100/100

Ömer Morgül NAME FAMILY NAME SECTION

EEE 443/543 Neural Networks
Final, Fall 2010

No credits will be given for unjustified answers.

Prob. 1 : (35 pt.s) Consider the Hopfield Neural Network given as :

$$o(k+1) = sgn(Wo(k) - \theta)$$

35/35

where sgn is the signum function given as

$$sgn(v) = \begin{cases} 1 & v \geq 0 \\ -1 & v < 0 \end{cases}$$

Here we consider a 2 dimensional case, where $o = \begin{pmatrix} o_1 \\ o_2 \end{pmatrix}$ is the output vector, $\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}$ is the threshold vector, and W is a 2×2 weight matrix, $k = 1, 2, \ldots$ is the iteration index. The cost $E(o)$ associated with Hopfield Network is given as $E(o) = -0.5o'Wo + \theta'o$, where the superscript $'$ denotes transpose. Assume that W and θ are given as :

$$W = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \qquad \theta = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

i : Find the fixed point(s) (i.e. stored patterns) of this network. (Note that fixed point of an iteration $x(k+1) = f(x(k))$ satisfies $x = f(x)$).

ii : Find the cost(s) associated with the fixed point(s).

iii : Let $o(1) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ be given. By using synchronous update rule, find $o(2)$ and $o(3)$, and find the costs associated with $o(1), o(2), o(3)$. Predict the future iterations, if you can.

iv : Let $o(1)$ be given as in iii. By using asynchronous update (i.e. update the first output only, then update the second output only...) update each output. Continue till the output converges to a stored pattern. (If the output does not converge, stop at 3rd update cycle). At each update step, compute the cost associated with the last updated output vector at that step.

v : Compare the results you obtained in iii and iv.

i) Fixed points $\Rightarrow$ $o(k) \to o(k) = o(k+1) = o(k+2) = \cdots = o^*$

Then $o^* = sgn(Wo^* - \theta)$ $\Rightarrow$ Fixed point equation where o^* is a fixed point

10/10

$$o^* = \begin{bmatrix} o_1^* \\ o_2^* \end{bmatrix} = sgn\left(\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} o_1^* \\ o_2^* \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) = sgn\left(\begin{bmatrix} -o_2^* - 1 \\ -o_1^* - 1 \end{bmatrix} \right)$$

Then $\left.\begin{aligned} o_1^* &= sgn(-o_2^* - 1) \\ o_2^* &= sgn(-o_1^* - 1) \end{aligned}\right\}$ Among 4 possibilities of $o = \begin{pmatrix} o_1 \\ o_2 \end{pmatrix}$: $\begin{bmatrix} +1 \\ +1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \end{bmatrix}$

X $o_1 \neq sgn(-o_2 - 1)$ ✓ ✓ X $o_1 \neq sgn(-o_2 - 1)$

There are 2 fixed points $\tilde{o}_1^* = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\tilde{o}_2^* = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

ii) $E(o) = -\frac{1}{2} o^T W o + \theta^T o$

05/05

$$E(\tilde{o}_1^*) = -\frac{1}{2} \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$= -\frac{1}{2} \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} + 0 = -\frac{1}{2} \cdot 2 = -1$$ ✓

$E(\tilde{o}_1^{*}) = -\frac{1}{2}\begin{bmatrix}-1 & 1\end{bmatrix}\begin{bmatrix}0 & -1\\ -1 & 0\end{bmatrix}\begin{bmatrix}-1\\ 1\end{bmatrix} + \begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}-1\\ 1\end{bmatrix}$

$= -\frac{1}{2}\begin{bmatrix}-1 & 1\end{bmatrix}\begin{bmatrix}-1\\ 1\end{bmatrix} + 0 = -\frac{1}{2}\cdot(2) = -1$ ✓

iii) $o(1) = \begin{bmatrix}1\\ 1\end{bmatrix}$, $o(2) = \mathrm{sgn}(W o(1) - \theta) = \mathrm{sgn}\left(\begin{bmatrix}0 & -1\\ -1 & 0\end{bmatrix}\begin{bmatrix}1\\ 1\end{bmatrix} - \begin{bmatrix}1\\ 1\end{bmatrix}\right) = \mathrm{sgn}\left(\begin{bmatrix}-2\\ -2\end{bmatrix}\right) = \begin{bmatrix}-1\\ -1\end{bmatrix}$

