INTRODUCTION TO APPLIED STATISTICAL SIGNAL ANALYSIS: GUIDE TO BIOMEDICAL AND ELECTRICAL ENGINEERING APPLICATIONS
INTRODUCTION TO
APPLIED STATISTICAL
SIGNAL ANALYSIS: GUIDE
TO BIOMEDICAL AND
ELECTRICAL ENGINEERING
APPLICATIONS

Richard Shiavi

Vanderbilt University
Nashville, TN
CONTENTS

* = new section
PREFACE xiii
DEDICATION xv
ACKNOWLEDGMENTS xvii
LIST OF SYMBOLS xix

I
INTRODUCTION AND TERMINOLOGY

1.1 Introduction 1
1.2 Signal Terminology 3
  1.2.1 Domain Types 3
  1.2.2 Amplitude Types 5
  1.2.3 Basic Signal Forms 6
  1.2.4 The Transformed Domain—The Frequency Domain 8
  1.2.5 General Amplitude Properties 9
1.3 Analog to Digital Conversion 10
1.4 Measures of Signal Properties 11
  1.4.1 Time Domain 11
  1.4.2 Frequency Domain 12
References 13
2
EMPIRICAL MODELING AND APPROXIMATION

2.1 Introduction 15
2.2 Model Development 16
2.3 Generalized Least Squares 21
2.4 Generalities 23
2.5 Models from Linearization 24
2.6 Orthogonal Polynomials 28
2.7 Interpolation and Extrapolation 33
  2.7.1 Lagrange Polynomials 34
  2.7.2 Spline Interpolation 38
2.8 Overview 43
References 43
Exercises 44

3
FOURIER ANALYSIS

3.1 Introduction 51
3.2 Review of Fourier Series 53
  3.2.1 Definition 53
  3.2.2 Convergence 60
3.3 Overview of Fourier Transform Relationships 61
  3.3.1 Continuous versus Discrete Time 61
  3.3.2 Discrete Time and Frequency 63
3.4 Discrete Fourier Transform 64
  3.4.1 Definition Continued 64
  3.4.2 Partial Summary of DFT Properties and Theorems 65
3.5 Fourier Analysis 68
  3.5.1 Frequency Range and Scaling 69
  3.5.2 The Effect of Discretizing Frequency 70
  3.5.3 The Effect of Truncation 73
  3.5.4 Windowing 77
  3.5.5 Resolution 79
  3.5.6 Detrending 82
3.6 Procedural Summary 82
3.7 Selected Applications 82
References 86
Exercises 87
Appendix 3.1: DFT of Ionosphere Data 92
Appendix 3.2: Review of Properties of Orthogonal Functions 93
4 PROBABILITY CONCEPTS AND SIGNAL CHARACTERISTICS

4.1 Introduction 101
4.2 Introduction to Random Variables 102
  4.2.1 Probability Descriptors 102
  4.2.2 Moments of Random Variables 108
  4.2.3 Gaussian Random Variable 110
4.3 Joint Probability 112
  4.3.1 Bivariate Distributions 112
  4.3.2 Moments of Bivariate Distributions 113
4.4 Concept of Sampling and Estimation 115
  4.4.1 Sample Moments 115
  4.4.2 Significance of the Estimate 119
4.5 Density Function Estimation 122
  4.5.1 General Principle for $\chi^2$ Approach 122
  4.5.2 Detailed Procedure for $\chi^2$ Approach 124
  *4.5.3 Quantile-Quantile Approach 127
4.6 Correlation and Regression 130
  *4.6.1 Estimate of Correlation 130
  *4.6.2 Simple Regression Model 132
4.7 General Properties of Estimators 136
  4.7.1 Convergence 136
  4.7.2 Recursion 137
  *4.7.3 Maximum Likelihood Estimation 138
4.8 Random Numbers and Signal Characteristics 139
  4.8.1 Random Number Generation 140
  4.8.2 Change of Mean and Variance 141
  4.8.3 Density Shaping 142

References 145
Exercises 146
Appendix 4.1: Plots and Formulas for Five Probability Density Functions 154

