Latent Variables

Principal Components, Factor Analysis, and the Measurement Model

Steven M. Boker

Department of Psychology
University of Virginia

Structural Equation Modeling
Psyc–8501–001

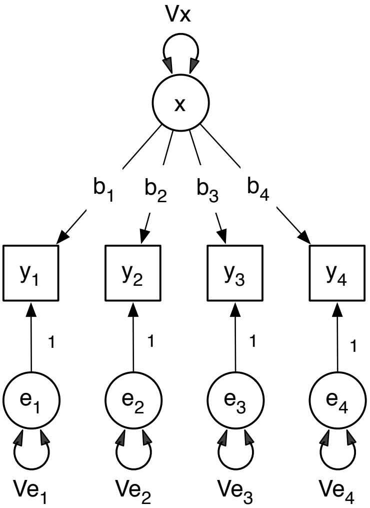
Overview

- Common Factor as a Latent Variable
- Transformations
- Eigenvalues and Eigenvectors
- Principal Components
- Covariance for PCA and FA
- Identification
- Latent Variables, Indicators, and Meaning
- Latent Structure.
- Examples of Latent Structure.
  1. Correlated Factors
  2. Structural Regression Models.
- Constraints.

Common Factor as a Latent Variable

- We have used latent variables to represent residuals.
- Latent variables can also be other unmeasured variables.
- Common factors can be thought of as latent variables.

Common Factor as a Latent Variable



Common Factor as a Latent Variable

- A common factor latent variable is the shared variance between a set of indicators.
- But how do we estimate the regression coefficients?
- This is where SEM really gets interesting.
- In order to understand it, we'll first need to review matrix transformations, eigenvalue–eigenvector decomposition, and principal components analysis.

Linear Transformations

- A transformation matrix, $\mathbf{T}$, is square and non-singular.
- Any matrix $\mathbf{A}$ can be transformed into a matrix $\mathbf{B}$ by multiplication by a transformation matrix $\mathbf{T}$. If $\mathbf{A}$ is premultiplied by $\mathbf{T}$

$$\mathbf{TA} = \mathbf{B}$$

  a row transformation has been performed and if $\mathbf{A}$ is postmultiplied by $\mathbf{T}$

$$\mathbf{AT} = \mathbf{B}$$

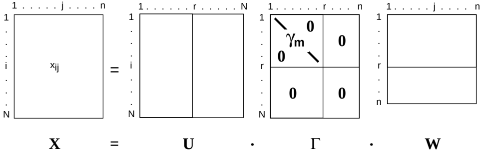
  a column transformation has been performed.
- A group of vectors produced by a transformation of another group of vectors will span the identical subspace spanned by the original group of vectors.

Linear Transformations

- An orthogonal transformation matrix $\mathbf{U}$ is a transformation matrix such that
$$\mathbf{U}' = \mathbf{U}^{-1}.$$
- The row vectors (or column vectors) contained in an orthogonal transformation matrix, $\mathbf{U}$, are of unit length and are mutually orthogonal so that:
$$\mathbf{U}\mathbf{U}' = \mathbf{U}'\mathbf{U} = \mathbf{I}.$$
- A columns (or rows) transformation by an orthogonal transformation matrix of a matrix of row (or column) vectors will produce an orthogonal rotation of axes.
- An orthogonal rotation of axes will not alter:
  - the lengths of the vectors in the space;
  - the inner product between vectors in the space;
  - the dimensionality of the subspace spanned by a group of vectors in the space.

Principal Roots and Vectors

- Any rectangular matrix, say $\mathbf{X}$, of order $N \times n$, can be expressed as the product of three matrices of the following form:



where $\mathbf{U}$ and $\mathbf{W}$ are orthogonal matrices so that

$$\mathbf{U}^{-1} = \mathbf{U}'$$
$$\mathbf{W}^{-1} = \mathbf{W}'$$

and $\mathbf{\Gamma}$ is zero except for a diagonal matrix of order r in the upper left corner with diagonal entries γ_m such that

$$\gamma_1 \geq \gamma_2 \geq \gamma_3 \geq \ldots \geq \gamma_r \geq 0$$

where $r \leq N$ and $r \leq n$.

Principal Roots and Vectors

- The columns of $\mathbf{U}$ will be called the **left principal vectors** of $\mathbf{X}$ and be designated $\mathbf{U}_m$; similarly, the rows of $\mathbf{W}$ will be called the **right principal vectors** of $\mathbf{X}$ and be designated $\mathbf{W}_m$ and, correspondingly, the γ_m will be called the **principal roots** of $\mathbf{X}$.
- For each $m = \{1 \ldots r\}$ the product of the corresponding left principal vector, principal root, and right principal vector will produce a matrix $\mathbf{Y}_{(m)}$, of order $N \times n$, which will be designated the mth **principal component matrix** of $\mathbf{X}$.

$$\mathbf{Y}_{(m)} = \mathbf{U}_m \gamma_m \mathbf{W}_m.$$

Elements of $\mathbf{Y}_{(m)}$ will be designated $y_{(m)ij}$.

