Løsningsforslag, STK4040, uke 44

Bjørn-Helge Mevik

4. november 2005

Oppgave 1

Vi har at

$$\text{MSE}(\hat{y}) = \mathsf{E}[(\hat{y} - Y)^2] \tag{1}$$
$$= \mathsf{E}[\hat{y}^2] - 2\,\mathsf{E}[\hat{y}Y] + \mathsf{E}[Y^2] \tag{2}$$
$$= \mathsf{E}[\hat{y}^2] - 2\,\mathsf{E}\,\hat{y}\,\mathsf{E}\,Y + \mathsf{E}[Y^2], \tag{3}$$

hvor siste overgang gjelder fordi Y og $\hat{y}$ er uavhengige, gitt $\boldsymbol{x}$. Videre er

$$\text{Bias}^2(\hat{y}) = (E[\hat{y} - Y])^2 = (\mathsf{E}\,\hat{y} - \mathsf{E}\,Y)^2 \tag{4}$$
$$= (\mathsf{E}\,\hat{y})^2 - 2\,\mathsf{E}\,\hat{y}\,\mathsf{E}\,Y + (\mathsf{E}\,Y)^2 \tag{5}$$

og

$$\text{Var}(\hat{y} - Y) = \text{Var}(\hat{y}) - 2\,\text{Cov}(\hat{y}, Y) + \text{Var}(Y) \tag{6}$$
$$= \text{Var}(\hat{y}) + \text{Var}(Y) \tag{7}$$
$$= \mathsf{E}[(\hat{y} - \mathsf{E}\,\hat{y})^2] + \mathsf{E}[(Y - \mathsf{E}\,Y)^2] \tag{8}$$
$$= \mathsf{E}[\hat{y}^2] - (\mathsf{E}\,\hat{y})^2 + \mathsf{E}[Y^2] - (\mathsf{E}\,Y)^2. \tag{9}$$

Vi ser da at $\text{MSE}(\hat{y}) = \text{Bias}^2(\hat{y}) + \text{Var}(\hat{y} - Y)$.

Oppgave 2

Det er en trykkfeil i oppgaveteksten: residualene e_i er lik $y_i - \hat{y}_i$, ikke $\hat{y}_i - y_i$.

a)

For å beregne $e_{(i)}$, trenger vi $\boldsymbol{b}_{(i)}$. Vi har at $\boldsymbol{b}_{(i)} = (\mathrm{X}_{(i)}^t \mathrm{X}_{(i)})^{-1} \mathrm{X}_{(i)}^t \boldsymbol{y}_{(i)}$. Første del av dette uttrykket (inversen) kjenner vi fra oppgaveteksten. Siste del er

$$\mathrm{X}_{(i)}^t \boldsymbol{y}_{(i)} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{i-1}, \boldsymbol{x}_{i+1}, \ldots, \boldsymbol{x}_n) \begin{pmatrix} y_1 \\ \vdots \\ y_{i-1} \\ y_{i+1} \\ \vdots \\ y_n \end{pmatrix} \tag{10}$$

$$= \boldsymbol{x}_1 y_1 + \cdots + \boldsymbol{x}_{i-1} y_{i-1} + \boldsymbol{x}_{i+1} y_{i+1} + \cdots, \boldsymbol{x}_n y_n \tag{11}$$

$$= \sum_{j \neq i} \boldsymbol{x}_j y_j = \sum_{j=1}^{n} \boldsymbol{x}_j y_j - \boldsymbol{x}_i y_i \tag{12}$$

$$= \mathrm{X}^t \boldsymbol{y} - \boldsymbol{x}_i y_i, \tag{13}$$

så vi får

$$\boldsymbol{b}_{(i)} = \left((\mathrm{X}^t\mathrm{X})^{-1} + \frac{(\mathrm{X}^t\mathrm{X})^{-1} \boldsymbol{x}_i \boldsymbol{x}_i^t (\mathrm{X}^t\mathrm{X})^{-1}}{1 - h_i} \right) (\mathrm{X}^t \boldsymbol{y} - \boldsymbol{x}_i y_i) \tag{14}$$

$$= (\mathrm{X}^t\mathrm{X})^{-1} \mathrm{X}^t \boldsymbol{y} - (\mathrm{X}^t\mathrm{X})^{-1} \boldsymbol{x}_i y_i \tag{15}$$

$$+ \frac{(\mathrm{X}^t\mathrm{X})^{-1} \boldsymbol{x}_i \boldsymbol{x}_i^t (\mathrm{X}^t\mathrm{X})^{-1} \mathrm{X}^t \boldsymbol{y}}{1 - h_i} - \frac{(\mathrm{X}^t\mathrm{X})^{-1} \boldsymbol{x}_i \boldsymbol{x}_i^t (\mathrm{X}^t\mathrm{X})^{-1} \boldsymbol{x}_i y_i}{1 - h_i}. \tag{16}$$

