Ömer Morgül

$\overline{NAME}$ $\overline{FAMILYNAME}$ $\overline{SECTION}$

EEE 443/543 Neural Networks
Final, Fall 2011

No credits will be given for unjustified answers.

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

$$o(k+1) = sgn(Wo(k))$$

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, the threshold θ is taken as $\theta = 0$, 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 $E(o) = -0.5o'Wo$, where the superscript $'$ denotes transpose. Assume that W is given as :

$$W = \begin{pmatrix} 0 & \omega \\ \omega & 0 \end{pmatrix} \quad .$$

Here ω is a real constant.

i : Find the range of ω so that the vector $o^* = (-1\ 1)'$ becomes a fixed point(i.e. a stored pattern) of this network. (Note that fixed point of an iteration $x(k+1) = f(x(k))$ satisfies $x = f(x)$).

ii : By using outer-product rule, find a weight matrix W_* to store the given pattern o^*. Show that with W_*, o^* becomes a fixed point. Does W_* satisfy the condition found in **i**?

iii : Assume $w = -2$. Find the fixed point(s) (i.e. stored patterns) of this network. Find the cost(s) associated with the fixed point(s).

iv : Consider the network given in **iii**. Let $o(1) = (1\ 1)'$. Find the cost associated with $o(1)$. By using synchronous update, find $o(2)$ and $o(3)$. Can you predict the following iterations? If yes, give an expression for future iterations. Also find the costs associated with $o(2)$ and $o(3)$.

v : Let the network and $o(1)$ be given as in **iv**. By using asynchronous update, find $o(2)$ and $o(3)$. Can you predict the following iterations? If yes, give an expression for future iterations. Also find the costs associated with $o(2)$ and $o(3)$.

Prob. 2 : (30 pt.s) Let x_1, x_2 denote two scalar inputs, and set $z = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbf{R}^2$. Consider the single neuron given below; as usual the scalar output o is given by $o = sgn(w_1 x_1 + w_2 x_2 - \theta)$; here w_1, w_2 are scalar weights, θ is is the scalar threshold value and and sgn is the signum function given as $f(v) = 1$ when $v \geq 0$ and $f(v) = -1$ when $v < 0$. Let us define the training set vectors as :

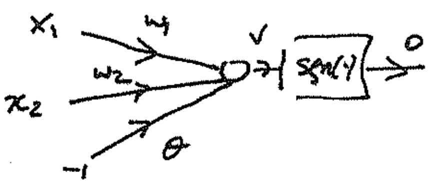
$$z_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad z_2 = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \quad z_3 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad z_4 = \begin{pmatrix} -2 \\ 0 \end{pmatrix}.$$

Consider the classes $C_1 = \{z_1,\ z_2\}$, $C_2 = \{z_3,\ z_4\}$; the desired outputs are $d = 1$ for C_1 and $d = -1$ for C_2.

i : Using only geometrical arguments, find the best separating line (i.e. support vector machine) for these classes. Determine the resulting weights $w_1,\ w_2$ and threshold θ.

ii : By using mathematical analysis (e.g. Lagrangian formulation), find the support vectors and support vector machine for these classes. Determine the resulting weights $w_1,\ w_2$ and threshold θ.

iii : Let z_i denote the support vector(s) in C_1 and z_j denote the support vector(s) in C_2. Show that $w'z_i - \theta = 1$ and $w'z_j - \theta = -1$. (Here $'$ denotes transpose and $w' = (w_1\ w_2)$.)



Prob. 3 : (35 pt.s) This problem is related to Kohonen Network and winner-take-all algorithm. Consider the neural network given below, where nonlinearity $f(\cdot)$ is bipolar sigmoidal function. The neuron weights are as indicated on the figure, and the thresholds are taken as 0. The input to the network is given as $z = (x_1 \ x_2)'$, where $'$ is the transpose. Consider the training set given as $S = \{z_1 \ , \ z_2 \ , \ z_3 \ , \ z_4\}$ where the patterns are given as

$$z_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \ , \ z_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \ , \ z_3 = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \ , \ z_4 = \begin{pmatrix} -1 \\ -1 \end{pmatrix} .$$

