1



# Cross Validation and PRESS

$$\tilde{e}_i = y_i - x_i^T \hat{\beta}_{-i} \qquad \hat{\beta}_{-i} = (X_{-i}^T X_{-i})^{-1} X_{-i}^T y_{-i}$$

$$\hat{e}_i = y_i - x_i^T \hat{\beta} \quad , \quad \hat{\beta} = (X^T X)^{-1} X^T y$$

$$RSS = \sum \hat{e}_i^2 \quad , \qquad PRESS = \sum \tilde{e}_i^2 \approx E[PSS]$$

Thm: let $H = X(X^T X)^{-1} X^T$ ("Hat matrix" or Projection matrix)

then $\tilde{e}_i = e_i / (1 - h_{ii})$

Pf: two steps - ① use matrix identity $(A + bc)^{-1} = A^{-1} - \frac{A^{-1} b c^T A^{-1}}{1 + c^T A^{-1} b}$

② simplify result

to calc. $\tilde{e}_i$, we need $\hat{\beta}_{-i}$, which needs $(X_{-i}^T X_{-i})^{-1}$

$$\underset{p\times n\ n\times p}{X^T X} = \sum_{i=1}^{n} \underset{p\times 1\ 1\times p}{x_i x_i^T} = \sum_{j \neq i} x_j x_j^T + x_i x_i^T$$

$$= X_{-i}^T X_{-i} + x_i x_i^T$$

Use identity: 
$$(X_{-i}^T X_{-i})^{-1} = (X^T X - x_i x_i^T)^{-1}$$
$$= S^{-1} + \frac{S^{-1} x_i x_i^T S^{-1}}{1 - x_i^T S^{-1} x_i}$$

$x_i^T (X^T X)^{-1} x_i = (X (X^T X)^{-1} X^T)_{ii}$

$$= S^{-1} + S^{-1} x_i x_i^T S^{-1} / (1 - h_{ii})$$

$$\downarrow$$

$$x_i^T \hat{\beta}_{-i} = x_i^T \Big( \qquad \Big) X_{-i}^T y_{-i}$$
$$= x_i^T \Big( \qquad \Big) [X^T y - x_i y_i]$$

$$= x_i^T \underbrace{\boxed{S^{-1} X^T y}}_{\hat{\beta}} + \frac{x_i^T S^{-1} x_i x_i^T \overbrace{\boxed{S^{-1} X^T y}}^{\hat{\beta}}}{1 - h_{ii}} - x_i^T S^{-1} x_i y_i - \frac{x_i^T S^{-1} x_i x_i^T S^{-1} x_i y_i}{1 - h_{ii}}$$

$$= x_i^T \hat{\beta} + \frac{h_{ii}}{1 - h_{ii}} x_i^T \hat{\beta} - \left[ h_{ii} y_i + \frac{h_{ii} (h_{ii} y_i)}{1 - h_{ii}} \right]$$

(2)

$$= \frac{x_i^T \hat{\beta}}{1-h_{ii}} - \frac{h_{ii} y_i}{1-h_{ii}}$$

$$\tilde{e}_i = y_i - x_i^T \hat{\beta}_{-i} = \frac{(1-h_{ii})y_i - x_i^T \hat{\beta} + h_{ii} y_i}{1-h_{ii}}$$

$$= \frac{y_i - x_i^T \beta}{1-h_{ii}} = \underline{\underline{\frac{\hat{e}_i}{1-h_{ii}}}}$$

$$\text{PRESS} = \sum \tilde{e}_i^2 = \sum \left(\hat{e}_i / 1-h_{ii}\right)^2 \approx E[PSS]$$