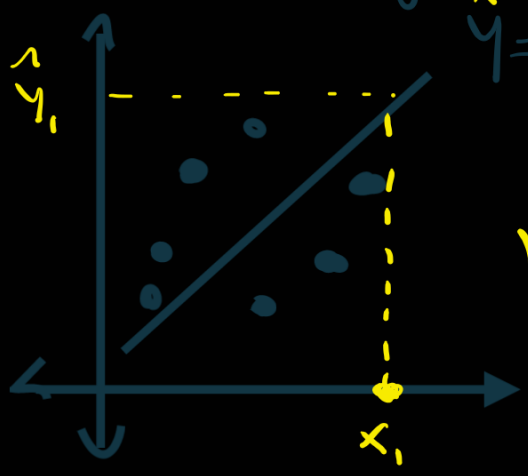
→ ML

→ Linear Regression



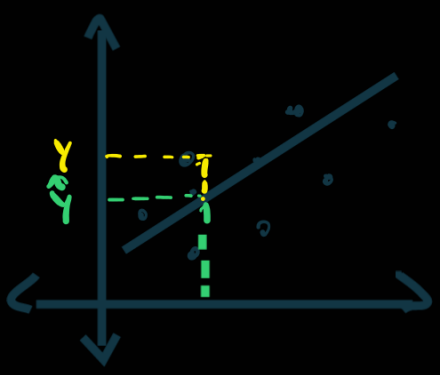
$\hat{y} = mx + c$ ← How do we know that this is the best prediction curve!

We need a measure!

→ Conditions :-

1. Should be able to formulate a linear relation.
2. Mean of residual should be zero.
3. Error terms are not supposed to be co-related.
4. X & residual must not be co-related.
5. Error term must showcase constant variance.
6. Error terms are supposed to be normally distributed.

→ What is Residual?



Residual is basically the perp. distance b/w the actual & pred. value.

i.e $|y - \hat{y}|$

$r = |y - (mx + c)|$

→ Since accuracy depends on each "individual" element we must eliminate any sort of sign in residuals.

$\therefore r^2 = \{y - (mx+c)\}^2$ **Residual Square.**

→ for all data points $(i=1 \dots m)$

$$\sum_{i=1}^{m} r^2 = \sum_{i=1}^{m} (y - (mx+c))^2$$

→ In order to minimise residual :-

$$\sum_{i=1}^{m} (y - (mx+c))^2$$

$$= \sum_{i=1}^{m} y^2 + (mx+c)^2 - 2y(mx+c)$$

$$\sum_{i=1}^{m} r^2 = \sum_{i=1}^{m} y^2 + m^2x^2 + c^2 - 2mxy - 2cy$$

→ for a particular point our aim is $r=0$.

Hence :- $\partial \Sigma r^2 / \partial m = 0$ $\partial \Sigma r^2 / \partial c = 0$ $(\Sigma r^2 = R)$

$$\frac{dr^2}{dm} = \frac{\partial R}{\partial m} = \Sigma 0 + 2mx^2 + 0 + 2xc - 2xy = 0$$

$$= \Sigma 2mx^2 + 2xc - 2xy = 0$$

$$= \Sigma 2x(mx + c - y) = 0 \quad \text{——— (1)}$$

$$\frac{\partial R}{\partial c} = \Sigma 2(c + mx - y) = 0 \quad \text{——— (2)}$$

$$\therefore \frac{\partial R}{\partial m} = \Sigma 2 \times mx + \Sigma 2xc - \Sigma 2xy = 0$$

$$\& \frac{\partial R}{\partial c} = \Sigma 2c + \Sigma 2mx - \Sigma 2y = 0$$

→ But how do we solve it?

$$m_{new} = m_{old} - \eta \frac{1}{m} \underbrace{\left(\sum_{i=1}^{m} (y - \hat{y}) \right)}_{\Sigma \sigma}$$

$$c_{new} = c_{old} - \eta \frac{1}{m} \left(\sum_{i=1}^{m} (y - \hat{y}) \right)$$

→ Simply :-

$$m_{new} = m_{old} - \eta \nabla E_m$$
$$c_{new} = c_{old} - \eta \nabla E_c$$

} You are seeing this on a graph.

→ Suppose :-

	w	h
①	60	5.1
②	62	5.3
③	65	5.5
④	72	6.1
⑤	40	3.1

} $m = ?$ $c = ?$ for these 2.

$\nabla E_m = mx^2 + xc - xy$

$\nabla E_c = c + mx - y$

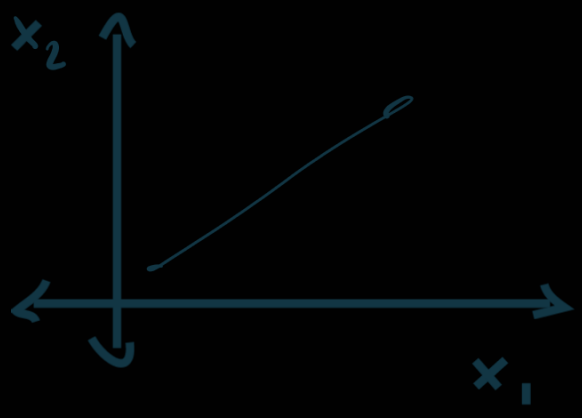
→ Calculate ∇E_m & ∇E_c for 3rd datapoint (Just a representational way of finding residual).

→ Now iteratively calculate m_{new} & c_{new} & repeat the steps.

→ What is multicolinearity ?

Remember MTH 113 ?

→ The dataset itself is related !
(No Use Basically)



→ How is a model tested ?

Done through R^2 Statistics.

Accuracy ⟶ 0 to 1.

$$\therefore R^2 = \left(1 - \frac{RSS}{TSS}\right)$$

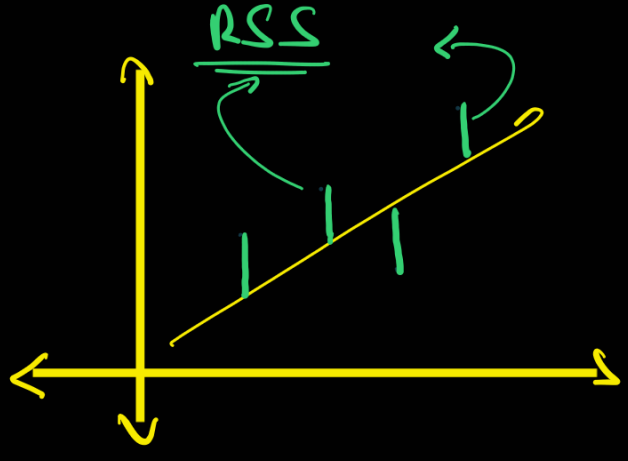
RSS :- Residual Summation of Squares.

TSS :- Total Summation of Squares. → Distance b/w y_i & $\hat{y}$

RSS ↓ R^2 ↑

TSS :- $\hat{y} = \frac{1}{m}\sum_{i=1}^{m} y_i$

↳ ⇒ $\Sigma|y_i - \hat{y}|$



→