How to interpolate geographically located curves: a precise and straightforward kriging approach to function-valued data

Gilberto Pereira Sassi (corresponding author)

Department of Statistics

Institute of Mathematics and Statistics

Federal University of Bahia, Salvador, Bahia – Brazil

Chang Chiann

Department of Statistics

Institute of Mathematics and Statistics

University of São Paulo, São Paulo, São Paulo – Brazil

Abstract

Functional Data Analysis has been highlighted for its in several fields of science and functional datasets can be spatially indexed curves, which are denominated Spatial Functional Data Analysis. The main goal of this paper is to supply a straightforward and precise approach to interpolate curves, i.e., the aim is to estimate a curve at an unmonitored location. It has been proved that the best linear unbiased estimator for this unsampled curve is the solution of a linear system, where the coefficients and the constant terms of the system are formed using a function called trace-variogram. In this paper, we propose the use of Legendre-Gauss quadrature to estimate the trace-variogram. Excellent numerical properties of this estimator were showed in simulation studies for normality dataset and non-normality dataset: smaller mean square error compared with the established estimation procedure. The novel estimation methodology is illustrated with a real dataset of temperature curves from 35 weather stations of Canada.

Key Words: Trace Variogram; Kriging; Functional Data Analysis; Fourier Series; Spatial Analysis.

1 Introduction

Functional Data Analysis is concerned with analyzing data presented in the form of curves or functions. It can be argued that the concepts of Functional Data Analysis were formalized and introduced by Ramsay and Dalzell (1991), and it is suitable for applications where we need to analyze an observation from a family $\{X(t_j)\}_{j=1,\ldots,J}$ where $t_1,\ldots,t_J$ are equally spaced and $t_j \in (t_{min}, t_{max}), j = 1,\ldots,J$. When the interval between t_j and t_{j+1} gets smaller, we could consider this observation as sampled from a random continuous family $\boldsymbol{\chi} = \{X(t) \mid t \in (t_{min}; t_{max})\}$. Furthermore, there are cases where the underlying data are clearly a continuous function even when the sample is scattered, for example, the child growth curve and the electrical consumption curve (Ferraty and Vieu 2006). Since the seminal paper of Ramsay Ramsay and Dalzell (1991), Functional Data Analysis has been widely developed and applied in various branches of Statistics, such as geostatistics, linear model, item response theory and others. More recently, Fang et al. (2020) considered the functional linear regression for multivariate responses and developed a locally sparse estimation for the functional coefficients. Chen, Goldsmith, and Ogden (2019) proposed a method to model dynamic positron emission tomography (PET) data from multiple subjects simultaneously, where impulse response functions (IRF) are estimated using linear mixed effects functional data model. Beyaztas and Shang (2019) proposed a robust method to forecast functional time series based on the minimum density power divergence estimator of Basu et al. (1998). Lee et al. (2019), used a Bayesian functional mixed model to analyze a glaucoma scleral strain dataset, where they could pick up nonparametric covariate effects, serial and nested interfunctional correlation. Zamani, Hashemi, and Haghbin (2019) proposed a improvement in the Portmanteau test of test of Gabrys and Kokoszka (2007) of functional observations, which is specially suited to small samples.

BIBLIOGRAPHY

Basu, Ayanendranath, Ian R Harris, Nils L Hjort, and MC Jones. 1998. "Robust and Efficient Estimation by Minimising a Density Power Divergence." *Biometrika* 85 (3): 549–59.

Beyaztas, Ufuk, and Han Lin Shang. 2019. "Forecasting Functional Time Series Using Weighted Likelihood Methodology." *Journal of Statistical Computation and Simulation* 89 (16): 3046–60.

Chen, Yakuan, Jeff Goldsmith, and R. Todd Ogden. 2019. "Functional Data Analysis of Dynamic Pet Data." *Journal of the American Statistical Association* 114 (526): 595–609.

Fang, Kuangnan, Xiaochen Zhang, Shuangge Ma, and Qingzhao Zhang. 2020. "Smooth and Locally Sparse Estimation for Multiple-Output Functional Linear Regression." *Journal of Statistical Computation and Simulation* 90 (2): 341–54.

Ferraty, Frédéric, and Philippe Vieu. 2006. *Nonparametric Functional Data Analysis: Theory and Practice*. Springer-Verlag.

Gabrys, Robertas, and Piotr Kokoszka. 2007. "Portmanteau Test of Independence for

Functional Observations." *Journal of the American Statistical Association* 102 (480): 1338–48.

Lee, Wonyul, Michelle F Miranda, Philip Rausch, Veerabhadran Baladandayuthapani, Massimo Fazio, J Crawford Downs, and Jeffrey S Morris. 2019. "Bayesian Semiparametric Functional Mixed Models for Serially Correlated Functional Data, with Application to Glaucoma Data." *Journal of the American Statistical Association* 114 (526): 495–513.

Ramsay, James O, and CJ Dalzell. 1991. "Some Tools for Functional Data Analysis." *Journal of the Royal Statistical Society B* 53 (3): 539–72.

Zamani, Atefeh, Maryam Hashemi, and Hossein Haghbin. 2019. "Improved Functional Portmanteau Tests." *Journal of Statistical Computation and Simulation* 89 (8): 1423–36.