Question 3. In the two-variable model

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i, \quad i = 1, \ldots, 11$$

Suppose that $x_1'x_1 = 2, x_2'x_2 = 2, x_1'x_2 = 1, x_1'y = 2, x_2'y = 1, y'y = 7/3$ where x_1, x_2 and y are the column vectors with typical elements x_{1i}, x_{2i} and y_i respectively.

Assume $\epsilon_i \sim$ i.i.d. $N\left(0, \sigma_\epsilon^2\right)$. Suppose you would like to make out-of-sample predictions about the left-hand-side (dependent) variable for two hypothetical observations with the following characteristics:

Obs.	x_1	x_2
12	5	-2
13	3	-7

sol)

$$x_1'x_1 = \begin{bmatrix} x_{11} & \cdots & x_{1i} \end{bmatrix} \begin{bmatrix} x_{11} \\ \vdots \\ x_{1i} \end{bmatrix} = \sum x_{1i}^2$$

$$x_1'y = \begin{bmatrix} x_{11} & \cdots & x_{1i} \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_i \end{bmatrix} = \sum x_{1i} y_i$$

$$y'y = \begin{bmatrix} y_1 & \cdots & y_i \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_i \end{bmatrix} = \sum y_i^2$$

$$x_2'x_2 = \sum x_{2i}^2$$

$$x_2'y = \sum x_{2i} y_i$$

$$x_1'x_2 = \sum x_{1i} x_{2i}$$

Denote

$$X = \begin{bmatrix} x_{11} & x_{21} \\ \vdots & \vdots \\ x_{1i} & x_{2i} \end{bmatrix} \quad X' = \begin{bmatrix} x_{11} & \cdots & x_{1i} \\ x_{21} & \cdots & x_{2i} \end{bmatrix} \quad y = \begin{bmatrix} y_1 \\ \vdots \\ y_i \end{bmatrix}$$

Then

$$X^T X = \begin{bmatrix} x_{11} & \cdots & x_{1i} \\ x_{21} & \cdots & x_{2i} \end{bmatrix} \begin{bmatrix} x_{11} & x_{21} \\ \vdots & \vdots \\ x_{1i} & x_{2i} \end{bmatrix} = \begin{bmatrix} \sum_i x_{1i}^2 & \sum_1 x_{1i} x_{2i} \\ \sum_i x_{2i} x_{1i} & \sum_i x_{2i}^2 \end{bmatrix} = \begin{bmatrix} x_1'x_1 & x_1'x_2 \\ x_1'x_2 & x_2'x_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$$\left(X^T X\right)^{-1} = \frac{1}{2 \cdot 2 - 1 \cdot 1} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

$$X'y = \begin{bmatrix} x_{11} & \cdots & x_{1i} \\ x_{21} & \cdots & x_{2i} \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_i \end{bmatrix} = \begin{bmatrix} \Sigma_i x_{1i} y_i \\ \sum_i x_{2i} y_i \end{bmatrix} = \begin{bmatrix} x_1'y \\ x_2'y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Using the formula

$$\begin{aligned} \hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} &= \left(X'X\right)^{-1} X'y \\ &= \frac{1}{3} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{aligned}$$

$$\therefore \hat{\beta}_1 = 1, \ \hat{\beta}_2 = 0$$

Insert the observations in the formula

$$\begin{aligned}
&\hat{y}_{12} = \hat{\beta}_1 x_{1,12} + \hat{\beta}_2 x_{2,12} = 1 \cdot 5 + 0 \cdot (-2) = 5 \\
&\hat{y}_{13} = \hat{\beta}_1 x_{1,12} + \hat{\beta}_2 x_{2,12} = 1 \cdot 3 + 0 \cdot (-7) = 3 \\
&\therefore \hat{y} = x_{1i}
\end{aligned}$$

$$\begin{aligned}
\hat{\sigma}^2 &= \frac{1}{n-2}\sum \hat{\epsilon}_i^2 = \frac{1}{n-2}\sum (y_i - \hat{y})^2 = \frac{1}{n-2}\sum (y_i - x_{1i})^2 \\
&= \frac{1}{n-2}\left(\sum y_i^2 - 2\Sigma y_i x_{1i} + \sum x_{1i}^2\right) = \frac{1}{11-2}\left(\frac{7}{3} - 2 \cdot 2 + 2\right) = \frac{1}{27}
\end{aligned}$$

1. Construct 80% prediction interval for the dependent variable y_{12} and y_{13}.

Let X^* be a row vector containing the values of the future observation.

$$y_f = X\hat{\beta} + \hat{\varepsilon} \quad , \quad \hat{\beta} = \left(X'X\right)^{-1} X'y$$

