Solution hints, STK4040, week 45

Bjørn-Helge Mevik

November 5, 2007

Exercise 1

We have that

$$\begin{aligned} \mathrm{MSE}(\hat{y}) &= \mathsf{E}[(\hat{y} - Y)^2] && (1) \\ &= \mathsf{E}[\hat{y}^2] - 2\,\mathsf{E}[\hat{y}Y] + \mathsf{E}[Y^2] && (2) \\ &= \mathsf{E}[\hat{y}^2] - 2\,\mathsf{E}\,\hat{y}\,\mathsf{E}\,Y + \mathsf{E}[Y^2], && (3) \end{aligned}$$

where the last equality is true because Y and $\hat{y}$ are independent, given $\boldsymbol{x}$. Further,

$$\begin{aligned} \mathrm{Bias}^2(\hat{y}) &= (E[\hat{y} - Y])^2 = (\mathsf{E}\,\hat{y} - \mathsf{E}\,Y)^2 && (4) \\ &= (\mathsf{E}\,\hat{y})^2 - 2\,\mathsf{E}\,\hat{y}\,\mathsf{E}\,Y + (\mathsf{E}\,Y)^2 && (5) \end{aligned}$$

and

$$\begin{aligned} \mathrm{Var}(\hat{y} - Y) &= \mathrm{Var}(\hat{y}) - 2\,\mathrm{Cov}(\hat{y}, Y) + \mathrm{Var}(Y) && (6) \\ &= \mathrm{Var}(\hat{y}) + \mathrm{Var}(Y) && (7) \\ &= \mathsf{E}[(\hat{y} - \mathsf{E}\,\hat{y})^2] + \mathsf{E}[(Y - \mathsf{E}\,Y)^2] && (8) \\ &= \mathsf{E}[\hat{y}^2] - (\mathsf{E}\,\hat{y})^2 + \mathsf{E}[Y^2] - (\mathsf{E}\,Y)^2. && (9) \end{aligned}$$

We then see that $\mathrm{MSE}(\hat{y}) = \mathrm{Bias}^2(\hat{y}) + \mathrm{Var}(\hat{y} - Y)$.

Exercise 2a)

To calculate $e_{(i)}$, we need $\boldsymbol{b}_{(i)}$. We have that $\boldsymbol{b}_{(i)} = (\mathrm{X}_{(i)}^T \mathrm{X}_{(i)})^{-1} \mathrm{X}_{(i)}^T \boldsymbol{y}_{(i)}$. The first part of this exression (the inverse) is given in the exercise. The last part is

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

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

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

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

so we get

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

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

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

Now $(\mathrm{X}^T\mathrm{X})^{-1}\mathrm{X}^T\boldsymbol{y} = \boldsymbol{b}$ and $\boldsymbol{x}_i^T(\mathrm{X}^T\mathrm{X})^{-1}\boldsymbol{x}_i = h_i$, so

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

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

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

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

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

$$\tag{22}$$

Therefore

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

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

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

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

which is what we wanted.

Exercise 2b)

The point is to show that $h_i = 1/n + (\boldsymbol{x}_i - \bar{\boldsymbol{x}})^T(\mathrm{X}_c^T\mathrm{X}_c)^{-1}(\boldsymbol{x}_i - \bar{\boldsymbol{x}})$, where $\bar{\boldsymbol{x}}$ is the coloumn means of X_1. $g_i = h_i - 1/n$ can thereby be interpreted as the leverages of a regression with centered variables.

Writing the observations of the uncentered model as $\begin{pmatrix} 1 & \boldsymbol{x}_i \end{pmatrix}$, we get

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

We have that

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

By the help of A.2.4(VII) we then get that

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

It is easy to show that $\mathrm{X}_1^T\mathrm{X}_1 - n\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}^T = \mathrm{X}_c^T\mathrm{X}_c$, so

$$\begin{aligned}
(\mathrm{X}^T\mathrm{X})^{-1} &= \begin{pmatrix} \frac{1}{n} & \mathbf{0}^T \\ \mathbf{0} & 0 \end{pmatrix} + \begin{pmatrix} -\bar{\boldsymbol{x}}^T \\ \mathrm{I} \end{pmatrix} (\mathrm{X}_c^T\mathrm{X}_c)^{-1} \begin{pmatrix} -\bar{\boldsymbol{x}} & \mathrm{I} \end{pmatrix} & (30) \\
&= \begin{pmatrix} \frac{1}{n} & \mathbf{0}^T \\ \mathbf{0} & 0 \end{pmatrix} + \begin{pmatrix} -\bar{\boldsymbol{x}}^T(\mathrm{X}_c^T\mathrm{X}_c)^{-1} \\ (\mathrm{X}_c^T\mathrm{X}_c)^{-1} \end{pmatrix} \begin{pmatrix} -\bar{\boldsymbol{x}} & \mathrm{I} \end{pmatrix} & (31) \\
&= \begin{pmatrix} \frac{1}{n} + \bar{\boldsymbol{x}}^T(\mathrm{X}_c^T\mathrm{X}_c)^{-1}\bar{\boldsymbol{x}} & -\bar{\boldsymbol{x}}^T(\mathrm{X}_c^T\mathrm{X}_c)^{-1} \\ -(\mathrm{X}_c^T\mathrm{X}_c)^{-1}\bar{\boldsymbol{x}} & (\mathrm{X}_c^T\mathrm{X}_c)^{-1} \end{pmatrix}. & (32)
\end{aligned}$$

This gives

$$\begin{aligned}
h_i &= \begin{pmatrix} 1 & \boldsymbol{x}_i \end{pmatrix} (\mathrm{X}^T\mathrm{X})^{-1} \begin{pmatrix} 1 \\ \boldsymbol{x}_i \end{pmatrix} & (33) \\
&= \frac{1}{n} + \bar{\boldsymbol{x}}^T(\mathrm{X}_c^T\mathrm{X}_c)^{-1}\bar{\boldsymbol{x}} - \boldsymbol{x}_i^T(\mathrm{X}_c^T\mathrm{X}_c)^{-1}\bar{\boldsymbol{x}} - \bar{\boldsymbol{x}}^T(\mathrm{X}_c^T\mathrm{X}_c)^{-1}\boldsymbol{x}_i + \boldsymbol{x}_i^T(\mathrm{X}_c^T\mathrm{X}_c)^{-1}\boldsymbol{x}_i & (34) \\
&= \frac{1}{n} - (\boldsymbol{x}_i - \bar{\boldsymbol{x}})^T(\mathrm{X}_c^T\mathrm{X}_c)^{-1}\bar{\boldsymbol{x}} + (\boldsymbol{x}_i - \bar{\boldsymbol{x}})^T(\mathrm{X}_c^T\mathrm{X}_c)^{-1}\boldsymbol{x}_i & (35) \\
&= \frac{1}{n} + (\boldsymbol{x}_i - \bar{\boldsymbol{x}})^T(\mathrm{X}_c^T\mathrm{X}_c)^{-1}(\boldsymbol{x}_i - \bar{\boldsymbol{x}}). & (36)
\end{aligned}$$

∎