STAT W4201 001, Homework 7

Brian Weinstein (bmw2148)

Mar 28, 2016

Code is attached here and also posted at https://github.com/BrianWeinstein/advanced-data-analysis. Where relevant, code snippets and output are are included in-line.

Problem 1: Ramsey 10.21

Let

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1p} \\ 1 & X_{21} & X_{22} & \cdots & X_{2p} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{np} \end{bmatrix}, \quad \text{and} \;\; \mathbf{b} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}.$$

By deconstructing $(\mathbf{X}^T\mathbf{X})\mathbf{b} = \mathbf{X}^T\mathbf{Y}$ we find the set of normal equations:

$$(\mathbf{X}^T\mathbf{X})\mathbf{b} = \mathbf{X}^T\mathbf{Y}$$

$$\Rightarrow \left(\begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_{11} & X_{21} & \cdots & X_{n1} \\ X_{12} & X_{22} & \cdots & X_{n2} \\ \vdots & \vdots & & \vdots \\ X_{1p} & X_{2p} & \cdots & X_{np} \end{bmatrix} \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1p} \\ 1 & X_{21} & X_{22} & \cdots & X_{2p} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{np} \end{bmatrix} \right) \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_{11} & X_{21} & \cdots & X_{n1} \\ X_{12} & X_{22} & \cdots & X_{n2} \\ \vdots & \vdots & & \vdots \\ X_{1p} & X_{2p} & \cdots & X_{np} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} n & \sum X_{i1} & \sum X_{i2} & \cdots & \sum X_{ip} \\ \sum X_{i1} & \sum X_{i1}^2 & \sum X_{i1}X_{i2} & \cdots & \sum X_{i1}X_{ip} \\ \sum X_{i2} & \sum X_{i2}X_{i1} & \sum X_{i2}^2 & \cdots & \sum X_{i2}X_{ip} \\ \vdots & \vdots & \vdots & & \vdots \\ \sum X_{ip} & \sum X_{ip}X_{i1} & \sum X_{ip}X_{i2} & \cdots & \sum X_{ip}^2 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} = \begin{bmatrix} \sum Y_i \\ \sum X_{i1}Y_i \\ \sum X_{i2}Y_i \\ \vdots \\ \sum X_{ip}Y_i \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} \beta_0 n + \beta_1 \sum X_{i1} + \beta_2 \sum X_{i2} + \cdots + \beta_p \sum X_{ip} \\ \beta_0 \sum X_{i1} + \beta_1 \sum X_{i1}^2 + \beta_2 \sum X_{i1}X_{i2} + \cdots + \beta_p \sum X_{i1}X_{ip} \\ \beta_0 \sum X_{i2} + \beta_1 \sum X_{i2}X_{i1} + \beta_2 \sum X_{i2}^2 + \cdots + \beta_p \sum X_{i2}X_{ip} \\ \vdots \\ \beta_0 \sum X_{ip} + \beta_1 \sum X_{ip}X_{i1} + \beta_2 \sum X_{ip}X_{i2} + \cdots + \beta_p \sum X_{ip}^2 \end{bmatrix} = \begin{bmatrix} \sum Y_i \\ \sum X_{i1}Y_i \\ \sum X_{i2}Y_i \\ \vdots \\ \sum X_{ip}Y_i \end{bmatrix},$$

where each $\sum$ indicates summation over all cases $(i = 1, 2, \ldots, n)$.

As long as $\mathbf{X}^T\mathbf{X}$ is invertible, the least squares solution is $\mathbf{b} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$. By the invertible matrix theorem, $\mathbf{X}^T\mathbf{X}$ is invertible whenever the columns of $\mathbf{X}$ are linearly independent.

Problem 2: Ramsey 10.22

Problem 3: Ramsey 10.26

Problem 4: Ramsey 11.8

Problem 5: Ramsey 11.16

Problem 6: Ramsey 11.20