Accepted Manuscript

Local and global temporal correlations for longitudinal data

Yang Zhou, Shu-Chin Lin, Jane-Ling Wang

PII: S0047-259X(17)30341-X
DOI: https://doi.org/10.1016/j.jmva.2018.03.015
Reference: YJMVA 4347

To appear in: *Journal of Multivariate Analysis*

Received date : 3 June 2017

Please cite this article as: Y. Zhou, S. Lin, J.-L. Wang, Local and global temporal correlations for longitudinal data, *Journal of Multivariate Analysis* (2018), https://doi.org/10.1016/j.jmva.2018.03.015

Local and global temporal correlations for longitudinal data

Yang Zhou[a], Shu-Chin Lin[a], Jane-Ling Wang[a,*]

[a]Department of Statistics, University of California, Davis, CA, USA 95616

Abstract

Quantifying the association between a pair of random functions recorded over a time period is of general interest. Several methods have been proposed in the literature but they either suffer from an ill-posed problem intrinsic to functional data or are suitable only for intensely recorded longitudinal data. In this paper we provide new methods that overcome these challenges by investigating the temporal Pearson correlation between paired random functions. We investigate both a local temporal correlation measure and a global summary measure of the dynamic temporal correlations and propose a nonparametric estimation method that covers both intensely observed and sparsely observed longitudinal data. Asymptotic results of the estimators are derived under mild conditions and the method is illustrated via simulations and a benchmark data set.

Keywords: Functional data, Functional correlation, Nonparametric smoothing, Pearson correlation, Rate of convergence

1. Introduction

Functional data analysis (FDA) is an area of research with rising interest owing to the rapid advancement of modern technology. A key ingredient of functional data is that data are in the form of random functions with one function per subject. These random functions are often regarded as realizations of an underlying smooth stochastic process. Much of the development in FDA has been for independent random curves [7, 13, 18, 20, 22] but there is also substantial interest for dependent functional data [1, 2, 8, 11, 14]. One issue of interest is the construction of measures to quantify the association between paired random functions. However, this topic received little attention so far. An early application is in Leurgans et al. [15] who used FDA approaches to model the relation between knee and hip movement for gait data.

Since functional data are often considered multivariate data extended to the infinite-dimensional setting, a natural approach to develop functional correlations is to extend Hotelling's canonical correlation (CC) [12] between random vectors to functional data. This concept of CC had been previously extended to multivariate time series by Brillinger [3] under the stationarity assumption and to functional data in Leurgans et al. [15]. A major drawback of the functional CC is that it is an ill-posed problem as it involves the inversion of covariance operators, which are compact linear operators in Hilbert space. Leurgans et al. [15] pointed out the need for regularization and implemented it through a smoothed and penalized version of functional CC. He et al. [9] took a different approach by imposing restrictions on the functional data so that a generalized inverse could be well defined and Lian [17] later derived the minimax convergence rate of the estimated functional CC under both the L^2 distance and the distance defined by a prediction risk. Cupidon et al. [4] proposed another regularization of the inverse covariance operator that is similar to ridge regression and Eubank and Hsing [6] and Shin and Lee [21] extended functional CC to reproducing kernel Hilbert spaces. Yang et al. [23] took yet another approach to avoid the inverse problem of functional CC through a functional singular component decomposition.

Two additional approaches that are fundamentally different from the above approaches aim at bypassing the inverse problem altogether. One is the dynamic correlation proposed by Dubin and Müller [5] which can be interpreted as

*Corresponding author
Email addresses: (Yang Zhou), (Shu-Chin Lin), `jlwang.ucdavis@gmail.com` (Jane-Ling Wang)