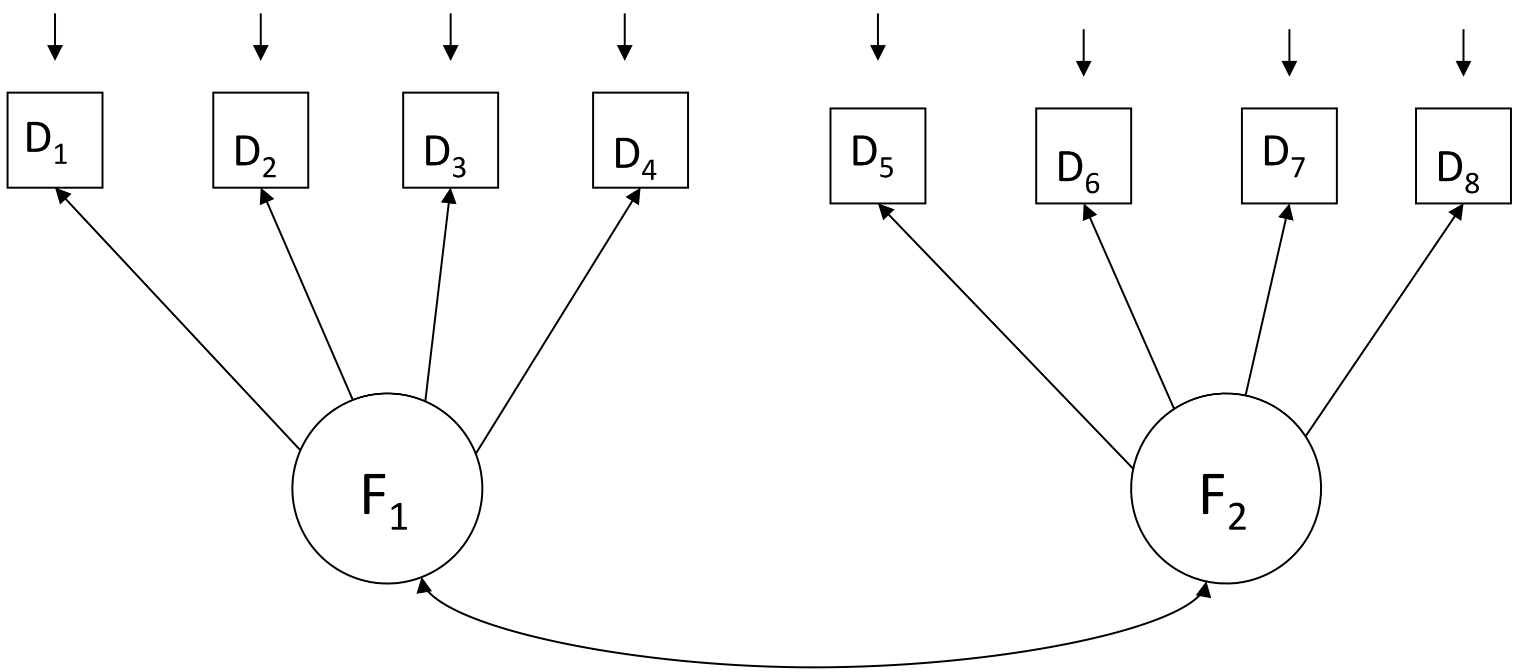
Exploratory Factor Analysis

STA2053: Fall 2022

Factor Analysis: The Measurement Model

$$\mathbf{D}_i = \mathbf{\Lambda}\mathbf{F}_i + \mathbf{e}_i$$



Example with 2 factors and 8 observed variables

$$\mathbf{D}_i = \mathbf{\Lambda}\mathbf{F}_i + \mathbf{e}_i$$

$$\begin{pmatrix} D_{i,1} \\ D_{i,2} \\ D_{i,3} \\ D_{i,4} \\ D_{i,5} \\ D_{i,6} \\ D_{i,7} \\ D_{i,8} \end{pmatrix} = \begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \\ \lambda_{31} & \lambda_{32} \\ \lambda_{41} & \lambda_{42} \\ \lambda_{51} & \lambda_{52} \\ \lambda_{61} & \lambda_{62} \\ \lambda_{71} & \lambda_{27} \\ \lambda_{81} & \lambda_{82} \end{pmatrix} \begin{pmatrix} F_{i,1} \\ F_{i,2} \end{pmatrix} + \begin{pmatrix} e_{i,1} \\ e_{i,2} \\ e_{i,3} \\ e_{i,4} \\ e_{i,5} \\ e_{i,6} \\ e_{i,7} \\ e_{i,8} \end{pmatrix}$$

$$\begin{aligned} D_{i,1} &= \lambda_{11}F_{i,1} + \lambda_{12}F_{i,2} + e_{i,1} \\ D_{i,2} &= \lambda_{21}F_{i,1} + \lambda_{22}F_{i,2} + e_{i,2} \ \text{ etc.} \end{aligned}$$

The lambda values are called **factor loadings.**

Terminology

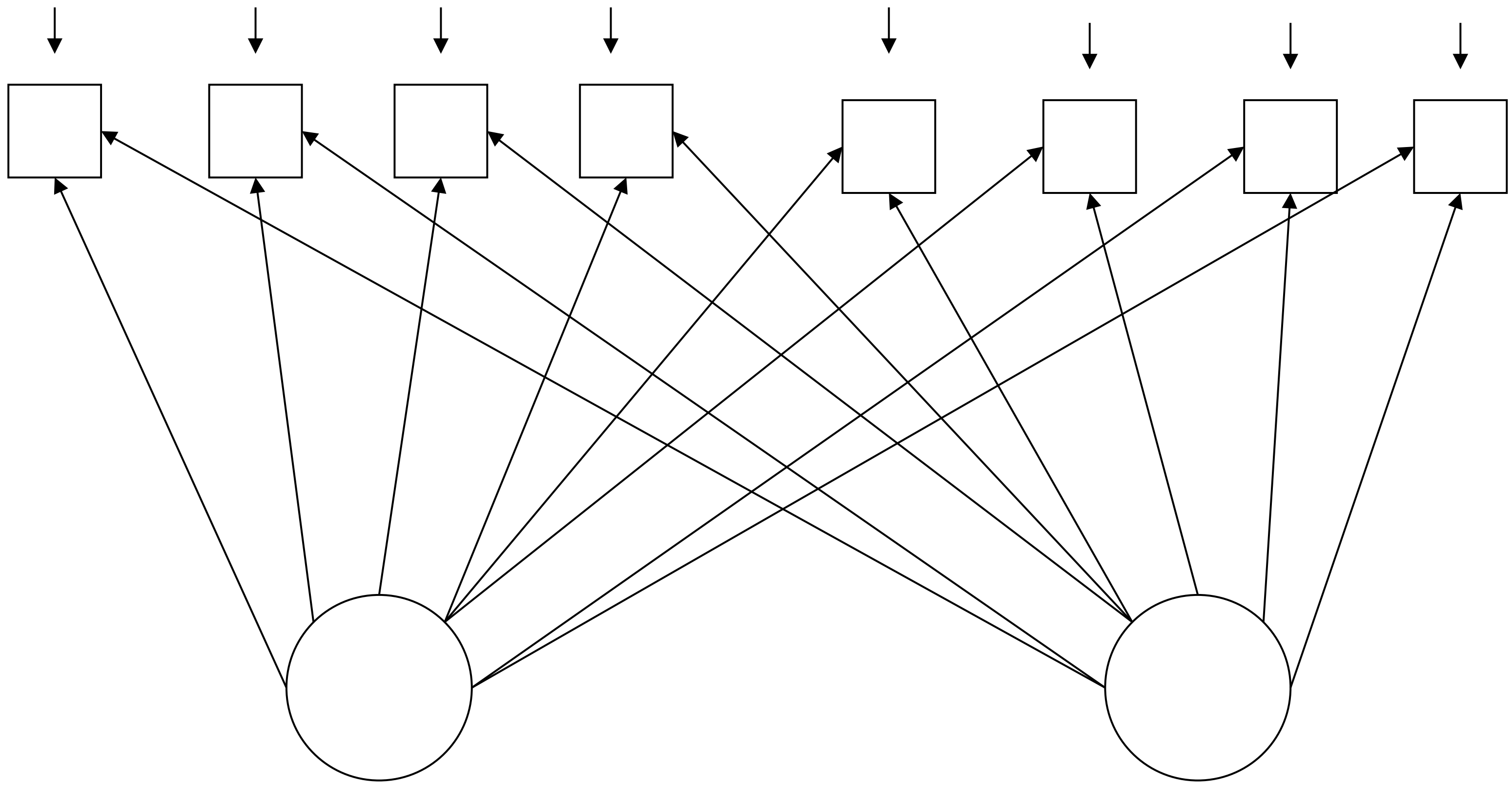
$$\begin{aligned} D_{i,1} &= \lambda_{11} F_{i,1} + \lambda_{12} F_{i,2} + e_{i,1} \\ D_{i,2} &= \lambda_{21} F_{i,1} + \lambda_{22} F_{i,2} + e_{i,2} \quad \text{etc.} \end{aligned}$$

