LINEAR REGRESSION

Training Data

x Time spent	Marks y
1	4
3	7
10	8
20	10

$x \in R$

$y \in R$

Test Data

8 hrs? marks?

?

Pattern?

x	x_1 Time spent	x_2 Attendance	x_3 Class Perf.	Marks y

input has 3 features

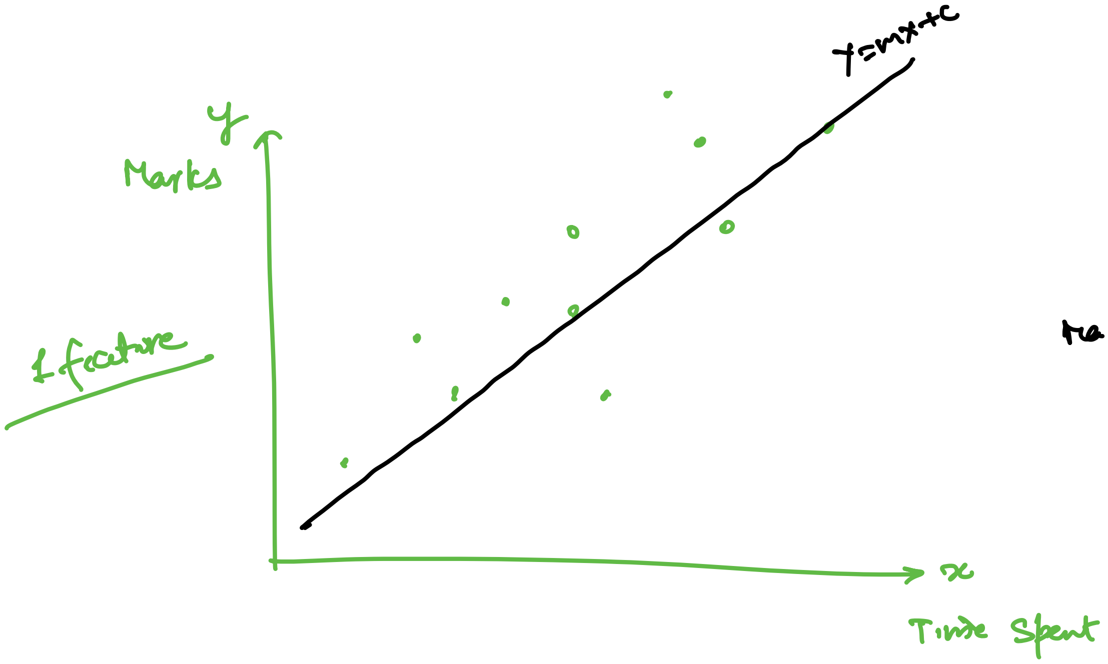
$x \in R^3$

$y \in R$

$x \in R^n$

n input features

$y \in R$



$Y = mx + C$

Marks

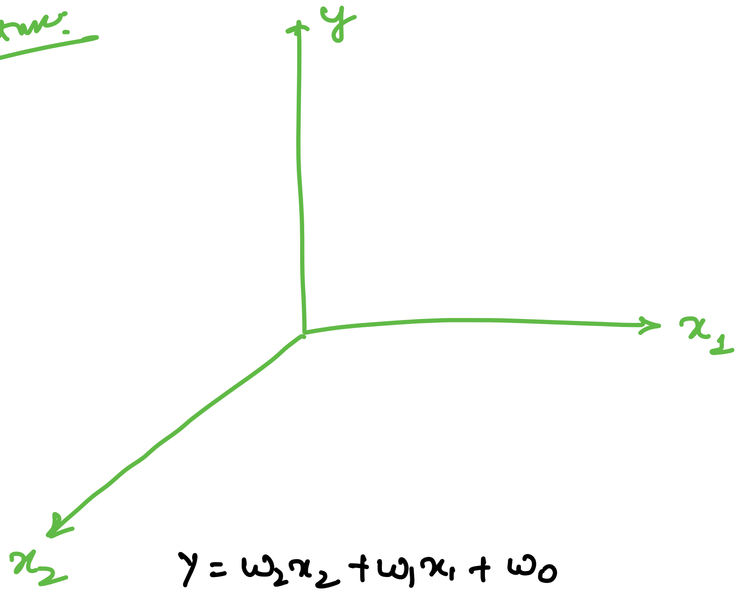
Time Spent

m, c ?

