Regularized Gaussian Psychological Networks: Brief Report on the Performance of Extended BIC Model Selection

Epskamp, S.

Citation for published version (APA):
Recent years have seen a emergence of the network conceptualisation of psychology, in which relationships between attitudes, moods, clinical symptoms and other observed psychological variables are seen as interacting components in a dynamical system, rather than indicators of one or more latent constructs\cite{Cramer2010, Borsboom2011, Schmittmann2013}. The models proposed take the form of networks, in which nodes represent observed variables which are connected by edges representing statistical relationships between these variables\cite{Epskamp2012}. These models strikingly differ from typically used network models such as social networks\cite{Wasserman1994} or transportation networks\cite{Newman2010}, in that variables are not static entities (e.g., people or cities) but random variables, and links are not observed (e.g., friendships or roads) but need to be estimated\cite{Epskamp2016}.

When data is assumed multivariate normal distributed, a prominent, interpretable and easy to use network model is the Gaussian Graphical Model (GGM; Lauritzen, 1996), a network in which edges represent partial correlations between two variables after conditioning on all other variables in the network. Such networks are extensively being applied to psychological datasets (e.g., McNally et al., 2015; Kossakowski et al., 2015; Ivoranu et al., 2016; Fried et al., 2016; van Borkulo et al., 2015). To control for spurious relationships a regularization technique called the ‘least absolute shrinkage and selection operator’ (LASSO; Tibshirani, 1996) is often used\cite{Costantini2015}. The graphical LASSO (glasso; Friedman et al., 2008) is a particularly fast variant of the LASSO that only requires a covariance matrix. As psychological data are often ordinal, an estimate of the covariance matrix can be obtained by computing polychoric and polyserial correlations\cite{Olsson1979, Olsson1982}, which can be used in the glasso algorithm.

LASSO regularization utilizes a tuning parameter, $\lambda$, which controls the sparsity of the network. Typically, a range of networks is estimated under different values of $\lambda$ (Zhao & Yu, 2006). The value for $\lambda$ under which no edges are retained (the empty network), $\lambda_{\text{max}}$, is set to the largest absolute correlation\cite{Zhao2015}. Next, a minimum value can be chosen by multiplying some ratio $R$ (typically set to 0.01 or 0.1) with this maximum value:

$$\lambda_{\text{min}} = R\lambda_{\text{max}}.$$  

A logarithmically spaced range of tuning parameters (typically 100 different values), ranging from $\lambda_{\text{min}}$ to $\lambda_{\text{max}}$, can be used to estimate different networks. Subsequently, an optimal network with many true connections and few spurious connections can be obtained through model selection\cite{Drton2004}. The network that has the least cross-validation prediction error or the lowest value of some information criterion is often the selected network. The extended Bayesian Information Criterion (EBIC; Chen and Chen, 2008; Foygel and Drton, 2010) adds an extra penalty for model complexity to the typical BIC and has been shown to work well in high-dimensional network model selection\cite{Foygel2010, Barber2015, vanBorkulo2014}. The EBIC uses a hyperparameter, $\gamma$, which controls the extra penalization; $\gamma = 0$ leads to the EBIC reducing to the BIC, and higher values of $\gamma$ lead to more penalization. Typically, $\gamma$ is set between 0 and 1. I will shorten EBIC selection of GGM models using LASSO regularization via the glasso algorithm to GeLasso\footnote{The term GeLasso is line with van Borkulo et al. (2014), who use eLasso in the context of estimating a pairwise markov random field for binary variables.}.

While GeLasso has already been shown to work well in retrieving the GGM structure\cite{Foygel2010, Drton2010} who suggest $\gamma = 0.5$, it has not been validated in plausible scenarios for psychological networks. In addition, no simulation study has assessed the performance of using a poly-
can be interpreted as partial correlation coefficients. Green edges indicate positive partial correlations, red edges indicate negative partial correlations and the wider and more saturated the edge the stronger the correlation [Epskamp et al., 2012]. The network was obtained by computing the (unregularized) sample partial correlation network and removing all absolute edges below 0.05.

