

Applied Nonparametric Regression is the first book to bring together in one place the techniques for regression curve smoothing involving more than one variable. The computer and the development of interactive graphics programs have made curve estimation popular. This volume focuses on the applications and practical problems of two central aspects of curve smoothing – the choice of smoothing parameters and the construction of confidence bounds.

Härdle argues that all smoothing methods are based on a local averaging mechanism, and thus can be seen as essentially equivalent to kernel smoothing. To simplify the exposition kernel smoothers are introduced and discussed in great detail. Building on this exposition various other smoothing methods (among them splines and orthogonal polynomials) are presented and their merits discussed. All the methods presented can be understood on an intuitive level; however, exercises and supplemental materials are provided for those readers desiring a deeper understanding of the techniques.

The methods covered in this text have numerous applications in many areas using statistical analysis. Motivating examples stem from economics such as the estimation of Engel curves as well as biomedical and engineering problems. For practical applications of the methods a computing environment for eXploratory Regression – XploRe – is described.

“Nonparametric regression analysis has become central to economic theory. Härdle, by writing the first comprehensive and accessible book on the subject, has contributed enormously to making nonparametric regression equally central to econometric practice.” –Charles F. Manski, University of Wisconsin, Madison

“This book represents an optimally estimated common thread for the numerous topics and results in the fast-growing area of nonparametric regression. The user-friendly approach taken by the author has successfully smoothed out most of the formidable asymptotic elaboration in developing the theory. This is an excellent collection for both beginners and experts.”

–Ker-Chau Li, University of California, Los Angeles

“This monograph on nonparametric regression presents a particularly clear and balanced view of the methodology and practice of this very important subject, and so is of use to theoreticians and practitioners alike.”

–Peter Hall, University of Glasgow

Cambridge
University
Press

ISBN 0-521-4295



9 780521 429511

Contents

<i>Preface</i>	page xi
<i>Symbols and notation</i>	xiv
 PART I. Regression smoothing	
Chapter 1. Introduction	3
1.1 Motivation	6
1.2 Scope of this book	12
Chapter 2. Basic idea of smoothing	14
2.1 The stochastic nature of the observations	21
2.2 Hurdles for the smoothing process	22
Complements	23
Chapter 3. Smoothing techniques	24
3.1 Kernel smoothing	24
Kernel estimators are local polynomial fits	30
Kernel estimators of derivatives	32
Computational aspects of kernel smoothing	35
Exercises	37
Complements	39
Proof of Proposition 3.1.1	39
Sketch of Proof for Proposition 3.1.2	41
3.2 k -nearest neighbor estimates	42
Computational aspects of k -NN smoothing	48
Exercises	49
3.3 Orthogonal series estimators	50
Exercises	55
Complements	55
Sketch of Proof for Proposition 3.3.1	55
3.4 Spline smoothing	56
Exercises	64

viii	Contents	
	Complements	64
	Proof of Proposition 3.4.2	64
	Proof of equivalence of the Silverman and Huber approximation	64
3.5	An overview of various smoothers	65
	Recursive techniques	66
	The regressogram	67
	Convolution smoothing	68
	Delta function sequence estimators	68
	Median smoothing	69
	Split linear fits	70
	Empirical regression	72
	Exercises	73
	Complements	74
3.6	A comparison of kernel, k -NN and spline smoothers	75
	Exercises	84
	Complements	85
	PART II. The kernel method	87
	Chapter 4. How close is the smooth to the true curve?	89
4.1	The speed at which the smooth curve converges	91
	Rates of convergence for nonquadratic distance measures	95
	Exercises	96
	Complements	97
4.2	Pointwise confidence intervals	98
	Exercises	108
	Complements	109
4.3	Variability bands for functions	110
	Connected error bars	111
	Smooth confidence bands	114
	Bootstrap bands	118
	Exercises	128
	Complements	129
4.4	Behavior at the boundary	130
	Exercises	132
4.5	The accuracy as a function of the kernel	133
	Exercises	137
	Complements	140
4.6	Bias reduction techniques	141
	Exercises	146

Chapter 5. Choosing the smoothing parameter	147
5.1 Cross-validation, penalizing functions and the plug-in method	148
Leave-one-out method, cross-validation	152
Penalizing functions	154
The plug-in method	155
Bandwidth choice for derivative estimation	160
The dependence of the smoothing parameter on the weight function	162
Exercises	162
Complements	163
5.2 Which selector should be used?	165
Exercises	173
Complements	175
5.3 Local adaptation of the smoothing parameter	177
Improving the smooth locally by bootstrapping	177
The supersmoother	181
Exercises	184
5.4 Comparing bandwidths between laboratories (canonical kernels)	184
Exercises	189
Chapter 6. Data sets with outliers	190
6.1 Resistant smoothing techniques	191
LOcally WEighted Scatter plot Smoothing (LOWESS)	192
<i>L</i> -smoothing	193
<i>R</i> -smoothing	194
<i>M</i> -smoothing	195
Exercises	200
Complements	201
Chapter 7. Nonparametric regression techniques for correlated data	203
7.1 Nonparametric prediction of time series	204
7.2 Smoothing with dependent errors	209
Exercises	215
Complements	215
Chapter 8. Looking for special features and qualitative smoothing	217
8.1 Monotonic and unimodal smoothing	218
Pool-adjacent-violators (PAV) Algorithm-8.1.1	218

x	Contents	
	Exercises	223
	Complements	224
8.2	Estimation of zeros and extrema	225
	Exercises	230
	Complements	230
Chapter 9.	Incorporating parametric components	232
9.1	Partial linear models	233
	Exercises	236
9.2	Shape-invariant modeling	237
	Exercises	243
	Complements	243
9.3	Comparing nonparametric and parametric curves	244
	Exercises	249
	Complements	252
PART III.	Smoothing in high dimensions	255
Chapter 10.	Investigating multiple regression by additive models	257
10.1	Regression trees	259
	Exercises	264
10.2	Projection pursuit regression	265
	Exercises	268
10.3	Alternating conditional expectations	269
	Exercises	273
10.4	Average derivative estimation	274
	Exercises	280
10.5	Generalized additive models	281
	Exercises	286
Appendix 1.	A desirable computing environment	291
Appendix 2.	Tables	301
<i>References</i>		305
<i>Name index</i>		325
<i>Subject index</i>		329