McGraw Hill Education

国外电子信息类系列教材

Statistical and Adaptive Signal processing

统计与自适应信号处理

（英文改编版）

[美] Dimitris G.Manolakis
Vinay K.Ingle 著
Stephen M.Kogon

阚永红 改编

Electronic Information

McGraw Hill

西安电子科技大学出版社
http://www.xduph.com

国外电子信息类系列教材

Statistical and Adaptive Signal Processing

统计与自适应信号处理(英文改编版)

Spectral Estimation, Signal Modeling and Adaptive Filtering

Dimitris G.Manolakis

Massachusetts Institute of Technology
Lincoln Laboratory

Vinay K.Ingle

Northeastern University

Stephen M.Kogon

Massachusetts Institute of Technology
Lincoln Laboratory

阎永红 改编

西安电子科技大学出版社

图书在版编目(CIP)数据

统计与自适应信号处理=Statistical and Adaptive Signal processing：**英文名/(美)马诺拉可斯**(Manolakis, D.G.),

(美)英格勒(Ingle, V.K.), **(美)哥根**(Kogon, S.M.)**著**. —西安：西安电子科技大学出版社，2012.8

ISBN 978-7-5606-2848-6

Ⅰ. ① 统… Ⅱ. ① 马… ② 英… ③ 哥… Ⅲ. ① 统计信号—信号处理—英文

② 自适应控制—信号处理—英文 Ⅳ. ① TN911.72

中国版本图书馆 CIP 数据核字(2012)第 149062 号

Dimitris G .Manolakis, Vinay K.Ingle, Stephen M.Kogen

Statistical and Adaptive Signal Processing, Spectral Estimation, Signal Modeling ,Adaptive Filtering and Array Processing.

ISBN：0-07-040051-2

策　　划　曹媛媛

责任编辑　曹媛媛　马晓娟

出版发行　西安电子科技大学出版社(西安市太白南路 2 号)

电　　话　(029)88242885　88201467　　邮　　编　710071

网　　址　www.xduph.com　　电子邮箱　xdupfxb001@163.com

经　　销　新华书店

印刷单位　西安文化彩印厂

版　　次　2012 年 8 月第 1 版　　2012 年 8 月第 1 次印刷

开　　本　880 毫米×1230 毫米　1/16　印　张　24.75

字　　数　645 千字

印　　数　1～3000 册

定　　价　56.00 元

ISBN 978-7-5606-2848-6/TN・0664

XDUP　3140001 -1

如有印装问题可调换

本社图书封面为激光防伪覆膜，谨防盗版。

ABOUT THE AUTHORS

DIMITRIS G.MANOLAKIS, a native of Greece, received his education (B.S. in physics and Ph.D. in electrical engineering) from the University of Athens, Greece. He is currently a member of the technical staff and MIT Lincoln Laboratory, in Lexington, Massachusetts. Previously, he was a Principal Member, Research Staff, at Riverside Research Institute. Dr. Manolakis has taught at the University of Athens, Northeastern University, Boston College, and Worcester Polytechnic Institure; and he is coauthor of the textbook *Digital Signal Processing: Principles, Algorithms, and Applications* (Prentice-Hall, 1996, 3d ed.). His research experience and interests include the areas of digital signal processing, adaptive filtering, array processing, pattern recognition, and radar systems.

VINAY K.INGLE is Associate Professor of Electrical and Computer Engineering at Northeastern University. He received his Ph.D. in electrical and computer engineering from Rensselaer Polytechnic Institute in 1981. He has broad research experience and has taught courses on topics including signal and image processing, stochastic processes, and estimation theory. Professor Ingle is coauthor of the textbooks *DSP Laboratory Using the ADSP-2101 Microprocessor* (Prentice-Hall, 1991) and *DSP Using Matlab* (PWS Publishing Co., Boston, 1996).

STEPHEN M.KOGON received the Ph.D. degree in electrical engineering from Georgia Institute of Technology. He is currently a member of the technical staff at MIT Lincoln Laboratory in Lexington, Massachusetts. Previously, he has been associated with Raytheon Co., Boston College, and Georgia Tech Research Institute. His research interests are in the areas of adaptive processing, array signal processing, radar, and statistical signal modeling.

One must learn by doing the thing;
for though you think you know it
You have no certainty, until you try.
—Sophocles, *Trachiniae*

PREFACE

The principal goal of this book is to provide a unified introduction to the theory, implementation, and applications of statistical and adaptive signal processing methods. We have focused on the key topics of spectral estimation, signal modeling, adaptive filtering, whose selection was based on the grounds of theoretical value and practical importance. The book has been primarily written with students and instructors in mind. The principal objectives are to provide an introduction to basic concepts and methodologies that can provide the foundation for further study, research, and application to new problems. To achieve these goals, we have focused on topics that we consider fundamental and have either multiple or important applications.

