Solutions for exam in Artificial Neural Networks

October 28, 2019

1) $w_{12} = 2$

$w_{21} = -1$

a) $S_2' = \mathrm{sgn}(w_{21} S_1) = -S_1 \quad (*)$

$$H = -\frac{w_{12}+w_{21}}{2} S_1 S_2$$

$$H' = -\frac{w_{12}+w_{21}}{2} S_1 S_2'$$

$$\Delta H = H' - H = -\frac{w_{12}+w_{21}}{2}(S_1 S_2' - S_1 S_2)$$

$$\Delta H = -\frac{w_{12}+w_{21}}{2} S_1 (S_2' - S_2) = -\frac{2-1}{2} S_1 (S_2' - S_2)$$

$$\Delta H = -\frac{1}{2} S_1 (S_2' - S_2)$$

If $S_2' = S_2 \Rightarrow \Delta H = 0$

Otherwise: $S_2' = -S_2 = -S_1$ (from $*$)

$$\Rightarrow \Delta H = -\frac{1}{2} S_1 (-S_1 - S_1) = \frac{1}{2} S_1 \cdot 2 S_1 = S_1^2 = 1$$

$\Rightarrow \boxed{\Delta H > 0 \text{, or } \Delta H = 0}$ Thus, the energy can increase upon updating the second neuron.

For updating the first neuron, find instead

$S_1' = \mathrm{sgn}(w_{12} S_2) = S_2 \quad (**)$

$$\Delta H = -\frac{w_{12}+w_{21}}{2}(S_1' S_2 - S_1 S_2) = -\frac{1}{2} S_2 (S_1' - S_1)$$

If $S_1' = S_1 \Rightarrow \Delta H = 0$

Otherwise $S_1' = -S_1 = S_2$ (from $(**)$)

It follows:

$\Delta H = -\frac{1}{2} S_2 \; (+S_2 + S_2) = -\frac{1}{2} S_2 \cdot 2 S_2 = -S_2^{\,2} < 0$

$\Rightarrow \boxed{\Delta H < 0 \text{ or } \Delta H = 0}$ Thus, the energy either stays constant or decreases upon updating the ~~second~~ first neuron.

b) $S_i' = \operatorname{sgn}(w_{ij} S_j)$

$S_1' = \operatorname{sgn}(w_{12} S_2) = S_2$

$S_2' = \operatorname{sgn}(w_{21} S_1) = -S_1$

$$\Delta H = -\frac{w_{12}+w_{21}}{2} \; S_1' S_2' + \frac{w_{12}+w_{21}}{2} S_1 S_2 =$$

$$= -\frac{w_{12}+w_{21}}{2} \big(\underbrace{S_1'}_{=S_2} \underbrace{S_2'}_{=-S_1} - S_1 S_2 \big) =$$

$$= -\frac{w_{12}+w_{21}}{2} (-S_1 S_2 - S_1 S_2)$$

$$= -\frac{2-1}{2} \cdot (-2) S_1 S_2 =$$

$\Delta H = S_1 S_2$

If $S_1 = S_2 \Rightarrow \Delta H = 1$

Otherwise, $S_1 = -S_2 \Rightarrow \Delta H = -1$

Thus, under synchronous update, the energy either increases or decreases. In other words, it cannot stay constant.

c) $H = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij} n_i n_j + \sum_{i=1}^{N} \mu_i n_i$

$n_m' = \Theta_H(b_m) \quad , b_m = \sum_{j=1}^{N} w_{mj} n_j - \mu_m$

$\Theta_H(b_m) = \begin{cases} 1 & , b_m > 0 \\ 0 & , \text{otherwise} \end{cases}$

$H' = -\frac{1}{2} w_{mm} n_m' n_m' - \frac{1}{2}\sum_{i\neq m}\sum_{j\neq m} w_{ij} n_i n_j \overbrace{- \frac{1}{2}\sum_{\substack{i=1 \\ i\neq m}}^{N} w_{im} n_i n_m'}^{\text{use } w_{im}=w_{mi} \text{, change } i\to j}$

$- \frac{1}{2}\sum_{j\neq m} w_{mj} n_m' n_j + \mu_m n_m' + \underbrace{\sum_{i\neq m} \mu_i n_i}_{\text{change } i\to j}$

