Economics 361
Problem Set #4

Jun Ishii *
Department of Economics
Amherst College

Fall 2022

Question 1: Ordinary Least Squares (OLS) Model

For a size N sample of the population (X, Y)

$$\{ \ (X_1 = x_1, Y_1 = y_1) \ (X_2 = x_2, Y_2 = y_2) \ \cdots \ (X_N = x_N, Y_N = y_N) \ \}$$

consider the bivariate ordinary least squares (OLS) model: $\hat{Y}(X_i) = a_{ols} + b_{ols} X_i$ where

$$(a_{ols}, b_{ols}) \quad = \quad \text{argmin}_{a,b} \qquad \underbrace{\sum_{i=1}^{N} \ [\ Y_i - \hat{Y}(X_i) \ ]^2}_{\text{Sum of Squared Residuals (SSR)}}$$

(a_{ols}, b_{ols}) are the values of (a, b) that minimizes the sum of squared residuals (SSR). Residuals refer to the difference between the values Y_i and their OLS predicted values $\hat{Y}(X_i)$.

Note that the SSR can be re-written as

$$\sum_{i=1}^{N} \ [\ Y_i - \hat{Y}(X_i) \ ]^2 \quad = \quad \sum_{i=1}^{N} \ [\ Y_i - (a + b X_i) \ ]^2$$

In lecture, I showed that

$$a_{ols} \quad = \quad \frac{1}{N} \ \sum_{i=1}^{N} Y_i - b_{ols} \ \frac{1}{N} \ \sum_{i=1}^{N} X_i \qquad b_{ols} \quad = \quad \frac{\frac{1}{N} \ \sum_{i=1}^{N} (X_i Y_i) - \left(\frac{1}{N} \ \sum_{i=1}^{N} X_i \right) \left(\frac{1}{N} \ \sum_{i=1}^{N} Y_i \right)}{\frac{1}{N} \ \sum_{i=1}^{N} (X_i^2) - \left(\frac{1}{N} \ \sum_{i=1}^{N} X_i \right)^2}$$

I rationalized the OLS estimator as the coefficients of the BLP of Y given X under the Mean Squared Error criterion, with the relevant population moments replaced by their sample analog.

In this question, we explore another moment-based rationalization for the OLS estimator.

*Office: Converse Hall 315 Phone: (413) 542-2901 E-mail: jishii@amherst.edu

In lecture, I asserted that

$$\begin{aligned}\text{Sample Covariance of } X \text{ and } Y &= \frac{1}{N}\sum_{i=1}^{N}(X_iY_i) - \left(\frac{1}{N}\sum_{i=1}^{N}X_i\right)\left(\frac{1}{N}\sum_{i=1}^{N}Y_i\right)\\ \text{Sample Variance of } X &= \frac{1}{N}\sum_{i=1}^{N}(X_i^2) - \left(\frac{1}{N}\sum_{i=1}^{N}X_i\right)^2\end{aligned}$$

Technically,

$$\begin{aligned}\text{Sample Covariance of } X \text{ and } Y &= \frac{1}{N}\sum_{i=1}^{N}(X_i - \bar{X}_N)(Y_i - \bar{Y}_N)\\ \text{Sample Variance of } X &= \frac{1}{N}\sum_{i=1}^{N}(X_i - \bar{X}_N)^2\end{aligned}$$

where $\bar{X}_N = \frac{1}{N}\sum_{i=1}^{N}X_i$ and $\bar{Y}_N = \frac{1}{N}\sum_{i=1}^{N}Y_i$.

(a) Show that

$$\frac{1}{N}\sum_{i=1}^{N}(X_i - \bar{X}_N)(Y_i - \bar{Y}_N) = \frac{1}{N}\sum_{i=1}^{N}(X_iY_i) - \left(\frac{1}{N}\sum_{i=1}^{N}X_i\right)\left(\frac{1}{N}\sum_{i=1}^{N}Y_i\right)$$

and

$$\frac{1}{N}\sum_{i=1}^{N}(X_i - \bar{X}_N)^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i^2) - \left(\frac{1}{N}\sum_{i=1}^{N}X_i\right)^2$$

(b) Let $Z \equiv Y - BLP_{MSE}(Y|X)$. Show that $E[Z] = 0$ and $\text{Cov}(X, Z) = 0$.

