SEM 1: Confirmatory Factor Analysis

Week 4 - Exploratory Factor Analysis

Sacha Epskamp

2020

Exploratory Factor Analysis (EFA)

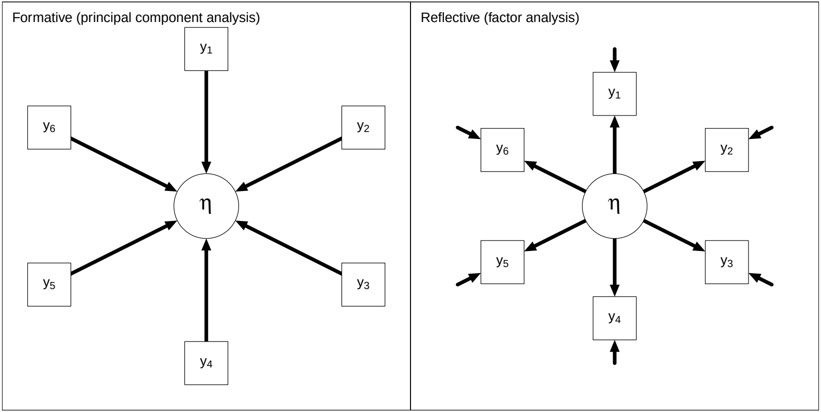
Exploratorily estimate $\boldsymbol{\Lambda}$ (no free elements in $\boldsymbol{\Lambda}$):

$$\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Psi}\boldsymbol{\Lambda}^\top + \boldsymbol{\Theta}$$

Very close, but not the same (!!) as principal component analysis (PCA):

$$\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Psi}\boldsymbol{\Lambda}^\top$$

Very different interpretation. EFA *measures* latents (there is measurement error) and captures *common variance*, PCA only *summarizes* the data and captures *total variance*.


Formative (principal component analysis)
y1
y2
y3
y4
y5
y6
η
Reflective (factor analysis)
y1
y2
y3
y4
y5
y6
η

Exploratory Factor Analysis (EFA)

If $\boldsymbol{\Lambda}$ is not somehow constrained, latent variable variance is not identified. We can arbitrarily add rotation matrices $\boldsymbol{R}$ and not change the decomposition:

$$\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{R}\boldsymbol{R}^{-1}\boldsymbol{\Psi}\boldsymbol{R}^{-1\top}\boldsymbol{R}^{\top}\boldsymbol{\Lambda}^{\top} + \boldsymbol{\Theta}$$

Can be seen as a different factor model with $\boldsymbol{\Lambda}^* = \boldsymbol{\Lambda}\boldsymbol{R}$ and $\boldsymbol{\Psi}^* = \boldsymbol{R}^{-1}\boldsymbol{\Psi}\boldsymbol{R}^{-1\top}$. To this end, in estimation one can assume uncorrelated factors, $\boldsymbol{\Psi} = \boldsymbol{I}$. Afterwards, rotation methods can be used to obtain simple structure for $\boldsymbol{\Lambda}$ while possibly allowing factors to correlate:

- orthogonal (varimax): axes remain orthogonal, independent
- oblique (promax/oblimin): axes become correlated

I often use promax rotation.

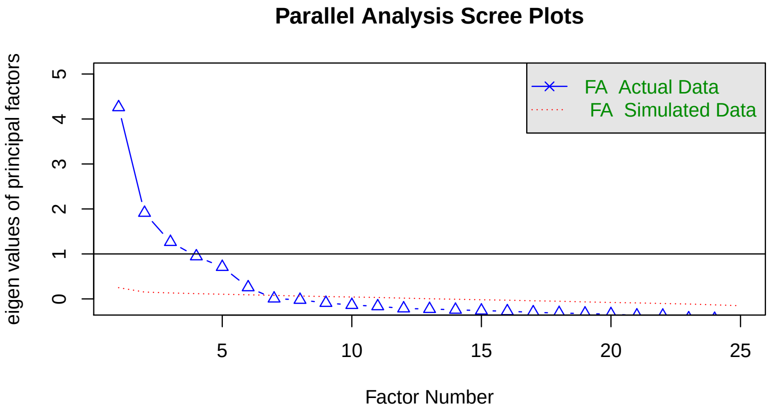
Choosing the number of Factors is a bit more involved than PCA

