcal{H}'\mathbf{f}_i - \mathbf{e}_i) \\ &= \mathbf{f}_i'\mathcal{R}\mathbf{f}_i + \lambda_i'(\mathcal{H}'\mathbf{f}_i - \mathbf{e}_i)\end{aligned}\quad (9-14)$$

where λ_i is the i th vector of the Lagrange multipliers, that is associated with the i th unbiasedness constraint. Our objective now is to minimize J_i with respect to \mathbf{f}_i and λ_i ($i = 1, 2, \dots, N$).

DERIVATION OF ESTIMATOR

A necessary condition for minimizing $J_i(\mathbf{f}_i, \lambda_i)$ is $\partial J_i(\mathbf{f}_i, \lambda_i) / \partial \mathbf{f}_i = \mathbf{0}$ ($i = 1, 2, \dots, n$); hence,

$$2\mathcal{R}\mathbf{f}_i + \mathcal{H}\lambda_i = \mathbf{0} \quad (9-15)$$

from which we determine \mathbf{f}_i , as

$$\mathbf{f}_i = -\frac{1}{2}\mathcal{R}^{-1}\mathcal{H}\lambda_i \quad (9-16)$$

For (9-16) to be valid, \mathcal{R}^{-1} must exist. Any noise $\mathcal{V}(k)$ whose covariance matrix \mathcal{R} is positive definite qualifies. Of course, if $\mathcal{V}(k)$ is white, then \mathcal{R} is diagonal (or block diagonal) and \mathcal{R}^{-1} exists. This may also be true if $\mathcal{V}(k)$ is not white. A second necessary condition for minimizing $J_i(\mathbf{f}_i, \lambda_i)$ is $\partial J_i(\mathbf{f}_i, \lambda_i) / \partial \lambda_i = \mathbf{0}$ ($i = 1, 2, \dots, n$), which gives us the unbiasedness constraints

$$\mathcal{H}'\mathbf{f}_i = \mathbf{e}_i \quad i = 1, 2, \dots, n \quad (9-17)$$

To determine λ_i , substitute (9-16) into (9-17). Doing this, we find that

$$\lambda_i = -2(\mathcal{H}'\mathcal{R}^{-1}\mathcal{H})^{-1}\mathbf{e}_i \quad (9-18)$$

whereupon

$$\mathbf{f}_i = \mathcal{R}^{-1}\mathcal{H}(\mathcal{H}'\mathcal{R}^{-1}\mathcal{H})^{-1}\mathbf{e}_i \quad (9-19)$$

($i = 1, 2, \dots, n$). Matrix $\mathbf{F}(k)$ is reconstructed from $\mathbf{f}_i(k)$, as follows:

$$\begin{aligned} \mathbf{F}'(k) &= (\mathbf{f}_1|\mathbf{f}_2| \dots |\mathbf{f}_n) \\ &= \mathcal{R}^{-1}\mathcal{H}(\mathcal{H}'\mathcal{R}^{-1}\mathcal{H})^{-1}(\mathbf{e}_1|\mathbf{e}_2| \dots |\mathbf{e}_n) \\ &= \mathcal{R}^{-1}\mathcal{H}(\mathcal{H}'\mathcal{R}^{-1}\mathcal{H})^{-1} \end{aligned} \quad (9-20)$$

Hence

$$\mathbf{F}_{BLU}(k) = [\mathcal{H}'(k)\mathcal{R}^{-1}(k)\mathcal{H}(k)]^{-1}\mathcal{H}'(k)\mathcal{R}^{-1}(k) \quad (9-21)$$

which means that

$$\hat{\theta}_{BLU}(k) = [\mathcal{H}'(k)\mathcal{R}^{-1}(k)\mathcal{H}(k)]^{-1}\mathcal{H}'(k)\mathcal{R}^{-1}(k)\mathcal{Z}(k) \quad (9-22)$$

COMPARISON OF $\hat{\theta}_{BLU}(k)$ AND $\hat{\theta}_{WLS}(k)$

We are struck by the close similarity between $\hat{\theta}_{BLU}(k)$ and $\hat{\theta}_{WLS}(k)$.

Theorem 9-1. *The BLUE of θ is the special case of the WLSE of θ when*

$$\mathbf{W}(k) = \mathcal{R}^{-1}(k) \quad (9-23)$$

If $\mathbf{W}(k)$ is diagonal, then (9-23) requires $\mathcal{V}(k)$ to be white.

Proof. Compare the formulas for $\hat{\theta}_{BLU}(k)$ in (9-22) and $\hat{\theta}_{WLS}(k)$ in (10) of Lesson 3. If $\mathbf{W}(k)$ is a diagonal matrix, then $\mathcal{R}(k)$ is a diagonal matrix only if $\mathcal{V}(k)$ is white. \square

Matrix $\mathcal{R}^{-1}(k)$ weights the contributions of precise measurements heavily and deemphasizes the contributions of imprecise measurements. The best linear unbiased estimation design technique has led to a weighting matrix that is quite sensible. See Problem 9-2.

Corollary 9-1. *All results obtained in Lessons 3, 4, and 5 for $\hat{\theta}_{WLS}(k)$ can be applied to $\hat{\theta}_{BLU}(k)$ by setting $\mathbf{W}(k) = \mathcal{R}^{-1}(k)$. \square*

We leave it to the reader to explore the full implications of this important corollary, by reexamining the wide range of topics, which were discussed in Lessons 3, 4 and 5.

Theorem 9-2 (Gauss-Markov Theorem). *If $\mathcal{R}(k) = \sigma_v^2 \mathbf{I}$, then $\hat{\theta}_{BLU}(k) = \hat{\theta}_{LS}(k)$.*

Proof. Using (9-22) and the fact that $\mathcal{R}(k) = \sigma_v^2 \mathbf{I}$, we find that

$$\hat{\theta}_{BLU}(k) = \left(\frac{1}{\sigma_v^2} \mathcal{H}'\mathcal{H}\right)^{-1} \mathcal{H}' \frac{1}{\sigma_v^2} \mathcal{Z} = (\mathcal{H}'\mathcal{H})^{-1}\mathcal{H}'\mathcal{Z} = \hat{\theta}_{LS}(k) \quad \square$$

Why is this a very important result? We have connected two seemingly different estimators, one of which— $\hat{\theta}_{BLU}(k)$ —has the properties of unbiased and minimum variance by design; hence, in this case $\hat{\theta}_{LS}(k)$ inherits these properties. Remember though that the derivation of $\hat{\theta}_{BLU}(k)$ required $\mathcal{H}(k)$ to be deterministic; thus, *Theorem 9-2 is “conditioned” on $\mathcal{H}(k)$ being deterministic*.

SOME PROPERTIES OF $\hat{\theta}_{BLU}(k)$

To begin, we direct our attention at the covariance matrix of parameter estimation error $\tilde{\theta}_{BLU}(k)$.

Theorem 9-3. *If $\mathcal{V}(k)$ is zero mean, then*

$$\text{cov}[\tilde{\theta}_{BLU}(k)] = [\mathcal{H}'(k)\mathcal{R}^{-1}(k)\mathcal{H}(k)]^{-1} \quad (9-24)$$

Proof. We apply Corollary 9-1 to $\text{cov}[\tilde{\theta}_{WLS}(k)]$ [given in (18) of Lesson 8] for the case when $\mathcal{H}(k)$ is deterministic, to see that

$$\begin{aligned} \text{cov}[\tilde{\theta}_{BLU}(k)] &= \text{cov}[\tilde{\theta}_{WLS}(k)]|_{\mathbf{W}(k) = \mathcal{R}^{-1}(k)} \\ &= (\mathcal{H}'\mathcal{R}^{-1}\mathcal{H})^{-1}\mathcal{H}'\mathcal{R}^{-1}\mathcal{R}\mathcal{R}^{-1}\mathcal{H}(\mathcal{H}'\mathcal{R}^{-1}\mathcal{H})^{-1} \\ &= (\mathcal{H}'\mathcal{R}^{-1}\mathcal{H})^{-1} \quad \square \end{aligned}$$

Observe the great simplification of the expression for $\text{cov}[\tilde{\theta}_{WLS}(k)]$, when $\mathbf{W}(k) = \mathcal{R}^{-1}(k)$. Note, also, that the error variance of $\hat{\theta}_{i,BLU}(k)$ is given by the i th diagonal element of $\text{cov}[\tilde{\theta}_{BLU}(k)]$.

Corollary 9-2. *When $\mathbf{W}(k) = \mathcal{R}^{-1}(k)$ then matrix $\mathbf{P}(k)$, which appears in the recursive WLSE of θ equals $\text{cov}[\tilde{\theta}_{BLU}(k)]$, i.e.,*

$$\mathbf{P}(k) = \text{cov}[\tilde{\theta}_{BLU}(k)] \quad (9-25)$$

Proof. Recall Equation (13) of Lesson 4, that

$$\mathbf{P}(k) = [\mathcal{H}'(k)\mathbf{W}(k)\mathcal{H}(k)]^{-1} \quad (9-26)$$

When $\mathbf{W}(k) = \mathcal{R}^{-1}(k)$, then

$$\mathbf{P}(k) = [\mathcal{H}'(k)\mathcal{R}^{-1}(k)\mathcal{H}(k)]^{-1} \quad (9-27)$$

hence, $\mathbf{P}(k) = \text{cov}[\hat{\theta}_{\text{BLU}}(k)]$ because of (9-24). \square

Soon we will examine a recursive BLUE. Matrix $\mathbf{P}(k)$ will have to be calculated, just as it has to be calculated for the recursive WLSE. Every time $\mathbf{P}(k)$ is calculated in our recursive BLUE, we obtain a quantitative measure of how well we are estimating θ . Just look at the diagonal elements of $\mathbf{P}(k)$, $k = 1, 2, \dots$. The same statement cannot be made for the meaning of $\mathbf{P}(k)$ in the recursive WLSE. In the recursive WLSE, $\mathbf{P}(k)$ has no special meaning.

Next, we examine $\text{cov}[\hat{\theta}_{\text{BLU}}(k)]$ in more detail.

Theorem 9-4. $\hat{\theta}_{\text{BLU}}(k)$ is a most efficient estimator of θ within the class of all unbiased estimators that are linearly related to the measurements $\mathfrak{X}(k)$.

Proof (Mendel, 1973, pp. 155–156). According to Definition 6-3, we must show that

$$\Sigma = \text{cov}[\hat{\theta}_a(k)] - \text{cov}[\hat{\theta}_{\text{BLU}}(k)] \quad (9-28)$$

is positive semidefinite. In (9-28), $\hat{\theta}_a(k)$ is the error associated with an arbitrary linear unbiased estimate of θ . For convenience, we write Σ as

$$\Sigma = \Sigma_a - \Sigma_{\text{BLU}} \quad (9-29)$$

In order to compute Σ_a we use the facts that

$$\hat{\theta}_a(k) = \mathbf{F}_a(k)\mathfrak{X}(k) \quad (9-30)$$

and

$$\mathbf{F}_a(k)\mathcal{H}(k) = \mathbf{I} \quad (9-31)$$

thus,

$$\begin{aligned} \Sigma_a &= E\{(\theta - \mathbf{F}_a\mathfrak{X})(\theta - \mathbf{F}_a\mathfrak{X})'\} \\ &= E\{(\theta - \mathbf{F}_a\mathcal{H}\theta - \mathbf{F}_a\mathcal{V})(\theta - \mathbf{F}_a\mathcal{H}\theta - \mathbf{F}_a\mathcal{V})'\} \\ &= E\{(\mathbf{F}_a\mathcal{V})(\mathbf{F}_a\mathcal{V})'\} \\ &= \mathbf{F}_a\mathcal{R}\mathbf{F}_a' \end{aligned} \quad (9-32)$$

Because $\hat{\theta}_{\text{BLU}}(k) = \mathbf{F}(k)\mathfrak{X}(k)$ and $\mathbf{F}(k)\mathcal{H}(k) = \mathbf{I}$,

$$\Sigma_{\text{BLU}} = \mathbf{F}\mathcal{R}\mathbf{F}' \quad (9-33)$$

Some Properties of $\hat{\theta}_{\text{BLU}}(k)$

Substituting (9-32) and (9-33) into (9-29), and making repeated use of the unbiasedness constraints $\mathcal{H}\mathbf{F}' = \mathcal{H}\mathbf{F}_a' = \mathbf{F}_a\mathcal{H}_a = \mathbf{I}$, we find that

$$\begin{aligned} \Sigma &= \mathbf{F}_a\mathcal{R}\mathbf{F}_a' - \mathbf{F}\mathcal{R}\mathbf{F}' \\ &= \mathbf{F}_a\mathcal{R}\mathbf{F}_a' - \mathbf{F}\mathcal{R}\mathbf{F}' + 2(\mathcal{H}\mathcal{R}^{-1}\mathcal{H})^{-1} - (\mathcal{H}\mathcal{R}^{-1}\mathcal{H})^{-1} \\ &\quad - (\mathcal{H}\mathcal{R}^{-1}\mathcal{H})^{-1} \\ &= \mathbf{F}_a\mathcal{R}\mathbf{F}_a' - \mathbf{F}\mathcal{R}\mathbf{F}' + 2(\mathcal{H}\mathcal{R}^{-1}\mathcal{H})^{-1}(\mathcal{H}\mathcal{R}^{-1}\mathcal{R}\mathbf{F}') \\ &\quad - (\mathcal{H}\mathcal{R}^{-1}\mathcal{H})^{-1}(\mathcal{H}\mathcal{R}^{-1}\mathcal{R}\mathbf{F}_a') \\ &\quad - (\mathbf{F}_a\mathcal{R}\mathcal{R}^{-1}\mathcal{H})(\mathcal{H}\mathcal{R}^{-1}\mathcal{H})^{-1} \end{aligned} \quad (9-34)$$

Making use of the structure of $\mathbf{F}(k)$, given in (9-21), we see that Σ can also be written as

$$\begin{aligned} \Sigma &= \mathbf{F}_a\mathcal{R}\mathbf{F}_a' - \mathbf{F}\mathcal{R}\mathbf{F}' + 2\mathcal{R}\mathbf{F}' - \mathbf{F}\mathcal{R}\mathbf{F}_a' - \mathbf{F}_a\mathcal{R}\mathbf{F}' \\ &= (\mathbf{F}_a - \mathbf{F})\mathcal{R}(\mathbf{F}_a - \mathbf{F})' \end{aligned} \quad (9-35)$$

In order to investigate the definiteness of Σ , consider the definiteness of $\mathbf{a}'\Sigma\mathbf{a}$, where \mathbf{a} is an arbitrary nonzero vector,

$$\mathbf{a}'\Sigma\mathbf{a} = [(\mathbf{F}_a - \mathbf{F})'\mathbf{a}]\mathcal{R}[(\mathbf{F}_a - \mathbf{F})'\mathbf{a}] \quad (9-36)$$

Matrix \mathbf{F} (i.e., \mathbf{F}_{BLU}) is unique; therefore $(\mathbf{F}_a - \mathbf{F})'\mathbf{a}$ is a nonzero vector, unless $\mathbf{F}_a - \mathbf{F}$ and \mathbf{a} are orthogonal, which is a possibility that cannot be excluded. Because matrix \mathcal{R} is positive definite, $\mathbf{a}'\Sigma\mathbf{a} \geq 0$, which means that Σ is positive semidefinite. \square

These results serve as further confirmation that designing $\mathbf{F}(k)$ as we have done, by minimizing only the diagonal elements of $\text{cov}[\hat{\theta}_{\text{BLU}}(k)]$, is sound.

Corollary 9-3. If $\mathcal{R}(k) = \Sigma_a^2 \mathbf{I}$, then $\hat{\theta}_{\text{LS}}(k)$ is a most efficient estimator of θ .

The proof of this result is a direct consequence of Theorems 9-2 and 9-4. \square

At the end of Lesson 3 we noted that $\hat{\theta}_{\text{WLS}}(k)$ may not be invariant under scale changes. We demonstrate next that $\hat{\theta}_{\text{BLU}}(k)$ is invariant to such changes.

Theorem 9-5. $\hat{\theta}_{\text{BLU}}(k)$ is invariant under changes of scale.

Proof (Mendel, 1973, pp. 156–157). Assume that observers A and B are observing a process; but, observer A reads the measurements in one set of units and B in another. Let \mathbf{M} be a symmetric matrix of scale factors relating A to B (e.g., 5,280 ft/mile, 454 g/lb, etc.), and $\mathfrak{X}_A(k)$ and $\mathfrak{X}_B(k)$ denote the total measurement vectors of A and B , respectively. Then

$$\mathfrak{X}_B(k) = \mathcal{H}_B(k)\theta + \mathcal{V}_B(k) = \mathbf{M}\mathfrak{X}_A(k) = \mathbf{M}\mathcal{H}_A(k)\theta + \mathbf{M}\mathcal{V}_A(k) \quad (9-37)$$

which means that

$$\mathcal{H}_B(k) = M\mathcal{H}_A(k) \quad (9-38)$$

$$\mathcal{V}_B(k) = MV_A(k) \quad (9-39)$$

and

$$\mathcal{R}_B(k) = M\mathcal{R}_A(k)M' = M\mathcal{R}_A(k)M \quad (9-40)$$

Let $\hat{\theta}_{A,\text{BLU}}(k)$ and $\hat{\theta}_{B,\text{BLU}}(k)$ denote the BLUE's associated with observers A and B, respectively; then,

$$\begin{aligned} \hat{\theta}_{B,\text{BLU}}(k) &= (\mathcal{H}'_B\mathcal{R}_B^{-1}\mathcal{H}_B)^{-1}\mathcal{H}'_B\mathcal{R}_B^{-1}\mathcal{L}_B \\ &= [\mathcal{H}_A M(M\mathcal{R}_A M)^{-1}M\mathcal{H}_A]^{-1}\mathcal{H}_A M(M\mathcal{R}_A M)^{-1}M\mathcal{L}_A \\ &= (\mathcal{H}_A M M^{-1}\mathcal{R}_A^{-1}M^{-1}M\mathcal{H}_A)^{-1}\mathcal{H}_A M M^{-1}\mathcal{R}_A^{-1}M^{-1}M\mathcal{L}_A \quad (9-41) \\ &= (\mathcal{H}_A\mathcal{R}_A^{-1}\mathcal{H}_A)^{-1}\mathcal{H}_A\mathcal{R}_A^{-1}\mathcal{L}_A \\ &= \hat{\theta}_{A,\text{BLU}}(k) \quad \square \end{aligned}$$

RECURSIVE BLUES

Because of Corollary 9-1, we obtain recursive formulas for the BLUE of θ by setting $1/w(k+1) = r(k+1)$ in the recursive formulas for the WLSEs of θ which are given in Lesson 4. In the case of a vector of measurements, we set (see Table 5-1) $w^{-1}(k+1) = R(k+1)$.

Theorem 9-6 (Information Form of Recursive BLUE). *A recursive structure for $\hat{\theta}_{\text{BLU}}(k)$ is:*

$$\hat{\theta}_{\text{BLU}}(k+1) = \hat{\theta}_{\text{BLU}}(k) + K_B(k+1)[z(k+1) - h'(k+1)\hat{\theta}_{\text{BLU}}(k)] \quad (9-42)$$

where

$$K_B(k+1) = P(k+1)h(k+1)r^{-1}(k+1) \quad (9-43)$$

and

$$P^{-1}(k+1) = P^{-1}(k) + h(k+1)r^{-1}(k+1)h'(k+1) \quad (9-44)$$

These equations are initialized by $\hat{\theta}_{\text{BLU}}(n)$ and $P^{-1}(n)$ [where $P(k)$ is $\text{cov}[\hat{\theta}_{\text{BLU}}(k)]$, given in (9-31)] and, are used for $k = n, n+1, \dots, N-1$. These equations can also be used for $k = 0, 1, \dots, N-1$ as long as $\hat{\theta}_{\text{BLU}}(0)$ and $P^{-1}(0)$ are chosen using Equations (21) and (20) in Lesson 4, respectively, in which $w(0)$ is replaced by $r^{-1}(0)$. \square

Theorem 9-7 (Covariance Form of Recursive BLUE). *Another recursive structure for $\hat{\theta}_{\text{BLU}}(k)$ is (9-42) in which*

$$K_B(k+1) = P(k)h(k+1)[h'(k+1)P(k)h(k+1) + r(k+1)]^{-1} \quad (9-45)$$

and

$$P(k+1) = [I - K_B(k+1)h'(k+1)]P(k) \quad (9-46)$$

These equations are initialized by $\hat{\theta}_{\text{BLU}}(n)$ and $P(n)$, and are used for $k = n, n+1, \dots, N-1$. They can also be used for $k = 0, 1, \dots, N-1$ as long as $\hat{\theta}_{\text{BLU}}(0)$ and $P(0)$ are chosen using Equations (4-21) and (4-20), respectively, in which $w(0)$ is replaced by $r^{-1}(0)$. \square

Recall that, in best-linear unbiased estimation, $P(k) = \text{cov}[\hat{\theta}_{\text{BLU}}(k)]$. Observe, in Theorem 9-7, that we compute $P(k)$ recursively, and not $P^{-1}(k)$. This is why the results in Theorem 9-7 (and, subsequently, Theorem 4-2) are referred to as the *covariance form* of recursive BLUE.

PROBLEMS

9-1. (Mendel, 1973, Exercise 3-2, pg. 175). Assume $\mathcal{H}(k)$ is random, and that $\hat{\theta}_{\text{BLU}}(k) = F(k)\mathcal{L}(k)$.

- (a) Show that unbiasedness of the estimate is attained when $E\{F(k)\mathcal{H}(k)\} = I$.
- (b) At what point in the derivation of $\hat{\theta}_{\text{BLU}}(k)$ do the computations break down because $\mathcal{H}(k)$ is random?

9-2. Here we examine the situation when $V(k)$ is colored noise and how to use a model to compute $R(k)$. Now our linear model is

$$z(k+1) = \mathcal{H}(k+1)\theta + v(k+1)$$

where $v(k)$ is colored noise modeled as

$$v(k+1) = A_v v(k) + \xi(k)$$

We assume that deterministic matrix A_v is known and that $\xi(k)$ is zero-mean white noise with covariance $R_\xi(k)$. Working with the *measurement difference* $z^*(k+1) = z(k+1) - A_v z(k)$ write down the formula for $\hat{\theta}_{\text{BLU}}(k)$ in batch form. Be sure to define all concatenated quantities.

9-3. (Sorenson, 1980, Exercise 3-15, pg. 130). Suppose $\hat{\theta}_1$ and $\hat{\theta}_2$ are unbiased estimators of θ with $\text{var}(\hat{\theta}_1) = \sigma_1^2$ and $\text{var}(\hat{\theta}_2) = \sigma_2^2$. Let $\hat{\theta}_3 = \alpha\hat{\theta}_1 + (1-\alpha)\hat{\theta}_2$.

- (a) Prove that $\hat{\theta}_3$ is unbiased.
- (b) Assume that $\hat{\theta}_1$ and $\hat{\theta}_2$ are statistically independent, and find the mean-squared error of $\hat{\theta}_3$.
- (c) What choice of α minimizes the mean-squared error?

9-4. (Mendel, 1973, Exercise 3-12, pp. 176-177). A series of measurements $z(k)$ are made, where $z(k) = H\theta + v(k)$. H is an $m \times n$ constant matrix, $E\{v(k)\} = 0$, and $\text{cov}[v(k)] = R$ is a constant matrix.

- (a) Using the two formulations of the recursive BLUE show that (Ho, 1963, pp. 152-154):

- (i) $P(k+1)H' = P(k)H'[HP(k)H' + R]^{-1}R$, and
- (ii) $HP(k) = R[HP(k-1)H' + R]^{-1}HP(k-1)$.

- (b) Next, show that
- $\mathbf{P}(k)\mathbf{H}' = \mathbf{P}(k-2)\mathbf{H}'[2\mathbf{H}\mathbf{P}(k-2)\mathbf{H}' + \mathbf{R}]^{-1}\mathbf{R}$;
 - $\mathbf{P}(k)\mathbf{H}' = \mathbf{P}(k-3)\mathbf{H}'[3\mathbf{H}\mathbf{P}(k-3)\mathbf{H}' + \mathbf{R}]^{-1}\mathbf{R}$; and
 - $\mathbf{P}(k)\mathbf{H}' = \mathbf{P}(0)\mathbf{H}'[k\mathbf{H}\mathbf{P}(0)\mathbf{H}' + \mathbf{R}]^{-1}\mathbf{R}$.
- (c) Finally, show that the asymptotic form ($k \rightarrow \infty$) for the BLUE of θ is (Ho, 1963, pp. 152–154)

$$\hat{\theta}_{BLU}(k+1) = \hat{\theta}_{BLU}(k) + \frac{1}{(k+1)} \mathbf{P}(0)\mathbf{H}'[\mathbf{H}\mathbf{P}(0)\mathbf{H}']^{-1}[\mathbf{z}(k+1) - \mathbf{H}(k+1)\hat{\theta}_{BLU}(k)]$$

This equation, with its $1/(k+1)$ weighting function, represents a form of multidimensional stochastic approximation.

Lesson 10

Likelihood

INTRODUCTION

This lesson provides background material for the method of maximum-likelihood. It explains the relationship of *likelihood* to *probability*, and when the terms *likelihood* and *likelihood ratio* can be used interchangeably. The major reference for this lesson is Edwards (1972), a most delightful book.

LIKELIHOOD DEFINED

To begin, we define what is meant by an hypothesis, H , and results (of an experiment), R . Suppose scalar parameter θ can assume only two values, 0 or 1; then, we say that there are two hypotheses associated with θ , namely H_0 and H_1 , where, for H_0 , $\theta = 0$, and for H_1 , $\theta = 1$. This is the situation of a *binary hypothesis*. Suppose next that scalar parameter θ can assume ten values, $a, b, c, d, e, f, g, h, i, j$; then, we say there are ten hypotheses associated with θ , namely $H_1, H_2, H_3, \dots, H_{10}$, where, for H_1 , $\theta = a$, for H_2 , $\theta = b, \dots$, and for H_{10} , $\theta = j$. Parameter θ may also assume values from an interval, i.e., $a \leq \theta \leq b$. In this case, we have an infinite, uncountable number of hypotheses about θ , each one associated with a real number in the interval $[a, b]$. Finally, we may have a vector of parameters each one of whose elements has a collection of hypotheses associated with it. For example, suppose that each one of the n elements of θ is either 0 or 1. Vector θ is then characterized by 2^n hypotheses.

Results, R , are the outputs of an experiment. In our work on parameter estimation for the linear model $\mathcal{L}(k) = \mathcal{H}(k)\theta + \mathcal{V}(k)$, the “results” are the data in $\mathcal{L}(k)$ and $\mathcal{H}(k)$.

We let $P(R|H)$ denote the probability of obtaining results R given hypothesis H according to some probability model, e.g., $p[z(k)|\theta]$. In probability, $P(R|H)$ is always viewed as a function of R for *fixed* values of H . Usually the explicit dependence of P on H is not shown. In order to understand the differences between probability and likelihood, it is important to show the explicit dependence of P on H .

Example 10-1

Random number generators are often used to generate a sequence of random numbers that can then be used as the input sequence to a dynamical system, or as an additive measurement noise sequence. To run a random number generator, you must choose a probability model. The Gaussian model is often used; however, it is characterized by two parameters, mean μ and variance σ^2 . In order to obtain a stream of Gaussian random numbers from the random number generator, you must fix μ and σ^2 . Let μ_T and σ_T^2 denote (true) values chosen for μ and σ^2 . The Gaussian probability density function for the generator is $p[z(k)|\mu_T, \sigma_T^2]$, and the numbers we obtain at its output, $z(1), z(2), \dots$, are of course quite dependent on the hypothesis $H_T = (\mu_T, \sigma_T^2)$. \square

For fixed H we can apply the axioms of probability (e.g., see Papoulis, 1965). If, for example, results R_1 and R_2 are mutually exclusive, then $P(R_1 \text{ or } R_2|H) = P(R_1|H) + P(R_2|H)$.

Definition 10-1 (Edwards, 1972, pg. 9). *Likelihood, $L(H|R)$, of the hypothesis H given the results R and a specific probability model is proportional to $P(R|H)$, the constant of proportionality being arbitrary, i.e.,*

$$L(H|R) = cP(R|H) \quad \square \quad (10-1)$$

For likelihood R is fixed (i.e., given ahead of time) and H is variable. *There are no axioms of likelihood.* Likelihood cannot be compared using different data sets (i.e., different results, say, R_1 and R_2) unless the data sets are statistically independent.

Example 10-2

Suppose we are given a sequence of Gaussian random numbers, using the random number generator that was described in Example 10-1, but, we do not know μ_T and σ_T^2 . Is it possible to infer (i.e., estimate) what the values of μ and σ^2 were that most likely generated the given sequence? The method of maximum-likelihood, which we study in Lesson 11, will show us how to do this. The starting point for the estimation of μ and σ^2 will be $p[z(k)|\mu, \sigma^2]$, where now $z(k)$ is fixed and μ and σ^2 are treated as variables. \square

Example 10-3 (Edwards, 1972, pg. 10)

To further illustrate the difference between $P(R|H)$ and $L(H|R)$, we consider the following binomial model which we assume describes the occurrence of boys and girls in a family of two children:

$$P(R|p) = \frac{(m+f)!}{m! f!} p^m (1-p)^f \quad (10-2)$$

where p denotes the probability of a male child, m equals the number of male children, f equals the number of female children, and, in this example,

$$m + f = 2 \quad (10-3)$$

Our objective is to determine p ; but, to do this we need some results. Knocking on neighbor's doors and conducting a simple survey, we establish two data sets:

$$R_1 = \{\text{1 boy and 1 girl}\} \Rightarrow m = 1 \quad \text{and} \quad f = 1 \quad (10-4)$$

$$R_2 = \{\text{2 boys}\} \Rightarrow m = 2 \quad \text{and} \quad f = 0 \quad (10-5)$$

In order to keep the determination of p simple for this meager collection of data, we shall only consider two values for p , i.e., two hypotheses,

$$\begin{cases} H_1 : p = 1/4 \\ H_2 : p = 1/2 \end{cases} \quad (10-6)$$

To begin, we create a *table of probabilities*, in which the entries are $P(R_i|H, \text{fixed})$, where this is computed using (10-2). For H_1 (i.e., $p = 1/4$), $P(R_1|1/4) = 3/8$ and $P(R_2|1/4) = 1/16$; for H_2 (i.e., $p = 1/2$), $P(R_1|1/2) = 1/2$ and $P(R_2|1/2) = 1/4$. These results are collected together in Table 10-1.

TABLE 10-1 $P(R_i|H, \text{fixed})$

	R_1	R_2
H_1	3/8	1/16
H_2	1/2	1/4

Next, we create a *table of likelihoods*, using (10-1). In this table (Table 10-2) the entries are $L(H_i|R, \text{fixed})$. Constants c_1 and c_2 are arbitrary and c_1 , for example, appears in each one of the table entries in the R_1 column.

TABLE 10-2 $L(H_i|R, \text{fixed})$

	R_1	R_2
H_1	$3/8 c_1$	$1/16 c_2$
H_2	$1/2 c_1$	$1/4 c_2$

What can we conclude from the table of likelihoods? First, for data R_1 , the likelihood of H_1 is $3/4$ the likelihood of H_2 . The number $3/4$ was obtained by taking the

ratio of likelihoods $L(H_1|R_1)$ and $L(H_2|R_1)$. Second, on data R_2 , the likelihood of H_1 is $1/4$ the likelihood of H_2 [note that $1/4 = 1/16 c_2/1/4 c_2$]. Finally, we conclude that, even from our two meager results, the value $p = 1/2$ appears to be more likely than the value $p = 1/4$, which, of course, agrees with our intuition. \square

LIKELIHOOD RATIO

In the preceding example we were able to draw conclusions about the likelihood of one hypothesis versus a second hypothesis by comparing ratios of likelihood, defined, of course, on the same set of data. Forming the ratios of likelihood, we obtain the *likelihood ratio*.

Definition 10-2 (Edwards, 1972, pg. 10). *The likelihood ratio of two hypotheses on the same data is the ratio of the likelihoods on the data. Let $L(H_1, H_2|R)$ denote likelihood ratio; then,*

$$L(H_1, H_2|R) = \frac{L(H_1|R)}{L(H_2|R)} = \frac{P(R|H_1)}{P(R|H_2)} \quad \square \quad (10-7)$$

Observe that likelihood ratio statements do not depend on the arbitrary constant c which appears in the definition of likelihood, because c cancels out of the ratio $cP(R|H_1)/cP(R|H_2)$.

Theorem 10-1 (Edwards, 1972, pg. 11). *Likelihood ratios of two hypotheses on statistically independent sets of data may be multiplied together to form the likelihood ratio of the combined data.*

Proof. Let $L(H_1, H_2|R_1 \& R_2)$ denote the likelihood ratio of H_1 and H_2 on the combined data $R_1 \& R_2$, i.e.,

$$L(H_1, H_2|R_1 \& R_2) = \frac{L(H_1|R_1 \& R_2)}{L(H_2|R_1 \& R_2)} = \frac{P(R_1 \& R_2|H_1)}{P(R_1 \& R_2|H_2)} \quad (10-8)$$

Because $R_1 \& R_2$ are statistically independent data, $P(R_1 \& R_2|H_i) = P(R_1|H_i)P(R_2|H_i)$; hence,

$$\begin{aligned} L(H_1, H_2|R_1 \& R_2) &= \frac{P(R_1|H_1)}{P(R_1|H_2)} \times \frac{P(R_2|H_1)}{P(R_2|H_2)} \\ &= L(H_1, H_2|R_1) \times L(H_1, H_2|R_2) \quad \square \end{aligned} \quad (10-9)$$

Example 10-4

We can now state the conclusions that are given at the end of Example 10-3 more formally. From Table 10-2, we see that $L(1/4, 1/2|R_1) = 3/4$ and $L(1/4, 1/2|R_2) = 1/4$. Additionally, because R_1 and R_2 are data from independent experiments, $L(1/4, 1/2|R_1 \& R_2) = 3/4 \times 1/4 = 3/16$. This reinforces our intuition that $p = 1/2$ is much more likely than $p = 1/4$. \square

$R_1 \& R_2 = 3/4 \times 1/4 = 3/16$. This reinforces our intuition that $p = 1/2$ is much more likely than $p = 1/4$. \square

RESULTS DESCRIBED BY CONTINUOUS DISTRIBUTIONS

Suppose the results R have a continuous distribution; then, we know from probability theory that the probability of obtaining a result that lies in the interval $(R, R + dR)$ is $P(R|R)dR$, as $dR \rightarrow 0$. $P(R|R)$ is then a probability density. In this case, $L(H|R) = cP(R|H)dR$; but, cdR can be defined as a new arbitrary constant, c_1 , so that $L(H|R) = c_1P(R|H)$. In likelihood ratio statements $c_1 = cdR$ disappears entirely; thus, likelihood and likelihood ratio are unaffected by the nature of the distribution of R .

Recall, that a transformation of variables greatly affects probability because of the dR which appears in the probability formula. *Likelihood and likelihood ratio, on the other hand, are unaffected by transformations of variables*, because of the absorption of dR into c_1 .

MULTIPLE HYPOTHESES

Thus far, all of our attention has been directed at the case of two hypotheses. In order to apply likelihood and likelihood ratio concepts to parameter estimation problems, where parameters take on more than two values, we must extend our preceding results to the case of multiple hypotheses. As stated by Edwards (1972, pg. 11), “Instead of forming all the pairwise likelihood ratios it is simpler to present the same information in terms of the likelihood ratios for the several hypotheses versus one of their number, which may be chosen quite arbitrarily for this purpose.”

Our extensions rely very heavily on the results for the two hypotheses case, because of the convenient introduction of an arbitrary comparison hypothesis, H^* . Let H_i denote the i th hypothesis; then,

$$L(H_i, H^*|R) = \frac{L(H_i|R)}{L(H^*|R)} \quad (10-10)$$

Observe, also, that

$$\frac{L(H_i, H^*|R)}{L(H_j, H^*|R)} = \frac{L(H_i|R)}{L(H_j|R)} = L(H_i, H_j|R) \quad (10-11)$$

which means that we can compute the likelihood-ratio between *any* two hypotheses H_i and H_j if we can compute the likelihood-ratio function $L(H_k, H^*|R)$.

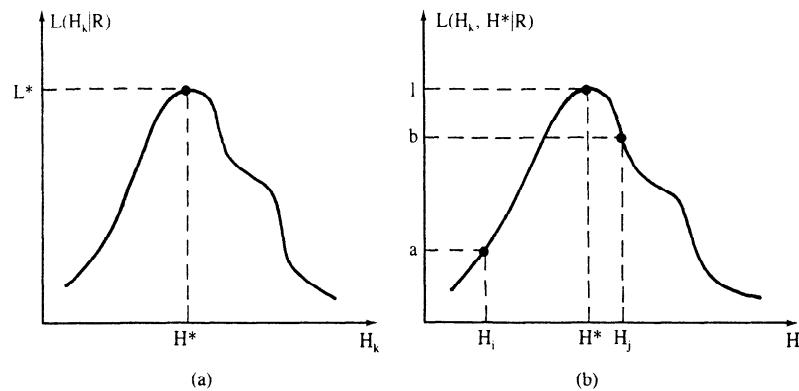


Figure 10-1 Multiple hypotheses case: (a) likelihood $L(H_k|R)$ versus H_k , and (b) likelihood ratio $L(H_k, H^*|R)$ versus H_k . Comparison hypothesis H^* has been chosen to be the hypothesis associated with the maximum value of $L(H_k|R)$, L^* .

Figure 10-1(a) depicts a likelihood function $L(H_k|R)$. Any value of H_k can be chosen as the comparison hypothesis. We choose H^* as the hypothesis associated with the maximum value of $L(H_k|R)$, so that the maximum value of $L(H_k, H^*|R)$ will be normalized to unity. The likelihood ratio function $L(H_k, H^*|R)$, depicted in Figure 10-1(b), was obtained from Figure 10-1(a) and (10-10). In order to compute $L(H_i, H_j|R)$, we determine from Figure 10-1(b), that $L(H_i, H^*|R) = a$ and $L(H_j, H^*|R) = b$, so that $L(H_i, H_j|R) = a/b$.

Is it really necessary to know H^* in order to carry out the normalization of $L(H_k, H^*|R)$ depicted in Figure 10-1(b)? No, because H^* can be eliminated by a clever “conceptual” choice of constant c . We can choose c , such that

$$L(H^*|R) = L^* \triangleq 1 \quad (10-12)$$

According to (10-1), this means that

$$c = 1/P(R|H^*) \quad (10-13)$$

If c is “chosen” in this manner, then

$$L(H_k, H^*|R) = \frac{L(H_k|R)}{L(H^*|R)} = L(H_k|R) \quad (10-14)$$

which means that, *in the case of multiple hypotheses, likelihood and likelihood ratio can be used interchangeably*. This helps to explain why authors use different names for the function that is the starting point in the method of maximum likelihood, including *likelihood function* and *likelihood-ratio function*.

P R O B L E M S

- 10-1.** A test is said to be a *likelihood-ratio test* if there is a number c such that this test leads to (d_i refers to the i th decision)

$$\begin{aligned} d_1 &\quad \text{if } L(H_1, H_2|R) > c \\ d_2 &\quad \text{if } L(H_1, H_2|R) < c \\ d_1 \text{ or } d_2 &\quad \text{if } L(H_1, H_2|R) = c \end{aligned}$$

Consider a sequence of n tosses of a coin of which m are heads, i.e., $P(R|p) = p^m(1-p)^{n-m}$. Let $H_1 = p_1$ and $H_2 = p_2$, where $p_1 > p_2$ and p represents the probability of a head. Show that $L(H_1, H_2|R)$ increases as m increases. Then show that $L(H_1, H_2|R)$ increases when $\hat{p} = m/n$ increases, and that the likelihood-ratio test which consists of accepting H_1 if $L(H_1, H_2|R)$ is larger than some constant is equivalent to accepting H_1 if \hat{p} is larger than some related constant.

- 10-2.** Consider n independent observations x_1, x_2, \dots, x_n which are normally distributed with mean μ and known standard deviation σ . Our two hypotheses are: $H_1: \mu = \mu_1$ and $H_2: \mu = \mu_2$. Using the sample mean as an estimator of μ , show that the likelihood-ratio test (defined in Problem 10-1) consists of accepting H_1 if the sample mean exceeds some constant.
- 10-3.** Suppose that the data consists of a single observation z with Cauchy density given by

$$p(z|\theta) = \frac{1}{\pi[1 + (z - \theta)^2]}$$

Test (see Problem 10-1) $H_1: \theta = \theta_1 = 1$ versus $H_2: \theta = \theta_2 = -1$ when $c = 1/2$, i.e., show that for $c = 1/2$, we accept H_1 when $z > -0.35$ or when $z < -5.65$.

Maximum-Likelihood Estimation

Likelihood*

Let us consider a vector of unknown parameters θ that describes a collection of N independent identically distributed observations $z(k)$, $k = 1, 2, \dots, N$. We collect these measurements into an $N \times 1$ vector $\mathbf{z}(N)$, (\mathbf{z} for short),

$$\mathbf{z} = \text{col}(z(1), z(2), \dots, z(N)) \quad (11-1)$$

The likelihood of θ , given the observations \mathbf{z} , is defined to be proportional to the value of the probability density function of the observations given the parameters

$$l(\theta|\mathbf{z}) \propto p(\mathbf{z}|\theta) \quad (11-2)$$

where l is the likelihood function and p is the conditional joint probability density function. Because $z(i)$ are independent and identically distributed,

$$l(\theta|\mathbf{z}) \propto p(z(1)|\theta)p(z(2)|\theta) \cdots p(z(N)|\theta) \quad (11-3)$$

In many applications $p(\mathbf{z}|\theta)$ is exponential (e.g., Gaussian). It is easier then to work with the natural logarithm of $l(\theta|\mathbf{z})$ than with $l(\theta|\mathbf{z})$. Let

$$L(\theta|\mathbf{z}) = \ln l(\theta|\mathbf{z}) \quad (11-4)$$

Quantity L is sometimes referred to as the log-likelihood function, the support function (Kmenta, 1971), the likelihood function [Mehra (1971), Schweppe

(1965), and Stepner and Mehra (1973), for example], or the conditional likelihood function (Nahi, 1969). We shall use these terms interchangeably.

MAXIMUM-LIKELIHOOD METHOD AND ESTIMATES*

The *maximum-likelihood method* is based on the relatively simple idea that different populations generate different samples and that any given sample is more likely to have come from some populations than from others. The maximum-likelihood estimate (MLE) $\hat{\theta}_{ML}$ is the value of θ that maximizes l or L for a particular set of measurements \mathbf{z} . The logarithm of l is a monotonic transformation of l (i.e., whenever l is decreasing or increasing, $\ln l$ is also decreasing or increasing); therefore, the point corresponding to the maximum of l is also the point corresponding to the maximum of $\ln l = L$.

Obtaining an MLE involves specifying the likelihood function and finding those values of the parameters that give this function its maximum value. It is required that, if L is differentiable, the partial derivative of L (or l) with respect to each of the unknown parameters $\theta_1, \theta_2, \dots, \theta_n$ equal zero:

$$\frac{\partial L(\theta|\mathbf{z})}{\partial \theta_i} \Big|_{\theta = \hat{\theta}_{ML}} = 0 \quad \text{for all } i = 1, 2, \dots, n \quad (11-5)$$

To be sure that the solution of (11-5) gives, in fact, a maximum value of $L(\theta|\mathbf{z})$, certain second-order conditions must be fulfilled. Consider a Taylor series expansion of $L(\theta|\mathbf{z})$ about $\hat{\theta}_{ML}$, i.e.,

$$L(\theta|\mathbf{z}) = L(\hat{\theta}_{ML}|\mathbf{z}) + \sum_{i=1}^n \frac{\partial L(\hat{\theta}_{ML}|\mathbf{z})}{\partial \theta_i} (\theta_i - \hat{\theta}_{i,ML}) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 L(\hat{\theta}_{ML}|\mathbf{z})}{\partial \theta_i \partial \theta_j} (\theta_i - \hat{\theta}_{i,ML})(\theta_j - \hat{\theta}_{j,ML}) + \dots \quad (11-6)$$

where $\partial L(\hat{\theta}_{ML}|\mathbf{z}) / \partial \theta_i$, for example, is short for $[\partial L(\theta|\mathbf{z}) / \partial \theta_i]_{\theta = \hat{\theta}_{ML}}$. Because $\hat{\theta}_{ML}$ is the MLE of θ , the second term on the right-hand side of (11-6) is zero [by virtue of (11-5)]; hence,

$$L(\theta|\mathbf{z}) = L(\hat{\theta}_{ML}|\mathbf{z}) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 L(\hat{\theta}_{ML}|\mathbf{z})}{\partial \theta_i \partial \theta_j} (\theta_i - \hat{\theta}_{i,ML})(\theta_j - \hat{\theta}_{j,ML}) + \dots \quad (11-7)$$

Recognizing that the second term in (11-7) is a quadratic form, we can write it in a more compact notation as $1/2(\theta - \hat{\theta}_{ML})' J_o(\hat{\theta}_{ML}|\mathbf{z})(\theta - \hat{\theta}_{ML})$, where $J_o(\hat{\theta}_{ML}|\mathbf{z})$, the *observed Fisher information matrix* [see Equation (6-46)], is

$$J_o(\hat{\theta}_{ML}|\mathbf{z}) = \left(\frac{\partial^2 L(\theta|\mathbf{z})}{\partial \theta_i \partial \theta_j} \right) \Big|_{\theta = \hat{\theta}_{ML}} \quad i, j = 1, 2, \dots, n \quad (11-8)$$

* The material in this section is taken from Mendel (1983b, pp. 95–98).

hence,

$$L(\theta|\mathcal{Z}) = L(\hat{\theta}_{ML}|\mathcal{Z}) + \frac{1}{2}(\theta - \hat{\theta}_{ML})' J_o(\hat{\theta}_{ML}|\mathcal{Z})(\theta - \hat{\theta}_{ML}) + \dots \quad (11-9)$$

Now let us examine sufficient conditions for the likelihood function to be maximum. We assume that, close to θ , $L(\theta|\mathcal{Z})$ is approximately quadratic, in which case

$$L(\theta|\mathcal{Z}) \approx L(\hat{\theta}_{ML}|\mathcal{Z}) + \frac{1}{2}(\theta - \hat{\theta}_{ML})' J_o(\hat{\theta}_{ML}|\mathcal{Z})(\theta - \hat{\theta}_{ML}) \quad (11-10)$$

From vector calculus, it is well known that a sufficient condition for a function of n variables to be maximum is that the matrix of second partial derivatives of that function, evaluated at the extremum, must be negative definite. For $L(\theta|\mathcal{Z})$ in (11-10), this means that *a sufficient condition for $L(\theta|\mathcal{Z})$ to be maximized is*

$$J_o(\hat{\theta}_{ML}|\mathcal{Z}) < 0 \quad (11-11)$$

Example 11-1

This is a continuation of Example 10-2. We observe a random sample $\{z(1), z(2), \dots, z(N)\}$ at the output of a Gaussian random number generator and wish to find the maximum-likelihood estimators of μ and σ^2 .

The Gaussian density function $p(z|\mu, \sigma^2)$ is

$$p(z|\mu, \sigma^2) = (2\pi\sigma^2)^{-1/2} \exp\left\{-\frac{1}{2}[(z - \mu)/\sigma]^2\right\} \quad (11-12)$$

Its natural logarithm is

$$\ln p(z|\mu, \sigma^2) = -\frac{1}{2}\ln(2\pi\sigma^2) - \frac{1}{2}[(z - \mu)/\sigma]^2 \quad (11-13)$$

The likelihood function is

$$l(\mu, \sigma^2) = p(z(1)|\mu, \sigma^2)p(z(2)|\mu, \sigma^2) \cdots p(z(N)|\mu, \sigma^2) \quad (11-14)$$

and its logarithm is

$$L(\mu, \sigma^2) = \sum_{i=1}^N \ln p(z(i)|\mu, \sigma^2) \quad (11-15)$$

Substituting for $\ln p(z(i)|\mu, \sigma^2)$ gives

$$\begin{aligned} L(\mu, \sigma^2) &= \sum_{i=1}^N \left[-\frac{1}{2}\ln(2\pi\sigma^2) - \frac{1}{2}\left(\frac{z(i) - \mu}{\sigma}\right)^2 \right] \\ &= -\frac{N}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^N [z(i) - \mu]^2 \end{aligned} \quad (11-16)$$

There are two unknown parameters in L , μ and σ^2 . Differentiating L with respect to each of them gives

$$\frac{\partial L}{\partial \mu} = \frac{1}{\sigma^2} \sum_{i=1}^N [z(i) - \mu] \quad (11-17a)$$

$$\frac{\partial L}{\partial (\sigma^2)} = -\frac{N}{2} \frac{1}{\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^N [z(i) - \mu]^2 \quad (11-17b)$$

Equating these partial derivatives to zero, we obtain

$$\frac{1}{\hat{\sigma}_{ML}^2} \sum_{i=1}^N [z(i) - \hat{\mu}_{ML}] = 0 \quad (11-18a)$$

$$-\frac{N}{2} \frac{1}{\hat{\sigma}_{ML}^2} + \frac{1}{2\hat{\sigma}_{ML}^4} \sum_{i=1}^N [z(i) - \hat{\mu}_{ML}]^2 = 0 \quad (11-18b)$$

For $\hat{\sigma}_{ML}^2$ different from zero (11-18a) reduces to

$$\sum_{i=1}^N [z(i) - \hat{\mu}_{ML}] = 0$$

giving

$$\hat{\mu}_{ML} = \frac{1}{N} \sum_{i=1}^N z(i) = \bar{z} \quad (11-19)$$

Thus the MLE of the mean of a Gaussian population is equal to the sample mean \bar{z} . Once again we see that the sample mean is an optimal estimator.

Observe that (11-18a) and (11-18b) can be solved for $\hat{\sigma}_{ML}^2$, the MLE of σ^2 . Multiplying Eq. (11-18b) by $2\hat{\sigma}_{ML}^4$ leads to

$$-N\hat{\sigma}_{ML}^2 + \sum_{i=1}^N [z(i) - \hat{\mu}_{ML}]^2 = 0$$

Substituting \bar{z} for $\hat{\mu}_{ML}$ and solving for $\hat{\sigma}_{ML}^2$ gives

$$\hat{\sigma}_{ML}^2 = \frac{1}{N} \sum_{i=1}^N [z(i) - \bar{z}]^2 \quad (11-20)$$

Thus, the MLE of the variance of a Gaussian population is simply equal to the sample variance. \square

PROPERTIES OF MAXIMUM-LIKELIHOOD ESTIMATES

The importance of maximum-likelihood estimation is that it produces estimates that have very desirable properties.

Theorem 11-1. Maximum-likelihood estimates are: (1) consistent, (2) asymptotically Gaussian with mean θ and covariance matrix $\frac{1}{N} \mathbf{J}^{-1}$, in which \mathbf{J}

is the Fisher Information Matrix [Equation (6-45)], and, (3) asymptotically efficient.

Proof. For proofs of consistency, asymptotic normality and asymptotic efficiency, see Sorenson (1980, pp. 187–190; 190–192; and 192–193, respectively). These proofs, though somewhat heuristic, convey the ideas needed to prove the three parts of this theorem. More detailed analyses can be found in Cramer (1946) and Zacks (1971). See also, Problems 11-13, 11-14, and 11-15. \square

Theorem 11-2 (Invariance Property of MLE's). *Let $g(\theta)$ be a vector function mapping θ into an interval in r -dimensional Euclidean space. Let $\hat{\theta}_{ML}$ be a MLE of θ ; then $\hat{g}(\hat{\theta}_{ML})$ is a MLE of $g(\theta)$; i.e.,*

$$\hat{g}(\hat{\theta})_{ML} = g(\hat{\theta}_{ML}) \quad (11-21)$$

Proof (See Zacks, 1971). Note that Zacks points out that in many books, this theorem is cited only for the case of one-to-one mappings, $g(\theta)$. His proof does not require $g(\theta)$ to be one-to-one. Note, also, that the proof of this theorem is related to the “consistency carry-over” property of a consistent estimator, which was discussed in Lesson 7. \square

Example 11-2

We wish to obtain a MLE of the variance σ_1^2 in our linear model $z(k) = h'(k)\theta + v(k)$. One approach is to let $\theta_1 = \sigma_1^2$, establish the log-likelihood function for θ_1 and maximize it, to determine $\hat{\theta}_{1,ML}$. Usually, mathematical programming (i.e., search techniques) must be used to determine $\hat{\theta}_{1,ML}$. Here is where a difficulty can occur, because θ_1 (a variance) is known to be positive; thus, $\hat{\theta}_{1,ML}$ must be constrained to be positive. Unfortunately, constrained mathematical programming techniques are more difficult than unconstrained ones.

A second approach is to let $\theta_2 = \sigma_1$, establish the log-likelihood function for θ_2 (it will be the same as the one for θ_1 , except that θ_1 will be replaced by θ_2^2), and maximize it to determine $\hat{\theta}_{2,ML}$. Because θ_2 is a standard deviation, which can be positive or negative, unconstrained mathematical programming can be used to determine $\hat{\theta}_{2,ML}$. Finally, we use the Invariance Property of MLE's to compute $\hat{\theta}_{1,ML}$, as

$$\hat{\theta}_{1,ML} = (\hat{\theta}_{2,ML})^2 \quad \square \quad (11-22)$$

THE LINEAR MODEL ($\mathcal{H}(k)$ deterministic)

We return now to the linear model

$$\mathcal{Z}(k) = \mathcal{H}(k)\theta + \mathcal{V}(k) \quad (11-23)$$

in which θ is an $n \times 1$ vector of deterministic parameters, $\mathcal{H}(k)$ is deterministic, and $\mathcal{V}(k)$ is zero mean white noise, with covariance matrix $\mathcal{R}(k)$. This is precisely the same model that was used to derive the BLUE of θ , $\hat{\theta}_{BLU}(k)$. Our

objectives in this paragraph are twofold: (1) to derive the MLE of θ , $\hat{\theta}_{ML}(k)$, and (2) to relate $\hat{\theta}_{ML}(k)$ to $\hat{\theta}_{BLU}(k)$ and $\hat{\theta}_{LS}(k)$.

In order to derive the MLE of θ , we need to determine a formula for $p(\mathcal{Z}(k)|\theta)$. To proceed, we assume that $\mathcal{V}(k)$ is Gaussian, with multivariate density function $p(\mathcal{V}(k))$, where

$$p(\mathcal{V}(k)) = \frac{1}{\sqrt{(2\pi)^N |\mathcal{R}(k)|}} \exp \left[-\frac{1}{2} \mathcal{V}'(k) \mathcal{R}^{-1}(k) \mathcal{V}(k) \right] \quad (11-24)$$

Recall (e.g., Papoulis, 1965) that linear transformations on, and linear combinations of, Gaussian random vectors are themselves Gaussian random vectors. For this reason, it is clear that when $\mathcal{V}(k)$ is Gaussian, $\mathcal{Z}(k)$ is as well. The multivariate Gaussian density function of $\mathcal{Z}(k)$, derived from $p(\mathcal{V}(k))$, is

$$p(\mathcal{Z}(k)|\theta) = \frac{1}{\sqrt{(2\pi)^N |\mathcal{R}(k)|}} \exp \left\{ -\frac{1}{2} [\mathcal{Z}(k) - \mathcal{H}(k)\theta]' \mathcal{R}^{-1}(k) [\mathcal{Z}(k) - \mathcal{H}(k)\theta] \right\} \quad (11-25)$$

Theorem 11-3. *When $p(\mathcal{Z}|\theta)$ is multivariate Gaussian and $\mathcal{H}(k)$ is deterministic, then the principle of ML leads to the BLUE of θ , i.e.,*

$$\hat{\theta}_{ML}(k) = \hat{\theta}_{BLU}(k) \quad (11-26)$$

Proof. We must maximize $p(\mathcal{Z}|\theta)$ with respect to θ . This can be accomplished by minimizing the argument of the exponential in (11-25); hence, $\hat{\theta}_{ML}(k)$ is the solution of

$$\frac{d}{d\theta} (\mathcal{Z} - \mathcal{H}\theta)' \mathcal{R}^{-1} (\mathcal{Z} - \mathcal{H}\theta) \Big|_{\theta = \hat{\theta}_{ML}} = 0 \quad (11-27)$$

This equation can also be expressed as $dJ[\hat{\theta}_{ML}]/d\theta = 0$ where $J[\hat{\theta}_{ML}] = \tilde{\mathcal{Z}}'(k) \mathcal{R}^{-1}(k) \tilde{\mathcal{Z}}(k)$ and $\tilde{\mathcal{Z}}(k) = \mathcal{Z}(k) - \mathcal{H}\hat{\theta}_{ML}(k)$. Comparing this version of (11-27) with Equation (3-4) and the subsequent derivation of the WLSE of θ , we conclude that

$$\hat{\theta}_{ML}(k) = \hat{\theta}_{WLS}(k) \Big|_{\mathcal{W}(k) = \mathcal{R}^{-1}(k)} \quad (11-28)$$

but, we also know, from Lesson 9, that

$$\hat{\theta}_{WLS}(k) \Big|_{\mathcal{W}(k) = \mathcal{R}^{-1}(k)} = \hat{\theta}_{BLU}(k) \quad (11-29)$$

From (11-28) and (11-29), we conclude that

$$\hat{\theta}_{ML}(k) = \hat{\theta}_{BLU}(k) \quad \square \quad (11-30)$$

We now suggest a reason why Theorem 9-6 (and, subsequently, Theorem 4-1) is referred to as the “information form” of recursive BLUE. From

Theorems 11-1 and 11-3 we know that, when $\mathcal{H}(k)$ is deterministic, $\hat{\theta}_{BLU} \sim N(\hat{\theta}; \theta, 1/N \mathbf{J}^{-1})$. This means that $\mathbf{P}(k) = \text{cov}[\hat{\theta}_{BLU}(k)]$ is proportional to \mathbf{J}^{-1} . Observe, in Theorem 9-6, that we compute \mathbf{P}^{-1} recursively, and not \mathbf{P} . Because \mathbf{P}^{-1} is proportional to the Fisher information matrix \mathbf{J} , the results in Theorem 9-6 (and 4-1) are therefore referred to as the “information form” of recursive BLUE.

A second more pragmatic reason is due to the fact that the inverse of any covariance matrix is known as an information matrix (e.g., Anderson and Moore, 1979, pg. 138). Consequently, any algorithm that is in terms of information matrices is known as an “information form” algorithm.

Corollary 11-1. *If $p[\mathcal{Z}(k)|\theta]$ is multivariate Gaussian, $\mathcal{H}(k)$ is deterministic, and $\mathcal{R}(k) = \sigma_v^2 \mathbf{I}$, then*

$$\hat{\theta}_{ML}(k) = \hat{\theta}_{BLU}(k) = \hat{\theta}_{LS}(k) \quad (11-31)$$

These estimators are: unbiased, most efficient (within the class of linear estimators), consistent, and Gaussian.

Proof. To obtain (11-31), combine the results in Theorems 11-3 and 9-2. The estimators are:

1. unbiased, because $\hat{\theta}_{BLU}(k)$ is unbiased;
2. most efficient, because $\hat{\theta}_{BLU}(k)$ is most efficient;
3. consistent, because $\hat{\theta}_{ML}(k)$ is consistent; and
4. Gaussian, because they depend linearly upon $\mathcal{Z}(k)$ which is Gaussian. \square

Observe that, when $\mathcal{Z}(k)$ is Gaussian, this corollary permits us to make statements about small-sample properties of MLE's. Usually, we cannot make such statements.

A LOG-LIKELIHOOD FUNCTION FOR AN IMPORTANT DYNAMICAL SYSTEM

In practice there are two major problems in obtaining MLE's of parameters in models of dynamical systems

1. obtaining an expression for $L\{\theta|\mathcal{Z}\}$, and
2. maximizing $L\{\theta|\mathcal{Z}\}$ with respect to θ .

In this section we direct our attention to the first problem for a linear, time-invariant dynamical system that is excited by a known forcing function, has deterministic initial conditions, and has measurements that are corrupted by

additive white Gaussian noise. This system is described by the following state-equation model:

$$\mathbf{x}(k+1) = \Phi \mathbf{x}(k) + \Psi \mathbf{u}(k) \quad (11-32)$$

and

$$\mathbf{z}(k+1) = \mathbf{H} \mathbf{x}(k+1) + \mathbf{v}(k+1) \quad k = 0, 1, \dots, N-1 \quad (11-33)$$

In this model, $\mathbf{u}(k)$ is known ahead of time, $\mathbf{x}(0)$ is deterministic, $E\{\mathbf{v}(k)\} = \mathbf{0}$, $\mathbf{v}(k)$ is Gaussian, and $E\{\mathbf{v}(k)\mathbf{v}'(j)\} = \mathbf{R} \delta_{kj}$.

To begin, we must establish the parameters that constitute θ . In theory θ could contain all of the elements in Φ , Ψ , \mathbf{H} and \mathbf{R} . In practice, however, these matrices are never completely unknown. State equations are either derived from physical principles (e.g., Newton's laws) or associated with a canonical model (e.g., controllability canonical form); hence, we usually know that certain elements in Φ , Ψ and \mathbf{H} are identically zero or are known constants.

Even though all of the elements in Φ , Ψ , \mathbf{H} and \mathbf{R} will not be unknown in an application, there still may be more unknowns present than can possibly be identified. How many parameters can be identified, and which parameters can be identified by maximum-likelihood estimation (or, for that matter, by any type of estimation) is the subject of *identifiability of systems* (Stepner and Mehra, 1973). We shall assume that θ is identifiable. Identifiability is akin to “existence.” When we assume θ is identifiable, we assume that it is possible to identify θ by ML methods. This means that all of our statements are predicated by the statement: If θ is identifiable, then

We let

$$\theta = \text{col}(\text{elements of } \Phi, \Psi, \mathbf{H}, \text{ and } \mathbf{R}) \quad (11-34)$$

Example 11-3

The “controllable canonical form” state-variable representation for the discrete-time autoregressive moving average (ARMA) model

$$H(z) = \frac{\beta_1 z^{n-1} + \beta_2 z^{n-2} + \dots + \beta_{n-1} z + \beta_n}{z^n + \alpha_1 z^{n-1} + \dots + \alpha_{n-1} z + \alpha_n} \quad (11-35)$$

which implies the ARMA difference equation (z denotes the unit advance operator)

$$\begin{aligned} y(k+n) + \alpha_1 y(k+n-1) + \dots + \alpha_n y(k) \\ = \beta_1 u(k+n-1) + \dots + \beta_n u(k) \end{aligned} \quad (11-36)$$

is

$$\begin{pmatrix} x_1(k+1) \\ x_2(k+1) \\ \vdots \\ x_n(k+1) \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\alpha_n & -\alpha_{n-1} & -\alpha_{n-2} & \cdots & -\alpha_1 \end{pmatrix}}_{\Phi} \begin{pmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_n(k) \end{pmatrix} + \underbrace{\begin{pmatrix} 0 \\ 0 \\ \vdots \\ i \end{pmatrix}}_{\Psi} u(k) \quad (11-37)$$

and

$$y(k) = \underbrace{(\beta_n, \beta_{n-1}, \dots, \beta_1)}_{\mathbf{H}} \mathbf{x}(k) \quad (11-38)$$

For this model there are no unknown parameters in matrix Ψ , and, Φ and \mathbf{H} each contain exactly n (unknown) parameters. Matrix Φ contains the $n \alpha$ -parameters which are associated with the poles of $H(z)$, whereas matrix \mathbf{H} contains the $n \beta$ -parameters which are associated with the zero of $H(z)$. In general, an n th-order, single-input single-output system is completely characterized by $2n$ parameters. \square

Our objective now is to determine $L(\theta|\mathcal{Z})$ for the system in (11-32) and (11-33). To begin, we must determine $p(\mathcal{Z}|\theta) = p(\mathbf{z}(1), \mathbf{z}(2), \dots, \mathbf{z}(N)|\theta)$. This is easy to do, because of the whiteness of noise $\mathbf{v}(k)$, i.e., $p(\mathbf{z}(1), \mathbf{z}(2), \dots, \mathbf{z}(N)) = p(\mathbf{z}(1))p(\mathbf{z}(2)) \dots p(\mathbf{z}(N))$; thus,

$$L(\theta|\mathcal{Z}) = \ln \left[\prod_{i=1}^N p(\mathbf{z}(i)|\theta) \right] \quad (11-39)$$

From the Gaussian nature of $\mathbf{v}(k)$ and the linear measurement model in (11-33), we know that

$$p(\mathbf{z}(i)|\theta) = \frac{1}{\sqrt{(2\pi)^m |\mathbf{R}|}} \exp \left\{ -\frac{1}{2} [\mathbf{z}(i) - \mathbf{H}\mathbf{x}(i)]' \mathbf{R}^{-1} [\mathbf{z}(i) - \mathbf{H}\mathbf{x}(i)] \right\} \quad (11-40)$$

thus,

$$\begin{aligned} L(\theta|\mathcal{Z}) &= -\frac{1}{2} \sum_{i=1}^N [\mathbf{z}(i) - \mathbf{H}\mathbf{x}(i)]' \mathbf{R}^{-1} [\mathbf{z}(i) - \mathbf{H}\mathbf{x}(i)] \\ &\quad - \frac{N}{2} \ln |\mathbf{R}| - \frac{N}{2} m \ln 2\pi \end{aligned} \quad (11-41)$$

The log-likelihood function $L(\theta|\mathcal{Z})$ is a function of θ . To indicate which quantities on the right-hand side of (11-41) may depend on θ , we subscript all such quantities with θ . Additionally, because $\frac{N}{2} m \ln 2\pi$ does not depend on θ we neglect it in subsequent discussions. Our final log-likelihood function is:

$$L(\theta|\mathcal{Z}) = -\frac{1}{2} \sum_{i=1}^N [\mathbf{z}(i) - \mathbf{H}_\theta \mathbf{x}_\theta(i)]' \mathbf{R}_\theta^{-1} [\mathbf{z}(i) - \mathbf{H}_\theta \mathbf{x}_\theta(i)] - \frac{N}{2} \ln |\mathbf{R}_\theta| \quad (11-42)$$

Observe that θ occurs explicitly and implicitly in $L(\theta|\mathcal{Z})$. Matrices \mathbf{H}_θ and \mathbf{R}_θ contain the explicit dependence of $L(\theta|\mathcal{Z})$ on θ , whereas state vector $\mathbf{x}_\theta(i)$ contains the implicit dependence of $L(\theta|\mathcal{Z})$ on θ . In order to numerically calculate the right-hand side of (11-42), we must solve state equation (11-32). This can be done when values of the unknown parameters which appear in Φ and Ψ are given specific values; for, then (11-32) becomes

$$\mathbf{x}_\theta(k+1) = \Phi_\theta \mathbf{x}_\theta(k) + \Psi_\theta \mathbf{u}(k) \quad \mathbf{x}_\theta(0) \text{ known} \quad (11-43)$$

In essence, then, state equation (11-43) is a constraint that is associated with the computation of the log-likelihood function.

How do we determine $\hat{\theta}_{ML}$ for $L(\theta|\mathcal{Z})$ given in (11-42) [subject to its constraint in (11-43)]? No simple closed-form solution is possible, because θ enters into $L(\theta|\mathcal{Z})$ in a complicated nonlinear manner. The only way presently known for obtaining $\hat{\theta}_{ML}$ is by means of mathematical programming, which is beyond the scope of this course (e.g., see Mendel, 1983, pp. 141–142).

Comment. This completes our studies of methods for estimating unknown deterministic parameters. Prior to studying methods for estimating unknown random parameters, we pause to review an important body of material about multivariate Gaussian random variables.

PROBLEMS

- 11-1. If \hat{a} is the MLE of a , what is the MLE of $a^{1/a}$? Explain your answer.
- 11-2. (Sorenson, 1980, Theorem 5.1, pg. 185). If an estimator exists such that equality is satisfied in the Cramer-Rao inequality, prove that it can be determined as the solution of the likelihood equation.
- 11-3. Consider a sequence of independently distributed random variables x_1, x_2, \dots, x_N , having the probability density function $\theta^2 x_i e^{-\theta x_i}$, where $\theta > 0$.
 - (a) Derive $\hat{\theta}_{ML}(N)$.
 - (b) You want to study whether or not $\hat{\theta}_{ML}(N)$ is an unbiased estimator of θ . Explain (without working out all the details) how you would do this.
- 11-4. Consider a random variable z which can take on the values $z = 0, 1, 2, \dots$. This variable is Poisson distributed, i.e., its probability function $P(z)$ is $P(z) = \mu^z e^{-\mu} / z!$ Let $z(1), z(2), \dots, z(N)$ denote N independent observations of z . Find $\hat{\mu}_{ML}(N)$.
- 11-5. If $p(z) = \theta e^{-\theta z}$, $z > 0$ and $p(z) = 0$, otherwise, find $\hat{\theta}_{ML}$ given a sample of N independent observations.
- 11-6. Find the maximum-likelihood estimator of θ from a sample of N independent observations that are uniformly distributed over the interval $(0, \theta)$.
- 11-7. Derive the maximum-likelihood estimate of the signal power, defined as ζ^2 , in the signal $z(k) = \zeta s(k)$, $k = 0, 1, \dots, N$, where $s(k)$ is a scalar, stationary, zero-mean Gaussian random sequence with autocorrelation $E[s(i)s(j)] = \phi(j-i)$.
- 11-8. Suppose x is a binary variable that assumes a value of 1 with probability a and a value of 0 with probability $(1-a)$. The probability distribution of x can be described by

$$P(x) = (1-a)^{(1-x)} a^x$$
 Suppose we draw a random sample of N values $\{x_1, x_2, \dots, x_N\}$. Find the MLE of a .
- 11-9. (Mendel, 1973, Exercise 3-15, pg. 177). Consider the linear model for which $\mathbf{V}(k)$ is Gaussian with zero mean and $\mathbf{R}(k) = \sigma^2 \mathbf{I}$.

- (a) Show that the maximum-likelihood estimator of σ^2 , denoted $\hat{\sigma}_{ML}^2$, is

$$\hat{\sigma}_{ML}^2 = \frac{(\mathcal{Z} - \mathcal{H}\hat{\theta}_{ML})'(\mathcal{Z} - \mathcal{H}\hat{\theta}_{ML})}{N}$$

where $\hat{\theta}_{ML}$ is the maximum-likelihood estimator of θ .

- (b) Show that $\hat{\sigma}_{ML}^2$ is biased, but that it is asymptotically unbiased.

- 11-10. We are given N independent samples $\mathbf{z}(1), \mathbf{z}(2), \dots, \mathbf{z}(N)$ of the identically distributed two-dimensional random vector $\mathbf{z}(i) = \text{col}[z_1(i), z_2(i)]$, with the Gaussian density function

$$p(z_1(i), z_2(i)|\rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{z_1^2(i) - 2\rho z_1(i)z_2(i) + z_2^2(i)}{2(1-\rho^2)}\right\}$$

where ρ is the correlation coefficient between z_1 and z_2 .

- (a) Determine $\hat{\rho}_{ML}$ [Hint: You will obtain a cubic equation for $\hat{\rho}_{ML}$ and must show that $\hat{\rho}_{ML} = r$, where r equals the sample correlation coefficient].
 (b) Determine the Cramer-Rao bound for $\hat{\rho}_{ML}$.

- 11-11. It is well known, from linear system theory, that the choice of state variables used to describe a dynamical system is not unique. Consider the n th-order difference equation

$$\begin{aligned} y(k+n) + a_1y(k+n-1) + \cdots + a_ny(k) \\ = b_n^0m(k+n-1) + b_{n-1}^0m(k+n-2) + \cdots + b_1^0m(k). \end{aligned}$$

A state-variable model for this system is

$$\begin{pmatrix} x_1(k+1) \\ x_2(k+1) \\ \vdots \\ x_n(k+1) \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{pmatrix} \begin{pmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_n(k) \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} m(k)$$

and

$$y(k) = (1 \ 0 \ \dots \ 0)x(k)$$

where

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ a_1 & 1 & 0 & \cdots & 0 & 0 \\ a_2 & a_1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n-2} & a_{n-3} & a_{n-4} & \cdots & 1 & 0 \\ a_{n-1} & a_{n-2} & a_{n-3} & \cdots & a_1 & 1 \end{pmatrix} \begin{pmatrix} b_1^0 \\ b_2^0 \\ \vdots \\ b_n^0 \end{pmatrix}$$

Suppose maximum-likelihood estimates have been determined for the a_i - and b_i -parameters. How does one compute the maximum-likelihood estimates of the b_i^0 -parameters?

- 11-12. Prove that, for a random sample of measurements, a maximum-likelihood estimator is a consistent estimator [Hints: (1) Show that $E\{\partial \ln p(\mathcal{Z}|\theta)/\partial\theta|\theta\} = 0$; (2) Expand $\partial \ln p(\mathcal{Z}|\hat{\theta}_{ML})/\partial\theta$ in a Taylor series about θ , and show that $\partial \ln p(\mathcal{Z}|\theta)/\partial\theta = -(\hat{\theta}_{ML} - \theta)' \partial^2 \ln p(\mathcal{Z}|\theta^*)/\partial\theta^2$, where $\theta^* = \lambda\theta + (1 - \lambda)\hat{\theta}_{ML}$ and $0 \leq \lambda \leq 1$; (3) Show that $\partial \ln p(\mathcal{Z}|\theta)/\partial\theta = \sum_{i=1}^N \partial \ln p(z(i)|\theta)/\partial\theta$ and $\partial^2 \ln p(z(i)|\theta)/\partial\theta^2 = \sum_{i=1}^N \partial^2 \ln p(z(i)|\theta)/\partial\theta^2$].

$p(\mathcal{Z}|\theta)/\partial\theta^2 = \sum_{i=1}^N \partial^2 \ln p(z(i)|\theta)/\partial\theta^2$; (4) Using the strong law of large numbers to assert that, with probability one, sample averages converge to ensemble averages, and assuming that $E\{\partial^2 \ln p(z(i)|\theta)/\partial\theta^2\}$ is negative definite, show that $\hat{\theta}_{ML} \rightarrow \theta$ with probability one; thus, $\hat{\theta}_{ML}$ is a consistent estimator of θ . The steps in this proof have been taken from Sorenson, 1980, pp. 187-191.]

- 11-13. Prove that, for a random sample of measurements, a maximum-likelihood estimator is asymptotically Gaussian with mean value θ and covariance matrix $(N\mathbf{J})^{-1}$ where \mathbf{J} is the Fisher information matrix for a single measurement $z(i)$. [Hints: (1) Expand $\partial \ln p(\mathcal{Z}|\hat{\theta}_{ML})/\partial\theta$ in a Taylor series about θ and neglect second- and higher-order terms; (2) Show that

$$\frac{1}{N} \frac{\partial}{\partial\theta} \ln p(\mathcal{Z}|\theta) = (\hat{\theta}_{ML} - \theta)' \left[-\frac{1}{N} \frac{\partial^2}{\partial\theta^2} \ln p(\mathcal{Z}|\theta) \right]$$

(3) Let $\mathbf{s}(\theta) = \partial \ln p(\mathcal{Z}|\theta)/\partial\theta$ and show that $\mathbf{s}(\theta) = \sum_{i=1}^N \mathbf{s}_i(\theta)$, where $\mathbf{s}_i(\theta) = \partial \ln p(z(i)|\theta)/\partial\theta$; (4) Let \bar{s} denote the sample mean of $\mathbf{s}_i(\theta)$, and show that the distribution of \bar{s} asymptotically converges to a Gaussian distribution having mean zero and covariance \mathbf{J}/N ; (5) Using the strong law of large numbers we know that $\frac{1}{N} \partial^2 \ln p(\mathcal{Z}|\theta)/\partial\theta^2 \rightarrow E\{\partial^2 \ln p(z(i)|\theta)/\partial\theta^2|\theta\}$; consequently, show that $(\hat{\theta}_{ML} - \theta)' \mathbf{J}$ is asymptotically Gaussian with zero mean and covariance \mathbf{J}/N ; (6) Complete the proof of the theorem. The steps in this proof have been taken from Sorenson, 1980, pg. 192.]

- 11-14. Prove that, for a random sample of measurements, a maximum-likelihood estimator is asymptotically efficient [Hints: (1) Let $\mathbf{S}(\theta) = -E\{\partial^2 \ln p(\mathcal{Z}|\theta)/\partial\theta^2|\theta\}$ and show that $\mathbf{S}(\theta) = N\mathbf{J}$ where \mathbf{J} is the Fisher information matrix for a single measurement $z(i)$; (2) Use the result stated in Problem 11-13 to complete the proof. The steps in this proof have been taken from Sorenson, 1980, pg. 193.]

Elements of Multivariate Gaussian Random Variables

INTRODUCTION

Gaussian random variables are important and widely used for at least two reasons. First, they often provide a model that is a reasonable approximation to observed random behavior. Second, if the random phenomenon that we observe at the macroscopic level is the superposition of an arbitrarily large number of independent random phenomena, which occur at the microscopic level, the macroscopic description is justifiably Gaussian.

Most (if not all) of the material in this lesson should be a review for a reader who has had a course in probability theory. We collect a wide range of facts about multivariate Gaussian random variables here, in one place, because they are often needed in the remaining lessons.

UNIVARIATE GAUSSIAN DENSITY FUNCTION

A random variable y is said to be distributed as the *univariate Gaussian distribution* with mean m_y and variance σ_y^2 [i.e., $y \sim N(y; m_y, \sigma_y^2)$] if the density function of y is given as

$$p(y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp[-(y - m_y)^2/2\sigma_y^2] \quad -\infty < y < \infty \quad (12-1)$$

For notational simplicity throughout this chapter, we do not condition density functions on their parameters; this conditioning is understood. Density $p(y)$ is the familiar bell-shaped curve, centered at $y = m_y$.

MULTIVARIATE GAUSSIAN DENSITY FUNCTION

Let y_1, y_2, \dots, y_m be random variables, and $\mathbf{y} = \text{col}(y_1, y_2, \dots, y_m)$. The function

$$p(y_1, y_2, \dots, y_m) = p(\mathbf{y}) = \frac{1}{\sqrt{(2\pi)^m |\mathbf{P}_y|}} \exp\left[-\frac{1}{2} (\mathbf{y} - \mathbf{m}_y)' \mathbf{P}_y^{-1} (\mathbf{y} - \mathbf{m}_y)\right] \quad (12-2)$$

is said to be a *multivariate (m-variate) Gaussian density function* [i.e., $\mathbf{y} \sim N(\mathbf{y}; \mathbf{m}_y, \mathbf{P}_y)$]. In (12-2),

$$\mathbf{m}_y = \mathbf{E}\{\mathbf{y}\} \quad (12-3)$$

and

$$\mathbf{P}_y = \mathbf{E}\{(\mathbf{y} - \mathbf{m}_y)(\mathbf{y} - \mathbf{m}_y)'\} \quad (12-4)$$

Note that, although we refer to $p(\mathbf{y})$ as a density function, it is actually a *joint density function* between the random variables y_1, y_2, \dots, y_m . If \mathbf{P}_y is not positive definite, it is more convenient to define the multivariate Gaussian distribution by its characteristic function. We will not need to do this.

JOINTLY GAUSSIAN RANDOM VECTORS

Let \mathbf{x} and \mathbf{y} individually be n - and m -dimensional Gaussian random vectors, i.e., $\mathbf{x} \sim N(\mathbf{x}; \mathbf{m}_x, \mathbf{P}_x)$ and $\mathbf{y} \sim N(\mathbf{y}; \mathbf{m}_y, \mathbf{P}_y)$. Let \mathbf{P}_{xy} and \mathbf{P}_{yx} denote the cross-covariance matrices between \mathbf{x} and \mathbf{y} , i.e.,

$$\mathbf{P}_{xy} = \mathbf{E}\{(\mathbf{x} - \mathbf{m}_x)(\mathbf{y} - \mathbf{m}_y)'\} \quad (12-5)$$

and

$$\mathbf{P}_{yx} = \mathbf{E}\{(\mathbf{y} - \mathbf{m}_y)(\mathbf{x} - \mathbf{m}_x)'\} \quad (12-6)$$

We are interested in the joint density between \mathbf{x} and \mathbf{y} , $p(\mathbf{x}, \mathbf{y})$. Vectors \mathbf{x} and \mathbf{y} are jointly Gaussian if

$$p(\mathbf{x}, \mathbf{y}) = \frac{1}{\sqrt{(2\pi)^{n+m} |\mathbf{P}_z|}} \exp\left\{-\frac{1}{2} (\mathbf{z} - \mathbf{m}_z)' \mathbf{P}_z^{-1} (\mathbf{z} - \mathbf{m}_z)\right\} \quad (12-7)$$

where

$$\mathbf{z} = \text{col}(\mathbf{x}, \mathbf{y}) \quad (12-8)$$

$$\mathbf{m}_z = \text{col}(\mathbf{m}_x, \mathbf{m}_y) \quad (12-9)$$

and

$$\mathbf{P}_z = \begin{pmatrix} \mathbf{P}_x & \mathbf{P}_{xy} \\ \mathbf{P}_{yx} & \mathbf{P}_y \end{pmatrix} \quad (12-10)$$

Note that if \mathbf{x} and \mathbf{y} are jointly Gaussian then they are marginally (i.e., individually) Gaussian. The converse is true if \mathbf{x} and \mathbf{y} are independent, but it is not necessarily true if they are not independent (Papoulis, 1965, pg. 184).

In order to evaluate $p(\mathbf{x}, \mathbf{y})$ in (12-7), we need $|\mathbf{P}_z|$ and \mathbf{P}_z^{-1} . It is straightforward to compute $|\mathbf{P}_z|$ once the given values for \mathbf{P}_x , \mathbf{P}_y , \mathbf{P}_{xy} , and \mathbf{P}_{yx} are substituted into (12-10). It is often useful to be able to express the components of \mathbf{P}_z^{-1} directly in terms of the components of \mathbf{P}_z . It is a straightforward exercise in algebra (just form $\mathbf{P}_z \mathbf{P}_z^{-1} = \mathbf{I}$ and equate elements on both sides) to show that

$$\mathbf{P}_z^{-1} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}' & \mathbf{C} \end{pmatrix} \quad (12-11)$$

where

$$\mathbf{A} = (\mathbf{P}_x - \mathbf{P}_{xy}\mathbf{P}_y^{-1}\mathbf{P}_{yx})^{-1} = \mathbf{P}_x^{-1} + \mathbf{P}_x^{-1}\mathbf{P}_{xy}\mathbf{C}\mathbf{P}_{yx}\mathbf{P}_x^{-1} \quad (12-12)$$

$$\mathbf{B} = -\mathbf{A}\mathbf{P}_{xy}\mathbf{P}_y^{-1} = -\mathbf{P}_x^{-1}\mathbf{P}_{xy}\mathbf{C} \quad (12-13)$$

and

$$\mathbf{C} = (\mathbf{P}_y - \mathbf{P}_{yx}\mathbf{P}_x^{-1}\mathbf{P}_{xy})^{-1} = \mathbf{P}_y^{-1} + \mathbf{P}_y^{-1}\mathbf{P}_{yx}\mathbf{A}\mathbf{P}_{xy}\mathbf{P}_y^{-1} \quad (12-14)$$

THE CONDITIONAL DENSITY FUNCTION

One of the most important density functions we will be interested in is the conditional density function $p(\mathbf{x}|\mathbf{y})$. Recall, from probability theory (e.g., Papoulis, 1965), that

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{x}, \mathbf{y})}{p(\mathbf{y})} \quad (12-15)$$

Theorem 12-1. Let \mathbf{x} and \mathbf{y} be n - and m -dimensional vectors that are jointly Gaussian. Then

$$p(\mathbf{x}|\mathbf{y}) = \frac{1}{\sqrt{(2\pi)^n |\mathcal{Q}|}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \mathbf{m})' \mathcal{Q}^{-1} (\mathbf{x} - \mathbf{m}) \right\} \quad (12-16)$$

where

$$\mathbf{m} = \mathbf{E}[\mathbf{x}|\mathbf{y}] = \mathbf{m}_x + \mathbf{P}_{xy}\mathbf{P}_y^{-1}(\mathbf{y} - \mathbf{m}_y) \quad (12-17)$$

and

$$\mathcal{Q} = \mathbf{A}^{-1} = \mathbf{P}_x - \mathbf{P}_{xy}\mathbf{P}_y^{-1}\mathbf{P}_{yx} \quad (12-18)$$

This means that $p(\mathbf{x}|\mathbf{y})$ is also multivariate Gaussian with (conditional) mean \mathbf{m} and covariance \mathcal{Q} .

