98/100

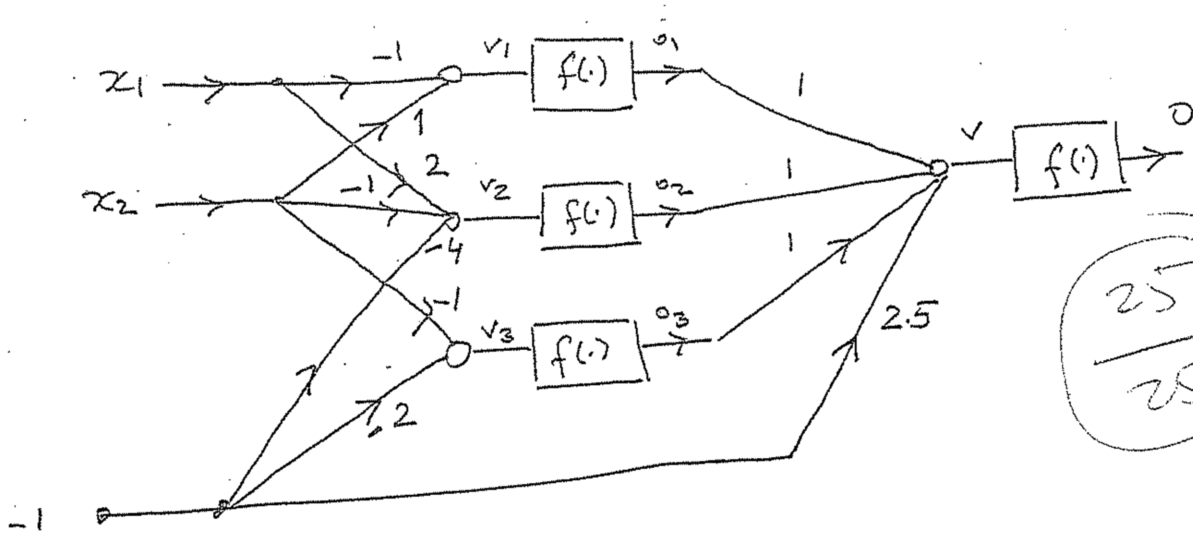
Ömer Morgül

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

EEE 443/543 Neural Networks
Midterm, Fall 2010

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 treshold function, i.e. $f(v) = 1$ when $v \geq 0$ and $f(v) = 0$ when $v < 0$. Find the region in $x_1 - x_2$ plane at which $o = 1$ (you may indicate this region geometrically). The weights are indicated on the network.



25/25

$o_1 = f(v_1)$, $v_1 = -x_1 + x_2$

$o_2 = f(v_2)$, $v_2 = 2x_1 - x_2 + 4$

$o_3 = f(v_3)$, $v_3 = -x_2 + 2$

$o = f(v)$, $v = o_1 + o_2 + o_3 - 2.5$

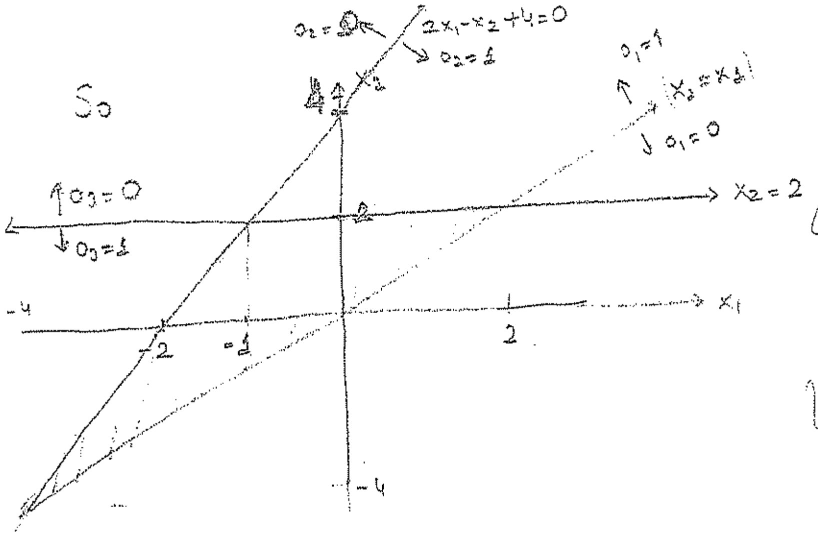
First we notice that $o = o_1 \wedge o_2 \wedge o_3$ because $v \geq 0$ only if $o_1 = o_2 = o_3 = 1$, so $o = 1$ if $o_1 = o_2 = o_3 = 1$, and it is zero otherwise,

hence $o = o_1 \wedge o_2 \wedge o_3$.

