

Zhang, Jin-Ting; Chen, Jianwei

Statistical inferences for functional data. (English) Zbl 1129.62029
Ann. Stat. 35, No. 3, 1052-1079 (2007).

Summary: With modern technology development, functional data are being observed frequently in many scientific fields. A popular method for analyzing such functional data is “smoothing first, then estimation”. That is, statistical inference, such as estimation and hypothesis testing, about functional data is conducted based on the substitution of the underlying individual functions by their reconstructions obtained by one smoothing technique or another. However, little is known about this substitution effect on functional data analysis.

In this paper this problem is investigated when the local polynomial kernel (LPK) smoothing technique is used for individual function reconstructions. We find that under some mild conditions the substitution effect can be ignored asymptotically. Based on this, we construct LPK reconstruction-based estimators for the mean, covariance and noise variance functions of a functional data set and derive their asymptotics. We also propose a generalized cross-validation (GCV) rule for selecting good bandwidths for the LPK reconstructions. When the mean function also depends on some time-independent covariates, we consider a functional linear model where the mean function is linearly related to the covariates but the covariate effects are functions of time.

The LPK reconstruction-based estimators for the covariate effects and the covariance function are also constructed and their asymptotics are derived. Moreover, we propose an L^2 -norm-based global test statistic for a general hypothesis testing problem about the covariate effects and derive its asymptotic random expression. The effect of the bandwidths selected by the proposed GCV rule on the accuracy of the LPK reconstructions and the mean function estimator is investigated via a simulation study. The proposed methodologies are illustrated via an application to a real functional data set collected in climatology.

MSC:

- 62G07 Density estimation
- 62G20 Asymptotic properties of nonparametric inference
- 62P12 Applications of statistics to environmental and related topics
- 62G10 Nonparametric hypothesis testing
- 62G05 Nonparametric estimation
- 62H12 Estimation in multivariate analysis

Cited in **1** Review
Cited in **54** Documents

Keywords:

asymptotic Gaussian process; asymptotic normal distribution; functional data; hypothesis test; local polynomial smoothing; nonparametric estimation; reconstructed individual functions; root- n consistent

Software:

[fda](#) (R); [KernSmooth](#)

Full Text: [DOI](#) [arXiv](#)

References:

- [1] Besse, P. (1992). PCA stability and choice of dimensionality. *\StatProb* 13 405–410. · [Zbl 0743.62046](#) · [doi:10.1016/0167-7152\(92\)90115-L](#)
- [2] Besse, P., Cardot, H. and Ferraty, F. (1997). Simultaneous nonparametric regressions of unbalanced longitudinal data. *\Comp* 24 255–270. · [Zbl 0900.62199](#) · [doi:10.1016/S0167-9473\(96\)00067-9](#)
- [3] Besse, P. and Ramsay, J. O. (1986). Principal components analysis of sampled functions. *Psychometrika* 51 285–311. · [Zbl 0623.62048](#) · [doi:10.1007/BF02293986](#)
- [4] Brumback, B. and Rice, J. A. (1998). Smoothing spline models for the analysis of nested and crossed samples of curves (with discussion). *\JASA* 93 961–994. *JSTOR*: · [Zbl 1064.62515](#) · [doi:10.2307/2669837](#) · [links.jstor.org](#)

