Simple Linear regression

Supervised ML → Regression

Data set

Independent / Future

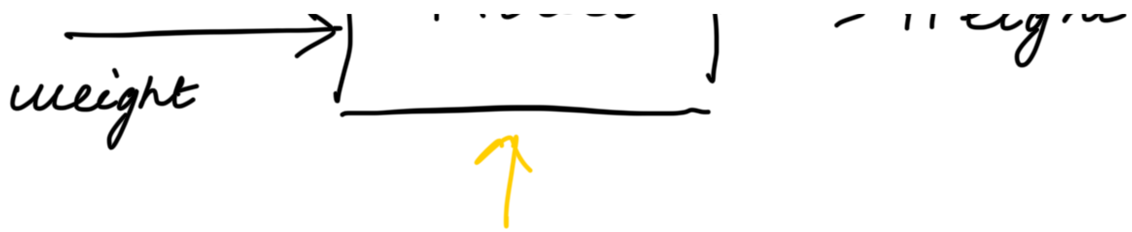
dependent / Target

weight	height
74	170 cm
80	180 cm
75	175.5cm
-	-

Our aim is to build a model, for any new weight, should be able to predict the height

Train

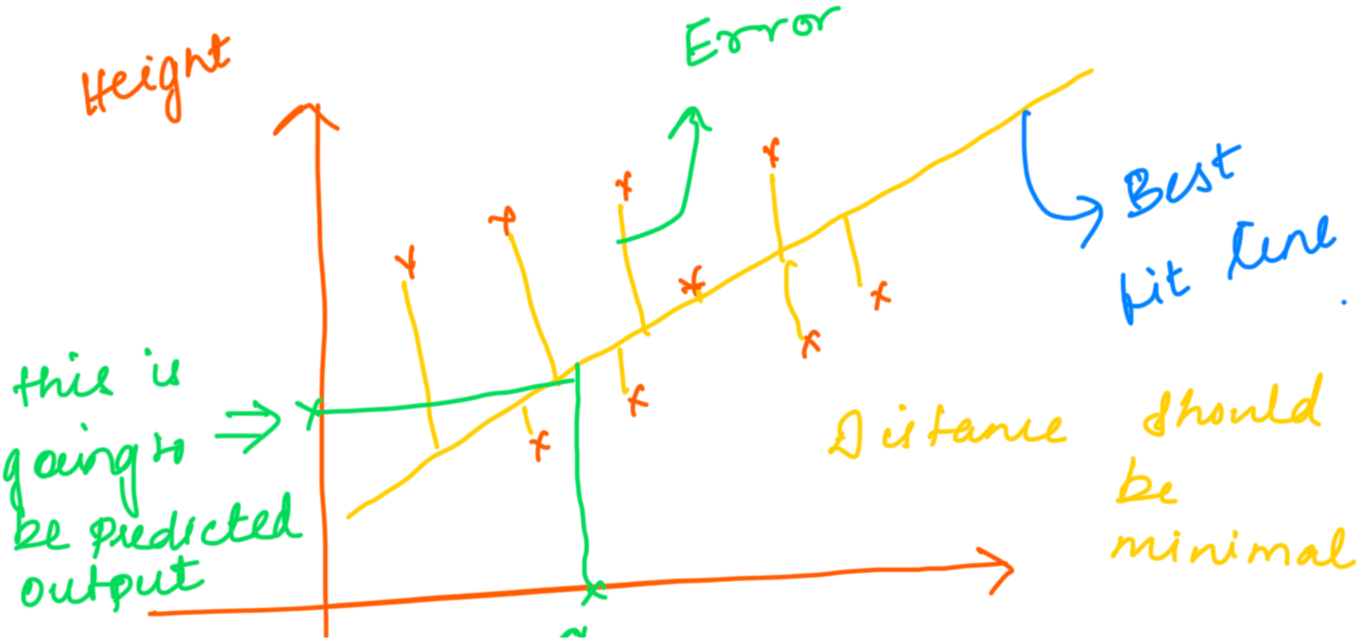
New → Model → Height



Simple Linear regression

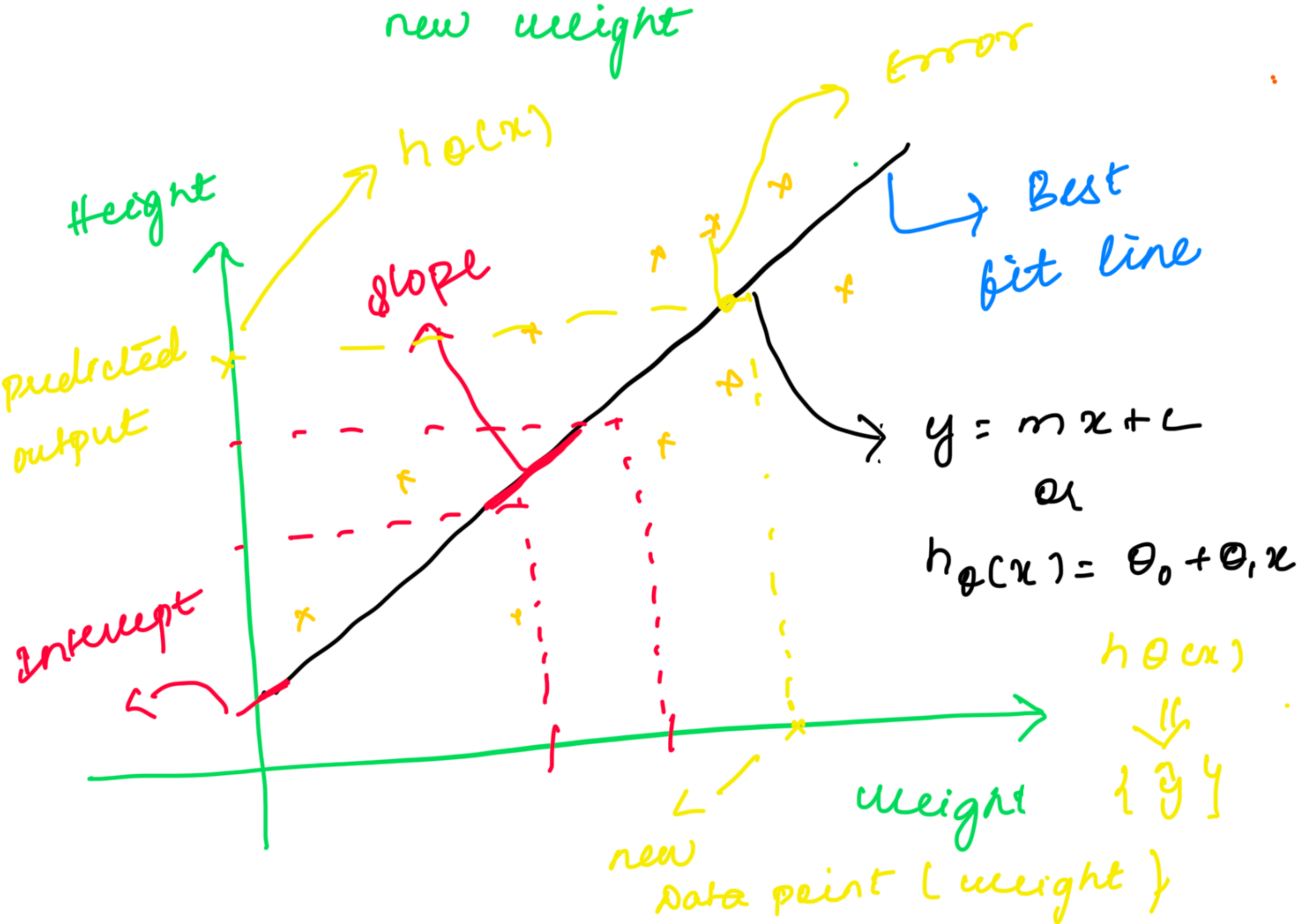
why we call it as simple Linear regression?

Because, here we are dealing with one Independent variable or one Feature Input variable and on Dependent variable or Target output variable.



| ↑ weight

new weight



$$h_\theta(x) = \theta_0 + \theta_1 x$$

θ_0 = Intercept

θ_1 = slope / coefficient

if $x = 0$

$h_\theta(x) = \theta_0$

↓

when there is a unit

movement in x axis, what would the movement with respect to y axis that is called as slope / coefficient.

