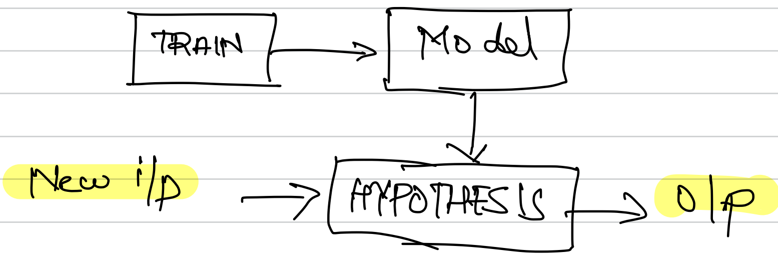
Linear Regression

Best fit line → The data points on this line shows the predicted value of Height for respective weight

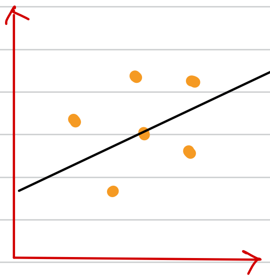
Actual data points

Predicted point



Aim :- To find the best fit line with minimum error b/w actual & predicted points

⇒ Exploring Maths how to find best fit line



Best fit line

$$\Rightarrow \hat{y} = mx + c$$

m → slope or coefficient

c → intercept

x → input

$\hat{y}$ → predicted point

Slope ⇒ If there is change of 1 unit in x, how much change is there in y unit

→ Creating hypothesis to get best fit line using single i/p feature

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

→ since we are using 1 i/p feature

⇒ Cost function –

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

$\theta_0 + \theta_1 x^{(i)}$

Predicted value

Actual value

Minimize this value by changing θ_0 and θ_1.

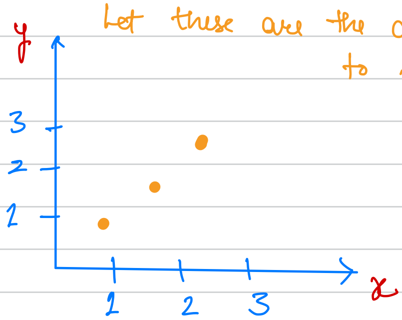
$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x)^{(i)} - y^{(i)} \right)^2$$

⇒ Comparison of cost func^n & value of θ_0, θ_1

We will make certain assumptions here –

① We are dealing in 2-D points ↳ just 1 input & 1 output

② $\theta_0 = 0$



Let these are the data points. Now, we will try to fit the line & find MSE

Let the value of $\theta_0 = 0$, $\theta_1 = 1$

$h_\theta(x^{(i)}) = \theta_1 x_i$

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

Acc. to data points

$x_1 = 1$, $x_2 = 2$, $x_3 = 3$

$h_\theta(x)^1 = 1 \times 1 = 1$, $h_\theta(x)^2 = 1 \times 2 = 2$

$h_\theta(x)^3 = 1 \times 3 = 3$.

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

$$= \frac{1}{6}(0)$$

$$J(\theta_1) = 0$$

$J(\theta_1)$	θ_1
0	1
0.58	0.5
2.105	0.05
0.58	1.5
2.33	2

Acc. to our data points
When $\theta_1 = 1$, $J(\theta_1) = 0$

$\Rightarrow$ We will try some more values for θ_1 such as 0.5, 0.05, 1.5, 2

If $\theta_1 = 0.5$

$$J(\theta_1) = \frac{1}{6}\left[(0.5\times 1 - 1)^2 + (0.5\times 2 - 2)^2 + (0.5\times 3 - 3)^2\right]$$

$$= \frac{1}{6}\left[(0.5)^2 + (-1)^2 + (-1.5)^2\right]$$

$$J(\theta_1) = 0.58$$

If $\theta_1 = 0.05$

$h(\theta)^1 = 0.05 \times 1$, $h(\theta)^2 = 0.05 \times 2$, $h(\theta)^3 = 0.05 \times 3$

$= 0.05$, $= 0.1$, $= 0.15$

$J(\theta_1) = \frac{1}{6}[(0.05-1)^2 + (0.1-2)^2 + (0.15-3)^2]$

$= 2.105$

If $\theta_1 = 1.5$

$h(\theta)^1 = 1.5 \times 1$, $h(\theta)^2 = 1.5 \times 2$, $h(\theta)^3 = 1.5 \times 3$

$= 1.5$, $= 3$, $= 4.5$

$J(\theta_1) = \frac{1}{6}[(1.5-1)^2 + (3-2)^2 + (4.5-3)^2]$

$= \frac{1}{6}[0.25 + 1 + 2.25]$

$= 0.58$

If $\theta_1 = 2$

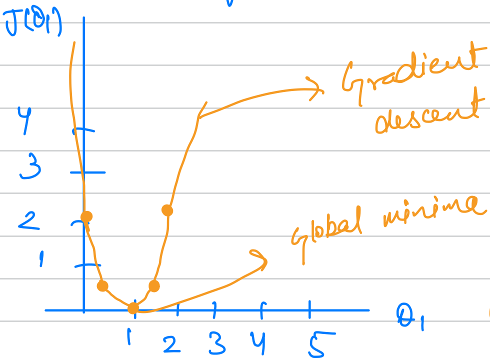
$h(\theta)^1 = 2 \times 1$, $h(\theta)^2 = 2 \times 2$, $h(\theta)^3 = 2 \times 3$

$= 2$, $= 4$, $= 6$

$J(\theta_1) = \frac{1}{6}[(2-1)^2 + (4-2)^2 + (6-3)^2]$

$= 2.33$

Let us plot graph b/w $J(\theta)$ Vs θ_1

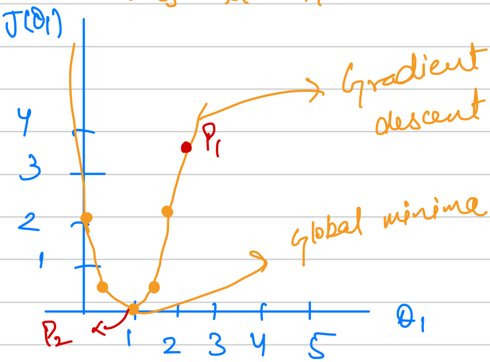


Our main is to find Global minima.

* THERE WON'T BE ANY CASE WHERE WE WILL HAVE 0 AS GM_1 AIM IS TO GET MIN. VALUE OF THAT

⇒ Another question that arises here is that how we'll reach from P_1 to P_2, imagining our first $J(\theta_1)$ value was at P_1

Let



The answer to above question is Repeat convergence theorem

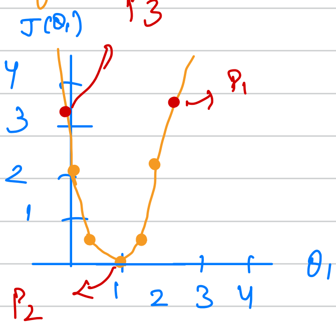
⇒ Repeat convergence theorem

$$\theta_j := \theta_j - \alpha \boxed{\frac{\partial \left(J(\theta_1) \right)}{\partial \theta_j}} \rightarrow \text{slope}$$

lel

learning rate

decide the size of steps of points



If slope is +ve

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

The updated value will ↓ses & move towards P_2

Let instead of P_1 we get P_3 as 1st value. Now

Slope is −ve

$$\theta_j := \theta_j - \alpha(-ve) \rightarrow \text{Updated value will} \uparrow \text{se.}$$