100/100

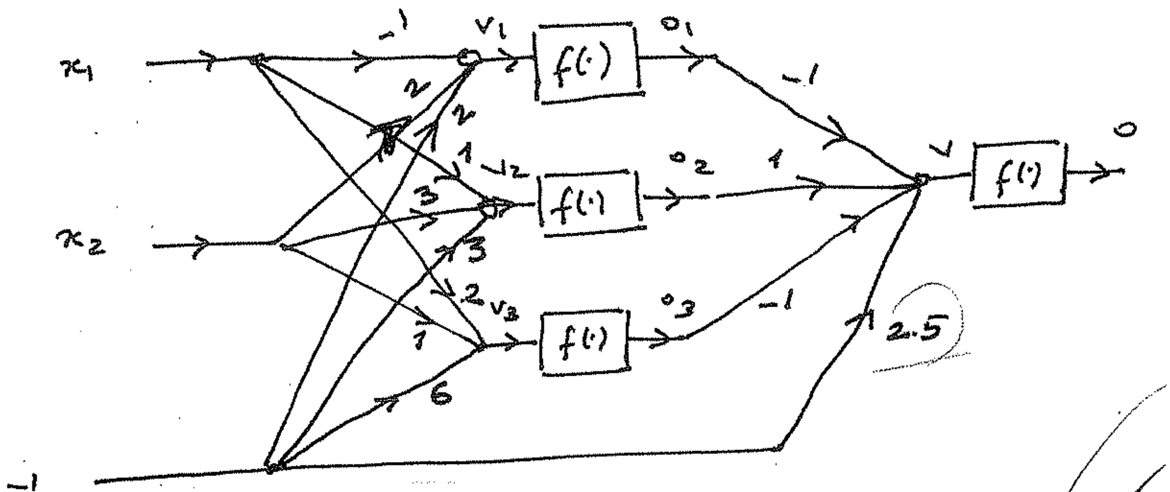
Ömer Morgül

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

EEE 443/543 Neural Networks
Midterm, Fall 2011

No credits will be given for unjustified answers.

Prob. 1 : (25 pt.s) Consider the following feedforward network. Note that as usual, $v_i = w_{i1}x_1 + w_{i2}x_2 - \theta_i$, and the nonlinear function $f(v)$ is the signum function, i.e. $f(v) = 1$ when $v \geq 0$ and $f(v) = -1$ when $v < 0$. Find the region at which $o = 1$, and indicate this region geometrically in $x_1 - x_2$ plane. The weights are indicated on the network.



25/25

$v_1 = -x_1 + 2x_2 - 2 \Rightarrow o_1 = f(v_1)$

$v_2 = x_1 + 3x_2 - 3 \Rightarrow o_2 = f(v_2)$

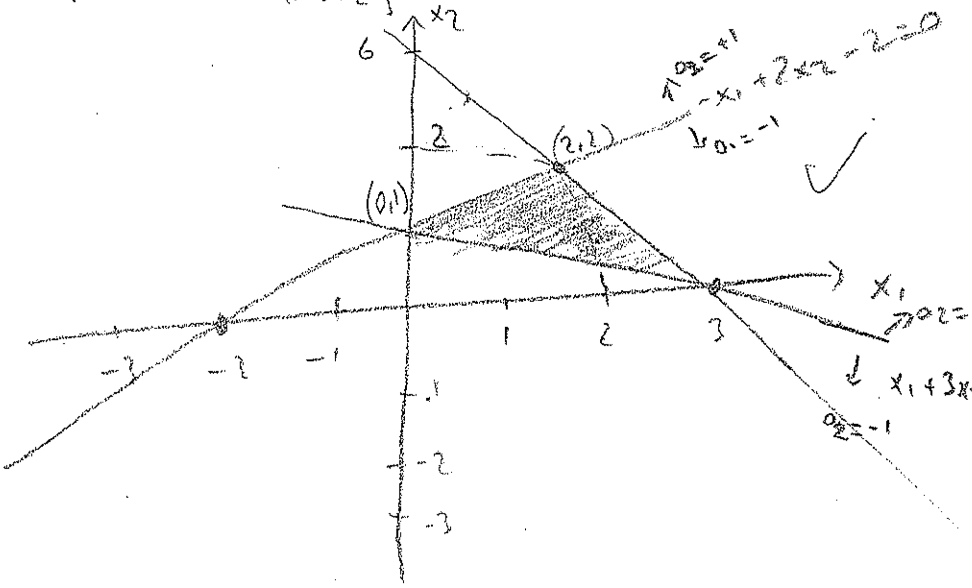
$v_3 = 2x_1 + x_2 - 6 \Rightarrow o_3 = f(v_3)$

