Contents

Preface to the Second Edition v

1 Introduction 1

1.1 What are functional data? 1
1.2 Functional models for nonfunctional data 5
1.3 Some functional data analyses 5
1.4 The goals of functional data analysis 9
1.5 The first steps in a functional data analysis 11
1.5.1 Data representation: smoothing and interpolation 11
1.5.2 Data registration or feature alignment 12
1.5.3 Data display 13
1.5.4 Plotting pairs of derivatives 13
1.6 Exploring variability in functional data 15
1.6.1 Functional descriptive statistics 15
1.6.2 Functional principal components analysis 15
1.6.3 Functional canonical correlation 16
1.7 Functional linear models 16
1.8 Using derivatives in functional data analysis 17
1.9 Concluding remarks 18

2 Tools for exploring functional data 19

2.1 Introduction 19
2.2 Some notation 20
2.2.1 Scalars, vectors, functions and matrices 20

2.2.2 Derivatives and integrals 20
2.2.3 Inner products 21
2.2.4 Functions of functions 21
2.3 Summary statistics for functional data 22
2.3.1 Functional means and variances 22
2.3.2 Covariance and correlation functions 22
2.3.3 Cross-covariance and cross-correlation functions . 24
2.4 The anatomy of a function 26
2.4.1 Functional features 26
2.4.2 Data resolution and functional dimensionality . . 27
2.4.3 The size of a function 28
2.5 Phase-plane plots of periodic effects 29
2.5.1 The log nondurable goods index 29
2.5.2 Phase–plane plots show energy transfer 30
2.5.3 The nondurable goods cycles 33
2.6 Further reading and notes 34

3 From functional data to smooth functions 37
3.1 Introduction 37
3.2 Some properties of functional data 38
3.2.1 What makes discrete data functional? 38
3.2.2 Samples of functional data 39
3.2.3 The interplay between smooth and noisy variation 39
3.2.4 The standard model for error and its limitations . 40
3.2.5 The resolving power of data 41
3.2.6 Data resolution and derivative estimation 41
3.3 Representing functions by basis functions 43
3.4 The Fourier basis system for periodic data 45
3.5 The spline basis system for open-ended data 46
3.5.1 Spline functions and degrees of freedom 47
3.5.2 The B-spline basis for spline functions 49
3.6 Other useful basis systems 53
3.6.1 Wavelets 53
3.6.2 Exponential and power bases 54
3.6.3 Polynomial bases 54
3.6.4 The polygonal basis 55
3.6.5 The step-function basis 55
3.6.6 The constant basis 55
3.6.7 Empirical and designer bases 56
3.7 Choosing a scale for t 56
3.8 Further reading and notes 57

4 Smoothing functional data by least squares 59
4.1 Introduction 59
4.2 Fitting data using a basis system by least squares 59

4.2.1 Ordinary or unweighted least squares fits 60
4.2.2 Weighted least squares fits 61
4.3 A performance assessment of least squares smoothing . . 62
4.4 Least squares fits as linear transformations of the data . 63
4.4.1 How linear smoothers work 64
4.4.2 The degrees of freedom of a linear smooth 66
4.5 Choosing the number K of basis functions 67
4.5.1 The bias/variance trade-off 67
4.5.2 Algorithms for choosing K 69
4.6 Computing sampling variances and confidence limits . . 70
4.6.1 Sampling variance estimates 70
4.6.2 Estimating $\mathbf{\Sigma}_e$ 71
4.6.3 Confidence limits 72
4.7 Fitting data by localized least squares 73
4.7.1 Kernel smoothing 74
4.7.2 Localized basis function estimators 76
4.7.3 Local polynomial smoothing 77
4.7.4 Choosing the bandwidth h 78
4.7.5 Summary of localized basis methods 78
4.8 Further reading and notes 79

5 Smoothing functional data with a roughness penalty 81
5.1 Introduction 81
5.2 Spline smoothing 82
5.2.1 Two competing objectives in function estimation 83
5.2.2 Quantifying roughness 84
5.2.3 The penalized sum of squared errors fitting criterion 84
5.2.4 The structure of a smoothing spline 85
5.2.5 How spline smooths are computed 86
5.2.6 Spline smoothing as a linear operation 87
5.2.7 Spline smoothing as an augmented least squares problem 89
5.2.8 Estimating derivatives by spline smoothing 90
5.3 Some extensions 91
5.3.1 Roughness penalties with fewer basis functions . . 91
5.3.2 More general measures of data fit 92
5.3.3 More general roughness penalties 92
5.3.4 Computing the roughness penalty matrix 93
5.4 Choosing the smoothing parameter 94
5.4.1 Some limits imposed by computational issues . . 94
5.4.2 The cross-validation or CV method 96
5.4.3 The generalized cross-validation or GCV method 97
5.4.4 Spline smoothing the simulated growth data . . . 99
5.5 Confidence intervals for function values and functional probes 100

