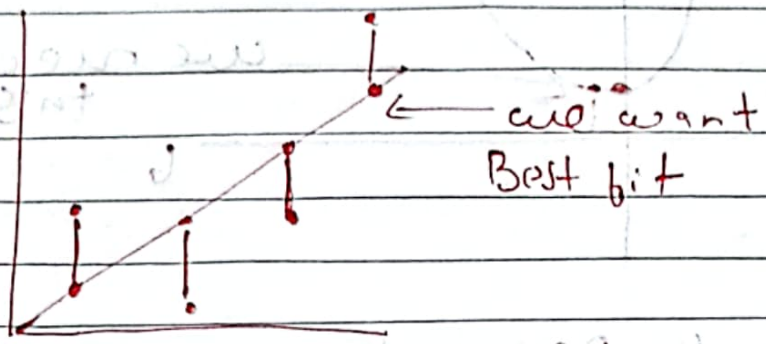
Gradient Descent

Intution



- Loss fon $L = \sum_{i=1}^{4} (y_i - \hat{y}_i)^2$

we know that $\hat{y}_i = mx_i + b$

- $L = \sum_{i=1}^{n} (y_i - mx_i - b)^2$

so Loss fonc is dependent on $\boxed{L(m, b)}$

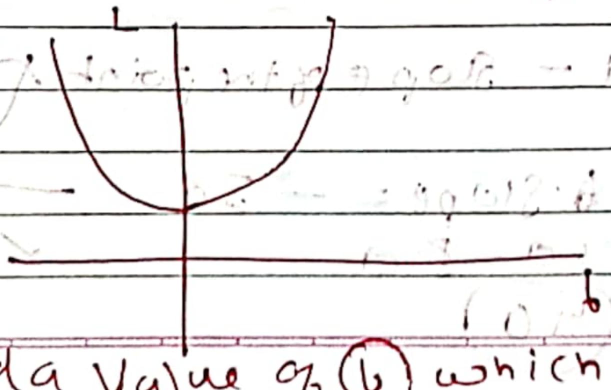
slope intercept

Cgpa	Lpa
–	–
–	–
–	–
–	–

for understanding Lets say $m = 78.35$

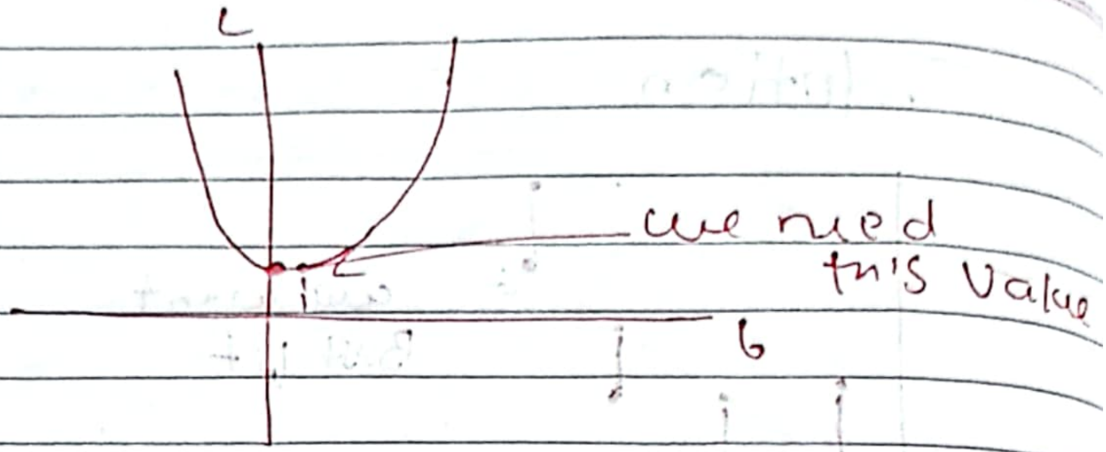
$L = \sum_{i=1}^{4} (y_i - 78.35 * x_i - b)^2$

we need the value of (b) for (L)



