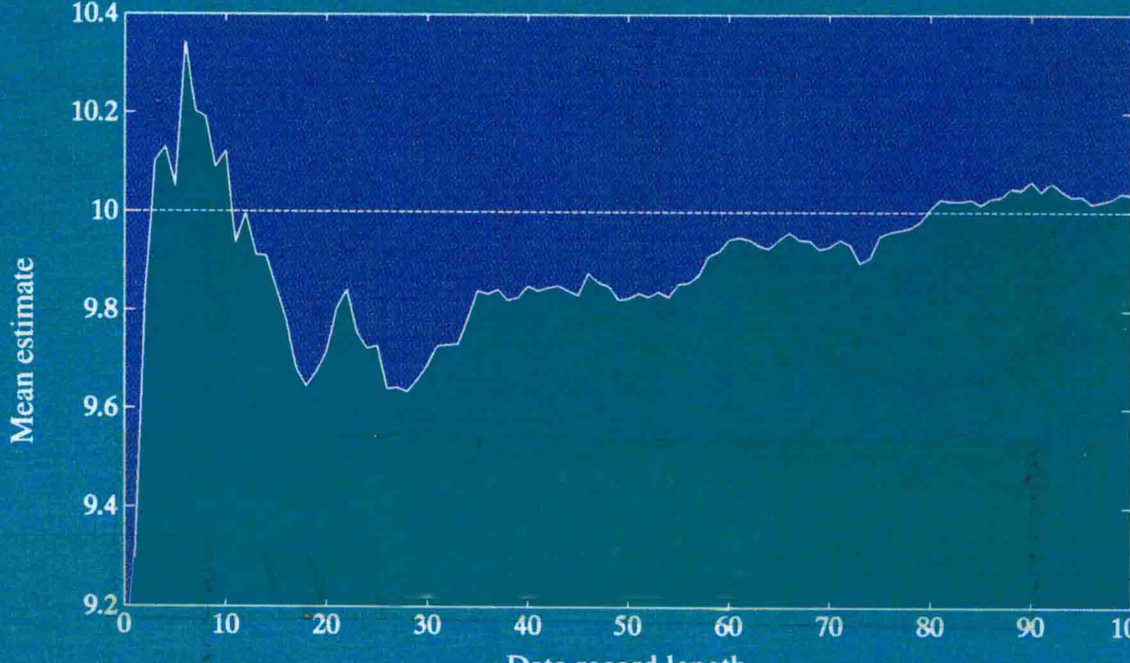
VOLUME I

FUNDAMENTALS OF STATISTICAL SIGNAL PROCESSING

ESTIMATION THEORY



STEVEN M. KAY

PRENTICE HALL SIGNAL PROCESSING SERIES
ALAN V. OPPENHEIM, SERIES EDITOR

Fundamentals of Statistical Signal Processing: Estimation Theory

Steven M. Kay
University of Rhode Island

Prentice Hall PTR
Upper Saddle River, New Jersey 07458

Library of Congress Cataloging-in-Publication Data

Kay, Steven M.
Fundamentals of statistical signal processing : estimation theory / Steven M. Kay.
p. cm. — (PH signal processing series)
Includes bibliographical references and index.
ISBN 0-13-345711-7
1. Signal processing—Statistical methods. 2. Estimation theory.
I. Title. II. Series: Prentice-Hall signal processing series.
TK5102.5.K379 1993
621.382'2—dc20 92-29495
CIP

Acquisitions Editor: Karen Gettman
Editorial Assistant: Barbara Alfieri
Prepress and Manufacturing Buyer: Mary E. McCartney
Cover Design: Wanda Lubelska
Cover Design Director: Eloise Starkweather

Prentice-Hall, Inc.
A Pearson Education Company
Upper Saddle River, New Jersey 07458

ISBN 0-13-345711-7

Text printed in the United States at Hamilton in Castleton, New York.
Nineteenth printing, August 2011

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*
Editora Prentice-Hall do Brasil, Ltda., *Rio de Janeiro*

Fundamentals of Statistical Signal Processing: Estimation Theory

PRENTICE HALL SIGNAL PROCESSING SERIES

Alan V. Oppenheim, *Series Editor*

BRIGHAM *The Fast Fourier Transform and Its Applications (AOD)*

BUCK, DANIEL & SINGER *Computer Explorations in Signals and Systems Using MATLAB*

