

Lecture Notes in Statistics

129

**Wolfgang Härdle Gerard Kerkycharian
Dominique Picard Alexander Tsybakov**

Wavelets, Approximation, and Statistical Applications



seminaire Paris-Berlin
Seminar Berlin-Paris



Springer

Wolfgang Härdle
Gerard Kerkyacharian
Dominique Picard
Alexander Tsybakov

Wavelets, Approximation, and Statistical Applications



Springer

W. Härdle
Humboldt-Universität zu Berlin
Wirtschaftswissenschaftliche Fakultät
Institut für Statistik und Ökonometrie
Spandauer Strasse 1
D 10178 Berlin

G. Kerkycharian
Université Paris X
URA CNRS 1321 Modal
200, av. De la République
92001 Nanterre Cedex
France

D. Picard
Université Paris VII
UFR Mathématique
URA CNRS 1321
2, Place Jussieu
F 75252 Paris cedex 5
France

A. B. Tsybakov
Université Paris VI
Laboratoire de Probabilités
B.P. 188, 4, Place Jussieu
F 75252 Paris cedex 5
France

Library of Congress Cataloging-in-Publication Data

Wavelets, approximation, and statistical applications / Wolfgang

Härdle ... [et al.].

p. cm. -- (Lecture notes in statistics ; 129)

Includes bibliographical references and indexes.

ISBN 0-387-98453-4 (softcover : alk. paper)

1. Wavelets (Mathematics) 2. Approximation theory.

3. Nonparametric statistics. I. Härdle, Wolfgang. II. Series:

Lecture notes in statistics (Springer-Verlag) ; v. 129.

QA403.3.W363 1998

515'.2433--dc21

97-48855

Printed on acid-free paper.

© 1998 Springer-Verlag New York, Inc.

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer-Verlag New York, Inc., 175 Fifth Avenue, New York, NY 10010, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use of general descriptive names, trade names, trademarks, etc., in this publication, even if the former are not especially identified, is not to be taken as a sign that such names, as understood by the Trade Marks and Merchandise Marks Act, may accordingly be used freely by anyone.

Camera ready copy provided by the authors.

Printed and bound by Braun-Brumfield, Ann Arbor, MI.

Printed in the United States of America.

9 8 7 6 5 4 3 2 1

ISBN 0-387- 98453-4 Springer-Verlag New York Berlin Heidelberg SPIN 10659526

Lecture Notes in Statistics

129

Edited by P. Bickel, P. Diggle, S. Fienberg, K. Krickeberg,
I. Olkin, N. Wermuth, S. Zeger

Springer

New York

Berlin

Heidelberg

Barcelona

Budapest

Hong Kong

London

Milan

Paris

Santa Clara

Singapore

Tokyo

Preface

The mathematical theory of *ondelettes* (wavelets) was developed by Yves Meyer and many collaborators about 10 years ago. It was designed for approximation of possibly irregular functions and surfaces and was successfully applied in data compression, turbulence analysis, image and signal processing. Five years ago wavelet theory progressively appeared to be a powerful framework for nonparametric statistical problems. Efficient computational implementations are beginning to surface in this second lustrum of the nineties. This book brings together these three main streams of wavelet theory. It presents the theory, discusses approximations and gives a variety of statistical applications. It is the aim of this text to introduce the novice in this field into the various aspects of wavelets. Wavelets require a highly interactive computing interface. We present therefore all applications with software code from an interactive statistical computing environment.

Readers interested in theory and construction of wavelets will find here in a condensed form results that are somewhat scattered around in the research literature. A practitioner will be able to use wavelets via the available software code. We hope therefore to address both theory and practice with this book and thus help to construct bridges between the different groups of scientists.

This text grew out of a French-German cooperation (*Séminaire Paris-Berlin, Seminar Berlin-Paris*). This seminar brings together theoretical and applied statisticians from Berlin and Paris. This work originates in the first of these seminars organized in Garchy, Burgundy in 1994. We are confident that there will be future research work originating from this yearly seminar.