because $L \rightarrow b^2$

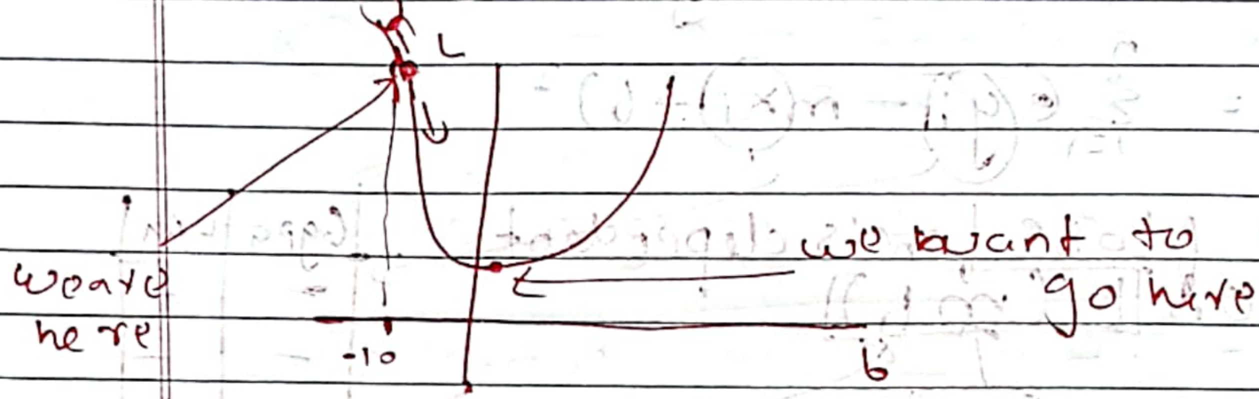
we need a value of (b) which make (L) minimum



In Gradient Descent

Step (1): Select a Random (b)

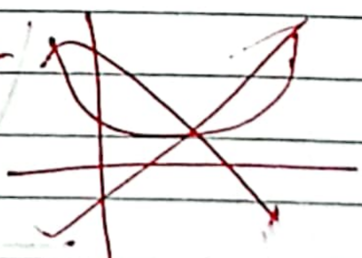
Let say $b = -10$



So How would a Machine will Know if we have to Move forward or backward.
Answer is: By using Slope

If Slope is +ve we have to decrease (b)
If slope is -ve we have to increase (b)

$b_{new} = b_{old}$ − slope of the point

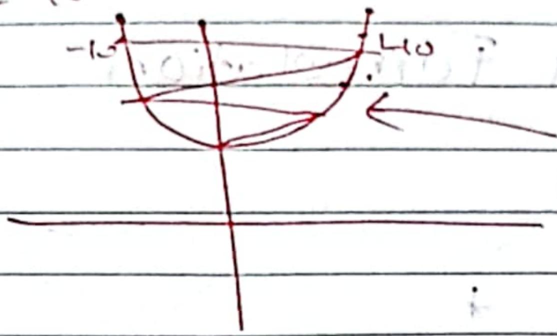


Let say slope = −50

$b_{new} = -10 - 50$

$b_{new} = 40$

$b_{new} = 40$



we use

Learning rate

because we don't want this to happen.