Proof. From (12-15), (12-7), and (12-2), we find that

$$p(\mathbf{x}|\mathbf{y}) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{P}_z|}} \exp \left[-\frac{1}{2} (\mathbf{z} - \mathbf{m}_z)' \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}' & \mathbf{C} - \mathbf{P}_y^{-1} \end{pmatrix} (\mathbf{z} - \mathbf{m}_z) \right] \quad (12-19)$$

The Conditional Density Function

Taking a closer look at the quadratic exponent, which we denote $E(\mathbf{x}, \mathbf{y})$, we find that

$$\begin{aligned} E(\mathbf{x}, \mathbf{y}) &= (\mathbf{x} - \mathbf{m}_x)' \mathbf{A} (\mathbf{x} - \mathbf{m}_x) + 2(\mathbf{x} - \mathbf{m}_x)' \mathbf{B} (\mathbf{y} - \mathbf{m}_y) \\ &\quad + (\mathbf{y} - \mathbf{m}_y)' (\mathbf{C} - \mathbf{P}_y^{-1}) (\mathbf{y} - \mathbf{m}_y) \\ &= (\mathbf{x} - \mathbf{m}_x)' \mathbf{A} (\mathbf{x} - \mathbf{m}_x) - 2(\mathbf{x} - \mathbf{m}_x)' \mathbf{A} \mathbf{P}_{xy} \mathbf{P}_y^{-1} (\mathbf{y} - \mathbf{m}_y) \\ &\quad + (\mathbf{y} - \mathbf{m}_y)' \mathbf{P}_{yx} \mathbf{A} \mathbf{P}_{xy} \mathbf{P}_y^{-1} (\mathbf{y} - \mathbf{m}_y) \end{aligned} \quad (12-20)$$

In obtaining (12-20) we have used (12-13) and (12-14). We now recognize that (12-20) looks like a quadratic expression in $\mathbf{x} - \mathbf{m}_x$ and $\mathbf{P}_{xy} \mathbf{P}_y^{-1} (\mathbf{y} - \mathbf{m}_y)$, and express it in factored form, as

$$E(\mathbf{x}, \mathbf{y}) = [(\mathbf{x} - \mathbf{m}_x) - \mathbf{P}_{xy} \mathbf{P}_y^{-1} (\mathbf{y} - \mathbf{m}_y)]' \mathbf{A} [(\mathbf{x} - \mathbf{m}_x) - \mathbf{P}_{xy} \mathbf{P}_y^{-1} (\mathbf{y} - \mathbf{m}_y)] \quad (12-21)$$

Defining \mathbf{m} and \mathcal{Q} as in (12-18) we see that

$$E(\mathbf{x}, \mathbf{y}) = (\mathbf{x} - \mathbf{m})' \mathcal{Q}^{-1} (\mathbf{x} - \mathbf{m}) \quad (12-22)$$

If we can show that $|\mathbf{P}_z|/|\mathbf{P}_y| = |\mathcal{Q}|$, then we will have shown that $p(\mathbf{x}|\mathbf{y})$ is given by (12-16), which will mean that $p(\mathbf{x}|\mathbf{y})$ is multivariate Gaussian with mean \mathbf{m} and covariance \mathcal{Q} .

It is not at all obvious that $|\mathcal{Q}| = |\mathbf{P}_z|/|\mathbf{P}_y|$. We shall reexpress matrix \mathbf{P}_z so that $|\mathbf{P}_z|$ can be determined by appealing to the following theorems (Graybill, 1961, pg. 6):

i. If \mathbf{V} and \mathbf{G} are $n \times n$ matrices, then $|\mathbf{VG}| = |\mathbf{V}| |\mathbf{G}|$.

ii. If \mathbf{M} is a square matrix such that

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{21} & \mathbf{M}_{22} \end{pmatrix}$$

where \mathbf{M}_{11} and \mathbf{M}_{22} are square matrices, and if $\mathbf{M}_{12} = \mathbf{0}$ or $\mathbf{M}_{21} = \mathbf{0}$, then $|\mathbf{M}| = |\mathbf{M}_{11}| |\mathbf{M}_{22}|$.

We now show that two matrices, \mathbf{L} and \mathbf{N} , can be found so that

$$\mathbf{P}_z = \begin{pmatrix} \mathbf{P}_x & \mathbf{P}_{xy} \\ \mathbf{P}_{yx} & \mathbf{P}_y \end{pmatrix} = \begin{pmatrix} \mathbf{L} & \mathbf{P}_{xy} \\ \mathbf{0} & \mathbf{P}_y \end{pmatrix} \begin{pmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{N} & \mathbf{I}_m \end{pmatrix} \quad (12-23)$$

Multiplying the two matrices on the right-hand side of (12-23), and equating the 1-1 and 2-1 components from both sides of the resulting equation, we find that

$$\mathbf{P}_x = \mathbf{L} + \mathbf{P}_{xy}\mathbf{N} \quad (12-24)$$

and

$$\mathbf{P}_{yx} = \mathbf{P}_y\mathbf{N} \quad (12-25)$$

from which it follows that

$$\mathbf{N} = \mathbf{P}_y^{-1}\mathbf{P}_{yx} \quad (12-26)$$

and

$$\mathbf{L} = \mathbf{P}_x - \mathbf{P}_{xy}\mathbf{P}_y^{-1}\mathbf{P}_{yx} \quad (12-27)$$

From (12-23), we see that

$$|\mathbf{P}_z| = |\mathbf{L}| |\mathbf{P}_y|$$

or,

$$|\mathbf{L}| = |\mathbf{P}_z|/|\mathbf{P}_y| \quad (12-28)$$

Comparing the equations for \mathbf{L} and \mathbf{Q} , we find they are the same; thus, we have proven that

$$|\mathbf{Q}| = |\mathbf{P}_z|/|\mathbf{P}_y| \quad (12-29)$$

which completes the proof of this theorem. \square

PROPERTIES OF MULTIVARIATE GAUSSIAN RANDOM VARIABLES

From the preceding formulas for $p(\mathbf{y})$, $p(\mathbf{x},\mathbf{y})$ and $p(\mathbf{x}|\mathbf{y})$, we see that *multivariate Gaussian probability density functions are completely characterized by their first two moments*, i.e., their mean vector and covariance matrix. All other moments can be expressed in terms of their first two moments (see, e.g., Papoulis, 1965).

From probability theory, we also recall the following two important facts about Gaussian random variables:

1. Statistically independent Gaussian random variables are uncorrelated and vice-versa; thus, \mathbf{P}_{xy} and \mathbf{P}_{yx} are both zero matrices.
2. Linear (or affine) transformations on and linear (or affine) combinations of Gaussian random variables are themselves Gaussian random variables; thus, if \mathbf{x} and \mathbf{y} are jointly Gaussian, then $\mathbf{z} = \mathbf{Ax} + \mathbf{By} + \mathbf{c}$ is also Gaussian. We refer to this property as the *linearity property*.

PROPERTIES OF CONDITIONAL MEAN

We learned, in Theorem 12-1, that

$$\mathbf{E}\{\mathbf{x}|\mathbf{y}\} = \mathbf{m}_x + \mathbf{P}_{xy}\mathbf{P}_y^{-1}(\mathbf{y} - \mathbf{m}_y) \quad (12-30)$$

Because $\mathbf{E}\{\mathbf{x}|\mathbf{y}\}$ depends on \mathbf{y} , which is random, it is also random.

Theorem 12-2. *When \mathbf{x} and \mathbf{y} are jointly Gaussian, $\mathbf{E}\{\mathbf{x}|\mathbf{y}\}$ is multivariate Gaussian, and is an affine combination of the elements of \mathbf{y} .*

Properties of Conditional Mean

Proof. That $\mathbf{E}\{\mathbf{x}|\mathbf{y}\}$ is Gaussian follows from the linearity property applied to (12-30). An affine transformation of \mathbf{y} has the structure $T\mathbf{y} + \mathbf{f}$. $\mathbf{E}\{\mathbf{x}|\mathbf{y}\}$ has this structure; thus, it is an affine transformation. Note that if $\mathbf{m}_x = \mathbf{0}$ and $\mathbf{m}_y = \mathbf{0}$, then $\mathbf{E}\{\mathbf{x}|\mathbf{y}\}$ is a linear transformation of \mathbf{y} . \square

Theorem 12-3. *Let \mathbf{x} , \mathbf{y} , and \mathbf{z} be $n \times 1$, $m \times 1$ and $r \times 1$ jointly Gaussian random vectors. If \mathbf{y} and \mathbf{z} are statistically independent, then*

$$\mathbf{E}\{\mathbf{x}|\mathbf{y}, \mathbf{z}\} = \mathbf{E}\{\mathbf{x}|\mathbf{y}\} + \mathbf{E}\{\mathbf{x}|\mathbf{z}\} - \mathbf{m}_x \quad (12-31)$$

Proof. Let $\xi = \text{col}(\mathbf{y}, \mathbf{z})$; then

$$\mathbf{E}\{\mathbf{x}|\xi\} = \mathbf{m}_x + \mathbf{P}_{x\xi}\mathbf{P}_\xi^{-1}(\xi - \mathbf{m}_\xi) \quad (12-32)$$

We leave it to the reader to show that

$$\mathbf{P}_{x\xi} = (\mathbf{P}_{xy} | \mathbf{P}_{xz}) \quad (12-33)$$

and

$$\mathbf{P}_\xi = \begin{pmatrix} \mathbf{P}_y & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_z \end{pmatrix} \quad (12-34)$$

(the off-diagonal elements in \mathbf{P}_ξ are zero if \mathbf{y} and \mathbf{z} are statistically independent; this is also true if \mathbf{y} and \mathbf{z} are uncorrelated, because \mathbf{y} and \mathbf{z} are jointly Gaussian); thus,

$$\begin{aligned} \mathbf{E}\{\mathbf{x}|\xi\} &= \mathbf{m}_x + \mathbf{P}_{xy}\mathbf{P}_y^{-1}(\mathbf{y} - \mathbf{m}_y) + \mathbf{P}_{xz}\mathbf{P}_z^{-1}(\mathbf{z} - \mathbf{m}_z) \\ &= \mathbf{E}\{\mathbf{x}|\mathbf{y}\} + \mathbf{E}\{\mathbf{x}|\mathbf{z}\} - \mathbf{m}_x \quad \square \end{aligned}$$

In our developments of recursive estimators, \mathbf{y} and \mathbf{z} (which will be associated with a partitioning of measurement vectors \mathbf{x}) are not necessarily independent. The following important generalization of Theorem 12-3 will be needed.

Theorem 12-4. *Let \mathbf{x} , \mathbf{y} , and \mathbf{z} be $n \times 1$, $m \times 1$ and $r \times 1$ jointly Gaussian random vectors. If \mathbf{y} and \mathbf{z} are not necessarily statistically independent, then*

$$\mathbf{E}\{\mathbf{x}|\mathbf{y}, \mathbf{z}\} = \mathbf{E}\{\mathbf{x}|\mathbf{y}, \bar{\mathbf{z}}\} \quad (12-35)$$

where

$$\bar{\mathbf{z}} = \mathbf{z} - \mathbf{E}\{\mathbf{z}|\mathbf{y}\} \quad (12-36)$$

so that

$$\mathbf{E}\{\mathbf{x}|\mathbf{y}, \mathbf{z}\} = \mathbf{E}\{\mathbf{x}|\mathbf{y}\} + \mathbf{E}\{\mathbf{x}|\bar{\mathbf{z}}\} - \mathbf{m}_x \quad (12-37)$$

Proof (Mendel, 1983b, pg. 53). The proof proceeds in two stages: (a) assume (12-35) is true and demonstrate the truth of (12-37), and (b) demonstrate the truth of (12-35).

- a. If we can show that y and \tilde{z} are statistically independent, then (12-37) follows from Theorem 12-3. For Gaussian random vectors, however, uncorrelatedness implies independence.

To begin, we assert that y and \tilde{z} are jointly Gaussian, because $\tilde{z} = z - E\{z|y\} = z - m_z - P_y P_y^{-1}(y - m_y)$ depends on y and z , which are jointly Gaussian.

Next, we show that \tilde{z} is zero mean. This follows from the calculation

$$m_{\tilde{z}} = E\{z - E\{z|y\}\} = E\{z\} - E\{E\{z|y\}\} \quad (12-38)$$

where the outer expectation in the second term on the right-hand side of (12-38) is with respect to y . From probability theory (Papoulis, 1965, pg. 208), $E\{z\}$ can be expressed as

$$E\{z\} = E\{E\{z|y\}\} \quad (12-39)$$

From (12-38) and (12-39), we see that $m_{\tilde{z}} = 0$.

Finally, we show that y and \tilde{z} are uncorrelated. This follows from the calculation

$$\begin{aligned} E\{(y - m_y)(\tilde{z} - m_{\tilde{z}})\}' &= E\{(y - m_y)\tilde{z}'\} = E\{y\tilde{z}'\} \\ &= E\{yz'\} - E\{yE\{z'|y\}\} \quad (12-40) \\ &= E\{yz'\} - E\{yz'\} = 0 \end{aligned}$$

- b. A detailed proof is given by Meditch (1969, pp. 101–102). The idea is to (1) compute $E\{x|y,z\}$ in expanded form, (2) compute $E\{x|y,\tilde{z}\}$ in expanded form, using \tilde{z} given in (12-36), and (3) compare the results from (1) and (2) to prove the truth of (12-35). \square

Equation (12-35) is very important. It states that, when z and y are dependent, conditioning on z can always be replaced by conditioning on another Gaussian random vector \tilde{z} , where \tilde{z} and y are statistically independent.

The results in Theorems 12-3 and 12-4 depend on all random vectors being jointly Gaussian. Very similar results, which are distribution free, are described in Problem 13-4; however, these results are restricted in yet another way.

PROBLEMS

- 12-1.** Fill in all of the details required to prove part b of Theorem 12-4.
- 12-2.** Let x , y and z be jointly distributed random vectors; c and h fixed constants; and $g(\cdot)$ a scalar-valued function. Assume $E\{x\}$, $E\{z\}$, and $E\{g(y)x\}$ exist. Prove the following useful properties of conditional expectation:

- (a) $E\{x|y\} = E\{x\}$ if x and y are independent
- (b) $E\{g(y)x|y\} = g(y)E\{x|y\}$
- (c) $E\{c|y\} = c$
- (d) $E\{g(y)|y\} = g(y)$
- (e) $E\{cx + hz|y\} = cE\{x|y\} + hE\{z|y\}$
- (f) $E\{x\} = E\{E\{x|y\}\}$ where the outer expectation is with respect to y
- (g) $E\{g(y)x\} = E\{g(y)E\{x|y\}\}$ where the outer expectation is with respect to y .

- 12-3.** Prove that the cross-covariance matrix of two uncorrelated random vectors is zero.

Lesson 13

Estimation of Random Parameters: General Results

INTRODUCTION

In Lesson 2 we showed that state estimation and deconvolution can be viewed as problems in which we are interested in estimating a vector of random parameters. For us, state estimation and deconvolution serve as the primary motivation for studying methods for estimating random parameters; however, the statistics literature is filled with other applications for these methods.

We now view θ as an $n \times 1$ vector of random unknown parameters. The information available to us are measurements $z(1), z(2), \dots, z(k)$, which are assumed to depend upon θ . In this lesson, we do not begin by assuming a specific structural dependency between $z(i)$ and θ . This is quite different than what we did in WLSE and BLUE. Those methods were studied for the linear model $z(i) = H(i)\theta + v(i)$, and, closed-form solutions for $\hat{\theta}_{\text{WLS}}(k)$ and $\hat{\theta}_{\text{BLU}}(k)$ could not have been obtained had we not begun by assuming the linear model. We shall study the estimation of random θ for the linear model in Lesson 14.

In this lesson we examine two methods for estimating a vector of random parameters. The first method is based upon minimizing the mean-squared error between θ and $\hat{\theta}(k)$. The resulting estimator is called a *mean-squared estimator*, and is denoted $\hat{\theta}_{\text{MS}}(k)$. The second method is based upon maximizing an unconditional likelihood function, one that not only requires knowledge of $p(\mathcal{L}|\theta)$ but also of $p(\theta)$. The resulting estimator is called a *maximum a posteriori estimator*, and is denoted $\hat{\theta}_{\text{MAP}}(k)$.

MEAN-SQUARED ESTIMATION

Many different measures of parameter estimation error can be minimized in order to obtain an estimate of θ [see Jazwinski (1970), Meditch (1969), van Trees (1968), and Sorenson (1980), for example]; but, by far the most widely studied measure is the mean-squared error.

Objective Function and Problem Statement

Given measurements $z(1), \dots, z(k)$, we shall determine an estimator of θ , namely

$$\hat{\theta}_{\text{MS}}(k) = \phi[z(i), i = 1, 2, \dots, k] \quad (13-1)$$

such that the mean-squared error

$$J[\hat{\theta}_{\text{MS}}(k)] = E\{\tilde{\theta}_{\text{MS}}(k)\tilde{\theta}_{\text{MS}}(k)\} \quad (13-2)$$

is minimized. In (13-2), $\tilde{\theta}_{\text{MS}}(k) = \theta - \hat{\theta}_{\text{MS}}(k)$.

The right-hand side of (13-1) means that we have some arbitrary and as yet unknown function of all the measurements. The n components of $\hat{\theta}_{\text{MS}}(k)$ may each depend differently on the measurements. The function $\phi[z(i), i = 1, 2, \dots, k]$ may be nonlinear or linear. Its exact structure will be determined by minimizing $J[\hat{\theta}_{\text{MS}}(k)]$. If perchance $\hat{\theta}_{\text{MS}}(k)$ is a linear estimator, then

$$\hat{\theta}_{\text{MS}}(k) = \sum_{i=1}^k A(i)z(i) \quad (13-3)$$

We now show that the notion of conditional expectation is central to the calculation of $\hat{\theta}_{\text{MS}}(k)$. As usual, we let

$$\mathcal{L}(k) = \text{col}[z(k), z(k-1), \dots, z(1)] \quad (13-4)$$

The underlying random quantities in our estimation problem are θ and $\mathcal{L}(k)$. We assume that their joint density function $p[\theta, \mathcal{L}(k)]$ exists, so that

$$E\{\tilde{\theta}_{\text{MS}}(k)\tilde{\theta}_{\text{MS}}(k)\} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \tilde{\theta}_{\text{MS}}(k)\tilde{\theta}_{\text{MS}}(k)p[\theta, \mathcal{L}(k)]d\theta d\mathcal{L}(k) \quad (13-5)$$

where $d\theta = d\theta_1 d\theta_2 \dots d\theta_n$, $d\mathcal{L}(k) = dz_1(1) \dots dz_1(k)dz_2(1) \dots dz_2(k) \dots dz_m(1) \dots dz_m(k)$, and there are $n + km$ integrals. Using the fact that

$$p[\theta, \mathcal{L}(k)] = p[\theta | \mathcal{L}(k)]p[\mathcal{L}(k)] \quad (13-6)$$

we rewrite (13-5) as

$$\begin{aligned} E\{\tilde{\theta}_{\text{MS}}(k)\tilde{\theta}_{\text{MS}}(k)\} &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \tilde{\theta}_{\text{MS}}(k)\tilde{\theta}_{\text{MS}}(k)p[\theta | \mathcal{L}(k)]d\theta \right\} \\ &\quad p[\mathcal{L}(k)]d\mathcal{L}(k) \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} E\{\tilde{\theta}_{\text{MS}}(k)\tilde{\theta}_{\text{MS}}(k) | \mathcal{L}(k)\}p[\mathcal{L}(k)]d\mathcal{L}(k) \end{aligned} \quad (13-7)$$

From this equation we see that minimizing the conditional expectation $E\{\hat{\theta}'_{MS}(k)\hat{\theta}_{MS}(k)|\mathfrak{L}(k)\}$ with respect to $\hat{\theta}_{MS}(k)$ is equivalent to our original objective of minimizing the total expectation $E\{\hat{\theta}'_{MS}(k)\hat{\theta}_{MS}(k)\}$. Note that the integrals on the right-hand side of (13-7) remove the dependency of the integrand on the data $\mathfrak{L}(k)$.

In summary, we have the following *mean-squared estimation problem*: Given the measurements $\mathbf{z}(1), \mathbf{z}(2), \dots, \mathbf{z}(k)$, determine an estimator of θ , namely,

$$\hat{\theta}_{MS}(k) = \phi[\mathbf{z}(i), i = 1, 2, \dots, k]$$

such that the conditional mean-squared error

$$J_1[\hat{\theta}_{MS}(k)] = E\{\hat{\theta}'_{MS}(k)\hat{\theta}_{MS}(k)|\mathbf{z}(1), \dots, \mathbf{z}(k)\} \quad (13-8)$$

is minimized.

Derivation of Estimator

The solution to the mean-squared estimation problem is given in Theorem 13-1, which is known as the *Fundamental Theorem of Estimation Theory*.

Theorem 13-1. The estimator that minimizes the mean-squared error is

$$\hat{\theta}_{MS}(k) = E\{\theta|\mathfrak{L}(k)\} \quad (13-9)$$

Proof (Mendel, 1983b). In this proof we omit all functional dependences on k , for notational simplicity. Our approach is to substitute $\hat{\theta}_{MS}(k) = \theta - \hat{\theta}'_{MS}(k)$ into (13-8) and to complete the square, as follows:

$$\begin{aligned} J_1[\hat{\theta}_{MS}(k)] &= E\{(\theta - \hat{\theta}_{MS})'(\theta - \hat{\theta}_{MS})|\mathfrak{L}\} \\ &= E\{\theta'\theta - \theta'\hat{\theta}_{MS} - \hat{\theta}'_{MS}\theta + \hat{\theta}'_{MS}\hat{\theta}_{MS}|\mathfrak{L}\} \\ &= E\{\theta'\theta|\mathfrak{L}\} - E\{\theta'|\mathfrak{L}\}\hat{\theta}_{MS} - \hat{\theta}'_{MS}E\{\theta|\mathfrak{L}\} + \hat{\theta}'_{MS}\hat{\theta}_{MS} \\ &= E\{\theta'\theta|\mathfrak{L}\} + [\hat{\theta}_{MS} - E\{\theta|\mathfrak{L}\}]'[\hat{\theta}_{MS} - E\{\theta|\mathfrak{L}\}] \\ &\quad - E\{\theta'|\mathfrak{L}\}E\{\theta|\mathfrak{L}\} \end{aligned} \quad (13-10)$$

To obtain the third line we used the fact that $\hat{\theta}_{MS}$, by definition, is a function of \mathfrak{L} ; hence, $E\{\hat{\theta}_{MS}|\mathfrak{L}\} = \hat{\theta}_{MS}$. The first and last terms in (13-10) do not depend on $\hat{\theta}_{MS}$; hence, the smallest value of $J_1[\hat{\theta}_{MS}(k)]$ is obviously attained by setting the bracketed terms equal to zero. This means that $\hat{\theta}_{MS}$ must be chosen as in (13-9). \square

Let $J_1^*[\hat{\theta}_{MS}(k)]$ denote the minimum value of $J_1[\hat{\theta}_{MS}(k)]$. We see, from (13-10) and (13-9), that

$$J_1^*[\hat{\theta}_{MS}(k)] = E\{\theta'\theta|\mathfrak{L}\} - \hat{\theta}'_{MS}(k)\hat{\theta}_{MS}(k) \quad (13-11)$$

As it stands, (13-9) is not terribly useful for computing $\hat{\theta}_{MS}(k)$. In general, we must first compute $p[\theta|\mathfrak{L}(k)]$ and then perform the requisite number of integrations of $\theta p[\theta|\mathfrak{L}(k)]$ to obtain $\hat{\theta}_{MS}(k)$. In the special but important case

when θ and $\mathfrak{L}(k)$ are jointly Gaussian we have a very important and practical corollary to Theorem 13-1.

Corollary 13-1. When θ and $\mathfrak{L}(k)$ are jointly Gaussian, the estimator that minimizes the mean-squared error is

$$\hat{\theta}_{MS}(k) = \mathbf{m}_\theta + \mathbf{P}_{\theta\mathfrak{L}}(k)\mathbf{P}_{\mathfrak{L}}^{-1}(k)[\mathfrak{L}(k) - \mathbf{m}_{\mathfrak{L}}(k)] \quad (13-12)$$

Proof. When θ and $\mathfrak{L}(k)$ are jointly Gaussian then $E\{\theta|\mathfrak{L}(k)\}$ can be evaluated using (12-17) of Lesson 12. Doing this we obtain (13-12). \square

Corollary 13-1 gives us an explicit structure for $\hat{\theta}_{MS}(k)$. We see that $\hat{\theta}_{MS}(k)$ is an affine transformation of $\mathfrak{L}(k)$. If $\mathbf{m}_\theta = \mathbf{0}$ and $\mathbf{m}_{\mathfrak{L}}(k) = \mathbf{0}$, then $\hat{\theta}_{MS}(k)$ is a linear transformation of $\mathfrak{L}(k)$.

In order to compute $\hat{\theta}_{MS}(k)$ using (13-12), we must know \mathbf{m}_θ and $\mathbf{m}_{\mathfrak{L}}(k)$ and we must first compute $\mathbf{P}_{\theta\mathfrak{L}}(k)$ and $\mathbf{P}_{\mathfrak{L}}^{-1}(k)$. We perform these computations in Lesson 14 for the linear model, $\mathfrak{L}(k) = \mathcal{H}(k)\theta + \mathcal{V}(k)$.

Corollary 13-2. Suppose θ and $\mathfrak{L}(k)$ are not necessarily jointly Gaussian, and that we know \mathbf{m}_θ , $\mathbf{m}_{\mathfrak{L}}(k)$, $\mathbf{P}_{\mathfrak{L}}(k)$ and $\mathbf{P}_{\theta\mathfrak{L}}(k)$. In this case, the estimator that is constrained to be an affine transformation of $\mathfrak{L}(k)$, and that minimizes the mean-squared error is also given by (13-12).

Proof. This corollary can be proved in a number of ways. A direct proof begins by assuming that $\hat{\theta}_{MS}(k) = \mathbf{A}(k)\mathfrak{L}(k) + \mathbf{b}(k)$ and choosing $\mathbf{A}(k)$ and $\mathbf{b}(k)$ so that $\hat{\theta}_{MS}(k)$ is an unbiased estimator of θ and $E\{\hat{\theta}'_{MS}(k)\hat{\theta}_{MS}(k)\} = \text{trace } E\{\hat{\theta}_{MS}(k)\hat{\theta}'_{MS}(k)\}$ is minimized. We leave the details of this direct proof to the reader.

A less direct proof is based upon the following Gedanken experiment. Using known first and second moments of θ and $\mathfrak{L}(k)$, we can conceptualize unique Gaussian random vectors that have these same first and second moments. For these statistically-equivalent (through second-order moments) Gaussian vectors, we know, from Corollary 13-1, that the mean-squared estimator is given by the affine transformation of $\mathfrak{L}(k)$ in (13-12). \square

Corollaries 13-1 and 13-2, as well as Theorem 13-1, provide us with the answer to the following important question: *When is the linear (affine) mean-squared estimator the same as the mean-squared estimator?* The answer is, when θ and $\mathfrak{L}(k)$ are jointly Gaussian. If θ and $\mathfrak{L}(k)$ are not jointly Gaussian, then $\hat{\theta}_{MS}(k) = E\{\theta|\mathfrak{L}(k)\}$, which, in general, is a nonlinear function of measurements $\mathfrak{L}(k)$, i.e., it is a nonlinear estimator.

Corollary 13-3 (Orthogonality Principle). Suppose $\mathbf{f}[\mathfrak{L}(k)]$ is any function of the data $\mathfrak{L}(k)$. Then the error in the mean-squared estimator is orthogonal to $\mathbf{f}[\mathfrak{L}(k)]$ in the sense that

$$E\{[\theta - \hat{\theta}_{MS}(k)]\mathbf{f}'[\mathfrak{L}(k)]\} = \mathbf{0} \quad (13-13)$$

Proof (Mendel, 1983b, pp. 46–47). We use the following result from probability theory (Papoulis, 1965; see, also, Problem 12-2(g)). Let α and β be jointly distributed random vectors and $g(\beta)$ be a scalar-valued function; then

$$E\{\alpha g(\beta)\} = E\{E\{\alpha|\beta\} g(\beta)\} \quad (13-14)$$

where the outer expectation on the right-hand side is with respect to β . We proceed as follows (again, omitting the argument k):

$$\begin{aligned} E\{(\theta - \hat{\theta}_{MS})f'(\mathcal{Z})\} &= E\{E\{(\theta - \hat{\theta}_{MS})|\mathcal{Z}\}f'(\mathcal{Z})\} \\ &= E\{(\hat{\theta}_{MS} - \hat{\theta}_{MS})f'(\mathcal{Z})\} = 0 \end{aligned}$$

where we have used the facts that $\hat{\theta}_{MS}$ is no longer random when \mathcal{Z} is specified and $E\{\theta|\mathcal{Z}\} = \hat{\theta}_{MS}$. \square

A frequently encountered special case of (13-13) occurs when $f[\mathcal{Z}(k)] = \hat{\theta}_{MS}(k)$; then Corollary 13-13 can be written as

$$E\{\tilde{\theta}_{MS}(k)\hat{\theta}'_{MS}(k)\} = 0 \quad (13-15)$$

Properties of Mean-Squared Estimates When θ and $\mathcal{Z}(k)$ are Gaussian

In this section we present a collection of important and useful properties associated with $\hat{\theta}_{MS}(k)$ for the case when θ and $\mathcal{Z}(k)$ are jointly Gaussian. In this case $\hat{\theta}_{MS}(k)$ is given by (13-12).

Property 1 (Unbiasedness). The mean-squared estimator, $\hat{\theta}_{MS}(k)$ in (13-12), is unbiased.

Proof. Taking the expected value of (13-12), we see that $E\{\hat{\theta}_{MS}(k)\} = \mathbf{m}_\theta$; thus, $\hat{\theta}_{MS}(k)$ is an unbiased estimator of θ . \square

Property 2 (Minimum Variance). Dispersion about the mean value of $\hat{\theta}_{MS}(k)$ is measured by the error variance $\sigma_{\hat{\theta}_{MS}}^2(k)$, where $i = 1, 2, \dots, n$. An estimator that has the smallest error variance is a minimum-variance estimator (an MVE). The mean-squared estimator in (13-12) is an MVE.

Proof. From Property 1 and the definition of error variance, we see that

$$\sigma_{\hat{\theta}_{i,MS}}^2(k) = E\{\hat{\theta}_{i,MS}^2(k)\} - i = 1, 2, \dots, n \quad (13-16)$$

Our mean-squared estimator was obtained by minimizing $J[\hat{\theta}_{MS}(k)]$ in (13-2), which can now be expressed as

$$J[\hat{\theta}_{MS}(k)] = \sum_{i=1}^n \sigma_{\hat{\theta}_{i,MS}}^2(k) \quad (13-17)$$

Mean-Squared Estimation

Because variances are always positive the minimum value of $J[\hat{\theta}_{MS}(k)]$ must be achieved when each of the n variances is minimized; hence, our mean-squared estimator is equivalent to an MVE. \square

Property 3 (Linearity). $\hat{\theta}_{MS}(k)$ in (13-12) is a “linear” (i.e., affine) estimator.

Proof. This is obvious from the form of (13-12). \square

Linearity of $\hat{\theta}_{MS}(k)$ permits us to infer the following very important property about both $\hat{\theta}_{MS}(k)$ and $\tilde{\theta}_{MS}(k)$.

Property 4 (Gaussian). Both $\hat{\theta}_{MS}(k)$ and $\tilde{\theta}_{MS}(k)$ are multivariate Gaussian.

Proof. We use the linearity property of jointly Gaussian random vectors stated in Lesson 12. Estimator $\hat{\theta}_{MS}(k)$ in (13-12) is an affine transformation of Gaussian random vector $\mathcal{Z}(k)$; hence, $\hat{\theta}_{MS}(k)$ is multivariate Gaussian. Estimation error $\tilde{\theta}_{MS}(k) = \theta - \hat{\theta}_{MS}(k)$ is an affine transformation of jointly Gaussian vectors θ and $\mathcal{Z}(k)$; hence, $\tilde{\theta}_{MS}(k)$ is also Gaussian. \square

Estimate $\hat{\theta}_{MS}(k)$ in (13-12) is itself random, because measurements $\mathcal{Z}(k)$ are random. To characterize it completely in a statistical sense, we must specify its probability density function. Generally, this is very difficult to do, and often requires that the probability density function of $\hat{\theta}_{MS}(k)$ be approximated using many moments (in theory an infinite number are required). In the Gaussian case, we have just learned that the structure of the probability density function for $\hat{\theta}_{MS}(k)$ [and $\tilde{\theta}_{MS}(k)$] is known. Additionally, we know that a Gaussian density function is completely specified by exactly two moments, its mean and covariance; thus, tremendous simplifications occur when θ and $\mathcal{Z}(k)$ are jointly Gaussian.

Property 5 (Uniqueness). Mean-squared estimator $\hat{\theta}_{MS}(k)$, in (13-12), is unique.

The proof of this property is not central to our developments; hence, it is omitted.

Generalizations

Many of the results presented in this section are applicable to objective functions other than the mean-squared objective function in (13-2). See Meditch (1969) for discussions on a wide number of objective functions that lead to $E\{\theta|\mathcal{Z}(k)\}$ as the optimal estimator of θ .

MAXIMUM A POSTERIORI ESTIMATION

Recall Bayes's rule (Papoulis, 1965, pg. 39):

$$p(\theta|\mathcal{L}(k)) = p(\mathcal{L}(k)|\theta)p(\theta)/p(\mathcal{L}(k)) \quad (13-18)$$

in which density function $p(\theta|\mathcal{L}(k))$ is known as the a posteriori (or posterior) conditional density function, and $p(\theta)$ is the prior probability density function for θ . Observe that $p(\theta|\mathcal{L}(k))$ is related to likelihood function $l(\theta|\mathcal{L}(k))$, because $l(\theta|\mathcal{L}(k)) \propto p(\mathcal{L}(k)|\theta)$. Additionally, because $p(\mathcal{L}(k))$ does not depend on θ ,

$$p(\theta|\mathcal{L}(k)) \propto p(\mathcal{L}(k)|\theta)p(\theta) \quad (13-19)$$

In maximum a posteriori (MAP) estimation, values of θ are found that maximize $p(\theta|\mathcal{L}(k))$ in (13-19); such estimates are known as MAP estimates, and will be denoted as $\hat{\theta}_{\text{MAP}}(k)$.

If $\theta_1, \theta_2, \dots, \theta_n$ are uniformly distributed, then $p(\theta|\mathcal{L}(k)) \propto p(\mathcal{L}(k)|\theta)$, and the MAP estimator of θ equals the ML estimator of θ . Generally, MAP estimates are quite different from ML estimates. For example, the invariance property of MLE's usually does not carry over to MAP estimates. One reason for this can be seen from (13-19). Suppose, for example, that $\phi = g(\theta)$ and we want to determine $\hat{\phi}_{\text{MAP}}$ by first computing $\hat{\theta}_{\text{MAP}}$. Because $p(\theta)$ depends on the Jacobian matrix of $g^{-1}(\phi)$, $\hat{\phi}_{\text{MAP}} \neq g(\hat{\theta}_{\text{MAP}})$. Kashyap and Rao (1976, pg. 137) note, "the two estimates are usually asymptotically identical to one another since in the large sample case the knowledge of the observations *swamps that of the prior distribution*." For additional discussions on the asymptotic properties of MAP estimators, see Zacks (1971).

Quantity $p(\theta|\mathcal{L}(k))$ in (13-19) is sometimes called an *unconditional likelihood function*, because the random nature of θ has been accounted for by $p(\theta)$. Density $p(\mathcal{L}(k)|\theta)$ is then called a *conditional likelihood function* (Nahi, 1969).

Obtaining a MAP estimate involves specifying both $p(\mathcal{L}(k)|\theta)$ and $p(\theta)$ and finding those values of θ that maximize $p(\theta|\mathcal{L}(k))$, or $\ln p(\theta|\mathcal{L}(k))$. Generally speaking, mathematical programming must be used to compute $\hat{\theta}_{\text{MAP}}(k)$. When $\mathcal{L}(k)$ is related to θ by our linear model, $\mathcal{L}(k) = \mathcal{H}(k)\theta + \mathcal{V}(k)$, then it may be possible to obtain $\hat{\theta}_{\text{MAP}}(k)$ in closed form. We examine this situation in some detail in Lesson 14.

Example 13-1

This is a continuation of Example 11-1. We observe a random sample $\{z(1), z(2), \dots, z(N)\}$ at the output of a Gaussian random number generator, i.e., $z(i) \sim N(z(i); \mu, \sigma_z^2)$. Now, however, μ is a random variable with prior distribution $N(\mu; 0, \sigma_\mu^2)$. Both σ_z^2 and σ_μ^2 are assumed known, and we wish to determine the MAP estimator of μ .

We can view this random number generator as a cascade of two random number generators. The first is characterized by $N(\mu; 0, \sigma_\mu^2)$ and provides at its output a single realization for μ , say μ_R . The second is characterized by $N(z(i); \mu_R, \sigma_z^2)$. Observe that

Maximum a Posteriori Estimation

μ_R , which is unknown to us in this example, is transferred from the first random number generator to the second one before we can obtain the given random sample $\{z(1), z(2), \dots, z(N)\}$.

Using the facts that

$$p(z(i)|\mu) = (2\pi\sigma_z^2)^{-1/2} \exp\left\{-\frac{1}{2}[z(i) - \mu]^2/\sigma_z^2\right\} \quad (13-20)$$

$$p(\mu) = (2\pi\sigma_\mu^2)^{-1/2} \exp\left\{-\frac{1}{2}\mu^2/\sigma_\mu^2\right\} \quad (13-21)$$

and

$$p(\mathcal{L}(N)|\mu) = \prod_{i=1}^N p(z(i)|\mu) \quad (13-22)$$

we find

$$\begin{aligned} p(\mu|\mathcal{L}(N)) &\propto (2\pi\sigma_z^2)^{-N/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^N [z(i) - \mu]^2/\sigma_z^2\right\} \\ &\cdot (2\pi\sigma_\mu^2)^{-1/2} \exp\left\{-\frac{1}{2}\mu^2/\sigma_\mu^2\right\} \end{aligned} \quad (13-23)$$

Taking the logarithm of (13-23) and neglecting the terms which do not depend upon μ , we obtain

$$L_{\text{MAP}}(\mu|\mathcal{L}(N)) = -\frac{1}{2}\sum_{i=1}^N \{[z(i) - \mu]^2/\sigma_z^2\} - \frac{1}{2}\mu^2/\sigma_\mu^2 \quad (13-24)$$

Setting $\partial L_{\text{MAP}}/\partial \mu = 0$, and solving for $\hat{\mu}_{\text{MAP}}(N)$, we find that

$$\hat{\mu}_{\text{MAP}}(N) = \frac{\sigma_\mu^2}{\sigma_z^2 + N\sigma_\mu^2} \sum_{i=1}^N z(i) \quad (13-25)$$

Next, we compare $\hat{\mu}_{\text{MAP}}(N)$ and $\hat{\mu}_{\text{ML}}(N)$, where [see (19) from Lesson 11]

$$\hat{\mu}_{\text{ML}}(N) = \frac{1}{N} \sum_{i=1}^N z(i) \quad (13-26)$$

In general $\hat{\mu}_{\text{MAP}}(N) \neq \hat{\mu}_{\text{ML}}(N)$. If, however, no a priori information about μ is available, then we let $\sigma_\mu^2 \rightarrow \infty$, in which case $\hat{\mu}_{\text{MAP}}(N) = \hat{\mu}_{\text{ML}}(N)$. Observe, also that, as $N \rightarrow \infty$ $\hat{\mu}_{\text{MAP}}(N) = \hat{\mu}_{\text{ML}}(N)$, which implies (Sorenson, 1980) that the influence of the prior knowledge about μ [i.e., $\mu \sim N(\mu; 0, \sigma_\mu^2)$] diminishes as the number of measurements increase. \square

Theorem 13-2. *If $\mathcal{L}(k)$ and θ are jointly Gaussian, then $\hat{\theta}_{\text{MAP}}(k) = \hat{\theta}_{\text{MS}}(k)$.*

Proof. If $\mathcal{L}(k)$ and θ are jointly Gaussian, then (see Theorem 12-1)

$$\begin{aligned} p(\theta|\mathcal{L}(k)) &= \frac{1}{\sqrt{(2\pi)^n |\mathcal{L}(k)|}} \exp\left\{-\frac{1}{2}[\theta - \mathbf{m}(k)]' \mathcal{L}^{-1}(k)[\theta - \mathbf{m}(k)]\right\} \end{aligned} \quad (13-27)$$

where

$$\mathbf{m}(k) = \mathbb{E}\{\boldsymbol{\theta}|\mathfrak{Z}(k)\} \quad (13-28)$$

$\hat{\boldsymbol{\theta}}_{\text{MAP}}(k)$ is found by maximizing $p(\boldsymbol{\theta}|\mathfrak{Z}(k))$, or equivalently by minimizing the argument of the exponential in (13-27). The minimum value of $[\boldsymbol{\theta} - \mathbf{m}(k)]' \mathfrak{Z}^{-1}(k) [\boldsymbol{\theta} - \mathbf{m}(k)]$ is zero, and this occurs when

$$[\boldsymbol{\theta} - \mathbf{m}(k)] \Big|_{\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}_{\text{MAP}}(k)} = \mathbf{0} \quad (13-29)$$

i.e.,

$$\hat{\boldsymbol{\theta}}_{\text{MAP}}(k) = \mathbb{E}\{\boldsymbol{\theta}|\mathfrak{Z}(k)\} \quad (13-30)$$

Comparing (13-30) and (13-9) we conclude that $\hat{\boldsymbol{\theta}}_{\text{MAP}}(k) = \hat{\boldsymbol{\theta}}_{\text{MS}}(k)$. \square

The result in Theorems 13-2 is true regardless of the nature of the model relating $\boldsymbol{\theta}$ to $\mathfrak{Z}(k)$. Of course, in order to use it, we must first establish that $\mathfrak{Z}(k)$ and $\boldsymbol{\theta}$ are jointly Gaussian. Except for the linear model, which we examine in Lesson 14, this is very difficult to do.

PROBLEMS

- 13-1. Prove that $\hat{\boldsymbol{\theta}}_{\text{MS}}(k)$, given in (13-12), is unique.
- 13-2. Prove Corollary 13-2 by means of a direct proof.
- 13-3. Let $\boldsymbol{\theta}$ and $\mathfrak{Z}(N)$ be zero mean $n \times 1$ and $N \times 1$ random vectors, respectively, with known second-order statistics, $\mathbf{P}_{\boldsymbol{\theta}}$, $\mathbf{P}_{\mathfrak{Z}}$, $\mathbf{P}_{\boldsymbol{\theta}\mathfrak{Z}}$, and $\mathbf{P}_{\mathfrak{Z}\boldsymbol{\theta}}$. View $\mathfrak{Z}(N)$ as a vector of measurements. It is desired to determine a linear estimator of $\boldsymbol{\theta}$,

$$\hat{\boldsymbol{\theta}}(N) = \mathbf{K}_L(N) \mathfrak{Z}(N)$$

where $\mathbf{K}_L(N)$ is an $n \times N$ matrix that is chosen to minimize the mean-squared error $\mathbb{E}\{[\boldsymbol{\theta} - \mathbf{K}_L(N) \mathfrak{Z}(N)]' [\boldsymbol{\theta} - \mathbf{K}_L(N) \mathfrak{Z}(N)]\}$.

- (a) Show that the gain matrix, which minimizes the mean-squared error, is $\mathbf{K}_L(N) = \mathbb{E}\{\boldsymbol{\theta} \mathfrak{Z}'(N)\} [\mathbb{E}\{\mathfrak{Z}(N) \mathfrak{Z}'(N)\}]^{-1} = \mathbf{P}_{\boldsymbol{\theta}\mathfrak{Z}} \mathbf{P}_{\mathfrak{Z}}^{-1}$.
- (b) Show that the covariance matrix, $\mathbf{P}(N)$, of the estimation error, $\hat{\boldsymbol{\theta}}(N) - \boldsymbol{\theta}$, is

$$\mathbf{P}(N) = \mathbf{P}_{\boldsymbol{\theta}} - \mathbf{P}_{\boldsymbol{\theta}\mathfrak{Z}} \mathbf{P}_{\mathfrak{Z}}^{-1} \mathbf{P}_{\mathfrak{Z}\boldsymbol{\theta}}$$

- (c) Relate the results obtained in this problem to those in Corollary 13-2.
- 13-4. For random vectors $\boldsymbol{\theta}$ and $\mathfrak{Z}(k)$, the *linear projection*, $\boldsymbol{\theta}^*(k)$, of $\boldsymbol{\theta}$ on a Hilbert space spanned by $\mathfrak{Z}(k)$ is defined as $\boldsymbol{\theta}^*(k) = \mathbf{a} + \mathbf{B} \mathfrak{Z}(k)$, where $\mathbb{E}\{\boldsymbol{\theta}^*(k)\} = \mathbb{E}\{\boldsymbol{\theta}\}$ and $\mathbb{E}\{[\boldsymbol{\theta} - \boldsymbol{\theta}^*(k)] \mathfrak{Z}'(k)\} = \mathbf{0}$. We denote the linear projection, $\boldsymbol{\theta}^*(k)$, as $\mathbb{E}\{\boldsymbol{\theta}|\mathfrak{Z}(k)\}$.
 - (a) Prove that the linear (i.e., affine), unbiased mean-squared estimator of $\boldsymbol{\theta}$, $\hat{\boldsymbol{\theta}}(k)$, is the linear projection of $\boldsymbol{\theta}$ on $\mathfrak{Z}(k)$.
 - (b) Prove that the linear projection, $\boldsymbol{\theta}^*(k)$, of $\boldsymbol{\theta}$ on the Hilbert space spanned by

$\mathfrak{Z}(k)$ is uniquely equal to the linear (i.e., affine), unbiased mean-squared estimator of $\boldsymbol{\theta}$, $\hat{\boldsymbol{\theta}}(k)$.

- (c) For random vectors \mathbf{x} , \mathbf{y} , \mathbf{z} , where \mathbf{y} and \mathbf{z} are uncorrelated, prove that

$$\hat{\mathbb{E}}\{\mathbf{x}|\mathbf{y}, \mathbf{z}\} = \hat{\mathbb{E}}\{\mathbf{x}|\mathbf{y}\} + \hat{\mathbb{E}}\{\mathbf{x}|\mathbf{z}\} - \mathbf{m}_{\mathbf{x}}$$

- (d) For random vectors \mathbf{x} , \mathbf{y} , \mathbf{z} , where \mathbf{y} and \mathbf{z} are correlated, prove that

$$\hat{\mathbb{E}}\{\mathbf{x}|\mathbf{y}, \mathbf{z}\} = \hat{\mathbb{E}}\{\mathbf{x}|\mathbf{y}, \hat{\mathbf{z}}\}$$

where

$$\hat{\mathbf{z}} = \mathbf{z} - \hat{\mathbb{E}}\{\mathbf{z}|\mathbf{y}\}$$

so that

$$\hat{\mathbb{E}}\{\mathbf{x}|\mathbf{y}, \mathbf{z}\} = \hat{\mathbb{E}}\{\mathbf{x}|\mathbf{y}\} + \hat{\mathbb{E}}\{\mathbf{x}|\hat{\mathbf{z}}\} - \mathbf{m}_{\mathbf{x}}$$

Parts (c) and (d) show that the results given in Theorems 12-3 and 12-4 are "distribution free" within the class of linear (i.e., affine), unbiased, mean-squared estimators.

- 13-5. Consider the linear model $\mathbf{z}(k) = 2\boldsymbol{\theta} + \mathbf{n}(k)$, where

$$p[n(k)] = \begin{cases} 1 - |n(k)|, & |n(k)| \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

and

$$p(\boldsymbol{\theta}) = \begin{cases} \frac{1}{2}, & |\boldsymbol{\theta}| \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

A random sample of N measurements is available. Explain how to find the ML and MAP estimators of $\boldsymbol{\theta}$, and be sure to list all of the assumptions needed to obtain a solution (they are not all given).

Lesson 14

Estimation of Random Parameters: The Linear and Gaussian Model

INTRODUCTION

In this lesson we begin with the linear model

$$\mathcal{Z}(k) = \mathcal{H}(k)\theta + \mathcal{V}(k) \quad (14-1)$$

where θ is an $n \times 1$ vector of random unknown parameters, $\mathcal{H}(k)$ is deterministic, and $\mathcal{V}(k)$ is white Gaussian noise with known covariance matrix $\mathcal{R}(k)$. We also assume that θ is multivariate Gaussian with known mean, \mathbf{m}_θ , and covariance, \mathbf{P}_θ , i.e.,

$$\theta \sim N(\theta; \mathbf{m}_\theta, \mathbf{P}_\theta) \quad (14-2)$$

and, that θ and $\mathcal{V}(k)$ are Gaussian and mutually uncorrelated. Our main objectives are to compute $\hat{\theta}_{MS}(k)$ and $\hat{\theta}_{MAP}(k)$ for this linear Gaussian model, and to see how these estimators are related. We shall also illustrate many of our results for the deconvolution and state estimation examples which were described in Lesson 2.

MEAN-SQUARED ESTIMATOR

Because θ and $\mathcal{V}(k)$ are Gaussian and mutually uncorrelated, they are jointly Gaussian. Consequently, $\mathcal{Z}(k)$ and θ are also jointly Gaussian; thus (Corollary 13-1),

$$\hat{\theta}_{MS}(k) = \mathbf{m}_\theta + \mathbf{P}_{\theta\mathcal{Z}}(k)\mathbf{P}_{\mathcal{Z}}^{-1}(k)[\mathcal{Z}(k) - \mathbf{m}_{\mathcal{Z}}(k)] \quad (14-3)$$

Mean-Squared Estimator

119

It is straightforward, using (14-1) and (14-2), to show that

$$\mathbf{m}_{\mathcal{Z}}(k) = \mathcal{H}(k)\mathbf{m}_\theta \quad (14-4)$$

$$\mathbf{P}_{\mathcal{Z}}(k) = \mathcal{H}(k)\mathbf{P}_\theta\mathcal{H}'(k) + \mathcal{R}(k) \quad (14-5)$$

and

$$\mathbf{P}_{\theta\mathcal{Z}}(k) = \mathbf{P}_\theta\mathcal{H}'(k) \quad (14-6)$$

consequently,

$$\hat{\theta}_{MS}(k) = \mathbf{m}_\theta + \mathbf{P}_\theta\mathcal{H}'(k)[\mathcal{H}(k)\mathbf{P}_\theta\mathcal{H}'(k) + \mathcal{R}(k)]^{-1}[\mathcal{Z}(k) - \mathcal{H}(k)\mathbf{m}_\theta] \quad (14-7)$$

Observe that $\hat{\theta}_{MS}(k)$ depends on all the given information, namely, $\mathcal{Z}(k)$, $\mathcal{H}(k)$, \mathbf{m}_θ and \mathbf{P}_θ .

Next, we compute the error-covariance matrix, $\mathbf{P}_{MS}(k)$, that is associated with $\hat{\theta}_{MS}(k)$ in (14-7). Because $\hat{\theta}_{MS}(k)$ is unbiased, we know that $E\{\hat{\theta}_{MS}(k)\} = \theta$ for all k ; thus,

$$\mathbf{P}_{MS}(k) = E\{\hat{\theta}_{MS}(k)\hat{\theta}_{MS}'(k)\} \quad (14-8)$$

From (14-7), and the fact that $\hat{\theta}_{MS}(k) = \theta - \tilde{\theta}_{MS}(k)$, we see that

$$\tilde{\theta}_{MS}(k) = (\theta - \mathbf{m}_\theta) - \mathbf{P}_\theta\mathcal{H}'(\mathcal{H}\mathbf{P}_\theta\mathcal{H}' + \mathcal{R})^{-1}(\mathcal{Z} - \mathcal{H}\mathbf{m}_\theta) \quad (14-9)$$

From (14-9), (14-8) and (14-1) it is a straightforward exercise to show that

$$\mathbf{P}_{MS}(k) = \mathbf{P}_\theta - \mathbf{P}_\theta\mathcal{H}'(k)[\mathcal{H}(k)\mathbf{P}_\theta\mathcal{H}'(k) + \mathcal{R}(k)]^{-1}\mathcal{H}(k)\mathbf{P}_\theta \quad (14-10)$$

Applying Matrix Inversion Lemma 4-1 to (14-10), we obtain the following alternate formula for $\mathbf{P}_{MS}(k)$,

$$\mathbf{P}_{MS}(k) = [\mathbf{P}_\theta^{-1} + \mathcal{H}'(k)\mathcal{R}^{-1}(k)\mathcal{H}(k)]^{-1} \quad (14-11)$$

Next, we express $\hat{\theta}_{MS}(k)$ as an explicit function of $\mathbf{P}_{MS}(k)$. To do this, we note that

$$\begin{aligned} \mathbf{P}_\theta\mathcal{H}'(\mathcal{H}\mathbf{P}_\theta\mathcal{H}' + \mathcal{R})^{-1} &= \mathbf{P}_\theta\mathcal{H}'(\mathcal{H}\mathbf{P}_\theta\mathcal{H}' + \mathcal{R})^{-1} \\ &\quad (\mathcal{H}\mathbf{P}_\theta\mathcal{H}' + \mathcal{R} - \mathcal{H}\mathbf{P}_\theta\mathcal{H}')\mathcal{R}^{-1} \\ &= \mathbf{P}_\theta\mathcal{H}'[\mathbf{I} - (\mathcal{H}\mathbf{P}_\theta\mathcal{H}' + \mathcal{R})^{-1}\mathcal{H}\mathbf{P}_\theta\mathcal{H}']\mathcal{R}^{-1} \\ &= [\mathbf{P}_\theta - \mathbf{P}_\theta\mathcal{H}'(\mathcal{H}\mathbf{P}_\theta\mathcal{H}' + \mathcal{R})^{-1}\mathcal{H}\mathbf{P}_\theta]\mathcal{H}\mathcal{R}^{-1} \\ &= \mathbf{P}_{MS}\mathcal{H}'\mathcal{R}^{-1} \end{aligned} \quad (14-12)$$

hence,

$$\hat{\theta}_{MS}(k) = \mathbf{m}_\theta + \mathbf{P}_{MS}(k)\mathcal{H}'(k)\mathcal{R}^{-1}(k)[\mathcal{Z}(k) - \mathcal{H}(k)\mathbf{m}_\theta] \quad (14-13)$$

Theorem 14-1. If $\mathbf{P}_\theta^{-1} = \mathbf{0}$, and, $\mathcal{H}(k)$ is deterministic, then

$$\hat{\theta}_{MS}(k) = \hat{\theta}_{BLU}(k) \quad (14-14)$$

Proof. Set $\mathbf{P}_\theta^{-1} = \mathbf{0}$ in (14-11), to see that

$$\mathbf{P}_{MS}(k) = [\mathcal{H}'(k)\mathcal{R}^{-1}(k)\mathcal{H}(k)]^{-1} \quad (14-15)$$

and, therefore,

$$\hat{\theta}_{MS}(k) = [\mathcal{H}'(k)\mathcal{R}^{-1}(k)\mathcal{H}(k)]^{-1}\mathcal{H}'(k)\mathcal{R}^{-1}(k)\mathcal{L}(k) \quad (14-16)$$

Compare (14-16) and (9-22), to conclude that $\hat{\theta}_{MS}(k) = \hat{\theta}_{BLU}(k)$. \square

One of the most startling aspects of Theorem 14-1 is that it shows us that *BLU estimation applies to random parameters as well as to deterministic parameters*. We return to a reexamination of BLUE below.

What does the condition $\mathbf{P}_\theta^{-1} = \mathbf{0}$, given in Theorem 14-1, mean? Suppose, for example, that the elements of θ are uncorrelated; then, \mathbf{P}_θ is a diagonal matrix, with diagonal elements $\sigma_{i,\theta}^2$. When all of these variances are very large, then $\mathbf{P}_\theta^{-1} = \mathbf{0}$. A large variance for θ_i means we have no idea where θ_i is located about its mean value.

Example 14-1 (Minimum-Variance Deconvolution)

In Example 2-6 we showed that, for the application of deconvolution, our linear model is

$$\mathcal{Z}(N) = \mathcal{H}(N-1)\mu + \mathcal{V}(N) \quad (14-17)$$

We shall assume that μ and $\mathcal{V}(N)$ are jointly Gaussian, and, that $\mathbf{m}_\mu = \mathbf{0}$ and $\mathbf{m}_\mathcal{V} = \mathbf{0}$; hence, $\mathbf{m}_\mathcal{Z} = \mathbf{0}$. Additionally, we assume that $\text{cov}[\mathcal{V}(N)] = \rho\mathbf{I}$. From (14-7), we determine the following formula for $\hat{\mu}_{MS}(N)$,

$$\hat{\mu}_{MS}(N) = \mathbf{P}_\mu \mathcal{H}'(N-1)[\mathcal{H}(N-1)\mathbf{P}_\mu \mathcal{H}'(N-1) + \rho\mathbf{I}]^{-1}\mathcal{Z}(N) \quad (14-18)$$

Recall, from Example 2-6, that when $\mu(k)$ is described by the product model $\mu(k) = q(k)r(k)$, then

$$\mu = \mathbf{Q}_q \mathbf{r} \quad (14-19)$$

where

$$\mathbf{Q}_q = \text{diag}[q(1), q(2), \dots, q(N)] \quad (14-20)$$

and

$$\mathbf{r} = \text{col}(r(1), r(2), \dots, r(N)) \quad (14-21)$$

In the product model, $r(k)$ is white Gaussian noise with variance σ_r^2 , and $q(k)$ is a Bernoulli sequence. Obviously, if we know \mathbf{Q}_q then μ is Gaussian, in which case

$$\mathbf{P}_\mu = \mathbf{Q}_q^2 \sigma_r^2 = \mathbf{Q}_q \sigma_r^2 \quad (14-22)$$

where we have used the fact that $\mathbf{Q}_q^2 = \mathbf{Q}_q$, because $q(k) = 0$ or 1. When \mathbf{Q}_q is known, (14-18) becomes

$$\hat{\mu}_{MS}(N|\mathbf{Q}_q) = \sigma_r^2 \mathbf{Q}_q \mathcal{H}'(N-1)[\sigma_r^2 \mathcal{H}(N-1) \mathbf{Q}_q \mathcal{H}'(N-1) + \rho\mathbf{I}]^{-1} \mathcal{Z}(N) \quad (14-23)$$

Although $\hat{\mu}_{MS}$ is a mean-squared estimator, so that it enjoys all of the properties of such an estimator (e.g., unbiased, minimum-variance, etc.), $\hat{\mu}_{MS}$ is not a consistent estimator of μ . Consistency is a large sample property of an estimator: however, as N increases, the dimension of μ increases, because μ is $N \times 1$. Consequently, we cannot prove consistency of $\hat{\mu}_{MS}$ (recall that, in all other problems, θ is $n \times 1$, where n is data independent; in these problems we can study consistency of $\hat{\theta}$).

Equations (14-18) and (14-23) are not very practical for actually computing $\hat{\mu}_{MS}$, because both require the inversion of an $N \times N$ matrix, and N can become quite large (it equals the number of measurements). We return to a more practical way for computing $\hat{\mu}_{MS}$ in Lesson 22. \square

BEST LINEAR UNBIASED ESTIMATION, REVISITED

In Lesson 9 we derived the BLUE of θ for the linear model (14-1), under the following assumptions about this model:

1. θ is a deterministic but unknown vector of parameters,
2. $\mathcal{H}(k)$ is deterministic, and
3. $\mathcal{V}(k)$ is zero-mean noise with covariance matrix $\mathcal{R}(k)$.

We assumed that $\hat{\theta}_{BLU}(k) = \mathbf{F}(k)\mathcal{Z}(k)$ and chose $\mathbf{F}_{BLU}(k)$ so that $\hat{\theta}_{BLU}(k)$ is an unbiased estimator of θ , and the error variance for each one of the n elements of θ is minimized. The reader should return to the derivation of $\hat{\theta}_{BLU}(k)$ to see that the assumption “ θ is deterministic” is never needed, either in the derivation of the unbiasedness constraint [see the proof of Theorem 6-1, in which Equation (6-9) becomes $[\mathbf{I} - \mathbf{F}(k)\mathcal{H}(k)]\mathbf{E}\{\theta\} = \mathbf{0}$, if θ is random], or in the derivation of $J_i(\mathbf{f}_i, \lambda_i)$ in Equation (9-14) (due to some remarkable cancellations); thus, $\hat{\theta}_{BLU}(k)$, given in (9-22), is applicable to random as well as deterministic parameters in our linear model (14-1); and, because the BLUE of θ is the special case of the WLSE when $\mathbf{W}(k) = \mathcal{R}^{-1}(k)$, $\hat{\theta}_{WLSE}(k)$, given in (3-10), is also applicable to random as well as deterministic parameters in our linear model.

Theorem 14-1 relates $\hat{\theta}_{MS}(k)$ and $\hat{\theta}_{BLU}(k)$ under some very stringent conditions that are needed in order to remove the dependence of $\hat{\theta}_{MS}$ on the a priori statistical information about θ (i.e., \mathbf{m}_θ and \mathbf{P}_θ), because this information was never used in the derivation of $\hat{\theta}_{BLU}(k)$.

Next, we derive a different BLUE of θ , one that incorporates the a priori statistical information about θ . To do this (Sorenson, 1980, pg. 210), we treat \mathbf{m}_θ as an additional measurement which will be augmented to $\mathcal{Z}(k)$. Our additional measurement equation is obtained by adding and subtracting θ in the identity $\mathbf{m}_\theta = \mathbf{m}_\theta$, i.e.,

$$\mathbf{m}_\theta = \theta + (\mathbf{m}_\theta - \theta) \quad (14-24)$$

Quantity $\mathbf{m}_\theta - \theta$ is now treated as zero-mean noise with covariance matrix \mathbf{P}_θ . Our augmented linear model is

$$\underbrace{\left(\frac{\mathcal{L}(k)}{\mathbf{m}_\theta} \right)}_{\mathcal{Z}_a} = \underbrace{\left(\frac{\mathcal{H}(k)}{\mathbf{I}} \right)}_{\mathcal{H}_a} \theta + \underbrace{\left(\frac{\mathcal{V}(k)}{\mathbf{m}_\theta - \theta} \right)}_{\mathcal{V}_a} \quad (14-25)$$

which can be written, as

$$\mathcal{Z}_a(k) = \mathcal{H}_a(k)\theta + \mathcal{V}_a(k) \quad (14-26)$$

where $\mathcal{Z}_a(k)$, $\mathcal{H}_a(k)$, and $\mathcal{V}_a(k)$ are defined in (14-25). Additionally,

$$\mathbf{E}\{\mathcal{V}_a(k)\mathcal{V}'_a(k)\} \triangleq \mathcal{R}_a(k) = \begin{pmatrix} \mathcal{R}(k) & \mathbf{0} \\ \mathbf{0} & |\mathbf{P}_\theta| \end{pmatrix} \quad (14-27)$$

We now treat (14-26) as the starting point for derivation of a BLUE of θ , which we denote $\hat{\theta}_{BLU}^a(k)$. Obviously,

$$\hat{\theta}_{BLU}^a(k) = [\mathcal{H}'_a(k)\mathcal{R}_a^{-1}(k)\mathcal{H}_a(k)]^{-1}\mathcal{H}'_a(k)\mathcal{R}_a^{-1}(k)\mathcal{Z}_a(k) \quad (14-28)$$

Theorem 14-2. For the linear Gaussian model, when $\mathcal{H}(k)$ is deterministic it is always true that

$$\hat{\theta}_{MS}(k) = \hat{\theta}_{BLU}^a(k) \quad (14-29)$$

Proof. To begin, substitute the definitions of \mathcal{H}_a , \mathcal{R}_a , and \mathcal{Z}_a into (14-28), and multiply out all the partitioned matrices, to show that

$$\hat{\theta}_{BLU}^a(k) = (\mathbf{P}_\theta^{-1} + \mathcal{H}'\mathcal{R}^{-1}\mathcal{H})^{-1}(\mathbf{P}_\theta^{-1}\mathbf{m}_\theta + \mathcal{H}'\mathcal{R}^{-1}\mathcal{Z}) \quad (14-30)$$

From (14-11), we recognize $(\mathbf{P}_\theta^{-1} + \mathcal{H}'\mathcal{R}^{-1}\mathcal{H})^{-1}$ as \mathbf{P}_{MS} ; hence,

$$\hat{\theta}_{BLU}^a(k) = \mathbf{P}_{MS}(k)[\mathbf{P}_\theta^{-1}\mathbf{m}_\theta + \mathcal{H}'(k)\mathcal{R}^{-1}(k)\mathcal{Z}(k)] \quad (14-31)$$

Next, we rewrite $\hat{\theta}_{MS}(k)$ in (14-13), as

$$\hat{\theta}_{MS}(k) = [\mathbf{I} - \mathbf{P}_{MS}\mathcal{H}'\mathcal{R}^{-1}\mathcal{H}]\mathbf{m}_\theta + \mathbf{P}_{MS}\mathcal{H}'\mathcal{R}^{-1}\mathcal{Z} \quad (14-32)$$

however,

$$\mathbf{I} - \mathbf{P}_{MS}\mathcal{H}'\mathcal{R}^{-1}\mathcal{H} = \mathbf{P}_{MS}(\mathbf{P}_{MS}^{-1} - \mathcal{H}'\mathcal{R}^{-1}\mathcal{H}) = \mathbf{P}_{MS}\mathbf{P}_\theta^{-1} \quad (14-33)$$

where we have again used (14-11). Substitute (14-33) into (14-32) to see that

$$\hat{\theta}_{MS}(k) = \mathbf{P}_{MS}(k)[\mathbf{P}_\theta^{-1}\mathbf{m}_\theta + \mathcal{H}'(k)\mathcal{R}^{-1}(k)\mathcal{Z}(k)] \quad (14-34)$$

hence, (14-29) follows when we compare (14-31) and (14-34). \square

To conclude this section, we note that the weighted-least squares objective function that is associated with $\hat{\theta}_{BLU}^a(k)$ is

$$\begin{aligned} J_a[\hat{\theta}^a(k)] &= \tilde{\mathcal{Z}}'_a(k)\mathcal{R}_a^{-1}(k)\tilde{\mathcal{Z}}_a(k) \\ &= (\mathbf{m}_\theta - \theta)' \mathbf{P}_\theta^{-1}(\mathbf{m}_\theta - \theta) + \tilde{\mathcal{Z}}'(k)\mathcal{R}^{-1}(k)\tilde{\mathcal{Z}}(k) \end{aligned} \quad (14-35)$$

The first term in (14-35) contains all the a priori information about θ . Quantity $\mathbf{m}_\theta - \theta$ is treated as the difference between “measurement” \mathbf{m}_θ and its noise-free model, θ .

MAXIMUM A POSTERIORI ESTIMATOR

In order to determine $\hat{\theta}_{MAP}(k)$ for the linear model in (14-1) we first need to determine $p(\theta|\mathcal{Z}(k))$. Using the facts that $\theta \sim N(\theta; \mathbf{m}_\theta, \mathbf{P}_\theta)$ and $\mathcal{V}(k) \sim N(\mathcal{V}(k); \mathbf{0}, \mathcal{R}(k))$, it follows that

$$p(\theta) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{P}_\theta|}} \exp\left\{-\frac{1}{2}(\theta - \mathbf{m}_\theta)' \mathbf{P}_\theta^{-1}(\theta - \mathbf{m}_\theta)\right\} \quad (14-36)$$

and

$$\begin{aligned} p(\mathcal{Z}(k)|\theta) &= \frac{1}{\sqrt{(2\pi)^N |\mathcal{R}(k)|}} \\ &\exp\left\{-\frac{1}{2}[\mathcal{Z}(k) - \mathcal{H}(k)\theta]' \mathcal{R}^{-1}(k)[\mathcal{Z}(k) - \mathcal{H}(k)\theta]\right\} \end{aligned} \quad (14-37)$$

hence,

$$\ln p(\theta|\mathcal{Z}) \propto -\frac{1}{2}(\theta - \mathbf{m}_\theta)' \mathbf{P}_\theta^{-1}(\theta - \mathbf{m}_\theta) - \frac{1}{2}[\mathcal{Z} - \mathcal{H}\theta]' \mathcal{R}^{-1}(\mathcal{Z} - \mathcal{H}\theta) \quad (14-38)$$

To find $\hat{\theta}_{MAP}(k)$ we must maximize $\ln p(\theta|\mathcal{Z})$ in (14-38). Note that to find $\hat{\theta}_{BLU}^a(k)$ we had to maximize $J_a[\hat{\theta}^a(k)]$ in (14-35); but,

$$\ln p(\theta|\mathcal{Z}) \propto -J_a[\hat{\theta}^a(k)] \quad (14-39)$$

This means that maximizing $\ln p(\theta|\mathcal{Z})$ leads to the same value of $\hat{\theta}$ as does minimizing $J_a[\hat{\theta}^a(k)]$. We have, therefore, shown that $\hat{\theta}_{MAP}(k) = \hat{\theta}_{BLU}^a(k)$.

Theorem 14-3. For the linear Gaussian model, when $\mathcal{H}(k)$ is deterministic it is always true that

$$\hat{\theta}_{MAP}(k) = \hat{\theta}_{BLU}^a(k) \quad \square \quad (14-40)$$

Combining the results in Theorems 14-2 and 14-3, we have the very important result, that for the linear Gaussian model

$$\hat{\theta}_{MS}(k) = \hat{\theta}_{BLU}^a(k) = \hat{\theta}_{MAP}(k) \quad (14-41)$$

Put another way, for the linear Gaussian model, all roads lead to the same estimator.

Of course, the fact that $\hat{\theta}_{\text{MS}}(k) = \hat{\theta}_{\text{MAP}}(k)$ should not come as any surprise, because we already established it (in a model-free environment) in Theorem 13-2.

Example 14-2 (Maximum-Likelihood Deconvolution)

As in Example 14-1, we begin with the deconvolution linear model

$$\mathbf{z}(N) = \mathcal{H}(N-1)\mu + \mathbf{v}(N) \quad (14-42)$$

Now, however, we use the product model for μ , given in (14-19), to express $\mathbf{z}(N)$ as

$$\mathbf{z}(N) = \mathcal{H}(N-1)\mathbf{Q}_q\mathbf{r} + \mathbf{v}(N) \quad (14-43)$$

For notational convenience, let

$$\mathbf{q} = \text{col}(q(1), q(2), \dots, q(N)) \quad (14-44)$$

Our objectives in this example are to obtain MAP estimators for both \mathbf{q} and \mathbf{r} . In the literature on maximum-likelihood deconvolution (e.g., Mendel, 1983) these estimators are referred to as unconditional ML estimators, and are denoted $\hat{\mathbf{r}}$ and $\hat{\mathbf{q}}$. We denote these estimators as $\hat{\mathbf{r}}_{\text{MAP}}$ and $\hat{\mathbf{q}}_{\text{MAP}}$, in order to be consistent with this book's notation.

The starting point for determining $\hat{\mathbf{q}}_{\text{MAP}}$ and $\hat{\mathbf{r}}_{\text{MAP}}$ is the joint density function $p(\mathbf{r}, \mathbf{q} | \mathbf{z}(N))$, where

$$p(\mathbf{r}, \mathbf{q} | \mathbf{z}(N)) \propto p(\mathbf{z}(N) | \mathbf{r}, \mathbf{q})p(\mathbf{r}, \mathbf{q}) \quad (14-45)$$

but

$$p(\mathbf{r}, \mathbf{q}) = p(\mathbf{r} | \mathbf{q})\text{Pr}(\mathbf{q}) \quad (14-46)$$

Equation (14-46) uses a probability function for \mathbf{q} rather than a probability density function, because $q(k)$ takes on only two discrete values, 0 and 1. Substituting (14-46) into (14-45), we find that

$$p(\mathbf{r}, \mathbf{q} | \mathbf{z}(N)) \propto p(\mathbf{z}(N) | \mathbf{r}, \mathbf{q})p(\mathbf{r} | \mathbf{q})\text{Pr}(\mathbf{q}) \quad (14-47)$$

Note, also, that

$$\begin{aligned} p(\mathbf{r}, \mathbf{z}(N) | \mathbf{q}) &= p(\mathbf{z}(N) | \mathbf{r}, \mathbf{q})p(\mathbf{r} | \mathbf{q}) \\ &= p(\mathbf{z}(N) | \mathbf{r}, \mathbf{q})p(\mathbf{r}) \end{aligned} \quad (14-48)$$

where we have used the fact that \mathbf{r} and \mathbf{q} are statistically independent. Substituting (14-48) for $p(\mathbf{z}(N) | \mathbf{r}, \mathbf{q})$ into (14-47), we see that

$$p(\mathbf{r}, \mathbf{q} | \mathbf{z}(N)) \propto p(\mathbf{r}, \mathbf{z}(N) | \mathbf{q})\text{Pr}(\mathbf{q}) \quad (14-49)$$

Observe that the \mathbf{r} dependence of the MAP-likelihood function is completely contained in $p(\mathbf{r}, \mathbf{z}(N) | \mathbf{q})$. Additionally, when \mathbf{q} is given, the only remaining random quantities are the zero-mean Gaussian quantities \mathbf{r} and $\mathbf{z}(N)$; hence, $p(\mathbf{r}, \mathbf{z}(N) | \mathbf{q})$ is multivariate Gaussian. Letting

$$\mathbf{x} = \text{col}(\mathbf{r}, \mathbf{z}(N)) \quad (14-50)$$

and

$$\mathbf{P}_x = \mathbf{E}(\mathbf{x}\mathbf{x}' | \mathbf{q}) \quad (14-51)$$

then

$$p(\mathbf{r}, \mathbf{z}(N) | \mathbf{q}) = (2\pi)^{-N}|\mathbf{P}_x|^{-1/2} \exp\left(-\frac{1}{2} \mathbf{x}' \mathbf{P}_x^{-1} \mathbf{x}\right) \quad (14-52)$$

We leave, as an exercise for the reader, the maximization of $p(\mathbf{r}, \mathbf{q} | \mathbf{z}(N))$, from which it follows (Mendel, 1983b, pp. 112–114) that

$$\mathbf{r}^*(N | \mathbf{q}) = \sigma_r^2 \mathbf{Q}_q \mathcal{H}'(N-1)[\sigma_r^2 \mathcal{H}(N-1) \mathbf{Q}_q \mathcal{H}'(N-1) + \rho \mathbf{I}]^{-1} \mathbf{z}(N) \quad (14-53)$$

$\hat{\mathbf{q}}_{\text{MAP}}$ can be found by maximizing

$$\begin{aligned} p(\mathbf{r}^*, \mathbf{q} | \mathbf{z}(N)) &\propto p(\mathbf{r}^*, \mathbf{z}(N) | \mathbf{q})\text{Pr}(\mathbf{q}) = (2\pi)^{-N}(\sigma_r^2 \rho)^{-N/2} \\ &\quad \exp\left\{-\frac{1}{2} \mathbf{z}'(N) [\sigma_r^2 \mathcal{H}(N-1) \mathbf{Q}_q \mathcal{H}'(N-1) + \rho \mathbf{I}]^{-1} \mathbf{z}(N)\right\} \text{Pr}(\mathbf{q}) \end{aligned} \quad (14-54)$$

where

$$\text{Pr}(\mathbf{q}) = \prod_{k=1}^N \text{Pr}[q(k)] = \lambda^{m_q} (1-\lambda)^{N-m_q} \quad (14-55)$$

and

$$m_q = \sum_{k=1}^N q(k) \quad (14-56)$$

and finally,

$$\hat{\mathbf{r}}_{\text{MAP}} = \mathbf{r}^*(N | \hat{\mathbf{q}}_{\text{MAP}}) \quad (14-57)$$

Equations (14-54) and (14-57) are quite interesting results. Observe that we are permitted first to direct our attention to finding $\hat{\mathbf{q}}_{\text{MAP}}$ and then to finding $\hat{\mathbf{r}}_{\text{MAP}}$. Observe, also, that $\hat{\mathbf{r}}_{\text{MAP}} = \hat{\mu}_{\text{MS}}(N | \mathbf{Q}_q)$ [compare (14-53) and (14-23)].

There is no simple solution for determining $\hat{\mathbf{q}}_{\text{MAP}}$. Because the elements of \mathbf{q} in $p(\mathbf{r}^*, \mathbf{q} | \mathbf{z}(N))$ have nonlinear interactions, and because the elements of \mathbf{q} are constrained to take on binary values, it is necessary to evaluate $p(\mathbf{r}^*, \mathbf{q} | \mathbf{z}(N))$ for every possible \mathbf{q} sequence to find the \mathbf{q} for which $p(\mathbf{r}^*, \mathbf{q} | \mathbf{z}(N))$ is a global maximum. Because $q(k)$ can take on one of two possible values, there are 2^N possible sequences where N is the number of elements in \mathbf{q} . For reasonable values of N (such as $N = 400$), finding the global maximum of $p(\mathbf{r}^*, \mathbf{q} | \mathbf{z}(N))$ would require several centuries of computer time.

We can always design a method for detecting significant values of $\mu(k)$ so that the resulting $\hat{\mathbf{q}}$ will be nearly as likely as the unconditional maximum-likelihood estimate, $\hat{\mathbf{q}}_{\text{MAP}}$. Two MAP detectors for accomplishing this are described in Mendel (1983b, pp. 127–137). [Note: The reader who is interested in more details about these MAP detectors should first review Chapter 5 in Mendel, 1983b, because they are designed using a slightly different likelihood function than the one in (14-54).] \square

Example 14-3 (State-Estimation)

In Example 2-4 we showed that, for the application of state estimation, our linear model is

$$\mathcal{Z}(N) = \mathcal{H}(N, k_1)\mathbf{x}(k_1) + \mathcal{V}(N, k_1) \quad (14-58)$$

From (2-17), we see that

$$\mathbf{x}(k_1) = \Phi^{k_1}\mathbf{x}(0) + \sum_{i=1}^{k_1} \Phi^{k_1-i} \gamma u(i-1) = \Phi^{k_1}\mathbf{x}(0) + \mathbf{L}\mathbf{u} \quad (14-59)$$

where

$$\mathbf{u} = \text{col}(u(0), u(1), \dots, u(N)) \quad (14-60)$$

and \mathbf{L} is an $n \times (N+1)$ matrix, the exact structure of which is not important for this example. Additionally, from (2-21), we see that

$$\mathcal{V}(N, k_1) = \mathbf{M}(N, k_1)\mathbf{u} + \mathbf{v} \quad (14-61)$$

where

$$\mathbf{v} = \text{col}(v(1), v(2), \dots, v(N)) \quad (14-62)$$

Observe that both $\mathbf{x}(k_1)$ and $\mathcal{V}(N, k_1)$ can be viewed as linear functions of $\mathbf{x}(0)$, \mathbf{u} and \mathbf{v} .

We now assume that $\mathbf{x}(0)$, \mathbf{u} and \mathbf{v} are jointly Gaussian. Reasons for doing this are discussed in Lesson 15. Then, because $\mathbf{x}(k_1)$ and $\mathcal{V}(N, k_1)$ are linear functions of $\mathbf{x}(0)$, \mathbf{u} and \mathbf{v} , $\mathbf{x}(k_1)$ and $\mathcal{V}(N, k_1)$ are jointly Gaussian (Papoulis, 1965, pg. 253); hence,

$$\begin{aligned} \hat{\mathbf{x}}_{\text{MS}}(k_1|N) &= \mathbf{m}_{\mathbf{x}(k_1)} + \mathbf{P}_{\mathbf{x}(k_1)}\mathcal{H}'(N, k_1)[\mathcal{H}(N, k_1)\mathbf{P}_{\mathbf{x}(k_1)}\mathcal{H}'(N, k_1) \\ &\quad + \mathcal{R}(N, k_1)]^{-1}[\mathcal{Z}(N) - \mathcal{H}(N, k_1)\mathbf{m}_{\mathbf{x}(k_1)}] \end{aligned} \quad (14-63)$$

and

$$\boxed{\hat{\mathbf{x}}_{\text{MAP}}(k_1|N) = \hat{\mathbf{x}}_{\text{MS}}(k_1|N)} \quad (14-64)$$

In order to evaluate $\hat{\mathbf{x}}_{\text{MS}}(k_1|N)$ we must first compute $\mathbf{m}_{\mathbf{x}(k_1)}$ and $\mathbf{P}_{\mathbf{x}(k_1)}$. We show how to do this in Lesson 15.

Formula (14-63) is very cumbersome. It appears that its right-hand side changes as a function of k_1 (and N). We conjecture, however, that it ought to be possible to express $\hat{\mathbf{x}}_{\text{MS}}(k_1|N)$ as an affine transformation of $\hat{\mathbf{x}}_{\text{MS}}(k_1-1|N)$, because $\mathbf{x}(k_1)$ is an affine transformation of $\mathbf{x}(k_1-1)$, i.e., $\mathbf{x}(k_1) = \Phi\mathbf{x}(k_1-1) + \gamma\mathbf{u}(k_1-1)$.

Because of the importance of state estimation in many different fields (e.g., control theory, communication theory, signal processing, etc.) we shall examine it in great detail in many of our succeeding lessons. \square

14-3. x and v are independent Gaussian random variables with zero means and variances σ_x^2 and σ_v^2 , respectively. We observe the single measurement $z = x + v = 1$.

- (a) Find \hat{x}_{ML} .
- (b) Find \hat{x}_{MAP} .

14-4. For the linear Gaussian model in which $\mathcal{H}(k)$ is deterministic, prove that $\hat{\theta}_{\text{MAP}}(k)$ is a most efficient estimator of θ . Do this in two different ways. Is $\hat{\theta}_{\text{MS}}(k)$ a most efficient estimator of θ ?

PROBLEMS

14-1. Derive Eq. (14-10) for $\mathbf{P}_{\text{MS}}(k)$.

14-2. Show that $\mathbf{r}^*(N|\mathbf{q})$ is given by Eq. (14-53), and that $\hat{\mathbf{q}}_{\text{MAP}}$ can be found by maximizing (14-54).

Elements of Discrete-Time Gauss-Markov Random Processes

INTRODUCTION

Lessons 13 and 14 have demonstrated the importance of Gaussian random variables in estimation theory. In this lesson we extend some of the basic concepts that were introduced in Lesson 12, for Gaussian random variables, to indexed random variables, namely random processes. These extensions are needed in order to develop state estimators.

DEFINITIONS AND PROPERTIES OF DISCRETE-TIME GAUSS-MARKOV RANDOM PROCESSES

Recall that a random process is a collection of random variables in which the notion of time plays a role.

