# RECAP

Dataset :

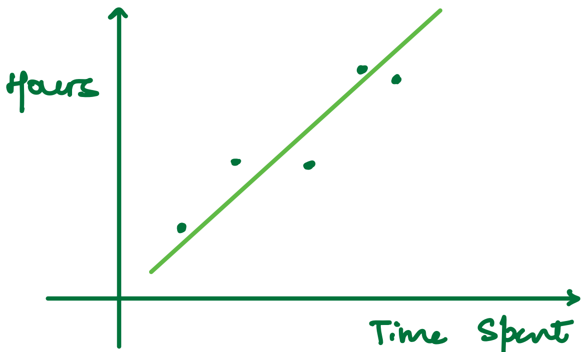
Training Data

| Time Spent ($x$) | Hours ($y$) |
|---|---|
| 1 | 4 |
| 3 | 7 |
| 10 | 8 |
| 20 | 10 |

Continuous value: Task of Regression

Test Data

| Time Spent ($x$) | Hours ($y$) |
|---|---|
| 15 hrs | ? |




Try to fit a line

Equation of line: $y = mx + c$

Need to find out $m$ & $c$.

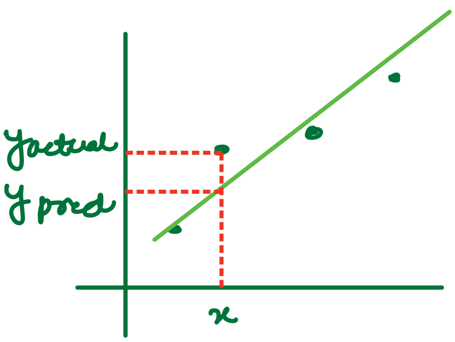
$$\underbrace{y = \theta_1 x + \theta_0}_{\text{hypothesis}}$$

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

how to find $\theta_0$ and $\theta_1$ ?

- Start with random value of $\theta_0$ and $\theta_1$
- Check how good your $\theta$ is ?
- Update $\theta$ so that it becomes a better $\theta$.

To check how good our $\theta$ is we use ERROR.




ERROR $= y_{pred} - y_{actual}$

+ve and -ve error cancels out. Use modulus

$= |y_{pred} - y_{actual}|$

for 1 datapoint

$$\text{Total Error} = \sum_{i=1}^{m} \left| y_{pred}^{(i)} - y_{actual}^{(i)} \right|$$

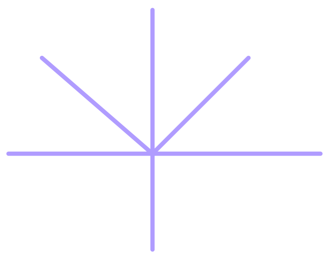
Sum of error for all datapoints

$$\text{Mean Absolute Error} = \frac{1}{m}\sum_{i=1}^{m} \left| y_{pred}^{(i)} - y_{actual}^{(i)} \right|$$

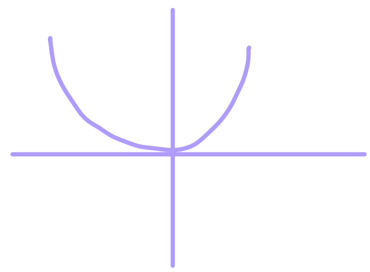
Avg error for 1 datapoint

**PROBLEM:** not differentiable

$|x|$



$x^2$



$$\underbrace{\text{Mean Squared Error (MSE)}}_{\text{Loss or Error function}} = \frac{1}{m}\sum_{i=1}^{m} \left( y_{pred}^{(i)} - y_{actual}^{(i)} \right)^2$$

$$MSE = \frac{1}{m}\sum_{i=1}^{m} \left( \hat{y}^{(i)} - y^{(i)} \right)$$

Predicted

Actual

$$\hat{y}^{(i)} = h_\theta(x^{(i)}) = \underbrace{\theta_1 x^{(i)} + \theta_0}_{\text{Equation of line}}$$

Prediction

Hypothesis for datapoint $x^{(i)}$

Loss or Error function

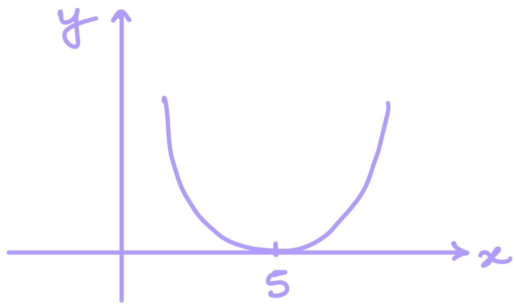
$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m} \left( \theta_1 x^{(i)} + \theta_0 - y^{(i)} \right)$$

