

Contents

1	Functional data structures	1
1.1	Examples of functional data	1
1.2	Detection of abnormal NO_x pollution levels	6
1.3	Prediction of the volume of credit card transactions	9
1.4	Classification of temporal gene expression data	12
1.5	Statistical packages, bases, and functional objects	15
Part I Independent functional observations		
2	Hilbert space model for functional data	21
2.1	Operators in a Hilbert space	21
2.2	The space L^2	22
2.3	Random elements in L^2 and the covariance operator	23
2.4	Estimation of mean and covariance functions	25
2.5	Estimation of the eigenvalues and the eigenfunctions	31
2.6	Asymptotic normality of the eigenfunctions	32
2.7	Proof of Theorem 2.7	34
2.8	Bibliographical notes	35
3	Functional principal components	37
3.1	A maximization problem	37
3.2	Optimal empirical orthonormal basis	38
3.3	Functional principal components	39
3.4	Computation of functional principal components	42
3.5	Bibliographical notes	43
4	Canonical correlation analysis	45
4.1	Multivariate canonical components	45
4.2	Functional canonical components	48
4.3	Sample functional canonical components	49
4.4	Functional canonical correlation analysis of a magnetometer data	54

4.5	Square root of the covariance operator	55
4.6	Existence of the functional canonical components	58
5	Two sample inference for the mean and covariance functions	65
5.1	Equality of mean functions	65
5.2	A simulation study and an application	69
5.3	Proofs of the theorems of Section 5.1	71
5.4	Equality of covariance operators	73
5.5	Bibliographical notes	76
6	Detection of changes in the mean function	79
6.1	Introduction	79
6.2	Notation and assumptions	82
6.3	Detection procedure	84
6.4	Finite sample performance and application to temperature data ..	88
6.5	An approximation theorem for functional observations and proofs of Theorems 6.1 and 6.2	94
6.6	Bibliographical notes	104
7	Portmanteau test of independence	105
7.1	Test procedure	106
7.2	Finite sample performance	108
7.3	Application to credit card transactions and diurnal geomagnetic variation	110
7.4	Proofs of the results of Section 7.1	113
7.5	Auxiliary lemmas for H -valued random elements	117
7.6	Limit theory for sample autocovariance matrices	118
 Part II The functional linear model		
8	Functional linear models	127
8.1	Introduction	127
8.2	Standard linear model and normal equations	129
8.3	The fully functional model	130
8.4	The scalar response model	134
8.5	The functional response model	139
8.6	Evaluating the goodness-of-fit	139
8.7	Bibliographical notes	144
9	Test for lack of effect in the functional linear model	147
9.1	Introduction and motivation	147
9.2	The test procedure	149
9.3	A small simulation study	151
9.4	Application to magnetometer data	153
9.5	Asymptotic theory	162
9.6	Proofs of Theorems 9.1 and 9.2	164

10 Two sample inference for regression kernels	169
10.1 Motivation and introduction	169
10.2 Derivation of the test for scalar responses	172
10.3 Derivation for functional responses	176
10.4 Summary of the testing procedures	178
10.5 A small simulation study	180
10.6 Application to medfly and magnetometer data	181
10.7 Asymptotic theory	184
11 Tests for error correlation in the functional linear model	191
11.1 Motivation and background	191
11.2 Least squares estimation	195
11.3 Description of the test procedures	197
11.4 A simulation study	200
11.5 Application to space physics and high-frequency financial data	202
11.6 Asymptotic theory	210
11.7 Proof of Theorem 11.2	213
11.8 Bibliographical notes	223
12 A test of significance in functional quadratic regression	225
12.1 Testing procedure	225
12.2 Application to spectral data	229
12.3 Outline for the Proof of Theorem 12.1	230
12.4 Outline for the Proof of Theorem 12.2	232
Part III Dependent functional data	
13 Functional autoregressive model	235
13.1 Existence	235
13.2 Estimation	238
13.3 Prediction	240
13.4 Predictive factors	246
13.5 The trace class and the polar and singular decompositions	250
14 Change point detection in the functional autoregressive process	253
14.1 Introduction	253
14.2 Testing procedure	254
14.3 Application to credit card transactions and Eurodollar futures	260
14.4 Asymptotic results	264
14.5 Proof of Proposition 14.1	268
14.6 Proof of Proposition 14.2	270
15 Determining the order of the functional autoregressive model	277
15.1 Representation of an $FAR(p)$ process as a functional linear model	278
15.2 Order determination procedure	282
15.3 Finite sample performance and application to financial data	286

