Homework Assignment 1 - MGSC 695

Lakshya Agarwal

January 9, 2024

1 Problem 1

1.1 Part 1

The dimensions of each matrix are as follows:

- Y is a $n \times 1$ matrix
- X is a $n \times (k+1)$ matrix
- β is a $(k+1) \times 1$ matrix
- ϵ is a $n \times 1$ matrix

Therefore, the dimensions of Equation (1) are as follows:

$$X_{n\times(k+1)}\beta_{(k+1)\times 1} + \epsilon_{n\times 1} = Y_{n\times 1}$$

It is clear that the left-hand side of the above equation results in a $n \times 1$ matrix, which is the same as the dimensions of Y. Therefore, the dimensions of the matrices on both sides of Equation (1) match.

1.2 Part 2

$$\begin{aligned}
Y &= X\beta + \epsilon \\
\epsilon &= Y - X\beta \\
\epsilon &= \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} - \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_k \end{bmatrix}
\end{aligned}$$

The i^{th} element of the ϵ matrix is given by:

$$\epsilon_{ij} = y_i - \sum_{l=0}^{(k+1)} x_{il}\beta_l$$

Then, the sum of squared residuals is given by:

$$\sum_{i=1}^{n} \epsilon_{ij}^2 = \sum_{i=1}^{n} (y_i - \sum_{l=0}^{(k+1)} x_{il}\beta_l)^2$$

In matrix notation, this can be written as:

$$\begin{aligned}
MSE &= \epsilon^T \epsilon \\
&= (Y - X\beta)^T (Y - X\beta)
\end{aligned}$$

Here, $\epsilon^T \epsilon$ is a dot product of two vectors, and therefore, it is a scalar value.

1.3 Part 3

To calculate the partial derivative of the sum of squared residuals with respect to β, we first expand the matrix multiplication in the above equation:

$$\begin{aligned} MSE &= (Y^T - \beta^T X^T)(Y - X\beta) \\ &= Y^T Y - Y^T X\beta - \beta^T X^T Y + \beta^T X^T X\beta \end{aligned}$$

Since $Y^T X\beta$ is a scalar, it is equal to its transpose. Therefore, $Y^T X\beta = \beta^T X^T Y$.

$$MSE = Y^T Y - 2\beta^T X^T Y + \beta^T X^T X\beta$$

Then, the partial derivative of the sum of squared residuals with respect to β is given by:

$$\frac{\partial MSE}{\partial \beta} = -2X^T Y + 2X^T X\beta$$

Setting the above equation to zero, we get:

$$\begin{aligned} -2X^T Y + 2X^T X\beta^* &= 0 \\ X^T Y &= X^T X\beta^* \\ (X^T X)\beta^* &= X^T Y \end{aligned}$$

1.4 Part 4

Since $X^T X$ is a square matrix of size $(k+1)$, it is invertible. Therefore, we can multiply both sides of the above equation by $(X^T X)^{-1}$ to get:

$$(X^T X)^{-1}(X^T X)\beta^* = (X^T X)^{-1} X^T Y$$

Since $A^{-1}A = I$, where I is the identity matrix, we get:

$$\begin{aligned} I\beta^* &= (X^T X)^{-1} X^T Y \\ \beta^* &= (X^T X)^{-1}(X^T Y) \end{aligned}$$

2 Problem 2

Using the `advertising.csv` dataset, the parameter estimates obtained from Equation (2) and using the `scikit-learn` library are as follows:

$$\begin{aligned} \text{Matrix algebra} &: \begin{bmatrix} 2.939 & 0.046 & 0.189 & -0.001 \end{bmatrix} \\ \text{scikit-learn} &: \begin{bmatrix} 2.939 & 0.046 & 0.189 & -0.001 \end{bmatrix} \end{aligned}$$

The parameter estimates obtained from the procedures are the same. Please run the attached `python` script to verify the results.