1. Let the three (x_i, y_i) points be $\{(1,4), (2,4), (3,8)\}$. We want to fit the best possible linear map between x and y. Let it be $y = mx + b$. Let the error be $\sum_i (y_i - (mx_i + b))^2$. (i) Starting from the initial solution $(m, b) = (1, 0)$, find the next solution as per the gradient descent method with $\eta = 0.01$, (ii) Starting from the initial solution $(m, b) = (0, 2)$ find the next solution using the Newton's descent method. {Note, you can write your answer on both sides of this sheet. Write appropriate intermediate steps also. Carefully read the question and consider the numbers given. Considering wrong numbers and solving the problem will not give any marks.} (5+15 = 20 Marks).

Given $(x_i, y_i) = \{(1, 4), (2, 4), (3, 8)\}$.

$$y = mx + b$$

error $\sum_i (y_i - (mx_i + b))^2$

$$f = \sum_i (y_i - (mx_i + b))^2$$

$$\frac{\partial f}{\partial m} = \frac{\partial}{\partial x}\left(\sum_i (y_i - (mx_i + b))^2 \right)$$

$$= 2 \sum_i (y_i - (mx_i + b))\ (-x_i)$$

$$= -2 \sum_i (x_i)(y_i - (mx_i + b))$$

$$\frac{\partial f}{\partial b} = \frac{\partial}{\partial b}\left(\sum_i (y_i - (mx_i + b))^2 \right)$$

$$= -2 \sum_i (y_i - (mx_i + b))$$

$$\frac{\partial^2 f}{\partial m^2} = \frac{\partial}{\partial m}\left(-2 \sum_i (x_i)(y_i - (mx_i + b))\right)$$

$$= 2 \sum_i x_i^2$$

$$\frac{\partial^2 f}{\partial b^2} = \frac{\partial}{\partial b}\left(-2 \sum_i (y_i - (mx_i + b))\right)$$

$$= 2 \sum_i 1$$

$$= 6$$

$$\frac{\partial^2 f}{\partial m\, \partial b} = \frac{\partial}{\partial m}\left(-2 \sum_i (y_i - (mx_i + b))\right)$$

$$= 2 \sum_i x_i$$

1. **GRADIENT DECENT METHOD:**

Given (m,b)=(1,0)

$$\eta = 0.01$$

$$a_{k+1} = a_k - \eta\, f'(a_k)$$

$$a_1 = a_0 - \eta\, f'(a_0)$$

$$f'(a_0) = \begin{pmatrix} \frac{\partial f}{\partial m} \\ \frac{\partial f}{\partial b} \end{pmatrix}_{\binom{1}{0}} = \begin{pmatrix} -44 \\ -20 \end{pmatrix}$$

$$a_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} - 0.01\begin{pmatrix} -44 \\ -20 \end{pmatrix}$$

$$a_1 = \begin{pmatrix} 1.44 \\ 0.2 \end{pmatrix}$$

2. **NEWTON'S DESCENT METHOD:**

$$(m, b) = (0,2)$$

$$a_{k+1} = a_k - \frac{1}{f''(a_k)} f'(a_k)$$

Here $\frac{1}{f''(a_k)} = H^{-1}$

$$a_1 = a_0 - H^{-1} f'(a_0)$$

Calculating H^{-1}

$$H = \begin{pmatrix} \frac{\partial^2 f}{\partial m^2} & \frac{\partial^2 f}{\partial m\, \partial b} \\ \frac{\partial^2 f}{\partial b\, \partial m} & \frac{\partial^2 f}{\partial b^2} \end{pmatrix}$$

$$= \begin{pmatrix} 2\ \sum_i x_i^2 & 2\ \sum_i x_i \\ 2\ \sum_i x_i & 6 \end{pmatrix} = \begin{pmatrix} 28 & 12 \\ 12 & 6 \end{pmatrix}$$

$$H^{-1} = \frac{1}{(28)(6)-(12)(12)} \begin{pmatrix} 6 & -12 \\ -12 & 28 \end{pmatrix}$$

$$= \frac{1}{24} \begin{pmatrix} 6 & -12 \\ -12 & 28 \end{pmatrix}$$

$$H^{-1} = \begin{pmatrix} \frac{1}{4} & \frac{-1}{2} \\ \frac{-1}{2} & \frac{7}{6} \end{pmatrix}$$

$$f'\begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} -48 \\ -20 \end{pmatrix}$$

$$a_1 = a_0 - H^{-1} f'(a_0)$$

$$a_1 = \binom{1}{1} - \begin{pmatrix} \frac{1}{4} & \frac{-1}{2} \\ \frac{-1}{2} & \frac{7}{6} \end{pmatrix} \binom{-48}{-20}$$

$$= \binom{2}{1.33}$$

$$a_1 = \binom{2}{1.33}$$