Curve Prediction and Clustering with Mixtures of Gaussian Process Functional Regression Models

Jian Qing Shi
B. Wang
School of Mathematics & Statistics, University of Newcastle, NE1 7RU, UK

Shi et al. (2006) proposed a Gaussian process functional regression (GPFR) model to model functional response curves with a set of functional covariates. Two main problems are addressed by this method: modeling nonlinear and nonparametric regression relationship and modeling covariance structure and mean structure simultaneously. The method gives very good results for curve fitting and prediction but side-steps the problem of heterogeneity. In this paper we present a new method for modeling functional data with 'spatially' indexed data, i.e., the heterogeneity is dependent on factors such as region and individual patient's information. For data collected from different sources, we assume that the data corresponding to each curve (or batch) follows a Gaussian process functional regression model as a lower-level model, and introduce an allocation model for the latent indicator variables as a higher-level model. This higher-level model is dependent on the information related to each batch. This method takes advantage of both GPFR and mixture models and therefore improves the accuracy of predictions. The mixture model has also been used for curve clustering, but focusing on the problem of clustering functional relationships between response curve and covariates. The model is examined on simulated data and real data.

[Jian Qing Shi, School of Mathematics & Statistics, University of Newcastle, NE1 7RU, UK; j.q.shi@ncl.ac.uk]