So

$b_{new} = b_{old} - \alpha * \text{slope (partial derivative)}$

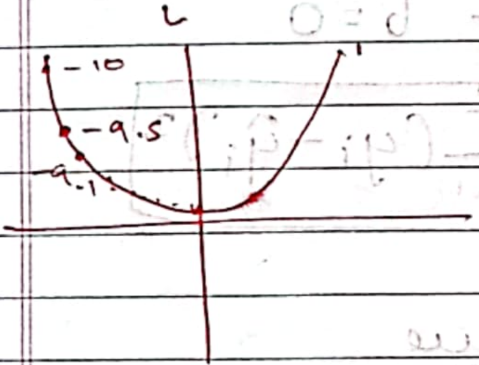
Learning rate

$\alpha = 0.01$

$b_{new} = -10 - (0.01 \times -50)$

$-10 + (0.01 \times 50)$

$= \boxed{9.5}$



let say slope at -9.5 is -40

$b_{new} = (-9.5) - (0.01 \times -40$

$= -9.5 + (0.01 \times 40)$

$= -9.5 - 0.4$

$= \boxed{-9.1}$

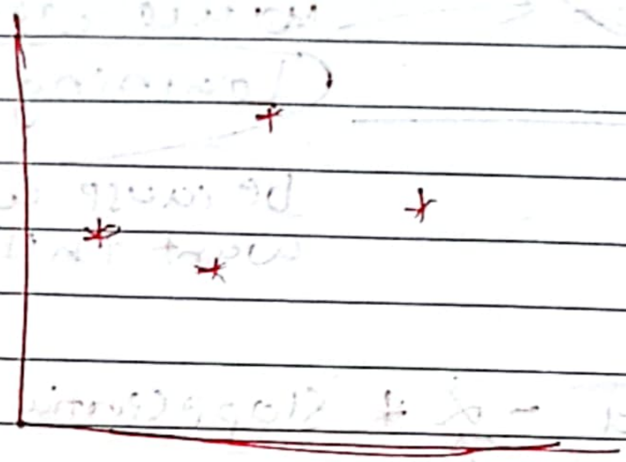
* When to stop because it is a Iterative process

① Approach ① diff b/w bold and bnew should be => 0.000

② Iteration limit

Epochs

* Mathematical Formulation



Lets say,

$m = 78.35$

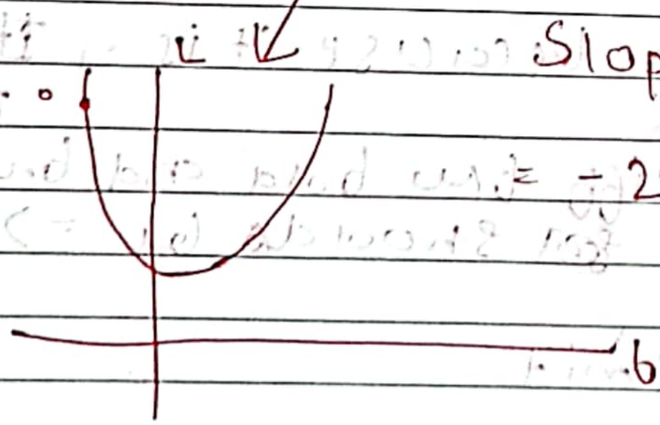
Step (1) Start with Random Value of (b)

Lets say $b = 0$

How to find slope at $b=0$

as we know $\boxed{\text{Loss func} = \sum_{i=1}^{n}(y_i-\hat{y}_i)^2}$

when we plot this



Slope eq is

$$= -2\sum_{i=1}^{n}(y_i - mx_i - b)$$

Page No.
Date

$b = 0$

$$\text{Slope} = -2 \sum_{i=1}^{n} (y_i - mx_i - b)$$

(Partial derivative)

$$= -2 \sum_{i=1} (y_i - mx_i - 0)$$

$$\text{Slope} = -2 \sum_{i=1}^{n} (y_i - 78.35x_i - 0)$$

Every target value

Every independent value

Then we will find b_{new} using

$$b_{new} = b_{old} - \alpha(\text{slope})$$

$$= 0 - 0.01 \text{ output}$$

b_{new} = Value

from this value we will find ~~new~~ bnew slope then again bnew until the given (epochs)

iteration

Lets say epoch = 10
then we will do this 10 time

Full intution

Now ~~we~~ we find m and b both

Step ① :- Initialize random value for m and b

lets ~~s~~ $m = 1$ and $b = 0$

Epochs = 100 , $\alpha = 0.01$

for i in epochs
  $b = b - \alpha \text{ slope}$
  $m = m - \alpha \text{ slope}$