16	Functional time series	289
16.1	Approximable functional time series	291
16.2	Convergence of sample eigenfunctions and a central limit theorem	296
16.3	The long-run variance	297
16.4	Estimation of the long-run covariance matrix under weak assumptions	302
16.5	Change point detection	306
16.6	Self-normalized statistics	310
16.7	Functional linear model with dependent regressors	317
16.8	Proofs of the results of Sections 16.2 and 16.3	319
16.9	Proof of Theorem 16.6	327
16.10	Proofs of Theorems 16.7 and 16.8	335
16.11	Proof of Theorem 16.9	338
17	Spatially distributed functional data	343
17.1	Introduction	344
17.2	Essentials of spatial statistics	347
17.3	Estimation of the mean function	350
17.4	Estimation of the principal components	354
17.5	Finite sample performance of the estimators	357
17.6	Testing for correlation of two spatial fields	361
17.7	Modeling and estimation of the covariance tensor	364
17.8	Size and power of the correlation test	366
17.9	Application to critical ionospheric frequency and magnetic curves	368
18	Consistency of the simple mean and the empirical functional principal components for spatially distributed curves	375
18.1	Motivating examples	376
18.2	Models and Assumptions	382
18.3	Regular spatial designs	386
18.4	Consistency of the sample mean function	389
18.5	Consistency of the empirical covariance operator	393
18.6	Inconsistent empirical functional principal components	396
18.7	Proofs of the results of Sections 18.4, 18.5 and 18.6	399
	References	405
	Index	419

Inference for Functional Data with Applications

This book presents recently developed statistical methods and theory required for the application of the tools of functional data analysis to problems arising in geosciences, finance, economics and biology. It is concerned with inference based on second order statistics, especially those related to the functional principal component analysis. While it covers inference for independent and identically distributed functional data, its distinguishing feature is an in depth coverage of dependent functional data structures, including functional time series and spatially indexed functions. Specific inferential problems studied include two sample inference, change point analysis, tests for dependence in data and model residuals and functional prediction. All procedures are described algorithmically, illustrated on simulated and real data sets, and supported by a complete asymptotic theory.

The book can be read at two levels. Readers interested primarily in methodology will find detailed descriptions of the methods and examples of their application. Researchers interested also in mathematical foundations will find carefully developed theory. The organization of the chapters makes it easy for the reader to choose an appropriate focus. The book introduces the requisite, and frequently used, Hilbert space formalism in a systematic manner. This will be useful to graduate or advanced undergraduate students seeking a self-contained introduction to the subject. Advanced researchers will find novel asymptotic arguments.

Lajos Horváth is Professor of Mathematics at the University of Utah. He has served on the editorial boards of *Statistics & Probability Letters*, *Journal of Statistical Planning and Inference* and *Journal of Time Series Econometrics*. He has coauthored more than 250 research papers and 3 books, including *Weighted Approximations in Probability and Statistics* and *Limit Theorems in Change-Point Analysis* (both with Miklós Csörgö).

Piotr Kokoszka is Professor of Statistics at Colorado State University. He has served on the editorial boards of the journals *Statistical Modelling* and *Computational Statistics*. He has coauthored over 100 papers in areas of statistics and its applications focusing on dependent data.

Statistics

ISBN 978-1-4614-3654-6



9 781461 436546

► springer.com