Error = $(y - \hat{y})$

Best Fit line should be summation of all the errors should be minimal.

cost function

Predicted points

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(\underbrace{h_\theta(x)^i - y^i}_{\text{Error value}} \right)^2$$

True o/p

Error value

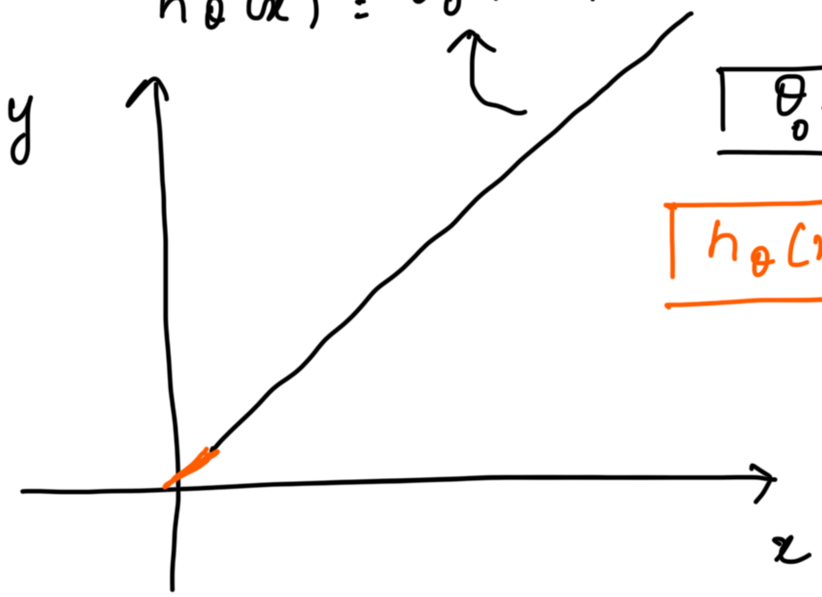
Mean squared Error

Final Aim what we need to solve

Minimize $J(\theta_0, \theta_1) = \frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta(x)^i - y^i\right)^2$