CASTLEMAN *Digital Image Processing*

COHEN *Time-Frequency Analysis*

CROCHIERE & RABINER *Multirate Digital Signal Processing (AOD)*

JOHNSON & DUDGEON *Array Signal Processing (AOD)*

KAY *Fundamentals of Statistical Signal Processing, Vols. I & II*

KAY *Modern Spectral Estimation (AOD)*

LIM *Two-Dimensional Signal and Image Processing*

MCCLELLAN, BURRUS, OPPENHEIM, PARKS, SCHAFER & SCHUESSLER *Computer-Based Exercises for Signal Processing Using MATLAB Ver. 5*

MENDEL *Lessons in Estimation Theory for Signal Processing, Communications, and Control, 2/e*

NIKIAS & PETROPULU *Higher Order Spectra Analysis*

OPPENHEIM & SCHAFER *Digital Signal Processing*

OPPENHEIM & SCHAFER *Discrete-Time Signal Processing, 2/e*

OPPENHEIM & WILLSKY, WITH NAWAB *Signals and Systems, 2/e*

ORFANIDIS *Introduction to Signal Processing*

PHILLIPS & NAGLE *Digital Control Systems Analysis and Design, 3/e*

QUATIERI *Discrete-Time Speech Signal Processing: Principles and Practice*

RABINER & JUANG *Fundamentals of Speech Recognition*

RABINER & SCHAFER *Digital Processing of Speech Signals*

TEKALP *Digital Video Processing*

VAIDYANATHAN *Multirate Systems and Filter Banks*

WANG, OSTERMANN & ZHANG *Video Processing and Communications*

WIDROW & STEARNS *Adaptive Signal Processing*

In memory of Jean Adler

and to my parents
Phyllis and Jack

and to my family
Cindy, Lisa, and Ashley

Preface

This text is the first volume of a series of books addressing statistical signal processing. It describes the application of statistical parameter estimation to extraction of information from received signals in noise. The second volume, entitled *Fundamentals of Statistical Signal Processing: Detection Theory* (Prentice Hall PTR, 1998, ISBN: 0-13-504135-X), is the application of statistical hypothesis testing to the detection of signals in noise. The series has been written to provide the reader with a broad introduction to the theory and application of statistical signal processing.

Parameter estimation is a subject that is standard fare in the many books available on statistics. These books range from the highly theoretical expositions written by statisticians to the more practical treatments contributed by the many users of applied statistics. This text is an attempt to strike a balance between these two extremes. The particular audience we have in mind is the community involved in the design and implementation of signal processing algorithms. As such, the primary focus is on obtaining optimal estimation algorithms that may be implemented on a digital computer. The data sets are therefore assumed to be samples of a continuous-time waveform or a sequence of data points. The choice of topics reflects what we believe to be the important approaches to obtaining an optimal estimator and analyzing its performance. As a consequence, some of the deeper theoretical issues have been omitted with references given instead.

It is the author's opinion that the best way to assimilate the material on parameter estimation is by exposure to and working with good examples. Consequently, there are numerous examples that illustrate the theory and others that apply the theory to actual signal processing problems of current interest. Additionally, an abundance of homework problems have been included. They range from simple applications of the theory to extensions of the basic concepts. A solutions manual is available from the publisher. To aid the reader, summary sections have been provided at the beginning of each chapter. Also, an overview of all the principal estimation approaches and the rationale for choosing a particular estimator can be found in Chapter 14. Classical estimation is first discussed in Chapters 2–9, followed by Bayesian estimation in Chapters 10–13. This delineation will, hopefully, help to clarify the basic differences between these two principal approaches. Finally, again in the interest of clarity, we present the estimation principles for scalar parameters first, followed by their vector extensions. This is because the matrix algebra required for the vector estimators can sometimes obscure the main concepts.

This book is an outgrowth of a one-semester graduate level course on estimation theory given at the University of Rhode Island. It includes somewhat more material than can actually be covered in one semester. We typically cover most of Chapters 1–12, leaving the subjects of Kalman filtering and complex data/parameter extensions to the student. The necessary background that has been assumed is an exposure to the basic theory of digital signal processing, probability and random processes, and linear and matrix algebra. This book can also be used for self-study and so should be useful to the practicing engineer as well as the student.

