Week Three

Monday, 1 February 2021 08:22

2.9 IML: $(A+BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}$ (A.1)

Sherman-morrison formula.

b and d are vectors, and $c = 1$.

$$(A + bd^T)^{-1} = A^{-1} - A^{-1}b\underbrace{(1 + d^TA^{-1}b)^{-1}}_{\text{scalar}} d^TA^{-1}$$

d was column matrix

Let dimensions be $n \times 1$ for d, so

d^T is a row vector of dimension $1 \times n$

$$\boxed{\underbrace{[d_1\, d_2 \cdots d_n]}_{1\times n}\underbrace{\begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn}\end{bmatrix}^{-1}}_{n \times n}\underbrace{\begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}}_{n\times 1} = \text{scalar}}$$

Thus, we get

$$(A + bd^T)^{-1} = A^{-1} - \frac{1}{1 + d^TA^{-1}b} A^{-1}bd^TA^{-1}$$

which is equation (A.2). ∎

$d^TA^{-1}b = e$ which is number ($\neq 0$, $\neq -1$)

$$(1 + d^TA^{-1}b)^{-1} = (1+e)^{-1} = \frac{1}{1+e}$$

$$e \cdot e^{-1} = e\frac{1}{e} = 1$$

2.11

We use this equation to obtain recursive least squares equations. See page 32, in textbook.

* $\hat{\beta}_{(n+1)} = \hat{\beta}_{(n)} + k_n e_{n+1}$ ← see eq (2.23)
* $k_n = h V_{(n)} \tilde{x}_{n+1}$
* $e_{n+1} = (y_{n+1} - \tilde{x}^T_{n+1}\hat{\beta}_{(n)})$
* $V_{(n)} = W^{-1}_{(n)} = (X^T_{(n)} X_{(n)})^{-1}$
* $X_{(n)} = \begin{bmatrix} x_1^T \\ \vdots \\ x_n^T \end{bmatrix} = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix}$

We are going to prove (2.24).

$$\boxed{Q_{n+1}(\hat{\beta}_{(n+1)}) = Q_n(\hat{\beta}_{(n)}) + he^2_{n+1}.}$$

Q_{n+1} we get (2.10)

$$Q_{n+1}(\hat{\beta}_{(n+1)}) = \|Y_{n+1} - \hat{Y}_{n+1}\|^2$$
$$= (Y_{n+1} - \hat{Y}_{n+1})^T(Y_{n+1} - \hat{Y}_{n+1})$$

Side note

$$Y_{n+1} = \begin{bmatrix} y_1 \\ \vdots \\ y_{n+1} \end{bmatrix} \qquad \hat{Y}_{n+1} = \begin{bmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_{n+1} \end{bmatrix}$$

$$= (Y_{n+1} - X_{n+1}\hat{\beta}_{(n+1)})^T(Y_{n+1} - \hat{Y}_{n+1})$$

Side note:

$$Y_{n+1} = \begin{bmatrix} Y_n \\ y_{n+1} \end{bmatrix}, \quad Y_n = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \qquad X_{n+1} = \underbrace{\begin{bmatrix} X_n \\ x^T_{n+1} \end{bmatrix}}_{(n+1)\times p}$$

X_n: a matrix $n \times p$; x^T_{n+1}: a row vector $1 \times p$

$$\|X - Y\|^2 = (X - Y)^T(X - Y)$$
$$= \left(\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} - \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}\right)^T\left(\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} - \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}\right)$$
$$= \begin{bmatrix} X_1 - Y_1 \\ X_2 - Y_2 \end{bmatrix}^T \begin{bmatrix} X_1 - Y_1 \\ X_2 - Y_2 \end{bmatrix}$$
$$= [(X_1 - Y_1)^T \; (X_2 - Y_2)^T]\begin{bmatrix} X_1 - Y_1 \\ X_2 - Y_2 \end{bmatrix}$$
$$= (X_1 - Y_1)^T(X_1 - Y_1) + (X_2 - Y_2)^T(X_2 - Y_2)$$
$$= \|X_1 - Y_1\|^2 + \|X_2 - Y_2\|^2$$

done with side note

$$(Y_{n+1} - X_{n+1}\hat{\beta}_{(n+1)})^T(Y_{n+1} - X_{n+1}\hat{\beta}_{(n+1)})$$
$$= \left(\begin{bmatrix} Y_n \\ y_{n+1} \end{bmatrix} - \begin{bmatrix} X_n \\ \tilde{x}^T_{n+1} \end{bmatrix}\hat{\beta}_{(n+1)}\right)^T\left(\begin{bmatrix} Y_n \\ y_{n+1} \end{bmatrix} - \begin{bmatrix} X_n \\ \tilde{x}^T_{n+1} \end{bmatrix}\hat{\beta}_{(n+1)}\right)$$
$$= (Y_n - X_n\hat{\beta}_{(n+1)})^T(Y_n - X_n\hat{\beta}_{(n+1)})$$
$$+ \underbrace{(y_{n+1} - \tilde{x}^T_{n+1}\hat{\beta}_{(n+1)})^T}_{\text{scalars}}\underbrace{(y_{n+1} - \tilde{x}^T_{n+1}\hat{\beta}_{(n+1)})}_{\text{scalars}}$$

