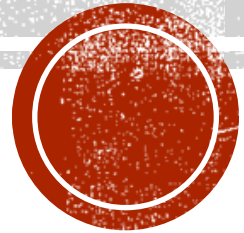
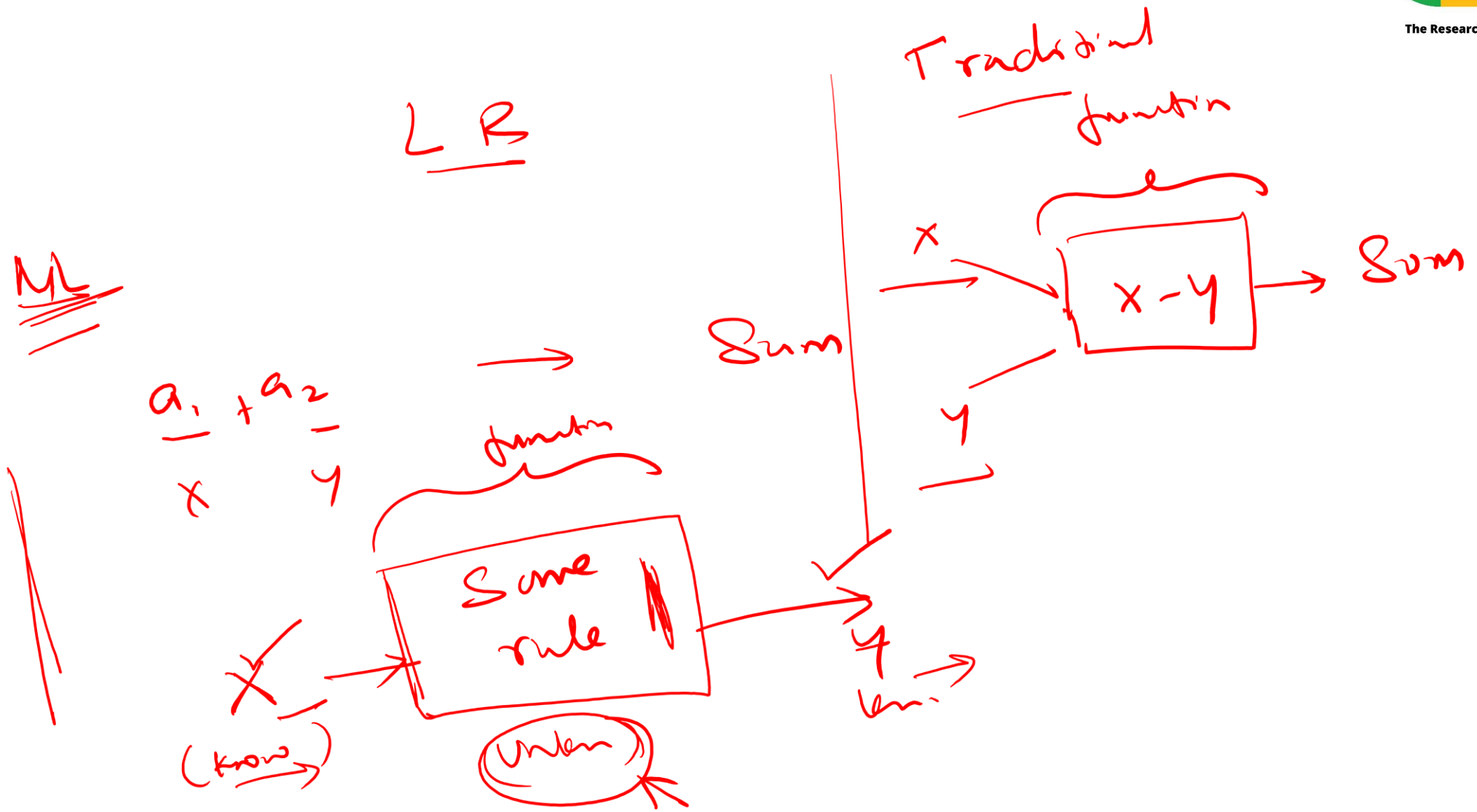
MACHINE LEARNING 35 – SIMPLE LINEAR REGRESSION - INTUITION AND SKLEARN MODEL



The Research Foundation

Dr. Ajantha Devi, Research Head


LR
ML
Traditional
function
x
x - y
Sum
Sum
$a_1 + a_2$
x
y
function
Some rule
y
(known)
Unknown

