Answer key for Q1 - Homework 2

Simon Jean

November 2023

Q1

Write a simple linear model in matrix form. Express the OLS estimates using matrices

Let $\mathbf{Y}$ be the $N \times 1$ vector of outcome variables such that :

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ ... \\ Y_N \end{pmatrix}$$

Let X be the $N \times 2$ matrix of observables, including a constant so $\mathbf{X} = [\mathbf{1}, \tilde{\mathbf{X}}]$ (where $\tilde{\mathbf{X}}$ is a $N \times 1$ vector). Eventually, let ϵ be a $N \times 1$ random error vector.

A matrix model for the linear regression can be written using β as 2×1 coefficient vector :

$$\mathbf{Y} = \mathbf{X}\beta + \epsilon \tag{1}$$

$$\begin{pmatrix} Y_1 \\ Y_2 \\ ... \\ Y_N \end{pmatrix} = \begin{pmatrix} 1 & X_1 \\ 1 & X_2 \\ ... & \\ 1 & X_N \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ ... \\ \epsilon_N \end{pmatrix} \tag{2}$$

Using this formulation, the minimization of the squared residuals is :

$$\begin{aligned} & \min_{\beta_0,\beta_1} (\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta) \\ \iff & \min_{\beta_0,\beta_1} (\mathbf{Y}' - \beta'\mathbf{X}')(\mathbf{Y} - \mathbf{X}\beta) \\ \iff & \min_{\beta_0,\beta_1} (\mathbf{Y}\mathbf{Y}' - \beta'\mathbf{X}'\mathbf{Y} - \mathbf{Y}'\mathbf{X}\beta - \beta'\mathbf{X}\mathbf{X}'\beta) \end{aligned}$$

Using the fact that $(A - B)' = A' - B'$ and $(AB)' = B'A'$ to go from line 1 to 2. Now notice that $\mathbf{Y}'\mathbf{X}\beta$ is $(1 \times N) \times (N \times 2) \times (2 \times 1)$ i.e, it's of dimension 1. Therefore, $\mathbf{Y}'\mathbf{X}\beta = (\mathbf{Y}'\mathbf{X}\beta)' = \beta'\mathbf{X}'\mathbf{Y}$. In the end, the objective function is :

$$\min_{\beta_0,\beta_1} (\mathbf{Y}\mathbf{Y}' - 2\beta'\mathbf{X}'\mathbf{Y} - \beta'\mathbf{X}\mathbf{X}'\beta) \tag{3}$$

You did not have to do that, but for illustration purposes using the rules of matrix differentiation i.e $\frac{\partial(b'a)}{\partial b} = a$ and $\frac{\partial b'Ub}{\partial b} = 2Ub$, we can differentiate the objective function with respect to β, and have the (necessary and sufficient, as we minimize a convex function, we have a concave problem) first order conditions:

$$\frac{\partial}{\partial \beta}(\mathbf{YY}' - 2\beta'\mathbf{X}'\mathbf{Y} - \beta'\mathbf{XX}'\beta) = 0$$
$$\iff 0 - 2\mathbf{X}'\mathbf{Y} - 2\mathbf{XX}'\beta = 0$$

$$\Rightarrow \beta = (\mathbf{XX}')^{-1}\mathbf{X}'\mathbf{Y} \tag{4}$$

Express the variance of the OLS estimates using matrices

Here, we assume that $\mathbf{X}$ is non random, and we are just interested in finding the sample variance, under our assumptions. Note that in the more general case, where $\mathbf{X}$ is also random (i.e, before sampling the data, we know we will not get the whole population, so we still need to account for randomness in $\mathbf{X}$), you would compute the variance of the estimator conditional on X, i.e, knowing what you know about $\mathbf{X}$[1]

So we want to know $\mathbf{V}(\beta)$:

$$\begin{aligned}
\mathbf{V}(\beta) &= \mathbf{V}(\mathbf{XX}')^{-1}\mathbf{X}'\mathbf{Y}) \\
&= \mathbf{V}((\mathbf{XX}')^{-1}\mathbf{X}'(\mathbf{X}\beta + \epsilon)) \\
&= \mathbf{V}(\beta + (\mathbf{XX}')^{-1}\mathbf{X}'\epsilon) \\
&= \mathbf{V}((\mathbf{XX}')^{-1}\mathbf{X}'\epsilon) \\
&= (\mathbf{XX}')^{-1}\mathbf{X}'\mathbf{V}(\epsilon)((\mathbf{XX}')^{-1}\mathbf{X}'))'
\end{aligned}$$

$$\Rightarrow \mathbf{V}(\beta) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1} \tag{5}$$

Where we use the linear model assumption to go from 1 to 2 (i.e, that $\mathbf{Y}$ and $\mathbf{X}$ are linearly related). To go from 3 to 4, use the fact that β here is not a random variable, it's a known value, and the variance of a known value is null (e.g, the variance of a random variable that is always equal to the same value is 0, it never changes). To go from 4 to 5, use $\mathbf{V}(\mathbf{AX}) = \mathbf{AV}(\mathbf{X})\mathbf{A}'$, the *homoskedasticity assumption* i.e $\mathbf{V}(\epsilon) = \sigma^2$ and finally reduce using the properties of the transpose operator to obtain equation 5

Use the above formula to express the standard erros of both the intecept and the slope estimators in terms of $\bar{x}, \sigma, S_{xx}$

Let $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} X_i$, and $S_{xx} = \sum_{i=1}^{N}(X_i - \bar{X})^2$. We will now focus on $(\mathbf{X}'\mathbf{X})^{-1}$.