This text would not have been possible without discussion and encouragement from colleagues in France and Germany. We would like to thank in particular Lucien Birgé, Christian Gourieroux, Yuri Golubev, Marc Hoffmann, Sylvie Huet, Emmanuel Jolivet, Oleg Lepski, Enno Mammen, Pascal

Massart, Michael Nussbaum, Michael Neumann, Volodja Spokoiny, Karine Tribouley. The help of Yuri Golubev was particularly important. Our Sections 11.5 and 12.5 are inspired by the notes that he kindly provided. The implementation in XploRe was professionally arranged by Sigbert Klinke and Clementine Dalelane. Steve Marron has established a fine set of test functions that we used in the simulations. Michael Kohler and Marc Hoffmann made many useful remarks that helped in improving the presentation. We had strong help in designing and applying our \LaTeX macros from Wolfram Kempe, Anja Bardeleben, Michaela Draganska, Andrea Tiersch and Kerstin Zanter. Un très grand merci!

Berlin-Paris, September 1997

Wolfgang Härdle
Gerard Kerkycharian,
Dominique Picard
Alexander Tsybakov

Symbols and Notation

φ	father wavelet
ψ	mother wavelet
$S1, S2, \dots$	symmlets
$D1, D2, \dots$	Daubechies wavelets
$C1, C2, \dots$	Coiflets
ISE	integrated squared error
$MISE$	mean integrated squared error
\mathbb{R}	the real line
\mathbb{Z}	set of all integers in \mathbb{R}
l_p	space of p -summable sequences
$L_p(\mathbb{R})$	space of p -integrable functions
$W_p^m(\mathbb{R})$	Sobolev space
$B_p^{sq}(\mathbb{R})$	Besov space
$\mathcal{D}(\mathbb{R})$	space of infinitely many times differentiable compactly supported functions
$S'(\mathbb{R})$	Schwartz space
H^λ	Hölder smoothness class with parameter λ
(f, g)	scalar product in $L_2(\mathbb{R})$
$\ f\ _p$	norm in $L_p(\mathbb{R})$
$\ a\ _{l_p}$	norm in l_p
$\ f\ _{spq}$	norm in $B_p^{sq}(\mathbb{R})$
ONS	orthonormal system
ONB	orthonormal basis
MRA	multiresolution analysis
RHS	right hand side
LHS	left hand side
DWT	discrete wavelet transform

$f * g$	convolution of f and g
$I\{A\}$	indicator function of a set A
a.e.	almost everywhere
$\text{supp } f$	support of function f
ess sup	essential supremum
$f^{(m)}$	m -th derivative
$\tau_h f(x) = f(x - h)$	shift operator
$\omega_p^1(f, t)$	modulus of continuity in the L_p norm
$K(x, y)$	kernel
δ_{jk}	Kronecker's delta
\simeq	asymptotic identical rate
\sum_k	sum over all $k \in \mathbb{Z}$
$\text{card } \Omega$	cardinality of a set Ω

Contents

Preface	v
List of Figures	xi
List of Tables	xv
Symbols and Notation	xvii
1 Wavelets	1
1.1 What can wavelets offer?	1
1.2 General remarks	7
1.3 Data compression	8
1.4 Local adaptivity	13
1.5 Nonlinear smoothing properties	15
1.6 Synopsis	15
2 The Haar basis wavelet system	17
3 The idea of multiresolution analysis	25
3.1 Multiresolution analysis	25
3.2 Wavelet system construction	27
3.3 An example	29
4 Some facts from Fourier analysis	31

5 Basic relations of wavelet theory	35
5.1 When do we have a wavelet expansion?	35
5.2 How to construct mothers from a father	42
5.3 Additional remarks	44
6 Construction of wavelet bases	47
6.1 Construction starting from Riesz bases	47
6.2 Construction starting from m_0	54
7 Compactly supported wavelets	59
7.1 Daubechies' construction	59
7.2 Coiflets	63
7.3 Symmlets	66
8 Wavelets and Approximation	71
8.1 Introduction	71
8.2 Sobolev Spaces	72
8.3 Approximation kernels	75
8.4 Approximation theorem in Sobolev spaces	76
8.5 Periodic kernels and projection operators	81
8.6 Moment condition for projection kernels	84
8.7 Moment condition in the wavelet case	90
9 Wavelets and Besov Spaces	101
9.1 Introduction	101
9.2 Besov spaces	101
9.3 Littlewood-Paley decomposition	106
9.4 Approximation theorem in Besov spaces	116
9.5 Wavelets and approximation in Besov spaces	117
10 Statistical estimation using wavelets	125
10.1 Introduction	125
10.2 Linear wavelet density estimation	126
10.3 Soft and hard thresholding	138
10.4 Linear versus nonlinear wavelet density estimation	146
10.5 Asymptotic properties of wavelet thresholding estimates . . .	163
10.6 Some real data examples	171