- [5] Buckley, M. J. and Eagleson, G. K. (1988). An approximation to the distribution of quadratic forms in normal random variables. \AJStat 30A 150–159. · [Zbl 0652.62017](#)
- [6] Canadian Climate Program (1982). Canadian Climate Normals 1981–1980 . Environment Canada, Ottawa.
- [7] Eubank, R. L. (1999). Nonparametric Regression and Spline Smoothing , 2nd ed. Dekker, New York. · [Zbl 0936.62044](#)
- [8] Fan, J. (1992). Design-adaptive nonparametric regression. \JASA 87 998–1004. JSTOR: · [Zbl 0850.62354](#) · [doi:10.2307/2290637](#) · [links.jstor.org](#)
- [9] Fan, J. (1993). Local linear regression smoothers and their minimax efficiencies. \AnnStat 21 196–216. · [Zbl 0773.62029](#) · [doi:10.1214/aos/1176349022](#) ·
- [10] Fan, J. and Gijbels, I. (1996). Local Polynomial Modelling and Its Applications . Chapman and Hall, London. · [Zbl 0873.62037](#)
- [11] Fan, J. and Lin, S.-K. (1998). Test of significance when data are curves. \JASA 93 1007–1021. JSTOR: · [Zbl 1064.62525](#) · [doi:10.2307/2669845](#) · [links.jstor.org](#)
- [12] Fan, J. and Yao, Q. (1998). Efficient estimation of conditional variance functions in stochastic regression. \Biometk 85 645–660. JSTOR: · [Zbl 0918.62065](#) · [doi:10.1093/biomet/85.3.645](#) · [www3.oup.co.uk](#)
- [13] Faraway, J. J. (1997). Regression analysis for a functional response. Technometrics 39 254–261. JSTOR: · [Zbl 0891.62027](#) · [doi:10.2307/1271130](#) · [links.jstor.org](#)
- [14] Green, P. J. and Silverman, B. W. (1994). Nonparametric Regression and Generalized Linear Models. A Roughness Penalty Approach . Chapman and Hall, London. · [Zbl 0832.62032](#)
- [15] Hall, P. and Marron, J. S. (1990). On variance estimation in nonparametric regression. \Biometk 77 415–419. JSTOR: · [Zbl 0711.62035](#) · [doi:10.1093/biomet/77.2.415](#) · [links.jstor.org](#)
- [16] Hart, J. D. and Wehrly, T. E. (1986). Kernel regression estimation using repeated measurements data. \JASA 81 1080–1088. JSTOR: · [Zbl 0635.62030](#) · [doi:10.2307/2289087](#) · [links.jstor.org](#)
- [17] Kneip, A. (1994). Nonparametric estimation of common regressors for similar curve data. \AnnStat 22 1386–1427. · [Zbl 0817.62029](#) · [doi:10.1214/aos/1176325634](#) ·
- [18] Kneip, A. and Engel, J. (1995). Model estimation in nonlinear regression under shape invariance. \AnnStat 23 551–570. · [Zbl 0828.62052](#) · [doi:10.1214/aos/1176324535](#) ·
- [19] Kneip, A. and Gasser, T. (1992). Statistical tools to analyze data representing a sample of curves. \AnnStat 20 1266–1305. · [Zbl 0785.62042](#) · [doi:10.1214/aos/1176348769](#) ·
- [20] Ramsay, J. O. (1995). Some tools for the multivariate analysis of functional data. In Recent Advances in Descriptive Multivariate Analysis (W. Krzanowski, ed.) 269–282. Oxford Univ. Press, New York.
- [21] Ramsay, J. O. and Dalzell, C. J. (1991). Some tools for functional data analysis (with discussion). \JRSSB 53 539–572. JSTOR: · [Zbl 0800.62314](#) · [links.jstor.org](#)
- [22] Ramsay, J. O. and Li, X. (1998). Curve registration. J. R. Stat. Soc. Ser. B Stat. Methodol. 60 351–363. JSTOR: · [Zbl 0909.62033](#) · [doi:10.1111/1467-9868.00129](#) · [links.jstor.org](#)
- [23] Ramsay, J. O. and Silverman, B. W. (1997). Functional Data Analysis . Springer, New York. · [Zbl 0882.62002](#)
- [24] Ramsay, J. O. and Silverman, B. W. (2002). Applied Functional Data Analysis . Methods and Case Studies . Springer, New York. · [Zbl 1011.62002](#) · [doi:10.1007/b98886](#)
- [25] Rice, J. A. and Silverman, B. W. (1991). Estimating the mean and covariance structure nonparametrically when the data are curves. \JRSSB 53 233–243. JSTOR: · [Zbl 0800.62214](#) · [links.jstor.org](#)
- [26] Silverman, B. W. (1995). Incorporating parametric effects into principal components analysis. \JRSSB 57 673–689. JSTOR: · [Zbl 0827.62051](#) · [links.jstor.org](#)
- [27] Wahba, G. (1990). Spline Models for Observational Data . SIAM, Philadelphia. · [Zbl 0813.62001](#)
- [28] Wand, M. P. and Jones, M. C. (1995). Kernel Smoothing . Chapman and Hall, London. · [Zbl 0854.62043](#)
- [29] Zhang, J.-T. (2005). Approximate and asymptotic distributions of chi-squared-type mixtures with applications. \JASA 100 273–285. · [Zbl 1117.62460](#) · [doi:10.1198/016214504000000575](#) · [masetto.asa.catchword.org](#)

This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.