ECON 124: Problem Set #3

Due on Jun 5, 2025

Dr. Deniz Baglan

Alejandro Ouslan

Problem 1

The following sample moments for $x = [1, x_1, x_2, x_3]$ were computed from 100 observations produced using a random number generator:

$$X'X \begin{bmatrix} 100 & 123 & 96 & 109 \\ 123 & 252 & 125 & 189 \\ 96 & 125 & 167 & 146 \\ 109 & 189 & 146 & 168 \end{bmatrix} \quad X'y = \begin{bmatrix} 460 \\ 810 \\ 615 \\ 712 \end{bmatrix} \quad y'y = 3924$$

The true model underlying these data is $y = x_1 + x_2 + x_3 + \epsilon$.

1. Compute the simple correlation among the regressors.
2. Compute the ordinary least squares coefficients in the regression of y on a constant, x_1, x_2, and x_3.
3. Compute the ordinary least squares coefficients in the regression of y on a constant x_1 and x_2, on a constant, X_1 and x_3, and on a constant, x_2 and x_3.
4. Compute the variance inflation factor associated with each variable.
5. The regressors are obviously badly collinear, Which is the problem variable? Explain

Problem 2

A multiple regression of y on a constant x_1 and x_2 produces the following resutls:

$$\hat{y} = 4 + 0.4x_1 + 0.9x_2 \quad R^2 = \frac{8}{60} \quad e'e = 520, \quad n = 29,$$

$$X'X = \begin{bmatrix} 29 & 0 & 0 \\ 0 & 50 & 10 \\ 0 & 10 & 80 \end{bmatrix}$$

Test the hypothesis that the two slopes sum to 1

Problem 3

The application in Chapter 3 used 15 of the 19,919 observations in Koop and Tobias's (2004) study of the relationship between wages and education, ability, and family characteristics. (See Appendix Table F3.2.) We will use the full data set for this exercise. The data may be downloaded from the *Journal of Applied Econometrics* data archive at **link**. The data file is in two parts. The fist file contains the panel of 19,919 obseervations on variables:
To create the data set for this exercise, it is necessary to merge these two data files. The ith observations in the first file will be replicated T_i times for the set of T_i observations in the first file. The *person id* variable indicates which rows must contain the data from the second file. (How this preparation is carried out will) vary from one computer package to another.) (*Note:* We are not attempting to replicate the data set.) Let

$$X_1 = [constant, education, experience, ability]$$

$$X_2 = [mother'seducation, father'seducation, brokenhome, numberofsiblings]$$

1. compute the full regression of $(\ln wage \sim X_1)$ and $(\ln wage \sim X_2)$
2. Use the F test to test the hypothesis that all coefficients except the constant term are zero.
3. Use the F statistic to test the joint hypothesis that the coefficient on the four household variables in X_2 are zero
4. Use a Wald test to carry out the test in part c.

Problem 4

In a paper in 1963, Mare Nerlove analyzed a cost function for 145 American electric companies. The attached data file, contains the data and the description file. Nerlove was interested in estimating a cost function: $TC = f(Q, PL, PF, PK)$.

1. First estimate an unrestricted Cobb-Douglas specification

$$\log TC_i = \beta_1 + \beta_2 \log Q_i + \beta_3 \log PL_i + \beta_4 \log PK_i + \beta_5 \log PF_i + \epsilon_i$$

   Report parameter estimates and standard errors.
2. What is the economic meaning of the restriction $H_0 : \beta_3 + \beta_4 + \beta_5 = 1$?
3. Estimate the regression in (a) by constrained least squares $\beta_3 + \beta_4 + \beta_5 = 1$. Report your parameter estimates and standard errors.
4. Test $H_0 : \beta_3 + \beta_4 + \beta_5 = 1$ using a Wald statistic.

Problem 5

Replicate Example 7.12 income elasticity of credit card expenditures in Green's textbook.the data set can be downloaded form the link below: