Markov Random Fields II

Network Analysis 2017

Sacha Epskamp

20-11-2017

Markov random fields

Markov random fields are undirected networks, of which the edges can be interpreted in several ways:

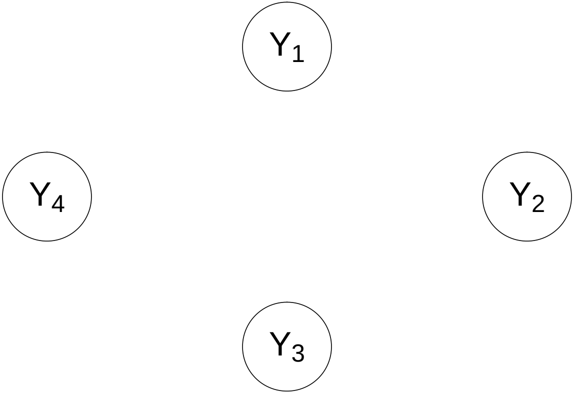
1. A representation of conditional independence relationships
2. Pairwise interactions
3. Highlight potential causal pathways
4. Highlights latent variables as clusters
5. Predictive effects
6. Genuine symmetric relationships between nodes
    - Ising Model

Gaussian graphical models and Ising models

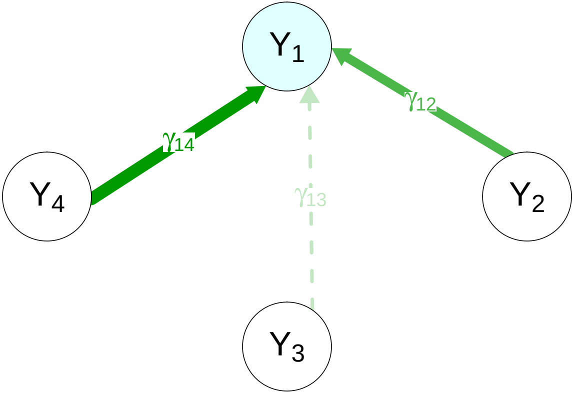
We discussed two main types of MRFs:

- Gaussian graphical model
  - Continuous (normal) variables
  - Edge-weights equal to partial correlation coefficients
  - Obtained from:
    - Inverting and standardizing the variance–covariance matrix
    - Multiple regression models for each variable
- Ising Model
  - Binary variables
  - Edge-weights also indicate conditional association
  - Obtained from multiple logistic regression models for each variable

Gausian Graphical Model

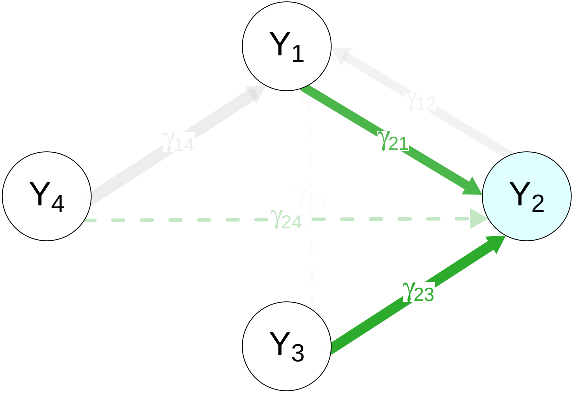


Gausian Graphical Model



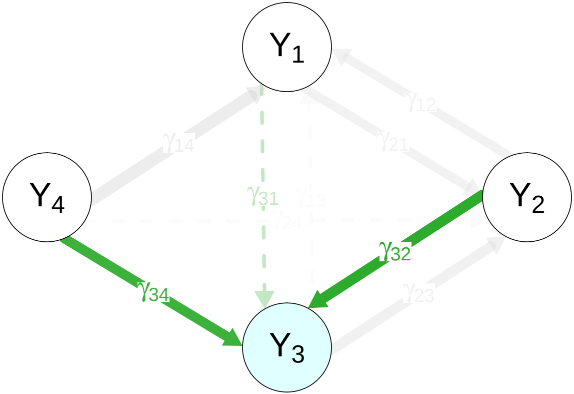
$$y_1 = \tau_1 + \gamma_{12} y_2 + \gamma_{13} y_3 + \gamma_{14} y_4 + \varepsilon_1$$

Gausian Graphical Model



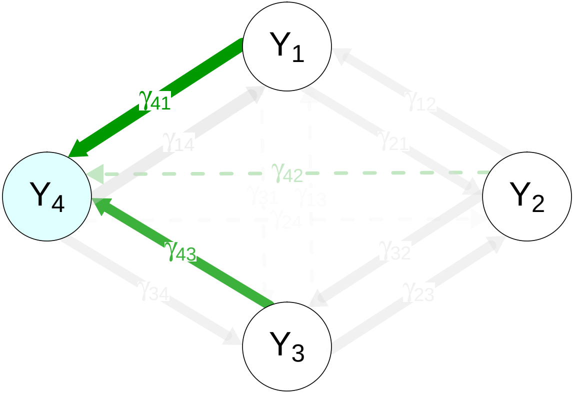
$$y_2 = \tau_2 + \gamma_{21} y_1 + \gamma_{23} y_3 + \gamma_{24} y_4 + \varepsilon_2$$