$o(3) = \mathrm{sgn}(W o(2) - \theta) = \mathrm{sgn}\left(\begin{bmatrix}0 & -1\\ -1 & 0\end{bmatrix}\begin{bmatrix}-1\\ -1\end{bmatrix} - \begin{bmatrix}1\\ 1\end{bmatrix}\right) = \mathrm{sgn}\left(\begin{bmatrix}0\\ 0\end{bmatrix}\right) = \begin{bmatrix}1\\ 1\end{bmatrix}$ ✓

so output will oscillate between $\begin{bmatrix}-1\\ -1\end{bmatrix}$ and $\begin{bmatrix}1\\ 1\end{bmatrix}$ ✓

$o(1): \begin{bmatrix}1\\ 1\end{bmatrix} \to \begin{bmatrix}-1\\ -1\end{bmatrix} \to \begin{bmatrix}1\\ 1\end{bmatrix} \to \begin{bmatrix}-1\\ -1\end{bmatrix} \to \ldots$

08/08

$E(o(1)) = -\frac{1}{2}\begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}0 & -1\\ -1 & 0\end{bmatrix}\begin{bmatrix}1\\ 1\end{bmatrix} + \begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}1\\ 1\end{bmatrix}$

$= -\frac{1}{2}\begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}-1\\ -1\end{bmatrix} + 2 = -\frac{1}{2}\cdot(-2) + 2 = 3$ ✓

$E(o(2)) = -\frac{1}{2}\begin{bmatrix}-1 & -1\end{bmatrix}\begin{bmatrix}0 & -1\\ -1 & 0\end{bmatrix}\begin{bmatrix}-1\\ -1\end{bmatrix} + \begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}-1\\ -1\end{bmatrix}$ ✓

$= -\frac{1}{2}\begin{bmatrix}-1 & -1\end{bmatrix}\begin{bmatrix}1\\ 1\end{bmatrix} - 2 = -\frac{1}{2}\cdot(-2) - 2 = -1$

$E(o(3)) = E(o(1)) = 3$

iv) $o(1) = \begin{bmatrix}1\\ 1\end{bmatrix}$ $W o(1) - \theta = \begin{bmatrix}-2\\ -2\end{bmatrix}$ → only update first output, $\mathrm{sgn}(-2) = -1$,

fixed point ← $o(2) = \begin{bmatrix}-1\\ 1\end{bmatrix}$

Check

$W o(2) - \theta = \begin{bmatrix}0 & -1\\ -1 & 0\end{bmatrix}\begin{bmatrix}-1\\ 1\end{bmatrix} - \begin{bmatrix}1\\ 1\end{bmatrix} = \begin{bmatrix}-1\\ 1\end{bmatrix} - \begin{bmatrix}1\\ 1\end{bmatrix} = \begin{bmatrix}-2\\ 0\end{bmatrix}$ only update second $\mathrm{sgn}(0) = 1$

$o(3) = \begin{bmatrix}-1\\ 1\end{bmatrix}$ $o(3) = \begin{bmatrix}-1\\ 1\end{bmatrix} = o(2)$

07/07

$o(4) = \begin{bmatrix}-1\\ 1\end{bmatrix}$ $W o(3) - \theta = \begin{bmatrix}0 & -1\\ -1 & 0\end{bmatrix}\begin{bmatrix}-1\\ 1\end{bmatrix} - \begin{bmatrix}1\\ 1\end{bmatrix} = \begin{bmatrix}-2\\ 0\end{bmatrix}$ → only update first, $\mathrm{sgn}(-2) = -1$

$o(4) = \begin{bmatrix}-1\\ 1\end{bmatrix} = o(3)$ ... ($\begin{bmatrix}-1\\ 1\end{bmatrix}$ is fixed point)

$E(o(1)) = 3$, as calculated in iii) $E(o(2)) = E(o(3)) = \ldots = -1$, as calculated in part ii)

v) With synchronous update rule, output oscillates between 2 different levels and it does not converge to a fixed point, however with asynchronous update, output converges to a stored point $\begin{bmatrix}-1\\ 1\end{bmatrix}$.

