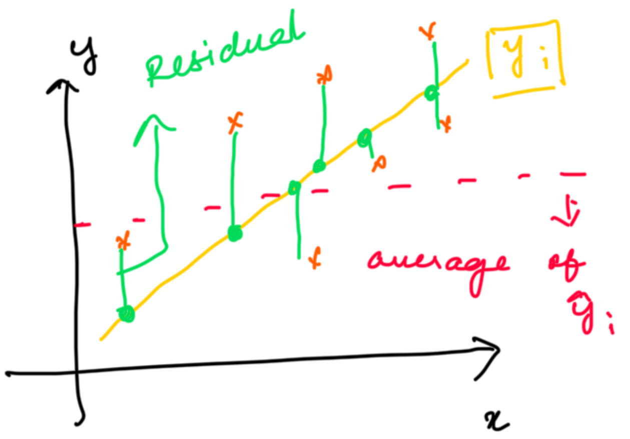
Performance metrics used in Linear regression

① R Square

$$R\text{squared} = 1 - \frac{SS_{Res}}{SS_{total}}$$



● → Predicted points

$\hat{y}_i$

therefore Rsquared can be denoted as

$$R\text{squared} = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y}_i)^2}$$

Basically, this will result is

$$= 1 - \frac{\text{Smaller number}}{\text{Bigger number}}$$

② Adjusted R square

Let say, computed R square value is

75% = 0.75

and let us assume there is an Input feature which is not at all corelated with the output feature but the R square value still increases in terms of accuracy.

This is the problem of Rsquared.

to prevent this we use adjusted R square.

It penalizes the respective Rsquared that are not corelated in input feature.

= 1 - smaller number

= 1 (approximately)

$$0 \xrightarrow[\text{accuracy}]{} 1$$

meaning, more the value towards 1, more the accuracy the model is.

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

N = No of data points

P = No of Independent features.