10.7 Comparison with kernel estimates	176
10.8 Regression estimation	183
10.9 Other statistical models	188
11 Wavelet thresholding and adaptation	193
11.1 Introduction	193
11.2 Different forms of wavelet thresholding	193
11.3 Adaptivity properties of wavelet estimates	198
11.4 Thresholding in sequence space	201
11.5 Adaptive thresholding and Stein's principle	205
11.6 Oracle inequalities	210
11.7 Bibliographic remarks	212
12 Computational aspects and software	215
12.1 Introduction	215
12.2 The cascade algorithm	216
12.3 Discrete wavelet transform	220
12.4 Statistical implementation of the DWT	222
12.5 Translation invariant wavelet estimation	228
12.6 Main wavelet commands in XploRe	231
A Tables	237
A.1 Wavelet Coefficients	237
A.2	239
B Software Availability	240
C Bernstein and Rosenthal inequalities	241
D A Lemma on the Riesz basis	246
Bibliography	248
Subject Index	260
Author Index	263

List of Figures

1.1	Bid-Ask spreads for one year of the DEM-USD FX-rate.	2
1.2	The first quarter of the DEM-USD FX rate.	2
1.3	The first 1024 points (2 weeks) of the DEM-USD FX rate with a location - frequency plot.	4
1.4	Distribution of coefficients for weeks 3-4.	4
1.5	The first 2 weeks of the YENDEM FX-rate.	5
1.6	The weeks 3 - 4 of the YENDEM FX-rate.	5
1.7	The smoothed periodogram of the YENDEM series.	6
1.8	Binned Belgian household data at x -axis. Wavelet density (solid) and kernel density (dashed).	6
1.9	Sine and cosine waves and wavelets (D10).	9
1.10	The step function and the Fourier series approximation with 5 terms.	9
1.11	The step function and the Fourier series with 50 terms.	11
1.12	Two waves with different frequency.	11
1.13	Location - frequency plot for the curve in Figure 1.12	12
1.14	The wavelet approximation (with its location - frequency plot) for the curve of Figure 1.13	12
1.15	The seminar label computed from 500 coefficients.	14
2.1	The sine example with a coarse Haar approximation.	23
6.1	The first 2 elements of the B-spline Riesz basis.	50
6.2	Battle-Lemarié father wavelet ($N=2$).	53
6.3	Battle-Lemarié mother wavelet ($N=2$).	54
7.1	Daubechies' wavelets $D2-D8$	64
7.2	Coiflets in order $C1$ to $C4$	67
7.3	Four symmlets $S4-S7$	68
9.1	Typical example of the Fourier transforms $\hat{\gamma}$, $A = 1$	108

