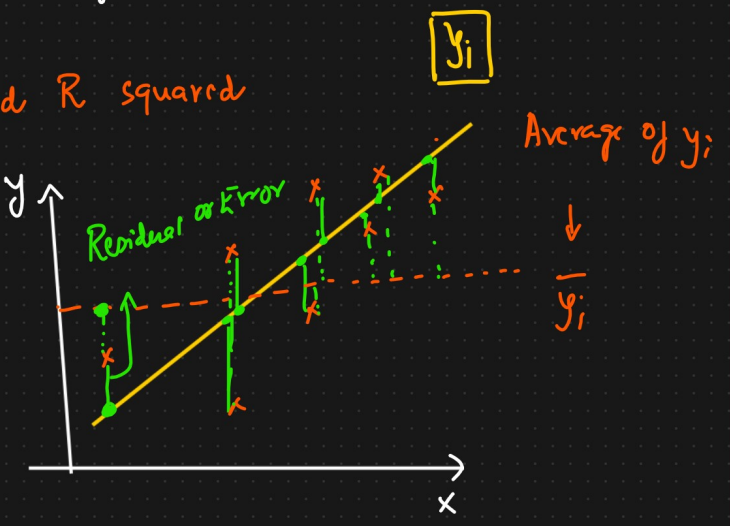
Performance Metrics Used In Linear Regression

① R squared ② Adjusted R squared



$$R^2 = 1 - \frac{SS_{Res}}{SS_{TOTAL}}$$

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

$$= 1 - \frac{\text{Small number}}{\text{Bigg number}}$$

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

$$\approx 1$$

$0.70 \Rightarrow 70\%$

$0.85 \Rightarrow 85\%$

$0.90 \Rightarrow 90\%$

{Overfitting, Underfitting}

1 ⇓

Accurate My Model Is?

R squared ↑↑↑

Size of the house ↑ Price ↑

+ve correlation

② Adjusted R Squared

Dataset → (Price)

Gender | Size of the house | No. of bedrooms | Location | Price

No. of bedroom ↑ Price ↑

+ve correlation

R squared = 75% ⇒ 0.75

This is the problem of R Squared

R squared => 80% => 0.80

R squared => 85% => 0.85

R squared => 87% => 0.87

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

- $P=2$ $R^2 = 90\%$ $R^2_{adjusted} = 86\%$
- $P=3$ $R^2 = 92\%$ $R^2_{adjusted} = 82\%$

N = No. of data points

P = No. of Independent features