Supervised cluster analysis of non-Gaussian functional data

Jiakun Jiang

Tsinghua University
E-mail: jiakunj@tsinghua.edu.cn

Abstract: In this paper we study cluster analysis of functional regression with a ran-dom response curve and vector covariates. We propose a mixed transformation functional regression model with an unknown number of clusters. Compared to the existing cluster analysis of functional regression, our model has several advantages. First, our model is free from normality assumption. Second, it is supervised, in that the clustering is based on the relationship between the functional response and covariates. Finally, we allow the number of clusters to be unknown a priori. We propose a combination of penalized likelihood and estimating equation methods to estimate the number of clusters, regression parameters and transformation function simultaneously. We establish theoret-ical properties, including sqrt(n) consistency and asymptotic normality, for the proposed estimators. Extensive simulation results show that the proposed esti-mation procedure works very well. The proposed method is utilized to analyze housing market conditions in China from 2007 to 2014, which leads to some interesting ndings.