$$o_1 = \begin{cases} 1 & -x_1 + x_2 \geq 0 \Rightarrow x_2 \geq x_1 \\ 0 & \text{o.w.} \end{cases}, \quad o_2 = \begin{cases} 1 & 2x_1 - x_2 + 4 \geq 0 \\ 0 & \text{o.w} \end{cases}$$

$$o_3 = \begin{cases} 1 & -x_2 + 2 \geq 0 \Rightarrow x_2 \leq 2 \\ 0 & \text{o.w} \end{cases}$$

So



So in order to have $o = 1$, we need $o_1 = o_2 = o_3 = 1$

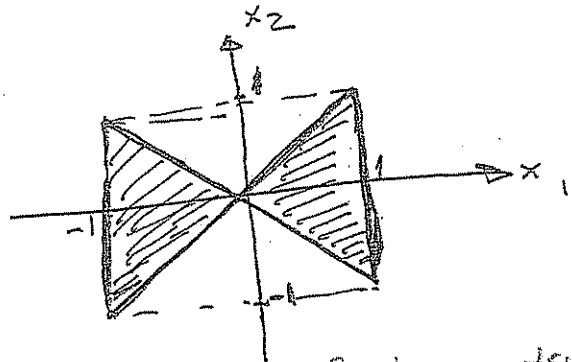
So in the dashed region which is a triangular with corners at $(2,2)$, $(-1,2)$, $(-4,-4)$, o_1 and o_2 and o_3 are 1 So output will be 1 in that region and it will be 0 at other regions.

Prob. 2 : (25 pt.s) Let x_1, x_2 denote two scalar inputs, and set $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$. Basic neuron unit which has x as (unextended) input vector and a 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 and θ is the scalar threshold value and sgn is the signum function given as

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

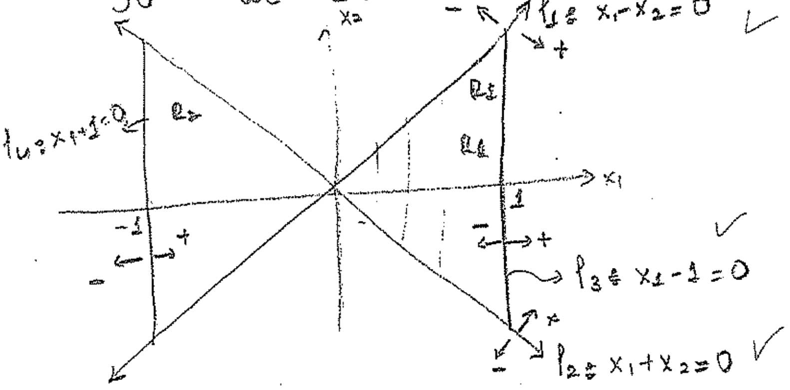
Design a multi-layer neural network whose output o has the value $o = 1$ when x is in the shaded region, and $o = -1$ otherwise.

25/25



1'st layer → forming lines
2nd layer → forming regions
3rd layer → combine "

So we should have a 3 layer structure.