In the simulation study, data was generated based on the network of Figure 1. Following, the network was estimated using the EBICglasso function in the qgraph package [Epskamp et al., 2012]. Sample size was varied between 50, 100, 250, 500, 1,000, and 2,500, $\gamma$ was varied between 0, 0.25, 0.5, 0.75, and 1, and $R$ was varied between 0.001, 0.01 and 0.1. The data was either simulated to be multivariate normal, in which case Pearson correlations were used in estimation, or ordinal, in which case polychoric correlations were used in the estimation. Ordinal data was created by sampling four thresholds for every variable from the standard normal distribution, and next using these thresholds to cut each variable in five levels. To compute polychoric correlations, the cor_auto function was used, which uses the lavaan function of the lavaan package [Rosseel, 2012]. The number of different $\lambda$ values used in generating networks was set to 100 (the default in qgraph).

For each simulation, in addition to the correlation between estimated and true edge weights, the sensitivity and specificity were computed [van Borkulo et al., 2014]. The sensitivity, also termed the true-positive rate, indicates the proportion of edges in the true network that were estimated to be nonzero:

$$\text{sensitivity} = \frac{\text{# true positives}}{\text{# true positives} + \text{# of false negatives}}.$$ 

Specificity, also termed the true negative rate, indicates the proportion of true missing edges that were also estimated to be missing:

$$\text{specificity} = \frac{\text{# true negatives}}{\text{# true negatives} + \text{# false positives}}.$$ 

When specificity is high, there are not many false positives (edges detected to be nonzero that are zero in the true network) in the estimated network.

**Results**

Each of the conditions was replicated 1,000 times, leading to 180,000 simulated datasets. Figure 2 shows the sensitivity of the analyses. This figure shows that sensitivity increases with sample size and is high for large sample sizes. When $\gamma > 0$, small sample sizes are likely to result in empty networks (no edges), indicating a sensitivity of 0. When ordinal data is used, small sample sizes (50 and 100) resulted in far too densely connected networks that are hard to interpret.
Setting γ to be higher remediated this by estimating empty networks. At higher sample sizes, γ does not play a role and sensitivity is comparable in all conditions. Using R = 0.1 remediates the poor performance of polychoric correlations in lower sample sizes, but also creates an upper bound to sensitivity at higher sample sizes.

Figure 2 shows the specificity of the analyses, which was all-around high except for the lower sample sizes in ordinal data using R = 0.01 or R = 0.001. Some outliers indicate that fully connected networks were estimated in ordinal data even when setting γ = 0.25 in small sample sizes. In all other conditions specificity was comparably high, with higher γ values only performing slightly better. Figure 3 shows the correlation between true and estimated edge weights. This figure shows a comparable good performance from sample sizes of 250 and higher in all conditions, with γ values up to 0.5 outperforming the higher γ values. It should be noted that the correlation was set to zero if the estimated network had no edges (all edge weights were then zero).

Conclusion

In this brief report I assessed the performance of GeLasso in 180,000 simulated datasets using a plausible psychological network structure. Results indicate that GeLasso performs well in estimating psychological networks using both Pearson correlations or polychoric correlations. The default setup of qgraph uses γ = 0.5 and R = 0.01, which are shown to work well in all conditions. Setting γ = 0.25 improved the detection rate, but sometimes led to poorly estimated networks based on polychoric correlations. γ can be set to 0 to err more on the side of discovery (Dziak et al. 2012), but should be done with care in low sample polychoric correlation matrices. All conditions showed increasing sensitivity with sample size and a high specificity all-around. This is comparable to other network estimation techniques (van Borkulo et al. 2014), and shows that even though a network does not contain all true edges, the edges that are returned can usually be expected to be genuine. The high correlation furthermore indicated that the strongest true edges are usually estimated to be strong as well.

The estimation of psychological networks is a rapidly evolving field of research. In addition to the GeLasso method many other network analysis methods exists (e.g., Zhao et al., 2015 [Kraemer et al., 2009] [Kalsch et al., 2012]. When variables are binary, a more appropriate model to use is the Ising Model (van Borkulo et al., 2014). In addition, new and promising methods have been developed for estimating network structures with mixed continuous and categorical vari-
Figure 3. The specificity of the simulated datasets. When specificity is high, there are not many edges in the estimated network that are not present in the true network. See caption of Figure 2 for more details.

Figure 4. Correlation between true edge weights and estimated edge weights. See caption of Figure 2 for more details.
For a detailed tutorial on both using the GeLasso method and on assessing the stability of such network structures, I refer the reader to Epskamp et al. (2016).

Acknowledgements

I would like to thank Lourens J. Waldorp, Eiko I. Fried and Adela M. Isvoranu for their helpful comments and feedback.

References


