620-371: Linear Models

Practice Class 4

24th March, 2009

This practice class will use a single example. We model an individual's income at age 30 against the number of years of formal education (with a linear model). The following data is collected:

Years of formal education (x)	Income (\$k) (y)
8	8
12	15
14	16
16	20
16	25
20	40

1. Write down the linear model in matrix form.

   Solution: $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, where

$$\mathbf{y} = \begin{bmatrix} 8 \\ 15 \\ 16 \\ 20 \\ 25 \\ 40 \end{bmatrix}, X = \begin{bmatrix} 1 & 8 \\ 1 & 12 \\ 1 & 14 \\ 1 & 16 \\ 1 & 16 \\ 1 & 20 \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{bmatrix}.$$

2. Find the normal equations for this model.

   Solution:

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 8 & 12 & 14 & 16 & 16 & 20 \end{bmatrix} \begin{bmatrix} 1 & 8 \\ 1 & 12 \\ 1 & 14 \\ 1 & 16 \\ 1 & 16 \\ 1 & 20 \end{bmatrix} = \begin{bmatrix} 6 & 86 \\ 86 & 1316 \end{bmatrix}.$$

$$X^T \mathbf{y} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 8 & 12 & 14 & 16 & 16 & 20 \end{bmatrix} \begin{bmatrix} 8 \\ 15 \\ 16 \\ 20 \\ 25 \\ 40 \end{bmatrix} = \begin{bmatrix} 124 \\ 1988 \end{bmatrix}.$$

The normal equations are

$$\begin{bmatrix} 6 & 86 \\ 86 & 1316 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 124 \\ 1988 \end{bmatrix}.$$

3. Estimate the parameters using the least squares method.

   Solution:

$$\begin{aligned} (X^T X)^{-1} &= \frac{1}{6 \times 1316 - 86^2} \begin{bmatrix} 1316 & -86 \\ -86 & 6 \end{bmatrix} = \frac{1}{250} \begin{bmatrix} 658 & -43 \\ -43 & 3 \end{bmatrix}. \\ \mathbf{b} &= \frac{1}{250} \begin{bmatrix} 658 & -43 \\ -43 & 3 \end{bmatrix} \begin{bmatrix} 124 \\ 1988 \end{bmatrix} = \frac{1}{125} \begin{bmatrix} -1946 \\ 316 \end{bmatrix}. \end{aligned}$$

4. This model is a simple linear regression model. Use the standard linear regression formulae,

$$b_1 = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2}, b_0 = \bar{y} - b_1 \bar{x},$$

   to estimate the parameters again (where the bar indicates the mean). Check that you have the same answers as in (c).

   Solution:

$$\begin{aligned} \overline{xy} &= \frac{994}{3} \\ \overline{x^2} &= \frac{658}{3} \\ \bar{y} &= \frac{62}{3} \\ \bar{x} &= \frac{43}{3} \end{aligned}$$

   This gives

$$b_1 = \frac{316}{125}, b_0 = -\frac{1946}{125}.$$

5. Estimate the income of a person who has had 18 years of formal education.

   Solution: $\begin{bmatrix} 1 & 18 \end{bmatrix} \frac{1}{125} \begin{bmatrix} -1946 \\ 316 \end{bmatrix} = \frac{3742}{125} = 29.9.$

6. We know that the least squares estimator $\mathbf{b}$ is an unbiased estimator for $\boldsymbol{\beta}$. Show that $\mathbf{t}^T\mathbf{b}$ is an unbiased estimator for $\mathbf{t}^T\boldsymbol{\beta}$, where $\mathbf{t}$ is a vector of constants.

   Solution: By assumption $E[\mathbf{b}] = \boldsymbol{\beta}$. Therefore $E[\mathbf{t}^T\mathbf{b}] = \mathbf{t}^T E[\mathbf{b}] = \mathbf{t}^T\boldsymbol{\beta}$, and $\mathbf{t}^T\mathbf{b}$ is an unbiased estimator for $\mathbf{t}^T\boldsymbol{\beta}$.