Gausian Graphical Model



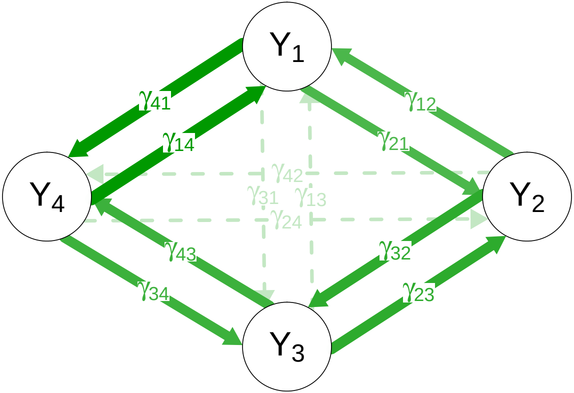
$$y_3 = \tau_3 + \gamma_{31} y_1 + \gamma_{32} y_2 + \gamma_{34} y_4 + \varepsilon_3$$

Gausian Graphical Model

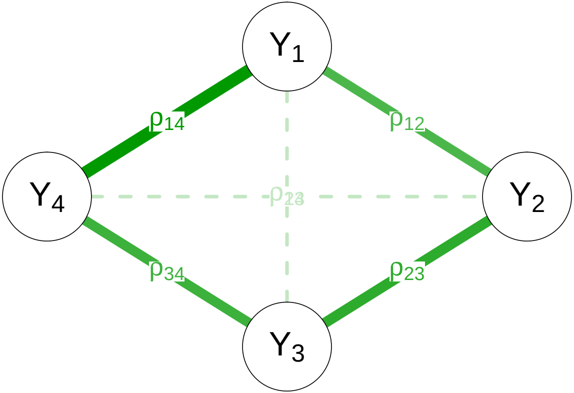


$$y_4 = \tau_4 + \gamma_{41} y_1 + \gamma_{42} y_2 + \gamma_{43} y_3 + \varepsilon_4$$

Gausian Graphical Model

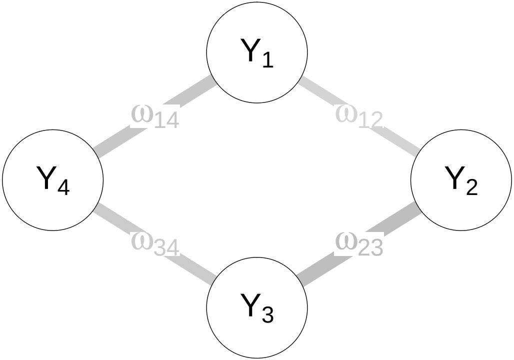


Gausian Graphical Model

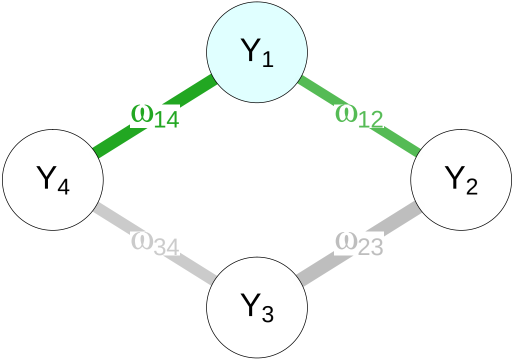


$$\rho_{ij} = \frac{\gamma_{ij}\mathrm{SD}\left(\varepsilon_j\right)}{\mathrm{SD}\left(\varepsilon_i\right)} = \frac{\gamma_{ji}\mathrm{SD}\left(\varepsilon_i\right)}{\mathrm{SD}\left(\varepsilon_j\right)} = -\frac{\kappa_{ij}}{\sqrt{\kappa_{ii}}\sqrt{\kappa_{jj}}}$$

Ising Model

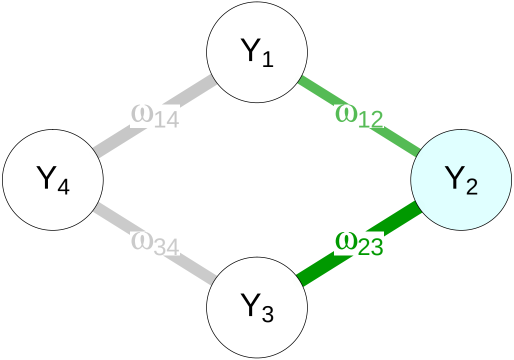


Ising Model



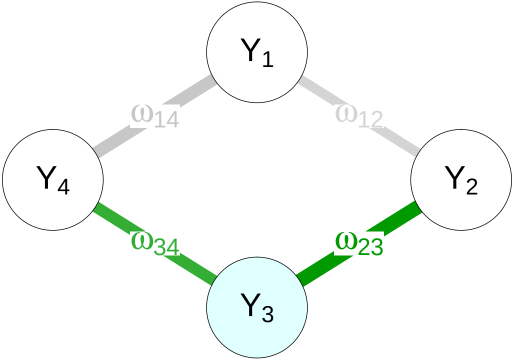
$$\Pr\left(X_1 = 1\right) \propto \exp\left(\tau_1 + \omega_{12}x_2 + \omega_{14}x_4\right)$$

Ising Model



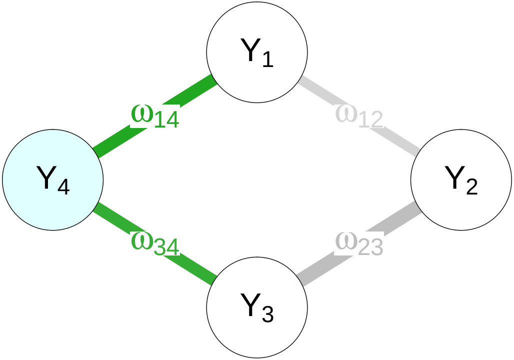
$$\Pr\left(X_2 = 1\right) \propto \exp\left(\tau_2 + \omega_{12} x_1 + \omega_{23} x_3\right)$$

