

Statistical Modeling by Wavelets

Brani Vidakovic

A Wiley-Interscience Publication

JOHN WILEY & SONS, INC.

New York / Chichester / Weinheim / Brisbane / Singapore / Toronto

Contents

<i>Preface</i>	<i>xi</i>
<i>Acknowledgments</i>	<i>xiii</i>
<i>1 Introduction</i>	<i>1</i>
1.1 <i>Wavelet Evolution</i>	<i>1</i>
1.2 <i>Wavelet Revolution</i>	<i>6</i>
1.3 <i>Wavelets and Statistics</i>	<i>16</i>
1.4 <i>An Appetizer: California Earthquakes</i>	<i>19</i>
<i>2 Prerequisites</i>	<i>23</i>
2.1 <i>General</i>	<i>23</i>
2.2 <i>Hilbert Spaces</i>	<i>24</i>
2.2.1 <i>Projection Theorem</i>	<i>25</i>
2.2.2 <i>Orthonormal Sets</i>	<i>26</i>
2.2.3 <i>Reproducing Kernel Hilbert Spaces</i>	<i>29</i>
2.3 <i>Fourier Transformation</i>	<i>29</i>
2.3.1 <i>Basic Properties</i>	<i>30</i>
2.3.2 <i>Poisson Summation Formula and Sampling Theorem</i>	<i>32</i>

vi CONTENTS

2.3.3	<i>Fourier Series</i>	33
2.3.4	<i>Discrete Fourier Transform</i>	34
2.4	<i>Heisenberg's Uncertainty Principle</i>	35
2.5	<i>Some Important Function Spaces</i>	36
2.6	<i>Fundamentals of Signal Processing</i>	39
2.7	<i>Exercises</i>	40
3	<i>Wavelets</i>	43
3.1	<i>Continuous Wavelet Transformation</i>	43
3.1.1	<i>Basic Properties</i>	44
3.1.2	<i>Wavelets for Continuous Transformations</i>	48
3.2	<i>Discretization of the Continuous Wavelet Transform</i>	50
3.3	<i>Multiresolution Analysis</i>	51
3.3.1	<i>Derivation of a Wavelet Function</i>	57
3.4	<i>Some Important Wavelet Bases</i>	60
3.4.1	<i>Haar's Wavelets</i>	60
3.4.2	<i>Shannon's Wavelets</i>	63
3.4.3	<i>Meyer's Wavelets</i>	65
3.4.4	<i>Franklin's Wavelets</i>	70
3.4.5	<i>Daubechies' Compactly Supported Wavelets</i>	73
3.5	<i>Some Extensions</i>	80
3.5.1	<i>Regularity of Wavelets</i>	80
3.5.2	<i>The Least Asymmetric Daubechies' Wavelets: Symmlets</i>	86
3.5.3	<i>Approximations and Characterizations of Functional Spaces</i>	87
3.5.4	<i>Daubechies-Lagarias Algorithm</i>	89
3.5.5	<i>Moment Conditions</i>	91
3.5.6	<i>Interpolating (Cardinal) Wavelets</i>	93
3.5.7	<i>Pollen-Type Parameterization of Wavelets</i>	94
3.6	<i>Exercises</i>	96
4	<i>Discrete Wavelet Transformations</i>	101
4.1	<i>Introduction</i>	101
4.2	<i>The Cascade Algorithm</i>	104
4.3	<i>The Operator Notation of DWT</i>	109
4.3.1	<i>Discrete Wavelet Transformations as Linear Transformations</i>	115
4.4	<i>Exercises</i>	117

5	<i>Some Generalizations</i>	119
5.1	<i>Coiflets</i>	120
5.1.1	<i>Construction of Coiflets</i>	121
5.2	<i>Biorthogonal Wavelets</i>	122
5.2.1	<i>Construction of Biorthogonal Wavelets</i>	125
5.2.2	<i>B-Spline Wavelets</i>	128
5.3	<i>Wavelet Packets</i>	133
5.3.1	<i>Basic Properties of Wavelet Packets</i>	135
5.3.2	<i>Wavelet Packet Tables</i>	138
5.4	<i>Best Basis Selection</i>	140
5.4.1	<i>Some Cost Measures and the Best Basis Algorithm</i>	141
5.5	<i>ϵ-Decimated and Stationary Wavelet Transformations</i>	145
5.5.1	<i>ϵ-Decimated Wavelet Transformation</i>	145
5.5.2	<i>Stationary (Non-Decimated) Wavelet Transformation</i>	147
5.6	<i>Periodic Wavelet Transformations</i>	150
5.7	<i>Multivariate Wavelet Transformations</i>	153
5.8	<i>Discussion</i>	160
5.9	<i>Exercises</i>	163
6	<i>Wavelet Shrinkage</i>	167
6.1	<i>Shrinkage Method</i>	168
6.2	<i>Linear Wavelet Regression Estimators</i>	170
6.2.1	<i>Wavelet Kernels</i>	171
6.2.2	<i>Local Constant Fit Estimators</i>	172
6.3	<i>The Simplest Non-Linear Wavelet Shrinkage: Thresholding</i>	175
6.3.1	<i>Variable Selection and Thresholding</i>	177
6.3.2	<i>Oracular Risk for Thresholding Rules</i>	177
6.3.3	<i>Why the Wavelet Shrinkage Works</i>	179
6.3.4	<i>Almost Sure Convergence of Wavelet Shrinkage Estimators</i>	184
6.4	<i>General Minimax Paradigm</i>	184
6.4.1	<i>Translation of Minimaxity Results to the Wavelet Domain</i>	186
6.5	<i>Thresholding Policies and Thresholding Rules</i>	188
6.5.1	<i>Exact Risk Analysis of Thresholding Rules</i>	188
6.5.2	<i>Large Sample Properties of \hat{f}</i>	189

