1.

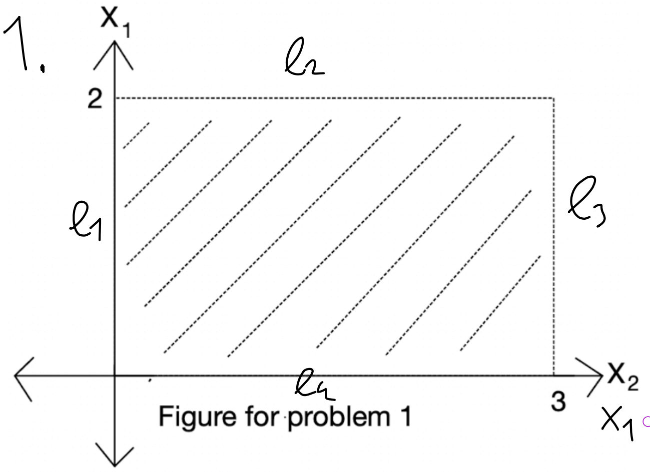


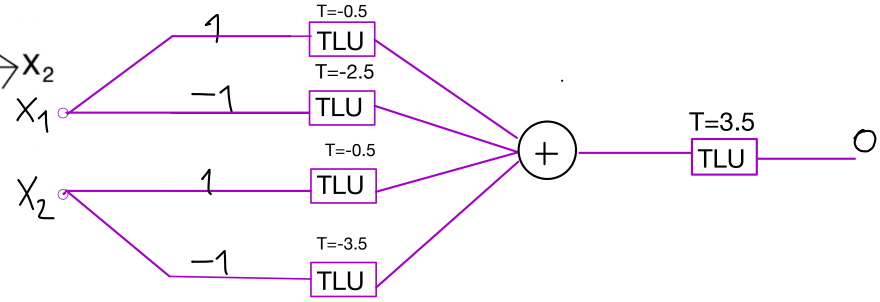
Figure for problem 1

$\ell_1: x_2 = 0 \rightarrow x_2 > 0$

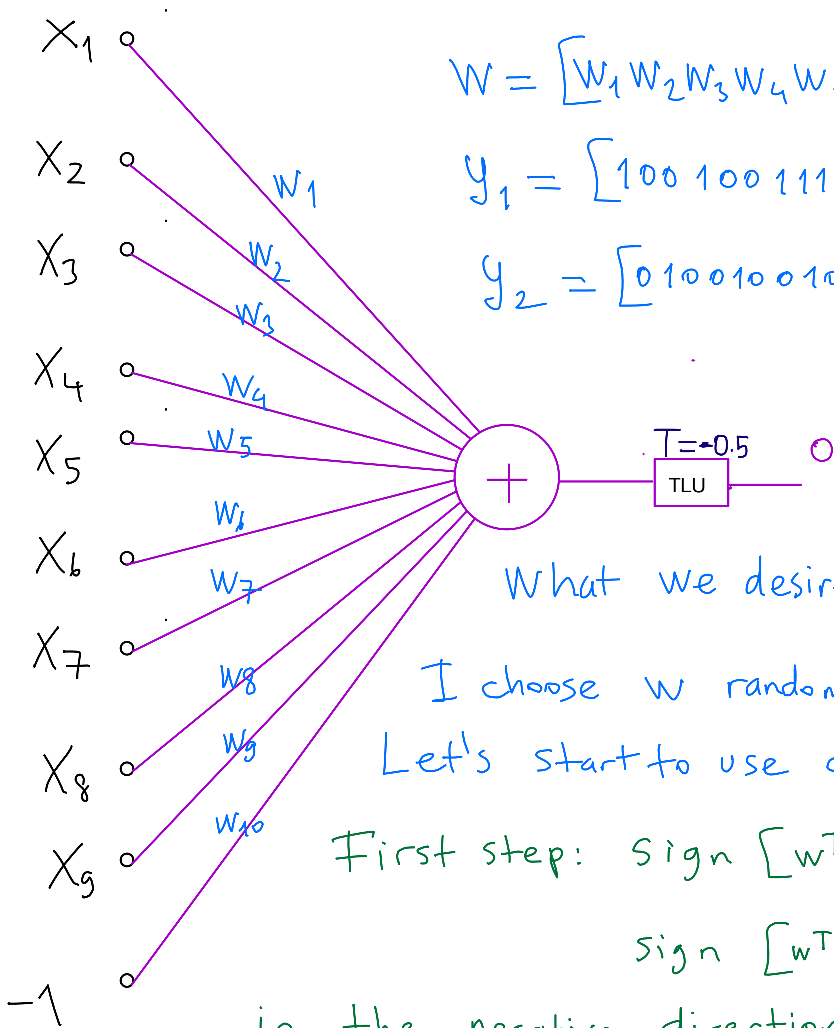
$\ell_2: x_1 = 2 \rightarrow x_1 < 2$

$\ell_3: x_2 = 3 \rightarrow x_2 < 3$

$\ell_4: x_1 = 0 \rightarrow x_1 > 0$



2.



