The `qgraph` package for network visualizations of psychometric data in R

Sacha Epskamp

University of Amsterdam
Department of Psychological Methods

AmstRdam
Outline

Introduction
What is qgraph?
Graphs

Creating graphs
Input modes
Layout modes

Visualizing statistics
Correlation matrices
Factor loadings
Confirmatory Factor Analysis

Concluding comments
References
qgraph

A R package (CRAN link)

Can be used to plot various types of graphs

Different from other R packages (e.g. igraph

Csardi & Nepusz, 2006) in:

Focus on Weighted Graphs

Intended for visualization of data as graphs

Optimized for vector-type image files (e.g. PDF, SVG)

Aims in qgraph

Simple input

Summarize a large amount of statistics without needing data reduction methods.

Visualize relations between variables

Main idea: Show variables as nodes, relationships as edges
qgraph

- A R package (CRAN link)

  Focus on Weighted Graphs
  Intended for visualization of data as graphs
  Optimized for vector-type image files (e.g. PDF, SVG)

  Aims in qgraph
  Simple input
  Summarize a large amount of statistics without needing data reduction methods.
  Visualize relations between variables
  Main idea: Show variables as nodes, relationships as edges
qgraph

- A R package (CRAN link)
- Can be used to plot various types of graphs
qgraph

- A R package (CRAN link)
- Can be used to plot various types of graphs
- Different from other R packages (e.g. igraph Csardi & Nepusz, 2006) in:
  - Focus on Weighted Graphs
  - Intended for visualization of data as graphs
  - Optimized for vector-type image files (e.g. PDF, SVG)
  - Simple input
  - Summarize a large amount of statistics without needing data reduction methods.
  - Visualize relations between variables
  - Main idea: Show variables as nodes, relationships as edges
qgraph

- A R package (CRAN link)
- Can be used to plot various types of graphs
- Different from other R packages (e.g. \texttt{igraph} Csardi & Nepusz, 2006) in:
  - Focus on \textit{Weighted Graphs}
qgraph

- A R package (CRAN link)
- Can be used to plot various types of graphs
- Different from other R packages (e.g. igraph Csardi & Nepusz, 2006) in:
  - Focus on *Weighted Graphs*
  - Intended for visualization of data as graphs
qgraph

- A R package (CRAN link)
- Can be used to plot various types of graphs
- Different from other R packages (e.g. igraph Csardi & Nepusz, 2006) in:
  - Focus on *Weighted Graphs*
  - Intended for visualization of data as graphs
  - Optimized for vector-type image files (e.g. PDF, SVG)
qgraph

- A R package (CRAN link)
- Can be used to plot various types of graphs
- Different from other R packages (e.g. igraph Csardi & Nepusz, 2006) in:
  - Focus on *Weighted Graphs*
  - Intended for visualization of data as graphs
  - Optimized for vector-type image files (e.g. PDF, SVG)
- Aims in qgraph
qgraph

- A R package (CRAN link)
- Can be used to plot various types of graphs
- Different from other R packages (e.g. igraph Csardi & Nepusz, 2006) in:
  - Focus on *Weighted Graphs*
  - Intended for visualization of data as graphs
  - Optimized for vector-type image files (e.g. PDF, SVG)
- Aims in qgraph
  - Simple input
qgraph

- A R package (CRAN link)
- Can be used to plot various types of graphs
- Different from other R packages (e.g. igraph Csardi & Nepusz, 2006) in:
  - Focus on *Weighted Graphs*
  - Intended for visualization of data as graphs
  - Optimized for vector-type image files (e.g. PDF, SVG)
- Aims in **qgraph**
  - Simple input
  - Summarize a large amount of statistics without needing data reduction methods.
qgraph

- A R package (CRAN link)
- Can be used to plot various types of graphs
- Different from other R packages (e.g. \texttt{igraph} Csardi & Nepusz, 2006) in:
  - Focus on \textit{Weighted Graphs}
  - Intended for visualization of data as graphs
  - Optimized for vector-type image files (e.g. PDF, SVG)
- Aims in \texttt{qgraph}
  - Simple input
  - Summarize a large amount of statistics without needing data reduction methods.
  - Visualize \textit{relations between variables}
qgraph

- A R package (CRAN link)
- Can be used to plot various types of graphs
- Different from other R packages (e.g. igraph Csardi & Nepusz, 2006) in:
  - Focus on *Weighted Graphs*
  - Intended for visualization of data as graphs
  - Optimized for vector-type image files (e.g. PDF, SVG)
- Aims in *qgraph*
  - Simple input
  - Summarize a large amount of statistics without needing data reduction methods.
  - Visualize *relations between variables*
- Main idea: Show variables as nodes, relationships as edges
Graphs