Approach and prerequisites

The adopted approach is intended to help both students and practicing engineers understand the fundamental mathematical principles underlying the operation of a method, appreciate its inherent limitations, and provide sufficient details for its practical implementation. The academic flavor of this book has been influenced by our teaching whereas its practical character has been shaped by our research and development activities in both academia and industry. The mathematical treatment throughout this book has been kept at a level that is within the grasp of upper-level undergraduate students, graduate students, and practicing electrical engineers with a background in digital signal processing, probability theory, and linear algebra.

Organization of the book

Chapter 1 introduces the basic concepts and applications of statistical and adaptive signal processing and provides an overview of the book. Chapters 2 introduce some basic concepts of estimation theory. Chapter 3 provides a treatment of parametric linear signal models (both deterministic and stochastic) in the time and frequency domains. Chapter 4 presents the most practical methods for the estimation of correlation and spectral densities. Chapter 5 provides a detailed study of the theoretical properties of optimum filters, assuming that the relevant signals can be modeled as stochastic processes with known statistical properties; and Chapter 6 contains algorithms and structures for optimum filtering, signal modeling, and prediction. Chapter 7 introduces the principle of least-squares estimation and its application to the design of practical filters and predictors. Chapters 8 and 9 use the theoretical work in Chapters 3, 5 and 6 and the practical methods in Chapter 7 to develop, evaluate, and apply practical techniques for signal modeling, adaptive filtering,

Theory and practice

It is our belief that sound theoretical understanding goes hand-in-hand with practical implementation and application to real-world problems. Therefore, the book includes a large number of computer experiments that illustrate important concepts and help the reader to easily implement the various methods. Every chapter includes examples, problems, and computer experiments that facilitate the comprehension of the material. To help the reader understand the theoretical basis and limitations of the various methods and apply them to real-world problems, we provide MATLAB functions for all major algorithms and examples illustrating their use.

Feedback

Although we are fully aware that there always exists room for improvement, we believe that this book is a big step forward for an introductory textbook in statistical and adaptive signal processing. However, as engineers, we know that every search for the optimum requires the will to change and quest for additional improvement. Thus, we would appreciate feedback from teachers, students, and engineers using this book for self-study at `vingle@lynx.neu.edu`.

Acknowledgments

We are indebted to a number of individuals who have contributed in different, but important, ways to the shaping of our knowledge in general and the preparation of this book in particular. In this respect we wish to extend our appreciation to C. Caroubalos, G. Carayannis, J. Makhoul (DGM), M. Schetzen (VKI), and J. Holder, S. Krich, and D. Williams (SMK). We are grateful to E. Baranoski, G. Borsari, J. Schodorf, S. Smith, A. Steinhardt, and J. Ward for helping us to improve the presentation in various parts of the book.

We express our sincere gratitude to Kevin Donohue, University of Kentucky; Amro El-Jaroudi, University of Pittsburgh; Edward A. Lee, University of California-Berkeley; Ray Liu, University of Maryland; Randy Moses, The Ohio State University; and Kristina Ropella, Marquette University, for their constructive and helpful reviews.

Lynn Cox persuaded us to choose McGraw-Hill, Inc., as our publisher, and we have not regretted that decision. We are grateful to Lynn for her enthusiasm and her influence in shaping the scope and the objectives of our book. The fine team at McGraw-Hill, including Michelle Flomenhoft, Catherine Fields, Betsy Jones, and Nina Kreiden, has made the publication of this book an exciting and pleasant experience. We also thank N. Bulock and Mathworks, Inc., for promptly providing various versions of MATLAB and A. Turcotte for helping with some of the drawings in the book.

Last, but not least, we would like to express our sincere appreciation to our families for their full-fledged support and understanding over the past several years. We fully realize that the completion of such a project would not be possible without their continual sustenance and encouragement.

Dimitris G. Manolakis
Vinay K. Ingle
Stephen M. Kogon

Contents

CHAPTER 1

Introduction

This book is an introduction to the theory and algorithms used for the analysis and processing of random signals and their applications to real-world problems. The fundamental characteristic of random signals is captured in the following statement: Although random signals are evolving in time in an unpredictable manner, their average statistical properties exhibit considerable regularity. This provides the ground for the description of random signals using statistical averages instead of explicit equations. When we deal with random signals, the main objectives are the statistical description, modeling, and exploitation of the dependence between the values of one or more discrete-time signals and their application to theoretical and practical problems.