(c) Let $\tilde{Z} \equiv Y - \tilde{a} - \tilde{b}X$. Solve for the $(\tilde{a}, \tilde{b})$ that set (i) the sample mean of $\tilde{Z}$ equal to zero and (ii) the sample covariance between X and $\tilde{Z}$ equal to zero. Note that

$$\begin{aligned}\text{Sample Mean of } \tilde{Z} &= \frac{1}{N}\sum_{i=1}^{N}\tilde{Z}_i = \tilde{Z}_N\\ \text{Sample Covariance of } X \text{ and } \tilde{Z} &= \frac{1}{N}\sum_{i=1}^{N}(X_i - \bar{X}_N)(\tilde{Z}_i - \tilde{Z}_N)\\ &\quad \text{where } \tilde{Z}_i = Y_i - \tilde{a} - \tilde{b}X_i\end{aligned}$$

(d) Use the above results to explain the following:

> The OLS estimator is the moment-based estimator that equates the population moments $E[Z] = 0$ and $\text{Cov}(X, Z) = 0$ to their sample analog.

(e) Show that $E[XZ] = 0$

HINT: See **(b)**

(f) Let $\tilde{Z}$ be defined as in **(c)**. This time, solve for the $(\tilde{a}, \tilde{b})$ that set (i) the sample mean of $\tilde{Z}$ equal to zero and (ii) the sample mean of $X\tilde{Z}$ equal to zero. Note that

$$\text{Sample Mean of } X\tilde{Z} \quad = \quad \frac{1}{N}\sum_{i=1}^{N} X_i \tilde{Z}_i$$

HINT: You can answer this problem fairly simply by using **(a)** and **(c)**.

(g) Use the above results to explain the following:

> The OLS estimator is the moment-based estimator that equates the population moments $E[Z] = 0$ and $E[XZ] = 0$ to their sample analog.

Moment conditions of the form $E[XZ] = 0$ are known as **orthogonality conditions**. We will be exploring estimators based on other orthogonality conditions later in the course.

Above, we defined Z as the difference between Y and the $BLP_{MSE}(Y|X)$ and rationalized the OLS estimator using the orthogonality condition between X and this Z. However, OLS is often rationalized using the orthogonality condition between X and the difference between Y and the $BP_{MSE}(Y|X)$. We will not fully develop this particular rationalization here. But we can do the first two steps ... Note: $BP_{MSE}(Y|X)$ is not necessarily the same as $BLP_{MSE}(Y|X)$

(h) Let $W \equiv Y - E[Y|X]$. Show that $E[W] = 0$ and $\text{Cov}(X, W) = 0$.

(i) Show that $E[XW] = 0$

HINT: See **(e)**

Violation of the orthogonality condition $E[XW] = 0$ is often dubbed an "**endogeneity** problem." We will explore this issue later in the course.

Question 2: A Production Example of OLS in Action

Suppose that firms in an industry produced output (Y) using two types of input, capital (K) and labor (L). The relationship between the amount of inputs used and the amount of output produced is given by the following **production function**:

$$Y = AK^{\beta}L^{\gamma} \tag{1}$$

This particular function is known as the **Cobb-Douglas** function, named after Charles Cobb and Paul Douglas, two former Amherst College professors.[1] The Cobb-Douglas function is one of the most widely used and known functions in economics.

The function consists of two key parameters $\{\beta, \gamma\}$ that are of economic interests. In this question, we show how one might use the ordinary least squares (OLS) model to estimate those unknown parameters using only a sample of $\{Y, K, L\}$. In a sense, we are re-tracing the steps of Paul Douglas.

Consider the following linear equation involving Y, K, L

$$ln(Y) = ln(A) + \alpha_1\ ln(K) + \alpha_2\ ln(L) \tag{2}$$

where $ln(\cdot)$ refers to the natural log transformation and $\{\alpha_1, \alpha_2\}$ are parameters

(a) What is the relationship between $\{\alpha_1, \alpha_2\}$ and $\{\beta, \gamma\}$?

Suppose you are given the following *random* sample of size $N > 3$ from a population (Y, A, K, L) determined by the above Cobb-Douglas production function

$$\{\ (Y_1 = y_1, A_1 = a_1, K_1 = k_1, L_1 = l_1) \cdots (Y_N = y_N, A_N = a_N, K_N = k_N, L_N = l_N)\ \}$$

(b) Explain how you *may* be able to **solve** for $\{\alpha_1, \alpha_2\}$. What must be true of the random sample in order for you to be able to solve for $\{\alpha_1, \alpha_2\}$?