The author would like to acknowledge the contributions of the many people who over the years have provided stimulating discussions of research problems, opportunities to apply the results of that research, and support for conducting research. Thanks are due to my colleagues L. Jackson, R. Kumaresan, L. Pakula, and D. Tufts of the University of Rhode Island, and L. Scharf of the University of Colorado. Exposure to practical problems, leading to new research directions, has been provided by H. Woodsum of Sonetech, Bedford, New Hampshire, and by D. Mook, S. Lang, C. Myers, and D. Morgan of Lockheed-Sanders, Nashua, New Hampshire. The opportunity to apply estimation theory to sonar and the research support of J. Kelly of the Naval Undersea Warfare Center, Newport, Rhode Island, J. Salisbury of Analysis and Technology, Middletown, Rhode Island (formerly of the Naval Undersea Warfare Center), and D. Sheldon of the Naval Undersea Warfare Center, New London, Connecticut, are also greatly appreciated. Thanks are due to J. Sjogren of the Air Force Office of Scientific Research, whose continued support has allowed the author to investigate the field of statistical estimation. A debt of gratitude is owed to all my current and former graduate students. They have contributed to the final manuscript through many hours of pedagogical and research discussions as well as by their specific comments and questions. In particular, P. Djurić of the State University of New York proofread much of the manuscript, and V. Nagesha of the University of Rhode Island proofread the manuscript and helped with the problem solutions.

Steven M. Kay
University of Rhode Island
Kingston, RI 02881
kay@ele.uri.edu

Contents

Chapter 1

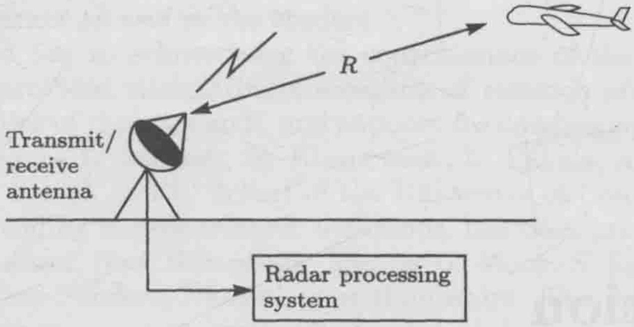
Introduction

1.1 Estimation in Signal Processing

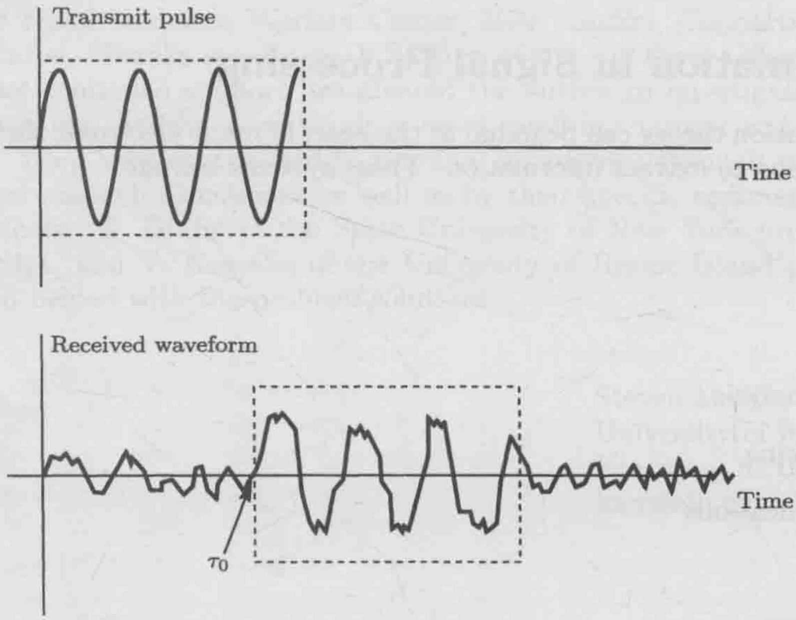
Modern estimation theory can be found at the heart of many electronic signal processing systems designed to extract information. These systems include