$y = mx + b$

m → slop
b → int

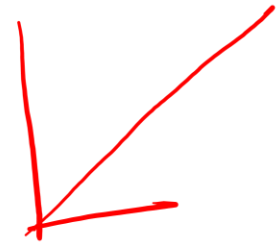
dep
indep

① Sum of Sq. dev → LS
② Correla coff → −1 to 1

Solutions
The Research Foundation

Assumptions of LR

1. Linearity ⇒

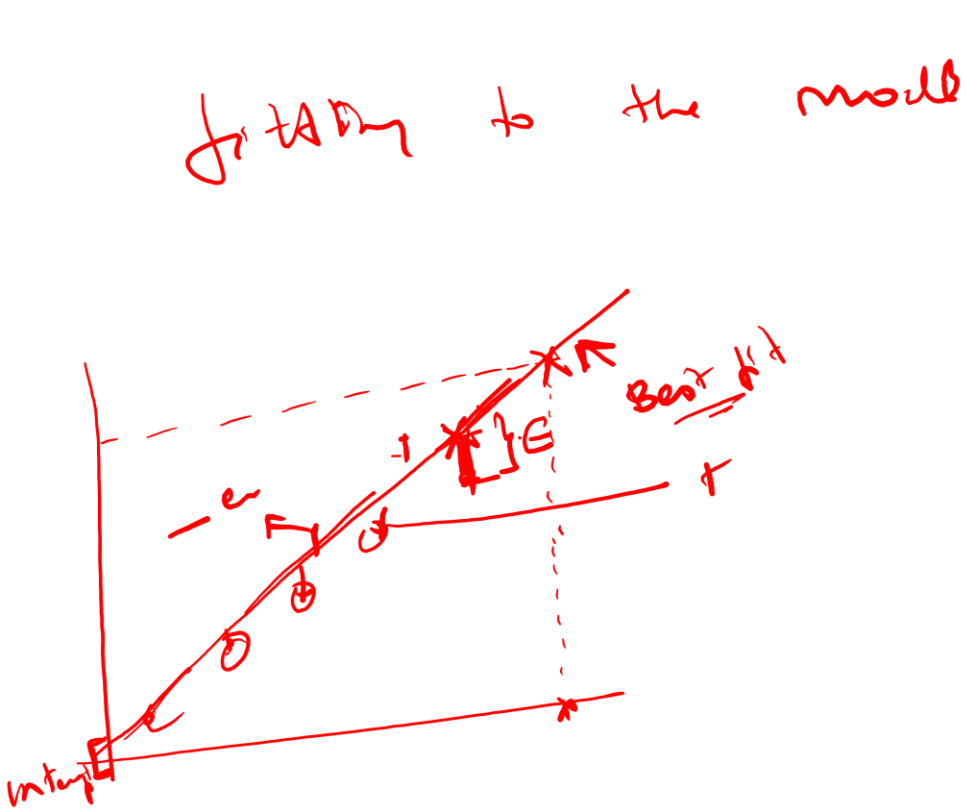


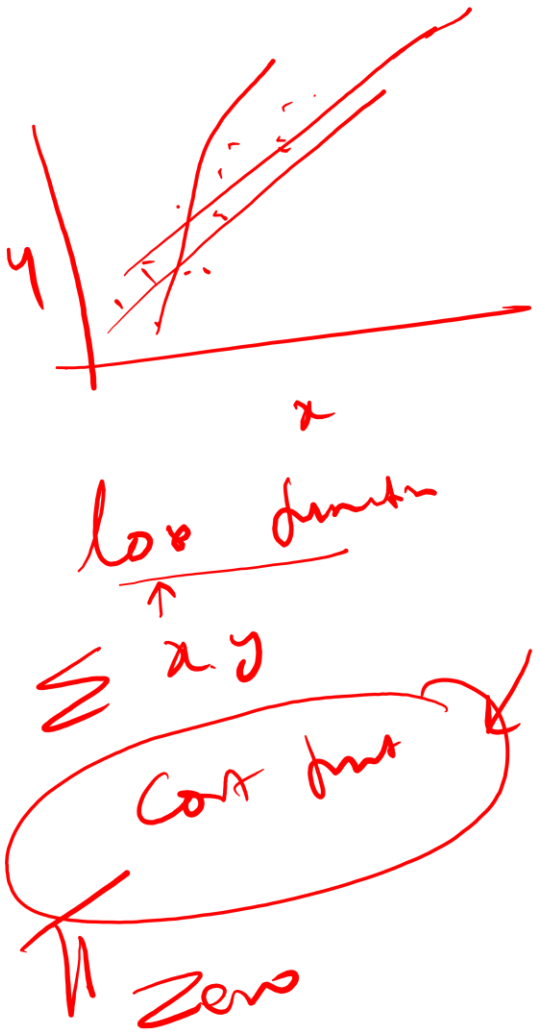
2. Independence

3. Homoscedasticity

4. Normality

5. No Multicollinearity





$y = mx + b$

- $m \Rightarrow \Delta y / \Delta x$
- $c/b \Rightarrow$

x is zero

hours Cost

x	y	xy	x^2
2	65	130	4
3	70	210	9
4	75	300	16
5	80	400	25
6	85	510	36
$\Sigma x = 20$	$\Sigma y = 375$	$\Sigma xy = 1550$	$\Sigma x^2 = 90$