5 INTRODUCTION TO RANDOM PROCESSES AND SIGNAL PROPERTIES

5.1 Introduction 155
5.2 Definition of Stationarity 156
5.3 Definition of Moment Functions 159
  5.3.1 General Definitions 159
  5.3.2 Moments of Stationary Processes 160
5.4 Time Averages and Ergodicity 162
5.5 Estimating Correlation Functions 166
  5.5.1 Estimator Definition 166
  5.5.2 Estimator Bias 168
  5.5.3 Consistency and Ergodicity 168
  5.5.4 Sampling Properties 170
  5.5.5 Asymptotic Distributions 171
5.6 Correlation and Signal Structure 176
  5.6.1 General Moving Average 176
  5.6.2 First-Order MA 177
  5.6.3 Second-Order MA 181
  5.6.4 Overview 181
5.7 Assessing Stationarity of Signals 182
  *5.7.1 Multiple Segments—Parametric 184
  *5.7.2 Multiple Segments—Nonparametric 189
References 193
Exercises 194
Appendix 5.1: Variance of Autocovariance Estimate 197
Appendix 5.2: Stationarity Tests 198

6
RANDOM SIGNALS, LINEAR SYSTEMS,
AND POWER SPECTRA

6.1 Introduction 201
6.2 Power Spectra 201
  *6.2.1 Empirical Approach 201
  *6.2.2 Theoretical Approach 203
6.3 System Definition Review 205
  6.3.1 Basic Definitions 205
  6.3.2 Relationships between Input and Output 208
6.4 Systems and Signal Structure 210
  6.4.1 Moving Average Process 210
  6.4.2 Structure with Autoregressive Systems 211
  6.4.3 Higher-Order AR Systems 215
6.5 Time Series Models for Spectral Density 219
References 225
Exercises 226


7

SPECTRAL ANALYSIS FOR RANDOM SIGNALS:
NONPARAMETRIC METHODS

7.1 Spectral Estimation Concepts 229
   7.1.1 Developing Procedures 233
   7.1.2 Sampling Moments of Estimators 234
7.2 Sampling Distribution for Spectral Estimators 239
   7.2.1 Spectral Estimate for White Noise 239
   7.2.2 Sampling Properties for General Random Processes 242
7.3 Consistent Estimators—Direct Methods 244
   7.3.1 Spectral Averaging 224
   7.3.2 Confidence Limits 248
   7.3.3 Summary of Procedure for Spectral Averaging 258
   7.3.4 Welch Method 259
   7.3.5 Spectral Smoothing 259
   7.3.6 Additional Applications 263
7.4 Consistent Estimators—Indirect Methods 264
   7.4.1 Spectral and Lag Windows 264
   7.4.2 Important Details for Using FFT Algorithms 266
   7.4.3 Statistical Characteristics of BT Approach 267
7.5 Autocorrelation Estimation 275

References 277
Exercises 278

Appendix 7.1: Variance of Periodogram 281
Appendix 7.2: Proof of Variance of BT Spectral Smoothing 283
Appendix 7.3: Window Characteristics 284
Appendix 7.4: Lag Window Functions 285
Appendix 7.5: Spectral Estimates from Smoothing 286

8

RANDOM SIGNAL MODELING AND PARAMETRIC
SPECTRAL ESTIMATION

8.1 Introduction 287
8.2 Model Development 288
8.3 Random Data Modeling Approach 293
   8.3.1 Basic Concepts 293
   8.3.2 Solution of General Model 296
9 THEORY AND APPLICATION OF CROSS CORRELATION AND COHERENCE

9.1 Introduction 331
9.2 Properties of Cross Correlation Functions 333
  9.2.1 Theoretical Function 333
  9.2.2 Estimators 334
9.3 Detection of Time-Limited Signals 339
  9.3.1 Basic Concepts 340
  9.3.2 Application of Pulse Detection 342
  9.3.3 Random Signals 343
  9.3.4 Time Difference of Arrival 345
  9.3.5 Marine Seismic Signal Analysis 347
  9.3.6 Procedure for Estimation 347
9.4 Cross Spectral Density Functions 349
  9.4.1 Definition and Properties 349
  9.4.2 Properties of Cross Spectral Estimators 351
9.5 Applications 354
9.6 Tests for Correlation between Time Series 355
  9.6.1 Coherence Estimators 355
  9.6.2 Statistical Properties of Estimators 358
  9.6.3 Confidence Limits 359
  9.6.4 Procedure for Estimation 362
  9.6.5 Application 362
References 364
Exercises 365


CONTENTS

*10

ENVELOPES AND KERNEL FUNCTIONS

10.1 The Hilbert Transform and Analytic Functions 367
10.1.1 Introduction 367
10.1.2 Hilbert Transform 368
10.1.3 Analytic Signal 370
10.1.4 Discrete Hilbert Transform 373
10.2 Point Processes and Continuous Signals via Kernel Functions 375
10.2.1 Concept 375
10.2.2 Nerve Activity and the Spike Density Function 378