1. Radar
2. Sonar
3. Speech
4. Image analysis
5. Biomedicine
6. Communications
7. Control
8. Seismology,

and all share the common problem of needing to estimate the values of a group of parameters. We briefly describe the first three of these systems. In radar we are interested in determining the position of an aircraft, as for example, in airport surveillance radar [Skolnik 1980]. To determine the range R we transmit an electromagnetic pulse that is reflected by the aircraft, causing an echo to be received by the antenna τ_0 seconds later, as shown in Figure 1.1a. The range is determined by the equation $\tau_0 = 2R/c$, where c is the speed of electromagnetic propagation. Clearly, if the round trip delay τ_0 can be measured, then so can the range. A typical transmit pulse and received waveform are shown in Figure 1.1b. The received echo is decreased in amplitude due to propagation losses and hence may be obscured by environmental noise. Its onset may also be perturbed by time delays introduced by the electronics of the receiver. Determination of the round trip delay can therefore require more than just a means of detecting a jump in the power level at the receiver. It is important to note that a typical modern



(a) Radar

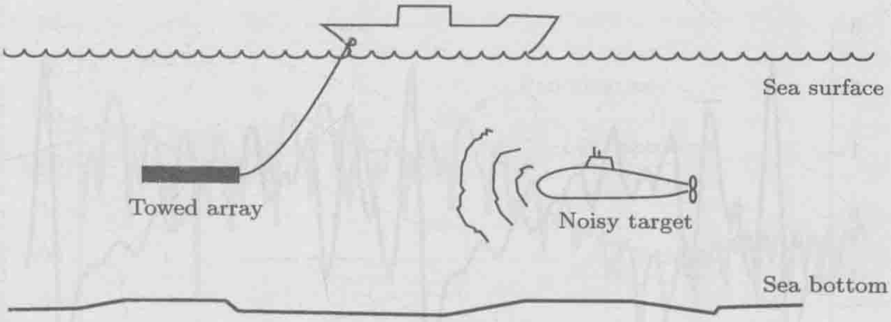


(b) Transmit and received waveforms

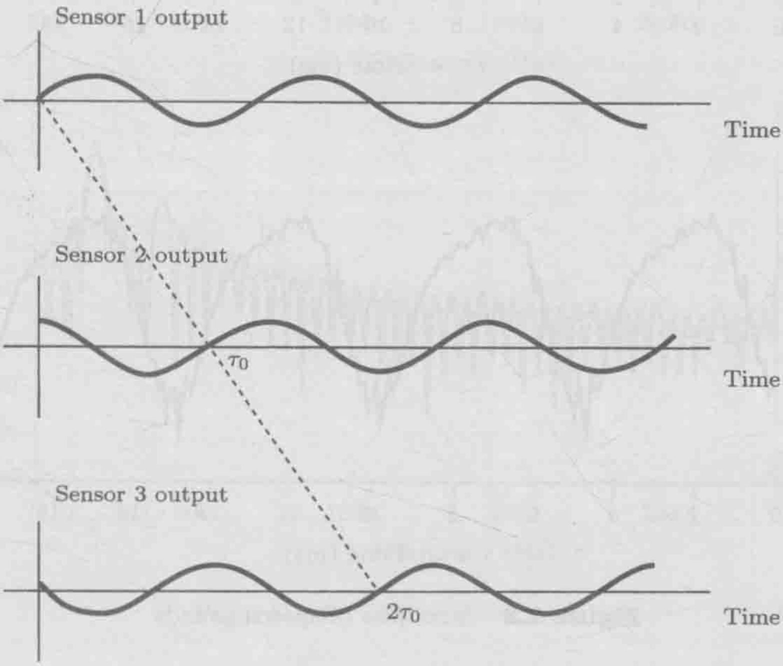
Figure 1.1 Radar system

radar system will input the received continuous-time waveform into a digital computer by taking samples via an analog-to-digital convertor. Once the waveform has been sampled, the data compose a *time series*. (See also Examples 3.13 and 7.15 for a more detailed description of this problem and optimal estimation procedures.)