$\Delta H = -\frac{1}{2} w_{mm} n_m' n_m' + \frac{1}{2} w_{mm} n_m n_m$

$\cancel{- \frac{1}{2}\sum_{i\neq m}\sum_{j\neq m} w_{ij} n_i n_j} \cancel{+ \frac{1}{2}\sum_{i\neq m}\sum_{j\neq m} w_{ij} n_i n_j}$

$- \frac{1}{2}\cdot 2\sum_{j\neq m} w_{mj} n_m' n_j + \frac{1}{2}\cdot 2\sum_{j\neq m} w_{mj} n_m n_j$

$+ \mu_m n_m' - \mu_m n_m \cancel{+ \sum_{j\neq m} \mu_j n_j} \cancel{- \sum_{j\neq m} \mu_j n_j}$

$= -\frac{1}{2} w_{mm} \underbrace{(n_m'^2 - n_m^2)}_{= (n_m' - n_m)(n_m' + n_m)}$

$- \sum_{j\neq m} w_{mj} n_j (n_m' - n_m)$

$+ \mu_m (n_m' - n_m)$

$\Delta H = -(n_m' - n_m)\Big[\sum_{j\neq m} w_{mj} n_j + w_{mm} n_m - w_{mm} n_m + (-\mu_m)$

$+ \frac{1}{2} w_{mm} (n_m' + n_m)\Big]$

$$\Delta H = -(n_m' - n_m)\Big[\underbrace{\sum_{j=1}^{N} w_{mj} n_j - \mu_m - w_{mm} n_m}_{b_m} + \frac{1}{2} w_{mm}(n_m' + n_m)$$

$$\Delta H = -(n_m' - n_m)\left[b_m + \frac{w_{mm}}{2}(n_m' - n_m)\right]$$

1) For $n_m' = n_m \Rightarrow \Delta H = 0$

2) Otherwise $n_m' = 1 - n_m$

A) Consider first $n_m' = 1$, $n_m = 0$. (note: $b_m > 0$)

$$\Rightarrow \Delta H = -\Big(\underbrace{b_m}_{>0} + \underbrace{\frac{w_{mm}}{2}}_{>0}\Big)$$

$$\Rightarrow \boxed{\Delta H < 0}$$

B) Now consider $n_m' = 0$ $(b_m < 0)$, $n_m = 1$

$$\Rightarrow \Delta H = + \left[b_m - \frac{w_{mm}}{2}\right]$$

$$\Delta H = \underbrace{b_m}_{<0} + \Big(\underbrace{-\frac{w_{mm}}{2}}_{<0}\Big)$$

$$\Rightarrow \boxed{\Delta H < 0}$$

Thus, the energy either stays constant or decreases, which was to be shown.

2) $w_{ij} = \frac{1}{N} \sum_{\mu=1}^{2} x_i^{(\mu)} x_j^{(\mu)}$, $i,j = 1,2,\ldots,N$

$N = 35$

$S_i \leftarrow \mathrm{sgn}\left(\sum_{j=1}^{N} w_{ij} S_j\right)$

a) $Q^{(\mu,\nu)} = \sum_{j=1}^{N} x_j^{(\mu)} x_j^{(\nu)}$, $\nu = 1,2,3,4,5$
$\mu = 1,2$

$Q^{(1,1)} = N = 35$

$Q^{(1,2)} = 35 - 2 \cdot 10 = 15$

$Q^{(1,3)} = 35 - 2 \cdot 8 = 19$

$Q^{(1,4)} = 35 - 2 \cdot 11 = 13$

$Q^{(1,5)} = 35 - 2 \cdot 35 = -35$

inverse of $\underline{1}$

$Q^{(2,1)} = Q^{(1,2)} = 15$

$Q^{(2,2)} = N = 35$

$Q^{(2,3)} = N - 2 \cdot 14 = 35 - 28 = 7$

$Q^{(2,4)} = 35 - 2 \cdot 3 = 29$

$Q^{(2,5)} = 35 - 2 \cdot 25 = -15$

b) $b_i^{(\nu)} = \sum_{j=1}^{N} w_{ij} x_j^{(\nu)} = \sum_{j=1}^{N} \frac{1}{N} x_i^{(1)} x_j^{(1)} x_j^{(\nu)} +$
$+ \sum_{j=1}^{N} \frac{1}{N} x_i^{(2)} x_j^{(2)} x_j^{(\nu)} =$