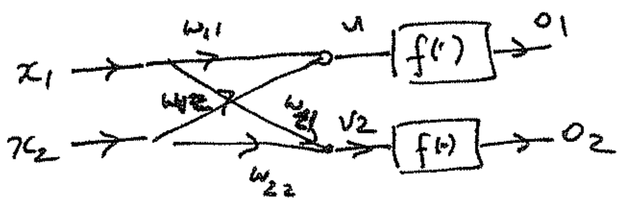
i : Initial weights are given as $w_1(0) = (1 \ \ 0)'$, $w_2(0) = (0 \ -1)'$. By using $\alpha = 0.5$, and without using normalization, apply the winner-take-all learning algorithm for one epoch. Find the last updated weights at the end of epoch. For the rest of this problem, use these last updated weights.

ii : 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 (standard maximum) selection scheme, determine which pattern belongs to which class.

iii : Let w_1 and w_2 denote the last updated weights. Determine the cluster centers z_{1c} and z_{2c} of C_1 and C_2 (i.e arithmetic average of patterns in each class). Find the distances between the weights and cluster centers $\|w_i - z_{jc}\|$ for $i = 1, 2$, $j = 1, 2$. Determine the cluster center which is close to w_1; and the cluster center which is close to w_2.

iv : Consider the neural network given below with the last updated weights. Find a separating line of the form $a_1x_1 + a_2x_2 = 0$ such that for the patterns $z = (x_1 \ x_2)'$ above the line (i.e $a_1x_1 + a_2x_2 \ > \ 0$), neuron 1 wins, and for the patterns below the line, neuron 2 wins.

iv : Let $z = (0 \ 1)'$ be given. Compute v_1, v_2 and determine to which class z belongs.



$$W_1 = \begin{pmatrix} w_{11} \\ w_{12} \end{pmatrix} \qquad W_2 = \begin{pmatrix} w_{21} \\ w_{22} \end{pmatrix}$$

FALL 2011 — NEURAL NETWORKS FINAL SOLUTIONS

P1.

i) (05) $o^* = \text{sgn}(w o^*)$ (01) $\Rightarrow \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \text{sgn}(w o^*) = \text{sgn}\begin{pmatrix} w \\ -w \end{pmatrix}$ (02) $\Rightarrow$ $-1 = \text{sgn}(w) \Rightarrow w < 0$; $1 = \text{sgn}(-w) \Rightarrow -w \geq 0 \Rightarrow w < 0$

$\Rightarrow \boxed{w < 0}$ (03)

ii) (05) $W_* = o^* \cdot o^{*T} - I = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \begin{pmatrix} -1 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

$\Rightarrow W_* = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$ (03) $\Rightarrow$ same form with $\boxed{w = -1}$ (02)

iii) (08) $w = -2 \Rightarrow$ $o_1^* = \text{sgn}(-2 o_2^*) = \text{sgn}(-o_2^*)$, $o_2^* = \text{sgn}(-2 o_1^*) = \text{sgn}(-o_1^*)$ (02) $\Rightarrow o_A^* = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ (02) $o_B^* = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ (02)

$E(o_A^*) = -\frac{1}{2} \begin{pmatrix} -1 & 1 \end{pmatrix} \begin{pmatrix} -2 \\ 2 \end{pmatrix} = -\frac{1}{2}(2+2) = -2$ (01)

$E(o_B^*) = -\frac{1}{2} \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} 2 \\ -2 \end{pmatrix} = -\frac{1}{2}(2+2) = -2$ (01)

iv) (08) $o(1) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rightarrow E(o(1)) = -\frac{1}{2} \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} -2 \\ -2 \end{pmatrix} = -\frac{1}{2}(-4) = 2$ (02)

$o(2) = \text{sgn}(W o(1)) = \text{sgn}\begin{pmatrix} -2 \\ -2 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \end{pmatrix} \rightarrow E(o(2)) = 2$ (02)

$o(3) = \text{sgn}(W o(2)) = \text{sgn}\begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rightarrow E(o(3)) = +2$ (02)