Ising Model



$$\Pr\left(X_3 = 1\right) \propto \exp\left(\tau_3 + \omega_{23}x_2 + \omega_{34}x_4\right)$$

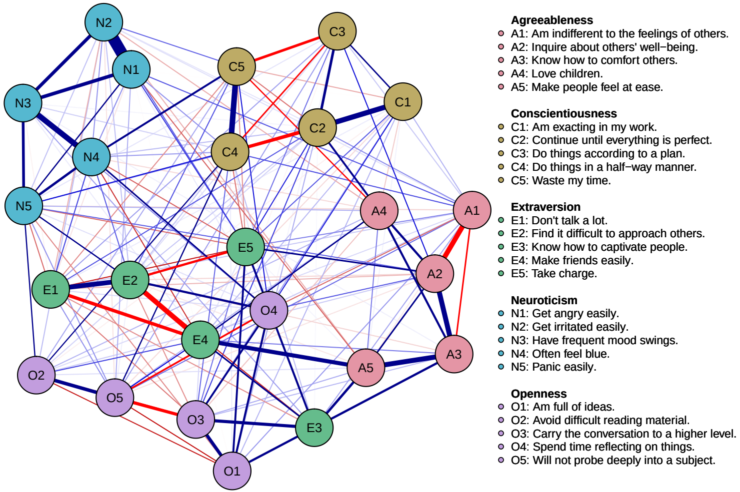
Ising Model



$$\Pr(X_4 = 1) \propto \exp(\tau_4 + \omega_{14} x_1 + \omega_{34} x_3)$$

Today

- Control for spurious connections
  - LASSO regularization
  - EBIC model selection
- Assess stability and accuracy of results
  - Bootstrap
- Compare networks
  - Permutation test (NetworkComparisonTest)
- Handle non-normal data
  - Non-normal continuous data
  - Ordinal data
  - Categorical data


N2
N1
N3
N4
N5
C3
C5
C1
C2
C4
A4
A1
A2
A3
A5
E5
E2
E1
E4
E3
O4
O2
O5
O3
O1
Agreeableness
A1: Am indifferent to the feelings of others.
A2: Inquire about others' well−being.
A3: Know how to comfort others.
A4: Love children.
A5: Make people feel at ease.
Conscientiousness
C1: Am exacting in my work.
C2: Continue until everything is perfect.
C3: Do things according to a plan.
C4: Do things in a half−way manner.
C5: Waste my time.
Extraversion
E1: Don't talk a lot.
E2: Find it difficult to approach others.
E3: Know how to captivate people.
E4: Make friends easily.
E5: Take charge.
Neuroticism
N1: Get angry easily.
N2: Get irritated easily.
N3: Have frequent mood swings.
N4: Often feel blue.
N5: Panic easily.
Openness
O1: Am full of ideas.
O2: Avoid difficult reading material.
O3: Carry the conversation to a higher level.
O4: Spend time reflecting on things.
O5: Will not probe deeply into a subject.

Not hard in R...

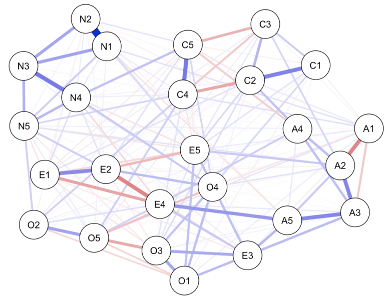
Console | Compile PDF

~/Dropbox/Work2/R/SummerSchool2017/

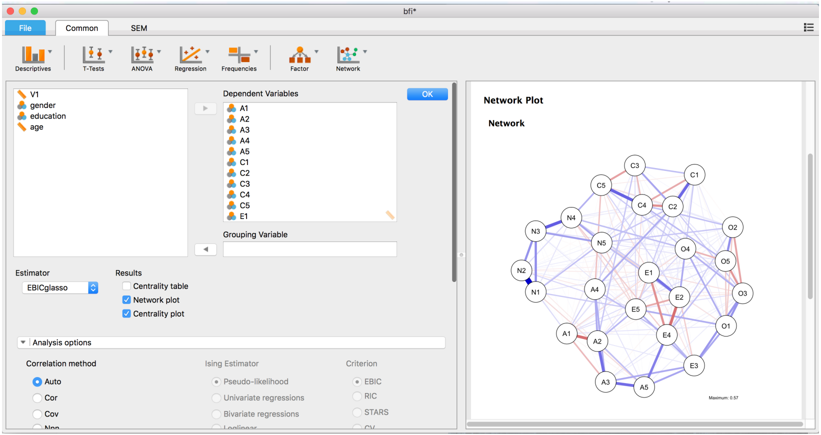
```
>
>
>
>
>
> # Load bootnet:
> library("bootnet")
> # Estimate network:
> Results <- estimateNetwork(Data, default = "EBICglasso"
)
Estimating Network. Using package::function:
  - qgraph::EBICglasso for EBIC model selection
    - using glasso::glasso
  - qgraph::cor_auto for correlation computation
    - using lavaan::lavCor
Variables detected as ordinal: A1; A2; A3; A4; A5; C1; C2
; C3; C4; C5; E1; E2; E3; E4; E5; N1; N2; N3; N4; N5; O1;
O2; O3; O4; O5
> # Plot network:
> plot(Results, cut = 0, theme = "colorblind", layout = "
spring")
>
```

