→ R^2 :- Evaluation metrices which is used to measure the performance of linear regression. Also known as coeff of determination

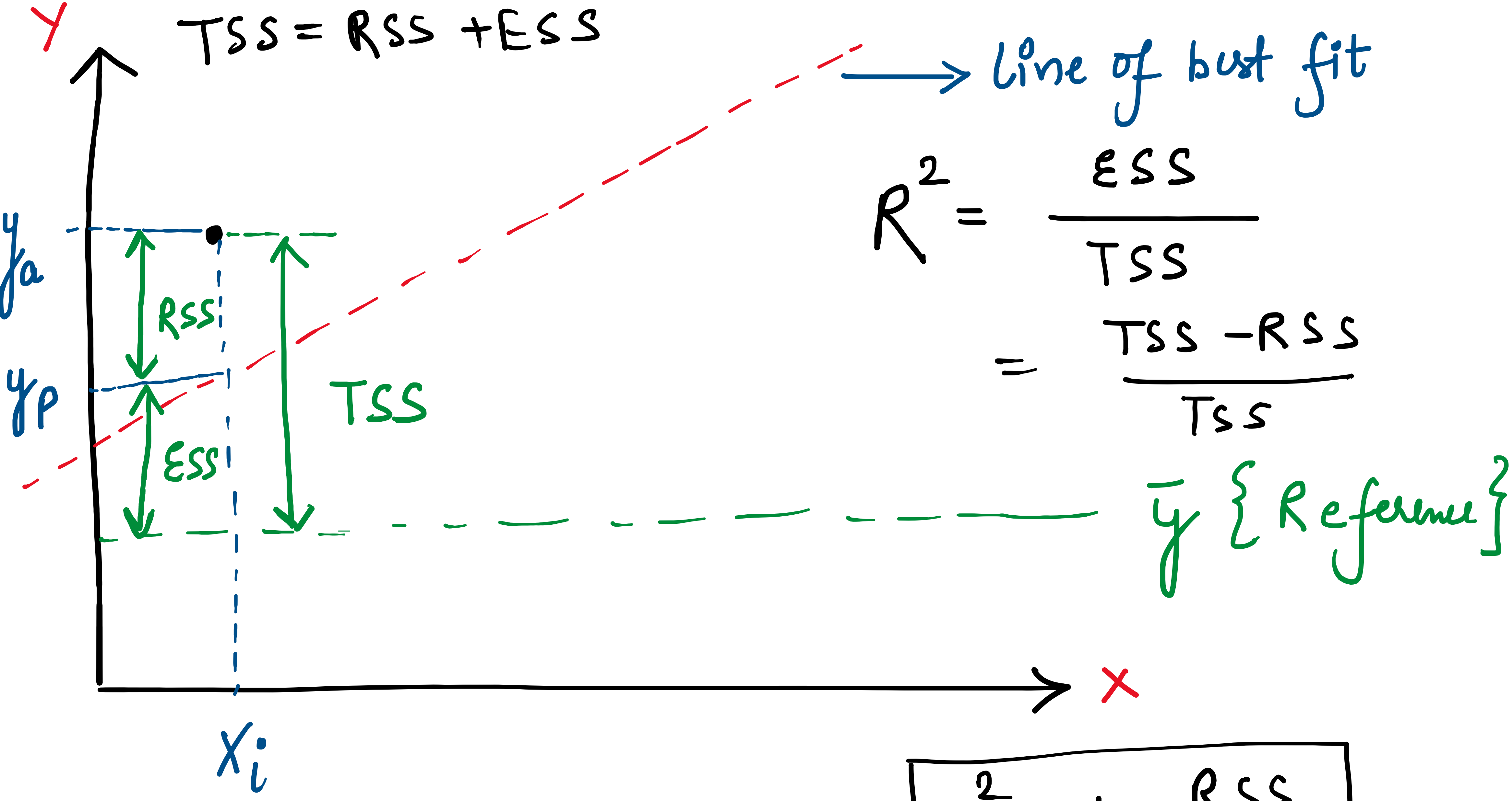
Mathematically,

$$\boxed{\uparrow R^2 = 1 - \frac{RSS \downarrow}{TSS}}$$

RSS → Residual sum of squares $= \sum (y_a - y_p)^2$

TSS → Total sum of squares $= \sum (y_a - \bar{y})^2$

$$TSS = RSS + ESS$$



y

$$R^2 = \frac{ESS}{TSS}$$

$$= \frac{TSS - RSS}{TSS}$$

$$\boxed{R^2 = 1 - \frac{RSS}{TSS}}$$

$$\left\{ y = 0.8X + 1.6 \right.\ (\text{LOBF})$$

Calculate R^2.