$$= (Y_n - X_n[\hat{\beta}_{(n)} + k_ne_{n+1}])^T(Y_n - X_n[\hat{\beta}_{(n)} + k_ne_{n+1}])$$
$$+ (y_{n+1} - \tilde{x}^T_{n+1}[\hat{\beta}_{(n)} + k_ne_{n+1}])^2$$
$$= ((Y_n - X_n\hat{\beta}_{(n)}) - X_nk_ne_{n+1})^T((Y_n - X_n\hat{\beta}_{(n)}) - X_nk_ne_{n+1})$$
$$+ ((y_{n+1} - \tilde{x}^T_{n+1}\hat{\beta}_{(n)}) - \tilde{x}^T_{n+1}k_ne_{n+1})^2$$
$$= (Y_n - X_n\hat{\beta}_{(n)})^T(Y_n - X_n\hat{\beta}_{(n)}) - \underbrace{(Y_n - X_n\hat{\beta}_{(n)})^TX_n}_{0 \text{ (see p. 247)}}k_ne_{n+1}$$
$$- (X_nk_ne_{n+1})^T(Y_n - X_n\hat{\beta}_{(n)}) + (X_nk_ne_{n+1})^TX_nk_ne_{n+1}$$
$$+ \underbrace{(y_{n+1} - \tilde{x}^T_{n+1}\hat{\beta}_{(n)})^2}_{e^2_{n+1}} - 2\underbrace{(y_{n+1} - \tilde{x}^T_{n+1}\hat{\beta}_{(n)})}_{e_{n+1}}(\tilde{x}^T_{n+1}k_ne_{n+1})$$
$$+ (\tilde{x}^T_{n+1}k_n)^2e^2_{n+1}$$
$$= Q_n(\hat{\beta}_{(n)}) + e_{n+1}k_n^TX_n^TX_nk_ne_{n+1} + e^2_{n+1} - 2e_{n+1}\tilde{x}^T_{n+1}k_ne_{n+1}$$
$$+ (\tilde{x}^T_{n+1}k_n)^2e^2_{n+1}$$
$$= Q_n(\hat{\beta}_{(n)}) + k_n^TX_n^TX_nk_ne^2_{n+1} + e^2_{n+1} - 2\tilde{x}^T_{n+1}k_ne^2_{n+1}$$
$$+ (\tilde{x}^T_{n+1}k_n)^2e^2_{n+1}$$
$$= Q_n(\hat{\beta}_{(n)}) + \underbrace{[k_n^TX_n^TX_nk_n + 1 - 2\tilde{x}^T_{n+1}k_n + (\tilde{x}^T_{n+1}k_n)^2]}_{=h}e^2_{n+1}$$

Side note:

$k_n = hV_n\tilde{x}_{n+1}$ (see p. 32)

$$= Q_n(\hat{\beta}_{(n)}) + [(hV_n\tilde{x}_{n+1})^TX_n^TX_nhV_n\tilde{x}_{n+1} + 1 - 2\tilde{x}^T_{n+1}hV_n\tilde{x}_{n+1}$$
$$+ (\tilde{x}^T_{n+1}hV_n\tilde{x}_{n+1})^2]e^2_{n+1}$$
$$= Q_n(\hat{\beta}_{(n)}) + [h^2\tilde{x}^T_{n+1}\cancel{V_n^T}\cancel{X_n^TX_n}V_n\tilde{x}_{n+1} + 1 - 2h\tilde{x}^T_{n+1}V_n\tilde{x}_{n+1}$$
$$+ h^2(\tilde{x}^T_{n+1}V_n\tilde{x}_{n+1})^2]e^2_{n+1}.$$

Side note,

Recall: $V_n = (W_n)^{-1} = (X_n^TX_n)^{-1}$

$V_n^T = (W_n^{-1})^T = (W_n^T)^{-1} = W_n^{-1} = V_n$

$$= Q_n(\hat{\beta}_{(n)}) + [h^2\tilde{x}^T_{n+1}V_n\tilde{x}_{n+1} + 1 - 2h\tilde{x}^T_{n+1}V_n\tilde{x}_{n+1} + h^2(\tilde{x}^T_{n+1}V_n\tilde{x}_{n+1})^2]e^2_{n+1}$$

