High-Dimensional Data Analysis Homework #2

102225014 陳愛群 Ai-Chun Chen

Question:

In Best-subset Selection, $\mathrm{Y} = \mathrm{IQ}$, $\mathrm{X}_1 = \mathrm{Gender}$, $\mathrm{X}_2 = \mathrm{Height}$ and $\mathrm{X}_3 = \mathrm{Age}$.
Consider one of the candidate model when $k = 2$, $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$.

Compute the residual sum of square $\mathrm{RSS}^2(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2)$.

Solution:

$$\mathrm{RSS}^2(\beta_0, \beta_1, \beta_2) = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2})^2$$

$$\frac{\partial}{\partial \beta_0} \mathrm{RSS}^2(\beta_0, \beta_1, \beta_2) = -2\sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2}) \overset{\text{let}}{=} 0$$

$$\Rightarrow n\bar{y} - n\beta_0 - n\beta_1 \bar{x}_{\cdot 1} - n\beta_2 \bar{x}_{\cdot 2} = 0 \quad \cdots \quad (1)$$

$$\frac{\partial}{\partial \beta_1} \mathrm{RSS}^2(\beta_0, \beta_1, \beta_2) = -2\sum_{i=1}^{n} x_{i1}(y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2}) \overset{\text{let}}{=} 0$$

$$\Rightarrow \langle \mathbf{x}_1, \mathbf{y} \rangle - n\beta_0 \bar{x}_{\cdot 1} - \beta_1 \langle \mathbf{x}_1, \mathbf{x}_1 \rangle - \beta_2 \langle \mathbf{x}_1, \mathbf{x}_2 \rangle = 0 \quad \cdots \quad (2)$$

$$\frac{\partial}{\partial \beta_2} \mathrm{RSS}^2(\beta_0, \beta_1, \beta_2) = -2\sum_{i=1}^{n} x_{i2}(y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2}) \overset{\text{let}}{=} 0$$

$$\Rightarrow \langle \mathbf{x}_2, \mathbf{y} \rangle - n\beta_0 \bar{x}_{\cdot 2} - \beta_1 \langle \mathbf{x}_1, \mathbf{x}_2 \rangle - \beta_2 \langle \mathbf{x}_2, \mathbf{x}_2 \rangle = 0 \quad \cdots \quad (3)$$

$(2) - (1) \times \bar{x}_{\cdot 1}$: (eliminate β_0)

$$\left(\langle \mathbf{x}_1, \mathbf{y} \rangle - n\bar{y}\bar{x}_{\cdot 1}\right) - \beta_1 \left(\langle \mathbf{x}_1, \mathbf{x}_1 \rangle - n\bar{x}_{\cdot 1}^2\right) - \beta_2 \left(\langle \mathbf{x}_1, \mathbf{x}_2 \rangle - n\bar{x}_{\cdot 1}\bar{x}_{\cdot 2}\right) = 0 \quad \cdots \quad (4)$$

$(3) - (1) \times \bar{x}_{\cdot 2}$: (eliminate β_0)

$$\left(\langle \mathbf{x}_2, \mathbf{y} \rangle - n\bar{y}\bar{x}_{\cdot 2}\right) - \beta_1 \left(\langle \mathbf{x}_1, \mathbf{x}_2 \rangle - n\bar{x}_{\cdot 1}\bar{x}_{\cdot 2}\right) - \beta_2 \left(\langle \mathbf{x}_2, \mathbf{x}_2 \rangle - n\bar{x}_{\cdot 2}^2\right) = 0 \quad \cdots \quad (5)$$

(4) can be written as $\mathrm{SS}_{\mathbf{x}_1\mathbf{y}} - \beta_1 \mathrm{SS}_{\mathbf{x}_1} - \beta_2 \mathrm{SS}_{\mathbf{x}_1\mathbf{x}_2} = 0$

(5) can be written as $\mathrm{SS}_{\mathbf{x}_2\mathbf{y}} - \beta_1 \mathrm{SS}_{\mathbf{x}_1\mathbf{x}_2} - \beta_2 \mathrm{SS}_{\mathbf{x}_2} = 0$

$$(4) \times \mathrm{SS}_{\mathbf{x}_1\mathbf{x}_2} - (5) \times \mathrm{SS}_{\mathbf{x}_1} : \left(\mathrm{SS}_{\mathbf{x}_1\mathbf{y}}\mathrm{SS}_{\mathbf{x}_1\mathbf{x}_2} - \mathrm{SS}_{\mathbf{x}_2\mathbf{y}}\mathrm{SS}_{\mathbf{x}_1}\right) - \beta_2 \left(\mathrm{SS}_{\mathbf{x}_1\mathbf{x}_2}^2 - \mathrm{SS}_{\mathbf{x}_1}\mathrm{SS}_{\mathbf{x}_2}\right) = 0$$

$$\Rightarrow \hat{\beta}_2 = \frac{\mathrm{SS}_{\mathbf{x}_1\mathbf{y}}\mathrm{SS}_{\mathbf{x}_1\mathbf{x}_2} - \mathrm{SS}_{\mathbf{x}_2\mathbf{y}}\mathrm{SS}_{\mathbf{x}_1}}{\mathrm{SS}_{\mathbf{x}_1\mathbf{x}_2}^2 - \mathrm{SS}_{\mathbf{x}_1}\mathrm{SS}_{\mathbf{x}_2}}$$

Similarly, $\hat{\beta}_1 = \dfrac{\mathrm{SS}_{\mathbf{x}_1\mathbf{y}}\mathrm{SS}_{\mathbf{x}_2} - \mathrm{SS}_{\mathbf{x}_2\mathbf{y}}\mathrm{SS}_{\mathbf{x}_1\mathbf{x}_2}}{\mathrm{SS}_{\mathbf{x}_1}\mathrm{SS}_{\mathbf{x}_2} - \mathrm{SS}_{\mathbf{x}_1\mathbf{x}_2}^2}$ and $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}_{\cdot 1} - \hat{\beta}_2 \bar{x}_{\cdot 2}$

Thus, $\text{RSS}^2(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2) = \sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2})^2$.