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

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

First layer

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

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

l_3: $v_3 = w_{31} x_1 + w_{32} x_2 - \theta_3 = 0$, $v_3 = x_1 - 1$, so $o_3 = f(v_3) = sgn(v_3)$

$w_{31} = 1$, $w_{32} = 0$, $\theta_3 = 1$

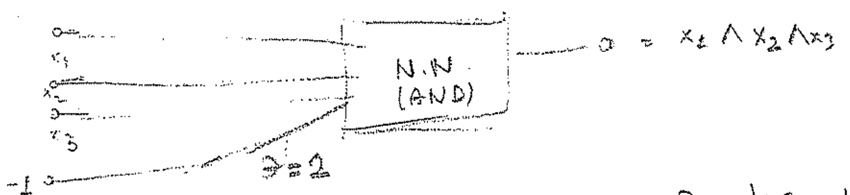
l_4: $v_4 = w_{41} x_1 + w_{42} x_2 - \theta_4 = 0$, $v_4 = x_1 + 1$, so $o_4 = f(v_4) = sgn(v_4)$

$w_{41} = 1$, $w_{42} = 0$, $\theta_4 = -1$

$R1: l_1 \wedge l_2 \wedge \overline{l_3}$ ✓

$R2: \overline{l_1} \wedge \overline{l_2} \wedge l_4$ ✓

So we know that for bipolar case,
AND operation threshold $\theta = n-1$ if there are n inputs
(choose $\theta = 3-1=2$. if $n=3$)



$$v = x_1 + x_2 + x_3 - 2$$
$$o = x_1 \wedge x_2 \wedge x_3$$

[Bipolar OR : $v = x_1 + x_2 + 1$, $o = x_1 \vee x_2$, Bipolar NOT $v = -x_1$, $o = \bar{x}_1$]

R1: $l_1 \wedge l_2 \wedge \bar{l}_3 \Rightarrow v_5 = o_1 + o_2 - o_3 - 2$, $o_5 = f(v_5) = sgn(v_5)$

R2: $\bar{l}_1 \wedge \bar{l}_2 \wedge l_4 \Rightarrow v_6 = -o_1 - o_2 + o_4 - 2$, $o_6 = f(v_6) = sgn(v_6)$

econd layer

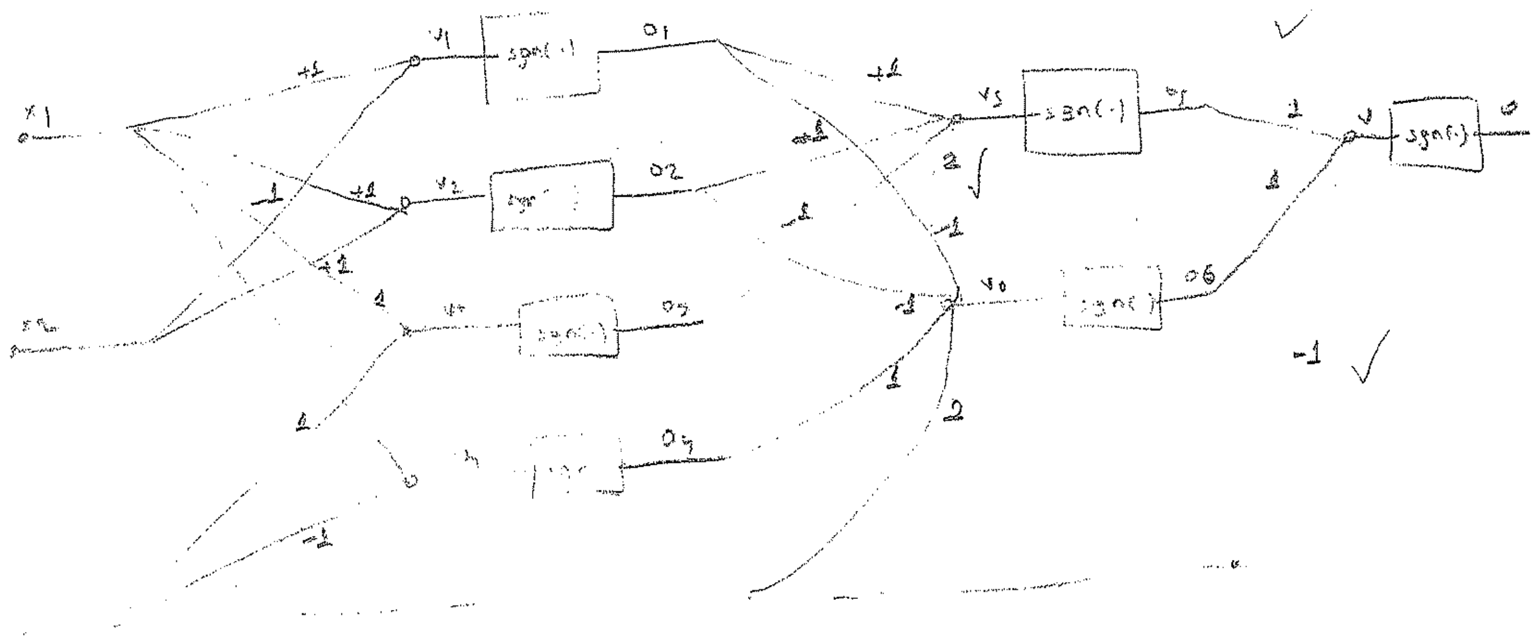
Finally $o = \begin{cases} 1, & x_1 \in R_1 \text{ or } R_2 \\ -1, & o.w \end{cases}$

