Assignment 1

SEM 1: Confirmatory Factor Analysis

Please hand in a `.pdf` file containing your answers and, if applicable, a `.R`, `.Rmd` (compile to PDF) or `Rnw` file containing your codes. The deadline of this assignment is Thursday April 12 15:00.

Questions

Given the general CFA framework:

$$\begin{aligned} \boldsymbol{y}_i &= \boldsymbol{\Lambda}\boldsymbol{\eta}_i + \boldsymbol{\varepsilon}_i \\ \boldsymbol{y} &\sim N(\boldsymbol{0}, \boldsymbol{\Sigma}) \\ \boldsymbol{\eta} &\sim N(\boldsymbol{0}, \boldsymbol{\Psi}) \\ \boldsymbol{\varepsilon} &\sim N(\boldsymbol{0}, \boldsymbol{\Theta}), \end{aligned}$$

Let:

$$\boldsymbol{\Lambda} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}, \boldsymbol{\Theta} = \begin{bmatrix} 3 & & \\ 0 & 7 & \\ 0 & 0 & 9 \end{bmatrix}$$

(note that upper-triangular elements of $\boldsymbol{\Theta}$ are also zero, but not shown for notational clarity because $\boldsymbol{\Theta}$ is symmetrical). Suppose the implied variance–covariance matrix equals:

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

Question 1 (1 point) Compute $\boldsymbol{\Sigma}$ by hand. Verify your work in R. ▪

Question 2 (1 point) What is the value of the implied variance, $\boldsymbol{\Psi}$, of the single latent variable in this model? To what form of *scaling* does this value relate? ▪

Question 3 (1 point) How much *variance* is explained by the latent variable in the first variable y_1? ▪

Question 4 (1 point) What is the model-implied *correlation* (not covariance) between y_1 and y_3? ▪

Now let:

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_{11} \\ \lambda_{21} \\ \lambda_{31} \\ \lambda_{41} \end{bmatrix}$$

$$\boldsymbol{\Theta} = \begin{bmatrix} \theta_{11} & & & \\ 0 & \theta_{22} & & \\ 0 & 0 & \theta_{33} & \\ 0 & 0 & 0 & \theta_{44} \end{bmatrix}$$

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

Question 5 (1 point) Draw a path diagram for this model. ▪

Question 6 (2 points) Derive an expression for each element of $\boldsymbol{\Sigma}$. ▪

Question 7 (1 point) How many degrees of freedom does this model have? Is it identified? ▪

Now let:

$$\mathbf{\Lambda} = \begin{bmatrix} 0.7 & 0 \\ 0.7 & 0 \\ 0.7 & 0 \\ 0 & 0.7 \\ 0 & 0.7 \\ 0 & 0.7 \end{bmatrix}$$

$$\mathbf{\Psi} = \begin{bmatrix} 1 & 0.3 \\ 0.3 & 1 \end{bmatrix}$$

$$\mathbf{\Theta} = \begin{bmatrix} 0.51 & & & & & \\ 0 & 0.51 & & & & \\ 0 & 0 & 0.51 & & & \\ 0 & 0 & 0 & 0.51 & & \\ 0 & 0 & 0 & 0 & 0.51 & \\ 0 & 0 & 0 & 0 & 0 & 0.51 \end{bmatrix}$$

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

Question 8 (1 point) Compute $\mathbf{\Sigma}$.

Now let:

$$\mathbf{\Lambda} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

$$\mathbf{\Psi} = \begin{bmatrix} 0.490 & 0.147 \\ 0.147 & 0.490 \end{bmatrix}$$

$$\mathbf{\Theta} = \begin{bmatrix} 0.51 & & & & & \\ 0 & 0.51 & & & & \\ 0 & 0 & 0.51 & & & \\ 0 & 0 & 0 & 0.51 & & \\ 0 & 0 & 0 & 0 & 0.51 & \\ 0 & 0 & 0 & 0 & 0 & 0.51 \end{bmatrix}$$

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

Question 9 (1 point) Recompute $\mathbf{\Sigma}$ and compare your answer to the previous question. Can you explain the result?

