Solutions of Homework 1

Wang Xin, Zheng Xuzhu

June 14, 2022

1 Question 1

1.1 Derive $\hat{\beta}$

From the regression model, we have

$$u = y - X\beta$$

Let J be the sum of residuals, then

$$J = u'u = (y - X\beta)'(y - X\beta) = y'y + \beta'X'X\beta - y'X\beta - \beta'X'y$$

Since $y'X\beta = (\beta'X'y)'$ and $\beta'X'y$ is a scalar,

$$J = y'y + \beta'X'X\beta - 2\beta'X'y. \tag{1}$$

The extreme value of J can be calculated by taking the first order condition(FOC).

$$\frac{\partial J}{\partial \hat{\beta}} = 2X'X\beta - 2X'y \tag{2}$$

$$\frac{\partial^2 J}{\partial^2 \beta} = 2X'X$$

The second order condition(SOC) of equation (1) is greater than 0, therefore J has the minimum value when equation (2) equals to 0.

Here $\hat{\beta}$ is the OLS estimator,

$$\frac{\partial J}{\partial \beta} = 2X'X\hat{\beta} - 2X'y = 0$$

$$\hat{\beta} = (X'X)^{-1}X'y \tag{3}$$

1.2 Derive mean and variance of $\hat{\beta}$.

From equation (3), we have

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

As set in question, $u \overset{\text{iid}}{\sim} N(0, \sigma^2)$, then we can derive the conditional expectation and variance for both sides:

$$E[\hat{\beta}|X] = E[\beta + (X'X)^{-1}X'u|X] = \beta + (X'X)^{-1}X'E[u|X] = \beta \tag{4}$$

$$\begin{aligned}V[\hat{\beta}|X] &= V[\beta + (X'X)^{-1}X'u|X] \\ &= (X'X)^{-1}X'V[u|X]((X'X)^{-1}X')' \\ &= \sigma^2(X'X)^{-1}X'X(X'X)^{-1} \\ &= \sigma^2(X'X)^{-1}\end{aligned} \tag{5}$$

1.3 Derive a distribution of $\hat{\beta}$, using the moment-generating function.

The moment generating function of $\widehat{\beta}$ is

$$\begin{aligned}M_{\widehat{\beta}}(\theta) &= \mathrm{E}(\exp\{\theta'(\beta + (X'X)^{-1}X'u)\}) \\ &= \exp(\theta'\beta)\mathrm{E}(\exp(\theta'(X'X)^{-1}X'u)).\end{aligned}$$

Since the moment generating function of $u \sim N(0, \sigma^2 I_n)$ is

$$M_u(\theta) = \mathrm{E}(\exp(\theta'u)) = \exp\{\frac{\sigma^2\theta'\theta}{2}\},$$

we can rewrite

$$\begin{aligned}M_{\widehat{\beta}}(\theta) &= \exp(\theta'\beta)M_u(\theta'(X'X)^{-1}X') \\ &= \exp(\theta'\beta)\exp(\frac{\sigma^2}{2}\theta'(X'X)^{-1}\theta) \\ &= \exp(\theta'\beta + \frac{\sigma^2}{2}\theta'(X'X)^{-1}\theta),\end{aligned}$$

which indicates that $\widehat{\beta} \sim N(\beta, \sigma^2(X'X)^{-1})$.

1.4 Show that $s^2 = \dfrac{1}{n-k}(y - X\hat{\beta})'(y - X\hat{\beta})'$ is an unbiased estimator of σ^2.

Since $\widehat{\beta} = \beta + (X'X)^{-1}X'u$, then

$$
\begin{aligned}
y - X\widehat{\beta} &= y - X(\beta + (X'X)^{-1}X'u) \\
&= (y - X\beta) - X(X'X)^{-1}X'u \\
&= \underbrace{(I_n - X(X'X)^{-1}X')}_{\text{idempotent and symmetric}} u
\end{aligned}
$$

