1060640?1
蔡昀霖 (HW2)

1. $k=1 \Rightarrow y_i=1$, $k=2 \Rightarrow y_i=0$

$$\hat{\mu}_1 = \frac{1}{396}\left(0.05\times 31 + 0.1\times 5) + 0.15\times 68 + 0.2\times 121 + 0.3\times 144\right) = 0.202$$

$$\hat{\mu}_2 = \frac{1}{604}\cdot\left(0.05\times 169 + 0.1\times 148 + 0.15\times 132 + 0.2\times 89 + 0.3\times 36\right) = 0.132$$

$$\hat{\pi}_1 = \frac{396}{1200} \quad \hat{\pi}_2 = \frac{604}{1200}$$

$$\hat{\sigma}_1^2 = \frac{1}{396-1}\cdot\sum_{y_i=1}(x_i-0.202)^2 = \frac{2.824}{395} = 0.0071$$

$$\hat{\sigma}_2^2 = \frac{1}{604-1}\sum_{y_i=0}(x_i-0.132)^2 = \frac{3.368}{603} = 0.0055$$

$$\hat{\delta}_k(x) = \frac{-x^2}{2\sigma_k^2} + \frac{x\mu_k}{\sigma_k^2} - \frac{\mu_k^2}{2\sigma_k^2} + \log\pi_k$$

$5\% = \delta_1(0.05) \Rightarrow 0.17$
$10\% : \delta_1(0.1) \Rightarrow 0.23$
$15\% = \delta_1(0.15) \Rightarrow 0.33$
$20\% = \delta_1(0.2) \Rightarrow 0.46$
$25\% = \delta_1(0.25) \Rightarrow 0.63$
$30\% = \delta_1(0.3) \Rightarrow 0.78$ #

← 查折扣所對應的兌換率 (代入公式算得)

---

2. $\hat{\sigma}^2 = 25$. $(\bar{x}_Y = 10,\ \bar{x}_N = 0)$

$$p_{Yes}(4) = \frac{0.8\cdot\exp\left(-\frac{1}{2\times 25}\cdot(4-10)^2\right)\cdot\frac{1}{5}}{\frac{1}{5}\cdot 0.8\times\exp\left(-\frac{1}{2\times 25}\cdot(4-10)^2\right) + (1-0.8)\cdot\exp\left(-\frac{1}{2\times 36}\times(4-0)^2\right)\frac{1}{6}} = 0.774 \ \#$$

---

3. MSE $CV(n) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i-\hat{y}_i}{1-h_i}\right)^2$, $h_i = x_i^T(X^TX)^{-1}x_i$

proof ① $\hat{y} = X(X^TX)^{-1}X^Ty \Rightarrow \hat{y}_i = x_i^T(X^TX^{-1})X^Ty$

② $\hat{y}_i^{(-i)} = x_i^T\left(X^{(-i)T}X^{(-i)}\right)^{-1}X^{(-i)T}\cdot y^{(-i)}$ ← Removing the ith term.

$$= x_i^T\left(X^TX - x_ix_i^T\right)^{-1}\left(X^TY - x_iy_i\right)$$

hint $(A - XX^T)^{-1} = A^{-1} + \frac{A^{-1}XX^TA^{-1}}{1-X^TA^{-1}X}$ $\Rightarrow$ $= x_i^T\left[\underbrace{(X^TX)^{-1} + \frac{(X^TX)^{-1}x_ix_i^T(X^TX)^{-1}}{1-x_i^T(X^TX)^{-1}x_i}}_{\text{let } A}\right](X^TY - x_iy_i)$

① $\Rightarrow A\cdot x_i = (X^TX)^{-1}x_i + (X^TX)^{-1}x_i\left(\frac{x_i^T(X^TX)^{-1}x_i}{1-x_i^T(X^TX)^{-1}x_i}\right)$

$$= \frac{1}{1-h_i}\cdot(X^TX)^{-1}x_i \Rightarrow x_i^T\cdot\frac{(X^TX)^{-1}x_i}{1-h_i}\cdot y_i = \frac{h_iy_i}{1-h_i} \ \text{— ①}$$