at the end of process output is $o = \bar{o}_1 \wedge o_2 \wedge \bar{o}_3$. this is "and" operation since when $o_1 = o_3 = -1$ and $o_2 = 1$ the output is 1, ow o=-

$$o_1 = \begin{cases} 1 & -x_1 + 2x_2 - 2 \geq 0 \\ -1 & \text{ow} \end{cases} \qquad o_2 = \begin{cases} 1 & x_1 + 3x_2 - 3 \geq 0 \\ -1 & \text{ow} \end{cases}$$

$$o_3 = \begin{cases} 1 & 2x_1 + x_2 - 6 \geq 0 \\ -1 & \text{ow} \end{cases}$$

in the $x_1 - x_2$ plane we have



$-x_1 + 2x_2 - 2 = 0$

$x_1 = 0 \quad x_2 = -2$

$x_1 + 3x_2 - 3 = 0$

$x_1 = 0 \quad x_2 = 1$

$x_2 = 0 \quad x_1 = 3$

$2x_1 + x_2 - 6 = 0$

$x_1 = 0 \quad x_2 = 6$

$o_3 = 1 \quad x_2 = 0 \quad x_1 = 3$

$o_3 = -1$

$2x_1 + x_2 - 6 = 0$

$o = \bar{o}_1 \wedge o_2 \wedge \bar{o}_3$

The black region, which is a triangle with end points $(0,1)$ $(3,0)$ and $(2,2)$, gives output as 1

$2x_1 + x_2 - 6 = 0$

$-x_1 + 2x_2 - 2 = 0$

$5x_2 - 10 = 0$

$x_2 = 2$

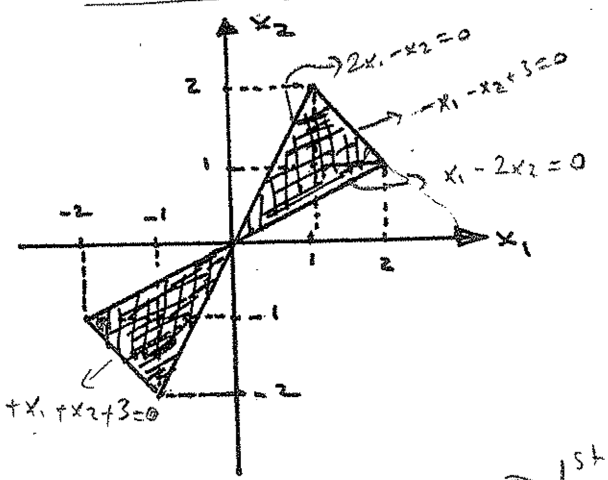
$x_1 = 2$

25/25

$-1(x_1-2) = x_2-1$

$-x_1-x_2+3=0$

Prob. 2 : (25 pt.s) Let x_1, x_2 denote two scalar inputs, and set $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbf{R}^2$. Basic neuron unit which has x as (unextended) input vector and a scalar output o is given by $o = f(w_1x_1 + w_2x_2 - \theta)$; here w_1, w_2 are scalar weights and θ is the scalar threshold value and $f(v)$ is the threshold function given as $f(v) = 1$ when $v \geq 0$ and $f(v) = 0$ when $v < 0$. Design a multi-layer neural network whose output o has the value $o = 1$ when x is in the shaded region, and $\boxed{o = 0 \text{ otherwise.}}$

