Bayesian Lasso Factor Analysis Models with Ordered Categorical Data

Junhao Pan[1], Lijin Zhang[1] , and Edward Ip[2]

1 Department of Psychology, Sun Yat-sen University, Guangzhou, China
2 Wake Forest University School of Medicine, USA

***Abstract*:** Due to the design of questionnaires and the nature of problems in behavioral sciences, researchers often encounter ordered categorical data. We extend previous work on Bayesian Lasso confirmatory factor analysis (CFA) model to account for the local dependency among observed data coming from ordered categorical variables. We treated this kind of discrete data as observations that are coming from a hidden continuous normal distribution with a threshold specification, and proposed a Lasso prior that enables the entire inverse residual covariance matrix of the underlying normal variables to be modeled as a sparse positive definite matrix that contains only a few off-diagonal elements bounded away from zero. Based on parameter expansion and Markov Chain Monte Carlo procedures, we developed Bayesian method that not only achieves model parsimony and generally fits the data better, but also keeps the factor structure (that is, the number of factors and how the observed indicators are loaded on the factors) intact. Both simulated and real data sets were analyzed to evaluate the validity and practical usefulness of the proposed procedure. Due to the similarity of CFA with ordered categorical data and Item Response Theory (IRT), the proposed methodology has potential to deal the local dependence problem in IRT analysis.

***Keywords*:** Confirmatory factor analysis, Ordered categorical data, Bayesian Lasso prior, parameter expansion, MCMC