Random signals are described mathematically by using the theory of probability, random variables, and stochastic processes. However, in practice we deal with random signals by using statistical techniques. Within this framework we can develop, at least in principle, theoretically *optimum signal processing* methods that can inspire the development and can serve to evaluate the performance of practical *statistical signal processing* techniques. The area of *adaptive signal processing* involves the use of optimum and statistical signal processing techniques to design signal processing systems that can modify their characteristics, during normal operation (usually in real time), to achieve a clearly predefined application-dependent objective.

The purpose of this chapter is twofold: to illustrate the nature of random signals with some typical examples and to introduce the three major application areas treated in this book: *spectral estimation*, *signal modeling and adaptive filtering*. Throughout the book, the emphasis is on the application of techniques to actual problems in which the theoretical framework provides a foundation to motivate the selection of a specific method.

1.1 Random Signals

A *discrete-time signal* or *time series* is a set of observations taken sequentially in time, space, or some other independent variable. Examples occur in various areas, including engineering, natural sciences, economics, social sciences, and medicine.

A discrete-time signal $x(n)$ is basically a sequence of real or complex numbers called samples. Although the integer index n may represent any physical variable (e.g., time, distance), we shall generally refer to it as *time*. Furthermore, in this book we consider only time series with observations occurring at equally spaced intervals of time.

Discrete-time signals can arise in several ways. Very often, a discrete-time signal is obtained by periodically sampling a continuous-time signal, that is, $x(n) = x_c(nT)$, where $T = 1/F_s$ (seconds) is the sampling period and F_s (samples per second or hertz) is the sampling frequency. At other times, the samples of a discrete-time signal are obtained by accumulating some quantity (which does not have an instantaneous value) over equal intervals of time, for example, the number of cars per day traveling on a certain road. Finally, some signals are inherently discrete-time, for example, daily stock market prices. *Throughout the book, except if otherwise stated, the terms* signal, time series, *or* sequence *will be used to refer to a discrete-time signal.*

The key characteristics of a time series are that the observations are *ordered* in time and that adjacent observations are *dependent* (related). To see graphically the relation between the samples of a signal that are l sampling intervals away, we plot the points $\{x(n), x(n+l)\}$ for $0 \le n \le N-1-l$, where N is the length of the data record. The resulting graph is known as the *l lag scatter plot*. This is illustrated in Figure 1.1, which shows a speech signal and two scatter plots that demonstrate the correlation between successive samples. We note that for adjacent samples the data points fall close to a straight line with a positive slope. This implies high correlation because every sample is followed by a sample with about the same amplitude. In contrast, samples that are 20 sampling intervals apart are much less correlated because the points in the scatter plot are randomly spread.

When successive observations of the series are dependent, we may use past observations to predict future values. If the prediction is exact, the series is said to be *deterministic*. However, in most practical situations we cannot predict a time series exactly. Such time series are called *random* or *stochastic*, and the degree of their predictability is determined by the dependence between consecutive observations. The ultimate case of randomness occurs when every sample of a random signal is independent of all other samples. Such a signal, which is completely unpredictable, is known as *white noise* and is used as a building block to simulate random signals with different types of dependence. To summarize, the fundamental characteristic of a random signal is the inability to precisely specify its values. In other words, a random signal is not predictable, it never repeats itself, and we cannot find a mathematical formula that provides its values as a function of time. As a result, random signals can only be mathematically described by using the theory of stochastic processes.

