STA 431s13 Assignment Nine[1]

For the SAS question, please bring your log and list files to the quiz. Do not write anything on the printouts except your name and student number. The other questions are just practice for the quiz on Friday March 22nd, and are not to be handed in.

1. The general structural equation model (with zero expected values and no intercepts) is

$$\begin{aligned} \mathbf{Y}_i &= \boldsymbol{\beta}\mathbf{Y}_i + \boldsymbol{\Gamma}\mathbf{X}_i + \boldsymbol{\epsilon}_i \\ \mathbf{F}_i &= \begin{pmatrix} \mathbf{X}_i \\ \mathbf{Y}_i \end{pmatrix} \\ \mathbf{D}_i &= \boldsymbol{\Lambda}\mathbf{F}_i + \mathbf{e}_i \end{aligned}$$

   More details are given in the lecture slides.

   (a) Calculate $V(\mathbf{Y}_i)$ in terms of $\boldsymbol{\beta}$, $\boldsymbol{\Gamma}$, $\boldsymbol{\Phi}_{11}$ and $\boldsymbol{\Psi}$. Show your work.

   (b) Calculate $C(\mathbf{X}_i, \mathbf{Y}_i)$ in terms of $\boldsymbol{\beta}$, $\boldsymbol{\Gamma}$, $\boldsymbol{\Phi}_{11}$ and $\boldsymbol{\Psi}$. Not all these matrices appear in the answer. Show your work.

   (c) Calculate $\boldsymbol{\Sigma} = V(\mathbf{D}_i)$ in terms of $\boldsymbol{\Lambda}$, $\boldsymbol{\Phi}$ and $\boldsymbol{\Omega}$. Show your work.

2. Suppose this is the latent variable part of a larger structural equation model, and all error terms are independent. Give the matrices $\boldsymbol{\beta}$ and $\boldsymbol{\Psi}$.

$$\begin{aligned} Y_1 &= \beta_3 Y_3 + \epsilon_1 \\ Y_2 &= \beta_1 Y_1 + \epsilon_2 \\ Y_3 &= \beta_2 Y_2 + \epsilon_3 \end{aligned}$$

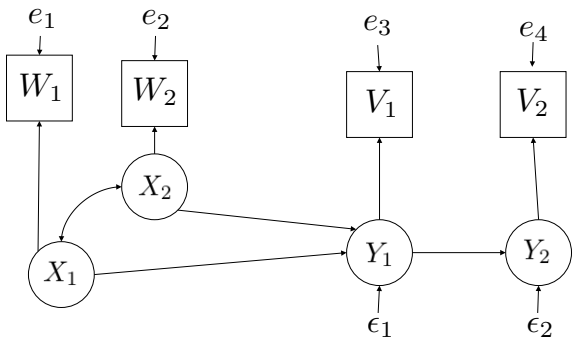
3. The following model has all expected values zero and zero covariance between all pairs of exogenous variables.

$$\begin{aligned} Y_1 &= \gamma_1 X_1 + \gamma_2 X_2 + \epsilon_1 \\ Y_2 &= \beta Y_1 + \gamma_3 X_1 + \epsilon_2 \\ W_1 &= \lambda_1 X_1 + e_1 \\ W_2 &= \lambda_2 X_2 + e_2 \\ V_1 &= \lambda_3 Y_1 + e_3 \\ V_2 &= \lambda_4 Y_2 + e_4 \end{aligned}$$

   Write the model equations in matrix form. Also give the matrices $\boldsymbol{\Phi}$, $\boldsymbol{\Psi}$ and $\boldsymbol{\Omega}$.

[1] Copyright information is at the end of the last page.

4. The following model has no intercepts, and all expected values equal zero.



   (a) Write coefficients on all the arrows in the latent variable model, except those coming from error terms. The coefficients on all the arrows from latent to observable variables are equal to one.

   (b) Write the model equations in scalar form.

   (c) Write the model equations in matrix form.

   (d) What is the matrix $\mathbf{\Lambda}$?

   (e) What is the matrix $\mathbf{\Phi}_{11}$? (Give your notation.)

   (f) What is the matrix $\mathbf{\Phi}$?

   (g) Does this model pass the test of the parameter count rule? Answer Yes or No and give *both* numbers.

