①

$$X = \begin{bmatrix} 0,1 & 1 \\ 0,4 & 1 \\ 0,8 & 1 \\ 1,2 & 1 \\ 2,0 & 1 \\ 2,5 & 1 \\ 3,0 & 1 \end{bmatrix} \qquad y = \begin{pmatrix} 0,5 \\ 0,2 \\ 0,7 \\ 2,5 \\ 1,5 \\ 1,0 \\ 4,0 \end{pmatrix} \qquad \theta = \begin{pmatrix} 3,0 \\ 1,5 \end{pmatrix}$$

a) $\nabla f(\theta) = \frac{1}{2}(X^TX)\theta - X^Ty = \frac{1}{2}X^T(X\theta - y) = \frac{1}{2}(X^TX\theta - X^Ty)$

$$X^TX = \begin{bmatrix} 0,1 & 0,4 & 0,8 & 1,2 & 2,0 & 2,5 & 3,0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0,1 & 1 \\ 0,4 & 1 \\ 0,8 & 1 \\ 1,2 & 1 \\ 2,0 & 1 \\ 2,5 & 1 \\ 3,0 & 1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 21,5 & 10 \\ 10 & 7 \end{bmatrix}$$

$$X^Ty = \begin{bmatrix} 21,19 \\ 10,4 \end{bmatrix}$$

$$d_0 = -\frac{1}{2} \cdot \left(\begin{bmatrix} 21,5 & 10 \\ 10 & 7 \end{bmatrix} \cdot \begin{bmatrix} 3,0 \\ 1,5 \end{bmatrix} - \begin{bmatrix} 21,19 \\ 10,4 \end{bmatrix} \right)$$

$$= -\frac{1}{2} \left(\begin{bmatrix} 79,5 \\ 40,5 \end{bmatrix} - \begin{bmatrix} 21,19 \\ 10,4 \end{bmatrix} \right) = -\frac{1}{2} \begin{pmatrix} 58,31 \\ 30,1 \end{pmatrix}$$

$$= \begin{pmatrix} -29,155 \\ -15,05 \end{pmatrix}$$

$$\theta_1 = \theta_0 + \alpha \cdot d_0 = \begin{pmatrix} 3,0 \\ 1,5 \end{pmatrix} + 0,3 \cdot \begin{pmatrix} -29,155 \\ -15,05 \end{pmatrix}$$

$$= \begin{pmatrix} 3,0 \\ 1,5 \end{pmatrix} + \begin{pmatrix} -8,75 \\ -4,52 \end{pmatrix}$$

$$= \begin{pmatrix} -5,75 \\ -3,02 \end{pmatrix}$$

$$d_1 = -\frac{1}{2} \left(\begin{bmatrix} 21,5 & 10 \\ 10 & 7 \end{bmatrix} \cdot \begin{pmatrix} -5,75 \\ -3,02 \end{pmatrix} - \begin{pmatrix} 21,19 \\ 10,4 \end{pmatrix} \right)$$

$$= -\frac{1}{2} \begin{pmatrix} -175,015 \\ -89,04 \end{pmatrix} = \begin{pmatrix} 87,51 \\ 44,52 \end{pmatrix}$$

$\Theta_2 = \Theta_1 + \alpha \cdot d_1$

$= \begin{pmatrix} -5,75 \\ -3,02 \end{pmatrix} + 0,3 \cdot \begin{pmatrix} 87,51 \\ 44,52 \end{pmatrix}$

$\approx \begin{pmatrix} 20,50 \\ 10,34 \end{pmatrix}$

b) Batch $1 = \begin{pmatrix} 0,1,2 \end{pmatrix}$
    indices of Datapoints

Batch $2 = (3,4)$
Batch $3 = (5,6)$

$\hookrightarrow$ Justification:
All Batches should be of equals size, since 7 can't be divided into 3 equaly sized Batches one must have a Datapoint more than the others.

c) Batch 1:

$X^T X = \begin{bmatrix} 0,1 & 0,4 & 0,8 \\ 1 & 1 & 1 \end{bmatrix}^T \cdot \begin{bmatrix} 0,1 & 1 \\ 0,4 & 1 \\ 0,8 & 1 \end{bmatrix}$

$\Rightarrow \begin{bmatrix} 0,81 & 1,3 \\ 1,3 & 3 \end{bmatrix}$

$X^T y = \begin{bmatrix} 0,1 & 0,4 & 0,8 \\ 1 & 1 & 1 \end{bmatrix}^T \cdot \begin{pmatrix} 0,5 \\ 0,2 \\ 0,7 \end{pmatrix}$

$= \begin{pmatrix} 0,69 \\ 1,4 \end{pmatrix}$

$d_0 = -\frac{1}{2} \left(\begin{bmatrix} 0,81 & 1,3 \\ 1,3 & 3 \end{bmatrix} \cdot \begin{pmatrix} 3,0 \\ 1,5 \end{pmatrix} - \begin{pmatrix} 0,69 \\ 1,4 \end{pmatrix} \right)$

$= -\frac{1}{2} \begin{pmatrix} 3,69 \\ 7 \end{pmatrix} = \begin{pmatrix} -1,845 \\ -3,5 \end{pmatrix}$

$\Theta_1 = \begin{pmatrix} 3 \\ 1,5 \end{pmatrix} + 0,3 \cdot \begin{pmatrix} -1,845 \\ -3,5 \end{pmatrix}$

$= \begin{pmatrix} 2,45 \\ 0,45 \end{pmatrix}$

$d_1 = -\frac{1}{2} \left(\begin{bmatrix} 0,81 & 1,3 \\ 1,3 & 3 \end{bmatrix} \cdot \begin{pmatrix} 2,45 \\ 0,45 \end{pmatrix} - \begin{pmatrix} 0,69 \\ 1,4 \end{pmatrix} \right)$

$= -\frac{1}{2} \begin{pmatrix} 1,88 \\ 3,14 \end{pmatrix} = \begin{pmatrix} -0,94 \\ -1,57 \end{pmatrix}$

$\Theta_2 = \begin{pmatrix} 2,45 \\ 0,45 \end{pmatrix} + 0,3 \cdot \begin{pmatrix} -0,94 \\ -1,57 \end{pmatrix}$

$= \begin{pmatrix} 2,17 \\ -0,02 \end{pmatrix}$