Nå er $(\mathrm{X}^t\mathrm{X})^{-1} \mathrm{X}^t \boldsymbol{y} = \boldsymbol{b}$ og $\boldsymbol{x}_i^t (\mathrm{X}^t\mathrm{X})^{-1} \boldsymbol{x}_i = h_i$, så

$$\boldsymbol{b}_{(i)} = \boldsymbol{b} - (\mathrm{X}^t\mathrm{X})^{-1} \boldsymbol{x}_i y_i + \frac{(\mathrm{X}^t\mathrm{X})^{-1} \boldsymbol{x}_i \boldsymbol{x}_i^t \boldsymbol{b}}{1 - h_i} - \frac{(\mathrm{X}^t\mathrm{X})^{-1} \boldsymbol{x}_i h_i y_i}{1 - h_i} \tag{17}$$

$$= \boldsymbol{b} - \frac{(1 - h_i)(\mathrm{X}^t\mathrm{X})^{-1} \boldsymbol{x}_i y_i - (\mathrm{X}^t\mathrm{X})^{-1} \boldsymbol{x}_i \boldsymbol{x}_i^t \boldsymbol{b} + h_i (\mathrm{X}^t\mathrm{X})^{-1} \boldsymbol{x}_i y_i}{1 - h_i} \tag{18}$$

$$= \boldsymbol{b} - \frac{(\mathrm{X}^t\mathrm{X})^{-1} \boldsymbol{x}_i y_i - (\mathrm{X}^t\mathrm{X})^{-1} \boldsymbol{x}_i \boldsymbol{x}_i^t \boldsymbol{b}}{1 - h_i} \tag{19}$$

$$= \boldsymbol{b} - \frac{(\mathrm{X}^t\mathrm{X})^{-1} \boldsymbol{x}_i (y_i - \boldsymbol{x}_i^t \boldsymbol{b})}{1 - h_i} \tag{20}$$

$$= \boldsymbol{b} - \frac{(\mathrm{X}^t\mathrm{X})^{-1} \boldsymbol{x}_i}{1 - h_i} e_i. \tag{21}$$

$$\tag{22}$$

Dermed blir

$$e_{(i)} = y_i - \boldsymbol{x}_i^t \boldsymbol{b}_{(i)} \tag{23}$$

$$= y_i - \boldsymbol{x}_i^t \boldsymbol{b} + \frac{\boldsymbol{x}_i^t (\mathrm{X}^t\mathrm{X})^{-1} \boldsymbol{x}_i}{1 - h_i} e_i \tag{24}$$

$$= e_i + \frac{h_i}{1 - h_i} e_i \tag{25}$$

$$= \frac{1}{1 - h_i} e_i, \tag{26}$$

hvilket skulle bevises.

b)

Denne oppgaven var kanskje litt uklart formulert. Poenget var å vise at $h_i = 1/n + (\boldsymbol{x}_i - \bar{\boldsymbol{x}})^t (\mathrm{X}_c^t \mathrm{X}_c)^{-1} (\boldsymbol{x}_i - \bar{\boldsymbol{x}})$, hvor $\bar{\boldsymbol{x}}$ er kolonnesnittene til X_1. $g_i = h_i - 1/n$ kan dermed tolkes som leverage'ne til en regresjon med sentrerte variabler.

Siden vi har separert ut konstantleddet fra $\boldsymbol{x}$, så er

$$h_i = \begin{pmatrix} 1 & \boldsymbol{x}_i \end{pmatrix} (\mathrm{X}^t\mathrm{X})^{-1} \begin{pmatrix} 1 \\ \boldsymbol{x}_i \end{pmatrix}. \tag{27}$$

Vi har at

$$\mathrm{X}^t\mathrm{X} = \begin{pmatrix} \boldsymbol{1}^t \\ \mathrm{X}_1^t \end{pmatrix} \begin{pmatrix} \boldsymbol{1} & \mathrm{X}_1 \end{pmatrix} = \begin{pmatrix} n & \boldsymbol{1}^t \mathrm{X}_1 \\ \mathrm{X}_1^t \boldsymbol{1} & \mathrm{X}_1^t \mathrm{X}_1 \end{pmatrix} = \begin{pmatrix} n & n\bar{\boldsymbol{x}}^t \\ n\bar{\boldsymbol{x}} & \mathrm{X}_1^t \mathrm{X}_1 \end{pmatrix}. \tag{28}$$