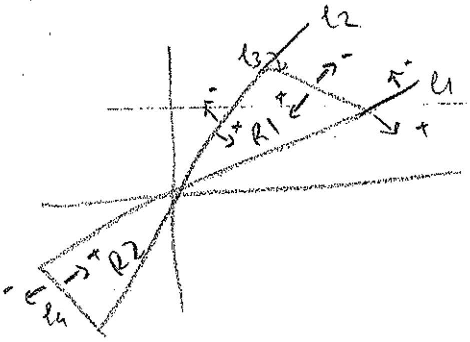


$o = f(v) = 1$ when $v \geq 0$

$o = f(v) = 0$ $\quad v < 0$

unipolar case

We should have 3 layer structure
- 1st → forming lines
- 2nd → forming regions
- 3rd → combine regions



l_1: $x_1 - 2x_2 = 0$ $(l1)$

l_2: $2x_1 - x_2 = 0$ $(l2)$

l_3: $-x_1 - x_2 + 3 = 0$ $(l3)$

l_4: $x_1 + x_2 + 3 = 0$ $(l4)$

First layer

l_1: $v_1 = w_{11}x_1 + w_{12}x_2 - \theta_1 = 0$, $\quad v_1 = x_1 - 2x_2$, $\quad o_1 = f(v_1)$

$\quad w_{11} = 1$, $w_{12} = -2$, $\theta_1 = 0$

l_2: $v_2 = w_{21}x_1 + w_{22}x_2 - \theta_2 = 0$, $\quad v_2 = 2x_1 - x_2$, $\quad o_2 = f(v_2)$

$\quad w_{21} = 2$, $w_{22} = -1$, $\theta_2 = 0$

l_3: $v_3 = w_{31}x_1 + w_{32}x_2 - \theta_3 = 0$, $\quad v_3 = -x_1 - x_2 + 3 = 0$, $\quad o_3 = f(v_3)$

$\quad w_{31} = -1$, $w_{32} = -1$, $\theta_3 = -3$

l_4: $v_4 = w_{41}x_1 + w_{42}x_2 - \theta_4 = 0$, $\quad v_4 = x_1 + x_2 + 3 = 0$, $\quad o_4 = f(v_4)$

$\quad w_{41} = 1$, $w_{42} = 1$, $\theta_4 = -3$

So R1: $\bar{l_1} \wedge l_2 \wedge l_3$

R2: $l_1 \wedge \bar{l_2} \wedge l_4$

for AND operation we use unipolar case which is
$o = 1$ when $v \geq 0$ and $o = 0$ when $v < 0$

$$\theta = n - 1/2 = 3 - 1/2 = 2.5$$

2nd layer

$R1: \bar{l}_1 \wedge l_2 \wedge l_3 \Rightarrow v_5 = -o_1 + o_2 + o_3 - 2.5 \quad , \quad o_5 = f(v_5)$

$R2: l_1 \wedge \bar{l}_2 \wedge l_4 \Rightarrow v_6 = o_1 - o_2 + o_3 - 2.5 \quad , \quad o_6 = f(v_6)$

