Two-step Sparse Boosting for High-Dimensional Longitudinal Data with Varying Coefficients

Ming-Yen Cheng

Hong Kong Baptist University
E-mail: chengmingyen@hkbu.edu.hk

Abstract: Varying-coefficient models are useful for analyzing longitudinal data measured repeatedly over time. Research on variable selection has a new focus on the analysis of high-dimensional longitudinal data. We propose a novel two-step sparse boosting approach, for varying-coefficient model with longitudinal data to carry out the variable selection and the model-based prediction. In the first step, we use the sparse boosting technique to yield an estimate of the correlation structure and in the second step, we take into account of the within-subject correlation structure and conduct variable selection and estimation by sparse boosting again. Extensive simulation studies are conducted to demonstrate the validity of the two-step sparse boosting method. We further demonstrate the proposed methodology by an empirical analysis of yeast cell cycle gene expression data.