$= \frac{1}{N} Q^{(1,\nu)} x_i^{(1)} + \frac{1}{N} Q^{(2,\nu)} x_i^{(2)}$

$b_i^{(1)} = x_i^{(1)} + \frac{15}{35} x_i^{(2)}$

$b_i^{(2)} = \frac{15}{35} x_i^{(1)} + x_i^{(2)}$

$b_i^{(3)} = \frac{19}{35} x_i^{(1)} + \frac{7}{35} x_i^{(2)}$

$b_i^{(4)} = \frac{13}{35} x_i^{(1)} + \frac{29}{35} x_i^{(2)}$

$b_i^{(5)} = -x_i^{(1)} - \frac{15}{35} x_i^{(2)}$

c) $\mathrm{sgn}(b_i^{(1)}) = \mathrm{sgn}(x_i^{(1)}) = x_i^{(1)}$ unchanged

$\mathrm{sgn}(b_i^{(2)}) = x_i^{(2)}$ → unchanged

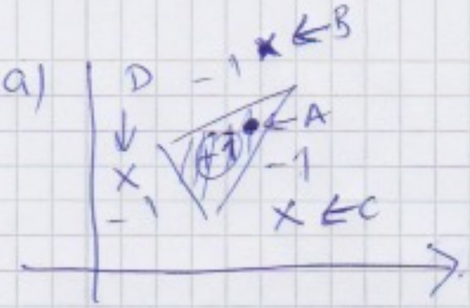
$\mathrm{sgn}(b_i^{(3)}) = \mathrm{sgn}(x_i^{(1)}) = x_i^{(1)}$ changes

$\mathrm{sgn}(b_i^{(4)}) = \mathrm{sgn}(x_i^{(2)}) = x_i^{(2)}$ changes

$\mathrm{sgn}(b_i^{(5)}) = -x_i^{(1)}$ → remains

Patterns $\underline{1},\underline{2}$ and $\underline{5}$ remain the same.

3) a)



Choose points A, B, C, D.
There is no line that can separate these points so that B, D and C lie on one side of the line and A is on the other: at least one point will be misclassified.
→ Problem not linearly separable.

b) Hidden node 1: use points $\begin{pmatrix}-1\\-1\end{pmatrix}$ and $\begin{pmatrix}-3\\2\end{pmatrix}$.

It follows:
$$\left.\begin{aligned} -\theta_1 + (w_{11}\ \ w_{12})\begin{pmatrix}-1\\-1\end{pmatrix} &= 0 \\ \text{and } -\theta_1 + (w_{11}\ \ w_{12})\begin{pmatrix}-3\\2\end{pmatrix} &= 0 \end{aligned}\right\}$$

$$\Rightarrow \left.\begin{aligned} -\theta_1 - w_{11} - w_{12} &= 0 \\ -\theta_1 - 3w_{11} + 2w_{12} &= 0 \end{aligned}\right\} \ominus \Rightarrow 2w_{11} - 3w_{12} = 0 \quad \boxed{w_{12} = \frac{2}{3}w_{11}}$$

substract second from the first

Given the orientation of $\underline{w}_1$, we can choose

$w_{11} = 1, \rightarrow w_{12} = \frac{2}{3}$

$\theta_1 = -w_{11} - w_{12} = -1 - \frac{2}{3} = -\frac{5}{3}$

$$\Rightarrow \boxed{\underline{w}_1 = \begin{pmatrix}1\\ \frac{2}{3}\end{pmatrix}, \ \theta_1 = -\frac{5}{3}}$$

Second hidden node. Choose points $\begin{pmatrix}-3\\2\end{pmatrix}$ and $\begin{pmatrix}2\\3\end{pmatrix}$.