$Y = m.8 + c$

$Y = w_1 \cdot x + w_0$

2 feature:



$y = w_2x_2 + w_1x_1 + w_0$

3 features

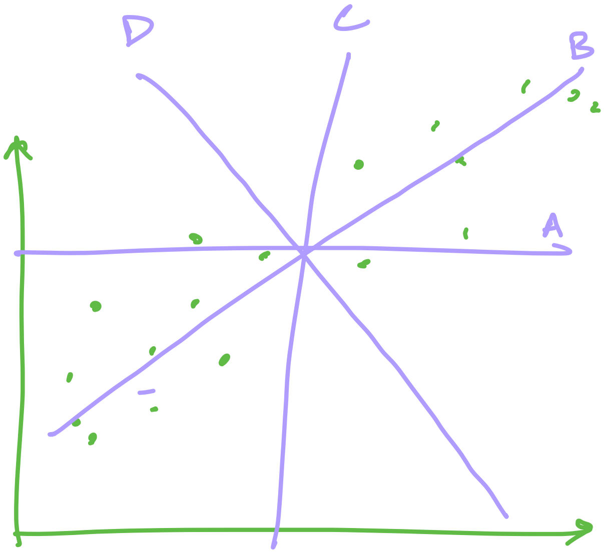
4D

$Y = \omega_3 x_3 + \omega_2 x_2 + \omega_1 x_1 + \omega_0$

$Y = w_1 x + w_0 \rightarrow$ prediction

$$\omega = \begin{bmatrix} w_0 \\ w_1 \end{bmatrix}$$

Hypothesis: $h_w(x) = w_1 x + w_0$



Aim:

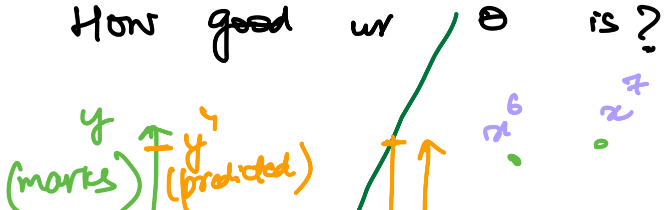
To learn best line which fits through data points?

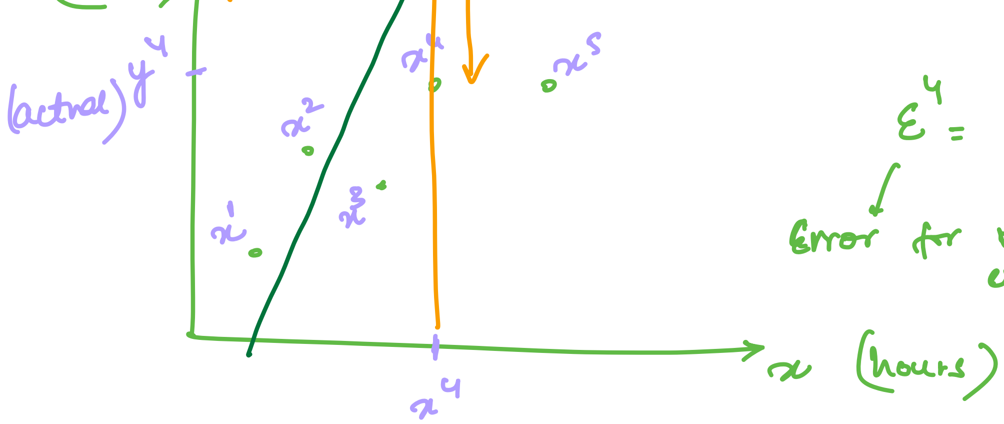
ⓐ Random value of $\boxed{w_0, w_1} \rightarrow$ line

ⓑ How good the line is? =

ⓒ w_0, w_1 update good performance =

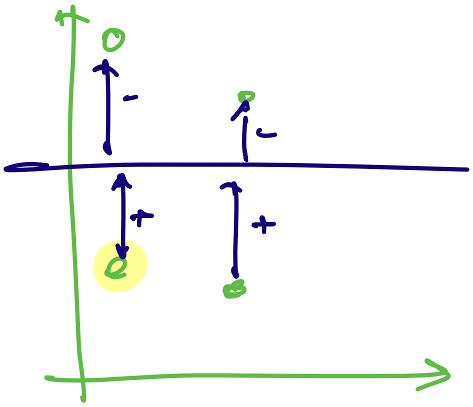
How good ur θ is?





$$\varepsilon^{4} = |y^{4}_{pred} - y^{4}_{actual}|$$

Error for 4th example



$$\varepsilon^{(i)} = |y^{(i)}_{pred} - y^{(i)}_{actual}|$$

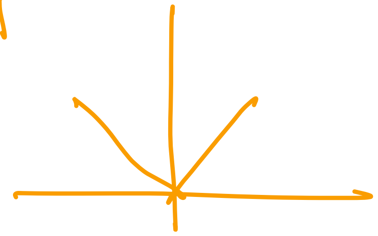
Total error for all data points = $\sum_{i=1}^{m} |y^{(i)}_{pred} - y^{(i)}_{actual}|$