Suppose that you did **not** observe $\{A_i\}_{i=1}^{N}$. Each A_i is now an unobserved "productivity shock" and A effectively a random variable. So you have a sample from (Y, K, L)

(c) How does the above affect your ability to **solve** for $\{\alpha_1, \alpha_2\}$ using the random sample?

Suppose you know that $E[ln(A) \mid ln(K), ln(L)] = \mu$, where μ is some parameter; the conditional mean of $ln(A)$ given $(ln(K), ln(L))$ does not vary with $(ln(K), ln(L))$.

(d) What is the $E[ln(Y) \mid ln(K), ln(L)]$ in terms of $\{\mu, \alpha_1, \alpha_2, ln(K), ln(L)\}$?

(e) Using your answers in **(a)-(d)**, explain how you may properly apply the ordinary least squares (OLS) model to your sample and obtain estimates of $\{\mu, \alpha_1, \alpha_2\}$.

[1] For the history, see "The Cobb-Douglas Production Function Once Again: Its History, Its Testing, and Some New Empirical Values," by P.H. Douglas, *Journal of Political Economy*, October 1976, pp.903-915.

Define the following matrices concerning your $N = 10$ sized random sample

$$L = \begin{pmatrix} ln(y_1) \\ \vdots \\ ln(y_{10}) \end{pmatrix} \qquad H = \begin{pmatrix} 1 & ln(k_1) & ln(l_1) \\ \vdots & \vdots & \vdots \\ 1 & ln(k_{10}) & ln(l_{10}) \end{pmatrix}$$

L is a (10×1) column vector and H is a (10×3) matrix.

Your trusty research assistant, Sam, has done the following calculations for you:

$$(H'H) = \begin{pmatrix} 10.00000 & 51.68789 & 31.29016 \\ 51.68789 & 269.3835 & 159.9216 \\ 31.29016 & 159.9216 & 113.8035 \end{pmatrix} \qquad (H'L) = \begin{pmatrix} 33.84597 \\ 174.65849 \\ 117.20837 \end{pmatrix}$$

$$(H'H)^{-1} = \begin{pmatrix} 15.87792 & -2.74415 & -0.50941 \\ -2.74415 & 0.496662 & 0.056572 \\ -0.50941 & 0.056572 & 0.069352 \end{pmatrix}$$

(f) Use the above calculations provided by Sam to calculate the OLS estimate of $\{\mu, \alpha_1, \alpha_2\}$

An important concept in economics concerning production is **economies of scale.** If output doubles when we double the amount of capital and labor, our production function exhibits **constant returns to scale**. If output increases by less than double, it exhibits **decreasing returns to scale**. And if output increases by more than double, it exhibits **increasing returns to scale.**

(g) Based on your answer in **(f)**, do you think the production function that generated your random sample exhibits constant, decreasing, or increasing returns to scale?

You are told that the $\text{Var}(ln(A) \mid ln(K), ln(L)) = \frac{1}{12}$.

(h) What is the $\text{Var}(ln(Y) \mid ln(K), ln(L))$?

(i) Based on the matrices provided above and your answer to part **(h)**, what is the variance of your OLS estimate for α_1? What is the covariance between your OLS estimate for α_1 and your OLS estimate for α_2?

Let $e_i \equiv ln(y_i) - (\hat{\mu} + \hat{\alpha}_1 ln(k_i) + \hat{\alpha}_2 ln(l_i))$ where $\{\hat{\mu}, \hat{\alpha}_1, \hat{\alpha}_2\}$ are the OLS estimates for $\{\mu, \alpha_1, \alpha_2\}$ you obtained in **(f)**. So e_i is the OLS residual for the ith observation.

(j) Show that the sample average of the OLS residuals equals zero: $\frac{1}{10} \sum_{i=1}^{10} e_i = 0$

HINT: $\sum_i e_i = \sum_i ln(y_i) - (\sum_i \hat{\mu} + \hat{\alpha}_1 \sum_i ln(k_i) + \hat{\alpha}_2 \sum_i ln(l_i))$

(k) Briefly explain why the result in **(j)** is not a fluke, not a chance happening for the particular realized sample you were given.