**Update θ so that it becomes a better θ.**

## GRADIENT DESCENT (in General)

**Way 1:**

$y=(x-5)^2$

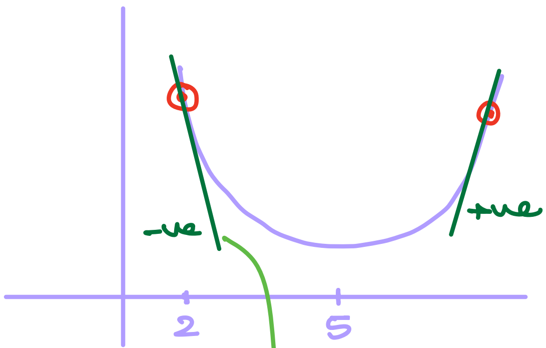



For what value of x, y is minimum?

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

$$x=5$$

**Way 2:**




① magnitude $\eta$

② dir$^n$ which is minimizing y $\frac{dy}{dx}$

learning rate

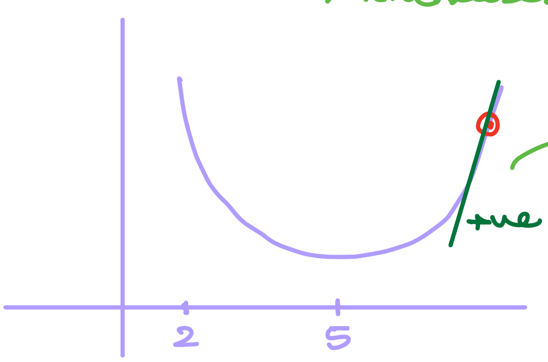
$$x = x - \eta \frac{\partial y}{\partial x}$$

Gradient / Slope

-ve +ve

$x = x +$ something

x increases




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

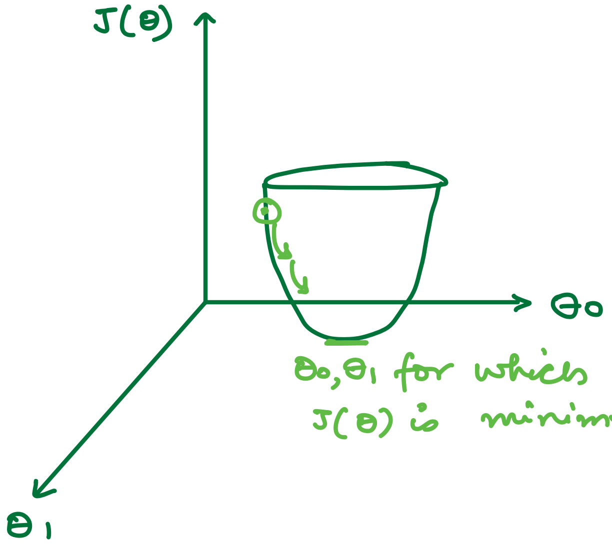
+ve

-ve

x decreases

Applying similar logic for $J(\theta)$

$$J(\theta) = \frac{1}{m} \sum_{i=1}^{m} \left(\theta_0 + \theta_1 x^{(i)} - y^{(i)}\right)^2 \qquad \theta = \begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix}$$




$$\left.\theta = \theta - \eta \frac{\partial J(\theta)}{\partial \theta}\right\} \text{Gradient Descent Equation}$$

$$\frac{\partial J(\theta)}{\partial \theta} \rightarrow \frac{\partial J(\theta)}{\partial \theta_0}, \quad \frac{\partial J(\theta)}{\partial \theta_1}$$

1 $$\frac{\partial J(\theta)}{\partial \theta_0} = \frac{\partial}{\partial \theta_0} \frac{1}{m} \sum_{i=1}^{m} \left[\left(\theta_0 + \theta_1 x^{(i)}\right) - y^{(i)}\right]^2$$

$$= \frac{1}{m} \sum_{i=1}^{m} \frac{\partial}{\partial \theta_0} \left[\left(\theta_0 + \theta_1 x^{(i)}\right) - y^{(i)}\right]^2$$

$$= \frac{1}{m} \sum_{i=1}^{m} 2\left[\left(\theta_0 + \theta_1 x^{(i)}\right) - y^{(i)}\right]$$

