

Contents

Preface	vii
1 Why Is Nonparametric Regression Important?	1
1.1 Regression Analysis and L_2 Risk	1
1.2 Regression Function Estimation and L_2 Error	2
1.3 Practical Applications	4
1.4 Application to Pattern Recognition	6
1.5 Parametric versus Nonparametric Estimation	9
1.6 Consistency	12
1.7 Rate of Convergence	13
1.8 Adaptation	14
1.9 Fixed versus Random Design Regression	15
1.10 Bibliographic Notes	16
Problems and Exercises	16
2 How to Construct Nonparametric Regression Estimates?	18
2.1 Four Related Paradigms	18
2.2 Curse of Dimensionality	23
2.3 Bias–Variance Tradeoff	24
2.4 Choice of Smoothing Parameters and Adaptation	26
2.5 Bibliographic Notes	28
Problems and Exercises	29

3	Lower Bounds	31
	3.1 Slow Rate	31
	3.2 Minimax Lower Bounds	36
	3.3 Individual Lower Bounds	43
	3.4 Bibliographic Notes	50
	Problems and Exercises	50
4	Partitioning Estimates	52
	4.1 Introduction	52
	4.2 Stone's Theorem	55
	4.3 Consistency	60
	4.4 Rate of Convergence	64
	4.5 Bibliographic Notes	67
	Problems and Exercises	68
5	Kernel Estimates	70
	5.1 Introduction	70
	5.2 Consistency	71
	5.3 Rate of Convergence	77
	5.4 Local Polynomial Kernel Estimates	80
	5.5 Bibliographic Notes	82
	Problems and Exercises	82
6	k-NN Estimates	86
	6.1 Introduction	86
	6.2 Consistency	88
	6.3 Rate of Convergence	93
	6.4 Bibliographic Notes	96
	Problems and Exercises	97
7	Splitting the Sample	100
	7.1 Best Random Choice of a Parameter	100
	7.2 Partitioning, Kernel, and Nearest Neighbor Estimates	105
	7.3 Bibliographic Notes	108
	Problems and Exercises	109
8	Cross-Validation	112
	8.1 Best Deterministic Choice of the Parameter	112
	8.2 Partitioning and Kernel Estimates	113
	8.3 Proof of Theorem 8.1	115
	8.4 Nearest Neighbor Estimates	126
	8.5 Bibliographic Notes	127
	Problems and Exercises	127
9	Uniform Laws of Large Numbers	130

9.1	Basic Exponential Inequalities	131
9.2	Extension to Random L_1 Norm Covers	134
9.3	Covering and Packing Numbers	140
9.4	Shatter Coefficients and VC Dimension	143
9.5	A Uniform Law of Large Numbers	153
9.6	Bibliographic Notes	156
	Problems and Exercises	156
10	Least Squares Estimates I: Consistency	158
10.1	Why and How Least Squares?	158
10.2	Consistency from Bounded to Unbounded Y	165
10.3	Linear Least Squares Series Estimates	170
10.4	Piecewise Polynomial Partitioning Estimates	174
10.5	Bibliographic Notes	180
	Problems and Exercises	180
11	Least Squares Estimates II: Rate of Convergence	183
11.1	Linear Least Squares Estimates	183
11.2	Piecewise Polynomial Partitioning Estimates	194
11.3	Nonlinear Least Squares Estimates	197
11.4	Preliminaries to the Proof of Theorem 11.4	203
11.5	Proof of Theorem 11.4	210
11.6	Bibliographic Notes	219
	Problems and Exercises	220
12	Least Squares Estimates III: Complexity Regularization	222
12.1	Motivation	222
12.2	Definition of the Estimate	225
12.3	Asymptotic Results	227
12.4	Piecewise Polynomial Partitioning Estimates	232
12.5	Bibliographic Notes	233
	Problems and Exercises	234
13	Consistency of Data-Dependent Partitioning Estimates	235
13.1	A General Consistency Theorem	235
13.2	Cubic Partitions with Data-Dependent Grid Size	241
13.3	Statistically Equivalent Blocks	243
13.4	Nearest Neighbor Clustering	245
13.5	Bibliographic Notes	250
	Problems and Exercises	251
14	Univariate Least Squares Spline Estimates	252
14.1	Introduction to Univariate Splines	252
14.2	Consistency	267
14.3	Spline Approximation	273

