Econometrics I: Solutions of Homework 5

Hiroki Kato *

May 20, 2020

Contents

1 Solutions

1.1 Question 1

We will show that $E(s^2) = \sigma^2$. The OLS estimator of β is $\hat{\beta} = (X'X)^{-1}X'y$. Substituting $y = X\beta + u$ into $\hat{\beta}$ yields

$$\hat{\beta} = (X'X)^{-1}X'(X\beta + u) = \beta + (X'X)^{-1}X'u.$$

Then, we will obtain

*e-mail: vge008kh@student.econ.osaka-u.ac.jp. Room 503. If you have any errors in handouts and materials, please contact me via e-mail.

$$
\begin{aligned}
y - X\hat{\beta} &= y - X(\beta + (X'X)^{-1}X'u) \\
&= (y - X\beta) + X(X'X)^{-1}X'u \\
&= (I_T - X(X'X)^{-1}X')u. \qquad (1)
\end{aligned}
$$

Let $P \equiv X(X'X)^{-1}X'$. The matrix P is called the *projection matrix*, which maps the vectors of response values (dependent variable) to the vector of fitted values. On the other hand, Define $M \equiv I_T - P$, which maps to vectors of response values to the vector of residual values. The matrix P and M are idempotent and symmetric, that is, $P^2 = P$, $P' = P$, $M^2 = M$ and $M' = M$ (we will review later).

Using equation (1), the estimator of σ^2 is

$$
\begin{aligned}
s^2 &= \frac{1}{T-k}(Mu)'Mu \\
&= \frac{1}{T-k}u'MMu \\
&= \frac{1}{T-k}u'Mu. \qquad (2)
\end{aligned}
$$

$u'Mu$ is scalar because u and M are $T \times 1$ and $T \times T$ matrices. Using properties of trace (see the lecture note), we obtain

$$
\begin{aligned}
u'Mu &= tr(u'Mu) \\
&= tr(Muu') \\
&= tr((I_T - (X'X)^{-1}X'X)uu') \\
&= tr((I_T - I_k)uu'). \qquad (3)
\end{aligned}
$$

Finally, the expectation of s^2 is

$$
\begin{aligned}
E(s^2) &= \frac{1}{T-k}E[tr((I_T - I_k)uu')] \\
&= \frac{1}{T-k}tr((I_T - I_k)E(uu')) \\
&= \frac{1}{T-k}\sigma^2(tr(I_T) - tr(I_k)) \\
&= \frac{1}{T-k}\sigma^2(T-k) \\
&= \sigma^2.
\end{aligned}
$$

1.2 Question 2

From the previous question, $(T-k)s^2$ yields

$$(T-k)s^2 = (y - X\hat{\beta})'(y - X\hat{\beta}) = u'Mu,$$

Since M is symmetric and idempotent, $rank(M)$ is equivalent to the value of trace, which leads to $tr(M) = T-k$. By the assumption that u is normally distributed,

$$\frac{(T-k)s^2}{\sigma^2} = \frac{u'Mu}{\sigma^2} \sim \chi^2(T-k) \tag{4}$$

1.3 Question 3

To show that OLS estimator is BLUE (i.e. best linear unbiased estimator), we need to prove that other linear unbiased estimators have larger variances than the OLS estimator, that is, $V(\tilde{\beta}) - V(\hat{\beta}) \geq 0$ where $\tilde{\beta}$ is other linear unbiased estimator.

The first step is to construct a linear unbiased estimator, $\tilde{\beta}$. Since a linear estimator is a function of dependent variable, y, define $\tilde{\beta} = Cy$ where C is a $k \times T$ matrix. Then, the expectation of $\tilde{\beta}$ is

$$E(\tilde{\beta}) = E(C(X\beta + u)) = CX\beta.$$

If $\tilde{\beta}$ is an unbiased estimator, it must hold that

$$CX = I_k, \tag{5}$$

where I_k is $k \times k$ identity matrix.

The second step is to derive the variance-covariance matrix of $\tilde{\beta}$, $V(\tilde{\beta})$. As in the lecture note, you can assume $C = D + (X'X)^{-1}X'$ without loss of generality, and calculate its variance-covariance matrix. In this material, we derive the variance-covariance matrix without assuming the matrix form of C. Assuming $CX = I_k$, we derive the variance-covariance matrix of $\tilde{\beta}$ as follows:

$$E[(\tilde{\beta} - \beta)(\tilde{\beta} - \beta)'] = E[Cu(Cu)'] = E[Cuu'C'] = CE(uu')C' = \sigma^2 CC'.$$

