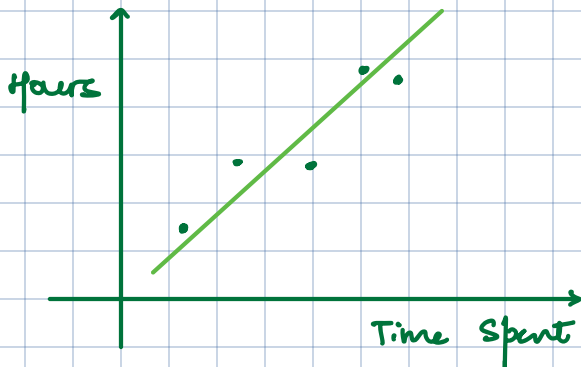


RECAP

Dataset:	x Time Spent	y Hours
Training Data	1	4
	3	7
	10	8
	20	10
Test Data	15 hrs	?

continuous value:
Task of Regression



Try to fit a line

Equation of line: $y = mx + c$

Need to find out m & c .

$$y = \theta_1 x + \theta_0$$

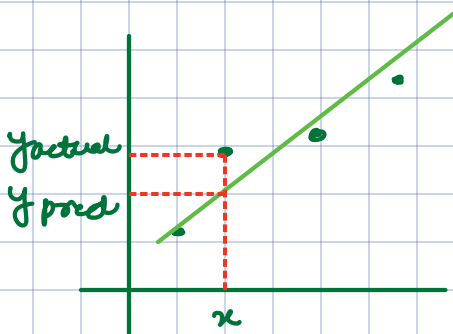
hypothesis

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

how to find θ_0 and θ_1 ?

- Start with random value of θ_0 and θ_1
- Check how good your θ is?
- update θ so that it becomes a better θ .

To check how good our θ is we use ERROR.



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

+ve and -ve error cancels out. Use modulus

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

for 1 datapoint

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

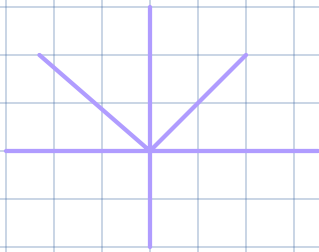
Sum of error for all datapoints

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

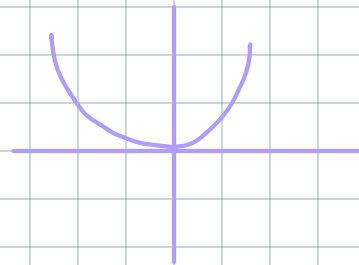
Avg error for 1 datapoint

PROBLEM: not differentiable

$|x|$



x^2



$$\text{Mean Squared Error (MSE)} = \frac{1}{m} \sum_{i=1}^m (y_{\text{pred}}^{(i)} - y_{\text{actual}}^{(i)})^2$$

loss or Error function

$$\text{MSE} = \frac{1}{m} \sum_{i=1}^m (\hat{y}^{(i)} - y^{(i)})^2$$

Predicted Actual

$$\hat{y}^{(i)} = h_{\theta}(x^{(i)}) = \theta_1 x^{(i)} + \theta_0$$

Prediction Hypothesis for datapoint $x^{(i)}$ Equation of line

loss or Error function

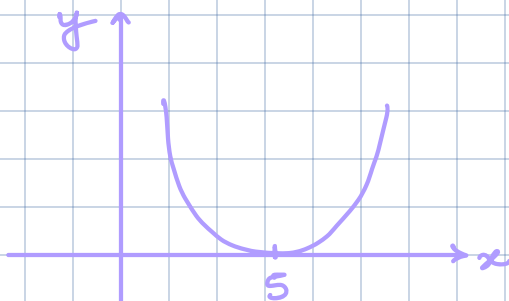
$$J(\theta) = \frac{1}{m} \sum_{i=1}^m (\theta_1 x^{(i)} + \theta_0 - y^{(i)})^2$$

update θ so that it becomes a better θ .