9.2	Fourier transform $\hat{\beta}$, $A = 1$	109
10.1	Uniform mixture random variables ($n = 500$) with density and a Haar wavelet estimate with $j_1 = 1$	136
10.2	The same variables as in Figure 10.1 and a Haar wavelet estimate with $j_1 = 2$	136
10.3	The same variables as in Figure 10.1 and a Haar wavelet density estimate with $j_1 = 4$	137
10.4	Haar wavelet density estimate with $j_1 = 6$	137
10.5	Soft and hard thresholding.	139
10.6	A sample of $n = 500$ points with uniform mixture density and a Haar wavelet density estimate. The hard threshold was set to $0.4 \max_{j,k} \hat{\beta}_{jk} $	140
10.7	A sample of $n = 500$ points with density and a Haar wavelet density estimate. The hard threshold was set to $0.6 \max_{j,k} \hat{\beta}_{jk} $	141
10.8	A sample of $n = 500$ points with density and a Haar wavelet density estimate. The hard threshold was set to $0.8 \max_{j,k} \hat{\beta}_{jk} $	142
10.9	Soft thresholding with data from Figure 10.6. Threshold value 0.8.	143
10.10	Soft thresholding with data from Figure 10.6. Threshold value 0.6.	144
10.11	Soft thresholding with data from Figure 10.6. Threshold value 0.4.	145
10.12	Translation invariant hard thresholding with data from Figure 10.6. Adaptive Threshold chosen.	147
10.13	Classification of optimal rates of convergence for linear and non-linear estimates.	153
10.14	Density estimate of IBM returns. Soft thresholding, $t = 0.6 \max_{j,k} \hat{\beta}_{jk}$	173
10.15	Density estimate of equally weighted portfolio. Soft thresholding, $t = 0.6 \max_{j,k} \hat{\beta}_{jk} $	174
10.16	A comparison of density estimates.	175
10.17	FES Income densities 1969-1972.	176
10.18	FES Income densities 1973-1976.	177
10.19	FES Income densities 1977-1980.	177
10.20	FES Income densities 1981-1983, 1969-1983.	178
10.21	A trimodal density and $n = 500$ data points.	179
10.22	The density with two kernel density estimates.	180
10.23	The density, a kernel estimate and a wavelet estimate with hard thresholding ($S4$, $j_1 = 8$, $t = 0.4 \max_{j,k} \hat{\beta}_{jk} $).	181
10.24	The density, a kernel estimate and a wavelet estimate with soft thresholding ($S4$, $j_1 = 8$, $t = 0.4 \max_{j,k} \hat{\beta}_{jk} $).	182

10.25 Data and regression curve. 185

10.26 Linear wavelet estimator and true curve, $j_1 = 8$ 186

10.27 Linear wavelet estimator and true curve, with $j_1 = 5$ 186

10.28 Wavelet smoother with soft threshold $0.2 \max_{jk} |\hat{\beta}_{jk}|$ 187

10.29 Wavelet regression with boundary effect. 187

10.30 Wavelet regression estimator after mirroring. 189

11.1 Regression function and the noisy data. 211

11.2 SURE regression estimator and the regression function. 211

12.1 ISE for $S8$ as a function of bin size 224

12.2 Translation invariant density estimation with $S8$ and hard threshold
 $0.25 \max |\hat{\beta}_{jk}|$ 232

List of Tables

10.1	ISE values for different density estimates	135
10.2	<i>ISE</i> for different threshold values, $j_1 = 8$, Haar wavelet.	146
10.3	Different bandwidth selectors for data of Figure 10.21	179
12.1	ISE values for different bin sizes	223
12.2	The coefficients for $S7$	233
A1.	The 296 coefficients for the wavelet construction.	237
A2.	The indices for the selected wavelets.	239

Chapter 1

Wavelets

1.1 What can wavelets offer?

A wavelet is, as the name suggests, a small wave. Many statistical phenomena have wavelet structure. Often small bursts of high frequency wavelets are followed by lower frequency waves or vice versa. The theory of wavelet reconstruction helps to localize and identify such accumulations of small waves and helps thus to better understand reasons for these phenomena. Wavelet theory is different from Fourier analysis and spectral theory since it is based on a local frequency representation.

Let us start with some illustrative examples of wavelet analysis for financial time series data. Figure 1.1 shows the time series of $25434 \log(\text{ask}) - \log(\text{bid})$ spreads of the DeutschMark (DEM) - USDollar (USD) exchange rates during the time period of October 1, 1992 to September 30, 1993. The series consists of offers (bids) and demands (asks) that appeared on the FXFX page of the Reuters network over the entire year, see Bossaerts, Hafner & Härdle (1996), Ghysels, Gouriéroux & Jasiak (1995). The graph shows the bid - ask spreads for each quarter of the year on the vertical axis. The horizontal axis denotes time for each quarter.

The quarterly time series show local bursts of different size and frequency. Figure 1.2 is a zoom of the first quarter. One sees that the bid-ask spread varies dominantly between 2 - 3 levels, has asymmetric behavior with thin but high rare peaks to the top and more oscillations downwards. Wavelets provide a way to quantify this phenomenon and thereby help to detect mechanisms