OK

hird layer

Hence Region 1 or Region 2: $\Rightarrow V = o_5 + o_6 + 1$, $o = f(v) = sgn(v)$

So if x in shaded region, $\Rightarrow o = 1$, if not. $o = -1$



28/30

Prob. 3 : (30 pt.s) Let x_1, x_2 denote two scalar inputs, and set $z = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$. Consider the single neuron given below; as usual the scalar output o is given by $o = sgn(w_1 x_1 + w_2 x_2)$; here w_1, w_2 are scalar weights, threshold is assumed to be zero, and sgn is the signum function given as

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

Let us define the input pattern vectors as :

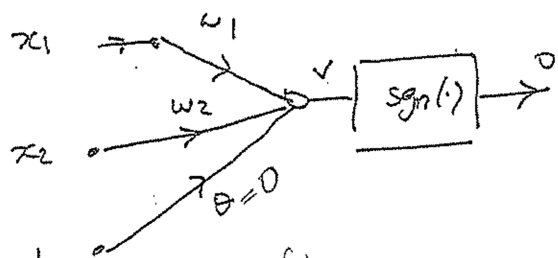
$$z_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad z_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad z_3 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad z_4 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}.$$

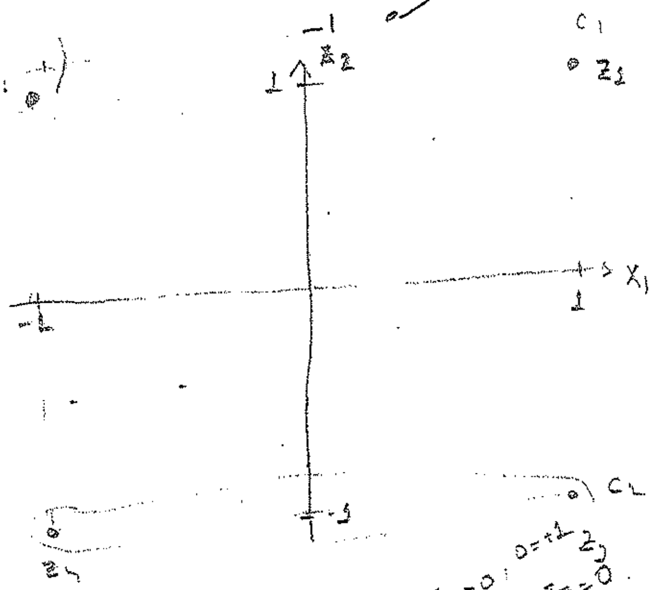
Consider the classes $C_1 = \{z_1,\ z_2\}$, $C_2 = \{z_3,\ z_4\}$; the desired outputs are $d = 1$ for C_1 and $d = -1$ for C_2.

i : Can we find a weights w_1 and w_2 such that the neuron given above classifies the indicated classes correctly ? (Don't try to prove your result, you may justify your answer by referring to known results)

ii : If the problem given above is solvable, find the set of all weights w_1 and w_2 which solve the problem given above. Indicate this set geometrically on $w_1 - w_2$ plane.

iii : Consider the perceptron training algorithm with initial weights $w_1(0) = 1$, $w_2(0) = 0$, and use the learning coefficient $c = 1$. By using the perceptron training algorithm find the last updated weights at the end of 3rd epoch. Does the final weights solve the classification problem given above?