5. Consider the general factor analysis model

$$\mathbf{D} = \mathbf{\Lambda F} + \mathbf{e},$$

   where $\mathbf{\Lambda}$ is a $k \times p$ matrix of factor loadings, the vector of factors $\mathbf{F}$ is a $p \times 1$ multivariate normal with expected value zero and covariance matrix $\mathbf{\Phi}$, and $\mathbf{e}$ is multivariate normal with expected value zero and covariance matrix $\mathbf{\Omega}$. All covariance matrices are positive definite.

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

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

   (c) Any positive definite matrix can be written as $\mathbf{SS}'$ (the matrix $\mathbf{S}$ is called the *square root matrix*). Using the square root matrix of $\mathbf{\Phi}$, show that the parameters of the general factor analysis model are not identifiable.

(d) In an attempt to obtain a model whose parameters can be successfully estimated, let $\mathbf{\Omega}$ be diagonal (errors are uncorrelated) and set $\mathbf{\Phi}$ to the identity matrix (standardizing the factors). Show that the parameters of this revised model are still not identifiable.

6. This is the SAS part of the farm co-op manager study described in the last question of Assignment 8. The data are given in the file `manager.data`. There is a link on the course web page in case the one in this document does not work. There are 98 co-ops, so please make sure you are reading the correct number of cases.

   (a) Using proc calis, fit the appropriate model. For comparison, my value of Akaike's Information Criterion (which will not be on the quiz) is 72.9046. Using your list file when necessary, be ready to answer questions like the following on the quiz.

      i. There is one manifest exogenous variable. What is it?

      ii. There is one latent endogenous variable. What is it?

      iii. Based on the number of covariance structure equations and the number of unknown paramters, how many equality restrictions should the model impose on the covariance matrix? The answer is a single number; you need not say what they all are.

      iv. Does your model fit the data adequately? Answer Yes or No and give three numbers: a chisquare statistic, the degrees of freedom, and a p-value.

      v. Controlling for knowledge, profit-loss orientation and job satisfaction, is there evidence that formal education is related to job performance? Answer Yes or No and give the value of a test statistic (actually it's a Z) that supports your conclusion. Of course in all these questions you are using the $\alpha = 0.05$ significance level and a 2-sided test.

      vi. Controlling for formal education, knowledge and profit-loss orientation, is there evidence that job satisfaction is related to job performance? Answer Yes or No and give the value of a test statistic (actually it's a Z) that supports your conclusion. If the answer is Yes, say whether satisfaction is positively related to performance, or negatively related.

      vii. Controlling for job satisfaction, formal education and knowledge, is there evidence that profit-loss orientation is related to job performance? Answer Yes or No and give the value of a test statistic (actually it's a Z) that supports your conclusion. If the answer is Yes, say whether profit-loss orientation is positively related to performance, or negatively related.

      viii. Carry out a Wald test of all the regression coefficients at once; use the `simtests` command. Be able to give the value of the chi-squared test statistic, the degrees of freedom, and the p-value – all numbers from your printout. Using the usual $\alpha = 0.05$ significance level, is there evidence that at least one regression coefficient must be non-zero?

ix. Estimate the reliability of Knowledge measurement 1. Your answer is a number. You will need a calculator.

x. Estimate the reliability of Knowledge measurement 2. Your answer is a number. You will need a calculator.

xi. Show that the reliabilities of Knowledge Measurements 1 and 2 are equal if and only if the variances of the two measurement error terms are equal. This is is a paper-and-pencil calculation. It is the basis of the next (and last) you are asked to carry out.

(b) Carry out a Wald (not likelihood ratio) test of the null hypothesis that the variances of the two measurement error terms for Knowledge measurements 1 and 2 are equal. By the last calculation you did, this is equivalent to testing whether the two reliabilities are equal.

i. What is the value of the chi-squared statistic? The answer is a number.

ii. What are the degrees of freedom? The answer is a number.

iii. What is the p-value? The answer is a number.

iv. Do you reject the null hypothesis at $\alpha = 0.05$? Answer Yes or No.

v. What do you conclude about the reliabilties of the two measurements? Do you have sufficient evidence to conclude that they are different? Answer Yes or No. If the answer is Yes, say which one seems to be more reliable. Of course the answer may not be Yes. If the answer is No, do *not* draw any conclusions about which measurement is more reliable.