Principal Roots and Vectors

- Any two different principal component matrices of $\mathbf{X}$, say $\mathbf{Y}_{(m)}$ and $\mathbf{Y}_{(m^\star)}$, where m and $m^\star$ are unequal, are orthogonal in that every row vector in one is orthogonal to every row vector in the other and every column vector in one is orthogonal to every column vector in the other; thus:

$$\mathbf{Y}_{(m)}\mathbf{Y}'_{(m^\star)} = 0$$

$$\mathbf{Y}'_{(m)}\mathbf{Y}_{(m^\star)} = 0$$

- The matrix $\mathbf{X}$ is the sum of its principal component matrices,

$$\mathbf{X} = \sum_{m=1}^{r} \mathbf{Y}_{(m)}$$

Principal Roots and Vectors

- The square of the norm of a principal component matrix equals the square of the corresponding principal root.

$$[Norm(\mathbf{Y}_{(m)})]^2 = \gamma_m^2.$$

- The square of the norm of the matrix **X** is equal to the sum of the squares of the norm of the principal component matrices and to the sum of the squares of the principal roots.

$$[Norm(\mathbf{X})]^2 = \sum_{m=1}^{r} [Norm(\mathbf{Y}_{(m)})]^2 = \sum_{m=1}^{r} \gamma_m^2.$$

- The rank of matrix **X** equals r, the number of non-zero principal roots of **X**.

Principal Roots and Vectors

- Let ρ be an integer in the range

$$0 \leq \rho \leq r,$$

the ρth approximation to $\mathbf{X}$ is designated by $\hat{\mathbf{X}}_{(\rho)}$ and is defined as the sum of the first ρ principal component matrices:

$$\hat{\mathbf{X}}_{(\rho)} = \sum_{m=1}^{\rho} \mathbf{Y}_{(m)}.$$

- The error of approximating $\mathbf{X}$ by $\hat{\mathbf{X}}_{(\rho)}$ is designated by $\mathbf{E}_{(\rho)}$ so that

$$\mathbf{X} = \hat{\mathbf{X}}_{(\rho)} + \mathbf{E}_{(\rho)}.$$

which implies

$$\mathbf{E}_{(\rho)} = \sum_{m=\rho+1}^{r} \mathbf{Y}_{(m)}$$

Principal Roots and Vectors

- The square of the norm of $\mathbf{X}$ is resolved into the squares of the norms of $\mathbf{X}_{(\rho)}$ and $\mathbf{E}_{(\rho)}$ in a manner parallel to the resolution of $\mathbf{X}$ into $\mathbf{X}_{(\rho)}$ and $\mathbf{E}_{(\rho)}$ above; thus:

$$
\begin{aligned}
[Norm(\hat{\mathbf{X}}_{(\rho)})]^2 &= \sum_{m=1}^{\rho} \gamma_m^2 \\
[Norm(\mathbf{E}_{(\rho)})]^2 &= \sum_{m=\rho+1}^{r} \gamma_m^2 \\
[Norm(\mathbf{X})]^2 &= [Norm(\hat{\mathbf{X}}_{(\rho)})]^2 + [Norm(\mathbf{E}_{(\rho)})]^2
\end{aligned}
$$

- The matrix $\mathbf{X}_{(\rho)}$ provides a least squares approximation matrix of rank ρ to the given matrix $\mathbf{X}$.

Principal Roots and Vectors

- The characteristic roots and vectors of the columns product matrix of $\mathbf{X}$,

$$\mathbf{P}_{(j)} = \mathbf{X}'\mathbf{X}$$

are related to the principal roots and right principal vectors of $\mathbf{X}$ in the following manner.

$$\mathbf{P}_{(j)} = \mathbf{W}'\beta_{(n)}\mathbf{W}$$

where $\beta_{\mathbf{n}}$ is a diagonal matrix of order n containing the characteristic roots of $\mathbf{P}_{(j)}$ and $\mathbf{W}'$ is the orthogonal matrix containing the characteristic vectors of $\mathbf{P}_{(j)}$, then

$$\begin{aligned} \beta_m = \gamma_m^2 \quad &\text{for} \quad m = 1, 2, 3, \ldots, r; \\ \beta_{m^{**}} = 0 \quad &\text{for} \quad m^{**} = (r+1), \ldots, n; \end{aligned}$$

and $\mathbf{W}$ contains the right principal vectors of $\mathbf{X}$.

Principal Roots and Vectors

- An orthogonal rotation of column axes of $\mathbf{X}$ by the matrix $\mathbf{W}'$ produces a matrix $\mathbf{Z}_{(c)}$,

$$\mathbf{X}\mathbf{W}' = \mathbf{Z}_c,$$

such that

$$\mathbf{Z}_c = \mathbf{U}\boldsymbol{\Gamma}$$

and the columns product matrix of $\mathbf{Z}_{(c)}$ equals the diagonal matrix $\beta_{(n)}$,

$$\mathbf{Z}'_{(c)}\mathbf{Z}_{(c)} = \beta_{(n)}.$$

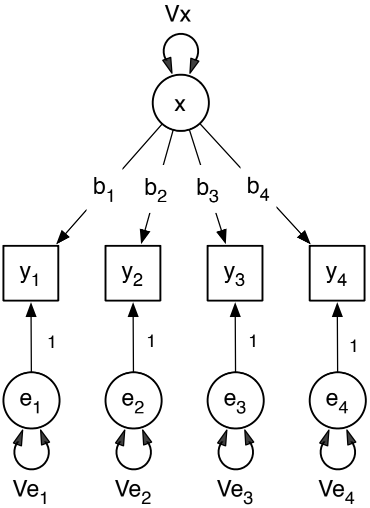
Principal Roots and Vectors

- A similar orthogonal rotation of row axes of $\mathbf{X}$ may be performed with matrix $\mathbf{U}'$ to $\mathbf{Z}_{(\mathbf{r})}$ so that

$$\begin{aligned} \mathbf{U}'\mathbf{X} &= \mathbf{Z}_{(r)}, \\ \mathbf{Z}_{(r)} &= \mathbf{\Gamma}\mathbf{W}, \\ \mathbf{Z}_{(r)}\mathbf{Z}'_{(r)} &= \beta_{(N)}\ . \end{aligned}$$

- The m^{**} columns of $\mathbf{Z}_{(c)}$, $m^{**} = (r+1), ..., n$, will contain all zero coordinates of the row vectors of $\mathbf{X}$. (Similarly, the m^{*} rows of $\mathbf{Z}_{(r)}$, $m^{*} = (r+1), ..., N$, will contain all zero coordinates of the column vectors of $\mathbf{X}$.)

