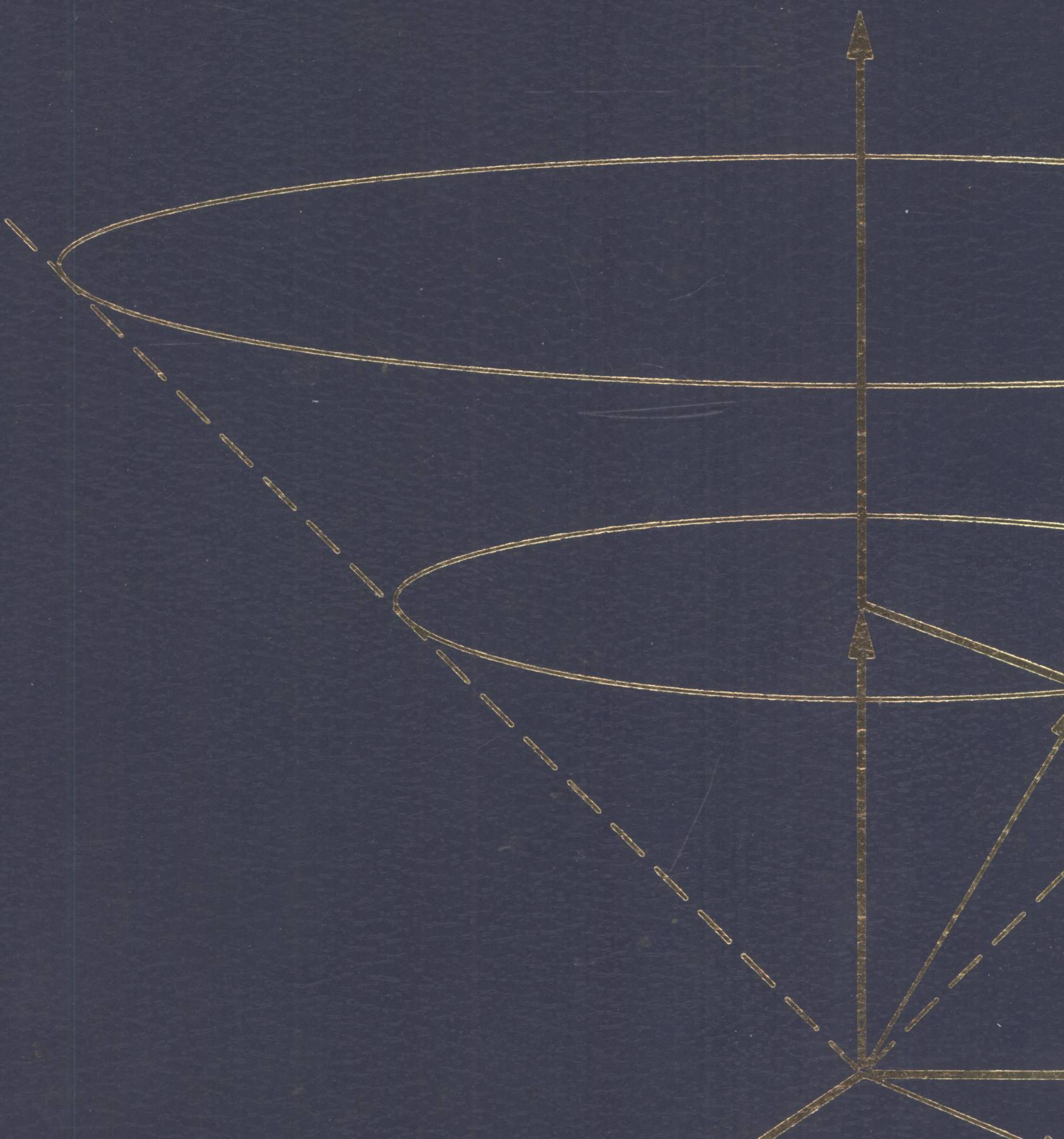
Statistical Signal Processing

Detection, Estimation, and Time Series Analysis

Louis L. Scharf



Statistical Signal Processing

Detection, Estimation, and Time Series Analysis

Louis L. Scharf
University of Colorado at Boulder

with Cédric Demeure collaborating on Chapters 10 and 11

ADDISON-WESLEY PUBLISHING COMPANY
Reading, Massachusetts • Menlo Park, California • New York
Don Mills, Ontario • Wokingham, England • Amsterdam
Bonn • Sydney • Singapore • Tokyo • Madrid • San Juan

This book is in the Addison-Wesley Series in Electrical and Computer Engineering: Digital Signal Processing

Consulting Editor: Richard Roberts

Library of Congress Cataloging-in-Publication Data

Scharf, Louis L.
Statistical signal processing : detection, estimation, and time series analysis / Louis L. Scharf.
p. cm.
Includes bibliographical references and index.
ISBN 0-201-19038-9
1. Signal processing—Statistical methods. I. Title.
TK5102.5.S3528 1990
621.382'2—dc20 90-42747
CIP

Reprinted with corrections July, 1991

2 3 4 5 6 7-MA-939291

To Carol, Greg, and Heidi

Preface

The field of statistical signal processing embraces the many mathematical procedures that engineers and statisticians use to draw inference from imperfect or incomplete measurements. The major domains of the field are detection, estimation, and time series analysis.

