Ordinary Least Squares

For Simple Regression

- We have the Residual Sum of Squares and we use the OLS to find the Coefficients β_0, β_1 estimates $\hat{\beta}_0, \hat{\beta}_1$ to minimize The RSS
- Computing Partial Derivatives
- Finding the estimates for $\hat{\beta}_0$

$$\frac{\partial RSS}{\partial \hat{\beta}_0} = \sum (y_i - \hat{y})^2 = 2 \sum (y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)) = 0$$
$$\frac{\partial RSS}{\partial \hat{\beta}_0} = \sum y_i - \hat{\beta}_0 \sum 1 - \hat{\beta}_1 \sum x_i = 0$$

With :

$$\sum_{i=1}^{n} y_i = n\bar{y} \;\; \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$$

- $\bar{y}$ is the mean on the y axis (same goes for $\bar{x}$)

$$\bar{y}n - n\hat{\beta}_0 - \hat{\beta}_1 \bar{x} n = 0$$
$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

Now for $\hat{\beta}_1$:

$$\frac{\partial \text{RSS}}{\partial \hat{\beta}_1} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = -2 \sum_{i=1}^{n} (y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)) x_i = 0$$
$$\sum_{i=1}^{n} x_i \left(\sum_{i=1}^{n} y_i - \hat{\beta}_0 \sum_{i=1}^{n} 1 - \hat{\beta}_1 \sum_{i=1}^{n} x_i \right) = 0$$
$$\sum_{i=1}^{n} y_i x_i - \hat{\beta}_0 \sum_{i=1}^{n} x_i - \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 = 0$$

- Now we substitute $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$ in the equation
- And get $\sum y_i x_i$ to the left side of the equation
- Divide both sides with $\sum x_i$
- Factor by $\hat{\beta}_1$

$$\sum_{i=1}^{n} y_i x_i - (\bar{y} - \hat{\beta}_1 \bar{x}) \sum x_i - \hat{\beta}_1 \sum x_i^2 = 0$$
$$\bar{y} \sum x_i - \hat{\beta}_1 \bar{x} \sum x_i + \hat{\beta}_1 \sum x_i^2 = \sum_{i=1}^{n} y \sum x_i$$
$$\bar{y} - \hat{\beta}_1 \bar{x} + \hat{\beta}_1 \sum x_i = \sum y_i$$
$$\hat{\beta}_1 \left(\sum_{i=1}^{n} x_i - \bar{x} \right) = \sum_{i=1}^{n} y_i - \bar{y}$$
$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} yi - \bar{y}}{\sum_{i=1}^{n} x_i - \bar{x}}$$

- It can also be written as : (For ease of calculation)

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

For Multiple Regression :

Recall Multiple Regression can be written as: (matrix form)

$$Y = X\beta + \varepsilon$$

- Goal : Finding The Coefficients to **minimize** the Residual Sum of Squares as much as possible
- $e = Y - \hat{Y}$

$$\hat{Y} = X\hat{\beta} + \varepsilon$$

The Residual Sum of Squares **RSS** :

$$\min \sum_{i=1}^{n} e_i^2 = \min e^T e = \min\, (y-\hat{y})^T(y-\hat{y}) = (y-x\hat{\beta})^T(y-x\hat{\beta})$$

Note : $x^2 = x^T x$

- Partial Derivation for $\hat{\beta}$:

$$\frac{\partial RSS}{\partial \hat{\beta}} = (y-x\hat{\beta})^T(y-x\hat{\beta}) = 0$$

$$\frac{\partial RSS}{\partial \hat{\beta}} = (y^T-\hat{\beta}^T x^T)(y-x\hat{\beta}) = y^T y - y^T x\hat{\beta} - \hat{\beta}^T x^T y + \hat{\beta}^T x^T Y x\hat{\beta} = 0$$

Note :

$$y^T x\hat{\beta} = (1\times p)(n\times p)(p\times 1) = 1$$

$$\hat{\beta}^T x^T y = (1\times p)(p\times n)(n\times 1) = 1$$

Now we derive with respect to $\hat{\beta}$

$$\frac{\partial RSS}{\partial \hat{\beta}} = y^t y - 2\hat{\beta}^T x^T y + \hat{\beta}^T x^T x\hat{\beta}$$

$$\frac{\partial}{\partial \hat{\beta}} = -2x^T y + 2x^T x\hat{\beta} = 0$$

$$x^T x\hat{\beta} = x^T y$$

Note : If symmetric

$$(x^T x) = (x^T x)^T$$

Adding $(x^T x)^{-1}$ on both sides

$$(x^T x)^{-1}(x^T x)\hat{\beta} = (x^T x)^{-1} x^T y$$

The Least Squares Estimator Coefficient (β Estimate)

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