Multiple Regression Analysis in Matrix Form

Define three vectors (**y**, **b** and **e**) and a matrix (**X**) as follows:

$$\underset{(N\times 1)}{\mathbf{y}} = \begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_N \end{bmatrix}, \quad \underset{(N\times(p+1))}{\mathbf{X}} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & x_{N1} & x_{N2} & \cdots & x_{Np} \end{bmatrix}, \quad \underset{((p+1)\times 1)}{\mathbf{b}} = \begin{bmatrix} b_0 \\ b_1 \\ \cdots \\ b_p \end{bmatrix} \text{ and } \underset{(N\times 1)}{\mathbf{e}} = \begin{bmatrix} e_1 \\ e_2 \\ \cdots \\ e_N \end{bmatrix}.$$

The standard multiple (linear) regression equation with p predictor variables and N observations

$$y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \ldots + b_p x_{ip} + e_i, \text{ where } i = 1, \ldots, N,$$

can then be rewritten using these vectors and the matrix as

$$\underset{(N\times 1)}{\mathbf{y}} = \underset{(N\times(p+1))}{\mathbf{X}} \underset{((p+1)\times 1)}{\mathbf{b}} + \underset{(N\times 1)}{\mathbf{e}},$$

or, in a more compact way that still makes explicit the dimensions of the vectors and the matrix, (and also illustrates their conformability for matrix multiplication)

$${}_N\mathbf{y}_1 = {}_N\mathbf{X}_{p+1}\mathbf{b}_1 + {}_N\mathbf{e}_1,$$

or even more compactly (dropping the explicit indication of the size of the vectors and matrix) as

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}.$$

The optimization problem in regression analysis is to minimize the residual sum of squares S:

$$\begin{aligned} \text{Min } S &= \sum_{i=1}^{n} e_i^2 = \mathbf{e}'\mathbf{e} = (\mathbf{y} - \mathbf{X}\mathbf{b})'(\mathbf{y} - \mathbf{X}\mathbf{b}) \\ &= \mathbf{y}'\mathbf{y} - \mathbf{b}'\mathbf{X}'\mathbf{y} - \mathbf{y}'\mathbf{X}\mathbf{b} + \mathbf{b}'\mathbf{X}'\mathbf{X}\mathbf{b} \\ &= \mathbf{y}'\mathbf{y} - 2\mathbf{b}'\mathbf{X}'\mathbf{y} + \mathbf{b}'\mathbf{X}'\mathbf{b}. \end{aligned}$$

S can be minimized by setting the following partial derivative to 0:

$$\frac{\partial S}{\partial \mathbf{b}} = -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\mathbf{b} = 0.$$

Rearranging, $\mathbf{X}'\mathbf{X}\mathbf{b} = \mathbf{X}'\mathbf{y}$, and by multiplying both sides of this equation by $(\mathbf{X}'\mathbf{X})^{-1}$, i.e., $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$, the regression coefficients, $\mathbf{b}$, can be obtained as follows:

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}.$$