Let $M \equiv I_n - X(X'X)^{-1}X'$, then $M^2 = M, M' = M$. Thus, s^2 can be rewritten as

$$
s^2 = \frac{1}{n-k}(Mu)'(Mu) = \frac{1}{n-k}u'MMu = \frac{1}{n-k}\underbrace{u'Mu}_{\text{scalar}}.
$$

Since for scalar $u'Mu$, $\text{tr}(u'Mu) = u'Mu$,

$$
\begin{aligned}
\text{E}(s^2) &= \frac{1}{n-k}\text{E}[\text{tr}(Muu')] \\
&= \frac{1}{n-k}\text{tr}(M\text{E}(uu')) \\
&= \frac{1}{n-k}\sigma^2(\text{tr}(I_n) - \text{tr}((X'X)^{-1}X'X)) \\
&= \frac{1}{n-k}(\text{tr}(I_n) - \text{tr}(I_k)) \\
&= \frac{1}{n-k}\sigma^2(n-k) \\
&= \sigma^2.
\end{aligned}
$$

1.5 Show that $\frac{(n-k)s^2}{\sigma^2}$ is distributed as a χ^2 random variavble with n - k degrees of freedom.

Since $s^2 = \frac{1}{n-k}u'Mu$, under the assumption that $u \sim N(0, \sigma^2 I_n)$, the distribution of s^2 denotes

$$
\frac{(n-k)s^2}{\sigma^2} = \frac{u'Mu}{\sigma^2} \sim \chi^2(\text{Rank}(M)).
$$

Note that for the idempotent and symmetric matrix M, $\text{Rank}(M) = \text{tr}(M) = n - k$.

1.6 Show that $\hat{\beta}$ is a best linear unbiased estimator.

Consider the alternative linear unbiased estimator $\widetilde{\beta}$ as follows:

$$
\widetilde{\beta} = \underbrace{C}_{k \times n} y = C(X\beta + u) = CX\beta + Cu.
$$

Then

$$
\text{E}(\widetilde{\beta}) = CX\beta + C\text{E}(u) = CX\beta
$$

Since $\widetilde{\beta}$ is assumed to be unbiased, $\mathrm{E}(\widetilde{\beta}) = \beta$ holds under the condition:

$$CX = I_k.$$

Then we derive that $V(\widetilde{\beta}) = \mathrm{E}(\widetilde{\beta}-\beta)(\widetilde{\beta}-\beta)' = \mathrm{E}[Cu(Cu)'] = \mathrm{E}(Cuu'C') = C\mathrm{E}(uu')C = \sigma^2 CC'$. Defining $C = D + (X'X)^{-1}X'$, $V(\widetilde{\beta})$ can be rewritten as

$$V(\widetilde{\beta}) = \sigma^2 (D + (X'X)^{-1}X')(D + (X'X)^{-1}X')'.$$

In addition, because of the unbiasedness of $\widetilde{\beta}$, $CX = (D+(X'X)^{-1}X')X = DX + I_k = I_k$ indicates that $DX = 0$. Thus,

$$V(\widetilde{\beta}) = \sigma^2 (X'X)^{-1} + \sigma^2 DD' = V(\widehat{\beta}) + \sigma^2 DD'.$$

$V(\widetilde{\beta}) - V(\widehat{\beta}) = \sigma^2 DD'$ is positive semidefinite matrix. $V(\widetilde{\beta}) \geq V(\widehat{\beta})$ holds.

1.7 Show that $\dfrac{(\hat{\beta}-\beta)'X'X(\hat{\beta}-\beta)}{\sigma^2} \sim \chi^2(k)$.

$$\begin{aligned}(\widehat{\beta}-\beta)'X'X(\widehat{\beta}-\beta) &= ((X'X)^{-1}X'u)'X'X(X'X)^{-1}X'u \\ &= u'X(X'X)^{-1}X'X(X'X)^{-1}X'u = u' \underbrace{X(X'X)^{-1}X'}_{\text{idempotent and symmetric}} u\end{aligned}$$

with $u \sim N(0, \sigma^2 I_n)$,

$$\frac{u'X(X'X)^{-1}X'u}{\sigma^2} \sim \chi^2(\mathrm{tr}(X(X'X)^{-1}X')).$$

Note that $\mathrm{tr}(X(X'X)^{-1}X') = \mathrm{tr}((X'X)^{-1}X'X) = \mathrm{tr}(I_k) = k$. Thus the degree of freedom is k.