$W = [w_1\, w_2\, w_3\, w_4\, w_5\, w_6\, w_7\, w_8\, w_9\, w_{10}]^T$

$y_1 = [1\,0\,0\,1\,0\,0\,1\,1\,1\,-1]$

$y_2 = [0\,1\,0\,0\,1\,0\,0\,1\,0\,-1]$

What we desire is $w^T . y_1 > 0$ and $w^T . y_2 < 0$

I choose w randomly as $w = [1\,1\,1\,1\,1\,1\,1\,1\,1\,1]$

Let's start to use algorithm. Take $c = 1$

First step: $\text{sign}[w^T . y_1] = 1$ ✓

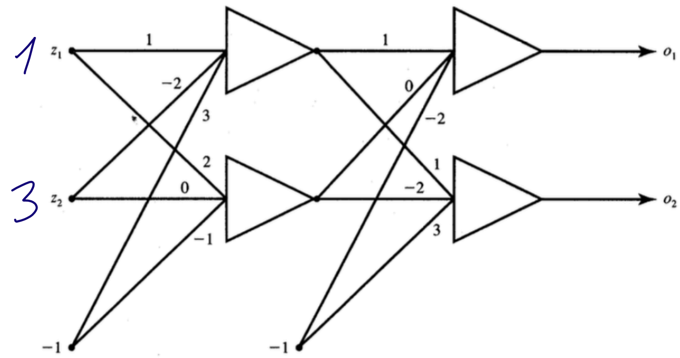
$\text{sign}[w^T y_2] = 1$ X we update w in the negative direction of y_2

$w_1^T = w_0^T - 1 . y_2 = [1\,0\,1\,1\,0\,1\,1\,0\,1\,2]$

Second step: $\text{sign}[w_1^T . y_1] = 1$ ✓ } no more update required

$\text{sign}[w_1^T . y_2] = -1$ ✓ } so we found $w = w_1$

3. i. $v = \begin{bmatrix} 1 & 2 \\ -2 & 0 \\ 3 & -1 \end{bmatrix}^T$, $w = \begin{bmatrix} 1 & 1 \\ 0 & -2 \\ -2 & 3 \end{bmatrix}^T$



ii. $net_{j_1} = -8$ $\quad$ $net_{j_2} = 3$

$y_1 = \dfrac{1}{1+e^{8}} \cong 3.35 \times 10^{-4}$ $\quad$ $y_2 = \dfrac{1}{1+e^{-3}} \cong 0.952$

$net_{k_1} = \dfrac{1}{1+e^{8}} + 2$ $\quad$ $net_{k_2} = \dfrac{1}{1+e^{8}} - \dfrac{2}{1+e^{-3}} - 3$

$\cong 2$ $\quad$ $\cong -4.9$

$o_1 = \dfrac{1}{1+e^{-2}}$ $\quad$ $o_2 = \dfrac{1}{1+e^{4.9}}$

$\cong 0.88$ $\quad$ $\cong 7.39 \times 10^{-3}$

iii. $f'(net_{j_1}) = y_1(1-y_1) = 3.35 \times 10^{-4}$

$f'(net_{j_2}) = y_2(1-y_2) = 4.51 \times 10^{-2}$

$f'(net_{j}) = \begin{bmatrix} 3.35 \times 10^{-4} \\ 4.51 \times 10^{-2} \end{bmatrix}$

$f'(net_{k_1}) = o_1(1-o_1) = 1.04 \times 10^{-1}$

$f'(net_{k_2}) = o_2(1-o_2) = 7.33 \times 10^{-3}$

$f'(net_{k}) = \begin{bmatrix} 1.04 \times 10^{-1} \\ 7.33 \times 10^{-3} \end{bmatrix}$

eq 4.29

iv) $\delta_{o_1} = (d_1 - o_1)\, o_1(1-o_1) = (\overset{d_1}{0.95} - \overset{o_1}{0.88}) \cdot \overbrace{1.04 \times 10^{-1}}^{o_1 \cdot (1-o_1)} = 7.28 \times 10^{-3}$