5.5.1 Linear functional probes 101
5.5.2 Two linear mappings defining a probe value . . . 102
5.5.3 Computing confidence limits for function values . 103
5.5.4 Confidence limits for growth acceleration 104
5.6 A bi-resolution analysis with smoothing splines 104
5.6.1 Complementary bases 105
5.6.2 Specifying the roughness penalty 106
5.6.3 Some properties of the estimates 107
5.6.4 Relationship to the roughness penalty approach . 108
5.7 Further reading and notes 109

6 Constrained functions **111**
6.1 Introduction 111
6.2 Fitting positive functions 111
6.2.1 A positive smoothing spline 113
6.2.2 Representing a positive function by a differential equation 114
6.3 Fitting strictly monotone functions 115
6.3.1 Fitting the growth of a baby's tibia 115
6.3.2 Expressing a strictly monotone function explicitly 115
6.3.3 Expressing a strictly monotone function as a differential equation 116
6.4 The performance of spline smoothing revisited 117
6.5 Fitting probability functions 118
6.6 Estimating probability density functions 119
6.7 Functional data analysis of point processes 121
6.8 Fitting a linear model with estimation of the density of residuals 123
6.9 Further notes and readings 126

7 The registration and display of functional data **127**
7.1 Introduction 127
7.2 Shift registration 129
7.2.1 The least squares criterion for shift alignment . . 131
7.3 Feature or landmark registration 132
7.4 Using the warping function h to register x 137
7.5 A more general warping function h 137
7.6 A continuous fitting criterion for registration 138
7.7 Registering the height acceleration curves 140
7.8 Some practical advice 142
7.9 Computational details 142
7.9.1 Shift registration by the Newton-Raphson algorithm 142
7.10 Further reading and notes 144

8 Principal components analysis for functional data 147
8.1 Introduction 147
8.2 Defining functional PCA 148
8.2.1 PCA for multivariate data 148
8.2.2 Defining PCA for functional data 149
8.2.3 Defining an optimal empirical orthonormal basis . 151
8.2.4 PCA and eigenanalysis 152
8.3 Visualizing the results 154
8.3.1 Plotting components as perturbations of the mean 154
8.3.2 Plotting principal component scores 156
8.3.3 Rotating principal components 156
8.4 Computational methods for functional PCA 160
8.4.1 Discretizing the functions 161
8.4.2 Basis function expansion of the functions 161
8.4.3 More general numerical quadrature 164
8.5 Bivariate and multivariate PCA 166
8.5.1 Defining multivariate functional PCA 167
8.5.2 Visualizing the results 168
8.5.3 Inner product notation: Concluding remarks . . . 170
8.6 Further readings and notes 171

9 Regularized principal components analysis 173
9.1 Introduction 173
9.2 The results of smoothing the PCA 175
9.3 The smoothing approach 177
9.3.1 Estimating the leading principal component . . . 177
9.3.2 Estimating subsequent principal components . . . 177
9.3.3 Choosing the smoothing parameter by CV 178
9.4 Finding the regularized PCA in practice 179
9.4.1 The periodic case 179
9.4.2 The nonperiodic case 181
9.5 Alternative approaches 182
9.5.1 Smoothing the data rather than the PCA 182
9.5.2 A stepwise roughness penalty procedure 184
9.5.3 A further approach 185

10 Principal components analysis of mixed data 187
10.1 Introduction 187
10.2 General approaches to mixed data 189
10.3 The PCA of hybrid data 190
10.3.1 Combining function and vector spaces 190
10.3.2 Finding the principal components in practice . . . 191
10.3.3 Incorporating smoothing 192
10.3.4 Balance between functional and vector variation . 192
10.4 Combining registration and PCA 194

10.4.1 Expressing the observations as mixed data 194
10.4.2 Balancing temperature and time shift effects . . . 194
10.5 The temperature data reconsidered 195
10.5.1 Taking account of effects beyond phase shift . . . 195
10.5.2 Separating out the vector component 198

