STA 312f07 Assignment 11

Do this assignment in preparation for the quiz on Friday, Nov. 30th. The questions are practice for the quiz, and are not to be handed in.

1. Consider the general factor analysis model

$$\mathbf{X} = \boldsymbol{\Lambda}\mathbf{F} + \mathbf{e},$$

where $\boldsymbol{\Lambda}$ is a p by m matrix of factor loadings, the vector of factors $\mathbf{F}$ is multivariate normal with expected value zero and covariance matrix $\mathbf{I}_m$ (the identity), and $\mathbf{e}$ is multivariate normal with expected value zero and covariance matrix $\boldsymbol{\Psi}$, a p by p *diagonal* matrix of error variances, all strictly greater than zero.

   (a) Calculate the matrix of covariances between the observable variables $\mathbf{X}$ and the underlying factors $\mathbf{F}$.

   (b) Give the covariance matrix of $\mathbf{X}$. Show your work.

   (c) Is this model identified? Answer Yes or No and prove your answer.

2. Here is a factor analysis model in which all the observed variables are *standardized*. That is, they are divided by their standard deviations as well as having the means subtracted off. This gives them mean zero and variance one. Therefore, we work with a correlation matrix rather than a covariance matrix; that's the classical way to do factor analysis.

$$\begin{aligned} X_1 &= \lambda_1 F_1 + e_1 \\ X_2 &= \lambda_2 F_2 + e_2 \\ X_3 &= \lambda_3 F_3 + e_3, \end{aligned}$$

where F_1, F_2 and F_3 are independent $N(0,1)$, e_1, e_2 and e_3 are normal and independent of each other and of F_1, F_2 and F_3, $V(X_1) = V(X_2) = V(X_3) = 1$, and λ_1, λ_2 and λ_3 are nonzero constants. The expected values of all random variables equal zero.

   (a) What is $V(e_1)$? $V(e_2)$? $V(e_3)$?

   (b) Give the communality of each observed variable. Recall that the communality is the proportion of variance explained by the common factor(s).

   (c) Give the variance-covariance matrix of the observed variables. It is a correlation matrix because the variances of all the observed variables equal one. (Recall $Corr(X,Y) = \frac{Cov(X,Y)}{SD(X)SD(Y)}$).

(d) What is $Corr(F_1, X_1)$?

(e) Is the model identified? Answer Yes or No and prove your answer.

3. Here is another factor analysis model. This one has a single underlying factor. Again, all the observed variables are standardized.

$$
\begin{aligned}
X_1 &= \lambda_1 F + e_1 \\
X_2 &= \lambda_2 F + e_2 \\
X_3 &= \lambda_3 F + e_3,
\end{aligned}
$$

where $F \sim N(0,1)$, e_1, e_2 and e_3 are normal and independent of F and each other with expected value zero, $V(X_1) = V(X_2) = V(X_3) = 1$, and λ_1, λ_2 and λ_3 are nonzero constants with $\lambda_1 > 0$.

(a) What is $V(e_1)$? $V(e_2)$? $V(e_3)$?

(b) Give the communality of each observed variable.

(c) Give the variance-covariance (correlation) matrix of the observed variables.

(d) What is $Corr(F, X_1)$?

(e) Is the model identified? Answer Yes or No and prove your answer.

4. In this factor analysis model, the observed variables are *not* standardized.

$$
\begin{aligned}
X_1 &= F + e_1 \\
X_2 &= \lambda_2 F + e_2 \\
X_3 &= \lambda_3 F + e_3,
\end{aligned}
$$

where $F \sim N(0, \phi)$, e_1, e_2 and e_3 are normal and independent of F and each other with expected value zero, $V(e_1) = \psi_1, V(e_2) = \psi_2, V(e_3) = \psi_3$, and λ_2 and λ_3 are nonzero constants.

(a) Give the variance-covariance (correlation) matrix of the observed variables.

(b) What is $Corr(F, X_1)$? What is $Corr(F, X_1)^2$? Does this look familiar from an early assignment?

(c) What is $Corr(F, X_2)^2$? Again, it's the proportion of the observed variable's variance that is *not* error.

(d) Is the model identified? Answer Yes or No and prove your answer.

5. We now extend the preceding model by adding another factor. Let

$$
\begin{aligned}
X_1 &= F_1 + e_1 \\
X_2 &= \lambda_2 F_1 + e_2 \\
X_3 &= \lambda_3 F_1 + e_3 \\
X_4 &= F_2 + e_4 \\
X_5 &= \lambda_5 F_2 + e_5 \\
X_6 &= \lambda_6 F_2 + e_6,
\end{aligned}
$$

where all expected values are zero, $V(e_i) = \psi_i$ for $i = 1, \ldots, 6$,

$$
V \left[\begin{array}{c} F_1 \\ F_2 \end{array} \right] = \left[\begin{array}{cc} \phi_{11} & \phi_{12} \\ \phi_{12} & \phi_{22} \end{array} \right],
$$

and $\lambda_2, \lambda_3, \lambda_5$ and λ_6 are nonzero constants.

(a) Give the covariance matrix of the observable variables. Show the necessary work. A lot of the work has already been done in Question 4.

(b) Is this model identified? Answer Yes or No and prove your answer.

6. What do you think would happen if we added a third factor to the model of Question 5? Would it be identified? You don't have to do any calculations; just think about it and see the pattern.