Gradient DESCENT ALGORITHM

x_1	x_2	y
1	2	4
2	3	5
3	4	6

$\theta_0 = 0$
$\theta_1 = 0$
$\theta_2 = 0$

$\alpha = 0.01$

$$y = \theta_0 + \theta_1 x_1 + \theta_2 x_2$$

Hypothesis Function :- $h(x) = \theta_0 x_0 + \theta_1 x_1 + \theta_2 x_2 =$

$h(1,2) = \theta_0 x_0 + (0)(1) + (0)(2) = 0$

$h(2,3) = 0 + 2 \times 1 + 3 \times 0 = 0$

$h(3,4) = 0 + 3 \times 1 + 4 \times 0 = 0.$

Cost function. $\Rightarrow J(\theta) = \frac{1}{2} \sum (h(x) - y)^2$

$= \frac{1}{2}\left[(0-4)^2 + (0-5)^2 + (0-6)^2\right]$

$= \frac{1}{2}[16 + 25 + 36]$

$= \frac{1}{2}[77] = 77/2 = \underline{38.5}$

Gradient Descent $\theta_0 : \theta_0 - \alpha \cdot \frac{\partial J}{\partial \theta_0} \Rightarrow \frac{\partial J}{\partial \theta}[h(x) - y]$

$\theta_1 = 0 - 0.01[0-4] \times x_0 + (0-5) \times (x_1) + (0-6) x_2$

$\theta_1 = 0 - 0.01[-4] + [5] + [-6]$

$\theta_1 = 0 - 0.01$

$\boxed{\theta_0 = 0.015}$

interation 1

$\theta_0 = 0.015$

$\theta_1 = 0.32$

$\theta_2 = 0.47.$

$\theta_1 = 0 - 0.01[0-4] 1 + [0.5] 2 + [0-6] 3$

$= 0 - 0.01[4] 1 + [-5] 2 + [-6] 3$

$= 0 - 0.01[4] - 10 - 18$

$= 0 - 0.01[4] - 10 - 18.$

$\boxed{\theta_1 = 0.32}$

$\theta_2 = 0 - 0.01[-4 - 10 - 18]$

$= 0 - 0.01[4] 2 - [5] 3 - 4[6]$

$= 0 - 0.01[-8] - 15 - 24$ $= \cancel{0 \cdot 0.01[-24]}$

$= 0 - 0.01[-8 - 15 - 24] = 0 \cdot 0.01[-47] = \underline{0.47}$

$\theta_2 =$

$h_\theta(x^{(1)}) = 0.015 + 0.32(2) + 0.47(3) = 2.065$ ⑤

$h(2,3) = 0.015 + 0.32(1) + 0.\cancel{0933}^{47}(2) = 1.275$ ① Second iteration

$h(3,4) = 0.015 + 0.32(3) + 0.\cancel{0933}_{47}(4) = 2.855$ ③

Cost function: $= \frac{1}{2}\Sigma[h(x) - y]^2$

$= \frac{1}{2}\left[(2.065-4)^2 + (1.275-5)^2 + (2.85-6)^2\right]$

$= \frac{1}{2}[3.74 + 13.87 + 9.9225]$

$= \frac{1}{2}[27.532]$

$J[\theta] = 13.766$

GA $\cdot$ $\theta - \alpha\left[\frac{\partial J}{\partial \theta}\right]$ $\quad \left[\frac{\partial J}{\partial \theta}\right] = \left[h_0(x) - y\right]$

$\cdot$ ~~0-0.01~~ [

$= \theta_0 - \alpha\left[\frac{\partial J}{\partial \theta}\right] \quad \left[\frac{\partial J}{\partial \theta}\right] = \left[h_0(x) - y\right] x_i$