Files | Plots | Packages | Help | Viewer

Zoom | Export | Publish



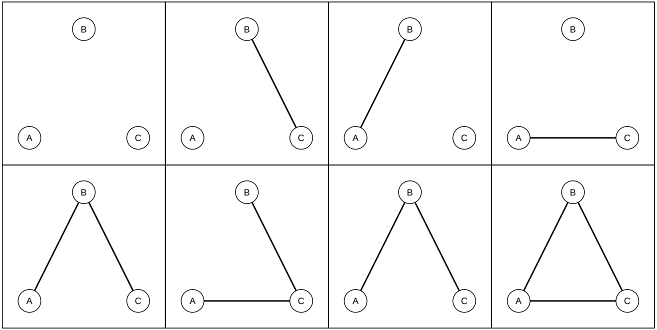
And becoming easier in Jasp! (see `jasp-stats.org/`)



Bootnet estimation

Default	Model	Data	Estimation	Penalty	Selection	Package
pcor	GGM	Continuous, ordinal	Inverse	None	None, sig, FDR	qgraph
EBICglasso	GGM	Continuous, ordinal	Inverse	glasso	EBIC	qgraph
IsingFit	Ising	Binary	Nodewise	LASSO	EBIC	IsingFit
IsingSampler	Ising	Binary	Nodewise, MLE	None	None	IsingSampler
adalasso	GGM	Continuous	Nodewise	Adaptive LASSO	CV	parcor
huge	GGM	Continuous	Inverse	glasso	EBIC	huge
mgm	MGM	Continuous, count, binary, categorical	Nodewise	LASSO	EBIC, CV	mgm

CV = Cross validation; Inverse = Inverse covariance matrix; MLE = Maximum (pseudo) likelihood

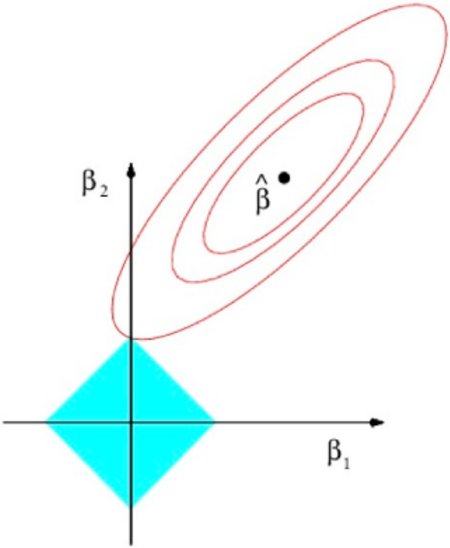


- How to select the best model?

A well accepted approach for jointly model selection and parameter estimation is the least absolute shrinkage and selection operator (LASSO)

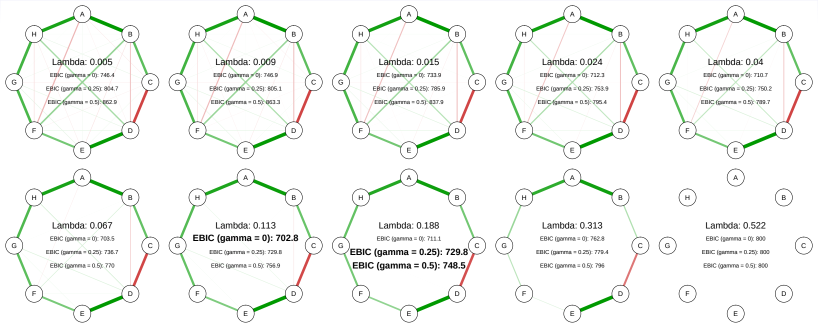
- Limits the sum of absolute regression weights, which causes insignificant edges to shrink to zero
  - Regularization
- Useful in generalized linear regression model!
  - Multiple linear regression
  - Multiple logistic regression
- Even usable when you have more variables then measures!
- Requires a tuning parameter, λ that specifies the sparsity, which need to be selected carefully

LASSO Estimation

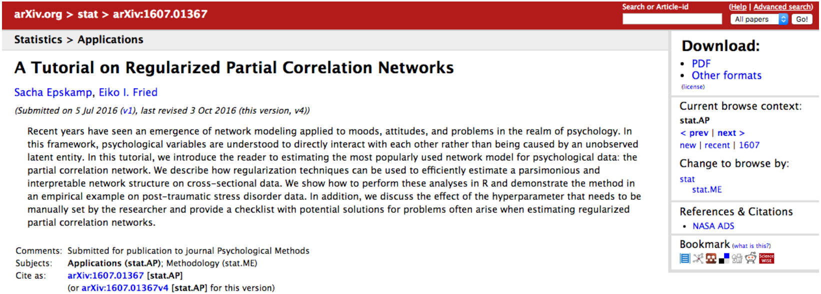


LASSO estimation

- *Nodewise estimation*
  - Perform regularized (logistic) regressions for each node
  - Two (non-identical) estimates per edge, average to obtain single estimate
  - AND-rule or OR-rule: include edge if one or both estimates are nonzero
- The GGM can also be estimated in a single model using the *graphical LASSO* (glasso)
  - Directly penalizes elements of inverse variance–covariance matrix



- Varying λ leads to a range of networks: model selection needed
  - Minimize cross-validation prediction error
  - Minimize an information criterion
- We often minimize the extended BIC (EBIC)
  - Hyperparameter γ needed to set between 0 (err on side of discovery) or 0.5 (err on side of caution). Commonly set to 0.25 or 0.5.



arXiv.org > stat > arXiv:1607.01367

Search or Article-id (Help | Advanced search)

All papers Go!