$\delta_{o_2} = (d_2 - o_2)\, o_2(1-o_2) = (0.05 - 0.0073) \cdot 7.33 \times 10^{-3} = 3.12 \times 10^{-4}$

eq. 4.17

$\delta_{y_1} = 3.35 \times 10^{-4} \cdot [\delta_{o_1}\ \delta_{o_2}] \begin{bmatrix} w_{11} \\ w_{21} \end{bmatrix} = 2.54 \times 10^{-6}$

$\delta_{y_2} = 4.51 \times 10^{-2} \cdot [\delta_{o_1} \cdot \delta_{o_2}] \begin{bmatrix} w_{12} \\ w_{22} \end{bmatrix} = -2.81 \times 10^{-5}$

v) $\Delta V = \eta \cdot \delta_y \cdot z^t = 1 \cdot \begin{bmatrix} 2.54 \times 10^{-6} \\ -2.81 \times 10^{-5} \end{bmatrix} \cdot \begin{bmatrix} 1 & 3 & -1 \end{bmatrix}$

Take $\eta = 1$

$$= \begin{bmatrix} 2.54 \times 10^{-6} & 7.62 \times 10^{-6} & -2.54 \times 10^{-6} \\ -2.81 \times 10^{-5} & 8.43 \times 10^{-5} & 2.81 \times 10^{-5} \end{bmatrix}$$

$$\Delta W = \eta \cdot \delta_o \cdot y^t = 1 \cdot \begin{bmatrix} 7.28 \times 10^{-3} \\ 3.12 \times 10^{-4} \end{bmatrix} \cdot \begin{bmatrix} 3.35 \times 10^{-4} & 9{,}52 \times 10^{-1} & -1 \end{bmatrix}$$

$$= \begin{bmatrix} 2.43 \times 10^{-6} & 6.93 \times 10^{-3} & 7.28 \times 10^{-3} \\ 1.04 \times 10^{-6} & 2.97 \times 10^{-4} & 3.12 \times 10^{-4} \end{bmatrix}$$

4 i. $x_1^1 = \begin{bmatrix} 0.707 \\ 0.707 \end{bmatrix}, \ x_2 = \begin{bmatrix} -0.707 \\ -0.707 \end{bmatrix}, \ x_3 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \ x_4 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$

$\hat{w}_1^{0T} \cdot x_1 = 0.707$ $\quad w_1^1 = w_1^0 + 0.5 \begin{bmatrix} -0.293 \\ 0.707 \end{bmatrix} = \begin{bmatrix} 0.853 \\ +0.353 \end{bmatrix}$

$\hat{w}_2^{0T} \cdot x_1 = -0.707$ $\quad \rightarrow \quad \hat{w}_2^1 = \hat{w}_2^0$

w_1^1 is needed to be normalized $\quad w_1^1 = \begin{bmatrix} 0.924 \\ +0.382 \end{bmatrix}$

$\hat{w}_1^{1T} \cdot x_2 = +0.383$ $\quad \rightarrow \quad \hat{w}_1^2 = \hat{w}_1^1$

$\hat{w}_2^{1T} \cdot x_2 = 0.707$ $\quad w_2^2 = \hat{w}_1^1 + 0.5 \cdot \begin{bmatrix} -0.707 \\ 0.293 \end{bmatrix} = \begin{bmatrix} -0.353 \\ -0.853 \end{bmatrix}$

$\hat{w}_2^2$ is needed to be normalized $\quad \hat{w}_2^2 = \begin{bmatrix} -0.382 \\ -0.924 \end{bmatrix}$

$\hat{w}_1^{2T} x_3 = +0.382$ $\quad \rightarrow \quad \hat{w}_1^3 = \hat{w}_1^2 + 0.5 \cdot \begin{bmatrix} -0.924 \\ 0.618 \end{bmatrix} = \begin{bmatrix} 0.462 \\ 0.691 \end{bmatrix}$

$\hat{w}_2^{2T} x_3 = -0.924$ $\quad \hat{w}_2^3 = w_2^2$

$\hat{w}_1^3$ is needed to be normalized $\quad \hat{w}_1^3 = \begin{bmatrix} 0.555 \\ 0.831 \end{bmatrix}$