Also at asynchronous upd., cost at each iteration decreases or remains same ($3 \to -1 \to -1 \to \ldots$), however at synch. update it decreases and increases in a repetitive fashion.

05/05

The reason of that is "change of cost per iteration" contains some higher order terms and they are zero if asynchronous update is used and diagonal terms of W = 0. So cost will only decrease or remain same ✓ in part iv) in asynchronous update case.

30/30

Prob. 2 : (30 pt.s) This problem is related to the relation between the minimum distance classifiers and Support Vector Machines. Let the two classes C_1 and C_2 be given by single patterns $x_1 \in R^n$ and $x_2 \in R^n$, i.e. $C_1 = \{x_1\}$ and $C_2 = \{x_2\}$. We want to design a single neuron classifier in the form $o = sgn(w'x - \theta)$, where $w \in R^n$ is the weight vector, $x \in R^n$ is the input vector, θ is the threshold, and the sgn function is the signum function as given in Problem 1. The superscript $'$ denotes the transpose. We want $o = 1$ when $x = x_1$, (i.e. $x \in C_1$), and $o = -1$ when $x = x_2$, (i.e. $x \in C_2$). The minimum distance classifier gives a solution as $w = (x_1 - x_2)$, and $\theta = 0.5(\|x_1\|^2 - \|x_2\|^2)$. Let us call this as the first classifier.

Obviously, for any $\alpha > 0$, we may use $\hat{w} = \alpha w$, and $\hat{\theta} = \alpha\theta$ and $o = sgn(\hat{w}'x - \hat{\theta})$ is also a valid classifier. Let us call this as second classifier.

i : For the first classifier compute $v_1 = w'x_1 - \theta$ and $v_2 = w'x_2 - \theta$, and show that $v_1 = -v_2$, with $v_1 > 0$.

ii : Show that when $\alpha = \frac{2}{\|x_1 - x_2\|^2}$, the second classifier given above becomes a support vector machine classifier for C_1 and C_2. (Do not give a geometrical argument, use either Karush-Kuhn-Tucker conditions or the Lagrangian formulation given in the notes).

iii : With the α, w and θ given above, show that $\alpha w'x_1 - \alpha\theta = 1$ and $\alpha w'x_2 - \alpha\theta = -1$, i.e. x_1 and x_2 are support vectors.

For this problem the following equalities might be helpful : $\|x\|^2 = x'x$, $x'y = y'x$, $\|x - y\|^2 = (x - y)'(x - y) = \|x\|^2 - 2x'y + \|y\|^2$.

i) $w^T x_1 - \theta = v_1 = (x_1 - x_2)^T x_1 - 0.5\left(\|x_1\|^2 - \|x_2\|^2\right)$

$= \|x_1\|^2 - x_1^T x_2 - 0.5\|x_1\|^2 + 0.5\|x_2\|^2$

$= \frac{1}{2}\cdot\left(\|x_2\|^2 - 2x_1^T x_2 + \|x_1\|^2\right) = \frac{1}{2}\|x_1 - x_2\|^2 > 0$

06/06

$w^T x_2 - \theta = v_2 = (x_1 - x_2)^T x_2 - 0.5\left(\|x_1\|^2 - \|x_2\|^2\right)$

$= x_1^T x_2 - \|x_2\|^2 - \frac{1}{2}\|x_1\|^2 + \frac{1}{2}\|x\|^2$

$= -\frac{1}{2}\left(\|x_1\|^2 - 2x_1^T x_2 + \|x_2\|^2\right) = -\frac{1}{2}\|x_1 - x_2\|^2 < 0$

so $v_1 = -v_2 = \frac{1}{2}\|x_1 - x_2\|^2$.

ii) if $\alpha = \frac{2}{\|x_1 - x_2\|^2}$ Then $\hat{w} = \alpha w = \frac{2(x_1 - x_2)}{\|x_1 - x_2\|^2}$, $\hat{\theta} = \alpha\theta = \frac{\|x_1\|^2 - \|x_2\|^2}{\|x_1 - x_2\|^2}$ we expect to get this after svm calculations.