$n = 5$

$$m = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2}$$

$$= \frac{5 \times 1550 - 20 \times 375}{5 \times (90) - (20)^2}$$

$$= \frac{7750 - 7500}{450 - 400}$$

$$= \frac{250}{50} = 5$$

m

$$b = \frac{375 - 5 \times 20}{5} = 55$$

Slope $m = 5$

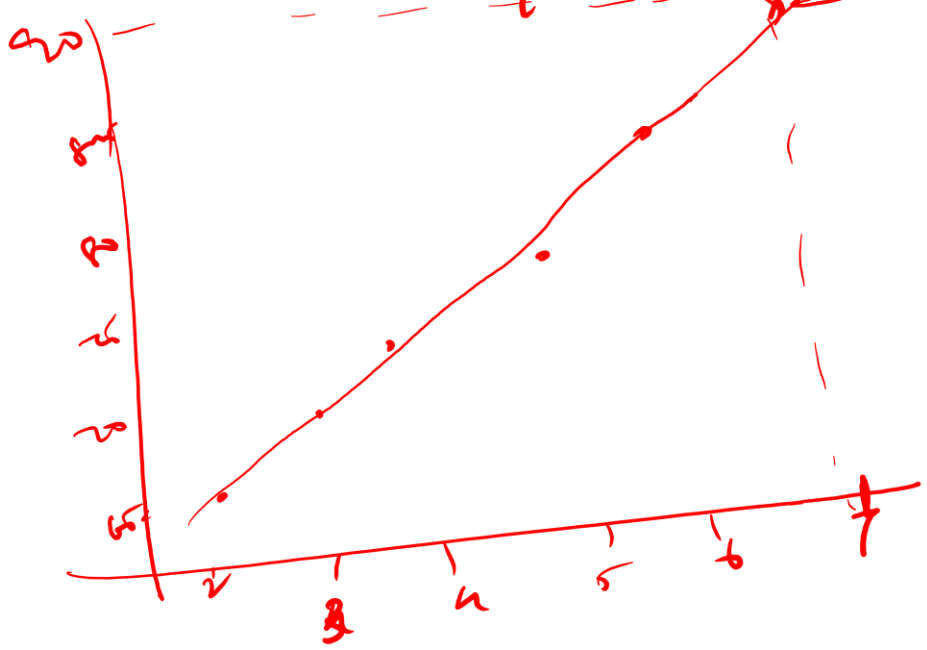
$b = 55$

$y = mx + b$

$y = 5 \times 7 + 55$

$= 90$

$x = 7$

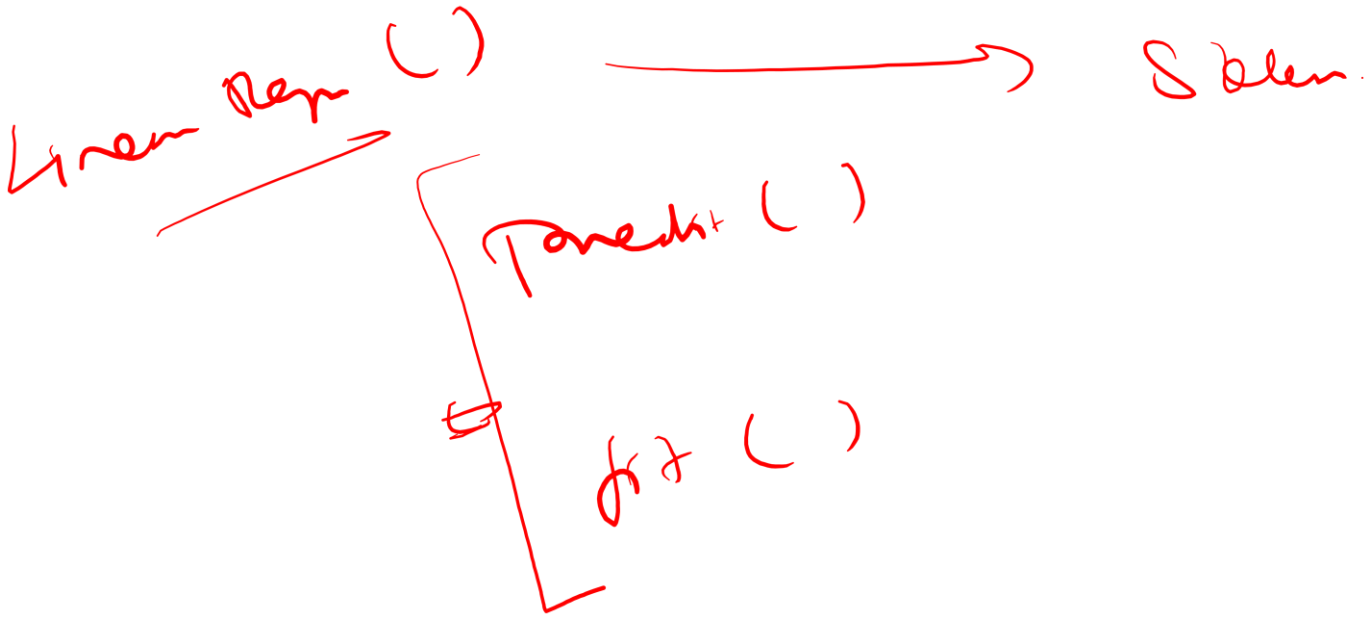


The Research Foundation

① R-Sq → Not Bunny Sq. R / Sq Sq Tm. 0-1

② RMSE $\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$ ↓

③ MAE $\frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$


Solution.