11 Canonical correlation and discriminant analysis 201
11.1 Introduction 201
11.1.1 The basic problem 201
11.2 Principles of classical CCA 204
11.3 Functional canonical correlation analysis 204
11.3.1 Notation and assumptions 204
11.3.2 The naive approach does not give meaningful results 205
11.3.3 Choice of the smoothing parameter 206
11.3.4 The values of the correlations 207
11.4 Application to the study of lupus nephritis 208
11.5 Why is regularization necessary? 209
11.6 Algorithmic considerations 210
11.6.1 Discretization and basis approaches 210
11.6.2 The roughness of the canonical variates 211
11.7 Penalized optimal scoring and discriminant analysis . . . 213
11.7.1 The optimal scoring problem 213
11.7.2 The discriminant problem 214
11.7.3 The relationship with CCA 214
11.7.4 Applications 215
11.8 Further readings and notes 215

12 Functional linear models 217
12.1 Introduction 217
12.2 A functional response and a categorical independent variable 218
12.3 A scalar response and a functional independent variable 219
12.4 A functional response and a functional independent variable 220
12.4.1 Concurrent 220
12.4.2 Annual or total 220
12.4.3 Short-term feed-forward 220
12.4.4 Local influence 221
12.5 What about predicting derivatives? 221
12.6 Overview 222

13 Modelling functional responses with multivariate covariates 223
13.1 Introduction 223
13.2 Predicting temperature curves from climate zones 223
13.2.1 Fitting the model 225

13.2.2 Assessing the fit 225
13.3 Force plate data for walking horses 229
13.3.1 Structure of the data 229
13.3.2 A functional linear model for the horse data . . . 231
13.3.3 Effects and contrasts 233
13.4 Computational issues 235
13.4.1 The general model 235
13.4.2 Pointwise minimization 236
13.4.3 Functional linear modelling with regularized basis expansions 236
13.4.4 Using the Kronecker product to express $\hat{\mathbf{B}}$ 238
13.4.5 Fitting the raw data 239
13.5 Confidence intervals for regression functions 239
13.5.1 How to compute confidence intervals 239
13.5.2 Confidence intervals for climate zone effects . . . 241
13.5.3 Some cautions on interpreting confidence intervals 243
13.6 Further reading and notes 244

14 Functional responses, functional covariates and the concurrent model 247
14.1 Introduction 247
14.2 Predicting precipitation profiles from temperature curves 248
14.2.1 The model for the daily logarithm of rainfall . . . 248
14.2.2 Preliminary steps 248
14.2.3 Fitting the model and assessing fit 250
14.3 Long-term and seasonal trends in the nondurable goods index 251
14.4 Computational issues 255
14.5 Confidence intervals 257
14.6 Further reading and notes 258

15 Functional linear models for scalar responses 261
15.1 Introduction 261
15.2 A naive approach: Discretizing the covariate function . . 262
15.3 Regularization using restricted basis functions 264
15.4 Regularization with roughness penalties 266
15.5 Computational issues 268
15.5.1 Computing the regularized solution 269
15.5.2 Computing confidence limits 270
15.6 Cross-validation and regression diagnostics 270
15.7 The direct penalty method for computing β 271
15.7.1 Functional interpolation 272
15.7.2 The two-stage minimization process 272
15.7.3 Functional interpolation revisited 273
15.8 Functional regression and integral equations 275

15.9 Further reading and notes 276

16 Functional linear models for functional responses 279
16.1 Introduction: Predicting log precipitation from temperature 279
16.1.1 Fitting the model without regularization 280
16.2 Regularizing the fit by restricting the bases 282
16.2.1 Restricting the basis $\eta(s)$ 282
16.2.2 Restricting the basis $\theta(t)$ 283
16.2.3 Restricting both bases 284
16.3 Assessing goodness of fit 285
16.4 Computational details 290
16.4.1 Fitting the model without regularization 291
16.4.2 Fitting the model with regularization 292
16.5 The general case 293
16.6 Further reading and notes 295

17 Derivatives and functional linear models 297
17.1 Introduction 297
17.2 The oil refinery data 298
17.3 The melanoma data 301
17.4 Some comparisons of the refinery and melanoma analyses 305

18 Differential equations and operators 307
18.1 Introduction 307
18.2 Exploring a simple linear differential equation 308
18.3 Beyond the constant coefficient first-order linear equation 310
18.3.1 Nonconstant coefficients 310
18.3.2 Higher order equations 311
18.3.3 Systems of equations 312
18.3.4 Beyond linearity 313
18.4 Some applications of linear differential equations and operators 313
18.4.1 Differential operators to produce new functional observations 313
18.4.2 The gross domestic product data 314
18.4.3 Differential operators to regularize or smooth models 316
18.4.4 Differential operators to partition variation 317
18.4.5 Operators to define solutions to problems 319
18.5 Some linear differential equation facts 319
18.5.1 Derivatives are rougher 319
18.5.2 Finding a linear differential operator that annihilates known functions 320
18.5.3 Finding the functions ξ_j satisfying $L\xi_j = 0$ . . . 322

