

Generalized Semiparametric Latent Variable Models for Analyzing Multidimensional and Mixed Mode Data

Xinyuan Song

The Chinese University of Hong Kong, Hong Kong xysong@sta.cuhk.edu.hk

In the medical, behavioral, social, and psychological sciences, latent variables represent unobservable traits that are measured by multiple observed variables. Latent variable models (LVMs) are useful tools for assessing the interrelationships among latent and observed variables. Due to their wide applications, LVMs have attracted significant attention from various fields. However, the majority of existing LVMs are parametric and have mainly considered continuous, homogeneous, and cross-sectional data. Given that mixed data types - continuous, count, ordinal, and nominal data, and multidimensional heterogeneous and longitudinal data are common in practice, and considering that the parametric LVMs are not adequate to reveal the true relationships between latent and observed variables, the development of more comprehensive LVMs to effectively deal with the abovementioned complex data is both necessary and important. In this study I will discuss a variety of generalized semiparametric LVMs. A Bayesian approach, together with advanced computational techniques, is developed to estimate unknown functions, model parameters, and latent variables in the models. The developed methodologies were applied to several medical and behavioral studies.

Key Words: Bayesian P-splines, latent variables, MCMC methods, parametric links.