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

① ? ②

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

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

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

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

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

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

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

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

learning rate

$$w_0 = w_0 - \frac{\eta}{m} \sum_{i=1}^{m} 2 \left[\hat{y}^{(i)} - y^{(i)} \right]$$

$$w_1 = w_1 - \frac{\eta}{m} \sum_{i=1}^{m} 2 \left[\hat{y}^{(i)} - y^{(i)} \right] x^{(i)}$$

Algo :

w_1, w_0 random value

do

{

loss/error fx^n → how good w_0, w_1 is ? → $\frac{1}{m}\sum_{i=1}^{m}(w_0+w_1x^{(i)} - y^{(i)})^2$

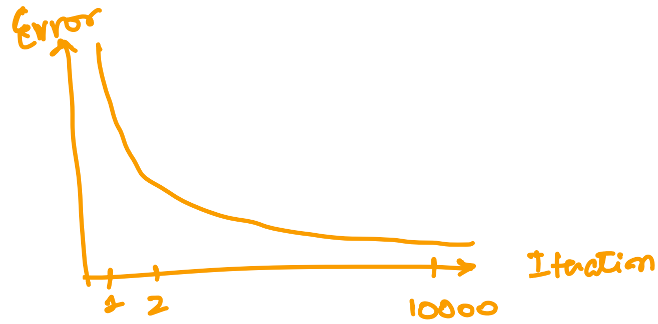
update w_0, w_1

} while (convergence)

?

① iterations fix

② error/loss fx^n plot



Test :

→ actually X

Hours : 8 hrs? Marks ?

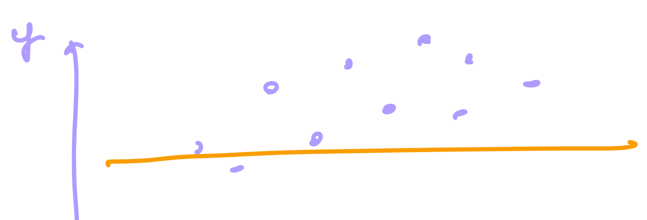
w_0, w_1 $\quad w_0 + w_1 \times 8$ = ______ → marks → S1, S2, S3

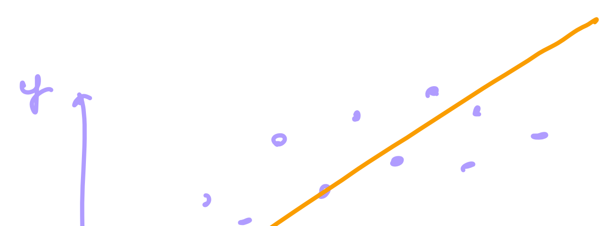
Metric

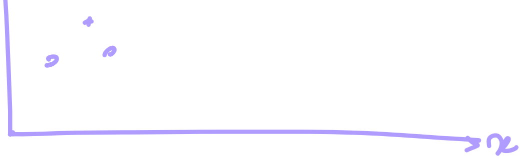
R2 Score

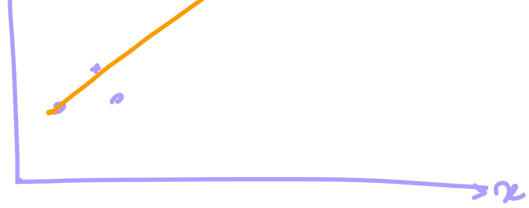
(R squared or coefficient of Determination)

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









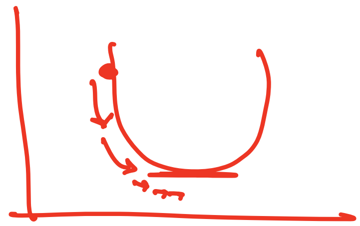
$\hat{y}^{(i)} = y^{(i)}_{avg}$

R2 Score = 1-1=0

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

R2 Score = 1-0=1

adaptive l.r



A travel agency wants an automated system to predict travel costs. The agency has the following data available with it.

(x) Table II (y)

	S. No.	Distance (in Km)	Travelling Cost (in Rupees)
x^1	1	1	2.75
x^2	2	2	3.5
x^3	3	3	4.25
x^4	4	4	5
x^5	5	5	5.75

Formulate the above problem as a linear model $h(x) = w_0 + w_1 x$ to predict the travelling cost for a given distance. The parameter w_0 is 2 (optimal). Apply gradient descent algorithm to find optimal parameter w_1. The learning rate for the first epoch is 0.073, and for the second epoch and later, the learning rate is 0.091. Let the initial value of w_1 is 0.5.