- The lambda values are called **factor loadings.**
- F_1 and F_2 are sometimes called **common factors**, because they influence all the observed variables.
- Error terms e_1, …, e_8 are sometimes called **unique factors**, because each one influences only a single observed variable.

Factor Analysis can be

- **Exploratory**: The goal is to describe and summarize the data by explaining a large number of observed variables in terms of a smaller number of latent variables (factors). The factors are the reason the observable variables have the correlations they do.
- **Confirmatory**: Statistical estimation and testing as usual.

Unconstrained (Exploratory) Factor Analysis



- Arrows from all factors to all observed variables.
- Massively non-identifiable.
- Reasonable, been going on for around 70-100 years, and completely DOOMED TO FAILURE as a method of statistical estimation.

Calculate the covariance matrix

$$\begin{aligned}\mathbf{D}_i &= \mathbf{\Lambda}\mathbf{F}_i + \mathbf{e}_i \\ cov(\mathbf{F}_i) &= \mathbf{\Phi} \\ cov(\mathbf{e}_i) &= \mathbf{\Omega}\end{aligned}$$

$\mathbf{F}_i$ and $\mathbf{e}_i$ independent (multivariate normal)

$$cov(\mathbf{D}_i) = \mathbf{\Sigma} = \mathbf{\Lambda}\mathbf{\Phi}\mathbf{\Lambda}^\top + \mathbf{\Omega}$$

A Re-parameterization

$$\mathbf{\Sigma} = \mathbf{\Lambda}\mathbf{\Phi}\mathbf{\Lambda}^\top + \mathbf{\Omega}$$

Square root matrix: $\mathbf{\Phi} = \mathbf{S}\mathbf{S} = \mathbf{S}\mathbf{S}^\top$, so

$$\begin{aligned}
\mathbf{\Lambda}\mathbf{\Phi}\mathbf{\Lambda}^\top &= \mathbf{\Lambda}\mathbf{S}\mathbf{S}^\top\mathbf{\Lambda}^\top \\
&= (\mathbf{\Lambda}\mathbf{S})\mathbf{I}(\mathbf{S}^\top\mathbf{\Lambda}^\top) \\
&= (\mathbf{\Lambda}\mathbf{S})\mathbf{I}(\mathbf{\Lambda}\mathbf{S})^\top \\
&= \mathbf{\Lambda}_2\mathbf{I}\mathbf{\Lambda}_2^\top
\end{aligned}$$