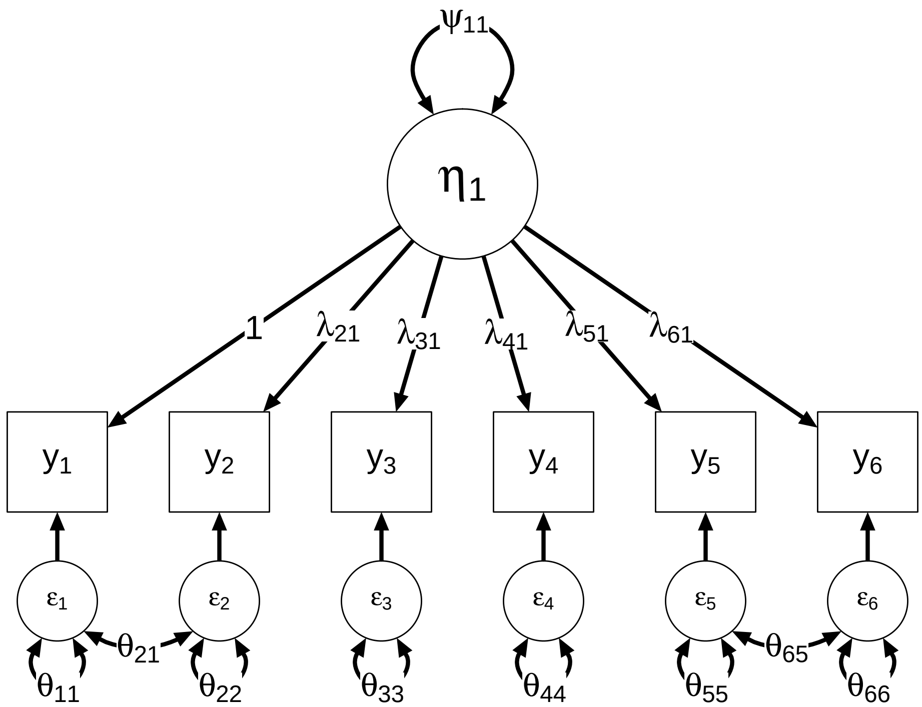
Suppose we observe:

$$\boldsymbol{S} = \begin{bmatrix} 1.03 & & & & & \\ 0.49 & 1.02 & & & & \\ 0.52 & 0.49 & 1.05 & & & \\ 0.63 & 0.14 & 0.15 & 0.97 & & \\ 0.10 & 0.18 & 0.13 & 0.45 & 0.97 & \\ 0.15 & 0.18 & 0.13 & 0.49 & 0.50 & 1.02 \end{bmatrix}$$

Question 10 (2 points) Compute F_{ML}. Tips: ln is the natural logarithm operator. In R, `log(x)` gives the natural logarithm of `x`, `det(x)` gives the determinent of `x` and `sum(diag(x))` gives the trace of `x`.

Question 11 (2 points) Compare your computed $\boldsymbol{\Sigma}$ to the observed $\boldsymbol{S}$. Do you think the model fits well? Is there a parameter you can think of to add to improve model fit? ▪

Given the following path diagram:



Question 12 (1 point) Is this model identified? How many degrees of freedom are there? ▪

Question 13 (1 point) Write down $\boldsymbol{\Lambda}$, $\boldsymbol{\Theta}$ and $\boldsymbol{\Psi}$ ▪

Question 14 (2 points) What is the model implied $\text{Cov}(y_1, y_2)$? ▪

Challenge Question

This question is **bonus**, and can lead to one bonus point! Consider the following model:

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_{11} \\ \lambda_{21} \\ \lambda_{31} \end{bmatrix}$$
$$\mathbf{\Psi} = \begin{bmatrix} 1 \end{bmatrix}$$
$$\mathbf{\Theta} = \begin{bmatrix} \theta_{11} & & \\ 0 & \theta_{22} & \\ 0 & 0 & \theta_{33} \end{bmatrix}$$
$$\mathbf{\Sigma} = \mathbf{\Lambda}\mathbf{\Lambda}^\top + \mathbf{\Theta}$$

This model is already identified. Suppose we observed the *maximum likelihood estimator* sample variance–covariance matrix $\boldsymbol{S}$:

$$\boldsymbol{S} = \begin{bmatrix} 1.01 & 0.50 & 0.50 \\ 0.50 & 1.01 & 0.47 \\ 0.50 & 0.47 & 0.95 \end{bmatrix}$$

Question 15 (1 point) Use `optim()` in R to estimate the $\mathbf{\Lambda}$ and $\mathbf{\Theta}$ matrices by minimizing F_{ML}. You can assume all parameters are positive. Lavaan gives the following solution to the parameters:

lhs	op	rhs	est	se	z	pvalue	ci.lower	ci.upper
Eta	=~	Y1	0.730	0.034	21.490	0	0.663	0.796
Eta	=~	Y2	0.685	0.034	20.347	0	0.619	0.751
Eta	=~	Y3	0.690	0.033	20.963	0	0.625	0.754
Y1	~~	Y1	0.477	0.036	13.086	0	0.406	0.548
Y2	~~	Y2	0.538	0.035	15.169	0	0.468	0.607
Y3	~~	Y3	0.478	0.034	14.081	0	0.412	0.545
Eta	~~	Eta	1.000	0.000	NA	NA	1.000	1.000

But Lavaan uses a different optimizer, so your results should be comparable, but may not be identical. ■