ECON 5350 Solutions to the Midterm Exam – Fall 2023

Consider the following regression model without a constant

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i;\ i = 1, ..., n$$

where the error terms are i.i.d., $\bar{x}_1 = \bar{x}_2 = 0$ and $corr(x_{1i}, x_{2i}) = 0$. Use this information to answer the following equally weighted questions.

1. Using summation notation, find the normal equations for the ordinary least squares (OLS) estimator of β_1 and β_2. Check the second-order conditions.

   Answer. The least squares objective function is

$$S(\beta_1, \beta_2) = \sum_{i=1}^{n} (y_i - \beta_1 x_{1i} - \beta_2 x_{2i})^2.$$

   The normal equations are the FOCs for minimization of $S(\beta_1, \beta_2)$:

$$\begin{aligned} \frac{\partial S(\beta_1, \beta_2)}{\partial \beta_1} &= -2\sum_{i=1}^{n} \epsilon_i x_{1i} = 0 \\ \frac{\partial S(\beta_1, \beta_2)}{\partial \beta_2} &= -2\sum_{i=1}^{n} \epsilon_i x_{2i} = 0. \end{aligned}$$

   The second-order conditions simplify to:

$$\begin{aligned} \frac{\partial^2 S(\beta_1, \beta_2)}{\partial \beta_1^2} &= 2\sum_{i=1}^{n} x_{1i}^2 > 0 \\ \frac{\partial^2 S(\beta_1, \beta_2)}{\partial \beta_2^2} &= 2\sum_{i=1}^{n} x_{2i}^2 > 0. \end{aligned}$$

2. Solve the normal equations to find the OLS estimator of β_1 and β_2.

   Answer. The normal equations can be rewritten as

$$\begin{aligned} \sum_{i=1}^{n} (y_i - \beta_1 x_{1i} - \beta_2 x_{2i}) x_{1i} &= 0 \\ \sum_{i=1}^{n} (y_i - \beta_1 x_{1i} - \beta_2 x_{2i}) x_{2i} &= 0. \end{aligned}$$

   Using the fact that $corr(x_{1i}, x_{2i}) = 0$, we can write

$$\begin{aligned} \sum_{i=1}^{n} (y_i - \beta_1 x_{1i}) x_{1i} &= 0 \\ \sum_{i=1}^{n} (y_i - \beta_2 x_{2i}) x_{2i} &= 0. \end{aligned}$$

This implies that the OLS estimators are

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n y_i x_{1i}}{\sum_{i=1}^n x_{1i}^2}$$
$$\hat{\beta}_2 = \frac{\sum_{i=1}^n y_i x_{2i}}{\sum_{i=1}^n x_{2i}^2}.$$

3. Write out the regression model in matrix notation. Clearly define all the matrices with appropriate dimensions.

   Answer. The regression model in matrix form is $Y = X\beta + \epsilon$, where

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_{n\times 1} \qquad X = \begin{bmatrix} x_{11} & x_{21} \\ x_{12} & x_{22} \\ \vdots & \vdots \\ x_{1n} & x_{2n} \end{bmatrix}_{n\times 2} \qquad \text{and } \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}_{n\times 1}.$$

4. Show the equivalence of the matrix-based OLS formula with the results from question #2.

   Answer. The matrix-based OLS formula is $b = (X'X)^{-1}(X'Y)$. Writing this out, we get

$$b = \begin{bmatrix} \sum_{i=1}^n x_{1i}^2 & \sum_{i=1}^n x_{1i}x_{2i} \\ \sum_{i=1}^n x_{2i}x_{1i} & \sum_{i=1}^n x_{2i}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^n y_i x_{1i} \\ \sum_{i=1}^n y_i x_{2i} \end{bmatrix}.$$

   Then using the fact that $corr(x_{1i}, x_{2i}) = 0$, we get

$$b = \begin{bmatrix} \frac{1}{\sum_{i=1}^n x_{1i}^2} & 0 \\ 0 & \frac{1}{\sum_{i=1}^n x_{2i}^2} \end{bmatrix} \begin{bmatrix} \sum_{i=1}^n y_i x_{1i} \\ \sum_{i=1}^n y_i x_{2i} \end{bmatrix},$$

   which shows the equivalence of b and $(\hat{\beta}_1, \hat{\beta}_2)$.