Common Factor as a Latent Variable



Common Factor as a Latent Variable

- Suppose we have an $N \times 4$ data matrix $\mathbf{Y}$ where each column has a mean of zero.
- Construct a multivariate linear combination such that

$$\begin{aligned} y_{i1} &= b_{11} x_{i1} + e_{i1} \\ y_{i2} &= b_{12} x_{i1} + e_{i2} \\ y_{i3} &= b_{13} x_{i1} + e_{i3} \\ y_{i4} &= b_{14} x_{i1} + e_{i4} \end{aligned}$$

which we write in matrix format as

$$\mathbf{Y} = \mathbf{X}\mathbf{B} + \mathbf{E}$$

where

$$\mathbf{B} = \left[\begin{array}{cccc} b_{11} & b_{12} & b_{13} & b_{14} \end{array} \right]$$

and $\mathbf{E}$ is an $N \times 4$ matrix of residuals with zero mean.

Common Factor as a Latent Variable

- So, since

$$\mathbf{Y} = \mathbf{XB} + \mathbf{E}$$

let's try to substitute and work out the covariance of $\mathbf{Y}$.

$$\begin{aligned}
\mathbf{C_{YY}} &= 1/N(\mathbf{Y}'\mathbf{Y}) \\
&= 1/N((\mathbf{XB}+\mathbf{E})'(\mathbf{XB}+\mathbf{E})) \\
&= 1/N(((\mathbf{XB})'+\mathbf{E}')(\mathbf{XB}+\mathbf{E})) \\
&= 1/N((\mathbf{B}'\mathbf{X}'+\mathbf{E}')(\mathbf{XB}+\mathbf{E})) \\
&= 1/N(\mathbf{B}'\mathbf{X}'\mathbf{XB}+\mathbf{B}'\mathbf{X}'\mathbf{E}+\mathbf{E}'\mathbf{XB}+\mathbf{E}'\mathbf{E}) \\
&= 1/N(\mathbf{B}'\mathbf{X}'\mathbf{XB}+\mathbf{B}'0+0\mathbf{B}+\mathbf{E}'\mathbf{E}) \\
&= 1/N(\mathbf{B}'\mathbf{X}'\mathbf{XB}+\mathbf{E}'\mathbf{E}) \\
&= \mathbf{B}'(1/N\mathbf{X}'\mathbf{X})\mathbf{B}+1/N\mathbf{E}'\mathbf{E} \\
&= \mathbf{B}'\mathbf{C_{XX}}\mathbf{B}+\mathbf{C_{EE}}
\end{aligned}$$

Principal Component as a Latent Variable

- Everything on the right hand side is unknown!
- This is impossible to estimate, right?
- But we do know some things about the right hand side:
  - The covariance of the residuals is a diagonal matrix.
  - The covariance matrix of $\mathbf{X}$ is just Vx.
  - The covariance matrix of $\mathbf{Y}$ is symmetric.
- Let's try making some assumptions
  - $\mathbf{C_{EE}} = 0$
  - $\mathbf{C_{XX}} = \mathbf{I}$
- So, now we have

$$\begin{aligned}\mathbf{C_{YY}} &= \mathbf{B}'\mathbf{C_{XX}}\mathbf{B} + \mathbf{C_{EE}} \\ \mathbf{C_{YY}} &= \mathbf{B}'\mathbf{I}\mathbf{B} + 0 \\ \mathbf{C_{YY}} &= \mathbf{B}'\mathbf{I}\mathbf{B}\end{aligned}$$

Principal Component as a Latent Variable

- But this looks familiar from a few slides ago.
- If the matrix $\mathbf{C_{YY}}$ is symmetric, then the matrix $\mathbf{V}$ of characteristic vectors will be an orthogonal transformation so that

$$\mathbf{V}' = \mathbf{V}^{-1},$$

and so

$$\mathbf{C_{YY}} = \mathbf{V \Lambda V}'.$$

- So, if we can find a transformation matrix that will turn $\mathbf{\Lambda}$ into $\mathbf{I}$ we could use eigenvalue eigenvector decomposition to find the values for $\mathbf{B}$.