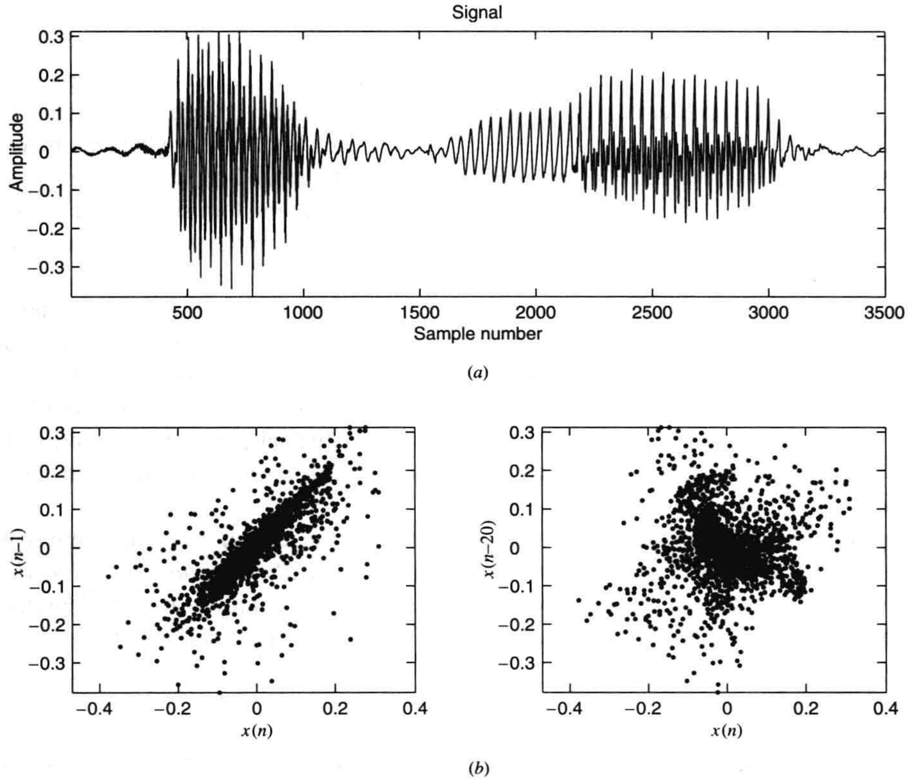

FIGURE 1.1
(*a*) The waveform for the speech signal "signal"; (*b*) two scatter plots for successive samples and samples separated by 20 sampling intervals.

This book provides an introduction to the fundamental theory and a broad selection of algorithms widely used for the processing of discrete-time random signals. Signal processing techniques, dependent on their main objective, can be classified as follows (see Figure 1.2):

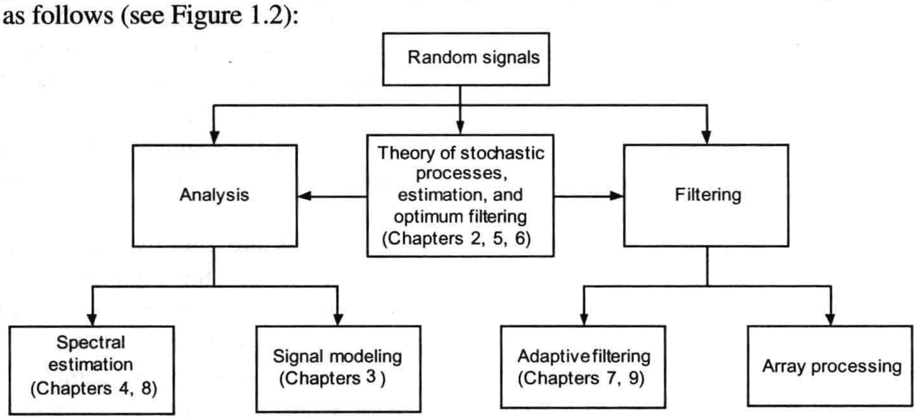

FIGURE 1.2
Classification of methods for the analysis and processing of random signals.

- **Signal analysis.** The primary goal is to extract useful information that can be used to understand the signal generation process or extract features that can be used for signal classification purposes. Most of the methods in this area are treated under the disciplines of *spectral estimation* and *signal modeling*. Typical applications include detection and classification of radar and sonar targets, speech and speaker recognition, detection and classification of natural and artificial seismic events, event detection and classification in biological and financial signals, efficient signal representation for data compression, etc.

- **Signal filtering.** The main objective of signal filtering is to improve the quality of a signal according to an acceptable criterion of performance. Signal filtering can be subdivided into the areas of frequency selective filtering and adaptive filtering. Typical applications include noise and interference cancelation, echo cancelation, channel equalization, seismic deconvolution, active noise control, etc.

We conclude this section with some examples of signals occurring in practical applications. Although the description of these signals is far from complete, we provide sufficient information to illustrate their random nature and significance in signal processing applications.

Speech signals. Figure 1.3 shows the spectrogram and speech waveform corresponding to the utterance "signal." The spectrogram is a visual representation of the distribution of the signal energy as a function of time and frequency. We note that the speech signal has significant changes in both amplitude level and spectral content across time. The waveform contains segments of voiced (quasi-periodic) sounds, such as "e," and unvoiced or fricative (noiselike) sounds, such as "g."