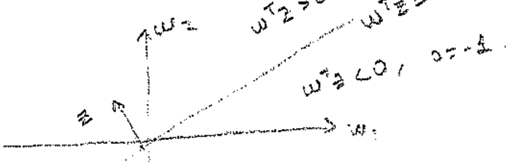
$C_1 = \{z_1, z_2\}$, $C_2 = \{z_3, z_4\}$

$d=1$ $\quad$ $d=-1$

passing this output?

i) Yes we can perform classification operation with this single neuron as two classes i.e C_1 and C_2 are linearly seperable Hence we can find w_1 and w_2 to classify input patterns

OK



ii) So

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

So we should have $\quad w^T z_1 \geq 0$ $\quad w^T z_2 \geq 0$ and

$w^T z_3 < 0$, $\quad w^T z_4 < 0$

$w^T z_1 = w_1 + w_2 \geq 0$, $w^T z_2 = -w_1 + w_2 \geq 0$,

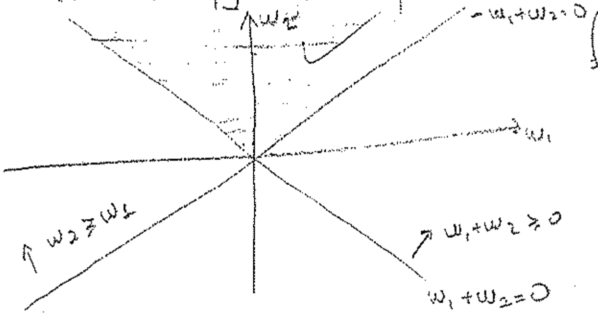
$w^T z_3 = w_1 - w_2 \leq 0$, $w^T z_4 = -w_1 - w_2 \leq 0$

$\Big\}$ $w_1 + w_2 \geq 0$, $-w_1 + w_2 \geq 0$

In fact we have only 2 constraints instead of 4.

So our weight vector should satisfy above constraints to classify given patterns correctly

5/5



So if my given weight vector $\begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ is in the dashed region, then problem is solved.!

iii) $w_1(0) = 1$, $w_2(0) = 0$, $c = 1$ (Initially wrong weight vectors)

$\underline{w}(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\underline{z} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ $\underline{z}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\underline{z}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$, $\underline{z}_3 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ $\underline{z}_4 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$

First epoch / Take $\underline{z}_1 \rightarrow$ $v = \underline{w}(0)^T \cdot \underline{z}_1 = 1 > 0$, $o = 1$, $d = 1$ $\rightarrow$ no update, $\underline{w}(1) = \underline{w}$

" $\underline{z}_2 \rightarrow$ $v = \underline{w}(1)^T \underline{z}_2 = -1 < 0$ $o = -1$, $d = 1$ $\rightarrow$ $w(2) = w(1) + \underline{z}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

" $\underline{z}_3 \rightarrow$ $v = \underline{w}(2)^T \cdot \underline{z}_3$ $-1 < 0$. $o = -1$, $d = -1$ $\rightarrow$ $w(3) = w(2)$.

" $\underline{z}_4 \rightarrow$ $v = \underline{w}(3)^T \cdot \underline{z}_4 = -1 < 0$, $o = -1$, $d = -1$ $\rightarrow$ $\underline{w}(4) = \underline{w}(3)$

Second epoch / " $z_1 \rightarrow$ $v = w(4)^T z_1 = 1 \geq 0$, $o = 1$, $d = 1$, $\Rightarrow$ $\underline{w}(5) = \underline{w}(4)$

" $\underline{z}_2 \rightarrow$ $v = w(5)^T z_2 = 1 \geq 0$, $o = 1$, $d = 1$ $\Rightarrow$ $\underline{w}(6) = \underline{w}(5)$

" $z_3 \rightarrow$ $v = w(6)^T z_3 = -1 < 0$, $o = -1$, $d = -1$ $\Rightarrow$ $w(7) = w(6)$

" $z_4 \rightarrow$ $v = w(7)^T z_4 = -1 < 0$, $o = -1$, $d = -1$ $\Rightarrow$ $w(8) = w(7)$

As we notice there is no update at the end of second epoch

This means, there will be no update at the 3rd epoch too

similarly, $\underline{w}(9) = \underline{w}(8)$ $\underline{w}(10) = w(9)$, $w(11) = w(10)$ $w(12) = w(11)$.