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

(2) $$\frac{\partial J(\theta)}{\partial \theta_1} = \frac{\partial}{\partial \theta_1} \frac{1}{m} \sum_{i=1}^{m} \left[ \theta_0 + \theta_1 x^{(i)} - y^{(i)} \right]^2$$

$$= \frac{1}{m} \sum_{i=1}^{m} 2 \left[ \theta_0 + \theta_1 x^{(i)} - y^{(i)} \right] x^{(i)} \rightarrow \text{grad 1}$$

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

lr.

$$\theta_0 = \theta_0 - \frac{\eta}{m} \sum_{i=1}^{m} 2 \left[ \hat{y}^{(i)} - y^{(i)} \right]$$

$$\theta_1 = \theta_1 - \frac{\eta}{m} \sum_{i=1}^{m} 2 \left[ \hat{y}^{(i)} - y^{(i)} \right] x^{(i)}$$

## Algo:

$\theta_0, \theta_1$ random value

do

{

loss/error $fx^n$ $\longrightarrow \frac{1}{m} \sum_{i=1}^{m} \left( \theta_0 + \theta_1 x^{(i)} - y^{(i)} \right)^2$

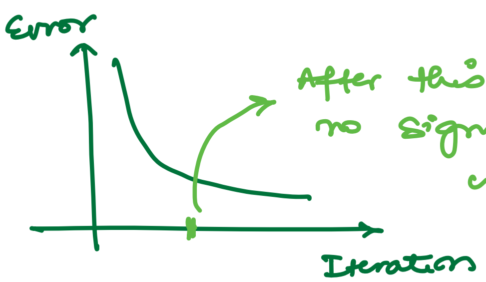
update $\theta_0, \theta_1$

} while (convergence)

?

(1) iterations fix

(2) error / loss $fx^n$ plot




After this iteration no significant decrease in loss. So stop.

## Test Time: Student studied for 15 hrs, marks?

Using the above algo we found the value of $\theta_0, \theta_1$

we have the eq$^n$ of line : $\hat{y} = h_\theta(x) = \theta_0 + \theta_1 x$

here $x$ is 15

just put $x = 15$ in eq$^n$ and you will get $\hat{y}$.

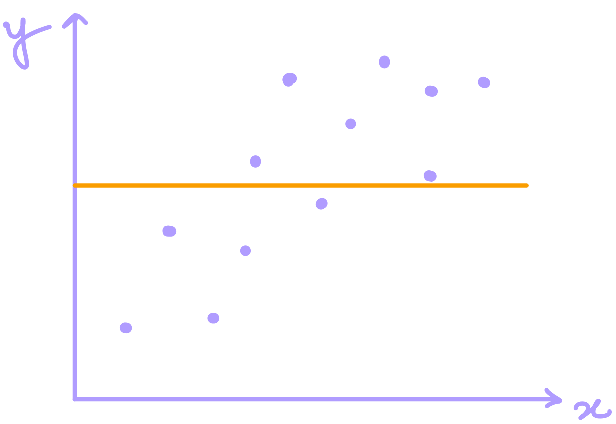
## METRIC

### R2 Score:

(R Squared or Coefficient of Determination)

$$R2\ Score = 1 - \frac{\sum_{i=1}^{m} \left(y^{(i)} - \hat{y}^{(i)}\right)^2}{\sum_{i=1}^{m} \left(y^{(i)} - y^{(i)}_{avg}\right)^2}$$

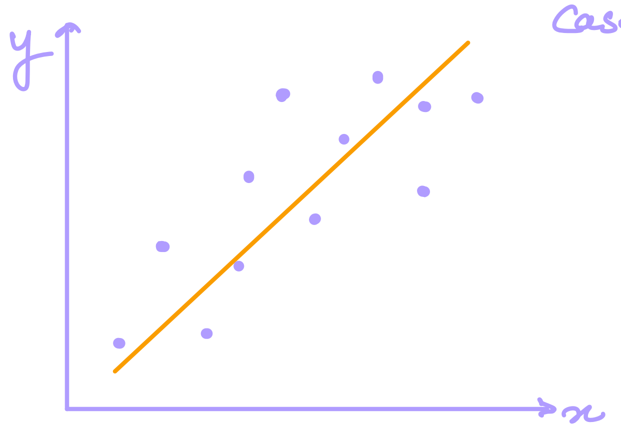
Worst Case




$\hat{y}^{(i)} = y^{(i)}_{avg}$

R2 Score = 1-1=0

Best Case




$\hat{y}^{(i)} = y^{(i)}$

R2 Score = 1-0=1