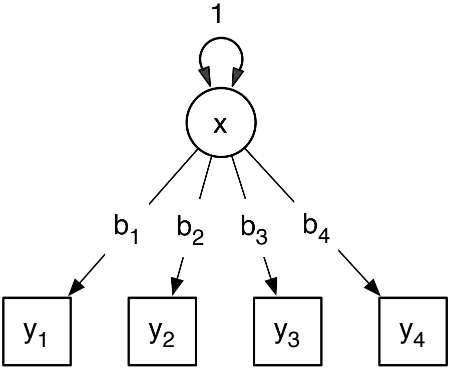
Principal Component as a Latent Variable

- Let $\mathbf{TT}' = \mathbf{\Lambda}$.
- This is easy to calculate if $\mathbf{\Lambda}$ is a diagonal matrix with nonnegative elements: just take the square root of each element down the diagonal.
- Now

$$
\begin{aligned}
\mathbf{C_{YY}} &= \mathbf{V\Lambda V}' \\
\mathbf{C_{YY}} &= \mathbf{VTT}'\mathbf{V}' \\
\mathbf{C_{YY}} &= \mathbf{VTIT}'\mathbf{V}' \\
\mathbf{C_{YY}} &= (\mathbf{VT})\mathbf{I}(\mathbf{VT})'
\end{aligned}
$$

- So if $\mathbf{B}'$ is the first column of $\mathbf{VT}$, our expected covariance matrix is the first principal component matrix of $\mathbf{Y}$.

Principal Component as a Latent Variable



Principal Component as a Latent Variable

- So, what is the expected covariance matrix for one principal component of four variables?

$$\mathbf{C_{YY}} = \mathbf{B'IB}$$

and so

$$\mathbf{C_{YY}} = \begin{bmatrix} (b_1)^2 & b_1 b_2 & b_1 b_3 & b_1 b_4 \\ b_1 b_2 & (b_2)^2 & b_2 b_3 & b_2 b_4 \\ b_1 b_3 & b_2 b_3 & (b_3)^2 & b_3 b_4 \\ b_1 b_4 & b_2 b_4 & b_3 b_4 & (b_4)^2 \end{bmatrix}$$

Common Factor as a Latent Variable

- How many degrees of freedom do we have?
  - How many statistics?

$$\frac{4 \times 5}{2} = 10$$

  - We are estimating 4 parameters so

$$10 - 4 = 6$$

- One assumption with principal components analysis is that there is no residual.
- That seems unlikely.
- How can we allow for unique factors as well as common factors?

Common Factor as a Latent Variable

- Let's try making some assumptions
  - $\mathbf{C_{EE}}$ =squared multiple correlation.
  - $\mathbf{C_{XX}} = \mathbf{I}$
- So, now we have

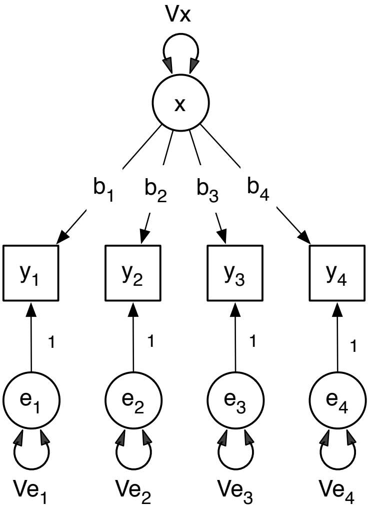
$$\begin{aligned} \mathbf{C_{YY}} &= \mathbf{B'C_{XX}B} + \mathbf{C_{EE}} \\ \mathbf{C_{YY}} &= \mathbf{B'IB} + \mathbf{SMC} \\ \mathbf{C_{YY}} - \mathbf{SMC} &= \mathbf{B'IB} \end{aligned}$$

- If we do the same eigenvalue eigenvector decomposition trick to find $\mathbf{B}$, we have what is known as the "principal axis solution to the common factor problem", which is usually written as

$$\mathbf{R} - \mathbf{U}^2 = \mathbf{ALA'}$$

- In SEM we use iterative methods to give better estimates of both $\mathbf{B}$ and $\mathbf{C_{EE}}$.

Common Factor as a Latent Variable



Common Factor as a Latent Variable

- So, what is the expected covariance matrix for one principal component of four variables?

$$\mathbf{C_{YY}} = \mathbf{B'IB} + \mathbf{C_{EE}}$$

and so

$$\mathbf{C_{YY}} = \begin{bmatrix} (b_1)^2 + Ve_1 & b_1 b_2 & b_1 b_3 & b_1 b_4 \\ b_1 b_2 & (b_2)^2 + Ve_2 & b_2 b_3 & b_2 b_4 \\ b_1 b_3 & b_2 b_3 & (b_3)^2 + Ve_3 & b_3 b_4 \\ b_1 b_4 & b_2 b_4 & b_3 b_4 & (b_4)^2 + Ve_4 \end{bmatrix}$$

Identification

- Note that I can transform to an equally good solution by.

$$\begin{aligned} \mathbf{C_{YY}} &= \mathbf{B'C_{XX}B} + \mathbf{C_{EE}} \\ \mathbf{C_{YY}} &= \mathbf{B'T'T'^{-1}C_{XX}T^{-1}TB} + \mathbf{C_{EE}} \\ \mathbf{C_{YY}} &= \mathbf{(TB)'(T'^{-1}C_{XX}T^{-1})(TB)} + \mathbf{C_{EE}} \end{aligned}$$

- So, if

$$\mathbf{B}^\star = \mathbf{TB}$$

  and

$$\mathbf{C_{XX}}^\star = \mathbf{T'^{-1}C_{XX}T^{-1}}$$

  then

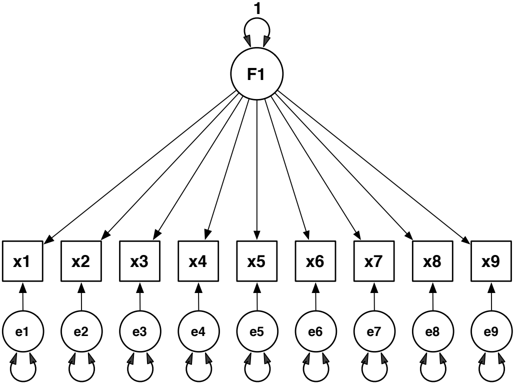
$$\mathbf{C_{YY}} = (\mathbf{B}^\star)'\mathbf{C_{XX}}^\star\mathbf{B}^\star + \mathbf{C_{EE}}$$

