The Journal of Open Source Software

minorbsem: An R package for structural equation models that account for the influence of minor factors

James Ohisei Uanhoro[1]

1 Department of Educational Psychology, University of North Texas, USA

DOI: 10.21105/joss.05292

Software

- Review
- Repository
- Archive

Editor: Samuel Forbes

Reviewers:

- @smasongarrison
- @aaronpeikert

Submitted: 08 March 2023
Published: 21 June 2023

License
Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License (CC BY 4.0).

Summary

minorbsem is an R package that allows users to fit Bayesian structural equation models (SEMs) under the assumption that minor factors influence the relationship between observed indicators. SEMs are models that capture the inter-relations between factors (latent variables) and the relations between these factors and observed items. However, SEMs are rarely able to reject the null hypothesis that there is no misspecification in the hypothesized model. One explanation for this problem is that covariance structures are influenced by major factors that we can hypothesize about, and minor factors that we cannot predict a-priori (MacCallum & Tucker, 1991).

minorbsem accounts for the influence of minor factors by estimating all residual covariances between the observed items (Uanhoro, 2023a). These residual covariances are estimated with priors that shrink the residual covariances towards zero. Additionally, the model returns the magnitude of the influence of minor factors. By estimating the influence of minor factors simultaneously with major factors, the parameters of interest (related to major factors) in the SEM will reflect the uncertainty due to the influence of minor factors.

minorbsem also fits random-effects meta-analytic confirmatory factor analysis models (CFAs) that capture the influence of minor factors (Uanhoro, 2023b).

Features

- Fits SEMs allowing latent regressions, cross-loadings, and correlated error terms.
- Fits CFAs allowing cross-loadings, correlated error terms, and fully oblique or orthogonal factors.
- Can estimate the size of the influence of minor factors on the SEM.
- Provides several priors to shrink residual covariances towards zero.
- Fits random-effects meta-analytic CFAs that capture the influence of minor factors.

Limitations

minorbsem assumes multivariate normal data and only fits a limited number of model configurations. At the moment, it does not support multiple-indicator multiple-cause models, multi-group models, multilevel models, or models with specially constrained parameters.

Statement of need

B. O. Muthén & Asparouhov (2012) showed the flexibility of Bayesian SEMs, and sparked interest in Bayesian SEMs that account for the influence of minor factors. The work of Wu

& Browne (2015) in frequentist SEMs also addresses the same substantive problem from a different angle. However, there are limited software implementations of these ideas.

Mplus (L. K. Muthén & Muthén, 1998--2017) is a proprietary software package that is also able to estimate Bayesian models that account for the influence of minor factors. While Mplus can estimate a very large number of model configurations, the user has to specify the size of the influence of minor factors manually, which can lead to inadequate uncertainty estimates for parameters of interest (Uanhoro, 2023a). Additionally, Mplus is not a free software option.

`LAWBL` (Chen, 2022) is an R package with similar functionality to minorbsem. LAWBL accommodates more data types than minorbsem e.g. binary data models are available. However, LAWBL only estimates CFA models and its syntax is not a commonly used syntax. Additionally, LAWBL does not return the magnitude of the influence of minor factors.

Finally, neither Mplus nor LAWBL allow the range of options for modeling the influence of minor factors that are implemented in minorbsem. Nor do they offer options for fitting random-effects meta-analytic SEMs that also capture the influence of minor factors.

Dependencies

minorbsem relies on the Stan project (Carpenter et al., 2017) for fitting Bayesian models, and interfaces with Stan via the `cmdstanr` package. Installation instructions for both Stan and minorbsem are available in the minorbsem README.

minorbsem relies on the `lavaan` package (Rosseel, 2012) for data processing. Moreover, minorbsem models are specified using lavaan-style syntax which allows for easy model specification.

Additionally, it relies on `Rcpp` to speed up costly operations (Eddelbuettel & François, 2011), the `huxtable` package for printing results (Hugh-Jones, 2022), the `posterior` package for accessing Stan model results (Bürkner et al., 2023), and `ggplot2` for plotting (Wickham, 2016).

References

Bürkner, P.-C., Gabry, J., Kay, M., & Vehtari, A. (2023). *Posterior: Tools for working with posterior distributions*. https://mc-stan.org/posterior/

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*(1). https://doi.org/10.18637/jss.v076.i01

Chen, J. (2022). *LAWBL: Latent (variable) analysis with Bayesian learning*. https://CRAN.R-project.org/package=LAWBL

Eddelbuettel, D., & François, R. (2011). Rcpp: Seamless R and C++ integration. *Journal of Statistical Software*, *40*(8), 1–18. https://doi.org/10.18637/jss.v040.i08

Hugh-Jones, D. (2022). *Huxtable: Easily create and style tables for LaTeX, HTML and other formats*. https://CRAN.R-project.org/package=huxtable

MacCallum, R. C., & Tucker, L. R. (1991). Representing sources of error in the common-factor model: Implications for theory and practice. *Psychological Bulletin*, *109*(3), 502–511. https://doi.org/10.1037/0033-2909.109.3.502

Muthén, B. O., & Asparouhov, T. (2012). Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychological Methods*, *17*(3), 313–335. https://doi.org/10.1037/a0026802

Muthén, L. K., & Muthén, B. O. (1998--2017). *Mplus user's guide* (Eighth, p. 950). Muthén & Muthén.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Uanhoro, J. O. (2023a). Modeling misspecification as a parameter in Bayesian structural equation modeling. *Educational and Psychological Measurement*, *0*(0). https://doi.org/10.1177/00131644231165306

Uanhoro, J. O. (2023b). Hierarchical covariance estimation approach to meta-analytic structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *30*(4), 532–546. https://doi.org/10.1080/10705511.2022.2142128

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. ISBN: 978-3-319-24277-4

Wu, H., & Browne, M. W. (2015). Quantifying adventitious error in a covariance structure as a random effect. *Psychometrika*, *80*(3), 571–600. https://doi.org/10.1007/s11336-015-9451-3