6.5.3	<i>Some Other Shrinkage Rules</i>	191
6.6	<i>How to Select a Threshold</i>	194
6.6.1	<i>Mallat's Model and Induced Percentile Thresholding</i>	194
6.6.2	<i>Universal Threshold</i>	195
6.6.3	<i>A Threshold Based on Stein's Unbiased Estimator of Risk</i>	199
6.6.4	<i>Cross-Validation</i>	200
6.6.5	<i>Thresholding as a Testing Problem</i>	202
6.6.6	<i>Lorentz Curve Thresholding</i>	203
6.6.7	<i>Block Thresholding Estimators</i>	206
6.7	<i>Other Methods and References</i>	207
6.8	<i>Exercises</i>	213
7	<i>Density Estimation</i>	217
7.1	<i>Orthogonal Series Density Estimators</i>	217
7.2	<i>Wavelet Density Estimation</i>	219
7.2.1	<i>δ-Sequence Density Estimators</i>	219
7.2.2	<i>Bias and Variance of Linear Wavelet Density Estimators</i>	223
7.2.3	<i>Linear Wavelet Density Estimators in a More General Setting</i>	224
7.3	<i>Non-Linear Wavelet Density Estimators</i>	225
7.3.1	<i>Global Thresholding Estimator</i>	227
7.4	<i>Non-Negative Density Estimators</i>	228
7.4.1	<i>Estimating the Square Root of a Density</i>	229
7.4.2	<i>Density Estimation by Non-Negative Wavelets</i>	234
7.5	<i>Other Methods</i>	236
7.5.1	<i>Multivariate Wavelet Density Estimators</i>	236
7.5.2	<i>Density Estimation as a Regression Problem</i>	239
7.5.3	<i>Cross-Validation Estimator</i>	240
7.5.4	<i>Multiscale Estimator</i>	240
7.5.5	<i>Estimation of a Derivative of a Density</i>	241
7.6	<i>Exercises</i>	243
8	<i>Bayesian Methods in Wavelets</i>	247
8.1	<i>Motivational Examples</i>	247
8.2	<i>Smooth Shrinkage</i>	250
8.3	<i>Bayesian Thresholding</i>	255

8.4	<i>MAP-Principle</i>	257
8.5	<i>Density Estimation Problem</i>	259
8.6	<i>Full Bayesian Model</i>	262
8.7	<i>Discussion and References</i>	265
8.8	<i>Exercises</i>	267
9	<i>Wavelets and Random Processes</i>	271
9.1	<i>Stationary Time Series</i>	272
9.2	<i>Wavelets and Stationary Processes</i>	273
9.2.1	<i>Wavelet Transformations of Stationary Processes</i>	273
9.2.2	<i>Whitening of Stationary Processes</i>	274
9.2.3	<i>Karhunen-Loève-Like Expansions</i>	276
9.3	<i>Estimation of Spectral Densities</i>	279
9.3.1	<i>Gao's Algorithm</i>	280
9.3.2	<i>Non-Gaussian Stationary Processes</i>	284
9.4	<i>Wavelet Spectrum</i>	286
9.4.1	<i>Wavelet Spectrum of a Stationary Time Series</i>	286
9.4.2	<i>Scalogram and Periodicities</i>	289
9.5	<i>Long-Memory Processes</i>	290
9.5.1	<i>Wavelets and Fractional Brownian Motion</i>	290
9.5.2	<i>Estimating Spectral Exponents in Self-Similar Processes</i>	292
9.5.3	<i>Quantifying the Whitening Property of Wavelet Transformations for fBm Processes</i>	295
9.6	<i>Discussion and References</i>	296
9.7	<i>Exercises</i>	297
10	<i>Wavelet-Based Random Variables and Densities</i>	299
10.1	<i>Scaling Function as a Density</i>	299
10.2	<i>Wavelet-Based Random Variables</i>	300
10.3	<i>Random Densities via Wavelets</i>	308
10.3.1	<i>Tree Algorithm</i>	308
10.4	<i>Properties of Wavelet-Based Random Densities</i>	310
10.5	<i>Random Densities With Constraints</i>	312
10.5.1	<i>Smoothness Constraints</i>	312
10.5.2	<i>Constraints on Symmetry</i>	312
10.5.3	<i>Constraints on Modality</i>	313
10.5.4	<i>Skewed Random Densities</i>	313