5. Rework question #2 with $corr(x_{1i}, x_{2i}) = 1$ and the normalization $var(x_{1i}) = var(x_{2i}) = 1$. Comment on the results.

   Answer. If x_{1i} and x_{2i} are perfectly correlated, the normal equations become

$$\sum_{i=1}^n (y_i - \beta_1 x_{1i} - \beta_2 x_{2i}) x_{1i} = 0$$
$$\sum_{i=1}^n (y_i - \beta_1 x_{1i} - \beta_2 x_{2i}) x_{2i} = 0.$$

   Using the fact that

$$\frac{1}{n}\sum_{i=1}^n x_{1i}^2 = \frac{1}{n}\sum_{i=1}^n x_{2i}^2 = \frac{1}{n}\sum_{i=1}^n x_{1i}x_{2i} = 1,$$

the normal equations reduce to

$$\begin{aligned}\frac{1}{n}\sum\nolimits_{i=1}^{n} y_i x_{1i} &= \beta_1 + \beta_2 \\ \frac{1}{n}\sum\nolimits_{i=1}^{n} y_i x_{2i} &= \beta_1 + \beta_2.\end{aligned}$$

It is not possible to solve for β_1 and β_2 due to the perfect collinearity between x_{1i} and x_{2i}. In matrix terms, this would imply that the $X'X$ matrix was singular.

6. Find the pdf of each error term, $f_i(\epsilon_i)$, if the cumulative distribution is $F_i(\epsilon_i) = \frac{3}{4}\epsilon_i(1 - \frac{1}{3}\epsilon_i^2) + \frac{1}{2}$ with support $(-1, 1)$.

   Answer. The pdf is found by taking the derivative of $F_i(\epsilon_i)$ with respect to ϵ_i, which gives $f_i(\epsilon_i) = \frac{3}{4}(1 - \epsilon_i^2)$.

7. Find the mean (μ) and variance (σ^2) of the errors. Sketch the pdf of the errors to verify your answer.

   Answer. The mean and variance are $\mu = 0$ and $\sigma^2 = 1/5$.

8. What is the limiting and asymptotic distribution of the sample mean, $\bar{\epsilon}_n$.

   Answer. The limiting distribution of $\bar{\epsilon}_n$ is a spike at $\mu = 0$. Using the central limit theorem, the asymptotic distribution is $\bar{\epsilon}_n \sim N(0, \sigma^2/n)$.

9. Find the joint pdf for the errors, $f(\epsilon_1, ..., \epsilon_n)$.

   Answer. Given independence, the joint pdf of the errors is

$$f(\epsilon_1, ..., \epsilon_n) = \prod\nolimits_{i=1}^{n} f_i(\epsilon_i) = \left(\frac{3}{4}\right)^n (1 - \epsilon_1^2)(1 - \epsilon_2^2)\cdots(1 - \epsilon_n^2).$$

10. Use $f(\epsilon_1, ..., \epsilon_n)$ to find the first-order condition for the maximum (log) likelihood estimate of β_1 and β_2. Is it possible to solve for β_1 and β_2? Comment.

    Answer. Ignoring the scale term, the (log) likelihood function is

$$\ln L(\beta_1, \beta_2) = \sum\nolimits_{i=1}^{n} \ln\left[1 - (y_i - \beta_1 x_{1i} - \beta_2 x_{2i})^2\right]$$

with FOCs

$$\begin{aligned}\frac{\partial \ln L(\beta_1, \beta_2)}{\partial \beta_1} &= 2\sum\nolimits_{i=1}^{n} (1 - \epsilon_i^2)^{-1}\epsilon_i x_{1i} = 0 \\ \frac{\partial \ln L(\beta_1, \beta_2)}{\partial \beta_2} &= 2\sum\nolimits_{i=1}^{n} (1 - \epsilon_i^2)^{-1}\epsilon_i x_{2i} = 0.\end{aligned}$$

These are two nonlinear equations in β_1 and β_2, which will require numerical methods to solve.