$\hat{w}_1^{3T} x_4 = -0.555$ $\quad \hat{w}_1^4 = \hat{w}_1^3$

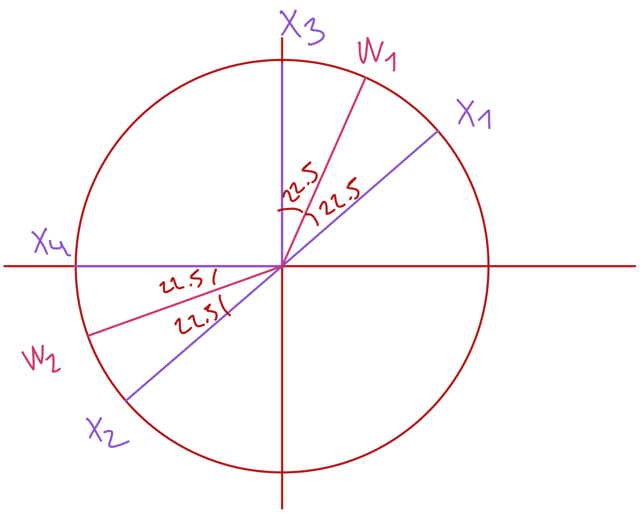
$\hat{w}_2^{3T} . x_4 = 0.353 \rightarrow \hat{w}_2^4 = \hat{w}_2^3 + 0.5 \begin{bmatrix} -0.618 \\ 0.924 \end{bmatrix} = \begin{bmatrix} -0.691 \\ -0.462 \end{bmatrix}$

$\hat{w}_2^4$ is needed to be normalized $\quad \hat{w}_2^4 = \begin{bmatrix} -0.831 \\ -0.555 \end{bmatrix}$

Here we completed the first cycle of training.

After the first cycle of training, we cannot say that $\hat{w}_1^4$ and $\hat{w}_2^4$ are the final weights. Final weights can only be found by matlab. However we can guess where it should end up

Here is our 4 vectors.



x_1 and x_3 will belong to cluster W_1

x_2 and x_4 will belong to cluster W_2

for now $\quad \hat{w}_1^4 = 1 \angle 56°$

$\hat{w}_2^4 = 1 \angle -213°$

$\hat{w}_2^4$ is close to its end value

$\hat{w}_1^4$ is a bit away from its end value

$\hat{w}_1^f$ will be $\begin{bmatrix} 0.383 \\ 0.924 \end{bmatrix}$

$\hat{w}_2^f$ will be $\begin{bmatrix} -0.924 \\ -0.383 \end{bmatrix}$

ii) As we found clusters that x_1, x_2, x_3, x_4 belong to, we can design a discrete perceptron with thresholds.

Since $\quad \hat{w}_1^{fT} . x_1 > w_2^{fT} x_1 \quad$ and $\quad w_2^{fT} x_2 > w_1^{fT} x_2$

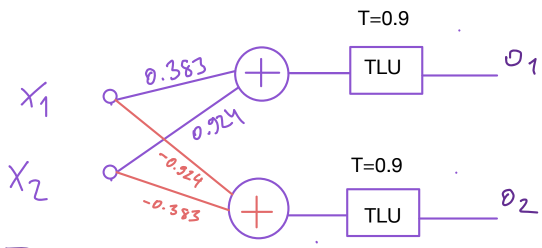
$\hat{w}_1^{fT} . x_3 > w_2^{fT} . x_3 \qquad\qquad w_2^{fT} x_4 > w_2^{fT} x_4$

It is proper to calculate $\hat{w}_1^{fT} x_1$ and $w_1^{fT} x_3$. Using this value as a threshold will help us distinguish x_1, x_3 from x_2, x_4

To distinguish x_2, x_4 from $x_1 x_4$, we can use $w_2^{f^T} x_2$ and $w_2^{f^T} x_4$ as threshold values. Our design becomes

$$w_1^{f^T} \cdot x_1 = w_1^{f^T} \cdot x_3 = 0.924$$

$$w_2^{f^T} \cdot x_2 = w_2^{f^T} \cdot x_4 = 0.924$$



This design will accept vectors around $\begin{bmatrix} 0.383 \\ 0.924 \end{bmatrix}$, will result $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

and accept vectors around $\begin{bmatrix} -0.924 \\ -0.383 \end{bmatrix}$, will result $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$

Otherwise $\begin{bmatrix} o_1 \\ o_2 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$.