Abstract:
In many applications, non-Gaussian data such as binary or count are observed over a continuous domain and there exists a smooth underlying structure for describing such data. We develop a new functional data method to deal with this kind of data when the data are regularly spaced on the continuous domain. Our method, referred to as Exponential Family Principal Component Analysis (EFPCA), assumes the data are generated from an exponential family distribution, and the matrix of the canonical parameters has a low-rank structure. The proposed method flexibly accommodates not only the standard one-way functional data, but also two-way (or bivariate) functional data. In addition, we introduce a new cross validation method for estimating the latent rank of a generalized data matrix. We demonstrate the efficacy of the proposed methods using a comprehensive simulation study. The proposed method is also applied to a real application of the UK mortality study, where data are binomially distributed and two-way functional across age groups and calendar years. The results offer novel insights into the underlying mortality pattern.