Definition 15-1 (Meditch, 1969, pg. 106). *A vector random process is a family of random vectors $\{s(t), t \in \mathcal{S}\}$ indexed by a parameter t all of whose values lie in some appropriate index set \mathcal{S} . When $\mathcal{S} = \{k: k = 0, 1, \dots\}$ we have a discrete-time random process.* \square

Definition 15-2 (Meditch, 1969, pg. 117). *A vector random process $\{s(t), t \in \mathcal{S}\}$ is defined to be multivariate Gaussian if, for any ℓ time points t_1, t_2, \dots, t_ℓ in \mathcal{S} where ℓ is an integer, the set of ℓ random n vectors $s(t_1), s(t_2), \dots, s(t_\ell)$ is jointly Gaussian distributed.* \square

Let $\mathcal{G}(l)$ be defined as

$$\mathcal{G}(l) = \text{col}(s(t_1), s(t_2), \dots, s(t_l)) \quad (15-1)$$

then, Definition 15-2 means that

$$p[\mathcal{G}(l)] = (2\pi)^{-nl/2} |\mathbf{P}_g|^{-1/2} \exp \left\{ -\frac{1}{2} [\mathcal{G}(l) - \mathbf{m}_g(l)]' \mathbf{P}_g^{-1}(l) [\mathcal{G}(l) - \mathbf{m}_g(l)] \right\} \quad (15-2)$$

in which

$$\mathbf{m}_g(l) = \mathbb{E}\{\mathcal{G}(l)\} \quad (15-3)$$

and $\mathbf{P}_g(l)$ is the $nl \times nl$ matrix $\mathbb{E}\{[\mathcal{G}(l) - \mathbf{m}_g(l)][\mathcal{G}(l) - \mathbf{m}_g(l)]'\}$ with elements $\mathbf{P}_g(i, j)$, where

$$\begin{aligned} \mathbf{P}_g(i, j) &= \mathbb{E}\{[s(t_i) - \mathbf{m}_s(t_i)][s(t_j) - \mathbf{m}_s(t_j)]'\} \\ i, j &= 1, 2, \dots, l. \end{aligned} \quad (15-4)$$

Definition 15-3 (Meditch, 1969, pg. 118). *A vector random process $\{s(t), t \in \mathcal{S}\}$ is a Markov process, if, for any m time points $t_1 < t_2 < \dots < t_m$ in \mathcal{S} , where m is any integer, it is true that*

$$\begin{aligned} \Pr[s(t_m) \leq S(t_m) | s(t_{m-1})] &= S(t_{m-1}), \dots, s(t_1) = S(t_1)] \\ &= \Pr[s(t_m) \leq S(t_m) | s(t_{m-1})] = S(t_{m-1})] \end{aligned} \quad (15-5)$$

For continuous random variables, this means that

$$p[s(t_m) | s(t_{m-1}), \dots, s(t_1)] = p[s(t_m) | s(t_{m-1})] \quad \square \quad (15-6)$$

Note that, in (15-5), $s(t_m) \leq S(t_m)$ means $s_i(t_m) \leq S_i(t_m)$ for $i = 1, 2, \dots, n$. If we view time point t_m as the present time and time points t_{m-1}, \dots, t_1 as the past, then a Markov process is one whose probability law (e.g., probability density function) depends only on the immediate past value, t_{m-1} . This is often referred to as the *Markov property* for a vector random process. Because the probability law depends only on the immediate past value we often refer to such a process as a *first-order Markov process* (if it depended on the immediate two past values it would be a second-order Markov process).

Theorem 15-1. *Let $\{s(t), t \in \mathcal{S}\}$ be a first-order Markov process, and $t_1 < t_2 < \dots < t_m$ be any time points in \mathcal{S} , where m is an integer. Then*

$$\begin{aligned} p[s(t_m), s(t_{m-1}), \dots, s(t_1)] &= p[s(t_m) | s(t_{m-1})] p[s(t_{m-1}) | s(t_{m-2})] \dots \\ &\quad p[s(t_2) | s(t_1)] p[s(t_1)] \end{aligned} \quad (15-7)$$

Proof. From probability theory (e.g., Papoulis, 1965) and the Markov property of $s(t)$, we know that

$$\begin{aligned} p[s(t_m), s(t_{m-1}), \dots, s(t_1)] \\ = p[s(t_m)|s(t_{m-1}), \dots, s(t_1)]p[s(t_{m-1}), \dots, s(t_1)] \\ = p[s(t_m)|s(t_{m-1})]p[s(t_{m-1}), \dots, s(t_1)] \end{aligned} \quad (15-8)$$

In a similar manner, we find that

$$\left. \begin{aligned} p[s(t_{m-1}), \dots, s(t_1)] &= p[s(t_{m-1})|s(t_{m-2})]p[s(t_{m-2}), \dots, s(t_1)] \\ p[s(t_{m-2}), \dots, s(t_1)] &= p[s(t_{m-2})|s(t_{m-3})]p[s(t_{m-3}), \dots, s(t_1)] \\ &\dots \\ p[s(t_2), s(t_1)] &= p[s(t_2)|s(t_1)]p[s(t_1)] \end{aligned} \right\} \quad (15-9)$$

Equation (15-7) is obtained by successively substituting each one of the equations in (15-9) into (15-8). \square

Theorem 15-1 demonstrates that a first-order Markov process is completely characterized by two probability density functions, namely, the *transition probability density function*, $p[s(t_i)|s(t_{i-1})]$, and the initial (prior) *probability density function* $p[s(t_1)]$. Note that generally the transition probability density functions can all be different, in which case they should be subscripted [e.g., $p_m[s(t_m)|s(t_{m-1})]$ and $p_{m-1}[s(t_{m-1})|s(t_{m-2})]$].

Theorem 15-2. *For a first-order Markov process,*

$$E\{s(t_m)|s(t_{m-1}), \dots, s(t_1)\} = E\{s(t_m)|s(t_{m-1})\} \quad \square \quad (15-10)$$

We leave the proof of this useful result as an exercise.

A vector random process that is both Gaussian and a first-order Markov process will be referred to in the sequel as a Gauss-Markov process.

Definition 15-4. *A vector random process $\{s(t), t \in \mathcal{T}\}$ is said to be a Gaussian white process if, for any m time points t_1, t_2, \dots, t_m in \mathcal{T} , where m is any integer, the m random vectors $s(t_1), s(t_2), \dots, s(t_m)$ are uncorrelated Gaussian random vectors.* \square

White noise is zero mean, or else it cannot have a flat spectrum. For white noise

$$E\{s(t_i)s'(t_j)\} = \mathbf{0} \quad \text{for all } i \neq j \quad (15-11)$$

Additionally, for Gaussian white noise

$$p[\mathcal{G}(l)] = p[s(t_1)]p[s(t_2)] \dots p[s(t_l)] \quad (15-12)$$

[because uncorrelatedness implies statistical independence (see Lesson 12)] where $p[s(t_i)]$ is a multivariate Gaussian probability density function.

Theorem 15-3. *A vector Gaussian white process $s(t)$ can be viewed as a first-order Gauss-Markov process for which*

$$p[s(t)|s(\tau)] = p[s(t)] \quad (15-13)$$

for all $t, \tau \in \mathcal{T}$ and $t \neq \tau$.

Proof. For a Gaussian white process, we know, from (15-12), that

$$p[s(t), s(\tau)] = p[s(t)]p[s(\tau)] \quad (15-14)$$

but, we also know that

$$p[s(t), s(\tau)] = p[s(t)|s(\tau)]p[s(\tau)] \quad (15-15)$$

Equating (15-14) and (15-15), we obtain (15-13). \square

Theorem 15-3 means that past and future values of $s(t)$ in no way help determine present values of $s(t)$. For Gaussian white processes, the transition probability density function equals the marginal density function, $p[s(t)]$, which is multivariate Gaussian. Additionally,

$$E\{s(t_m)|s(t_{m-1}), \dots, s(t_1)\} = E\{s(t_m)\} \quad (15-16)$$

A BASIC STATE-VARIABLE MODEL

In succeeding lessons we shall develop a variety of state estimators for the following basic linear, (possibly) time-varying, discrete-time dynamical system (our *basic state-variable model*), which is characterized by $n \times 1$ state vector $\mathbf{x}(k)$ and $m \times 1$ measurement vector $\mathbf{z}(k)$:

$$\boxed{\mathbf{x}(k+1) = \Phi(k+1, k)\mathbf{x}(k) + \Gamma(k+1, k)\mathbf{w}(k) + \Psi(k+1, k)\mathbf{u}(k)} \quad (15-17)$$

and

$$\boxed{\mathbf{z}(k+1) = \mathbf{H}(k+1)\mathbf{x}(k+1) + \mathbf{v}(k+1)} \quad (15-18)$$

where $k = 0, 1, \dots$. In this model $\mathbf{w}(k)$ and $\mathbf{v}(k)$ are $p \times 1$ and $m \times 1$ mutually uncorrelated (possibly nonstationary) jointly Gaussian white noise sequences; i.e.,

$$\boxed{E\{\mathbf{w}(i)\mathbf{w}'(j)\} = \mathbf{Q}(i)\delta_{ij}} \quad (15-19)$$

$$\boxed{E\{\mathbf{v}(i)\mathbf{v}'(j)\} = \mathbf{R}(i)\delta_{ij}} \quad (15-20)$$

and

$$E\{w(i)v'(j)\} = S = \mathbf{0} \quad \text{for all } i \text{ and } j \quad (15-21)$$

Covariance matrix $Q(i)$ is positive semidefinite and $R(i)$ is positive definite [so that $R^{-1}(i)$ exists]. Additionally, $u(k)$ is an $l \times 1$ vector of known system inputs, and initial state vector $x(0)$ is multivariate Gaussian, with mean $m_x(0)$ and covariance $P_x(0)$, i.e.,

$$x(0) \sim N(x(0); m_x(0), P_x(0)) \quad (15-22)$$

and, $x(0)$ is not correlated with $w(k)$ and $v(k)$. The dimensions of matrices Φ , Γ , Ψ , H , Q and R are $n \times n$, $n \times p$, $n \times l$, $m \times n$, $p \times p$, and $m \times m$, respectively.

Disturbance $w(k)$ is often used to model the following types of uncertainty:

1. disturbance forces acting on the system (e.g., wind that buffets an airplane);
2. errors in modeling the system (e.g., neglected effects); and
3. errors, due to actuators, in the translation of the known input, $u(k)$, into physical signals.

Vector $v(k)$ is often used to model the following types of uncertainty:

1. errors in measurements made by sensing instruments;
2. unavoidable disturbances that act directly on the sensors; and
3. errors in the realization of feedback compensators using physical components [this is valid only when the measurement equation contains a direct throughput of the input $u(k)$, i.e., when $z(k+1) = H(k+1)x(k+1) + G(k+1)u(k+1) + v(k+1)$; we shall examine this situation in Lesson 23].

Of course, not all dynamical systems are described by this basic model. In general, $w(k)$ and $v(k)$ may be correlated, some measurements may be made so accurate that, for all practical purposes, they are "perfect" (i.e., there is no measurement noise associated with them), and either $u(k)$ or $v(k)$, or both, may be colored noise processes. We shall consider the modification of our basic state-variable model for each of these important situations in Lesson 23.

PROPERTIES OF THE BASIC STATE-VARIABLE MODEL

In this section we state and prove a number of important statistical properties for our basic state-variable model.

Theorem 15-4. *When $x(0)$ and $w(k)$ are jointly Gaussian then $\{x(k), k = 0, 1, \dots\}$ is a Gauss-Markov sequence.*

Note that if $x(0)$ and $w(k)$ are individually Gaussian and statistically independent (or uncorrelated), then they will be jointly Gaussian (Papoulis, 1965).

Proof

- a. Gaussian Property [assuming $u(k)$ nonrandom]. Because $u(k)$ is non-random, it has no effect on determining whether $x(k)$ is Gaussian; hence, for this part of the proof we assume $u(k) = \mathbf{0}$. The solution to (15-17) is

$$x(k) = \Phi(k,0)x(0) + \sum_{i=1}^k \Phi(k,i)\Gamma(i,i-1)w(i-1) \quad (15-23)$$

where

$$\Phi(k,i) = \Phi(k,k-1)\Phi(k-1,k-2)\dots\Phi(i+1,i) \quad (15-24)$$

Observe that $x(k)$ is a linear transformation of jointly Gaussian random vectors $x(0), w(0), w(1), \dots, w(k-1)$; hence, $x(k)$ is Gaussian.

- b. Markov Property. This property does not require $x(k)$ or $w(k)$ to be Gaussian. Because x satisfies state equation (15-17), we see that $x(k)$ depends only on its immediate past value; hence, $x(k)$ is Markov. \square

We have been able to show that our dynamical system is Markov because we specified a model for it. Without such a specification, it would be quite difficult (or impossible) to test for the Markov nature of a random process.

By stacking up $x(1), x(2), \dots$ into a supervector it is easily seen that this supervector is just a linear transformation of jointly Gaussian quantities $x(0), w(0), w(1), \dots$; hence, $x(1), x(2), \dots$ are themselves *jointly Gaussian*.

A Gauss-Markov sequence can be completely characterized in two ways:

1. specify the marginal density of the initial state vector, $p[x(0)]$, and the transition density $p[x(k+1)|x(k)]$, or
2. specify the mean and covariance of the state vector sequence. The second characterization is a complete one because Gaussian random vectors are completely characterized by their means and covariances (Lesson 12). We shall find the second characterization more useful than the first.

The Gaussian density function for state vector $\mathbf{x}(k)$ is

$$p[\mathbf{x}(k)] = [(2\pi)^n |\mathbf{P}_x(k)|]^{-1/2} \exp\left\{-\frac{1}{2} [\mathbf{x}(k) - \mathbf{m}_x(k)]' \mathbf{P}_x^{-1}(k) [\mathbf{x}(k) - \mathbf{m}_x(k)]\right\} \quad (15-25)$$

where

$$\mathbf{m}_x(k) = \mathbf{E}\{\mathbf{x}(k)\} \quad (15-26)$$

and

$$\mathbf{P}_x(k) = \mathbf{E}\{[\mathbf{x}(k) - \mathbf{m}_x(k)][\mathbf{x}(k) - \mathbf{m}_x(k)]'\} \quad (15-27)$$

We now demonstrate that $\mathbf{m}_x(k)$ and $\mathbf{P}_x(k)$ can be computed by means of recursive equations.

Theorem 15-5. *For our basic state-variable model,*

- a. $\mathbf{m}_x(k)$ can be computed from the vector recursive equation

$$\mathbf{m}_x(k + 1) = \Phi(k + 1, k)\mathbf{m}_x(k) + \Psi(k + 1, k)\mathbf{u}(k) \quad (15-28)$$

where $k = 0, 1, \dots$, and $\mathbf{m}_x(0)$ initializes (15-28),

- b. $\mathbf{P}_x(k)$ can be computed from the matrix recursive equation

$$\begin{aligned} \mathbf{P}_x(k + 1) &= \Phi(k + 1, k)\mathbf{P}_x(k)\Phi'(k + 1, k) \\ &\quad + \Gamma(k + 1, k)\mathbf{Q}(k)\Gamma'(k + 1, k) \end{aligned} \quad (15-29)$$

where $k = 0, 1, \dots$, and $\mathbf{P}_x(0)$ initializes (15-29), and

- c. $\mathbf{E}\{[\mathbf{x}(i) - \mathbf{m}_x(i)][\mathbf{x}(j) - \mathbf{m}_x(j)]'\} \triangleq \mathbf{P}_x(i, j)$ can be computed from

$$\mathbf{P}_x(i, j) = \begin{cases} \Phi(i, j)\mathbf{P}_x(j) & \text{when } i > j \\ \mathbf{P}_x(i)\Phi'(j, i) & \text{when } i < j \end{cases} \quad (15-30)$$

Proof

- a. Take the expected value of both sides of (15-17), using the facts that expectation is a linear operation (Papoulis, 1965) and $\mathbf{w}(k)$ is zero mean, to obtain (15-28).
- b. For notational simplicity, we omit the temporal arguments of Φ and Γ in this part of the proof. Using (15-17) and (15-28), we obtain

$$\begin{aligned} \mathbf{P}_x(k + 1) &= \mathbf{E}\{[\mathbf{x}(k + 1) - \mathbf{m}_x(k + 1)][\mathbf{x}(k + 1) - \mathbf{m}_x(k + 1)]'\} \\ &= \mathbf{E}\{[\Phi[\mathbf{x}(k) - \mathbf{m}_x(k)] + \Gamma\mathbf{w}(k)]' \\ &\quad [\Phi[\mathbf{x}(k) - \mathbf{m}_x(k)] + \Gamma\mathbf{w}(k)]\} \\ &= \Phi\mathbf{P}_x(k)\Phi' + \Gamma\mathbf{Q}(k)\Gamma' + \Phi\mathbf{E}\{[\mathbf{x}(k) - \mathbf{m}_x(k)]\mathbf{w}'(k)\}\Gamma' \\ &\quad + \Gamma\mathbf{E}\{\mathbf{w}(k)[\mathbf{x}(k) - \mathbf{m}_x(k)]'\}\Phi' \end{aligned} \quad (15-31)$$

Because $\mathbf{m}_x(k)$ is not random and $\mathbf{w}(k)$ is zero mean, $\mathbf{E}\{\mathbf{m}_x(k)\mathbf{w}'(k)\} =$

$\mathbf{m}_x(k)\mathbf{E}\{\mathbf{w}'(k)\} = \mathbf{0}$, and $\mathbf{E}\{\mathbf{w}(k)\mathbf{m}_x'(k)\} = \mathbf{0}$. State vector $\mathbf{x}(k)$ depends at most on random input $\mathbf{w}(k - 1)$ [see (15-17)]; hence,

$$\mathbf{E}\{\mathbf{x}(k)\mathbf{w}'(k)\} = \mathbf{E}\{\mathbf{x}(k)\}\mathbf{E}\{\mathbf{w}'(k)\} = \mathbf{0} \quad (15-32)$$

and $\mathbf{E}\{\mathbf{w}(k)\mathbf{x}'(k)\} = \mathbf{0}$ as well. The last two terms in (15-31) are therefore equal to zero, and the equation reduces to (15-29).

- c. We leave the proof of (15-30) as an exercise. Observe that once we know covariance matrix $\mathbf{P}_x(k)$ it is an easy matter to determine any cross-covariance matrix between state $\mathbf{x}(k)$ and $\mathbf{x}(i)$ ($i \neq k$). The Markov nature of our basic state-variable model is responsible for this. \square

Observe that mean vector $\mathbf{m}_x(k)$ satisfies a deterministic vector state equation, (15-28), covariance matrix $\mathbf{P}_x(k)$ satisfies a deterministic matrix state equation, (15-29), and (15-28) and (15-29) are easily programmed for digital computation.

Next we direct our attention to the statistics of measurement vector $\mathbf{z}(k)$.

Theorem 15-6. *For our basic state-variable model, when $\mathbf{x}(0)$, $\mathbf{w}(k)$ and $\mathbf{v}(k)$ are jointly Gaussian, then $\{\mathbf{z}(k), k = 1, 2, \dots\}$ is Gaussian, and*

$$\mathbf{m}_z(k + 1) = \mathbf{H}(k + 1)\mathbf{m}_x(k + 1) \quad (15-33)$$

and

$$\mathbf{P}_z(k + 1) = \mathbf{H}(k + 1)\mathbf{P}_x(k + 1)\mathbf{H}'(k + 1) + \mathbf{R}(k + 1) \quad (15-34)$$

where $\mathbf{m}_x(k + 1)$ and $\mathbf{P}_x(k + 1)$ are computed from (15-28) and (15-29), respectively. \square

We leave the proof as an exercise for the reader. Note that if $\mathbf{x}(0)$, $\mathbf{w}(k)$ and $\mathbf{v}(k)$ are statistically independent and Gaussian they will be jointly Gaussian.

Example 15-1

Consider the simple single-input single-output first-order system

$$x(k + 1) = \frac{1}{2}x(k) + w(k) \quad (15-35)$$

$$z(k + 1) = x(k + 1) + v(k + 1) \quad (15-36)$$

where $w(k)$ and $v(k)$ are wide-sense stationary white noise processes, for which $q = 20$ and $r = 5$. Additionally, $m_x(0) = 4$ and $p_x(0) = 10$.

The mean of $x(k)$ is computed from the following homogeneous equation

$$m_x(k + 1) = \frac{1}{2}m_x(k) \quad m_x(0) = 4 \quad (15-37)$$

and, the variance of $x(k)$ is computed from Equation (15-29), which in this case simplifies to

$$p_x(k+1) = \frac{1}{4} p_x(k) + 20 \quad p_x(0) = 10 \quad (15-38)$$

Additionally, the mean and variance of $z(k)$ are computed from

$$m_z(k+1) = m_x(k+1) \quad (15-39)$$

and

$$p_z(k+1) = p_x(k+1) + 5 \quad (15-40)$$

Figure 15-1 depicts $m_x(k)$ and $p_x^{1/2}(k)$. Observe that $m_x(k)$ decays to zero very rapidly and that $p_x^{1/2}(k)$ approaches a steady-state value, $\bar{p}_x^{1/2} = 5.163$. This steady-state value can be computed from equation (15-38) by setting $p_x(k) = p_x(k+1) = \bar{p}_x$. The existence of \bar{p}_x is guaranteed by our first-order system being stable.

Although $m_x(k) \rightarrow 0$ there is a lot of uncertainty about $x(k)$, as evidenced by the large value of \bar{p}_x . There will be an even larger uncertainty about $z(k)$, because $\bar{p}_z \rightarrow 31.66$. These large values for \bar{p}_z and \bar{p}_x are due to the large values of q and r . In many practical applications, both q and r will be much less than unity in which case \bar{p}_z and \bar{p}_x will be quite small. \square

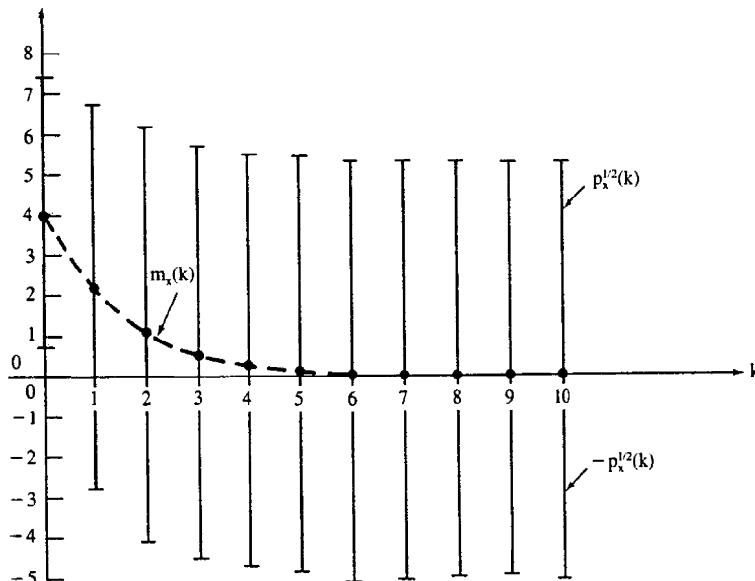


Figure 15-1 Mean (dashed) and standard deviation (bars) for first-order system (15-35) and (15-36).

If our basic state-variable model is time-invariant and stationary, and, if Φ is associated with an asymptotically stable system (i.e., one whose poles all lie within the unit circle), then (Anderson and Moore, 1979) matrix $\mathbf{P}_x(k)$ reaches a limiting (steady-state) solution $\bar{\mathbf{P}}_x$, i.e.,

$$\lim_{k \rightarrow \infty} \mathbf{P}_x(k) = \bar{\mathbf{P}}_x \quad (15-41)$$

Matrix $\bar{\mathbf{P}}_x$ is the solution of the following steady-state version of (15-29),

$$\bar{\mathbf{P}}_x = \Phi \bar{\mathbf{P}}_x \Phi' + \Gamma Q \Gamma' \quad (15-42)$$

This equation is called a *discrete-time Lyapunov equation*. See Laub (1979) for an excellent numerical method that can be used to solve (15-42) for $\bar{\mathbf{P}}_x$.

SIGNAL-TO-NOISE RATIO

In this section we simplify our basic state-variable model (15-17) and (15-18) to a time-invariant, stationary, single-input single-output model:

$$\mathbf{x}(k+1) = \Phi \mathbf{x}(k) + \gamma w(k) + \psi u(k) \quad (15-43)$$

$$z(k+1) = \mathbf{h}' \mathbf{x}(k+1) + v(k+1) \quad (15-44)$$

Measurement $z(k)$ is of the classical form of signal plus noise, where "signal" $s(k) = \mathbf{h}' \mathbf{x}(k)$.

The *signal-to-noise ratio* is an often-used measure of quality of measurement $z(k)$. Here we define that ratio, denoted by $\text{SNR}(k)$, as

$$\text{SNR}(k) = \frac{\sigma_s^2(k)}{r} \quad (15-45)$$

From preceding analyses, we see that

$$\text{SNR}(k) = \frac{\mathbf{h}' \mathbf{P}_x(k) \mathbf{h}}{r} \quad (15-46)$$

Because $\mathbf{P}_x(k)$ is in general a function of time, $\text{SNR}(k)$ is also a function of time. If, however, Φ is associated with an asymptotically stable system then (15-41) is true. In this case we can use $\bar{\mathbf{P}}_x$ in (15-46) to provide us with a single number, $\overline{\text{SNR}}$, for the signal-to-noise ratio, i.e.,

$$\overline{\text{SNR}} = \frac{\mathbf{h}' \bar{\mathbf{P}}_x \mathbf{h}}{r} \quad (15-47)$$

Finally, we demonstrate that $\text{SNR}(k)$ (or $\overline{\text{SNR}}$) can be computed without knowing q and r explicitly; all that is needed is the ratio q/r . Multiplying and dividing the right-hand side of (15-46) by q , we find that

$$\text{SNR}(k) = \left[\mathbf{h}' \frac{\mathbf{P}_x(k)}{q} \mathbf{h} \right] (q/r) \quad (15-48)$$

Scaled covariance matrix $\mathbf{P}_x(k)/q$ is computed from the following version of (15-29)

$$\frac{\mathbf{P}_x(k+1)}{q} = \Phi \frac{\mathbf{P}_x(k)}{q} \Phi' + \gamma \gamma' \quad (15-49)$$

One of the most useful ways for using (15-48) is to compute q/r for a given signal-to-noise ratio $\overline{\text{SNR}}$, i.e.,

$$\frac{q}{r} = \frac{\overline{\text{SNR}}}{\mathbf{h}' \frac{\mathbf{P}_x}{q} \mathbf{h}} \quad (15-50)$$

In Lesson 18 we show that q/r can be viewed as an estimator tuning parameter; hence, signal-to-noise ratio, SNR , can also be treated as such a parameter.

Example 15-2 (Mendel, 1981)

Consider the first-order system

$$x(k+1) = \phi x(k) + \gamma w(k) \quad (15-51)$$

$$z(k+1) = h x(k+1) + v(k) \quad (15-52)$$

In this case, it is easy to solve (15-49), to show that

$$\overline{P}_x/q = \gamma^2/(1 - \phi^2) \quad (15-53)$$

hence,

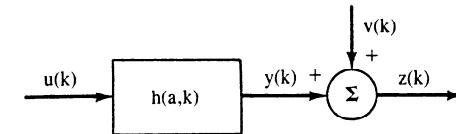
$$\overline{\text{SNR}} = \left(\frac{h^2 \gamma^2}{1 - \phi^2} \right) \left(\frac{q}{r} \right) \quad (15-54)$$

Observe that, if $h^2 \gamma^2 = 1 - \phi^2$, then $\overline{\text{SNR}} = q/r$. The condition $h^2 \gamma^2 = 1 - \phi^2$ is satisfied if, for example, $\gamma = 1$, $\phi = 1/\sqrt{2}$ and $h = 1/\sqrt{2}$. \square

PROBLEMS

- 15-1. Prove Theorem 15-2, and then show that for Gaussian white noise $E\{s(t_m)|s(t_{m-1}), \dots, s(t_1)\} = E\{s(t_m)\}$.
- 15-2. Derive the formula for the cross-covariance of $x(k)$, $\mathbf{P}_x(i, j)$, given in (15-30).
- 15-3. Derive the first- and second-order statistics of measurement vector $z(k)$, that are summarized in Theorem 15-6.
- 15-4. Reconsider the basic state-variable model when $x(0)$ is correlated with $w(0)$, and $w(k)$ and $v(k)$ are correlated [$E\{w(k)v'(k)\} = S(k)\}$.
 - (a) Show that the covariance equation for $z(k)$ remains unchanged.
 - (b) Show that the covariance equation for $x(k)$ is changed, but only at $k = 1$.
 - (c) Compute $E\{z(k+1)z'(k)\}$.

15-5.



In this problem, assume that $u(k)$ and $v(k)$ are individually Gaussian and uncorrelated. Impulse response h depends on parameter a , where a is a Gaussian random variable that is statistically independent of $u(k)$ and $v(k)$.

- (a) Evaluate $E\{z(k)\}$.
- (b) Explain whether or not $y(k)$ is Gaussian.

State Estimation: Prediction

INTRODUCTION

We have mentioned, a number of times in this book, that in state estimation three situations are possible depending upon the relative relationship of total number of available measurements, N , and the time point, k , at which we estimate state vector $\mathbf{x}(k)$, namely: prediction ($N < k$), filtering ($N = k$), and smoothing ($N > k$). In this lesson we develop algorithms for mean-squared predicted estimates, $\hat{\mathbf{x}}_{\text{MS}}(k|j)$, of state $\mathbf{x}(k)$. In order to simplify our notation, we shall abbreviate $\hat{\mathbf{x}}_{\text{MS}}(k|j)$ as $\hat{\mathbf{x}}(k|j)$. (Just in case you have forgotten what the notation $\hat{\mathbf{x}}(k|j)$ stands for, see Lesson 2.) Note that, in prediction, $k > j$.

SINGLE-STAGE PREDICTOR

The most important predictor of $\mathbf{x}(k)$ for our future work on filtering and smoothing is the *single-stage predictor* $\hat{\mathbf{x}}(k|k - 1)$. From the Fundamental Theorem of Estimation Theory (Theorem 13-1), we know that

$$\hat{\mathbf{x}}(k|k - 1) = \mathbf{E}\{\mathbf{x}(k)|\mathcal{Z}(k - 1)\} \quad (16-1)$$

where

$$\mathcal{Z}(k - 1) = \text{col}(\mathbf{z}(1), \mathbf{z}(2), \dots, \mathbf{z}(k - 1)) \quad (16-2)$$

It is very easy to derive a formula for $\hat{\mathbf{x}}(k|k - 1)$ by operating on both sides of the state equation

$$\begin{aligned} \mathbf{x}(k) &= \Phi(k, k - 1)\mathbf{x}(k - 1) + \Gamma(k, k - 1)\mathbf{w}(k - 1) \\ &\quad + \Psi(k, k - 1)\mathbf{u}(k - 1) \end{aligned} \quad (16-3)$$

with the linear expectation operator $\mathbf{E}\{\cdot|\mathcal{Z}(k - 1)\}$. Doing this, we find that

$$\begin{aligned} \hat{\mathbf{x}}(k|k - 1) &= \Phi(k, k - 1)\hat{\mathbf{x}}(k - 1|k - 1) \\ &\quad + \Psi(k, k - 1)\mathbf{u}(k - 1) \end{aligned} \quad (16-4)$$

where $k = 1, 2, \dots$. To obtain (16-4) we have used the facts that $\mathbf{E}\{\mathbf{w}(k - 1)\} = \mathbf{0}$ and $\mathbf{u}(k - 1)$ is deterministic.

Observe, from (16-4), that the single-stage predicted estimate, $\hat{\mathbf{x}}(k|k - 1)$, depends on the filtered estimate, $\hat{\mathbf{x}}(k - 1|k - 1)$, of the preceding state vector $\mathbf{x}(k - 1)$. At this point, (16-4) is an interesting theoretical result; but, there is nothing much we can do with it, because we do not as yet know how to compute filtered state estimates. In Lesson 17 we shall begin our study into filtered state estimates, and shall learn that such estimates of $\mathbf{x}(k)$ depend on predicted estimates of $\mathbf{x}(k)$, just as predicted estimates of $\mathbf{x}(k)$ depend on filtered estimates of $\mathbf{x}(k - 1)$; thus, *filtered and predicted state estimates are very tightly coupled together*.

Let $\mathbf{P}(k|k - 1)$ denote the error-covariance matrix that is associated with $\hat{\mathbf{x}}(k|k - 1)$, i.e.,

$$\begin{aligned} \mathbf{P}(k|k - 1) &= \mathbf{E}\{[\hat{\mathbf{x}}(k|k - 1) \\ &\quad - \mathbf{m}_{\hat{\mathbf{x}}}(k|k - 1)][\hat{\mathbf{x}}(k|k - 1) - \mathbf{m}_{\hat{\mathbf{x}}}(k|k - 1)]'\} \end{aligned} \quad (16-5)$$

where

$$\hat{\mathbf{x}}(k|k - 1) = \mathbf{x}(k) - \hat{\mathbf{x}}(k|k - 1) \quad (16-6)$$

Additionally, let $\mathbf{P}(k - 1|k - 1)$ denote the error-covariance matrix that is associated with $\hat{\mathbf{x}}(k - 1|k - 1)$, i.e.,

$$\begin{aligned} \mathbf{P}(k - 1|k - 1) &= \mathbf{E}\{[\hat{\mathbf{x}}(k - 1|k - 1) \\ &\quad - \mathbf{m}_{\hat{\mathbf{x}}}(k - 1|k - 1)][\hat{\mathbf{x}}(k - 1|k - 1) - \mathbf{m}_{\hat{\mathbf{x}}}(k - 1|k - 1)]'\} \end{aligned} \quad (16-7)$$

For our basic state-variable model (see Property 1 of Lesson 13), $\mathbf{m}_{\hat{\mathbf{x}}}(k|k - 1) = \mathbf{0}$ and $\mathbf{m}_{\hat{\mathbf{x}}}(k - 1|k - 1) = \mathbf{0}$, so that

$$\mathbf{P}(k|k - 1) = \mathbf{E}\{\hat{\mathbf{x}}(k|k - 1)\hat{\mathbf{x}}'(k|k - 1)\} \quad (16-8)$$

and

$$\mathbf{P}(k - 1|k - 1) = \mathbf{E}\{\hat{\mathbf{x}}(k - 1|k - 1)\hat{\mathbf{x}}'(k - 1|k - 1)\} \quad (16-9)$$

Combining (16-3) and (16-4), we see that

$$\begin{aligned} \hat{\mathbf{x}}(k|k - 1) &= \Phi(k, k - 1)\hat{\mathbf{x}}(k - 1|k - 1) \\ &\quad + \Gamma(k, k - 1)\mathbf{w}(k - 1) \end{aligned} \quad (16-10)$$

A straightforward calculation leads to the following formula for $\mathbf{P}(k|k-1)$,

$$\boxed{\mathbf{P}(k|k-1) = \Phi(k,k-1)\mathbf{P}(k-1|k-1)\Phi'(k,k-1) + \Gamma(k,k-1)\mathbf{Q}(k-1)\Gamma'(k,k-1)} \quad (16-11)$$

where $k = 1, 2, \dots$

Observe, from (16-4) and (16-11), that $\hat{\mathbf{x}}(0|0)$ and $\mathbf{P}(0|0)$ initialize the single-stage predictor and its error-covariance. Additionally,

$$\hat{\mathbf{x}}(0|0) = \mathbf{E}\{\mathbf{x}(0)|\text{no measurements}\} = \mathbf{m}_x(0) \quad (16-12)$$

and

$$\begin{aligned} \mathbf{P}(0|0) &= \mathbf{E}\{\hat{\mathbf{x}}(0|0)\hat{\mathbf{x}}'(0|0)\} \\ &= \mathbf{E}\{[\mathbf{x}(0) - \mathbf{m}_x(0)][\mathbf{x}(0) - \mathbf{m}_x(0)]'\} = \mathbf{P}_x(0) \end{aligned} \quad (16-13)$$

Finally, recall (Property 4 of Lesson 13) that both $\hat{\mathbf{x}}(k|k-1)$ and $\tilde{\mathbf{x}}(k|k-1)$ are Gaussian.

A GENERAL STATE PREDICTOR

In this section we generalize the results of the preceding section so as to obtain predicted values of $\mathbf{x}(k)$ that look further into the future than just one step. We shall determine $\hat{\mathbf{x}}(k|j)$ where $k > j$ under the assumption that filtered state estimate $\hat{\mathbf{x}}(j|j)$ and its error-covariance matrix $\mathbf{E}\{\hat{\mathbf{x}}(j|j)\hat{\mathbf{x}}'(j|j)\} = \mathbf{P}(j|j)$ are known for some $j = 0, 1, \dots$

Theorem 16-1

- a. If input $\mathbf{u}(k)$ is deterministic, or does not depend on any measurements, then the mean-squared predicted estimator of $\mathbf{x}(k)$, $\hat{\mathbf{x}}(k|j)$, is given by the expression

$$\begin{aligned} \hat{\mathbf{x}}(k|j) &= \Phi(k,j)\hat{\mathbf{x}}(j|j) \\ &\quad + \sum_{i=j+1}^k \Phi(k,i)\Psi(i,i-1)\mathbf{u}(i-1) \quad k > j \end{aligned} \quad (16-14)$$

- b. The vector random process $\{\tilde{\mathbf{x}}(k|j), k = j+1, j+2, \dots\}$ is:

- i. zero mean,
- ii. Gaussian, and
- iii. first-order Markov, and
- iv. its covariance matrix is governed by

$$\begin{aligned} \mathbf{P}(k|j) &= \Phi(k,k-1)\mathbf{P}(k-1|j)\Phi'(k,k-1) \\ &\quad + \Gamma(k,k-1)\mathbf{Q}(k-1)\Gamma'(k,k-1) \end{aligned} \quad (16-15)$$

Before proving this theorem, let us observe that the prediction formula (16-14) is intuitively what one would expect. Why is this so? Suppose we have processed all of the measurements $\mathbf{z}(1), \mathbf{z}(2), \dots, \mathbf{z}(j)$ to obtain $\hat{\mathbf{x}}(j|j)$ and are asked to predict the value of $\mathbf{x}(k)$, where $k > j$. No additional measurements can be used during prediction. All that we can therefore use is our dynamical state equation. When that equation is used for purposes of prediction we neglect the random disturbance term, because the disturbances are not measurable. We can only use measured quantities to assist our prediction efforts. The simplified state equation is

$$\mathbf{x}(k+1) = \Phi(k+1,k)\mathbf{x}(k) + \Psi(k+1,k)\mathbf{u}(k) \quad (16-16)$$

a solution of which is

$$\mathbf{x}(k) = \Phi(k,j)\mathbf{x}(j) + \sum_{i=j+1}^k \Phi(k,i)\Psi(i,i-1)\mathbf{u}(i-1) \quad (16-17)$$

Substituting $\hat{\mathbf{x}}(j|j)$ for $\mathbf{x}(j)$, we obtain the predictor in (16-14). In our proof of Theorem 16-1 we establish (16-14) in a more rigorous manner.

Proof

- a. The solution to state equation (16-3), for $\mathbf{x}(k)$, can be expressed in terms of $\mathbf{x}(j)$, where $j < k$, as

$$\begin{aligned} \mathbf{x}(k) &= \Phi(k,j)\mathbf{x}(j) + \sum_{i=j+1}^k \Phi(k,i)[\Gamma(i,i-1)\mathbf{w}(i-1) \\ &\quad + \Psi(i,i-1)\mathbf{u}(i-1)] \end{aligned} \quad (16-18)$$

We apply the Fundamental Theorem of Estimation Theory to (16-18) by taking the conditional expectation with respect to $\mathcal{E}(j)$ on both sides of it. Doing this, we find that

$$\begin{aligned} \hat{\mathbf{x}}(k|j) &= \Phi(k,j)\hat{\mathbf{x}}(j|j) + \sum_{i=j+1}^k \Phi(k,i)[\Gamma(i,i-1)\mathbf{E}\{\mathbf{w}(i-1)|\mathcal{E}(j)\} \\ &\quad + \Psi(i,i-1)\mathbf{E}\{\mathbf{u}(i-1)|\mathcal{E}(j)\}] \end{aligned} \quad (16-19)$$

Note that $\mathcal{E}(j)$ depends at most on $\mathbf{x}(j)$ which, in turn, depends at most on $\mathbf{w}(j-1)$. Consequently,

$$\begin{aligned} \mathbf{E}\{\mathbf{w}(i-1)|\mathcal{E}(j)\} &= \mathbf{E}\{\mathbf{w}(i-1)|\mathbf{w}(0), \mathbf{w}(1), \dots, \mathbf{w}(j-1)\} \end{aligned} \quad (16-20)$$

where $i = j+1, j+2, \dots, k$. Because of this range of values on argument i , $\mathbf{w}(i-1)$ is never included in the conditioning set of values $\mathbf{w}(0), \mathbf{w}(1), \dots, \mathbf{w}(j-1)$; hence,

$$\mathbf{E}\{\mathbf{w}(i-1)|\mathcal{E}(j)\} = \mathbf{E}\{\mathbf{w}(i-1)\} = \mathbf{0} \quad (16-21)$$

for all $i = j+1, j+2, \dots, k$.

Note, also, that

$$\mathbf{E}\{\mathbf{u}(i-1)|\mathcal{E}(j)\} = \mathbf{E}\{\mathbf{u}(i-1)\} = \mathbf{u}(i-1) \quad (16-22)$$

because we have assumed that $u(i-1)$ does not depend on any of the measurements. Substituting (16-21) and (16-22) into (16-19), we obtain the prediction formula (16-14).

- b-i. and b-ii. have already been proved in Properties 1 and 4 of Lesson 13.
- b-iii. Starting with $\tilde{x}(k|j) = x(k) - \hat{x}(k|j)$, and substituting (16-18) and (16-14) into this relation, we find that

$$\tilde{x}(k|j) = \Phi(k,j)\tilde{x}(j|j) + \sum_{i=j+1}^k \Phi(k,i)\Gamma(i,i-1)w(i-1) \quad (16-23)$$

This equation looks quite similar to the solution of state equation (16-3), when $u(k) = 0$ (for all k), e.g., see (16-18). In fact, $\tilde{x}(k|j)$ also satisfies the state equation

$$\tilde{x}(k|j) = \Phi(k,k-1)\tilde{x}(k-1|j) + \Gamma(k,k-1)w(k-1) \quad (16-24)$$

Because $\tilde{x}(k|j)$ depends only upon its previous value, $\tilde{x}(k-1|j)$, it is first-order Markov.

- b-iv. We derived a recursive covariance equation for $x(k)$ in Theorem 15-5. That equation is (15-29). Because $\tilde{x}(k|j)$ satisfies the state equation for $x(k)$, its covariance $P(k|j)$ is also given by (15-29). We have rewritten this equation as in (16-15). \square

Observe that by setting $j = k-1$ in (16-14) we obtain our previously derived single-stage predictor $\hat{x}(k|k-1)$.

Theorem 16-1 is quite limited because presently the only values of $\hat{x}(j|j)$ and $P(j|j)$ that we know are those at $j=0$. For $j=0$, (16-14) becomes

$$\hat{x}(k|0) = \Phi(k,0)m_x(0) + \sum_{i=1}^k \Phi(k,i)\Psi(i,i-1)u(i-1) \quad (16-25)$$

The reader might feel that this predictor of $x(k)$ becomes poorer and poorer as k gets farther and farther away from zero. The following example demonstrates that this is not necessarily true.

Example 16-1

Let us examine prediction performance, as measured by $p(k|0)$, for the first-order system

$$x(k+1) = \frac{1}{\sqrt{2}}x(k) + w(k) \quad (16-26)$$

where $q = 25$ and $p(0)$ is variable. Quantity $p(k|0)$, which in the case of a scalar state vector is a variance, is easily computed from the recursive equation

$$p(k|0) = \frac{1}{2}p(k-1|0) + 25 \quad (16-27)$$

for $k = 1, 2, \dots$. Two cases are summarized in Figure 16-1. When $p(0) = 6$ we have

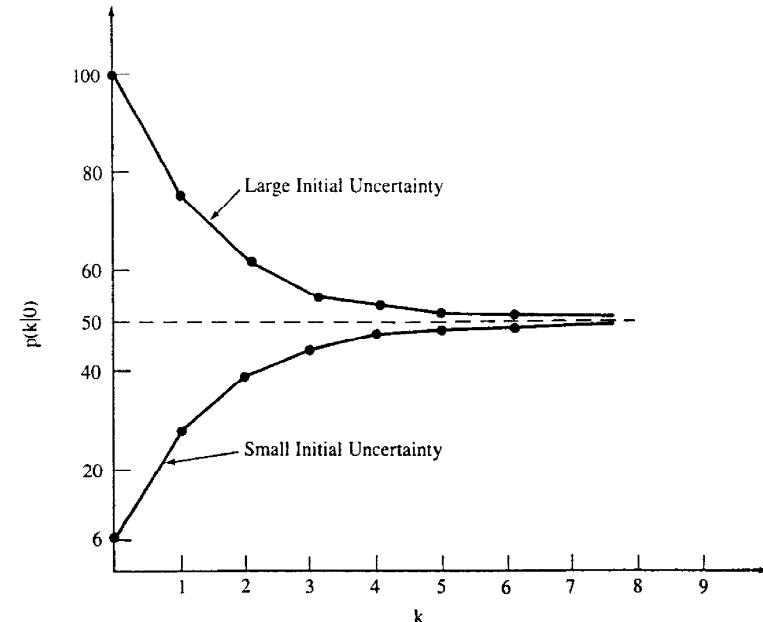


Figure 16-1 Prediction error variance $p(k|0)$.

relatively small uncertainty about $\hat{x}(0|0)$, and as we expected, our predictions of $x(k)$ for $k \geq 1$ do become worse, because $p(k|0) > 6$ for all $k \geq 1$. After a while $p(k|0)$ reaches a limiting value equal to 50. When this occurs we are estimating $\hat{x}(k|0)$ by a number that is very close to zero, because $\hat{x}(k|0) = \left(\frac{1}{\sqrt{2}}\right)^k \hat{x}(0|0)$, and $\left(\frac{1}{\sqrt{2}}\right)^k$ approaches zero for large values of k .

When $p(0) = 100$ we have large uncertainty about $\hat{x}(0|0)$, and, perhaps to our surprise, our predictions of $x(k)$ for $k \geq 1$ improve in performance, because $p(k|0) < 100$ for all $k \geq 1$. In this case the predictor discounts the large initial uncertainty; however, as in the former case, $p(k|0)$ again reaches the limiting value of 50.

For suitably large values of k , the predictor is completely insensitive to $p(0)$. It reaches a steady-state level of performance equal to 50, which can be predetermined by setting $p(k|0)$ and $p(k-1|0)$ equal to \bar{p} , in (16-27), and solving the resulting equation for \bar{p} . \square

Prediction is possible only because we have a known dynamical model, namely, our state-variable model. Without a model, prediction is dubious at best (e.g., try predicting tomorrow's price of a stock listed on any stock exchange using today's closing price).

THE INNOVATIONS PROCESS

Suppose we have just computed the single-stage, predicted estimate of $\mathbf{x}(k+1)$, $\hat{\mathbf{x}}(k+1|k)$. Then the single-stage predicted estimate of $\mathbf{z}(k+1)$, $\hat{\mathbf{z}}(k+1|k)$, is

$$\hat{\mathbf{z}}(k+1|k) = \mathbf{E}\{\mathbf{z}(k+1)|\mathcal{Z}(k)\} = \mathbf{E}\{[\mathbf{H}(k+1)\mathbf{x}(k+1) + \mathbf{v}(k+1)]|\mathcal{Z}(k)\}$$

or

$$\hat{\mathbf{z}}(k+1|k) = \mathbf{H}(k+1)\hat{\mathbf{x}}(k+1|k) \quad (16-28)$$

The error between $\mathbf{z}(k+1)$ and $\hat{\mathbf{z}}(k+1|k)$ is $\tilde{\mathbf{z}}(k+1|k)$, i.e.,

$$\tilde{\mathbf{z}}(k+1|k) = \mathbf{z}(k+1) - \hat{\mathbf{z}}(k+1|k) \quad (16-29)$$

Signal $\tilde{\mathbf{z}}(k+1|k)$ is often referred to either as the *innovations process*, *prediction error process*, or *measurement residual process*. We shall refer to it as the innovations process, because this is most commonly done in the estimation theory literature (e.g., Kailath, 1968). The innovations process plays a very important role in mean-squared filtering and smoothing. We summarize important facts about it in the following:

Theorem 16-2 (Innovations)

- a. The following representations of the innovations process $\tilde{\mathbf{z}}(k+1|k)$ are equivalent:

$$\tilde{\mathbf{z}}(k+1|k) = \mathbf{z}(k+1) - \hat{\mathbf{z}}(k+1|k) \quad (16-30)$$

$$\tilde{\mathbf{z}}(k+1|k) = \mathbf{z}(k+1) - \mathbf{H}(k+1)\hat{\mathbf{x}}(k+1|k) \quad (16-31)$$

or

$$\tilde{\mathbf{z}}(k+1|k) = \mathbf{H}(k+1)\hat{\mathbf{x}}(k+1|k) + \mathbf{v}(k+1) \quad (16-32)$$

- b. The innovations is a zero-mean Gaussian white noise sequence, with

$$\begin{aligned} \mathbf{E}\{\tilde{\mathbf{z}}(k+1|k)\tilde{\mathbf{z}}'(k+1|k)\} &= \mathbf{P}_{\tilde{\mathbf{z}}}(k+1|k) \\ &= \mathbf{H}(k+1)\mathbf{P}(k+1|k)\mathbf{H}'(k+1) + \mathbf{R}(k+1) \end{aligned} \quad (16-33)$$

Proof (Mendel, 1983b)

- a. Substitute (16-28) into (16-29) in order to obtain (16-31). Next, substitute the measurement equation $\mathbf{z}(k+1) = \mathbf{H}(k+1)\mathbf{x}(k+1) + \mathbf{v}(k+1)$ into (16-31), and use the fact that $\hat{\mathbf{x}}(k+1|k) = \mathbf{x}(k+1) - \hat{\mathbf{x}}(k+1|k)$, to obtain (16-32).
- b. Because $\hat{\mathbf{x}}(k+1|k)$ and $\mathbf{v}(k+1)$ are both zero mean, $\mathbf{E}\{\tilde{\mathbf{z}}(k+1|k)\} = \mathbf{0}$. The innovations is Gaussian because $\mathbf{z}(k+1)$ and $\hat{\mathbf{x}}(k+1|k)$ are Gaussian, and, therefore, $\tilde{\mathbf{z}}(k+1|k)$ is a linear transformation of Gaussian

random vectors. To prove that $\tilde{\mathbf{z}}(k+1|k)$ is white noise we must show that

$$\mathbf{E}\{\tilde{\mathbf{z}}(i+1|i)\tilde{\mathbf{z}}'(j+1|j)\} = \mathbf{P}_{\tilde{\mathbf{z}}}(i+1|i)\delta_{ij} \quad (16-34)$$

We shall consider the cases $i > j$ and $i = j$, leaving the case $i < j$ as an exercise for the reader. When $i > j$,

$$\begin{aligned} \mathbf{E}\{\tilde{\mathbf{z}}(i+1|i)\tilde{\mathbf{z}}'(j+1|j)\} &= \mathbf{E}\{[\mathbf{H}(i+1)\hat{\mathbf{x}}(i+1|i) + \mathbf{v}(i+1)] \\ &\quad [\mathbf{H}(j+1)\hat{\mathbf{x}}(j+1|j) + \mathbf{v}(j+1)]'\} \\ &= \mathbf{E}\{\mathbf{H}(i+1)\hat{\mathbf{x}}(i+1|i) \\ &\quad [\mathbf{H}(j+1)\hat{\mathbf{x}}(j+1|j) + \mathbf{v}(j+1)]'\} \end{aligned}$$

because $\mathbf{E}\{\mathbf{v}(i+1)\mathbf{v}'(j+1)\} = \mathbf{0}$ and $\mathbf{E}\{\mathbf{v}(i+1)\hat{\mathbf{x}}'(j+1|j)\} = \mathbf{0}$. The latter is true because, for $i > j$, $\hat{\mathbf{x}}(j+1|j)$ does not depend on measurement $\mathbf{z}(i+1)$; hence, for $i > j$, $\mathbf{v}(i+1)$ and $\hat{\mathbf{x}}(j+1|j)$ are independent, so that $\mathbf{E}\{\mathbf{v}(i+1)\hat{\mathbf{x}}'(j+1|j)\} = \mathbf{E}\{\mathbf{v}(i+1)\}\mathbf{E}\{\hat{\mathbf{x}}'(j+1|j)\} = \mathbf{0}$. We continue, as follows:

$$\begin{aligned} \mathbf{E}\{\tilde{\mathbf{z}}(i+1|i)\tilde{\mathbf{z}}'(j+1|j)\} \\ &= \mathbf{H}(i+1)\mathbf{E}\{\hat{\mathbf{x}}(i+1|i)[\mathbf{z}(j+1) - \mathbf{H}(j+1)\hat{\mathbf{x}}(j+1|j)]'\} = \mathbf{0} \end{aligned}$$

by repeated application of the orthogonality principle (Corollary 13-3).

When $i = j$,

$$\begin{aligned} \mathbf{P}_{\tilde{\mathbf{z}}}(i+1|i) &= \mathbf{E}\{[\mathbf{H}(i+1)\hat{\mathbf{x}}(i+1|i) + \mathbf{v}(i+1)][\mathbf{H}(i+1)\hat{\mathbf{x}}(i+1|i) \\ &\quad + \mathbf{v}(i+1)]'\} = \mathbf{H}(i+1)\mathbf{P}(i+1|i)\mathbf{H}'(i+1) + \mathbf{R}(i+1) \end{aligned}$$

because, once again $\mathbf{E}\{\mathbf{v}(i+1)\hat{\mathbf{x}}'(i+1|i)\} = \mathbf{0}$, and $\mathbf{P}(i+1) = \mathbf{E}\{\hat{\mathbf{x}}(i+1|i)\hat{\mathbf{x}}'(i+1|i)\}$. \square

In Lesson 17 the inverse of $\mathbf{P}_{\tilde{\mathbf{z}}}(k+1|k)$ is needed; hence, we shall assume that $\mathbf{H}(k+1)\mathbf{P}(k+1|k)\mathbf{H}'(k+1) + \mathbf{R}(k+1)$ is nonsingular. This is usually true and will always be true if, as in our basic state-variable model, $\mathbf{R}(k+1)$ is positive definite.

PROBLEMS

- 16-1. Develop the counterpart to Theorem 16-1 for the case when input $\mathbf{u}(k)$ is random and independent of $\mathcal{Z}(j)$. What happens if $\mathbf{u}(k)$ is random and dependent upon $\mathcal{Z}(j)$?
- 16-2. For the innovations process $\tilde{\mathbf{z}}(k+1|k)$, prove that $\mathbf{E}\{\tilde{\mathbf{z}}(i+1|i)\tilde{\mathbf{z}}'(j+1|j)\} = \mathbf{0}$ when $i < j$.
- 16-3. In the proof of part (b) of Theorem 16-2 we make repeated use of the orthogonality principle, stated in Corollary 13-3. In the latter corollary $\mathbf{f}[\mathcal{Z}(k)]$

appears to be a function of *all* of the measurements used in $\hat{\theta}_{MS}(k)$. In the expression $E\{\tilde{x}(i+1|i)z'(j+1)\}, i > j$, $z'(j+1)$ certainly is not a function of all the measurements used in $\tilde{x}(i+1|i)$. What is $f[\cdot]$, when we apply the orthogonality principle to $E\{\tilde{x}(i+1|i)z'(j+1)\}, i > j$, to conclude that this expectation is zero?

- 16-4. Refer to Problem 15-5. Assume that $u(k)$ can be measured [e.g., $u(k)$ might be the output of a random number generator], and that $a = a[z(1), z(2), \dots, z(k-1)]$. What is $\hat{z}(k|k-1)$?
- 16-5. Consider the following autoregressive model, $z(k+n) = a_1z(k+n-1) - a_2z(k+n-2) - \dots - a_nz(k) + w(k+n)$ in which $w(k)$ is white noise. Measurements $z(k), z(k+1), \dots, z(k+n-1)$ are available.
- Compute $\hat{z}(k+n|k+n-1)$.
 - Explain why the result in (a) is the overall mean-squared prediction of $z(k+n)$ even if $w(k+n)$ is non-Gaussian.

Lesson 17

State Estimation: Filtering (the Kalman Filter)

INTRODUCTION

In this lesson we shall develop the Kalman filter, which is a recursive mean-squared error filter for computing $\hat{x}(k+1|k+1)$, $k = 0, 1, 2, \dots$. As its name implies, this filter was developed by Kalman [circa 1959 (Kalman, 1960)].

From the Fundamental Theorem of Estimation Theory, Theorem 13-1, we know that

$$\hat{x}(k+1|k+1) = E\{x(k+1)|\mathcal{Z}(k+1)\} \quad (17-1)$$

Our approach to developing the Kalman filter is to partition $\mathcal{Z}(k+1)$ into two sets of measurements, $\mathcal{Z}(k)$ and $z(k+1)$, and to then expand the conditional expectation in terms of data sets $\mathcal{Z}(k)$ and $z(k+1)$, i.e.,

$$\hat{x}(k+1|k+1) = E\{x(k+1)|\mathcal{Z}(k), z(k+1)\} \quad (17-2)$$

What complicates this expansion is the fact that $\mathcal{Z}(k)$ and $z(k+1)$ are statistically dependent. Measurement vector $\mathcal{Z}(k)$ depends on state vectors $x(1), x(2), \dots, x(k)$, because $z(j) = H(j)x(j) + v(j)$ ($j = 1, 2, \dots, k$). Measurement vector $z(k+1)$ also depends on state vector $x(k)$, because $z(k+1) = H(k+1)x(k+1) + v(k+1)$ and $x(k+1) = \Phi(k+1,k)x(k) + \Gamma(k+1,k)w(k) + \Psi(k+1,k)u(k)$. Hence $\mathcal{Z}(k)$ and $z(k+1)$ both depend on $x(k)$ and are, therefore, dependent.

Recall that $\mathbf{x}(k+1)$, $\mathbf{z}(k)$ and $\mathbf{z}(k+1)$ are jointly Gaussian random vectors; hence, we can use Theorem 12-4 to express (17-2) as

$$\hat{\mathbf{x}}(k+1|k+1) = \mathbf{E}\{\mathbf{x}(k+1)|\mathbf{z}(k), \tilde{\mathbf{z}}\} \quad (17-3)$$

where

$$\tilde{\mathbf{z}} = \mathbf{z}(k+1) - \mathbf{E}\{\mathbf{z}(k+1)|\mathbf{z}(k)\} \quad (17-4)$$

We immediately recognize $\tilde{\mathbf{z}}$ as the innovations process $\tilde{\mathbf{z}}(k+1|k)$ [see (16-29)]; thus, we rewrite (17-3) as

$$\hat{\mathbf{x}}(k+1|k+1) = \mathbf{E}\{\mathbf{x}(k+1)|\mathbf{z}(k), \tilde{\mathbf{z}}(k+1|k)\} \quad (17-5)$$

Applying (12-37) to (17-5), we find that

$$\begin{aligned} \hat{\mathbf{x}}(k+1|k+1) &= \mathbf{E}\{\mathbf{x}(k+1)|\mathbf{z}(k)\} \\ &\quad + \mathbf{E}\{\mathbf{x}(k+1)|\tilde{\mathbf{z}}(k+1|k)\} - \mathbf{m}_x(k+1) \end{aligned} \quad (17-6)$$

We recognize the first term on the right-hand side of (17-6) as the single-stage predicted estimator of $\mathbf{x}(k+1)$, $\hat{\mathbf{x}}(k+1|k)$; hence,

$$\begin{aligned} \hat{\mathbf{x}}(k+1|k+1) &= \hat{\mathbf{x}}(k+1|k) \\ &\quad + \mathbf{E}\{\mathbf{x}(k+1)|\tilde{\mathbf{z}}(k+1|k)\} - \mathbf{m}_x(k+1) \end{aligned} \quad (17-7)$$

This equation is the starting point for our derivation of the Kalman filter.

Before proceeding further, we observe, upon comparison of (17-2) and (17-5), that our original conditioning on $\mathbf{z}(k+1)$ has been replaced by conditioning on the innovations process $\tilde{\mathbf{z}}(k+1|k)$. One can show that $\tilde{\mathbf{z}}(k+1|k)$ is computable from $\mathbf{z}(k+1)$, and that $\mathbf{z}(k+1)$ is computable from $\tilde{\mathbf{z}}(k+1|k)$; hence, it is said that $\mathbf{z}(k+1)$ and $\tilde{\mathbf{z}}(k+1|k)$ are *causally invertible* (Anderson and Moore, 1979). We explain this statement more carefully at the end of this lesson.

A PRELIMINARY RESULT

In our derivation of the Kalman filter, we shall determine that

$$\hat{\mathbf{x}}(k+1|k+1) = \hat{\mathbf{x}}(k+1|k) + \mathbf{K}(k+1)\tilde{\mathbf{z}}(k+1|k) \quad (17-8)$$

where $\mathbf{K}(k+1)$ is an $n \times m$ (Kalman) gain matrix. We will calculate the optimal gain matrix in the next section.

Here let us view (17-8) as the structure of an arbitrary recursive linear filter, which is written in so-called *predictor-corrector* format; i.e., the filtered estimate of $\mathbf{x}(k+1)$ is obtained by a predictor step, $\hat{\mathbf{x}}(k+1|k)$, and a corrector step, $\mathbf{K}(k+1)\tilde{\mathbf{z}}(k+1|k)$. The predictor step uses information from the state equation, because $\hat{\mathbf{x}}(k+1|k) = \Phi(k+1,k)\hat{\mathbf{x}}(k|k) + \Psi(k+1,k)\mathbf{u}(k)$. The corrector step uses the new measurement available at t_{k+1} . The correction is proportional to the difference between that measurement and its best pre-

dicted value, $\hat{\mathbf{z}}(k+1|k)$. The following result provides us with the means for evaluating $\hat{\mathbf{x}}(k+1|k+1)$ in terms of its error-covariance matrix $\mathbf{P}(k+1|k+1)$.

Preliminary Result. *Filtering error-covariance matrix $\mathbf{P}(k+1|k+1)$ for the arbitrary linear recursive filter (17-8) is computed from the following equation:*

$$\begin{aligned} \mathbf{P}(k+1|k+1) &= [\mathbf{I} - \mathbf{K}(k+1)\mathbf{H}(k+1)]\mathbf{P}(k+1|k)[\mathbf{I} - \mathbf{K}(k+1)\mathbf{H}(k+1)]' \\ &\quad + \mathbf{K}(k+1)\mathbf{R}(k+1)\mathbf{K}'(k+1) \end{aligned} \quad (17-9)$$

Proof. Substitute (16-32) into (17-8) and then subtract the resulting equation from $\mathbf{x}(k+1)$ in order to obtain

$$\begin{aligned} \tilde{\mathbf{x}}(k+1|k+1) &= [\mathbf{I} - \mathbf{K}(k+1)\mathbf{H}(k+1)]\tilde{\mathbf{x}}(k+1|k) \\ &\quad - \mathbf{K}(k+1)\mathbf{v}(k+1) \end{aligned} \quad (17-10)$$

Substitute this equation into $\mathbf{P}(k+1|k+1) = \mathbf{E}\{\tilde{\mathbf{x}}(k+1|k+1)\tilde{\mathbf{x}}'(k+1|k+1)\}$ to obtain equation (17-9). As in the proof of Theorem 16-2, we have used the fact that $\tilde{\mathbf{x}}(k+1|k)$ and $\mathbf{v}(k+1)$ are independent to show that $\mathbf{E}\{\tilde{\mathbf{x}}(k+1|k)\mathbf{v}'(k+1)\} = 0$. \square

The state prediction-error covariance matrix $\mathbf{P}(k+1|k)$ is given by equation (16-11). Observe that (17-9) and (16-11) can be computed recursively, once gain matrix $\mathbf{K}(k+1)$ is specified, as follows: $\mathbf{P}(0|0) \rightarrow \mathbf{P}(1|0) \rightarrow \mathbf{P}(1|1) \rightarrow \mathbf{P}(2|1) \rightarrow \mathbf{P}(2|2) \rightarrow \dots$ etc.

It is important to reiterate the fact that (17-9) is true for *any* gain matrix, including the optimal gain matrix given next in Theorem 17-1.

THE KALMAN FILTER

Theorem 17-1

- a. *The mean-squared filtered estimator of $\mathbf{x}(k+1)$, $\hat{\mathbf{x}}(k+1|k+1)$, written in predictor-corrector format, is*

$$\hat{\mathbf{x}}(k+1|k+1) = \hat{\mathbf{x}}(k+1|k) + \mathbf{K}(k+1)\tilde{\mathbf{z}}(k+1|k) \quad (17-11)$$

for $k = 0, 1, \dots$, where $\hat{\mathbf{x}}(0|0) = \mathbf{m}_x(0)$, and $\tilde{\mathbf{z}}(k+1|k)$ is the innovations process [$\tilde{\mathbf{z}}(k+1|k) = \mathbf{z}(k+1) - \mathbf{H}(k+1)\hat{\mathbf{x}}(k+1|k)$].

- b. *$\mathbf{K}(k+1)$ is an $n \times m$ matrix (commonly referred to as the Kalman gain matrix, or weighting matrix) which is specified by the set of relations*

$$\begin{aligned} \mathbf{K}(k+1) &= \mathbf{P}(k+1|k)\mathbf{H}'(k+1)[\mathbf{H}(k+1)\mathbf{P}(k+1|k)\mathbf{H}'(k+1) \\ &\quad + \mathbf{R}(k+1)]^{-1} \end{aligned} \quad (17-12)$$

$$\begin{aligned} \mathbf{P}(k+1|k) &= \Phi(k+1,k)\mathbf{P}(k|k)\Phi'(k+1,k) \\ &\quad + \Gamma(k+1,k)\mathbf{Q}(k)\Gamma'(k+1,k) \end{aligned} \quad (17-13)$$

and

$$\mathbf{P}(k+1|k+1) = [\mathbf{I} - \mathbf{K}(k+1)\mathbf{H}(k+1)]\mathbf{P}(k+1|k) \quad (17-14)$$

for $k = 0, 1, \dots$, where \mathbf{I} is the $n \times n$ identity matrix, and $\mathbf{P}(0|0) = \mathbf{P}_x(0)$.

- c. The stochastic process $\{\hat{\mathbf{x}}(k+1|k+1), k = 0, 1, \dots\}$, which is defined by the filtering error relation

$$\tilde{\mathbf{x}}(k+1|k+1) = \mathbf{x}(k+1) - \hat{\mathbf{x}}(k+1|k+1) \quad (17-15)$$

$k = 0, 1, \dots$, is a zero-mean Gauss-Markov sequence whose covariance matrix is given by (17-14).

Proof (Mendel, 1983b, pp. 56–57)

- a. We begin with the formula for $\hat{\mathbf{x}}(k+1|k+1)$ in (17-7). Recall that $\mathbf{x}(k+1)$ and $\mathbf{z}(k+1)$ are jointly Gaussian. Because $\mathbf{z}(k+1)$ and $\tilde{\mathbf{z}}(k+1|k)$ are causally invertible, $\mathbf{x}(k+1)$ and $\tilde{\mathbf{z}}(k+1|k)$ are also jointly Gaussian. Additionally, $\mathbf{E}\{\tilde{\mathbf{z}}(k+1|k)\} = \mathbf{0}$; hence,

$$\begin{aligned} \mathbf{E}\{\mathbf{x}(k+1)|\tilde{\mathbf{z}}(k+1|k)\} &= \mathbf{m}_x(k+1) \\ &+ \mathbf{P}_{xz}(k+1,k+1|k)\mathbf{P}_{zz}^{-1}(k+1|k)\tilde{\mathbf{z}}(k+1|k) \end{aligned} \quad (17-16)$$

We define gain matrix $\mathbf{K}(k+1)$ as

$$\mathbf{K}(k+1) = \mathbf{P}_{xz}(k+1,k+1|k)\mathbf{P}_{zz}^{-1}(k+1|k) \quad (17-17)$$

Substituting (17-16) and (17-17) into (17-7) we obtain the Kalman filter equation (17-11). Because $\hat{\mathbf{x}}(k+1|k) = \Phi(k+1,k)\hat{\mathbf{x}}(k|k) + \Psi(k+1,k)\mathbf{u}(k)$, equation (17-11) must be initialized by $\hat{\mathbf{x}}(0|0)$, which we have shown must equal $\mathbf{m}_x(0)$ [see Equation (16-12)].

- b. In order to evaluate $\mathbf{K}(k+1)$ we must evaluate \mathbf{P}_{xz} and \mathbf{P}_{zz}^{-1} . Matrix \mathbf{P}_{zz} has been computed in (16-33). By definition of cross-covariance,

$$\begin{aligned} \mathbf{P}_{xz} &= \mathbf{E}\{[\mathbf{x}(k+1) - \mathbf{m}_x(k+1)]\tilde{\mathbf{z}}'(k+1|k)\} \\ &= \mathbf{E}\{\mathbf{x}(k+1)\tilde{\mathbf{z}}'(k+1|k)\} \end{aligned} \quad (17-18)$$

because $\tilde{\mathbf{z}}(k+1|k)$ is zero-mean. Substituting (16-32) into this expression, we find that

$$\mathbf{P}_{xz} = \mathbf{E}\{\mathbf{x}(k+1)\tilde{\mathbf{x}}'(k+1|k)\}\mathbf{H}'(k+1) \quad (17-19)$$

because $\mathbf{E}\{\mathbf{x}(k+1)\mathbf{v}'(k+1)\} = \mathbf{0}$. Finally, expressing $\mathbf{x}(k+1)$ as $\tilde{\mathbf{x}}(k+1|k) + \hat{\mathbf{x}}(k+1|k)$ and applying the orthogonality principle (13-15), we find that

$$\mathbf{P}_{xz} = \mathbf{P}(k+1|k)\mathbf{H}'(k+1) \quad (17-20)$$

Combining equations (17-20) and (16-33) into (17-17), we obtain equation (17-12) for the Kalman gain matrix.

State prediction-error covariance matrix $\mathbf{P}(k+1|k)$ was derived in Lesson 16.

State filtering-error covariance matrix $\mathbf{P}(k+1|k+1)$ is obtained by substituting (17-12) for $\mathbf{K}(k+1)$ into (17-9), as follows:

$$\begin{aligned} \mathbf{P}(k+1|k+1) &= (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}(\mathbf{I} - \mathbf{K}\mathbf{H})' + \mathbf{K}\mathbf{R}\mathbf{K}' \\ &= (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P} - \mathbf{P}\mathbf{H}'\mathbf{K}' + \mathbf{K}\mathbf{H}\mathbf{P}'\mathbf{K}' + \mathbf{K}\mathbf{R}\mathbf{K}' \\ &= (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P} - \mathbf{P}\mathbf{H}'\mathbf{K}' + \mathbf{K}(\mathbf{H}\mathbf{P}' + \mathbf{R})\mathbf{K}' \\ &= (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P} - \mathbf{P}\mathbf{H}'\mathbf{K}' + \mathbf{P}\mathbf{H}'\mathbf{K}' \\ &= (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P} \end{aligned} \quad (17-21)$$

- c. The proof that $\tilde{\mathbf{x}}(k+1|k+1)$ is zero-mean, Gaussian and Markov is so similar to the proof of part b of Theorem 16-1 that we omit its details [see Meditch (1969, pp. 181–182)]. \square

OBSERVATIONS ABOUT THE KALMAN FILTER

1. Figure 17-1 depicts the interconnection of our basic dynamical system [equations (15-17) and (15-18)] and Kalman filter system. The feedback nature of the Kalman filter is quite evident. Observe, also, that the Kalman filter contains within its structure a model of the plant.

The feedback nature of the Kalman filter manifests itself in two different ways, namely in the calculation of $\hat{\mathbf{x}}(k+1|k+1)$ and also in the calculation of the matrix of gains, $\mathbf{K}(k+1)$, both of which we shall explore below.

2. The predictor-corrector form of the Kalman filter is illuminating from an information-usage viewpoint. Observe that the predictor equations, which compute $\hat{\mathbf{x}}(k+1|k)$ and $\mathbf{P}(k+1|k)$, use information only from the state equation, whereas the corrector equations, which compute $\mathbf{K}(k+1)$, $\hat{\mathbf{x}}(k+1|k+1)$ and $\mathbf{P}(k+1|k+1)$, use information only from the measurement equation.
3. Once the gain matrix is computed, then (17-11) represents a *time-varying recursive digital filter*. This is seen more clearly when equations (16-4) and (16-31) are substituted into (17-11). The resulting equation can be rewritten as

$$\begin{aligned} \hat{\mathbf{x}}(k+1|k+1) &= [\mathbf{I} - \mathbf{K}(k+1)\mathbf{H}(k+1)]\Phi(k+1,k)\hat{\mathbf{x}}(k|k) \\ &+ \mathbf{K}(k+1)\mathbf{z}(k+1) \\ &+ [\mathbf{I} - \mathbf{K}(k+1)\mathbf{H}(k+1)]\Psi(k+1,k)\mathbf{u}(k) \end{aligned} \quad (17-22)$$

for $k = 0, 1, \dots$. This is a state equation for state vector $\hat{\mathbf{x}}$, whose time-varying plant matrix is $[\mathbf{I} - \mathbf{K}(k+1)\mathbf{H}(k+1)]\Phi(k+1,k)$. Equation (17-22) is time-varying even if our dynamical system in equations (15-17) and (15-18) is time-invariant and stationary, because gain matrix $\mathbf{K}(k+1)$ is still time-varying in that case. It is possible,

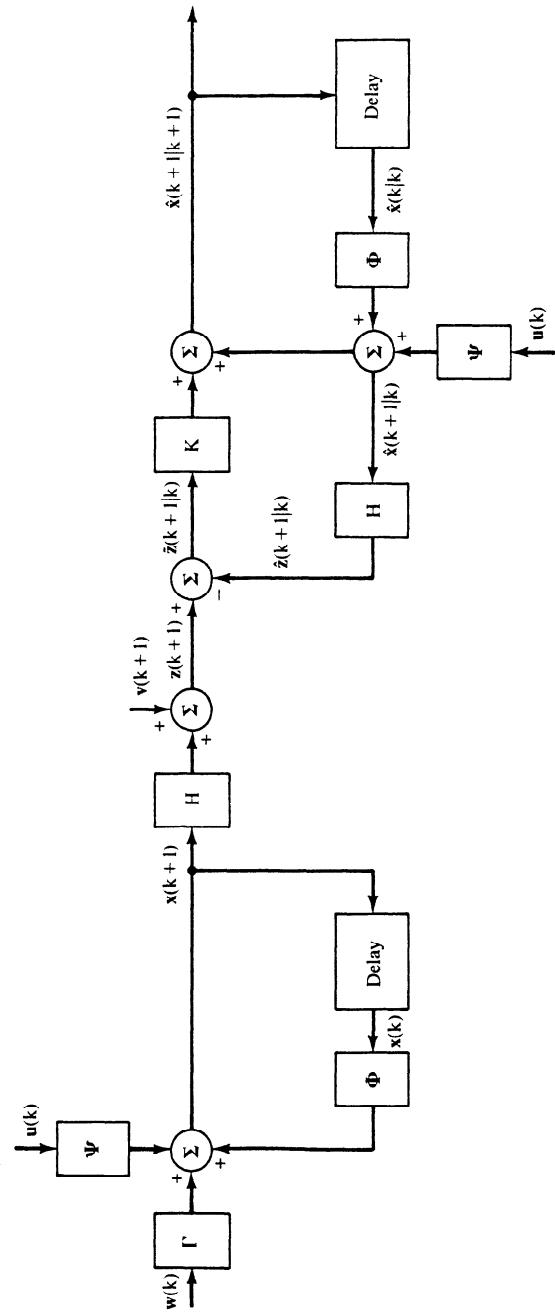


Figure 17-1 Interconnection of system and Kalman filter. (Mendel, 1983b, © 1983, Academic Press, Inc.).

Observations About the Kalman Filter

155

however, for $K(k+1)$ to reach a limiting value (i.e., a steady-state value, \bar{K}), in which case (17-22) reduces to a recursive constant coefficient filter. We will have more to say about this important steady-state case in Lesson 19.

Equation (17-22) is in a *recursive filter form*, in that it relates the filtered estimate of $x(k+1)$, $\hat{x}(k+1|k+1)$, to the filtered estimate of $x(k)$, $\hat{x}(k|k)$. Using substitutions similar to those used in the derivation of (17-22), one can also obtain the following *recursive predictor form* of the Kalman filter (details left as an exercise),

$$\begin{aligned}\hat{x}(k+1|k) = & \Phi(k+1,k)[I - K(k)H(k)]\hat{x}(k|k-1) \\ & + \Phi(k+1,k)K(k)z(k) + \Psi(k+1,k)u(k)\end{aligned}\quad (17-23)$$

Observe that in (17-23) the predicted estimate of $x(k+1)$, $\hat{x}(k+1|k)$, is related to the predicted estimate of $x(k)$, $\hat{x}(k|k-1)$. Interestingly enough, the recursive predictor (17-23), and not the recursive filter (17-22), plays an important role in mean-squared smoothing, as we shall see in Lesson 21.

The structures of (17-22) and (17-23) are summarized in Figure 17-2. This figure supports the claim made in Lesson 1 that our recursive estimators are nothing more than time-varying digital filters that operate on random (and also deterministic) inputs.

4. Embedded within the recursive Kalman filter equations is another set of recursive equations—(17-12), (17-13) and (17-14). Because $P(0|0)$ initializes these calculations, these equations must be ordered as follows: $P(k|k) \rightarrow P(k+1|k) \rightarrow K(k+1) \rightarrow P(k+1|k+1) \rightarrow \text{etc.}$

By combining these three equations it is possible to get a matrix recursive equation for $P(k+1|k)$ as a function of $P(k|k-1)$, or a similar equation for $P(k+1|k+1)$ as a function of $P(k|k)$. These equations are nonlinear and are known as *matrix Riccati equations*. For example, the matrix Riccati equation for $P(k+1|k)$ is

$$\begin{aligned}P(k+1|k) = & \Phi P(k|k-1) \{ I - H' [H P(k|k-1) H' + R]^{-1} \\ & H P(k|k-1) \} \Phi' + \Gamma Q \Gamma'\end{aligned}\quad (17-24)$$

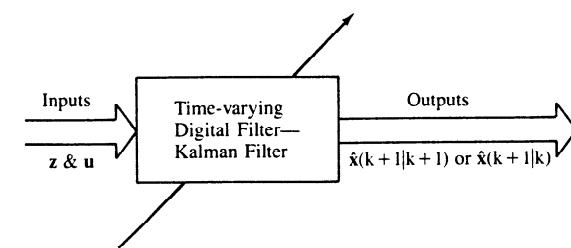


Figure 17-2 Input-output interpretation of the Kalman filter.

where we have omitted the temporal arguments on Φ , Γ , \mathbf{H} , \mathbf{Q} and \mathbf{R} for notational simplicity [their correct arguments are $\Phi(k+1|k)$, $\Gamma(k+1|k)$, $\mathbf{H}(k)$, $\mathbf{Q}(k)$, and $\mathbf{R}(k)$]. The matrix Riccati equation for $\mathbf{P}(k+1|k+1)$ is obtained by substituting (17-12) into (17-14), and then (17-13) into the resulting equation. We leave its derivation as an exercise.

5. A measure of recursive predictor performance is provided by matrix $\mathbf{P}(k+1|k)$. This covariance matrix can be calculated prior to any processing of real data, using its matrix Riccati Equation (17-24) or Equations (17-13), (17-14), and (17-12). A measure of recursive filter performance is provided by matrix $\mathbf{P}(k+1|k+1)$, and this covariance matrix can also be calculated prior to any processing of real data. Note that $\mathbf{P}(k+1|k+1) \neq \mathbf{P}(k+1|k)$. These calculations are often referred to as *performance analyses*. It is indeed interesting that the Kalman filter utilizes a measure of its mean-squared error during its real-time operation.
6. Two formulas are available for computing $\mathbf{P}(k+1|k+1)$, namely (17-14), which is known as the *standard form*, and (17-9), which is known as the *stabilized form*. Although the stabilized form requires more computations than the standard form, it is much less sensitive to numerical errors from the prior calculation of gain matrix $\mathbf{K}(k+1)$ than is the standard form. In fact, one can show that first-order errors in the calculation of $\mathbf{K}(k+1)$ propagate as first-order errors in the calculation of $\mathbf{P}(k+1|k+1)$, when the standard form is used, but only as second-order errors in the calculation of $\mathbf{P}(k+1|k+1)$ when the stabilized form is used. This is why (17-9) is called the stabilized form [for detailed derivations, see Aoki (1967), Jazwinski (1970), or Mendel (1973)].
7. On the subject of computation, the calculation of $\mathbf{P}(k+1|k)$ is the most costly one for the Kalman filter, because of the term $\Phi(k+1|k)\mathbf{P}(k|k)\Phi'(k+1|k)$, which entails two multiplications of two $n \times n$ matrices [i.e., $\mathbf{P}(k|k)\Phi'(k+1|k)$ and $\Phi(k+1|k)(\mathbf{P}(k|k)\Phi'(k+1|k))$]. Total computation for the two matrix multiplications is on the order of $2n^3$ multiplications and $2n^3$ additions [for more detailed computation counts, including storage requirements, for all of the Kalman filter equations, see Mendel (1971), Gura and Bierman (1971), and Bierman (1973a)].

One must be very careful to code the standard or stabilized forms for $\mathbf{P}(k+1|k+1)$ so that $n \times n$ matrices are never multiplied. We leave it to the reader to show that the standard algorithm can be coded in such a manner that it only requires on the order of $\frac{1}{2}mn^2$ multiplications, whereas the stabilized algorithm can be coded in such a manner that it only requires on the order of $\frac{5}{2}mn^2$ multiplications. In many applications, system order n is larger than number of measurements m , so that $mn^2 \ll n^3$. Usually, computation is most sensitive to system order; so, whenever possible, use low-order (but adequate) models.

8. Because of the equivalence between mean-squared, best-linear unbiased, and weighted-least squares filtered estimates of our state vector $\mathbf{x}(k)$ (see Lesson 14), we must realize that our Kalman filter equations are just a recursive solution to a system of normal equations (see Lesson 3). Other implementations of the Kalman filter that solve the normal equations using stable algorithms from numerical linear algebra (see, e.g., Bierman, 1977) and involve orthogonal transformations, have better numerical properties than (17-11)–(17-14). We reiterate, however, that because this is a book on estimation *theory*, theoretical formulas such as those in (17-11)–(17-14) are appropriate.
9. In Lesson 4 we developed two forms for a recursive least-squares estimator, namely, the covariance and information forms. Compare $\mathbf{K}(k+1)$ in (17-12) with $\mathbf{K}_w(k+1)$ in (4-25) to see they have the same structure; hence, our formulation of the Kalman filter is often known as the *covariance formulation*. We leave it to the reader to show that $\mathbf{K}(k+1)$ can also be computed as

$$\mathbf{K}(k+1) = \mathbf{P}(k+1|k+1)\mathbf{H}'(k+1)\mathbf{R}^{-1}(k+1) \quad (17-25)$$

where $\mathbf{P}(k+1|k+1)$ is computed as

$$\begin{aligned} \mathbf{P}^{-1}(k+1|k+1) &= \mathbf{P}^{-1}(k+1|k) \\ &\quad + \mathbf{H}'(k+1)\mathbf{R}^{-1}(k+1)\mathbf{H}(k+1) \end{aligned} \quad (17-26)$$

When these equations are used along with (17-11) and (17-13) we have the *information formulation* of the Kalman filter. Of course, the ordering of the computations in these two formulations of the Kalman filter are different. See Lessons 4 and 9 for related discussions.

10. In Lesson 14 we showed that $\hat{\mathbf{x}}_{\text{MAP}}(k_1|N) = \hat{\mathbf{x}}_{\text{MS}}(k_1|N)$; hence,

$$\hat{\mathbf{x}}_{\text{MAP}}(k|k) = \hat{\mathbf{x}}_{\text{MS}}(k|k) \quad (17-27)$$

This means that the Kalman filter also gives MAP estimates of the state vector $\mathbf{x}(k)$ for our basic state-variable model.

11. At the end of the introduction section in this lesson we mentioned that $\mathbf{z}(k+1)$ and $\bar{\mathbf{z}}(k+1|k)$ are *causally invertible*. This means that we can compute one from the other using a causal (i.e., realizable) system. For example, when the measurements are available, then $\bar{\mathbf{z}}(k+1|k)$ can be obtained from Equations (17-23) and (16-31), which we repeat here for the convenience of the reader:

$$\begin{aligned} \hat{\mathbf{x}}(k+1|k) &= \Phi(k+1,k)[\mathbf{I} - \mathbf{K}(k)\mathbf{H}(k)]\hat{\mathbf{x}}(k|k-1) \\ &\quad + \Psi(k+1,k)\mathbf{u}(k) + \Phi(k+1,k)\mathbf{K}(k)\mathbf{z}(k) \end{aligned} \quad (17-28)$$

and

$$\bar{\mathbf{z}}(k+1|k) = -\mathbf{H}(k+1)\hat{\mathbf{x}}(k+1|k) + \mathbf{z}(k+1) \quad k = 0, 1, \dots \quad (17-29)$$

We refer to (17-28) and (17-29) in the rest of this book as the *Kalman innovations system*. It is initialized by $\hat{x}(0|0) = 0$.