GRADIENT DESCENT (in General)

Way 1:

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

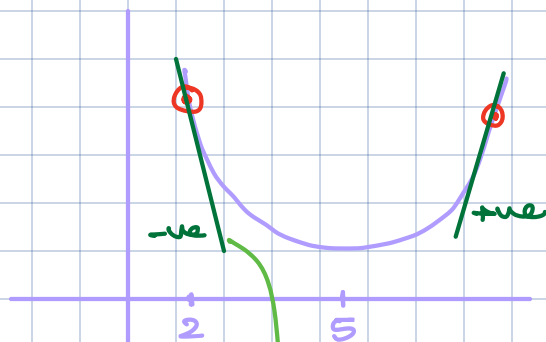


For what value of x , y is minimum?

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

$$x = 5$$

Way 2:



① magnitude η

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

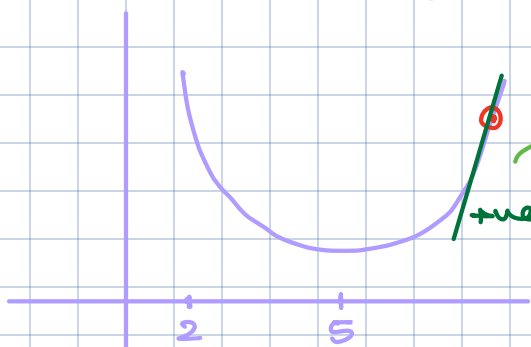
learning rate

$$x = x - \eta \left(\frac{dy}{dx} \right)$$

Gradient / slope

-ve +ve

$x = x + \text{something}$
 x increases



$$x = x - \eta \left(\frac{dy}{dx} \right)$$

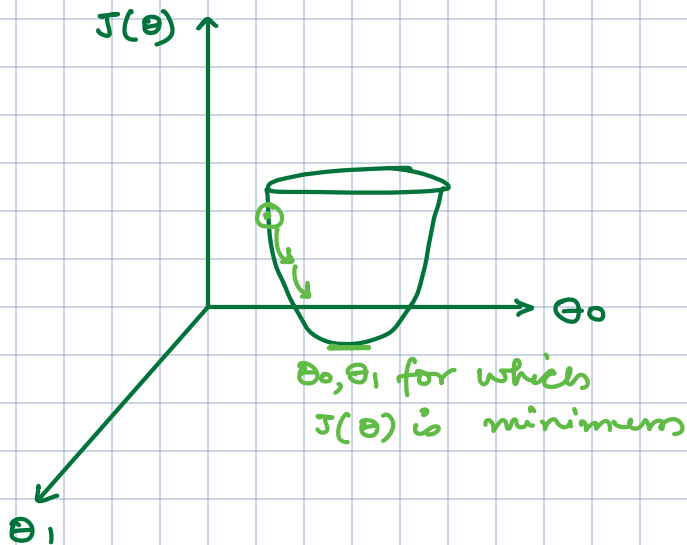
+ve -ve

x decreases

Applying similar logic for $J(\theta)$

$$J(\theta) = \frac{1}{n} \sum_{i=1}^n (\theta_0 + \theta_1 x^{(i)} - y^{(i)})^2$$

$$\theta = \begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix}$$



$$\theta = \theta - \eta \frac{\partial J(\theta)}{\partial \theta} \quad \left. \vphantom{\frac{\partial J(\theta)}{\partial \theta}} \right\} \begin{array}{l} \text{Gradient} \\ \text{Descent} \\ \text{Equation} \end{array}$$

$$\begin{array}{c} \frac{\partial J(\theta)}{\partial \theta} \\ \swarrow \quad \searrow \\ \frac{\partial J(\theta)}{\partial \theta_0} \quad \frac{\partial J(\theta)}{\partial \theta_1} \end{array}$$

$$\begin{aligned} \textcircled{1} \quad \frac{\partial J(\theta)}{\partial \theta_0} &= \frac{\partial}{\partial \theta_0} \frac{1}{n} \sum_{i=1}^n [(\theta_0 + \theta_1 x^{(i)}) - y^{(i)}]^2 \\ &= \frac{1}{n} \sum_{i=1}^n \frac{\partial}{\partial \theta_0} [(\theta_0 + \theta_1 x^{(i)}) - y^{(i)}]^2 \\ &= \frac{1}{n} \sum_{i=1}^n 2 [(\theta_0 + \theta_1 x^{(i)}) - y^{(i)}] \\ &= \frac{1}{n} \sum_{i=1}^n 2 [\hat{y}^{(i)} - y^{(i)}] \end{aligned}$$

