Applied Machine Learning.

Assignment 1.

Name: Muhammad Hassham

Roll No.: 19L-1380

Section: 8A

Submission Date: Feb 20, 2022

Q: Consider the following - - - least square problem.

d] Assume that the - - - using MATLAB or PYTHON.

Ans: Given,

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^{m} w^{(i)} \left(y^{(i)} - (\phi(x^{(i)}))^T \theta \right)^2 \quad -②$$

also given polynomial order is 1 and hence

$$\phi(x^i) = \begin{bmatrix} 1 \\ x^i \end{bmatrix}$$

$$\phi(x^i)^T = \begin{bmatrix} 1 & x^i \end{bmatrix}$$

we know, $\theta = \begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix}$

$$\phi(x^{(i)})^T \theta = \begin{bmatrix} 1 & x^i \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix} = \theta_0 + \theta_1 x^i \quad -①$$

Let $x = x_1^i$:

Now substituting ① in ②

and given that $w = \frac{1}{8}$ for all instances

$$J(\theta) = \frac{1}{2} \sum_{i=1}^{m} \frac{1}{8} \left(y^i - (\theta_0 + \theta_1 x^i) \right)^2$$

Taking $\frac{1}{8}$ common

$$J(\theta) = \frac{1}{2(8)} \sum_{i=1}^{m} \left(y^i - (\theta_0 + \theta_1 x^i) \right)^2$$

$$J(\theta) = \frac{1}{16} \sum_{i=}^{m} \left(y^i - (\theta_0 + \theta_1 x^i) \right)^2 \quad -③$$

To find θ_0 and θ_1, vector update rule for gradient descent can be used:

$$\theta_j = \theta_j - \alpha \frac{\partial}{\partial \theta} J(\theta) \quad -④$$

Substituting ③ in ④ for θ_0:

$$\theta_j = \theta_j - \alpha \frac{\partial}{\partial \theta_0}\left[\frac{1}{16}\sum_{i=1}^{m}\left(y^i - (\theta_0 + \theta_1 x^i)\right)^2\right]$$

$$\theta_j = \theta_j - \frac{\alpha}{16}\cdot 2\left(\sum_{i=1}^{m}\left(y^i - (\theta_0 + \theta_1 x^i)\right)^2 \cdot (-1)\right)$$

$$\theta_j = \theta_j + \frac{\alpha}{8}\sum_{i=1}^{m}\left(y^{(i)} - (\theta_0 + \theta_1 x^i)\right)^2 \quad -⑤$$

similary for θ_1:

$$\theta_1 = \theta_1 - \alpha \frac{\partial}{\partial \theta_1}\left[\frac{1}{16}\sum_{i=1}^{m}\left(y^i - (\theta_0 + \theta_1 x^i)\right)^2\right]$$

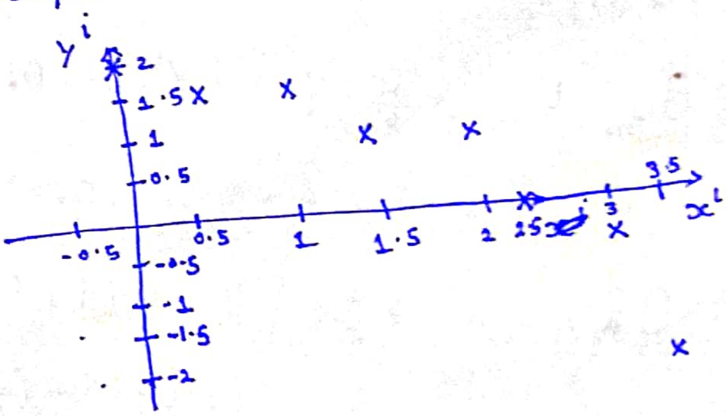
$$\theta_1 = \theta_1 + \frac{\alpha}{8}\sum_{i=1}^{m}\left(y^i - (\theta_0 + \theta_1 x^i)\right)\cdot x^i \quad -⑥$$

Now, to proceed further assumptions for α, θ_0 and θ_1, are need. To Do that let's observe data:

Tabular:

i	x^i	y^i
1	0	2
2	0.5	1.5
3	1	1.5
4	1.5	1
5	2	1
6	2.5	0
7	3	-0.5
8	3.5	-2

Graphical



As, it can be seen the regression (linear) line will have a negive slope and some positive slope, so assuming $\theta_0 = 1$ and $\theta_1 = -1$.

Let α be 0.02.

So, in the 1st iteration, we'll get:

For θ_0:

$$\theta_0 = \theta_0 + \frac{\alpha}{8} \sum_{i=1}^{m} \left[y^i - (\theta_0 + \theta_1 x^i) \right]$$

$m = 8$, $\alpha = 0.02$, $\theta_0 = 1$ and $\theta_1 = -1$

$$\theta_0 = 1 + \frac{0.02}{8} \Big[(2 - (1 + (-1)(0))) + (1.5 - (1 + (-1)(0.5))) + (1.5 - (1 + (-1)(1))) + (1 - (1 - (1.5)(1))) + (1 - (1 - 2(1))) + (0 - (1 + (-1)(2.5))) + (-0.5 - (1 + (-1)(3))) + (-2 - (1 + (-1 \times 3.5))) \Big]$$

$$\boxed{\theta_0 = 1.02625}$$

For θ_1:

$$\theta_1 = \theta_1 + \frac{\alpha}{8} \sum_{i=1}^{m} (y^i - (\theta_0 + \theta_1 x^i)) \cdot x^i$$