On the other hand, if the innovations are given a priori, then $z(k+1)$ can be obtained from

$$\begin{aligned}\hat{x}(k+1|k) &= \Phi(k+1,k)\hat{x}(k|k-1) + \Psi(k+1,k)u(k) \\ &\quad + \Phi(k+1,k)K(k)\bar{z}(k|k-1)\end{aligned}\quad (17-30)$$

and

$$\begin{aligned}z(k+1) &= H(k+1)\hat{x}(k+1|k) + \bar{z}(k+1|k) \\ k &= 0, 1, \dots\end{aligned}\quad (17-31)$$

Equation (17-30) was obtained by rearranging the terms in (17-28), and (17-31) was obtained by solving (17-29) for $z(k+1)$. Equation (17-30) is again initialized by $\hat{x}(0|0) = 0$.

Note that (17-30) and (17-31) are equivalent to our basic state-variable model in (15-17) and (15-18), from an input-output point of view. Consequently, model (17-30) and (17-31) is often the starting point for important problems such as the stochastic realization problem (e.g., Faurre, 1976).

PROBLEMS

- 17-1. Prove that $\hat{x}(k+1|k+1)$ is zero mean, Gaussian and first-order Markov.
- 17-2. Derive the recursive predictor form of the Kalman filter, given in (17-23).
- 17-3. Derive the matrix Riccati equation for $P(k+1|k+1)$.
- 17-4. Show that gain matrix $K(k+1)$ can also be computed using (17-25).
- 17-5. Suppose a small error δK is made in the computation of the Kalman filter gain $K(k+1)$.
 - (a) Show that when $P(k+1|k+1)$ is computed from the “standard form” equation, then to first-order terms,
$$\delta P(k+1|k+1) = -\delta K(k+1)H(k+1)P(k+1|k)$$
 - (b) Show that when $P(k+1|k+1)$ is computed from the “stabilized form” equation, then to first-order terms
$$\delta P(k+1|k+1) \approx 0$$
- 17-6. Consider the basic scalar system $x(k+1) = \phi x(k) + w(k)$ and $z(k+1) = x(k+1) + v(k+1)$.
 - (a) Show that $p(k+1|k) \geq q$, which means that the variance of the system disturbance sets the performance limit on prediction accuracy.
 - (b) Show that $0 \leq K(k+1) \leq 1$.
 - (c) Show that $0 \leq p(k+1|k+1) \leq r$.
- 17-7. An RC filter with time constant τ is excited by Gaussian white noise, and the output is measured every T seconds. The output at the sample times obeys the

equation $x(k) = e^{-T/\tau}x(k-1) + w(k-1)$, where $E\{x(0)\} = 1$, $E\{x^2(0)\} = 2$, $q = 2$, and $T = \tau = 0.1$ sec. The measurements are described by $z(k) = x(k) + v(k)$, where $v(k)$ is a white but *non-Gaussian* noise sequence for which $E\{v(k)\} = 0$ and $E\{v^2(k)\} = 4$. Additionally, $w(k)$ and $v(k)$ are uncorrelated. Measurements $z(1) = 1.5$ and $z(2) = 3.0$.

- (a) Find the best linear estimate of $x(1)$ based on $z(1)$.
- (b) Find the best linear estimate of $x(2)$ based on $z(1)$ and $z(2)$.

- 17-8. Table 17-1 lists multiplications and additions for all the basic matrix operations used in a Kalman filter. Using the formulas for the Kalman filter given in Theorem 17-1, establish the number of multiplications and additions required to compute $\hat{x}(k+1|k)$, $P(k+1|k)$, $K(k+1)$, $\hat{x}(k+1|k+1)$, and $P(k+1|k+1)$.

TABLE 17-1 Operation Characteristics

Name	Function	Multiplications	Additions
Matrix addition	$C_{MN} = A_{MN} + B_{MN}$	—	MN
Matrix subtraction	$C_{MN} = A_{MN} - B_{MN}$	—	MN
Matrix multiply	$C_{ML} = A_{MN}B_{NL}$	MNL	$ML(N-1)$
Matrix transpose			
multiply	$C_{MN} = A_{ML}(B'_{NL})$	MNL	$ML(N-1)$
Matrix inversion	$A_{NN} \rightarrow A_{NN}^{-1}$	αN^3	βN^3
Scalar-vector product	$C_{NI} = pA_{NI}$	N	—

- 17-9. Show that the standard algorithm for computing $P(k+1|k+1)$ only requires on the order of $\frac{1}{2}mn^2$ multiplications, whereas the stabilized algorithm requires on the order of $\frac{5}{2}mn^2$ multiplications (use Table 17-1). Note that this last result requires a very clever coding of the stabilized algorithm.

Lesson 18

State Estimation: Filtering Examples

INTRODUCTION

In this lesson (which is an excellent one for self-study) we present five examples, which illustrate some interesting numerical and theoretical aspects of Kalman filtering.

EXAMPLES

Example 18-1

In Lesson 17 we learned that the Kalman filter is a dynamical feedback system. Its gain matrix and predicted- and filtering-error covariance matrices comprise a matrix feedback system operating within the Kalman filter. Of course, these matrices can be calculated prior to processing of data, and such calculations constitute a *performance analysis* of the Kalman filter. Here we examine the results of these calculations for two second-order systems, $H_1(z) = 1/(z^2 - 1.32z + 0.875)$ and $H_2(z) = 1/(z^2 - 1.36z + 0.923)$. The second system is less damped than the first. Impulse responses of both systems are depicted in Figure 18-1.

In Figure 18-1 we also depict $p_{11}(k|k)$, $p_{22}(k|k)$, $K_1(k)$ and $K_2(k)$ versus k for both systems. In both cases $P(0|0)$ was set equal to the zero matrix. For system 1, $q = 1$ and $r = 5$, whereas for system 2 $q = 1$ and $r = 20$. Observe that the error variances and Kalman gains exhibit a transient response as well as a steady-state response; i.e., after a certain value of k ($k \approx 10$ for system-1 and $k \approx 15$ for system-2) $p_{11}(k|k)$, $p_{22}(k|k)$, $K_1(k)$, and $K_2(k)$ reaching limiting values. These limiting values do not depend on $P(0|0)$, as can be seen from Figure 18-2. The Kalman filter is initially influenced by its initial conditions, but eventually ignores them, paying much greater attention to model

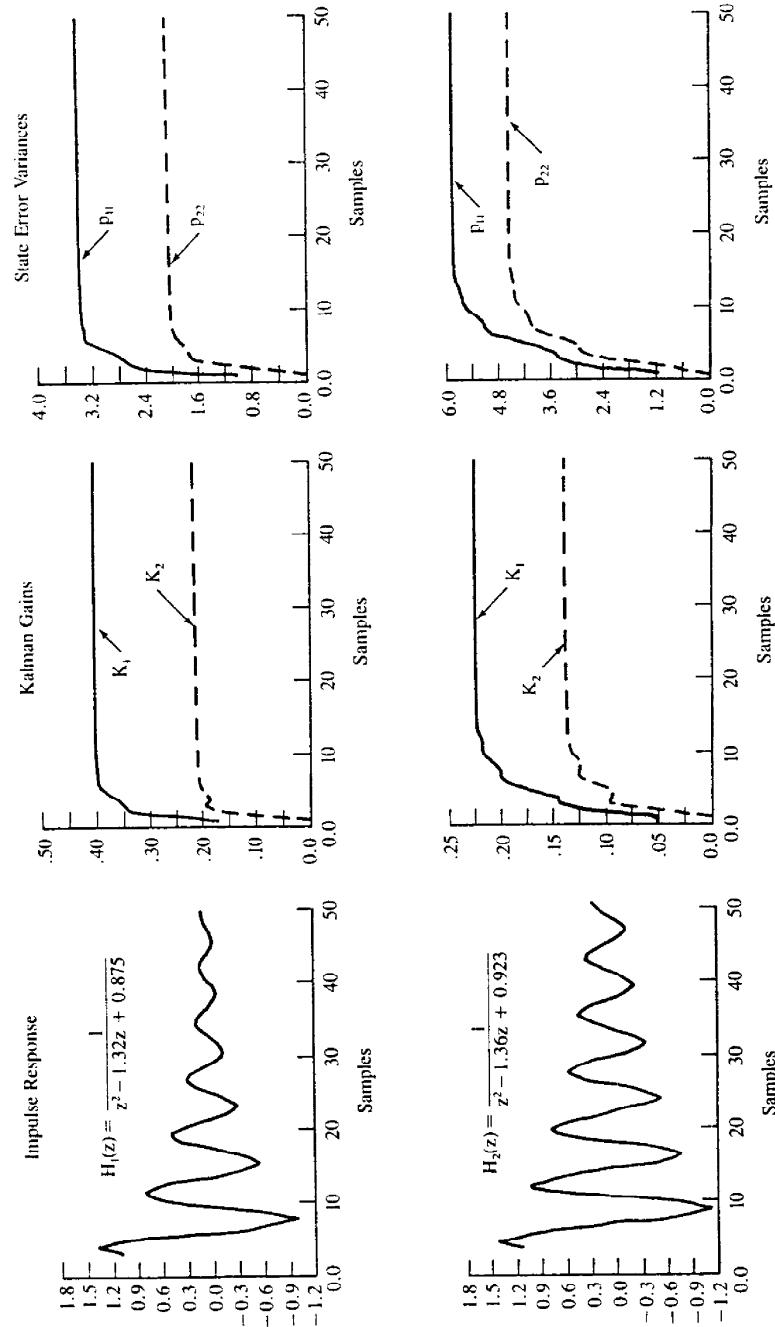


Figure 18-1 Kalman gains and filtering error variances for two second-order systems.

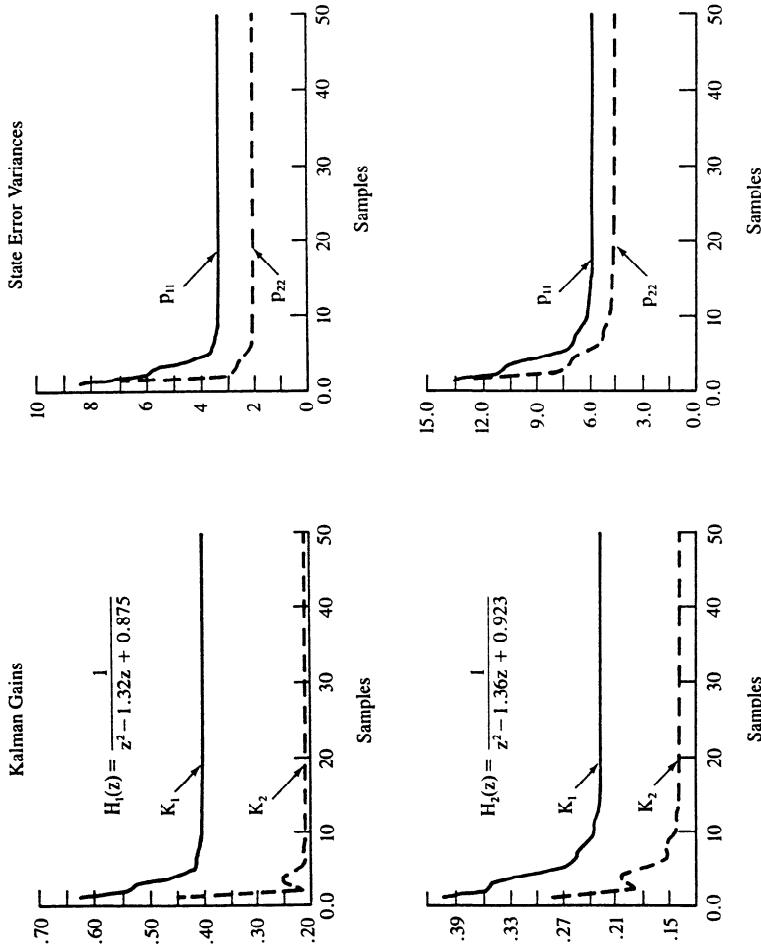


Figure 18-2 Same as Figure 18-1, except for a different choice of $P(0|0)$.

Examples

163

parameters and the measurements. The relatively large steady-state values for p_{11} and p_{22} are due to the large values of r . \square

Example 18-2

A state estimate is implicitly conditioned on knowing the true values for all system parameters (i.e., Φ , Γ , Ψ , H , Q , and R). Sometimes we do not know these values exactly; hence, it is important to learn how *sensitive* (i.e., robust) the Kalman filter is to parameter errors. Many references which treat Kalman filter sensitivity issues can be found in Mendel and Giesecking (1971) under category 2f, "State Estimation: Sensitivity Considerations."

Let θ denote any parameter that may appear in Φ , Γ , Ψ , H , Q , or R . In order to determine the sensitivity of the Kalman filter to *small* variations in θ one computes $\partial\hat{x}(k+1|k+1)/\partial\theta$, whereas, for large variations in θ one computes $\Delta\hat{x}(k+1|k+1)/\Delta\theta$. An analysis of $\Delta\hat{x}(k+1|k+1)/\Delta\theta$, for example, reveals the interesting chain of events, that: $\Delta\hat{x}(k+1|k+1)/\Delta\theta$ depends on $\Delta K(k+1)/\Delta\theta$, which in turn depends on $\Delta P(k+1|k)/\Delta\theta$, which in turn depends on $\Delta P(k|k)/\Delta\theta$. Hence, for each variable parameter, θ , we have a *Kalman filter sensitivity system* comprised of equations from which we compute $\Delta\hat{x}(k+1|k+1)/\Delta\theta$ (see also, Lesson 26). An alternative to using these equations is to perform a computer perturbation study. For example,

$$\frac{\Delta K(k+1)}{\Delta\theta} = \frac{K(k+1)|_{\theta=\theta_N+\Delta\theta} - K(k+1)|_{\theta=\theta_N}}{\Delta\theta} \quad (18-1)$$

where θ_N denotes a nominal value of θ .

We define *sensitivity* as the ratio of the percentage change in a function [e.g., $K(k+1)$] to the percentage change in a parameter, θ . For example,

$$S_{\theta}^{K_{ij}(k+1)} = \frac{\frac{\Delta K_{ij}(k+1)}{\Delta\theta}}{\frac{K_{ij}(k+1)}{\theta}} \quad (18-2)$$

denotes the sensitivity of the ij th element of matrix $K(k+1)$ with respect to parameter θ . All other sensitivities, such as $S_{\theta}^{P_{ij}(k|k)}$, are defined similarly.

Here we present some numerical sensitivity results for the simple first-order system

$$x(k+1) = a x(k) + b w(k) \quad (18-3)$$

$$z(k) = h x(k) + n(k) \quad (18-4)$$

where $a_N = 0.7$, $b_N = 1.0$, $h_N = 0.5$, $q_N = 0.2$, and $r_N = 0.1$. Figure 18-3 depicts $S_a^{K(k+1)}$, $S_b^{K(k+1)}$, $S_h^{K(k+1)}$, $S_q^{K(k+1)}$, and $S_r^{K(k+1)}$ for parameter variations of $\pm 5\%$, $\pm 10\%$, $\pm 20\%$, and $\pm 50\%$ about nominal values, a_N , b_N , h_N , q_N , and r_N .

Observe that the sensitivity functions vary with time and that they all reach steady-state values. Table 18-1 summarizes the steady-state sensitivity coefficients $\bar{S}_\theta^{K(k+1)}$, $\bar{S}_\theta^{P(k|k)}$ and $\bar{S}_\theta^{P(k+1|k)}$.

Some conclusions that can be drawn from these numerical results are: (1) $K(k+1)$, $P(k|k)$ and $P(k+1|k)$ are most sensitive to changes in parameter b , and

(2) $\bar{S}_{\theta}^{K(k+1)} = \bar{S}_{\theta}^{P(k|k)}$ for $\theta = a, b$, and q . This last observation could have been foreseen, because of our alternate equation for $K(k+1)$ [Equation (17-25)].

$$K(k+1) = P(k+1|k+1)H(k+1)R^{-1}(k+1) \quad (18-5)$$

This expression shows that if h and r are fixed then $K(k+1)$ varies exactly the same way as $P(k+1|k+1)$.

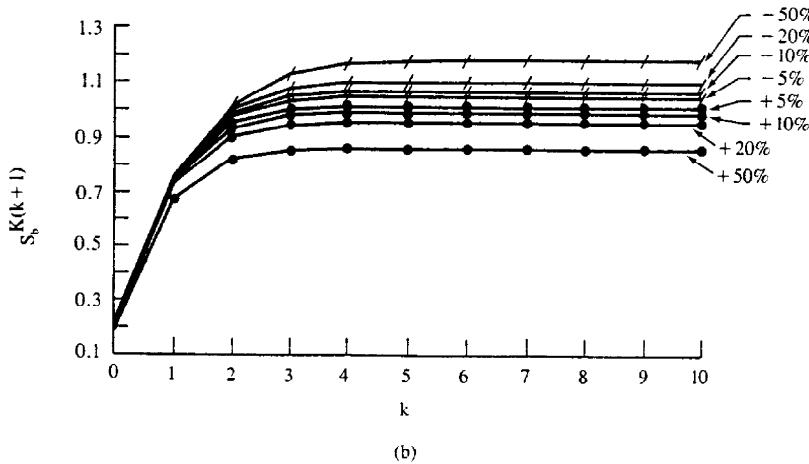
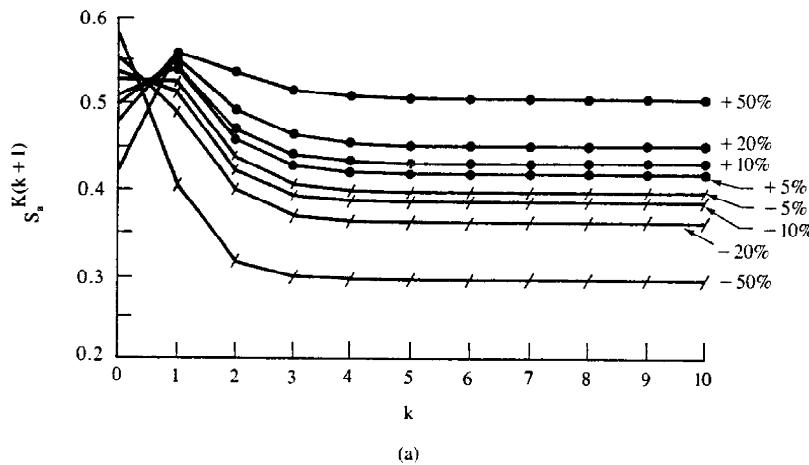


Figure 18-3 Sensitivity plots.

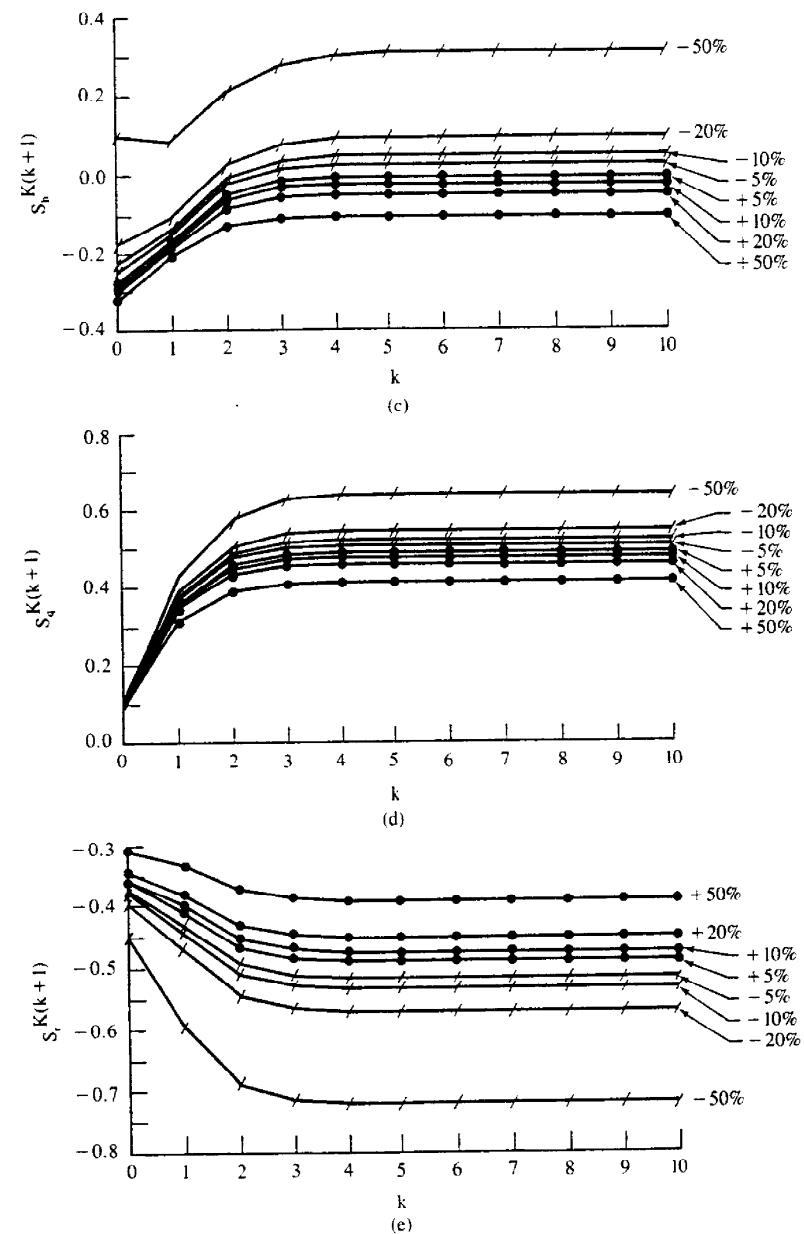


Figure 18-3 Cont.

TABLE 18-1 Steady State Sensitivity Coefficients

		Percentage Change in θ							
		+50	+20	+10	+5	-5	-10	-20	-50
(a) $\bar{S}_\theta^{P(k k)}$	θ	.506	.451	.430	.419	.396	.384	.361	.294
a		.838	.937	.972	.989	1.025	1.042	1.077	1.158
b		-.108	-.053	-.026	-.010	.026	.047	.096	.317
h		.417	.465	.483	.493	.515	.526	.551	.648
q		-.393	-.452	-.476	-.490	-.518	-.534	-.570	-.717
(b) $\bar{S}_\theta^{P(k+1 k)}$	θ	.506	.451	.430	.419	.396	.384	.361	.294
a		.838	.937	.972	.989	1.025	1.042	1.077	1.158
b		-.739	-.877	-.932	-.962	-.1025	-.1059	-.1130	-.1367
h		.417	.465	.483	.493	.515	.526	.551	.648
q		.410	.457	.476	.486	.507	.519	.544	.642
(c) $\bar{S}_\theta^{P(k+1 k)}$	θ	.1047	.820	.754	.723	.664	.637	.585	.453
a		2.022	1.836	1.775	1.745	1.684	1.653	1.591	1.402
b		-.213	-.253	-.268	-.277	-.295	-.305	-.325	-.393
h		.832	.846	.851	.854	.860	.864	.871	.899
q		.118	.131	.137	.140	.146	.149	.157	.185

Here is how to use the results in Table 18-1. From Equation (18-2), for example, we see that

$$\frac{\Delta K_{ij}(k+1)}{K_{ij}(k+1)} = \left(\frac{\Delta \theta}{\theta} \right) S_{\theta}^{K_{ij}(k+1)} \quad (18-6)$$

or

$$\% \text{ Change in } K_{ij}(k+1) = (\% \text{ Change in } \theta) \times S_{\theta}^{K_{ij}(k+1)} \quad (18-7)$$

From Table 18-1 and this formula we see, for example, that a 20% change in a produces a $(20)(.451) = 9.02\%$ change in K , whereas a 20% change in h only produces a $(20)(-.053) = -1.06\%$ change in K , etc. \square

Example 18-3

In the single-channel case, when $w(k)$, $z(k)$ and $v(k)$ are scalars, then $K(k+1)$, $P(k+1|k)$ and $P(k+1|k+1)$ do not depend on q and r separately. Instead, as we demonstrate next, they depend only on the ratio q/r . In this case, $H = h'$, $\Gamma = \gamma$, $Q = q$ and $R = r$, and Equations (17-12), (17-13) and (17-14) can be expressed as

$$K(k+1) = \frac{P(k+1|k)}{r} h' \left[h' \frac{P(k+1|k)}{r} h' + 1 \right]^{-1} \quad (18-8)$$

$$\frac{P(k+1|k)}{r} = \Phi \frac{P(k|k)}{r} \Phi' + \frac{q}{r} \gamma \gamma' \quad (18-9)$$

Examples

and

$$\frac{P(k+1|k+1)}{r} = [I - K(k+1)h'] \frac{P(k+1|k)}{r} \quad (18-10)$$

Observe that, given ratio q/r , we can compute $P(k+1|k)/r$, $K(k+1)$, and $P(k+1|k+1)/r$. We refer to Equations (18-8), (18-9), and (18-10) as the *scaled Kalman filter equations*.

Ratio q/r can be viewed as a *filter tuning parameter*. Recall that, in Lesson 15, we showed q/r is related to signal-to-noise ratio; thus, using (15-50) for example, we can also view signal-to-noise ratio as a (single-channel) Kalman filter tuning parameter. Suppose, for example, that data quality is quite poor, so that signal-to-noise ratio, as measured by $\overline{\text{SNR}}$, is very small. Then q/r will also be very small, because $q/r = \overline{\text{SNR}}/h'(\bar{P}_x/q)h$. In this case the Kalman filter rejects the low-quality data, i.e., the Kalman gain matrix approaches the zero matrix, because

$$\frac{P(k+1|k)}{r} \approx \Phi \frac{P(k|k)}{r} \Phi' \rightarrow 0 \quad (18-11)$$

The Kalman filter is therefore quite sensitive to signal-to-noise ratio, as indeed are most digital filters. Its dependence on signal-to-noise ratio is complicated and nonlinear. Although signal-to-noise ratio (or q/r) enters quite simply into the equation for $P(k+1|k)/r$, it is transformed in a nonlinear manner via (18-8) and (18-10). \square

Example 18-4

A recursive unbiased minimum-variance estimator (BLUE) of a random parameter vector θ can be obtained from the Kalman filter equations in Theorem 17-1 by setting $x(k) = \theta$, $\Phi(k+1,k) = I$, $\Gamma(k+1,k) = 0$, $\Psi(k+1,k) = 0$, and $Q(k) = 0$. Under these conditions we see that $w(k) = 0$ for all k , and

$$x(k+1) = x(k)$$

which means, of course, that $x(k)$ is a vector of constants, namely θ . The Kalman filter equations reduce to

$$\hat{\theta}(k+1|k+1) = \hat{\theta}(k|k) + K(k+1)[z(k+1) - H(k+1)\hat{\theta}(k|k)], \quad (18-12)$$

$$P(k+1|k) = P(k|k)$$

$$K(k+1) = P(k|k)H'(k+1)[H(k+1)P(k|k)H'(k+1) + R(k+1)]^{-1} \quad (18-13)$$

and

$$P(k+1|k+1) = [I - K(k+1)H(k+1)]P(k|k) \quad (18-14)$$

Note that it is no longer necessary to distinguish between filtered and predicted quantities, because $\hat{\theta}(k+1|k) = \hat{\theta}(k|k)$ and $P(k+1|k) = P(k|k)$; hence, the notation $\hat{\theta}(k|k)$ can be simplified to $\hat{\theta}(k)$, for example. Equations (18-12), (18-13), and (18-14) were obtained earlier in Lesson 9 (see Theorem 9-7) for the case of scalar measurements. \square

Example 18-5

This example illustrates the *divergence phenomenon*, which often occurs when either process noise or measurement noise or both are small. We shall see that the Kalman filter locks onto wrong values for the state, but believes them to be the true values, i.e., it “learns” the wrong state too well.

Our example is adapted from Jazwinski (1970, pp. 302–303). We begin with the following simple first-order system

$$x(k+1) = x(k) + b \quad (18-15)$$

$$z(k+1) = x(k+1) + v(k+1) \quad (18-16)$$

where b is a very small bias, so small that, when we design our Kalman filter, we choose to neglect it. Our Kalman filter is based on the following model

$$x_m(k+1) = x_m(k) \quad (18-17)$$

$$z(k+1) = x_m(k+1) + v(k+1) \quad (18-18)$$

Using this model we estimate $x(k)$ and $\hat{x}_m(k|k)$, where it is straightforward to show that

$$\hat{x}_m(k+1|k+1) = \hat{x}_m(k|k) + \underbrace{\frac{p(0)}{(k+1)p(0)+r}}_{K(k+1)} [z(k+1) - \hat{x}_m(k|k)] \quad (18-19)$$

Observe that, as $k \rightarrow \infty$, $K(k+1) \rightarrow 0$ so that $\hat{x}_m(k+1|k+1) \rightarrow \hat{x}_m(k|k)$. The Kalman filter is rejecting the new measurements because it believes (18-17) to be the true model for $x(k)$; but, of course, it is not the true model.

The Kalman filter computes the error variance, $p_m(k|k)$, between $\hat{x}_m(k|k)$ and $x_m(k)$. The true error variance is associated with $\tilde{x}(k|k)$, where

$$\tilde{x}(k|k) = x(k) - \hat{x}_m(k|k) \quad (18-20)$$

We leave it to the reader to show that

$$\begin{aligned} \tilde{x}(k|k) &= \frac{r}{kp(0)+r} \tilde{x}(0|0) \\ &\quad - \frac{p(0)}{kp(0)+r} \sum_{i=1}^k v(i) + \frac{[(k-1)k/2]p(0) + kr}{kp(0)+r} b \end{aligned} \quad (18-21)$$

As $k \rightarrow \infty$, $\tilde{x}(k|k) \rightarrow \infty$ because the third term on the right-hand side of (18-21) diverges to infinity. This term contains the bias b that was neglected in the model used by the Kalman filter. Note also that $\hat{x}_m(k|k) = x_m(k) - \tilde{x}_m(k|k) \rightarrow 0$ as $k \rightarrow \infty$; thus, the Kalman filter has locked onto the wrong state and is unaware that the true error variance is diverging.

A number of different remedies have been proposed for controlling divergence effects, including:

1. adding fictitious process noise,
2. finite-memory filtering, and
3. fading memory filtering.

Fictitious process noise, which appears in the state equation, can be used to account for neglected modeling effects that enter into the state equation (e.g., truncation of second- and higher-order effects when a nonlinear state equation is linearized, as described in Lesson 24). This process noise introduces Q into the Kalman filter equations. Observe, in our first-order example, that Q does not appear in the equa-

tions for $\hat{x}_m(k+1|k+1)$ or $\tilde{x}(k|k)$, because state equation (18-17) contains no process noise.

Divergence is a large-sample property of the Kalman filter. Finite-memory and fading memory filtering control divergence by not letting the Kalman filter get into its “large sample” regime. Finite-memory filtering (Jazwinski, 1970) uses a finite window of measurements (of fixed length W) to estimate $x(k)$. As we move from $t = k_1$ to $t = k_1 + 1$, we must account for two effects, namely, the new measurement at $t = k_1 + 1$ and a discarded measurement at $t = k_1 - W$. Fading-memory filtering, due to Sorenson and Sacks (1971) exponentially ages the measurements, weighting the recent measurement most heavily and past measurements much less heavily. It is analogous to weighted least squares, as described in Lesson 3.

Fading-memory filtering seems to be the most successful and popular way to control divergence effects. \square

PROBLEMS

18-1. Derive the equations for $\hat{x}_m(k+1|k+1)$ and $\tilde{x}(k|k)$, in (18-19) and (18-21), respectively.

18-2. In Lesson 5 we described cross-sectional processing for weighted least-squares estimates. Cross-sectional (also known as sequential) processing can be performed in Kalman filtering. Suppose $\mathbf{z}(k+1) = \text{col}(\mathbf{z}_1(k+1), \mathbf{z}_2(k+1), \dots, \mathbf{z}_q(k+1))$, where $\mathbf{z}_i(k+1) = \mathbf{H}_i \mathbf{x}(k+1) + \mathbf{v}_i(k+1)$, $\mathbf{v}_i(k+1)$ are mutually uncorrelated for $i = 1, 2, \dots, q$, $\mathbf{z}_i(k+1)$ is $m_i \times 1$, and $m_1 + m_2 + \dots + m_q = m$. Let $\hat{\mathbf{x}}(k+1|k+1)$ be a “corrected” estimate of $\mathbf{x}(k+1)$ that is associated with processing $\mathbf{z}_i(k+1)$.

(a) Using the Fundamental Theorem of Estimation Theory, prove that a cross-sectional structure for the corrector equation of the Kalman filter is:

$$\begin{aligned} \hat{\mathbf{x}}_1(k+1|k+1) &= \hat{\mathbf{x}}(k+1|k) + \mathbf{E}\{\mathbf{x}(k+1)|\hat{\mathbf{z}}_1(k+1|k)\} \\ \hat{\mathbf{x}}_2(k+1|k+1) &= \hat{\mathbf{x}}_1(k+1|k+1) + \mathbf{E}\{\mathbf{x}(k+1)|\hat{\mathbf{z}}_2(k+1|k)\} \\ &\vdots \\ \hat{\mathbf{x}}_q(k+1|k+1) &= \hat{\mathbf{x}}_{q-1}(k+1|k+1) + \mathbf{E}\{\mathbf{x}(k+1)|\hat{\mathbf{z}}_q(k+1|k)\} \\ &= \hat{\mathbf{x}}(k+1|k+1) \end{aligned}$$

(b) Provide equations for computing $\mathbf{E}\{\mathbf{x}(k+1)|\hat{\mathbf{z}}_i(k+1|k)\}$.

18-3. (Project) Choose a second-order system and perform a thorough sensitivity study of its associated Kalman filter. Do this for various nominal values and for both small and large variations of the system’s parameters. You will need a computer for this project. Present the results both graphically and tabularly, as in Example 18-2. Draw as many conclusions as possible.

Lesson 19

State Estimation: Steady-State Kalman Filter and Its Relationship to a Digital Wiener Filter

INTRODUCTION

In this lesson we study the steady-state Kalman filter from different points of view, and we then show how it is related to a digital Wiener filter.

STEADY-STATE KALMAN FILTER

For time-invariant and stationary systems, if $\lim_{k \rightarrow \infty} P(k+1|k) = \bar{P}$ exists, then $\lim_{k \rightarrow \infty} K(k) \rightarrow \bar{K}$ and the Kalman filter (17-11) becomes a constant coefficient filter. Because $P(k+1|k)$ and $P(k|k)$ are intimately related, then, if \bar{P} exists, $\lim_{k \rightarrow \infty} P(k|k) \triangleq \bar{P}_1$ also exists. We have already observed limiting behaviors for $K(k+1)$ and $P(k|k)$ in Example 18-1. The following theorem, which is adopted from Anderson and Moore (1979, pg. 77), tells us when \bar{P} exists, and assures us that the steady-state Kalman filter will be asymptotically stable.

Theorem 19-1 (Steady-State Kalman Filter). *If our dynamical model in Equations (15-17) and (15-18) is time-invariant, stationary, and asymptotically stable (i.e., all the eigenvalues of Φ lie inside the unit circle), then:*

- For any nonnegative-definite symmetric initial condition $P(0|−1)$, one has $\lim_{k \rightarrow \infty} P(k+1|k) = \bar{P}$ with \bar{P} independent of $P(0|−1)$ and satisfying the following steady-state version of Equation (17-24):*

$$\bar{P} = \Phi \bar{P} [\mathbf{I} - \mathbf{H}'(\mathbf{H}\bar{P}\mathbf{H}' + \mathbf{R})^{-1}\mathbf{H}\bar{P}] \Phi' + \Gamma Q \Gamma' \quad (19-1)$$

Equation (19-1) is often referred to either as a steady-state or algebraic Riccati equation.

- The eigenvalues of the steady-state Kalman filter, $\lambda[\Phi - \bar{K}\Phi]$, all lie within the unit circle, so that the filter is asymptotically stable; i.e.,*

$$|\lambda[\Phi - \bar{K}\Phi]| < 1 \quad (19-2)$$

If our dynamical model in Equations (15-17) and (15-18) is time-invariant and stationary, but is not necessarily asymptotically stable, then points (a) and (b) still hold as long as the system is completely stabilizable and detectable. \square

A proof of this theorem is beyond the scope of this textbook. It can be found in Anderson and Moore, pp. 78–82 (1979). For definitions of the system-theoretic terms *stabilizable* and *detectable*, the reader should consult a textbook on linear systems, such as Kailath (1980) or Chen (1970). By *completely detectable* and *completely stabilizable*, we mean that (Φ, H) is completely detectable and $(\Phi, \Gamma Q_1)$ is completely stabilizable, where $Q = Q_1 Q_1'$. Additionally, any asymptotically stable model is always completely stabilizable and detectable.

Probably the most interesting case of a system that is not asymptotically stable, for which we want to design a steady-state Kalman filter, is one that has a pole on the unit circle.

Example 19-1

In this example (which is similar to Example 5.4 in Meditch, 1969, pp. 189–190) we consider the scalar system

$$x(k+1) = x(k) + w(k) \quad (19-3)$$

$$z(k+1) = x(k+1) + v(k+1) \quad (19-4)$$

It has a pole on the unit circle. When $q = 20$ and $r = 5$, Equations (17-13), (17-12) and (17-14) reduce to:

$$p(k+1|k) = p(k|k) + 20 \quad (19-5)$$

$$K(k+1) = \frac{p(k|k) + 20}{p(k|k) + 25} \quad (19-6)$$

and

$$p(k+1|k+1) = \frac{5[p(k|k) + 20]}{p(k|k) + 25} = 5 K(k+1) \quad (19-7)$$

Starting with $p(0|0) = p_s(0) = 50$, it is a relatively simple matter to compute $p(k+1|k)$, $K(k+1)$ and $p(k+1|k+1)$ for $k = 0, 1, \dots$. The results for the first few iterations are given in Table 19-1.

TABLE 19-1 Kalman Filter Quantities

k	$p(k k-1)$	$K(k)$	$p(k k)$
0	50
1	70	0.933	4.67
2	24.67	0.831	4.16
3	24.16	0.829	4.14
4	24.14	0.828	4.14

The steady-state value of $p(k|k)$ is obtained by setting $p(k+1|k+1) = p(k|k) = \bar{p}_1$ in the last of the above three relations to obtain

$$\bar{p}_1^2 + 20\bar{p}_1 - 100 = 0 \quad (19-8)$$

Because \bar{p}_1 is a variance it must be nonnegative; hence, only the solution $\bar{p}_1 = 4.14$ is valid. Comparing this result with $p(3|3)$ we see that the filter is in the steady state to within the indicated computational accuracy after processing just three measurements. In steady-state, the Kalman filter equation is

$$\begin{aligned} \hat{x}(k+1|k+1) &= \hat{x}(k|k) + 0.828[z(k+1) - \hat{x}(k|k)] \\ &= 0.172\hat{x}(k|k) + 0.828z(k+1) \end{aligned} \quad (19-9)$$

Observe that the filter's pole at 0.172 lies inside the unit circle. \square

Many ways have been reported for solving the algebraic Riccati equation (19-1) [see Laub (1979), for example], ranging from direct iteration of the matrix Riccati Equation (17-24) until $\bar{P}(k+1|k)$ does not change appreciably from $\bar{P}(k|k-1)$, to solving the nonlinear algebraic Riccati equation via an iterative Newton-Raphson procedure, to solving that equation in one shot by the Schur method. Iterative methods are quite sensitive to error accumulation. The one-shot Schur method possesses a high degree of numerical integrity, and appears to be one of the most successful ways for obtaining \bar{P} . For details about this method, see Laub (1979).

In summary, then, to design a steady-state Kalman filter:

1. given $(\Phi, \Gamma, \Psi, H, Q, R)$, compute \bar{P} , the positive definite solution of (19-1);
2. compute \bar{K} , as

$$\bar{K} = \bar{P}H'(\bar{P}H' + R)^{-1} \quad (19-10)$$

3. use (19-10) in

$$\begin{aligned} \hat{x}(k+1|k+1) &= \Phi\hat{x}(k|k) + \Psi u(k) + \bar{K}\hat{z}(k+1|k) \\ &= (\mathbf{I} - \bar{K}H)\Phi\hat{x}(k|k) + \bar{K}z(k+1) \\ &\quad + (\mathbf{I} - \bar{K}H)\Psi u(k) \end{aligned} \quad (19-11)$$

SINGLE-CHANNEL STEADY-STATE KALMAN FILTER

The steady-state Kalman filter can be viewed outside of the context of estimation theory as a recursive digital filter. As such, it is sometimes useful to be able to compute its impulse response, transfer function, and frequency response. In this section we restrict our attention to the single-channel steady-state Kalman filter. From (19-11) we observe that this filter is excited by two inputs, $z(k+1)$ and $u(k)$ [in the single-channel case, $z(k+1)$ and $u(k)$]. In this section we shall only be interested in transfer functions which are associated with the effect of $z(k+1)$ on the filter; hence, we set $u(k) = 0$. Additionally, we shall view the signal component of measurement $z(k)$ as our desired output. The signal component of $z(k)$ is $h'\hat{x}(k)$.

Let $H_f(z)$ denote the z -transform of the impulse response of the steady-state filter; i.e.,

$$H_f(z) = \mathcal{Z}\{\hat{z}(k|k) \text{ when } z(k) = \delta(k)\} \quad (19-12)$$

This transfer function is found from the following *steady-state filter system*:

$$\hat{x}(k|k) = [\mathbf{I} - \bar{K}h']\Phi\hat{x}(k-1|k-1) + \bar{K}\delta(k) \quad (19-13)$$

$$\hat{z}(k|k) = h'\hat{x}(k|k) \quad (19-14)$$

Taking the z -transform of (19-13) and (19-14), it follows that

$$H_f(z) = h'(\mathbf{I} - \Phi_f z^{-1})^{-1} \bar{K} \quad (19-15)$$

where

$$\Phi_f = [\mathbf{I} - \bar{K}h']\Phi \quad (19-16)$$

Equation (19-15) can also be written as

$$H_f(z) = h_f(0) + h_f(1)z^{-1} + h_f(2)z^{-2} + \dots \quad (19-17)$$

where the filter's coefficients (i.e., Markov parameters), $h_f(j)$, are

$$h_f(j) = h' \Phi_f^j \bar{K} \quad (19-18)$$

for $j = 0, 1, \dots$

In our study of mean-squared smoothing, in Lessons 20, 21, and 22, we will see that the steady-state predictor system plays an important role. The *steady-state predictor*, obtained from (17-23), is given by

$$\hat{x}(k+1|k) = \Phi_p\hat{x}(k|k-1) + \gamma_p z(k) \quad (19-19)$$

where

$$\Phi_p = \Phi(\mathbf{I} - \bar{K}h') \quad (19-20)$$

and

$$\gamma_p = \Phi \bar{K} \quad (19-21)$$

Let $H_p(z)$ denote the z -transform of the impulse response of the steady-state predictor, i.e.,

$$H_p(z) = \mathcal{Z}\{\hat{z}(k|k-1)\} \text{ when } z(k) = \delta(k) \quad (19-22)$$

This transfer function is found from the following *steady-state predictor system*:

$$\hat{x}(k+1|k) = \Phi_p \hat{x}(k|k-1) + \gamma_p \delta(k) \quad (19-23)$$

$$\hat{z}(k|k-1) = \mathbf{h}' \hat{x}(k|k-1) \quad (19-24)$$

Taking the z -transform of these equations, we find that

$$H_p(z) = \mathbf{h}' (\mathbf{I} - \Phi_p z^{-1})^{-1} z^{-1} \gamma_p \quad (19-25)$$

which can also be written as

$$H_p(z) = h_p(1)z^{-1} + h_p(2)z^{-2} + \dots \quad (19-26)$$

where the predictor's coefficients, $h_p(j)$, are

$$h_p(j) = \mathbf{h}' \Phi_p^{j-1} \gamma_p \quad (19-27)$$

for $j = 1, 2, \dots$.

Although $H_f(z)$ and $H_p(z)$ contain an infinite number of terms, they can usually be truncated, because both are associated with asymptotically stable filters.

Example 19-2

In this example we examine the impulse response of the steady-state predictor for the first-order system

$$x(k+1) = \frac{1}{\sqrt{2}} x(k) + w(k) \quad (19-28)$$

$$z(k+1) = \frac{1}{\sqrt{2}} x(k+1) + v(k+1) \quad (19-29)$$

In Example 15-2 of Lesson 15, we showed that ratio q/r is proportional to signal-to-noise ratio, $\overline{\text{SNR}}$. For this example, $\overline{\text{SNR}} = q/r$. Our numerical results, given below, are for $\overline{\text{SNR}} = 20, 5, 1$ and use the scaled steady-state prediction-error variance \bar{p}/r , which can be solved for from (19-1) when it is written as:

$$\frac{\bar{p}}{r} = \frac{1}{2} \frac{\bar{p}}{r} \left[1 - \frac{1}{2} \left(\frac{1}{2} \frac{\bar{p}}{r} + 1 \right)^{-1} \frac{\bar{p}}{r} \right] + \overline{\text{SNR}} \quad (19-30)$$

Note that we could just as well have solved (19-1) for \bar{p}/q but the structure of its equation is more complicated than (19-30). The positive solution of (19-30) is

$$\frac{\bar{p}}{r} = \frac{1}{2} (\overline{\text{SNR}} - 1) + \frac{1}{2} \sqrt{(\overline{\text{SNR}} - 1)^2 + 8 \overline{\text{SNR}}} \quad (19-31)$$

Additionally,

$$\bar{K} = \frac{1}{\sqrt{2}} \frac{\bar{p}}{r} \left(\frac{1}{2} \frac{\bar{p}}{r} + 1 \right)^{-1} \quad (19-32)$$

Table 19-2 summarizes \bar{p}/r , \bar{K} , ϕ_p and γ_p , quantities which are needed to compute the impulse response $h_p(k)$ for $k \geq 0$, which is depicted in Figure 19-1. Observe that all three responses peak at $j = 1$; however, the decay time for $\overline{\text{SNR}} = 20$ is quicker than

TABLE 19-2 Steady-State Predictor Quantities

$\overline{\text{SNR}}$	\bar{p}/r	\bar{K}	ϕ_p	γ_p
20	20.913	1.287	0.064	0.910
5	5.742	1.049	0.183	0.742
1	1.414	0.586	0.414	0.414

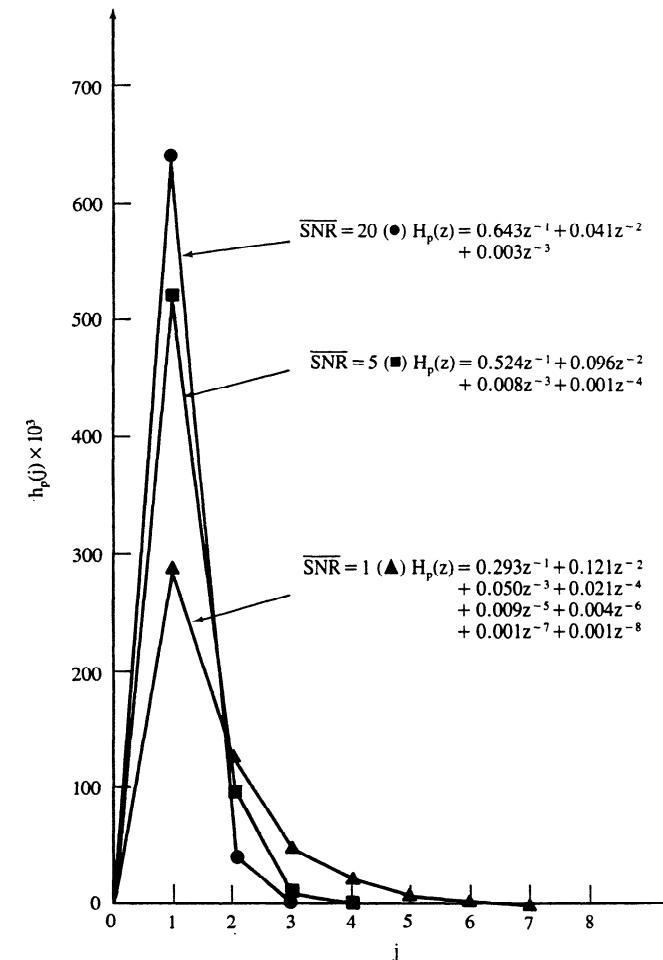


Figure 19-1 Impulse response of steady-state predictor for $\overline{\text{SNR}} = 20, 5$ and 1 . $H_p(z)$ is shown to three significant figures (Mendel, 1981, © 1981, IEEE).

the decay times for lower \overline{SNR} values, and the peak amplitude is larger for higher \overline{SNR} values. The steady-state predictor tends to reject measurements which have a low signal-to-noise ratio (the same is true for the steady-state filter). \square

Example 19-3

In this example we present impulse response and frequency response plots for the steady-state predictors associated with the systems

$$H_1(z) = \frac{1}{z^2 - z + 0.5} \quad (19-33)$$

and

$$H_2(z) = \frac{-0.688z^3 + 1.651z^2 - 1.221z + 0.25}{z^4 - 2.586z^3 + 2.489z^2 - 1.033z + 0.168} \quad (19-34)$$

Matrices Φ , γ and \mathbf{h} for these systems are:

$$\Phi_1 = \begin{pmatrix} 0 & 1 \\ -1/2 & 1 \end{pmatrix} \quad \gamma_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \mathbf{h}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and

$$\Phi_2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -0.168 & +1.033 & -2.489 & +2.586 \end{pmatrix} \quad \gamma_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad \mathbf{h}_2 = \begin{pmatrix} 0.250 \\ -1.221 \\ 1.651 \\ -0.688 \end{pmatrix}$$

Figure 19-2 depicts $h_1(k)$, $|H_1(j\omega)|$ (in db) and $\angle H_1(j\omega)$ as well as $h_{p1}(k)$, $|H_{p1}(j\omega)|$, and $\angle H_{p1}(j\omega)$ for $\overline{SNR} = 1, 5$, and 20 . Figure 19-3 depicts comparable quantities for the second system. Observe that, as signal-to-noise ratio decreases, the steady-state predictor rejects the measurements; for the amplitudes of $h_{p1}(k)$ and $h_{p2}(k)$ become smaller as \overline{SNR} becomes smaller. It also appears, from examination of $|H_{p1}(j\omega)|$ and $|H_{p2}(j\omega)|$, that at high signal-to-noise ratios the steady-state predictor behaves like a high-pass filter for system-1 and a bandpass filter for system-2. On the other hand, at low signal-to-noise ratios it appears to behave like a band-pass filter.

The steady-state predictor appears to be quite dependent on system dynamics at high signal-to-noise ratios, but is much less dependent on the system dynamics at low signal-to-noise ratios. At low signal-to-noise ratios the predictor is rejecting the measurements regardless of system dynamics.

Just because the steady-state predictor behaves like a high-pass filter at high signal-to-noise ratios does not mean that it passes a lot of noise through it, because *high signal-to-noise ratio* means that measurement noise level is quite low. Of course, a spurious burst of noise would pass through this filter quite easily. \square

RELATIONSHIPS BETWEEN THE STEADY-STATE KALMAN FILTER AND A FINITE IMPULSE RESPONSE DIGITAL WIENER FILTER

The steady-state Kalman filter is a recursive digital filter with filter coefficients equal to $h_f(j)$, $j = 0, 1, \dots$ [see (19-18)]. Quite often $h_f(j) \rightarrow 0$ for $j \geq J$, so that $H_f(z)$ can be truncated, i.e.,

$$H_f(z) = h_f(0) + h_f(1)z^{-1} + \dots + h_f(J)z^{-J} \quad (19-35)$$

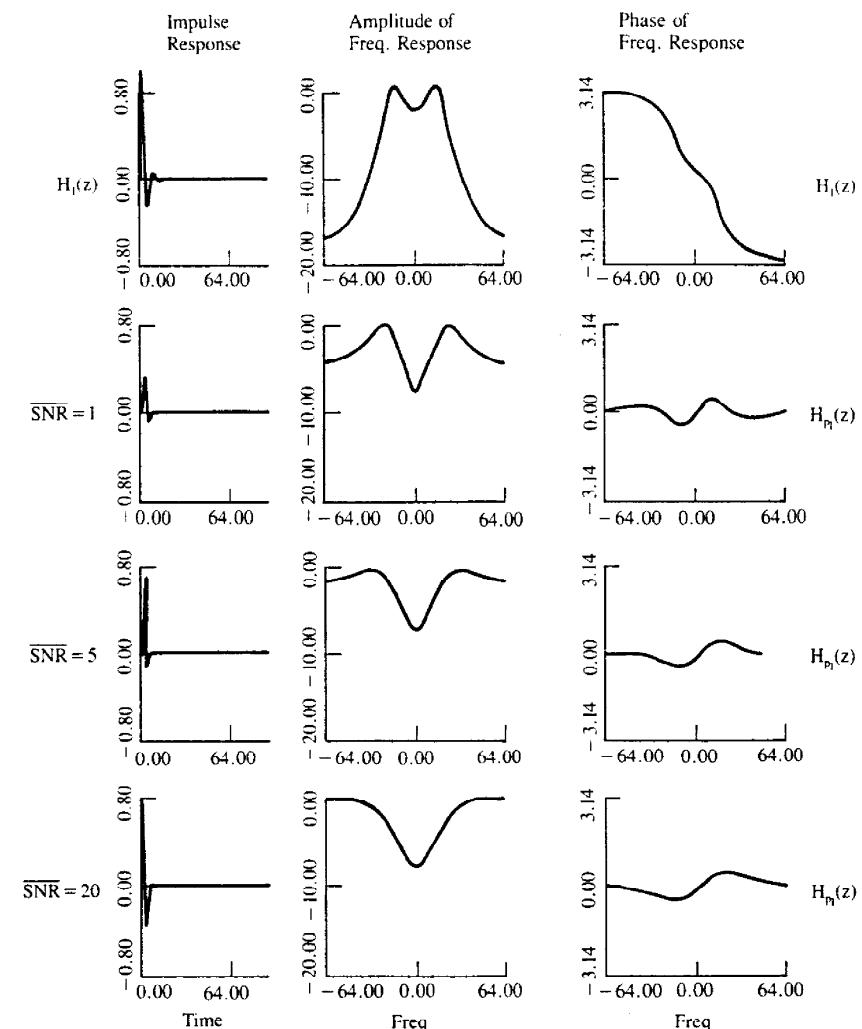


Figure 19-2 Impulse and frequency responses for system $H_1(z)$ and its associated steady-state predictor.

The *truncated steady-state Kalman filter* can then be implemented as a finite-impulse response (FIR) digital filter.

There is a more direct way for designing a FIR minimum mean-squared error filter, i.e., a *digital Wiener filter*, as we describe next.

Consider the situation depicted in Figure 19-4. We wish to design digital filter $F(z)$'s coefficients $f(0), f(1), \dots, f(\eta)$ so that the filter's output, $y(k)$, is

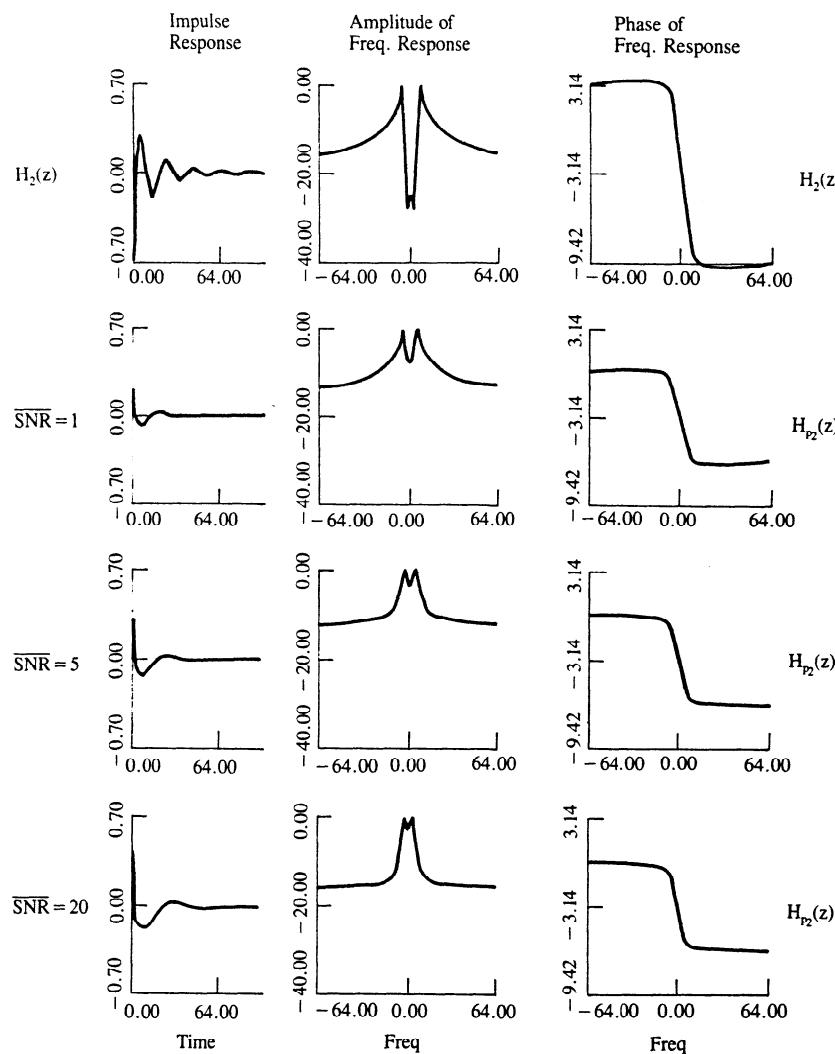


Figure 19-3 Impulse and frequency responses for system $H_2(z)$ and its associated steady-state predictor.

close, in some sense, to a desired signal $d(k)$. In a digital Wiener filter design, $f(0), f(1), \dots, f(\eta)$ are obtained by minimizing the following mean-squared error

$$I(f) = E\{[d(k) - y(k)]^2\} = E\{e^2(k)\} \quad (19-36)$$

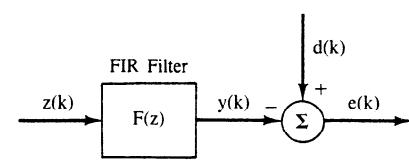


Figure 19-4 Starting point for design of FIR digital Wiener filter.

Using the fact that

$$y(k) = f(k)*z(k) = \sum_{i=0}^{\eta} f(i)z(k-i) \quad (19-37)$$

we see that

$$I(f) = E\{[d(k) - \sum_{i=0}^{\eta} f(i)z(k-i)]^2\} \quad (19-38)$$

The filter coefficients that minimize $I(f)$ are found by setting $\partial I(f) / \partial f(j) = 0$ for $j = 0, 1, \dots, \eta$. We leave it as an exercise for the reader to show that doing this results in the following linear system of $\eta + 1$ equations in the $\eta + 1$ unknown filter coefficients,

$$\sum_{i=0}^{\eta} f(i)\phi_{zz}(i-j) = \phi_{zd}(j) \quad j = 0, 1, \dots, \eta \quad (19-39)$$

where $\phi_{zz}(\cdot)$ is the auto-correlation function of filter input $z(k)$, and $\phi_{zd}(\cdot)$ is the cross-correlation function between $z(k)$ and $d(k)$. Equations (19-39) are known as the *discrete-time Wiener-Hopf equations*. They are a system of normal equations, and can be solved in many different ways, the fastest of which is by the Levinson Algorithm (Treitel and Robinson, 1966).

The minimum mean-squared error, $I^*(f)$, can be shown to be given by

$$I^*(f) = \phi_{dd}(0) - \sum_{i=0}^{\eta} f(i)\phi_{zd}(i) \quad (19-40)$$

One property of the digital Wiener filter is that $I^*(f)$ becomes smaller as η , the number of filter coefficients, increases. In general, $I^*(f)$ approaches a nonzero limiting value, a value that is often reached for modest values of η .

In order to relate this FIR Wiener filter to the truncated steady-state Kalman filter, we must first assume a signal-plus-noise model for $z(k)$, i.e., (see Figure 19-5),

$$z(k) = s(k) + v(k) = h(k)*w(k) + v(k) \quad (19-41)$$

where $h(k)$ is the impulse response of a linear time-invariant system, and, as in our basic state-variable model (Lesson 15), $w(k)$ and $v(k)$ are mutually uncorrelated (stationary) white noise processes with variances q and r , respectively. We must also specify an explicit form for "desired signal" $d(k)$. We shall require that

$$d(k) = s(k) = h(k)*w(k) \quad (19-42)$$

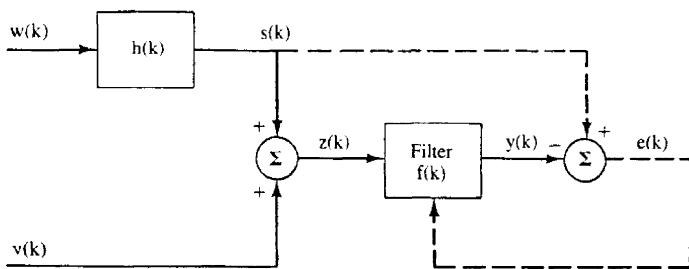


Figure 19-5 Signal-plus-noise model incorporated into the design of a FIR digital Wiener filter. The dashed lines denote paths that exist only during the design stage; these paths do not exist in the implementation of the filter.

which means, of course, that we want the output of the FIR digital Wiener filter to be as close as possible to signal $s(k)$.

Using (19-41) and (19-42), it is another straightforward exercise to compute $\phi_{zz}(i - j)$ and $\phi_{zd}(j)$. The resulting discrete-time Wiener-Hopf equations are

$$\sum_{i=0}^{\eta} f(i) \left[\frac{q}{r} \phi_{hh}(j-i) + \delta(i-j) \right] = \frac{q}{r} \phi_{hh}(j) \quad j = 0, 1, \dots, \eta \quad (19-43)$$

where

$$\phi_{hh}(i) = \sum_{l=0}^{\infty} h(l)h(l+i) \quad (19-44)$$

Observe, that, as in the case of a single-channel Kalman filter, our single channel Wiener filter only depends on the ratio q/r (and subsequently on SNR).

Theorem 19-2. *The steady-state Kalman filter is an infinite-length digital Wiener filter. The truncated steady-state Kalman filter is a FIR digital Wiener filter.*

Proof (heuristic). The digital Wiener filter has constant coefficients as does the steady-state Kalman filter. Both filters minimize error variances. The infinite-length Wiener filter has the smallest error variance of all Wiener filters, as does the steady-state Kalman filter; hence, the steady-state Kalman filter is an infinite length digital Wiener filter.

The second part of the theorem is proved in a similar manner. \square

Using Theorem 19-2, we suggest the following procedure for designing a recursive Wiener filter (i.e., a steady-state Kalman filter):

1. obtain a state-variable representation of $h(k)$;
2. determine the steady-state Kalman gain matrix, \bar{K} , as described above;
3. implement the steady-state Kalman filter (19-11);
4. compute the estimate of desired signal $s(k)$, as

$$\hat{s}(k|k) = h' \hat{x}(k|k) \quad (19-45)$$

COMPARISONS OF KALMAN AND WIENER FILTERS

We conclude this lesson with a brief comparison of Kalman and Wiener filters. First of all, the filter designs use different types of modeling information. Auto-correlation information is used for FIR digital Wiener filtering, whereas a combination of auto-correlation and difference equation information is used in Kalman filtering. In order to compute $\phi_{hh}(l)$, $h(k)$ must be known; it is always possible to go directly from $h(k)$ to the parameters in a state-variable model [i.e., (Φ, γ, h)], using an approximate realization procedure (e.g., Kung, 1978).

Because the Kalman filter is recursive, it is an infinite-length filter; hence, unlike the FIR Wiener filter, where filter length is a design variable, the filter's length is not a design variable in the Kalman filter.

Our derivation of the digital Wiener filter was for a single-channel model, whereas the derivation of the Kalman filter was for a more general multichannel model (i.e., the case of vector inputs and outputs). For a derivation of a multichannel digital Wiener filter, see Treitel (1970).

There is a conceptual difference between Kalman and FIR Wiener filtering. A Wiener filter acts directly on the measurements to provide a signal $\hat{s}(k|k)$ that is "close" to $s(k)$. The Kalman filter does not do this directly. It provides us with such a signal in two steps. In the first step, a signal is obtained that is "close" to $x(k)$, and in the second step, a signal is obtained that is "close" to $s(k)$. The first step provides the optimal filtered value of $x(k)$, $\hat{x}(k|k)$; the second step provides the optimal filtered value of $s(k)$, because $\hat{s}(k|k) = h' \hat{x}(k|k)$.

Finally, we can picture a diagram similar to the one in Figure 19-5 for the Kalman filter, except that the dashed lines are solid ones. In the Kalman filter the filter coefficients are usually time-varying and are affected in real time by a measure of error $e(k)$ (i.e., by error-covariance matrices), whereas Wiener filter coefficients are constant and have only been affected by a measure of $e(k)$ during their design.

In short, a Kalman filter is a generalization of the FIR digital Wiener filter to include time-varying filter coefficients which incorporate the effects of error in an active manner. Additionally, a Kalman filter is applicable to either time-varying or nonstationary systems or both, whereas the digital Wiener filter is not.

PROBLEMS

- 19-1. Derive the discrete-time Wiener-Hopf equations given in (19-39).
- 19-2. Prove that the minimum mean-squared error, $I^*(\mathbf{f})$, given in (19-40), becomes smaller as η , the number of filter coefficients, increases.
- 19-3. Consider the basic state-variable model, $x(k+1) = \frac{2}{\sqrt{3}} x(k) + w(k)$ and $z(k+1) = x(k+1) + v(k+1)$, where $q = 2$, $r = 1$, $m_x(0) = 0$ and $E\{x^2(0)\} = 2$.
- (a) Specify $\hat{x}(0|0)$ and $p(0|0)$.
 - (b) Give the recursive predictor for this system.
 - (c) Obtain the steady-state predictor.
 - (d) Suppose $z(5)$ is not provided (i.e., there is a gap in the measurements at $k = 5$). How does this affect $p(6|5)$ and $p(100|99)$?
- 19-4. Consider the basic scalar system $x(k+1) = \phi x(k) + w(k)$ and $z(k+1) = x(k+1) + v(k+1)$. Assume that $q = 0$ and let $\lim_{k \rightarrow \infty} p(k|k) = \bar{p}_1$.
- (a) Show that $\bar{p}_1 = 0$ and $\bar{p}_1 = (\phi^2 - 1)r / \phi^2$.
 - (b) Which of the two solutions in (a) is the correct one when $\phi^2 < 1$?

Lesson 20

State Estimation: Smoothing

THREE TYPES OF SMOOTHERS

Recall that in smoothing we obtain mean-squared estimates of state vector $\mathbf{x}(k)$, $\hat{\mathbf{x}}(k|j)$ for which $k < j$. From the Fundamental Theorem of Estimation Theory, Theorem 13-1, we know that the structure of the mean-squared smoother is

$$\hat{\mathbf{x}}(k|j) = E\{\mathbf{x}(k)|\mathcal{Z}(j)\} \quad \text{where} \quad k < j \quad (20-1)$$

In this lesson we shall develop recursive smoothing algorithms that are comparable to our recursive Kalman filter algorithm.

Smoothing is much more complicated than filtering, primarily because we are using future measurements to obtain estimates at earlier time points. Because there is some flexibility associated with how we choose to process future measurements, it is convenient to distinguish between three types of smoothing (Meditch, 1969, pp. 204–208): fixed-interval, fixed-point, and fixed-lag.

The *fixed-interval smoother* is $\hat{\mathbf{x}}(k|N)$, $k = 0, 1, \dots, N - 1$, where N is a fixed positive integer. The situation here is as follows: with an experiment completed, we have measurements available over the fixed interval $1 \leq k \leq N$. For each time point within this interval we wish to obtain the optimal estimate of state vector $\mathbf{x}(k)$, which is based on *all* the available measurement data $\{\mathbf{z}(j), j = 1, 2, \dots, N\}$. Fixed-interval smoothing is very useful in signal processing, where the processing is done after all the data is

collected. It cannot be carried out on-line during an experiment, as filtering can be. Because all the available data is used, one cannot hope to do better (by other forms of smoothing) than by fixed-interval smoothing.

The *fixed-point* smoothed estimate is $\hat{x}(k|j)$, $j = k + 1, k + 2, \dots$ where k is a fixed positive integer. Suppose we want to improve our estimate of a state at a specific time by making use of future measurements. Let this time be \bar{k} . Then fixed-point smoothed estimates of $x(\bar{k})$ will be $\hat{x}(\bar{k}|\bar{k}+1)$, $\hat{x}(\bar{k}|\bar{k}+2), \dots$, etc. The last possible fixed-point estimate is $\hat{x}(\bar{k}|N)$, which is the same as the fixed-interval estimate of $x(\bar{k})$. Fixed-point smoothing can be carried out on-line, if desired, but the calculation of $\hat{x}(\bar{k}|\bar{k}+d)$ is subject to a delay of dT sec.

The *fixed-lag* smoothed estimate is $\hat{x}(k|k+L)$, $k = 0, 1, \dots$, where L is a fixed positive integer. In this case, the point at which we seek the estimate of the system's state lags the time point of the most recent measurement by a fixed interval of time, L ; i.e., $t_{k+L} - t_k = L$ which is a positive constant for all $k = 0, 1, \dots$. This type of estimator can be used where a constant lag between measurements and state estimates is permissible. Fixed-lag smoothing can be carried out on-line, if desired, but the calculation of $\hat{x}(k|k+L)$ is subject to an LT sec delay.

APPROACHES FOR DERIVING SMOOTHERS

The literature on smoothing is filled with many different approaches for deriving recursive smoothers. By augmenting suitably defined states to (15-17) and (15-18), one can reduce the derivation of smoothing formulas to a Kalman filter for the augmented state-variable model (Anderson and Moore, 1979). The "filtered" estimates of the newly-introduced states turn out to be equivalent to smoothed values of $x(k)$. We shall examine this augmentation approach in Lesson 21. A second approach is to use the orthogonality principle to derive a discrete-time Wiener-Hopf equation, which can then be used to establish the smoothing formulas. We do not treat this approach in this book.

A third approach, the one we shall follow in this lesson, is based on the causal invertibility between the innovations process $\tilde{z}(k+j|k+j-1)$ and measurement $z(k+j)$, and repeated applications of Theorem 12-4.

A SUMMARY OF IMPORTANT FORMULAS

The following formulas, which have been derived in earlier lessons, are used so frequently in this lesson, as well as in Lesson 21, that we collect them here for the convenience of the reader:

Single-Stage Smoother

$$\tilde{z}(k+1|k) = z(k+1) - H(k+1)\hat{x}(k+1|k) \quad (20-2)$$

$$\tilde{z}(k+1|k) = H(k+1)\tilde{x}(k+1|k) + v(k+1) \quad (20-3)$$

$$P_{\tilde{z}\tilde{z}}(k+1|k) = H(k+1)P(k+1|k)H'(k+1) + R(k+1) \quad (20-4)$$

$$\tilde{x}(k+1|k) = \Phi(k+1,k)\tilde{x}(k|k) + \Gamma(k+1,k)w(k) \quad (20-5)$$

and

$$\begin{aligned} \tilde{x}(k+1|k+1) &= [I - K(k+1)H(k+1)]\tilde{x}(k+1|k) \\ &\quad - K(k+1)v(k+1) \end{aligned} \quad (20-6)$$

SINGLE-STAGE SMOOTHER

As in our study of prediction, it is useful first to develop a single-stage smoother and then to obtain more general smoothers.

Theorem 20-1. *The single-stage mean-squared smoothed estimator of $x(k)$, $\hat{x}(k|k+1)$, is given by the expression*

$$\hat{x}(k|k+1) = \hat{x}(k|k) + M(k|k+1)\tilde{z}(k+1|k) \quad (20-7)$$

where single-stage smoother gain matrix, $M(k|k+1)$, is

$$\begin{aligned} M(k|k+1) &= P(k|k)\Phi'(k+1,k)H'(k+1) \\ &\quad [H(k+1)P(k+1|k)H'(k+1) + R(k+1)]^{-1} \end{aligned} \quad (20-8)$$

Proof. From the Fundamental Theorem of Estimation Theory and Theorem 12-4, we know that

$$\begin{aligned} \hat{x}(k|k+1) &= E\{x(k)|\mathcal{Z}(k+1)\} \\ &= E\{x(k)|\mathcal{Z}(k), z(k+1)\} \\ &= E\{x(k)|\mathcal{Z}(k), \tilde{z}(k+1|k)\} \\ &= E\{x(k)|\mathcal{Z}(k)\} + E\{x(k)|\tilde{z}(k+1|k)\} - m_s(k) \end{aligned} \quad (20-9)$$

which can also be expressed as (see Corollary 13-1)

$$\hat{x}(k|k+1) = \hat{x}(k|k) + P_{\tilde{z}\tilde{z}}(k,k+1|k)P_{\tilde{z}\tilde{z}}^{-1}(k+1|k)\tilde{z}(k+1|k) \quad (20-10)$$

Defining single-stage smoother gain matrix $M(k|k+1)$ as

$$M(k|k+1) = P_{\tilde{z}\tilde{z}}(k,k+1|k)P_{\tilde{z}\tilde{z}}^{-1}(k+1|k) \quad (20-11)$$

(20-10) reduces to (20-7).

Next, we must show that $M(k|k+1)$ can also be expressed as in (20-8). We already have an expression for $P_{\tilde{z}\tilde{z}}(k+1|k)$, namely (20-4); hence, we

must compute $\mathbf{P}_{\bar{x}}(k, k+1|k)$. To do this, we make use of (20-3), (20-5) and the orthogonality principle, as follows:

$$\begin{aligned}\mathbf{P}_{\bar{x}}(k, k+1|k) &= \mathbb{E}\{\mathbf{x}(k)\bar{\mathbf{z}}'(k+1|k)\} \\ &= \mathbb{E}\{\mathbf{x}(k)[\mathbf{H}(k+1)\bar{\mathbf{x}}(k+1|k) + \mathbf{v}(k+1)]'\} \\ &= \mathbb{E}\{\mathbf{x}(k)\bar{\mathbf{x}}'(k+1|k)\}\mathbf{H}'(k+1) \\ &= \mathbb{E}\{\mathbf{x}(k)[\Phi(k+1,k)\bar{\mathbf{x}}(k|k) \\ &\quad + \Gamma(k+1,k)\mathbf{w}(k)']'\}\mathbf{H}'(k+1) \\ &= \mathbb{E}\{\mathbf{x}(k)\bar{\mathbf{x}}'(k|k)\}\Phi'(k+1,k)\mathbf{H}'(k+1) \\ &= \mathbb{E}\{[\hat{\mathbf{x}}(k|k) + \bar{\mathbf{x}}(k|k)]\bar{\mathbf{x}}'(k|k)\}\Phi'(k+1,k)\mathbf{H}'(k+1) \\ &= \mathbf{P}(k|k)\Phi'(k+1,k)\mathbf{H}'(k+1)\end{aligned}\quad (20-12)$$

Substituting (20-12) and (20-4) into (20-11), we obtain (20-8). \square

For future reference, we record the following fact,

$$\mathbb{E}\{\mathbf{x}(k)\bar{\mathbf{x}}'(k+1|k)\} = \mathbf{P}(k|k)\Phi'(k+1,k) \quad (20-13)$$

This is obtained by comparing the third and last lines of (20-12).

Observe that the structure of the single-stage smoother is quite similar to that of the Kalman filter. The Kalman filter obtains $\hat{\mathbf{x}}(k+1|k+1)$ by adding a correction that depends on the most recent innovations, $\bar{\mathbf{z}}(k+1|k)$, to the *predicted* value of $\mathbf{x}(k+1)$. The single-stage smoother, on the other hand, obtains $\hat{\mathbf{x}}(k|k+1)$ by adding a correction that also depends on $\bar{\mathbf{z}}(k+1|k)$, to the *filtered* value of $\mathbf{x}(k)$. We see that filtered estimates are required to obtain smoothed estimates.

Corollary 20-1. *Kalman gain matrix $\mathbf{K}(k+1)$ is a factor in $\mathbf{M}(k|k+1)$, i.e.,*

$$\mathbf{M}(k|k+1) = \mathbf{A}(k)\mathbf{K}(k+1) \quad (20-14)$$

where

$$\mathbf{A}(k) \triangleq \mathbf{P}(k|k)\Phi'(k+1,k)\mathbf{P}^{-1}(k+1|k) \quad (20-15)$$

Proof. Using the fact that [Equation (17-12)]

$$\begin{aligned}\mathbf{K}(k+1) &= \mathbf{P}(k+1|k)\mathbf{H}'(k+1) \\ &= [\mathbf{H}(k+1)\mathbf{P}(k+1|k)\mathbf{H}'(k+1) + \mathbf{R}(k+1)]^{-1}\end{aligned}\quad (20-16)$$

we see that

$$\begin{aligned}\mathbf{H}'(k+1)[\mathbf{H}(k+1)\mathbf{P}(k+1|k)\mathbf{H}'(k+1) + \mathbf{R}(k+1)]^{-1} \\ &= \mathbf{P}^{-1}(k+1|k)\mathbf{K}(k+1)\end{aligned}\quad (20-17)$$

When (20-17) is substituted into (20-8), we obtain (20-14). \square

Corollary 20-2. *Another way to express $\hat{\mathbf{x}}(k|k+1)$ is*

$$\hat{\mathbf{x}}(k|k+1) = \hat{\mathbf{x}}(k|k) + \mathbf{A}(k)[\hat{\mathbf{x}}(k+1|k+1) - \hat{\mathbf{x}}(k+1|k)] \quad (20-18)$$

Proof. Substitute (20-14) into (20-7), to see that

$$\hat{\mathbf{x}}(k|k+1) = \hat{\mathbf{x}}(k|k) + \mathbf{A}(k)\mathbf{K}(k+1)\bar{\mathbf{z}}(k+1|k) \quad (20-19)$$

but [see (17-11)]

$$\mathbf{K}(k+1)\bar{\mathbf{z}}(k+1|k) = \hat{\mathbf{x}}(k+1|k+1) - \hat{\mathbf{x}}(k+1|k) \quad (20-20)$$

Substitute (20-20) into (20-19) to obtain the desired result in (20-18). \square

Formula (20-7) is useful for computational purposes, whereas (20-18) is most useful for theoretical purposes. These facts will become more clear when we examine double-stage smoothing in our next section.

Whereas the structure of the single-stage smoother is similar to that of the Kalman filter, we see that $\mathbf{M}(k|k+1)$ does not depend on single-stage smoothing error-covariance matrix $\mathbf{P}(k|k+1)$. Kalman gain $\mathbf{K}(k+1)$, of course, does depend on $\mathbf{P}(k+1|k)$ [or $\mathbf{P}(k|k)$]. In fact, $\mathbf{P}(k|k+1)$ does not appear at all in the smoothing equations and must be computed (if one desires to do so) separately. We address this calculation in Lesson 21.

DOUBLE-STAGE SMOOTHER

Instead of immediately generalizing the single-stage smoother to an N stage smoother, we first present results for the double-stage smoother. We will then be able to write down the general results (almost) by inspection of the single- and double-stage results.

Theorem 20-2. *The double-stage mean-squared smoothed estimator of $\mathbf{x}(k)$, $\hat{\mathbf{x}}(k|k+2)$, is given by the expression*

$$\hat{\mathbf{x}}(k|k+2) = \hat{\mathbf{x}}(k|k+1) + \mathbf{M}(k|k+2)\bar{\mathbf{z}}(k+2|k+1) \quad (20-21)$$

where double-stage smoother gain matrix, $\mathbf{M}(k|k+2)$, is

$$\begin{aligned}\mathbf{M}(k|k+2) &= \mathbf{P}(k|k)\Phi'(k+1,k)[\mathbf{I} - \mathbf{K}(k+1) \\ &\quad \mathbf{H}(k+1)']'\Phi'(k+2,k+1) \\ &= \mathbf{H}'(k+2)[\mathbf{H}(k+2)\mathbf{P}(k+2|k+1) \\ &\quad \mathbf{H}'(k+2) + \mathbf{R}(k+2)]^{-1}\end{aligned}\quad (20-22)$$

Proof. From the Fundamental Theorem of Estimation Theory, Theorem 12-4, and Corollary 13-1, we know that

$$\begin{aligned}\hat{\mathbf{x}}(k|k+2) &= \mathbb{E}\{\mathbf{x}(k)|\mathcal{Z}(k+2)\} \\ &= \mathbb{E}\{\mathbf{x}(k)|\mathcal{Z}(k+1), \mathbf{z}(k+2)\} \\ &= \mathbb{E}\{\mathbf{x}(k)|\mathcal{Z}(k+1), \bar{\mathbf{z}}(k+2|k+1)\} \\ &= \mathbb{E}\{\mathbf{x}(k)|\mathcal{Z}(k+1)\} + \mathbb{E}\{\mathbf{x}(k)|\bar{\mathbf{z}}(k+2|k+1)\} - \mathbf{m}_s(k)\end{aligned}\quad (20-23)$$

which can also be expressed as

$$\hat{x}(k|k+2) = \hat{x}(k|k+1) + P_{\hat{x}\hat{x}}(k, k+2|k+1)P_{\hat{z}\hat{z}}^{-1}(k+2|k+1)\hat{z}(k+2|k+1) \quad (20-24)$$

Defining double-stage smoother gain matrix $M(k|k+2)$ as

$$M(k|k+2) = P_{\hat{x}\hat{x}}(k, k+2|k+1)P_{\hat{z}\hat{z}}^{-1}(k+2|k+1) \quad (20-25)$$

(20-24) reduces to (20-21).

In order to show that $M(k|k+2)$ in (20-25) can be expressed as in (20-22), one proceeds as in our derivation of $M(k|k+1)$ in (20-12); however, the details are lengthier because $\hat{z}(k+2|k+1)$ involves quantities that are two time units away from $x(k)$, whereas $\hat{z}(k+1|k)$ involves quantities that are only one time unit away from $x(k)$. Equation (20-13) is used during the derivation. We leave the detailed derivation of (20-22) as an exercise for the reader. \square

Whereas (20-22) is a computationally useful formula, it is not useful from a theoretical viewpoint; i.e., when we examine $M(k|k+1)$ in (20-8) and $M(k|k+2)$ in (20-22), it is not at all obvious how to generalize these formulas to $M(k|k+N)$, or even to $M(k|k+3)$. The following result for $M(k|k+2)$ is easily generalized.

Corollary 20-3. *Kalman gain matrix $K(k+2)$ is a factor in $M(k|k+2)$, i.e.,*

$$M(k|k+2) = A(k)A(k+1)K(k+2) \quad (20-26)$$

where $A(k)$ is defined in (20-15).

Proof. Increment k to $k+1$ in (20-17) to see that

$$\begin{aligned} H'(k+2)[H(k+2)P(k+2|k+1)H'(k+2) + R(k+2)]^{-1} \\ = P^{-1}(k+2|k+1)K(k+2) \end{aligned} \quad (20-27)$$

Next, note from (17-14), that

$$I - K(k+1)H(k+1) = P(k+1|k+1)P^{-1}(k+1|k) \quad (20-28)$$

hence,

$$[I - K(k+1)H(k+1)]' = P^{-1}(k+1|k)P(k+1|k+1) \quad (20-29)$$

Substitute (20-27) and (20-29) into (20-22) to see that

$$\begin{aligned} M(k|k+2) &= P(k|k)\Phi'(k+1,k)P^{-1}(k+1|k) \\ &\quad P(k+1|k+1)\Phi'(k+2,k+1) \\ &\quad P^{-1}(k+2|k+1)K(k+2) \end{aligned} \quad (20-30)$$

Using the definition of matrix $A(k)$ in (20-15), we see that (20-30) can be expressed as in (20-26). \square

Corollary 20-4. *Two other ways to express $\hat{x}(k|k+2)$ are:*

$$\begin{aligned} \hat{x}(k|k+2) &= \hat{x}(k|k+1) \\ &\quad + A(k)A(k+1)[\hat{x}(k+2|k+2) - \hat{x}(k+2|k+1)] \end{aligned} \quad (20-31)$$

and

$$\hat{x}(k|k+2) = \hat{x}(k|k) + A(k)[\hat{x}(k+1|k+2) - \hat{x}(k+1|k)] \quad (20-32)$$

Proof. The derivation of (20-31) follows exactly the same path as the derivation of (20-18), and is therefore left as an exercise for the reader. Equation (20-31) is the starting place for the derivation of (20-32). Observe, from (20-18), that

$$\begin{aligned} A(k+1)[\hat{x}(k+2|k+2) - \hat{x}(k+2|k+1)] \\ = \hat{x}(k+1|k+2) - \hat{x}(k+1|k+1) \end{aligned} \quad (20-33)$$

thus, (20-31) can be written as

$$\begin{aligned} \hat{x}(k|k+2) &= \hat{x}(k|k+1) \\ &\quad + A(k)[\hat{x}(k+1|k+2) - \hat{x}(k+1|k+1)] \end{aligned} \quad (20-34)$$

Substituting (20-18) into (20-34), we obtain the desired result in (20-32). \square

The alternate forms we have obtained for both $\hat{x}(k|k+1)$ and $\hat{x}(k|k+2)$ will suggest how we can generalize our single- and double-stage smoothers to N -stage smoothers.