Use Lagrangian formulation. $\max_{\alpha} L(\alpha)$ s.t. ① $\alpha_1 \geq 0$, $\alpha_2 \geq 0$ ② $\sum_{i=1}^{2} \alpha_i d_i = 0 \Rightarrow \alpha_1 - \alpha$ $\boxed{\alpha_1 = \alpha_2}$

$\begin{pmatrix} d_1 = +1 \\ d_2 = -1 \end{pmatrix}$

$L(\alpha) = \sum_{i=1}^{2} \alpha_i - \frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2} \alpha_i \alpha_j d_i d_j x_i^T x_j$,

$= \alpha_1 + \alpha_2 - \frac{1}{2}\left(\alpha_1^2 \cdot \|x_1\|^2 + 2\cdot\alpha_1\alpha_2\cdot(-1)\cdot x_1^T x_2 + \alpha_2^2 \cdot \|x_2\|^2\right)$

$= \alpha_1 + \alpha_2 - \frac{1}{2}\alpha_1^2\|x_1\|^2 - \frac{1}{2}\alpha_2^2\|x_2\|^2 + \alpha_1\alpha_2 x_1^T x_2$ $(\alpha_1 = \alpha_2)$

$= 2\alpha_1 - \left(\frac{\|x_1\|^2 + \|x_2\|^2}{2}\right)\alpha_1^2 + \alpha_1^2 x_1^T x_2$

$\frac{\partial L(\alpha)}{\partial \alpha_1} = 2 - \left(\|x_1\|^2 + \|x_2\|^2\right)\cdot\alpha_1 + 2\alpha_1 x_1^T x_2 = 0 \Rightarrow \alpha_1 = \alpha_2 = \frac{2}{\|x_1\|^2 + \|x_2\|^2 - 2x_1^T x_2} = \frac{2}{\|x_1 -}$

$\left(\frac{\partial^2 L(\alpha_1)}{\partial \alpha_1^2} = -(\|x_1 - x_2\|) < 0\right)$ so this will give the maximum

Then $w_{svm} = \sum_{i=1}^{2} \alpha_i d_i x_i = \alpha_1 x_1 - \alpha_2 x_2 = \alpha_1 (x_1 - x_2)$

$$= \frac{2}{\|x_1 - x_2\|^2} \cdot (x_1 - x_2) = \frac{2(x_1 - x_2)}{\|x_1 - x_2\|^2}$$

$$\theta_{svm} = \frac{1}{2} \sum_{i=1}^{2} (w^T x_i - d_i) = \frac{1}{2} \cdot \left[\frac{2(x_1 - x_2)^T x_1}{\|x_1 - x_2\|^2} \cancel{-1} + \frac{2(x_1 - x_2)^T x_2}{\|x_1 - x_2\|^2} \cancel{+1} \right]$$

$$= \frac{1}{2} \cdot \frac{2\|x_1\|^2 - \cancel{2 x_1^T x_2} + \cancel{2 x_1^T x_2} - 2\|x_2\|^2}{\|x_1 - x_2\|^2}$$

$$= \frac{\|x_1\|^2 - \|x_2\|^2}{\|x_1 - x_2\|^2}$$

20/20

So $\left.\begin{array}{l} w_{svm} = \hat{w} \\ \theta_{svm} = \hat{\theta} \end{array}\right\}$ second classifier is support vector machine classifier in fac

iii) $\alpha w^T x_1 - \alpha\theta = \alpha \underbrace{(w^T x_1 - \theta)}_{1/2 \|x_1 - x_2\|^2} \rightarrow$ part i) , $\alpha = \frac{2}{\|x_1 - x_2\|^2}$

$$= \frac{1}{2}\|x_1 - x_2\|^2 \cdot \frac{2}{\|x_1 - x_2\|^2} = 1$$

$\alpha w^T x_2 - \alpha\theta = \alpha \cdot \underbrace{(w^T x_2 - \theta)}_{-\frac{1}{2}\|x_1 - x_2\|^2} \rightarrow$ part ii), $\alpha = \frac{2}{\|x_1 - x_2\|^2}$

$$= \frac{2}{\|x_1 - x_2\|^2} \cdot \left(-\frac{1}{2}\right) \|x_1 - x_2\|^2 = -1$$