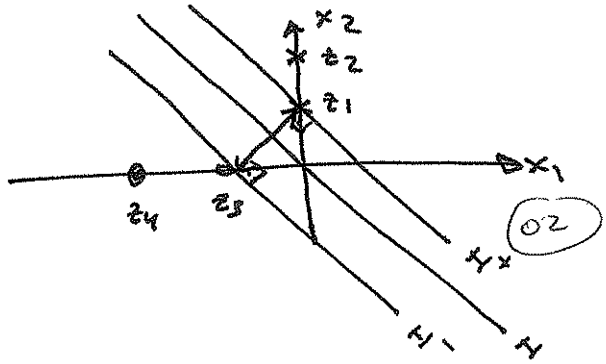
$\Rightarrow o(1) = o(3) = \dots = o(2k+1) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $o(2) = o(4) = \dots = o(2k) = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$ (02)

v) (07/07) $o(1) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} \text{sgn}(-2) \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} -1 \\ \text{sgn}(2) \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ (02)

$\Downarrow$ (02) $E = -2$; $\Downarrow$ $E = -2$ (01)

$\rightarrow$ o repeats $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ in future iterations. (02)

(P2) i) (06)



Smallest distance is between z_1 & z_3 $\Rightarrow$ z_1 & z_3 are support vectors (02)

The line passing in between

$\alpha(x_1+x_2)=0 \quad \alpha>0$

$\Rightarrow w_1=\alpha \quad w_2=\alpha \quad \alpha>0.$

To get $\alpha(x_1+x_2)=1$ at H^+ $\Rightarrow$ $\alpha.1=1 \Rightarrow \boxed{\alpha=1} \Rightarrow \boxed{w_1=1,\ w_2=1,\ \theta=0}$ (02)

ii) (20/20)

$z_1^Tz_1=1,\quad z_2^Tz_1=2,\quad z_3^Tz_1=0,\quad z_4^Tz_1=0$

$z_1^Tz_2=2,\quad z_2^Tz_2=4,\quad z_3^Tz_2=0,\quad z_4^Tz_2=0$

$z_1^Tz_3=0,\quad z_2^Tz_3=0,\quad z_3^Tz_3=1,\quad z_4^Tz_3=2$

$z_1^Tz_4=0,\quad z_2^Tz_4=0,\quad z_3^Tz_4=2,\quad z_4^Tz_4=4$ (02)

$$\Rightarrow L(\alpha)=\sum_{i=1}^{4}\alpha_i-\frac{1}{2}\sum_{j=1}^{4}\sum_{i=1}^{4}\alpha_i\alpha_j d_i d_j z_i^Tz_j$$

$$=(\alpha_1+\alpha_2+\alpha_3+\alpha_4)-\frac{1}{2}\left(\alpha_1^2+4\alpha_1\alpha_2+4\alpha_2^2+\alpha_3^2+4\alpha_3\alpha_4+4\alpha_4^2\right)$$

$$=\underbrace{\left(\alpha_1+\alpha_2-\frac{1}{2}\alpha_1^2-2\alpha_1\alpha_2-2\alpha_2^2\right)}_{L_1}\ (02)+\underbrace{\left(\alpha_3+\alpha_4-\frac{1}{2}\alpha_3^2-2\alpha_3\alpha_4-2\alpha_4^2\right)}_{L_2}\ (02)$$

$L_1(\alpha_1,\alpha_2)=\alpha_1+\alpha_2-\frac{1}{2}\alpha_1^2-2\alpha_1\alpha_2-2\alpha_2^2 \Rightarrow \max_{\alpha_1,\alpha_2} L_1$

$\Rightarrow \frac{\partial L_1}{\partial\alpha_1}=1-\alpha_1-2\alpha_2=0 \Rightarrow \alpha_1=1-2\alpha_2$ (1) (02)

$\frac{\partial L_1}{\partial\alpha_2}=1-4\alpha_2-2\alpha_1=0 \Rightarrow \alpha_1=\frac{1}{2}-2\alpha_2$ (2) (02)

use (1) in L_1 $\Rightarrow$ $L_1=(1-2\alpha_2)+\alpha_2-\frac{1}{2}(1-2\alpha_2)^2-2(1-2\alpha_2)\alpha_2-2\alpha_2^2$

