Computational Statistics and Data Analysis 120 (2018) 25–41

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Fast symmetric additive covariance smoothing

Jona Cederbaum, Fabian Scheipl, Sonja Greven *

Department of Statistics, Faculty of Mathematics, Computer Science and Statistics, LMU Munich, Germany

ARTICLE INFO

Article history:
Received 22 September 2016
Received in revised form 2 November 2017
Accepted 3 November 2017
Available online 20 November 2017

Keywords:
Functional data
Longitudinal data
Functional principal components
Penalized splines

ABSTRACT

A fast bivariate smoothing approach for symmetric surfaces is proposed that has a wide range of applications. It is shown how it can be applied to estimate the covariance function in longitudinal data as well as multiple additive covariances in functional data with complex correlation structures. The proposed symmetric smoother can handle (possibly noisy) data sampled on a common, dense grid as well as irregularly or sparsely sampled data. Estimation is based on bivariate penalized spline smoothing using a mixed model representation and the symmetry is used to reduce computation time compared to the usual non-symmetric smoothers. The application of the approach in functional principal component analysis for very general functional linear mixed models is outlined and its practical value is demonstrated in two applications. The approach is evaluated in extensive simulations. Documented open source software is provided that implements the fast symmetric bivariate smoother building on established algorithms for additive models.

1. Introduction

Covariance functions play a central role in many areas of statistics. They summarize the dependency between stochastic observations and encode smoothness assumptions about (observed or latent) random processes. We propose a fast bivariate smoothing approach for symmetric surfaces which can estimate covariance functions in a wide range of data situations. Our approach can handle dependent processes based on an additive decomposition of the covariance function and is also applicable to processes that are observed on irregular or sparse grids.

In functional data analysis (FDA; see, e.g., Ramsay and Silverman, 2005), covariance functions are at the heart of functional principal component analysis (FPCA), a key tool for dimension reduction based on an eigen analysis of the covariance operator of a random process. FPCA is commonly used to estimate the model parameters in functional predictor and functional response regression models (see Morris, 2015 for an overview). Other examples that are based on covariance functions include functional discriminant analysis (James and Hastie, 2001) and functional canonical correlation analysis (Leurgans et al., 1993). In longitudinal data analysis (LDA), where measurements are frequently recorded at irregularly spaced time points, the correct specification of the covariance benefits the estimation efficiency of the fixed effects and improves the individual predictions (cf. Fan et al., 2007). The covariance is also a crucial ingredient in time series analysis, e.g., in risk models and portfolio allocation (cf. Tai, 2009). The interest commonly lies in a single time series in contrast to FDA (and LDA) where multiple curves are observed, e.g., over time. In principle, our symmetric smoothing approach is also applicable to time series which is, however, not the focus in this paper.

* Correspondence to: Department of Statistics, Faculty of Mathematics, Computer Science and Statistics, LMU Munich, Ludwigstr. 33, 80539 Munich, Germany.

E-mail address: Sonja.Greven@stat.uni-muenchen.de (S. Greven).

https://doi.org/10.1016/j.csda.2017.11.002
0167-9473/

Covariance functions are commonly assumed to be smooth. Thus, when the observed curves are not sufficiently smooth (i.e., observed with error) or not measured on a common dense grid, smoothing becomes necessary at some point during covariance estimation. Directly smoothing the observed curves (see, e.g., Besse and Ramsay, 1986), however, is very difficult or impossible for sparsely observed data which are frequently recorded both in FDA and LDA (Yao et al., 2005). Moreover, pre-smoothing the observed curves removes the measurement error, which is not accounted for in subsequent estimation steps. We pursue an alternative approach and apply bivariate smoothing to the sample covariance of the observed data points.

Most existing work on non-parametric covariance estimation is either restricted to independent functional (or longitudinal) observations and/or only applies to data sampled on a common grid. Furthermore, most bivariate smoothing approaches are not specifically designed for covariances. They do not exploit the symmetry of the estimated surface and thus use redundant information in the available data. To the best of our knowledge, previous approaches have never addressed these issues simultaneously. They can be divided according to three main criteria: (1) the generality of the assumed correlation structure in the data, (2) the generality of possible sampling grids, and (3) the estimation procedure including the selection of the degree of smoothing.

