

Contents

<i>Preface</i>	<i>page</i> xiii
<i>Guide to Notation</i>	xv
1 Introduction	1
1.1 Assessing the Carcinogenicity of Phenolphthalein	3
1.2 Salinity and Fishing in North Carolina	4
1.3 Management of a Retirement Fund	5
1.4 Biomonitoring of Airborne Mercury	7
1.5 Term Structure of Interest Rates	7
1.6 Air Pollution and Mortality in Milan: The Harvesting Effect	11
2 Parametric Regression	15
2.1 Introduction	15
2.2 Linear Regression Models	15
2.3 Regression Diagnostics	20
2.4 Inference	28
2.5 Parametric Additive Models	36
2.6 Model Selection	44
2.7 Polynomial Regression Models	46
2.8 Nonlinear Regression	48
2.9 Transformations in Regression	51
2.10 Bibliographic Notes	55
2.11 Summary of Formulas	55
3 Scatterplot Smoothing	57
3.1 Introduction	57
3.2 Preliminary Ideas	58
3.3 Practical Implementation	62
3.4 Automatic Knot Selection	64
3.5 Penalized Spline Regression	65
3.6 Quadratic Spline Bases	67
3.7 Other Spline Models and Bases	69
3.8 Other Penalties	74
3.9 General Definition of a Penalized Spline	75
3.10 Linear Smoothers	76
3.11 Error of a Smoother	76
	vii

3.12 Rank of a Smoother	78
3.13 Degrees of Freedom of a Smoother	80
3.14 Residual Degrees of Freedom	82
3.15 Other Approaches to Scatterplot Smoothing	84
3.16 Choosing a Scatterplot Smoother	87
3.17 Bibliographical Notes	88
3.18 Summary of Formulas	89
4 Mixed Models	91
4.1 Introduction	91
4.2 Mixed Models	91
4.3 Prediction	95
4.4 The Linear Mixed Model (LMM)	98
4.5 Estimation and Prediction in LMM	98
4.6 Estimated BLUP (EBLUP)	101
4.7 Standard Error Estimation	102
4.8 Hypothesis Testing	104
4.9 Penalized Splines as BLUPs	108
4.10 Bibliographical Notes	110
4.11 Summary of Formulas	110
5 Automatic Scatterplot Smoothing	112
5.1 Introduction	112
5.2 The Likelihood Approach	113
5.3 The Model Selection Approach	114
5.4 Caveats of Automatic Parameter Selection	120
5.5 Choosing the Knots and Basis Functions	123
5.6 Automatic Selection of the Number of Knots	127
5.7 Bibliographical Notes	131
5.8 Summary of Formulas	131
6 Inference	133
6.1 Introduction	133
6.2 Variability Bands	133
6.3 Confidence and Prediction Intervals	135
6.4 Inference for Penalized Splines	137
6.5 Simultaneous Confidence Bands	142
6.6 Testing the Adequacy of Parametric Models	145
6.7 Testing for No Effect	149
6.8 Inference Using First Derivatives	151
6.9 Testing for Existence of a Feature	156
6.10 Bibliographical Notes	158
6.11 Summary of Formulas	159
7 Simple Semiparametric Models	161
7.1 Introduction	161
7.2 Beyond Scatterplot Smoothing	161