Total error for all data points = $\sum_{i=1}^{m} |\hat{y}^{(i)} - y^{(i)}|$

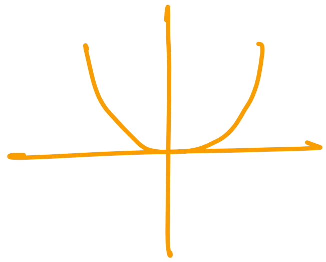
Average Error = $\frac{1}{m}\sum_{i=1}^{m} |\hat{y}^{(i)} - y^{(i)}|$

Average Absolute Error

$|x|$



x^2



---

mean squared error (MSE) $= \frac{1}{m}\sum_{i=1}^{m}\left[\underbrace{\hat{y}^{(i)}}_{\text{predicted}} - \underbrace{y^{(i)}}_{\text{actual}}\right]^2$

m = # data points

minimize

Objective / Loss function
J

$$J = \frac{1}{m}\sum_{i=1}^{m}\left[\hat{y}^{(i)} - y^{(i)}\right]^2$$

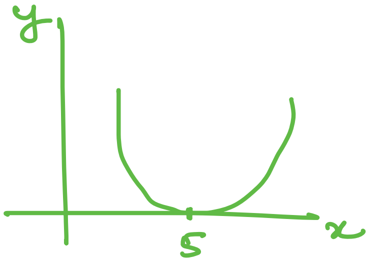
$$J(w) = \frac{1}{m}\sum_{i=1}^{m}\left[w_1 x^{(i)} + w_0 - y^{(i)}\right]^2$$

make updation to w_0, w_1, so that it becomes better.

GRADIENT DESCENT (in General)

Way1: $y = (x-5)^2$

for what value of x y is min?



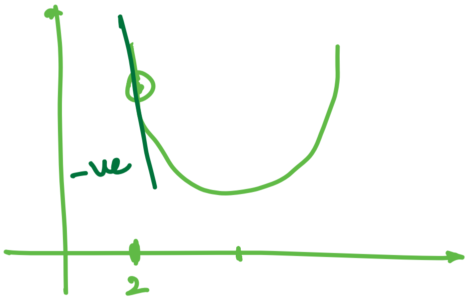
$$\frac{dy}{dx} = 0$$

$$\frac{d(x-5)^2}{dx} = 0$$

$$2(x-5) = 0$$

$$x = 5$$

Way2:



(learning rate)

⊙ magnitude η

⊙ drⁿ which is minimizing y $\frac{dy}{dx}$

Gradient Descent

$$x = x - \eta \frac{dy}{dx}$$

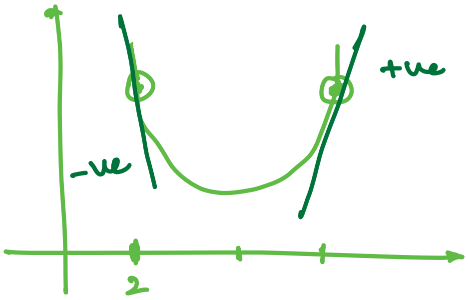
→ Gradient / slope

-ve

+ve

$x = x +$ something

x increase



$$x = x - \eta \boxed{\frac{dy}{dx}}$$

+ve

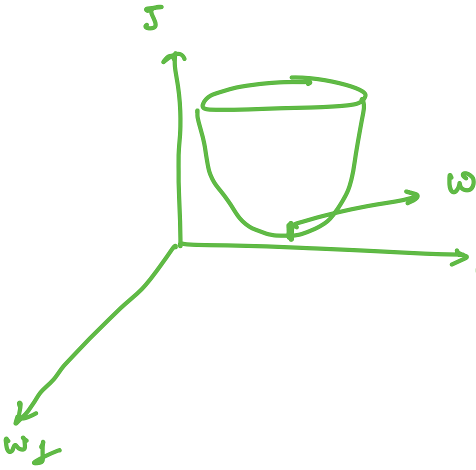
-ve

$x = x -$ something

x decreases.

$$J(w) = \frac{1}{m} \sum_{i=1}^{m} \left[w_1 x^{(i)} + w_0 - y^{(i)} \right]^2$$

$$w = \begin{bmatrix} w_0 \\ w_1 \end{bmatrix}$$



w_0, w_1 such that J is minimum

$$w_0 = w_0 - \eta \frac{\partial J(w)}{\partial w_0} \qquad w_1 = w_1 - \eta \frac{\partial J(w)}{\partial w_1}$$