$=\frac{1}{2}-\alpha_2 \Rightarrow$ max at $\alpha_2=0$, $\alpha_1=1$ $\quad L_{1max}=1/2$ (02)

use (2) in L_1 $\Rightarrow$ $L_1=\left(\frac{1}{2}-2\alpha_2\right)+\alpha_2-\frac{1}{2}\left(\frac{1}{2}-2\alpha_2\right)^2-2\left(\frac{1}{2}-2\alpha_2\right)\alpha_2-2\alpha_2^2$

$=\frac{3}{8}-\alpha_2 \Rightarrow$ max at $\alpha_2=0$, $\alpha_1=\frac{1}{2}$ $\rightarrow L_{1max}=3/8$ (02)

$\frac{1}{2}$ is BIGGER $\Rightarrow$ $\alpha_1=1$, $\alpha_2=0$ $\Rightarrow$ z_1 is SUPPORT VECTOR (01)

L_2 is similar $\Rightarrow$ $\alpha_3=1$, $\alpha_4=0$ $\Rightarrow$ z_3 is SUPPORT VECTOR (01)

$$\Rightarrow \quad w = \sum_{i=1}^{4} \alpha_i d_i z_i = \alpha_1 z_1 + \alpha_3 z_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = w$$

$\Rightarrow$ $\boxed{w_1 = w_2 = 1}$ (02) $\quad w^T z_1 - \theta = 1 \Rightarrow$ (02) $1 - \theta = 1 \Rightarrow$ $\theta = 0$

$w^T z_3 - \theta = -1 \Rightarrow \quad -1 - \theta = -1 \Rightarrow$ $\theta = 0$

iii) Already shown above.

$w^T z_1 - \theta = (1 \quad 1) \begin{pmatrix} 0 \\ 1 \end{pmatrix} - 0 = 1$ (02)

$w^T z_3 - \theta = (1 \quad 1) \begin{pmatrix} -1 \\ 0 \end{pmatrix} - 0 = -1$ (02)

Support vector for $C_1 \Rightarrow z_1$

Support vector for $C_2 \Rightarrow z_3$

(04/04)

P3: i) (10)

$z_1 \rightarrow \quad v_1 = w_1^T z_1 = 1 \Rightarrow \text{wins} \qquad w_1 \leftarrow w_1 + \alpha(z_1 - w_1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 0.5 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0.5 \end{pmatrix}$

$v_2 = w_2^T z_1 = -1$ (02)

$z_2 \rightarrow \quad v_1 = w_1^T z_2 = 0.5$

$v_2 = w_2^T z_2 = 1 \Rightarrow \text{wins}$ (02) $\qquad w_2 \leftarrow w_2 + \alpha(z_2 - w_2) = \begin{pmatrix} 0 \\ -1 \end{pmatrix} + 0.5 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.5 \\ -1 \end{pmatrix}$

$z_3 \rightarrow \quad v_1 = w_1^T z_3 = -0.5 \Rightarrow \text{wins} \qquad w_1 \leftarrow w_1 + \alpha(z_3 - w_1) = \begin{pmatrix} 1 \\ 0.5 \end{pmatrix} + 0.5 \begin{pmatrix} -2 \\ 0.5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0.75 \end{pmatrix}$

$v_2 = w_2^T z_3 = -1.5$ (02)

$z_4 \rightarrow \quad v_1 = w_1^T z_4 = -0.75$ (02)

$v_2 = w_2^T z_4 = 0.5 \Rightarrow \text{wins} \qquad w_2 \leftarrow w_2 + \alpha(z_4 - w_2) = \begin{pmatrix} 0.5 \\ -1 \end{pmatrix} + 0.5 \begin{pmatrix} -1.5 \\ 0 \end{pmatrix} = \begin{pmatrix} -0.25 \\ -1 \end{pmatrix}$