Side notes

p. 32, $h = \dfrac{1}{1 + \tilde{x}^T_{n+1}V_n\tilde{x}_{n+1}}$ $\Big| \cdot (1 + \tilde{x}^T_{n+1}V_n\tilde{x}_{n+1})$

$$\cancel{h} + h\tilde{x}^T_{n+1}V_n\tilde{x}_{n+1} = 1 - h$$

$$\tilde{x}^T_{n+1}V_n\tilde{x}_{n+1} = \frac{1-h}{h}$$

$$= Q_n(\hat{\beta}_{(n)}) + \left[h^{\cancel{2}}\frac{1-h}{\cancel{h}} + 1 - 2\cancel{h}\frac{1-h}{\cancel{h}} + \cancel{h^2}\left(\frac{1-h}{\cancel{h}}\right)^2\right]e^2_{n+1}$$
$$= Q_n(\hat{\beta}_{(n)}) + [h - h^2 + 1 - 2 + 2h + (1-h)^2]e^2_{n+1}$$
$$= Q_n(\hat{\beta}_{(n)}) + [h - \cancel{h^2} + \cancel{1} - \cancel{2} + \cancel{2h} + \cancel{1} - \cancel{2h} + \cancel{h^2}]e^2_{n+1}$$
$$= Q_n(\hat{\beta}_{(n)}) + he^2_{n+1}.$$

$$\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n}(y_i - \tilde{x}_i^T\hat{\beta}_{(n)})^2 \quad ∎$$

$$h = \frac{1}{1 + \tilde{x}^T_{n+1}V_n\tilde{x}_{n+1}}$$

$$e_{n+1} = (y_{n+1} - \tilde{x}^T_{n+1}\hat{\beta}_{(n)})$$

$$Q_{n+1}(\hat{\beta}_{(n+1)}) = \sum_{i=1}^{n+1}(y_i - x_i^T\hat{\beta}_{(n+1)})^2$$

3.1 Prove (3.2).

Want to show that

$$E[(\hat{Y} - f(x'))^2] = [E[\hat{Y}] - f(x')]^2 + Var(\hat{Y})$$

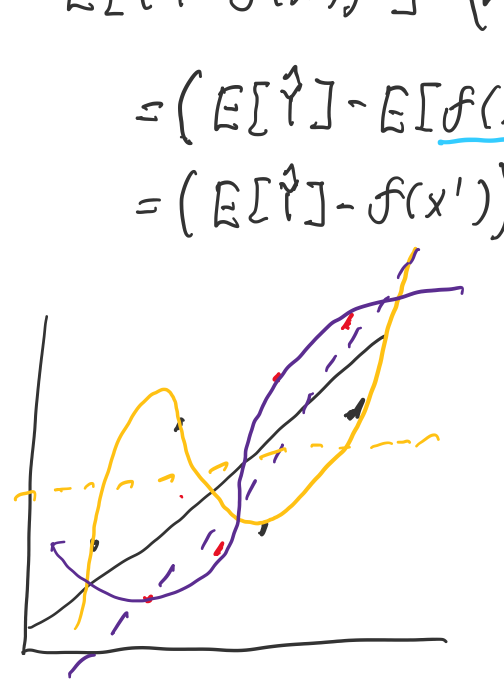
Recall: RV X then $Var(X) = E[X^2] - E[X]^2$

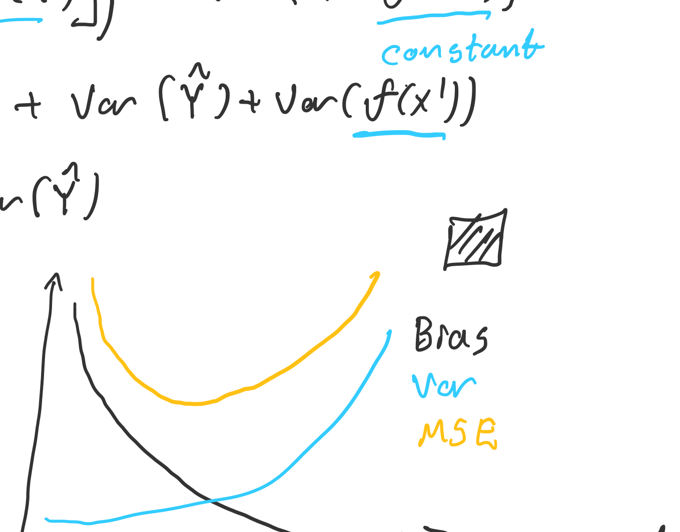
$\Leftrightarrow E[X^2] = E[X]^2 + Var(X)$

So

$$E[(\hat{Y} - f(x'))^2] = (E[\hat{Y} - f(x')])^2 + Var(\hat{Y} - f(x'))$$
$$= (E[\hat{Y}] - E[f(x')])^2 + Var(\hat{Y}) + Var(\underbrace{f(x')}_{\text{constant}})$$
$$= (E[\hat{Y}] - f(x'))^2 + Var(\hat{Y})$$

∎





LOOCV-exercise