$$\left.\begin{aligned} -\theta_2 + (w_{21}\ \ w_{22})\begin{pmatrix}-3\\2\end{pmatrix} &= 0 \\ -\theta_2 + (w_{21}\ \ w_{22})\begin{pmatrix}2\\3\end{pmatrix} &= 0 \end{aligned}\right\} \left.\begin{aligned} -\theta_2 - 3w_{21} + 2w_{22} &= 0 \\ -\theta_2 + 2w_{21} + 3w_{22} &= 0 \end{aligned}\right\}$$

$$\Rightarrow 5w_{21} + w_{22} = 0 \Rightarrow \boxed{w_{22} = -5w_{21}}$$

Given the orientation of $\underline{w}_2$, we can choose:

$w_{21} = -1 \Rightarrow w_{22} = 5$

$\theta_2 = -3w_{21} + 2w_{22} = 3 + 10 = 13$

$$\boxed{\underline{w}_2 = \begin{pmatrix} -1 \\ 5 \end{pmatrix}, \quad \theta_2 = 13}$$

Third hidden node. Choose points $\begin{pmatrix} -1 \\ -1 \end{pmatrix}$, and $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$

$$\left.\begin{aligned} -\theta_3 + (w_{31} \; w_{32})\begin{pmatrix} -1 \\ -1 \end{pmatrix} &= 0 \\ -\theta_3 + (w_{31} \; w_{32})\begin{pmatrix} 2 \\ 3 \end{pmatrix} &= 0 \end{aligned}\right\} \left.\begin{aligned} -\theta_3 - w_{31} - w_{32} &= 0 \\ -\theta_3 + 2w_{31} + 3w_{32} &= 0 \end{aligned}\right\} \rightsquigarrow$$

$3w_{31} + 4w_{32} = 0 \Rightarrow w_{32} = -\frac{3}{4} w_{31}$

Given the orientation of $\underline{w}_3$, we may choose $w_{31} = 1$

$\Rightarrow w_{32} = -\frac{3}{4}$

$\theta_3 = -w_{31} - w_{32} = -1 + \frac{3}{4} = -\frac{1}{4}$

$$\Rightarrow \boxed{\underline{w}_3 = \begin{pmatrix} 1 \\ -\frac{3}{4} \end{pmatrix}, \quad \theta_3 = -\frac{1}{4}}$$

c) In the hidden space, the problem is 3-dimensional. It sufficies to find where the region inside the triangle maps to in the hidden space.

Given the orientations of the weight vectors $\underline{w}_1, \underline{w}_2, \underline{w}_3$, the region inside the triangle maps to $\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$.

// We can check by taking any point within the triangle, say $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

$V_1 = \text{sgn}\left[+ \frac{5}{3} + \left(1 \;\; \frac{2}{3}\right)\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right] = \text{sgn}\left(\frac{5}{3} + \frac{2}{3}\right) = 1$

$$V_2 = \mathrm{sgn}\left[-13 + (-1 \quad 5)\begin{pmatrix}0\\1\end{pmatrix}\right] = \mathrm{sgn}(-13+5) = -1$$

$$V_3 = \mathrm{sgn}\left[\frac{1}{4} + \left(1 \quad -\frac{3}{4}\right)\begin{pmatrix}0\\1\end{pmatrix}\right] = \mathrm{sgn}\left(\frac{1}{4} - \frac{3}{4}\right) = -1 \quad //$$

Therefore, we illustrate the problem in the hidden space as follows:

● → output 1

○ → output −1

(though one white ball is not present; this doesn't make a difference. We only need to separate the black ball from the remainder of the cube.

A plane that separates the black point from the rest is indicated in the figure. We choose the following three points that lie on this plane:

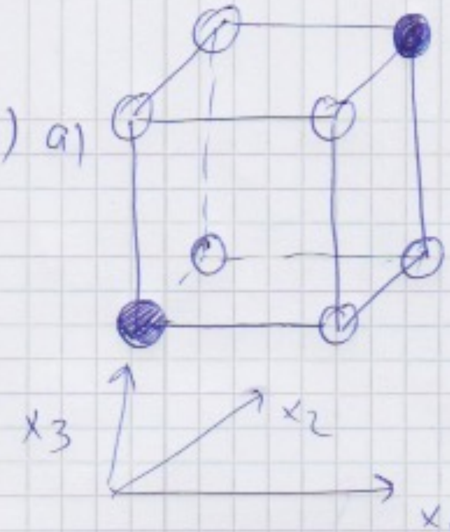
$$\begin{pmatrix}0\\-1\\-1\end{pmatrix}, \begin{pmatrix}1\\0\\-1\end{pmatrix}, \text{ and } \begin{pmatrix}1\\-1\\0\end{pmatrix}$$