14.4	Rate of Convergence	277
14.5	Bibliographic Notes	281
	Problems and Exercises	281
15	Multivariate Least Squares Spline Estimates	283
15.1	Introduction to Tensor Product Splines	283
15.2	Consistency	290
15.3	Rate of Convergence	294
15.4	Bibliographic Notes	296
	Problems and Exercises	296
16	Neural Networks Estimates	297
16.1	Neural Networks	297
16.2	Consistency	300
16.3	Rate of Convergence	315
16.4	Bibliographic Notes	326
	Problems and Exercises	328
17	Radial Basis Function Networks	329
17.1	Radial Basis Function Networks	329
17.2	Consistency	332
17.3	Rate of Convergence	340
17.4	Increasing Kernels and Approximation	348
17.5	Bibliographic Notes	350
	Problems and Exercises	350
18	Orthogonal Series Estimates	353
18.1	Wavelet Estimates	353
18.2	Empirical Orthogonal Series Estimates	356
18.3	Connection with Least Squares Estimates	358
18.4	Empirical Orthogonalization of Piecewise Polynomials	361
18.5	Consistency	366
18.6	Rate of Convergence	372
18.7	Bibliographic Notes	378
	Problems and Exercises	378
19	Advanced Techniques from Empirical Process Theory	380
19.1	Chaining	380
19.2	Extension of Theorem 11.6	385
19.3	Extension of Theorem 11.4	390
19.4	Piecewise Polynomial Partitioning Estimates	397
19.5	Bibliographic Notes	404
	Problems and Exercises	405
20	Penalized Least Squares Estimates I: Consistency	407

20.1	Univariate Penalized Least Squares Estimates	408
20.2	Proof of Lemma 20.1	414
20.3	Consistency	418
20.4	Multivariate Penalized Least Squares Estimates	425
20.5	Consistency	427
20.6	Bibliographic Notes	429
	Problems and Exercises	429
21	Penalized Least Squares Estimates II: Rate of Convergence	433
21.1	Rate of Convergence	433
21.2	Application of Complexity Regularization	440
21.3	Bibliographic notes	446
	Problems and Exercises	447
22	Dimension Reduction Techniques	448
22.1	Additive Models	449
22.2	Projection Pursuit	451
22.3	Single Index Models	456
22.4	Bibliographic Notes	457
	Problems and Exercises	457
23	Strong Consistency of Local Averaging Estimates	459
23.1	Partitioning Estimates	459
23.2	Kernel Estimates	479
23.3	k-NN Estimates	486
23.4	Bibliographic Notes	491
	Problems and Exercises	491
24	Semirecursive Estimates	493
24.1	A General Result	493
24.2	Semirecursive Kernel Estimate	496
24.3	Semirecursive Partitioning Estimate	507
24.4	Bibliographic Notes	510
	Problems and Exercises	511
25	Recursive Estimates	512
25.1	A General Result	512
25.2	Recursive Kernel Estimate	517
25.3	Recursive Partitioning Estimate	518
25.4	Recursive NN Estimate	518
25.5	Recursive Series Estimate	520
25.6	Pointwise Universal Consistency	526
25.7	Bibliographic Notes	537
	Problems and Exercises	537

26 Censored Observations	540
26.1 Right Censoring Regression Models	540
26.2 Survival Analysis, the Kaplan-Meier Estimate	541
26.3 Regression Estimation for Model A	548
26.4 Regression Estimation for Model B	555
26.5 Bibliographic Notes	563
Problems and Exercises	563
27 Dependent Observations	564
27.1 Stationary and Ergodic Observations	565
27.2 Dynamic Forecasting: Autoregression	568
27.3 Static Forecasting: General Case	572
27.4 Time Series Problem: Cesàro Consistency	576
27.5 Time Series Problem: Universal Prediction	576
27.6 Estimating Smooth Regression Functions	582
27.7 Bibliographic Notes	587
Problems and Exercises	588
Appendix A: Tools	589
A.1 A Denseness Result	589
A.2 Inequalities for Independent Random Variables	592
A.3 Inequalities for Martingales	598
A.4 Martingale Convergences	601
Problems and Exercises	607
Notation	609
Bibliography	612
Author Index	639
Subject Index	644