A graph is a network that consists of $n$ nodes (or vertices) that are connected with $m$ edges. Each edge has a weight indicating the strength of that connection. An edge can be directed (have an arrow) or undirected.
Graphs

- A graph is a *network* that consists of *n* nodes (or vertices) that are connected with *m* edges.
Graphs

- A graph is a *network* that consists of $n$ nodes (or vertices) that are connected with $m$ edges.
- Each edge has a *weight* indicating the strength of that connection.
Graphs

- A graph is a network that consists of $n$ nodes (or vertices) that are connected with $m$ edges.
- Each edge has a weight indicating the strength of that connection.
- An edge can be directed (have an arrow) or undirected.
Unweighted graph
Weighted graph
Weighted graph
Directed graph
Outline

Introduction
  What is qgraph?
  Graphs

Creating graphs
  Input modes
  Layout modes

Visualizing statistics
  Correlation matrices
  Factor loadings
  Confirmatory Factor Analysis

Concluding comments
References
The qgraph() function

- The main function in `qgraph` is `qgraph()`

Usage:

```
qgraph( adj, ... )
```
The `qgraph()` function

- The main function in `qgraph` is `qgraph()`
  - Most other functions are either wrapping functions using `qgraph()` or functions used in `qgraph()`

Usage:

```r
qgraph( adj, ... )
```
The `qgraph()` function

- The main function in `qgraph` is `qgraph()`
  - Most other functions are either wrapping functions using `qgraph()` or functions used in `qgraph()`
- The `qgraph()` function requires only one argument (adj)

**Usage:**

```r
qgraph( adj, ... )
```
The `qgraph()` function

- The main function in `qgraph` is `qgraph()`
  - Most other functions are either wrapping functions using `qgraph()` or functions used in `qgraph()`
- The `qgraph()` function requires only one argument (`adj`)
- A lot of other arguments can be specified, but these are all optional

Usage:

```r
qgraph(adj, ...)  
```
The adjacency matrix

- The adj argument is the input. This can be an adjacency matrix.
The adjacency matrix

- The adj argument is the input. This can be an adjacency matrix.

- An adjacency matrix is a square $n$ by $n$ matrix in which each element indicates the relationship between two variables.
The adjacency matrix

- The `adj` argument is the input. This can be an *adjacency matrix*.
- An adjacency matrix is a square $n$ by $n$ matrix in which each element indicates the relationship between two variables.
- Any relationship can be used as long as:
  - A 0 indicates no relationship.
  - Absolute negative values are similar in strength to positive values.
  - Examples:
    - A 1 indicating a connection (unweighted graphs).
    - Correlations.
    - Regression parameters.
    - Factor loadings.

Adjacency matrices occur naturally in statistics!
The adjacency matrix

- The `adj` argument is the input. This can be an *adjacency matrix*.
- An adjacency matrix is a square $n$ by $n$ matrix in which each element indicates the relationship between two variables.
- Any relationship can be used as long as:
  - A 0 indicates no relationship.
The adjacency matrix

- The `adj` argument is the input. This can be an *adjacency matrix*.
- An adjacency matrix is a square $n$ by $n$ matrix in which each element indicates the relationship between two variables.
- Any relationship can be used as long as:
  - A 0 indicates no relationship.
  - Absolute negative values are similar in strength to positive values.

Adjacency matrices occur naturally in statistics!
The adjacency matrix

- The adj argument is the input. This can be an adjacency matrix.
- An adjacency matrix is a square $n$ by $n$ matrix in which each element indicates the relationship between two variables.
- Any relationship can be used as long as:
  - A 0 indicates no relationship.
  - Absolute negative values are similar in strength to positive values.
- Examples:
The adjacency matrix

- The adj argument is the input. This can be an *adjacency matrix*.
- An adjacency matrix is a square $n$ by $n$ matrix in which each element indicates the relationship between two variables.
- Any relationship can be used as long as:
  - A 0 indicates no relationship.
  - Absolute negative values are similar in strength to positive values.
- Examples:
  - A 1 indicating a connection (unweighted graphs).
The adjacency matrix