Statistics > Applications

A Tutorial on Regularized Partial Correlation Networks

Sacha Epskamp, Eiko I. Fried

(Submitted on 5 Jul 2016 (v1), last revised 3 Oct 2016 (this version, v4))

Recent years have seen an emergence of network modeling applied to moods, attitudes, and problems in the realm of psychology. In this framework, psychological variables are understood to directly interact with each other rather than being caused by an unobserved latent entity. In this tutorial, we introduce the reader to estimating the most popularly used network model for psychological data: the partial correlation network. We describe how regularization techniques can be used to efficiently estimate a parsimonious and interpretable network structure on cross-sectional data. We show how to perform these analyses in R and demonstrate the method in an empirical example on post-traumatic stress disorder data. In addition, we discuss the effect of the hyperparameter that needs to be manually set by the researcher and provide a checklist with potential solutions for problems often arise when estimating regularized partial correlation networks.

Comments: Submitted for publication to journal Psychological Methods
Subjects: **Applications (stat.AP)**; Methodology (stat.ME)
Cite as: arXiv:1607.01367 [stat.AP]
(or arXiv:1607.01367v4 [stat.AP] for this version)

Download:

- PDF
- Other formats

(license)

Current browse context:
stat.AP
< prev | next >
new | recent | 1607

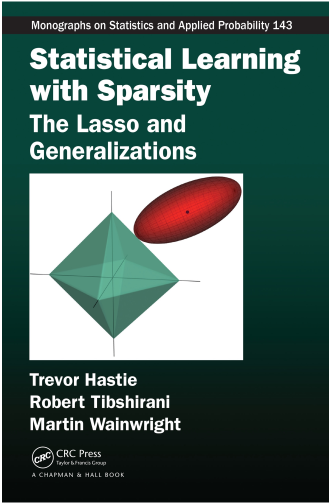
Change to browse by:
stat
stat.ME

References & Citations
- NASA ADS

Bookmark (what is this?)

- https://arxiv.org/abs/1607.01367

More on LASSO estimation

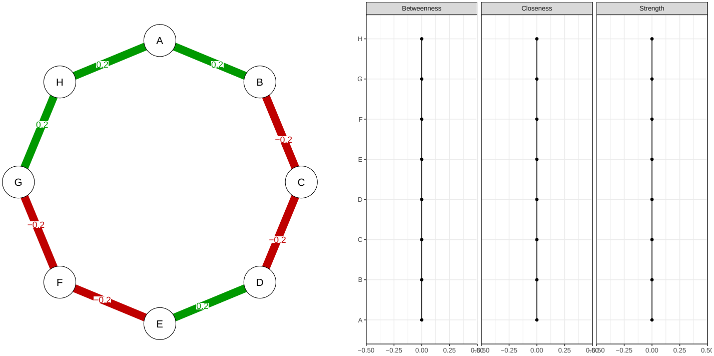


`https://web.stanford.edu/~hastie/StatLearnSparsity/`

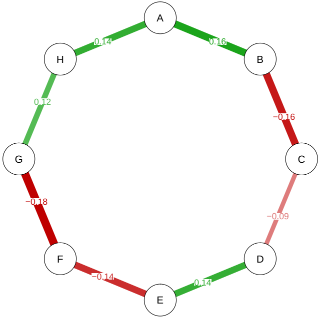
The bet on Sparsity

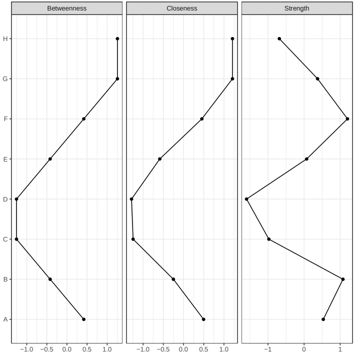
- The optimal Markov random field describing the data can be sparse
  - Contain elements that are zero
- Thus, estimating a sparse simplifies the model
- In high-dimensional cases, a variance–covariance matrix can *not* be inverted, but a sparse inverse *can* be obtained!
- This is crucial for many high-dimensional computations
- Likewise, high-dimensional Ising models allow for a powerful characterization of the joint likelihood of binary variables without evoking latent variables
- LASSO searches such a sparse model, but relies on an assumption that the true model is sparse: the bet on sparsity

True model:


A
B
C
D
E
F
G
H
0.2
−0.2
Betweenness
Closeness
Strength
−0.50
−0.25
0.00
0.25
0.50

Based on sample of $N = 500$:





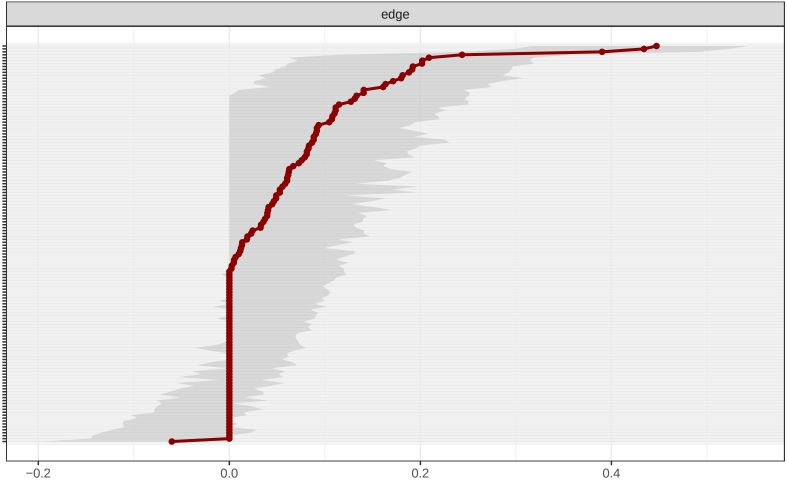
Accuracy and Stability

- Estimated network structures are subject to sampling variation
- Thus, care needs to be taken in interpreting differences between edges or descriptive measures (e.g., centrality)
- We propose bootstrapping methods to gain insight in the stability of parameter estimates
  - Epskamp, S., Borsboom, D., & Fried, E. I. (2017). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*. doi:10.3758/s13428
- Two bootstraps
  - Nonparametric bootstrap (re-sampling datasets of same N with replacement)
  - Subset bootstrap (sampling subsets of cases without replacement)