② $\Rightarrow (X^TX)^{-1}\cdot X^T + (X^TX)^{-1}x_i\cdot\left(\frac{x_i^T(X^TX)^{-1}X^T}{1-h_i}\right) = (X^TX)^{-1}\cdot\left[\frac{1\cdot(1-h_i)}{1-h_i} + \frac{x_ix_i^T(X^TX)^{-1}}{1-h_i}\right]\cdot X^T = \frac{(X^TX)^{-1}X^T}{1-h_i}$

$$\Rightarrow x_i^T\frac{(X^TX)^{-1}X^T}{1-h_i}\cdot Y = \frac{\hat{y}_i}{1-h_i} \ \text{— ②} \quad \Rightarrow \ ② - ① = \frac{\hat{y}_i - h_iy_i}{1-h_i} = \hat{y}_i^{(-i)}$$

③ MSE is

$$CV(n) = \frac{1}{n}\cdot\sum_{i=1}^{n}\left(y_i - \hat{y}_i^{(-i)}\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i-\hat{y}_i}{1-h_i}\right)^2 \ \#$$

4.(a) $L = \|Y - XB\|^2 + \lambda\|B\|^2$

$= (Y-XB)^T(Y-XB) + \lambda B^TB = Y^TY - Y^TXB - B^TX^TY + B^TX^TXB + \lambda B^TB$

对 B 偏微：

得 $= 0 - X^TY - X^TY + (X^TX + (X^TX)^T)B + 2\lambda B$

$= 2X^TXB - 2X^TY + 2\lambda B = 0$

$\Rightarrow 2X^T(XB - Y) + 2\lambda B = 0$

$\Rightarrow \hat{B}_\lambda = (X^TX + \lambda I)^{-1}X^Ty$ #

$\because \lambda \geq 0$ $\therefore$ 當 $\lambda > 0$ 時，如同分母變大 則 B 会縮小.

$\Rightarrow \|B_\lambda\| |_{\lambda>0} \leq \|B_\lambda\| |_{\lambda=0}$ #

(b) $err = \frac{1}{n} \times (y - X(X^TX+\lambda I)^{-1}X^Ty)^2$

$= \frac{1}{n} \times (I_m y - X(X^TX+\lambda I)^{-1}X^Ty)^2$

$= \frac{1}{n} \times \left[(I_m - X(X^TX+\lambda I)^{-1}X^T) \cdot Y\right]^2$

$= \frac{1}{n} \cdot Y^T \cdot (I_m - X(X^TX+\lambda I)^{-1}X^T)^2 \cdot Y$ #

$\because \lambda \geq 0$ $\therefore$ 當 λ 增大時 則扣会變小. $\Rightarrow I_m -$ 變小 則變大.

$\Rightarrow err$ 遞增 #

5.

H别

6 $(10,3,5) \longrightarrow (2.846, 1.192)$

2 $(10,3,3) \longrightarrow (2.37, 1.14)$

2 $(10,5,5) \longrightarrow (1, 1.14)$

2 $(3,10,10) \longrightarrow (3.57, 1.14)$

2 $(3,5,5) \longrightarrow (3.5, 0.5)$

2 $(3,10,10) \longrightarrow (1, 1.14)$

2 $(5,3,3) \longrightarrow (5.5, 0.5)$

共20種

加權平均

$\bar{\beta} = (3.229, 1.058)$

(a)

$RSS = 4.35 + 0.269 + 1.41$

$RSE = \sqrt{\frac{RSS}{7-2}} = 1.434 = \hat{\sigma}$

$SE(\beta_0) = \sqrt{(X^TX)^{-1}\hat{\sigma}^2} = \sqrt{3.467} = 1.86$

$SE(\beta_1) = \sqrt{(X^TX)^{-1}\hat{\sigma}^2} = \sqrt{0.077} = 0.27$

(b)

$\beta_0 = 2.84$

$\beta_1 = 1.192$

$SE(\beta_0) = 1.33$

$SE(\beta_1) = 0.2$. #

5.