04/04

OK

Hence $\left.\begin{array}{l} \alpha w^T x_1 - \alpha\theta = \hat{w}^T x_1 - \hat{\theta} = 1 \rightarrow x_1 \text{ is on } H^+ \\ \alpha w^T x_2 - \alpha\theta = \hat{w}^T x_2 - \hat{\theta} = -1 \rightarrow x_2 \text{ is on } H^- \end{array}\right\}$ x_1 and x_2 are support vectors

Note that this can be deduced from part 2) as well. We know that in support vec. mach. classification, $\alpha_i > 0$ only if x_i is s.v. and $\alpha_i = 0$ o.w. So $\alpha_1 > 0$, $\alpha_2 > 0$ in part ii), then x_1 and x_2 are support vectors

35/35

Prob. 3 : (35 pt.s) This problem is related to Kohonen Network and winner-take-all algorithm. Consider the 2×2 pixel template given as picture p below. The pattern $z = (x_1\ x_2\ x_3\ x_4)'$ is computed as usual : $x_i = +1$ if i^{th} pixel is black, and $x_i = -1$ if i^{th} pixel is white. The superscript $'$ denotes the transpose. Now consider the pictures p_1, p_2, p_3, p_4 given below.

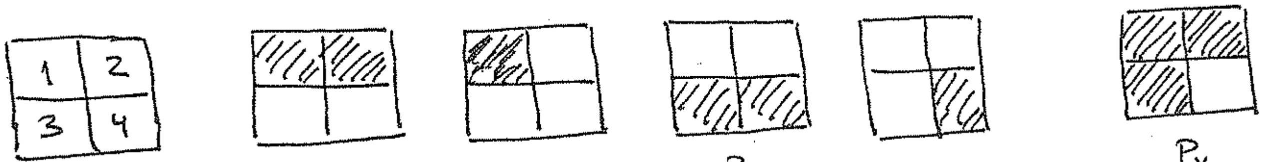
i : According to the coding given above, find the patterns z_1, z_2, z_3, z_4.

ii : Now consider a 4-input, 2-output Kohonen network given below. Linear sums v_1 and v_2 are given as $v_1 = w_1' z$, $v_2 = w_2' z$; here z is the input pattern, w_1 and w_2 are the weight vectors for first and second neurons, thresholds are taken as zero, and the nonlinearity is the standard bipolar sigmoidal function. Initial weights are given as $w_1(0) = (1\ \ 0\ \ -1\ \ -1)'$, $w_2(0) = (-1\ \ -1\ \ 0\ \ 1)'$. By using $\alpha = 1$, and without using normalization, apply the winner-take-all learning algorithm for one epoch. Find the last updated weights at the end of epoch. For the rest of this problem, use these last updated weights.

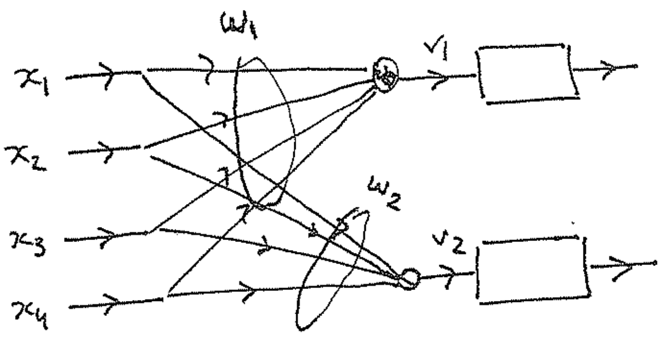
(Note that in actual algorithm, $\alpha < 1$, and normalization should be used. Here we choose $\alpha = 1$ and no normalization just to make computations simple).

iii : Let us define the classes C_1 and C_2 as $z_j \in C_i$ if v_i is maximum, $i = 1, 2$, $j = 1, 2, 3, 4$. According to this selection scheme, determine which pattern belongs to which class.

iv : Now consider a new picture p_* as given below. Find its pattern z_*, compute v_1, v_2 and determine to which class this picture belongs.