Another common application is in sonar, in which we are also interested in the position of a target, such as a submarine [Knight et al. 1981, Burdic 1984] . A typical passive sonar is shown in Figure 1.2a. The target radiates noise due to machinery on board, propellor action, etc. This noise, which is actually the *signal* of interest, propagates through the water and is received by an array of sensors. The sensor outputs



(a) Passive sonar



(b) Received signals at array sensors

Figure 1.2 Passive sonar system

are then transmitted to a tow ship for input to a digital computer. Because of the positions of the sensors relative to the arrival angle of the target signal, we receive the signals shown in Figure 1.2b. By measuring τ_0, the delay between sensors, we can determine the bearing β from the expression

$$\beta = \arccos\left(\frac{c\tau_0}{d}\right) \tag{1.1}$$

where c is the speed of sound in water and d is the distance between sensors (see Examples 3.15 and 7.17 for a more detailed description). Again, however, the received

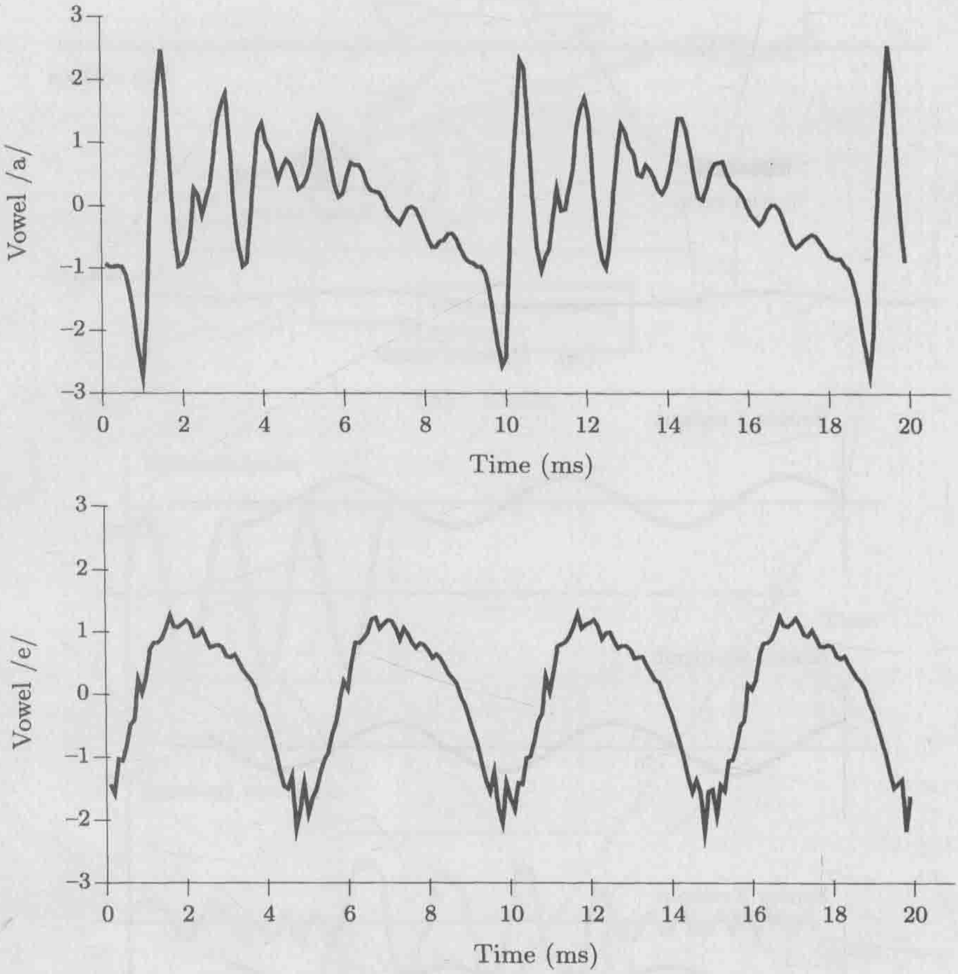


Figure 1.3 Examples of speech sounds

waveforms are not "clean" as shown in Figure 1.2b but are embedded in noise, making the determination of τ_0 more difficult. The value of β obtained from (1.1) is then only an estimate.