(1) Equation of Straight Line

$h_\theta(x) = \theta_0 + \theta_1 x$



$\boxed{\theta_0 = 0}$

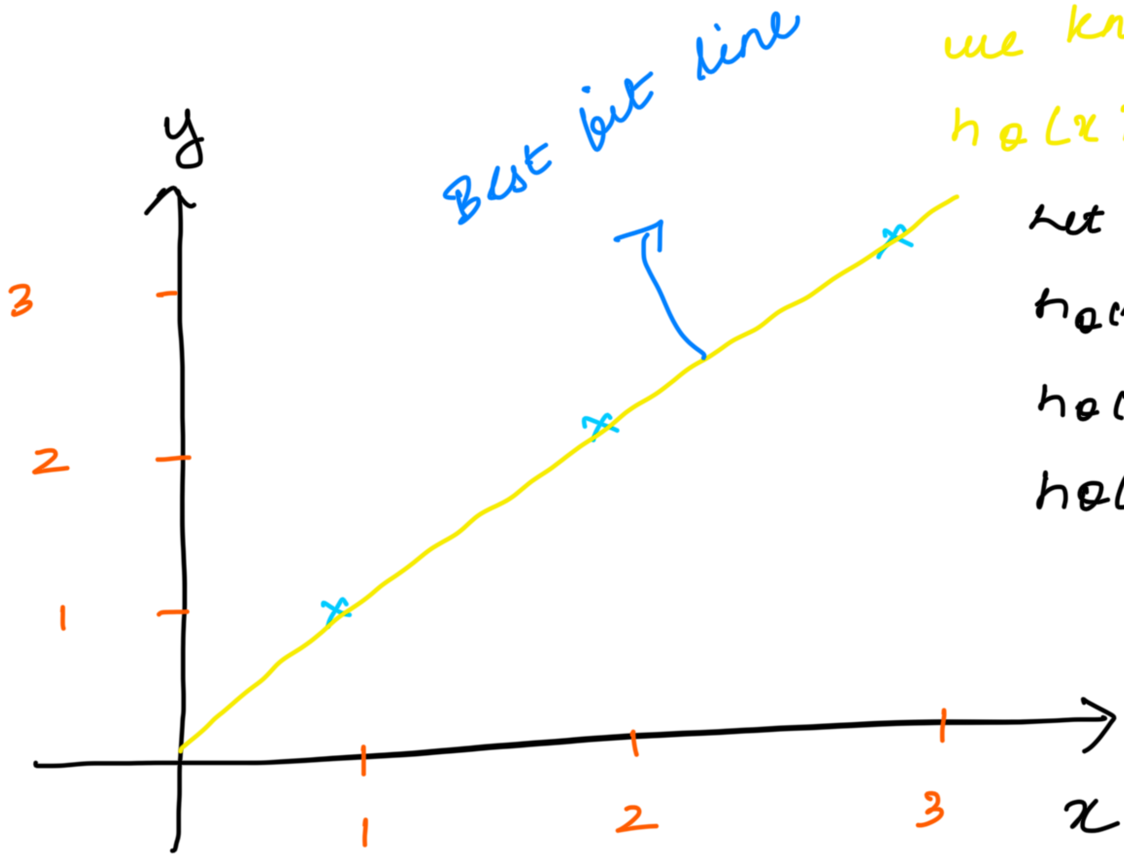
$\boxed{h_\theta(x) = \theta_1 x}$

Lets consider the below dataset

x	y
1	1
2	2

2 2

3 3



we know,

$h_\theta(x) = \theta_1 x$

Let $\theta_1 = 1$

$h_\theta(x) = 1 \mid x = 1$

$h_\theta(x) = 2 \mid x = 2$

$h_\theta(x) = 3 \mid x = 3$

Cost Function

$$J(\theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x)^i - y^i \right)^2$$