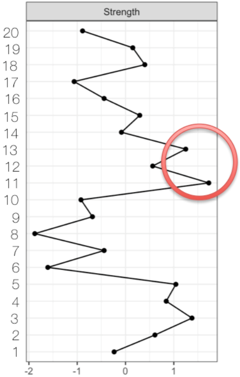
Non-parametric bootstrap

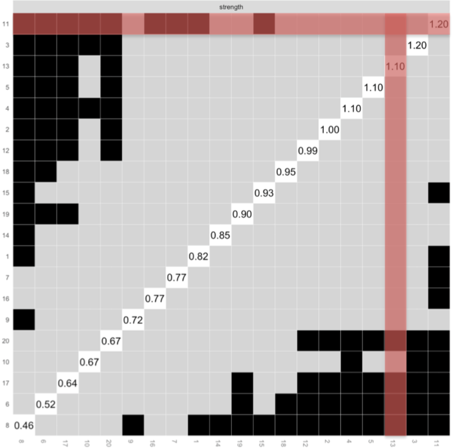
- The non-parametric *bootstrap* is a well-known data-driven approach to investigate sampling variation
  - Efron, B. (1992). Bootstrap methods: another look at the jackknife. In Breakthroughs in statistics (pp. 569-593). Springer, New York, NY.
- 1. Compute some statistic from your data (e.g., edge-weight)
- 2. Generate a new dataset by sampling cases from your original data *with* replacement
- 3. Use these new datasets to estimate a range of the statistic
- 4. Use these ranges to draw confidence intervals
- The bootstrap samples can also be used to test for *differences* between parameters

Confidence-intervals



Differences

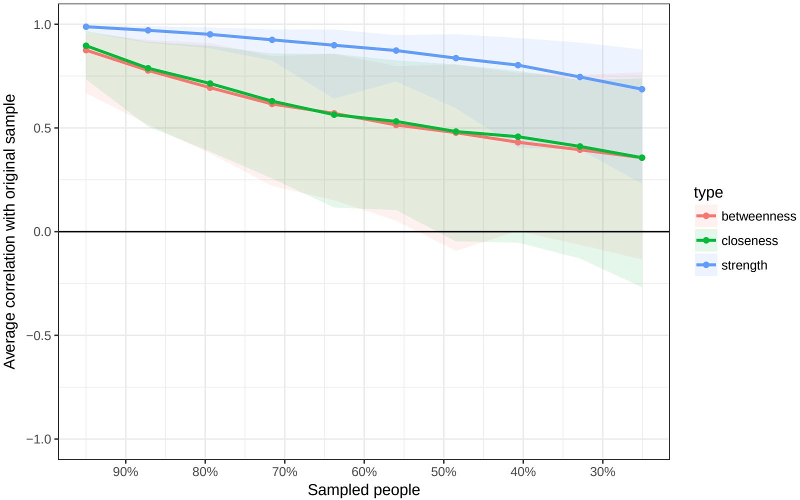




Case-drop bootstrap

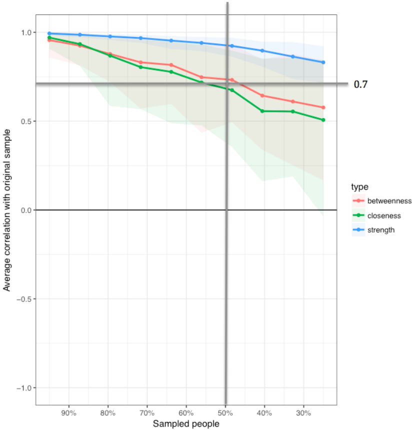
- CIs cannot be formed for centrality indices using the nonparametric bootstrap
- We proposed case-dropping bootstrap:
  - Drop x% of the cases (people) at random
  - Compute a network and derive centrality indices
  - Correlate obtained centrality indices with the original centrality indices
- Ideally, we would want centrality to remain comparable to the original network even after dropping many cases from the dataset!

Stability

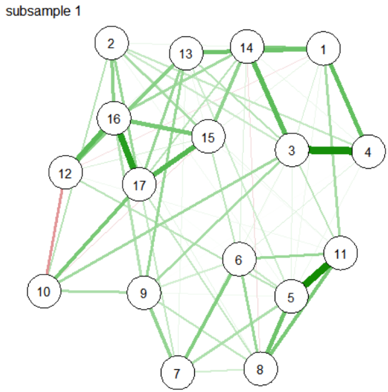

1.0
0.5
0.0
−0.5
−1.0
Average correlation with original sample
90%
80%
70%
60%
50%
40%
30%
Sampled people
type
betweenness
closeness
strength

CS-coefficient

- The *correlation stability coefficient* (CS-coefficient) can be used to quantify the case-drop bootstrap
- the proportion of data that can be dropped to retain with 95% certainty a correlation of at least 0.7 with the original centrality coefficients
- Preferably above 0.5, and should not be below 0.25
  - Although these recommendations are just as arbitrary as $\alpha < 0.05$



Network Comparison



Do these networks differ?