Parameters are not identifiable

$$\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}^\top + \boldsymbol{\Omega} \; = \; \boldsymbol{\Lambda}_2\mathbf{I}\boldsymbol{\Lambda}_2^\top + \boldsymbol{\Omega}$$

- Two distinct (Lambda, Phi, Omega) sets give the same Sigma, and hence the same distribution of the data (under normality).
- Actually, there are infinitely many. Let **Q** be an arbitrary covariance matrix for **F**.

$$\begin{aligned}
\boldsymbol{\Lambda}_2\mathbf{I}\boldsymbol{\Lambda}_2^\top &= \boldsymbol{\Lambda}_2\mathbf{Q}^{-\frac{1}{2}}\mathbf{Q}\mathbf{Q}^{-\frac{1}{2}}\boldsymbol{\Lambda}_2^\top \\
&= (\boldsymbol{\Lambda}_2\mathbf{Q}^{-\frac{1}{2}})\mathbf{Q}(\mathbf{Q}^{-\frac{1}{2}\top}\boldsymbol{\Lambda}_2^\top) \\
&= (\boldsymbol{\Lambda}_2\mathbf{Q}^{-\frac{1}{2}})\mathbf{Q}(\boldsymbol{\Lambda}_2\mathbf{Q}^{-\frac{1}{2}})^\top \\
&= \boldsymbol{\Lambda}_3\mathbf{Q}\boldsymbol{\Lambda}_3^\top
\end{aligned}$$

Restrict the model

$$\mathbf{\Lambda\Phi\Lambda}^\top = \mathbf{\Lambda}_2\mathbf{I}\mathbf{\Lambda}_2^\top$$

- Set Phi = the identity, so cov(**F**) = I
- All the factors are standardized, as well as independent.
- Justify this on the grounds of simplicity.
- Say the factors are “orthogonal” (at right angles, uncorrelated).

Another Source of non-identifiability

R is an orthoganal (rotation) matrix

$$
\begin{aligned}
\boldsymbol{\Sigma} &= \boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top + \boldsymbol{\Omega} \\
&= \boldsymbol{\Lambda}\mathbf{R}\mathbf{R}^\top\boldsymbol{\Lambda}^\top + \boldsymbol{\Omega} \\
&= (\boldsymbol{\Lambda}\mathbf{R})(\mathbf{R}^\top\boldsymbol{\Lambda}^\top) + \boldsymbol{\Omega} \\
&= (\boldsymbol{\Lambda}\mathbf{R})(\boldsymbol{\Lambda}\mathbf{R})^\top + \boldsymbol{\Omega} \\
&= \boldsymbol{\Lambda}_2\boldsymbol{\Lambda}_2^\top + \boldsymbol{\Omega}
\end{aligned}
$$

Infinitely many rotation matrices produce the same Sigma.

A Solution

- Place some restrictions on the factor loadings, so that the only rotation matrix that preserves the restrictions is the identity matrix. For example, $\lambda_{ij} = 0$ for $j>i$
- There are other sets of restrictions that work.
- Generally, they result in a set of factor loadings that are impossible to interpret. Don't worry about it.
- Estimate the loadings by maximum likelihood. Other methods are possible but used much less than in the past.
- All (orthoganal) rotations result in the same value of the likelihood function (the maximum is not unique).
- Rotate the factors (that is, post-multiply the estimated loadings by a rotation matrix) so as to achieve a simple pattern that is easy to interpret.
- The result is often satisfying, but has no necessary connection to reality.

Consulting advice

- When a non-statistician claims to have done a "factor analysis," ask what kind.
- Usually it was a principal components analysis.
- Principal components are linear combinations of the observed variables. They come from the observed variables by direct calculation.
- In true factor analysis, it's the observed variables that arise from the factors.
- So principal components analysis is kind of like backwards factor analysis, though the spirit is similar.
- Most factor analysis software (SAS, SPSS, etc.) does principal components analysis by default.

Copyright Information