A clustering approach with variable selection for longitudinal data

Marie Denis and Mahlet G. Tadesse

In medicine or agronomy, longitudinal studies are conducted to understand dynamic processes, such as disease progress or growth. The identification of groups of individuals with similar profiles over time along with their associated genetic markers can help in the development of more effective therapeutic strategies in human disease and in gaining insights into the adaptation of plants to climate change. Most existing statistical methods do not allow the simultaneous analysis of longitudinal outcomes and the selection of relevant markers in high-dimensional data. In this talk, I will present a Bayesian approach that combines mixture of mixed effects models and variable selection to identify groups of individuals with similar longitudinal response profiles and their associated subsets of covariates with time varying effects. I will illustrate the performance of the approach with simulated data and yeast cell-cycle gene expression data.

Variable selection, clustering, longitudinal data