[1] Taking an overall estimator of the variance of the estimator requires further information on how we get $\mathbf{X}$

First, focus on $\mathbf{X'X}$:

$$\mathbf{X'X} = \begin{pmatrix} 1 & 1 & ... & 1 \\ X_1 & X_2 & ... & X_N \end{pmatrix} \times \begin{pmatrix} 1 & X_1 \\ 1 & X_2 \\ ... & \\ 1 & X_N \end{pmatrix}$$
$$= \begin{pmatrix} n & \sum_{i=1}^N X_i \\ \sum_{i=1}^N X_i & \sum_{i=1}^N X_i^2 \end{pmatrix}$$

Inverting a 2×2 matrix is easy using the determinant formula, provided $ad - bc \neq 0$:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \Rightarrow A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Using this method :

$$(\mathbf{X'X})^{-1} = \frac{1}{n \sum_{i=1}^N X_i^2 - \left(\sum_{i=1}^N X_i\right)^2} \begin{pmatrix} \sum_{i=1}^N X_i^2 & -\sum_{i=1}^N X_i \\ -\sum_{i=1}^N X_i & n \end{pmatrix}$$

Now, notice that:

$$S_{xx} = \sum_{i=1}^N \left(X_i - \bar{X})\right)^2$$
$$= \sum_{i=1}^N \left(X_i^2 - 2X_i\bar{X} + \bar{X}^2\right)$$
$$= \sum_{i=1}^N X_i^2 - 2\left(\sum_{i=1}^N X_i\right)\bar{X} + \sum_{i=1}^N \bar{X}^2$$
$$= \sum_{i=1}^N X_i^2 - 2\left(\sum_{i=1}^N X_i\right)\left(\frac{\sum_{i=1}^N X_i}{n}\right) + n \times \left(\frac{\sum_{i=1}^N X_i}{n}\right)^2$$
$$= \sum_{i=1}^N X_i^2 - 2\frac{\left(\sum_{i=1}^N X_i\right)^2}{n} + \frac{\left(\sum_{i=1}^N X_i\right)^2}{n}$$

$$S_{xx} = \sum_{i=1}^N X_i^2 - \frac{1}{n}\left(\sum_{i=1}^N X_i\right)^2 \quad (6)$$

Hence, using equation 6 back to invering $\mathbf{X'X}$:

$$
\begin{aligned}
(\mathbf{X'X})^{-1} &= \frac{1}{n\sum_{i=1}^N X_i^2 - \left(\sum_{i=1}^N X_i\right)^2} \begin{pmatrix} \sum_{i=1}^N X_i^2 & -\sum_{i=1}^N X_i \\ -\sum_{i=1}^N X_i & n \end{pmatrix} \\
&= \frac{\frac{1}{n}}{\sum_{i=1}^N X_i^2 - \frac{1}{n}\left(\sum_{i=1}^N X_i\right)^2} \begin{pmatrix} \sum_{i=1}^N X_i^2 & -\sum_{i=1}^N X_i \\ -\sum_{i=1}^N X_i & n \end{pmatrix} \\
&= \frac{\frac{1}{n}}{S_{xx}} \begin{pmatrix} \sum_{i=1}^N X_i^2 & -\sum_{i=1}^N X_i \\ -\sum_{i=1}^N X_i & n \end{pmatrix}
\end{aligned}
$$

Now, we can find the *variance-covariance matrix* of the estimators β_0 and β_1 using $\sigma^2\mathbf{X'X}^{-1}$. Use the diagonal elements to find the variances of the coefficients :

$$
\begin{aligned}
\mathbf{V}(\beta_0) &= \frac{\frac{\sigma^2}{n}}{S_{xx}} \sum_{i=1}^N X_i^2 \\
&= \frac{\sigma^2}{nS_{xx}} \left(S_{xx} + \frac{1}{n}\left(\sum_{i=1}^N X_i\right)^2 \right) \\
&= \frac{\sigma^2}{nS_{xx}} \left(S_{xx} + n\bar{X}^2 \right) \\
&= \sigma^2 \left(\frac{1}{n} + \frac{\bar{X}^2}{S_{xx}} \right) \\
&\Rightarrow SE(\beta_0) = \sigma \sqrt{\left(\frac{1}{n} + \frac{\bar{X}^2}{S_{xx}} \right)}
\end{aligned}
$$

I use equation 6 to go from line 1 to 2, the fact that $\frac{1}{n}\left(\sum_{i=1}^N X_i\right)^2 = n\bar{X}^2$ to go from 2 to 3.

For β_1 :

$$
\begin{aligned}
\mathbf{V}(\beta_1) &= \frac{\frac{1}{n}}{S_{xx}} n\sigma^2 \\
&= \frac{\sigma^2}{S_{xx}} \\
&\Rightarrow SE(\beta_1) = \frac{\sigma}{\sqrt{S_{xx}}}
\end{aligned}
$$