SINGLE- AND DOUBLE-STAGE SMOOTHERS AS GENERAL SMOOTHERS

At the beginning of this lesson we described three types of smoothers, namely: fixed-interval, fixed-point, and fixed-lag smoothers. Table 20-1 shows how our single- and double-stage smoothers fit into these three categories.

In order to obtain the fixed-interval smoother formulas, given for both the single- and double-stage smoothers, set $k+1=N$ in (20-18) and $k+2=N$ in (20-32), respectively. Doing this forces the left-hand side of both equations to be conditioned on data length N . Observe that, before we can compute $\hat{x}(N-1|N)$ or $\hat{x}(N-2|N)$, we must run a Kalman filter on all of the data in order to obtain $\hat{x}(N|N)$. This last filtered state estimate initializes the backward running fixed-interval smoother. Observe, also, that we must compute $\hat{x}(N-1|N)$ before we can compute $\hat{x}(N-2|N)$. Clearly, the limitation of our results so far is that we can only perform fixed-interval smoothing for

TABLE 20-1 Smoothing Interrelationships

	Single-Stage	Double-Stage
Fixed-Interval	$\hat{x}(N - l N) = \hat{x}(N - 1 N - 1) + A(N - 1)$ [$\hat{x}(N N) - \hat{x}(N N - 1)$]	$\hat{x}(N - 2 N) = \hat{x}(N - 2 N - 2) + A(N - 2)$ [$\hat{x}(N - 1 N) - \hat{x}(N - 1 N - 2)$]
	<p>Solution proceeds in reverse time, from N to $N - 1$, where N is fixed.</p>	<p>Solution proceeds in reverse time, from N to $N - 1$, to $N - 2$, where N is fixed.</p>
Fixed-Point	$\hat{x}(\bar{k} \bar{k} + 1) = \hat{x}(\bar{k} \bar{k}) + A(\bar{k})$ $[\hat{x}(\bar{k} + 1 \bar{k} + 1) - \hat{x}(\bar{k} + 1 \bar{k})]$ <i>k fixed at \bar{k}</i>	$\hat{x}(\bar{k} \bar{k} + 2) = \frac{\hat{x}(\bar{k} \bar{k} + 1) + A(\bar{k})A(\bar{k} + 1)}{[\hat{x}(\bar{k} + 2 \bar{k} + 2) - \hat{x}(\bar{k} + 2 \bar{k} + 1)]}$ $\hat{x}(\bar{k} \bar{k})$ $\hat{x}(\bar{k} + 1 \bar{k} + 1)$ <i>k fixed at \bar{k}</i>
	<p>Solution proceeds in forward time from \bar{k} to $\bar{k} + 1$ on the filtering time scale and then back to \bar{k} on the smoothing time scale. A one-unit time delay is present.</p>	<p>Solution proceeds in forward time. Results from single-stage smoother as well as optimal filter are required at $g = \bar{k} + 1$, whereas at $g = \bar{k}$ only results from optimal filter are required. A two-unit time delay is present.</p>

Table 20-1 Cont.

	Single-Stage	Double-Stage
Fixed-Lag	$\hat{x}(k k + 1) = \hat{x}(k k) + A(k)$ [$\hat{x}(k + 1 k + 1) - \hat{x}(k + 1 k)$]	$\hat{x}(k k + 2) = \hat{x}(k k + 1) + A(k)A(k + 1)$ [$\hat{x}(k + 2 k + 2) - \hat{x}(k + 2 k + 1)$]
	<p><i>k</i> variable</p> <p>Picture and discussion same as in Fixed-Point case, replacing \bar{k} by k. Here we have a one-unit window from k to $k + 1$.</p>	<p><i>k</i> variable</p> <p>Picture and discussion same as in Fixed-Point case, replacing \bar{k} by k. Here we have a two-unit window from k to $k + 2$.</p>

$k = N - 1$ and $N - 2$. More general results, that will permit us to perform fixed-interval smoothing for any $k < N$, are described in Lesson 21.

In order to obtain the fixed-point smoother formulas, given for both the single- and double-stage smoothers, set $k = \bar{k}$ in (20-18) and (20-31), respectively. Time point \bar{k} represents the fixed point in our smoother formulas. As noted at the beginning of this lesson, fixed-point smoothing can be carried out on-line, but it is subject to a delay. From an information availability point of view, a one-unit time delay is present in the single-stage fixed-point smoother, because our smoothed estimate at $k = \bar{k}$ uses the filtered estimate computed at $k = \bar{k} + 1$. A two-unit time delay is present in the double-stage fixed-point smoother, because our smoothed estimate at $\mathcal{S} = \bar{k} + 1$ uses the filtered estimate computed at $k = \bar{k} + 1$ and $\bar{k} + 2$.

The fixed-lag smoother formulas look just like the fixed-point formulas, except that k is a variable in the former and is fixed in the latter. Observe that the windows of measurements used in the single-stage fixed-lag smoother are nonoverlapping, whereas they are overlapping in the case of the double-stage fixed-lag smoother. As in the case of fixed-point smoothing, fixed-interval smoothing can be carried out on line, subject, of course, to a fixed delay equal to the "lag" of the smoother.

PROBLEMS

- 20-1. Derive the formula for the double-stage smoother gain matrix $M(k|k+2)$, given in (20-22).
- 20-2. Derive the alternative expression for $\hat{x}(k|k+2)$ given in (20-31).
- 20-3. Using the Fundamental Theorem of Estimation Theory, derive expressions for $\hat{w}(k|k+1)$ and $\hat{w}(k|k+2)$. These represent single- and double-stage estimators of disturbance $w(k)$.

Lesson 21

State-Estimation: Smoothing (General Results)

INTRODUCTION

In Lesson 20 we introduced three general types of smoothers, namely, fixed-interval, fixed-point, and fixed-lag smoothers. We also developed formulas for single-stage and double-stage smoothers and showed how these specialized smoothers fit into the three general categories. In this lesson we shall develop general formulas for fixed-interval, fixed-point, and fixed-lag smoothers.

FIXED-INTERVAL SMOOTHERS

In this section we develop a number of algorithms for the fixed-interval smoothed estimator of $x(k)$, $\hat{x}(k|N)$. Not all of them will be useful, because we will not be able to compute with some. Only those with which we can compute will be considered useful.

Recall that

$$\hat{x}(k|N) = E\{x(k)|\mathcal{E}(N)\} \quad (21-1)$$

It is straightforward to proceed as we did in Lesson 20, by showing first that

$$\begin{aligned} \hat{x}(k|N) &= E\{x(k)|\mathcal{E}(N-1), \hat{z}(N|N-1)\} \\ &= E\{x(k)|\mathcal{E}(N-1)\} + E\{x(k)|\hat{z}(N|N-1)\} - m_x(k) \quad (21-2) \\ &= \hat{x}(k|N-1) + M(k|N)\hat{z}(N|N-1) \end{aligned}$$

where

$$\mathbf{M}(k|N) = \mathbf{P}_{\dot{\mathbf{x}}\dot{\mathbf{x}}}(k, N|N-1)\mathbf{P}_{\ddot{\mathbf{z}}\ddot{\mathbf{z}}}^{-1}(N|N-1) \quad (21-3)$$

and then that

$$\mathbf{M}(k|N) = \mathbf{A}(k)\mathbf{A}(k+1)\dots\mathbf{A}(N-1)\mathbf{K}(N) \quad (21-4)$$

and, finally, that other algorithms for $\hat{\mathbf{x}}(k|N)$ are

$$\hat{\mathbf{x}}(k|N) = \hat{\mathbf{x}}(k|N-1) + \left[\prod_{i=k}^{N-1} \mathbf{A}(i) \right] [\hat{\mathbf{x}}(N|N) - \hat{\mathbf{x}}(N|N-1)] \quad (21-5)$$

and

$$\hat{\mathbf{x}}(k|N) = \hat{\mathbf{x}}(k|k) + \mathbf{A}(k)[\hat{\mathbf{x}}(k+1|N) - \hat{\mathbf{x}}(k+1|k)] \quad (21-6)$$

A detailed derivation of these results that uses a mathematical induction proof, can be found in Meditch (1969, pp. 216–220). The proof relies, in part, on our previously derived results in Lesson 20, for the single- and double-stage smoothers. The reader, however, ought to be able to obtain (21-4), (21-5), and (21-6) directly from the rules of formation that can be inferred from the Lesson 20 formulas for $\mathbf{M}(k|k+1)$ and $\mathbf{M}(k|k+2)$, and $\hat{\mathbf{x}}(k|k+1)$ and $\hat{\mathbf{x}}(k|k+2)$.

In order to compute $\hat{\mathbf{x}}(k|N)$ for specific values of k , all of the terms that appear on the right-hand side of an algorithm must be available. We have three possible algorithms for $\hat{\mathbf{x}}(k|N)$; however, as we explain next, only one is “useful.”

The first value of k for which the right-hand side of (21-2) is fully available is $k = N - 1$. By running a Kalman filter over all of the measurements, we are able to compute $\hat{\mathbf{x}}(N-1|N-1)$ and $\hat{\mathbf{z}}(N|N-1)$, so that we can compute $\hat{\mathbf{x}}(N-1|N)$. We can now try to iterate (21-2) in the backward direction, to see if we can compute $\hat{\mathbf{x}}(N-2|N)$, $\hat{\mathbf{x}}(N-3|N)$, etc. Setting $k = N - 2$ in (21-2), we obtain

$$\hat{\mathbf{x}}(N-2|N) = \hat{\mathbf{x}}(N-2|N-1) + \mathbf{M}(N-2|N)\hat{\mathbf{z}}(N|N-1) \quad (21-7)$$

Unfortunately, $\hat{\mathbf{x}}(N-2|N-1)$, which is a single-stage smoothed estimate of $\mathbf{x}(N-2)$, has not been computed; hence, (21-7) is not useful for computing $\hat{\mathbf{x}}(N-2|N)$. We, therefore, reject (21-2) as a useful fixed-interval smoother.

A similar argument can be made against (21-5); thus, we also reject (21-5) as a useful fixed-interval smoother.

Equation (21-6) is a useful fixed-interval smoother, because all of its terms on its right-hand side are available when we iterate it in the backward direction. For example,

$$\begin{aligned} \hat{\mathbf{x}}(N-1|N) &= \hat{\mathbf{x}}(N-1|N-1) \\ &\quad + \mathbf{A}(N-1)[\hat{\mathbf{x}}(N|N) - \hat{\mathbf{x}}(N|N-1)] \end{aligned} \quad (21-8)$$

and

$$\begin{aligned} \hat{\mathbf{x}}(N-2|N) &= \hat{\mathbf{x}}(N-2|N-2) \\ &\quad + \mathbf{A}(N-2)[\hat{\mathbf{x}}(N-1|N) - \hat{\mathbf{x}}(N-1|N-2)] \end{aligned} \quad (21-9)$$

Observe how $\hat{\mathbf{x}}(N-1|N)$ is used in the calculation of $\hat{\mathbf{x}}(N-2|N)$. Equation (21-6) was developed by Rauch, et al. (1965).

Theorem 21-1. A useful mean-squared fixed-interval smoothed estimator of $\mathbf{x}(k)$, $\hat{\mathbf{x}}(k|N)$, is given by the expression

$$\hat{\mathbf{x}}(k|N) = \hat{\mathbf{x}}(k|k) + \mathbf{A}(k)[\hat{\mathbf{x}}(k+1|N) - \hat{\mathbf{x}}(k+1|k)] \quad (21-10)$$

where matrix $\mathbf{A}(k)$ is defined in (20-15), and $k = N - 1, N - 2, \dots, 0$. Additionally, the error-covariance matrix associated with $\hat{\mathbf{x}}(k|N)$, $\mathbf{P}(k|N)$, is given by

$$\mathbf{P}(k|N) = \mathbf{P}(k|k) + \mathbf{A}(k)[\mathbf{P}(k+1|N) - \mathbf{P}(k+1|k)]\mathbf{A}'(k) \quad (21-11)$$

where $k = N - 1, N - 2, \dots, 0$.

Proof. We have already derived (21-10); hence, we direct our attention at the derivation of the algorithm for $\mathbf{P}(k|N)$. Our derivation of (21-11) follows the one given in Meditch, 1969, pp. 222–224. To begin, we know that

$$\hat{\mathbf{x}}(k|N) = \mathbf{x}(k) - \hat{\mathbf{x}}(k|N) \quad (21-12)$$

Substitute (21-10) into (21-12) to see that

$$\hat{\mathbf{x}}(k|N) = \hat{\mathbf{x}}(k|k) - \mathbf{A}(k)[\hat{\mathbf{x}}(k+1|N) - \hat{\mathbf{x}}(k+1|k)] \quad (21-13)$$

Collecting terms conditioned on N on the left-hand side, and terms conditioned on k on the right-hand side, (21-13) becomes

$$\hat{\mathbf{x}}(k|N) + \mathbf{A}(k)\hat{\mathbf{x}}(k+1|N) = \hat{\mathbf{x}}(k|k) + \mathbf{A}(k)\hat{\mathbf{x}}(k+1|k) \quad (21-14)$$

Treating (21-14) as an identity, we see that the covariance of its left-hand side must equal the covariance of its right-hand side; hence, [in order to obtain (21-15) we use the orthogonality principle to eliminate all the cross-product terms]

$$\begin{aligned} \mathbf{P}(k|N) + \mathbf{A}(k)\mathbf{P}_{\dot{\mathbf{x}}\dot{\mathbf{x}}}(k+1|N)\mathbf{A}'(k) \\ = \mathbf{P}(k|k) + \mathbf{A}(k)\mathbf{P}_{\dot{\mathbf{x}}\dot{\mathbf{x}}}(k+1|k)\mathbf{A}'(k) \end{aligned} \quad (21-15)$$

or

$$\mathbf{P}(k|N) = \mathbf{P}(k|k) + \mathbf{A}(k)[\mathbf{P}_{\dot{\mathbf{x}}\dot{\mathbf{x}}}(k+1|k) - \mathbf{P}_{\dot{\mathbf{x}}\dot{\mathbf{x}}}(k+1|N)]\mathbf{A}'(k) \quad (21-16)$$

We must now evaluate the two covariances $\mathbf{P}_{\dot{\mathbf{x}}\dot{\mathbf{x}}}(k+1|k)$ and $\mathbf{P}_{\dot{\mathbf{x}}\dot{\mathbf{x}}}(k+1|N)$.

Recall that $\mathbf{x}(k+1) = \hat{\mathbf{x}}(k+1|k) + \tilde{\mathbf{x}}(k+1|k)$; thus,

$$\mathbf{P}_{\dot{\mathbf{x}}\dot{\mathbf{x}}}(k+1) = \mathbf{P}_{\dot{\mathbf{x}}\dot{\mathbf{x}}}(k+1|k) + \mathbf{P}(k+1|k) \quad (21-17)$$

Additionally, $\mathbf{x}(k+1) = \hat{\mathbf{x}}(k+1|N) + \tilde{\mathbf{x}}(k+1|N)$, so that

$$\mathbf{P}_x(k+1) = \mathbf{P}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}(k+1|N) + \mathbf{P}(k+1|N) \quad (21-18)$$

Equating (21-17) and (21-18) we find that

$$\mathbf{P}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}(k+1|k) - \mathbf{P}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}(k+1|N) = \mathbf{P}(k+1|N) - \mathbf{P}(k+1|k) \quad (21-19)$$

Finally, substituting (21-19) into (21-16), we obtain the desired expression for $\mathbf{P}(k|N)$, (21-11). \square

We leave proof of the fact that $\{\tilde{\mathbf{x}}(k|N), k = N, N-1, \dots, 0\}$ is a zero-mean second-order Gauss-Markov process as an exercise for the reader.

Example 21-1

In order to illustrate fixed-interval smoothing and obtain at the same time a comparison of the relative accuracies of smoothing and filtering, we return to Example 19-1. To review briefly, we have the scalar system $x(k+1) = x(k) + w(k)$ with the scalar measurement $z(k+1) = x(k+1) + v(k+1)$ and $p(0) = 50$, $q = 20$, and $r = 5$. In this example (which is similar to Example 6.1 in Meditch, 1969, pg. 225) we choose $N = 4$ and compute quantities that are associated with $\hat{x}(k|4)$, where from (21-10)

$$\hat{x}(k|4) = \hat{x}(k|k) + A(k)[\hat{x}(k+1|4) - \hat{x}(k+1|k)] \quad (21-20)$$

$k = 3, 2, 1, 0$. Because $\Phi = 1$ and $p(k+1|k) = p(k|k) + 20$

$$A(k) = p(k|k)\Phi'p^{-1}(k+1|k) = \frac{p(k|k)}{p(k|k) + 20} \quad (21-21)$$

and, therefore,

$$p(k|4) = p(k|k) + \left[\frac{p(k|k)}{p(k|k) + 20} \right]^2 [p(k+1|4) - p(k+1|k)] \quad (21-22)$$

Utilizing these last two expressions, we compute $A(k)$ and $p(k|4)$ for $k = 3, 2, 1, 0$ and present them, along with the results summarized in Table 19-1, in Table 21-1. The three estimation error variances are given in adjacent columns for ease in comparison.

TABLE 21-1 Kalman Filter and Fixed-interval Smoother Quantities

k	$p(k k-1)$	$p(k k)$	$p(k 4)$	$K(k)$	$A(k)$
0	...	50	16.31	...	0.714
1	70	4.67	3.92	0.933	0.189
2	24.67	4.16	3.56	0.831	0.172
3	24.16	4.14	3.74	0.829	0.171
4	24.14	4.14	4.14	0.828	0.171

Observe the large improvement (percentage wise) of $p(k|4)$ over $p(k|k)$. Improvement seems to get larger the farther away we get from the end of our data; thus, $p(0|4)$ is more than three times as small as $p(0|0)$. Of course, it should be, because

Fixed-Interval Smoothers

$p(0|0)$ is an *initial condition* that is data independent, whereas $p(0|4)$ is a result of processing $z(1), z(2), z(3)$, and $z(4)$. In essence, fixed-interval smoothing has let us look into the future and reflect the future back to time zero.

Finally, note that, for large values of k , $A(k)$ reaches a steady-state value, \bar{A} , where in this example

$$\bar{A} = \bar{p}_1 / (\bar{p}_1 + 20) = 0.171 \quad (21-23)$$

This steady-state value is achieved for $k = 3$. \square

Equation (21-10) requires the multiplication of $3 n \times n$ matrices as well as a matrix inversion at each iteration; hence, it is somewhat limited for practical computing purposes. The following results, which are due to Bryson and Frazier (1963) and Bierman (1973b), represent the most practical way for computing $\hat{x}(k|N)$ and also $\mathbf{P}(k|N)$.

Theorem 21-2. (a) A most useful mean-squared fixed-interval smoothed estimator of $\mathbf{x}(k)$, $\hat{\mathbf{x}}(k|N)$, is

$$\hat{\mathbf{x}}(k|N) = \hat{\mathbf{x}}(k|k-1) + \mathbf{P}(k|k-1)\mathbf{r}(k|N) \quad (21-24)$$

where $k = N-1, N-2, \dots, 1$, and $n \times 1$ vector \mathbf{r} satisfies the backward-recursive equation

$$\begin{aligned} \mathbf{r}(j|N) &= \Phi_p'(j+1,j)\mathbf{r}(j+1|N) \\ &\quad + \mathbf{H}'(j)[\mathbf{H}(j)\mathbf{P}(j|j-1)\mathbf{H}'(j) + \mathbf{R}(j)]^{-1}\tilde{\mathbf{z}}(j|j-1) \end{aligned} \quad (21-25)$$

where $j = N, N-1, \dots, 1$ and $\mathbf{r}(N+1|N) = \mathbf{0}$.

(b) The smoothing error-covariance matrix, $\mathbf{P}(k|N)$, is

$$\mathbf{P}(k|N) = \mathbf{P}(k|k-1) - \mathbf{P}(k|k-1)\mathbf{S}(k|N)\mathbf{P}(k|k-1) \quad (21-26)$$

where $k = N-1, N-2, \dots, 1$, and $n \times n$ matrix $\mathbf{S}(j|N)$, which is the covariance matrix of $\mathbf{r}(j|N)$, satisfies the backward-recursive equation

$$\begin{aligned} \mathbf{S}(j|N) &= \Phi_p'(j+1,j)\mathbf{S}(j+1|N)\Phi_p(j+1,j) \\ &\quad + \mathbf{H}'(j)[\mathbf{H}(j)\mathbf{P}(j|j-1)\mathbf{H}'(j) + \mathbf{R}(j)]^{-1}\mathbf{H}(j) \end{aligned} \quad (21-27)$$

where $j = N, N-1, \dots, 1$ and $\mathbf{S}(N+1|N) = \mathbf{0}$. Matrix Φ_p is defined in (21-33).

Proof. (Mendel, 1983b, pp. 64–65). (a) Substitute the Kalman filter equation (17-11) for $\hat{\mathbf{x}}(k|k)$ into Equation (21-10), to show that

$$\begin{aligned} \hat{\mathbf{x}}(k|N) &= \hat{\mathbf{x}}(k|k-1) + \mathbf{K}(k)\tilde{\mathbf{z}}(k|k-1) \\ &\quad + \mathbf{A}(k)[\hat{\mathbf{x}}(k+1|N) - \hat{\mathbf{x}}(k+1|k)] \end{aligned} \quad (21-28)$$

Residual state vector $\mathbf{r}(k|N)$ is defined as

$$\mathbf{r}(k|N) = \mathbf{P}^{-1}(k|k-1)[\hat{\mathbf{x}}(k|N) - \hat{\mathbf{x}}(k|k-1)] \quad (21-29)$$

Next, substitute $\mathbf{r}(k|N)$ and $\mathbf{r}(k+1|N)$, using (21-29), into (21-28), to show that

$$\begin{aligned}\mathbf{r}(k|N) &= \mathbf{P}^{-1}(k|k-1)[\mathbf{K}(k)\tilde{\mathbf{z}}(k|k-1) \\ &\quad + \mathbf{P}(k|k)\Phi'(k+1,k)\mathbf{r}(k+1|N)]\end{aligned}\quad (21-30)$$

From (17-12) and (17-13) and the symmetry of covariance matrices, we find that

$$\mathbf{P}^{-1}(k|k-1)\mathbf{K}(k) = \mathbf{H}'(k)[\mathbf{H}(k)\mathbf{P}(k|k-1)\mathbf{H}'(k) + \mathbf{R}(k)]^{-1} \quad (21-31)$$

and

$$\mathbf{P}^{-1}(k|k-1)\mathbf{P}(k|k) = [\mathbf{I} - \mathbf{K}(k)\mathbf{H}(k)]' \quad (21-32)$$

Substituting (21-31) and (21-32) into Equation (21-30), and defining

$$\Phi_p(k+1,k) = \Phi(k+1,k)[\mathbf{I} - \mathbf{K}(k)\mathbf{H}(k)] \quad (21-33)$$

we obtain Equation (21-25). Setting $k = N+1$ in (21-29), we establish $\mathbf{r}(N+1|N) = \mathbf{0}$. Finally, solving (21-29) for $\hat{\mathbf{x}}(k|N)$ we obtain Equation (21-24).

(b) The orthogonality principle in Corollary 13-3 leads us to conclude that

$$\mathbf{E}\{\tilde{\mathbf{x}}(k|N)\mathbf{r}'(k|N)\} = \mathbf{0} \quad (21-34)$$

because $\mathbf{r}(k|N)$ is simply a linear combination of all the observations $\mathbf{z}(1), \mathbf{z}(2), \dots, \mathbf{z}(N)$. From (21-24) we find that

$$\tilde{\mathbf{x}}(k|k-1) = \tilde{\mathbf{x}}(k|N) - \mathbf{P}(k|k-1)\mathbf{r}(k|N) \quad (21-35)$$

and, therefore, using (21-34), we find that

$$\mathbf{P}(k|k-1) = \mathbf{P}(k|N) - \mathbf{P}(k|k-1)\mathbf{S}(k|N)\mathbf{P}(k|k-1) \quad (21-36)$$

where

$$\mathbf{S}(k|N) = \mathbf{E}\{\mathbf{r}(k|N)\mathbf{r}'(k|N)\} \quad (21-37)$$

is the covariance-matrix of $\mathbf{r}(k|N)$ [note that $\mathbf{r}(k|N)$ is zero mean]. Equation (21-36) is solved for $\mathbf{P}(k|N)$ to give the desired result in Equation (21-26).

Because the innovations process is uncorrelated, (21-27) follows from substitution of (21-25) into (21-37). Finally, $\mathbf{S}(N+1|N) = \mathbf{0}$ because $\mathbf{r}(N+1|N) = \mathbf{0}$. \square

Equations (21-24) and (21-25) are very efficient; they require no matrix inversions or multiplications of $n \times n$ matrices. The calculation of $\mathbf{P}(k|N)$ does require multiplications of $n \times n$ matrices.

Matrix $\Phi_p(k+1,k)$ in (21-33) is the plant matrix of the recursive predictor (Lesson 16). It is interesting that the recursive predictor and not the recursive filter plays the predominant role in fixed-interval smoothing. This is further borne out by the appearance of predictor quantities on the right-hand side of (21-24). Observe that (21-25) looks quite similar to a recursive predic-

tor which is excited by the innovations—one which is running in a backward direction.

Finally, note that (21-24) can also be used for $k = N$, in which case its right-hand side reduces to $\hat{\mathbf{x}}(N|N-1) + \mathbf{K}(N)\tilde{\mathbf{z}}(N|N-1)$, which, of course, is $\hat{\mathbf{x}}(N|N)$.

FIXED-POINT SMOOTHING

A fixed-point smoother, $\hat{\mathbf{x}}(k|j)$ where $j = k+1, k+2, \dots$, can be obtained in exactly the same manner as we obtained fixed-interval smoother (21-5). It is obtained from this equation by setting N equal to j and then letting $j = k+1, k+2, \dots$; thus,

$$\hat{\mathbf{x}}(k|j) = \hat{\mathbf{x}}(k|j-1) + \mathbf{B}(j)[\hat{\mathbf{x}}(j|j) - \hat{\mathbf{x}}(j|j-1)] \quad (21-38)$$

where

$$\mathbf{B}(j) = \prod_{i=k}^{j-1} \mathbf{A}(i) \quad (21-39)$$

and $j = k+1, k+2, \dots$. Additionally, one can show that the fixed-point smoothing error-covariance matrix, $\mathbf{P}(k|j)$, is computed from

$$\mathbf{P}(k|j) = \mathbf{P}(k|j-1) + \mathbf{B}(j)[\mathbf{P}(j|j) - \mathbf{P}(j|j-1)]\mathbf{B}'(j) \quad (21-40)$$

where $j = k+1, k+2, \dots$.

Equation (21-38) is impractical from a computational viewpoint, because of the many multiplications of $n \times n$ matrices required first to form the $\mathbf{A}(i)$ matrices and then to form the $\mathbf{B}(j)$ matrices. Additionally, the inverse of matrix $\mathbf{P}(i+1|i)$ is needed in order to compute matrix $\mathbf{A}(i)$. The following results present a “fast” algorithm for computing $\hat{\mathbf{x}}(k|j)$. It is fast in the sense that no multiplications of $n \times n$ matrices are needed to implement it.

Theorem 21-3. A most useful mean-squared fixed-point smoothed estimator of $\mathbf{x}(k)$, $\hat{\mathbf{x}}(k|k+l)$ where $l = 1, 2, \dots$, is given by the expression

$$\begin{aligned}\hat{\mathbf{x}}(k|k+l) &= \hat{\mathbf{x}}(k|k+l-1) \\ &\quad + \mathbf{N}_x(k|k+l)[\mathbf{z}(k+l) - \mathbf{H}(k+l)\hat{\mathbf{x}}(k+l|k+l-1)]\end{aligned}\quad (21-41)$$

where

$$\begin{aligned}\mathbf{N}_x(k|k+l) &= \mathcal{D}_x(k,l)\mathbf{H}'(k+l) \\ &\quad [\mathbf{H}(k+l)\mathbf{P}(k+l|k+l-1)\mathbf{H}'(k+l) + \mathbf{R}(k+l)]^{-1}\end{aligned}\quad (21-42)$$

and

$$\begin{aligned}\mathcal{D}_x(k,l) &= \mathcal{D}_x(k,l-1) \\ &\quad [\mathbf{I} - \mathbf{K}(k+l-1)\mathbf{H}(k+l-1)']\Phi'(k+l,k+l-1)\end{aligned}\quad (21-43)$$

Equations (21-41) and (21-43) are initialized by $\hat{x}(k|k)$ and $\mathcal{D}_x(k,1) = P(k|k)\Phi'(k+1,k)$, respectively. Additionally,

$$P(k|k+l) = P(k|k+l-1) - N_x(k|k+l)[H(k+l)P(k+l|k+l-1) \\ H'(k+l) + R(k+l)]N_x(k|k+l) \quad (21-44)$$

which is initialized by $P(k|k)$. \square

We leave the proof of this useful theorem, which is similar to a result given by Fraser (1967), as an exercise for the reader.

Example 21-2

Here we consider the problem of fixed-point smoothing to obtain a refined estimate of the initial condition for the system described in Example 21-1. Recall that $p(0|0) = 50$ and that by fixed-interval smoothing we had obtained the result $p(0|4) = 16.31$, which is a significant reduction in the uncertainty associated with the initial condition.

Using Equation (21-40) or (21-44) we compute $p(0|1)$, $p(0|2)$, and $p(0|3)$ to be 16.69, 16.32, and 16.31, respectively. Observe that a major reduction in the smoothing error variance occurs as soon as the first measurement is incorporated, and that the improvement in accuracy thereafter is relatively modest. This seems to be a general trait of fixed-point smoothing. \square

Another way to derive fixed-point smoothing formulas is by the following *state augmentation procedure* (Anderson and Moore, 1979). We assume that for $k \geq j$

$$\mathbf{x}_a(k) \triangleq \mathbf{x}(j) \quad (21-45)$$

The state equation for state vector $\mathbf{x}_a(k)$ is

$$\mathbf{x}_a(k+1) = \mathbf{x}_a(k) \quad k \geq j \quad (21-46)$$

It is initialized at $k = j$ by (21-45). Augmenting (21-46) to our basic state-variable model in (15-17) and (15-18), we obtain the following *augmented basic state-variable model*:

$$\begin{pmatrix} \mathbf{x}(k+1) \\ \mathbf{x}_a(k+1) \end{pmatrix} = \begin{pmatrix} \Phi(k+1,k) & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{x}(k) \\ \mathbf{x}_a(k) \end{pmatrix} \\ + \begin{pmatrix} \Psi(k+1,k) \\ \mathbf{0} \end{pmatrix} \mathbf{u}(k) + \begin{pmatrix} \Gamma(k+1,k) \\ \mathbf{0} \end{pmatrix} \mathbf{w}(k) \quad (21-47)$$

and

$$\mathbf{z}(k+1) = (\mathbf{H}(k+1) \mathbf{0}) \begin{pmatrix} \mathbf{x}(k) \\ \mathbf{x}_a(k) \end{pmatrix} + \mathbf{v}(k+1) \quad (21-48)$$

The following two-step procedure can be used to obtain an algorithm for $\hat{x}(j|k)$:

1. Write down the Kalman filter equations for the augmented basic state-

variable model. Anderson and Moore (1979) give these equations for the recursive predictor; i.e., they find

$$\mathbf{E} \left\{ \begin{pmatrix} \mathbf{x}(k+1) \\ \mathbf{x}_a(k+1) \end{pmatrix} \middle| \mathcal{Z}(k) \right\} = \begin{pmatrix} \hat{\mathbf{x}}(k+1|k) \\ \hat{\mathbf{x}}_a(k+1|k) \end{pmatrix} = \begin{pmatrix} \hat{\mathbf{x}}(k+1|k) \\ \hat{\mathbf{x}}(j|k) \end{pmatrix} \quad (21-49)$$

where $k \geq j$. The last equality in (21-49) makes use of (21-46) and (21-45). Observe that $\hat{\mathbf{x}}(j|k)$, the fixed-point smoother of $\mathbf{x}(j)$, has been found as the second component of the recursive predictor for the augmented model.

2. The Kalman filter (or recursive predictor) equations are partitioned in order to obtain the explicit structure of the algorithm for $\hat{\mathbf{x}}(j|k)$. We leave the details of this two-step procedure as an exercise for the reader.

FIXED-LAG SMOOTHING

The earliest attempts to obtain a fixed-lag smoother $\hat{\mathbf{x}}(k|k+L)$ led to an algorithm (e.g., Meditch, 1969), which was later shown to be unstable (Kelly and Anderson, 1971). The following state augmentation procedure leads to a stable fixed-interval smoother for $\hat{\mathbf{x}}(k-L|k)$.

We introduce $L+1$ state vectors, as follows: $\mathbf{x}_1(k+1) = \mathbf{x}(k)$, $\mathbf{x}_2(k+1) = \mathbf{x}(k-1)$, $\mathbf{x}_3(k+1) = \mathbf{x}(k-2)$, ..., $\mathbf{x}_{L+1}(k+1) = \mathbf{x}(k-L)$ [i.e., $\mathbf{x}_i(k+1) = \mathbf{x}(k+1-i)$, $i = 1, 2, \dots, L+1$]. The state equations for these $L+1$ state vectors are

$$\left. \begin{array}{l} \mathbf{x}_1(k+1) = \mathbf{x}(k) \\ \mathbf{x}_2(k+1) = \mathbf{x}_1(k) \\ \mathbf{x}_3(k+1) = \mathbf{x}_2(k) \\ \vdots \\ \mathbf{x}_{L+1}(k+1) = \mathbf{x}_L(k) \end{array} \right\} \quad (21-50)$$

Augmenting (21-50) to our basic state-variable model in (15-17) and (15-18), we obtain yet another *augmented basic state-variable model*:

$$\begin{pmatrix} \mathbf{x}(k+1) \\ \mathbf{x}_1(k+1) \\ \mathbf{x}_2(k+1) \\ \vdots \\ \mathbf{x}_{L+1}(k+1) \end{pmatrix} = \begin{pmatrix} \Phi(k+1,k) & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{x}(k) \\ \mathbf{x}_1(k) \\ \mathbf{x}_2(k) \\ \vdots \\ \mathbf{x}_{L+1}(k) \end{pmatrix} \\ + \begin{pmatrix} \Psi(k+1,k) \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix} \mathbf{u}(k) + \begin{pmatrix} \Gamma(k+1,k) \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix} \mathbf{w}(k) \quad (21-51)$$

The following two-step procedure can be used to obtain an algorithm for $\hat{\mathbf{x}}(k - L|k)$:

1. Write down the Kalman filter equations for the augmented basic state-variable model. Anderson and Moore (1979) give these equations for the recursive predictor; i.e., they find

$$\begin{aligned} \mathbf{E}\{\text{col}(\mathbf{x}(k+1), \mathbf{x}_1(k+1), \dots, \mathbf{x}_{L+1}(k+1)|\mathcal{Z}(k)\}) \\ = \text{col}(\hat{\mathbf{x}}(k+1|k), \hat{\mathbf{x}}_1(k+1|k), \dots, \hat{\mathbf{x}}_{L+1}(k+1|k)) \\ = \text{col}(\hat{\mathbf{x}}(k+1|k), \hat{\mathbf{x}}(k|k), \hat{\mathbf{x}}(k-1|k), \dots, \hat{\mathbf{x}}(k-L|k)) \end{aligned} \quad (21-52)$$

The last equality in (21-52) makes use of the fact that $\mathbf{x}_i(k+1) = \mathbf{x}(k+1-i)$, $i = 1, 2, \dots, L+1$.

2. The Kalman filter (or recursive predictor) equations are partitioned in order to obtain the explicit structure of the algorithm for $\hat{\mathbf{x}}(k - L|k)$. The detailed derivation of the algorithm for $\hat{\mathbf{x}}(k - L|k)$ is left as an exercise for the reader (it can be found in Anderson and Moore, 1979, pp. 177–181).

Some aspects of this fixed-lag smoother are:

1. It is numerically stable, because its stability is determined by the stability of the recursive predictor (i.e., no new feedback loops are introduced into the predictor as a result of the augmentation procedure);
2. In order to compute $\hat{\mathbf{x}}(k - L|k)$, we must also compute the $L - 1$ fixed-lag estimates, $\hat{\mathbf{x}}(k-1|k)$, $\hat{\mathbf{x}}(k-2|k)$, \dots , $\hat{\mathbf{x}}(k-L+1|k)$; this may be costly to do from a computational point of view; and
3. Computation can be reduced by careful coding of the partitioned recursive predictor equations.

P R O B L E M S

- 21-1. Derive the formula for $\hat{\mathbf{x}}(k|N)$ in (21-5) using mathematical induction. Then derive $\hat{\mathbf{x}}(k|N)$ in (21-6).
- 21-2. Prove that $\{\hat{\mathbf{x}}(k|N), k = N, N-1, \dots, 0\}$ is a zero-mean, *second-order* Gauss-Markov process.
- 21-3. Derive the formula for the fixed-point smoothing error-covariance matrix, $\mathbf{P}(k|j)$, given in (21-40).
- 21-4. Prove Theorem 21-3, which gives formulas for a most useful mean-squared fixed-point smoother of $\mathbf{x}(k)$, $\hat{\mathbf{x}}(k|k+l)$, $l = 1, 2, \dots$
- 21-5. Using the two-step procedure described at the end of the section entitled Fixed-Point Smoothing, derive the resulting fixed-point smoother equations.
- 21-6. Using the two-step procedure described at the end of the section entitled

Fixed-Lag Smoothing, derive the resulting fixed-lag smoother equations. Show, by means of a block diagram, that this smoother is stable.

- 21-7. (Meditch, 1969, Exercise 6-13, pg. 245). Consider the scalar system $x(k+1) = 2^{-k}x(k) + w(k)$, $z(k+1) = x(k+1)$, $k = 0, 1, \dots$, where $x(0)$ has mean zero and variance σ_0^2 , and $w(k)$, $k = 0, 1, \dots$ is a zero mean Gaussian white sequence which is independent of $x(0)$ and has a variance equal to q .
 - (a) Assuming that optimal fixed-point smoothing is to be employed to determine $x(0|j)$, $j = 1, 2, \dots$, what is the equation for the appropriate smoothing filter?
 - (b) What is the limiting value of $p(0|j)$ as $j \rightarrow \infty$?
 - (c) How does this value compare with $p(0|0)$?

State Estimation: Smoothing Applications

INTRODUCTION

In this lesson we present some applications that illustrate interesting numerical and theoretical aspects of fixed-interval smoothing. These applications are taken from the field of digital signal processing.

MINIMUM-VARIANCE DECONVOLUTION (MVD)

Here, as in Examples 2-6 and 14-1, we begin with the convolutional model

$$z(k) = \sum_{i=1}^k \mu(i)h(k-i) + v(k), \quad k = 1, 2, \dots, N \quad (22-1)$$

Recall that deconvolution is the signal processing procedure for removing the effects of $h(j)$ and $v(j)$ from the measurements so that one is left with an estimate of $\mu(j)$. Here we shall obtain a useful algorithm for a mean-squared fixed-interval estimator of $\mu(j)$.

To begin, we must convert (22-1) into an equivalent state-variable model.

Theorem 22-1 (Mendel, 1983a, pp. 13–14). *The single-channel state-variable model*

$$\mathbf{x}(k+1) = \Phi\mathbf{x}(k) + \gamma\mu(k) \quad (22-2)$$

$$z(k) = \mathbf{h}'\mathbf{x}(k) + v(k) \quad (22-3)$$

Minimum-Variance Deconvolution (MVD)

is equivalent to the convolutional sum model in (22-1) when $\mathbf{x}(0) = \mathbf{0}$, $\mu(0) = 0$, $\mathbf{h}(0) = 0$, and

$$h(l) = \mathbf{h}'\Phi^{l-1}\gamma, \quad l = 1, 2, \dots \quad (22-4)$$

Proof. Iterate (22-2) and substitute the results into (22-3). Compare the resulting equation with (22-1) to see that, under the conditions $\mathbf{x}(0) = \mathbf{0}$, $\mu(0) = 0$ and $\mathbf{h}(0) = 0$, they are the same. \square

The condition $\mathbf{x}(0) = \mathbf{0}$ merely initializes our state-variable model. The condition $\mu(0) = 0$ means there is no input at time zero. The coefficients in (22-4) represent sampled values of the impulse response. If we are given impulse response data $\{h(1), h(2), \dots, h(L)\}$ then we can determine matrices Φ , γ and \mathbf{h} as well as system order n by applying an approximate realization procedure, such as Kung's (1978), to $\{h(1), h(2), \dots, h(L)\}$. Additionally, if $h(0) \neq 0$ it is simple to modify Theorem 22-1.

In Example 14-1 we obtained a rather unwieldy formula for $\hat{\mu}_{MS}(N)$. Note that, in terms of our conditioning notation, the elements of $\hat{\mu}_{MS}(N)$ are $\hat{\mu}_{MS}(k|N)$, $k = 1, 2, \dots, N$. We now obtain a very useful algorithm for $\hat{\mu}_{MS}(k|N)$. For notational convenience, we shorten $\hat{\mu}_{MS}$ to $\hat{\mu}$.

Theorem 22-2 (Mendel, 1983a, pp. 68–70)

- a. *A two-pass fixed-interval smoother for $\mu(k)$ is*

$$\hat{\mu}(k|N) = q(k)\gamma'r(k+1|N) \quad (22-5)$$

where $k = N - 1, N - 2, \dots, 1$.

- b. *The smoothing error variance, $\sigma_{\mu}^2(k|N)$, is*

$$\sigma_{\mu}^2(k|N) = q(k) - q(k)\gamma'S(k+1|N)\gamma q(k) \quad (22-6)$$

where $k = N - 1, N - 2, \dots, 1$. In these formulas $r(k|N)$ and $S(k|N)$ are computed using (21-25) and (21-27), respectively, and $E\{\mu^2(k)\} = q(k)$ [here $q(k)$ denotes the variance of $\mu(k)$, and should not be confused with the event sequence; which appears in the product model for $\mu(k)$].

Proof a. To begin, we apply the fundamental theorem of estimation theory, Theorem 13-1, to (22-2). We operate on both sides of that equation with $E\{\cdot|\mathcal{E}(N)\}$, to show that

$$\gamma\hat{\mu}(k|N) = \hat{\mathbf{x}}(k+1|N) - \Phi\hat{\mathbf{x}}(k|N) \quad (22-7)$$

By performing appropriate manipulations on this equation we can derive (22-5) as follows. Substitute $\hat{\mathbf{x}}(k|N)$ and $\hat{\mathbf{x}}(k+1|N)$ from Equation (21-24) into Equation (22-7), to see that

$$\begin{aligned}\gamma\hat{\mu}(k|N) &= \hat{x}(k+1|k) + P(k+1|k)r(k+1|N) \\ &\quad - \Phi[\hat{x}(k|k-1) + P(k|k-1)r(k|N)] \\ &= \hat{x}(k+1|k) - \Phi\hat{x}(k|k-1) \\ &\quad + P(k+1|k)r(k+1|N) - \Phi P(k|k-1)r(k|N)\end{aligned}\quad (22-8)$$

Applying (17-11) and (16-4) to the state-variable model in (22-2) and (22-3), it is straightforward to show that

$$\hat{x}(k+1|k) = \Phi\hat{x}(k|k-1) + \Phi K(k)\tilde{z}(k|k-1) \quad (22-9)$$

hence, (22-8) reduces to

$$\begin{aligned}\gamma\hat{\mu}(k|N) &= \Phi K(k)\tilde{z}(k|k-1) + P(k+1|k)r(k+1|N) \\ &\quad - \Phi P(k|k-1)r(k|N)\end{aligned}\quad (22-10)$$

Next, substitute (21-25) into (22-10), to show that

$$\begin{aligned}\gamma\hat{\mu}(k|N) &= \Phi K(k)\tilde{z}(k|k-1) + P(k+1|k)r(k+1|N) \\ &\quad - \Phi P(k|k-1)\Phi'_p(k+1,k)r(k+1|N) \\ &\quad - \Phi P(k|k-1)\mathbf{h}'[\mathbf{h}'P(k|k-1)\mathbf{h} + r]^{-1}\tilde{z}(k|k-1)\end{aligned}\quad (22-11)$$

Making use of Equation (17-12) for $K(k)$, we find that the first and last terms in Equation (22-11) are identical; hence,

$$\begin{aligned}\gamma\hat{\mu}(k|N) &= P(k+1|k)r(k+1|N) - \Phi P(k|k-1) \\ &\quad \Phi'_p(k+1,k)r(k+1|N)\end{aligned}\quad (22-12)$$

Combine Equations (17-13) and (17-14) to see that

$$P(k+1|k) = \Phi P(k|k-1)\Phi'_p(k+1,k) + \gamma q(k)\gamma' \quad (22-13)$$

Finally, substitute (22-13) into Equation (22-12) to observe that

$$\gamma\hat{\mu}(k|N) = \gamma q(k)\gamma' r(k+1|N)$$

which has the unique solution given by

$$\hat{\mu}(k|N) = q(k)\gamma' r(k+1|N) \quad (22-14)$$

which is Equation (22-5).

- b. To derive Equation (22-6) we use (22-5) and the definition of estimation error $\tilde{\mu}(k|N)$,

$$\tilde{\mu}(k|N) = \mu(k) - \hat{\mu}(k|N) \quad (22-15)$$

to form

$$\mu(k) = \tilde{\mu}(k|N) + q(k)\gamma' r(k+1|N) \quad (22-16)$$

Taking the variance of both sides of (22-16), and using the orthogonality condition

$$E\{\tilde{\mu}(k|N)r(k+1|N)\} = \mathbf{0} \quad (22-17)$$

we see that

$$q(k) = \sigma_\mu^2(k|N) + q(k)\gamma'S(k+1|N)\gamma q(k) \quad (22-18)$$

which is equivalent to (22-6). \square

Observe, from (22-5) and (22-6), that $\hat{\mu}(k|N)$ and $\sigma_\mu^2(k|N)$ are easily computed once $r(k|N)$ and $S(k|N)$ have been computed.

The extension of Theorem 22-2 to a time-varying state-variable model is straightforward and can be found in Mendel (1983).

Example 22-1

In this example we compute $\hat{\mu}(k|N)$, first for a broadband channel IR, $h_1(k)$, and then for a narrower-band channel IR, $h_2(k)$. The transfer functions of these channel models are

$$H_1(z) = \frac{-0.76286z^3 + 1.5884z^2 - 0.82356z + 0.000222419}{z^4 - 2.2633z^3 + 1.77734z^2 - 0.49803z + 0.045546} \quad (22-19)$$

and

$$H_2(z) = \frac{0.0378417z^3 - 0.0306517z}{z^4 - 3.4016497z^3 + 4.5113732z^2 - 2.7553363z + 0.6561} \quad (22-20)$$

respectively. Plots of these IR's and their squared amplitude spectra are depicted in Figures 22-1 and 22-2.

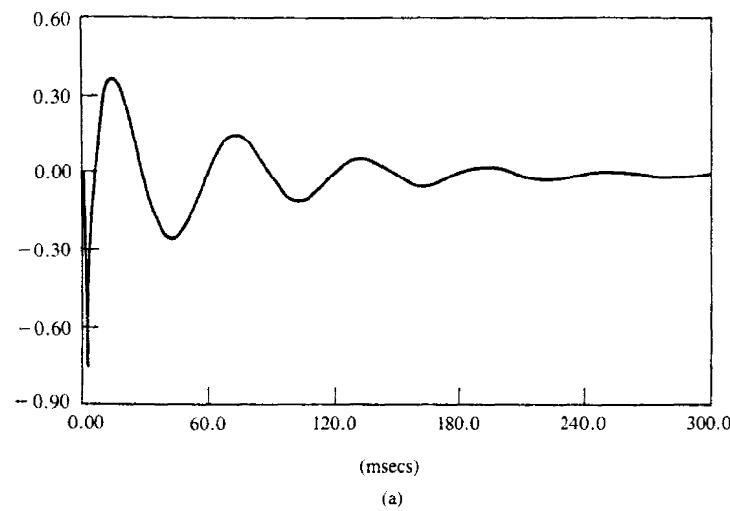
Measurements, $z(k)$ ($k = 1, 2, \dots, 250$, where $T = 3$ msec), were generated by convolving each of these IR's with a sparse spike train (i.e., a Bernoulli-Gaussian sequence) and then adding measurement noise to the results. These measurements, which, of course, represent the starting point for deconvolution, are depicted in Figure 22-3.

Figure 22-4 depicts $\hat{\mu}(k|N)$. Observe that much better results are obtained for the broadband channel than for the narrower-band channel, even though data quality, as measured by \overline{SNR} , is much lower in the former case. The MVD results for the narrower-band channel appear "smeared out," whereas the MVD results for the broadband channel are quite sharp. We provide a theoretical explanation for this effect below.

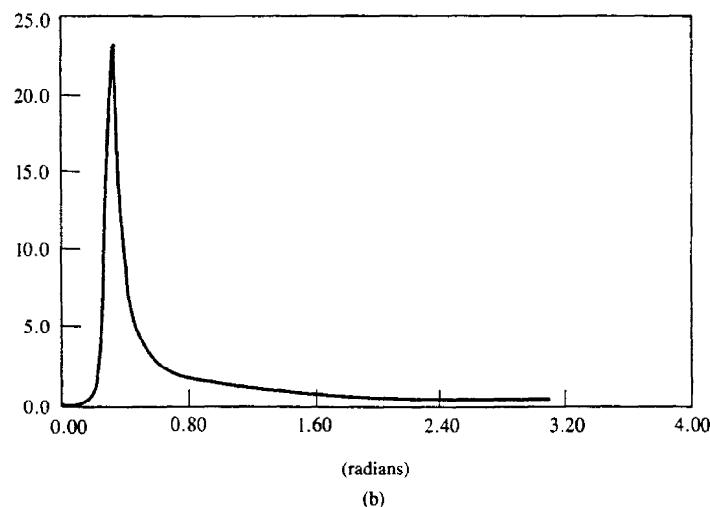
Observe, also, that $\hat{\mu}(k|N)$ tends to undershoot $\mu(k)$. See Chi and Mendel (1984) for a theoretical explanation about why this occurs. \square

STEADY-STATE MVD FILTER

For a time-invariant IR and stationary noises, the Kalman gain matrix, as well as the error-covariance matrices, will reach steady-state values. When this occurs, both the Kalman innovations filter and anticausal μ -filter [(22-5) and (21-25)] become time invariant, and we then refer to the MVD filter as a *steady-state MVD filter*. In this section we examine an important property of this steady-state filter.

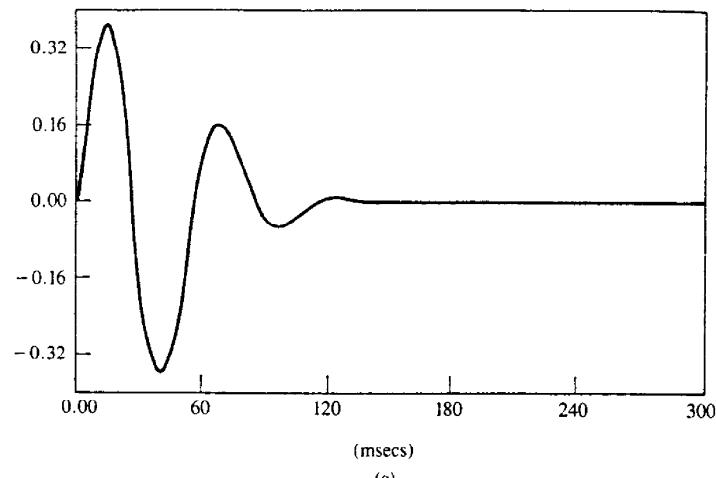


(a)

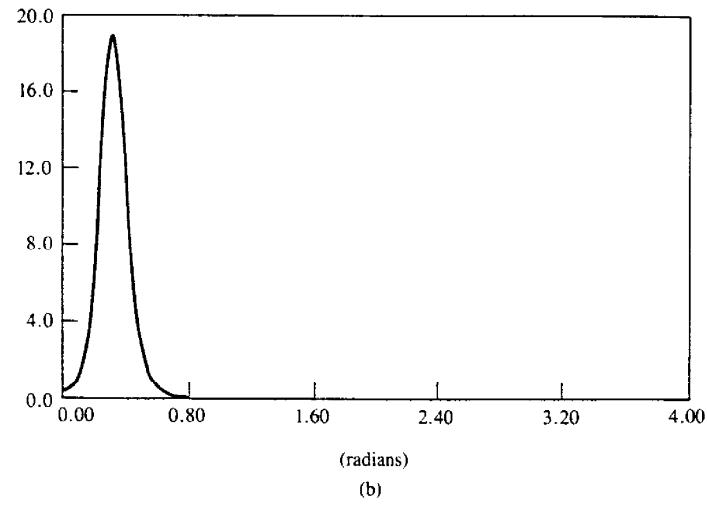


(b)

Figure 22-1 (a) Fourth-order broad-band channel IR, and (b) its squared amplitude spectrum (Chi, 1983).



(a)



(b)

Figure 22-2 (a) Fourth-order narrower-band channel, IR, and (b) its squared amplitude spectrum (Chi and Mendel, 1984, © 1984, IEEE).

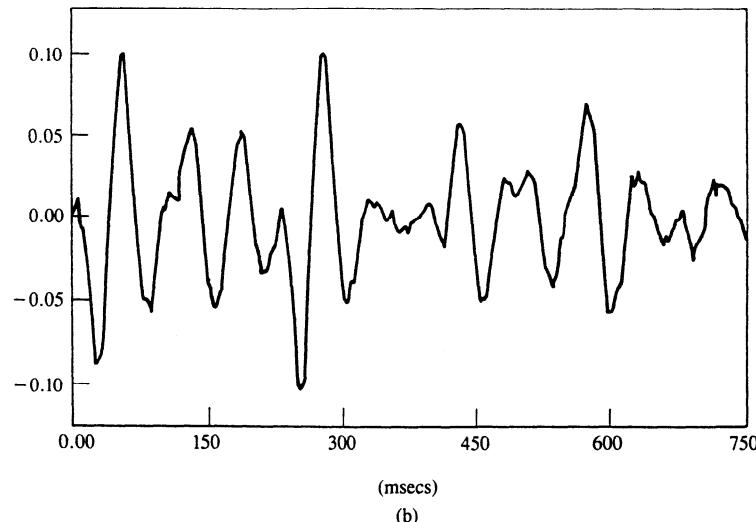
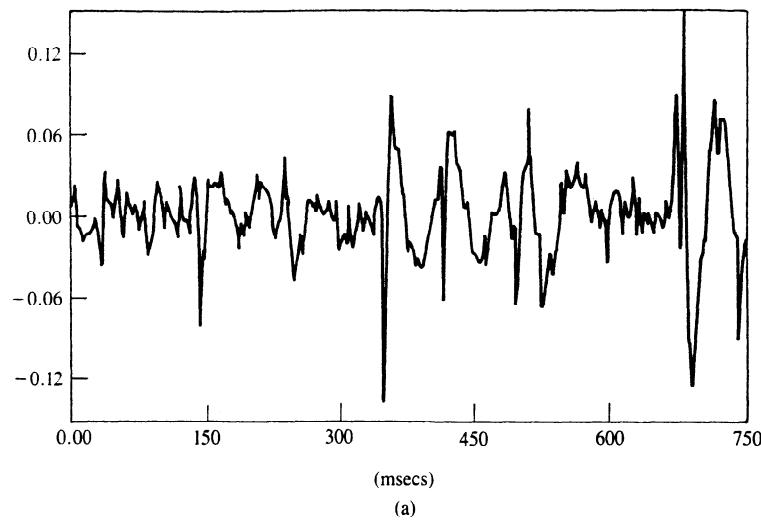


Figure 22-3 Measurements associated with (a) broad-band channel ($\bar{SNR} = 10$) and (b) narrower-band channel ($\bar{SNR} = 100$), (Chi and Mendel, 1984, © 1984, IEEE).

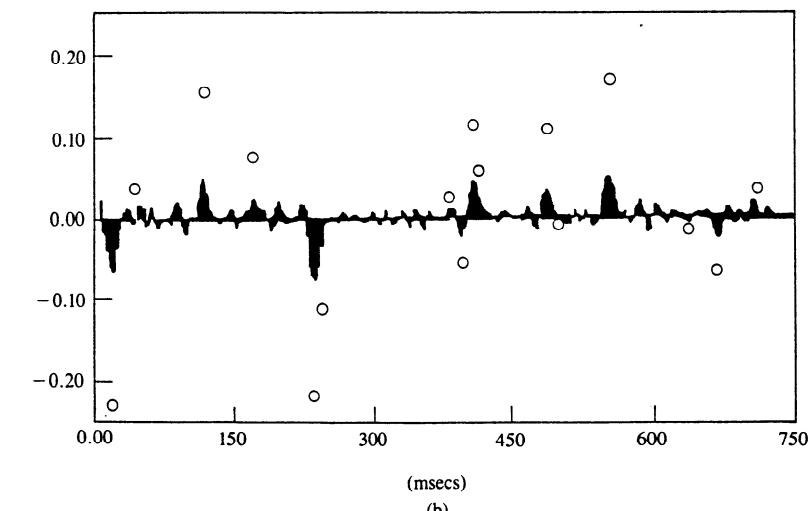
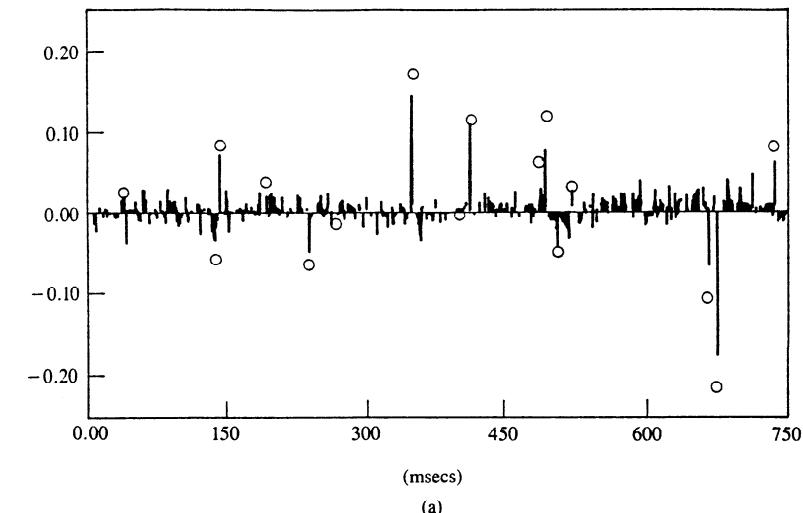


Figure 22-4 $\hat{\mu}(k|N)$ for (a) broadband channel ($\bar{SNR} = 10$) and (b) narrower-band channel ($\bar{SNR} = 100$). Circles depict true $\mu(k)$ and bars depict estimate of $\mu(k)$, (Chi and Mendel, 1984, © 1984, IEEE).

Let $h_i(k)$ and $h_\mu(k)$ denote the IR's of the steady-state Kalman innovations and anticausal μ -filters, respectively. Then,

$$\begin{aligned}\hat{\mu}(k|N) &= h_\mu(k)*\tilde{z}(k|k-1) \\ &= h_\mu(k)*h_i(k)*z(k) \\ &= h_\mu(k)*h_i(k)*h(k)*\mu(k) \\ &\quad + h_\mu(k)*h_i(k)*v(k)\end{aligned}\quad (22-21)$$

which can also be expressed as

$$\hat{\mu}(k|N) = \hat{\mu}_s(k|N) + n(k|N) \quad (22-22)$$

where the *signal component* of $\hat{\mu}(k|N)$, $\hat{\mu}_s(k|N)$, is

$$\hat{\mu}_s(k|N) = h_\mu(k)*h_i(k)*h(k)*\mu(k) \quad (22-23)$$

and the *noise component* of $\hat{\mu}(k|N)$, $n(k|N)$, is

$$n(k|N) = h_\mu(k)*h_i(k)*v(k) \quad (22-24)$$

We shall refer to $h_\mu(k)*h_i(k)$ as the IR of the MVD filter, $h_{MV}(k)$, i.e.,

$$h_{MV}(k) = h_\mu(k)*h_i(k) \quad (22-25)$$

The following result has been proven by Chi and Mendel (1984) for the slightly modified model $x(k+1) = \Phi x(k) + \gamma \mu(k+1)$ and $z(k) = \mathbf{h}'x(k) + v(k)$ [because of the $\mu(k+1)$ input instead of the $\mu(k)$ input, $h(0) \neq 0$].

Theorem 22-3. *In the stationary case:*

a. *The Fourier transform of $h_{MV}(k)$ is*

$$H_{MV}(\omega) = \frac{qH^*(\omega)}{q|H(\omega)|^2 + r} \quad (22-26)$$

where $H^*(\omega)$ denotes the complex conjugate of $H(\omega)$; and

b. *the signal component of $\hat{\mu}(k|N)$, $\hat{\mu}_s(k|N)$, is given by*

$$\hat{\mu}_s(k|N) = R(k)*\mu(k) \quad (22-27)$$

where $R(k)$ is the auto-correlation function

$$R(k) = \frac{q}{\eta} [h(k)*h_i(k)]*[h(-k)*h_i(-k)] \quad (22-28)$$

in which

$$\eta = \mathbf{h}'\bar{\mathbf{P}}\mathbf{h} + r \quad (22-29)$$

additionally,

$$R(\omega) = \frac{q|H(\omega)|^2}{q|H(\omega)|^2 + r} \quad \square \quad (22-30)$$

We leave the proof of this theorem as an exercise for the reader. Observe that part (b) of the theorem means that $\hat{\mu}_s(k|N)$ is a zero-phase wave-shaped version of $\mu(k)$. Observe, also, that $R(\omega)$ can be written as

$$R(\omega) = \frac{|H(\omega)|^2 q / r}{1 + |H(\omega)|^2 q / r} \quad (22-31)$$

which demonstrates that q/r , and subsequently $\overline{\text{SNR}}$, is an MVD filter tuning parameter. As $q/r \rightarrow \infty$, $R(\omega) \rightarrow 1$ so that $R(k) \rightarrow \delta(k)$; thus, for high signal-to-noise ratios $\hat{\mu}_s(k|N) \rightarrow \mu(k)$. Additionally, when $|H(\omega)|^2 q / r \gg 1$, $R(\omega) \rightarrow 1$, and once again $R(k) \rightarrow \delta(k)$. Broadband IRs often satisfy this condition. In general, however, $\hat{\mu}_s(k|N)$ is a smeared-out version of $\mu(k)$; however, the nature of the smearing is quite dependent on the bandwidth of $h(k)$ and $\overline{\text{SNR}}$.

Example 22-2

This example is a continuation of Example 22-1. Figure 22-5 depicts $R(k)$ for both the broadband and narrower-band IRs, $h_1(k)$ and $h_2(k)$, respectively. As predicted by (22-31), $R_1(k)$ is much spikier than $R_2(k)$, which explains why the MVD results for the broadband IR are quite sharp, whereas the MVD results for the narrower-band IR are smeared out. Note, also, the difference in peak amplitudes for $R_1(k)$ and $R_2(k)$. This explains why $\hat{\mu}(k|N)$ underestimates the true values of $\mu(k)$ by such large amounts in the narrower-band case (see Figs. 22-4a and b). \square

RELATIONSHIP BETWEEN STEADY-STATE MVD FILTER AND AN INFINITE IMPULSE RESPONSE DIGITAL WIENER DECONVOLUTION FILTER

We have seen that an MVD filter is a cascade of a causal Kalman innovations filter and an anticausal μ -filter; hence, it is a noncausal filter. Its impulse response extends from $k = -\infty$ to $k = +\infty$, and the IR of the steady-state MVD filter is given in the time-domain by $h_{MV}(k)$ in (22-25), or in the frequency domain by $H_{MV}(\omega)$ in (22-26).

There is a more direct way for designing an IIR minimum mean-squared error deconvolution filter, i.e., an *IIR digital Wiener deconvolution filter*, as we describe next.

We return to the situation depicted in Figure 19-4, but now we assume that: Filter $F(z)$ is an IIR filter, with coefficients $\{f(j), j = 0, \pm 1, \pm 2, \dots\}$;

$$d(k) = \mu(k) \quad (22-32)$$

where $\mu(k)$ is a white noise sequence; $\mu(k)$, $v(k)$, and $n(k)$ are stationary; and, $\mu(k)$ and $v(k)$ are uncorrelated.

In this case, (19-39) becomes

$$\sum_{i=-\infty}^{\infty} f(i)\phi_{zz}(i-j) = \phi_{z\mu}(j), \quad j = 0, \pm 1, \pm 2, \dots \quad (22-33)$$

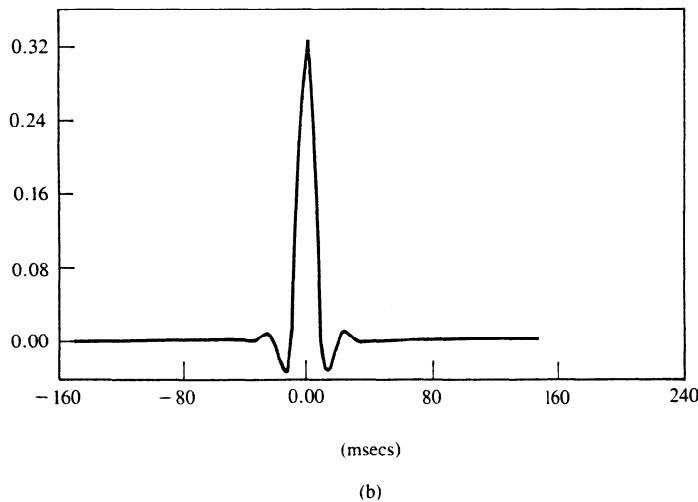
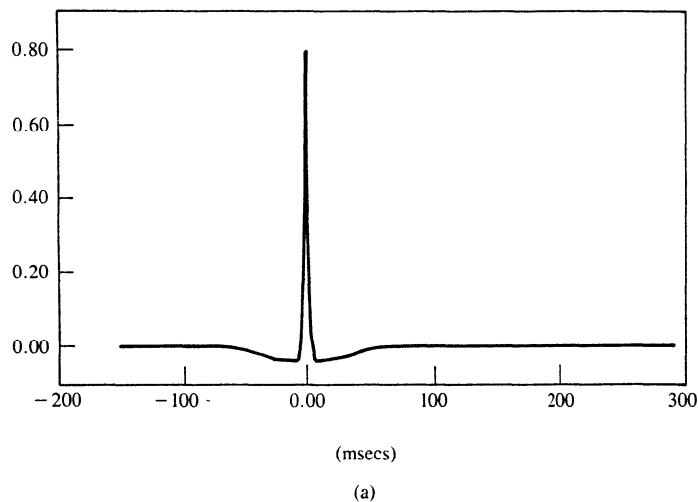


Figure 22-5 $R(k)$ for (a) broadband channel ($\overline{SNR} = 100$) and (b) narrower-band channel ($\overline{SNR} = 100$), (Chi and Mendel, 1984, © 1984, IEEE).

Using (22-1), the whiteness of $\mu(k)$ and the assumptions that $\mu(k)$ and $\nu(k)$ are uncorrelated and stationary, it is straightforward to show that

$$\phi_{z\mu}(j) = qh(-j) \quad (22-34)$$

Substituting (22-34) into (22-33), we have

$$\sum_{i=-\infty}^{\infty} f(i)\phi_{zz}(i-j) = qh(-j), \quad j = 0, \pm 1, \pm 2, \dots \quad (22-35)$$

Taking the discrete-time Fourier transform of (22-35), we see that

$$F(\omega)\Phi_{zz}(\omega) = qH^*(\omega) \quad (22-36)$$

but, from (22-1), we also know that

$$\Phi_{zz}(\omega) = q|H(\omega)|^2 + r \quad (22-37)$$

Substituting (22-37) into (22-36), we determine $F(\omega)$, as

$$F(\omega) = \frac{qH^*(\omega)}{q|H(\omega)|^2 + r} \quad (22-38)$$

This IIR digital Wiener deconvolution filter (i.e., two-sided least-squares inverse filter) was, to the best of our knowledge, first derived by Berkhouit (1977).

Theorem 22-4 (Chi and Mendel, 1984). *The steady-state MVD filter, whose IR is given by $h_{MV}(k)$, is exactly the same as Berkhouit's IIR digital Wiener deconvolution filter.* \square

The steady-state MVD filter is a recursive implementation of Berkhouit's infinite-length filter. Of course, the MVD filter is also applicable to time-varying and nonstationary systems, whereas his filter is not.

MAXIMUM-LIKELIHOOD DECONVOLUTION

In Example 14-2 we began with the deconvolution linear model $\mathcal{Z}(N) = \mathcal{H}(N-1)\mu + \mathcal{V}(N)$, used the product model for μ (i.e., $\mu = Q_q r$), and showed that a separation principle exists for the determination of \hat{r}_{MAP} and \hat{q}_{MAP} . We showed that first one must determine \hat{q}_{MAP} , after which \hat{r}_{MAP} can be computed using (14-57). We repeat (14-57) here for convenience,

$$\begin{aligned} \hat{r}_{MAP}(N|\hat{q}) &= \sigma_r^2 Q_q \mathcal{H}'(N-1) \\ &\quad [\sigma_r^2 \mathcal{H}(N-1) Q_q \mathcal{H}'(N-1) + \rho I]^{-1} \mathcal{Z}(N) \end{aligned} \quad (22-39)$$

where \hat{q} is short for \hat{q}_{MAP} . In terms of our conditioning notation used in state estimation, the elements of $\hat{r}_{MAP}(N|\hat{q})$ are $\hat{r}_{MAP}(k|N; \hat{q})$, $k = 1, 2, \dots, N$. Equation (22-39) is terribly unwieldy because of the $N \times N$ matrix,

$\sigma^2 \mathcal{H}(N-1) Q_q \mathcal{H}'(N-1) + \rho I$, that must be inverted. The following theorem provides a more practical way to compute $\hat{r}_{MAP}(k|N; \hat{\mathbf{q}})$.

Theorem 22-5 (Mendel, 1983b). *Unconditional maximum-likelihood (i.e., MAP) estimates of \mathbf{r} can be obtained by applying MVD formulas to the state-variable model*

$$\mathbf{x}(k+1) = \Phi \mathbf{x}(k) + \gamma \hat{q}_{MAP}(k) r(k) \quad (22-40)$$

$$z(k) = \mathbf{h}' \mathbf{x}(k) + v(k) \quad (22-41)$$

where $\hat{q}_{MAP}(k)$ is a MAP estimate of $q(k)$.

Proof. Example 14-2 showed that a MAP estimate of \mathbf{q} can be obtained prior to finding a MAP estimate of \mathbf{r} . By using the product model for $\mu(k)$, and \hat{q}_{MAP} , our state-variable model in (22-2) and (22-3) can be expressed as in (22-40) and (22-41). Applying (14-41) to this system, we see that

$$\hat{r}_{MAP}(k|N) = \hat{r}_{MS}(k|N) \quad (22-42)$$

but, by comparing (22-40) and (22-2), and (22-41) and (22-3), we see that $\hat{r}_{MS}(k|N)$ can be found from the MVD algorithm in Theorem 22-2 in which we replace $\mu(k)$ by $r(k)$ and set $q(k) = \sigma_r^2 \hat{q}_{MAP}(k)$. \square

RECURSIVE WAVESHAPING

In Lesson 19 we described the design of a FIR waveshaping filter (e.g., see Figure 19-4). In this section we shall develop a recursive waveshaping filter in the framework of state-variable models and mean-squared estimation theory. Other approaches to the design of recursive waveshaping filters have been given by Shanks (1967) and Aguilara, et al. (1970).

We direct our attention at the situation depicted in Figure 22-6. To

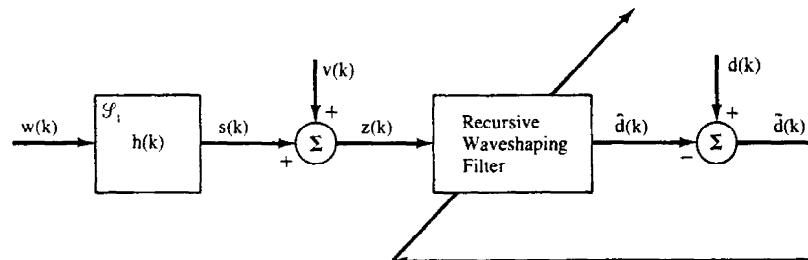


Figure 22-6 Waveshaping problem studied in this section. Information about $\tilde{d}(k)$, not necessarily $\tilde{d}(k)$ itself, is used to drive the (time-varying) recursive waveshaping filter. (Mendel, 1983a, © 1983, IEEE).

begin, we must obtain the following state-variable models for $h(k)$ and $d(k)$ [i.e., we must use an approximate realization procedure (Kung, 1978) or any other viable technique to map $\{h(i), i = 0, 1, \dots, I_1\}$ into $\{\Phi_1, \gamma_1, \mathbf{h}_1\}$, and $\{d(i), i = 0, 1, \dots, I_2\}$ into $\{\Phi_2, \gamma_2, \mathbf{h}_2\}$]:

$$\mathbf{x}_1(k+1) = \Phi_1 \mathbf{x}_1(k) + \gamma_1 \delta(k) \quad (22-43)$$

$$h(k) = \mathbf{h}'_1 \mathbf{x}_1(k) \quad (22-44)$$

and

$$\mathbf{x}_2(k+1) = \Phi_2 \mathbf{x}_2(k) + \gamma_2 \delta(k) \quad (22-45)$$

$$d(k) = \mathbf{h}'_2 \mathbf{x}_2(k) \quad (22-46)$$

State vectors \mathbf{x}_1 and \mathbf{x}_2 are $n_1 \times 1$ and $n_2 \times 1$, respectively. Signal $\delta(k)$ is the unit spike.

In the stochastic situation depicted in Figure 22-6, where $h(k)$ is excited by the white sequence $w(k)$ and noise, $v(k)$, corrupts $s(k)$, the best we can possibly hope to achieve by waveshaping is to make $z(k) = w(k)*h(k) + v(k)$ look like $w(k)*d(k)$ (Figure 22-7). This is because both $h(k)$ and $d(k)$ must be excited by the same random input, $w(k)$, for the waveshaping problem in this situation to be well posed. The state-variable model for this situation is

$$\mathcal{G}_1: \begin{cases} \mathbf{x}_1(k+1) = \Phi_1 \mathbf{x}_1(k) + \gamma_1 w(k) \\ z(k) = \mathbf{h}'_1 \mathbf{x}_1(k) + v(k) \end{cases} \quad (22-47)$$

$$(22-48)$$

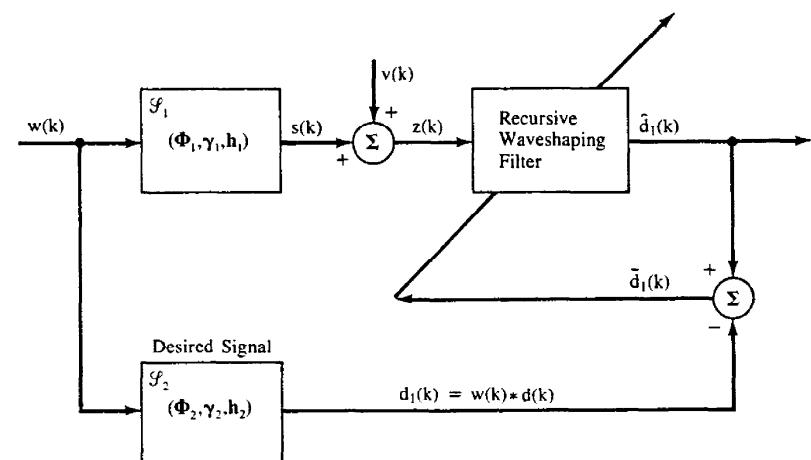


Figure 22-7 State-variable formulation of recursive waveshaping problem (Mendel, 1983a, © 1983, IEEE).

and

$$\begin{aligned} \mathcal{S}_2: \quad & \mathbf{x}_2(k+1) = \Phi_2 \mathbf{x}_2(k) + \gamma_2 w(k) \\ & d_1(k) = \mathbf{h}_2^T \mathbf{x}_2(k) \end{aligned} \quad (22-49) \quad (22-50)$$

Observe that both \mathcal{S}_1 and \mathcal{S}_2 are excited by the same input, $w(k)$. Additionally, $w(k)$ and $v(k)$ are zero-mean mutually uncorrelated white noise sequences, for which

$$E\{w^2(k)\} = q \quad (22-51)$$

and

$$E\{v^2(k)\} = r \quad (22-52)$$

We now proceed to formulate the recursive waveshaping filter design problem in the context of mean-squared estimation theory. Our *filter design problem* is: Given the measurements $z(1), z(2), \dots, z(j)$ determine an estimator $\hat{d}_1(k|j)$ such that the mean-squared error

$$J[\hat{d}_1(k|j)] = E\{\hat{d}_1(k|j)\}^2 \quad (22-53)$$

is minimized. The solution to this problem is given next.

Theorem 22-6 (Structure of Minimum-Variance Waveshaper, Dai and Mendel, 1986). *The Minimum-Variance Waveshaping filter consists of two components 1. stochastic inversion, and 2. waveshaping.*

Proof. According to the Fundamental Theorem of Estimation Theory, Theorem 13-1, the unbiased, minimum-variance estimator of $d_1(k)$ based on the measurements $\{z(1), z(2), \dots, z(j)\}$ is

$$\hat{d}_1(k|j) = E\{d_1(k)|\mathcal{Z}(j)\} \quad (22-54)$$

where $\mathcal{Z}(j) = \text{col}(z(1), z(2), \dots, z(j))$. Observe, from Figure 22-7, that

$$d_1(k) = w(k)*d(k) \quad (22-55)$$

hence,

$$\begin{aligned} \hat{d}_1(k|j) &= E\{w(k)*d(k)|\mathcal{Z}(j)\} \\ &= E\{w(k)|\mathcal{Z}(j)\}*d(k) = \hat{w}(k|j)*d(k) \end{aligned} \quad (22-56)$$

Equation (22-56) tells us that there are two steps to obtain $\hat{d}_1(k|j)$:

1. first obtain $\hat{w}(k|j)$, and
2. then convolve the desired signal with $\hat{w}(k|j)$.

Step 1 removes the effects of the original wavelet and noise from the mea-

surements. It is the problem of stochastic inversion and can be performed by means of minimum-variance deconvolution. \square

In this book we have only discussed fixed-interval MVD, from which we obtain $\hat{w}(k|N)$. Fixed point algorithms are also available (e.g., Mendel, 1983).

Theorem 22-7 (Recursive waveshaping, Mendel, 1983a, pg. 600). *Let $\hat{w}(k|N)$ denote the fixed-interval estimate of $w(k)$, which can be obtained from S_1 via minimum-variance deconvolution (MVD), (Theorem 22-2). Then $\hat{d}_1(k|N)$ is obtained from the waveshaping filter*

$$\hat{d}_1(k|N) = \mathbf{h}_2^T \hat{\mathbf{x}}_2(k|N) \quad (22-57)$$

$$\hat{\mathbf{x}}_2(k+1|N) = \Phi_2 \hat{\mathbf{x}}_2(k|N) + \gamma_2 \hat{w}(k|N) \quad (22-58)$$

where $k = 0, 1, \dots, N-1$, and $\hat{\mathbf{x}}_2(0|N) = \mathbf{0}$. \square

We leave the proof of this theorem as an exercise for the reader.

Some observations about Theorems 22-6 and 22-7 are in order. First, MVD is analogous to solving a stochastic inverse problem; hence, an MVD filter can be thought of as an optimal inverse filter. Second, if $w(k)$ was deterministic and $v(k) = 0$, then our intuition tells us that the recursive waveshaping filter should consist of the following two distinct components: an inverse filter, to remove the effects of $H(z)$, followed by a waveshaping filter, whose transfer function is $D(z)$. The transfer function of the recursive waveshaping filter would then be $\frac{1}{H(z)} D(z)$, so that $D(z) = H(z) \left[\frac{D(z)}{H(z)} \right] = D(z)$.

Finally, the results in these theorems support our intuition even in the stochastic case; for, $\hat{d}_1(k|N)$, for example, is also obtained in two steps. As shown in Figure 22-8, first $\hat{w}(k|N)$ is obtained via MVD. Then this signal is reshaped to give $\hat{d}_1(k|N)$.

Note, also, that in order to compute $\hat{d}_1(k|N)$ it is not really necessary to use the state-variable model for $\hat{d}_1(k|N)$. Observe, from (22-56), that

$$\hat{d}_1(k|N) = d(k)*\hat{w}(k|N) \quad (22-59)$$

Example 22-3

In this example we describe a simulation study for the Bernoulli-Gaussian input sequence [i.e., $w(k)$] depicted in Figure 22-9. When this sequence is convolved with the fourth-order IR depicted in Figure 22-1a and noise is added to the result we obtain measurements $z(k)$, $k = 1, 2, \dots, 1000$, depicted in Figure 22-10.

In Figure 22-11 we see $\hat{w}(k|N)$. Observe that the large spikes in $w(k)$ have been estimated quite well. When $\hat{w}(k|N)$ is convolved with a first-order decaying exponential [$d_1(t) = e^{-300t}$], we obtain the shaped signal depicted in Figure 22-12. Some “smoothing” of the data occurs when $\hat{w}(k|N)$ is convolved with the exponential $d_1(k)$.

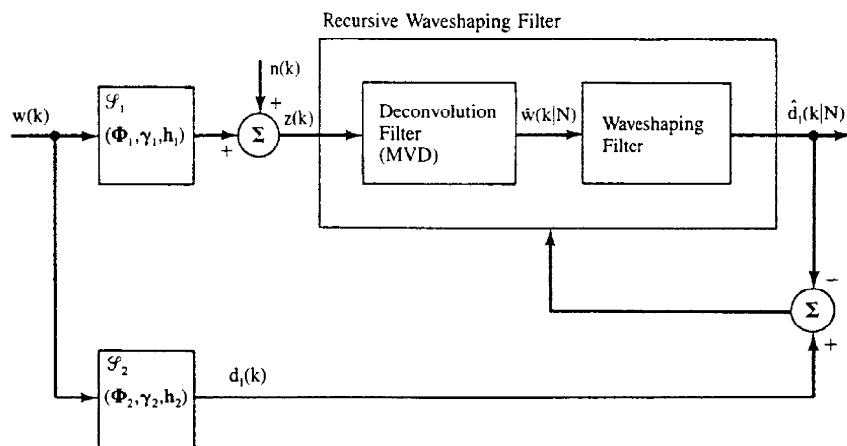


Figure 22-8 Details of recursive linear waveshaping filter. (Mendel, 1983a, © 1983 IEEE.)

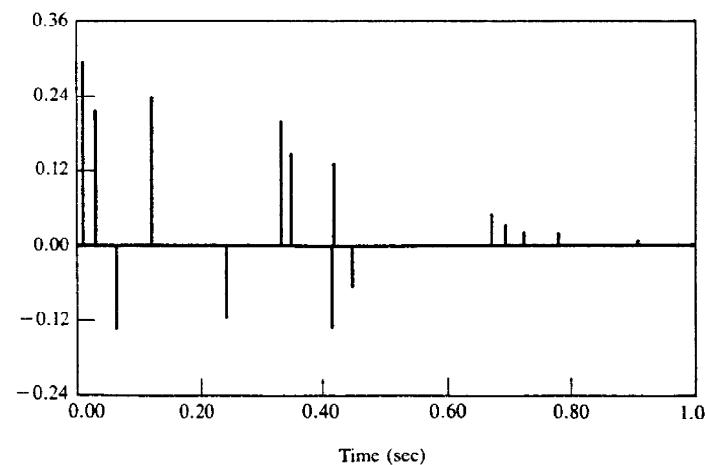


Figure 22-9 Bernoulli-Gaussian input sequence (Mendel, 1983a, © 1983, IEEE).

