Network Visualizations of Relationships in Psychometric Data and Structural Equation Models

Using the qgraph package for R

Sacha Epskamp

University of Amsterdam
Department of Psychological Methods

IMPS 2013
Look at your data...
Simulated data

<table>
<thead>
<tr>
<th>var 1</th>
<th>0.37</th>
<th>-0.26</th>
<th>0.25</th>
<th>-0.26</th>
<th>0.23</th>
<th>0.52</th>
<th>0.26</th>
<th>-0.46</th>
</tr>
</thead>
<tbody>
<tr>
<td>var 2</td>
<td>-0.17</td>
<td>0.24</td>
<td>0.29</td>
<td>-0.35</td>
<td>0.29</td>
<td>0.26</td>
<td>0.57</td>
<td>-0.29</td>
</tr>
<tr>
<td>var 3</td>
<td>-0.40</td>
<td>-0.29</td>
<td>0.23</td>
<td>-0.33</td>
<td>-0.29</td>
<td>-0.29</td>
<td>0.40</td>
<td></td>
</tr>
<tr>
<td>var 4</td>
<td>0.46</td>
<td>-0.25</td>
<td>0.34</td>
<td>0.33</td>
<td>0.33</td>
<td>0.26</td>
<td>-0.19</td>
<td></td>
</tr>
<tr>
<td>var 5</td>
<td>-0.42</td>
<td>0.38</td>
<td>0.33</td>
<td>0.15</td>
<td>-0.33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>var 6</td>
<td>-0.24</td>
<td>-0.37</td>
<td>-0.32</td>
<td>0.34</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>var 7</td>
<td>0.37</td>
<td>0.21</td>
<td>-0.31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>var 8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.30</td>
</tr>
<tr>
<td>var 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.36</td>
</tr>
<tr>
<td>var 10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Simulated data
Real data example

Included in qgraph is a dataset in which the Dutch translation of a commonly used personality test, the NEO-PI-R (Costa & McCrae, 1992; Hoekstra, de Fruyt, & Ormel, 2003), was administered to 500 first year psychology students (Dolan, Oort, Stoel, & Wicherts, 2009). The NEO-PI-R consists of 240 items designed to measure the five central personality factors:

- Neuroticism
- Extroversion
- Agreeableness
- Openness to Experience
- Conscientiousness
Neuroticism correlations

The image depicts a network of correlations between different variables labeled N1, N6, N11, ..., N236. The network appears to consist of a matrix with nodes connected by lines, indicating correlation strengths or relationships. The specific values or relationships are not legible in the image provided.
Big 5 correlations

-1.0  -0.5  0.0  0.5  1.0
When extended to the amount of variables commonly used in tests:
  - Scatterplot and ellipse visualizations become unreadable
  - Heatmaps plots do better, but only show trends
    - Very hard, if not impossible, to see important violations of trends

We need another way to look at our data
The Network Approach

Basic idea:

- Nodes represent variables
  - Possible to vary in color, shape, size and label to indicate different statistics
- Edges represent relationships
  - Green edges indicate positive relationships
  - Red edges indicate negative relationships
  - The wider and more saturated an edge, the stronger the absolute relationship
Implied correlations

CFA model

Covariance structure
Implied correlations

CFA model

\[ \psi_{11} \]
\[ \eta_1 \]
\[ \lambda_{11}(y) \]
\[ \lambda_{21}(y) \]
\[ \lambda_{31}(y) \]
\[ y_1 \]
\[ y_2 \]
\[ y_3 \]
\[ \theta_{11}^{(e)} \]
\[ \theta_{22}^{(e)} \]
\[ \theta_{33}^{(e)} \]