Another application is in speech processing systems [Rabiner and Schafer 1978]. A particularly important problem is speech recognition, which is the recognition of speech by a machine (digital computer). The simplest example of this is in recognizing individual speech sounds or *phonemes*. Phonemes are the vowels, consonants, etc., or the fundamental sounds of speech. As an example, the vowels /a/ and /e/ are shown in Figure 1.3. Note that they are periodic waveforms whose period is called the *pitch*. To recognize whether a sound is an /a/ or an /e/ the following simple strategy might be employed. Have the person whose speech is to be recognized say each vowel three times and store the waveforms. To recognize the spoken vowel, compare it to the stored vowels and choose the one that is closest to the spoken vowel or the one that

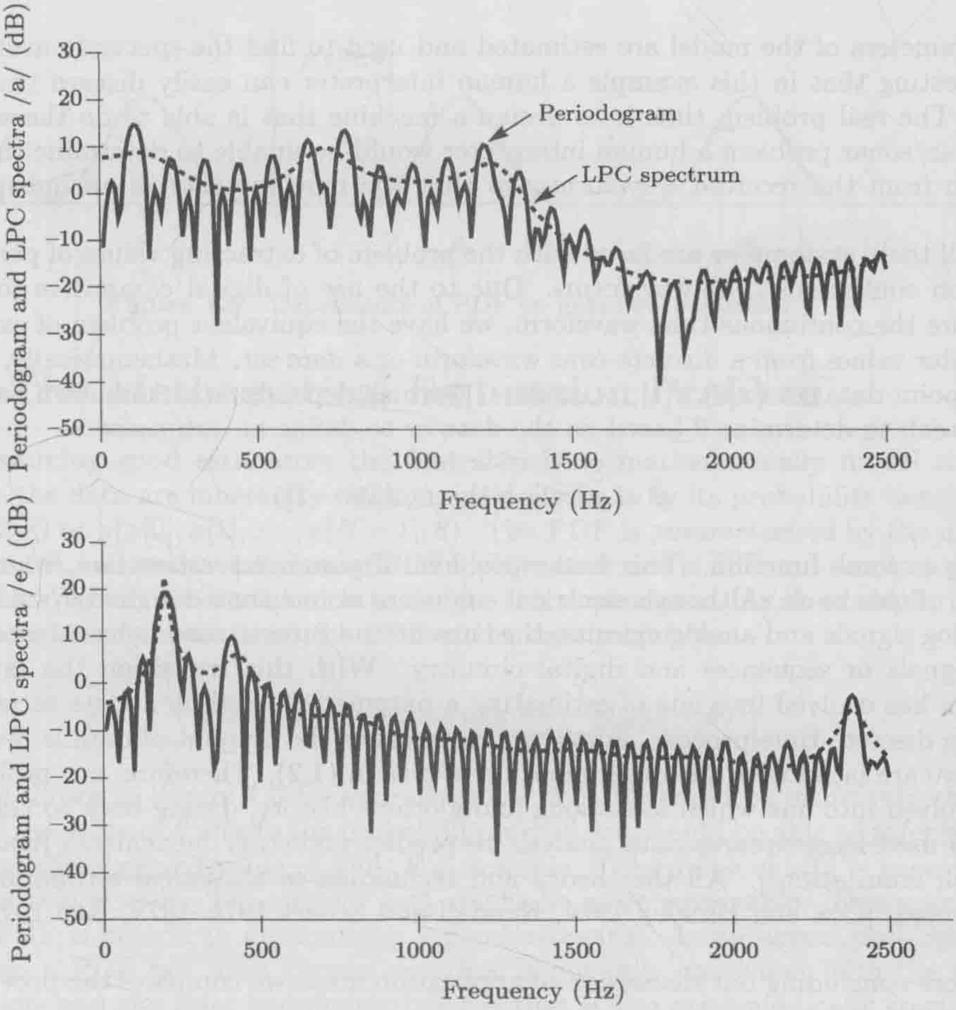


Figure 1.4 LPC spectral modeling

minimizes some distance measure. Difficulties arise if the pitch of the speaker's voice changes from the time he or she records the sounds (the training session) to the time when the speech recognizer is used. This is a natural variability due to the nature of human speech. In practice, attributes, other than the waveforms themselves, are used to measure distance. Attributes are chosen that are less susceptible to variation. For example, the spectral envelope will not change with pitch since the Fourier transform of a periodic signal is a sampled version of the Fourier transform of one period of the signal. The period affects only the spacing between frequency samples, not the values. To extract the spectral envelope we employ a model of speech called *linear predictive coding* (LPC). The parameters of the model determine the spectral envelope. For the speech sounds in Figure 1.3 the power spectrum (magnitude-squared Fourier transform divided by the number of time samples) or periodogram and the estimated LPC spectral envelope are shown in Figure 1.4. (See Examples 3.16 and 7.18 for a description of how