$$\left.\begin{aligned} -\theta + (w_1\ w_2\ w_3)\begin{pmatrix}0\\-1\\-1\end{pmatrix} &= 0 \\ -\theta + (w_1\ w_2\ w_3)\begin{pmatrix}1\\0\\-1\end{pmatrix} &= 0 \\ -\theta + (w_1\ w_2\ w_3)\begin{pmatrix}1\\-1\\0\end{pmatrix} &= 0 \end{aligned}\right\} \left\{\begin{aligned} -\theta - w_2 - w_3 &= 0 \\ -\theta + w_1 - w_3 &= 0 \\ -\theta + w_1 - w_2 &= 0 \end{aligned}\right\}$$

$$\left.\begin{aligned} \theta &= -w_2 - w_3 \\ w_2 + \cancel{w_3} + w_1 - \cancel{w_3} &= 0 \\ \cancel{w_2} + w_3 + w_1 - \cancel{w_2} &= 0 \end{aligned}\right\} \left.\begin{aligned} \theta &= -w_2 - w_3 \\ w_1 &= -w_2 \\ w_1 &= -w_3 \end{aligned}\right\} \begin{aligned} &w_2 = w_3 = -w_1 \\ &\theta = 2w_1 \\ &\text{Choose: } w_1 = 1 \end{aligned}$$

// Direction important! // $\Rightarrow \boxed{\underline{w} = \begin{pmatrix}1\\-1\\-1\end{pmatrix},\ \theta = 2}$

4) a)



$\circ \to t^{(\mu)} = -1$

$\bullet \to t^{(\mu)} = 1$

The two black points are on the diagonal of the cube: a single plane cannot separate the two from the other white balls.
Therefore, the problem is not linearly separable.

b) $\mu = 1$: $g_1(\underline{x}^{(1)}) = \exp(-|\begin{pmatrix}0\\1\\0\end{pmatrix} - \begin{pmatrix}0\\0\\0\end{pmatrix}|^2) = \exp(-1) \approx 0.37$

$g_2(\underline{x}^{(1)}) = \exp(-|\begin{pmatrix}0\\1\\0\end{pmatrix} - \begin{pmatrix}1\\1\\1\end{pmatrix}|^2) = \exp(-2) \approx 0.14$

$\mu = 2$: $g_1(\underline{x}^{(2)}) = \exp(-|\begin{pmatrix}1\\0\\1\end{pmatrix} - \begin{pmatrix}0\\0\\0\end{pmatrix}|^2) = \exp(-2) \approx 0.14$

$g_2(\underline{x}^{(2)}) = \exp(-|\begin{pmatrix}1\\0\\1\end{pmatrix} - \begin{pmatrix}1\\1\\1\end{pmatrix}|^2) = \exp(-1) \approx 0.37$

$\mu = 3$: $g_1(\underline{x}^{(3)}) = \exp(-|\begin{pmatrix}0\\0\\0\end{pmatrix} - \begin{pmatrix}0\\0\\0\end{pmatrix}|^2) = 1$

$g_2(\underline{x}^{(3)}) = \exp(-|\begin{pmatrix}0\\0\\0\end{pmatrix} - \begin{pmatrix}1\\1\\1\end{pmatrix}|^2) = \exp(-3) \approx 0.05$

$\mu = 4$: $g_1(\underline{x}^{(4)}) = \exp(-|\begin{pmatrix}1\\0\\0\end{pmatrix} - \begin{pmatrix}0\\0\\0\end{pmatrix}|^2) = \exp(-1) \approx 0.37$

$g_2(\underline{x}^{(4)}) = \exp(-|\begin{pmatrix}1\\0\\0\end{pmatrix} - \begin{pmatrix}1\\1\\1\end{pmatrix}|^2) = \exp(-2) \approx 0.14$