References 382
Exercises 383

APPENDICES

Table A Values of the Standardized Normal cdf \( \Phi(z) \)
Table B Student’s \( t \) Distribution
Table C Chi-Square Distribution
"Table D Critical Points for the Q-Q Plot Correlation Coefficient Test for Normality
Table E F Distribution Significance Limit for 97.5th Percentile
Table F Percentage Points of Run Distribution

INDEX 393
This book presents a practical introduction to signal analysis techniques that are commonly used in a broad range of engineering areas such as biomedical engineering, communications, geophysics, speech, etc. In order to emphasize the analytic approaches, a certain background is necessary. The book is designed for an individual who has a basic background in mathematics, science, and computer programming that is required in an undergraduate engineering curriculum. In addition one needs to have an introductory-level background in probability and statistics and discrete time systems.

The sequence of material begins with definitions of terms and symbols used for representing discrete data measurements and time series/signals and a presentation of techniques for polynomial modeling and data interpolation. Chapter 3 focuses on the windowing and the discrete Fourier transform. It is introduced by presenting first the various definitions of the Fourier transform and harmonic modeling using the Fourier series. The remainder of the book deals with signals having some random signal component and Chapter 4 reviews the aspects of probability theory and statistics needed for subsequent topics. In addition, histogram fitting, correlation, regression, and maximum likelihood estimation are presented. In the next chapter these concepts are extended to define random signals and introduce the estimation of correlation functions and tests of stationarity. Chapter 6 reviews linear systems and defines power spectra. Chapter 7 presents classical spectral analysis and its estimators. The periodogram and Blackman-Tukey methods are covered in detail. Chapter 8 covers autoregressive modeling of signals and parametric spectral estimation. Chapter 9 presents the classical uses of cross correlation and coherence functions. In particular, the practical techniques for estimating coherence function are presented in detail. Chapter 10 is a new chapter for the third edition and covers envelope estimation and kernel functions. Presentation of these topics is motivated by the growth in usage of these techniques. Envelope estimation is important not only for signals such as electromyograms but also when using high frequency carrier signals such as in ultrasound applications. The fundamentals of Hilbert transforms, analytic signals, and their estimation are
Kernel functions appear in the neuromuscular literature dealing with point processes such as action potentials. The main purpose is to create a continuous amplitude function from a point process. A summary of kernel functions and their implementation is presented.

The material in Chapter 10 is drawn with permission from the doctoral dissertation work of Robert Brychta and Melanie Bernard. They are both graduate students in Biomedical Engineering at Vanderbilt University. Robert’s research is being done in the General Clinical Research Center and Melanie’s is being done in the Visual System’s laboratory.

The presentation style is designed for the individual who wants a theoretical introduction to the basic principles and then the knowledge necessary to implement them practically. The mode of presentation is to: define a theoretical concept, show areas of engineering in which these concepts are useful, define the algorithms and assumptions needed to implement them, and then present detailed examples that have been implemented using FORTRAN and more recently MATLAB. The exposure to engineering applications will hopefully develop an appreciation for the utility and necessity of signal processing methodologies.

The exercises at the end of the chapters are designed with several goals. Some focus directly on the material presented and some extend the material for applications that are less often encountered. The degree of difficulty ranges from simple pencil and paper problems to computer implementation of simulations and algorithms for analysis. For an introductory course, the environment and software recommended are those that are not overly sophisticated and complex so that the student cannot comprehend the code or script. When used as a course textbook, most of the material can be studied in one semester in a senior undergraduate or first year graduate course. The topic selection is obviously the instructor’s choice.

Most of the examples and many of the exercises use measured signals, many from the biomedical domain. Copies of these are available from the publisher’s Website. Also available, for interactive learning, are a series of MATLAB notebooks that have been designed for interactive learning. These notebooks are written in the integrated environment of Microsoft Word and MATLAB. Each notebook presents a principle and demonstrates its implementation via script in MATLAB. The student is then asked to exercise other aspects of the principle interactively by making simple changes in the script. The student then receives immediate feedback concerning what is happening and can relate theoretical concepts to real effects upon a signal. The final one or two questions in the notebooks are more comprehensive and ask the student to make a full implementation of the technique or principle being studied. This requires understanding all of the previous material and selecting, altering, and then integrating parts of the MATLAB script previously used.