X_1

X_2 ✓	Y_a	y_p	$(y_a - y_p)^2$	$(y_a - \bar{y})^2$
1	2	2.4	.16	4
2	3	3.2	.04	1
3	5	4.0	1	1
4	4	4.8	.64	0
5	6	5.6	.16	4

$\Sigma 2$ → RSS

$\Sigma 10$ → TSS

$$\bar{y} = \frac{2+3+5+4+6}{5}$$

$$\bar{y} = 4$$

$$R^2 = 1 - \frac{RSS}{TSS}$$

$$R^2 = 1 - \frac{2}{10}$$

$$= 0.8$$

→ Adjusted R^2 :- .

N – No. of datapo

P – No. of predict

{ P = No. of feature (X)

N = No of datapoints

$$\text{Adj } R^2 = 1 - \frac{(1-R^2)(N-1)}{(N-P-1)}$$

$$\text{Adj } R^2 = 1 - \frac{(1-0.8)(5-1)}{5-1-1} = 1 - \frac{0.2 \times 4}{3}$$

$$= 1 - \frac{0.8}{3} = \frac{2.2}{3} = .73$$

→ **Cost function** :- Mathematical error f^n which is to be minimised to get the line of best fit.

i.e = Mean Square Error $\left(\frac{RSS}{N}\right)$

✓

✓ $$C = \frac{\sum (y_a - y_p)^2}{N}$$

✓ $$C = \frac{1}{N} \sum \left[y_a - (mx_i + c) \right]^2$$ ✓

→ Gradient Descent Algorithm :- It is an iterative algorithm to minimize the cost function to get the line of best fit. (m & c)

* Step 1 :- Initialising the value of m & c as zero.
* Step 2 :- Calculate the cost f^n $\downarrow J = \frac{1}{N} \sum [y_a - (mx + c)]^2$
* Step 3 :- $m' = m - \alpha \frac{\partial J}{\partial m}$ $c' = c - \alpha \frac{\partial J}{\partial c}$
* Step 4 :- Repeat 2 & 3 untill we got the minima value

$$m' = m - \alpha \underbrace{\frac{\partial J}{\partial m}}_{\text{Error gradient}}$$

$$c' = c - \alpha \underbrace{\frac{\partial J}{\partial c}}_{\text{Error gradient}}$$

$\boxed{\alpha = 0.001}$

$\alpha \rightarrow$ learning rate

$\rightarrow$ speed of learn

$$\overbrace{0 \quad 1 \quad 2 \quad 3 \quad 4 \quad 5}^{5}$$

$$\underbrace{0,\ 0.1,\ 0.2,\ 0.3 \ \ldots\ 5}_{50}$$

0

→ Calculation of $\frac{\partial J}{\partial m}$: (Not important)

$\frac{dJ}{dc}$

$x^n \to nx^{n-1}$

$$J = \frac{1}{N} \sum \left[y_a - (mx + c) \right]^2$$

$$\frac{\partial J}{\partial m} = \frac{1}{N} \times 2 \times \sum \left[y_a - (mx + c) \right]^{2-1} (0 - (x + $$

$$= \frac{-2}{N} \sum \left[y_a - (mx + c) \right] (x)$$

→ Calculation of $\frac{\partial J}{\partial c}$:

$$J = \frac{1}{N} \sum \left[y_a - (mx + c) \right]^2$$

$$\frac{\partial J}{\partial c} = \frac{1}{N} \times 2 \times \sum \left[y_a - (mx + c) \right]^{2-1} (0 - (0 + 1))$$

$$= \frac{-2}{N} \sum \left[y_a - (mx + c) \right]$$

→ Variance Inflation factor (VIF) :- Statistical metric to measure the multicollinearity.

$$VIF = \frac{1}{1-R^2}$$

$R^2 \rightarrow$ Corr b/w x_i

Ideally, in our LR model the $VIF \leq 5$.