- So, there are as many equally good solutions as there are transformation matrices! (hint: ∞)

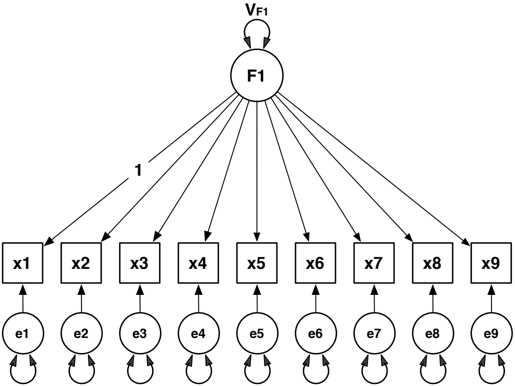
Identification

- That was the **identification problem** for latent variables.
  - Unless you specify something as fixed, there are an infinite number of solutions to latent variable regression coefficients.
- A simple rule that works (but sometimes is useful to break) is that you must have at least one of the following:
  1. At least one fixed value arrow leading away from the latent variable.
  2. A fixed value for the variance of the latent variable.
  3. A equality constraint on an arrow leading away from the variable that resolves itself to identification somewhere else in the model.
- This is often called **setting the scale** for the latent variable.

One Factor with Fixed Variance

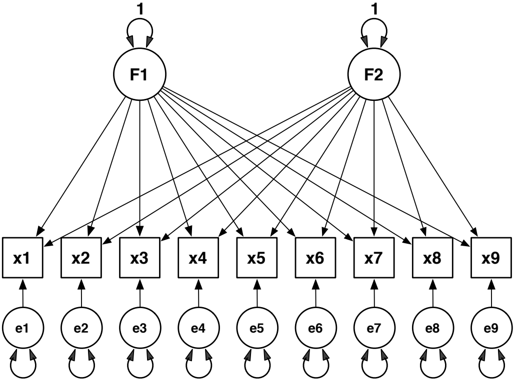


One Factor with Fixed Loading

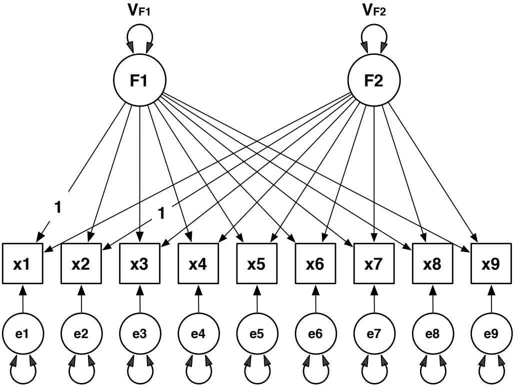


- Example One Factor Confirmatory Models:
  - OneFactorCov-OpenMx100221.R
  - OneFactorRaw-OpenMx100221.R.

Two Orthogonal Factors with Fixed Variance

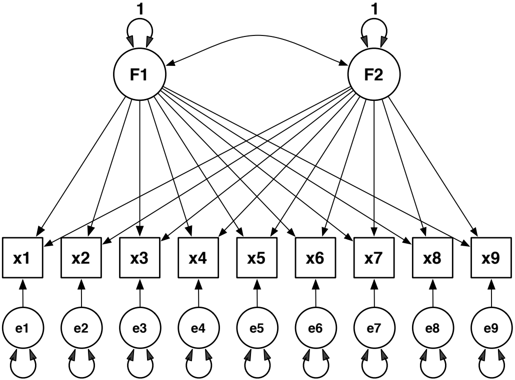


Two Orthogonal Factors with Fixed Loadings

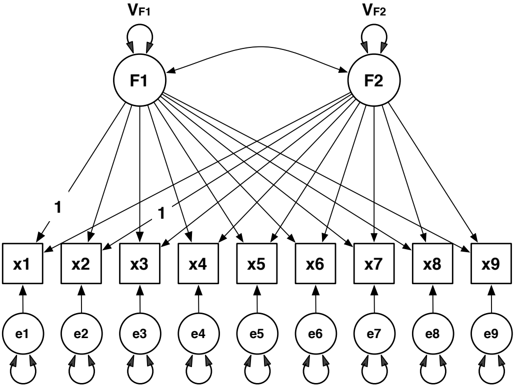


- Example Two Orthogonal Factors:
  - TwoFactorOrthoCov-OpenMx100221.R
  - TwoFactorOrthoRaw-OpenMx100221.R.