3 Http://books.elsevier.com/companions/9780120885817
This book is dedicated to my wife, Gloria, and to my parents who encouraged me and gave me the opportunity to be where I am today.
ACKNOWLEDGMENTS

The author of a textbook is usually helped significantly by the institution by which he is employed and through surrounding circumstances. In particular I am indebted to the Department of Biomedical Engineering and the School of Engineering at Vanderbilt University for giving me some released time and for placing a high priority on writing this book for academic purposes. This being the third edition, there have been three sets of reviewers. I would like to thank them because they have contributed to the book through suggestions of new topics and constructive criticism of the initial drafts. In addition, I am very grateful to Robert Brychta and Melanie Bernard, both graduate students in Biomedical Engineering at Vanderbilt University. Their doctoral research provided the basis for the topics in Chapter 10.
LIST OF SYMBOLS

ENGLISH

\(a(i), b(i)\), parameters of AR, MA, and ARMA models
\(A_m\), polynomial coefficient
\(B\), bandwidth
\(B_e\), equivalent bandwidth
\(c_x(k)\), sample covariance function
\(c_{xy}(k)\), sample cross covariance function
\(C_n\), coefficients of trigonometric Fourier series
\(C_x(k)\), autocovariance function
\(C_{xy}(k)\), cross covariance function
\(\text{Cov}[]\), covariance operator
\(d(n)\), data window
\(D(f)\), data spectral window
\(e_i\), error in polynomial curve fitting
\(E[]\), expectation operator
\(E_M\), sum of squared errors
\(E_{tot}\), total signal energy
\(f\), cyclic frequency
\(f_d\), frequency spacing
\(f_N\), folding frequency, highest frequency component
\(f_s\), sampling frequency
\(f(t)\), scaler function of variable t
**LIST OF SYMBOLS**

\( f_x(\alpha), f(x) \) probability density function
\( f_{xy}(\alpha, \beta), f(x, y) \) bivariate probability density function
\( F_x(\alpha), F(x) \) probability distribution function
\( F_{xy}(\alpha, \beta), F(x, y) \) bivariate probability distribution function
\( g \) loss coefficient
\( g_1 \) sample coefficient of skewness
\( h(t), h(n) \) impulse response
\( H(f), H(\omega) \) transfer function
\( I(f), I(m) \) periodogram
\( \text{Im}(\cdot) \) imaginary part of a complex function
\( \mathbb{N}[\cdot] \) imaginary operator
\( K^2(f), K^2(m) \) magnitude squared coherence function
\( L_r(x) \) Lagrange coefficient function
\( m \) mean
\( N \) number of points in a discrete time signal
\( p, q \) order of AR, MA, and ARMA processes
\( P \) signal power, or signal duration
\( P[\cdot] \) probability of [ ]
\( P_m(x) \) polynomial function
\( q-q \) quantile-quantile
\( r_Q \) correlation coefficient for q-q plot
\( \text{Re}(\cdot) \) real part of a complex function
\( R_x(k) \) autocorrelation function
\( R_{ys}(k) \) cross correlation function
\( \mathbb{N}[\cdot] \) real operator
\( s_p^2 \) variance of linear prediction error
\( S(f), S(m) \) power spectral density function
\( S_{ys}(f), S_{ys}(m) \) cross spectral density function
\( T \) sampling interval
\( U(t) \) unit step function
\( \text{Var}[\cdot] \) variance operator
\( w(k) \) lag window
\( W(f) \) lag spectral window
\( x(t), x(n) \) time function
\( X(f), X(m), X(\omega) \) Fourier transform
\( z_m \) coefficients of complex Fourier series

**GREEK**

\( \alpha \) significance level
\( \gamma_r(t_0, t_1), \gamma_{s(k)} \) ensemble autocovariance function
\( \gamma_1 \) coefficient of skewness
\( \delta(t) \) impulse function, dirac delta function
## LIST OF SYMBOLS