比例

⑥ (10,3,5) ⟶ (2.846, 1.192)

② (10,3,3) ⟶ (3.37, 1.14)

② (10,5,5) ⟶ (1, 1.4)

② (3,10,10) ⟶ (3.57, 1.14)

② (3,5,5) ⟶ (3.5, 0.5)

② (5,10,10) ⟶ (1, 1.4)

③ (5,3,3) ⟶ (3.5, 0.5)

(20種)

加權平均

$\bar{\beta} = (3.229, 1.058)$

(a)

$\widehat{SE}(\beta) = \left(\frac{1}{B-1} \sum_{r=1}^{B} (\hat{\beta}_r - \bar{\beta})^2 \right)^{0.5} \Big|_{B \to \infty}$

∵ $B \to \infty$，∴ 可用相同比例者值。

$\Rightarrow \widehat{SE}(\beta_0) = \sqrt{2.597} = 1.61$

$\widehat{SE}(\beta_1) = \sqrt{0.1132} = 0.33$ #

(b)

$\hat{\beta}_0 = 2.84$, RSS $= 1.038$.

$\hat{\beta}_1 = 1.192$

$\widehat{SE}(\hat{\beta}_0) = 1.33$

$\widehat{SE}(\hat{\beta}_1) = 0.2$. #

6. (a) $(y_1-\hat\beta_1 x_{11}-\hat\beta_2 x_{12})^2+(y_2-\hat\beta_1 x_{21}-\hat\beta_2 x_{22})^2+\lambda(\hat\beta_1^2+\hat\beta_2^2)$ # (Ridge)

(b) ① 展開上式：$= y_1^2+\hat\beta_1^2 x_{11}^2+\hat\beta_2^2 x_{12}^2-2\hat\beta_1 x_{11}y_1-2\hat\beta_2 x_{12}y_1+2\hat\beta_1\hat\beta_2 x_{11}x_{12}$
$+y_2^2+\hat\beta_1^2 x_{21}^2+\hat\beta_2^2 x_{22}^2-2\hat\beta_1 x_{21}y_2-2\hat\beta_2 x_{22}y_2+2\hat\beta_1\hat\beta_2 x_{21}x_{22}$

② $\frac{d}{d\hat\beta_1}=(2\hat\beta_1 x_{11}^2-2x_{11}y_1+2\hat\beta_2 x_{11}x_{12})+(2\hat\beta_1 x_{21}^2-2x_{21}y_2+2\hat\beta_2 x_{21}x_{22})+2\lambda\hat\beta_1=0$

③ $x_{11}=-x_{12}=a$ $\Rightarrow$ 上式 $=(\hat\beta_1 a^2-ay_1-\hat\beta_2 a^2)+(\hat\beta_1 b^2-by_2-\hat\beta_2 b^2)+\lambda\hat\beta_1=0$
$x_{21}=-x_{22}=b$
$=\hat\beta_1\cdot(a^2+b^2)-\hat\beta_2(a^2+b^2)+\lambda\hat\beta_1=ay_1+by_2$

④ Add $2\hat\beta_1 ab$, $-2\hat\beta_2 ab$

$\Rightarrow \hat\beta_1(a+b)^2-\hat\beta_2(a+b)^2+\lambda\hat\beta_1=ay_1+by_2+2\hat\beta_1 ab-2\hat\beta_2 ab$

$\Rightarrow \lambda\hat\beta_1=ay_1+by_2+2\hat\beta_1 ab-2\hat\beta_2 ab$

類似地可得 $\Rightarrow \lambda\hat\beta_2=-ay_1-by_2-2\hat\beta_1 ab+2\hat\beta_2 ab$

$\Rightarrow \hat\beta_1=-\hat\beta_2$ #

(c) Lasso $(y_1-\hat\beta_1 x_{11}-\hat\beta_2 x_{12})^2+(y_2-\hat\beta_1 x_{21}-\hat\beta_2 x_{22})^2+\lambda(|\hat\beta_1|+|\hat\beta_2|)$ #

(d) 僅變動正則化項而已.

$\Rightarrow \frac{d}{dB}\lambda|B|=\lambda\frac{|B|}{B} \Rightarrow \lambda\frac{|\hat\beta_1|}{\hat\beta_1}=-\lambda\frac{|\hat\beta_2|}{\hat\beta_2}$ #

(hint $\frac{d}{dx}|x|=\frac{|x|}{x}$).