x CONTENTS

10.6 Exercises	316
11 Miscellaneous Statistical Applications	319
11.1 Deconvolution Problems	319
11.2 Wavelet-Vaguelette Decompositions	323
11.3 Pursuit Methods	325
11.4 Moments of Order Statistics	329
11.5 Wavelets and Statistical Turbulence	333
11.5.1 K41 Theory	335
11.5.2 Townsend's Decompositions	336
11.6 Software and WWW Resources for Wavelet Analysis	338
11.6.1 Commercial Wavelet Software	338
11.6.2 Free Wavelet Software	340
11.6.3 Some WWW Resources	341
11.7 Exercises	342
References	345
Notation Index	370
Author Index	373
Subject Index	379

Preface

Just two months ago astronomers did not know about it. But now they are giving good odds that Hyakutake will be the most impressive comet since the invention of telescope 400 years ago. (*Herald Sun, Durham, NC, March 24, 1996.*)

One can trace the origins of wavelets back to the beginning of this century; however, wavelets, understood as a systematic way of producing local orthogonal bases, are a recent unification of existing theories in various fields and some important “discoveries.” They are mathematical objects that have interpretation and application in many scientific fields, most notably in the fields of signal processing, nonparametric function estimation, and data compression. In the early 1990s, a series of papers by Donoho and Johnstone and their coauthors demonstrated that wavelets are appropriate tools in problems of denoising, regression, and density estimation. The subsequent burgeoning wavelet research broadened to a wide range of statistical problems.

This book is aimed at graduate students in statistics and mathematics, practicing statisticians, and statistically curious engineers. It can serve as a text for an introductory wavelet course concerned with an interface of wavelet methods and statistical inference. The necessary mathematical background is proficiency in advanced calculus and algebra; consequently, this book should be useful to advanced undergraduate students as well as to graduate students in statistics, mathematics, and engineering.

This book originated from the class notes supporting the Special Topics Course on Multiscale Methods at Duke University. The content can be divided into two parts:

an introduction to wavelets (Chapters 1–5) and statistical modeling (Chapters 6–11). An introduction and some mathematical prerequisites are presented in Chapters 1 and 2. Continuous and discrete wavelet transformations are covered in Chapters 3 and 4. Some important generalizations (coiflets, biorthogonal wavelets, wavelet packets, stationary, periodized and multivariate wavelets) are covered in Chapter 5.

Chapters 6–11 are data-oriented. Chapter 6 is the crux of the book, covering the theory and practice of wavelet shrinkage. Important theoretical aspects of wavelet density estimation are covered in Chapter 7. Chapter 8 discusses Bayesian modeling in the wavelet domain. Time series are covered in Chapter 9, while Chapter 10 contains several probabilistic and simulational properties of wavelet-based random functions and densities. Chapter 11 gives some novel and important wavelet applications in statistics.

Instead of providing appendices with data sets and programs used in the book, I opted for a more modern style. The web page:

<http://www.isds.duke.edu/~brani/wiley.html>

is associated with the book. This page contains all data sets, functions, and programs referred to.

I hope the reader will find this book useful. All comments, suggestions, updates, and critiques will be appreciated.

BRANI VIDA KOVIC

*Institute of Statistics and Decision Sciences
Duke University
Durham, February 1999*

Acknowledgments

This book was made possible by the support of many individuals. First, I am grateful to colleagues from the Institute of Statistics and Decision Sciences (ISDS) – Duke University for supporting a wavelet-based statistics course. The attending students were patient and understood the difficulties of transferring research papers to a working copy of course notes. Their industry and enthusiasm is treasured. Duke University helped with this project through the Grant of Arts and Sciences Council 1997, and partial support was provided by the National Science Foundation Award DMS-9626159 at Duke University. Figure 1.10 is reproduced with the permission of the Salvador Dalí Museum, Inc. in St. Petersburg, Florida.

Many colleagues contributed to this project in different ways: Anestis Antoniadis, Tony Cai, Merlise Clyde, Lubo Dechevsky, Iain Johnstone, Gabriel Katul, Eric Kolarczyk, Pedro Morettin, Peter Müller, Giovanni Parmigiani, Marianna Pensky, David Rios, Fabrizio Ruggeri, Rainer von-Sacks, Naoki Saito, Yazhen Wang, and Gilbert Walter, to list a few. Collaboration with software gurus Hong-Ye Gao [TeraLogic Inc.] and Andrew Bruce [MathSoft Inc.] was fruitful. The S+Wavelets module (for S-Plus) was used for almost all of the computer examples, figures, and calculations. I am grateful to Alison Bory, Angiolina Loreda, and Steve Quigley from Wiley, for their enthusiastic assistance, and to Courtney Johnson, Michael Kozdron, and Kathy Zhou, doctoral students at Duke University, for their help in proofreading the manuscript.

And most of all, I am grateful to my family for their love and strong and continuous support.