1.8 Show that $\hat{\beta}$ is independent of $s^2 = \dfrac{1}{n-k}(y - X\hat{\beta})'(y - X\hat{\beta})$.

$\mathrm{Cov}(X, Y) = 0 \iff \mathrm{Cov}(g(X), h(Y)) = 0$.

$$s^2 = \frac{1}{n-k}e'e,$$

We prove the independence of $\widehat{\beta}$ to s^2 by deriving $\mathrm{Cov}(\widehat{\beta}, e) = 0$ because both $\widehat{\beta}$ and e are normal.

$$\begin{aligned}\mathrm{Cov}(\widehat{\beta}, e) &= \mathrm{E}(e(\widehat{\beta}-\beta)') = \mathrm{E}(Mu((X'X)^{-1}X'u)') \\ &= \mathrm{E}(Muu'X(X'X)^{-1}) = M\mathrm{E}(uu')X(X'X)^{-1} \\ &= M(\sigma^2 I_n)X(X'X)^{-1} = \sigma^2 MX(X'X)^{-1} \\ &= \sigma^2(X(X'X)^{-1} - X(X'X)^{-1}X'X(X'X)^{-1} \\ &= \sigma^2(X(X'X)^{-1} - X(X'X)^{-1}) \\ &= 0\end{aligned}$$

1.9 Show that $\frac{(\hat{\beta}-\beta)'X'X(\hat{\beta}-\beta)/k}{(y-X\hat{\beta})'(y-X\hat{\beta})/(n-k)} \sim F(k, n-k)$.

We have proven that

$$\frac{(\widehat{\beta}-\beta)'X'X(\widehat{\beta}-\beta)}{\sigma^2} = \frac{u'X(X'X)^{-1}X'u}{\sigma^2} \sim \chi^2(k)$$

in question 1.7 and

$$\frac{(y-X\widehat{\beta})'(y-X\widehat{\beta})}{\sigma^2} \sim \chi^2(n-k)$$

in question 1.4 and 1.5.

Since we proved that $\widehat{\beta}$ is independent of e, we have

$$\frac{\frac{(\widehat{\beta}-\beta)'X'X(\widehat{\beta}-\beta)}{\sigma^2}/k}{\frac{(y-X\widehat{\beta})'(y-X\widehat{\beta})}{\sigma^2}/(n-k)} \sim \mathrm{F}(k, n-k).$$

1.10 Show that $\sum_i (y_i - \bar{y})^2 = y'(I_n - ii')y$, where $y = (y_1, y_2, ..., y_n)'$ and $i = (1, 1, ..., 1)'$.

First, we can simplify the equation as

$$\sum_i (y_i-\bar{y})^2 = \sum_i y_i(y_i-\bar{y}) - \bar{y}\sum_i (y_i-\bar{y}) = \sum_i y_i(y_i-\bar{y}) = \sum_i y_i^2 - \bar{y}\sum_i y_i$$

Then each item on the right side can be rewrite as

$$\sum_i y_i^2 = y'y$$

$$\bar{y} = \frac{1}{n}i'y = \frac{1}{n}y'i$$

$$\sum_i y_i = i'y$$

Therefore, the above equation is

$$\sum_i (y_i-\bar{y})^2 = y'y - \frac{1}{n}y'ii'y = y'(I_n - \frac{1}{n}ii')y$$

1.11 Show that $I_n - \frac{1}{n}ii'$ is symmetric and idempotent.

From the definition of symmetric and idempotent matrix, first we can prove that

$$(I_n - \frac{1}{n}ii')(I_n - \frac{1}{n}ii') = I_nI_n + \frac{1}{n^2}ii'ii' - 2\frac{1}{n}I_nii' = I_n - \frac{1}{n}ii',$$

which means that $I_n - \frac{1}{n}ii'$ is idempotent.

