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 $\hat{\beta}_0$

$$egin{aligned} rac{\partial RSS}{\partial \hat{eta}_0} &= \sum (y_i - \hat{y})^2 = 2 \sum (y_i - (\hat{eta}_0 + \hat{eta}_1 x_i)) = 0 \ &rac{\partial RSS}{\partial \hat{eta}_0} &= \sum y_i - \hat{eta}_0 \sum 1 - \hat{eta}_1 \sum x_i = 0 \end{aligned}$$

With:

$$\sum_{i=1}^n y_i = n ar{y} \ \ ar{y} = rac{1}{n} \sum_{i=1}^n y_i$$

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

$$ar{y}n-n\hat{eta}_0-\hat{eta}_1ar{x}n=0 \ \hat{eta}_0=ar{y}-\hat{eta}_1ar{x}$$

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

$$egin{split} rac{\partial ext{RSS}}{\partial \hat{eta}_1} &= \sum_{i=1}^n (y_i - \hat{y}_i)^2 = -2 \sum_{i=1}^n (y_i - (\hat{eta}_0 + \hat{eta}_1 x_i)) x_i = 0 \ \sum_{i=1}^n x_i \left(\sum_{i=1}^n y_i - \hat{eta}_0 \sum_{i=1}^n 1 - \hat{eta}_1 \sum_{i=1}^n x_i
ight) = 0 \ \sum_{i=1}^n y_i x_i - \hat{eta}_0 \sum_{i=1}^n x_i - \hat{eta}_1 \sum_{i=1}^n x_i^2 = 0 \end{split}$$

- ullet Now we substitute $\hat{eta}_0 = ar{y} \hat{eta}_1ar{x}$ 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 $\hat{\beta}_1$

$$egin{aligned} \sum_{i=1}^n y_i x_i - (ar{y} - \hat{eta}_1 ar{x}) \sum x_i - \hat{eta}_1 \sum x_i^2 &= 0 \ ar{y} \sum x_i - \hat{eta}_1 ar{x} \sum x_i + \hat{eta}_1 \sum x_i^2 &= \sum_{i=1}^n y \sum x_i \ ar{y} - \hat{eta}_1 ar{x} + \hat{eta}_1 \sum x_i &= \sum y_i \ \hat{eta}_1 \left(\sum_{i=1}^n x_i - ar{x}
ight) &= \sum_{i=1}^n y_i - ar{y} \ \hat{eta}_1 &= rac{\sum_{i=1}^n y_i - ar{y}}{\sum_{i=1}^n x_i - ar{x}} \end{aligned}$$

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

$$\hat{eta}_1 = rac{\sum_{i=1}^n (x_i - ar{x})(y_i - ar{y})}{\sum_{i=1}^n (x_i - ar{x})^2}$$

For Multiple Regression:

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

$$Y=Xeta+arepsilon$$

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

$$\hat{Y} = X\hat{eta} + arepsilon$$

The Residual Sum of Squares **RSS**:

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

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

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

$$egin{aligned} rac{\partial RSS}{\partial \hat{eta}} &= (y-x\hat{eta})^T(y-x\hat{eta}) = 0 \ & rac{\partial RSS}{\partial \hat{eta}} &= (y^T-\hat{eta}^Tx^T)(y-x\hat{eta}) = y^Ty-y^Tx\hat{eta}-\hat{eta}^Tx^Ty+\hat{eta}^Tx^TYx\hat{eta} = 0 \end{aligned}$$

Note:

$$egin{aligned} y^Tx\hat{eta}&=(1 imes p)(n imes p)(p imes 1)=1\ \hat{eta}^Tx^Ty&=(1 imes p)(p imes n)(n imes 1)=1 \end{aligned}$$

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

$$egin{align} rac{\partial RSS}{\partial \hat{eta}} &= y^t y - 2\hat{eta}^T x^T y + \hat{eta}^T x^T x \hat{eta} \ & rac{\partial}{\partial \hat{eta}} &= -2 x^T y + 2 x^T x \hat{eta} = 0 \ & x^T x \hat{eta} &= x^T y \ \end{pmatrix}$$

Note: If symmetric

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

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

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

The Least Squares Estimator Coefficient (β Estimate)

$$\hat{eta} = (X^TX)^{-1}X^TY$$