Stat 462 HW2 Jiaqi Li

1. By using $\hat{\beta} = (X^TX)^{-1}X^Ty$,

$$\underbrace{\begin{bmatrix} 0.5 & 0 & -0.25 \\ 0 & 0.01 & 0 \\ -0.25 & 0 & 0.5 \end{bmatrix}}_{(X^TX)^{-1}} \underbrace{\begin{bmatrix} 12 \\ -50 \\ 20 \end{bmatrix}}_{X^Ty} = \begin{bmatrix} 1 \\ -0.5 \\ 2 \end{bmatrix} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix}$$

Then, $\hat{\beta}_0 = 1$, $\hat{\beta}_1 = -0.5$, $\hat{\beta}_2 = 2$

Thus, $\hat{y} = 1 - 0.5x_1 + 2x_2$

$$\text{Then } \hat{y} = \begin{bmatrix} 1 - 0.5\times 5 + 2\times 2 \\ 1 - 0.5\times 5 + 2\times 0 \\ 1 - 0.5\times(-5) + 2\times 2 \\ 1 - 0.5\times(-5) + 2\times 0 \end{bmatrix} = \begin{bmatrix} 2.5 \\ -1.5 \\ 7.5 \\ 3.5 \end{bmatrix}$$

Thus, $\hat{y}_1 = 2.5$, $\hat{y}_2 = -1.5$, $\hat{y}_3 = 7.5$, $\hat{y}_4 = 3.5$.

2.

① For $\hat{\beta}_1$, we have:

$$\hat{\beta}_1 = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2} = \frac{\sum_i (x_i - \bar{x}) y_i}{\sum_i (x_i - \bar{x})^2}.$$

Since $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$

$$\hat{\beta}_1 = \frac{\sum_i (x_i - \bar{x})(\beta_0 + \beta_1 x_i + \varepsilon_i)}{\sum_i (x_i - \bar{x})^2}$$

$$= \frac{\beta_0 \sum_i (x_i - \bar{x}) + \beta_1 \sum_i (x_i - \bar{x}) x_i + \sum_i (x_i - \bar{x})\varepsilon_i}{\sum_i (x_i - \bar{x})^2}$$

$$= \frac{\beta_1 \sum_i (x_i - \bar{x}) x_i + \sum_i (x_i - \bar{x})\varepsilon_i}{\sum_i (x_i - \bar{x})^2}$$

$$= \frac{\beta_1 \left[\sum_i (x_i - \bar{x})(x_i - \bar{x}) + \sum_i (x_i - \bar{x})\bar{x}\right] + \sum_i (x_i - \bar{x})\varepsilon_i}{\sum_i (x_i - \bar{x})^2}$$

1. $$\hat{\beta}_1 = \frac{\beta_1 \sum_i (x_i-\bar{x})^2 + \sum_i (x_i-\bar{x})\varepsilon_i}{\sum_i (x_i-\bar{x})^2}$$

$$= \beta_1 + \frac{\sum_i (x_i-\bar{x})\varepsilon_i}{\sum_i (x_i-\bar{x})^2}$$

$$Var(\hat{\beta}_1) = Var\left(\beta_1 + \frac{\sum_i (x_i-\bar{x})\varepsilon_i}{\sum_i (x_i-\bar{x})^2}\right)$$

$$= \frac{\sum_i (x_i-\bar{x})^2}{\left[\sum_i (x_i-\bar{x})^2\right]^2} \cdot \sigma^2 \quad \text{since } \varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$$

$$= \frac{\sigma^2}{\sum_i (x_i-\bar{x})^2}$$

Since we assumed $\hat{\beta}_1$ is an unbiased estimator, $E(\hat{\beta}_1) = \beta_1$

Thus, $\hat{\beta}_1 \sim N\left(\beta_1, \frac{\sigma^2}{\sum_i (x_i-\bar{x})^2}\right)$

② For $\hat{\beta}_0$, we have $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$

Since $\bar{y} = \beta_0 + \beta_1 \bar{x} + \frac{1}{n}\sum_i \varepsilon_i$, plug in and we have:

$$\hat{\beta}_0 = \left(\beta_0 + \beta_1\bar{x} + \frac{1}{n}\sum_i \varepsilon_i\right) - \hat{\beta}_1 \bar{x} = \left(\beta_0 + \beta_1\bar{x} + \frac{1}{n}\sum_i \varepsilon_i\right) - \left(\beta_1\bar{x} + \frac{\sum_i (x_i-\bar{x})\varepsilon_i}{\sum_i (x_i-\bar{x})^2}\bar{x}\right)$$

$$= \beta_0 + \frac{1}{n}\sum_i \varepsilon_i - \frac{\sum_i (x_i-\bar{x})\varepsilon_i}{\sum_i (x_i-\bar{x})^2}\bar{x} = \beta_0 + \sum_i \left(\varepsilon_i\left[\frac{1}{n} - \frac{(x_i-\bar{x})\bar{x}}{(x_i-\bar{x})^2}\right]\right)$$

$$Var(\hat{\beta}_0) = Var\left(\beta_0 + \sum_i \left(\varepsilon_i\left[\frac{1}{n} - \frac{(x_i-\bar{x})}{(x_i-\bar{x})^2}\right]\right)\right)$$

$$= \sum_i \left[\frac{1}{n} - \frac{(x_i-\bar{x})\bar{x}}{(x_i-\bar{x})^2}\right]^2 \sigma^2 = \left[\sum_i \frac{1}{n^2} - \frac{2\bar{x}}{n}\frac{\sum_i (x_i-\bar{x})}{\sum_i (x_i-\bar{x})^2} + \frac{\sum_i (x_i-\bar{x})^2 \bar{x}^2}{\left[\sum_i (x_i-\bar{x})^2\right]^2}\right]\sigma^2$$

$$= \left[\frac{1}{n} + \frac{\bar{x}^2}{\sum_i (x_i-\bar{x})^2}\right]\sigma^2$$

Since we assumed $\hat{\beta}_0$ is an unbiased estimator, $E(\hat{\beta}_0) = \beta_0$

Thus, $\hat{\beta}_0 \sim N\left(\beta_0, \left[\frac{1}{n} + \frac{\bar{x}^2}{\sum_i (x_i-\bar{x})^2}\right]\sigma^2\right)$.