$h(x) = 2 + w_1 x$

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

$$= w_1 - \frac{\eta}{m} \cdot 2 \sum_{i=1}^{m} (\hat{y}^{(i)} - y^{(i)})\, x^{(i)}$$

initialize $w_1 \rightarrow 0.5$

do

{

$J(w)$

update w → 1st epoch: $\eta = 0.073$

2nd epoch $\eta = 0.091$

}

while (convergence)

$x^{(i)}$	$h_w(x^{(i)})$ / $\hat{y}^{(i)}$	$y^{(i)}$	$\hat{y}^{(i)} - y^{(i)}$	$(\hat{y}^{(i)} - y^{(i)})x^{(i)}$
1	$0.5 * 1 + 2 = 2.5$	2.75	-0.25	-0.25
2	$0.5 * 2 + 2 = 3$	3.5	-0.5	-1
3	$0.5 * 3 + 2 = 3.5$	4.25	-0.75	-2.25
4	$0.5 * 4 + 2 = 4$	5	-1	-4
5	$0.5 * 5 + 2 = 4.5$	5.75	-1.25	-6.25

$\Sigma = -13.75$

$$\sum_{i=1}^{m} (\hat{y}^{(i)} - y^{(i)})x^{(i)}$$

$$\hat{y}^{(i)} = w_1 x^{(i)} + w_0$$

$$\hat{y}^{(i)} = 0.5\, x^{(i)} + 2$$

$$w_1 = w_1 - \frac{\eta * 2}{m} \sum_{i=1}^{m} (\hat{y}^{(i)} - y^{(i)})x^{(i)}$$

$$w_1 = 0.5 - \frac{0.073 * 2}{5} (-13.75)$$

$$w_1 = 0.9$$

2nd epoch

now $w_1 = 0.9$ instead of 0.5

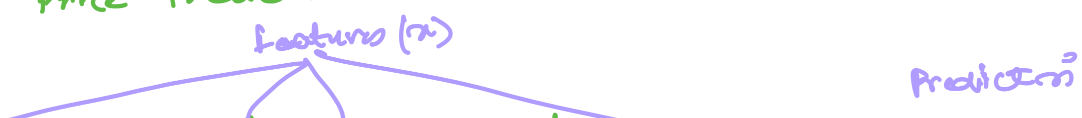
and $\eta = 0.091$ instead of 0.073

[

final w_1.

Linear Regression with multiple features:

Eg: House Price Prediction



x_0		Area	# Floors	# Bedrooms	# Age	Price (y)
1	x^1	200 x_1^1	2 x_2^1	3 x_3^1	10 x_4^1	2.5 y^1
1	x^2	100 x_1^2	3 x_2^2	2 x_3^2	5 x_4^2	3 y^2
⋮		⋮				
1	x^m					

$$\hat{y}^{(i)} = \hat{w}_0^{(i)} + w_1 x_1^{(i)} + w_2 x_2^{(i)} + w_3 x_3^{(i)} + w_4 x_4^{(i)} \ldots\ldots w_n x_n^{(i)}$$

$$\hat{y}^{(i)} = w_0 x_0^{(i)} + w_1 x_1^{(i)} + w_2 x_2^{(i)} + w_3 x_3^{(i)} + w_4 x_4^{(i)} \ldots\ldots w_n x_n^{(i)}$$

x_0 will always be 1

$$\hat{y}^{(i)} = \sum_{j=0}^{n} w_j x_j^{(i)}$$

$$\frac{\partial J(w)}{\partial w_0}, \quad \frac{\partial J(w)}{\partial w_1} \quad ?$$

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

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

$$= \frac{1}{m}\sum_{i=1}^{m} \frac{\partial}{\partial w_j}\left(w_0 x_0^{(i)} + w_1 x_1^{(i)} + \ldots\ldots w_j x_j^{(i)} + \ldots - w_n x_n^{(i)} - y^{(i)}\right)^2$$

$$= \frac{1}{m}\sum_{i=1}^{m} 2\left(w_0 x_0^{(i)} + w_1 x_1^{(i)} + \ldots\ldots w_j x_j^{(i)} + \ldots - w_n x_n^{(i)} - y^{(i)}\right) x_j^{(i)}$$

$$\frac{\partial J(w)}{\partial w_j} = \frac{1}{m}\sum_{i=1}^{m} 2\left(\hat{y}^{(i)} - y^{(i)}\right) x_j^{(i)}$$