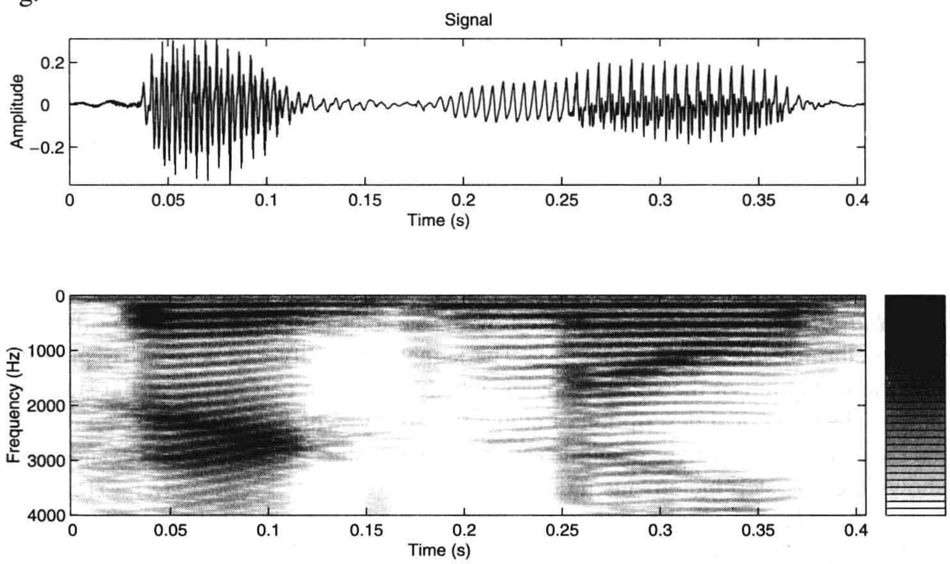

FIGURE 1.3
Spectrogram and acoustic waveform for the utterance "signal." The horizontal dark bands show the resonances of the vocal tract, which change as a function of time depending on the sound or phoneme being produced.

Speech production involves three processes: generation of the sound excitation, articulation by the vocal tract, and radiation from the lips and/or nostrils. If the excitation is a quasi-periodic train of air pressure pulses, produced by the vibration of the vocal cords, the result is a voiced sound. Unvoiced sounds are produced by first creating a constriction in the vocal tract, usually toward the mouth end. Then we generate turbulence by forcing air through the constriction at a sufficiently high velocity. The resulting excitation is a broadband noiselike waveform.

The spectrum of the excitation is shaped by the vocal tract tube, which has a frequency response that resembles the resonances of organ pipes or wind instruments. The resonant frequencies of the vocal tract tube are known as *formant frequencies*, or simply *formants*. Changing the shape of the vocal tract changes its frequency response and results in the generation of different sounds. Since the shape of the vocal tract changes slowly during continuous speech, we usually assume that it remains almost constant over intervals on the order of 10 ms. More details about speech signal generation and processing can be found in Rabiner and Schafer 1978; O'Shaughnessy 1987; and Rabiner and Juang 1993.

Other examples of the stochastic signals are Electrophysiological signals, geophysical signals, radar signals.

1.2 Spectral Estimation

The central objective of signal analysis is the development of quantitative techniques to study the properties of a signal and the differences and similarities between two or more signals from the same or different sources. The major areas of random signal analysis are (1) statistical analysis of signal amplitude (i.e., the sample values); (2) analysis and modeling of the correlation among the samples of an individual signal; and (3) joint signal analysis (i.e., simultaneous analysis of two signals in order to investigate their interaction or interrelationships). These techniques are summarized in Figure 1.4. The prominent tool in signal analysis is spectral estimation, which is a generic term for a multitude of techniques used to estimate the distribution of energy or power of a signal from a set of observations. Spectral estimation is a very complicated process that requires a deep understanding of the underlying theory and a great deal of practical experience. Spectral analysis finds many applications in areas such as medical diagnosis, speech analysis, seismology and geophysics, radar and sonar, nondestructive fault detection, testing of physical theories, and evaluating the predictability of time series.

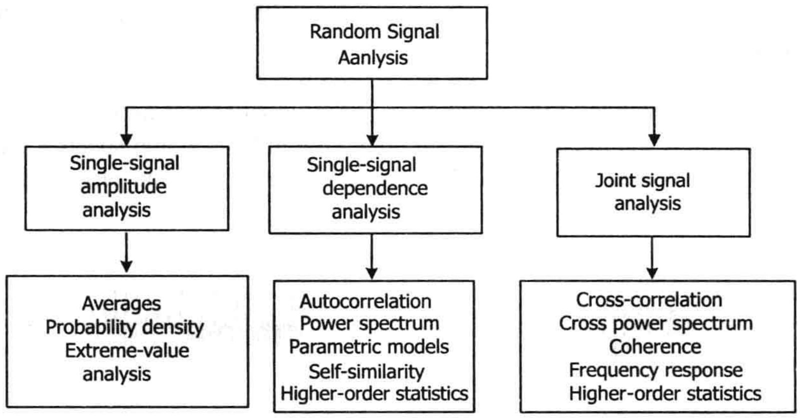

FIGURE 1.4
Summary of random signal analysis techniques.

Amplitude distribution. The range of values taken by the samples of a signal and how often the signal assumes these values together determine the signal variability. The signal variability can be seen by plotting the time series and is quantified by the histogram of the signal samples, which shows the percentage of the signal amplitude values within a certain range. The numerical description of signal variability, which depends only on the value of the signal samples and not on their ordering, involves quantities such as mean value, median, variance, and dynamic range.

Figure 1.5 shows the one-step increments, that is, the first difference $x_d(n) = x(n) - x(n-1)$ whereas Figure 1.6 shows their histograms. Careful examination of the shape of the histogram curves indicates that the second signal jumps quite frequently between consecutive samples with large steps. In other words, the probability of large increments is significant, as exemplified by the fat tails of the histogram in Figure 1.6(*b*). The knowledge of the probability of extreme values is essential in the design of detection systems for digital communications, military surveillance using infrared and radar sensors, and intensive care monitoring. In general, the shape of the histogram, or more precisely the probability density, is very important in applications such as signal coding and event detection. Although many practical signals follow a Gaussian distribution, many other signals of practical interest have distributions that are non-Gaussian. For example, speech signals have a probability density that can be reasonably approximated by a gamma distribution (Rabiner and Schafer 1978).

The significance of the Gaussian distribution in signal processing stems from the following facts. First, many physical signals can be described by Gaussian processes. Second, the central limit theorem states that any process that is the result of the combination of many elementary processes will tend, under quite general conditions, to be Gaussian. Finally, linear systems preserve the Gaussianity of their input signals. To understand the last two statements, consider N independent random quantities $x_1, x_2, \ldots, x_N$ with the same probability density $p(x)$ and pose the following question: When does the probability distribution $p_N(x)$ of their sum $x = x_1 + x_2 + \cdots + x_N$ have the same shape (within a scale factor) as the distribution $p(x)$ of the individual quantities? The standard answer is that $p(x)$ should be Gaussian, because the sum of N Gaussian random variables is again a Gaussian, but with variance equal to N times that of the individual signals. However, if we allow for distributions with infinite variance, additional solutions are possible. The resulting probability distributions, known as *stable* or *Levy distributions*, have infinite variance and are characterized by a thin main lobe and fat tails, resembling the shape of the histogram in Figure 1.6(*b*). Interestingly enough, the Gaussian distribution is a stable distribution with finite variance (actually the only one). Because Gaussian and stable non-Gaussian distributions are invariant under linear signal processing operations, they are very important in signal processing.

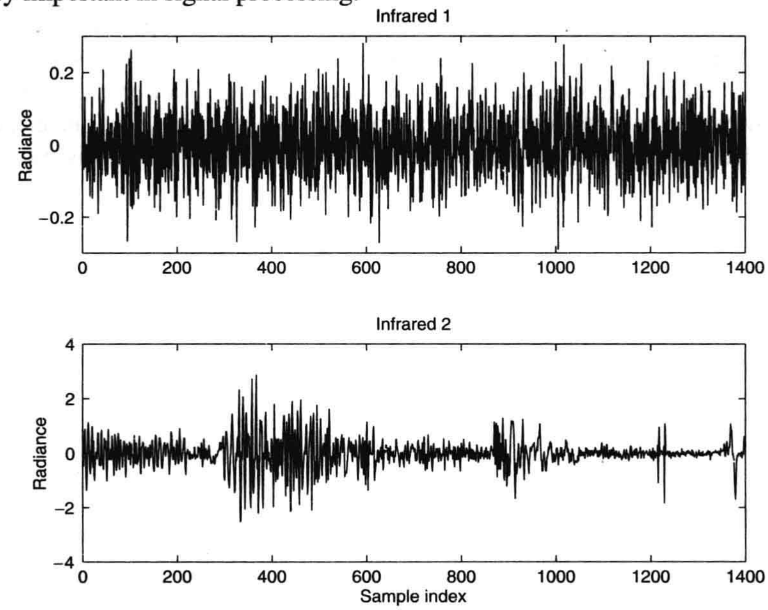

FIGURE 1.5

One-step-increment time series for the infrared data.