$m = 8$, $\alpha = 0.02$, $\theta_0 = 1$ and $\theta_1 = -1$

$$\theta_1 = -1 + \frac{0.02}{8} \Big[((2 - (1 + (-1(0))))0) + ((1.5 - (1 + (-1(0.5))))0.5) + ((1.5 - (1 + (-1(1))))1) + (1 - (1 + (-1(1.5))))(1.5) + (1 - (1 + (-2)))(2) + ((0 - (1 + (-2.5))(2.5)) + (-0.5 - (1 + (-1)(3)))(3) + (-2 - (1 + (-1 \times 3.5)))/3.5$$

$$\boxed{\theta_1 = -0.95}$$

Using the same method as above, in 2nd iteration:

For θ_0: $\theta_0 = \theta_0 + \frac{\alpha}{8} \sum_i^m \left[y^i - (\theta_0 + \theta_1 x^i) \right]$

where $\theta_0 = 1.02625$ and $\theta_1 = -0.95$, $\alpha = 0.02$, $m = 8$

$$\theta_0 = 1.02625 + \frac{0.02}{8}\Big[(2-(1.02625+(-0.95(0)))) + (1.5-(1.02625+(-0.95(0.5)))) + (1.5-(1.02625+(-0.95(1)))) + (1-(1.02625+(-0.95(1.5)))) + (1-(1.02625+(-0.95(2)))) + (0-(1.02625+(-0.95(2.5)))) + (-0.5-(1.02625+(-0.95(3)))) + (-2-(1.02625+(-0.95(3.5))))\Big]$$

$$\boxed{\theta_0 = 1.050219}$$

For θ_1:

$$\theta_1 = \theta_1 + \frac{\alpha}{8}\sum_{i=1}^{m}\left(y^i - (\theta_0 + \theta_1 x^i)\right] x^i$$

$\theta_0 = 1.02625$, $\theta_1 = -0.95$, $m = 8$, $\alpha = 0.02$

$$\theta_1 = -0.95 + \frac{0.02}{8}\Big[(2-(1.02625+(-0.95\times 0)))\times 0 + (1.5-(1.02625+(-0.95\times 0.5)))\times 0.5 + (1.5-(1.02625+(-0.95\times 1)))\times 1 + (1-(1.02625+(-0.95\times 1.5)))\times 1.5 + (1-(1.02625+(-0.95\times 2)))\times 2 + (0-(1.02625+(-0.95\times 2.5)))\times 2.5 + (-0.5-(1.02625+(-0.95\times 3)))\times 3 + (-2-(1.02625+(-0.95\times 3.5)))\times 3.5\Big]$$

$$\boxed{\theta_1 = -0.909688}$$

Repeating same process for iteration 3, we get

$\boxed{\theta_0 = 1.0723}$ $\boxed{\theta_1 = -0.874}$

Since it's same repetitive process, I have manually solved it for first 3 iterations only. However, for verification through python code (attached at end), I have ran it for 10000 iterations.

b) Now assume that the - - - last data point.

Ans: From part a, we know

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

$\theta_0 = 1$, $\theta_1 = -1$ and $\alpha = 0.02$

For vector update equations:

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

and

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

which gives, m=8

$$\theta_0 = \theta_0 + 0.02 \sum_{i=1}^{8} w^i [y^i - (\theta_0 + \theta_1 x^i)]$$

and

$$\theta_1 = \theta_1 + 0.02 \sum_i^8 w^i [y^i - (\theta_0 + \theta_1 x^i)] x^i$$

In first iteration, we'll get

$$\theta_0 = 1 + 0.02 [0.1 \times (2 - (1 + (-1 \times 0))) + (0.1 \times (1.5 - (1 + (-1 \times 0.5)))) + (0.1 \times (1.5 - (1 + (1 \times -1)))) + (0.1 \times (1 + (-1 \times 1.5)))) + (0.1 \times (1 - (1 + (-1 \times 2)))) + (0.1 \times (0 - (1 + (-1 \times 2.5)))) + (0.1 \times (-0.5 - (1 + (-1 \times 3)))) + (0.3 \times (-2 - (1 + (-1 \times 3.5))))]$$

$$\boxed{\theta_0 = 1.023}$$

For θ_1:

$\theta_0 = 1$, $\theta_1 = -1$, $\alpha = 0.02$ and $m = 8$

$$\theta_1 = -1 + 0.02[0.1[2 - (1 + (-1 \times 0))] \times 0 + 0.1[1.5 - (1 + (-1 \times 0.5))] \times 0.5 +$$
$$0.1 \times (1.5 - (1 + (-1 \times 1)) \times 1 + 0.1 \times (1 - (1 + (-1 \times 1.5)) \times 1.5 +$$
$$0.1 \times (1 - (1 + (-1 \times 2)) \times 2 + 0.1 \times (0 - (1 + (-1 \times 2.5)) \times 2.5 +$$
$$0.1 \times (-0.5 - (1 + (-1 \times 3)) \times 3 + 0.3 \times (-2 - (1 + (-1 \times 3.5)) \times 3.5)$$

$$\boxed{\theta_1 = -0.9565}$$

In iteration 2, repeating the same above steps with $\theta_0 = 1.023$ and $\theta_1 = -0.956$, we get

$$\boxed{\theta_0 = 1.04371}$$

$$\boxed{\theta_1 = -0.9191462}$$

Similary for iteration 3, with update θ_0 and θ_1, we get

$$\boxed{\theta_0 = 1.062422}$$

$$\boxed{\theta_1 = -0.88699}$$