Bayesian regularization in multiple indicators multiple causes models

Lijin Zhang[1], Xinya Liang[2]

1. Department of Psychology, Sun Yat-sen University, China

2. Department of Rehabilitation, Human Resources and Communication Disorders, University of Arkansas, USA

Bayesian methods for machine learning such as regularization are gaining popularity in structural equation modeling. The main purpose of regularization is to improve the generalizability of variable selection and model estimation. Bayesian regularization assigns penalty priors to sparse matrices in which most elements are zero while the remaining elements are non-zero, in an attempt to regularize the posterior distributions of model parameters. In this study, we propose to use Bayesian regularization to investigate the impact of covariates in multiple indicators, multiple causes models. Five Bayesian regularization methods: ridge, lasso, adaptive lasso, spike and slab, and horseshoe, are examined and compared for covariate selection and parameter estimation. Sparse solutions are developed for the structural coefficient matrix that contains only a small portion of non-zero structural coefficients characterizing the effects of selected covariates on latent factors. Results from both the simulation study and empirical example show that comparing to diffuse priors, regularization is advantageous in handling small sample sizes and collinearity among covariates. Shrinkage priors with only the global penalty tend to yield higher model convergence rates and power but also greater Type I error rates than priors with both the global and local penalties. Implications and practical guidelines will be discussed.

Keywords: regularization, Bayesian analysis, MIMIC, structural equation modeling