Linear Regression

Newton's three sisters

February 9, 2025

Department of Statistics
Sungshin Women's University

Outline

d

c

b

Outline

Outline

Outline

Hat matrix

- Hat matrix defined by

$$\hat{y} = X\hat{\beta} = X(X^TX)^{-1}X^Ty = Hy$$

$$H \triangleq X(X^TX)^{-1}X^T$$

- some properties

$$H^2 = X(X^TX)^{-1}X^T \cdot X(X^TX)^{-1}X^T = X(X^TX)^{-1}X^T = H$$

$$(I-H)^2 = I - 2H + H^2 = I - H$$

$$HX = X(X^TX)^{-1}X^T \cdot X = X$$

Residual sum of square

- RSS defined

$$\text{RSS} \triangleq ||y - \hat{y}||^2$$

- Using hat matrix

$$\begin{aligned} y - \hat{y} = y - Hy = (I - H)y &= (I - H)(X\beta + \varepsilon) \\ &= (X - HX)\beta + (I - H)\varepsilon = (I - H)\varepsilon \end{aligned}$$

$$RSS \triangleq ||y - \hat{y}||^2 = \{(I - H)\varepsilon\}^T (I - H)\varepsilon = \varepsilon^T (I - H)^2 \varepsilon = \varepsilon^T (I - H)\varepsilon$$

Eigenvalues of H and Null space of $(I - H)$

- Proof by contrapositive

$$Hx = x \Rightarrow (I - H)x = 0$$
$$(I - H)x = 0 \Rightarrow Hx = x$$

- Dimensions of the eigenspaces of H is $p + 1$

Proof using $\text{rank}(X) = p + 1$

$$\text{rank}(H) \leq \min\{\text{rank}(X(X^T X)^{-1}), \text{rank}(X)\} \leq \text{rank}(X) = p + 1$$
$$\text{rank}(H) \geq \text{rank}(HX) = \text{rank}(X) = p + 1$$

- Dimensions of the null space of $I - H$ is $N - (p + 1)$

$$P(I - H)P^T = \text{diag}(\underbrace{1, \ldots, 1}_{N-p-1}, \underbrace{0, \ldots, 0}_{p+1})$$

- We define $v = P\varepsilon \in \mathbb{R}^N$, then from $\varepsilon = P^T v$

$$\text{RSS} = \varepsilon^T (I - H)\varepsilon = (P^T v)^T (I - H) P^T v = v^T P(I - H)P^T v$$

$$= [v_1, \cdots, v_{N-p-1}, v_{N-p}, \cdots, v_n] \begin{bmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & 0 & \cdots & \cdots & \vdots \\ \vdots & 0 & 1 & \cdots & \cdots & 0 \\ 0 & 0 & 0 & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & \cdots & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_{N-p-1} \\ v_{N-p} \\ \vdots \\ v_N \end{bmatrix} = \sum_{i=1}^{N-p-1} v_i^2$$

- Let $w \in \mathbb{R}^{N-p-1}$
- Average

  $E[v] = E[P\varepsilon] = 0$

  $E[w] = 0$
- Covariance

  $E[vv^t] = E[P\varepsilon(P\varepsilon)^T] = PE[\varepsilon\varepsilon^t]P = P\sigma^2 IP^T = \sigma^2 I$

  $E[ww^T] = \sigma^2 I$
- We have RSS

$$\frac{RSS}{\sigma^2} \sim \chi^2_{N-p-1}$$

Outline

- We define a matrix $W \in \mathbb{R}^{N \times N}$ such that all the elements are $1/N$
- Explained sum of squres ESS

$$\mathrm{ESS} \triangleq ||\hat{y} - \bar{y}||^2 = ||\hat{y} - Wy||^2 = ||(H - W)y||^2$$

- Total sum of squres TSS

$$\mathrm{TSS} \triangleq ||y - \bar{y}||^2 = ||(I - W)y||^2$$

Outline

a

Q & A

Thank you :)