STA 313F2004 Assignment 6

The problems in this assignment are about model identification. In the classical structural equation models, where the variables are multivariate normal with expected value zero, a model is identified if and only if it is possible to solve uniquely for the set of model parameters $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$ in terms of the unique elements $\sigma_{i,j}$ of the variance-covariance matrix $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\theta})$ of the manifest variables. That is, the function $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ is a one-to-one function.

In practice, this means that you are dealing with a system of $\frac{p(p+1)}{2}$ equations in k unknowns, where there are p manifest variables and k parameters in the model. You are trying to either show that the model is identified, or that it is not identified.

- To prove that a model *is* identified, the technique you have available now is to explicitly solve for the θ values in terms of the $\sigma_{i,j}$ values. If you obtain a unique solution, you have proved that the model is identified.

- To prove that a model is *not* identified, there are two main techniques.
  - If at any point in the process, you find you have more unknowns than equations, a unique solution is impossible, and you are done. In particular, always compare k and $\frac{p(p+1)}{2}$ before doing anything else. Of course, $\frac{p(p+1)}{2} \geq k$ is just a necessary condition for identification; it's not sufficient.
  - To prove that a model not identified, it is enough to produce two distinct values of $\boldsymbol{\theta}$ that yield the same $\boldsymbol{\Sigma}$. A simple numerical example is best.

Do this assignment in preparation for the quiz, which may be on Friday, Nov. 26th, or possibly Monday Nov. 29th. The problems are not to be handed in; do them in preparation for the quiz.

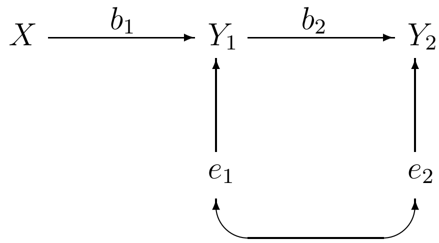
1. Consider the model $y = b_1 x_1 + b_2 x_2 + e$, where $x_1 \sim N(0, \sigma_1^2)$, $x_2 \sim N(0, \sigma_2^2)$, $Cov(x_1, x_2) = \sigma_{12}$, $e \sim N(0, \sigma_e^2)$, and e is independent of x_1 and x_2. Is this model identified? Prove your answer.

2. Consider this model:

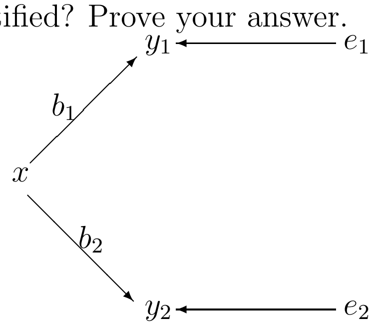
$$
\begin{aligned}
x_1 &= F_1 + e_1 \\
x_2 &= F_2 + e_2 \\
y &= b_1 F_1 + b_2 F_2 + e_3,
\end{aligned}
$$

where the variables are multivariate normal with expectation zero, F_1 and F_2 are correlated with each other but independent of e_1, e_2 and e_3, and the error terms e_1, e_2 and e_3 are independent of each other. Is this model identified? Prove your answer.

3. Is the following model identified? Prove your answer.



4. Is this model identified? Prove your answer.



5. Consider this model:

$$
\begin{aligned}
y_1 &= b_1x_1 + b_2x_2 + e_1 \\
y_2 &= b_3x_2 + b_4x_3 + e_2 \\
y_3 &= b_5y_1 + b_6y_2 + e_3,
\end{aligned}
$$

where all random variables have expected value zero, the vector $\mathbf{x} = (x_1, x_2, x_3)'$ is multivariate normal, and the error terms e_1, e_2 and e_3 are normal, and independent of $\mathbf{x}$ and of each other. As usual, the manifest variables are the x and y variables. Is the model identified? Prove your answer.

6. Consider the model $y = b_1x_1 + b_2x_2 + e$, where x_1 and x_2 are independent with $V(x_1) = V(x_2) = 1$, $V(e) = \sigma^2_e$, and x_1 is independent of e, but $Cov(x_2, e) = k \neq 0$. As usual, all random variables have expected value zero. Is the model identified? Prove your answer.

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

$$
\begin{aligned}
y_1 &= \gamma_1F_1 + e_1 \\
y_2 &= \gamma_2F_2 + e_2 \\
y_3 &= \gamma_3F_3 + e_3,
\end{aligned}
$$

where F_1, F_2 and F_3 are independent $N(0,1)$, e_1, e_2 and e_3 are normal and independent with expected value zero, $V(y_1) = V(y_2) = V(y_3) = 1$, and γ_1, γ_2 and γ_3 are nonzero constants.

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

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

(c) What is $Corr(F_1, y_1)$?

(d) Is the model identified? Prove your answer.

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

$$\begin{aligned} y_1 &= \gamma_1 F + e_1 \\ y_2 &= \gamma_2 F + e_2 \\ y_3 &= \gamma_3 F + e_3, \end{aligned}$$

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, $V(y_1) = V(y_2) = V(y_3) = 1$, and γ_1, γ_2 and γ_3 are nonzero constants with $\gamma_1 > 0$. Is the model identified? Prove your answer.

9. Consider this model:

$$\begin{aligned} y_1 &= b_1 x_1 + e_1 \\ y_2 &= b_2 x_2 + e_2, \end{aligned}$$

Where all the variables have expected value zero, the independent variables x_1 and x_2 are independent of the error terms e_1 and e_2, $Cov(x_1, x_2) = c \neq 0$, and $Cov(e_1, e_2) = k \neq 0$. Is the model identified? Prove your answer.