A number of approaches address covariance smoothing in LDA. They are restricted to independent curves but allow for general sampling grids. Smoothing is either accomplished by bivariate kernel smoothing (e.g., Staniswalis and Lee, 1998; Yao et al., 2003, 2005) or by bivariate (penalized) spline smoothing (e.g., Kauermann and Wegener, 2011). The degree of smoothing is either chosen by visual inspection (Staniswalis and Lee, 1998), different leave-one-curve-out cross-validation algorithms (e.g., Yao et al., 2003, 2005) or based on a mixed model representation (e.g., Kauermann and Wegener, 2011). These approaches do not account for the symmetry of the estimated surface. James et al. (2000) directly estimate the smooth eigenfunctions of the covariance function. They estimate a reduced rank mixed effects model via the EM algorithm and use B-spline basis functions to represent the eigenfunctions of the covariance operator. Peng and Paul (2009) estimate the same reduced rank model based on a more efficient Newton–Raphson procedure on the Stiefel manifold. The extension of these reduced rank methods to complex correlation structures is not straightforward. Xiao et al. (2017) recently proposed a bivariate smoother designed for covariance smoothing which can be used for sparsely observed, independent functions. They use bivariate penalized B-splines and enforce a symmetry constraint on the spline coefficients which we take up in our extension to correlated curves. Estimation is done by a three-step procedure which accounts for the covariance of the sample covariance. Their leave-one-curve-out cross-validation procedure for selecting the smoothing parameter is not applicable for correlated functional data, however.

Other covariance smoothing approaches can be applied to correlated functions but are restricted to functions sampled on a common grid and considerably simpler correlation structures than ours. Di et al. (2009) and (Greven et al., 2010) use bivariate penalized splines and select the smoothing parameter using restricted maximum likelihood (REML; Patterson and Thompson, 1971) estimation. Shou et al. (2015) apply a method of moments approach based on symmetric sums represented in a sandwich form. For smoothing, they propose to use an extension of the fast covariance estimation algorithm of Xiao et al. (2016b) to correlated functions. Di et al. (2014) extend the functional random intercept model of Di et al. (2009) to sparsely sampled functional data, but the correlation structure remains less general than ours and an extension is not straightforward. More general correlation structures are allowed in the approach of Cederbaum et al. (2016) that is also suitable for sparsely and irregularly sampled functional data. Their focus lies, however, on a model with crossed functional random effects and estimation is only discussed for this special case. Apart from considering less general correlation structures, all these approaches neither avoid the use of redundant information nor account for the symmetry of the smoothed surface.

We propose a fast symmetric bivariate smoothing approach that applies to data with a broad range of possible correlation structures, much broader than existing methods. Furthermore, our approach is well-suited for (possibly noisy) data sampled on a common, dense grid as well as for irregularly or sparsely sampled data. Strength is borrowed by pooling information across different curves, which is particularly important for curves observed on sparse, unequal grids. The smoothing approach we present is widely applicable: In this paper, we demonstrate how it can be applied to longitudinal data as a special case of independent functional data as well as to correlated functional data with very general and complex correlation structures. For the latter, we extend our bivariate smoothing approach to smoothing additive covariance functions. To the best of our knowledge, all previous proposals in this field have been restricted to estimating much less general dependency structures.

We estimate the covariance functions using a smooth method of moments approach represented as a bivariate additive varying coefficient model. The estimation is based on bivariate penalized splines. We choose the smoothing parameters using REML, which allows the direct extension to additive bivariate smoothing of a superposition of multiple covariance functions. This allows our method to be used for a broad range of complex real-word data settings. It also frees us from having to pre-specify a discrete grid of candidate values for the smoothing parameters that is required for cross-validation based approaches like (Xiao et al., 2017). Smoothing the sample covariance quickly becomes a high-dimensional problem as the number of elements in the sample covariance increases quadratically with the number of grid points. We take advantage of the symmetry of the sample covariance and only estimate the upper triangle of the surface including the diagonal. The estimates are then reflected across the diagonal to obtain the entire estimated covariance, which is continuous but not necessarily smooth across the diagonal. To avoid boundary effects on the diagonal and to ensure identifiability of our models, we enforce smoothness across the diagonal by imposing a symmetry constraint on the spline coefficients, which for the simplest case of independent curves reduces to that of Xiao et al. (2017). We show how the symmetry constraint can be