$= \cancel{\theta_0} -$

~~$= 0.015 - 0.01\left[(2.065 - 4)^2\right.$~~

$\theta_1 = 0.015 - 0.01\left[(1.275-4)^2 + (2.065-5)^2 + (2.855-3)^2\right]$

$= 0.005\left[(7.45) + 8.614 + 0.021\right]$

$= 0.005[16.085] = 0.08425.$

$\theta_2 : 0.32 - 0.01\left[(1.275-4)^2(1) + (2.065-5)^2(2) + (2.855-6)^2(3)\right]$

$= 0.31 \times [7.425 + 17.22 + 29.67]$

$\theta_3 = 0.\cancel{47} - 0.01\left[(1.275-4)^2(2) + (2.065-5)^2(3) + (2.855-6)^2(4)\right]$

$= 0.46\,[14.85 + 8.528 + 39.5641]$

$= 2.$

$\theta_1 = 0.08425$

$\theta_2 = 16.83$

$\theta_3 = 28.95$

Problem

x_1	x_2	y
2	3	10
5	7	21
8	9	30

$$X = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 5 & 7 \\ 1 & 8 & 9 \end{bmatrix}$$

$$Y = \begin{bmatrix} 10 \\ 21 \\ 30 \end{bmatrix}$$

$$\theta = (X^T X)^{-1} X^T y.$$

X tran $\quad X^T = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 5 & 8 \\ 3 & 7 & 9 \end{bmatrix}$

$$X^T \cdot X = \underset{3\times 3}{\begin{bmatrix} 1 & 1 & 1 \\ 2 & 5 & 8 \\ 3 & 7 & 9 \end{bmatrix}} \underset{3\times 3.}{\begin{bmatrix} 1 & 2 & 3 \\ 1 & 5 & 7 \\ 1 & 8 & 9 \end{bmatrix}}$$

$$= \begin{bmatrix} 1(1)+1(1)+1(1) & 1(2)+1(5)+1(8) & 1(3)1(7)1(9) \\ 2(1)+5(1)+8(1) & 2(2)+5(5)+8(8) & 2(3)+5(7)+8(9) \\ 3(1)+7(1)+9(1) & 3(3)+7(7)+9(9) & 3(2)+7(5)+9(8) \end{bmatrix}$$

$$X^T \cdot X^{-1} = \begin{bmatrix} 3 & 15 & 19 \\ 15 & 93 & 121 \\ 19 & 121 & 158 \end{bmatrix}$$

Compute $X^T y$ $\quad \underset{3\times 3}{\begin{bmatrix} 1 & 1 & 1 \\ 2 & 5 & 8 \\ 3 & 7 & 9 \end{bmatrix}} \underset{3\times 1}{\begin{bmatrix} 10 \\ 21 \\ 30 \end{bmatrix}} = \begin{bmatrix} 10 + 21 + 30 \\ 20 + 105 + 240 \\ 30 + 147 + 270 \end{bmatrix} = \begin{bmatrix} 61 \\ 365 \\ 447 \end{bmatrix}$

$$\theta = \begin{bmatrix} 3 & 15 & 19 \\ 15 & 93 & 121 \\ 19 & 121 & 158 \end{bmatrix}^{-1} \begin{bmatrix} 61 \\ 365 \\ 447 \end{bmatrix}$$

Step 4 $\quad (X^T \cdot X)^{-1}$ find inverse.

$$(X^T \cdot X)^{-1} = \begin{bmatrix} 3 & 15 & 19 \\ 15 & 93 & 121 \\ 19 & 121 & 158 \end{bmatrix}^{-1}$$

$$A^{-1} = \frac{1}{\det |A|} \times \text{Ad of } A.$$

$\boxed{\text{we can compute } A^{-1}}$

$$A^{-1} = \frac{1}{\det |A)} \text{adj}(A). \qquad = 159 - 1065 +$$

$$A^{-1} = 0$$

$$\det \begin{bmatrix} 3 & 15 & 19 \\ 15 & 93 & 121 \\ 19 & 121 & 158 \end{bmatrix} = \underline{6} \qquad \text{As } (A) \neq 0 \text{ [invertible}$$

