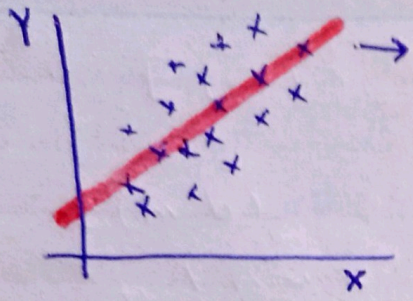
Simple Linear Regression | Mathematical Formulation



$\rightarrow y = mx + b$

here we need to find 'm' and 'b'

→ There are two ways to find 'm' and 'b'

① Closed form Solution
direct formula ⟷ OLS

② Non-Closed form Solution.
Gradient descent.

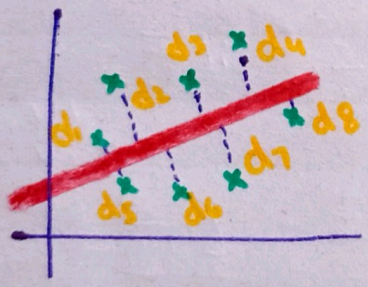
---

m and b → direct formula

$$b = \bar{y} - m\bar{x}$$

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

Formula's from Scratch.

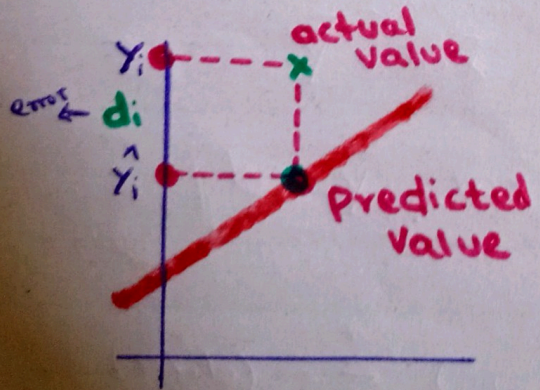


$$\text{Error} = d_1 + d_2 + d_3 + d_4 + \cdots + d_n$$

$$\text{Error} = d_1^2 + d_2^2 + d_3^2 + d_4^2 + \cdots + d_n^2$$

Square to get +ve result.

$$E = \sum_{i=1}^{n} d_i^2$$ → Error function
we can also represent E as (J)



$$d_i = Y_i - \hat{Y}_i$$

$$E = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$ ⇒ Total Error

$$E = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$ ⇒ Average error

$$\hat{y}_i = mx_i + b \Rightarrow \sum_{i=1}^{n} (y_i - mx_i - b)^2$$

Prediction Formula:
- Formula:

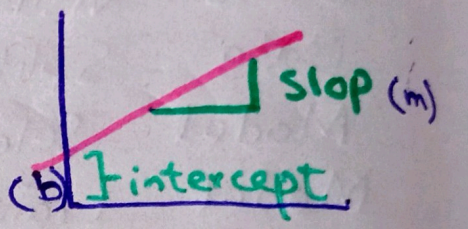
$$\hat{y} = m \cdot x + b$$

$$E(m,b) = \sum_{i=1}^{n} (y_i - mx_i - b)^2$$

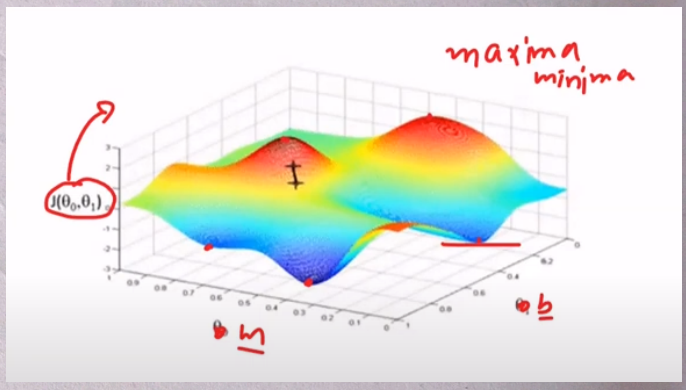
$E(m,b) \Rightarrow$ Error function depends on 'm' and 'b'
slop (m) and intercept (b)

$E(m,b)$
to find maxima and minima.

$$\frac{dE}{dm} = 0 \quad , \quad \frac{dE}{db} = 0$$



$$\frac{dE}{db} = \sum_{i=1}^{n} \frac{d}{db} (y_i - mx_i - b)^2$$



Chain Rule

$$\sum_{i=1}^{n} -2 (y_i - mx_i - b) =$$

$\div$ by -2

$$\sum_{i=1}^{n} (y_i - mx_i - b) = 0$$

$$\sum_{i=1}^{n} y_i - \sum_{i=1}^{n} mx_i - \sum b = 0$$

divided by n

$$\frac{\sum_{i=1}^{n} y_i}{n} - \frac{\sum_{i=1}^{n} mx_i}{n} - \frac{\sum b}{n} = \frac{0}{n}$$

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

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

$$\frac{\sum b}{n} = nb$$

$b+b+b+\cdots+b = nb$

$$\bar{y} - m\bar{x} - \frac{nb}{n} = 0$$

$$\bar{y} - m\bar{x} = b$$

$$\boxed{b = \bar{y} - m\bar{x}}$$

Intercept (b) Calculation Formula:

$$b = \bar{y} - m \cdot \bar{x}$$

$$E(m,b) = \sum_{i=1}^{n} (y_i - mx_i - \bar{y} + m\bar{x})^2$$

$$\frac{dE}{dm} = \sum_{i=1}^{n} \frac{d}{dm} (y_i - mx_i - \bar{y} + m\bar{x})^2 = 0$$

$$\sum_{i=1}^{n} \frac{d}{dm} (y_i - mx_i - \bar{y} + m\bar{x})^2 = 0$$

Chain Rule.

$$\Sigma 2 (y_i - mx_i - \bar{y} + m\bar{x})(-x + \bar{x})$$

$$\Sigma -2 (y_i - mx_i - \bar{y} + m\bar{x})(x_i - \bar{x}) = 0$$

divided by -2 (both side)

$$\Sigma (y_i - mx_i - \bar{y} + m\bar{x})(x_i - \bar{x}) = 0$$

$$\Sigma [(y_i - \bar{y}) - m(x_i - \bar{x})](x_i - \bar{x}) = 0$$

$$\Sigma (y_i - \bar{y})(x_i - \bar{x}) = m \Sigma (x_i - \bar{x})^2$$

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

Slope (m) Calculation Formula:

- Formula:

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