The projection matrix P under OLS estimator is $P = X(X'X)^{-1}X'$, which is a $T \times T$ matrix. Moreover, the matrix M that makes the vector of residuals is $M = I - P$. Thus, $P + M = I_T$. Inserting $P + M$ into the variance-covariance matrix of $\tilde{\beta}$ yields

$$\begin{aligned}
V(\tilde{\beta}) &= \sigma^2 C I_T C' \\
&= \sigma^2 C(P + M)C' \\
&= \sigma^2 [CPC' + CMC'] \\
&= \sigma^2 [CX(X'X)^{-1}X'C + CMC'] \\
&= \sigma^2 [I_k(X'X)^{-1}I_k + CMC'] \\
&= \sigma^2 (X'X)^{-1} + \sigma^2 CMC'.
\end{aligned}$$

Since the variance-covariance matrix of $\hat{\beta}$, OLS estimator, is $\hat{\beta} = \sigma^2(X'X)^{-1}$, we obtain

$$V(\tilde{\beta}) - V(\hat{\beta}) = \sigma^2 CMC'.$$

Because M is idempotent, M is positive-semidefinite. Since M is symmetric and positive-semidefinite, CMC' is also symmetric and positive-semidefinite [1] . Thus, $V(\tilde{\beta}) \geq V(\hat{\beta})$ holds.

[1]Let A be $m \times n$ matrix. $A'MA$ is symmetric and positive-semidefinite if M is $m \times m$ symmetric and positive semidefinite. The proof is straightforward. Define b as any $n \times 1$ vector. Then, $b'A'MAb = c'Mc$ where $c = Ab$ is larger than or equal to zero. By the defenition of positive-semidefinite matrix, $c'Mc \geq 0$. Hence, $b(A'MA)b \geq 0$, that is, $A'MA$ is positive-semidefinite

2 Review

2.1 Projection Matrix

Using the same notations as above, consider the regression model, $y = X\beta + u$. The OLS estimator of β is given by $\hat{\beta} = (X'X)^{-1}X'y$. Then, the fitted value of y is

$$\hat{y} = X\hat{\beta} = X(X'X)^{-1}X'y = P_X y$$

where $P_X \equiv X(X'X)^{-1}X'$. The matrix P is called the *projection matrix*. This matrix maps a vector of response values to a vector of its fitted values. Using the projection matrix, we can express residuals as follows:

$$y - \hat{y} = (I_T - P_X)y = M_X y$$

where $M_X = I_T - P_X = I_T - X(X'X)^{-1}X'$, and I_T is a $T \times T$ identity matrix. The matrix M maps a vector of response values to a vector of residual values. These two operators have the following properties:

1. P_X and M_X are idempotent and symmetric;
2. $P_X X = X$ and $M_X X = 0$;
3. $P_X M_X = M_X P_X = 0$

Proof of Statement 1: First, we will prove the statement that P_X and M_X are symmetric. About the projection matrix, P_X,

$$\begin{aligned} P_X' = (X(X'X)^{-1}X')' &= ((X'X)^{-1}X')'X' \\ &= X((X'X)^{-1})'X' \\ &= X((X'X)')^{-1}X' \\ &= X(X'X)^{-1}X' = P_X. \end{aligned}$$

Thus, we prove that $P_X' = P_X$. Using this, we derive $M_X' = M_X$ because

$$M'_X = (I_T - P_X)' = I_T - P'_X = I_T - P_X = M_X.$$

Second, we will prove the statement that P_X and M_X are idempotent. The matrix A is idempotent if and only if $A^n = A$ for $n \in \mathbb{Z}_{++}$. Note that $\mathbb{Z}_{++}$ is a set of strictly positive integers. Consider the projection matrix P_X. For the sufficiency for an idempotent matrix, prove the case of $n = 2$. Then,

$$P_X P_X = X(X'X)^{-1}X'X(X'X)^{-1}X' = X(X'X)^{-1}(X'X)(X'X)^{-1}X' = X(X'X)^{-1}X' = P_X.$$

Thus, we conclude sufficiency for an idempotent matrix. Next, prove the necessity for an idempotent matrix with mathematical induction. First, consider the case of $n = 1$. It is clear that the statement is true. Suppose that the statement is true for some $n \geq 2$. Clearly,

$$P_X^{n+1} = P_X^n P_X = P_X P_X = X(X'X)^{-1}X' = P_X.$$

Thus, the statement holds for any n. Note that you can prove that M_X is idempotent using the property that P_X is idempotent. (proof is omitted, but the procedure is same).

Proof of Statement 2: Clearly,

$$P_X X = (X(X'X)^{-1}X')X = X,$$
$$M_X X = (I_T - P_X)X = X - X = 0.$$

Proof of Statement 3: Clearly,

$$P_X M_X = P_X(I_T - P_X) = P_X - P_X = 0,$$
$$M_X P_X = (I_T - P_X)P_X = P_X - P_X = 0.$$

2.2 Property of Idempotent Matrix

Let **A** be a $N \times N$ idempotent matrix. An idempotent matrix has the following useful properties:

1. Eigenvalue of idempotent matrix $\mathbf{A}$ is 0 or 1.

2. An idempotent matrix $\mathbf{A}$ is positive-semidefinite.

3. $rank(\mathbf{A}) = tr(\mathbf{A})$

4. If an idempotent matrix $\mathbf{A}$ is symmetric, then $\mathbf{u}'\mathbf{A}\mathbf{u} \sim \chi^2(r)$ where $rank(\mathbf{A}) = r$ and $\mathbf{u} \sim N(0, \mathbf{I}_N)$.