If, $\mathrm{Var}(\hat{\varepsilon}) = \hat{\sigma}_{\varepsilon}^2 I,\ X$ is Full rank

$$\begin{aligned}
\mathrm{Var}\left(y_f\right) &= \mathrm{Var}\left(X^*\hat{\beta} + \hat{\varepsilon}\right) \\
&= \mathrm{Var}\left(X^*\left(X'X\right)^{-1} X'y\right) + \mathrm{Var}(\hat{\varepsilon}) \\
&= X^*\left(X'X\right)^{-1} X' \,\mathrm{Var}(y)\left[X^*\left(X'X\right)^{-1} X'\right]' + \hat{\sigma}^2 \\
&= \hat{\sigma}^2 X^*\left(X'X\right)^{-1} X'X\left(X'X\right)^{-1} X^{*'} + \hat{\sigma}^2 \\
&= \hat{\sigma}^2\left(X^*(X'X)^{-1}X^{*'} + 1\right)
\end{aligned}$$

Denote

$$\begin{aligned}
&X_{12} = \begin{bmatrix} 5 & -2 \end{bmatrix} \\
&X_{12}' = \begin{bmatrix} 5 \\ -2 \end{bmatrix} \\
&\left(X'X\right)^{-1} = \frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \\
&\hat{\sigma}^2 = \frac{1}{27}
\end{aligned}$$

$$\begin{aligned}
&\mathrm{Var}\left(y_{12}\right) = \frac{1}{27}\left(\begin{bmatrix} 5 & -2 \end{bmatrix}\frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 5 \\ -2 \end{bmatrix} + 1\right) = 1 \\
&\mathrm{Var}\left(y_{13}\right) = \frac{1}{27}\left(\begin{bmatrix} 3 & -7 \end{bmatrix}\frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 3 \\ -7 \end{bmatrix} + 1\right) = \frac{1}{27}\left(\frac{158}{3} + 1\right)
\end{aligned}$$

Thus

$$\begin{aligned}
&80\%\text{ PI for } y_{12} = \left[\hat{y}_{12} \pm t_{(0.1,9)}\sqrt{\mathrm{Var}\left(y_{12}\right)}\right] = 5 \pm 1.383 = [3.617, 6.383] \\
&80\%\text{ PI for } y_{13} = \left[\hat{y}_{13} \pm t_{(0.1,9)}\sqrt{\mathrm{Var}\left(y_{13}\right)}\right] = 3 \pm 1.383 \times 1.41 = [1.050, 4.949]
\end{aligned}$$

2. Construct 80% prediction interval for the expected value of y_{12} and y_{13}.

$$\begin{aligned}
Var\left(y_f \mid x\right) &= Var\left(X^*\hat{\beta}\right) \\
&= Var\left(X^*\left(X'X\right)^{-1}X'y\right) \\
&= X^*\left(X'X\right)^{-1}X'Var(y)\left[X^*\left(X'X\right)^{-1}X'\right]' \\
&= \hat{\sigma}^2 X^*\left(X'X\right)^{-1}X'X\left(X'X\right)^{-1}X^{*'} \\
&= \hat{\sigma}^2 X^*(X'X)^{-1}X^{*'}
\end{aligned}$$

$$\begin{aligned}
Var\left(y_{12} \mid x_{12}\right) &= \frac{1}{27}\left(\begin{bmatrix} 5 & -2 \end{bmatrix}\frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 5 \\ -2 \end{bmatrix}\right) = \frac{26}{27} \\
Var\left(y_{13} \mid x_{13}\right) &= \frac{1}{27}\left(\begin{bmatrix} 3 & -7 \end{bmatrix}\frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 3 \\ -7 \end{bmatrix}\right) = \frac{1}{27}\cdot\frac{158}{3} = 1.951
\end{aligned}$$

$$\begin{aligned}
\text{80\% PI for } E\left(y_{12} \mid x_{12}\right) &= \left[\hat{y}_{12} \pm t_{(0.1,9)}\cdot\sqrt{Var\left(y_{12} \mid x_{12}\right)}\right] = 5 \pm 1.383\cdot\frac{26}{27} = [3.642, 6.357] \\
\text{80\% PI for } E\left(y_{13} \mid x_{13}\right) &= \left[\hat{y}_{13} \pm t_{(0.1,9)}\cdot\sqrt{Var\left(y_{13} \mid x_{13}\right)}\right] = 3 \pm 1.383\cdot 1.397 = [1.068, 4.932]
\end{aligned}$$

3. Do the answers above differ? Why?

   Prediction intervals for y_{12} and y_{13} are wider than PI for $E\left[y_{12} \mid x_{12}\right]$ and $E\left[y_{13} \mid x_{13}\right]$. Because PI for y_{12} and y_{13} including the uncertainty (the error term ε).