STA 431s11 Assignment 5

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 dependent variables, so the regression can be multivariate; there are q dependent variables.

$\mathbf{X}_i$ is a $p \times 1$ multivariate normal observable random vector; there are p independent 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 dependent variable and one column for each independent 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) Are the parameters of this model identifiable? Answer Yes or No and prove your answer.

2. This question shows what can happen when errors of measurement have a non-zero covariance. Independently for $i = 1, \ldots, n$, let

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

where

- 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 $\omega_{1,1} > 0$
- $e_{i,2}$ is normally distributed with mean zero and variance $\omega_{2,2} > 0$
- $Cov(e_{i,1}, e_{i,2}) = \omega_{1,2}$. This is the unusual feature (unusual in statistical models – maybe not so unusual in reality).
- X_i and ϵ_i are independent of one another.
- X_i is independent of $e_{i,1}$ and $e_{i,2}$
- ϵ_i is independent of $e_{i,1}$ and $e_{i,2}$.

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

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

(c) Calculate the variance-covariance matrix of the observable variables. Remember that $Cov(e_{i,1}, e_{i,2}) \neq 0$. Show your work.

(d) There are 6 covariance structure equations in 6 unknowns. Try to solve them. If you can do it, you have proved that the parameter is identifiable, and you are done.

(e) If you cannot solve the covariance structure equations, try to prove that the parameter vector is *not* identifiable. To do this, you need a simple numerical example: two different $\boldsymbol{\theta}$ vectors that produce the same $\boldsymbol{\Sigma}$. To make it easier on yourself, let $\beta = 0$ in both vectors. Be sure to give the covariance matrix (a 3×3 matrix of numbers) that is produced by both sets of parameter values. In your example, make sure $|\boldsymbol{\Omega}| > 0$ (a point that is easy to overlook).

3. Consider the following simple regression through the origin with measurement error in both the independent and dependent variables. Independently for $i = 1, \ldots, n$,

$$
\begin{aligned}
Y_i &= \beta X_i + \epsilon_i \\
W_{i,1} &= X_i + e_{i,1} \\
V_{i,1} &= Y_i + e_{i,2} \\
W_{i,2} &= X_i + e_{i,3} \\
V_{i,2} &= Y_i + e_{i,4}
\end{aligned}
$$

where

- X_i and Y_i are latent variables
- $X_i \sim N(0, \phi)$
- $\epsilon_i \sim N(0, \psi)$
- $\mathbf{e}_i = (e_{i,1}, e_{i,2}, e_{i,3}, e_{i,4})'$
- X_i, ϵ_i and $\mathbf{e}_i$ are independent of one another
- $\mathbf{e}_i$ is multivariate normal with mean zero and covariance matrix

$$
\boldsymbol{\Omega} = \left[\begin{array}{cccc}
\omega_{1,1} & \omega_{1,2} & 0 & 0 \\
\omega_{1,2} & \omega_{2,2} & 0 & 0 \\
0 & 0 & \omega_{3,3} & \omega_{3,4} \\
0 & 0 & \omega_{3,4} & \omega_{4,4}
\end{array} \right].
$$

The pattern of zeros in the covariance matrix of the measurement errors is not arbitrary. It says that $W_{i,1}$ and $V_{i,1}$ form one set of measurements, while $W_{i,2}$ and $V_{i,2}$ form a second set. Measurement errors may be correlated within sets, but not between sets. The two sets of data would be collected at separate times and perhaps by separate methods.

(a) Calculate the variance-covariance matrix of the observable variables. Be careful; the measurement error terms are not all independent, and the expected value of the product is not always the product of expected values; Look at $\boldsymbol{\Omega}$ to tell. Show your work.

(b) Write down the moment structure equations.

(c) Are the parameters of this model identifiable? Answer Yes or No and prove your answer.

4. 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 $i = 1, \ldots, n$,

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

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[1] but it's a big chore. The primary interest is in the parameter $\beta_{1,1}$. Show that just this parameter is identifiable.

[1]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.