$$\det = \begin{bmatrix} 3 & 15 & 19 \\ 15 & 93 & 121 \\ 19 & 121 & 158 \end{bmatrix}^{-1}$$

$$= 8\begin{bmatrix} 93 & 121 \\ 121 & 158 \end{bmatrix} + \cancel{19}\begin{bmatrix} 19 & 121 \\ 19 & 158 \end{bmatrix} + \cancel{19}\begin{bmatrix} 15 & 93 \\ 19 & 121 \end{bmatrix}$$

$$= \cancel{15}\begin{bmatrix} 15 & 19 \\ 121 & 158 \end{bmatrix} + \cancel{93}\begin{bmatrix} 3 & 19 \\ 19 & 158 \end{bmatrix} + \cancel{121}\begin{bmatrix} 3 & 15 \\ 19 & 121 \end{bmatrix}$$

$$= 19\begin{bmatrix} 15 & 19 \\ 93 & 121 \end{bmatrix} + \cancel{121}\begin{bmatrix} 3 & 19 \\ 15 & 121 \end{bmatrix} + \cancel{158}\begin{bmatrix} 3 & 15 \\ 15 & 93 \end{bmatrix}$$

$$19\,[1815 - 1787] +$$

$$= 912 / + /9438 + 8532 = \begin{bmatrix} 53 & -71 & 48 \\ -71 & 113 & -78 \\ 48 & -78 & 54 \end{bmatrix}$$

$$\operatorname{adj}(A) = \begin{bmatrix} 53 & -71 & 48 \\ -71 & 113 & -78 \\ 48 & -78 & 54 \end{bmatrix}$$

$$= \frac{1}{\det|A|} \operatorname{adj}[A]$$

$$= \frac{1}{6}\begin{bmatrix} 53 & -71 & 48 \\ -71 & 113 & -78 \\ 48 & -78 & 54 \end{bmatrix}$$

$$= \begin{bmatrix} 53/6 & -71/6 & 48/6 \\ -71/6 & 113/6 & -78/6 \\ 48/6 & -78/6 & 54/6 \end{bmatrix}$$

$$A^{-1} = (X^T X)^{-1} = \begin{bmatrix} 8.83 & -11.83 & 8.00 \\ -11.83 & 11.83 & -13 \\ 8.00 & -13.00 & 9 \end{bmatrix}$$

$$= 538.63 + 4317.95 + 357.6$$

$$= -721.63 + 4317.95 + (-5841)$$

$$=$$

$$\theta = \begin{bmatrix} 158.0 \\ -252.0 \\ 175.00 \end{bmatrix}$$

$$\Uparrow$$

$$Y = \cancel{X^{1}} \quad \theta = (X^T X)^{-1} X^T y = \begin{bmatrix} 8.83 & -11.83 & 8.00 \\ -11.83 & 11.83 & -13 \\ 8.00 & -13.00 & 7 \end{bmatrix} \begin{bmatrix} 61 \\ 365 \\ 447 \end{bmatrix}$$

$$\boxed{\theta_1 = 158.0,\ \theta_2 = -252,\ \theta_3 = 175}$$

To check the predicted value (θ) is close to actual value.

Compute $h(x) = X\theta$.

$$X \begin{bmatrix} 1 & 2 & 3 \\ 1 & 5 & 7 \\ 1 & 8 & 9 \end{bmatrix} \begin{bmatrix} 158.04 \\ -252.04 \\ 175.00 \end{bmatrix}$$

Now $h(x) = (1 \times 158.04) + (2 \times (-252.04)) + 3(175.00) = 10.96.$

$$= 1 \times \begin{bmatrix} 10.96 \\ 21.16 \\ 30.72 \end{bmatrix}$$

$h(x) = y$. predicted θ values are close

11