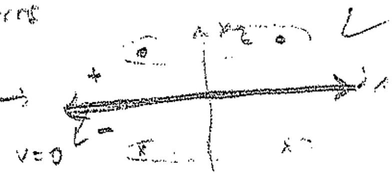
So final weight $\underline{w} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ $\Rightarrow$ $w_1^* = 0$, $w_2^* = 1$.

Inded this weight vector in the dashed region of part ii)

Also since there is no error in the last epoch (2nd and 3rd epochs indeed)

This weight vector correctly classifies all the patterns

$v = x_1 w_{11} + x_2 w_{12} = x_2$ so $o = \mathrm{sgn}(x_2)$ $\rightarrow$



15/15

20/20

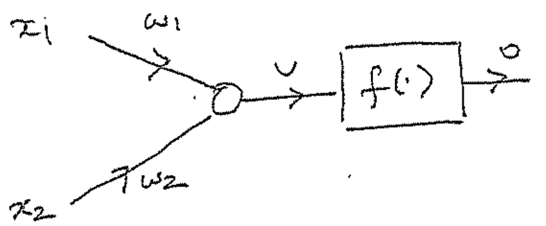
Prob. 4 : (20 pt.s) Consider the following neuron model. Here $v = w_1x_1 + w_2x_2$, $o = f(v)$, and f is given in the figure shown below; threshold is assumed to be zero. (Note that this function may be considered as a rough approximation of bipolar sigmoidal function). Let the training set be given as in Problem 3, namely let us define the input pattern vectors as :

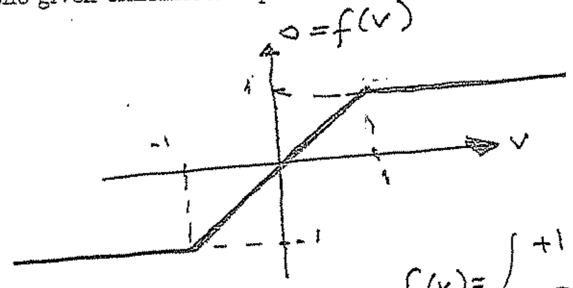
$$z_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad z_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad z_3 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad z_4 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}.$$

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 : Let the initial weights be given as $w_1(0) = w_2(0) = 0.25$, and the learning coefficient c as $c = 1$. By using the error back propagation algorithm modified for this neuron model find the last updated weights at the end of one epoch.

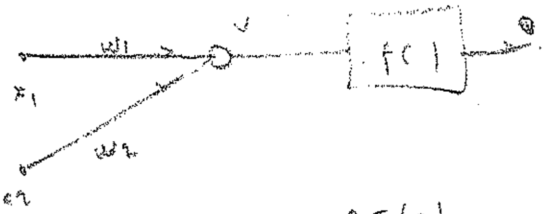
ii : Now consider the weight vector found above. Find the average error over the training set. Determine whether the neuron given above solve the given classification problem.





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

i)



$$w_i(k+1) = w_i(k) - \eta \cdot \frac{\partial E(n)}{\partial w_i}$$

i = input index, $i = 1, 2$.
k - iteration index
n = pattern index $n = 1, 2, 3, 4$.

