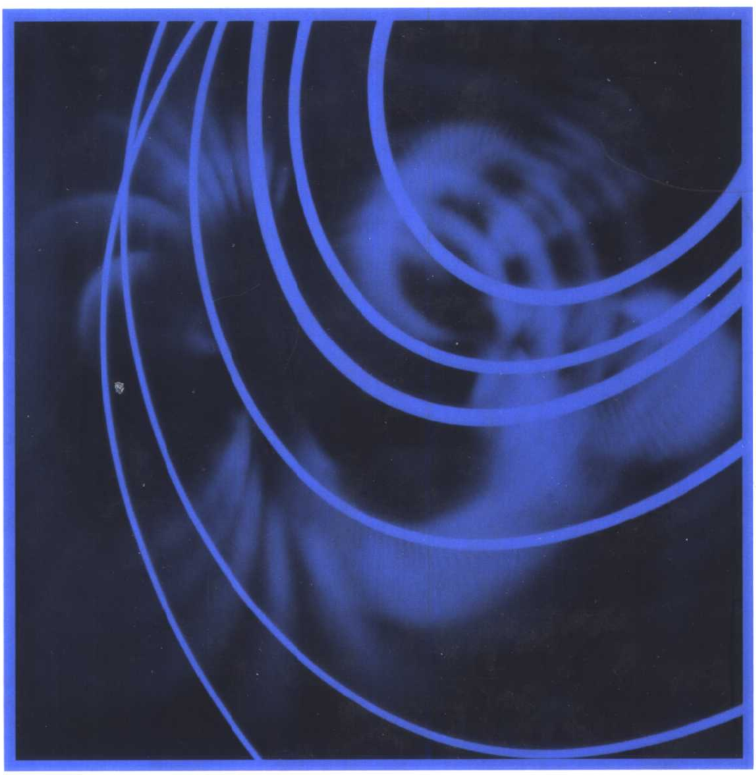
Algorithms for Statistical Signal Processing



John G. Proakis • Charles M. Rader
Fuyun Ling • Chrysostomos L. Nikias
Marc Moonen • Ian K. Proudler

ALGORITHMS FOR STATISTICAL SIGNAL PROCESSING

John G. Proakis

Charles M. Rader

Fuyun Ling

Chrysostomos L. Nikias

Marc Moonen

Ian K. Proudler

Prentice Hall

Upper Saddle River, New Jersey 07458

Library of Congress Cataloging-in-Publication Data

Proakis, John G.
Algorithms for statistical signal processing / John G. Proakis. . . [et al.]
p. cm.
Includes bibliographical references and index.
1. Signal Processing—Digital techniques—Mathematics. 2. Algorithms. I. Proakis, John G.
ISBN 0-13-062219-2

TK5102.9.A43 2001
621.382′2—dc21 2001036844

Vice President and Editorial Director, ECS: *Marcia J. Horton*
Publisher: *Tom Robbins*
Editorial Assistant: *Jessica Romeo*
Vice President and Director of Production and Manufacturing, ESM: *David W. Riccardi*
Executive Managing Editor: *Vince O'Brien*
Managing Editor: *David A. George*
Production Editor: *Chanda Wakefield*
Director of Creative Services: *Paul Belfanti*
Creative Director: *Carole Anson*
Art Director: *Jayne Conte*
Art Editor: *Greg Dulles*
Manufacturing Manager: *Trudy Pisciotti*
Manufacturing Buyer: *Lynda Castillo*
Marketing Manager: *Holly Stark*
Marketing Assistant: *Karen Moon*

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

ISBN 0-13-062219-2

Pearson Education Ltd., *London*
Pearson Education Australia Pty. Ltd., *Sydney*
Pearson Education Singapore, Pte. Ltd.
Pearson Education North Asia Ltd., *Hong Kong*
Pearson Education Canada, Inc., *Toronto*
Pearson Educacíon de Mexico, S.A. de C.V.
Pearson Education—Japan, *Tokyo*
Pearson Education Malaysia, Pte. Ltd.

Preface

The field of digital signal processing (DSP) has expanded rapidly over the past three decades. During the late sixties and seventies, we witnessed the development of the basic theory for digital filter design and the development of computationally efficient algorithms for evaluating the Fourier transform, convolution, and correlation. During the past two decades, we experienced an explosion in DSP applications spurred by significant advances in digital computer technology and integrated-circuit fabrication. In this period, the basic DSP theory has expanded to include parametric signal modeling, with applications to power spectrum estimation and system modeling, adaptive signal processing algorithms, multirate and multidimensional signal processing, and higher-order statistical methods for signal processing.