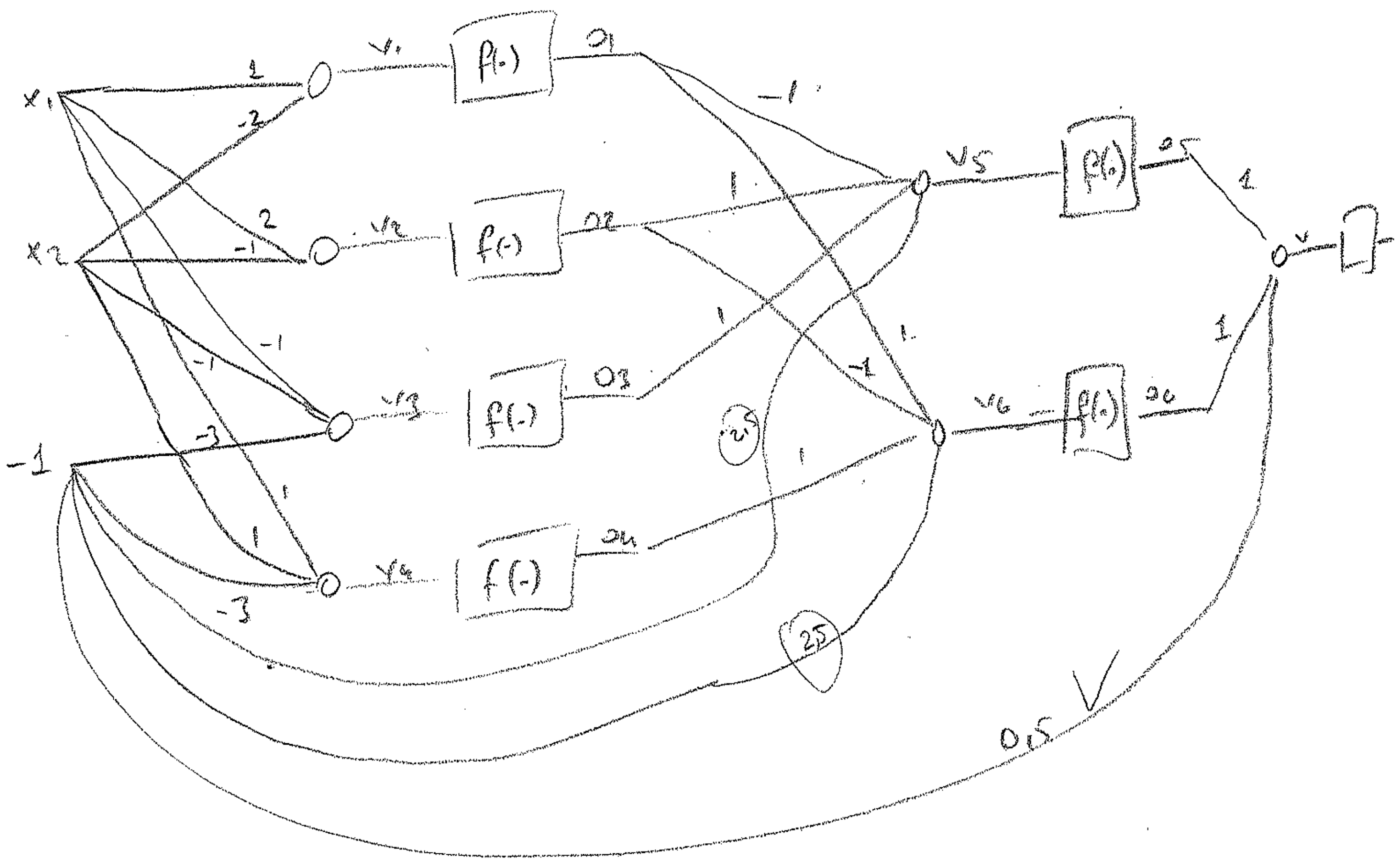
Third layer

$$o = \begin{cases} 1, & x \in R_1 \text{ or } R_2 \\ 0, & \text{ow} \end{cases}$$

$\theta = 0.5$ for or operation unipolar case

Region 1 or Region 2 $\quad v = o_5 + o_6 - 0.5 \quad , \quad o = f(v).$

so if x in shaded region, $o = 1$ ow $o = 0$.



30/30

Prob. 3 : (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 $o = sgn(w_1x_1 + w_2x_2 - \theta)$; 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 input pattern vectors as :

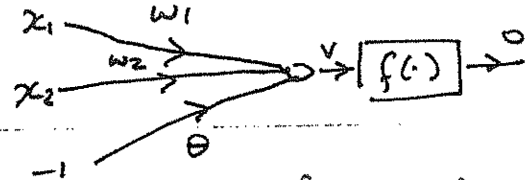
$$z_1 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad z_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad z_3 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad z_4 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Consider the classes C_1 , C_2, where each class contains only two input patterns; the desired outputs are $d = 1$ for C_1 and $d = -1$ for C_2.

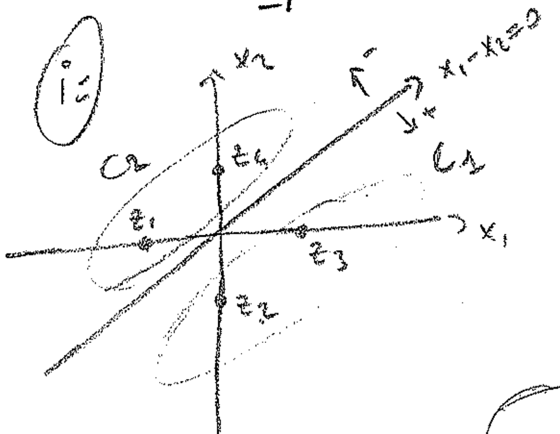
i : Give one example of C_1 , C_2 where perceptron learning algorithm converges and another example of C_1 , C_2 where perceptron learning algorithm does not converge. Briefly justify your choice. (Don't try to prove, you may justify your answer by referring to known results)

ii : Consider the case $C_1 = \{z_1,\ z_2\}$, $C_2 = \{z_3,\ z_4\}$; the desired outputs are $d = 1$ for C_1 and $d = -1$ for C_2. Show that for any $\theta \geq 0$, if we choose $w_1 \leq -\theta$, and $w_2 \leq -\theta$, the neuron with the selected weights solves the given classification problem.