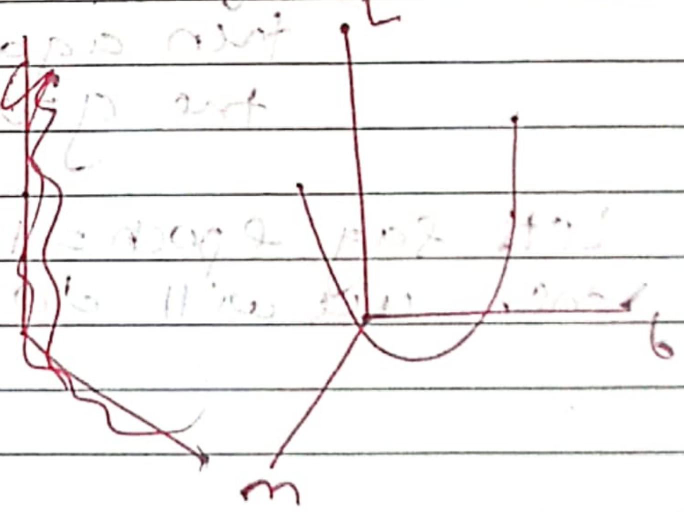
So,

$$L = \Sigma (y_i - \hat{y}_i)^2$$

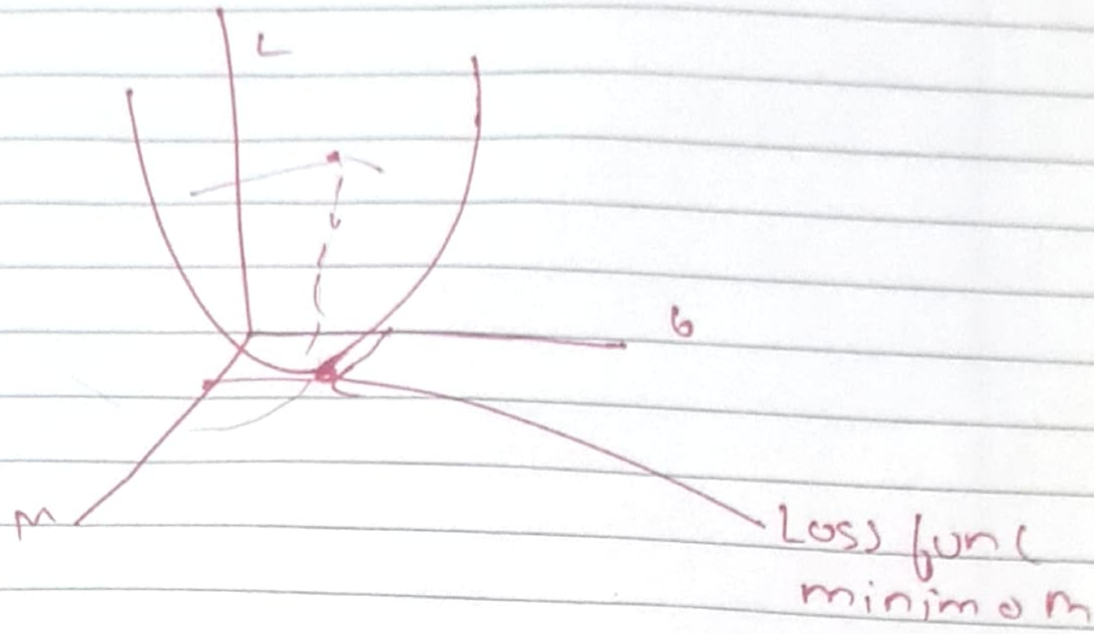
$$L = \Sigma (y_i - m x_i - b)^2$$

Now loss function depent on m and b

$L(m\ b)$



see 3d graph in github repo



Now slope formula will be

$$b\_slope = \frac{\partial L}{\partial b}$$

$$m\_slope = \frac{\partial L}{\partial m}$$

partial Darivate of the Loss function at Point m and b

$$\Sigma(y_i - mx_i - b)^2$$

Darivate ~~with~~ w.r.t 'b'	Darivate w.r.t 'm'
$\frac{\partial L}{\partial b} = \Sigma(y_i - mx_i - b)^2$	$\frac{\partial L}{\partial m} = -2\Sigma(y_i - mx_i - b$
$\frac{\partial L}{\partial b} = -2\Sigma(y_i - mx_i - b)$	$\frac{\partial L}{\partial m} = \boxed{-2\Sigma(y_i - mx_i - b)x_i}$