STA 431s13 Assignment Five[1]

All these questions are just practice for the quiz on Feb. 15th, and are not to be handed in. The first question is a deliberate repeat from last week. If you did it already, that's great.

1. Here is a multivariate regression model with no intercept and no measurement error. Independently for $i = 1, \ldots, n$,

$$\mathbf{Y}_i = \boldsymbol{\beta}\mathbf{X}_i + \boldsymbol{\epsilon}_i$$

where

> $\mathbf{Y}_i$ is an $q \times 1$ random vector of observable response variables, so the regression can be multivariate; there are q response variables.
>
> $\mathbf{X}_i$ is a $p \times 1$ multivariate normal observable random vector; there are p explanatory variables. $\mathbf{X}_i$ has expected value zero and variance-covariance matrix $\boldsymbol{\Phi}$, a $p \times p$ symmetric and positive definite matrix of unknown constants.
>
> $\boldsymbol{\beta}$ is a $q \times p$ matrix of unknown constants. These are the regression coefficients, with one row for each response variable and one column for each explanatory variable.
>
> $\boldsymbol{\epsilon}_i$ is the error term of the latent regression. It is an $q \times 1$ multivariate normal random vector with expected value zero and variance-covariance matrix $\boldsymbol{\Psi}$, a $q \times q$ symmetric and positive definite matrix of unknown constants. $\boldsymbol{\epsilon}_i$ is independent of $\mathbf{X}_i$.

(a) Calculate the variance-covariance matrix of the observable variables. It's a partitioned matrix. Show your work.

(b) Write down the moment structure equations. These are matrix equations.

(c) Does this problem pass the test of the Parameter Count Rule? Answer Yes or No and give the numbers.

(d) Are the parameters of this model identifiable? Answer Yes or No and prove your answer. If the answer is no, all you need is a simple numerical example of two different parameter vectors that yield the same probability distribution for the sample data.

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

2. Independently for $i = 1, \ldots, n$, let

$$\begin{aligned} Y_i &= \beta X_i + \epsilon_i \\ W_{i,1} &= X_i + e_{i,1} \\ W_{i,2} &= X_i + e_{i,2} \\ V_i &= Y_i + e_{i,3}, \end{aligned}$$

where

- X_i and Y_i are latent variables, while $W_{i,1}$, $W_{i,2}$ and V_i are observable.
- X_i is a normally distributed *latent* variable with mean zero and variance $\phi > 0$
- ϵ_i is normally distributed with mean zero and variance $\psi > 0$
- $e_{i,1}$ is normally distributed with mean zero and variance $\omega_1 > 0$
- $e_{i,2}$ is normally distributed with mean zero and variance $\omega_2 > 0$
- $e_{i,3}$ is normally distributed with mean zero and variance $\omega_3 > 0$
- X_i, ϵ_i, $e_{i,1}$, $e_{i,2}$ and $e_{i,3}$ are all independent of one another.

(a) What is the parameter vector $\boldsymbol{\theta}$ for this model?

(b) Does this problem pass the test of the Parameter Count Rule? Answer Yes or No and give the numbers.

(c) Calculate the variance-covariance matrix of the observable variables. Show your work.

(d) Is the parameter β identifiable? Answer Yes and prove it.

(e) Give a simple numerical example to show that the entire parameter vector is not identifiable.

(f) This model places one equality constraint on the covariance matrix. What is it?

(g) What null hypothesis could you test about the covariances σ_{ij} to challenge the model? What would you conclude if H_0 were rejected? Hint: the null hypothesis contains an $=$ sign.

(h) You can also use the model to deduce more than one testable *inequality* involving the variances and covariances. Give at least two.

3. Let

$$\begin{aligned} W_i &= X_i + e_i \\ Y_{i,1} &= \beta_1 X_i + \epsilon_{i,1} \\ Y_{i,2} &= \beta_2 X_i + \epsilon_{i,2} \end{aligned}$$

where X_i, e_i, $\epsilon_{i,1}$ and $\epsilon_{i,2}$ are all independent, $Var(X_i) = \phi$, $Var(e_i) = \omega$, $Var(\epsilon_{i,1}) = \psi_1$, $Var(\epsilon_{i,2}) = \psi_2$, and all the expected values are zero. The explanatory variable X_i is latent, while W_i, $Y_{i,1}$ and $Y_{i,2}$ are observable

(a) What is the parameter vector $\boldsymbol{\theta}$ for this model?

(b) Does this problem pass the test of the Parameter Count Rule? Answer Yes or No and give the numbers.

(c) Calculate the variance-covariance matrix of the observable variables. Show your work.

(d) The parameter of primary interest is β_1. Is β_1 identifiable at points in the parameter space where $\beta_1 = 0$? Why or why not?

(e) Give a simple numerical example to show that β_1 is not identifiable at points in the parameter space where $\beta_1 \neq 0$ and $\beta_2 = 0$.