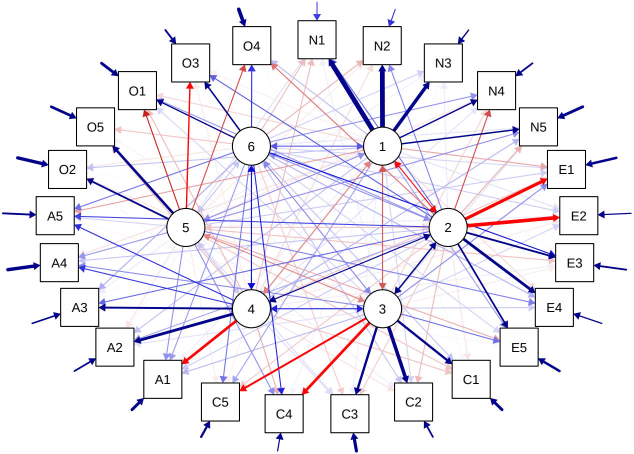
- One method involves checking how many eigenvalues in $\boldsymbol{S} - \hat{\boldsymbol{\Theta}}$ are above 0
  - $\hat{\boldsymbol{\Theta}}$ is then estimated using a 1-factor model
- Parallel analysis takes sampling variation into account, and checks how many eigenvalues are statistically above what can be expected given an independence model

Big 5 example

```r
library("psych")

# Load data:
data(bfi)
bfiSub <- bfi[,1:25]

# Correlations:
corMat <- cor(bfiSub, use = "pairwise.complete.obs")
N <- nrow(bfiSub)
```

```r
fa.parallel(corMat, N, fa = "fa")

## Parallel analysis suggests that the number of factors =  6  and the 
```





O4
N1
N2
O3
N3
O1
N4
O5
N5
6
1
O2
E1
A5
5
2
E2
A4
E3
A3
4
3
E4
A2
E5
A1
C1
C5
C2
C4
C3

Estimating the dimensionality of intelligence like data using Exploratory Graph Analysis

Hudson F. Golino [a], Andreas Demetriou [b,*]

[a] Universidade Salgado de Oliveira, Brazil
[b] University of Nicosia, Cyprus

A B S T R A C T

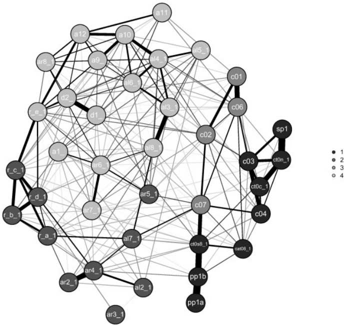
This study compared various exploratory and confirmatory factor methods for recovering factors of cognitive test-like data. We first note the problems encountered by several widely used methods, such as parallel analysis, minimum average partial procedure, and confirmatory factor analysis, in estimating the number of dimensions underlying performance on test batteries. We then argue that a new method, Exploratory Graph Analysis (EGA), can more accurately uncover underlying dimensions or factors and demonstrate how this method outperforms the other methods. We use several published data sets to demonstrate the advantages of EGA. We conclude that a combination of EGA and confirmatory factor analysis or structural equation modeling may be the ideal in precisely specifying latent factors and their relations.

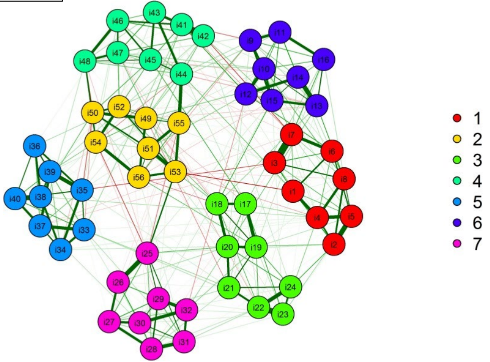
Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research

Hudson F. Golino[1,2]*, Sacha Epskamp[3]

1 Department of Psychology, University of Virginia, Charlottesville, VA, United States of America, **2** Graduate School of Psychology, Universidade Salgado de Oliveira, Rio de Janeiro, Brasil, **3** University of Amsterdam, Amsterdam, Netherlands

* hfgolino@gmail.com





Exploratory versus confirmatory research

The picture of the Texas sharpshooter is taken from an illustration by Dirk-Jan Hoek (CC-BY).

Exploratory versus confirmatory ~~research~~ factor analysis

- Perhaps poorly chosen terms!
- Exploratory factor analysis allows for confirmatory tests!
  - E.g., test if a 5-factor structure generally fits better than a 4-factor structure
- Confirmatory factor analysis on the other hand is often exploratory!
  - E.g., cherry picking a model, many different tests, model modifications with modification indices, etcetera
  - In fact, you could argue CFA suffers from many arbitrary researcher degrees of freedom
- EFA and CFA are generally just very different, and EFA should not always be followed by CFA
  - E.g., personality questionaires usually show *very* poor performance with CFA
  - An alternative is also exploratory SEM (ESEM), which puts EFA in a SEM framework