Abstractly, statistical signal processing is a theory for using experimental measurements to transform a prior model for a signal into a posterior model that may be used to make informed decisions. The quality of the decision is measured by a loss function that depends on ground truth and the decision taken. It is the intricate interplay between prior models, measurement schemes, loss functions, and decision rules that gives statistical signal processing its great variety.

ORGANIZATION AND USE

This book is my personal statement about the fundamental ideas in statistical signal processing. The book breaks down along four distinct topical lines: mathematical and statistical preliminaries; detection theory; estimation theory; and time series analysis. There is enough material to support a two-semester course in statistical signal processing, but the book may be used for separate one-semester courses in detection theory, estimation theory, or time series analysis. In a detection theory course, Chapters 1 through 5 may be covered in their entirety. In an estimation theory course, Chapters 1 through 3 and 6 through 8 may be covered. In a time series course, Chapters 1 through 3 and 9 through 11 are appropriate. Chapter 9 on least squares is a swing chapter that may be treated as a topic in estimation theory or time series analysis.

A GUIDED TOUR OF THE BOOK

Mathematical and Statistical Preliminaries

I begin in Chapter 2 with a fairly comprehensive review of linear algebra, matrix theory, and multivariate normal theory. Linear algebra, and the geometrical pictures that bring life to it, forms the basis for our prior structural information about a signal. Matrix theory provides an algebra for manipulating and composing linear transformations, and multivariate normal theory provides the statistical methodology for computing the distribution of linear and quadratic forms in normal random vectors. When teaching from this chapter, I pick and choose from the topics, making sure to develop the ideas of linear independence, subspaces and their spans, orthogonal subspaces, QR factorizations, projections, the singular value decomposition, the multivariate normal distribution, and the distribution of quadratic forms in projection operators. In Chapter 3, I develop the main results in the theory of sufficient statistics and show the fundamental role they play in the computation of minimum variance unbiased estimators.

Detection Theory

Chapters 4 and 5 are dedicated to detection theory. In Chapter 4 I treat the many faces of the Neyman-Pearson theory of hypothesis testing. I cover the rudiments of decision theory, discuss the roles of sufficiency and invariance in hypothesis testing, and develop the theory of uniformly most powerful tests. I then apply the principles of sufficiency and invariance to signal detection when the signal model, or the noise model, is incompletely known. This produces matched filters, CFAR matched filters, matched subspace filters, and CFAR matched subspace filters. The final two sections of Chapter 4 treat reduced-rank detectors and linear discriminant functions for detecting Gaussian signals in Gaussian noise. Chapter 4 is long, so some instructors may wish to omit the linear discriminant functions and, perhaps, the sections on matched subspace filtering, although the latter is very important in the modern study of detection theory. In Chapter 5, I treat the Bayesian theory of hypothesis testing, wherein a prior distribution is assigned to the hypotheses under test. Minimax tests are constructed as Bayes tests against least favorable priors. The study of M-orthogonal symbolling produces insights into channel capacity, and the study of composite matched filters produces insights into associative memories.

Estimation Theory

Chapters 6 through 9 are devoted to estimation theory. I begin in Chapter 6 with the maximum likelihood theory of parameter estimation, where I discuss the roles that sufficiency and invariance play in the maximum likelihood theory and discuss the Cramer-Rao bound for the variance of unbiased estimators. Nuisance parameters are explored in depth. In the last several sections of the chapter, I apply maximum likelihood theory to the identification of subspaces, to the identification of ARMA parameters, and to the identification of structured covariance matrices. Chapter 6 is long, so some instructors may want to give nuisance parameters a once-over-lightly, and select just a few of the applications. In Chapter 7 parameters are endowed with prior distributions, and the Bayes theory is developed for turning prior distributions into posterior distributions. The Bayes theory produces the Gauss-Markov theorem for multivariate normal parameters and measurements. I interpret the Gauss-Markov theorem by showing how it transforms a channel model for measurements into an inverse channel model, or estimator-plus-noise model. The Gauss-Markov theorem is applied to sequential Bayes estimators, the Kalman filter, and the Wiener filter. In Chapter 8 we explore in more detail minimum mean-squared error estimators and the role that the conditional mean estimator plays. I derive conditional mean estimators to solve a number of problems in signal processing: low-rank Wiener filters, linear predictors, Kalman filters, low-rank approximations of random vectors, scalar and block quantizers, and reduced-rank block quantizers. The last application produces rate-distortion formulas of the Shannon variety. In Chapter 9 we develop the theory of least squares for estimating parameters in the linear statistical model and stress the singular value decomposition for the insight that it brings to least squares problems. We then study the performance of least squares when errors are normally distributed. This study produces order selection rules for reducing the rank of least squares estimators. The middle sections of the chapter are devoted to special topics such as sequential, weighted, constrained, and total least squares. The last several sections are devoted to applications: inverse problems, mode identification, parameter estimation in ARMA systems, linear prediction, and the identification of structured covariance matrices. Chapter 9 is long, so some instructors may want to select just a few of the applications.