(f) Is β_1 identifiable at points in the parameter space where $\beta_2 \neq 0$? Answer Yes or No and prove your answer.

(g) Show that the entire parameter vector is identifiable at points in the parameter space where $\beta_1 \neq 0$ and $\beta_2 \neq 0$.

(h) It is reasonable to assume $\beta_2 \neq 0$, since Y_2 is an instrumental variable and we assume it's well chosen. So at points in the parameter space where $\beta_2 \neq 0$, what *two* equality constraints on the elements of $\boldsymbol{\Sigma}$ are implied by $H_0 : \beta_1 = 0$?

(i) Assuming $\beta_1 \neq 0$ and $\beta_2 \neq 0$, you can use the model to deduce more than one testable *inequality* involving the variances and covariances. Give at least one example.

4. This example shows that sometimes, another explanatory variable can be as useful as an instrumental variable. Independently for $i = 1, \ldots, n$, let

$$\begin{aligned} W_{i,1} &= \nu_1 + X_{i,1} + e_{i,1} \\ Y_{i,1} &= \alpha_1 + \beta_1 X_{i,1} + \epsilon_{i,1} \\ W_{i,2} &= \nu_2 + X_{i,2} + e_{i,2} \\ Y_{i,2} &= \alpha_2 + \beta_2 X_{i,2} + \epsilon_{i,2}, \end{aligned}$$

where $E(X_{i,j}) = \mu_j$, $e_{i,j}$ and $\epsilon_{i,j}$ are independent of one another and of $X_{i,j}$, $Var(e_{i,j}) = \omega_j$, $Var(\epsilon_{i,j}) = \psi_j$, and

$$V \left(\begin{array}{c} X_{i,1} \\ X_{i,1} \end{array} \right) = \left(\begin{array}{cc} \phi_{11} & \phi_{12} \\ \phi_{12} & \phi_{22} \end{array} \right).$$

(a) Calculate the variance-covariance matrix of the observable variables. Show your work.

(b) Show that β_1 and β_2 are identifiable provided $\phi_{12} \neq 0$.

(c) Are there any other points in the parameter space where β_1 is identifiable? β_2?

(d) Give one testable equality constraint (a statement about the σ_{ij} quantities) that is implied by the model. Is it still true with $\phi_{12} = 0$? $\beta_1 = 0$? $\beta_2 = 0$?

(e) Suppose you wanted to estimate β_1. Suggest a *statistic* (function of the sample data) to serve as an estimator.

(f) Is your estimator consistent? Under what circumstances? You don't have to prove anything in detail.

(g) If the primary interest is in β_1, do we really need the response variable $Y_{i,2}$?

5. In this problem, $Y_{i,1}$ is the dependent variable of primary interest, while $Y_{i,2}$ and $Y_{i,3}$ are instrumental variables. The point of the question is that the error terms of instrumental variables need not all be independent.

   Independently for $i = 1, \ldots, n$,

$$
\begin{array}{rcl}
Y_{i,1} & = & \beta_{0,1} + \beta_{1,1} X_i + \epsilon_{i,1} \\
Y_{i,2} & = & \beta_{0,2} + \beta_{1,2} X_i + \epsilon_{i,2} \\
Y_{i,3} & = & \beta_{0,3} + \beta_{1,3} X_i + \epsilon_{i,3} \\
W_i & = & X_i + e_i
\end{array}
$$

   where

   - $X_i \sim N(\mu_x, \phi)$ is a latent variable
   - $e_i \sim N(0, \omega)$
   - $\boldsymbol{\epsilon}_i = (\epsilon_{i,1}, \epsilon_{i,2}, \epsilon_{i,3})'$
   - X_i, e_i and $\boldsymbol{\epsilon}_i$ are independent of one another
   - $\boldsymbol{\epsilon}_i$ is multivariate normal with mean zero and covariance matrix

$$
\boldsymbol{\Psi} = \left[\begin{array}{ccc} \psi_{1,1} & \psi_{1,2} & 0 \\ \psi_{1,2} & \psi_{2,2} & \psi_{2,3} \\ 0 & \psi_{2,3} & \psi_{3,3} \end{array} \right].
$$

   (a) What is the parameter vector $\boldsymbol{\theta}$ for this model?

   (b) How many moment structure equations are there. You do not have to say what they are; just give a number. Don't forget the means.

   (c) Does this problem pass the test of the Counting Rule? Answer Yes or No.

   (d) Calculate the variance-covariance matrix of the observable variables. Remember that some covariances between errors are non-zero. Show your work.

   (e) Solving the complete set of moment structure equations can be done[2] but it's a big chore. The primary interest is in the parameter $\beta_{1,1}$. Show that just this parameter is identifiable.

---

[2]Even the intercepts are identifiable from the mean vector $\boldsymbol{\mu}$, because there is no measurement bias term in this model. That's unrealistic, of course.