With the expansion of basic DSP theory and the rapid growth in applications (spurred by the development of fast and inexpensive digital signal processors), there is a growing interest in advanced courses in DSP covering a variety of topics. This book was written with the goal of satisfying, in part, the resulting need for textbooks covering these advanced topics.

Most of the material contained in this book was first published in 1992 by the Macmillan Publishing Company, in a book entitled *Advanced Digital Signal Processing* (which went out of print in 1997). This new book differs from the earlier publication by the inclusion of a new chapter (Chapter 7) on QRD-based fast adaptive filter algorithms, and the deletion of a chapter on multirate signal processing. The other chapters have remained essentially the same.

The major focus of this book is on algorithms for statistical signal processing. Chapter 2 treats computationally efficient algorithms for convolution and for the computation of the discrete Fourier transform. Chapter 3 treats linear prediction and optimum

Wiener filters; included in this chapter is a description of the Levinson-Durbin and Schur algorithms. Chapter 4 considers the filter design problem based on the least-squares method and describes several methods for solving least squares problems, including the Givens transformation, the Householder transformation, and singular-value decomposition. Chapter 5 treats single-channel adaptive filters based on the LMS algorithm and on recursive least-squares algorithms. Chapter 6 describes computationally efficient recursive least-squares algorithms for multichannel signals. Chapter 7 is focused on the uses of signal flow graphs for deriving computationally efficient adaptive filter algorithms based on the QR decomposition. Chapter 8 deals with power spectrum estimation, including both parametric and nonparametric methods. Chapter 9 describes the use of higher-order statistical methods for signal modeling and system identification.

Although the material in this book was written by six different authors, we have tried very hard to maintain common notation throughout the book. We believe we have succeeded in developing a coherent treatment of the major topics outlined in the preceding overview. Chapter 1 provides an introduction to selected basic DSP material that is typically found in a first-level DSP text, and also serves to establish some of the notation used throughout the book.

In our treatment of the various topics covered herein, we generally assume that the reader has had a prior course on the fundamentals of digital signal processing. The fundamental topics assumed as background include the z-transform, the analysis and characterization of discrete-time systems, the Fourier transform, the discrete Fourier transform (DFT), and the design of FIR and IIR digital filters.

John G. Proakis
Charles M. Rader
Fuyun Ling
Chrysostomos L. Nikias
Marc Moonen
Ian K. Proudler

Contents

1

Introduction

In this chapter, we review some basic topics in digital signal processing (DSP) and, in the process, establish notation that is used throughout the text. We begin with the characterization of deterministic and random signals in Section 1.1. In Section 1.2, we describe the characterization of linear time-invariant systems, in both the time and frequency domains. Included in this treatment are definitions of basic properties of systems, such as causality, stability, minimum phase, maximum phase, mixed phase, all-pass, and bandpass. The response of linear time-invariant systems to random input signals is also derived.

The third major topic of this chapter is concerned with the sampling of signals. Conditions are derived for alias-free sampling of continuous-time signals. Also treated in this section is the discrete Fourier transform (DFT) for finite duration sequences.

Linear filtering methods based on the use of the DFT make up the fourth major topic of this chapter. The final topic of the chapter is a description of the complex cepstrum of a signal. The use of the complex cepstrum in performing signal deconvolution is also treated briefly.

The foregoing topics are usually covered in a first course in digital signal processing. Consequently, our treatment is intended to serve as a brief review. Our choice of review topics was influenced by the advanced topics treated in this book; we should emphasize, however, that many other important topics have been omitted. We assume that the reader is familiar with z-transforms and Fourier transforms, and their use in the analysis of linear time-invariant systems. We also assume that the reader is familiar with filter design methods, and design tools and algorithms for both analog and digital filters. The introductory texts by Mitra (1998), Oppenheim and Schafer (1989), and Proakis and Manolakis (1996) provide the necessary background material for the topics treated in this book.

1.1 CHARACTERIZATION OF SIGNALS

A signal is defined as any physical quantity that varies with time, space, or any other independent variable or variables. Mathematically, we describe a signal as a function of one or more independent variables. If the signal is a function of a single independent variable, the signal is said to be *one-dimensional.* On the other hand, a signal is *M-dimensional* (multidimensional) if it is a function of M independent variables.