- The adj argument is the input. This can be an *adjacency matrix*.
- An adjacency matrix is a square $n$ by $n$ matrix in which each element indicates the relationship between two variables.
- Any relationship can be used as long as:
  - A 0 indicates no relationship.
  - Absolute negative values are similar in strength to positive values.
- Examples:
  - A 1 indicating a connection (unweighted graphs).
  - Correlations.
The adjacency matrix

- The `adj` argument is the input. This can be an adjacency matrix.
- An adjacency matrix is a square $n$ by $n$ matrix in which each element indicates the relationship between two variables.
- Any relationship can be used as long as:
  - A 0 indicates no relationship
  - Absolute negative values are similar in strength to positive values
- Examples:
  - A 1 indicating a connection (unweighted graphs)
  - Correlations
  - Regression parameters
The adjacency matrix

- The `adj` argument is the input. This can be an adjacency matrix.
- An adjacency matrix is a square $n$ by $n$ matrix in which each element indicates the relationship between two variables.
- Any relationship can be used as long as:
  - A 0 indicates no relationship.
  - Absolute negative values are similar in strength to positive values.
- Examples:
  - A 1 indicating a connection (unweighted graphs).
  - Correlations.
  - Regression parameters.
  - Factor loadings.

Adjacency matrices occur naturally in statistics!
The adjacency matrix

- The adj argument is the input. This can be an adjacency matrix.
- An adjacency matrix is a square $n$ by $n$ matrix in which each element indicates the relationship between two variables.
- Any relationship can be used as long as:
  - A 0 indicates no relationship.
  - Absolute negative values are similar in strength to positive values.
- Examples:
  - A 1 indicating a connection (unweighted graphs).
  - Correlations.
  - Regression parameters.
  - Factor loadings.
- Adjacency matrices occur naturally in statistics!
\[
\begin{bmatrix}
[1,] & 0 & 1 & 1 \\
[2,] & 0 & 0 & 1 \\
[3,] & 0 & 0 & 0 \\
\end{bmatrix}
\]
The Big 5

Included is a dataset in which the Dutch translation of a commonly used personality test, the NEO-PI-R (Costa & McCrae, 1992; Hoekstra, 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
The Big 5

> data(big5)
> str(big5)

num [1:500, 1:240] 2 3 4 4 5 2 2 1 4 2 ...
  - attr(*, "dimnames")=List of 2
    ..$: NULL
    ..$: chr [1:240] "N1" "E2" "O3" "A4" ...
The Big 5

> cor(big5)[1:15, 1:3]

<table>
<thead>
<tr>
<th></th>
<th>N1</th>
<th>E2</th>
<th>O3</th>
</tr>
</thead>
<tbody>
<tr>
<td>N1</td>
<td>1.0000000e+00</td>
<td>-0.1855271104</td>
<td>0.095361021</td>
</tr>
<tr>
<td>E2</td>
<td>-1.855271e-01</td>
<td>1.0000000000</td>
<td>0.082879486</td>
</tr>
<tr>
<td>O3</td>
<td>9.536102e-02</td>
<td>0.0828794862</td>
<td>1.0000000000</td>
</tr>
<tr>
<td>A4</td>
<td>-1.388302e-01</td>
<td>0.2639947325</td>
<td>-0.077486908</td>
</tr>
<tr>
<td>C5</td>
<td>-6.958419e-02</td>
<td>0.0433073349</td>
<td>-0.019451268</td>
</tr>
<tr>
<td>N6</td>
<td>1.901235e-01</td>
<td>-0.1144449337</td>
<td>0.052585940</td>
</tr>
<tr>
<td>E7</td>
<td>-8.556772e-02</td>
<td>0.1968042360</td>
<td>-0.001360255</td>
</tr>
<tr>
<td>O8</td>
<td>3.599480e-02</td>
<td>0.0380351107</td>
<td>0.143715099</td>
</tr>
<tr>
<td>A9</td>
<td>3.775956e-02</td>
<td>0.0918642940</td>
<td>-0.147722360</td>
</tr>
<tr>
<td>C10</td>
<td>7.113742e-02</td>
<td>-0.1250087306</td>
<td>0.005001674</td>
</tr>
<tr>
<td>N11</td>
<td>3.491657e-01</td>
<td>-0.1085540815</td>
<td>0.072827947</td>
</tr>
<tr>
<td>E12</td>
<td>2.680803e-01</td>
<td>0.0456805867</td>
<td>-0.003996781</td>
</tr>
<tr>
<td>O13</td>
<td>9.768583e-02</td>
<td>-0.0002757219</td>
<td>0.218802592</td>
</tr>
<tr>
<td>A14</td>
<td>4.512596e-05</td>
<td>0.1133985066</td>
<td>0.009208309</td>
</tr>
<tr>
<td>C15</td>
<td>5.859827e-03</td>
<td>0.0988472658</td>
<td>-0.051836728</td>
</tr>
</tbody>
</table>
> qgraph(cor(big5), minimum = 0.2)
> data(big5groups)
> qgraph(cor(big5), groups = big5groups, minimum = 0.2)
Fruchterman-Reingold layout (20 iterations)
> data(big5groups)
> Q <- qgraph(cor(big5), groups = big5groups, 
+   minimum = 0.2, layout = "spring")
> Q <- qgraph(Q, legend = FALSE)
> Q <- qgraph(Q, vsize = 2, borders = FALSE, + vTrans = 150)
> Q <- qgraph(Q, overlay = TRUE)
> qgraph(Q, transparency = T, bg = T, bgcontrol = 5,
+     filetype = "png", filename = "bg", res = 144,
+     width = 7, height = 7)
Association