Two Oblique Factors with Fixed Variance

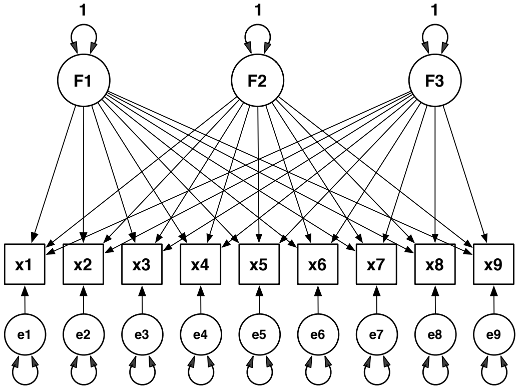


Two Oblique Factors with Fixed Loadings

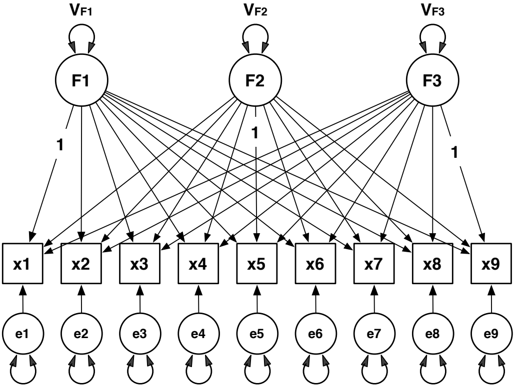


- Example Two Oblique Factors:
  - TwoFactorObliqueCov-OpenMx100221.R
  - TwoFactorObliqueRaw-OpenMx100221.R.

Three Orthogonal Factors with Fixed Variance

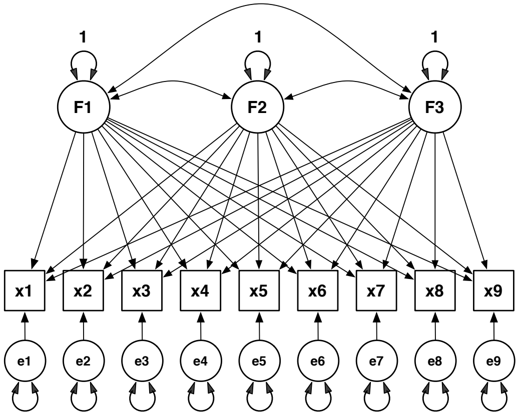


Three Orthogonal Factors with Fixed Loadings

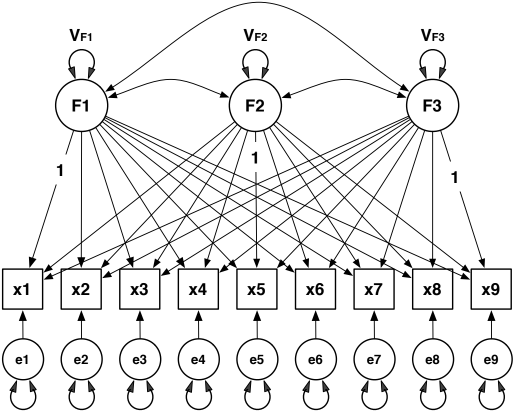


- Example Three Orthogonal Factors:
  - ThreeFactorOrthoCov-OpenMx100221.R
  - ThreeFactorOrthoRaw-OpenMx100221.R.

Three Oblique Factors with Fixed Variance



Three Oblique Factors with Fixed Loadings

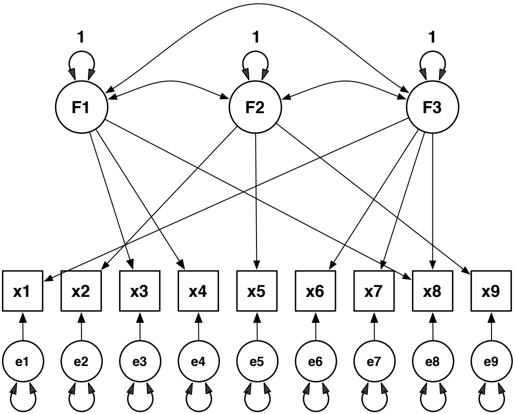


- Example Three Oblique Factors:
  - ThreeFactorObliqueCov-OpenMx100221.R
  - ThreeFactorObliqueRaw-OpenMx100221.R.

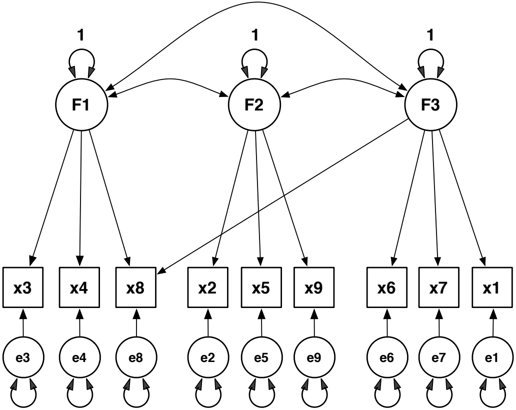
Simple Structure

- Let's try and modify ThreeFactorObliqueCov-OpenMx100221.R in order to find simple structure.

Simple Structure



Simple Structure



Meaning

- How can we name a latent variable?
- We certainly don't want meaningless latent variables in our models!
- The meaning of the variable can be found by asking yourself the question
  - "What is common to all of the variables to which the arrows from the latent variable point?"
- You find the answer to that question by understanding your **indicator variables** (the variables to which the latent variable points) and looking carefully at the **loading matrix** (the estimated values for the arrows pointing away from the latent variable).
- Altogether, this is called a **measurement model**.

Latent Structure

- Latent constructs are identified from manifest variables.
- Your theory is likely a set of relations between latent variables.
- These relations can have all of the characteristics we have covered for manifest variables.
  1. Covariances between latent variables.
  2. Regression coefficients.
  3. Additional measurement models: Latent variables can be indicators for other latent variables.
- All the same path tracing rules still apply.
- The same **A** and **S** matrices are constructed in order to fit the models.