In some applications, signals are generated by multiple sources or multiple sensors. Such signals can be represented in vector form, where each element of the vector is a signal from a single source or a single sensor. The signal vector is called a *multichannel signal.*

In this book we deal mainly with one-dimensional, single-channel or multichannel signals for which the independent variable is time. When the independent variable is continuous, the signal is called a *continuous-time signal* or an analog signal. On the other hand, when the independent variable is discrete, the signal is called a *discrete-time signal.*

1.1.1 Deterministic Signals

Let us consider a deterministic continuous-time signal $x(t)$, which may be real- or complex-valued. We assume that the signal has finite energy, defined as

$$\mathcal{E} = \int_{-\infty}^{\infty} |x(t)|^2 \, dt \tag{1.1.1}$$

Such a signal is represented in the frequency domain by its Fourier transform

$$X(F) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi F t} \, dt \tag{1.1.2}$$

where F is the frequency in cycles per second or hertz (Hz). From Parseval's theorem we have

$$\mathcal{E} = \int_{-\infty}^{\infty} |x(t)|^2 \, dt = \int_{-\infty}^{\infty} |X(F)|^2 \, dF \tag{1.1.3}$$

The quantity $|X(F)|^2$ represents the distribution of signal energy as a function of frequency and, hence, is called the *energy density spectrum.* It is denoted as

$$S_{xx}(F) = |X(F)|^2 \tag{1.1.4}$$

$S_{xx}(F)$ may also be viewed as the Fourier transform of another function, $r_{xx}(\tau)$, called the *autocorrelation function* of the finite energy signal $x(t)$, which is defined as

$$r_{xx}(\tau) = \int_{-\infty}^{\infty} x^*(t) x(t + \tau) \, dt \tag{1.1.5}$$

Indeed, it easily follows that

$$S_{xx}(F) = \int_{-\infty}^{\infty} r_{xx}(\tau)e^{-j2\pi F\tau}\,d\tau \qquad (1.1.6)$$

so that $S_{xx}(F)$ and $r_{xx}(\tau)$ are a Fourier transform pair.

Similar relationships hold for discrete-time signals, which often are the result of uniformly sampling continuous-time signals. To be specific, suppose that $x(n)$ is a real or complex-valued sequence, where n takes integer values. If $x(n)$ is deterministic and has finite energy, that is,

$$\mathcal{E} = \sum_{n=-\infty}^{\infty} |x(n)|^2 \qquad (1.1.7)$$

then $x(n)$ has the frequency domain representation

$$X(\omega) = \sum_{n=-\infty}^{\infty} x(n)e^{-j\omega n} \qquad (1.1.8)$$

or, equivalently,

$$X(f) = \sum_{n=-\infty}^{\infty} x(n)e^{-j2\pi fn} \qquad (1.1.9)$$

where $\omega = 2\pi f$. The units for the frequency variables ω and f are radians and cycles, respectively. (Or, they are radians per sample interval and cycles per sample interval, if the sequence $x(n)$ is obtained by sampling a continuous-time signal $x(t)$ at a rate of $F_s = 1/T$ samples per second, where T is the sample interval. Then, $\omega = \Omega T = 2\pi FT$ and $f = FT$).

We note that $X(\omega)$ is periodic with period $\omega_p = 2\pi$ and $X(f)$ is periodic with period $f_p = 1$. In fact, the Fourier transform relationship in (1.1.9) may be interpreted as a Fourier series representation of the periodic function $X(f)$, where the sequence $\{x(n)\}$ constitutes the set of Fourier coefficients. Thus,

$$\begin{aligned} x(n) &= \int_{-1/2}^{1/2} X(f)e^{j2\pi fn}\,df \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} X(\omega)e^{j\omega n}\,d\omega \end{aligned} \qquad (1.1.10)$$

This relationship may also be viewed as the inverse Fourier transform that yields the sequence $x(n)$ from $X(f)$ or, equivalently, from $X(\omega)$.

By applying Parseval's theorem, the energy of the discrete-time sequence is also given as

$$\mathcal{E} = \int_{-1/2}^{1/2} |X(f)|^2\,df \qquad (1.1.11)$$

The quantity $|X(f)|^2$ represents the distribution of signal energy as a function of frequency and, hence, is called the *energy density spectrum* of the discrete-time signal.