Last weights $w_1 = \begin{pmatrix} 0 \\ 0.75 \end{pmatrix} = \begin{pmatrix} 0 \\ 3/4 \end{pmatrix}$ (01) $\qquad w_2 = \begin{pmatrix} -0.25 \\ -1 \end{pmatrix} = \begin{pmatrix} -1/4 \\ -1 \end{pmatrix}$ (01)

ii) (05)

$z_1 \Rightarrow \quad v_1 = w_1^T z_1 = 0.75 \Rightarrow v_1 \text{ wins} \Rightarrow z_1 \in C_1$ (01)

$v_2 = w_2^T z_1 = -5/4$

$z_2 \Rightarrow \quad v_1 = w_1^T z_2 = -3/4$

$v_2 = w_2^T z_2 = 3/4 \Rightarrow v_2 \text{ wins} \Rightarrow z_2 \in C_2$ (01)

$z_3 \Rightarrow \quad v_1 = w_1^T z_3 = 3/4 \Rightarrow v_1 \text{ wins} \Rightarrow z_3 \in C_1$ (01)

$v_2 = w_2^T z_3 = -3/4$

$z_4 \Rightarrow \quad v_1 = w_1^T z_4 = -3/4$

$v_2 = w_2^T z_4 = 5/4 \Rightarrow v_2 \text{ wins} \Rightarrow z_4 \in C_2$ (01)

$C_1 = \{z_1, z_3\}$ (01)

$C_2 = \{z_2, z_4\}$

iii) (04)

$z_{1c} = \frac{1}{2}(z_1 + z_3) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ (01) $\qquad z_{2c} = \frac{1}{2}(z_2 + z_4) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$

$\|w_1 - z_{1c}\| = \left\| \begin{pmatrix} 0 \\ 1/4 \end{pmatrix} \right\| = \frac{1}{4}$ (01)

$\|w_1 - z_{2c}\| = \left\| \begin{pmatrix} 0 \\ 7/4 \end{pmatrix} \right\| = 7/4$ (01) $\Rightarrow$ w_1 IS CLOSER TO z_{1c} (01)

$\|w_2 - z_{1c}\| = \left\| \begin{pmatrix} -1/4 \\ -2 \end{pmatrix} \right\| = \sqrt{4 + \frac{1}{16}} = 2.01$ (01)

$\|w_2 - z_{2c}\| = \left\| \begin{pmatrix} -1/4 \\ 0 \end{pmatrix} \right\| = 1/4$ (01) $\Rightarrow$ w_2 IS CLOSER TO z_{2c}. (01)

iv) (10/10)

$v_1 = w_1^T z = \frac{3}{4} x_2$ (01)

$v_2 = w_2^T z = -1/4\, x_1 - x_2$ (02)

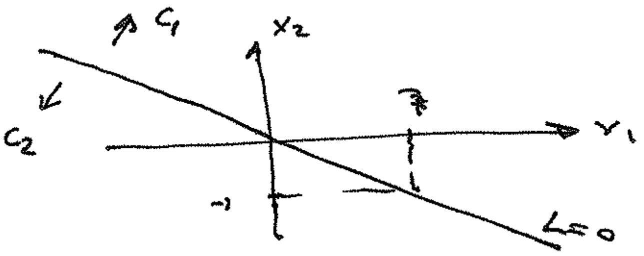
$z \in C_1$ if $v_1 > v_2$ (01) $\Rightarrow$ $\frac{3}{4} x_2 > -\frac{1}{4} x_1 - x_2$ (02) $\Rightarrow$ $\boxed{\frac{1}{4} x_1 + \frac{7}{4} x_2 > 0}$ (02)

$z \in C_2$ if $v_1 < v_2$ (01)

Line: $\boxed{\frac{1}{4} x_1 + \frac{7}{4} x_2 = 0}$ (02)

Above the line $\Rightarrow$ $z = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in C_1$

Below the line $\Rightarrow$ $z \in C_2$



v) $z = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \Rightarrow$ (03)

$v_1 = w_1^T z_1 = 3/4$ (01) $\rightarrow$ $\boxed{z \in C_1}$ (01)

$v_2 = w_2^T z_2 = -1$ (01)

(Actually ABOVE THE LINE)