$$= \frac{1}{2(3)} \left[(1-1)^2 + (2-2)^2 + (3-3)^2 \right]$$

$$J(\theta_1) = 0$$

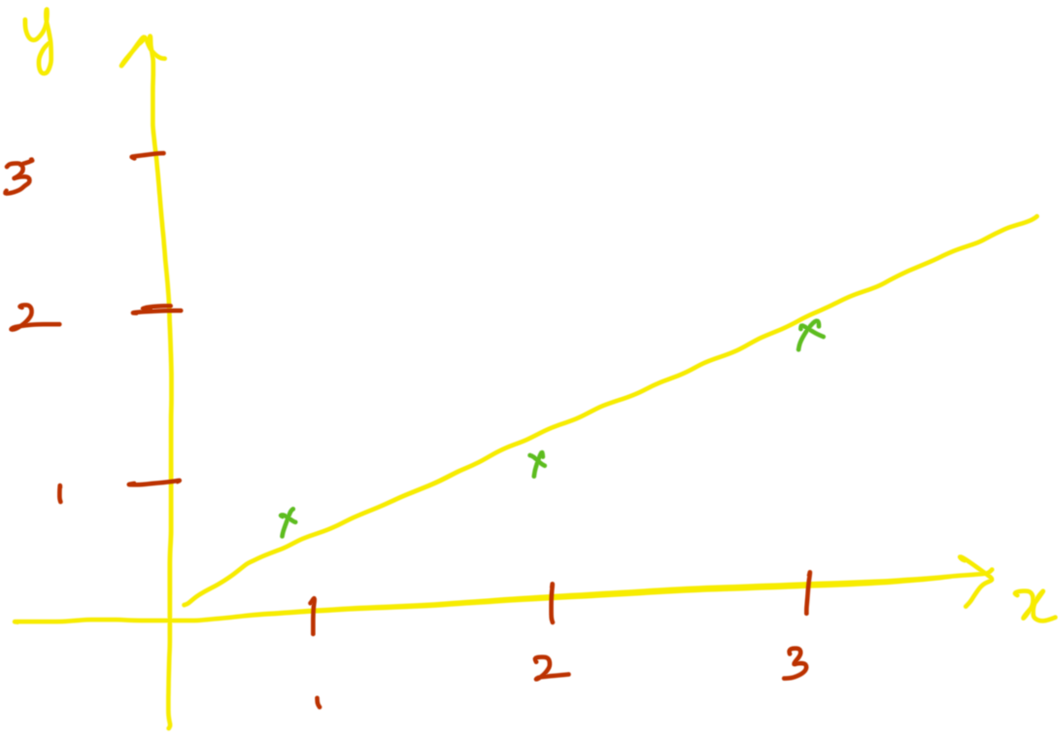
Let $\theta_1 = 0.5$

$h_\theta(x) = 0.5 \quad | \text{ if } x = 1$

$h_\theta(x) = 1 \quad | \text{ if } x = 2$

$h_\theta(x) = 1.5 \quad | \text{ if } x = 3$

Lets plot it in the graph

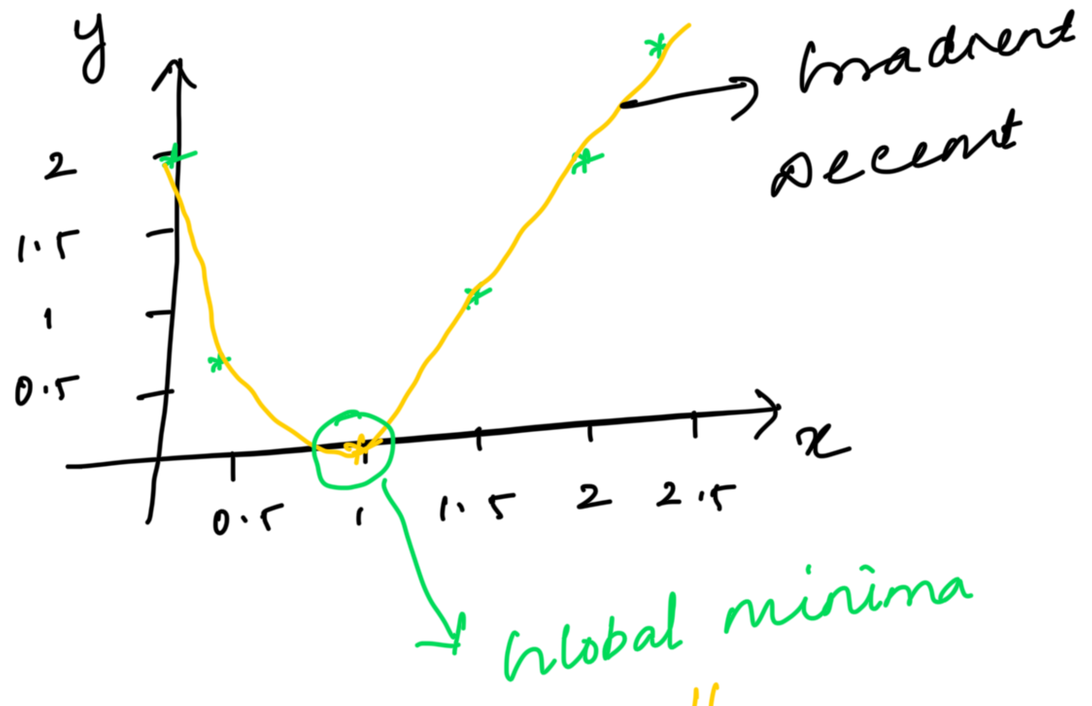


Applying cost function

$$J(\theta_1) = \frac{1}{2(3)}\left[(0.5-1)^2 + (1-2)^2 + (1.5-3)^2\right]$$

$$J(\theta_1) = 0.58$$

this cost function should applied on the various θ value to find out the best fit line



⇓

Error has been

minimized

we do this by changing θ_1 or θ_0 value.

Convergence Algorithm

This is super important as this the optimization technique

The main aim of convergence algorithm is to optimize the changes of θ_1 value.

What does convergence algorithm basically say?

Repeat until convergence

(Global minima)

$$\{$$

$$\theta_j = \theta_j - \alpha \boxed{\frac{\partial}{\partial \theta_j} J(\theta_j)}$$