Time Series Analysis

In Chapters 10 and 11 we cover linear prediction and modal analysis. In our treatment of linear prediction we begin with the classical stationary theory of Wold and Kolmogorov and establish the connection between linear prediction and maximum entropy extension of correlation sequences. We then develop the nonstationary theory of fitting order-increasing whiteners to finite covariance matrices, paying special attention to the Levinson and Schur recursions for computing the reflection coefficients that keep the recursions going. We study the least squares theory of linear prediction and derive the lattice recursions for computing reflection coefficients. Linear prediction in ARMA time series produces the MSK algorithms for fast Kalman filtering. In the last two sections of the chapter we apply linear prediction to the computation of likelihood and the design of a differential PCM system. In Chapter 11 we draw a distinction between linear prediction and modal analysis. We study Prony's method, exact least squares, the total least squares of Golub and Van Loan, the principal components of Tufts and Kumaresan, and MUSIC of Bienvenu and Schmidt as the most prominent techniques for estimating modes. We outline pencil methods and then present Kumaresan's procedure for estimating modes from frequency domain data.

ACKNOWLEDGMENTS

Since the early 70s I have had the opportunity to talk with Henry Cox and Ben Friedlander about many problems in signal processing. Both have generously volunteered interesting problems to explore, shared their own elegant solutions, and offered insight-

ful interpretations of others' solutions. In 1980 Claude Gueguen spent six months at Colorado State University, where he taught his course on Parametric Signal Processing. In this course, and in our collaborative research, I gained my first appreciation for the power of linear algebraic models as structural models for signals. I want to thank Claude for sharing with me his ideas and his deep understanding of system theory and signal processing. Since 1985 I have resided in the office adjacent to Tom Mullis. This has provided me with the opportunity to follow Tom's courses in linear algebra and spectrum analysis, to collaborate with him on several pieces of writing, and to attend his research seminars. I thank him for sharing his gift for clear thinking with me and for getting me out of a few tight spots in the final stages of writing this manuscript.

I first conceived this book in 1984, while teaching a graduate course in signal processing at the University of Rhode Island. During my stay at the University of Rhode Island I profited immensely from my association with Don Tufts. His insights profoundly influenced my own thinking, and I would like to acknowledge my intellectual debt to him. Ramdas Kumaresan improved my understanding of modal analysis, and Steve Kay directed me to a deeper understanding of signal detection in the linear statistical model. Many other people at the University of Rhode Island made my stay there a happy time of intellectual growth and professional development. I would like to acknowledge their collegiality.

Dick Roberts, late Professor of Electrical and Computer Engineering at the University of Colorado, reviewed my writing of several sections of the book. He encouraged me to complete it and recommended that I use his editor, Tom Robbins at Addison-Wesley. I followed Dick's advice, and as a consequence I have had the pleasure of working with Tom for the past 18 months, as I put the final touches on the manuscript. During this period, Lynn Kirlin used the manuscript at the University of Victoria and offered many improvements to the presentation.

For the past 20 years or so I have been able to maintain a single-minded interest in statistical signal processing because four program directors at the Office of Naval Research have supported our work. I thank Drs. Bruce McDonald, Doug Depriest, Ed Wegman, and Neil Gerr for their consistent management of ONR's programs in signal processing and mathematical statistics.

I thank my many students for their contributions to my understanding of signal processing. Cédric Demeure helped me write Chapters 10 and 11, so much so that they are as much his as mine. Finally, let me express my sincere thanks to Julie Fredlund, secretary to the group in digital signal processing at the University of Colorado, for her masterful and patient preparation of a demanding manuscript.

Louis Scharf
Boulder, CO

Contents

CHAPTER **5**

Bayes Detectors 179

CHAPTER 8

Minimum Mean-Squared Error Estimators 323

CHAPTER 9

Least Squares 359