Hed hjelp av A.2.4(VII) får vi da at

$$(\mathrm{X}^t\mathrm{X})^{-1} = \begin{pmatrix} \frac{1}{n} & \boldsymbol{0}^t \\ \boldsymbol{0} & 0 \end{pmatrix} + \begin{pmatrix} -\frac{1}{n} n\bar{\boldsymbol{x}}^t \\ \mathrm{I} \end{pmatrix} \left(\mathrm{X}_1^t \mathrm{X}_1 - n\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}^t \right)^{-1} \begin{pmatrix} -\frac{1}{n} n\bar{\boldsymbol{x}} & \mathrm{I} \end{pmatrix}. \tag{29}$$

Det er lett å vise at $\mathrm{X}_1^t \mathrm{X}_1 - n\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}^t = \mathrm{X}_c^t \mathrm{X}_c$, så

$$(\mathrm{X}^t\mathrm{X})^{-1} = \begin{pmatrix} \frac{1}{n} & \boldsymbol{0}^t \\ \boldsymbol{0} & 0 \end{pmatrix} + \begin{pmatrix} -\bar{\boldsymbol{x}}^t \\ \mathrm{I} \end{pmatrix} (\mathrm{X}_c^t \mathrm{X}_c)^{-1} \begin{pmatrix} -\bar{\boldsymbol{x}} & \mathrm{I} \end{pmatrix} \tag{30}$$

$$= \begin{pmatrix} \frac{1}{n} & \boldsymbol{0}^t \\ \boldsymbol{0} & 0 \end{pmatrix} + \begin{pmatrix} -\bar{\boldsymbol{x}}^t (\mathrm{X}_c^t \mathrm{X}_c)^{-1} \\ (\mathrm{X}_c^t \mathrm{X}_c)^{-1} \end{pmatrix} \begin{pmatrix} -\bar{\boldsymbol{x}} & \mathrm{I} \end{pmatrix} \tag{31}$$

$$= \begin{pmatrix} \frac{1}{n} + \bar{\boldsymbol{x}}^t (\mathrm{X}_c^t \mathrm{X}_c)^{-1} \bar{\boldsymbol{x}} & -\bar{\boldsymbol{x}}^t (\mathrm{X}_c^t \mathrm{X}_c)^{-1} \\ -(\mathrm{X}_c^t \mathrm{X}_c)^{-1} \bar{\boldsymbol{x}} & (\mathrm{X}_c^t \mathrm{X}_c)^{-1} \end{pmatrix}. \tag{32}$$

Dette gir

$$h_i = \begin{pmatrix} 1 & \boldsymbol{x}_i \end{pmatrix} (\mathrm{X}^t\mathrm{X})^{-1} \begin{pmatrix} 1 \\ \boldsymbol{x}_i \end{pmatrix} \tag{33}$$

$$= \frac{1}{n} + \bar{\boldsymbol{x}}^t(\mathrm{X}_c^t\mathrm{X}_c)^{-1}\bar{\boldsymbol{x}} - \boldsymbol{x}_i^t(\mathrm{X}_c^t\mathrm{X}_c)^{-1}\bar{\boldsymbol{x}} - \bar{\boldsymbol{x}}^t(\mathrm{X}_c^t\mathrm{X}_c)^{-1}\boldsymbol{x}_i + \boldsymbol{x}_i^t(\mathrm{X}_c^t\mathrm{X}_c)^{-1}\boldsymbol{x}_i \tag{34}$$

$$= \frac{1}{n} - (\boldsymbol{x}_i - \bar{\boldsymbol{x}})^t(\mathrm{X}_c^t\mathrm{X}_c)^{-1}\bar{\boldsymbol{x}} + (\boldsymbol{x}_i - \bar{\boldsymbol{x}})^t(\mathrm{X}_c^t\mathrm{X}_c)^{-1}\boldsymbol{x}_i \tag{35}$$

$$= \frac{1}{n} + (\boldsymbol{x}_i - \bar{\boldsymbol{x}})^t(\mathrm{X}_c^t\mathrm{X}_c)^{-1}(\boldsymbol{x}_i - \bar{\boldsymbol{x}}). \tag{36}$$