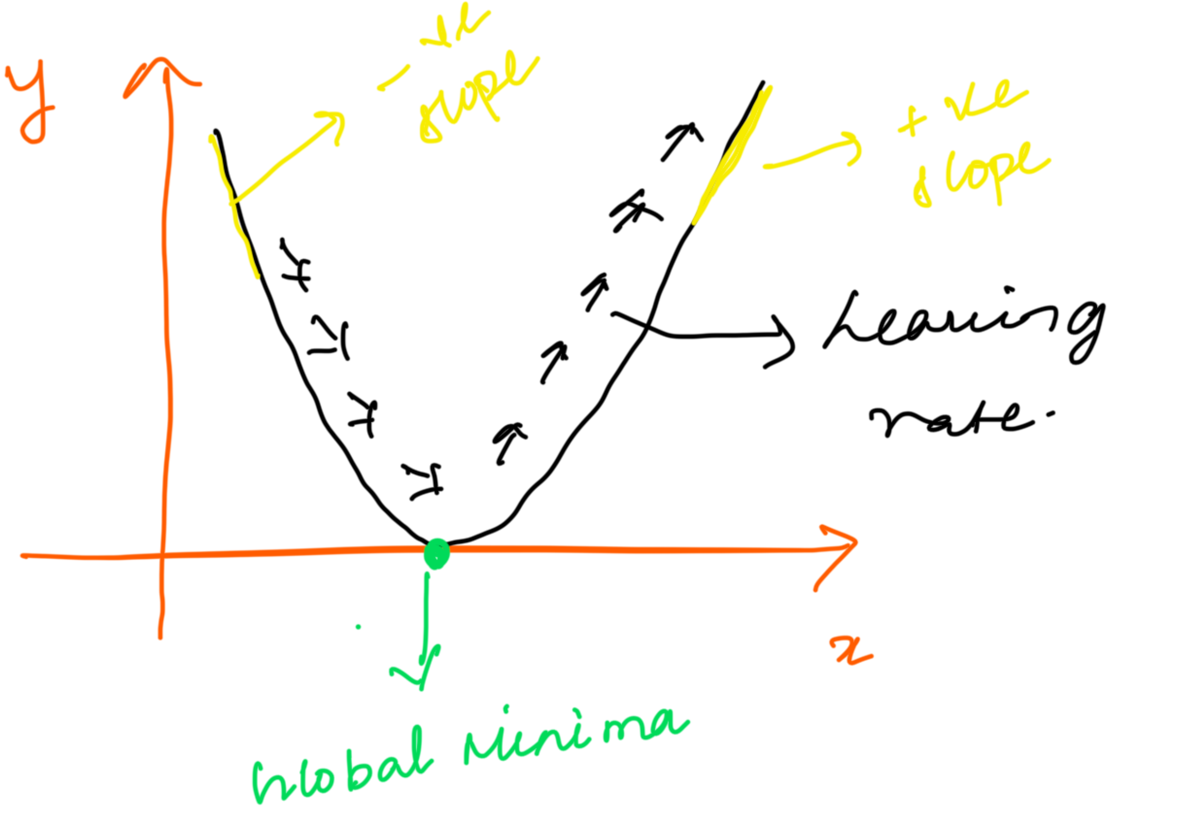
$$\}$$

$\Downarrow$

Derivative = slope.

$\alpha \Rightarrow$ Learning rate

For example



is negative

If the slope is negative

$$\theta_j := \theta_j - \alpha(-ve)$$

$$\theta_j := \theta_j + (+ve)$$

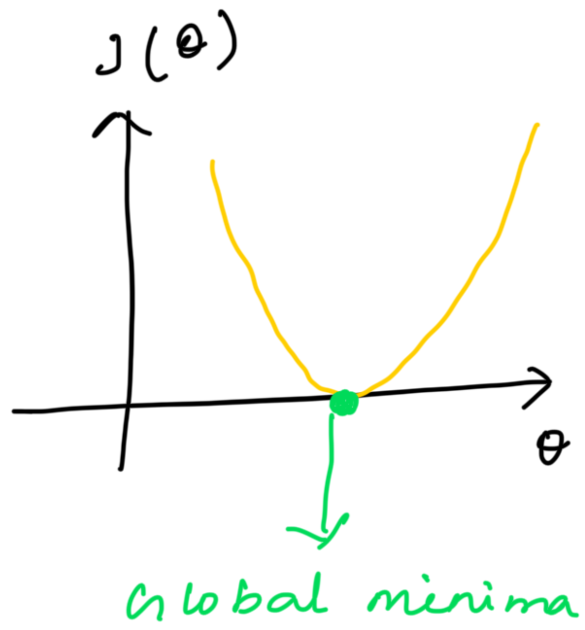
If the slope is positive

$$\theta_j := \theta_j - \alpha(+ve)$$

$$\theta_j := \theta_j - (+ve)$$

Final conclusion

Gradient Descent



Convergence Algorithm

repeat unit convergence

{

$$\theta_j := \theta_j - \alpha \boxed{\frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1)}$$

J = 0 and 1

}

$$\frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1) = \frac{\partial}{\partial \theta_j} \frac{1}{2m} \sum_{i=n}^{m} \left(h_\theta(x)^i - y^i \right)^2$$