Moreover, ii' and I_n are $n \times n$ symmetric matrices,

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \qquad i'i = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}$$

which can prove that $I_n - \frac{1}{n}ii'$ are symmetric.

$$I_n - \frac{1}{n}ii' = \begin{pmatrix} 1-\frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ -\frac{1}{n} & 1-\frac{1}{n} & \cdots & -\frac{1}{n} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{n} & -\frac{1}{n} & \cdots & 1-\frac{1}{n} \end{pmatrix}$$

2 Question 2

2.1 What is the distribution of $\frac{(X-\mu i)'(X-\mu i)}{\sigma^2}$?

Rewrite the equation into scalars, we can obtain

$$\frac{(X-\mu i)'(X-\mu i)}{\sigma^2} = \sum_{j=1}^{n}(\frac{X_j-\mu}{\sigma})^2$$

Since $X_i \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$, $\frac{X_i-\mu}{\sigma} \overset{\text{iid}}{\sim} N(0,1)$ can be easily obtained.

From the definition of chi-squared distribution, we know that Z^2 is chi-squared distributed with 1 degree of freedom if Z is of standard normal distribution.

Then, from the additivity of chi-squared distribution, the n random variables indicate that

$$\sum_{j=1}^{n}(\frac{X_j-\mu}{\sigma})^2 \overset{\text{iid}}{\sim} \chi^2(n)$$

In conclusion, $\frac{(X-\mu i)'(X-\mu i)}{\sigma^2}$ are of chi-squared distribution with n degree of freedom.

2.2 Show that $(X-\mu i)'(I_n - \frac{1}{n}ii')(X-\mu i) = \sum_{j=1}^{n}(X_j - \bar{X})$

For the right side, which is similar to 1.10, the below equation can be easily derived:

$$\sum_{j=1}^{n}(X_j - \bar{X}) = X'(I_n - \frac{1}{n}ii')X \tag{6}$$

As for the left side, it can be expand as

$$\begin{aligned} &(X-\mu i)'(I_n - \frac{1}{n}ii')(X-\mu i) \\ &= X'(I_n - \frac{1}{n}ii')(X-\mu i) - \mu i'(I_n - \frac{1}{n}ii')(X-\mu i) \end{aligned} \tag{7}$$

The second term can be simplified as

$$\mu i'(I_n - \frac{1}{n}ii')(X-\mu i) = (\mu i' - \frac{1}{n}\mu i'ii')(X-\mu i) = (\mu i' - \mu i')(X-\mu i) = 0$$

So equation (7) continued to be expanded as

$$\begin{aligned} &(X-\mu i)'(I_n - \frac{1}{n}ii')(X-\mu i) \\ &= X'(I_n - \frac{1}{n}ii')(X-\mu i) \\ &= X'(I_n - \frac{1}{n}ii')X - X'(I_n - \frac{1}{n}ii')\mu i \\ &= X'(I_n - \frac{1}{n}ii')X \end{aligned} \tag{8}$$

From equation (6) and equation (8), we can prove that

$$(X-\mu i)'(I_n - \frac{1}{n}ii')(X-\mu i) = \sum_{j=1}^{n}(X_j - \bar{X})$$

2.3 Show that $\frac{(X-\mu i)'(I_n - \frac{1}{n}ii')(X-\mu i)}{\sigma^2} \sim \chi^2(n-1)$.

Similar to 1.5 and 2.1, the distribution is derived as

$$\frac{(X-\mu i)'(I_n - \frac{1}{n}ii')(X-\mu i)}{\sigma^2} \sim \chi^2(tr(I_n - \frac{1}{n}ii'))$$

The properties of trace indicates that

$$tr(I_n - \frac{1}{n}ii') = tr(I_n) - tr(\frac{1}{n}ii') = n - tr(\frac{1}{n}i'i) = n-1$$

Consequently, we can prove that

$$\frac{(X-\mu i)'(I_n - \frac{1}{n}ii')(X-\mu i)}{\sigma^2} \sim \chi^2(n-1)$$