Question 3: Deviating from the Average

Consider a random experiment characterized by two random variables (X, Y)

- You are told that $E[Y|X] = \alpha + \beta X$ and $\text{Var}[Y|X] = \sigma^2$
- You know the value of σ^2 but not (α, β)

Consider two transformed random variables $W \equiv X - E[X]$ and $Z \equiv Y - E[Y]$

(a) Briefly explain why $E[Y|W] = \alpha + \beta X$ and $\text{Var}[Y|W] = \sigma^2$; i.e. why is conditioning on W the same as conditioning on X?

(b) Solve for $E[Z|W]$ and $\text{Var}[Z|W]$.

Now, suppose you are given a size N *random* sample from (X, Y)

$$\{ \ (X_1 = x_1, Y_1 = y_1) \cdots (X_N = x_N, Y_N = y_N) \ \}$$

Considered a transformed version of this random sample:

$$\{ \ (\tilde{X}_1 = \tilde{x}_1, \tilde{Y}_1 = \tilde{y}_1) \cdots (\tilde{X}_N = \tilde{x}_N, \tilde{Y}_N = \tilde{y}_N) \ \}$$

where $\tilde{X}_i = X_i - \bar{X}_N$ and $\tilde{Y}_i = Y_i - \bar{Y}_N$. In other words, variables in each sample observation are subtracted by their sample mean.

(c) Briefly explain why the transformed sample is random if the untransformed sample is random.

Let (a^*, b^*) be the values of (a, b) that minimizes $\sum_{i=1}^{N}(Y_i - a - bX_i)^2$ and $(\tilde{a}, \tilde{b})$ be the values of (a, b) that minimizes $\sum_{i=1}^{N}(\tilde{Y}_i - a - b\tilde{X}_i)^2$

(d) Solve for $(\tilde{a}, \tilde{b})$ in terms of (a^*, b^*).

HINT: What is the sample mean of $\tilde{X}$? Sample mean of $\tilde{Y}$?

(e) Use the analogy (moment) principle to provide intuition for the result obtained in (**d**).

Question 4: A Finance Example of OLS in Action

A key idea in the finance literature is the **Efficient Market Hypothesis (EMH)**. EMH can be expressed in many manners. One colloquial manner is as follows:

> **EMH**: Market prices sufficiently account for all publicly available information

In other words, no investor can "beat" the market using public information.

This hypothesis has been empirically "tested" time and time again, with mixed results. One clever test of the EMH involves the use of data from sports gambling involving "point spreads."

A brief introduction to "point spreads"

> In a sporting event – say American football – the result depends on the difference between the score of the "home" team and the score of the "visiting" team. If the home team scores more than the visiting team, then the home team wins. If the visiting team scores more than the home team, the visitors win.
>
> Often, the match-ups are lop-sided, with one team heavily favored to win over the other. Sports gambling works best when gamblers are asked to choose among equally likely (but mutually exclusive) outcomes. In order to account for one team being heavily favored, gambling markets (such as those run in Las Vegas) require the favored team to win by a certain amount, called a point spread.
>
> So if the home team is favored and the point spread is 3, a gambler betting on the home team wins his gamble only if the home team wins by 3 or more scores. If the home team wins by less than 3 scores (or loses the game), the gambler betting on the visiting team wins the gamble.

Read the following articles

- "Beating the Spread: Testing the Efficiency of Gambling Markets for National Football League Games," by R. Zuber, J. Gandar, & B. Bowers, *Journal of Political Economy*, August 1985, pp.800-806, **zgb85.pdf**
- "Hold Your Bets: Another Look at the Efficiency of the Gambling Markets for National Football League Games," by R. Sauer, V. Brajer, S. Ferris, & M. Marr, *Journal of Political Economy*, February 1988, pp.206-213, **sbfm88.pdf**

Think about how the above two papers relate to the idea of the OLS model as an estimator of the Best Predictor and/or Best Linear Predictor of the actual Point Difference given the Vegas Point Spread and other publicly available information. Do not worry about aspects of the OLS model (such as t-statistics and standard error) we have not yet covered.

There is no need to write anything down for this question – just do the reading and thinking. We will re-visit this paper in lecture and/or quiz.