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \delta(n) )</td>
<td>unit impulse, Kronecker delta function</td>
</tr>
<tr>
<td>( \varepsilon(n) )</td>
<td>linear prediction error</td>
</tr>
<tr>
<td>( \lambda_i )</td>
<td>energy in a function</td>
</tr>
<tr>
<td>( \Lambda_{yy}(f) )</td>
<td>co-spectrum</td>
</tr>
<tr>
<td>( \mu_k )</td>
<td>kth central moment</td>
</tr>
<tr>
<td>( \eta(n) )</td>
<td>white noise process</td>
</tr>
<tr>
<td>( \xi(\tau) )</td>
<td>ensemble normalized autocovariance function</td>
</tr>
<tr>
<td>( \Xi )</td>
<td>Gaussian probability distribution function</td>
</tr>
<tr>
<td>( \rho )</td>
<td>correlation coefficient</td>
</tr>
<tr>
<td>( \rho_x(k) )</td>
<td>normalized autocovariance function</td>
</tr>
<tr>
<td>( \rho_{yx}(k) )</td>
<td>normalized cross covariance function</td>
</tr>
<tr>
<td>( \sigma^2 )</td>
<td>variance</td>
</tr>
<tr>
<td>( \sigma_e )</td>
<td>standard error of estimate</td>
</tr>
<tr>
<td>( \sigma_{xy}^2 )</td>
<td>covariance</td>
</tr>
<tr>
<td>( \phi(f) )</td>
<td>phase response</td>
</tr>
<tr>
<td>( \phi_{yx}(f), \phi_{yx}(m) )</td>
<td>cross phase spectrum</td>
</tr>
<tr>
<td>( \Phi_{\eta(t)} )</td>
<td>orthogonal function set</td>
</tr>
<tr>
<td>( \varphi_x(t_0, t_1), \varphi_x(k) )</td>
<td>ensemble autocorrelation function</td>
</tr>
<tr>
<td>( \Psi_{yx}(f) )</td>
<td>quadrature spectrum</td>
</tr>
<tr>
<td>( \omega )</td>
<td>radian frequency</td>
</tr>
<tr>
<td>( \omega_d )</td>
<td>radian frequency spacing</td>
</tr>
</tbody>
</table>

## ACRONYMS

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACF</td>
<td>autocorrelation function</td>
</tr>
<tr>
<td>ACVF</td>
<td>autocovariance function</td>
</tr>
<tr>
<td>AIC</td>
<td>Akaike’s information criterion</td>
</tr>
<tr>
<td>AR</td>
<td>autoregressive</td>
</tr>
<tr>
<td>ARMA</td>
<td>autoregressive-moving average</td>
</tr>
<tr>
<td>BT</td>
<td>Blackman-Tukey</td>
</tr>
<tr>
<td>CCF</td>
<td>cross correlation function</td>
</tr>
<tr>
<td>CCCF</td>
<td>cross covariance function</td>
</tr>
<tr>
<td>cdf</td>
<td>cumulative distribution function</td>
</tr>
<tr>
<td>CF</td>
<td>correlation function</td>
</tr>
<tr>
<td>CSD</td>
<td>cross spectral density</td>
</tr>
<tr>
<td>CTFT</td>
<td>continuous time Fourier transform</td>
</tr>
<tr>
<td>DFT</td>
<td>discrete Fourier transform</td>
</tr>
<tr>
<td>DTFT</td>
<td>discrete time Fourier transform</td>
</tr>
<tr>
<td>E</td>
<td>energy</td>
</tr>
<tr>
<td>erf</td>
<td>error function</td>
</tr>
<tr>
<td>FPE</td>
<td>final prediction error</td>
</tr>
<tr>
<td>FS</td>
<td>Fourier series</td>
</tr>
</tbody>
</table>
LIST OF SYMBOLS

IDFT \hspace{1em} \text{inverse discrete Fourier transform}
IDTFT \hspace{1em} \text{inverse discrete time Fourier transform}
LPC \hspace{1em} \text{linear prediction coefficient}
MA \hspace{1em} \text{moving average}
MEM \hspace{1em} \text{maximum entropy method}
MLE \hspace{1em} \text{maximum likelihood estimator}
MSC \hspace{1em} \text{magnitude squared coherence}
MSE \hspace{1em} \text{mean square error}
NACF \hspace{1em} \text{normalized autocovariance function}
NCCF \hspace{1em} \text{normalized cross covariance function}
pdf \hspace{1em} \text{probability density function}
PL \hspace{1em} \text{process loss}
PSD \hspace{1em} \text{power spectral density}
PW \hspace{1em} \text{power}
TSE \hspace{1em} \text{total square error}
VR \hspace{1em} \text{variance reduction}
WN \hspace{1em} \text{white noise}
YW \hspace{1em} \text{Yule-Walker}

OPERATORS

\begin{align*}
X(t)^* & \hspace{1em} \text{conjugation} \\
x(n) * y(n) & \hspace{1em} \text{convolution} \\
\hat{S}(m) & \hspace{1em} \text{sample estimate} \\
\tilde{S}(m) & \hspace{1em} \text{smoothing} \\
\bar{x}(n) & \hspace{1em} \text{periodic repetition}
\end{align*}

FUNCTIONS

\begin{align*}
\text{sgn}(x) & = 1, \quad x > 0 \\
& = -1, \quad x < 0
\end{align*}