Implied covariance structure

\[ \lambda_{11}^2 \psi_{11} + \varepsilon_1 \]
\[ \lambda_{21} \lambda_{31} \psi_{11} \]
\[ \lambda_{31}^2 \psi_{11} + \varepsilon_3 \]
\[ \lambda_{31} \lambda_{21} \psi_{11} \]
\[ \lambda_{21}^2 \psi_{11} + \varepsilon_2 \]
Implied correlations

\[ \Sigma = \Lambda \psi \Lambda^T + \Theta \]

\[
\begin{bmatrix}
\sigma_1^2 & \sigma_{12} & \sigma_{13} \\
\sigma_{21} & \sigma_2^2 & \sigma_{23} \\
\sigma_{31} & \sigma_{32} & \sigma_3^2
\end{bmatrix}
= \begin{bmatrix}
\lambda_1 \\
\lambda_2 \\
\lambda_3
\end{bmatrix}
\begin{bmatrix}
\psi_{11}
\end{bmatrix}
\begin{bmatrix}
\lambda_1 & \lambda_2 & \lambda_3
\end{bmatrix}
+ \begin{bmatrix}
\theta_{11} & 0 & 0 \\
0 & \theta_{22} & 0 \\
0 & 0 & \theta_{33}
\end{bmatrix}
\]

\[
= \begin{bmatrix}
\lambda_1^2 \psi_{11} + \theta_{11} & \lambda_1 \lambda_2 \psi_{11} & \lambda_1 \lambda_3 \psi_{11} \\
\lambda_2 \lambda_1 \psi_{11} & \lambda_2^2 \psi_{11} + \theta_{22} & \lambda_2 \lambda_3 \psi_{11} \\
\lambda_3 \lambda_1 \psi_{11} & \lambda_3 \lambda_2 \psi_{11} & \lambda_3^2 \psi_{11} + \theta_{33}
\end{bmatrix}
\]
Implied correlations

\( \psi \) can be set to 1 without loss of information.

\[
\begin{pmatrix}
\lambda_1^2 + \theta_{11} & \lambda_1 \lambda_2 & \lambda_1 \lambda_3 \\
\lambda_2 \lambda_1 & \lambda_2^2 + \theta_{22} & \lambda_2 \lambda_3 \\
\lambda_3 \lambda_1 & \lambda_3 \lambda_2 & \lambda_3^2 + \theta_{33}
\end{pmatrix}
\]
Implied correlations

- The single factor model implies a very distinct correlational pattern:
  - Two variables are correlated highly if, and only if, they both load strongly on the factor
  - Because of this, if one factor underlies, for example, four variables, and we know there are strong correlations between two pairs of variables, then all other correlations between these four variables must be strong as well
- Multiple factors do not change this as long as there are no crossloadings
Implied correlations
Implied correlations
Factor loadings: EFA
Factor loadings: EFA crossloadings
CFA model
Correlations
qgraph

- qgraph is a network plotting package aimed at automatically visualizing fully connected weighted networks
- Can be used for many other kinds of networks as well
- qgraph objects can be exported to igraph objects
- Support for following classes:
  - "principal" (psych)
  - "loadings" (stats)
  - "factanal" (stats)
  - "graphNEL" (Rgraphviz)
  - "pcAlgo" (pcalg)
  - ("bn" (bnlearn))
  - ("bn.strength" (bnlearn))
- Easy to use!
Graphical options
semPlot
Currently supported SEM software packages:

▶ lavaan
▶ sem
▶ OpenMx (RAM only)
▶ MPlus (using MplusAutomation)
▶ LISREL (using LisrelToR)

Other R functions:

▶ lm
▶ factanal
▶ principal

Supported modelling frameworks:

▶ lavaan
▶ LISREL matrix model
▶ RAM matrix model
▶ Mplus matrix model
Holzinger and Swineford (1939) CFA example:
Holzinger and Swineford (1939) CFA example:
semPlot
Based on output file from Little (in press).

Within

BWITHDRA  BSELFCAR  BCOMPLIA  BANTISOC

Between

WITHDRAW  SELFCARE  COMPLIAN  ANTISOCI
semPlot
semPlot
Concluding comments

qgraph:
- Paper in JSS: http://www.jstatsoft.org/v48/i04
- Stable versions on CRAN:
  cran.r-project.org/package=qgraph
- developmental version on Github:
  https://github.com/SachaEpskamp/qgraph

semPlot:
- Paper in preparation
- Stable versions on CRAN:
  cran.r-project.org/package=semPlot
- developmental version on Github:
  https://github.com/SachaEpskamp/semPlot
- More info on my website: sachaepskamp.com
And more...

Netherlands:

Germany:
Thank you for your attention!
References