$$\begin{aligned}
 \textcircled{2} \quad \frac{\partial J(\theta)}{\partial \theta_1} &= \frac{\partial}{\partial \theta_1} \frac{1}{m} \sum_{i=1}^m \left[\theta_0 + \theta_1 x^{(i)} - y^{(i)} \right]^2 \\
 &= \frac{1}{m} \sum_{i=1}^m 2 \left[\theta_0 + \theta_1 x^{(i)} - y^{(i)} \right] x^{(i)} \rightarrow \text{grad} \\
 &= \frac{1}{m} \sum_{i=1}^m 2 \left[\hat{y}^{(i)} - y^{(i)} \right] x^{(i)}
 \end{aligned}$$

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

lr.

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

Algo:

θ_0, θ_1 random value

do
{

loss/error $f(x)$ $\rightarrow \frac{1}{m} \sum_{i=1}^m (\theta_0 + \theta_1 x^{(i)} - y^{(i)})^2$

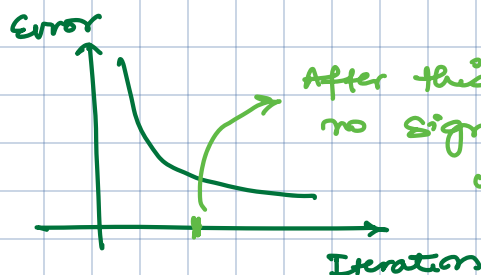
update θ_0, θ_1

} while (convergence)

?

① iterations fix

② error / loss $f(x)$ plot



Test Time: Student studied for 15 hrs, marks?

Using the above algo we found the value of Θ_0, Θ_1

we have the eqⁿ of line: $\hat{y} = h_{\Theta}(x) = \Theta_0 + \Theta_1 x$

here x is 15

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

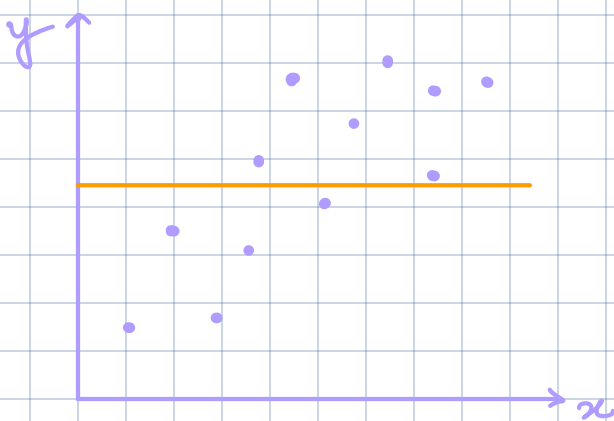
METRIC

R2 Score:

(R Squared or Coefficient of Determination)

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

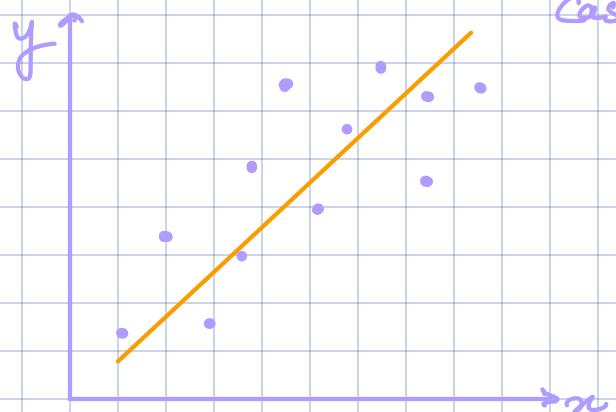
Worst Case



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

$$R^2 \text{ Score} = 1 - 1 = 0$$

Best Case



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

$$R^2 \text{ Score} = 1 - 0 = 1$$