The Journal of Open Source Software

GGMnonreg: Non-Regularized Gaussian Graphical Models in R

Donald R. Williams[1]

1 Department of Psychology, University of California, Davis

DOI: 10.21105/joss.03308

Software

- Review
- Repository
- Archive

Editor: Arfon Smith

Reviewers:

- @AlexChristensen
- @GiulioCostantini

Submitted: 03 April 2021
Published: 22 May 2021

Summary

Studying complex relations in multivariate datasets is a common task across the sciences. Cognitive neuroscientists model brain connectivity with the goal of unearthing functional and structural associations between cortical regions (Ortiz et al., 2015). In clinical psychology, researchers wish to better understand the intricate web of symptom interrelations that underlie mental health disorders (Borsboom et al., 2011; McNally, 2016). To this end, graphical modeling has emerged as an oft-used tool in the chest of scientific inquiry. The basic idea is to characterize multivariate relations by learning the conditional dependence structure. The cortical regions or symptoms are *nodes* and the featured connections linking nodes are *edges* that graphically represent the conditional dependence structure.

Graphical modeling is quite common in fields with wide data, that is, when there are more variables (p) than observations (n). Accordingly, many regularization-based approaches have been developed for those kinds of data. There are key drawbacks of regularization, including, but not limited to, the fact that obtaining a valid measure of parameter uncertainty is very (very) difficult (Bühlmann et al., 2014) and there can be an inflated false positive rate (see for example, Donald R. Williams et al., 2019).

More recently, graphical modeling has emerged in psychology (Epskamp et al. 2018), where the data is typically long or low-dimensional ($p < n$; Donald R. Williams et al. (2019), Donald R. Williams & Rast (2019)). The primary purpose of GGMnonreg is to provide methods that were specifically designed for low-dimensional data (e.g., those common in the social-behavioral sciences).

Statement of Need

The following were designed specifically for low-dimensional data, for which there is a dearth of methodology.

Supported Models

- Gaussian graphical model. The following data types are supported.
    - Gaussian
    - Ordinal
    - Binary
- Ising model (Marsman et al., 2017)
- Mixed graphical model

The Journal of Open Source Software

Additional methods

The following are also included

- Expected network replicability (Donald R. Williams, 2020)
- Compare Gaussian graphical models
- Measure of parameter uncertainty (Donald R. Williams et al., 2019)
- Edge inclusion "probabilities" (e.g., Figure 6.4 in Hastie et al., 2015)
- Network visualization
- Constrained precision matrix (the network, given an assumed graph, see p. 631 in Hastie et al., 2009)
- Predictability (variance explained for each node, Haslbeck & Waldorp, 2018)

Implementation

```
# make binary
Y <- ifelse(ptsd[,1:5] == 0, 0, 1)

# fit model
fit <- ising_search(Y, IC = "BIC",
                    progress = FALSE)

fit
#>          1        2        3        4        5
#> 1 0.000000 1.439583 0.000000 1.273379 0.000000
#> 2 1.439583 0.000000 1.616511 0.000000 1.182281
#> 3 0.000000 1.616511 0.000000 1.716747 1.077322
#> 4 1.273379 0.000000 1.716747 0.000000 1.662550
#> 5 0.000000 1.182281 1.077322 1.662550 0.000000
```

Note the same code, more or less, is also used for GGMs and mixed graphical models.

Network Visualization

A key aspect of graphical modeling is visualizing the conditional dependence structure. To this end, **GGMnonreg** makes network plots with **ggplot2** (Wickham, 2016).

```
plot(get_graph(fit),
     node_names = colnames(Y),
     edge_magnify = 2)
```