- Neuroticism
- Extraversion
- Openness
- Agreeableness
- Conscientiousness
Significance

\[ \texttt{qgraph(cor(big5), Q, graph = "sig")} \]
Factor loadings

- A factor loadings matrix can be visualized using `qgraph.loadings()`
Factor loadings

- A factor loadings matrix can be visualized using `qgraph.loadings()`
- There are two wrapper functions that perform an analysis and send the results to `qgraph.loadings()`:
  - `qgraph.efa()` performs an exploratory factor analysis (EFA) using `stats:::factanal`
  - `qgraph.pca()` performs a principal component analysis (PCA) using `psych:::principal` (Revelle, 2010)
  - These functions use a correlation or covariance matrix as input
A factor loadings matrix can be visualized using `qgraph.loadings()`.

There are two wrapper functions that perform an analysis and send the results to `qgraph.loadings()`:

- `qgraph.efa()` performs an exploratory factor analysis (EFA) using `stats:::factanal`
- `qgraph.pca()` performs a principal component analysis (PCA) using `psych:::principal` (Revelle, 2010)
Factor loadings

- A factor loadings matrix can be visualized using `qgraph.loadings()`
- There are two wrapper functions that perform an analysis and send the results to `qgraph.loadings()`:
  - `qgraph.efa()` performs an exploratory factor analysis (EFA) using `stats:::factanal`
  - `qgraph.pca()` performs a principal component analysis (PCA) using `psych:::principal` (Revelle, 2010)
Factor loadings

- A factor loadings matrix can be visualized using `qgraph.loadings()`
- There are two wrapper functions that perform an analysis and send the results to `qgraph.loadings()`:
  - `qgraph.efa()` performs an exploratory factor analysis (EFA) using `stats:::factanal`
  - `qgraph.pca()` performs a principal component analysis (PCA) using `psych:::principal` (Revelle, 2010)
- These functions use a correlation or covariance matrix as input
Factor loadings: EFA

```r
> qgraph.efa(big5, 5, groups = big5groups,
+    rotation = "promax", minimum = 0.2, cut = 0.4,
+    vsize = c(1, 15), borders = FALSE, asize = 0.07,
+    esize = 4, vTrans = 200)
```
Factor loadings: EFA crossloadings

```r
> qgraph.efa(big5, 5, groups = big5groups,
+       rotation = "promax", minimum = 0.2, cut = 0.4,
+       vsize = c(1, 15), borders = FALSE, asize = 0.07,
+       esize = 4, vTrans = 200, crossloadings = TRUE)
```
Factor loadings: PCA

```r
> qgraph.pca(cor(big5), 5, groups = big5groups,
+   rotation = "promax", minimum = 0.2, cut = 0.4,
+   vsize = c(1, 15), borders = FALSE, asize = 0.07,
+   esize = 4, vTrans = 200)
```
Confirmatory Factor Analysis

- `qgraph.cfa()` can be used to fit a simple confirmatory factor model
Confirmatory Factor Analysis

- `qgraph.cfa()` can be used to fit a simple confirmatory factor model
  - Each variable loads on only one factor
Confirmatory Factor Analysis

- `qgraph.cfa()` can be used to fit a simple confirmatory factor model
  - Each variable loads on only one factor
  - Factors are correlated

Returns a "sem" or `lavaan` object

Results can be sent to `qgraph.sem()` or `qgraph.lavaan()` for a full report
Confirmatory Factor Analysis

