Path Analysis with Latent Variables

In a fashion analogous to Path Analysis with manifest variables, Path Analysis with latent variables allows researchers to **test putative causal relationships among latent variables**. The structural equations are as follows:

$$\boldsymbol{\eta} = \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\Gamma}\boldsymbol{x} + \boldsymbol{\zeta} \qquad (1)$$

The **reduced form** of (1) which results from solving (1) for η is

$$\boldsymbol{\eta} = (\boldsymbol{I} - \mathbf{B})^{-1}\boldsymbol{\Gamma}\boldsymbol{x} + (\boldsymbol{I} - \mathbf{B})^{-1}\boldsymbol{\zeta} \,. \qquad (2)$$

The variables on the y-side are explained as follows:

$$\boldsymbol{y} = \boldsymbol{\Lambda}_y\boldsymbol{\eta} + \boldsymbol{\varepsilon} \qquad (3)$$

Inserting (2) into (3) yields

$$\boldsymbol{y} = \boldsymbol{\Lambda}_y(\boldsymbol{I} - \mathbf{B})^{-1}\boldsymbol{\Gamma}\boldsymbol{x} + \boldsymbol{\Lambda}_y(\boldsymbol{I} - \mathbf{B})^{-1}\boldsymbol{\zeta} + \boldsymbol{\varepsilon} \,. \qquad (4)$$

In words, the path model with latent variables involves

- **exogenous factors**, that is, factors used to explain other factors; these are the factors ‘on the x-side’
- **endogenous factors**, that is, factors explained (‘caused’) by exogenous factors; these are the factors ‘on the y-side’; and
- **mediating factors**, that is, factors between exogenous and endogenous factors; mediating factors can be either on the x-side or on the y-side

Notice that the only difference between manifest and latent variable path analysis is that whereas manifest variable path analysis relates observed variables to each other, **latent variable path analysis relates factors to each other**.

Matrices involved:

1. On the x-side of the model

- Λ_x: loadings of x-variables on ξ-factors (NX x NK)
- Θ_δ: variance/covariance matrix of residuals of x-variables (NX x NX)
- Φ: variance covariance matrix of ξ-factors (NK x NK)

2. On the y-side of the model

- Λ_y: loadings of y-variables on η-factors (NY x Ne)
- Θ_ϵ: variance/covariance matrix of residuals of y-variables (NY x NY)
- B: structural matrix of (asymmetric) relationships between η-factors (NE x NE)
- Ψ: variance/covariance matrix of η-factors (NE x NE)

3. Relating the x- and the y-side of the model

- Γ: structural matrix of (asymmetric) relationships between factors on the x-side (ξ-factors) and factors on the y-side (η-factors)

Thus, the path model with latent variables can be expressed using the following three terms:

1. **The measurement model on the x-side**

$$\boldsymbol{x} = \boldsymbol{\Lambda}_x \boldsymbol{\xi} + \boldsymbol{\delta}, \tag{5}$$

2. **The Structural Equation model**

$$\boldsymbol{\eta} = \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta} \tag{6}$$

3. **The measurement model on the y-side**

$$\boldsymbol{y} = \boldsymbol{\Lambda}_y \boldsymbol{\eta} + \boldsymbol{\varepsilon} \,. \tag{7}$$

Data Example (Forest Succession Simulation)

Sample: N = 89 forest sample plots

Variables: % overstory and % climax species, measured in four consecutive decades: O1, C1, O2, C2, O3, C3, O4, D4.

Model:

- the manifest variables constitute factors of physical change, each of which has two indicators: O. and C.

Results (all parts of model specified at the y-side)

- excellent fit in all respects