$\mu = 5$: $g_1(\underline{x}^{(5)}) = \exp(-|\begin{pmatrix}1\\1\\0\end{pmatrix} - \begin{pmatrix}0\\0\\0\end{pmatrix}|^2) = \exp(-2) \approx 0.14$

$g_2(\underline{x}^{(5)}) = \exp(-|\begin{pmatrix}1\\1\\0\end{pmatrix} - \begin{pmatrix}1\\1\\1\end{pmatrix}|^2) = \exp(-1) \approx 0.37$

$\mu = 6$: $g_1(\underline{x}^{(6)}) = \exp(-|\begin{pmatrix}0\\0\\1\end{pmatrix} - \begin{pmatrix}0\\0\\0\end{pmatrix}|^2) = \exp(-1) \approx 0.37$

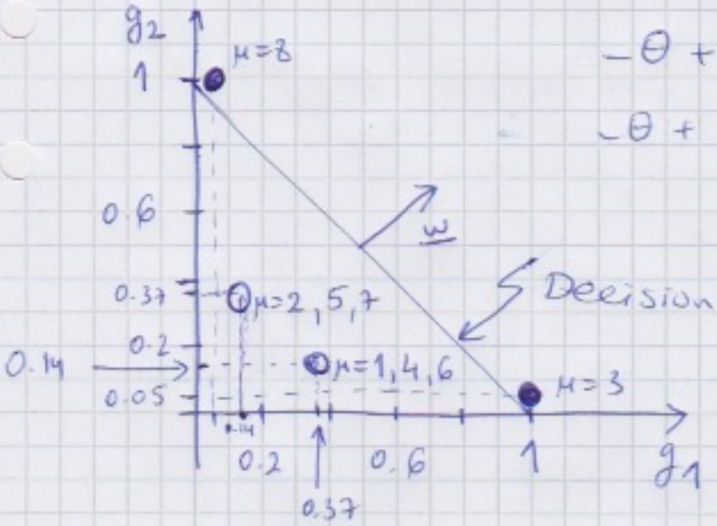
$g_2(\underline{x}^{(6)}) = \exp(-|\begin{pmatrix}0\\0\\1\end{pmatrix} - \begin{pmatrix}1\\1\\1\end{pmatrix}|^2) = \exp(-2) \approx 0.14$

$\mu = 7$ $\quad g_1(\underline{x}^{(7)}) = \exp\left(-\left|\begin{pmatrix}0\\1\\1\end{pmatrix} - \begin{pmatrix}0\\0\\0\end{pmatrix}\right|^2\right) = \exp(-2) \approx 0.14$

$\qquad g_2(\underline{x}^{(7)}) = \exp\left(-\left|\begin{pmatrix}0\\1\\1\end{pmatrix} - \begin{pmatrix}1\\1\\1\end{pmatrix}\right|^2\right) = \exp(-1) \approx 0.37$

$\mu = 8$ $\quad g_1(\underline{x}^{(8)}) = \exp\left(-\left|\begin{pmatrix}1\\1\\1\end{pmatrix} - \begin{pmatrix}0\\0\\0\end{pmatrix}\right|^2\right) = \exp(-3) \approx 0.05$

$\qquad g_2(\underline{x}^{(8)}) = \exp\left(-\left|\begin{pmatrix}1\\1\\1\end{pmatrix} - \begin{pmatrix}1\\1\\1\end{pmatrix}\right|^2\right) = \exp(0) = 1$



$$\left.\begin{aligned} -\theta + (w_1\ w_2)\begin{pmatrix}0\\1\end{pmatrix} &= 0 \\ -\theta + (w_1\ w_2)\begin{pmatrix}1\\0\end{pmatrix} &= 0 \end{aligned}\right\} \begin{aligned} -\theta + w_2 &= 0 \\ -\theta + w_1 &= 0 \end{aligned}$$

$\Rightarrow \boxed{w_1 = w_2 = \theta}$

$\underline{w}$ must point towards the region that maps to 1. Thus, we may chose

$\boxed{w_1 = w_2 = \theta = 1}$

Answer: $\boxed{\underline{w} = \begin{pmatrix}1\\1\end{pmatrix},\ \theta = 1}$

5)