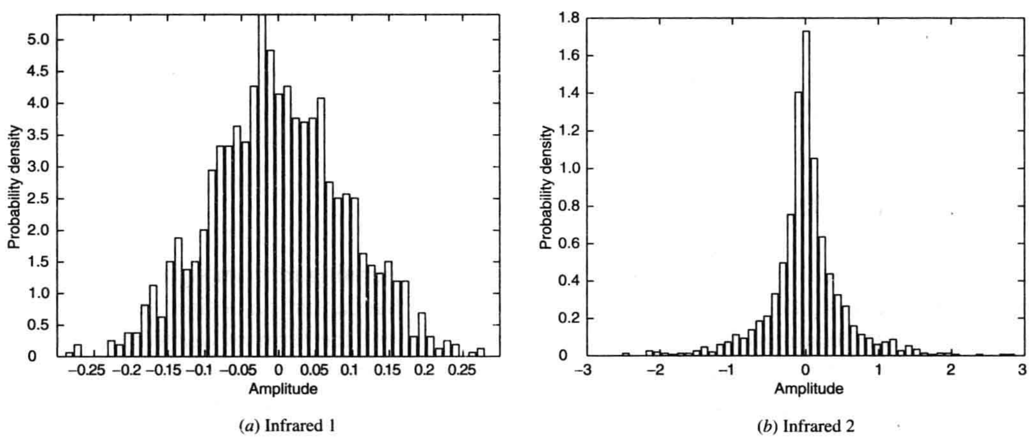

FIGURE 1.6

Histograms for the infrared increment signals.

Correlation and spectral analysis. Although scatter plots (see Figure 1.1) illustrate nicely the existence of correlation, to obtain quantitative information about the correlation structure of a time series $x(n)$ with zero mean value, we use the *empirical normalized autocorrelation sequence*

$$\hat{\rho}(l) = \frac{\sum_{n=l}^{N-1} x(n)x^*(n-l)}{\sum_{n=0}^{N-1} |x(n)|^2} \tag{1.2.1}$$

which is an estimate of the theoretical normalized autocorrelation sequence. For lag $l = 0$, the sequence is perfectly correlated with itself and we get the maximum value of 1. If the sequence does not change significantly from sample to sample, the correlation of the sequence with its shifted copies, though diminished, is still close to 1. Usually, the correlation decreases as the lag increases because distant samples become less and less dependent. Note that reordering the samples of a time series changes its autocorrelation but not its histogram.

We say that signals whose empirical autocorrelation decays fast, such as an exponential, have short-memory or short-range dependence. If the empirical autocorrelation decays very slowly, as a hyperbolic function does, we say that the signal has long-memory or long-range dependence. Furthermore, we shall see in the next section that effective modeling of time series with short or long memory requires different types of models.

The spectral density function shows the distribution of signal power or energy as a function of frequency (see Figure 1.7). The autocorrelation and the spectral density of a signal form a Fourier transform pair and hence contain the same information. However, they present this information in different forms, and one can reveal information that cannot be easily extracted from the other. It is fair to say that the spectral density is more widely used than the autocorrelation.

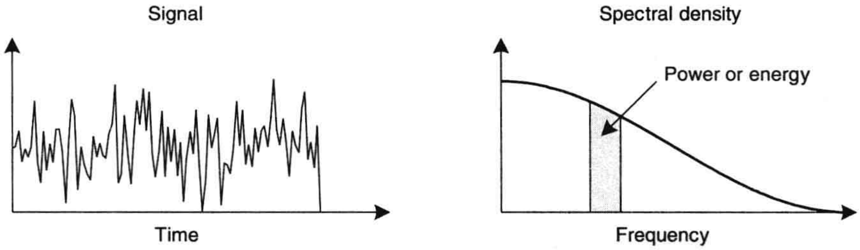

FIGURE 1.7
Illustration of the concept of power or energy spectral density function of a random signal.

Joint signal analysis. In many applications, we are interested in the relationship between two different random signals. There are two cases of interest. In the first case, the two signals are of the same or similar nature, and we want to ascertain and describe the similarity or interaction between them.

In the second case, we may have reason to believe that there is a *causal relationship* between the two signals. For example, one signal may be the input to a system and the other signal the output. The task in this case is to find an accurate description of the system, that is, a description that allows accurate estimation of future values of the output from the input. This process is known as *system modeling* or *identification* and has many practical applications, including understanding the operation of a system in order to improve the design of new systems or to achieve better control of existing systems.

1.3 Signal Modeling

In many theoretical and practical applications, we are interested in generating random signals with certain properties or obtaining an efficient representation of real-world random signals that captures a desired set of their characteristics (e.g., correlation or spectral features) in the best possible way. We use the term *model* to refer to a mathematical description that provides an efficient representation of the "essential" properties of a signal.