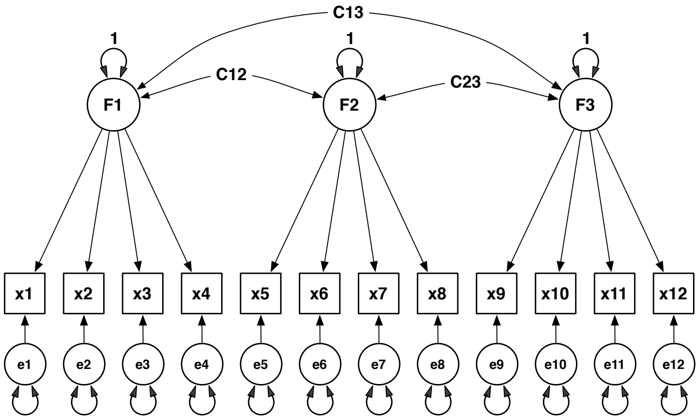
Latent Structure

- In order to make your theories into models.
  1. Decide on **several** ways your latent constructs may interrelate.
  2. Decide how your measured variables indicate the latent variables.
- Notice the emphasis on several models.
- This allows you to perform model comparisons.
- These comparisons are made between models with different constraints on the latent structure.
- A good strategy is to find a minimally complex and maximally complex model as well as the theories you are actually interested in.

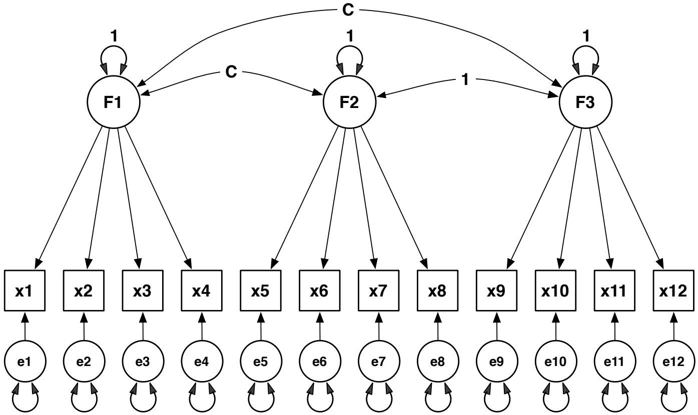
Latent Structure: Correlated Factors

- Suppose you have a multiple factor model.
- You may wonder how these factors are correlated.
- You may also wonder how many factors are sufficient.
- One way to examine this is by placing constraints on the factor intercorrelations.