P $\quad P_1 \quad P_2 \quad P_3 \quad P_4 \quad P_*$



i) $z_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}$, $z_2 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ -1 \end{bmatrix}$

05/05

$z_3 = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}$, $z_4 = \begin{bmatrix} -1 \\ -1 \\ -1 \\ 1 \end{bmatrix}$

ii) $v_1 = w_1^T z$, $\quad v_2 = w_2^T z$ $\qquad w_1(0) = \begin{bmatrix} 1 \\ 0 \\ -1 \\ -1 \end{bmatrix}$ $\quad w_2(0) = \begin{bmatrix} -1 \\ -1 \\ 0 \\ 1 \end{bmatrix}$

$\alpha = 1$.

$W(0) = \begin{bmatrix} w_1(0)^T \\ w_2(0)^T \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 & -1 \\ -1 & -1 & 0 & 1 \end{bmatrix}$, $\quad v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ (find the max entry at each time)

$v = Wz$

→ Start epoch

* $W(0) \cdot z_1 : \begin{bmatrix} 1 & 0 & -1 & -1 \\ -1 & -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ -3 \end{bmatrix}$ → winner neuron, update only the 1'st weight vector

$W_2(1) = W_2(0)$, $\quad W_1(1) = W_1(0) + \alpha\,(z_1 - W_1(0)) = z_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}$

* $W(1) \cdot z_2 = \begin{bmatrix} 1 & 1 & -1 & -1 \\ -1 & -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ → winner neuron, update only the 1'st wei

$W_2(2) = W_2(1)$

$W_1(2) = W_1(1) + z_2 - W_1(1) = \begin{bmatrix} 1 \\ -1 \\ -1 \\ -1 \end{bmatrix}$

* $W(2) \cdot z_3 = \begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ → winner neuron
update only the second weight vector

$W_1(3) = W_1(2)$

$W_2(3) = \cancel{W_2(2)} + z_3 - \cancel{W_2(2)} = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}$

* $W(3) \cdot z_4 = \begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$ → winner neuron, update only the second weigh vector

16/16

$W_1(4) = W_1(3)$

$W_2(4) = \cancel{W_2(3)} + z_4 - \cancel{W_2(3)} = \begin{bmatrix} -1 \\ -1 \\ -1 \\ 1 \end{bmatrix}$

→ End of epoch

Last updated weights $W_1^* = \begin{bmatrix} 1 \\ -1 \\ -1 \\ -1 \end{bmatrix}$, $W_2^* = \begin{bmatrix} -1 \\ -1 \\ -1 \\ 1 \end{bmatrix}$

iii) $W^* = \begin{bmatrix} W_1^{*T} \\ W_2^{*T} \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & -1 & -1 & 1 \end{bmatrix}$

$W^* z_1 = \begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$ → $v_1 = W_1^{*T} z_1$ is maximum so $z_1 \in C_1$

$W^* z_2 = \begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$ → $v_1 = W_1^{*T} z_2$ is maximum so $z_2 \in C_1$

09/09

$W^* z_3 = \begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$ → $v_2 = W_2^{*T} z_3$ is max. so $z_3 \in C_2$

$W^* z_4 = \begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \end{bmatrix}$ → $v_2 = W_2^{*T} z_4$ is max. so $z_4 \in C_2$

Hence $\{z_1, z_2\} \in C_1$ and $\{z_3, z_4\} \in C_2$

iv) $z_* = \begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix}$

$v_1 = W_1^{*T} z_* = \begin{bmatrix} 1 & -1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix} = 0$

05/05

$v_2 = W_2^{*T} z_* = \begin{bmatrix} -1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix} = -4$

So $v_1 > v_2 \Rightarrow z_*$ belongs to Class 1