For example, a finite segment $\{x(n)\}_{n=0}^{N-1}$ of any signal can be approximated by a linear combination of constant $(\lambda_k = 1)$ or exponentially fading $(0 < \lambda_k < 1)$ sinusoids

$$x(n) \simeq \sum_{k=1}^{M} a_k \lambda_k^n \cos(\omega_k n + \varphi_k) \tag{1.3.1}$$

where $\{a_k, \lambda_k, \omega_k, \varphi_k\}_{k=1}^{M}$ are the model parameters. A good model should provide an accurate description of the signal with $4M \ll N$ parameters. From a practical viewpoint, we are most interested in *parametric models*, which assume a given functional form completely specified by a finite number of parameters. In contrast, *nonparametric models* do not put any restriction on the functional form or the number of model parameters.

If any of the model parameters in (1.3.1) is random, the result is a random signal. The most widely used model is given by

$$x(n) = \sum_{k=1}^{M} a_k \cos(\omega_k n + \varphi_k)$$

where the amplitudes $\{a_k\}_1^N$ and the frequencies $\{\omega_k\}_1^N$ are constants and the phases $\{\varphi_k\}_1^N$ are random. This model is known as the *harmonic process model* and has many theoretical and practical applications.

Suppose next that we are given a sequence $\omega(n)$ of independent and identically distributed observations. We can create a time series $x(n)$ with dependent observations, by linearly combining the values of $\omega(n)$ as

$$x(n) = \sum_{k=-\infty}^{\infty} h(k)\omega(n-k) \tag{1.3.2}$$

which results in the widely used *linear random signal model*. The model specified by the convolution summation (1.3.2) is clearly nonparametric because, in general, it depends on an infinite number of parameters. Furthermore, the model is a linear, time-invariant system with impulse response $h(k)$ that determines the *memory* of the model and, therefore, the dependence properties of the output $x(n)$. By properly choosing the weights $h(k)$, we can generate a time series with almost any type of dependence among its samples.

In practical applications, we are interested in linear parametric models. As we will see, parametric models exhibit a dependence imposed by their structure. However, if the number of parameters approaches the range of the dependence (in number of samples), the model can mimic any form of dependence. The list of desired features for a good model includes these: (1) the number of model parameters should be as small as possible (*parsimony*), (2) estimation of the model parameters from the data should be easy, and (3) the model parameters should have a physically meaningful interpretation.

If we can develop a successful parametric model for the behavior of a signal, then we can use the model for various applications:

1. To achieve a better understanding of the physical mechanism generating the signal (e.g., earth structure in the case of seismograms).
2. To track changes in the source of the signal and help identify their cause (e.g., EEG).
3. To synthesize artificial signals similar to the natural ones (e.g., speech, infrared backgrounds, natural scenes, data network traffic).
4. To extract parameters for pattern recognition applications (e.g., speech and character recognition).
5. To get an efficient representation of signals for data compression (e.g., speech, audio, and video coding).
6. To forecast future signal behavior (e.g., stock market indexes) (Pindyck and Rubinfeld 1998).

In practice, signal modeling involves the following steps: (1) selection of an appropriate model, (2) selection of the "right" number of parameters, (3) fitting of the model to the actual data, and (4) model testing to see if the model satisfies the user requirements for the particular application. As we shall see in Chapter 8 this process is very complicated and depends heavily on the understanding of the theoretical model properties (see Chapter 3), the amount of familiarity with the particular application, and the experience of the user.

Rational or Pole-Zero Models

Suppose that a given sample $x(n)$, at time n, can be approximated by the previous sample weighted by a coefficient a, that is, $x(n) \approx ax(n-1)$, where a is assumed constant over the signal segment to be modeled. To make the above relationship exact, we add an excitation term $\omega(n)$, resulting in

$$x(n) = ax(n-1) + \omega(n) \tag{1.3.3}$$

where $\omega(n)$ is an excitation sequence. Taking the z-transform of both sides, we have

$$X(z) = az^{-1}X(z) + W(z) \tag{1.3.4}$$

which results in the following system function:

$$H(z) = \frac{X(z)}{W(z)} = \frac{1}{1-az^{-1}} \tag{1.3.5}$$

By using the identity

$$H(z) = \frac{1}{1-az^{-1}} = 1 + az^{-1} + a^2z^{-2} + \cdots \qquad -1 < a < 1 \tag{1.3.6}$$

the single-parameter model in (1.3.3) can be expressed in the following nonparametric form

$$x(n) = \omega(n) + a\omega(n-1) + a^2\omega(n-2) + \cdots \tag{1.3.7}$$

which clearly indicates that the model generates a time series with *exponentially* decaying dependence.

A more general model can be obtained by including a linear combination of the P previous values of the signal and of the Q previous values of the excitation in (1.3.3), that is,

$$x(n) = \sum_{k=1}^{P}(-a_k)x(n-k) + \sum_{k=0}^{Q} d_k\omega(n-k) \tag{1.3.8}$$

The resulting system function