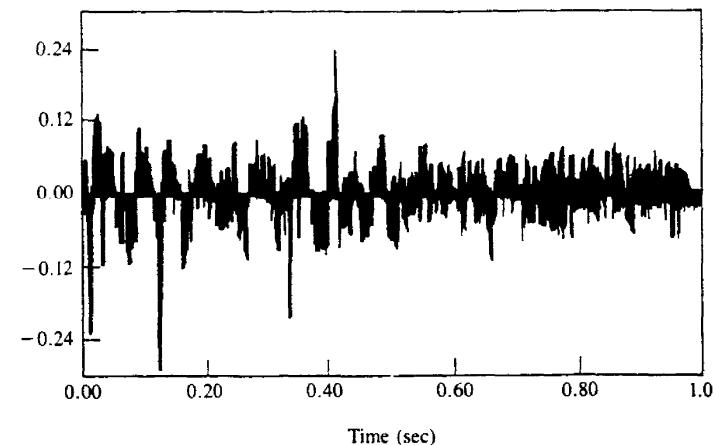


Figure 22-10 Noise-corrupted signal $z(k)$. Signal-to-noise ratio chosen equal to ten (Mendel, 1983a, © 1983, IEEE).

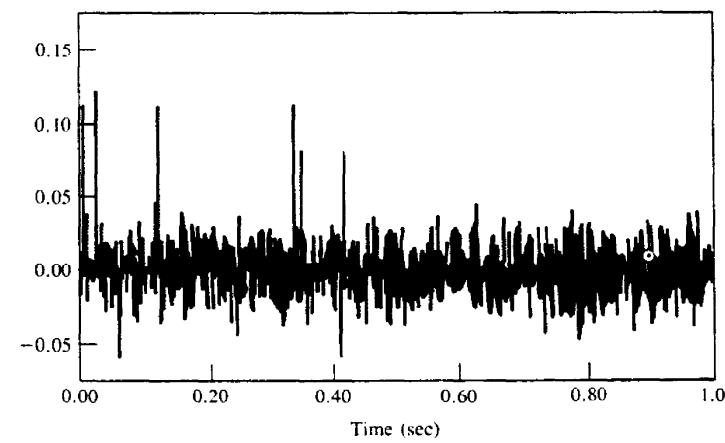


Figure 22-11 $\hat{w}(k|N)$. (Mendel, 1983a, © 1983, IEEE).

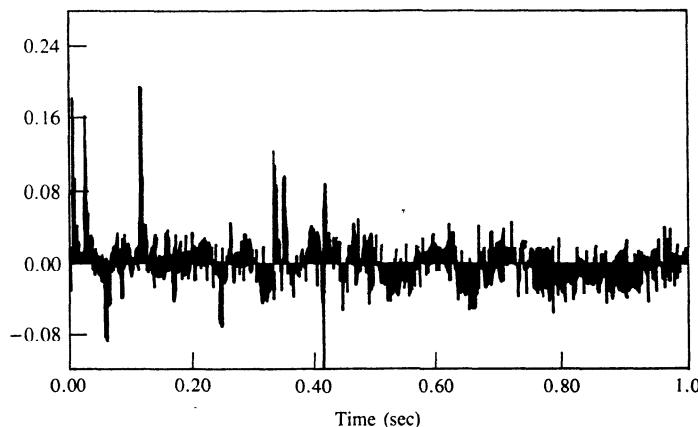


Figure 22-12 $\hat{d}_1(k|N)$ when $d_1(t) = e^{-300t}$ (Mendel 1983a, © 1983, IEEE).

More smoothing is achieved when $\hat{w}(k|N)$ is convolved with a zero-phase waveform. \square

Finally, note that, because of Theorem 22-3, we know that perfect waveshaping is not possible. For example, the signal component of $\hat{d}_1(k|N)$, $\hat{d}_{ls}(k|N)$, is given by the expression

$$\hat{d}_{ls}(k|N) = d(k)*R(k)*w(k) \quad (22-60)$$

How much the auto-correlation function $R(k)$ will distort $\hat{d}_{ls}(k|N)$ from $d(k)*w(k)$, depends, of course, on bandwidth and signal-to-noise ratio considerations.

PROBLEMS

- 22-1. Rederive the MVD algorithm for $\hat{\mu}(k|N)$, which is given in (22-5), from the Fundamental Theorem of Estimation Theory, i.e., $\hat{\mu}(k|N) = E\{\mu(k)|\mathcal{E}(N)\}$.
- 22-2. Prove Theorem 22-3. Explain why part (b) of the theorem means that $\hat{\mu}_s(k|N)$ is a zero-phase waveshaped version of $\mu(k)$.
- 22-3. This problem is a memory refresher. You probably have either seen or carried out the calculations asked for in a course on random processes.
 - (a) Derive Equation (22-34);
 - (b) Derive Equation (22-37).
- 22-4. Prove the recursive waveshaping Theorem 22-7.

Lesson 23

State Estimation for the Not-So-Basic State-Variable Model

INTRODUCTION

In deriving all of our state estimators we assumed that our dynamical system could be modeled as in Lesson 15, i.e., as our *basic state-variable model*. The results so obtained are applicable only for systems that satisfy all the conditions of that model: the noise processes $w(k)$ and $v(k)$ are both zero mean, white, and mutually uncorrelated, no known bias functions appear in the state or measurement equations, and no measurements are noise-free (i.e., perfect). The following cases frequently occur in practice:

1. either nonzero-mean noise processes or known bias functions or both in the state or measurement equations
2. correlated noise processes,
3. colored noise processes, and
4. some perfect measurements.

In this lesson we show how to modify some of our earlier results in order to treat these important special cases. In order to see the forest from the trees, we consider each of these four cases separately. In practice, some or all of them may occur together.

BIASES

Here we assume that our basic state-variable model, given in (15-17) and (15-18), has been modified to

$$\mathbf{x}(k+1) = \Phi(k+1,k)\mathbf{x}(k) + \Gamma(k+1,k)\mathbf{w}_1(k) + \Psi(k+1,k)\mathbf{u}(k) \quad (23-1)$$

$$\mathbf{z}(k+1) = \mathbf{H}(k+1)\mathbf{x}(k+1) + \mathbf{G}(k+1)\mathbf{u}(k+1) + \mathbf{v}_1(k+1) \quad (23-2)$$

where $\mathbf{w}_1(k)$ and $\mathbf{v}_1(k)$ are nonzero mean, individually and mutually uncorrelated Gaussian noise sequences, i.e.,

$$E\{\mathbf{w}_1(k)\} = \mathbf{m}_{\mathbf{w}_1}(k) \neq \mathbf{0} \quad \mathbf{m}_{\mathbf{w}_1}(k) \text{ known} \quad (23-3)$$

$$E\{\mathbf{v}_1(k)\} = \mathbf{m}_{\mathbf{v}_1}(k) \neq \mathbf{0} \quad \mathbf{m}_{\mathbf{v}_1}(k) \text{ known} \quad (23-4)$$

$E\{[\mathbf{w}_1(i) - \mathbf{m}_{\mathbf{w}_1}(i)][\mathbf{w}_1(j) - \mathbf{m}_{\mathbf{w}_1}(j)]'\} = \mathbf{Q}(i)\delta_{ij}$, $E\{[\mathbf{v}_1(i) - \mathbf{m}_{\mathbf{v}_1}(i)][\mathbf{v}_1(j) - \mathbf{m}_{\mathbf{v}_1}(j)]'\} = \mathbf{R}(i)\delta_{ij}$, and $E\{[\mathbf{w}_1(i) - \mathbf{m}_{\mathbf{w}_1}(i)][\mathbf{v}_1(j) - \mathbf{m}_{\mathbf{v}_1}(j)]'\} = \mathbf{0}$.

This case is handled by reducing (23-1) and (23-2) to our previous basic state-variable model, using the following simple transformations. Let

$$\mathbf{w}(k) \triangleq \mathbf{w}_1(k) - \mathbf{m}_{\mathbf{w}_1}(k) \quad (23-5)$$

and

$$\mathbf{v}(k) \triangleq \mathbf{v}_1(k) - \mathbf{m}_{\mathbf{v}_1}(k) \quad (23-6)$$

Observe that both $\mathbf{w}(k)$ and $\mathbf{v}(k)$ are zero-mean white noise processes, with covariances $\mathbf{Q}(k)$ and $\mathbf{R}(k)$, respectively. Adding and subtracting $\Gamma(k+1,k)\mathbf{m}_{\mathbf{w}_1}(k)$ in state equation (23-1) and $\mathbf{m}_{\mathbf{v}_1}(k+1)$ in measurement equation (23-2), these equations can be expressed as

$$\mathbf{x}(k+1) = \Phi(k+1,k)\mathbf{x}(k) + \Gamma(k+1,k)\mathbf{w}(k) + \mathbf{u}_1(k) \quad (23-7)$$

and

$$\mathbf{z}_1(k+1) = \mathbf{H}(k+1)\mathbf{x}(k+1) + \mathbf{v}(k+1) \quad (23-8)$$

where

$$\mathbf{u}_1(k) = \Psi(k+1,k)\mathbf{u}(k) + \Gamma(k+1,k)\mathbf{m}_{\mathbf{v}_1}(k) \quad (23-9)$$

and

$$\mathbf{z}_1(k+1) = \mathbf{z}(k+1) - \mathbf{G}(k+1)\mathbf{u}(k+1) - \mathbf{m}_{\mathbf{v}_1}(k+1) \quad (23-10)$$

Clearly, (23-7) and (23-8) is once again a basic state-variable model, one in which $\mathbf{u}_1(k)$ plays the role of $\Psi(k+1,k)\mathbf{u}(k)$ and $\mathbf{z}_1(k+1)$ plays the role of $\mathbf{z}(k+1)$.

Theorem 23-1. *When biases are present in a state-variable model, then that model can always be reduced to a basic state-variable model [e.g., (23-7) to (23-10)]. All of our previous state estimators can be applied to this basic state-variable model by replacing $\mathbf{z}(k)$ by $\mathbf{z}_1(k)$ and $\Psi(k+1,k)\mathbf{u}(k)$ by $\mathbf{u}_1(k)$. \square*

CORRELATED NOISES

Here we assume that our basic state-variable model is given by (15-17) and (15-18), except that now $\mathbf{w}(k)$ and $\mathbf{v}(k)$ are correlated, i.e.,

$$E\{\mathbf{w}(k)\mathbf{v}'(k)\} = \mathbf{S}(k) \neq \mathbf{0} \quad (23-11)$$

There are many approaches for treating correlated process and measurement noises, some leading to a recursive predictor, some to a recursive filter, and others to a filter in predictor-corrector form, as in the following:

Theorem 23-2. *When $\mathbf{w}(k)$ and $\mathbf{v}(k)$ are correlated, then a predictor-corrector form of the Kalman filter is*

$$\begin{aligned} \hat{\mathbf{x}}(k+1|k) &= \Phi(k+1,k)\hat{\mathbf{x}}(k|k) + \Psi(k+1,k)\mathbf{u}(k) \\ &+ \Gamma(k+1,k)\mathbf{S}(k)[\mathbf{H}(k)\mathbf{P}(k|k-1)\mathbf{H}'(k) + \mathbf{R}(k)]^{-1}\tilde{\mathbf{z}}(k|k-1) \end{aligned} \quad (23-12)$$

and

$$\hat{\mathbf{x}}(k+1|k+1) = \hat{\mathbf{x}}(k+1|k) + \mathbf{K}(k+1)\tilde{\mathbf{z}}(k+1|k) \quad (23-13)$$

where Kalman gain matrix, $\mathbf{K}(k+1)$, is given by (17-12), filtering-error covariance matrix, $\mathbf{P}(k+1|k+1)$, is given by (17-14), and, prediction-error covariance matrix, $\mathbf{P}(k+1|k)$, is given by

$$\mathbf{P}(k+1|k) = \Phi_1(k+1,k)\mathbf{P}(k|k)\Phi_1'(k+1,k) + \mathbf{Q}_1(k) \quad (23-14)$$

in which

$$\Phi_1(k+1,k) = \Phi(k+1,k) - \Gamma(k+1,k)\mathbf{S}(k)\mathbf{R}^{-1}(k)\mathbf{H}(k) \quad (23-15)$$

and

$$\begin{aligned} \mathbf{Q}_1(k) &= \Gamma(k+1,k)\mathbf{Q}(k)\Gamma'(k+1,k) \\ &- \Gamma(k+1,k)\mathbf{S}(k)\mathbf{R}^{-1}(k)\mathbf{S}'(k)\Gamma'(k+1,k) \end{aligned} \quad (23-16)$$

Observe that, if $\mathbf{S}(k) = \mathbf{0}$, then (23-12) reduces to the more familiar predictor equation (16-4), and (23-14) reduces to the more familiar (17-13).

Proof. The derivation of correction equation (23-13) is exactly the same, when $\mathbf{w}(k)$ and $\mathbf{v}(k)$ are correlated, as it was when $\mathbf{w}(k)$ and $\mathbf{v}(k)$ were assumed uncorrelated. See the proof of part (a) of Theorem 17-1 for the details.

In order to derive predictor equation (23-12), we begin with the Fundamental Theorem of Estimation Theory, i.e.,

$$\hat{\mathbf{x}}(k+1|k) = E\{\mathbf{x}(k+1)|\mathcal{E}(k)\} \quad (23-17)$$

Substitute state equation (15-17) into (23-17), to show that

$$\begin{aligned} \hat{\mathbf{x}}(k+1|k) &= \Phi(k+1,k)\hat{\mathbf{x}}(k|k) + \Psi(k+1,k)\mathbf{u}(k) \\ &+ \Gamma(k+1,k)E\{\mathbf{w}(k)|\mathcal{E}(k)\} \end{aligned} \quad (23-18)$$

Next, we develop an expression for $E\{\mathbf{w}(k)|\mathbf{z}(k)\}$.

Let $\mathbf{z}(k) = \text{col}(\mathbf{z}(k-1), \mathbf{z}(k))$; then,

$$\begin{aligned} E\{\mathbf{w}(k)|\mathbf{z}(k)\} &= E\{\mathbf{w}(k)|\mathbf{z}(k-1), \mathbf{z}(k)\} \\ &= E\{\mathbf{w}(k)|\mathbf{z}(k-1), \tilde{\mathbf{z}}(k|k-1)\} \\ &= E\{\mathbf{w}(k)|\mathbf{z}(k-1)\} + E\{\mathbf{w}(k)|\tilde{\mathbf{z}}(k|k-1)\} - E\{\mathbf{w}(k)\} \\ &= E\{\mathbf{w}(k)|\tilde{\mathbf{z}}(k|k-1)\} \end{aligned} \quad (23-19)$$

In deriving (23-19) we used the facts that $\mathbf{w}(k)$ is zero mean, and $\mathbf{w}(k)$ and $\mathbf{z}(k-1)$ are statistically independent. Because $\mathbf{w}(k)$ and $\tilde{\mathbf{z}}(k|k-1)$ are jointly Gaussian,

$$E\{\mathbf{w}(k)|\tilde{\mathbf{z}}(k|k-1)\} = \mathbf{P}_{w\tilde{z}}(k, k|k-1) \mathbf{P}_{\tilde{z}\tilde{z}}^{-1}(k|k-1) \tilde{\mathbf{z}}(k|k-1) \quad (23-20)$$

where $\mathbf{P}_{\tilde{z}\tilde{z}}$ is given by (16-33), and

$$\begin{aligned} \mathbf{P}_{w\tilde{z}}(k, k|k-1) &= E\{\mathbf{w}(k)\tilde{\mathbf{z}}'(k|k-1)\} \\ &= E\{\mathbf{w}(k)[\mathbf{H}(k)\tilde{\mathbf{x}}(k|k-1) + \mathbf{v}(k)]'\} \\ &= \mathbf{S}(k) \end{aligned} \quad (23-21)$$

In deriving (23-21) we used the facts that $\tilde{\mathbf{x}}(k|k-1)$ and $\mathbf{w}(k)$ are statistically independent, and $\mathbf{w}(k)$ is zero mean. Substituting (23-21) and (16-33) into (23-20), we find that

$$E\{\mathbf{w}(k)|\tilde{\mathbf{z}}(k|k-1)\} = \mathbf{S}(k)[\mathbf{H}(k)\mathbf{P}(k|k-1)\mathbf{H}'(k) + \mathbf{R}(k)]^{-1} \quad (23-22)$$

Substituting (23-22) into (23-19), and the resulting equation into (23-18) completes our derivation of the recursive predictor equation (23-12).

We leave the derivation of (23-14) as an exercise. It is straightforward but algebraically tedious. \square

Recall that the recursive predictor plays the predominant role in smoothing; hence, we present

Corollary 23-1. *When $\mathbf{w}(k)$ and $\mathbf{v}(k)$ are correlated, then a recursive predictor for $\mathbf{x}(k+1)$, is*

$$\begin{aligned} \hat{\mathbf{x}}(k+1|k) &= \Phi(k+1, k)\hat{\mathbf{x}}(k|k-1) \\ &\quad + \Psi(k+1, k)\mathbf{u}(k) + \mathbf{L}(k)\tilde{\mathbf{z}}(k|k-1) \end{aligned} \quad (23-23)$$

where

$$\begin{aligned} \mathbf{L}(k) &= [\Phi(k+1, k)\mathbf{P}(k|k-1)\mathbf{H}'(k) \\ &\quad + \Gamma(k+1, k)\mathbf{S}(k)][\mathbf{H}(k)\mathbf{P}(k|k-1)\mathbf{H}'(k) + \mathbf{R}(k)]^{-1} \end{aligned} \quad (23-24)$$

and

$$\begin{aligned} \mathbf{P}(k+1|k) &= [\Phi(k+1, k) - \mathbf{L}(k)\mathbf{H}(k)]\mathbf{P}(k|k-1) \\ &\quad + [\Phi(k+1, k) - \mathbf{L}(k)\mathbf{H}(k)]' \\ &\quad + \Gamma(k+1, k)\mathbf{Q}(k)\Gamma'(k+1, k) \\ &\quad - \Gamma(k+1, k)\mathbf{S}(k)\mathbf{L}'(k) \\ &\quad - \mathbf{L}(k)\mathbf{S}(k)\Gamma'(k+1, k) \\ &\quad + \mathbf{L}(k)\mathbf{R}(k)\mathbf{L}'(k) \end{aligned} \quad (23-25)$$

Proof. These results follow directly from Theorem 23-2; or, they can be derived in an independent manner, as explained in Problem 23-2. \square

Corollary 23-2. *When $\mathbf{w}(k)$ and $\mathbf{v}(k)$ are correlated, then a recursive filter for $\mathbf{x}(k+1)$ is*

$$\begin{aligned} \hat{\mathbf{x}}(k+1|k) &= \Phi_1(k+1, k)\hat{\mathbf{x}}(k|k) + \Psi(k+1, k)\mathbf{u}(k) \\ &\quad + \mathbf{D}(k)\mathbf{z}(k) + \mathbf{K}(k+1)\tilde{\mathbf{z}}(k+1|k) \end{aligned} \quad (23-26)$$

where

$$\mathbf{D}(k) = \Gamma(k+1, k)\mathbf{S}(k)\mathbf{R}^{-1}(k) \quad (23-27)$$

and all other quantities have been defined above.

Proof. Again, these results follow directly from Theorem 23-2; however, they can also be derived, in a much more elegant and independent manner, as described in Problem 23-3. \square

COLORED NOISES

Quite often, some or all of the elements of either $\mathbf{v}(k)$ or $\mathbf{w}(k)$ or both are colored (i.e., have finite bandwidth). The following three-step procedure is used in these cases:

1. model each colored noise by a low-order difference equation that is excited by white Gaussian noise;
2. augment the states associated with the step 1 colored noise models to the original state-variable model;
3. apply the recursive filter or predictor to the augmented system.

We try to model colored noise processes by low-order Markov processes, i.e., low-order difference equations. Usually, first- or second-order models

are quite adequate. Consider the following first-order model for colored noise process $\omega(k)$,

$$\omega(k+1) = \alpha\omega(k) + n(k) \quad (23-28)$$

In this model $n(k)$ is white noise with variance σ_n^2 ; thus, this model contains two parameters, α and σ_n^2 , which must be determined from a priori knowledge about $\omega(k)$. We may know the amplitude spectrum of $\omega(k)$, correlation information about $\omega(k)$, steady-state variance of $\omega(k)$, etc. Two independent pieces of information are needed in order to uniquely identify α and σ_n^2 .

Example 23-1

We are given the facts that scalar noise $w(k)$ is stationary with the properties $E\{w(k)\} = 0$ and $E\{w(i)w(j)\} = e^{-2|i-j|}$. A first-order Markov model for $w(k)$ can easily be obtained as

$$\xi(k+1) = e^{-2} \xi(k) + \sqrt{1 - e^{-4}} n(k) \quad (23-29)$$

$$w(k) = \xi(k) \quad (23-30)$$

where $E\{\xi(0)\} = 0$, $E\{\xi^2(0)\} = 1$, $E\{n(k)\} = 0$ and $E\{n(i)n(j)\} = \delta_{ij}$. \square

Example 23-2

Here we illustrate the state augmentation procedure for the first-order system

$$x(k+1) = a_1 x(k) + \omega(k) \quad (23-31)$$

$$z(k+1) = h x(k+1) + v(k+1) \quad (23-32)$$

where $\omega(k)$ is a first-order Markov process, i.e.,

$$\omega(k+1) = a_2 \omega(k) + n(k) \quad (23-33)$$

and $v(k)$ and $n(k)$ are white noise processes. We "augment" (23-33) to (23-31), as follows. Let

$$x(k) = \text{col}(\chi(k), \omega(k)) \quad (23-34)$$

then (23-31) and (23-33) can be combined, to give

$$\underbrace{\begin{pmatrix} \chi(k+1) \\ \omega(k+1) \end{pmatrix}}_{\mathbf{x}(k+1)} = \underbrace{\begin{pmatrix} a_1 & 1 \\ 0 & a_2 \end{pmatrix}}_{\Phi} \underbrace{\begin{pmatrix} \chi(k) \\ \omega(k) \end{pmatrix}}_{\mathbf{x}(k)} + \underbrace{\begin{pmatrix} 0 \\ 1 \end{pmatrix}}_{\gamma} n(k) \quad (23-35)$$

Equation (23-35) is our *augmented state equation*. Observe that it is once again excited by a white noise process, just as our basic state equation (15-17) is.

In order to complete the description of the augmented state-variable model, we must express measurement $z(k+1)$ in terms of the augmented state vector, $x(k+1)$, i.e.,

$$\underbrace{z(k+1)}_{\mathbf{H}} = \underbrace{(h \ 0)}_{\mathbf{H}} \underbrace{\begin{pmatrix} \chi(k+1) \\ \omega(k+1) \end{pmatrix}}_{\mathbf{x}(k+1)} + v(k+1) \quad (23-36)$$

Colored Noises

Equations (23-35) and (23-36) constitute the augmented state-variable model. We observe that, when the original process noise is colored and the measurement noise is white, the state augmentation procedure leads us once again to a basic (augmented) state-variable model, one that is of higher dimension than the original model because of the modeled colored process noise. Hence, in this case we can apply all of our state estimation algorithms to the augmented state-variable model. \square

Example 23-3

Here we consider the situation where the process noise is white but the measurement noise is colored, again for a first-order system,

$$\chi(k+1) = a_1 \chi(k) + \omega(k) \quad (23-37)$$

$$z(k+1) = h \chi(k+1) + v(k+1) \quad (23-38)$$

As in the preceding example, we model $v(k)$ by the following first-order Markov process

$$v(k+1) = a_2 v(k) + n(k) \quad (23-39)$$

where $n(k)$ is white noise. Augmenting (23-39) to (23-37) and reexpressing (23-38) in terms of the augmented state vector $x(k)$, where

$$x(k) = \text{col}(\chi(k), v(k)) \quad (23-40)$$

we obtain the following augmented state-variable model,

$$\underbrace{\begin{pmatrix} \chi(k+1) \\ v(k+1) \end{pmatrix}}_{\mathbf{x}(k+1)} = \underbrace{\begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix}}_{\Phi} \underbrace{\begin{pmatrix} \chi(k) \\ v(k) \end{pmatrix}}_{\mathbf{x}(k)} + \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{\Gamma} \underbrace{\begin{pmatrix} \omega(k) \\ n(k) \end{pmatrix}}_{\mathbf{w}(k)} \quad (23-41)$$

and

$$\underbrace{z(k+1)}_{\mathbf{H}} = \underbrace{(h \ 1)}_{\mathbf{H}} \underbrace{\begin{pmatrix} \chi(k+1) \\ v(k+1) \end{pmatrix}}_{\mathbf{x}(k+1)} \quad (23-42)$$

Observe that a vector process noise now excites the augmented state equation and that there is no measurement noise in the measurement equation. This second observation can lead to serious numerical problems in our state estimators, because it means that we must set $\mathbf{R} = \mathbf{0}$ in those estimators, and, when we do this, covariance matrices become and remain singular. \square

Let us examine what happens to $\mathbf{P}(k+1|k+1)$ when covariance matrix \mathbf{R} is set equal to zero. From (17-14) and (17-12) (in which we set $\mathbf{R} = \mathbf{0}$), we find that

$$\mathbf{P}(k+1|k+1) = \mathbf{P}(k+1|k) - \mathbf{P}(k+1|k) \mathbf{H}'(k+1) [\mathbf{H}(k+1) \mathbf{P}(k+1|k) \mathbf{H}'(k+1)]^{-1} \mathbf{H}(k+1) \mathbf{P}(k+1|k) \quad (23-43)$$

Multiplying both sides of (23-43) on the right by $\mathbf{H}'(k+1)$, we find that

$$\mathbf{P}(k+1|k+1) \mathbf{H}'(k+1) = \mathbf{0} \quad (23-44)$$

Because $\mathbf{H}'(k+1)$ is a nonzero matrix, (23-44) implies that $\mathbf{P}(k+1|k+1)$ must be a singular matrix. We leave it to the reader to show that once $\mathbf{P}(k+1|k+1)$ becomes singular it remains singular for all other values of k .

PERFECT MEASUREMENTS: REDUCED-ORDER ESTIMATORS

We have just seen that when $\mathbf{R} = \mathbf{0}$ (or, in fact, even if some, but not all, measurements are perfect) numerical problems can occur in the Kalman filter. One way to circumvent these problems is ad hoc, and that is to use small values for the elements of covariance matrix \mathbf{R} , even though measurements are thought to be perfect. Doing this has a stabilizing effect on the numerics of the Kalman filter.

A second way to circumvent these problems is to recognize that a set of “perfect” measurements reduces the number of states that have to be estimated. Suppose, for example, that there are l perfect measurements and that state vector $\mathbf{x}(k)$ is $n \times 1$. Then, we conjecture that we ought to be able to estimate $\mathbf{x}(k)$ by a Kalman filter whose dimension is no greater than $n - l$. Such an estimator will be referred to as a *reduced-order estimator*. The payoff for using a reduced-order estimator is fewer computations and less storage.

In order to illustrate an approach to designing a reduced-order estimator, we limit our discussions in this section to the following time-invariant and stationary basic state-variable model in which $\mathbf{u}(k) \triangleq \mathbf{0}$ and *all* measurements are perfect,

$$\mathbf{x}(k+1) = \Phi\mathbf{x}(k) + \Gamma\mathbf{w}(k) \quad (23-45)$$

$$\mathbf{y}(k+1) = \mathbf{H}\mathbf{x}(k+1) \quad (23-46)$$

In this model \mathbf{y} is $l \times 1$. What makes the design of a reduced-order estimator challenging is the fact that the l perfect measurements are linearly related to the n states, i.e., \mathbf{H} is rectangular.

To begin we introduce a *reduced-order state vector*, $\mathbf{p}(k)$, whose dimension is $(n - l) \times 1$; $\mathbf{p}(k)$ is assumed to be a linear transformation of $\mathbf{x}(k)$, i.e.,

$$\mathbf{p}(k) \triangleq \mathbf{C}\mathbf{x}(k) \quad (23-47)$$

Augmenting (23-47) to (23-46), we obtain

$$\begin{pmatrix} \mathbf{y}(k) \\ \mathbf{p}(k) \end{pmatrix} = \begin{pmatrix} \mathbf{H} \\ \mathbf{C} \end{pmatrix} \mathbf{x}(k) \quad (23-48)$$

Design matrix \mathbf{C} is chosen so that $\begin{pmatrix} \mathbf{H} \\ \mathbf{C} \end{pmatrix}$ is invertible. Of course, many different choices of \mathbf{C} are possible; thus, this first step of our reduced-order estimator design procedure is nonunique. Let

$$\begin{pmatrix} \mathbf{H} \\ \mathbf{C} \end{pmatrix}^{-1} = (\mathbf{L}_1 | \mathbf{L}_2) \quad (23-49)$$

where \mathbf{L}_1 is $n \times l$ and \mathbf{L}_2 is $n \times (n - l)$; thus,

$$\mathbf{x}(k) = \mathbf{L}_1\mathbf{y}(k) + \mathbf{L}_2\mathbf{p}(k) \quad (23-50)$$

In order to obtain a filtered estimate of $\mathbf{x}(k)$, we operate on both sides of (23-50) with $\mathbf{E}\{\cdot | \mathbf{y}(k)\}$, where

$$\mathbf{y}(k) = \text{col}(\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(k)) \quad (23-51)$$

Doing this, we find that

$$\hat{\mathbf{x}}(k|k) = \mathbf{L}_1\mathbf{y}(k) + \mathbf{L}_2\hat{\mathbf{p}}(k|k) \quad (23-52)$$

which is a *reduced-order estimator* for $\mathbf{x}(k)$. Of course, in order to evaluate $\hat{\mathbf{x}}(k|k)$ we must develop a reduced-order Kalman filter to estimate $\mathbf{p}(k)$. Knowing $\hat{\mathbf{p}}(k|k)$ and $\mathbf{y}(k)$, it is then a simple matter to compute $\hat{\mathbf{x}}(k|k)$, using (23-52).

In order to obtain $\hat{\mathbf{p}}(k|k)$, using our previously-derived Kalman filter algorithm, we first must establish a state-variable model for $\mathbf{p}(k)$. A state equation for \mathbf{p} is easily obtained, as follows:

$$\begin{aligned} \mathbf{p}(k+1) &= \mathbf{C}\mathbf{x}(k+1) = \mathbf{C}[\Phi\mathbf{x}(k) + \Gamma\mathbf{w}(k)] \\ &= \mathbf{C}\Phi[\mathbf{L}_1\mathbf{y}(k) + \mathbf{L}_2\mathbf{p}(k)] + \mathbf{C}\Gamma\mathbf{w}(k) \\ &= \mathbf{C}\Phi\mathbf{L}_2\mathbf{p}(k) + \mathbf{C}\Phi\mathbf{L}_1\mathbf{y}(k) + \mathbf{C}\Gamma\mathbf{w}(k) \end{aligned} \quad (23-53)$$

Observe that this state equation is driven by white noise $\mathbf{w}(k)$ and the known forcing function, $\mathbf{y}(k)$.

A measurement equation is obtained from (23-46), as

$$\begin{aligned} \mathbf{y}(k+1) &= \mathbf{H}\mathbf{x}(k+1) = \mathbf{H}[\Phi\mathbf{x}(k) + \Gamma\mathbf{w}(k)] \\ &= \mathbf{H}\Phi[\mathbf{L}_1\mathbf{y}(k) + \mathbf{L}_2\mathbf{p}(k)] + \mathbf{H}\Gamma\mathbf{w}(k) \\ &= \mathbf{H}\Phi\mathbf{L}_2\mathbf{p}(k) + \mathbf{H}\Phi\mathbf{L}_1\mathbf{y}(k) + \mathbf{H}\Gamma\mathbf{w}(k) \end{aligned} \quad (23-54)$$

At time $k + 1$ we know $\mathbf{y}(k)$; hence, we can reexpress (23-54) as

$$\mathbf{y}_1(k+1) = \mathbf{H}\Phi\mathbf{L}_2\mathbf{p}(k) + \mathbf{H}\Gamma\mathbf{w}(k) \quad (23-55)$$

where

$$\mathbf{y}_1(k+1) \triangleq \mathbf{y}(k+1) - \mathbf{H}\Phi\mathbf{L}_1\mathbf{y}(k) \quad (23-56)$$

Before proceeding any farther, we make some important observations about our state-variable model in (23-53) and (23-55). First, the new measurement $\mathbf{y}_1(k+1)$ represents a weighted difference between measurements $\mathbf{y}(k+1)$ and $\mathbf{y}(k)$. The technique for obtaining our reduced-order state-variable model is, therefore, sometimes referred to as a *measurement-differencing technique* (e.g., Bryson and Johansen, 1965). Because we have already used $\mathbf{y}(k)$ to reduce the dimension of $\mathbf{x}(k)$ from n to $n - l$, we cannot again use $\mathbf{y}(k)$ alone as the measurements in our reduced-order state-variable model. As we have just seen, we must use both $\mathbf{y}(k)$ and $\mathbf{y}(k+1)$.

Second, measurement equation (23-55) appears to be a combination of signal and noise. Unless $\mathbf{H}\Gamma = \mathbf{0}$, the term $\mathbf{H}\Gamma\mathbf{w}(k)$ will act as the measurement noise in our reduced-order state-variable model. Its covariance matrix is $\mathbf{H}\Gamma\mathbf{Q}\Gamma'\mathbf{H}'$. Unfortunately, it is possible for $\mathbf{H}\Gamma$ to equal the zero matrix. From linear system theory, we know that $\mathbf{H}\Gamma$ is the matrix of first Markov parameters for our original system in (23-45) and (23-46), and $\mathbf{H}\Gamma$ may equal zero. If this occurs, then we must repeat all of the above until we obtain a reduced-order state vector whose measurement equation is excited by white noise. We see, therefore, that depending upon system dynamics, it is possible to obtain a reduced-order estimator of $\mathbf{x}(k)$ that uses a reduced-order Kalman filter of dimension less than $n - l$.

Third, the noises, which appear in state equation (23-53) and measurement equation (23-55) are the same, namely $\mathbf{w}(k)$; hence, the reduced-order state-variable model involves the correlated noise case that we described before in this chapter in the section entitled Correlated Noises.

Finally, and most important, measurement equation (23-55) is non-standard, in that it expresses \mathbf{y}_1 at $k + 1$ in terms of \mathbf{p} at k rather than \mathbf{p} at $k + 1$. Recall that the measurement equation in our basic state-variable model is $\mathbf{z}(k + 1) = \mathbf{Hx}(k + 1) + \mathbf{v}(k + 1)$. We cannot immediately apply our Kalman filter equations to (23-53) and (23-55) until we express (23-55) in the standard way.

To proceed, we let

$$\zeta(k) \triangleq \mathbf{y}_1(k + 1) \quad (23-57)$$

so that

$$\zeta(k) = \mathbf{H}\Phi\mathbf{L}_2\mathbf{p}(k) + \mathbf{H}\Gamma\mathbf{w}(k) \quad (23-58)$$

Measurement equation (23-58) is now in the standard form; however, because $\zeta(k)$ equals a future value of \mathbf{y}_1 , namely $\mathbf{y}_1(k + 1)$, we must be very careful in applying our estimator formulas to our reduced-order model (23-53) and (23-58).

In order to see this more clearly, we define the following two data sets,

$$\mathcal{D}_{y_1}(k + 1) = \{\mathbf{y}_1(1), \mathbf{y}_1(2), \dots, \mathbf{y}_1(k + 1), \dots\} \quad (23-59)$$

and

$$\mathcal{D}_t(k) = \{\zeta(0), \zeta(1), \dots, \zeta(k), \dots\} \quad (23-60)$$

Obviously,

$$\mathcal{D}_{y_1}(k + 1) = \mathcal{D}_t(k) \quad (23-61)$$

Letting

$$\hat{\mathbf{p}}_{y_1}(k + 1|k + 1) = \mathbf{E}\{\mathbf{p}(k + 1)|\mathcal{D}_{y_1}(k + 1)\} \quad (23-62)$$

and

$$\hat{\mathbf{p}}_t(k + 1|k) = \mathbf{E}\{\mathbf{p}(k + 1)|\mathcal{D}_t(k)\} \quad (23-63)$$

we see that

$$\hat{\mathbf{p}}_{y_1}(k + 1|k + 1) = \hat{\mathbf{p}}_t(k + 1|k) \quad (23-64)$$

Equation (23-64) tells us to obtain a recursive filter for our reduced-order model, that is in terms of data set $\mathcal{D}_{y_1}(k + 1)$, we must first obtain a recursive predictor for that model, which is in terms of data set $\mathcal{D}_t(k)$. Then, wherever $\zeta(k)$ appears in the recursive predictor, it can be replaced by $\mathbf{y}_1(k + 1)$.

Using Corollary 23-1, applied to the reduced-order model in (23-53) and (23-58), we find that

$$\begin{aligned} \hat{\mathbf{p}}_t(k + 1|k) &= \mathbf{C}\Phi\mathbf{L}_2\hat{\mathbf{p}}_t(k|k - 1) + \mathbf{C}\Phi\mathbf{L}_1\mathbf{y}(k) \\ &\quad + \mathbf{L}(k)[\zeta(k) - \mathbf{H}\Phi\mathbf{L}_2\hat{\mathbf{p}}_t(k|k - 1)] \end{aligned} \quad (23-65)$$

thus,

$$\begin{aligned} \hat{\mathbf{p}}_{y_1}(k + 1|k + 1) &= \mathbf{C}\Phi\mathbf{L}_2\hat{\mathbf{p}}_{y_1}(k|k) + \mathbf{C}\Phi\mathbf{L}_1\mathbf{y}(k) \\ &\quad + \mathbf{L}(k)[\mathbf{y}_1(k + 1) - \mathbf{H}\Phi\mathbf{L}_2\hat{\mathbf{p}}_{y_1}(k|k)] \end{aligned} \quad (23-66)$$

Equation (23-66) is our reduced-order Kalman filter. It provides filtered estimates of $\mathbf{p}(k + 1)$ and is only of dimension $(n - l) \times 1$. Of course, when $\mathbf{L}(k)$ and $\mathbf{P}_{y_1}(k + 1|k + 1)$ are computed using (23-13) and (23-14), respectively, we must make the following substitutions: $\Phi(k + 1, k) \rightarrow \mathbf{C}\Phi\mathbf{L}_2$, $\mathbf{H}(k) \rightarrow \mathbf{H}\Phi\mathbf{L}_2$, $\Gamma(k + 1, k) \rightarrow \mathbf{C}\Gamma$, $\mathbf{Q}(k) \rightarrow \mathbf{Q}$, $\mathbf{S}(k) \rightarrow \mathbf{Q}\Gamma'\mathbf{H}'$, and $\mathbf{R}(k) \rightarrow \mathbf{H}\Gamma\mathbf{Q}\Gamma'\mathbf{H}'$.

FINAL REMARK

In order to see the forest from the trees, we have considered each of our special cases separately. In actual practice, some or all of them may occur simultaneously. The exercises at the end of this lesson will permit the reader to gain experience with such cases.

PROBLEMS

- 23-1. Derive the prediction-error covariance equation (23-14).
- 23-2. Derive the recursive predictor, given in (23-23), by expressing $\hat{\mathbf{x}}(k + 1|k)$ as $\mathbf{E}\{\mathbf{x}(k + 1)|\mathcal{D}(k)\} = \mathbf{E}\{\mathbf{x}(k + 1)|\mathcal{D}(k - 1), \hat{\mathbf{x}}(k|k - 1)\}$.
- 23-3. Here we derive the recursive filter, given in (23-26), by first adding a convenient form of zero to state equation (15-17), in order to decorrelate the process noise in this modified basic state-variable model from the measurement noise $\mathbf{v}(k)$. Add $\mathbf{D}(k)[\mathbf{z}(k) - \mathbf{H}(k)\mathbf{x}(k) - \mathbf{v}(k)]$ to (15-17). The process noise, $\mathbf{w}_1(k)$, in the modified basic state-variable model, is equal to $\Gamma(k + 1, k)\mathbf{w}(k) - \mathbf{D}(k)\mathbf{v}(k)$. Choose “decorrelation” matrix $\mathbf{D}(k)$ so that $\mathbf{E}\{\mathbf{w}_1(k)\mathbf{v}'(k)\} = \mathbf{0}$. Then complete the derivation of (23-26). Observe that (23-14) can be obtained by inspection, via this derivation.

- 23-4. In solving Problem 23-3, one arrives at the following predictor equation,

$$\dot{\hat{x}}(k+1|k) = \Phi_1(k+1,k)\dot{x}(k|k) + \Psi(k+1,k)\mathbf{u}(k) + \mathbf{D}(k)\mathbf{z}(k)$$

Beginning with this predictor equation, and corrector equation (23-13), derive the recursive predictor given in (23-23).

- 23-5. Show that once $\mathbf{P}(k+1|k+1)$ becomes singular it remains singular for all other values of k .

- 23-6. Assume that $\mathbf{R} = \mathbf{0}$, $\mathbf{H}\Gamma = \mathbf{0}$, and $\mathbf{H}\Phi\Gamma \neq \mathbf{0}$. Obtain the reduced-order estimator and its associated reduced-order Kalman filter for this situation. Contrast this situation with the case given in the text, for which $\mathbf{H}\Gamma \neq \mathbf{0}$.

- 23-7. Develop a reduced-order estimator and its associated reduced-order Kalman filter for the case when l measurements are perfect and $m - l$ measurements are noisy.

- 23-8. Consider the first-order system $x(k+1) = \frac{3}{4}x(k) + w_1(k)$ and $z(k+1) = x(k+1) + v(k+1)$, where $E\{w_1(k)\} = 3$, $E\{v(k)\} = 0$, $w_1(k)$ and $v(k)$ are both white and Gaussian, $E\{w_1^2(k)\} = 10$, $E\{v^2(k)\} = 2$, and, $w_1(k)$ and $v(k)$ are correlated, i.e., $E\{w_1(k)v(k)\} = 1$.

(a) Obtain the steady-state recursive Kalman filter for this system.

(b) What is the steady-state filter error variance, and how does it compare with the steady-state predictor error variance?

- 23-9. Consider the first-order system $x(k+1) = \frac{1}{2}x(k) + w(k)$ and $z(k+1) = x(k+1) + v(k+1)$, where $w(k)$ is a first-order Markov process and $v(k)$ is Gaussian white noise with $E\{v(k)\} = 4$ and $r = 1$.

(a) Let the model for $w(k)$ be $w(k+1) = \alpha w(k) + u(k)$, where $u(k)$ is a zero-mean white Gaussian noise sequence for which $E\{u^2(k)\} = \sigma_u^2$. Additionally, $E\{w(k)\} = 0$. What value must α have if $E\{w^2(k)\} = W$ for all k ?

(b) Suppose $W^2 = 2$ and $\sigma_u^2 = 1$. What are the Kalman filter equations for estimation of $x(k)$ and $w(k)$?

- 23-10. Consider the first-order system $x(k+1) = -\frac{1}{8}x(k) + w(k)$ and $z(k+1) = x(k+1) + v(k+1)$, where $w(k)$ is white and Gaussian [$w(k) \sim N(w(k); 0, 1)$] and $v(k)$ is also a noise process. The model for $v(k)$ is summarized in Figure P23-10.

(a) Verify that a correct state-variable model for $v(k)$ is,

$$x_1(k+1) = -\frac{1}{2}x_1(k) + \frac{3}{2}n(k)$$

$$v(k) = x_1(k) + n(k)$$

(b) Show that $v(k)$ is also a white process.

(c) Noise $n(k)$ is white and Gaussian [$n(k) \sim N(n(k); 0, \frac{1}{4})$].

What are the Kalman filter equations for finding $\hat{x}(k+1|k+1)$?

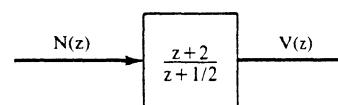


Figure P23-10

- 23-11. Obtain the equations from which we can find $\dot{x}_1(k+1|k+1)$, $\dot{x}_2(k+1|k+1)$ and $\dot{v}(k+1|k+1)$ for the following system:

$$x_1(k+1) = -x_1(k) + x_2(k)$$

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

$$z(k+1) = x_1(k+1) + v(k+1)$$

where $v(k)$ is a colored noise process, i.e.,

$$v(k+1) = -\frac{1}{2}v(k) + n(k)$$

Assume that $w(k)$ and $n(k)$ are white processes and are mutually uncorrelated, and, $\sigma_w^2(k) = 4$ and $\sigma_n^2(k) = 2$. Include a block diagram of the interconnected system and reduced-order KF.

- 23-12. Consider the system $x(k+1) = \Phi x(k) + \gamma\mu(k)$ and $z(k+1) = \mathbf{h}'x(k+1) + v(k+1)$, where $\mu(k)$ is a colored noise sequence and $v(k)$ is zero-mean white noise. What are the formulas for computing $\hat{\mu}(k|k+1)$? Filter $\hat{\mu}(k|k+1)$ is a deconvolution filter.

- 23-13. Consider the scalar moving average (MA) time-series model

$$z(k) = r(k) + r(k-1)$$

where $r(k)$ is a unit variance, white Gaussian sequence. Show that the optimal one-step predictor for this model is [assume $P(0|0) = 1$]

$$\hat{z}(k+1|k) = \frac{k}{k+1}[z(k) - \hat{z}(k|k-1)]$$

(Hint: Express the MA model in state-space form.)

- 23-14. Consider the basic state variable model for the stationary time-invariant case. assume also that $w(k)$ and $v(k)$ are correlated, i.e., $E\{w(k)v'(k)\} = S$.

- (a) Show, from first principles, that the single-stage smoother of $x(k)$, i.e., $\hat{x}(k|k+1)$ is given by

$$\hat{x}(k|k+1) = \hat{x}(k|k) + M(k|k+1)\hat{z}(k+1|k)$$

where $M(k|k+1)$ is an appropriate smoother gain matrix.

- (b) Derive a closed form solution for $M(k|k+1)$ as a function of the correlation matrix S and other quantities of the basic state-variable model.

Lesson 24

Linearization and Discretization of Nonlinear Systems

INTRODUCTION

Many real-world systems are continuous-time in nature and quite a few are also nonlinear. For example, the state equations associated with the motion of a satellite of mass m about a spherical planet of mass M , in a planet-centered coordinate system, are nonlinear, because the planet's force field obeys an inverse square law. Figure 24-1 depicts a situation where the measurement equation is nonlinear. The measurement is angle ϕ_i and is expressed in a rectangular coordinate system, i.e., $\phi_i = \tan^{-1}[y/(x - l_i)]$. Sometimes the state equation may be nonlinear and the measurement equation linear, or vice-versa, or they may both be nonlinear. Occasionally, the coordinate system in which one chooses to work causes the two former situations. For example, equations of motion in a polar coordinate system are nonlinear, whereas the measurement equations are linear. In a polar coordinate system, where ϕ is a state-variable, the measurement equation for the situation depicted in Figure 24-1 is $z_i = \phi_i$, which is linear. In a rectangular coordinate system, on the other hand, equations of motion are linear, but the measurement equations are nonlinear.

Finally, we may begin with a linear system that contains some unknown parameters. When these parameters are modeled as first-order Markov processes, and these models are augmented to the original system, the augmented model is nonlinear, because the parameters that appeared in the original "linear" model are treated as states. We shall describe this situation in much more detail in Lesson 25.

A Dynamical Model

237

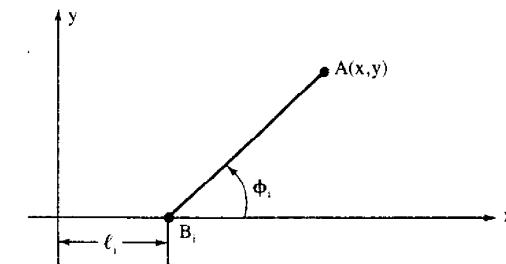


Figure 24-1 Coordinate system for an angular measurement between two objects A and B .

The purpose of this lesson is to explain how to linearize and discretize a nonlinear differential equation model. We do this so that we will be able to apply our digital estimators to the resulting discrete-time system.

A DYNAMICAL MODEL

The starting point for this lesson is the nonlinear state-variable model

$$\dot{\mathbf{x}}(t) = \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t] + \mathbf{G}(t)\mathbf{w}(t) \quad (24-1)$$

$$\mathbf{z}(t) = \mathbf{h}[\mathbf{x}(t), \mathbf{u}(t), t] + \mathbf{v}(t) \quad (24-2)$$

We shall assume that measurements are only available at specific values of time, namely at $t = t_i$, $i = 1, 2, \dots$; thus, our measurement equation will be treated as a discrete-time equation, whereas our state equation will be treated as a continuous-time equation. State vector $\mathbf{x}(t)$ is $n \times 1$; $\mathbf{u}(t)$ is an $l \times 1$ vector of known inputs; measurement vector $\mathbf{z}(t)$ is $m \times 1$; $\dot{\mathbf{x}}(t)$ is short for $d\mathbf{x}(t)/dt$; nonlinear functions \mathbf{f} and \mathbf{h} may depend both implicitly and explicitly on t , and we assume that both \mathbf{f} and \mathbf{h} are continuous and continuously differentiable with respect to all the elements of \mathbf{x} and \mathbf{u} ; $\mathbf{w}(t)$ is a continuous-time white noise process, i.e., $E\{\mathbf{w}(t)\} = \mathbf{0}$ and

$$E\{\mathbf{w}(t)\mathbf{w}'(\tau)\} = Q(t)\delta(t - \tau); \quad (24-3)$$

$\mathbf{v}(t_i)$ is a discrete-time white noise process, i.e., $E\{\mathbf{v}(t_i)\} = \mathbf{0}$ for $t = t_i$, $i = 1, 2, \dots$, and

$$E\{\mathbf{v}(t_i)\mathbf{v}'(t_j)\} = R(t_i)\delta_{ij}; \quad (24-4)$$

and, $\mathbf{w}(t)$ and $\mathbf{v}(t_i)$ are mutually uncorrelated at all $t = t_i$, i.e.,

$$E\{\mathbf{w}(t)\mathbf{v}'(t_i)\} = \mathbf{0} \quad \text{for } t = t_i \quad i = 1, 2, \dots \quad (24-5)$$

Example 24-1

Here we expand upon the previously mentioned satellite-planet example. Our example is taken from Meditch (1969, pp. 60–61), who states . . . “Assuming that the planet’s force field obeys an inverse square law, and that the only other forces present are the satellite’s two thrust forces $u_r(t)$ and $u_\theta(t)$ (see Figure 24-2), and that the satellite’s initial position and velocity vectors lie in the plane, we know from elementary particle mechanics that the satellite’s motion is confined to the plane and is governed by the two equations

$$\ddot{r} = r\dot{\theta}^2 - \frac{\gamma}{r^2} + \frac{1}{m} u_r(t) \quad (24-6)$$

and

$$\ddot{\theta} = -\frac{2\dot{r}\dot{\theta}}{r} + \frac{1}{m} u_\theta(t) \quad (24-7)$$

where $\gamma = GM$ and G is the universal gravitational constant.

“Defining $x_1 = r$, $x_2 = \dot{r}$, $x_3 = \theta$, $x_4 = \dot{\theta}$, $u_1 = u_r$, and $u_2 = u_\theta$, we have

$$\left. \begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_1 x_4^2 - \frac{\gamma}{x_1^2} + \frac{1}{m} u_1(t) \\ \dot{x}_3 &= x_4 \\ \dot{x}_4 &= -\frac{2x_2 x_4}{x_1} + \frac{1}{m} u_2(t) \end{aligned} \right\} \quad (24-8)$$

which is of the form in (24-1). . . Assuming . . . that the measurement made on the satellite during its motion is simply its distance from the surface of the planet, we have the scalar measurement equation

$$z(t) = r(t) - r_0 + v(t) = x_1(t) - r_0 + v(t) \quad (24-9)$$

where r_0 is the planet’s radius.”

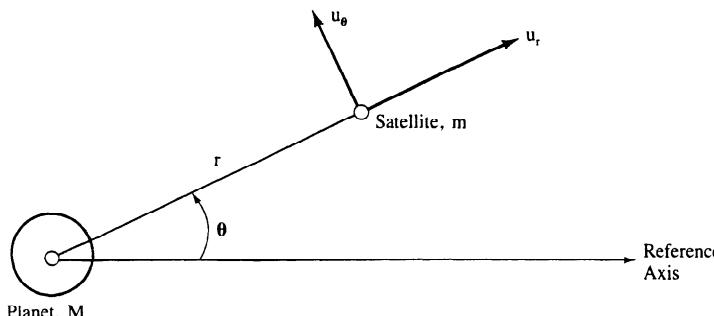


Figure 24-2 Schematic for satellite-planet system (Copyright 1969, McGraw-Hill).

Comparing (24-8) and (24-1), and (24-9) and (24-2), we conclude that

$$\mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t] = \text{col} \left[x_2, x_1 x_4^2 - \frac{\gamma}{x_1^2} + \frac{1}{m} u_1, x_4, -\frac{2x_2 x_4}{x_1} + \frac{1}{m} u_2 \right] \quad (24-10)$$

and

$$\mathbf{h}[\mathbf{x}(t), \mathbf{u}(t), t] = x_1 - r_0 \quad (24-11)$$

Observe that, in this example, only the state equation is nonlinear. \square

LINEAR PERTURBATION EQUATIONS

In this section we shall linearize our nonlinear dynamical model in (24-1) and (24-2) about nominal values of $\mathbf{x}(t)$ and $\mathbf{u}(t)$, $\mathbf{x}^*(t)$ and $\mathbf{u}^*(t)$, respectively. If we are given a nominal input, $\mathbf{u}^*(t)$, then $\mathbf{x}^*(t)$ satisfies the following nonlinear differential equation,

$$\dot{\mathbf{x}}^*(t) = \mathbf{f}[\mathbf{x}^*(t), \mathbf{u}^*(t), t] \quad (24-12)$$

and associated with $\mathbf{x}^*(t)$ and $\mathbf{u}^*(t)$ is the following nominal measurement, $\mathbf{z}^*(t)$, where

$$\mathbf{z}^*(t) = \mathbf{h}[\mathbf{x}^*(t), \mathbf{u}^*(t)] \quad t = t_i \quad i = 1, 2, \dots \quad (24-13)$$

Throughout this lesson, we shall assume that $\mathbf{x}^*(t)$ exists. We discuss two methods for choosing $\mathbf{x}^*(t)$ in Lesson 25. Obviously, one is just to solve (24-12) for $\mathbf{x}^*(t)$.

Note that $\mathbf{x}^*(t)$ must provide a good approximation to the actual behavior of the system. The approximation is considered good if the difference between the nominal and actual solutions can be described by a system of linear differential equations, called *linear perturbation equations*. We derive these equations next.

Let

$$\delta \mathbf{x}(t) = \mathbf{x}(t) - \mathbf{x}^*(t) \quad (24-14)$$

and

$$\delta \mathbf{u}(t) = \mathbf{u}(t) - \mathbf{u}^*(t) \quad (24-15)$$

then,

$$\begin{aligned} \frac{d}{dt} \delta \mathbf{x}(t) &= \delta \dot{\mathbf{x}}(t) = \dot{\mathbf{x}}(t) - \dot{\mathbf{x}}^*(t) = \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t] \\ &\quad + \mathbf{G}(t)\mathbf{w}(t) - \mathbf{f}[\mathbf{x}^*(t), \mathbf{u}^*(t), t] \end{aligned} \quad (24-16)$$

Fact 1. When $\mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t]$ is expanded in a Taylor series about $\mathbf{x}^*(t)$ and $\mathbf{u}^*(t)$, we obtain

$$\begin{aligned}\mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t] &= \mathbf{f}[\mathbf{x}^*(t), \mathbf{u}^*(t), t] + \mathbf{F}_x[\mathbf{x}^*(t), \mathbf{u}^*(t), t]\delta\mathbf{x}(t) \\ &\quad + \mathbf{F}_u[\mathbf{x}^*(t), \mathbf{u}^*(t), t]\delta\mathbf{u}(t) + \text{higher-order terms}\end{aligned}\quad (24-17)$$

where \mathbf{F}_x and \mathbf{F}_u are $n \times n$ and $n \times \ell$ Jacobian matrices, i.e.,

$$\mathbf{F}_x[\mathbf{x}^*(t), \mathbf{u}^*(t), t] = \begin{pmatrix} \frac{\partial f_1}{\partial x_1^*} & \cdots & \frac{\partial f_1}{\partial x_n^*} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1^*} & \cdots & \frac{\partial f_n}{\partial x_n^*} \end{pmatrix} \quad (24-18)$$

and

$$\mathbf{F}_u[\mathbf{x}^*(t), \mathbf{u}^*(t), t] = \begin{pmatrix} \frac{\partial f_1}{\partial u_1^*} & \cdots & \frac{\partial f_1}{\partial u_\ell^*} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial u_1^*} & \cdots & \frac{\partial f_n}{\partial u_\ell^*} \end{pmatrix} \quad (24-19)$$

In these expressions $\frac{\partial f_i}{\partial x_j^*}$ and $\frac{\partial f_i}{\partial u_l^*}$ are short for

$$\frac{\partial f_i}{\partial x_j^*} = \left. \frac{\partial f_i[\mathbf{x}(t), \mathbf{u}(t), t]}{\partial x_j(t)} \right|_{\mathbf{x}(t) = \mathbf{x}^*(t), \mathbf{u}(t) = \mathbf{u}^*(t)} \quad (24-20)$$

and

$$\frac{\partial f_i}{\partial u_l^*} = \left. \frac{\partial f_i[\mathbf{x}(t), \mathbf{u}(t), t]}{\partial u_l(t)} \right|_{\mathbf{x}(t) = \mathbf{x}^*(t), \mathbf{u}(t) = \mathbf{u}^*(t)} \quad (24-21)$$

Proof. The Taylor series expansion of the i th component of $\mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t]$ is

$$\begin{aligned}f_i[\mathbf{x}(t), \mathbf{u}(t), t] &= f_i[\mathbf{x}^*(t), \mathbf{u}^*(t), t] + \frac{\partial f_i}{\partial x_1^*}[x_1(t) - x_1^*(t)] + \cdots \\ &\quad + \frac{\partial f_i}{\partial x_n^*}[x_n(t) - x_n^*(t)] + \frac{\partial f_i}{\partial u_1^*}[u_1(t) - u_1^*(t)] + \cdots \\ &\quad + \frac{\partial f_i}{\partial u_\ell^*}[u_\ell(t) - u_\ell^*(t)] + \text{higher-order terms}\end{aligned}\quad (24-22)$$

where $i = 1, 2, \dots, n$. Collecting these n equations together in vector-matrix format, we obtain (24-17), in which \mathbf{F}_x and \mathbf{F}_u are defined in (24-18) and (24-19), respectively. \square

Substituting (24-17) into (24-16) and neglecting the “higher-order terms,” we obtain the following *perturbation state equation*

$$\begin{aligned}\delta\dot{\mathbf{x}}(t) &= \mathbf{F}_x[\mathbf{x}^*(t), \mathbf{u}^*(t), t]\delta\mathbf{x}(t) \\ &\quad + \mathbf{F}_u[\mathbf{x}^*(t), \mathbf{u}^*(t), t]\delta\mathbf{u}(t) + \mathbf{G}(t)\mathbf{w}(t)\end{aligned}\quad (24-23)$$

Observe that, even if our original nonlinear differential equation is not an explicit function of time [i.e., $\mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t] = \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t)]$], our perturbation state equation is always time-varying because Jacobian matrices \mathbf{F}_x and \mathbf{F}_u vary with time, because \mathbf{x}^* and \mathbf{u}^* vary with time.

Next, let

$$\delta\mathbf{z}(t) = \mathbf{z}(t) - \mathbf{z}^*(t) \quad (24-24)$$

Fact 2. When $\mathbf{h}[\mathbf{x}(t), \mathbf{u}(t), t]$ is expanded in a Taylor series about $\mathbf{x}^*(t)$ and $\mathbf{u}^*(t)$, we obtain

$$\begin{aligned}\mathbf{h}[\mathbf{x}(t), \mathbf{u}(t), t] &= \mathbf{h}[\mathbf{x}^*(t), \mathbf{u}^*(t), t] + \mathbf{H}_x[\mathbf{x}^*(t), \mathbf{u}^*(t), t]\delta\mathbf{x}(t) \\ &\quad + \mathbf{H}_u[\mathbf{x}^*(t), \mathbf{u}^*(t), t]\delta\mathbf{u}(t) + \text{higher-order terms}\end{aligned}\quad (24-25)$$

where \mathbf{H}_x and \mathbf{H}_u are $m \times n$ and $m \times \ell$ Jacobian matrices, i.e.,

$$\mathbf{H}_x[\mathbf{x}^*(t), \mathbf{u}^*(t), t] = \begin{pmatrix} \frac{\partial h_1}{\partial x_1^*} & \cdots & \frac{\partial h_1}{\partial x_n^*} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_m}{\partial x_1^*} & \cdots & \frac{\partial h_m}{\partial x_n^*} \end{pmatrix} \quad (24-26)$$

and

$$\mathbf{H}_u[\mathbf{x}^*(t), \mathbf{u}^*(t), t] = \begin{pmatrix} \frac{\partial h_1}{\partial u_1^*} & \cdots & \frac{\partial h_1}{\partial u_\ell^*} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_m}{\partial u_1^*} & \cdots & \frac{\partial h_m}{\partial u_\ell^*} \end{pmatrix} \quad (24-27)$$

In these expressions $\frac{\partial h_i}{\partial x_j^*}$ and $\frac{\partial h_i}{\partial u_l^*}$ are short for

$$\frac{\partial h_i}{\partial x_j^*} = \left. \frac{\partial h_i[\mathbf{x}(t), \mathbf{u}(t), t]}{\partial x_j(t)} \right|_{\mathbf{x}(t) = \mathbf{x}^*(t), \mathbf{u}(t) = \mathbf{u}^*(t)} \quad (24-28)$$

and

$$\frac{\partial h_i}{\partial u_l^*} = \left. \frac{\partial h_i[\mathbf{x}(t), \mathbf{u}(t), t]}{\partial u_l(t)} \right|_{\mathbf{x}(t) = \mathbf{x}^*(t), \mathbf{u}(t) = \mathbf{u}^*(t)} \quad (24-29)$$

We leave the derivation of this fact to the reader, because it is analogous to the derivation of the Taylor series expansion of $\mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t]$.

Substituting (24-25) into (24-24) and neglecting the “higher-order terms,” we obtain the following *perturbation measurement equation*

$$\begin{aligned}\delta\mathbf{z}(t) &= \mathbf{H}_x[\mathbf{x}^*(t), \mathbf{u}^*(t), t]\delta\mathbf{x}(t) \\ &\quad + \mathbf{H}_u[\mathbf{x}^*(t), \mathbf{u}^*(t), t]\delta\mathbf{u}(t) + \mathbf{v}(t) \quad t = t_i, \quad i = 1, 2, \dots\end{aligned}\quad (24-30)$$

Equations (24-23) and (24-30) constitute our linear perturbation equations, or our *linear perturbation state-variable model*.

Example 24-2

Returning to our satellite-planet Example 24-1, we find that

$$\mathbf{F}_x[\mathbf{x}^*(t), \mathbf{u}^*(t), t] = \mathbf{F}_x[\mathbf{x}^*(t)] = \begin{pmatrix} 0 & 1 & 0 & 0 \\ x_3^2(t) + \frac{2\gamma}{x_1^3(t)} & 0 & 0 & 2x_1(t)x_4(t) \\ 0 & 0 & 0 & 1 \\ \frac{2x_2(t)x_4(t)}{x_1^2(t)} & -\frac{2x_4(t)}{x_1(t)} & 0 & -\frac{2x_2(t)}{x_1(t)} \end{pmatrix}_*$$

$$\mathbf{F}_u[\mathbf{x}^*(t), \mathbf{u}^*(t), t] = \mathbf{F}_u = \begin{pmatrix} 0 & 0 \\ \frac{1}{m} & 0 \\ 0 & 0 \\ 0 & \frac{1}{m} \end{pmatrix}$$

$$\mathbf{H}_x[\mathbf{x}^*(t), \mathbf{u}^*(t), t] = \mathbf{H}_x = (1 \ 0 \ 0 \ 0)$$

and

$$\mathbf{H}_u[\mathbf{x}^*(t), \mathbf{u}^*(t), t] = \mathbf{0}$$

In the equation for $\mathbf{F}_x[\mathbf{x}^*(t)]$, the notation $(\)_*$ means that all $x_i(t)$ within the matrix are nominal values, i.e., $x_i(t) = x_i^*(t)$.

Observe that the linearized satellite-planet system is time-varying, because its linearized plant matrix, $\mathbf{F}_x[\mathbf{x}^*(t)]$, depends upon the nominal trajectory $\mathbf{x}^*(t)$. \square

DISCRETIZATION OF A LINEAR TIME-VARYING STATE-VARIABLE MODEL

In this section we describe how one discretizes the general linear, time-varying state-variable model

$$\dot{\mathbf{x}}(t) = \mathbf{F}(t)\mathbf{x}(t) + \mathbf{C}(t)\mathbf{u}(t) + \mathbf{G}(t)\mathbf{w}(t) \quad (24-31)$$

$$\mathbf{z}(t) = \mathbf{H}(t)\mathbf{x}(t) + \mathbf{v}(t) \quad t = t_i, \quad i = 1, 2, \dots \quad (24-32)$$

The application of this section's results to the perturbation state-variable model is given in the following section.

In (24-31) and (24-32), $\mathbf{x}(t)$ is $n \times 1$, control input $\mathbf{u}(t)$ is $l \times 1$, process noise $\mathbf{w}(t)$ is $p \times 1$, and $\mathbf{z}(t)$ and $\mathbf{v}(t)$ are each $m \times 1$. Additionally, $\mathbf{w}(t)$ is a continuous-time white noise process, $\mathbf{v}(t_i)$ is a discrete-time white noise process, and, $\mathbf{w}(t)$ and $\mathbf{v}(t_i)$ are mutually uncorrelated at all $t = t_i$, $i = 1, 2, \dots$, i.e., $\mathbf{E}\{\mathbf{w}(t)\} = \mathbf{0}$ for all t , $\mathbf{E}\{\mathbf{v}(t_i)\} = \mathbf{0}$ for all t_i , and

$$\mathbf{E}\{\mathbf{w}(t)\mathbf{w}'(\tau)\} = \mathbf{Q}(t)\delta(t - \tau) \quad (24-33)$$

$$\mathbf{E}\{\mathbf{v}(t_i)\mathbf{v}'(t_j)\} = \mathbf{R}(t_i)\delta_{ij} \quad (24-34)$$

and

$$\mathbf{E}\{\mathbf{w}(t)\mathbf{v}'(t_i)\} = \mathbf{0} \quad \text{for } t = t_i \quad i = 1, 2, \dots \quad (24-35)$$

Our approach to discretizing state equation (24-30) begins with the solution of that equation.

Theorem 24-1. *The solution to state equation (24-31) can be expressed as*

$$\mathbf{x}(t) = \Phi(t, t_0)\mathbf{x}(t_0) + \int_{t_0}^t \Phi(t, \tau)[\mathbf{C}(\tau)\mathbf{u}(\tau) + \mathbf{G}(\tau)\mathbf{w}(\tau)]d\tau \quad (24-36)$$

where state transition matrix $\Phi(t, \tau)$ is the solution to the following matrix homogeneous differential equation,

$$\left. \begin{aligned} \dot{\Phi}(t, \tau) &= \mathbf{F}(t)\Phi(t, \tau) \\ \Phi(t, t) &= \mathbf{I} \end{aligned} \right\} \quad \square \quad (24-37)$$

This result should be a familiar one to the readers of this book; hence, we omit its proof.

Next, we assume that $\mathbf{u}(t)$ is a piecewise constant function of time for $t \in [t_k, t_{k+1}]$, and set $t_0 = t_k$ and $t = t_{k+1}$ in (24-36), to obtain

$$\begin{aligned} \mathbf{x}(t_{k+1}) &= \Phi(t_{k+1}, t_k)\mathbf{x}(t_k) + \left[\int_{t_k}^{t_{k+1}} \Phi(t_{k+1}, \tau)\mathbf{C}(\tau)d\tau \right] \mathbf{u}(t_k) \\ &\quad + \int_{t_k}^{t_{k+1}} \Phi(t_{k+1}, \tau)\mathbf{G}(\tau)\mathbf{w}(\tau)d\tau \end{aligned} \quad (24-38)$$

which can also be written as

$$\mathbf{x}(k+1) = \Phi(k+1, k)\mathbf{x}(k) + \Psi(k+1, k)\mathbf{u}(k) + \mathbf{w}_d(k) \quad (24-39)$$

where

$$\Phi(k+1, k) = \Phi(t_{k+1}, t_k) \quad (24-40)$$

$$\Psi(k+1, k) = \int_{t_k}^{t_{k+1}} \Phi(t_{k+1}, \tau)\mathbf{C}(\tau)d\tau \quad (24-41)$$

and $\mathbf{w}_d(k)$ is a discrete-time white Gaussian sequence that is statistically equivalent to

$$\int_{t_k}^{t_{k+1}} \Phi(t_{k+1}, \tau)\mathbf{G}(\tau)\mathbf{w}(\tau)d\tau$$

The mean and covariance matrices of $\mathbf{w}_d(k)$ are

$$\mathbf{E}\{\mathbf{w}_d(k)\} = \mathbf{E}\left\{ \int_{t_k}^{t_{k+1}} \Phi(t_{k+1}, \tau)\mathbf{G}(\tau)\mathbf{w}(\tau)d\tau \right\} = \mathbf{0} \quad (24-42)$$

and

$$\begin{aligned} E\{\mathbf{w}_d(k)\mathbf{w}_d'(k)\} &\triangleq \mathbf{Q}_d(k) \\ &= E\left\{\int_{t_k}^{t_{k+1}} \Phi(t_{k+1}, \tau) \mathbf{G}(\tau) \mathbf{w}(\tau) d\tau \int_{t_k}^{t_{k+1}} \mathbf{w}'(\xi) \mathbf{G}'(\xi) \Phi'(t_{k+1}, \xi) d\xi\right\} \\ &= \int_{t_k}^{t_{k+1}} \Phi(t_{k+1}, \tau) \mathbf{G}(\tau) \mathbf{Q}(\tau) \mathbf{G}'(\tau) \Phi'(t_{k+1}, \tau) d\tau \end{aligned} \quad (24-43)$$

respectively.

Observe, from the right-hand side of Equations (24-40), (24-41), and (24-43), that these quantities can be computed from knowledge about $\mathbf{F}(t)$, $\mathbf{C}(t)$, $\mathbf{G}(t)$, and $\mathbf{Q}(t)$. In general, we must compute $\Phi(k+1, k)$, $\Psi(k+1, k)$, and $\mathbf{Q}_d(k)$ using numerical integration, and, these matrices change from one time interval to the next because $\mathbf{F}(t)$, $\mathbf{C}(t)$, $\mathbf{G}(t)$, and $\mathbf{Q}(t)$ usually change from one time interval to the next.

Because our measurements have been assumed to be available only at sampled values of t , namely at $t = t_i$, $i = 1, 2, \dots$, we can express (24-32) as

$$\mathbf{z}(k+1) = \mathbf{H}(k+1)\mathbf{x}(k+1) + \mathbf{v}(k+1) \quad (24-44)$$

Equations (24-39) and (24-44) constitute our discretized state-variable model.

Example 24-3

Great simplifications of the calculations in (24-40), (24-41) and (24-43) occur if $\mathbf{F}(t)$, $\mathbf{C}(t)$, $\mathbf{G}(t)$, and $\mathbf{Q}(t)$ are approximately constant during the time interval $[t_k, t_{k+1}]$, i.e., if

$$\left. \begin{array}{l} \mathbf{F}(t) \approx \mathbf{F}_k, \quad \mathbf{C}(t) \approx \mathbf{C}_k, \quad \mathbf{G}(t) \approx \mathbf{G}_k, \text{ and} \\ \mathbf{Q}(t) \approx \mathbf{Q}_k \quad \text{for } t \in [t_k, t_{k+1}] \end{array} \right\} \quad (24-45)$$

To begin, (24-37) is easily integrated to yield

$$\Phi(t, \tau) = e^{\mathbf{F}_k(t - \tau)} \quad (24-46)$$

hence,

$$\Phi(k+1, k) = e^{\mathbf{F}_k T} \quad (24-47)$$

where we have assumed that $t_{k+1} - t_k = T$. The matrix exponential is given by the infinite series

$$e^{\mathbf{F}_k T} = \mathbf{I} + \mathbf{F}_k T + \mathbf{F}_k^2 \frac{T^2}{2} + \mathbf{F}_k^3 \frac{T^3}{3!} + \dots \quad (24-48)$$

and, for sufficiently small values of T

$$e^{\mathbf{F}_k T} \approx \mathbf{I} + \mathbf{F}_k T \quad (24-49)$$

We use this approximation for $e^{\mathbf{F}_k T}$ in deriving simpler expressions for $\Psi(k+1, k)$ and $\mathbf{Q}_d(k)$. Comparable results can be obtained for higher-order truncations of $e^{\mathbf{F}_k T}$.

Substituting (24-46) into (24-41), we find that

$$\begin{aligned} \Psi(k+1, k) &= \int_{t_k}^{t_{k+1}} \Phi(t_{k+1}, \tau) \mathbf{C}_k d\tau = \int_{t_k}^{t_{k+1}} e^{\mathbf{F}_k(t_{k+1} - \tau)} \mathbf{C}_k d\tau \\ &\approx \int_{t_k}^{t_{k+1}} [\mathbf{I} + \mathbf{F}_k(t_{k+1} - \tau)] \mathbf{C}_k d\tau \\ &= \mathbf{C}_k T + \mathbf{F}_k \mathbf{C}_k t_{k+1} - \mathbf{F}_k \mathbf{C}_k \int_{t_k}^{t_{k+1}} \tau d\tau \\ &= \mathbf{C}_k T + \mathbf{F}_k \mathbf{C}_k \frac{T^2}{2} = \mathbf{C}_k T \end{aligned} \quad (24-50)$$

where we have truncated $\Psi(k+1, k)$ to its first-order term in T . Proceeding in a similar manner for $\mathbf{Q}_d(k)$, it is straightforward to show that

$$\mathbf{Q}_d(k) \approx \mathbf{G}_k \mathbf{Q}_k \mathbf{G}_k' \quad (24-51)$$

Note that (24-47), (24-49), (24-50), and (24-51), while much simpler than their original expressions, can change in values from one time-interval to another, because of their dependence upon k . \square

DISCRETIZED PERTURBATION STATE-VARIABLE MODEL

Applying the results of the preceding section to the perturbation state-variable model in (24-23) and (24-30), we obtain the following *discretized perturbation state-variable model*

$$\delta \mathbf{x}(k+1) = \Phi(k+1, k; *) \delta \mathbf{x}(k) + \Psi(k+1, k; *) \delta \mathbf{u}(k) + \mathbf{w}_d(k) \quad (24-52)$$

$$\begin{aligned} \delta \mathbf{z}(k+1) &= \mathbf{H}_x(k+1; *) \delta \mathbf{x}(k+1) \\ &\quad + \mathbf{H}_u(k+1; *) \delta \mathbf{u}(k+1) + \mathbf{v}(k+1) \end{aligned} \quad (24-53)$$

The notation $\Phi(k+1, k; *)$, for example, denotes the fact that this matrix depends on $\mathbf{x}^*(t)$ and $\mathbf{u}^*(t)$. More specifically,

$$\Phi(k+1, k; *) = \Phi(t_{k+1}, t_k; *) \quad (24-54)$$

where

$$\begin{aligned} \Phi(t, \tau; *) &= \mathbf{F}_x[\mathbf{x}^*(t), \mathbf{u}^*(t), t] \Phi(t, \tau; *) \\ \Phi(t, t; *) &= \mathbf{I} \end{aligned} \quad (24-55)$$

Additionally,

$$\Psi(k+1, k; *) = \int_{t_k}^{t_{k+1}} \Phi(t_{k+1}, \tau; *) \mathbf{F}_u[\mathbf{x}^*(\tau), \mathbf{u}^*(\tau), \tau] d\tau \quad (24-56)$$

and

$$\mathbf{Q}_d(k; *) = \int_{t_k}^{t_{k+1}} \Phi(t_{k+1}, \tau; *) \mathbf{G}(\tau) \mathbf{Q}(\tau) \mathbf{G}'(\tau) \Phi'(t_{k+1}, \tau; *) d\tau \quad (24-57)$$

P R O B L E M S

- 24-1. Derive the Taylor series expansion of $h[x(t), u(t), t]$ given in (24-25).
- 24-2. Derive the formula for $Q_d(k)$ given in (24-51).
- 24-3. Derive formulas for $\Psi(k+1, k)$ and $Q_d(k)$ that include first- and second-order effects of T , using the first three terms in the expansion of e^{fkT} .
- 24-4. Let a zero-mean stationary Gaussian random process $v(t)$ have the auto-correlation function $\phi_v(\tau)$ given by $\phi_v(\tau) = e^{-|\tau|} + e^{-2|\tau|}$.
 (a) Show that this colored-noise process can be generated by passing white noise $\mu(t)$ through the linear system whose transfer function is

$$\sqrt{6} \frac{s + \sqrt{2}}{(s + 1)(s + 2)}$$

- (b) Obtain a discrete-time state-variable model for this colored noise process (assume $T = 1$ msec).
 24-5. This problem presents a model for estimation of the altitude, velocity, and constant ballistic coefficient of a vertically falling body (Athans, et al., 1968). The measurements are made at discrete instants of time by a radar that measures range in the presence of discrete-time white Gaussian noise. The state equations for the falling body are

$$\begin{aligned}\dot{x}_1 &= -x_2 \\ \dot{x}_2 &= -e^{-\gamma x_1} x_2^2 x_3 \\ \dot{x}_3 &= 0\end{aligned}$$

where $\gamma = 5 \times 10^{-5}$, $x_1(t)$ is altitude, $x_2(t)$ is downward velocity, and x_3 is a constant ballistic parameter. Measured range is given by

$$z(k) = \sqrt{M^2 + [x_1(k) - H]^2} + v(k) \quad k = 1, 2, \dots,$$

where M is horizontal distance and H is radar altitude. Obtain the discretized perturbation state-variable model for this system.

- 24-6. Normalized equations of a stirred reactor are

$$\begin{aligned}\dot{x}_1 &= -(c_1 + c_4)x_1 + c_3(1 + x_4)^2 \exp[K_1 x_1 / (1 + x_1)] + c_4 x_2 + c_1 u_1 \\ \dot{x}_2 &= -(c_5 + c_6)x_2 + c_5 x_1 + c_6 x_3 \\ \dot{x}_3 &= -(c_7 + c_8)x_3 + c_8 x_2 + c_7 u_2 \\ \dot{x}_4 &= -c_1 x_4 - c_2(1 + x_4)^2 \exp[K_1 x_1 / (1 + x_1)] + c_1 u_3\end{aligned}$$

in which u_1 , u_2 , and u_3 are control inputs. Measurements are

$$z_i(k) = x_i(k) + v_i(k) \quad i = 1, 2, 3,$$

where $v_i(k)$ are zero-mean white Gaussian noises with variances r_i . Obtain the discretized perturbation state-variable model for this system.

- 24-7. Obtain the discretized perturbation-state equation for each of the following nonlinear systems:

- (a) Equation for the unsteady operation of a synchronous motor:

$$\ddot{x}(t) + C\dot{x}(t) + p \sin x(t) = L(t)$$

- (b) Duffing's equation:

$$\ddot{x}(t) + C\dot{x}(t) + \alpha x(t) + \beta x^3(t) = F \cos \omega t$$

- (c) Van der Pol's equation:

$$\ddot{x}(t) - \epsilon \dot{x}(t) \left[1 - \frac{1}{3} \dot{x}^2(t) \right] + x(t) = m(t)$$

- (d) Hill's equation:

$$\ddot{x}(t) - ax(t) + bp(t)x(t) = m(t)$$

where $p(t)$ is a known periodic function.

Iterated Least Squares and Extended Kalman Filtering

INTRODUCTION

This lesson is primarily devoted to the extended Kalman filter (EKF), which is a form of the Kalman filter “extended” to nonlinear dynamical systems of the type described in Lesson 24. We shall show that the EKF is related to the method of iterated least squares (ILS), the major difference being that the EKF is for dynamical systems whereas ILS is not.

ITERATED LEAST SQUARES

We shall illustrate the method of ILS for the nonlinear model described in Example 2-5 of Lesson 2, i.e., for the model

$$z(k) = f(\theta, k) + v(k) \quad (25-1)$$

where $k = 1, 2, \dots, N$.

Iterated least squares is basically a four step procedure.

1. Linearize $f(\theta, k)$ about a nominal value of θ, θ^* . Doing this, we obtain the *perturbation measurement equation*

$$\delta z(k) = F_\theta(k; \theta^*) \delta \theta + v(k) \quad k = 1, 2, \dots, N \quad (25-2)$$

where

$$\delta z(k) = z(k) - z^*(k) = z(k) - f(\theta^*, k) \quad (25-3)$$

$$\delta \theta = \theta - \theta^* \quad (25-4)$$

and

$$F_\theta(k; \theta^*) = \frac{\partial f(\theta, k)}{\partial \theta} \Big|_{\theta = \theta^*} \quad (25-5)$$

2. Concatenate (25-2) and compute $\hat{\delta}\theta_{WLS}(N)$ [or $\hat{\delta}\theta_{LS}(N)$] using our Lesson 3 formulas.
3. Solve the equation

$$\hat{\delta}\theta_{WLS}(N) = \hat{\theta}_{WLS}(N) - \theta^* \quad (25-6)$$

for $\hat{\theta}_{WLS}(N)$, i.e.,

$$\hat{\theta}_{WLS}(N) = \theta^* + \hat{\delta}\theta_{WLS}(N) \quad (25-7)$$

4. Replace θ^* with $\hat{\theta}_{WLS}(N)$ and return to Step 1. Iterate through these steps until convergence occurs. Let $\hat{\theta}_{WLS}^i(N)$ and $\hat{\theta}_{WLS}^{i+1}(N)$ denote estimates of θ obtained at the i th and $(i+1)$ st iterations, respectively. Convergence of the ILS method occurs when

$$|\hat{\theta}_{WLS}^{i+1}(N) - \hat{\theta}_{WLS}^i(N)| < \epsilon \quad (25-8)$$

where ϵ is a prespecified small positive number.

We observe, from this four-step procedure, that ILS uses the estimate obtained from the linearized model to generate the nominal value of θ about which the nonlinear model is *relinearized*. Additionally, in each complete cycle of this procedure, we use both the nonlinear and linearized models. The nonlinear model is used to compute $z^*(k)$, and subsequently $\delta z(k)$, using (25-3).

The notions of relinearizing about a filter output and using both the nonlinear and linearized models are also at the very heart of the EKF.

EXTENDED KALMAN FILTER

The nonlinear dynamical system of interest to us is the one described in Lesson 24. For convenience to the reader, we summarize aspects of that system next. The nonlinear state-variable model is

$$\dot{x}(t) = f[x(t), u(t), t] + G(t)w(t) \quad (25-9)$$

$$z(t) = h[x(t), u(t), t] + v(t) \quad t = t_i \quad i = 1, 2, \dots \quad (25-10)$$

Given a nominal input, $u^*(t)$, and assuming that a nominal trajectory, $x^*(t)$, exists, $x^*(t)$ and its associated nominal measurement satisfy the following nominal system model,

$$\dot{x}^*(t) = f[x^*(t), u^*(t), t] \quad (25-11)$$

$$z^*(t) = h[x^*(t), u^*(t), t] \quad t = t_i \quad i = 1, 2, \dots \quad (25-12)$$

Letting $\delta x(t) = x(t) - x^*(t)$, $\delta u(t) = u(t) - u^*(t)$, and $\delta z(t) = z(t) - z^*(t)$, we also have the following *discretized perturbation state-variable model* that is associated with a linearized version of the original nonlinear state-variable model,

$$\delta x(k+1) = \Phi(k+1, k; *) \delta x(k) + \Psi(k+1, k; *) \delta u(k) + w_d(k) \quad (25-13)$$

$$\begin{aligned} \delta z(k+1) &= H_x(k+1; *) \delta x(k+1) \\ &\quad + H_u(k+1; *) \delta u(k+1) + v(k+1) \end{aligned} \quad (25-14)$$

In deriving (25-13) and (25-14), we made the important assumption that higher-order terms in the Taylor series expansions of $f[x(t), u(t), t]$ and $h[x(t), u(t), t]$ could be neglected. Of course, this is only correct as long as $x(t)$ is “close” to $x^*(t)$ and $u(t)$ is “close” to $u^*(t)$.

