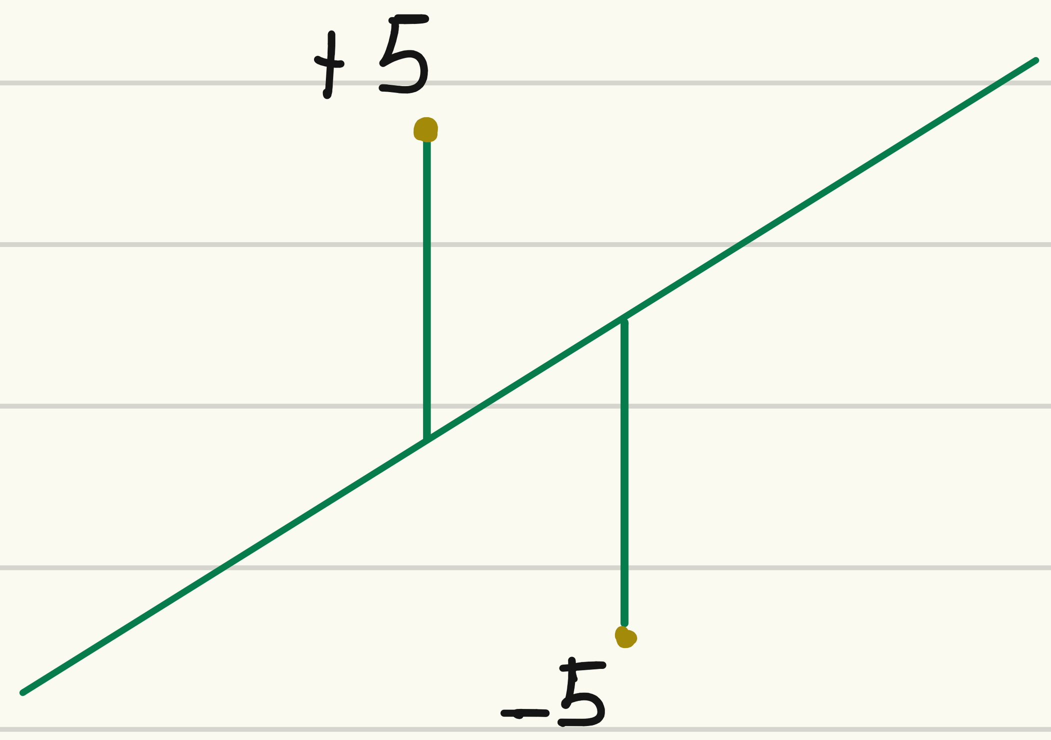
Linear Regression

$$\hat{y} = mx + b$$

$$E = \sum_{i=1}^{N} (y_1 - \hat{y}_i)^2$$

Why Squared ?



If $E = \sum_{i=1}^{N} (y_i - \hat{y}_i)$

$E = 0$

there is no error ?

$\Downarrow$

No

In $E = \sum_{i=1}^{N} (y_i - y_i)^2$;

$E \neq 0$

$$E = \sum_{i=1}^{N} (y_i - (ax+b))^2$$

$\therefore$ E changes wrt a & b.

dE/da & dE/db is needed to calculate.

$$\frac{dE}{da} = \sum_{i=1}^{N} 2\,(y_i - (ax+b))(-x)$$

$$dE/db = \sum_{i=1}^{N} 2\,(y_i - (ax+b))(-1)$$

Let dE/da & $dE/db = 0$ to find a, b.

$$\frac{dE}{da} = 0$$

$$\sum_{i=1}^{N} 2(y_i - (ax_i + b))(-x_i) = 0$$

$$\sum_{i=1}^{N} (-y_i x_i + a x_i^2 + b x_i) = 0$$

$$\left(a \sum_{i=1}^{N} x_i^2 + b \sum_{i=1}^{N} x_i = \sum_{i=1}^{N} y_i x_i \right)$$

$$a \sum x^2 + b \sum x = \sum yx$$

$$dE/db = 0$$

$$\sum_{i=1}^{N} 2(y_i - (ax+b))(-1) = 0$$

$$\sum_{i=1}^{N} y_i - \sum_{i=1}^{N} ax - \sum_{i=1}^{N} b = 0$$

$$\left(a \sum_{i=1}^{N} x + \sum_{i=1}^{N} b = \sum_{i=1}^{N} y_i \right) \times 1/N$$

$$a\bar{x} + b = \bar{y}$$

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

$$a \sum x^2 + (\bar{y} - a\bar{x}) \sum x = \sum yx$$

$$a \sum x^2 + \bar{y} \sum x - a\bar{x} \sum x = \sum yx$$

$$a\left(\sum x^2 - \bar{x} \sum x \right) = \sum yx - \bar{y} \sum x$$

$$\boxed{a = \frac{\sum yx - \bar{y}\sum x}{\sum x^2 - \bar{x}\sum x}}$$

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

$$= \bar{y} - \left(\frac{\sum yx - \bar{y}\sum x}{\sum x^2 - \bar{x}\sum x} \right) \bar{x}$$

$$= \frac{\bar{y}\sum x^2 - \cancel{\bar{y}\bar{x}\sum x} - \bar{x}\sum yx + \cancel{\bar{x}\bar{y}\sum x}}{\sum x^2 - \bar{x}\sum x}$$

$$\boxed{b = \frac{\bar{y}\sum x^2 - \bar{x}\sum yx}{\sum x^2 - \bar{x}\sum x}}$$

Mean

$$\text{Mean, } \bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

In Program Implementation,

$$\boxed{\begin{array}{lcl} \bar{x} & = & x \text{ mean ( )} \\ x & = & x\text{.sum ( )} \\ \sum x^2 & = & x\text{. dot } (x) \\ \sum xy & = & x\text{.dot } (y) \end{array}}$$

R - Squared

$$\sum (y - \hat{y})^2$$

$$R^2 = 1 - \frac{SS_{res}}{SS_{total}}$$

$$\sum (y - \bar{y})^2$$