It is denoted as

$$S_{xx}(f) = |X(f)|^2 \tag{1.1.12}$$

The energy density spectrum $S_{xx}(f)$ is related to the autocorrelation sequence

$$r_{xx}(m) = \sum_{n=-\infty}^{\infty} x^*(n)x(n+m) \tag{1.1.13}$$

via the Fourier transform. That is,

$$S_{xx}(f) = \sum_{m=-\infty}^{\infty} r_{xx}(m)e^{-j2\pi fm} \tag{1.1.14}$$

Two elementary deterministic signals that we will use frequently are the unit impulse and the unit step functions. In the continuous-time domain, the unit impulse may be defined by the property

$$\int_{-\infty}^{\infty} \delta(t)g(t)\,dt = g(0) \tag{1.1.15}$$

where $g(t)$ is an arbitrary function continuous at $t = 0$. Hence, its area is

$$\int_{-\infty}^{\infty} \delta(t)\,dt = 1 \tag{1.1.16}$$

The unit step function is defined as

$$u(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0 \end{cases} \tag{1.1.17}$$

In discrete-time, the unit sample, or unit impulse sequence, is defined as

$$\delta(n) = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases} \tag{1.1.18}$$

The unit step sequence is denoted as $u(n)$ and defined as

$$u(n) = \begin{cases} 1, & n \geq 0 \\ 0, & n < 0 \end{cases} \tag{1.1.19}$$

A continuous-time signal $x(t)$ may be represented in general as the convolution of itself with a unit impulse,

$$x(t) = \int_{-\infty}^{\infty} x(\tau)\delta(t-\tau)\,d\tau \tag{1.1.20}$$

for all t. Similarly, a sequence $x(n)$ may be represented as

$$x(n) = \sum_{k=-\infty}^{\infty} x(k)\delta(n-k) \tag{1.1.21}$$

The expression in (1.1.21) is basically a convolution of the sequence $x(n)$ with the unit sample sequence $\delta(n)$. Equivalently, (1.1.21) may be viewed as the superposition (sum over k) of unit sample sequences $\delta(n-k)$, scaled in amplitude by the corresponding values $x(k)$ of the sequence $x(n)$.

1.1.2 Random Signals, Correlation Functions, and Power Spectra

In this section we provide a brief review of the characterization of random signals in terms of statistical averages expressed in both the time and frequency domains. The reader is assumed to have a background in probability theory and random processes at least at the level given in the texts of Stark and Woods (1994), Leon-Garcia (1994), Helstrom (1991), Peebles (1987), Papoulis (1984), and Davenport (1970).

Random Processes. Many physical phenomena encountered in nature are best characterized in statistical terms. For example, meteorological phenomena such as air temperature and air pressure fluctuate randomly as a function of time. Thermal noise voltages generated in the resistors of an electronic device, such as a radio or television receiver, are also randomly fluctuating phenomena. These are just a few examples of random signals. Such signals are usually modeled as infinite-duration, infinite-energy signals.

Suppose that we take the set of waveforms corresponding to the air temperatures in different cities around the world. For each city there is a corresponding waveform that is a function of time, as illustrated in Fig. 1.1. The set of all possible waveforms is called an *ensemble* of time functions or, equivalently, a *random process*. The waveform for the temperature in any particular city is a *single realization* or a *sample function* of the random process. Similarly, the thermal noise voltage generated in a resistor is a single realization or a sample function of the random process that consists of all noise voltage waveforms generated by the set of all resistors.

The set (ensemble) of all possible waveforms of a random process is denoted as $X(t, S)$, where t represents the time index and S represents the set (sample space) of all possible sample functions. A single waveform in the ensemble is denoted by $x(t, s)$. Usually, we drop the variable s (or S) for notational convenience, so that the random process is denoted as $X(t)$ and a single realization is denoted as $x(t)$.

Having defined a random process $X(t)$ as an ensemble of sample functions, let us consider the values of the process for any set of time instants $t_1 > t_2 > t_3 > \cdots > t_n$, where n is any positive integer. In general, the samples $X_{t_i} \equiv X(t_i)$, $i = 1, 2, \ldots, n$, are n random variables characterized statistically by their joint probability density function (pdf), denoted as $p(x_{t_1}, x_{t_2}, \ldots, x_{t_n})$, and any n.

Stationary Random Process. Suppose that we have n samples of the random process $X(t)$ at $t = t_i, i = 1, 2, \ldots, n$, and another set of n samples displaced in time from the first set by an amount τ. Thus the second set of samples are $X_{t_i+\tau} = X(t_i + \tau)$, $i = 1, 2, \ldots, n$, as shown in Fig. 1.1. This second set of n random variables is characterized by the joint probability density function $p(x_{t_1+\tau}, \ldots, x_{t_n+\tau})$. The joint pdf's of the two sets of random variables may or may not be identical. When they are