Calculating coefficients using partial derivatives

Let us start with the basic regression line equation

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \qquad \text{---- (1)}$$

Where,

y_i is the actual value of dependent variable
x_i is the value of independent variable
β_0 is the intercept (value of y when x = 0)
β_1 is the slope of the line
ε_i is the residuals or errors

To find the best fit line, we minimize the squared sum of errors.

$$SSE = \sum_{i=1}^{n} \left(y_i - (\beta_0 + \beta_1 x_i)\right)^2 \qquad \text{---- (2)}$$

Minimize SSE with respect to β_0 and β_1. Hence, we will take the partial derivative w.r.t. β_0 and β_1 and set them to zero since the minimum value of error can be zero.

The question arises why we are using partial derivatives and not regular differentiation. Regular differentiation works when there's only one variable. In case of linear regression, the objective function depends on two (or more) variables (β_0, $\beta_1, \beta_2, \dots, \beta_n$). Partial derivatives allow us to analyze the effect of **each variable independently** while holding others constant.

Hence, the partial derivatives will let us minimize the function w.r.t. one variable at a time, holding the other variable constant.

<u>Step 1:</u> Partial derivative w.r.t. $\boldsymbol{\beta_0}$

$$\frac{\partial SSE}{\partial \beta_0} = \frac{\partial}{\partial \beta_0}\left(\sum_{i=1}^{n} \left(y_i - (\beta_0 + \beta_1 x_i)\right)^2\right)$$

$$= -2\sum_{i=1}^{n}\left(y_i - (\beta_0 + \beta_1 x_i)\right)$$

Setting $\frac{\partial SSE}{\partial \beta_0} = 0$

$$-2\sum_{i=1}^{n}\left(y_i - (\beta_0 + \beta_1 x_i)\right) = 0$$

$$=> \sum_{i=1}^{n}\left(y_i - (\beta_0 + \beta_1 x_i)\right) = 0$$

$$=> \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} \beta_0 - \sum_{i=1}^{n} \beta_1 x_i = 0$$

$$=> \sum_{i=1}^{n} y_i - n\beta_0 - \sum_{i=1}^{n} \beta_1 x_i = 0$$

$$=> n\beta_0 = \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} \beta_1 x_i$$

$$=> n\beta_0 = \sum_{i=1}^{n} y_i - \beta_1 \sum_{i=1}^{n} x_i$$

$$=> \beta_0 = \frac{\sum_{i=1}^{n} y_i}{n} - \beta_1 \frac{\sum_{i=1}^{n} x_i}{n}$$

$$=> \boldsymbol{\beta_0 = \bar{y} - \beta_1 \bar{x}} \qquad \text{---- (3)}$$

<u>Step 1:</u> Partial derivative w.r.t. $\boldsymbol{\beta_1}$

$$\frac{\partial SSE}{\partial \beta_1} = \frac{\partial}{\partial \beta_1}\left(\sum_{i=1}^{n}\big(y_i - (\beta_0 + \beta_1 x_i)\big)^2\right)$$

$$= -2\sum_{i=1}^{n}\big(y_i - (\beta_0 + \beta_1 x_i)\big)x_i$$

Setting $\frac{\partial SSE}{\partial \beta_1} = 0$

$$=> -2\sum_{i=1}^{n}\big(y_i - (\beta_0 + \beta_1 x_i)\big)x_i = 0$$

$$=> \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)x_i = 0$$

$$=> \sum_{i=1}^{n} y_i x_i - \beta_0 \sum_{i=1}^{n} x_i - \beta_1 \sum_{i=1}^{n} x_i^2 = 0$$

Substituting equation 3 $\boldsymbol{\beta_0 = \bar{y} - \beta_1 \bar{x}}$

$$=> \sum_{i=1}^{n} y_i x_i - (\bar{y} - \beta_1 \bar{x}) \sum_{i=1}^{n} x_i - \beta_1 \sum_{1=1}^{n} x_i^2 = 0$$

$$=> \sum_{i=1}^{n} y_i x_i - \bar{y}\sum_{i=1}^{n} x_i + \beta_1 \bar{x}\sum_{i=1}^{n} x_i - \beta_1 \sum_{i=1}^{n} x_i^2 = 0$$

$$=> \sum_{i=1}^{n} y_i x_i - \bar{y}\sum_{i=1}^{n} x_i - \beta_1 \left(\sum_{i=1}^{n} x_i^2 - \bar{x}\sum_{i=1}^{n} x_i\right) = 0$$

$$\beta_1\left(\sum_{i=1}^{n} x_i^2 - \bar{x}\sum_{i=1}^{n} x_i\right) = \sum_{i=1}^{n} y_i x_i - \bar{y}\sum_{i=1}^{n} x_i \quad \text{---- (4)}$$