The update formulae are

$$\begin{cases} w_{mn} \leftarrow w_{mn} + \Delta w_{mn} \\ \Theta_m \leftarrow \Theta_m + \Delta \Theta_m \\ W_{1m} \leftarrow W_{1m} + \Delta W_{1m} \\ \Theta_1 \leftarrow \Theta_1 + \Delta \Theta_1 \end{cases} \qquad (1)$$

with :

$$\begin{cases} \Delta w_{mn} = -\eta \dfrac{\partial H^{(n)}}{\partial w_{mn}} \\ \Delta \Theta_m = -\eta \dfrac{\partial H^{(n)}}{\partial \Theta_m} \\ \Delta W_{1m} = -\eta \dfrac{\partial H^{(n)}}{\partial W_{1m}} \\ \Delta \Theta_1 = -\eta \dfrac{\partial H^{(n)}}{\partial \Theta_1} \end{cases} \qquad (2)$$

In ②, $H^{(\mu)} = \frac{1}{2}(t_1^{(\mu)} - O_1^{(\mu)})^2$

Define

$$\begin{cases} b_m^{(\mu)} = \sum_{n=1}^{3} w_{mn} x_n^{(\mu)} - \theta_m \\ B_1^{(\mu)} = \sum_{m=1}^{5} W_{1m} V_m^{(\mu)} - \Theta_1 \end{cases}$$

Find:

$$\frac{\partial H^{(\mu)}}{\partial W_{1m}} = -(t_1^{(\mu)} - O_1^{(\mu)}) \frac{\partial}{\partial W_{1m}} g(B_1^{(\mu)})$$

$$= \underbrace{(O_1^{(\mu)} - t_1^{(\mu)}) g'(B_1^{(\mu)})}_{\equiv \Delta_1^{(\mu)}} \frac{\partial}{\partial W_{1m}} \left(\sum_{p=1}^{5} W_{1p} V_p^{(\mu)} - \Theta_1 \right)$$

$$\sum_{p} \delta_{pm} V_p$$

$$= \Delta_1^{(\mu)} V_m^{(\mu)}$$

$(*)$

$$\frac{\partial H^{(\mu)}}{\partial \Theta_1} = -\Delta_1^{(\mu)} \quad (*)$$

$$\frac{\partial H^{(\mu)}}{\partial w_{mn}} = \Delta_1^{(\mu)} \sum_{p=1}^{5} W_{1p} \frac{\partial V_p^{(\mu)}}{\partial w_{mn}}$$

$$= \Delta_1^{(\mu)} \sum_p W_{1p}\, g'(b_p^{(\mu)}) \frac{\partial}{\partial w_{mn}} \Big(\underbrace{\sum_{q=1}^{3} w_{pq} x_q^{(\mu)} - \Theta_p}_{} \Big)$$

$$= \sum_p \underbrace{\Delta_1^{(\mu)} W_{1p}\, g'(b_p^{(\mu)})}_{\equiv \delta_p^{(\mu)}} \sum_q \delta_{pm} \delta_{nq} x_q^{(\mu)}$$

$$= \delta_m^{(\mu)} x_n^{(\mu)} \quad (*)$$

$$\frac{\partial H^{(\mu)}}{\partial \Theta_m} = -\delta_m^{(\mu)}. \quad (*)$$

Update rules are given by (1) with (2) and the (*)'s inserted.

6) $x_1 = 31$

$y_1 = 31$

$z_1 = 3$

$x_2 = 15$

$y_2 = 15$

$z_2 = 10$

$x_3 = 3$

$y_3 = 3$

$z_3 = 10$

$y_4 = 10$

$y_5 = 5$

The number of weights:

1) into the convolution layer $= 3 \cdot 3 \cdot 3 \cdot 10 = \boxed{270}$

2) into the max-pooling layer $= \boxed{0}$

3) Into the fully connected layer $1 = 3 \cdot 3 \cdot 10 \cdot 10 = \boxed{900}$

4) Into the fully connected layer $2 = 10 \cdot 5 = \boxed{50}$

The number of biases:

1) Into the Convolution layer: $\boxed{10}$

2) Into the max-pooling layer: $\boxed{0}$.

3) Into the fully connected layer $1 = \boxed{10}$

4) Into the fully connected layer $2 = \boxed{5}$