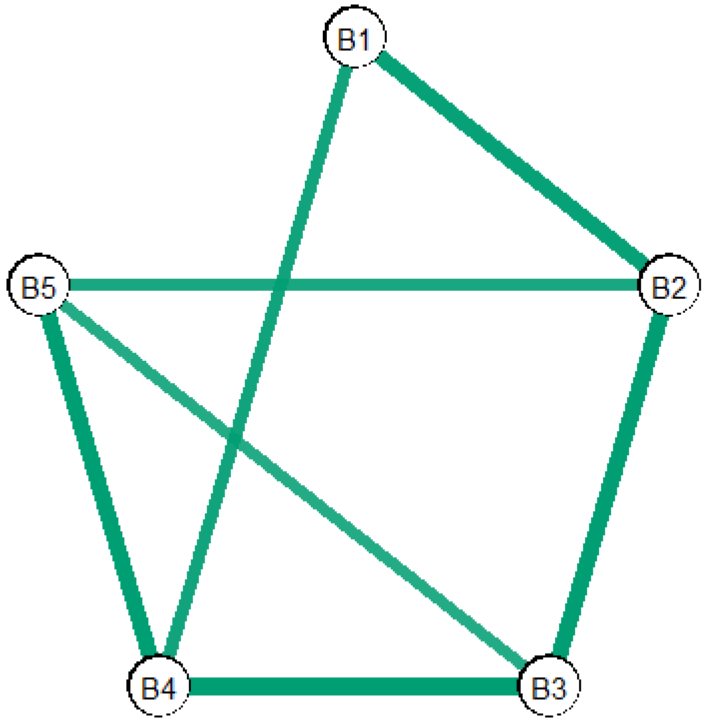

Figure 1: Conditional Dependence Structure

Acknowledgements

DRW was supported by a National Science Foundation Graduate Research Fellowship under Grant No. 1650042

References

Borsboom, D., Cramer, A. O. J., Schmittmann, V. D., Epskamp, S., & Waldorp, L. J. (2011). The Small World of Psychopathology. *PLoS ONE*, *6*(11), e27407. https://doi.org/10.1371/journal.pone.0027407

Bühlmann, P., Kalisch, M., & Meier, L. (2014). High-Dimensional Statistics with a View Toward Applications in Biology. *Annual Review of Statistics and Its Application*, *1*(1), 255–278. https://doi.org/10.1146/annurev-statistics-022513-115545

Haslbeck, J. M., & Waldorp, L. J. (2018). How well do network models predict observations? On the importance of predictability in network models. *Behavior Research Methods*, *50*(2), 853–861. https://doi.org/10.3758/s13428-017-0910-x

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. Springer Science & Business Media.

Hastie, T., Tibshirani, R., & Wainwright, M. (2015). *Statistical Learning with Sparsity: The Lasso and Generalizations* (p. 362). CRC Press. https://doi.org/10.1201/b18401

Marsman, M., Borsboom, D., Kruis, J., Epskamp, S., Bork, R. van, Waldorp, L. J., Taylor, B., Waldorp, L., J van der Maas, H. L., & Maris, G. (2017). An Introduction to Network

Psychometrics: Relating Ising Network Models to Item Response Theory Models. *Taylor & Francis*, *53*(1), 15–35. https://doi.org/10.1080/00273171.2017.1379379

McNally, R. J. (2016). *Can network analysis transform psychopathology?* (Vol. 86, pp. 95–104). Elsevier Ltd. https://doi.org/10.1016/j.brat.2016.06.006

Ortiz, A., Munilla, J., Álvarez-Illán, I., Górriz, J. M., & Ramírez, J. (2015). Exploratory graphical models of functional and structural connectivity patterns for Alzheimer's disease diagnosis. *Frontiers in Computational Neuroscience*, *9*(November), 1–18. https://doi.org/10.3389/fncom.2015.00132

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. ISBN: 978-3-319-24277-4

Williams, Donald R. (2020). Learning to live with sampling variability: Expected replicability in partial correlation networks. *PsyArXiv*. https://doi.org/10.31234/osf.io/fb4sa

Williams, Donald R., & Rast, P. (2019). Back to the basics: Rethinking partial correlation network methodology. *British Journal of Mathematical and Statistical Psychology*. https://doi.org/10.1111/bmsp.12173

Williams, Donald R., Rhemtulla, M., Wysocki, A. C., & Rast, P. (2019). On nonregularized estimation of psychological networks. *Multivariate Behavioral Research*, *54*(5), 719–750. https://doi.org/10.1080/00273171.2019.1575716