Getting a Step Ahead: Using the Regularized Horseshoe Prior to Select Cross-Loadings in Bayesian CFA

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: 695

Intented Journal of Publication: Structural Equation Modeling

FETC permission obtained on 27'th September 2021

Confirmatory Factor Analysis (CFA) is an essential tool for modeling measurement structures. Usually, all cross-loadings, factor loadings that relate items to factors that they theoretically do not belong to, are fixed to zero to identify the model. This often leads to poor model fit, and forces researchers to free some cross-loadings again to improve fit, a practice that is flawed for a variety of reasons, among which risking capitalization on chance (MacCallum, Roznowski, & Necowitz, 1992). As solution, Muthen and Asparouhov (2012) proposed that rather than fixing *all* cross-loadings to zero, one should should assume that *most* cross-loadings are zero. Formally, this is achieved by setting the so-called *Small Variance Normal Prior* for the cross-loadings, which is a normal distribution with mean zero and a very small variance (e.g.: $\sigma^2 = 0.01,\ \sigma^2 = 0.001$). This prior has a large peak at zero, and very thin tails (see Figure 1). Hence, it attaches large prior mass to cross-loadings of or near zero, while attaching almost no prior mass to cross-loadings further from zero. Consequently, all cross-loadings in the model are shrunken. The larger the prior's variance, the more admissive the model is in the amount of deviation from zero it allows.

An issue with Muthen and Asparouhov (2012)'s approach is that not only the cross-loadings close to zero that are considered irrelevant are shrunken to zero, as desired, but also the ones further from zero are shrunken towards zero, which introduces bias. The method thus requires a two-step approach. First, the model is estimated with the Small Variance Normal Prior. Then the model is re-estimated, with cross-loadings that have been shrunken to zero in the previous step fixed to zero, and the remaining cross-loadings estimated without shrinkage. This process is tedious and computationally expensive. Therefore, alternative priors need to be identified that can outperform the Small Variance Normal Prior *in a single step.*

One prior that appears as particularly promising is the *Regularized Horseshoe Prior* (Piironen & Vehtari, 2017). This prior still has a sharp peak at zero resulting in shrinkage to zero for small coefficients (see Figure 1). However, it has much heavier tails, resulting in practically no shrinkage of larger coefficients. This prior is thus the best of both worlds, by only shrinking irrelevant parameters to zero, within one step (Piironen & Vehtari, 2017).

While the Regularized Horseshoe Prior has been shown to perform well in the selection of relevant predictors in regression (Piironen & Vehtari, 2017; Van Erp, Oberski, & Mulder, 2019), no previous research has validated its performance in selecting relevant cross-loadings in CFA. To fill this gap, the aim of this study is to compare the Regularized Horseshoe Prior to the Small Variance Normal Prior in their performance in selecting the true factor structure in CFA.

Analytic Strategy

A Monte Carlo simulation study is conducted using stan (Stan Development Team, 2021). As main outcomes, we consider the bias in the estimated cross-loadings, and the trade-off in true positives vs. false positives in correctly identifying non-zero cross-loadings as non-zero (ROC curves). Hereby, the main criterion in selecting cross-loadings as non-zero is whether their 95% HPD interval contains zero (Zhang, Pan, & Ip, 2021).

We include two factor structures, one with a single, and one with three non-zero cross-loadings. Each model consists of three factors a three items. Factors are scaled by setting their means to zero and their variances to one. The correlations between all factors is set to 0.5, and the residual variance of all items to 0.3. We vary between three sample sizes (100, 200, 300), and three magnitudes of the cross-loadings (0.1, 0.2, 0.3).

Following Muthen and Asparouhov (2012) we consider three values for the hyperparameter of the Small Variance Normal Prior ($\sigma^2 = 0.001, 0.01, 0.1$). The Regularized Horseshoe Prior has five hyperparameters, and varying all of them broadly would lead to an unfeasible number of conditions. We will therefore conduct a pilot study on a single simulated dataset to identify combinations of hyperparameters that are worth to consider in the main study.

All models will be sampled using the No-U-Turn-Sampler (Homan & Gelman, 2014). We aim to identify a suitable chain length in the pilot study.

References

Homan, M. D., & Gelman, A. (2014). The No-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *The Journal of Machine Learning Research*, *15*(1), 1593–1623.

MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*, *111*(3), 490–504. https://doi.org/10.1037/0033-2909.111.3.490

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

Piironen, J., & Vehtari, A. (2017). Sparsity information and regularization in the horseshoe and other shrinkage priors. *Electronic Journal of Statistics*, *11*(2), 5018–5051. https://doi.org/10.1214/17-EJS1337SI

Stan Development Team. (2021). Stan User Guide. Retrieved from https://mc-stan.org/docs/2_27/stan-users-guide-2_27.pdf

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

Zhang, L., Pan, J., & Ip, E. H. (2021). Criteria for Parameter Identification in Bayesian Lasso Methods for Covariance Analysis: Comparing Rules for Thresholding, p -value, and Credible Interval. *Structural Equation Modeling: A Multidisciplinary Journal*, 1–10. https://doi.org/10.1080/10705511.2021.1945456

Additional References

Betancourt, Michael. 2018. "A Conceptual Introduction to Hamiltonian Monte Carlo." ArXiv:1701.02434 [Stat].

Carpenter, Bob, Andrew Gelman, Matthew D. Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2017. "Stan: A Probabilistic Programming Language." Journal of Statistical Software 76(1):1–32. doi: 10.18637/jss.v076.i01.

Carvalho, Carlos M., Nicholas G. Polson, and James G. Scott. 2010. "The Horseshoe Estimator for Sparse Signals." Biometrika 97(2):465–80. doi: 10.1093/biomet/asq017.