18.6 Initial conditions, boundary conditions and other constraints 323
18.6.1 Why additional constraints are needed to define a solution 323
18.6.2 How L and B partition functions 324
18.6.3 The inner product defined by operators L and B 325
18.7 Further reading and notes 325

19 Principal differential analysis **327**
19.1 Introduction 327
19.2 Defining the problem 328
19.3 A principal differential analysis of lip movement 329
19.3.1 The biomechanics of lip movement 330
19.3.2 Visualizing the PDA results 332
19.4 PDA of the pinch force data 334
19.5 Techniques for principal differential analysis 338
19.5.1 PDA by point-wise minimization 338
19.5.2 PDA using the concurrent functional linear model 339
19.5.3 PDA by iterating the concurrent linear model . . 340
19.5.4 Assessing fit in PDA 343
19.6 Comparing PDA and PCA 343
19.6.1 PDA and PCA both minimize sums of squared errors 343
19.6.2 PDA and PCA both involve finding linear operators 344
19.6.3 Differences between differential operators (PDA) and projection operators (PCA) 345
19.7 Further readings and notes 348

20 Green's functions and reproducing kernels **349**
20.1 Introduction 349
20.2 The Green's function for solving a linear differential equation 350
20.2.1 The definition of the Green's function 351
20.2.2 A matrix analogue of the Green's function 352
20.2.3 A recipe for the Green's function 352
20.3 Reproducing kernels and Green's functions 353
20.3.1 What is a reproducing kernel? 354
20.3.2 The reproducing kernel for $\ker B$ 355
20.3.3 The reproducing kernel for $\ker L$ 356
20.4 Further reading and notes 357

21 More general roughness penalties **359**
21.1 Introduction 359
21.1.1 The lip movement data 360
21.1.2 The weather data 361

21.2 The optimal basis for spline smoothing 363
21.3 An $O(n)$ algorithm for L-spline smoothing 364
21.3.1 The need for a good algorithm 364
21.3.2 Setting up the smoothing procedure 366
21.3.3 The smoothing phase 367
21.3.4 The performance assessment phase 367
21.3.5 Other $O(n)$ algorithms 369
21.4 A compact support basis for L-splines 369
21.5 Some case studies 370
21.5.1 The gross domestic product data 370
21.5.2 The melanoma data 371
21.5.3 The GDP data with seasonal effects 373
21.5.4 Smoothing simulated human growth data 374

22 Some perspectives on FDA **379**
22.1 The context of functional data analysis 379
22.1.1 Replication and regularity 379
22.1.2 Some functional aspects elsewhere in statistics . . 380
22.1.3 Functional analytic treatments 381
22.2 Challenges for the future 382
22.2.1 Probability and inference 382
22.2.2 Asymptotic results 383
22.2.3 Multidimensional arguments 383
22.2.4 Practical methodology and applications 384
22.2.5 Back to the data! 384

Appendix: Some algebraic and functional techniques **385**
A.1 Inner products $\langle x, y \rangle$ 385
A.1.1 Some specific examples 386
A.1.2 General properties 387
A.1.3 Descriptive statistics in inner product notation . 389
A.1.4 Some extended uses of inner product notation . . 390
A.2 Further aspects of inner product spaces 391
A.2.1 Projections 391
A.2.2 Quadratic optimization 392
A.3 Matrix decompositions and generalized inverses 392
A.3.1 Singular value decompositions 392
A.3.2 Generalized inverses 393
A.3.3 The QR decomposition 393
A.4 Projections 394
A.4.1 Projection matrices 394
A.4.2 Finding an appropriate projection matrix 395
A.4.3 Projections in more general inner product spaces 395
A.5 Constrained maximization of a quadratic function 396
A.5.1 The finite-dimensional case 396

A.5.2 The problem in a more general space 396
A.5.3 Generalized eigenproblems 397
A.6 Kronecker Products 398
A.7 The multivariate linear model 399
A.7.1 Linear models from a transformation perspective 399
A.7.2 The least squares solution for **B** 400
A.8 Regularizing the multivariate linear model 401
A.8.1 Definition of regularization 401
A.8.2 Hard-edged constraints 401
A.8.3 Soft-edged constraints 402

References **405**

Index **419**