iii : Consider the case $C_1 = \{z_1,\ z_2\}$, $C_2 = \{z_3,\ z_4\}$; the desired outputs are $d = 1$ for C_1 and $d = -1$ for C_2. Consider the perceptron training algorithm with initial weights $w_1(0) = 1$, $w_2(0) = 1$, $\theta(0) = 1$, and use the learning coefficient $c = 1$. By using the perceptron training algorithm find the last updated weights (including threshold) at the end of 2nd epoch. Does the final weights (including threshold) solve the classification problem given above?

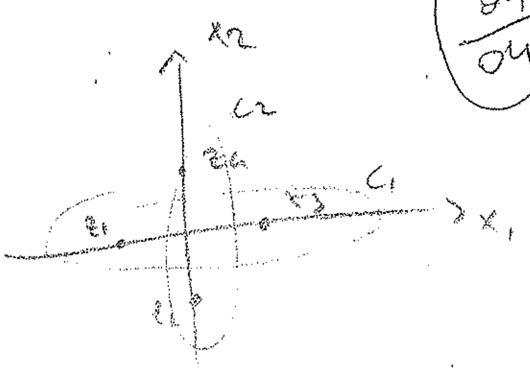


i:



for convergence taking classes should be linearly seperable. From the graph it is seen that $(z_2, z_3) \in C_1$ and $(z_1, z_4) \in C_2$. Taking these inputs like this we can seperate them by $x_1 - x_2 = 0$ line. So for C_1 desired output is 1, and for C_2 desired output is -1. Therefore perceptron algorithm converges.

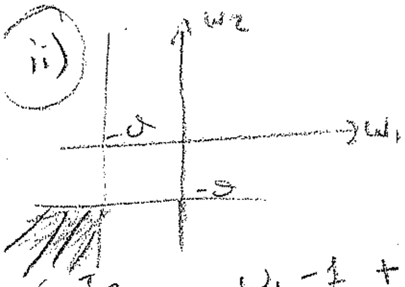
04/04



(2nd)

As seen from the figure, we cannot seperate these two classes linearly. So perceptron algorithm does not converge.

ii)



$C_1 = \{z_1, z_2\}$ $\begin{pmatrix}-1\\0\end{pmatrix}\begin{pmatrix}0\\-1\end{pmatrix}$

$C_2 = \{z_3, z_4\}$ $\begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}$

$w = \begin{bmatrix}w_1\\w_2\end{bmatrix}$

since for z_1, z_2 $d=1$

10/10

$C_1 \begin{cases} w^T z_1 = w_1 \cdot -1 + w_2 \cdot 0 - \theta \geq 0 \Rightarrow -\theta \geq w_1 \\ w^T z_2 = w_1 \cdot 0 + w_2 (-1) - \theta \geq 0 \Rightarrow -\theta \geq w_2 \end{cases}$

$C_2 \begin{cases} w^T z_3 = w_1 \cdot 1 + w_2 \cdot 0 - \theta < 0 \Rightarrow w_1 < \theta \\ w^T z_4 = w_1 0 + w_2 1 - \theta < 0 \Rightarrow w_2 < \theta \end{cases}$

for any $\theta \geq 0$ we have $w_1 \leq -\theta$, $w_2 \leq -\theta$

this is satisfied with 4 equations given above.

since w_2 w_1 $-\theta$ 0 θ OK!