Ghosh, Joyee, Yingbo Li, and Robin Mitra. 2018. "On the Use of Cauchy Prior Distributions for Bayesian Logistic Regression." Bayesian Analysis 13(2):359–83. doi: 10.1214/17-BA1051.

Hoerl, Arthur E., and Robert W. Kennard. 2000. "Ridge Regression: Biased Estimation for Nonorthogonal Problems." Technometrics 42(1):80–86. doi: 10.2307/1271436.

Hsiang, T. C. 1975. "A Bayesian View on Ridge Regression." Journal of the Royal Statistical Society. Series D (The Statistician) 24(4):267–68. doi: 10.2307/2987923.

Jacobucci, Ross, Kevin J. Grimm, and John J. McArdle. 2016. "Regularized Structural Equation Modeling." Structural Equation Modeling: A Multidisciplinary Journal 23(4):555–66. doi: 10.1080/10705511.2016.1154793.

Lu, Zhao-Hua, Sy-Miin Chow, and Eric Loken. 2016. "Bayesian Factor Analysis as a Variable-Selection Problem: Alternative Priors and Consequences." Multivariate Behavioral Research 51(4):519–39. doi: 10.1080/00273171.2016.1168279.

Merkle, Edgar C., Ellen Fitzsimmons, James Uanhoro, and Ben Goodrich. 2020. "Efficient Bayesian Structural Equation Modeling in Stan." ArXiv:2008.07733 [Stat].

Morris, Tim P., Ian R. White, and Michael J. Crowther. 2019. "Using Simulation Studies to Evaluate Statistical Methods." Statistics in Medicine 38(11):2074–2102.

doi: 10.1002/sim.8086.

Park, Trevor, and George Casella. 2008. “The Bayesian Lasso.” Journal of the American Statistical Association 103(482):681–86. doi: 10.1198/016214508000000337.

Tibshirani, Robert. 1996. “Regression Shrinkage and Selection Via the Lasso.” Journal of the Royal Statistical Society: Series B (Methodological) 58(1):267–88. doi: 10.1111/j.2517-6161.1996.tb02080.x.

Tibshirani, Robert. 2011. “Regression Shrinkage and Selection via the Lasso: A Retrospective.” Journal of the Royal Statistical Society. Series B (Statistical Methodology) 73(3):273–82.

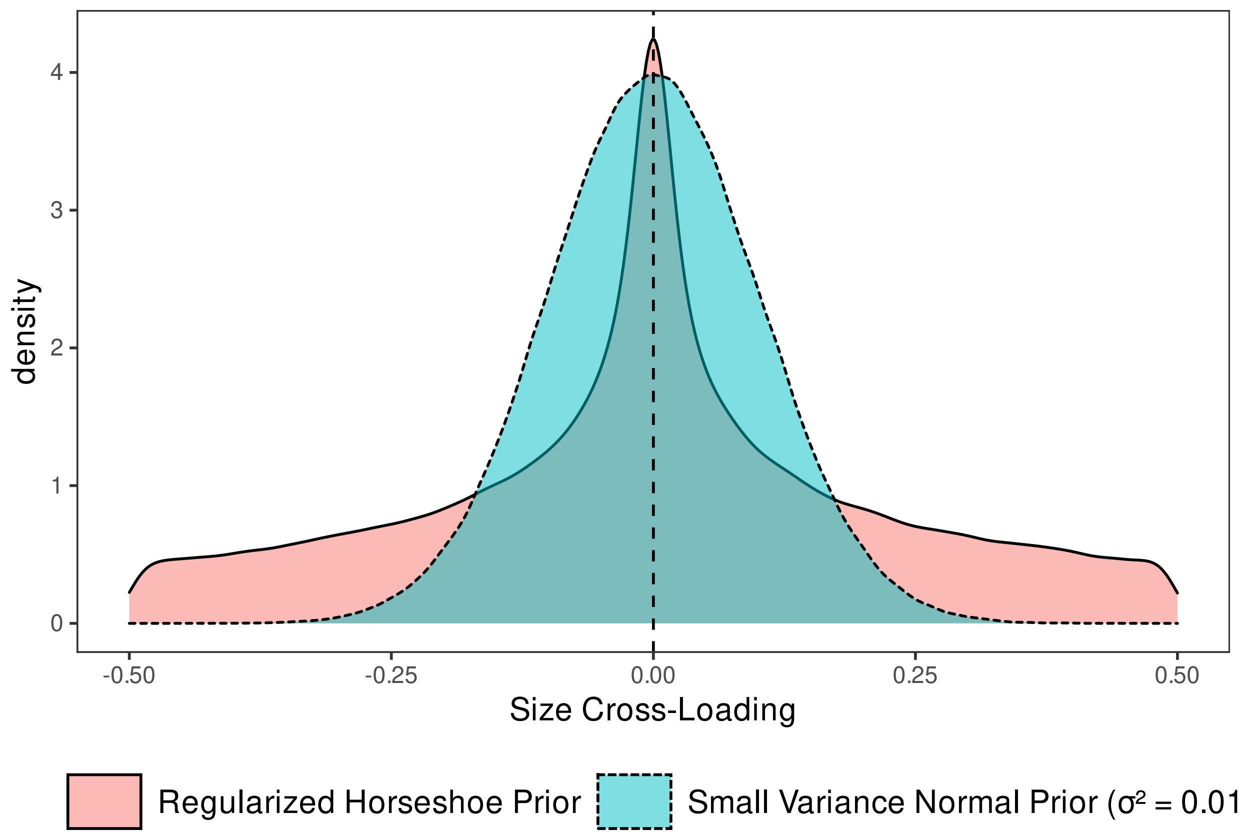

Figure 1. Density Plots of the Regularization Priors of Interest