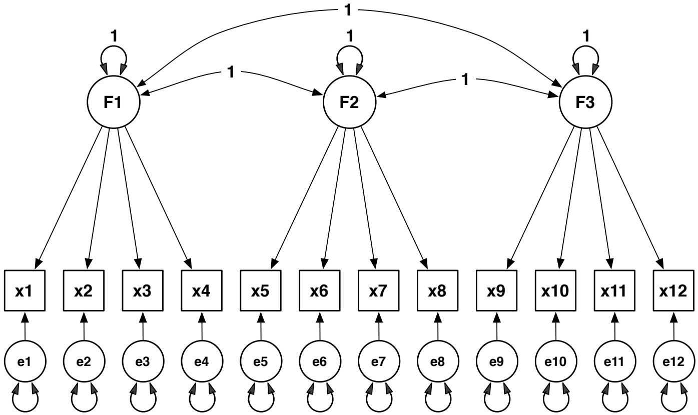
Latent Structure: Correlated Factors



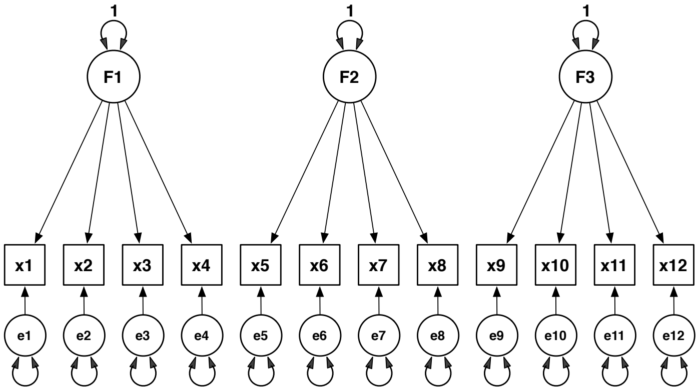
Latent Structure: Correlated Factors



Latent Structure: Correlated Factors



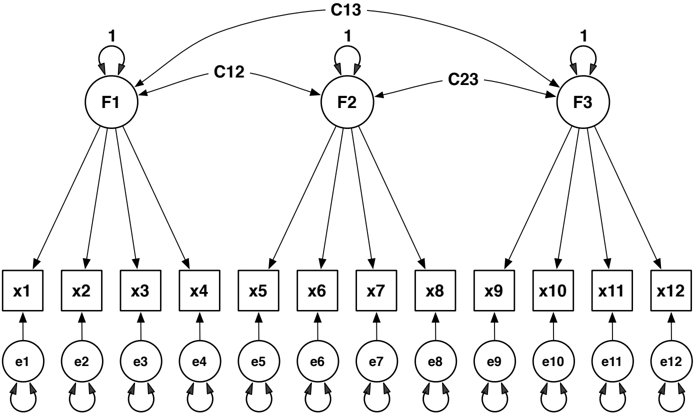
Latent Structure: Correlated Factors



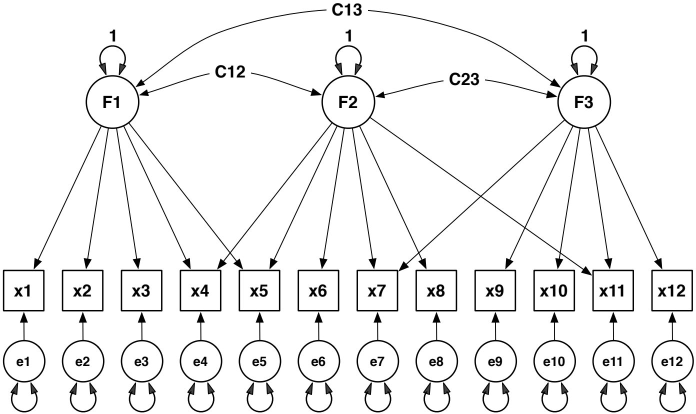
Latent Structure: Crossed Loadings

- Some times simple structure is not obtainable.
- In this case we might have *crossed loadings*.
- The idea is that some indicators have a special relationship with more than one factor.
- This can be specified, but it leads to more complicated patterns of covariance between indicators.

Latent Structure: Crossed Loadings



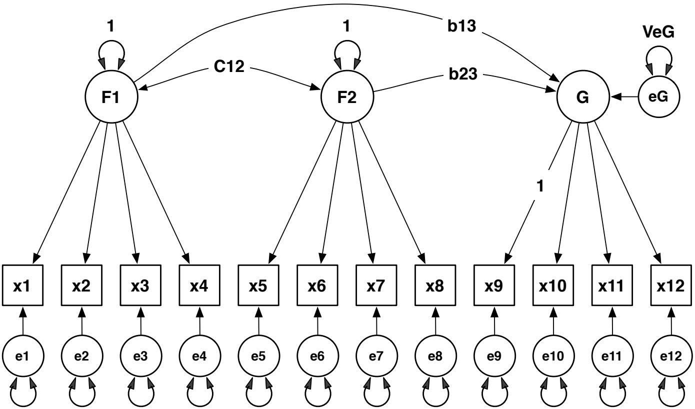
Latent Structure: Crossed Loadings



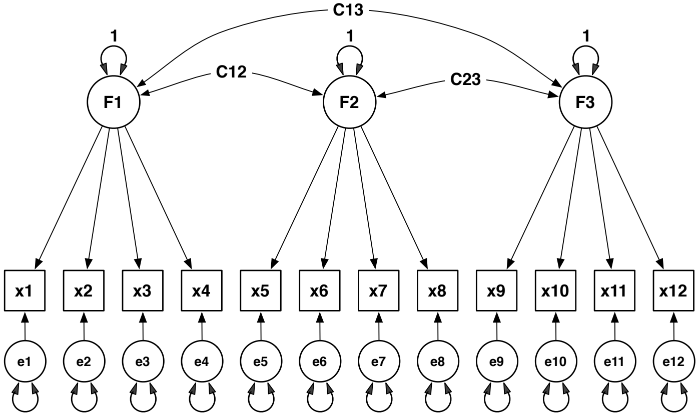
Latent Structure: Structural Regression Models

- One of the most common models is where one latent construct is predicted by one or more other latent constructs.
- There are many variants on this type of model.
- Let's consider latent structural multiple regression.

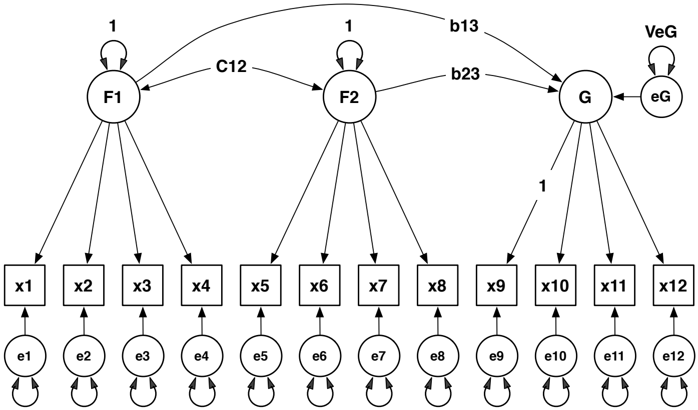
Latent Structure: Structural Regression Models



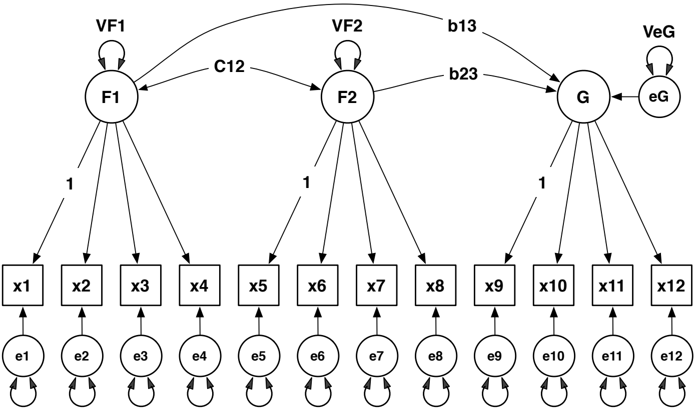
Latent Structure: Structural Regression Models



Latent Structure: Structural Regression Models



Latent Structure: Structural Regression Models



Next Lecture

- Maximum Likelihood.
- Fit Statistics.
- Assumptions Review.
- Diagnostics.