If $u(t)$ is an input derived from a feedback control law, so that $u(t) = u[x(t), t]$, then $u(t)$ can differ from $u^*(t)$, because $x(t)$ will differ from $x^*(t)$. On the other hand, if $u(t)$ does not depend on $x(t)$ then usually $u(t)$ is the same as $u^*(t)$, in which case $\delta u(t) = 0$. We see, therefore, that $x^*(t)$ is the critical quantity in the calculation of our discretized perturbation state-variable model.

Suppose $x^*(t)$ is given a priori; then we can compute predicted, filtered, or smoothed estimates of $\delta x(k)$ by applying all of our previously derived estimators to the discretized perturbation state-variable model in (25-13) and (25-14). We can precompute $x^*(t)$ by solving the nominal differential equation (25-11). The Kalman filter associated with using a precomputed $x^*(t)$ is known as a *relinearized KF*.

A relinearized KF usually gives poor results, because it relies on an open-loop strategy for choosing $x^*(t)$. When $x^*(t)$ is precomputed there is no way of forcing $x^*(t)$ to remain close to $x(t)$, and this must be done or else the perturbation state-variable model is invalid. Divergence of the relinearized KF often occurs; hence, we do not recommend the relinearized KF.

The relinearized KF is based only on the discretized perturbation state-variable model. It does not use the nonlinear nature of the original system in an active manner. The extended Kalman filter relinearizes the nonlinear system about each new estimate as it becomes available; i.e., at $k = 0$, the system is linearized about $\hat{x}(0|0)$. Once $z(1)$ is processed by the EKF, so that $\hat{x}(1|1)$ is obtained, the system is linearized about $\hat{x}(1|1)$. By “linearize about $\hat{x}(1|1)$,” we mean $\hat{x}(1|1)$ is used to calculate all the quantities needed to make the transition from $\hat{x}(1|1)$ to $\hat{x}(2|1)$, and subsequently $\hat{x}(2|2)$. This phrase will become clear below. The purpose of relinearizing about the filter’s output is to use a better reference trajectory for $x^*(t)$. Doing this, $\delta x = x - \hat{x}$ will be held as small as possible, so that our linearization assumptions are less likely to be violated than in the case of the relinearized KF.

The EKF is developed below in predictor-corrector format (Jazwinski, 1970). Its prediction equation is obtained by integrating the nominal differential equation for $x^*(t)$, from t_k to t_{k+1} . In order to do this, we need to know

how to choose $x^*(t)$ for the entire interval of time $t \in [t_k, t_{k+1}]$. Thus far, we have only mentioned how $x^*(t)$ is chosen at t_k , i.e., as $\hat{x}(k|k)$.

Theorem 25-1. *As a consequence of relinearizing about $\hat{x}(k|k)$ ($k = 0, 1, \dots$),*

$$\delta \hat{x}(t|t_k) = \mathbf{0} \quad \text{for all } t \in [t_k, t_{k+1}] \quad (25-15)$$

This means that

$$x^*(t) = \hat{x}(t|t_k) \text{ for all } t \in [t_k, t_{k+1}] \quad (25-16)$$

Before proving this important result, we observe that it provides us with a choice of $x^*(t)$ over the entire interval of time $t \in [t_k, t_{k+1}]$, and, it states that at the left-hand side of this time interval $x^*(t_k) = \hat{x}(k|k)$, whereas at the right-hand side of this time interval $x^*(t_{k+1}) = \hat{x}(k+1|k)$. The transition from $\hat{x}(k+1|k)$ to $\hat{x}(k+1|k+1)$ will be made using the EKF’s correction equation.

Proof. Let t_l be an arbitrary value of t lying in the interval between t_k and t_{k+1} (see Figure 25-1). For the purposes of this derivation, we can assume that $\delta u(k) = \mathbf{0}$ [i.e., perturbation input $\delta u(k)$ takes on no new values in the interval from t_k to t_{k+1} ; recall the piecewise-constant assumption made about $u(t)$ in the derivation of (24-37)], i.e.,

$$\delta x(k+1) = \Phi(k+1, k; *) \delta x(k) + w_d(k) \quad (25-17)$$

Using our general state-predictor results given in (16-14), we see that (remember that k is short for t_k , and that $t_{k+1} - t_k$ does not have to be a constant; this is true in all of our predictor, filter, and smoother formulas)

$$\begin{aligned} \delta \hat{x}(t_l|t_k) &= \Phi(t_l, t_k; *) \delta \hat{x}(t_k|t_k) \\ &= \Phi(t_l, t_k; *) [\hat{x}(k|k) - x^*(k)] \end{aligned} \quad (25-18)$$

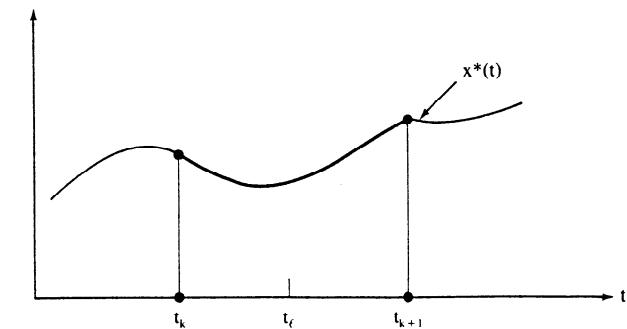


Figure 25-1 Nominal state trajectory $x^*(t)$.

In the EKF we set $\mathbf{x}^*(k) = \hat{\mathbf{x}}(k|k)$; thus, when this is done,

$$\delta\hat{\mathbf{x}}(t_l|t_k) = \mathbf{0} \quad (25-19)$$

and, because $t_l \in [t_k, t_{k+1}]$,

$$\delta\hat{\mathbf{x}}(t_l|t_k) = \mathbf{0} \quad \text{for all } t_l \in [t_k, t_{k+1}] \quad (25-20)$$

which is (25-15). Equation (25-16) follows from (25-20) and the fact that $\delta\hat{\mathbf{x}}(t_l|t_k) = \hat{\mathbf{x}}(t_l|t_k) - \mathbf{x}^*(t_l)$. \square

We are now able to derive the EKF. As mentioned above, the EKF must be obtained in predictor-corrector format. We begin the derivation by obtaining the predictor equation for $\hat{\mathbf{x}}(k+1|k)$.

Recall that $\mathbf{x}^*(t)$ is the solution of the nominal state equation (25-11). Using (25-16) in (25-11), we find that

$$\frac{d}{dt}\hat{\mathbf{x}}(t|t_k) = \mathbf{f}[\hat{\mathbf{x}}(t|t_k), \mathbf{u}^*(t), t] \quad (25-21)$$

Integrating this equation from $t = t_k$ to $t = t_{k+1}$, we obtain

$$\hat{\mathbf{x}}(k+1|k) = \hat{\mathbf{x}}(k|k) + \int_{t_k}^{t_{k+1}} \mathbf{f}[\hat{\mathbf{x}}(t|t_k), \mathbf{u}^*(t), t] dt \quad (25-22)$$

which is the *EKF prediction equation*. Observe that the nonlinear nature of the system's state equation is used to determine $\hat{\mathbf{x}}(k+1|k)$. The integral in (25-22) is evaluated by means of numerical integration formulas that are initialized by $\mathbf{f}[\hat{\mathbf{x}}(t_k|t_k), \mathbf{u}^*(t_k), t_k]$.

The corrector equation for $\hat{\mathbf{x}}(k+1|k+1)$ is obtained from the Kalman filter associated with the discretized perturbation state-variable model in (25-13) and (25-14), and is

$$\begin{aligned} \delta\hat{\mathbf{x}}(k+1|k+1) &= \delta\hat{\mathbf{x}}(k+1|k) + \mathbf{K}(k+1;*)[\delta\mathbf{z}(k+1) \\ &\quad - \mathbf{H}_x(k+1;*)\delta\hat{\mathbf{x}}(k+1|k) - \mathbf{H}_u(k+1;*)\delta\mathbf{u}(k+1)] \end{aligned} \quad (25-23)$$

As a consequence of relinearizing about $\hat{\mathbf{x}}(k|k)$, we know that

$$\delta\hat{\mathbf{x}}(k+1|k) = \mathbf{0} \quad (25-24)$$

$$\begin{aligned} \delta\hat{\mathbf{x}}(k+1|k+1) &= \hat{\mathbf{x}}(k+1|k+1) - \mathbf{x}^*(k+1) \\ &= \hat{\mathbf{x}}(k+1|k+1) - \hat{\mathbf{x}}(k+1|k) \end{aligned} \quad (25-25)$$

and

$$\begin{aligned} \delta\mathbf{z}(k+1) &= \mathbf{z}(k+1) - \mathbf{z}^*(k+1) \\ &= \mathbf{z}(k+1) - \mathbf{h}[\mathbf{x}^*(k+1), \mathbf{u}^*(k+1), k+1] \\ &= \mathbf{z}(k+1) - \mathbf{h}[\hat{\mathbf{x}}(k+1|k), \mathbf{u}^*(k+1), k+1] \end{aligned} \quad (25-26)$$

Substituting these three equations into (25-23) we obtain

$$\begin{aligned} \hat{\mathbf{x}}(k+1|k+1) &= \hat{\mathbf{x}}(k+1|k) + \mathbf{K}(k+1;*)[\mathbf{z}(k+1) \\ &\quad - \mathbf{h}[\hat{\mathbf{x}}(k+1|k), \mathbf{u}^*(k+1), k+1] - \mathbf{H}_u(k+1;*)\delta\mathbf{u}(k+1)] \end{aligned} \quad (25-27)$$

which is the *EKF correction equation*. Observe that the nonlinear nature of the system's measurement equation is used to determine $\hat{\mathbf{x}}(k+1|k+1)$. One usually sees this equation for the case when $\delta\mathbf{u} = \mathbf{0}$, in which case the last term on the right-hand side of (25-27) is not present.

In order to compute $\hat{\mathbf{x}}(k+1|k+1)$, we must compute the EKF gain matrix $\mathbf{K}(k+1;*)$. This matrix, as well as its associated $\mathbf{P}(k+1|k;*)$ and $\mathbf{P}(k+1|k+1;*)$ matrices, depends upon the nominal $\mathbf{x}^*(t)$ that results from prediction, namely $\hat{\mathbf{x}}(k+1|k)$. Observe, from (25-16), that $\mathbf{x}^*(k+1) = \hat{\mathbf{x}}(k+1|k)$, and that the argument of \mathbf{K} in the correction equation is $k+1$; hence, we are indeed justified to use $\hat{\mathbf{x}}(k+1|k)$ as the nominal value of \mathbf{x}^* during the calculations of $\mathbf{K}(k+1;*)$, $\mathbf{P}(k+1|k;*)$, and $\mathbf{P}(k+1|k+1;*)$. These three quantities are computed from

$$\begin{aligned} \mathbf{K}(k+1;*) &= \mathbf{P}(k+1|k;*)\mathbf{H}_x'(k+1;*)[\mathbf{H}_x(k+1;*) \\ &\quad \mathbf{P}(k+1|k;*)\mathbf{H}_x'(k+1;*) + \mathbf{R}(k+1)]^{-1} \end{aligned} \quad (25-28)$$

$$\mathbf{P}(k+1|k;*) = \Phi(k+1, k;*)\mathbf{P}(k|k;*)\Phi'(k+1, k;*) + \mathbf{Q}_d(k;*) \quad (25-29)$$

$$\mathbf{P}(k+1|k+1;*) = [\mathbf{I} - \mathbf{K}(k+1;*)\mathbf{H}_x(k+1;*)]\mathbf{P}(k+1|k;*) \quad (25-30)$$

Remember that in these three equations * denotes the use of $\hat{\mathbf{x}}(k+1|k)$.

The EKF is very widely used, especially in the aerospace industry; however, it does not provide an optimal estimate of $\mathbf{x}(k)$. The optimal estimate of $\mathbf{x}(k)$ is still $\mathbb{E}\{\mathbf{x}(k)|\mathcal{Z}(k)\}$, regardless of the linear or nonlinear nature of the system's model. The EKF is a first-order approximation of $\mathbb{E}\{\mathbf{x}(k)|\mathcal{Z}(k)\}$ that sometimes works quite well, but cannot be guaranteed always to work well. No convergence results are known for the EKF; hence, the EKF must be viewed as an ad hoc filter. Alternatives to the EKF, which are based on nonlinear filtering, are quite complicated and are rarely used.

The EKF is designed to work well as long as $\delta\mathbf{x}(k)$ is "small." The *iterated EKF* (Jazwinski, 1970), depicted in Figure 25-2, is designed to keep $\delta\mathbf{x}(k)$ as small as possible. The iterated EKF differs from the EKF in that it iterates the correction equation L times until $\|\hat{\mathbf{x}}_L(k+1|k+1) - \hat{\mathbf{x}}_{L-1}(k+1|k+1)\| \leq \epsilon$. Corrector #1 computes $\mathbf{K}(k+1;*)$, $\mathbf{P}(k+1|k;*)$, and $\mathbf{P}(k+1|k+1;*)$ using $\mathbf{x}^* = \hat{\mathbf{x}}(k+1|k)$; corrector #2 computes these quantities using $\mathbf{x}^* = \hat{\mathbf{x}}_1(k+1|k+1)$; corrector #3 computes these quantities using $\mathbf{x}^* = \hat{\mathbf{x}}_2(k+1|k+1)$; etc.

Often, just adding one additional corrector (i.e., $L = 2$) leads to substantially better results for $\hat{\mathbf{x}}(k+1|k+1)$ than are obtained using the EKF.

APPLICATION TO PARAMETER ESTIMATION

One of the earliest applications of the extended Kalman filter was to parameter estimation (Kopp and Orford, 1963). Consider the continuous-time linear system

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{w}(t) \quad (25-31a)$$

$$\mathbf{z}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{v}(t) \quad t = t_i \quad i = 1, 2, \dots \quad (25-31b)$$

Matrices \mathbf{A} and \mathbf{H} contain some unknown parameters, and our objective is to estimate these parameters from the measurements $\mathbf{z}(t_i)$ as they become available.

To begin, we assume differential equation models for the unknown parameters, i.e., either

$$\dot{a}_l(t) = 0 \quad l = 1, 2, \dots, l^* \quad (25-32a)$$

$$\dot{h}_j(t) = 0 \quad j = 1, 2, \dots, j^* \quad (25-32b)$$

or

$$\dot{a}_l(t) = c_l a_l(t) + n_l(t) \quad l = 1, 2, \dots, l^* \quad (25-33a)$$

$$\dot{h}_j(t) = d_j h_j(t) + \eta_j(t) \quad j = 1, 2, \dots, j^* \quad (25-33b)$$

In the latter models $n_l(t)$ and $\eta_j(t)$ are white noise processes, and one often chooses $c_l = 0$ and $d_j = 0$. The noises $n_l(t)$ and $\eta_j(t)$ introduce uncertainty about the “constancy” of the a_l and h_j parameters.

Next, we augment the parameter differential equations to (25-31a) and (25-31b). The resulting system is nonlinear, because it contains products of states [e.g., $a_l(t)x_i(t)$]. The augmented system can be expressed as in (25-9) and (25-10), which means we have reduced the problem of parameter estimation in a linear system to state estimation in a nonlinear system.

Finally, we apply the EKF to the augmented state-variable model to obtain $\hat{a}_l(k|k)$ and $\hat{h}_j(k|k)$.

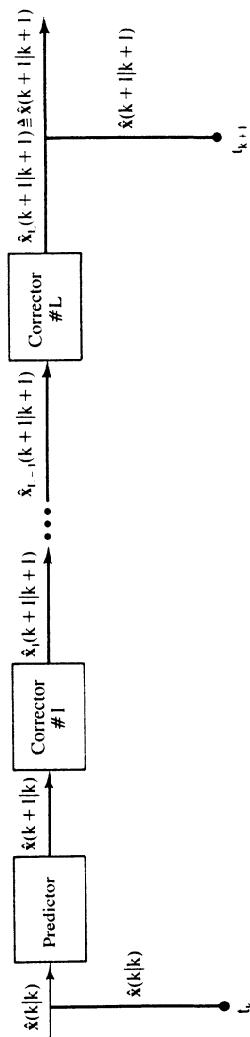
Ljung (1979) has studied the convergence properties of the EKF applied to parameter estimation, and has shown that parameter estimates do not converge to their true values. He shows that another term must be added to the EKF corrector equation in order to guarantee convergence. For details, see his paper.

Example 25-1

Consider the satellite and planet Example 24-1, in which the satellite's motion is governed by the two equations

$$\ddot{r} = r\dot{\theta}^2 - \frac{\gamma}{r^2} + \frac{1}{m} u_r(t) \quad (25-34)$$

Figure 25-2 Iterated EKF. All of the calculations provide us with a refined estimate of $\mathbf{x}(k+1)$, $\hat{\mathbf{x}}(k+1|k+1)$, starting with $\hat{\mathbf{x}}(k|k)$.



and

$$\ddot{\theta} = -\frac{2r\dot{\theta}}{r} + \frac{1}{m} u_\theta(t) \quad (25-35)$$

We shall assume that m and γ are unknown constants, and shall model them as

$$\dot{m}(t) = 0 \quad (25-36)$$

$$\dot{\gamma}(t) = 0 \quad (25-37)$$

Defining $x_1 = r$, $x_2 = \dot{r}$, $x_3 = \theta$, $x_4 = \dot{\theta}$, $x_5 = m$, $x_6 = \gamma$, $u_1 = u_r$, and $u_2 = u_\theta$, we have

$$\left. \begin{array}{l} \dot{x}_1 = x_2 \\ \dot{x}_2 = x_1 x_4^2 - \frac{x_6}{x_1^2} + \frac{1}{x_5} u_1(t) \\ \dot{x}_3 = x_4 \\ \dot{x}_4 = -\frac{2x_2 x_4}{x_1} + \frac{1}{x_5} u_2(t) \\ \dot{x}_5 = 0 \\ \dot{x}_6 = 0 \end{array} \right\} \quad (25-38)$$

Now,

$$\mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t] = \text{col} \left[x_2, x_1 x_4^2 - \frac{x_6}{x_1^2} + \frac{1}{x_5} u_1, x_4, -\frac{2x_2 x_4}{x_1} + \frac{1}{x_5} u_2, 0, 0 \right] \quad \square \quad (25-39)$$

We note, finally, that the modeling and augmentation approach to parameter estimation, described above, is not restricted to continuous-time linear systems. Additional situations are described in the exercises.

PROBLEMS

- 25-1.** In the first-order system, $x(k+1) = ax(k) + w(k)$, and $z(k+1) = x(k+1) + v(k+1)$, $k = 1, 2, \dots, N$, a is an unknown parameter that is to be estimated. Sequences $w(k)$ and $v(k)$ are, as usual, mutually uncorrelated and white, and, $w(k) \sim N(w(k); 0, 1)$ and $v(k) \sim N(v(k); 0, 1/2)$. Explain, using equations and a flowchart, how parameter a can be estimated using an EKF.
- 25-2.** Repeat the preceding problem where all conditions are the same except that now $w(k)$ and $v(k)$ are correlated, and $E\{w(k)v(k)\} = 1/4$.
- 25-3.** The system of differential equations describing the motion of an aerospace vehicle about its pitch axis can be written as (Kopp and Orford, 1963)

$$\begin{aligned} \dot{x}_1(t) &= x_2(t) \\ \dot{x}_2(t) &= a_1(t)x_2(t) + a_2(t)x_1(t) + a_3(t)u(t) \end{aligned}$$

where $x_1 = \dot{\theta}(t)$, which is the actual pitch rate. Sampled measurements are made of the pitch rate, i.e.,

$$z(t) = x_1(t) + v(t) \quad t = t_i \quad i = 1, 2, \dots, N$$

Noise $v(t_i)$ is white and Gaussian, and $\sigma_v^2(t_i)$ is given. The control signal $u(t)$ is the sum of a desired control signal $u^*(t)$ and additive noise, i.e.,

$$u(t) = u^*(t) + \delta u(t)$$

The additive noise $\delta u(t)$ is a normally distributed random variable modulated by a function of the desired control signal, i.e.,

$$\delta u(t) = S[u^*(t)]w_0(t)$$

where $w_0(t)$ is zero-mean white noise with intensity $\sigma_{w_0}^2$. Parameters $a_1(t)$, $a_2(t)$ and $a_3(t)$ may be unknown and are modeled as

$$\dot{a}_i(t) = \alpha_i(t)[a_i(t) - \bar{a}_i(t)] + w_i(t) \quad i = 1, 2, 3$$

In this model the parameters $\alpha_i(t)$ are assumed given, as are the a priori values of $a_i(t)$ and $\bar{a}_i(t)$, and, $w_i(t)$ are zero-mean white noises with intensities $\sigma_{w_i}^2$.

- (a) What are the EKF formulas for estimation of x_1 , x_2 , a_1 , and a_2 , assuming that a_3 is known?

- (b) Repeat (a) but now assume that a_3 is unknown.

- 25-4.** Suppose we begin with the nonlinear discrete-time system,

$$\mathbf{x}(k+1) = \mathbf{f}[\mathbf{x}(k), k] + \mathbf{w}(k)$$

$$\mathbf{z}(k) = \mathbf{h}[\mathbf{x}(k), k] + \mathbf{v}(k) \quad k = 1, 2, \dots$$

Develop the EKF for this system [Hint: expand $\mathbf{f}[\mathbf{x}(k), k]$ and $\mathbf{h}[\mathbf{x}(k), k]$ in Taylor series about $\hat{\mathbf{x}}(k|k)$ and $\hat{\mathbf{x}}(k|k-1)$, respectively].

- 25-5.** Refer to Problem 24-7. Obtain the EKF for

- (a) Equation for the unsteady operation of a synchronous motor, in which C and p are unknown;
 (b) Duffing's equation, in which C , α , and β are unknown;
 (c) Van der Pol's equation, in which ϵ is unknown; and
 (d) Hill's equation, in which a and b are unknown.

Lesson 26

Maximum-Likelihood State and Parameter Estimation

INTRODUCTION

In Lesson 11 we studied the problem of obtaining maximum-likelihood estimates of a collection of parameters, $\theta = \text{col}(\text{elements of } \Phi, \Psi, H, \text{ and } R)$, that appear in the state-variable model

$$x(k+1) = \Phi x(k) + \Psi u(k) \quad (26-1)$$

$$z(k+1) = Hx(k+1) + v(k+1) \quad k = 0, 1, \dots, N-1 \quad (26-2)$$

We determined the log-likelihood function to be

$$L(\theta|\mathcal{Z}) = -\frac{1}{2} \sum_{i=1}^N [z(i) - H_\theta x_\theta(i)]' R_\theta^{-1} [z(i) - H_\theta x_\theta(i)] - \frac{N}{2} \ln |R_\theta| \quad (26-3)$$

where quantities that are subscripted θ denote a dependence on θ . Finally, we pointed out that the state equation (26-1), written as

$$x_\theta(k+1) = \Phi_\theta x_\theta(k) + \Psi_\theta u(k) \quad x_\theta(0) \text{ known} \quad (26-4)$$

acts as a constraint that is associated with the computation of the log-likelihood function. Parameter vector θ must be determined by maximizing $L(\theta|\mathcal{Z})$ subject to the constraint (26-4). This can only be done using mathematical programming techniques (i.e., an optimization algorithm such as steepest descent or Marquardt-Levenberg).

In this lesson we study the problem of obtaining maximum-likelihood

A Log-Likelihood Function for the Basic State-Variable Model

259

estimates of a collection of parameters, also denoted θ , that appear in our basic state-variable model,

$$x(k+1) = \Phi x(k) + \Gamma w(k) + \Psi u(k) \quad (26-5)$$

$$z(k+1) = Hx(k+1) + v(k+1) \quad k = 0, 1, \dots, N-1 \quad (26-6)$$

Now, however,

$$\theta = \text{col}(\text{elements of } \Phi, \Gamma, \Psi, H, Q, \text{ and } R) \quad (26-7)$$

and, we assume that θ is $d \times 1$. As in Lesson 11, we shall assume that θ is identifiable.

Before we can determine $\hat{\theta}_{ML}$ we must establish the log-likelihood function for our basic state-variable model.

A LOG-LIKELIHOOD FUNCTION FOR THE BASIC STATE-VARIABLE MODEL

As always, we must compute $p(\mathcal{Z}|\theta) = p(z(1), z(2), \dots, z(N)|\theta)$. This is difficult to do for the basic state-variable model, because

$$p(z(1), z(2), \dots, z(N)|\theta) \neq p(z(1)|\theta)p(z(2)|\theta) \cdots p(z(N)|\theta) \quad (26-8)$$

The measurements are all correlated due to the presence of either the process noise, $w(k)$, or random initial conditions or both. This represents the major difference between our basic state-variable model, (26-5) and (26-6), and the state-variable model studied earlier, in (26-1) and (26-2). Fortunately, the measurements and innovations are causally invertible, and the innovations are all uncorrelated, so that it is still relatively easy to determine the log-likelihood function for the basic state-variable model.

Theorem 26-1. *The log-likelihood function for our basic state-variable model in (26-5) and (26-6) is*

$$L(\theta|\mathcal{Z}) = -\frac{1}{2} \sum_{j=1}^N [\tilde{z}_\theta(j|j-1) N_\theta^{-1}(j|j-1) \tilde{z}_\theta(j|j-1) + \ln N_\theta(j|j-1)] \quad (26-9)$$

where $\tilde{z}_\theta(j|j-1)$ is the innovations process, and $N_\theta(j|j-1)$ is the covariance of that process [in Lesson 16 we used the symbol $P_{zz}(j|j-1)$ for this covariance],

$$N_\theta(j|j-1) = H_\theta P_\theta(j|j-1) H_\theta' + R_\theta \quad (26-10)$$

This theorem is also applicable to either time-varying or nonstationary systems or both. Within the structure of these more complicated systems there must still be a collection of unknown but constant parameters. It is these parameters that are estimated by maximizing $L(\theta|\mathcal{Z})$.

Proof (Mendel, 1983b, pp. 101–103). We must first obtain the joint density function $p(\tilde{\mathbf{z}}|\theta) = p(\mathbf{z}(1), \dots, \mathbf{z}(N)|\theta)$. In Lesson 17 we saw that the innovations process $\tilde{\mathbf{z}}(i|i-1)$ and measurement $\mathbf{z}(i)$ are causally invertible; thus, the density function

$$\tilde{p}(\tilde{\mathbf{z}}(1|0), \tilde{\mathbf{z}}(2|1), \dots, \tilde{\mathbf{z}}(N|N-1)|\theta)$$

contains the same data information as $p(\mathbf{z}(1), \dots, \mathbf{z}(N)|\theta)$ does. Consequently, $L(\theta|\tilde{\mathbf{z}})$ can be replaced by $\tilde{L}(\theta|\tilde{\mathbf{z}})$, where

$$\tilde{\mathbf{z}} = \text{col}(\tilde{\mathbf{z}}(1|0), \dots, \tilde{\mathbf{z}}(N|N-1)) \quad (26-11)$$

and

$$\tilde{L}(\theta|\tilde{\mathbf{z}}) = \ln \tilde{p}(\tilde{\mathbf{z}}(1|0), \dots, \tilde{\mathbf{z}}(N|N-1)|\theta) \quad (26-12)$$

Now, however, we use the fact that the innovations process is Gaussian white noise to express $\tilde{L}(\theta|\tilde{\mathbf{z}})$ as

$$\tilde{L}(\theta|\tilde{\mathbf{z}}) = \ln \prod_{j=1}^N \tilde{p}_j(\tilde{\mathbf{z}}(j|j-1)|\theta) \quad (26-13)$$

For our basic state-variable model, the innovations are Gaussian distributed, which means that $\tilde{p}_j(\tilde{\mathbf{z}}(j|j-1)|\theta) = \tilde{p}(\tilde{\mathbf{z}}(j|j-1)|\theta)$ for $j = 1, \dots, N$; hence,

$$\tilde{L}(\theta|\tilde{\mathbf{z}}) = \ln \prod_{j=1}^N \tilde{p}(\tilde{\mathbf{z}}(j|j-1)|\theta) \quad (26-14)$$

From part (b) of Theorem 16-2 in Lesson 16 we know that

$$\begin{aligned} \tilde{p}(\tilde{\mathbf{z}}(j|j-1)|\theta) &= [(2\pi)^m |\mathcal{N}(j|j-1)|]^{-1/2} \\ &\exp\left[-\frac{1}{2}\tilde{\mathbf{z}}'(j|j-1)\mathcal{N}^{-1}(j|j-1)\tilde{\mathbf{z}}(j|j-1)\right] \end{aligned} \quad (26-15)$$

Substitute (26-15) into (26-14) to show that

$$\begin{aligned} \tilde{L}(\theta|\tilde{\mathbf{z}}) &= -\frac{1}{2} \sum_{j=1}^N [\tilde{\mathbf{z}}'(j|j-1)\mathcal{N}^{-1}(j|j-1)\tilde{\mathbf{z}}(j|j-1) \\ &\quad + \ln|\mathcal{N}(j|j-1)|] \end{aligned} \quad (26-16)$$

where by convention we have neglected the constant term $-\ln(2\pi)^{m/2}$ because it does not depend on θ .

Because $\tilde{p}(\#|\theta)$ and $p(\#|\theta)$ contain the same information about the data, $\tilde{L}(\theta|\tilde{\mathbf{z}})$ and $L(\theta|\tilde{\mathbf{z}})$ must also contain the same information about the data; hence, we can use $L(\theta|\tilde{\mathbf{z}})$ to denote the right-hand side of (26-16), as in (26-9). To indicate which quantities on the right-hand side of (26-9) may depend on θ , we have subscripted all such quantities with θ . \square

The innovations process $\tilde{\mathbf{z}}_\theta(j|j-1)$ can be generated by a Kalman filter; hence, the Kalman filter acts as a constraint that is associated with the computation of the log-likelihood function for the basic state-variable model.

In the present situation, where the true values of θ , θ_T , are not known but are being estimated, the estimate of $\mathbf{x}(j)$ obtained from a Kalman filter will be suboptimal due to wrong values of θ being used by that filter. In fact, we must use $\hat{\theta}_{ML}$ in the implementation of the Kalman filter, because $\hat{\theta}_{ML}$ will be the best information available about θ_T at t_j . If $\hat{\theta}_{ML} \rightarrow \theta_T$ as $N \rightarrow \infty$, then $\tilde{\mathbf{z}}_{\hat{\theta}_{ML}}(j|j-1) \rightarrow \tilde{\mathbf{z}}_{\theta_T}(j|j-1)$ as $N \rightarrow \infty$, and the suboptimal Kalman filter will approach the optimal Kalman filter. This result is about the best that one can hope for in maximum-likelihood estimation of parameters in our basic state-variable model.

Note also, that although we began with a parameter estimation problem, we wound up with a simultaneous state and parameter estimation problem. This is due to the uncertainties present in our state equation, which necessitated state estimation using a Kalman filter.

ON COMPUTING $\hat{\theta}_{ML}$

How do we determine $\hat{\theta}_{ML}$ for $L(\theta|\tilde{\mathbf{z}})$ given in (26-9) (subject to the constraint of the Kalman filter)? No simple closed-form solution is possible, because θ enters into $L(\theta|\tilde{\mathbf{z}})$ in a complicated *nonlinear* manner. The only way presently known to obtain $\hat{\theta}_{ML}$ is by means of mathematical programming.

The most effective optimization methods to determine $\hat{\theta}_{ML}$ require the computation of the gradient of $L(\theta|\tilde{\mathbf{z}})$ as well as the Hessian matrix, or a pseudo-Hessian matrix of $L(\theta|\tilde{\mathbf{z}})$. The Marquardt-Levenberg algorithm (also known as the Levenberg-Marquardt algorithm [Bard, 1970; Marquardt, 1963]), for example, has the form

$$\hat{\theta}_{ML}^{i+1} = \hat{\theta}_{ML}^i - (\mathbf{H}_i + \mathbf{D}_i)^{-1} \mathbf{g}_i \quad i = 0, 1, \dots \quad (26-17)$$

where \mathbf{g}_i denotes the gradient

$$\mathbf{g}_i = \frac{\partial L(\theta|\tilde{\mathbf{z}})}{\partial \theta} \Big|_{\theta = \hat{\theta}_{ML}^i} \quad (26-18)$$

\mathbf{H}_i denotes a pseudo-Hessian

$$\mathbf{H}_i \simeq \frac{\partial^2 L(\theta|\tilde{\mathbf{z}})}{\partial \theta^2} \Big|_{\theta = \hat{\theta}_{ML}^i} \quad (26-19)$$

and \mathbf{D}_i is a diagonal matrix chosen to force $\mathbf{H}_i + \mathbf{D}_i$ to be positive definite, so that $(\mathbf{H}_i + \mathbf{D}_i)^{-1}$ will always be computable.

We do not propose to discuss details of the Marquardt-Levenberg algorithm. The interested reader should consult the preceding references for general discussions, and Mendel (1983) or Gupta and Mehra (1974) for discussions directly related to the application of this algorithm to the present problem, maximization of $L(\theta|\tilde{\mathbf{z}})$.

We direct our attention to the calculations of \mathbf{g}_i and \mathbf{H}_i . The gradient of $L(\theta|\mathcal{Z})$ will require the calculations of

$$\frac{\partial \tilde{\mathbf{z}}_\theta(j|j-1)}{\partial \theta} \quad \text{and} \quad \frac{\partial \mathcal{N}_\theta(j|j-1)}{\partial \theta}$$

The innovations depend upon $\hat{\mathbf{x}}_\theta(j|j-1)$; hence, in order to compute $\partial \tilde{\mathbf{z}}_\theta(j|j-1)/\partial \theta$, we must compute $\partial \hat{\mathbf{x}}_\theta(j|j-1)/\partial \theta$. A Kalman filter must be used to compute $\hat{\mathbf{x}}_\theta(j|j-1)$; but this filter requires the following sequence of calculations: $\mathbf{P}_\theta(k|k) \rightarrow \mathbf{P}_\theta(k+1|k) \rightarrow \mathbf{K}_\theta(k+1) \rightarrow \hat{\mathbf{x}}_\theta(k+1|k) \rightarrow \dot{\hat{\mathbf{x}}}_\theta(k+1|k+1)$. Taking the partial derivative of the prediction equation with respect to θ_i we find that

$$\begin{aligned} \frac{\partial \hat{\mathbf{x}}_\theta(k+1|k)}{\partial \theta_i} &= \Phi_\theta \frac{\partial \hat{\mathbf{x}}_\theta(k|k)}{\partial \theta_i} + \frac{\partial \Phi_\theta}{\partial \theta_i} \hat{\mathbf{x}}_\theta(k|k) \\ &\quad + \frac{\partial \Psi_\theta}{\partial \theta_i} \mathbf{u}(k) \quad i = 1, 2, \dots, d \end{aligned} \quad (26-20)$$

We see that to compute $\partial \hat{\mathbf{x}}_\theta(k+1|k)/\partial \theta_i$, we must also compute $\partial \hat{\mathbf{x}}_\theta(k|k)/\partial \theta_i$. Taking the partial derivative of the correction equation with respect to θ_i , we find that

$$\begin{aligned} \frac{\partial \hat{\mathbf{x}}_\theta(k+1|k+1)}{\partial \theta_i} &= \frac{\partial \hat{\mathbf{x}}_\theta(k+1|k)}{\partial \theta_i} + \frac{\partial \mathbf{K}_\theta(k+1)}{\partial \theta_i} [\mathbf{z}(k+1) - \mathbf{H}_\theta \hat{\mathbf{x}}_\theta(k+1|k)] \\ &\quad - \mathbf{K}_\theta(k+1) \left[\frac{\partial \mathbf{H}_\theta}{\partial \theta_i} \hat{\mathbf{x}}_\theta(k+1|k) \right. \\ &\quad \left. + \mathbf{H}_\theta \frac{\partial \hat{\mathbf{x}}_\theta(k+1|k)}{\partial \theta_i} \right] \quad i = 1, 2, \dots, d \end{aligned} \quad (26-21)$$

Observe that to compute $\partial \hat{\mathbf{x}}_\theta(k+1|k+1)/\partial \theta_i$, we must also compute $\partial \mathbf{K}_\theta(k+1)/\partial \theta_i$. We leave it to the reader to show that the calculation of $\partial \mathbf{K}_\theta(k+1)/\partial \theta_i$ requires the calculation of $\partial \mathbf{P}_\theta(k+1|k)/\partial \theta_i$, which in turn requires the calculation of $\partial \mathbf{P}_\theta(k+1|k+1)/\partial \theta_i$.

The system of equations

$$\begin{aligned} \partial \hat{\mathbf{x}}_\theta(k+1|k)/\partial \theta_i &\quad \partial \hat{\mathbf{x}}_\theta(k+1|k+1)/\partial \theta_i \\ \partial \mathbf{K}_\theta(k+1)/\partial \theta_i &\quad \partial \mathbf{P}_\theta(k+1|k)/\partial \theta_i \quad \partial \mathbf{P}_\theta(k+1|k+1)/\partial \theta_i \end{aligned}$$

is called a Kalman filter sensitivity system. It is a linear system of equations, just as the Kalman filter, which is not only driven by measurements $\mathbf{z}(k+1)$ [e.g., see (26-21)] but is also driven by the Kalman filter [e.g., see (26-20) and (26-21)]. We need a total of d such sensitivity systems, one for each of the d unknown parameters in θ .

Each system of sensitivity equations requires about as much computation as a Kalman filter. Observe, however, that Kalman filter quantities

are used by the sensitivity equations; hence, the Kalman filter must be run together with the d sets of sensitivity equations. This procedure for recursively calculating the gradient $\partial L(\theta|\mathcal{Z})/\partial \theta$ therefore requires about as much computation as $d + 1$ Kalman filters. The sensitivity systems are totally uncoupled and lend themselves quite naturally to parallel processing (see Figure 26-1).

The Hessian matrix of $L(\theta|\mathcal{Z})$ is quite complicated, involving not only first derivatives of $\tilde{\mathbf{z}}_\theta(j|j-1)$ and $\mathcal{N}_\theta(j|j-1)$, but also their second derivatives. The pseudo-Hessian matrix of $L(\theta|\mathcal{Z})$ ignores all the second derivative terms; hence, it is relatively easy to compute because all the first derivative terms have already been computed in order to calculate the gradient of $L(\theta|\mathcal{Z})$. Justification for neglecting the second derivative terms is given by Gupta and Mehra (1974), who show that as $\hat{\theta}_{ML}$ approaches θ_T , the expected value of the dropped terms goes to zero.

The estimation literature is filled with many applications of maximum-likelihood state and parameter estimation. For example, Mendel (1983b) applies it to seismic data processing, Mehra and Tyler (1973) apply it to aircraft parameter identification and McLaughlin (1980) applies it to groundwater flow.

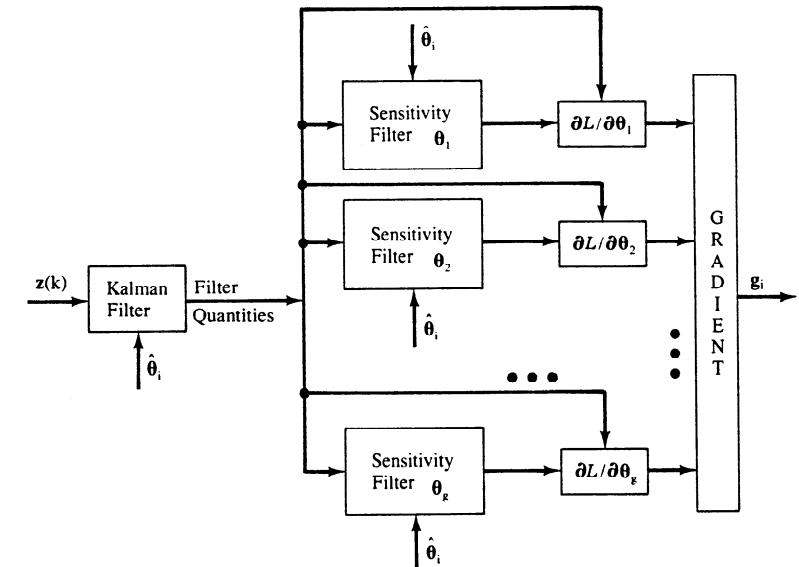


Figure 26-1 Calculations needed to compute gradient vector \mathbf{g}_i . Note that θ_i denotes the i th component of θ (Mendel, 1983b, © 1983, Academic Press, Inc.).

A STEADY-STATE APPROXIMATION

Suppose our basic state variable model is time-invariant and stationary so that $\bar{\mathbf{P}} = \lim_{j \rightarrow \infty} \mathbf{P}(j|j-1)$ exists. Let

$$\bar{\mathcal{N}} = \bar{\mathbf{H}}\bar{\mathbf{P}}\bar{\mathbf{H}}' + \mathbf{R} \quad (26-22)$$

and

$$\bar{L}(\theta|\mathfrak{L}) = -\frac{1}{2} \sum_{j=1}^N \bar{\mathbf{z}}'(j|j-1) \bar{\mathcal{N}}^{-1} \bar{\mathbf{z}}(j|j-1) - \frac{1}{2} N \ln |\bar{\mathcal{N}}| \quad (26-23)$$

Log-likelihood function $\bar{L}(\theta|\mathfrak{L})$ is a steady-state approximation of $L(\theta|\mathfrak{L})$. The steady-state Kalman filter used to compute $\bar{\mathbf{z}}(j|j-1)$ and $\bar{\mathcal{N}}$ is

$$\hat{\mathbf{x}}(k+1|k) = \Phi \hat{\mathbf{x}}(k|k) + \Psi \mathbf{u}(k) \quad (26-24)$$

$$\begin{aligned} \hat{\mathbf{x}}(k+1|k+1) &= \hat{\mathbf{x}}(k+1|k) + \bar{\mathbf{K}}[\mathbf{z}(k+1) \\ &\quad - \mathbf{H}\hat{\mathbf{x}}(k+1|k)] \end{aligned} \quad (26-25)$$

in which $\bar{\mathbf{K}}$ is the steady-state Kalman gain matrix.

Recall that

$$\theta = \text{col}(\text{elements of } \Phi, \Gamma, \Psi, \mathbf{H}, \mathbf{Q}, \text{ and } \mathbf{R}) \quad (26-26)$$

We now make the following transformations of variables:

$$\left. \begin{array}{l} \Phi \rightarrow \Phi \\ \Psi \rightarrow \Psi \\ \mathbf{H} \rightarrow \mathbf{H} \\ (\Phi, \Gamma, \Psi, \mathbf{H}, \mathbf{Q}, \mathbf{R}) \rightarrow \bar{\mathcal{N}} \\ (\Phi, \Gamma, \Psi, \mathbf{H}, \mathbf{Q}, \mathbf{R}) \rightarrow \bar{\mathbf{K}} \end{array} \right\} \quad (26-27)$$

We ignore Γ (initially) for reasons that are explained below Equation (26-32). Let

$$\phi = \text{col}(\text{elements of } \Phi, \Psi, \mathbf{H}, \bar{\mathcal{N}}, \text{ and } \bar{\mathbf{K}}) \quad (26-28)$$

where ϕ is $p \times 1$, and view \bar{L} as a function of ϕ , i.e.,

$$\bar{L}(\phi|\mathfrak{L}) = -\frac{1}{2} \sum_{j=1}^N \bar{\mathbf{z}}_\phi'(j|j-1) \bar{\mathcal{N}}_\phi^{-1} \bar{\mathbf{z}}_\phi(j|j-1) - \frac{1}{2} N \ln |\bar{\mathcal{N}}_\phi| \quad (26-29)$$

Instead of finding $\hat{\theta}_{ML}$ that maximizes $\bar{L}(\theta|\mathfrak{L})$, subject to the constraints of a full-blown Kalman filter, we now propose to find $\hat{\phi}_{ML}$ that maximizes $\bar{L}(\phi|\mathfrak{L})$, subject to the constraints of the following filter:

$$\hat{\mathbf{x}}_\phi(k+1|k) = \Phi_\phi \hat{\mathbf{x}}_\phi(k|k) + \Psi_\phi \mathbf{u}(k) \quad (26-30)$$

$$\hat{\mathbf{x}}_\phi(k+1|k+1) = \hat{\mathbf{x}}_\phi(k+1|k) + \bar{\mathbf{K}}_\phi \bar{\mathbf{z}}_\phi(k+1|k) \quad (26-31)$$

A Steady-State Approximation

and

$$\bar{\mathbf{z}}_\phi(k+1|k) = \mathbf{z}(k+1) - \mathbf{H}_\phi \hat{\mathbf{x}}_\phi(k+1|k) \quad (26-32)$$

Once we have computed $\hat{\phi}_{ML}$ we can compute $\hat{\theta}_{ML}$ by inverting the transformations in (26-27). Of course, when we do this, we are also using the invariance property of maximum-likelihood estimates.

Observe that $\bar{L}(\phi|\mathfrak{L})$ in (26-29) and the filter in (26-30)–(26-32) do not depend on Γ ; hence, we have not included Γ in any definition of ϕ . We explain how to reconstruct Γ from $\hat{\phi}_{ML}$ following Equation (26-44).

Because maximum-likelihood estimates are asymptotically efficient (Lesson 11), once we have determined $\hat{\mathbf{K}}_{ML}$, the filter in (26-30) and (26-31) will be the steady-state Kalman filter.

The major advantage of this steady-state approximation is that the filter sensitivity equations are greatly simplified. When $\bar{\mathbf{K}}$ and $\bar{\mathcal{N}}$ are treated as matrices of unknown parameters we do not need the predicted and corrected error-covariance matrices to “compute” $\bar{\mathbf{K}}$ and $\bar{\mathcal{N}}$. The sensitivity equations for (26-32), (26-30), and (26-31) are

$$\frac{\partial \bar{\mathbf{z}}_\phi(k+1|k)}{\partial \phi_i} = -\mathbf{H}_\phi \frac{\partial \hat{\mathbf{x}}_\phi(k+1|k)}{\partial \phi_i} - \frac{\partial \mathbf{H}_\phi}{\partial \phi_i} \hat{\mathbf{x}}_\phi(k+1|k) \quad (26-33)$$

$$\frac{\partial \hat{\mathbf{x}}_\phi(k+1|k)}{\partial \phi_i} = \Phi_\phi \frac{\partial \hat{\mathbf{x}}_\phi(k|k)}{\partial \phi_i} + \frac{\partial \Phi_\phi}{\partial \phi_i} \hat{\mathbf{x}}_\phi(k|k) + \frac{\partial \Psi_\phi}{\partial \phi_i} \mathbf{u}(k) \quad (26-34)$$

and

$$\begin{aligned} \frac{\partial \bar{\mathbf{z}}_\phi(k+1|k+1)}{\partial \phi_i} &= \frac{\partial \hat{\mathbf{x}}_\phi(k+1|k)}{\partial \phi_i} + \frac{\partial \bar{\mathbf{K}}_\phi}{\partial \phi_i} [\mathbf{z}(k+1) - \mathbf{H}_\phi \hat{\mathbf{x}}_\phi(k|k+1)] \\ &\quad - \bar{\mathbf{K}}_\phi \frac{\partial \hat{\mathbf{x}}_\phi(k+1|k)}{\partial \phi_i} \\ &\quad - \bar{\mathbf{K}}_\phi \frac{\partial \mathbf{H}_\phi}{\partial \phi_i} \hat{\mathbf{x}}_\phi(k+1|k) \end{aligned} \quad (26-35)$$

where $i = 1, 2, \dots, p$. Note that $\partial \bar{\mathbf{K}}_\phi / \partial \phi_i$ is zero for all ϕ_i not in $\bar{\mathbf{K}}_\phi$ and is a matrix filled with zeros and a single unity value for ϕ_i in $\bar{\mathbf{K}}_\phi$.

There are more elements in ϕ than in θ , because $\bar{\mathcal{N}}$ and $\bar{\mathbf{K}}$ have more unknown elements in them than do \mathbf{Q} and \mathbf{R} , i.e., $p > d$. Additionally, $\bar{\mathcal{N}}$ does not appear in the filter equations; it only appears in $\bar{L}(\phi|\mathfrak{L})$. It is, therefore, possible to obtain a closed-form solution for $\hat{\mathbf{N}}_{ML}$.

Theorem 26-2. A closed-form solution for matrix $\hat{\mathbf{N}}_{ML}$ is

$$\hat{\mathbf{N}}_{ML} = \frac{1}{N} \sum_{j=1}^N \bar{\mathbf{z}}_\phi(j|j-1) \bar{\mathbf{z}}_\phi'(j|j-1) \quad (26-36)$$

Proof. To determine $\hat{\bar{N}}_{ML}$ we must set $\partial \bar{L}(\phi|\mathcal{Z})/\partial \bar{N}_{\phi} = 0$ and solve the resulting equation for \bar{N}_{ML} . This is most easily accomplished by applying gradient matrix formulas to (26-29), that are given in Schweppe (1974). Doing this, we obtain

$$\frac{\partial \bar{L}(\phi|\mathcal{Z})}{\partial \bar{N}} = \sum_{j=1}^N [\bar{N}^{-1} \bar{z}_{\phi}(j|j-1) \bar{z}_{\phi}'(j|j-1) \bar{N}^{-1}]' - N(\bar{N}^{-1})' = 0 \quad (26-37)$$

whose solution is $\hat{\bar{N}}_{ML}$ in (26-36). \square

Observe that $\hat{\bar{N}}_{ML}$ is the sample steady-state-covariance matrix of \bar{z}_{ϕ} ; i.e., as

$$\hat{\Phi}_{ML} \rightarrow \hat{\Phi}_T \quad \hat{\bar{N}}_{ML} \rightarrow \lim_{j \rightarrow \infty} \text{cov}[\bar{z}(j|j-1)]$$

Suppose we are also interested in determining \hat{Q}_{ML} and \hat{R}_{ML} . How do we obtain these quantities from $\hat{\Phi}_{ML}$?

As in Lesson 19, we let \bar{K} , \bar{P} , and \bar{P}_1 denote the steady-state values of $K(k+1)$, $P(k+1|k)$, and $P(k|k)$, respectively, where

$$\bar{K} = \bar{P}H'(\bar{P}H' + R)^{-1} = \bar{P}H'\bar{N}^{-1} \quad (26-38)$$

$$\bar{P} = \Phi\bar{P}_1\Phi' + \Gamma Q\Gamma' \quad (26-39)$$

and

$$\bar{P}_1 = (I - \bar{K}H)\bar{P} \quad (26-40)$$

Additionally, we know that

$$\bar{N} = \bar{P}H' + R \quad ((26-41))$$

By the invariance property of maximum-likelihood estimates, we know that

$$\hat{\bar{N}}_{ML} = \widehat{(\bar{P}H' + R)}_{ML} = \hat{H}_{ML}\bar{P}\hat{H}'_{ML} + \hat{R}_{ML} \quad (26-42)$$

and

$$\hat{\bar{K}}_{ML} = \widehat{(\bar{P}H'\bar{N}^{-1})}_{ML} = \bar{P}\hat{H}'_{ML}(\hat{\bar{N}}_{ML})^{-1} \quad (26-43)$$

Solving (26-43) for $\bar{P}\hat{H}'_{ML}$ and substituting the resulting expression into (26-42), we obtain the following solution for \hat{R}_{ML} ,

$$\hat{R}_{ML} = (I - \hat{H}_{ML}\hat{\bar{K}}_{ML})\hat{\bar{N}}_{ML} \quad (26-44)$$

No closed-form solution exists for \hat{Q}_{ML} . Substituting $\hat{\Phi}_{ML}$, \hat{H}_{ML} , $\hat{\bar{K}}_{ML}$, and $\hat{\bar{N}}_{ML}$ into (26-38)–(26-40), and combining (26-39) and (26-40), we obtain the following two equations

$$\bar{P} = \hat{\Phi}_{ML}(I - \hat{\bar{K}}_{ML}\hat{H}_{ML})\bar{P}\hat{\Phi}'_{ML} + \widehat{(\Gamma Q\Gamma')}_{ML} \quad (26-45)$$

and

$$\bar{P}\hat{H}'_{ML} = \hat{\bar{K}}_{ML}\hat{\bar{N}}_{ML} \quad (26-46)$$

These equations must be solved simultaneously for \bar{P} and $\widehat{(\Gamma Q\Gamma')}_{ML}$ using iterative numerical techniques. For details, see Mehra (1970a).

Note finally, that the best for which we can hope by this approach is not $\hat{\bar{N}}_{ML}$ and \hat{Q}_{ML} , but only $\widehat{(\Gamma Q\Gamma')}_{ML}$. This is due to the fact that, when Γ and Q are both unknown, there will be an ambiguity in their determination, i.e., the term $\Gamma w(k)$ which appears in our basic state-variable model [for which $E\{w(k)w'(k)\} = Q$] cannot be distinguished from the term $w_1(k)$, for which

$$E\{w_1(k)w'_1(k)\} = Q_1 = \Gamma Q\Gamma' \quad (26-47)$$

This observation is also applicable to the original problem formulation wherein we obtained \hat{Q}_{ML} directly; i.e., when both Γ and Q are unknown, we should really choose

$$\theta = \text{col}(\text{elements of } \Phi, \Psi, H, \Gamma Q\Gamma', \text{ and } R) \quad (26-48)$$

In summary, when our basic state-variable model is time-invariant and stationary, we can first obtain $\hat{\Phi}_{ML}$ by maximizing $\bar{L}(\phi|\mathcal{Z})$ given in (26-29), subject to the constraints of the simple filter in (26-30), (26-31), and (26-32). A mathematical programming method must be used to obtain those elements of $\hat{\Phi}_{ML}$ associated with $\hat{\Phi}_{ML}$, $\hat{\Psi}_{ML}$, \hat{H}_{ML} , and $\hat{\bar{K}}_{ML}$. The closed-form solution, given in (26-36), is used for $\hat{\bar{N}}_{ML}$. Finally, if we want to reconstruct R_{ML} and $\widehat{(\Gamma Q\Gamma')}_{ML}$, we use (26-44) for the former and must solve (26-45) and (26-46) for the latter.

Example 26-1 (Mehra, 1971)

The following fourth-order system, which represents the short period dynamics and the first bending mode of a missile, was simulated:

$$\begin{pmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \\ x_4(k+1) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\alpha_1 & -\alpha_2 & -\alpha_3 & -\alpha_4 \end{pmatrix} \begin{pmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} w(k) \quad (26-49)$$

$$z(k+1) = x_1(k+1) + v(k+1) \quad (26-50)$$

For this model, it was assumed that $x(0) = 0$, $q = 1.0$, $r = 0.25$, $\alpha_1 = -0.656$, $\alpha_2 = 0.784$, $\alpha_3 = -0.18$, and $\alpha_4 = 1.0$.

Using measurements generated from the simulation, maximum-likelihood estimates were obtained for ϕ , where

$$\phi = \text{col}(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \bar{N}, \bar{k}_1, \bar{k}_2, \bar{k}_3, \bar{k}_4) \quad (26-51)$$

In (26-51), \bar{N} is a scalar because, in this example, $z(k)$ is a scalar. Additionally, it was assumed that $x(0)$ was known exactly. According to Mehra (1971, pg. 30) “The starting values for the maximum likelihood scheme were obtained using a correlation technique given in Mehra (1970b). The results of successive iterations are shown in Table 26-1. The variances of the estimates obtained from the matrix of second partial derivatives (i.e., the Hessian matrix of \bar{L}) are also given. For comparison purposes, results obtained by using 1000 data points and 100 data points are given.” \square

PROBLEMS

TABLE 26-1 Parameter Estimates for Missile Example

Iteration	$L \times 10^{-3}$	\bar{N}	α_1	α_2	α_3	α_4	\bar{k}_1	\bar{k}_2	\bar{k}_3	\bar{k}_4
⁰ Its from correlation technique, Mehra (1970b).										
1	-1.0706	2.3800	-0.5965	0.8029	-0.1360	0.8696	0.6830	0.2837	0.4191	0.820
estimates using 1000 points										
2	-1.0660	2.3811	-0.5938	0.8029	-0.1338	0.8759	0.6803	0.2840	0.4200	0.831
3	-1.0085	2.4026	-0.6054	0.7452	-0.1494	0.9380	0.6304	0.2888	0.4392	1.03
4	-0.9798	2.4409	-0.6036	0.8161	-0.1405	0.8540	0.6801	0.3210	0.6108	1.18
5	-0.9785	2.4412	-0.5999	0.8196	-0.1370	0.8580	0.6803	0.3214	0.6107	1.18
6	-0.9771	2.4637	-0.6014	0.8086	-0.1503	0.8841	0.7068	0.3479	0.6059	1.22
7	-0.9769	2.4603	-0.6023	0.8130	-0.1470	0.8773	0.7045	0.3429	0.6106	1.21
8	-0.9744	2.5240	-0.6313	0.8105	-0.1631	0.9279	0.7990	0.3756	0.6484	1.25
9	-0.9743	2.5241	-0.6306	0.8108	-0.1622	0.9296	0.7989	0.3749	0.6480	1.25
0	-0.9734	2.5270	-0.6374	0.7961	-0.1630	0.9505	0.7974	0.3568	0.6378	1.25
1	-0.9728	2.5313	-0.6482	0.7987	-0.1620	0.9577	0.8103	0.3443	0.6403	1.23
2	-0.9720	2.5444	-0.6602	0.7995	-0.1783	0.9866	0.8487	0.3303	0.6083	1.20
3	-0.9714	2.5600	-0.6634	0.7919	-0.2036	1.0280	0.8924	0.3143	0.6014	1.20
4	-0.9711	2.5657	-0.6624	0.7808	-0.2148	1.0491	0.9073	0.3251	0.6122	1.22
estimates using 100 points										
0	-0.9659	2.620	-0.6094	0.7663	-0.1987	1.0156	1.24	0.136	0.454	1.10
al values										
-0.94	2.557	-0.6560	0.7840	-0.1800	1.0000	0.8937	0.2957	0.6239	1.25	
nates of standard deviation using 1000 points										
0.0317	0.0277	0.0247	0.0275	0.0261	0.0302	0.0323	0.0302	0.029		
nates of standard deviation using 100 points										
0.149	0.104	0.131	0.084	0.184	0.303	0.092	0.082	0.09		

ce: Mehra (1971, pg. 30), © 1971, AIAA.

- 26-1. Obtain the sensitivity equations for $\partial \mathbf{K}_\theta(k+1)/\partial \theta_i$, $\partial \mathbf{P}_\theta(k+1|k)/\partial \theta_i$ and $\partial \mathbf{P}_\theta(k+1|k+1)/\partial \theta_i$. Explain why the sensitivity system for $\partial \hat{\mathbf{x}}_\theta(k+1|k)/\partial \theta_i$ and $\partial \dot{\mathbf{x}}_\theta(k+1|k+1)/\partial \theta_i$ is linear.

- 26-2. Compute formulas for \mathbf{g}_i and \mathbf{H}_i . Then simplify \mathbf{H}_i to a pseudo-Hessian.

- 26-3. In the first-order system $x(k+1) = ax(k) + w(k)$, and $z(k+1) = x(k+1) + v(k+1)$, $k = 1, 2, \dots, N$, a is an unknown parameter that is to be estimated. Sequences $w(k)$ and $v(k)$ are, as usual, mutually uncorrelated and white, and, $w(k) \sim N(w(k); 0, 1)$ and $v(k) \sim N(v(k); 0, 1/2)$. Explain, using equations and a flowchart, how parameter a can be estimated using a MLE.

- 26-4. Repeat the preceding problem where all conditions are the same except that now $w(k)$ and $v(k)$ are correlated, and $E[w(k)v(k)] = 1/4$.

- 26-5. We are interested in estimating the parameters a and r in the following first-order system:

$$\begin{aligned} x(k+1) + ax(k) &= w(k) \\ z(k) = x(k) + v(k) &\quad k = 1, 2, \dots, N \end{aligned}$$

Signals $w(k)$ and $v(k)$ are mutually uncorrelated, white, and Gaussian, and, $E[w^2(k)] = 1$ and $E[v^2(k)] = r$.

- (a) Let $\theta = \text{col}(a, r)$. What is the equation for the log-likelihood function?
 (b) Prepare a macro flow chart that depicts the sequence of calculations required to maximize $L(\theta|\mathcal{Z})$. Assume an optimization algorithm is used which requires gradient information about $L(\theta|\mathcal{Z})$.
 (c) Write out the Kalman filter sensitivity equations for parameters a and r .

- 26-6. Develop the sensitivity equations for the case considered in Lesson 11, i.e., for the case where the only uncertainty present in the state-variable model is measurement noise. Begin with $L(\theta|\mathcal{Z})$ in (11-42).

- 26-7. Refer to Problem 24-7. Explain, using equations and a flowchart, how to obtain MLE's of the unknown parameters for:

- (a) equation for the unsteady operation of a synchronous motor, in which C and p are unknown;
 (c) Duffing's equation, in which C , α , and β are unknown;
 (c) Van der Pol's equation, in which ϵ is unknown; and
 (d) Hill's equation, in which a and b are unknown.

Kalman-Bucy Filtering

INTRODUCTION

The Kalman-Bucy filter is the continuous-time counterpart to the Kalman filter. It is a continuous-time minimum-variance filter that provides state estimates for continuous-time dynamical systems that are described by linear, (possibly) time-varying, and (possibly) nonstationary ordinary differential equations.

The Kalman-Bucy filter (KBF) can be derived in a number of different ways, including the following three:

1. Use a formal limiting procedure to obtain the KBF from the KF (e.g., Meditch, 1969).
2. Begin by assuming the optimal estimator is a linear transformation of *all* measurements. Use a calculus of variations argument or the orthogonality principle to obtain the Wiener-Hopf integral equation. Embedded within this equation is the filter kernel. Take the derivative of the Wiener-Hopf equation to obtain a differential equation which is the KBF (Meditch, 1969).
3. Begin by assuming a linear differential equation structure for the KBF, one that contains an unknown time-varying gain matrix that weights the difference between the measurement made at time t and the estimate of that measurement. Choose the gain matrix that minimizes the mean-squared error (Athans and Tse, 1967).

We shall briefly describe the first and third approaches, but first we must define our continuous-time model and formally state the problem we wish to solve.

SYSTEM DESCRIPTION

Our continuous-time system is described by the following state-variable model,

$$\dot{\mathbf{x}}(t) = \mathbf{F}(t)\mathbf{x}(t) + \mathbf{G}(t)\mathbf{w}(t) \quad (27-1)$$

$$\mathbf{z}(t) = \mathbf{H}(t)\mathbf{x}(t) + \mathbf{v}(t) \quad (27-2)$$

where $\mathbf{x}(t)$ is $n \times 1$, $\mathbf{w}(t)$ is $p \times 1$, $\mathbf{z}(t)$ is $m \times 1$, and $\mathbf{v}(t)$ is $m \times 1$. For simplicity, we have omitted a known forcing function term in state equation (27-1). Matrices $\mathbf{F}(t)$, $\mathbf{G}(t)$, and $\mathbf{H}(t)$ have dimensions which conform to the dimensions of the vector quantities in this state-variable model. Disturbance $\mathbf{w}(t)$ and measurement noise $\mathbf{v}(t)$ are zero-mean white noise processes, which are assumed to be uncorrelated, i.e., $E\{\mathbf{w}(t)\} = \mathbf{0}$, $E\{\mathbf{v}(t)\} = \mathbf{0}$,

$$E\{\mathbf{w}(t)\mathbf{w}'(\tau)\} = \mathbf{Q}(t)\delta(t - \tau) \quad (27-3)$$

$$E\{\mathbf{v}(t)\mathbf{v}'(\tau)\} = \mathbf{R}(t)\delta(t - \tau) \quad (27-4)$$

and

$$E\{\mathbf{w}(t)\mathbf{v}'(\tau)\} = \mathbf{0} \quad (27-5)$$

Equations (27-3), (27-4), and (27-5) apply for $t \geq t_0$. Additionally, $\mathbf{R}(t)$ is continuous and positive definite, whereas $\mathbf{Q}(t)$ is continuous and positive semidefinite. Finally, we assume that the initial state vector $\mathbf{x}(t_0)$ may be random, and if it is, it is uncorrelated with both $\mathbf{w}(t)$ and $\mathbf{v}(t)$. The statistics of a random $\mathbf{x}(t_0)$ are

$$E\{\mathbf{x}(t_0)\} = \mathbf{m}_x(t_0) \quad (27-6)$$

and

$$\text{cov}\{\mathbf{x}(t_0)\} = \mathbf{P}_x(t_0) \quad (27-7)$$

Measurements $\mathbf{z}(t)$ are assumed to be made for $t_0 \leq t \leq \tau$.

If $\mathbf{x}(t_0)$, $\mathbf{w}(t)$, and $\mathbf{v}(t)$ are jointly Gaussian for all $t \in [t_0, \tau]$, then the KBF will be the optimal estimator of state vector $\mathbf{x}(t)$. We will not make any distributional assumptions about $\mathbf{x}(t_0)$, $\mathbf{w}(t)$, and $\mathbf{v}(t)$ in this lesson, being content to establish the *linear optimal estimator* of $\mathbf{x}(t)$.

NOTATION AND PROBLEM STATEMENT

Our notation for a continuous-time estimate of $\mathbf{x}(t)$ and its associated estimation error parallels our notation for the comparable discrete-time quantities, i.e., $\hat{\mathbf{x}}(t|t)$ denotes the optimal estimate of $\mathbf{x}(t)$ which uses all the measurements $\mathbf{z}(t)$, where $t \geq t_0$, and

$$\tilde{\mathbf{x}}(t|t) = \mathbf{x}(t) - \hat{\mathbf{x}}(t|t) \quad (27-8)$$

The mean-squared state estimation error is

$$J[\hat{x}(t|t)] = E\{\hat{x}'(t|t)\hat{x}(t|t)\} \quad (27-9)$$

We shall determine $\hat{x}(t|t)$ that minimizes $J[\hat{x}(t|t)]$, subject to the constraints of our state-variable model and data set.

THE KALMAN-BUCY FILTER

The solution to the problem stated in the preceding section is the Kalman-Bucy Filter, the structure of which is summarized in the following:

Theorem 27-1. *The KBF is described by the vector differential equation*

$$\dot{\hat{x}}(t|t) = F(t)\hat{x}(t|t) + K(t)[z(t) - H(t)\hat{x}(t|t)] \quad (27-10)$$

where $t \geq t_0$, $\hat{x}(t_0|t_0) = m_x(t_0)$,

$$K(t) = P(t|t)H'(t)R^{-1}(t) \quad (27-11)$$

and

$$\begin{aligned} \dot{P}(t|t) &= F(t)P(t|t) + P(t|t)F'(t) - P(t|t)H'(t)R^{-1}(t)H(t)P(t|t) \\ &\quad + G(t)Q(t)G'(t) \end{aligned} \quad (27-12)$$

Equation (27-12), which is a matrix Riccati differential equation, is initialized by $P(t_0|t_0) = P_x(t_0)$. \square

Matrix $K(t)$ is the *Kalman-Bucy gain matrix*, and $P(t|t)$ is the state-estimation-error covariance matrix, i.e.,

$$P(t|t) = E\{\hat{x}(t|t)\hat{x}'(t|t)\} \quad (27-13)$$

Equation (27-10) can be rewritten as

$$\dot{\hat{x}}(t|t) = [F(t) - K(t)H(t)]\hat{x}(t|t) + K(t)z(t) \quad (27-14)$$

which makes it very clear that the KBF is a time-varying filter that processes the measurements linearly to produce $\hat{x}(t|t)$.

The solution to (27-14) is

$$\hat{x}(t|t) = \Phi(t,t_0)\hat{x}(t_0|t_0) + \int_{t_0}^t \Phi(t,\tau)K(\tau)z(\tau)d\tau \quad (27-15)$$

where the state transition matrix $\Phi(t,\tau)$ is the solution to the matrix differential equation

$$\left. \begin{aligned} \dot{\Phi}(t,\tau) &= [F(t) - K(t)H(t)]\Phi(t,\tau) \\ \Phi(t,t) &= I \end{aligned} \right\} \quad (27-16)$$

If $\hat{x}(t_0|t_0) = 0$, then

$$\hat{x}(t|t) = \int_{t_0}^t \Phi(t,\tau)K(\tau)z(\tau)d\tau = \int_{t_0}^t A(t,\tau)z(\tau)d\tau \quad (27-17)$$

where the filter kernel $A(t,\tau)$ is

$$A(t,\tau) = \Phi(t,\tau)K(\tau) \quad (27-18)$$

The second approach to deriving the KBF, mentioned in the introduction to this chapter, begins by assuming that $\hat{x}(t|t)$ can be expressed as in (27-17) where $A(t,\tau)$ is unknown. The mean-squared estimation error is minimized to obtain the following Wiener-Hopf integral equation:

$$E\{x(t)z'(\sigma)\} - \int_{t_0}^t A(t,\tau)E\{z(\tau)z'(\sigma)\}d\tau = 0 \quad (27-19)$$

where $t_0 \leq \sigma \leq t$. When this equation is converted into a differential equation, one obtains the KBF described in Theorem 27-1. For the details of this derivation of Theorem 27-1 see Meditch, 1969, Chapter 8.

DERIVATION OF KBF USING A FORMAL LIMITING PROCEDURE

Kalman filter Equation (17-11), expressed as

$$\begin{aligned} \hat{x}(k+1|k+1) &= \Phi(k+1,k)\hat{x}(k|k) \\ &\quad + K(k+1)[z(k+1) - H(k+1)\Phi(k+1,k)\hat{x}(k|k)] \end{aligned} \quad (27-20)$$

can also be written as

$$\begin{aligned} \hat{x}(t+\Delta t|t+\Delta t) &= \Phi(t+\Delta t,t)\hat{x}(t|t) \\ &\quad + K(t+\Delta t)[z(t+\Delta t) - H(t+\Delta t)\Phi(t+\Delta t,t)\hat{x}(t|t)] \end{aligned} \quad (27-21)$$

where we have let $t_k = t$ and $t_{k+1} = t + \Delta t$. In Example 24-3 we showed that

$$\Phi(t+\Delta t,t) \approx I + F(t)\Delta t + O(\Delta t^2) \quad (27-22)$$

and

$$Q_d(t) \approx G(t)Q(t)G'(t)\Delta t + O(\Delta t^2) \quad (27-23)$$

Observe that $Q_d(t)$ can also be written as

$$Q_d(t) = [G(t)\Delta t] \left[\frac{Q(t)}{\Delta t} \right] [G(t)\Delta t]' + O(\Delta t^2) \quad (27-24)$$

and, if we express $w_d(k)$ as $\Gamma(k+1,k)w(k)$, so that $Q_d(k) = \Gamma(k+1,k)Q(k)\Gamma'(k+1,k)$, then

$$\Gamma(t+\Delta t,t) = G(t)\Delta t + O(\Delta t^2) \quad (27-25)$$

and

$$Q(k = t) \Rightarrow \frac{Q(t)}{\Delta t} \quad (27-26)$$

Equation (27-26) means that we replace $Q(k = t)$ in the KF by $Q(t)/\Delta t$. Note that we have encountered a bit of a notational problem here, because we have used $w(k)$ [and its associated covariance, $Q(k)$] to denote the disturbance in our discrete-time model, and $w(t)$ [and its associated intensity, $Q(t)$] to denote the disturbance in our continuous-time model.

Without going into technical details, we shall also replace $R(t + \Delta t)$ in the KF by $R(t + \Delta t)/\Delta t$, i.e.,

$$R(k + 1 = t + \Delta t) \Rightarrow R(t + \Delta t)/\Delta t \quad (27-27)$$

See Meditch (1969, pp. 139–142) for an explanation.

Substituting (27-22) into (27-21), and omitting all higher-order terms in Δt , we find that

$$\begin{aligned} \hat{x}(t + \Delta t|t + \Delta t) &= [\mathbf{I} + F(t)\Delta t]\hat{x}(t|t) + K(t + \Delta t)\{z(t + \Delta t) \\ &\quad - H(t + \Delta t)[\mathbf{I} + F(t)\Delta t]\hat{x}(t|t)\} \end{aligned} \quad (27-28)$$

from which it follows that

$$\begin{aligned} \lim_{\Delta t \rightarrow 0} \frac{\hat{x}(t + \Delta t|t + \Delta t) - \hat{x}(t|t)}{\Delta t} &= F(t)\hat{x}(t|t) \\ &\quad + \lim_{\Delta t \rightarrow 0} K(t + \Delta t)\{z(t + \Delta t) - H(t + \Delta t)[\mathbf{I} + F(t)\Delta t]\hat{x}(t|t)\}/\Delta t \end{aligned}$$

or

$$\begin{aligned} \dot{\hat{x}}(t|t) &= F(t)\hat{x}(t|t) + \lim_{\Delta t \rightarrow 0} K(t + \Delta t)\{z(t + \Delta t) \\ &\quad - H(t + \Delta t)[\mathbf{I} + F(t)\Delta t]\hat{x}(t|t)\}/\Delta t \end{aligned} \quad (27-29)$$

Under suitable regularity conditions, which we shall assume are satisfied here, we can replace the limit of a product of functions by the product of limits, i.e.,

$$\begin{aligned} \lim_{\Delta t \rightarrow 0} K(t + \Delta t)\{z(t + \Delta t) - H(t + \Delta t)[\mathbf{I} + F(t)\Delta t]\hat{x}(t|t)\}/\Delta t \\ = \lim_{\Delta t \rightarrow 0} \frac{K(t + \Delta t)}{\Delta t} \lim_{\Delta t \rightarrow 0} \{z(t + \Delta t) - H(t + \Delta t)[\mathbf{I} + F(t)\Delta t]\hat{x}(t|t)\} \end{aligned} \quad (27-30)$$

The second limit on the right-hand side of (27-30) is easy to evaluate, i.e.,

$$\lim_{\Delta t \rightarrow 0} \{z(t + \Delta t) - H(t + \Delta t)[\mathbf{I} + F(t)\Delta t]\hat{x}(t|t)\} = z(t) - H(t)\hat{x}(t|t) \quad (27-31)$$

In order to evaluate the first limit on the right-hand side of (27-30), we first substitute $R(t + \Delta t)/\Delta t$ for $R(k + 1 = t + \Delta t)$ in (17-12), to obtain

$$\begin{aligned} K(t + \Delta t) &= P(t + \Delta t|t)H'(t + \Delta t) \\ &= [H(t + \Delta t)P(t + \Delta t|t)H'(t + \Delta t)\Delta t + R(t + \Delta t)]^{-1}\Delta t \end{aligned} \quad (27-32)$$

Then, we substitute $Q(t)/\Delta t$ for $Q(k = t)$ in (17-13), to obtain

$$\begin{aligned} P(t + \Delta t|t) &= \Phi(t + \Delta t, t)P(t|t)\Phi'(t + \Delta t, t) \\ &\quad + \Gamma(t + \Delta t, t) \frac{Q(t)}{\Delta t} \Gamma'(t + \Delta t, t) \end{aligned} \quad (27-33)$$

We leave it to the reader to show that

$$\lim_{\Delta t \rightarrow 0} P(t + \Delta t|t) = P(t|t) \quad (27-34)$$

hence,

$$\lim_{\Delta t \rightarrow 0} \frac{K(t + \Delta t)}{\Delta t} = P(t|t)H'(t)\mathbf{R}^{-1}(t) \triangleq K(t) \quad (27-35)$$

Combining (27-29), (27-30), (27-31), and (27-35), we obtain the KBF in (27-10) and the KB gain matrix in (27-11).

In order to derive the matrix differential equation for $P(t|t)$, we begin with (17-14), substitute (27-33) along with the expansions of $\Phi(t + \Delta t, t)$ and $\Gamma(t + \Delta t, t)$ into that equation, and use the fact that $K(t + \Delta t)$ has no zero-order terms in Δt , to show that

$$\begin{aligned} P(t + \Delta t|t + \Delta t) &= P(t + \Delta t|t) - K(t + \Delta t)H(t + \Delta t)P(t + \Delta t|t) \\ &= P(t|t) + [F(t)P(t|t) + P(t|t)F'(t) \\ &\quad + G(t)Q(t)G'(t)]\Delta t \\ &\quad - K(t + \Delta t)H(t + \Delta t)P(t + \Delta t|t) \end{aligned} \quad (27-36)$$

Consequently,

$$\begin{aligned} \lim_{\Delta t \rightarrow 0} \frac{P(t + \Delta t|t + \Delta t) - P(t|t)}{\Delta t} &= \dot{P}(t|t) = F(t)P(t|t) + P(t|t)F'(t) \\ &\quad + G(t)Q(t)G'(t) - \lim_{\Delta t \rightarrow 0} \frac{K(t + \Delta t)H(t + \Delta t)P(t + \Delta t|t)}{\Delta t} \end{aligned} \quad (27-37)$$

or finally, using (27-35)

$$\begin{aligned} \dot{P}(t|t) &= F(t)P(t|t) + P(t|t)F'(t) + G(t)Q(t)G'(t) \\ &\quad - P(t|t)H'(t)\mathbf{R}^{-1}(t)H(t)P(t|t) \end{aligned} \quad (27-38)$$

This completes the derivation of the KBF using a formal limiting procedure. It is also possible to obtain continuous-time smoothers by means of this procedure (e.g., see Meditch, 1969, Chapter 7).

DERIVATION OF KBF WHEN STRUCTURE OF THE FILTER IS PRESPECIFIED

In this derivation of the KBF we begin by assuming that the filter has the following structure,

$$\dot{\hat{x}}(t|t) = F(t)\hat{x}(t|t) + K(t)[z(t) - H\hat{x}(t|t)] \quad (27-39)$$

Our objective is to find the matrix function $\mathbf{K}(\tau)$, $t_0 \leq t \leq \tau$, that minimizes the following mean-squared error,

$$J[\mathbf{K}(\tau)] = \mathbf{E}\{\mathbf{e}'(\tau)\mathbf{e}(\tau)\} \quad \tau \geq t_0 \quad (27-40)$$

where

$$\mathbf{e}(\tau) = \mathbf{x}(\tau) - \hat{\mathbf{x}}(\tau|t) \quad (27-41)$$

This optimization problem is a fixed-time, free-end-point (i.e., τ is fixed but $\mathbf{e}(\tau)$ is not fixed) problem in the calculus of variations (e.g., Kwakernaak and Sivan, 1972; Athans and Falb, 1965; and Bryson and Ho, 1969).

It is straightforward to show that $\mathbf{E}\{\mathbf{e}(\tau)\} = \mathbf{0}$, so that $\mathbf{E}\{\mathbf{e}(\tau) - \mathbf{E}\{\mathbf{e}(\tau)\}\} = \mathbf{E}\{\mathbf{e}(\tau)\mathbf{e}'(\tau)\}$. Letting

$$\mathbf{P}(t|t) = \mathbf{E}\{\mathbf{e}(t)\mathbf{e}'(t)\} \quad (27-42)$$

we know that $J[\mathbf{K}(\tau)]$ can be reexpressed as

$$J[\mathbf{K}(\tau)] = \text{tr } \mathbf{P}(\tau|\tau) \quad (27-43)$$

We leave it to the reader to derive the following state equation for $\mathbf{e}(t)$, and its associated covariance equation,

$$\dot{\mathbf{e}}(t) = [\mathbf{F}(t) - \mathbf{K}(t)\mathbf{H}(t)]\mathbf{e}(t) + \mathbf{G}(t)\mathbf{w}(t) - \mathbf{K}(t)\mathbf{v}(t) \quad (27-44)$$

and

$$\begin{aligned} \dot{\mathbf{P}}(t|t) &= [\mathbf{F}(t) - \mathbf{K}(t)\mathbf{H}(t)]\mathbf{P}(t|t) + \mathbf{P}(t|t)[\mathbf{F}(t) - \mathbf{K}(t)\mathbf{H}(t)]' \\ &\quad + \mathbf{G}(t)\mathbf{Q}(t)\mathbf{G}'(t) + \mathbf{K}(t)\mathbf{R}(t)\mathbf{K}'(t) \end{aligned} \quad (27-45)$$

where $\mathbf{e}(t_0) = \mathbf{0}$ and $\mathbf{P}(t_0|t_0) = \mathbf{P}_*(t_0)$.

Our optimization problem for determining $\mathbf{K}(t)$ is: given the matrix differential equation (27-45), satisfied by the error-covariance matrix $\mathbf{P}(t|t)$, a terminal time τ , and the cost functional $J[\mathbf{K}(\tau)]$ in (27-43), determine the matrix $\mathbf{K}(t)$, $t_0 \leq t \leq \tau$ that minimizes $J[\mathbf{K}(t)]$.

The elements $p_{ij}(t|t)$ of $\mathbf{P}(t|t)$ may be viewed as the state variables of a dynamical system, and the elements $k_{ij}(t)$ of $\mathbf{K}(t)$ may be viewed as the control variables in an optimal control problem. The cost functional is then a terminal time penalty function on the state variables $p_{ij}(t|t)$. Euler-Lagrange equations associated with a free end-point problem can be used to determine the optimal gain matrix $\mathbf{K}(t)$.

To do this we define a set of costate variables $\sigma_{ij}(t)$ that correspond to the $p_{ij}(t|t)$, $i, j = 1, 2, \dots, n$. Let $\Sigma(t)$ be an $n \times n$ costate matrix that is associated with $\mathbf{P}(t|t)$, i.e., $\Sigma(t) = (\sigma_{ij}(t))_{ij}$. Next, we introduce the Hamiltonian function $\mathcal{H}(\mathbf{K}, \mathbf{P}, \Sigma)$ where for notational convenience we have omitted the dependence of \mathbf{K} , \mathbf{P} , and Σ on t , and,

$$\mathcal{H}(\mathbf{K}, \mathbf{P}, \Sigma) = \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij}(t) \dot{p}_{ij}(t|t) = \text{tr} [\dot{\mathbf{P}}(t|t)\Sigma'(t)] \quad (27-46)$$

Substituting (27-45) into (27-46), we see that

$$\begin{aligned} \mathcal{H}(\mathbf{K}, \mathbf{P}, \Sigma) &= \text{tr} [\mathbf{F}(t)\mathbf{P}(t|t)\Sigma'(t)] \\ &\quad - \text{tr} [\mathbf{K}(t)\mathbf{H}(t)\mathbf{P}(t|t)\Sigma'(t)] \\ &\quad + \text{tr} [\mathbf{P}(t|t)\mathbf{F}'(t)\Sigma'(t)] \\ &\quad - \text{tr} [\mathbf{P}(t|t)\mathbf{H}'(t)\mathbf{K}'(t)\Sigma'(t)] \\ &\quad + \text{tr} [\mathbf{G}(t)\mathbf{Q}(t)\mathbf{G}'(t)\Sigma'(t)] \\ &\quad + \text{tr} [\mathbf{K}(t)\mathbf{R}(t)\mathbf{K}'(t)\Sigma'(t)] \end{aligned} \quad (27-47)$$

The Euler-Lagrange equations for our optimization problem are:

$$\frac{\partial \mathcal{H}(\mathbf{K}, \mathbf{P}, \Sigma)}{\partial \mathbf{K}} \Big|_* = \mathbf{0} \quad (27-48)$$

$$\dot{\Sigma}^*(t) = -\frac{\partial \mathcal{H}(\mathbf{K}, \mathbf{P}, \Sigma)}{\partial \mathbf{P}} \Big|_* \quad (27-49)$$

$$\dot{\mathbf{P}}^*(t|t) = \frac{\partial \mathcal{H}(\mathbf{K}, \mathbf{P}, \Sigma)}{\partial \Sigma} \Big|_* \quad (27-50)$$

and

$$\Sigma^*(\tau) = \frac{\partial}{\partial \mathbf{P}} \text{tr } \mathbf{P}(\tau|\tau) \quad (27-51)$$

In these equations starred quantities denote optimal quantities, and $|_*$ denotes the replacement of \mathbf{K} , \mathbf{P} , and Σ by \mathbf{K}^* , \mathbf{P}^* , and Σ^* after the appropriate derivative has been calculated. Note, also, that the derivatives of $\mathcal{H}(\mathbf{K}, \mathbf{P}, \Sigma)$ are derivatives of a scalar quantity with respect to a matrix (e.g., \mathbf{K} , \mathbf{P} , or Σ). The calculus of gradient matrices (e.g., Schweppe, 1974; or Athans and Schweppe, 1965) can be used to evaluate these derivatives. The results are:

$$-\Sigma^*\mathbf{P}^*\mathbf{H}' - \Sigma^*\mathbf{P}^*\mathbf{H}' + \Sigma^*\mathbf{K}^*\mathbf{R} + \Sigma^*\mathbf{K}^*\mathbf{R} = \mathbf{0} \quad (27-52)$$

$$\dot{\Sigma}^* = -\Sigma^*(\mathbf{F} - \mathbf{K}^*\mathbf{H}) - (\mathbf{F} - \mathbf{K}^*\mathbf{H})'\Sigma^* \quad (27-53)$$

$$\dot{\mathbf{P}}^* = (\mathbf{F} - \mathbf{K}^*\mathbf{H})\mathbf{P}^* + \mathbf{P}^*(\mathbf{F} - \mathbf{K}^*\mathbf{H})' + \mathbf{G}\mathbf{Q}\mathbf{G}' + \mathbf{K}^*\mathbf{R}\mathbf{K}^* \quad (27-54)$$

and

$$\Sigma^*(\tau) = \mathbf{I} \quad (27-55)$$

Our immediate objective is to obtain an expression for $\mathbf{K}^*(t)$.