So selecting $w_1 \leq -\theta$, and $w_2 \leq -\theta$, ~~the~~ problem is solved

iii) $w_1(0) = 1$, $w_2(0) = 1$, $\theta(0) = 1$, $c = 1$. (extended weight vector and input is used)

$\underline{w}_e(0) = \begin{pmatrix}1\\1\\1\end{pmatrix}$ $z_{1e} = \begin{pmatrix}-1\\0\\-1\end{pmatrix}$ $z_{2e} = \begin{pmatrix}0\\-1\\-1\end{pmatrix}$ $z_{3e} = \begin{pmatrix}1\\0\\-1\end{pmatrix}$, $z_{4e} = \begin{pmatrix}0\\1\\-1\end{pmatrix}$, $w(i+1) = w(i) + 0.5\, c\,(d(i) - o(i))$

First epoch Take $z_{1e} \Rightarrow$ $v = w_e^T(0)\, z_{1e} = -1 + 0 - 1 = -2 < 0$, $o = -1$, $d = 1$, update is needed $w_e(1) = w_e(0) + \begin{pmatrix}-1\\0\\-1\end{pmatrix} = \begin{pmatrix}0\\1\\0\end{pmatrix}$

take $z_{2e} \to v = w_e^T(1) \cdot z_{2e} = 0.0 - 1 + 0 = -1 < 0$, $o = -1$, $d = 1$, $w_e(2) = w_e(1) + \begin{pmatrix}0\\-1\\-1\end{pmatrix} = \begin{pmatrix}0\\0\\-1\end{pmatrix}$

take $z_{3e} \to v = w_e^T(2) \cdot z_{3e} = 0 + 0 + 1 = 1 > 0$, $o = 1$, $d = -1$, $w_e(3) = w_e(2) - \begin{pmatrix}1\\0\\-1\end{pmatrix} = \begin{pmatrix}-1\\0\\0\end{pmatrix}$

take $z_{4e} \to v = w_e^T(3)\, z_{4e} = 0 \Rightarrow o = 1$, $d = -1$, $w_e(4) = w_e(3) - \begin{pmatrix}0\\1\\-1\end{pmatrix} = \begin{pmatrix}-1\\-1\\+1\end{pmatrix}$

Second epoch

Take $z_{1e} \to v = w_e^T(4) \cdot z_{e1} = +1 + 0 - 1 = 0 \Rightarrow o = 1$, $d = 1$ no update $w_e(5) = w_e(4)$

// $z_{2e} \to v = w_e^T(5)\, z_{e2} = 0 + 1 - 1 = 0 \Rightarrow o = 1$, $d = 1$ no update $w_e(6) = w_e(5)$

// $z_{3e} \to v = w_e^T(6)\, z_{e3} = -1 + 0 - 1 = -2 \Rightarrow o = -1$, $d = -1$ no update $w_e(7) = w_e(6)$

// $z_{4e} \to v = w_e^T(7)\, z_{e4} = 0 - 1 - 1 = -2 \Rightarrow o = -1$, $d = -1$ no update $w_e(8) = w_e(7)$

At the second epoch the weight vector is not changed, so the final weight vector $\begin{pmatrix}-1\\-1\\+1\end{pmatrix}$ solves the classification problem. In fact, the weight vector is in the dashed line in the figure shown in part (ii). with $\theta = 1$, $w_1 = -1 = -\theta = -1 = w_2$

OK

16/16

20/20

Prob. 4 : (20 pt.s) Consider the following neuron model. Here $v = w_1 x_1 + w_2 x_2 - \theta$, $o = f(v)$, and f is given in the figure shown below. Let the training set be given as

$$z_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad z_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

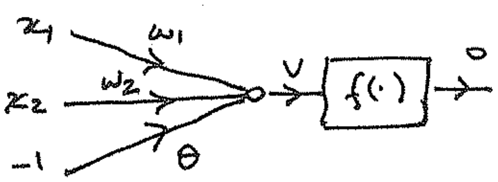
The desired outputs are $d = 1$ for z_1 and $d = -1$ for z_2. The cost function to be minimized is given as :

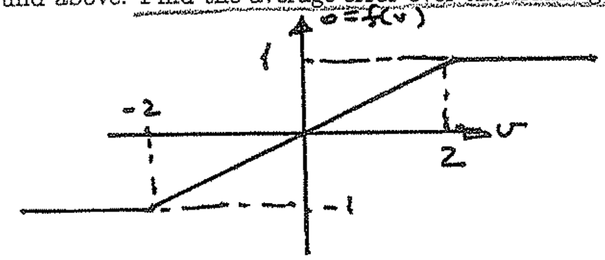
$$E = 0.5(d - o)^4$$

Note that this cost function is different from the standard quadratic cost function used in classical back propagation algorithm.

i : Let the initial weights be given as $w_1(0) = 1$, $w_2(0) = -1$, $\theta(0) = 1$, and the learning coefficient η as $\eta = 2$. Modify the back propagation algorithm (i.e. the gradient descent method) for the given cost function and the neuron model, update the weights (including threshold) for one epoch. determine the last updated weights (including threshold) at the end of one epoch.

ii : Now consider the weight vector found above. Find the average error over the training set.





