$$\frac{\partial MSE}{\partial W_{i,j}} = \frac{\partial MSE}{\partial Z_j} \frac{\partial Z_j}{\partial H_j} \frac{\partial H_j}{\partial W_{i,j}}.$$

$$MSE_i = \frac{1}{2}(\hat{R}_i - R_i)^2.$$

$$\frac{\partial MSE}{\partial Z_j} = \frac{\partial MSE_0}{\partial Z_j} + \frac{\partial MSE_1}{\partial Z_j} + \cdots + \frac{\partial MSE_q}{\partial Z_j}$$

$$= \frac{\partial MSE_0}{\partial O_0} \frac{\partial O_0}{\partial Z_j} + \cdots + \frac{\partial MSE_q}{\partial O_q} \frac{\partial O_q}{\partial Z_j}$$

$$= \frac{\partial MSE_0}{\partial R_0} \frac{\partial R_0}{\partial O_0} \frac{\partial O_0}{\partial Z_j} + \cdots + \frac{\partial MSE_q}{\partial R_q} \frac{\partial R_q}{\partial O_q} \frac{\partial O_q}{\partial Z_j}$$

$$= (\hat{R}_0 - R_0) \cdot R_0(1 - R_0) \cdot w_{j,0} + \cdots$$

$$+ (\hat{R}_q - R_q) \cdot R_q(1 - R_q) \cdot w_{j,q}.$$

$$\frac{\partial Z_j}{\partial H_j} = Z_j(1 - Z_j).$$

$$\frac{\partial H_j}{\partial W_{i,j}} = I_i.$$

Assn 3 Written assignment.

4-1. 1)

$$H_i = W_{0,i} I_0 + W_{1,i} I_1 + \cdots + W_{782,i} I_{782} + W_{783,i} I_{783} + b_i \quad \cdots ①$$

$i : 0 \sim 127.$

$$Z_i = \phi(H_i) \quad \cdots ②$$

$$O_j = W'_{0,j} Z_0 + W'_{1,j} Z_1 + \cdots + W'_{127,j} Z_{127} + b'_j \quad \cdots ③$$

$$R_j = \phi(O_j) \quad \cdots ④$$

$j : 0 \sim 9.$

$$MSE = \frac{1}{2}\sum_{k=0}^{9} (\hat{R}_k - R_k)^2 \quad \cdots ⑤$$

$\hat{R}_k = 0$ if label $== k$ else 0.

a) $$\frac{\partial MSE}{\partial W_{i,j}} = \frac{\partial MSE}{\partial Z_j} \cdot \frac{\partial Z_j}{\partial H_j} \cdot \frac{\partial H_j}{\partial W_{i,j}} \quad ⑥$$

i) Let's denote $MSE_k = \frac{1}{2}(\hat{R}_k - R_k)^2$ and $\therefore MSE = \sum_{k=0}^{9} MSE_i$.

Then $$\frac{\partial MSE}{\partial Z_j} = \frac{\partial MSE_0}{\partial Z_j} + \frac{\partial MSE_1}{\partial Z_j} + \cdots + \frac{\partial MSE_9}{\partial Z_j}$$

$$= \frac{\partial MSE_0}{\partial R_0}\frac{\partial R_0}{\partial O_0}\frac{\partial O_0}{\partial Z_j} + \cdots + \frac{\partial MSE_9}{\partial R_9}\frac{\partial R_9}{\partial O_9}\frac{\partial O_9}{\partial Z_j}$$

$$= (\hat{R}_0 - R_0) \cdot R_0(1-R_0) \cdot W_{j,0} + \cdots + (\hat{R}_9 - R_9) \cdot R_9(1-R_9) \cdot W_{j,9}.$$

(by ⑥, ④, ③.)

$$= \sum_{k=0}^{9} (\hat{R}_k - R_k) \cdot R_k(1-R_k) W_{j,k}.$$

ii) $$\frac{\partial Z_j}{\partial H_j} = Z_j(1-Z_j) \quad (\text{by } ②)$$

iii) $$\frac{\partial H_j}{\partial W_{i,j}} = I_i. \quad (\text{by } ①).$$

$$\therefore \frac{\partial MSE}{\partial W_{i,j}} = \sum_{k=0}^{q} (\hat{R}_k - R_k) \cdot R_k(1-R_k) W_{j,k} \cdot Z_j(1-Z_j) \cdot I_i.$$

b) $$\frac{\partial MSE}{\partial b_j} = \frac{\partial MSE}{\partial Z_j} \cdot \frac{\partial Z_j}{\partial H_j} \cdot \frac{\partial H_j}{\partial b_j}$$

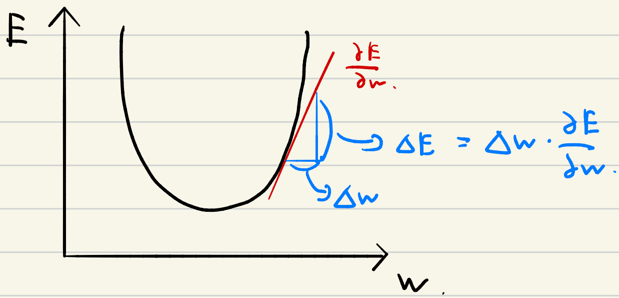
As $\frac{\partial H_j}{\partial b_j} = 1 \quad (\text{by } ①)$,

$$\frac{\partial MSE}{\partial b_j} = \sum_{k=0}^{q} (\hat{R}_k - R_k) \cdot R_k(1-R_k) W_{j,k} \cdot Z_j(1-Z_j)$$

(by a-i), a-ii)).

2) $w = w - \eta \frac{\partial MSE}{\partial w} = w + \Delta w$.

notice that $\Delta MSE = \frac{\partial E}{\partial w} \Delta w$.



as $\Delta w = -\eta \frac{\partial E}{\partial w}$,

$$\Delta E = \Delta w \cdot \frac{\partial E}{\partial w} = -\eta \left(\frac{\partial E}{\partial w} \right)^2 < 0.$$