Comparing networks

Steps to assessing if networks differ:

- Step 1: visually inspect the networks
  - Make sure the layouts are equal! (e.g., use `averageLayout`) from qgraph
  - Keep differences in sample size in mind, the less *n* the sparser the network!
- Step 2: correlate the weights matrices
- Step 3: permutation test
  - NetworkComparisonTest R package
  - Van Borkulo, C. D., Boschloo, L., Kossakowski, J., Tio, P., Schoevers, R., Borsboom, D., & Waldorp, L. (2016). Comparing network structures on three aspects: A permutation test.

Permutation Test

- A *permutation test* can be used to test if statistics from two groups differ
    1. Compute a statistic of interest in both groups
    2. Pool all cases in one large dataset
    3. Randomly create new groups by re-distributing the cases
    4. Compute the statistic in each new pair of groups to obtain a null-distribution
    5. Test if the observed difference is in the null-distribution
- `https://www.researchgate.net/publication/314750838_Comparing_network_structures_on_three_aspects_A_permutation_test`

Network Comparison Test

1. Network structure invariance hypothesis
   - Structure is completely identical across subpopulations
2. Global strength invariance hypothesis
   - Overall level of connectivity is identical across subpopulations
3. Edge strength invariance hypothesis
   - A specific edge is identical across subpopulations

Evidence That Psychopathology Symptom Networks Have Limited Replicability

Miriam K. Forbes
University of Minnesota

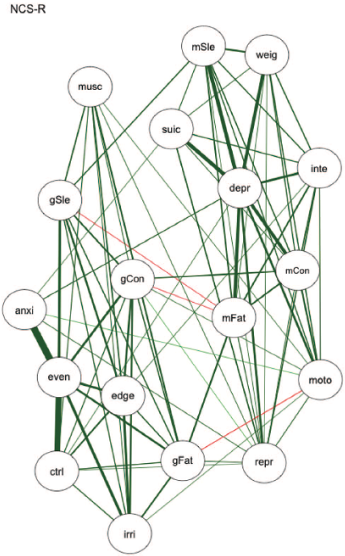
Aidan G. C. Wright
University of Pittsburgh

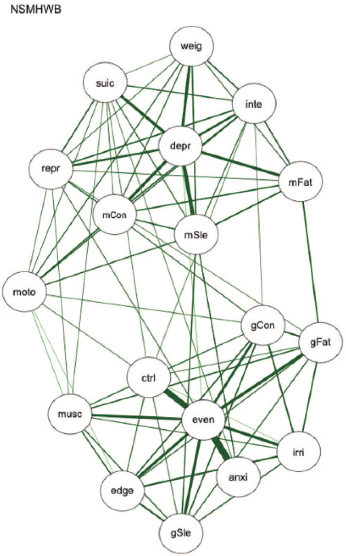
Kristian E. Markon
University of Iowa

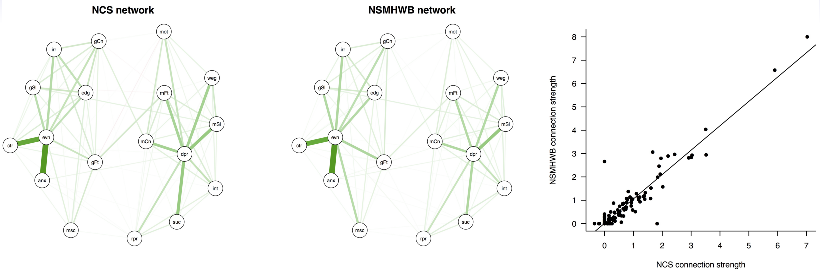
Robert F. Krueger
University of Minnesota

Network analysis is quickly gaining popularity in psychopathology research as a method that aims to reveal causal relationships among individual symptoms. To date, 4 main types of psychopathology networks have been proposed: (a) association networks, (b) regularized concentration networks, (c) relative importance networks, and (d) directed acyclic graphs. The authors examined the replicability of these analyses based on symptoms of major depression and generalized anxiety between and within 2 highly similar epidemiological samples (i.e., the National Comorbidity Survey—Replication [n = 9282] and the National Survey of Mental Health and Wellbeing [n = 8841]). Although association networks were stable, the 3 other types of network analysis (i.e., the conditional independence networks) had poor replicability between and within methods and samples. The detailed aspects of the models—such as the estimation of specific edges and the centrality of individual nodes—were particularly unstable. For example, 44% of the symptoms were estimated as the "most influential" on at least 1 centrality index across the 6 conditional independence networks in the full samples, and only 13–21% of the edges were consistently estimated across these networks. One of the likely reasons for the instability of the networks is the predominance of measurement error in the assessment of individual symptoms. The authors discuss the implications of these findings for the growing field of psychopathology network research, and conclude that novel results originating from psychopathology networks should be held to higher standards of evidence before they are ready for dissemination or implementation in the field.

osf.io/xcfdq/


NCS-R
mSle
weig
musc
suic
inte
depr
gSle
gCon
mCon
anxi
mFat
even
edge
moto
ctrl
gFat
repr
irri


NSMHWB
weig
suic
inte
depr
repr
mFat
mCon
mSle
moto
gCon
gFat
ctrl
musc
even
irri
anxi
edge
gSle



- Weights matrices correlate 0.95
- NetworkComparisonTest revealed no statistical differences between global network structure and individual edges
- Borsboom, D., Fried, E. I., Epskamp, S., Waldorp, L. J., Van Borkulo, C. D., Van der Maas, H. L. J., & Cramer, A. O. J. (in press). False Alarm? A comprehensive reanalysis of "Evidence that psychopathology symptom networks have limited replicability" by Forbes, Wright, Markon, and Krueger. *Journal of Abnormal Psychology.*
  - `psyarxiv.com/z49tk`