$W_1(0) = 1$ $\quad \eta = 2$

$W_2(0) = -1$

$\theta(0) = 1$

$$f(v) = \begin{cases} 1 & v \geq 2 \\ \frac{v}{2} & -2 \leq v \leq 2 \\ -1 & -2 \geq v \end{cases}$$

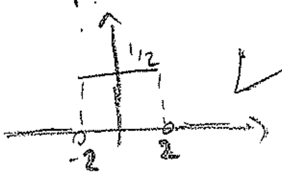
$E = 0.5(d-o)^4$, $e = (d-o) \Rightarrow E = 0.5 e^4$

07/07

$$\frac{\partial E}{\partial w_i} = \frac{\partial E}{\partial e} \frac{\partial e}{\partial o} \frac{\partial o}{\partial v} \frac{\partial v}{\partial w_i} = 2e^3 (-1) f'(v) x_i$$

$$W_i(k+1) = w_i(k) - \eta \frac{\partial E}{\partial w_i} = w_i(k) + 4e^3 f'(v) x_i$$

$f'(v)$



) take first pattern. $z_1 = \begin{pmatrix}1\\0\end{pmatrix}$ $z_2 = \begin{pmatrix}0\\1\end{pmatrix}$ $w_1(0)=1$ $w_2(0)=-1$ $\theta(0)=1$

$d=1$ $d=-1$

$V = w_1 z_1 + w_2 z_2 - \theta = 1.1 + 0.(-1) - 1 = 0 \Rightarrow o = f(v) = 0$ $d=1$

$e = 1-0 = 1$

$w_1(1) = w_1(0) + 4e^3 f'(v) z_1$

$= 1 + 4^2.1^3 \, 1/2 . 1 = 3$

$w_2(1) = w_2(0) + 4e^3 f'(v) z_2$

$= -1 + 4.1. \, 1/2 \, 0 = -1$

$\theta(1) = \theta(0) + 4e^3 f'(v) z_3$

$= 1 + 4^2.1 \, 1/2 .(-1) = -1$

take the second pattern

$V = w_1(1) z_1 + w_2(1) z_2 - \theta(1) = 3\,0 + (-1).1 - (-1) = 0 \Rightarrow o = 0$

$d=-1$

$e = -1-0 = -1$

08/09

$w_1(2) = w_1(1) + 4e^3 f'(v).z_1$

$= 3 + 4.(-1)^3 . \frac{1}{2} \, 0$

$= 3$

$w_2(2) = w_2(1) + 4e^3 f'(v) z_2$

$= -1 + 4.(-1)^3 \frac{1}{2}.1 = -3$

$\theta(2) = \theta(1) + 4e^3 f'(v) z_3 = -1 + 4.(-1)^3 . \frac{1}{2} .(-1) = 1$

the end of epoch 1 with final weights $w_1^* = 3$, $w_2^* = -3$, $\theta(2) = 1$

ii) $E_{ave} = \frac{1}{2} \sum_{i=1}^{2} E(n)$ where $E(n) = 0.5(d-o)^4 \Rightarrow e(n) = (d-o)$

1st pattern $\rightarrow$ $v = 3.1 + (-3).0 - 1 = 2 \Rightarrow f(2) = 1 = o(1)$

$d=1 \Rightarrow e(1) = 0 \Rightarrow E(1) = 0$

2nd pattern $\rightarrow$ $v = 3.0 + (-3).1 - 1 = -4 \Rightarrow f(-4) = -1 = o(2)$

$d = -1$

$\Rightarrow e(2) = 0 \Rightarrow E(2) = 0$

So $E_{ave} = \frac{1}{2} \sum_{i=1}^{2} E(n) = \frac{1}{2}(0+0) = \frac{0}{2}$

05/09