One- vs. Two-Step Approach in Regularized Bayesian Structural Equation Modeling (SEM)

Research Proposal

Michael Koch (6412157)

Methodology and Statistics for the Behavioral, Biomedical, and Social Sciences

Supervisor: Dr. Sara van Erp

Email: j.m.b.koch@students.uu.nl

Word Count: 750

Intented Journal of Publication: Structural Equation Modeling

Structural Equation Modeling (SEM,)...

- Identification constraints
- Cross-loadings fixed to zero: $\mathcal{N}(0, 0)$
- If bad fit: Adjust model based on modification indices (flawed)

Bayesian SEM

- Bayesian SEM (Muthen & Asparouhov, 2012)
  - Don't just assume that cross-loadings are zero
  - Instead make a more realistic and flexible assumption:
    * Small Variance Prior, e.g.: $\mathcal{N}(0, 0.01)$
    * allowing for a more *continuous* model identification and model selection

$$\mathcal{N}(\mu = 0,\ \sigma = 0.01)$$

Regularization

In Regression and Machine Learning...

The general approach is to apply a penalty the loss-function of a model that...

$$argmin\{(|\mathbf{Y} - \mathbf{X}\beta|)^{\mathbf{2}} + \mathbf{\Sigma^{p}_{j=1}}|\beta^{\mathbf{q}}_{\mathbf{j}}|\}$$

In a Bayesian context, so-called *shrinkage priors* are used to achieve the same end (see Van Erp, Oberski, & Mulder, 2019 for an overview).

Lu, Chow, and Loken (2016) pointed out that ...

- Original approach with the small variance prior requires a 2-step approach (Lu et al., 2016) Not only the parameters close to zero are shrunken to zero (as desired)...

…but also the parameters that are far from zero, and should not be shrunken (introducing bias!) More advanced priors need to be identified that can outperform the small-variance prior *in a single step*

Regularized Horseshoe prior

Horseshoe prior (Carvalho, Polson, & Scott, 2010): Practically no shrinkage for large coefficients, shrinkage to zero for small coefficients *Global* component λ, shrinking all parameters towards zero, *local* component τ_j allowing large parameters to escape shrinkage

The *Regularized Horseshoe* prior builds on the horseshoe prior, by…

$$\beta_j|\tau_j^2, \sim \mathcal{N}(0, \tau_j^2)$$

$$\tau_j \sim C^+(0, \lambda) \ for \ j = 1, \ ..., \ p$$

$$\lambda|\sigma \sim C^+(0, \sigma)$$

Analytic Strategy

A Monte Carlo simulation study is conducted using stan (Carpenter et al., 2017). True Positives vs. False Positives in estimating truly non-zero cross-loadings as non-zero are considered as main outcome (ROC). Conditions are based on (Lu et al., 2016) and will include two factor structures (1 non-zero cross-loading, several non-zero cross-loadings), three sample sizes (100, 200, 300), and three magnitudes of the cross-loadings (0.1, 0.2, 0.3).

References

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., … Riddell, A. (2017). Stan: A Probabilistic Programming Language. *Journal of Statistical Software*, *76*(1), 1–32. https://doi.org/10.18637/jss.v076.i01

Carvalho, C. M., Polson, N. G., & Scott, J. G. (2010). The horseshoe estimator for sparse signals. *Biometrika*, *97*(2), 465–480. https://doi.org/10.1093/biomet/asq017

Lu, Z.-H., Chow, S.-M., & Loken, E. (2016). Bayesian Factor Analysis as a Variable-Selection Problem: Alternative Priors and Consequences. *Multivariate Behavioral Research*, *51*(4), 519–539. https://doi.org/10.1080/00273171.2016.1168279

Muthen, B., & Asparouhov, T. (2012). Bayesian SEM: A more flexible representation of substantive theory, 78. https://doi.org/10.1037/a0026802

Van Erp, S., Oberski, D. L., & Mulder, J. (2019). Shrinkage priors for Bayesian penalized regression. *Journal of Mathematical Psychology*, *89*, 31–50. https://doi.org/10.1016/j.jmp.2018.12.004

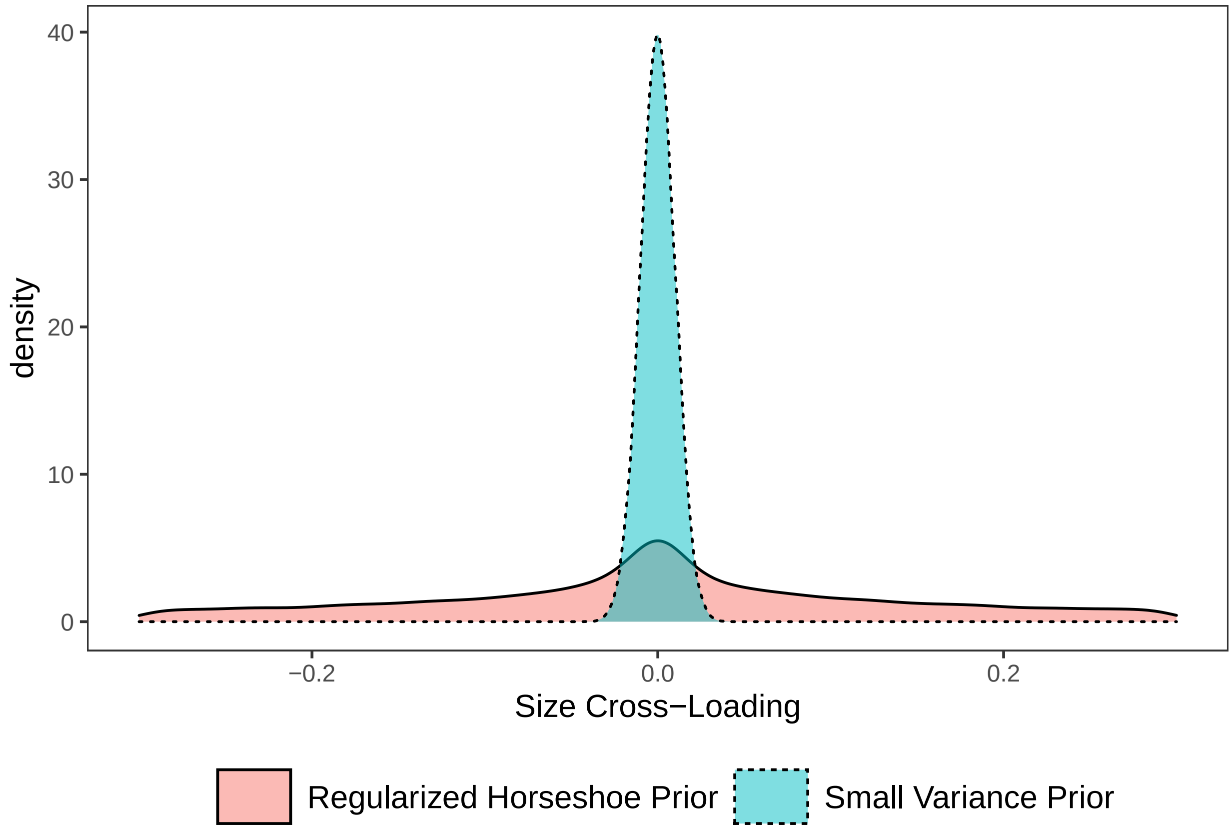

Figure 1. Density Plots of the Regularization Priors of Interest