- `qgraph.cfa()` can be used to fit a simple confirmatory factor model
  - Each variable loads on only one factor
  - Factors are correlated
  - Scaling by fixing first loading of each factor to 1

Returns a "sem" or `lavaan` object

Results can be sent to `qgraph.sem()` or `qgraph.lavaan()` for a full report
Confirmatory Factor Analysis

- `qgraph.cfa()` can be used to fit a simple confirmatory factor model
  - Each variable loads on only one factor
  - Factors are correlated
  - Scaling by fixing first loading of each factor to 1
- This can be done with the `sem` (Fox, 2006) or `lavaan` (Rosseel, 2011) packages
Confirmatory Factor Analysis

- `qgraph.cfa()` can be used to fit a simple confirmatory factor model
  - Each variable loads on only one factor
  - Factors are correlated
  - Scaling by fixing first loading of each factor to 1
- This can be done with the `sem` (Fox, 2006) or `lavaan` (Rosseel, 2011) packages
- Returns a "sem" or lavaan object
Confirmatory Factor Analysis

- `qgraph.cfa()` can be used to fit a simple confirmatory factor model
  - Each variable loads on only one factor
  - Factors are correlated
  - Scaling by fixing first loading of each factor to 1
- This can be done with the `sem` (Fox, 2006) or `lavaan` (Rosseel, 2011) packages
- Returns a "sem" or lavaan object
- Results can be send to `qgraph.sem()` or `qgraph.lavaan()` for a full report
Confirmatory Factor Analysis

> names(big5groups) <- strtrim(names(big5groups), + 1)
> names(big5) <- 1:ncol(big5)
> fit <- qgraph.cfa(cov(big5), nrow(big5),
+ big5groups, pkg = "lavaan", opts = list(se = "none"),
+ vsize.man = 1, vsize.lat = 6, edge.label.cex = 0.5)
> print(fit)

Lavaan (0.4-8) converged normally after 161 iterations

Number of observations 500

Estimator ML
Minimum Function Chi-square 60838.192
Degrees of freedom 28430
P-value 0.000
Confirmatory Factor Analysis

> pdf("big5cfaModel%03d.pdf", width = 7, height = 7,
+    onefile = FALSE)
> qgraph.lavaan(fit, filetype = "", include = 1:7,
+    vsize.man = 1, vsize.lat = 6, edge.label.cex = 0.5,
+    residSize = 0.1, groups = big5groups,
+    titles = FALSE)
> dev.off()

pdf
  2

> pdf("big5cfaRes%03d.pdf", width = 14, height = 7,
+    onefile = FALSE)
> qgraph.lavaan(fit, filetype = "", include = 8:12,
+    vsize.man = 1, vsize.lat = 6, edge.label.cex = 0.5,
+    residSize = 0.1, groups = big5groups,
+    titles = FALSE, minimum = 0.15)
> dev.off()

pdf
  2
Results
Results
Correlations
Outline

Introduction
  What is qgraph?
  Graphs

Creating graphs
  Input modes
  Layout modes

Visualizing statistics
  Correlation matrices
  Factor loadings
  Confirmatory Factor Analysis

Concluding comments
References
Concluding comments

- *qgraph* is still work in progress
Concluding comments

- qgraph is still work in progress
- Plans for the future:
Concluding comments

- qgraph is still work in progress
- Plans for the future:
  - More wrapper functions for different statistics (e.g. IRT)
Concluding comments

- **qgraph** is still work in progress
- Plans for the future:
  - More wrapper functions for different statistics (e.g. IRT)
  - More layout modes
Concluding comments

- **qgraph** is still work in progress
- Plans for the future:
  - More wrapper functions for different statistics (e.g. IRT)
  - More layout modes
  - Estimating and fitting causal models
Concluding comments

- **qgraph** is still work in progress
- Plans for the future:
  - More wrapper functions for different statistics (e.g. IRT)
  - More layout modes
  - Estimating and fitting causal models
- Some things I couldn’t describe...
Layout constraints
Grayscale colors

> Q <- qgraph(cor(big5), minimum = 0.25, cut = 0.4,
+       vsize = 2, groups = big5groups, legend = T,
+       borders = F, vTrans = 200, gray = TRUE)
Tooltips

Link
Modelling

\[ \xi \]

\[ \eta_1 \quad \eta_2 \quad \eta_3 \]

\[ y_1 \quad y_2 \quad y_3 \quad y_4 \quad y_5 \quad y_6 \quad y_7 \]

\[ x_1 \quad x_2 \quad x_3 \]
Concluding comments

Thank you for your attention!