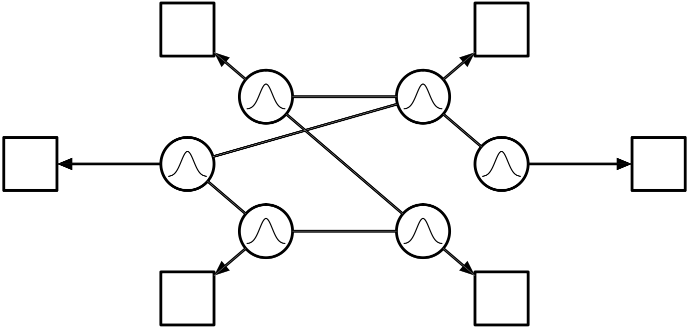
PSYCH NETWORKS

Organized Incoherence by Eiko Fried

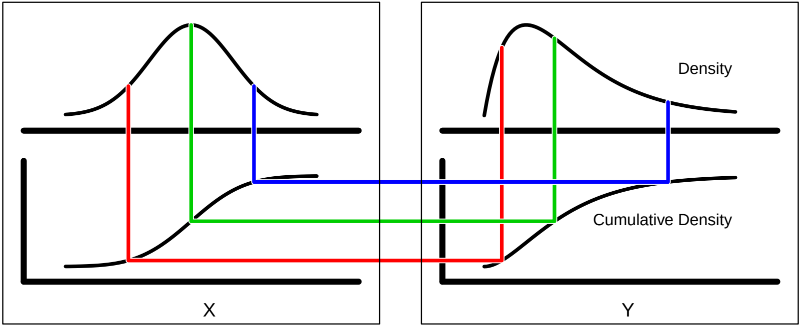
- Read the full story on Eiko's blog
  - `http://psych-networks.com/network-models-do-not-replicate-not/`
- Eiko also wrote a great paper that *actually* investigates network replicability!
  - `https://osf.io/2t7qp/`

Non-normal non-binary data

For categorical or count variables: mixed graphical models. For non-normal continuous and ordinal: transformation of GGM.

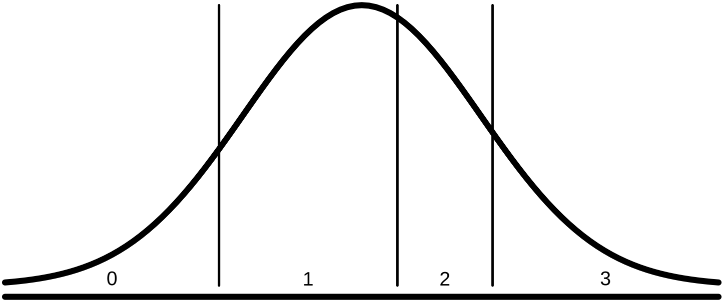


Non-normal continuous data



Non-paranormal transformation (`huge.npn` in huge package).

Ordinal data



Polychoric correlations as input to GGM (`lavCor` in lavaan package).

Frequent questions

- Network is *very* dense or contains ridiculously strong (possibly negative) edge weights
  - Check if polychoric correlation matrix is positive definite
- Negative edges where you expect positive ones
  - Could be real! Colliders in the data can make edges negative
  - Could also be spurious, did you condition on a function of the data? For example, do not split the data on the sumscore!
- Does a pretty network mean that the latent variable model is false?
  - No, especially clusters in a network can arise due to latents!

`arxiv.org/abs/1607.01367`

Estimation

- GGM can be computed using *qgraph*
- Ordinal data
  - Use polychoric correlations (`cor_auto` in *qgraph*) as input
- Non-normal continuous data
  - Transform variables first (`huge.npn` in *huge*)
- Binary data
  - Use *IsingFit*
- Mixed variables (Gaussian, Poisson, binary or categorical)
  - Use *mgm*

The *bootnet* function contains a wrapper function, `estimateNetwork` for these packages.

Estimating a MRF using `estimateNetwork` from *bootnet*:

```
# Load bootnet:
library("bootnet")

# Estimate network (see ?estimateNetwork):
Results <- estimateNetwork(Data, default = "...")

# Obtain weights matrix:
Results$graph

# Plot network (same arguments as qgraph):
plot(Results, layout = "spring")

# Centrality:
library("qgraph")
centralityPlot(Results)
```

Bootnet estimation

Default	Model	Data	Estimation	Penalty	Selection	Package
pcor	GGM	Continuous, ordinal	Inverse	None	None, sig, FDR	qgraph
EBICglasso	GGM	Continuous, ordinal	Inverse	glasso	EBIC	qgraph
IsingFit	Ising	Binary	Nodewise	LASSO	EBIC	IsingFit
IsingSampler	Ising	Binary	Nodewise, MLE	None	None	IsingSampler
adalasso	GGM	Continuous	Nodewise	Adaptive LASSO	CV	parcor
huge	GGM	Continuous	Inverse	glasso	EBIC	huge
mgm	MGM	Continuous, count, binary, categorical	Nodewise	LASSO	EBIC, CV	mgm

CV = Cross validation; Inverse = Inverse covariance matrix; MLE = Maximum (pseudo) likelihood

Take-home message

- Regularization controls for spurious connection
  - LASSO regularization
  - EBIC model selection
- Bootstrap methods assess accuracy and stability of results
  - Non-parametric bootstrap
  - Case-drop bootstrap
- Comparing networks takes three steps
  - Visually inspect; Correlate weights; Permutation test (NetworkComparisonTest)
- Non-normal data
  - Non-paranormal transformation
  - Polychoric correlations

Thank you for your attention!