Proof of Statement 1: Eigenvalues λ are defined by $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ where $\mathbf{x} \neq \mathbf{0}$ is a corresponding eigenvector. The definition of idempotent matrix yields

$$\begin{aligned}
\mathbf{A}\mathbf{x} &= \lambda\mathbf{x} \\
\mathbf{A}\mathbf{A}\mathbf{x} &= \lambda\mathbf{x} \\
\mathbf{A}(\lambda\mathbf{x}) &= \lambda\mathbf{x} \\
\lambda(\mathbf{A}\mathbf{x}) &= \lambda\mathbf{x} \\
\lambda^2\mathbf{x} &= \lambda\mathbf{x}
\end{aligned}$$

Therefore, we obtain $\lambda(\lambda - 1)\mathbf{x} = 0$. By $\mathbf{x} \neq \mathbf{0}$, we have $\lambda = 0, 1$.

Proof of Statement 2: The statement that $\mathbf{A}$ is positive-semidefinite is equivalent to the statement that all eigenvalues are non-negative. By statement 1, $\mathbf{A}$ is positive-semidefinite.

Proof of Statement 3: Suppose that the rank of $\mathbf{A}$ is r. There exists a $N \times r$ matrix $\mathbf{B}$ and a $r \times N$ matrix $\mathbf{L}$, each of rank R, such that $\mathbf{A} = \mathbf{B}\mathbf{L}$ [2] . Then,

$$\mathbf{B}\mathbf{L}\mathbf{B}\mathbf{L} = \mathbf{A}^2 = \mathbf{A} = \mathbf{B}\mathbf{L} = \mathbf{B}\mathbf{I}_r\mathbf{L},$$

where $\mathbf{I}_r$ is a $r \times r$ identity matrix. Thus, we obtain $\mathbf{L}\mathbf{B} = \mathbf{I}_r$. By the property of trace,

$$tr(\mathbf{A}) = tr(\mathbf{B}\mathbf{L}) = tr(\mathbf{L}\mathbf{B}) = tr(\mathbf{I}_r) = r = rank(\mathbf{A}).$$

Proof of Statement 4: By symmetric matrix, there exists an orthogonal matrix $\mathbf{C}$ such that $\mathbf{A} =$

[2]This decomposition is known as *rank factorization* (階数因数分解).

$\mathbf{C}\Lambda\mathbf{C}'$ where Λ is a diagonal matrix whose elements are eigenvalues λ_i, that is,

$$\Lambda = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \lambda_i & \vdots \\ 0 & \cdots & \lambda_N \end{pmatrix} = diag(\lambda_1, \cdots, \lambda_N).$$

By the statement 3,

$$rank(\mathbf{A}) = rank(\mathbf{C}\Lambda\mathbf{C}') = rank(\Lambda) = r, \tag{6}$$

$$rank(\mathbf{A}) = tr(\mathbf{A}) = tr(\mathbf{C}\Lambda\mathbf{C}') = tr(\Lambda\mathbf{C}'\mathbf{C}) = tr(\Lambda) = r. \tag{7}$$

For the equation (6), the third equality holds because $rank(EG) = rank(GE) = rank(G)$ where E is full-rank matrix, and an orthogonal matrix is full-rank. For the equation (7), the forth equality comes from the defenition of orthogonality, $\mathbf{C}'\mathbf{C} = \mathbf{I}_N$. By this result and the statement 1, without loss of generality, we can define $\lambda_i = 1$ for $i = 1, \ldots, r$, and $\lambda_i = 0$ for $i = r+1, \ldots, N$.

Next, let $\mathbf{z} = \mathbf{C}'\mathbf{u}$. Then, $E[\mathbf{z}] = 0$ and $E[\mathbf{z}\mathbf{z}'] = \mathbf{C}'\mathbf{I}_N\mathbf{C} = \mathbf{I}_N$ by the defenition of orthogonality, $\mathbf{C}'\mathbf{C} = \mathbf{I}_N$. This implies that $\mathbf{z} \sim N(0, \mathbf{I}_N)$.

Finally, we obtain

$$\mathbf{u}'\mathbf{A}\mathbf{u} = \mathbf{u}'\mathbf{C}\Lambda\mathbf{C}'\mathbf{u} = \mathbf{z}'\Lambda\mathbf{z} = \sum_{i=1}^{r} z_i^2,$$

where $\Lambda = diag(1, \ldots 1, 0, \ldots 0)$. By the defenition of chi-squared distribution, $\mathbf{u}'\mathbf{A}\mathbf{u} \sim \chi^2(r)$.