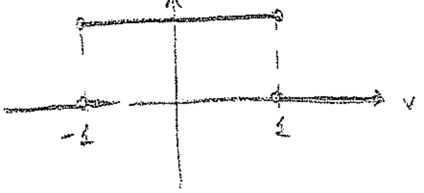
$$E(n) = \frac{1}{2} \cdot e(n)^2 \quad \text{where} \quad e(n) = (d(n) - o(n))$$

$$\frac{\partial E(n)}{\partial w_i} = \frac{\partial E(n)}{\partial e(n)} \cdot \frac{\partial e(n)}{\partial o} \cdot \frac{\partial o}{\partial v} \cdot \frac{\partial v}{\partial w_i}$$

$$= e(n) \cdot (-1) \cdot f'(v) \cdot x_i$$

Hence component-wise update $\Rightarrow$ $\boxed{w_i \longleftarrow w_i + \eta \cdot e(n) \cdot f'(v) \cdot x_i}$

$f'(v) =$



$$f'(v) = \begin{cases} 1, & -1 < v < 1 \\ 0 & \text{o.w} \end{cases}$$
(not defined in $v = \pm 1$)

$w_1(0) = 0.25$ $\quad z_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \; z_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \; z_3 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \; z_4 = \begin{pmatrix} -1 \\ -1 \end{pmatrix} \quad \eta = 1$

$w_2(0) = 0.25$

* Take 1st pattern

Forward prop. $\Rightarrow$ $v = w_1 z_1 + w_2 z_2 = 0.5$, $o = f(v) = 0.5$, $d = +1$

$e = 1 - 0.5 = \underline{\underline{0.5}}$

$w_1(1) = w_1(0) + 1 \cdot e \cdot \underbrace{f'(0.5)}_{1} \cdot 1$

$= 0.25 + 1 \cdot 0.5 \cdot 1 \cdot 1 = 0.75$

$w_2(1) = 0.25 + 1 \cdot 0.5 \cdot 1 \cdot 1 = 0.75$

* Take 2nd pattern:

Forward prop $\rightarrow$ $v = -0.75 + 0.75 = 0$ $\quad o = f(v) = 0$, $d = +1$

$e = 1 - 0 = \underline{1}$

$w_1(2) = w_1(1) + 1 \cdot e \cdot f'(0) \cdot (-1)$

$= 0.75 + 1 \cdot 1 \cdot 1 \cdot (-1) = -0.25$

$w_2(2) = 0.75 + 1 \cdot 1 \cdot 1 \cdot (+1) = 1.75$

* Take 3rd pattern

Forward prop $\longrightarrow$ $v = -0.25 - 1.75 = -2$ $\quad o = f(-2) = -1$, $d = -1$

$e = 0$

$w_1(3) = w_1(2)$

$w_2(3) = w_2(2)$.

* Take 4th pattern

Forward prop $\longrightarrow$ $v = +0.25 - 1.75 = -1.5$ $\quad o = f(-1.5) = -1$, $d = -1$

$e = 0$

$w_1(4) = w_1(3)$

$w_2(4) = w_2(3)$

So final weights are $\boxed{w_1^* = -0.25, \; w_2^* = 1.75}$

ii) $E_{ave} = \frac{1}{4} \sum_{n=1}^{4} E(n)$ where $E(n) = \frac{1}{2} \cdot e(n)^2 = \frac{1}{2} \left(d(n) - o(n) \right)^2$

1st pattern $\rightarrow$ $v = -0.25 \cdot 1 + 1.75 \cdot 1 = 1.5$, $f(1.5) = 1 = o(1)$, $d = 1$

$e(1) = d(1) - o(1) = 0$, $E(1)$

2nd pattern $\rightarrow$ $v = -0.25 \cdot (-1) + 1.75 \cdot 1 = 2$, $f(2) = 1 = o(2)$, $d = 1$

$e(2) = d(2) - o(2) = 0$, $E(2$

3rd pattern $\longrightarrow$ $v = -0.25 \cdot 1 - 1.75 \cdot = -2$, $o = f(-2) = -1$, $d = -1$

$e(3) = 0$, $E(3) = 0$

4th pattern $\rightarrow$ $v = +0.25 - 1.75 = -1.5$ $\quad o = f(-1.5) = -1$, $d = -1$

$e(4) = 0$, $E(4) = 0$

So $E_{ave} = \frac{1}{4} \cdot (E(1) + E(2) + E(3) + E(4))$

$= 0$. [illegible] solve the given classification problem.