Fact. Matrix $\Sigma^*(t)$ is symmetric and positive definite. \square

We leave the proof of this fact as an exercise for the reader. Using this fact, and the fact that covariance matrix $\mathbf{P}^*(t)$ is symmetric, we are able to express (27-52) as

$$2\Sigma^*(\mathbf{K}^*\mathbf{R} - \mathbf{P}^*\mathbf{H}') = \mathbf{0} \quad (27-56)$$

Because $\Sigma^* > 0$, $(\Sigma^*)^{-1}$ exists so that (27-56) has for its only solution

$$\mathbf{K}^*(t) = \mathbf{P}^*(t|t)\mathbf{H}'(t)\mathbf{R}^{-1}(t) \quad (27-57)$$

which is the Kalman-Bucy gain matrix stated in Theorem 27-1. In order to obtain the covariance equation associated with $\mathbf{K}^*(t)$, substitute (27-57) into (27-54). The result is (27-12).

This completes the derivation of the KBF when the structure of the filter is prespecified.

STEADY-STATE KBF

If our continuous-time system is time-invariant and stationary, then, when certain system-theoretic conditions are satisfied (see, e.g., Kwakernaak and Sivan, 1972), $\dot{\mathbf{P}}(t|t) \rightarrow \mathbf{0}$ in which case $\mathbf{P}(t|t)$ has a steady-state value, denoted $\bar{\mathbf{P}}$. In this case, $\mathbf{K}(t) \rightarrow \bar{\mathbf{K}}$, where

$$\bar{\mathbf{K}} = \bar{\mathbf{P}}\mathbf{H}'\mathbf{R}^{-1} \quad (27-58)$$

$\bar{\mathbf{P}}$ is the solution of the algebraic Riccati equation

$$\mathbf{F}\bar{\mathbf{P}} + \bar{\mathbf{P}}\mathbf{F}' - \bar{\mathbf{P}}\mathbf{H}'\mathbf{R}^{-1}\bar{\mathbf{H}}\bar{\mathbf{P}} + \mathbf{G}\mathbf{Q}\mathbf{G}' = \mathbf{0} \quad (27-59)$$

and the steady-state KBF is asymptotically stable, i.e., the eigenvalues of $\mathbf{F} - \bar{\mathbf{K}}\mathbf{H}$ all lie in the left-half of the complex s -plane.

Example 27-1

Here we examine the steady-state KBF for the simplest second-order system, the double integrator,

$$\ddot{x}(t) = w(t) \quad (27-60)$$

and

$$z(t) = x(t) + v(t) \quad (27-61)$$

in which $w(t)$ and $v(t)$ are mutually uncorrelated white noise processes, with intensities q and r , respectively. With $x_1(t) = x(t)$ and $x_2(t) = \dot{x}(t)$, this system is expressed in state-variable format as

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} w \quad (27-62)$$

$$z = (1 \ 0) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + v \quad (27-63)$$

The algebraic Riccati equation for this system is

$$\begin{aligned} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \bar{p}_{11} & \bar{p}_{12} \\ \bar{p}_{12} & \bar{p}_{22} \end{pmatrix} + \begin{pmatrix} \bar{p}_{11} & \bar{p}_{12} \\ \bar{p}_{12} & \bar{p}_{22} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \\ - \begin{pmatrix} \bar{p}_{11} & \bar{p}_{12} \\ \bar{p}_{12} & \bar{p}_{22} \end{pmatrix} \begin{pmatrix} 1 \\ r \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \bar{p}_{11} & \bar{p}_{12} \\ \bar{p}_{12} & \bar{p}_{22} \end{pmatrix} \\ + \begin{pmatrix} 0 \\ 1 \end{pmatrix} q \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned} \quad (27-64)$$

Steady-State KBF

which leads to the following three algebraic equations:

$$2\bar{p}_{12} - \frac{1}{r}\bar{p}_{11}^2 = 0 \quad (27-65a)$$

$$\bar{p}_{22} - \frac{1}{r}\bar{p}_{11}\bar{p}_{12} = 0 \quad (27-65b)$$

and

$$-\frac{1}{r}\bar{p}_{12}^2 + q = 0 \quad (27-65c)$$

It is straightforward to show that the unique solution of these nonlinear algebraic equations, for which $\bar{\mathbf{P}} > 0$, is

$$\bar{p}_{12} = (qr)^{1/2} \quad (27-66a)$$

$$\bar{p}_{11} = \sqrt{2} q^{1/4} r^{3/4} \quad (27-66b)$$

$$\bar{p}_{22} = \sqrt{2} q^{3/4} r^{1/4} \quad (27-66c)$$

The steady-state KB gain matrix is computed from (27-58), as

$$\bar{\mathbf{K}} = \bar{\mathbf{P}}\mathbf{H}' \frac{1}{r} = \frac{1}{r} \begin{pmatrix} \bar{p}_{11} \\ \bar{p}_{12} \end{pmatrix} = \begin{pmatrix} \sqrt{2} (q/r)^{1/4} \\ (q/r)^{1/2} \end{pmatrix} \quad (27-67)$$

Observe that, just as in the discrete-time case, the single-channel KBF depends only on the ratio q/r .

Although we only needed \bar{p}_{11} and \bar{p}_{12} to compute $\bar{\mathbf{K}}$, \bar{p}_{22} is an important quantity, because

$$\bar{p}_{22} = \lim_{t \rightarrow \infty} \mathbf{E}\{[\dot{x}(t) - \hat{x}(t|t)]^2\} \quad (27-68)$$

Additionally,

$$\bar{p}_{11} = \lim_{t \rightarrow \infty} \mathbf{E}\{[x(t) - \hat{x}(t|t)]^2\} \quad (27-69)$$

Using (27-66b) and (27-66c), we find that

$$\bar{p}_{22} = (q/r)^{1/2} \bar{p}_{11} \quad (27-70)$$

If $q/r > 1$ (i.e., $\overline{\text{SNR}}$ possibly greater than unity), we will always have larger errors in estimation of $\dot{x}(t)$ than in estimation of $x(t)$. This is not too surprising because our measurement depends only on $x(t)$, and both $w(t)$ and $v(t)$ affect the calculation of $\dot{x}(t|t)$.

The steady-state KBF is characterized by the eigenvalues of matrix $\mathbf{F} - \bar{\mathbf{K}}\mathbf{H}$, where

$$\mathbf{F} - \bar{\mathbf{K}}\mathbf{H} = \begin{pmatrix} -\sqrt{2} (q/r)^{1/4} & 1 \\ -(q/r)^{1/2} & 0 \end{pmatrix} \quad (27-71)$$

These eigenvalues are solutions of the equation

$$s^2 + \sqrt{2} (q/r)^{1/4} s + (q/r)^{1/2} = 0 \quad (27-72)$$

When this equation is expressed in the normalized form

$$s^2 + 2\zeta\omega_n s + \omega_n^2 = 0$$

we find that

$$\omega_n = (q/r)^{1/4} \quad (27-73)$$

and

$$\zeta = 0.707 \quad (27-74)$$

thus, the steady-state KBF for the simple double integrator system is damped at 0.707. The filter's poles lie on the 45° line depicted in Figure 27-1. They can be moved along this line by adjusting the ratio q/r ; hence, once again, we may view q/r as a filter tuning parameter. \square

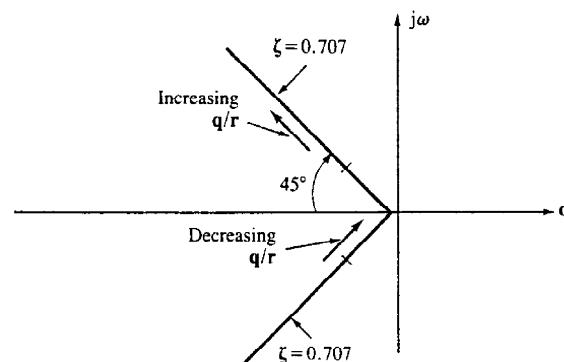


Figure 27-1 Eigenvalues of steady-state KBF lie along ± 45 degree lines. Increasing q/r moves them farther away from the origin, whereas decreasing q/r moves them closer to the origin.

AN IMPORTANT APPLICATION FOR THE KBF

Consider the system

$$\left. \begin{aligned} \dot{x}(t) &= F(t)x(t) + B(t)u(t) + G(t)w(t) \\ x(t_0) &= x_0 \end{aligned} \right\} \quad (27-75)$$

for $t \geq t_0$ where x_0 is a random initial condition vector with mean $m_x(t_0)$ and covariance matrix $P_x(t_0)$. Measurements are given by

$$z(t) = H(t)x(t) + v(t) \quad (27-76)$$

for $t \geq t_0$. The joint random process $\text{col}[w(t), v(t)]$ is a white noise process with intensity

$$\begin{pmatrix} Q(t) & 0 \\ 0 & R(t) \end{pmatrix} \quad t \geq t_0$$

The *controlled variable* can be expressed as

$$x(t) = D(t)u(t) \quad t \geq t_0 \quad (27-77)$$

The *stochastic linear optimal output feedback regulator problem* is the problem of finding the functional

$$u(t) = f[z(\tau)] \quad t_0 \leq \tau \leq t \quad (27-78)$$

for $t_0 \leq t \leq t_1$ such that the objective function

$$J[u] = E \left\{ \frac{1}{2} x'(t_1) W_1 x(t_1) + \frac{1}{2} \int_{t_0}^{t_1} [\xi'(\tau) W_3 \xi(\tau) + u'(\tau) W_2 u(\tau)] d\tau \right\} \quad (27-79)$$

is minimized. Here W_1 , W_2 , and W_3 are symmetric weighting matrices, and, $W_1 \geq 0$, $W_2 > 0$, and $W_3 > 0$ for $t_0 \leq t \leq t_1$.

In the control theory literature, this problem is also known as the linear-quadratic-Gaussian regulator problem (i.e., the LQG problem; see, Athans, 1971, for example). We state the structure of the solution to this problem, without proof, next.

The optimal control, $u^*(t)$, which minimizes $J[u]$ in (27-79) is

$$u^*(t) = -F^0(t)\hat{x}(t|t) \quad (27-80)$$

where $F^0(t)$ is an optimal gain matrix, computed as

$$F^0(t) = W_2^{-1}B(t)P_c(t) \quad (27-81)$$

where $P_c(t)$ is the solution of the control Riccati equation

$$\left. \begin{aligned} \dot{P}_c(t) &= F'(t)P_c(t) + P_c(t)F(t) - P_c(t)B(t)W_2^{-1}B'(t)P_c(t) \\ &\quad + D'(t)W_3D(t) \end{aligned} \right\} \quad (27-82)$$

$P_c(t_1)$ given

and $\hat{x}(t|t)$ is the output of a KBF, properly modified to account for the control term in the state equation, i.e.,

$$\dot{\hat{x}}(t|t) = F(t)\hat{x}(t|t) + B(t)u^*(t) + K(t)[z(t) - H(t)\hat{x}(t|t)] \quad (27-83)$$

We see that the KBF plays an essential role in the solution of the LQG problem.

PROBLEMS

27-1. Explain the replacement of covariance matrix $R(k+1=t+\Delta t)$ by $R(t+\Delta t)/\Delta t$ in (27-27).

27-2. Show that $\lim_{\Delta t \rightarrow 0} P(t+\Delta t|t) = P(t|t)$.

27-3. Derive the state equation for error $e(t)$, given in (27-44), and its associated covariance equation (27-45).

27-4. Prove that matrix $\Sigma^*(t)$ is symmetric and positive definite.

Sufficient Statistics and Statistical Estimation of Parameters

INTRODUCTION

In this lesson,* we discuss the usefulness of the notion of sufficient statistics in statistical estimation of parameters. Specifically, we discuss the role played by sufficient statistics and exponential families in maximum-likelihood and uniformly minimum-variance unbiased (UMVU) parameter estimation.

CONCEPT OF SUFFICIENT STATISTICS

The notion of a sufficient statistic can be explained intuitively (Ferguson, 1967), as follows. We observe $\mathbf{z}(N)$ (\mathbf{z} for short), where $\mathbf{z} = \text{col} [\mathbf{z}(1), \mathbf{z}(2), \dots, \mathbf{z}(N)]$, in which $\mathbf{z}(1), \dots, \mathbf{z}(N)$ are independent and identically distributed random vectors, each having a density function $p(\mathbf{z}(i)|\theta)$, where θ is unknown. Often the information in \mathbf{z} can be represented equivalently in a statistic, $T(\mathbf{z})$, whose dimension is independent of N , such that $T(\mathbf{z})$ contains all of the information about θ that is originally in \mathbf{z} . Such a statistic is known as a *sufficient statistic*.

Example A-1

Consider a sampled sequence of N manufactured cars. For each car we record whether it is defective or not. The observed sample can be represented as $\mathbf{z} = \text{col} [\mathbf{z}(1), \dots,$

* This lesson was written by Dr. Rama Chellappa, Department of Electrical Engineering-Systems, University of Southern California, Los Angeles, CA 90089.

$\mathbf{z}(N)]$, where $z(i) = 0$ if the i th car is not defective and $z(i) = 1$ if the i th car is defective. The total number of observed defective cars is

$$T(\mathbf{z}) = \sum_{i=1}^N z(i)$$

This is a statistic that maps many different values of $z(1), \dots, z(N)$ into the same value of $T(\mathbf{z})$. It is intuitively clear that, if one is interested in estimating the proportion θ of defective cars, nothing is lost by simply recording and using $T(\mathbf{z})$ in place of $z(1), \dots, z(N)$. The particular sequence of ones and zeros is irrelevant. Thus, as far as estimating the proportion of defective cars, $T(\mathbf{z})$ contains all the information contained in \mathbf{z} . \square

An advantage associated with the concept of a sufficient statistic is dimensionality reduction. In Example A-1, the dimensionality reduction is from N to 1.

Definition A-1. A statistic $T(\mathbf{z})$ is sufficient for vector parameter θ , if and only if the distribution of \mathbf{z} , conditioned on $T(\mathbf{z}) = t$, does not involve θ . \square

Example A-2

This example illustrates the application of Definition A-1 to identify a sufficient statistic for the model in Example A-1. Let θ be the probability that a car is defective. Then $z(1), z(2), \dots, z(N)$ is a record of N Bernoulli trials with probability θ ; thus, $Pr(t) = \theta^t(1 - \theta)^{N-t}$, $0 < \theta < 1$, where $t = \sum_{i=1}^N z(i)$, and $z(i) = 1$ or 0. The conditional distribution of $z(1), \dots, z(N)$, given $\sum_{i=1}^N z(i) = t$, is

$$P[\mathbf{z}|T = t] = \frac{P[\mathbf{z}, T = t]}{P(T = t)} = \frac{\theta^t(1 - \theta)^{N-t}}{\binom{N}{t} \theta^t(1 - \theta)^{N-t}} = \frac{1}{\binom{N}{t}}$$

which is independent of θ ; hence, $T(\mathbf{z}) = \sum_{i=1}^N z(i)$ is sufficient. Any one-to-one function of $T(\mathbf{z})$ is also sufficient. \square

This example illustrates that deriving a sufficient statistic using Definition A-1 can be quite difficult. An equivalent definition of sufficiency, which is easy to apply, is given in the following:

Theorem A-1 (Factorization Theorem). A necessary and sufficient condition for $T(\mathbf{z})$ to be sufficient for θ is that there exists a factorization

$$p(\mathbf{z}|\theta) = g(T(\mathbf{z}), \theta)h(\mathbf{z}) \quad (\text{A-1})$$

where the first factor in (A-1) may depend on θ , but depends on \mathbf{z} only through $T(\mathbf{z})$, whereas the second factor is independent of θ . \square

The proof of this theorem is given in Ferguson (1967) for the continuous case and Duda and Hart (1973) for the discrete case.

Example A-3 (Continuation of Example A-2)

In Example A-2, the probability distribution of samples $z(1), \dots, z(N)$ is

$$P[\mathcal{Z}|\theta] = \theta^t (1-\theta)^{N-t} \quad (\text{A-2})$$

where the total number of defective cars is $t = \sum_{i=1}^N z(i)$, and $z(i)$ is either 0 or 1. Equation (A-2) can be written equivalently as

$$P[\mathcal{Z}|\theta] = \exp \left[t \ln \frac{\theta}{1-\theta} + N \ln (1-\theta) \right] \quad (\text{A-3})$$

Comparing (A-3) with (A-1), we conclude that

$$\begin{aligned} h(\mathcal{Z}) &= 1 \\ g(T(\mathcal{Z}), \theta) &= \exp \left[t \ln \frac{\theta}{1-\theta} + N \ln (1-\theta) \right] \end{aligned}$$

and

$$T(\mathcal{Z}) = t = \sum_{i=1}^N z(i)$$

Using the Factorization Theorem, it was easy to determine $T(\mathcal{Z})$. \square

Example A-4

Let $\mathcal{Z} = \text{col}[z(1), \dots, z(N)]$ be a random sample drawn from a univariate Gaussian distribution, with unknown mean μ , and, known variance $\sigma^2 > 0$. Then

$$p(\mathcal{Z}) = \exp \left[\frac{\mu}{\sigma^2} \sum_{i=1}^N z(i) - \frac{N\mu^2}{2\sigma^2} \right] h(\mathcal{Z})$$

where

$$h(\mathcal{Z}) = \exp \left[-\frac{N}{2} \ln 2\pi\sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^N z^2(i) \right]$$

Based on the Factorization Theorem we identify $T(\mathcal{Z}) = \sum_{i=1}^N z(i)$ as a sufficient statistic for μ . \square

Because the concept of sufficient statistics involves reduction of data, it is worthwhile to know how far such a reduction can be done for a given problem. The dimension of the smallest set of statistics that is still sufficient for the parameters is called a *minimal sufficient statistic*. See Barankin (1959) and Datz (1959) for techniques useful in identifying a minimal sufficient statistic.

EXPONENTIAL FAMILIES OF DISTRIBUTIONS

It is of interest to study families of distributions, $p(z(i)|\theta)$ for which, irrespective of the sample size N , there exists a sufficient statistic of fixed dimension. The exponential families of distributions have this property. For exam-

ple, the family of normal distributions $N(\mu, \sigma^2)$, with σ^2 known and μ unknown, is an exponential family which, as we have seen in Example A-4, has a one-dimensional sufficient statistic for μ , that is equal to $\sum_{i=1}^N z(i)$. As Bickel and Doksum (1977) state,

Definition A-2 (Bickel and Doksum, 1977). *If there exist real-valued functions $a(\theta)$, and $b(\theta)$ on parameter space Θ , and real-valued functions $T(z)$ [z is short for $z(i)$] and $h(z)$ on \mathbb{R}^N , such that the density function $p(z|\theta)$ can be written as*

$$p(z|\theta) = \exp [a'(\theta)T(z) + b(\theta) + h(z)] \quad (\text{A-4})$$

then $p(z|\theta)$, $\theta \in \Theta$, is said to be a one-parameter exponential family of distributions. \square

The Gaussian, Binomial, Beta, Rayleigh, and Gamma distributions are examples of such one-parameter exponential families.

In a one parameter exponential family, $T(z)$ is sufficient for θ . The family of distributions obtained by sampling from one-parameter exponential families is also a one-parameter exponential family. For example, suppose that $z(1), \dots, z(N)$ are independent and identically distributed with common density $p(z|\theta)$; then,

$$p(\mathcal{Z}|\theta) = \exp \left[a'(\theta) \sum_{i=1}^N T(z(i)) + Nb(\theta) + \sum_{i=1}^N h(z(i)) \right] \quad (\text{A-5})$$

The sufficient statistic $T(\mathcal{Z})$ for this situation is

$$T(\mathcal{Z}) = \sum_{i=1}^N T(z(i))$$

Example A-5

Let $z(1), \dots, z(N)$ be a random sample from a multivariate Gaussian distribution with unknown $d \times 1$ mean vector μ and known covariance matrix P_μ . Then [z is short for $z(i)$]

$$p(z|\mu) = \exp [a'(\mu)T(z) + b(\mu) + h(z)]$$

where

$$a'(\mu) = \mu' P_\mu^{-1}$$

$$b(\mu) = -\frac{1}{2} \mu' P_\mu^{-1} \mu$$

$$T(z) = z$$

and

$$h(z) = \exp \left[-\frac{1}{2} z' P_\mu^{-1} z - \frac{d}{2} \ln 2\pi - \frac{1}{2} \ln \det P_\mu \right]$$

Additionally,

$$p(\mathbf{z}|\boldsymbol{\mu}) = \exp[\boldsymbol{\mu}' \mathbf{P}_{\boldsymbol{\mu}}^{-1} \mathbf{T}(\mathbf{z}) + Nb(\boldsymbol{\mu}) + h(\mathbf{z})]$$

where

$$\mathbf{T}(\mathbf{z}) = \sum_{i=1}^N \mathbf{z}(i)$$

and

$$h(\mathbf{z}) = \exp \left[-\frac{1}{2} \sum_{i=1}^N \mathbf{z}'(i) \mathbf{P}_{\boldsymbol{\mu}}^{-1} \mathbf{z}(i) - \frac{Nd}{2} \ln 2\pi - \frac{N}{2} \ln \det \mathbf{P}_{\boldsymbol{\mu}} \right] \quad \square$$

The notion of a one-parameter exponential family of distributions as stated in Bickel and Doksum (1977) can easily be extended to m parameters and vector observations in a straightforward manner.

Definition A-3. If there exist real matrices $\mathbf{A}_1, \dots, \mathbf{A}_m$, a real function b of $\boldsymbol{\theta}$, where $\boldsymbol{\theta} \in \Theta$, real matrices $\mathbf{T}_i(\mathbf{z})$ and a real function $h(\mathbf{z})$, such that the density function $p(\mathbf{z}|\boldsymbol{\theta})$ can be written as

$$p(\mathbf{z}|\boldsymbol{\theta}) = \exp \left\{ \text{tr} \left[\sum_{i=1}^m \mathbf{A}_i(\boldsymbol{\theta}) \mathbf{T}_i(\mathbf{z}) \right] \right\} \exp \{b(\boldsymbol{\theta}) + h(\mathbf{z})\} \quad (\text{A-6})$$

then $p(\mathbf{z}|\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Theta$ is said to be an m -parameter exponential family of distributions. \square

Example A-6

The family of d -variate normal distributions $N(\boldsymbol{\mu}, \mathbf{P}_{\boldsymbol{\mu}})$, where both $\boldsymbol{\mu}$ and $\mathbf{P}_{\boldsymbol{\mu}}$ are unknown, is an example of a 2-parameter exponential family in which $\boldsymbol{\theta}$ contains $\boldsymbol{\mu}$ and the elements of $\mathbf{P}_{\boldsymbol{\mu}}$. In this case

$$\mathbf{A}_1(\boldsymbol{\theta}) = \mathbf{a}_1(\boldsymbol{\theta}) = \mathbf{P}_{\boldsymbol{\mu}}^{-1} \boldsymbol{\mu}$$

$$\mathbf{T}_1(\mathbf{z}) = \mathbf{z}'$$

$$\mathbf{A}_2(\boldsymbol{\theta}) = -\frac{1}{2} \mathbf{P}_{\boldsymbol{\mu}}^{-1}$$

$$\mathbf{T}_2(\mathbf{z}) = \mathbf{z}\mathbf{z}'$$

$$b(\boldsymbol{\theta}) = -\frac{1}{2} \boldsymbol{\mu}' \mathbf{P}_{\boldsymbol{\mu}}^{-1} \boldsymbol{\mu} - \frac{d}{2} \ln 2\pi - \frac{1}{2} \ln \det \mathbf{P}_{\boldsymbol{\mu}}$$

and

$$h(\mathbf{z}) = 0 \quad \square$$

As is true for a one-parameter exponential family of distributions, if $\mathbf{z}(1), \dots, \mathbf{z}(N)$ are drawn randomly from an m -parameter exponential family, then $p[\mathbf{z}(1), \dots, \mathbf{z}(N)|\boldsymbol{\theta}]$ form an m -parameter exponential family with sufficient statistics $\mathbf{T}_1(\mathbf{z}), \dots, \mathbf{T}_m(\mathbf{z})$, where $\mathbf{T}_i(\mathbf{z}) = \sum_{j=1}^N \mathbf{T}_i(\mathbf{z}(j))$.

EXPONENTIAL FAMILIES AND MAXIMUM-LIKELIHOOD ESTIMATION

Let us consider a vector of unknown parameters $\boldsymbol{\theta}$ that describe a collection of N independent and identically distributed observations $\mathbf{z} = \text{col} [\mathbf{z}(1), \dots, \mathbf{z}(N)]$. The maximum-likelihood estimate (MLE) of $\boldsymbol{\theta}$ is obtained by maximizing the likelihood of $\boldsymbol{\theta}$ given the observations \mathbf{z} . Likelihood is defined in Lesson 11 to be proportional to the value of the probability density of the observations, given the parameters, i.e.,

$$l(\boldsymbol{\theta}|\mathbf{z}) \propto p(\mathbf{z}|\boldsymbol{\theta})$$

As discussed in Lesson 11, a sufficient condition for $l(\boldsymbol{\theta}|\mathbf{z})$ to be maximized is

$$J_0(\hat{\boldsymbol{\theta}}_{\text{ML}}|\mathbf{z}) < 0 \quad (\text{A-7})$$

where

$$J_0(\hat{\boldsymbol{\theta}}_{\text{ML}}|\mathbf{z}) = \left(\frac{\partial^2 L(\boldsymbol{\theta}|\mathbf{z})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}} \right)_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{\text{ML}}} \quad i, j = 1, 2, \dots, n$$

and $L(\boldsymbol{\theta}|\mathbf{z}) = \ln l(\boldsymbol{\theta}|\mathbf{z})$. Maximum-likelihood estimates of $\boldsymbol{\theta}$ are obtained by solving the system of n equations

$$\frac{\partial L(\boldsymbol{\theta}|\mathbf{z})}{\partial \boldsymbol{\theta}} = 0 \quad i = 1, 2, \dots, n \quad (\text{A-8})$$

for $\hat{\boldsymbol{\theta}}_{\text{ML}}$ and checking whether the solution to (A-8) satisfies (A-7).

When this technique is applied to members of exponential families, $\hat{\boldsymbol{\theta}}_{\text{ML}}$ can be obtained by solving a set of algebraic equations. The following theorem paraphrased from Bickel and Doksum (1977) formalizes this technique for vector observations.

Theorem A-2 (Bickel and Doksum, 1977). Let $p(\mathbf{z}|\boldsymbol{\theta}) = \exp[\mathbf{a}'(\boldsymbol{\theta})\mathbf{T}(\mathbf{z}) + b(\boldsymbol{\theta}) + h(\mathbf{z})]$ and let \mathcal{A} denote the interior of the range of $\mathbf{a}(\boldsymbol{\theta})$. If the equation

$$\mathbf{E}_{\boldsymbol{\theta}}\{\mathbf{T}(\mathbf{z})\} = \mathbf{T}(\mathbf{z}) \quad (\text{A-9})$$

has a solution $\hat{\boldsymbol{\theta}}(\mathbf{z})$ for which $\mathbf{a}[\hat{\boldsymbol{\theta}}(\mathbf{z})] \in \mathcal{A}$, then $\hat{\boldsymbol{\theta}}(\mathbf{z})$ is the unique MLE of $\boldsymbol{\theta}$. \square

The proof of this theorem can be found in Bickel and Doksum (1977).

Example A-7 (Continuation of Example A-5)

In this case

$$\mathbf{T}(\mathbf{z}) = \sum_{i=1}^N \mathbf{z}(i)$$

and

$$\mathbf{E}_{\boldsymbol{\mu}}\{\mathbf{T}(\mathbf{z})\} = N\boldsymbol{\mu}$$

hence (A-9) becomes

$$\sum_{i=1}^N \mathbf{z}(i) = N\boldsymbol{\mu}$$

whose solution, $\hat{\boldsymbol{\mu}}$, is

$$\hat{\boldsymbol{\mu}} = \frac{1}{N} \sum_{i=1}^N \mathbf{z}(i)$$

which is the well-known MLE of $\boldsymbol{\mu}$. \square

Theorem A-2 can be extended to the m -parameter exponential family case by using Definition A-3. We illustrate the applicability of this extension using the example given below.

Example A-8 (see, also, Example A-6)

Let $\mathbf{Z} = \text{col}[\mathbf{z}(1), \dots, \mathbf{z}(N)]$ be randomly drawn from $p(\mathbf{z}|\theta) = N(\boldsymbol{\mu}, \mathbf{P}_{\boldsymbol{\mu}})$, where both $\boldsymbol{\mu}$ and $\mathbf{P}_{\boldsymbol{\mu}}$ are unknown, so that θ contains $\boldsymbol{\mu}$ and the elements of $\mathbf{P}_{\boldsymbol{\mu}}$. Vector $\boldsymbol{\mu}$ is $d \times 1$ and matrix $\mathbf{P}_{\boldsymbol{\mu}}$ is $d \times d$, symmetric and positive definite. We express $p(\mathbf{Z}|\theta)$ as

$$\begin{aligned} p(\mathbf{Z}|\theta) &= (2\pi)^{-Nd/2} (\det \mathbf{P}_{\boldsymbol{\mu}})^{-N/2} \\ &\quad \exp \left[-\frac{1}{2} \sum_{i=1}^N (\mathbf{z}(i) - \boldsymbol{\mu})' \mathbf{P}_{\boldsymbol{\mu}}^{-1} (\mathbf{z}(i) - \boldsymbol{\mu}) \right] \\ &= (2\pi)^{-Nd/2} (\det \mathbf{P}_{\boldsymbol{\mu}})^{-N/2} \exp \left\{ -\frac{1}{2} \text{tr} \left[\mathbf{P}_{\boldsymbol{\mu}}^{-1} \right. \right. \\ &\quad \left. \left. \left(\sum_{i=1}^N (\mathbf{z}(i) - \boldsymbol{\mu})(\mathbf{z}(i) - \boldsymbol{\mu})' \right) \right] \right\} \\ &= (2\pi)^{-Nd/2} (\det \mathbf{P}_{\boldsymbol{\mu}})^{-N/2} \exp \left\{ -\frac{1}{2} \text{tr} \left[\mathbf{P}_{\boldsymbol{\mu}}^{-1} \sum_{i=1}^N \mathbf{z}(i) \mathbf{z}'(i) \right. \right. \\ &\quad \left. \left. - 2\boldsymbol{\mu}' \sum_{i=1}^N \mathbf{z}'(i) + N\boldsymbol{\mu}\boldsymbol{\mu}' \right] \right\} \end{aligned}$$

Using Theorem A-1 or Definition A-3 it can be seen that $\sum_{i=1}^N \mathbf{z}(i)$ and $\sum_{i=1}^N \mathbf{z}(i)\mathbf{z}'(i)$ are sufficient for $(\boldsymbol{\mu}, \mathbf{P}_{\boldsymbol{\mu}})$. Letting

$$\mathbf{T}_1(\mathbf{Z}) = \sum_{i=1}^N \mathbf{z}(i)$$

and

$$\mathbf{T}_2(\mathbf{Z}) = \sum_{i=1}^N \mathbf{z}(i)\mathbf{z}'(i)$$

we find that

$$\mathbf{E}_{\theta}[\mathbf{T}_1(\mathbf{Z})] = N\boldsymbol{\mu}$$

and

$$\mathbf{E}_{\theta}[\mathbf{T}_2(\mathbf{Z})] = N(\mathbf{P}_{\boldsymbol{\mu}} + \boldsymbol{\mu}\boldsymbol{\mu}')$$

Applying (A-9) to both $\mathbf{T}_1(\mathbf{Z})$ and $\mathbf{T}_2(\mathbf{Z})$, we obtain

$$N\boldsymbol{\mu} = \sum_{i=1}^N \mathbf{z}(i)$$

and

$$N(\mathbf{P}_{\boldsymbol{\mu}} + \boldsymbol{\mu}\boldsymbol{\mu}') = \sum_{i=1}^N \mathbf{z}(i)\mathbf{z}'(i)$$

whose solutions, $\hat{\boldsymbol{\mu}}$ and $\hat{\mathbf{P}}_{\boldsymbol{\mu}}$, are

$$\hat{\boldsymbol{\mu}} = \frac{1}{N} \sum_{i=1}^N \mathbf{z}(i)$$

and

$$\hat{\mathbf{P}}_{\boldsymbol{\mu}} = \frac{1}{N} \sum_{i=1}^N [\mathbf{z}(i) - \hat{\boldsymbol{\mu}}][\mathbf{z}(i) - \hat{\boldsymbol{\mu}}]'$$

which are the MLE's of $\boldsymbol{\mu}$ and $\mathbf{P}_{\boldsymbol{\mu}}$. \square

Example A-9 (Linear Model)

Consider the linear model

$$\mathbf{Z}(k) = \mathcal{H}(k)\theta + \mathcal{V}(k)$$

in which θ is an $n \times 1$ vector of deterministic parameters, $\mathcal{H}(k)$ is deterministic, and $\mathcal{V}(k)$ is a zero-mean white noise sequence, with known covariance matrix $\mathcal{R}(k)$. From (11-25) in Lesson 11, we can express $p(\mathbf{Z}(k)|\theta)$, as

$$p(\mathbf{Z}(k)|\theta) = \exp [a'(\theta)\mathbf{T}(\mathbf{Z}(k)) + b(\theta) + h(\mathbf{Z}(k))]$$

where

$$a'(\theta) = \theta'$$

$$\mathbf{T}(\mathbf{Z}(k)) = \mathcal{H}'(k)\mathcal{R}^{-1}(k)\mathcal{Z}(k)$$

$$b(\theta) = -\frac{N}{2} \ln 2\pi - \frac{1}{2} \ln \det \mathcal{R}(k) - \frac{1}{2} \theta' \mathcal{H}'(k)\mathcal{R}^{-1}(k)\mathcal{H}(k)\theta$$

and

$$h(\mathbf{Z}(k)) = -\frac{1}{2} \mathbf{Z}'(k)\mathcal{R}^{-1}(k)\mathbf{Z}(k)$$

Observe that

$$\mathbf{E}_{\theta}[\mathcal{H}(k)\mathcal{R}^{-1}(k)\mathcal{Z}(k)] = \mathcal{H}'(k)\mathcal{R}^{-1}(k)\mathcal{H}(k)\theta$$

hence, applying (A-9), we obtain

$$\mathcal{H}(k)\mathcal{R}^{-1}(k)\mathcal{H}(k)\hat{\theta} = \mathcal{H}'(k)\mathcal{R}^{-1}(k)\mathcal{Z}(k)$$

whose solution, $\hat{\theta}(k)$, is

$$\hat{\theta}(k) = [\mathcal{H}'(k)\mathcal{R}^{-1}(k)\mathcal{H}(k)]^{-1}\mathcal{H}'(k)\mathcal{R}^{-1}(k)\mathcal{Z}(k)$$

which is the well-known expression for the MLE of θ (see Theorem 11-3). The case when $\mathbf{R}(k) = \sigma^2\mathbf{I}$, where σ^2 is unknown can be handled in a manner very similar to that in Example A-8. \square

SUFFICIENT STATISTICS AND UNIFORMLY MINIMUM-VARIANCE UNBIASED ESTIMATION

In this section we discuss how sufficient statistics can be used to obtain uniformly minimum-variance unbiased (UMVU) estimates. Recall, from Lesson 6, that an estimate $\hat{\theta}$ of parameter θ is said to be unbiased if

$$E\{\hat{\theta}\} = \theta \quad (\text{A-10})$$

Among such unbiased estimates, we can often find one estimate, denoted θ^* , which improves all other estimates in the sense that

$$\text{var}(\theta^*) \leq \text{var}(\hat{\theta}) \quad (\text{A-11})$$

When (A-11) is true for all (admissible) values of θ , θ^* is known as the UMVU estimate of θ . The UMVU estimator is obtained by choosing the estimator which has the minimum variance among the class of unbiased estimators. If the estimator is constrained further to be a *linear* function of the observations, then it becomes the BLUE which was discussed in Lesson 9.

Suppose we have an estimate, $\hat{\theta}(\mathcal{Z})$, of parameter θ that is based on observations $\mathcal{Z} = \text{col}[z(1), \dots, z(N)]$. Assume further that $p(\mathcal{Z}|\theta)$ has a finite-dimensional sufficient statistic, $T(\mathcal{Z})$, for θ . Using $T(\mathcal{Z})$, we can construct an estimate $\theta^*(\mathcal{Z})$ which is at least as good as, or even better, than $\hat{\theta}$ by the celebrated Rao-Blackwell Theorem (Bickel and Doksum, 1977). We do this by computing the conditional expectation of $\hat{\theta}(\mathcal{Z})$, i.e.,

$$\theta^*(\mathcal{Z}) = E\{\hat{\theta}(\mathcal{Z})|T(\mathcal{Z})\} \quad (\text{A-12})$$

Estimate $\theta^*(\mathcal{Z})$ is “better than $\hat{\theta}$ ” in the sense that $E\{[\theta^*(\mathcal{Z}) - \theta]^2\} < E\{[\hat{\theta}(\mathcal{Z}) - \theta]^2\}$. Because $T(\mathcal{Z})$ is sufficient, the conditional expectation $E\{\hat{\theta}(\mathcal{Z})|T(\mathcal{Z})\}$ will not depend on θ ; hence, $\theta^*(\mathcal{Z})$ is a function of \mathcal{Z} only. Application of this conditioning technique can only improve an estimate such as $\hat{\theta}(\mathcal{Z})$; it does not guarantee that $\theta^*(\mathcal{Z})$ will be the UMVU estimate. To obtain the UMVU estimate using this conditioning technique, we need the additional concept of *completeness*.

Definition A-4 [Lehmann (1959; 1980); Bickel and Doksum (1977)]. *A sufficient statistic $T(\mathcal{Z})$ is said to be complete, if the only real-valued function, g , defined on the range of $T(\mathcal{Z})$, which satisfies $E_\theta\{g(T)\} = 0$ for all θ , is the function $g(T) = 0$. \square*

Completeness is a property of the family of distributions of $T(\mathcal{Z})$ generated as θ varies over its range. The concept of a complete sufficient statistic, as stated by Lehmann (1983), can be viewed as an extension of the notion of sufficient statistics in reducing the amount of useful information required for the estimation of θ . Although a sufficient statistic achieves data reduction, it may contain some additional information not required for the estimation of θ . For instance, it may be that $E_\theta[g(T(\mathcal{Z}))]$ is a constant independent of θ for some nonconstant function g . If so, we would like to have $E_\theta[g(T(\mathcal{Z}))] = c$, (constant independent of θ) imply that $g(T(\mathcal{Z})) = c$. By subtracting c from $E_\theta[g(T(\mathcal{Z}))]$, one arrives at Definition A-4. Proving completeness using Definition A-4 can be cumbersome. In the special case when $p(z(k)|\theta)$ is a one-parameter exponential family, i.e., when

$$p(z(k)|\theta) = \exp[a(\theta)T(z(k)) + b(\theta) + h(z(k))] \quad (\text{A-13})$$

the completeness of $T(z(k))$ can be verified by checking if the range of $a(\theta)$ has an open interval (Lehmann, 1959).

Example A-10

Let $\mathcal{Z} = \text{col}[z(1), \dots, z(N)]$ be a random sample drawn from a univariate Gaussian distribution whose mean μ is unknown, and whose variance $\sigma^2 > 0$ is known. From Example A-5, we know that the distribution of \mathcal{Z} forms a one-parameter exponential family, with $T(\mathcal{Z}) = \sum_{i=1}^N z(i)$ and $a(\mu) = \mu/\sigma^2$. Because $a(\mu)$ ranges over an open interval as μ varies from $-\infty$ to $+\infty$, $T(\mathcal{Z}) = \sum_{i=1}^N z(i)$ is complete and sufficient.

The same conclusion can be obtained using Definition A-4 as follows. We must show that the Gaussian family of probability distributions (with μ unknown and σ^2 fixed) is complete. Note that the sufficient statistic $T(\mathcal{Z}) = \sum_{i=1}^N z(i)$ (see Example A-5) is Gaussian with mean $N\mu$ and variance $N^2\sigma^2$. Suppose g is a function such that $E_\mu\{g(T)\} = 0$ for all $-\infty < \mu < \infty$; then,

$$\begin{aligned} & \int_{-\infty}^{\infty} \frac{g(T)}{N\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2 N^2}(T - N\mu)^2\right] dT \\ &= \int_{-\infty}^{\infty} \sqrt{2\pi} g(v\sigma N + N\mu) \exp\left(-\frac{v^2}{2}\right) dv = 0 \end{aligned} \quad (\text{A-14})$$

implies $g(\cdot) = 0$ for all values of the argument of g . \square

Other interesting examples that prove completeness for families of distributions are found in Lehmann (1959).

Once a complete and sufficient statistic $T(\mathcal{Z})$ is known for a given parameter estimation problem, the Lehmann-Scheffe Theorem, given next, can be used to obtain a unique UMVU estimate. This theorem is paraphrased from Bickel and Doksum (1977).

Theorem A-3 [Lehmann-Scheffe Theorem (e.g., Bickel and Doksum, 1977)]. *If a complete and sufficient statistic, $T(\mathcal{Z})$, exists for θ , and $\hat{\theta}$ is an unbiased estimator of θ , then $\theta^*(\mathcal{Z}) = E\{\hat{\theta}|T(\mathcal{Z})\}$ is an UMVU estimator of θ . If*

Variance $[\theta^*(\mathcal{L})] < \infty$ for all θ , then $\theta^*(\mathcal{L})$ is the unique UMVU estimate of θ . \square

A proof of this theorem can be found in Bickel and Doksum (1977).

This theorem can be applied in two ways to determine an UMVU estimator [Bickel and Doksum (1977), and Lehmann (1959)].

Method 1. Find a statistic of the form $h(T(\mathcal{L}))$ such that

$$E\{h(T(\mathcal{L}))\} = \theta \quad (\text{A-15})$$

where $T(\mathcal{L})$ is a complete and sufficient statistic for θ . Then, $h(T(\mathcal{L}))$ is an UMVU estimator of θ . This follows from the fact that

$$E\{h(T(\mathcal{L}))|T(\mathcal{L})\} = h(T(\mathcal{L}))$$

Method 2. Find an unbiased estimator, $\hat{\theta}$, of θ ; then, $E\{\hat{\theta}|T(\mathcal{L})\}$ is an UMVU estimator of θ for a complete and sufficient statistic $T(\mathcal{L})$.

Example A-11 (Continuation of Example A-10)

We know that $T(\mathcal{L}) = \sum_{i=1}^N z(i)$ is a complete and sufficient statistic for μ . Furthermore, $1/N \sum_{i=1}^N z(i)$ is an unbiased estimator of μ ; hence, we obtain the well-known result from Method 1, that the sample mean, $1/N \sum_{i=1}^N z(i)$, is an UMVU estimate of μ . Because this estimator is linear, it is also the BLUE of μ . \square

Example A-12 (Linear Model)

As in Example A-9, consider the linear model

$$\mathcal{Z}(k) = \mathcal{H}(k)\theta + \mathcal{V}(k) \quad (\text{A-16})$$

where θ is a deterministic but unknown $n \times 1$ vector of parameters, $\mathcal{H}(k)$ is deterministic, and $E\{\mathcal{V}(k)\} = 0$. Additionally, assume that $\mathcal{V}(k)$ is Gaussian with known covariance matrix $\mathcal{R}(k)$. Then, the statistic $T(\mathcal{Z}(k)) = \mathcal{H}'(k)\mathcal{R}^{-1}\mathcal{Z}(k)$ is sufficient (see Example A-9). That it is also complete can be seen by using Theorem A-4. To obtain UMVU estimate θ , we need to identify a function $h[T(\mathcal{Z}(k))]$ such that $E\{h[T(\mathcal{Z}(k))]\} = \theta$. The structure of $h[T(\mathcal{Z}(k))]$ is obtained by observing that

$$E\{T(\mathcal{Z}(k))\} = E\{\mathcal{H}'(k)\mathcal{R}^{-1}(k)\mathcal{Z}(k)\} = \mathcal{H}'(k)\mathcal{R}^{-1}(k)\mathcal{H}(k)\theta$$

hence,

$$[\mathcal{H}'(k)\mathcal{R}^{-1}(k)\mathcal{H}(k)]^{-1}E\{T(\mathcal{Z}(k))\} = \theta$$

Consequently, the UMVU estimator of θ is

$$[\mathcal{H}'(k)\mathcal{R}^{-1}(k)\mathcal{H}(k)]^{-1}\mathcal{H}'(k)\mathcal{R}^{-1}(k)\mathcal{Z}(k)$$

which agrees with Equation (9-26). \square

We now generalize the discussions given above to the case of an m -parameter exponential family, and scalar observations. This theorem is paraphrased from Bickel and Doksum (1977).

Theorem A-4 [Bickel and Doksum (1977) and Lehmann (1959)]. Let $p(z|\theta)$ be an m -parameter exponential family given by

$$p(z|\theta) = \exp \left[\sum_{i=1}^m a_i(\theta)T_i(z) + b(\theta) + h(z) \right]$$

where a_1, \dots, a_m and b are real-valued functions of θ , and T_1, \dots, T_m and h are real-valued functions of z . Suppose that the range of $\mathbf{a} = \text{col}[a_1(\theta), \dots, a_m(\theta)]$ has an open m -rectangle [if $(x_1, y_1), \dots, (x_m, y_m)$ are m open intervals, the set $\{(s_1, \dots, s_m) : x_i < s_i < y_i, 1 \leq i \leq m\}$ is called the open m -rectangle], then $T(z) = \text{col}[T_1(z), \dots, T_m(z)]$ is complete as well as sufficient. \square

Example A-13 (This example is taken from Bickel and Doksum, 1977, pp. 123–124)

As in Example A-6, let $\mathcal{Z} = \text{col}[z(1), \dots, z(N)]$ be a sample from a $N(\mu, \sigma^2)$ population where both μ and σ^2 are unknown. As a special case of Example A-6, we observe that the distribution of \mathcal{Z} forms a two-parameter exponential family where $\theta = \text{col}(\mu, \sigma^2)$. Because $\text{col}[a_1(\theta), a_2(\theta)] = \text{col}(\mu/\sigma^2, -1/2\sigma^2)$ ranges over the lower halfplane, as θ ranges over $\text{col}((-\infty, \infty), (0, \infty))$, the conditions of Theorem A-4 are satisfied. As a result, $T(\mathcal{Z}) = \text{col}[\sum_{i=1}^N z(i), \sum_{i=1}^N z^2(i)]$ is complete and sufficient. \square

Theorem A-3 also generalizes in a straightforward manner to:

Theorem A-5. If a complete and sufficient statistic $T(\mathcal{Z}) = \text{col}(T_1(\mathcal{Z}), \dots, T_m(\mathcal{Z}))$ exists for θ , and $\hat{\theta}$ is an unbiased estimator of θ , then $\theta^*(\mathcal{Z}) = E\{\hat{\theta}|T(\mathcal{Z})\}$ is an UMVU estimator of θ . If the elements of the covariance matrix of $\theta^*(\mathcal{Z})$ are $< \infty$ for all θ , then $\theta^*(\mathcal{Z})$ is the unique UMVU estimate of θ . \square

The proof of this theorem is a straightforward extension of the proof of Theorem A-3, which can be found in Bickel and Doksum (1977).

Example A-14 (Continuation of Example A-13)

In Example A-13 we saw that $\text{col}[T_1(\mathcal{Z}), T_2(\mathcal{Z})] = \text{col}[\sum_{i=1}^N z(i), \sum_{i=1}^N z^2(i)]$ is sufficient and complete for both μ and σ^2 . Furthermore, since

$$\bar{z} = \frac{1}{N} \sum_{i=1}^N z(i)$$

and

$$\bar{\sigma}^2 = \frac{1}{N-1} \sum_{i=1}^N [z(i) - \bar{z}]^2$$

are unbiased estimators of μ and σ^2 , respectively, we use the extension of Method 1 to the vector parameter case to conclude that \bar{z} and $\bar{\sigma}^2$ are UMVU estimators of μ and σ^2 . \square

It is not always possible to identify a function $h(T(\mathcal{Z}))$ that is an unbiased estimator of θ . Examples that use the conditioning Method 2 to obtain

UMVU estimators are found, for example, in Bickel and Doksum (1977) and Lehman (1980).

PROBLEMS

- A-1.** Suppose $z(1), \dots, z(N)$ are independent random variables, each uniform on $[0, \theta]$, where $\theta > 0$ is unknown. Find a sufficient statistic for θ .
- A-2.** Suppose we have two independent observations from the Cauchy distribution,

$$p(z) = \frac{1}{\pi} \frac{1}{1 + (z - \theta)^2} \quad -\infty < z < \infty$$

Show that no sufficient statistic exists for θ .

- A-3.** Let $z(1), z(2), \dots, z(N)$ be generated by the first-order auto-regressive process,
- $$z(i) = \theta z(i-1) + \sqrt{\beta} w(i)$$

where $\{w(i), i = 1, \dots, N\}$ is an independent and identically distributed Gaussian noise sequence with zero mean and unit variance. Find a sufficient statistic for θ and β .

- A-4.** Suppose that $T(\mathcal{L})$ is sufficient for θ , and that $\hat{\theta}(\mathcal{L})$ is a maximum-likelihood estimate of θ . Show that $\hat{\theta}(\mathcal{L})$ depends on \mathcal{L} only through $T(\mathcal{L})$.

- A-5.** Using Theorem A-2, derive the maximum-likelihood estimator of θ when observations $z(1), \dots, z(N)$ denote a sample from

$$p(z(i)|\theta) = \theta e^{-\theta z(i)} \quad z(i) \geq 0, \theta > 0$$

- A-6.** Show that the family of Bernoulli distributions, with unknown probability of success p ($0 \leq p \leq 1$), is complete.

- A-7.** Show that the family of uniform distributions on $(0, \theta)$, where $\theta > 0$ is unknown, is complete.

- A-8.** Let $z(1), \dots, z(N)$ be independent and identically distributed samples, where $p(z(i)|\theta)$ is a Bernoulli distribution with unknown probability of success p ($0 \leq p \leq 1$). Find a complete sufficient statistic, T ; the UMVU estimate $\phi(T)$ of p ; and, the variance of $\phi(T)$.

- A-9.** [Taken from Bickel and Doksum (1977)]. Let $z(1), z(2), \dots, z(N)$ be an independent and identically distributed sample from $N(\mu, 1)$. Find the UMVU estimator of $p_\mu[z(1) \geq 0]$.

- A-10.** [Taken from Bickel and Doksum (1977)]. Suppose that T_1 and T_2 are two UMVU estimates of θ with finite variances. Show that $T_1 = T_2$.

- A-11.** In Example A-12 prove that $T(\mathcal{L}(k))$ is complete.

Appendix A

Glossary of Major Results

Equations (3-10) and (3-11)	Batch formulas for $\hat{\theta}_{WLS}(k)$ and $\hat{\theta}_{LS}(k)$.
Theorem 4-1	Information form of recursive LSE.
Lemma 4-1	Matrix inversion lemma.
Theorem 4-2	Covariance form of recursive LSE.
Theorem 5-1	Multistage LSE.
Theorem 6-1	Necessary and sufficient conditions for a linear batch estimator to be unbiased.
Theorem 6-2	Sufficient condition for a linear recursive estimator to be unbiased.
Theorem 6-3	Cramer-Rao inequality for a scalar parameter.
Corollary 6-1	Achieving the Cramer-Rao lower bound.
Theorem 6-4	Cramer-Rao inequality for a vector of parameters.
Corollary 6-2	Inequality for error-variance of i th parameter.
Theorem 7-1	Mean-squared convergence implies convergence in probability.
Theorem 7-2	Conditions under which $\hat{\theta}(k)$ is a consistent estimator of θ .
Theorem 8-1	Sufficient conditions for $\hat{\theta}_{WLS}(k)$ to be an unbiased estimator of θ .
Theorem 8-2	A formula for $\text{cov}[\hat{\theta}_{WLS}(k)]$.

Corollary 8-1

Theorem 8-3

Theorem 8-4

Theorem 8-5

Equation (9-22)

Theorem 9-1

Corollary 9-1

Theorem 9-2

Theorem 9-3

Corollary 9-2

Theorem 9-4

Corollary 9-3

Theorem 9-5

Theorem 9-6

Theorem 9-7

Definition 10-1

Theorem 10-1

Theorem 11-1

Theorem 11-2

Theorem 11-3

Corollary 11-1

Theorem 12-1

Theorem 12-2

Theorem 12-3

Theorem 12-4

Glossary of Major Results

A formula for $\text{cov}[\hat{\theta}_{\text{LS}}(k)]$ under special conditions on the measurement noise.

An unbiased estimator of σ_v^2 .

Sufficient conditions for $\hat{\theta}_{\text{LS}}(k)$ to be a consistent estimator of θ .

Sufficient conditions for $\hat{\sigma}_v^2(k)$ to be a consistent estimator of σ_v^2 .

Batch formula for $\hat{\theta}_{\text{BLU}}(k)$.

The relationship between $\hat{\theta}_{\text{BLU}}(k)$ and $\hat{\theta}_{\text{WLS}}(k)$.

When all the results obtained in Lessons 3, 4 and 5 for $\hat{\theta}_{\text{WLS}}(k)$ can be applied to $\hat{\theta}_{\text{BLU}}(k)$.

When $\hat{\theta}_{\text{BLU}}(k)$ equals $\hat{\theta}_{\text{LS}}(k)$ (Gauss-Markov Theorem).

A formula for $\text{cov}[\hat{\theta}_{\text{BLU}}(k)]$.

The equivalence between $P(k)$ and $\text{cov}[\hat{\theta}_{\text{BLU}}(k)]$.

Most efficient estimator property of $\hat{\theta}_{\text{BLU}}(k)$.

When $\hat{\theta}_{\text{LS}}(k)$ is a most efficient estimator of θ .

Invariance of $\hat{\theta}_{\text{BLU}}(k)$ to scale changes.

Information form of recursive BLUE.

Covariance form of recursive BLUE.

Likelihood defined.

Likelihood ratio of combined data from statistically independent sets of data.

Large-sample properties of maximum-likelihood estimates.

Invariance property of MLE's.

Condition under which $\hat{\theta}_{\text{ML}}(k) = \hat{\theta}_{\text{BLU}}(k)$.

Conditions under which $\hat{\theta}_{\text{ML}}(k) = \hat{\theta}_{\text{BLU}}(k) = \hat{\theta}_{\text{LS}}(k)$, and, resulting estimator properties.

A formula for $p(\mathbf{x}|y)$ when \mathbf{x} and y are jointly Gaussian.

Properties of $E\{\mathbf{x}|y\}$ when \mathbf{x} and y are jointly Gaussian.

Expansion formula for $E\{\mathbf{x}|y,z\}$ when \mathbf{x} , y , and \mathbf{z} are jointly Gaussian, and y and \mathbf{z} are statistically independent.

Expansion formula for $E\{\mathbf{x}|y,z\}$ when \mathbf{x} , y and \mathbf{z} are jointly Gaussian and y and \mathbf{z} are not necessarily statistically independent.

Appendix A

Theorem 13-1

Corollary 13-1

Corollary 13-2

Corollary 13-3

Theorem 13-2

Theorem 14-1

Theorem 14-2

Theorem 14-3

Theorem 15-1

Theorem 15-2

Theorem 15-3

Equations (15-17) & (15-18)

Theorem 15-4

Theorem 15-5

Theorem 15-6

Equations (16-4) & (16-11)

Theorem 16-1

Theorem 16-2

Theorem 17-1

Theorem 19-1

Theorem 19-2

Theorem 20-1

Corollary 20-1

Corollary 20-2

A formula for $\hat{\theta}_{\text{MS}}(k)$ (The Fundamental Theorem of Estimation Theory).

A formula for $\hat{\theta}_{\text{MS}}(k)$ when θ and $\mathbf{Z}(k)$ are jointly Gaussian.

A linear mean-squared estimator of θ in the non-Gaussian case.

Orthogonality principle.

When $\hat{\theta}_{\text{MAP}}(k) = \hat{\theta}_{\text{MS}}(k)$.

Conditions under which $\hat{\theta}_{\text{MS}}(k) = \hat{\theta}_{\text{BLU}}(k)$.

Condition under which $\hat{\theta}_{\text{MS}}(k) = \hat{\theta}_{\text{BLU}}^e(k)$.

Condition under which $\hat{\theta}_{\text{MAP}}(k) = \hat{\theta}_{\text{BLU}}^e(k)$.

Expansion of a joint probability density function for a first-order Markov process.

Calculation of conditional expectation for a first-order Markov process.

Interpretation of Gaussian white noise as a special first-order Markov process.

The basic state-variable model.

Conditions under which $\mathbf{x}(k)$ is a Gauss-Markov sequence.

Recursive equations for computing $\mathbf{m}_x(k)$ and $\mathbf{P}_x(k)$.

Formulas for computing $\mathbf{m}_z(k)$ and $\mathbf{P}_z(k)$.

Single-stage predictor formulas for $\hat{\mathbf{x}}(k|k-1)$ and $\mathbf{P}(k|k-1)$.

Formula for and properties of general state predictor, $\hat{\mathbf{x}}(k|j)$, $k > j$.

Representations and properties of the innovations process.

Kalman filter formulas and properties of resulting estimates and estimation error.

Steady-state Kalman filter.

Equivalence of steady-state Kalman filter and infinite length digital Wiener filter.

Single-state smoother formula for $\hat{\mathbf{x}}(k|k+1)$.

Relationship between single-stage smoothing gain matrix and Kalman gain matrix.

Another way to express $\hat{\mathbf{x}}(k|k+1)$.

Theorem 20-2	Double-stage smoother formula for $\hat{x}(k k+2)$.
Corollary 20-3	Relationship between double-stage smoothing gain matrix and Kalman gain matrix.
Corollary 20-4	Two other ways to express $\hat{x}(k k+2)$.
Theorem 21-1	Formulas for a useful fixed-interval smoother of $x(k)$, $\hat{x}(k N)$, and its error-covariance matrix, $P(k N)$.
Theorem 21-2	Formulas for a most useful two-pass fixed-interval smoother of $x(k)$ and its associated error-covariance matrix.
Theorem 21-3	Formulas for a most useful fixed-point smoothed estimator of $x(k)$, $\hat{x}(k k+l)$ where $l = 1, 2, \dots$, and its associated error-covariance matrix, $P(k k+l)$.
Theorem 22-1	Conditions under which a single-channel state-variable model is equivalent to a convolutional sum model.
Theorem 22-2	Recursive minimum-variance deconvolution formulas.
Theorem 22-3	Steady-state MVD filter, and zero phase nature of $\hat{\mu}_e(k N)$.
Theorem 22-4	Equivalence between steady-state MVD filter and Berkhouit's infinite impulse response digital Wiener deconvolution filter.
Theorem 22-5	Maximum-likelihood deconvolution results.
Theorem 22-6	Structure of minimum-variance waveshaper.
Theorem 22-7	Recursive fixed-interval waveshaping results.
Theorem 23-1	How to handle biases that may be present in a state-variable model.
Theorem 23-2	Predictor-corrector Kalman filter for the correlated noise case.
Corollary 23-1	Recursive predictor formulas for the correlated noise case.
Corollary 23-2	Recursive filter formulas for the correlated noise case.
Equations (24-1) & (24-2)	Nonlinear state-variable model.
Equations (24-23) & (24-30)	Perturbation state-variable model.
Theorem 24-1	Solution to a time-varying continuous-time state equation.

Equations (24-39) & (24-44)	Discretized state-variable model.
Theorem 25-1	A consequence of relinearizing about $\hat{x}(k k)$.
Equations (25-22) & (25-27)	Extended Kalman filter prediction and correction equations.
Theorem 26-1	Formula for the log-likelihood function of the basic state-variable model.
Theorem 26-2	Closed-form formula for the maximum-likelihood estimate of the steady-state value of the innovation's covariance matrix.
Theorem 27-1	Kalman-Bucy filter equations.
Definition A-1	Sufficient statistic defined.
Theorem A-1	Factorization theorem.
Theorem A-2	A method for computing the unique maximum-likelihood estimator of θ that is associated with exponential families of distributions.
Theorem A-3	Lehmann-Scheffe Theorem. Provides a uniformly minimum-variance unbiased estimator of θ .
Theorem A-4	Method for determining whether or not $T(\mathbf{z})$ is complete as well as sufficient when $p(\mathbf{z} \theta)$ is an m -parameter exponential family.
Theorem A-5	Provides a uniformly minimum-variance unbiased estimator of vector θ .

References

301

- AGUILERA, R., J. A. DEBREMAECKER, and S. HERNANDEZ. 1970. "Design of recursive filters." *Geophysics*, Vol. 35, pp. 247-253.
- ANDERSON, B. D. O., and J. B. MOORE. 1979. *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall.
- AOKI, M. 1967. *Optimization of Stochastic Systems—Topics in Discrete-Time Systems*. NY: Academic Press.
- ÅSTRÖM, K. J. 1968. "Lectures on the identification problem—the least squares method." Rept. No. 6806, Lund Institute of Technology, Division of Automatic Control.
- ATHANS, M. 1971. "The role and use of the stochastic linear-quadratic-Gaussian problem in control system design." *IEEE Trans. on Automatic Control*, Vol. AC-16, pp. 529-552.
- ATHANS, M., and P. L. FALB. 1965. *Optimal Control: An Introduction to the Theory and Its Applications*. NY: McGraw-Hill.
- ATHANS, M., and F. SCHWEPPE. 1965. "Gradient matrices and matrix calculations." MIT Lincoln Labs., Lexington, MA, Tech. Note 1965-53.
- ATHANS, M., and E. TSE. 1967. "A direct derivation of the optimal linear filter using the maximum principle." *IEEE Trans. on Automatic Control*, Vol. AC-12, pp. 690-698.
- ATHANS, M., R. P. WISHNER, and A. BERTOLINI. 1968. "Suboptimal state estimation for continuous-time nonlinear systems from discrete noisy measurements." *IEEE Trans. on Automatic Control*, Vol. AC-13, pp. 504-514.
- BARANKIN, E. W., and M. KATZ, JR. 1959. "Sufficient Statistics of Minimal Dimension." *Sankhya*, Vol. 21, pp. 217-246.
- BARANKIN, E. W. 1961. "Application to Exponential Families of the Solution to the Minimal Dimensionality Problem for Sufficient Statistics." *Bull. Inst. Internat. Stat.*, Vol. 38, pp. 141-150.

References

- BARD, Y. 1970. "Comparison of gradient methods for the solution of nonlinear parameter estimation problems." *SIAM J. Numerical Analysis*, Vol. 7, pp. 157-186.
- * BERKHOUT, A. G. 1977. "Least-squares inverse filtering and wavelet deconvolution." *Geophysics*, Vol. 42, pp. 1369-1383.
- BICKEL, P. J., and K. A. DOKSUM. 1977. *Mathematical Statistics: Basic Ideas and Selected Topics*. San Francisco: Holden-Day, Inc.
- BIERMAN, G. J. 1973a. "A comparison of discrete linear filtering algorithms." *IEEE Trans. on Aerospace and Electronic Systems*, Vol. AES-9, pp. 28-37.
- BIERMAN, G. J. 1973b. "Fixed-interval smoothing with discrete measurements." *Int. J. Control*, Vol. 18, pp. 65-75.
- BIERMAN, G. J. 1977. *Factorization Methods for Discrete Sequential Estimation*. NY: Academic Press.
- BRYSON, A. E., JR., and M. FRAZIER. 1963. "Smoothing for linear and nonlinear dynamic systems." TDR 63-119, pp. 353-364, Aero. Sys. Div., Wright-Patterson Air Force Base, Ohio.
- BRYSON, A. E., JR., and D. E. JOHANSEN. 1965. "Linear filtering for time-varying systems using measurements containing colored noise." *IEEE Trans. on Automatic Control*, Vol. AC-10, pp. 4-10.
- BRYSON, A. E., JR., and Y. C. HO. 1969. *Applied Optimal Control*. Waltham, MA: Blaisdell.
- CHEN, C. T. 1970. *Introduction to Linear System Theory*. NY: Holt.
- CHI, C. Y. 1983. "Single-channel and multichannel deconvolution." Ph.D. dissertation, Univ. of Southern California, Los Angeles, CA.
- CHI, C. Y., and J. M. MENDEL. 1984. "Performance of minimum-variance deconvolution filter." *IEEE Trans. on Acoustics, Speech and Signal Processing*, Vol. ASSP-32, pp. 1145-1153.
- CRAMER, H. 1946. *Mathematical Methods of Statistics*. Princeton, NJ: Princeton Univ. Press.
- DAI, G-Z., and J. M. MENDEL. 1986. "General Problems of Minimum-Variance Recursive Waveshaping." *IEEE Trans. on Acoustics, Speech and Signal Processing*, Vol. ASSP-34.
- DONGARRA, J. J., J. R. BUNCH, C. B. MOLER, and G. W. STEWART. 1979. *LINPACK User's Guide*, Philadelphia: SIAM.
- DUDA, R. D., and P. E. HART. 1973. *Pattern Classification and Scene Analysis*. NY: Wiley Interscience.
- EDWARDS, A. W. F. 1972. *Likelihood*. London: Cambridge Univ. Press.
- FAURRE, P. L. 1976. "Stochastic Realization Algorithms." in *System Identification: Advances and Case Studies* (eds., R. K. Mehra and D. G. Lainiotis), pp. 1-25. NY: Academic Press.
- FERGUSON, T. S. 1967. *Mathematical Statistics: A Decision Theoretic Approach*. NY: Academic Press.
- FRASER, D. 1967. "Discussion of optimal fixed-point continuous linear smoothing (by J. S. Meditch)." *Proc. 1967 Joint Automatic Control Conf.*, p. 249, Univ. of PA, Philadelphia.

- GOLDBERGER, A. S. 1964. *Econometric Theory*. NY: John Wiley.
- GRAYBILL, F. A. 1961. *An Introduction to Linear Statistical Models*. Vol. 1, NY: McGraw-Hill.
- GUPTA, N. K., and R. K. MEHRA. 1974. "Computational aspects of maximum likelihood estimation and reduction of sensitivity function calculations." *IEEE Trans. on Automatic Control*, Vol. AC-19, pp. 774-783.
- GURA, I. A., and A. B. BIERMAN. 1971. "On computational efficiency of linear filtering algorithms." *Automatica*, Vol. 7, pp. 299-314.
- HAMMING, R. W. 1983. *Digital Filters*, 2nd Edition. Englewood Cliffs, NJ: Prentice-Hall.
- HO, Y. C. 1963. "On the stochastic approximation method and optimal filtering." *J. of Math. Anal. and Appl.*, Vol. 6, pp. 152-154.
- JAZWINSKI, A. H. 1970. *Stochastic Processes and Filtering Theory*. NY: Academic Press.
- KAILATH, T. 1968. "An innovations approach to least-squares estimation—Part 1: Linear filtering in additive white noise." *IEEE Trans. on Automatic Control*, Vol. AC-13, pp. 646-655.
- KAILATH, T. K. 1980. *Linear Systems*. Englewood Cliffs, NJ: Prentice-Hall.
- KALMAN, R. E. 1960. "A new approach to linear filtering and prediction problems." *Trans. ASME J. Basic Eng. Series D*, Vol. 82, pp. 35-46.
- KALMAN, R. E., and R. BUCY. 1961. "New results in linear filtering and prediction theory." *Trans. ASME J. Basic Eng., Series D*, Vol. 83, pp. 95-108.
- KASHYAP, R. L., and A. R. RAO. 1976. *Dynamic Stochastic Models from Empirical Data*. NY: Academic Press.
- KELLY, C. N., and B. D. O. ANDERSON. 1971. "On the stability of fixed-lag smoothing algorithms." *J. Franklin Inst.*, Vol. 291, pp. 271-281.
- KMENTA, J. 1971. *Elements of Econometrics*. NY: MacMillan.
- KOPP, R. E., and R. J. ORFORD. 1963. "Linear regression applied to system identification for adaptive control systems." *AIAA J.*, Vol. 1, pp. 2300.
- KUNG, S. Y. 1978. "A new identification and model reduction algorithm via singular value decomposition." Paper presented at the 12th Annual Asilomar Conference on Circuits, Systems, and Computers, Pacific Grove, CA.
- KWAKernaak, H., and R. SIVAN. 1972. *Linear Optimal Control Systems*. NY: Wiley-Interscience.
- LAUB, A. J. 1979. "A Schur method for solving algebraic Riccati equations." *IEEE Trans. on Automatic Control*, Vol. AC-24, pp. 913-921.
- LEHMANN, E. L. 1959. *Testing Statistical Hypotheses*. NY: John Wiley.
- LEHMANN, E. L. 1980. *Theory of Point Estimation*. NY: John Wiley.
- LJUNG, L. 1976. "Consistency of the Least-Squares Identification Method." *IEEE Trans. on Automatic Control*, Vol. AC-21, pp. 779-781.
- LJUNG, L. 1979. "Asymptotic behavior of the extended Kalman filter as a parameter estimator for linear systems." *IEEE Trans. on Automatic Control*, Vol. AC-24, pp. 36-50.

- MARQUARDT, D. W. 1963. "An algorithm for least-squares estimation of nonlinear parameters." *J. Soc. Indust. Appl. Math.*, Vol. 11, pp. 431-441.
- MCLOUGHLIN, D. B. 1980. "Distributed systems—notes." *Proc. 1980 Pre-JACC Tutorial Workshop on Maximum-Likelihood Identification*, San Francisco, CA.
- MEDITCH, J. S. 1969. *Stochastic Optimal Linear Estimation and Control*. NY: McGraw-Hill.
- MEHRA, R. K. 1970a. "An algorithm to solve matrix equations $\mathbf{PH}^T = \mathbf{G}$ and $\mathbf{P} = \Phi\mathbf{P}\Phi^T + \Gamma\Gamma^T$." *IEEE Trans. on Automatic Control*, Vol. AC-15.
- MEHRA, R. K. 1970b. "On-line identification of linear dynamic systems with applications to Kalman filtering." *Proc. Joint Automatic Control Conference*, Atlanta, GA, pp. 373-382.
- MEHRA, R. K. 1971. "Identification of stochastic linear dynamic systems using Kalman filter representation." *AIAA J.*, Vol. 9, pp. 28-31.
- MEHRA, R. K., and J. S. TYLER. 1973. "Case studies in aircraft parameter identification." *Proc. 3rd IFAC Symposium on Identification and System Parameter Estimation*, North Holland, Amsterdam.
- MENDEL, J. M. 1971. "Computational requirements for a discrete Kalman filter." *IEEE Trans. on Automatic Control*, Vol. AC-16, pp. 748-758.
- MENDEL, J. M. 1973. *Discrete Techniques of Parameter Estimation: the Equation Error Formulation*. NY: Marcel Dekker.
- MENDEL, J. M. 1975. "Multi-stage least squares parameter estimators." *IEEE Trans. on Automatic Control*, Vol. AC-20, pp. 775-782.
- MENDEL, J. M. 1981. "Minimum-variance deconvolution." *IEEE Trans. on Geoscience and Remote Sensing*, Vol. GE-19, pp. 161-171.
- MENDEL, J. M. 1983a. "Minimum-variance and maximum-likelihood recursive wave-shaping." *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. ASSP-31, pp. 599-604.
- MENDEL, J. M. 1983b. *Optimal Seismic Deconvolution: an Estimation Based Approach*. NY: Academic Press.
- MENDEL, J. M., and D. L. GIESEKING. 1971. "Bibliography on the linear-quadratic-gaussian problem." *IEEE Trans. on Automatic Control*, Vol. AC-16, pp. 847-869.
- MORRISON, N. 1969. *Introduction to Sequential Smoothing and Prediction*. NY: McGraw-Hill.
- NAH, N. E. 1969. *Estimation Theory and Applications*. NY: John Wiley.
- OPPENHEIM, A. V., and R. W. SCHAFER. 1975. *Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall.
- PAPOULIS, A. 1965. *Probability, Random Variables, and Stochastic Processes*. NY: McGraw-Hill.
- PELED, A., and B. LIU. 1976. *Digital Signal Processing: Theory, Design, and Implementation*. NY: John Wiley.
- RAUCH, H. E., F. TUNG, and C. T. STRIEBEL. 1965. "Maximum-likelihood estimates of linear dynamical systems." *AIAA J.*, Vol. 3, pp. 1445-1450.
- SCHWEPPE, F. C. 1965. "Evaluation of likelihood functions for gaussian signals." *IEEE Trans. on Information Theory*, Vol. IT-11, pp. 61-70.

- SCHWEPPE, F. C. 1973. *Uncertain Dynamic systems*. Englewood Cliffs, NJ: Prentice-Hall.
- SHANKS, J. L. 1967. "Recursion filters for digital processing." *Geophysics*, Vol. 32, pp. 32-51.
- SORENSEN, H. W. 1970. "Least-squares estimation: from Gauss to Kalman." *IEEE Spectrum*, Vol. 7, pp. 63-68.
- SORENSEN, H. W. 1980. *Parameter Estimation: Principles and Problems*. NY: Marcel Dekker.
- SORENSEN, H. W., and J. E. SACKS. 1971. "Recursive fading memory filtering." *Information Science*, Vol. 3, pp. 101-119.
- STEFANI, R. T. 1967. "Design and simulation of a high performance, digital, adaptive, normal acceleration control system using modern parameter estimation techniques." Rept. No. DAC-60637, Douglas Aircraft Co., Santa Monica, CA.
- STEPNER, D. E., and R. K. MEHRA. 1973. "Maximum likelihood identification and optimal input design for identifying aircraft stability and control derivatives." Ch. IV, NASA-CR-2200.
- STEWART, G. W. 1973. *Introduction to Matrix Computations*. NY: Academic Press.
- TREITEL, S. 1970. "Principles of digital multichannel filtering." *Geophysics*, Vol. XXXV, pp. 785-811.
- TREITEL, S., and E. A. ROBINSON. 1966. "The design of high-resolution digital filters." *IEEE Trans. on Geoscience and Electronics*, Vol. GE-4, pp. 25-38.
- TUCKER, H. G. 1962. *An Introduction to Probability and Mathematical Statistics*. NY: Academic Press.
- TUCKER, H. G. 1967. *A Graduate Course in Probability*. NY: Academic Press.
- VAN TREES, H. L. 1968. *Detection, Estimation and Modulation Theory*, Vol. 1. NY: John Wiley.
- ZACKS, S. 1971. *The Theory of Statistical Inference*. NY: John Wiley.

Index

- estimation; Maximum-likelihood estimation)
- Estimator properties (*see* Small sample properties of estimators; Large sample properties of estimators)
- Estimation of random parameters (*see* Mean-squared estimation; Maximum a posteriori estimator; Best linear unbiased estimator; Least-squares estimator)
- Estimation techniques, 3 (*see* Best linear unbiased estimator; Least-squares estimator; Maximum-likelihood estimators; Mean-squared estimators; Maximum a posteriori estimation; State estimation)
- Expanding memory estimator, 23
- Exponential family of distributions, 284-286
- Extended Kalman filter: application to parameter estimation, 255-256 correction equation, 253 derived, 249-253 iterated, 253-254 prediction equation, 252
- Factorization theorem, 283
- Filtering: computations, 156 covariance formulation, 157 divergence phenomenon, 167-169 examples, 160-169 Kalman filter derivation, 151-153 properties, 153-158 recursive filter, 153 relationship to Wiener filtering, 176-181 steady-state Kalman filter, 170-176
- Finite-difference equation coefficient identification (*see* Identification of)
- Fisher information matrix, 51
- Fixed-interval smoother, 183-184, 190, 193-199
- Fixed-lag smoother, 184, 191, 201-202
- Fixed memory estimator, 23
- Fixed-point smoother, 184, 190, 199-201
- Gauss-Markov random processes defined, 128-129
- Gaussian random processes (*see* Gauss-Markov random processes)
- Gaussian random variables (*see also* Conditional mean) conditional density function, 102-104 joint density function, 101-102 multivariate density function, 101 properties of, 104 univariate density function, 100
- Glossary of major results, 295-299
- Hypotheses: binary, 81 multiple, 85-86
- Identifiability, 95-96
- Identification of: coefficients of a finite-difference equation, 9, 65 impulse response, 8-9, 39-42, 64-65 initial condition vector in an un-
- forced state equation model, 10
- Impulse response identification (*see* Identification of)
- Information form of recursive least-squares estimator, 29, 94
- Initial condition identification (*see* Identification of)
- Innovations process: defined, 146 properties of, 146-147
- Instrument calibration example, 20-21, 35-36
- Invariance property of maximum-likelihood estimators (*see* Maximum-likelihood estimates)
- Iterated least squares, 248-249
- Kalman-Bucy filtering: derivation using a formal limiting procedure, 273-275 derivation when structure of filter is prespecified, 275-278 notation and problem statement for derivation, 271-272 optimal control application, 280-281 statement of, 272 steady-state, 278-280 system description for derivation, 271
- Kalman-Bucy gain matrix, 272
- Kalman filter (*see* Filtering; Steady-state Kalman filtering)
- Kalman filter sensitivity system, 163, 262-263
- Kalman filter tuning parameter, 167
- Kalman gain matrix, 151
- Large sample properties of estimators (*see also* Least-squares estimator): asymptotic efficiency, 60 asymptotic unbiasedness, 57 consistency, 57-60
- Least-squares estimation processing: batch, 17-24 cross-sectional, 37-38 recursive, 26-32, 35-42
- Least-squares estimator: derivation of batch form, 19-20 derivation of recursive covariance form, 30-31 derivation of recursive information form, 27-28 examples, 20-23, 35-36 initialization of recursive forms, 29 large sample properties, 68-70 multistage, 38-42 properties, 63-70 for random parameters, 121 small sample properties, 63-68 for vector measurements, 36
- Lehmann-Scheffe theorem, 291-293
- Likelihood: compared with probability, 81-82 conditional, 114 continuous distributions, 85, 88 defined, 82-84 unconditional, 114
- Likelihood ratio: defined for multiple hypotheses, 85-86 defined for two hypotheses, 84-86
- Linear model: defined, 7 examples of, 8-14 (*see also* Identification of; State estimation example; Deconvolution; Non-linear measurement model)

Linear projection, 116–117
 Log-likelihood:
 for the basic state-variable model, 259–261
 defined, 88–89
 for an important dynamical system, 94–97
 Markov process, 129–130
 Matrix inversion lemma, 30
 Matrix Riccati differential equation, 272
 Matrix Riccati equation, 155
 Maximum a posteriori estimation:
 comparison with best linear unbiased estimator, 123–124
 comparison with mean-squared estimator, 115–116
 Gaussian linear model, 123–124
 general case, 114–116
 Maximum-likelihood deconvolution (*see Deconvolution*)
 Maximum-likelihood estimation, 88–97
 Maximum-likelihood estimators:
 comparison with best linear unbiased estimator, 93–94
 comparison with least-squares estimator, 94
 for exponential families, 287–290
 the linear model, 92–94
 obtaining them, 89–91
 properties, 91–92
 Maximum-likelihood method, 89
 Maximum-likelihood state and parameter estimation:
 computing θ_{ML} , 261
 log-likelihood function for the basic state-variable model, 259–261
 steady-state approximation, 264–268
 Mean-squared convergence (*see Convergence in mean-square*)
 Mean-squared estimation, 109–113
 Mean-squared estimators:
 comparison with best linear unbiased estimator, 122–123
 comparison with maximum a posteriori estimator, 115–116
 derivation, 110
 Gaussian case, 111–113
 for linear and Gaussian model, 118–120
 properties of, 112–113
 Measurement differencing technique, 231
 Minimum-variance deconvolution (*see Deconvolution*)
 MLD (*see Deconvolution*)
 Modeling:
 estimation problem, 1
 measurement problem, 1
 representation problem, 1
 validation problem, 2
 Multistage least-squares (*see Least-squares estimator*)
 Multivariate Gaussian random variables (*see Gaussian random variables*)
 MVD (*see Deconvolution*)

Nonlinear dynamical systems:
 discretized perturbation state-variable model, 245
 linear perturbation equations, 239–242
 model, 237–239
 Nonlinear measurement model, 12
 Not-so-basic state-variable model:
 biases, 224
 colored noises, 227–230
 correlated noises, 225–227
 perfect measurements, 230–233
 Notation, 14–15
 Orthogonality principle, 111–112
 Parameter estimation (*see Extended Kalman filter*)
 Perfect measurements (*see Not-so-basic state-variable model*)
 Perturbation equations (*see Non-linear dynamical systems*)
 Philosophy:
 estimation theory, 6
 modeling, 5–6
 Prediction:
 general, 142–145
 recursive predictor, 155
 single-stage, 140–142
 steady-state predictor, 173
 Predictor-corrector form of Kalman filter, 151–152
 Properties of best linear unbiased estimators (*see Best linear unbiased estimator*)
 Properties of estimators (*see Small sample properties of estimators; Large sample properties of estimators*)
 Properties of least-squares estimator (*see Least-squares estimator*)
 Properties of maximum-likelihood estimators (*see Maximum-likelihood estimators*)
 Properties of mean-squared estimators (*see Mean-squared estimators*)
 Random processes (*see Gauss-Markov random processes*)
 Random variables (*see Gaussian random variables*)
 Recursive calculation of state covariance matrix, 134
 Recursive calculation of state mean vector, 134
 Recursive processing (*see Least-squares estimation processing; Best linear unbiased estimator*)
 Recursive waveshaping, 216–222
 Reduced-order Kalman filter, 231
 Reduced-order state estimator, 231
 Riccati equation (*see Matrix Riccati equation; Algebraic Riccati equation*)
 Sample mean as a recursive digital filter, 6
 Scale changes:
 best linear unbiased estimator, 77–78
 Zoom-in algorithm, 39–42

least squares, 23–24
 Sensitivity of Kalman filter, 163–166
 Signal-to-noise ratio, 137–138, 167
 Single-channel steady-state Kalman filter, 173–176
 Small sample properties of estimators (*see also Least-squares estimator*):
 efficiency, 46–52
 unbiasedness, 44–46
 Smoothing:
 applications, 204–222
 double-stage, 187–189
 fixed-interval, 190, 193–199
 fixed-lag, 191, 201–202
 fixed-point, 190, 199–201
 single-stage, 185–187
 three types, 183–184
 Stabilized form for computing $P(k+1|k+1)$, 156
 Standard form for computing $P(k+1|k+1)$, 156
 State estimation (*see Prediction; Filtering; Smoothing*)
 State estimation example, 10–12, 126
 State-variable model (*see Basic state-variable model; Not-so-basic state-variable model*)
 Steady-state approximation (*see Maximum-likelihood state and parameter estimation*)
 Steady-state filter system, 173
 Steady-state Kalman filter, 170–172
 Steady-state MVD filter:
 defined, 207
 properties of, 212–213
 relationship to IIR Wiener deconvolution filter, 213–215
 Steady-state predictor system, 173
 Stochastic linear optimal output feedback regulator problem, 281
 Sufficient statistics:
 defined, 282–284
 for exponential families of distributions, 285–286, 287–290
 and uniformly minimum-variance unbiased estimation, 290–294
 Unbiasedness (*see Small sample properties of estimators*)
 Uniformly minimum-variance unbiased estimation (*see Sufficient statistics*)
 Variance estimator, 67–68
 Waveshaping (*see Recursive waveshaping*)
 Weighted least-squares estimator (*see Least-squares estimator; Best linear unbiased estimator*)
 White noise discrete, 130–131
 Wiener filter:
 derivation, 178–179
 relation to Kalman filter, 180–181
 Wiener-Hopf equations, 179