Simplifying part of the left side of the above equation (4)

$$\left(\sum_{i=1}^{n} x_i^2 - \bar{x}\sum_{i=1}^{n} x_i\right)$$

Since, $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$

$$\sum_{i=1}^{n} x_i^2 - \bar{x}\sum_{i=1}^{n} x_i$$

$$= \sum_{i=1}^{n} x_i^2 - \left(\frac{\sum_{i=1}^{n} x_i}{n}\right)\sum_{i=1}^{n} x_i$$

$$= \sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2 \quad \text{--- (5)}$$

Sum of Squared Deviation (SSD) = $\sum_{i=1}^{n}(x_i - \bar{x})^2$ --- (6)

$$= \sum_{i=1}^{n} \left[x_i^2 - 2x_i\bar{x} + \bar{x} \cdot \bar{x} \right]$$

$$= \sum_{i \Rightarrow 1}^{n} x_i^2 - 2\bar{x} \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} \bar{x} \cdot \bar{x}$$

$$= \sum_{i=1}^{n} x_i^2 - \frac{2\sum_{i=1}^{n} x_i}{n} \left(\sum_{i=1}^{n} x_i \right) + n\bar{x} \cdot \bar{x}$$

$$= \sum_{i=1}^{n} x_i^2 - \frac{2}{n} \left(\sum_{i=1}^{n} x_i \right)^2 + n \left(\frac{\sum_{i=1}^{n} x_i}{n} \right)^2$$

$$= \sum_{i=1}^{n} x_i^2 - \frac{2}{n} \left(\sum_{i=1}^{n} x_i \right)^2 + \frac{1}{n} \left(\sum_{i=1}^{n} x_i \right)^2$$

$$= \boxed{\sum_{i=1}^{n} x_i^2 - \frac{1}{n} \left(\sum_{i=1}^{n} x_i \right)^2} \qquad \text{--- (7)}$$

We see that equation (5) and equation (7) are the same.
So, we can say that equation (5) is equal to equation (6).

$$\sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - \frac{1}{n} \left(\sum_{i=1}^{n} x_i \right)^2$$

Dividing both sides by n

$$\frac{1}{n}\left(\sum_{i=1}^{n}(x_i-\bar{x})^2\right)=\frac{1}{n}\left[\sum_{i=1}^{n}x_i^2-\frac{1}{n}\left(\sum_{i=1}^{n}x_i\right)^2\right]$$

Var(x) = $\frac{1}{n}\left[\sum_{i=1}^{n}x_i^2-\frac{1}{n}(\sum_{i=1}^{n}x_i)^2\right]$

n Var(x) = $\left[\sum_{i=1}^{n}x_i^2-\frac{1}{n}(\sum_{i=1}^{n}x_i)^2\right]$ --- (8)

Hence, using equation (8), we can write equation (4) as:

$\beta_1 n\,Var(x)=\sum_{i=1}^{n}y_i x_i-\bar{y}\sum_{i=1}^{n}x_i$ --- (9)

Solving the right side of equation (9)

$$\sum_{i=1}^{n}y_i x_i-\bar{y}\sum_{i=1}^{n}x_i$$

$$=\sum_{i=1}^{n}y_i x_i-\frac{1}{n}\left(\sum_{i=1}^{n}y_i\sum_{i=1}^{n}x_i\right)$$

We know that,

cov(x, y) = $\frac{1}{n}(\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}))$

$$=\frac{1}{n}\left(\sum_{i=1}^{n}x_i y_i-\bar{x}\sum_{i=1}^{n}y_i-\bar{y}\sum_{i=1}^{n}x_i+n\bar{x}\bar{y}\right)$$

$$= \frac{1}{n}\left(\sum_{i=i}^{n} x_i y_i - \frac{1}{n}\left(\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i\right) - \frac{1}{n}\left(\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i\right) + n\left(\frac{\sum_{i=1}^{n} x_i}{n}\right)\left(\frac{\sum_{i=1}^{n} y_i}{n}\right)\right)$$

$$= \frac{1}{n}\left(\sum_{i=1}^{n} x_i y_i - \frac{2}{n}\left(\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i\right) + \frac{1}{n}\left(\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i\right)\right)$$

$$= \frac{1}{n}\left(\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\left(\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i\right)\right)$$

Hence, we can solve this as:

n Cov(x,y) = $\left[\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\left(\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i\right)\right]$

n Cov(x,y) $= \sum_{i=1}^{n} y_i x_i - \bar{y}\sum_{i=1}^{n} x_i$ --- (10)

Combining, equation (9) and (10)

$$\beta_1 n\, Var(x) = n\, Cov(x, y)$$

$$\beta_1 = \frac{Cov(x,y)}{Var(x)}$$

$$\beta_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Since, $r = \frac{Cov(x,y)}{\sigma(x)\sigma(y)}$

and $\beta_1 = \frac{Cov(x,y)}{Var(x)}$

$$= \frac{r \cdot \sigma(x)\sigma(y)}{\left(\sigma(x)\right)^2}$$

$$\beta_1 = r \cdot \frac{\sigma(y)}{\sigma(x)}$$