<i>Contents</i>	ix
7.3 Semiparametric Binary Offset Model	162
7.4 Additivity and Interactions	164
7.5 General Parametric Component	164
7.6 Inference	167
7.7 Bibliographical Notes	168
8 Additive Models	170
8.1 Introduction	170
8.2 Fitting an Additive Model	171
8.3 Degrees of Freedom	174
8.4 Smoothing Parameter Selection	176
8.5 Hypothesis Testing	181
8.6 Model Selection	183
8.7 Bibliographical Notes	185
9 Semiparametric Mixed Models	186
9.1 Introduction	186
9.2 Additive Mixed Models	186
9.3 Subject-Specific Curves	191
9.4 Bibliographical Notes	192
10 Generalized Parametric Regression	194
10.1 Introduction	194
10.2 Binary Response Data	194
10.3 Logistic Regression	195
10.4 Other Generalized Linear Models	197
10.5 Iteratively Reweighted Least Squares	200
10.6 Hat Matrix, Degrees of Freedom, and Standard Errors	201
10.7 Overdispersion and Variance Functions: Pseudolikelihood	201
10.8 Generalized Linear Mixed Models	203
10.9 Deviance	209
10.10 Technical Details	210
10.11 Bibliographical Notes	213
11 Generalized Additive Models	214
11.1 Introduction	214
11.2 Generalized Scatterplot Smoothing	215
11.3 Generalized Additive Mixed Models	217
11.4 Degrees-of-Freedom Approximations	219
11.5 Automatic Smoothing Parameter Selection	220
11.6 Hypothesis Testing	220
11.7 Model Selection	221
11.8 Density Estimation	221
11.9 Bibliographical Notes	222
12 Interaction Models	223
12.1 Introduction	223
12.2 Binary-by-Continuous Interaction Models	224

x	<i>Contents</i>
12.3	Factor-by-Curve Interactions in Additive Models 226
12.4	Varying Coefficient Models 234
12.5	Continuous-by-Continuous Interactions 235
12.6	Bibliographical Notes 237
13	Bivariate Smoothing 238
13.1	Introduction 238
13.2	Choice of Bivariate Basis Functions 240
13.3	Kriging 242
13.4	General Radial Smoothing 248
13.5	Default Automatic Bivariate Smoother 256
13.6	Geoadditive Models 258
13.7	Additive Plus Interaction Models 259
13.8	Generalized Bivariate Smoothing 259
13.9	Appendix: Equivalence of BLUP using \mathbf{Z}_R and \mathbf{Z}_P 259
13.10	Bibliographical Notes 260
14	Variance Function Estimation 261
14.1	Introduction 261
14.2	Formulation 263
14.3	Application to the LIDAR Data 264
14.4	Quasilikelihood and Variance Functions 266
14.5	Bibliographical Notes 267
15	Measurement Error 268
15.1	Introduction 268
15.2	Formulation 269
15.3	The Expectation Maximization (EM) Algorithm 270
15.4	Simulated Example Revisited 273
15.5	Sensitivity Analysis Example 273
15.6	Bibliographical Notes 275
16	Bayesian Semiparametric Regression 276
16.1	Introduction 276
16.2	General Framework 277
16.3	Scatterplot Smoothing 279
16.4	Linear Mixed Models 285
16.5	Generalized Linear Mixed Models 288
16.6	Rao–Blackwellization 291
16.7	Bibliographical Notes 292
17	Spatially Adaptive Smoothing 293
17.1	Introduction 293
17.2	A Local Penalty Method 294
17.3	Completely Automatic Algorithm 295
17.4	Bayesian Inference 296

<i>Contents</i>	xi
17.5 Simulations	298
17.6 LIDAR Example	304
17.7 Additive Models	305
17.8 Bibliographical Notes	307
18 Analyses	308
18.1 Cancer Rates on Cape Cod	308
18.2 Assessing the Carcinogenicity of Phenolphthalein	308
18.3 Salinity and Fishing in North Carolina	308
18.4 Management of a Retirement Fund	313
18.5 Biomonitoring of Airborne Mercury	314
18.6 Term Structure of Interest Rates	315
18.7 Air Pollution and Mortality in Milan: The Harvesting Effect	319
19 Epilogue	320
19.1 Introduction	320
19.2 Minimalist Statistics	320
19.3 Some Omitted Topics	321
19.4 Future Research	325
A Technical Complements	326
A.1 Introduction	326
A.2 Matrix Definitions and Results	326
A.3 Linear Algebra	331
A.4 Probability Definitions and Results	333
A.5 Maximum Likelihood Estimation	335
A.6 Bibliographical Notes	335
B Computational Issues	336
B.1 Fast Computation of Penalized Spline Smooths	336
B.2 Computation of Covariance Matrix Estimators	351
B.3 Software	353
<i>Bibliography</i>	361
<i>Author Index</i>	375
<i>Notation Index</i>	380
<i>Example Index</i>	381
<i>Subject Index</i>	382