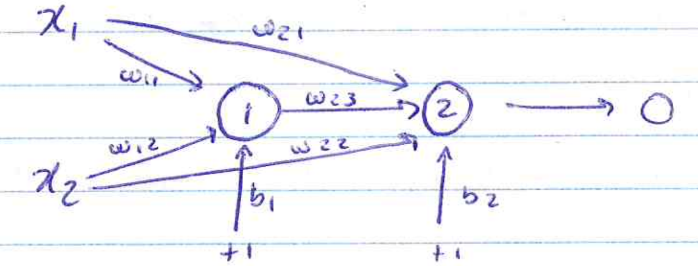
1.



$w_{11} = w_{12} = w_{21} = w_{22} = +1$; $w_{23} = -2$; $b_1 = -1.5$; $b_2 = -0.5$

1 a) Determine the separation planes.

First, we need to determine the expressions for the output of each neuron, using the step function as the activation function, φ:

$$\varphi(v_j) = \begin{cases} 0, & \text{if } v_j \leq 0 \\ 1, & \text{if } v_j > 1 \end{cases}$$

$$N_1 = \varphi\left(x_1 \overbrace{w_{11}}^{1} + x_2 \overbrace{w_{12}}^{1} + \overbrace{b_1}^{-1.5}\right) \Leftrightarrow$$

$$\Leftrightarrow N_1 = \varphi(x_1 + x_2 - 1.5)$$

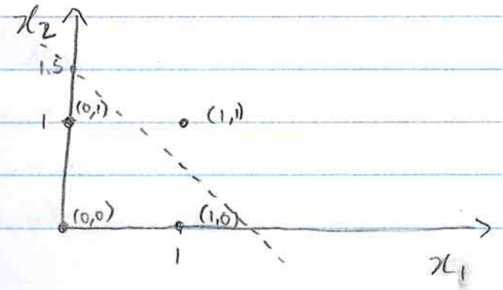
Since N_1 activation occurs at $v_1 = 0$, we can obtain the separation plane as follows:

$$v_1 = 0 \Leftrightarrow$$

$$\Leftrightarrow x_1 + x_2 - 1.5 = 0 \Leftrightarrow$$

$$\Leftrightarrow x_2 = -x_1 + 1.5$$

which can be plotted as follows (dashed line):



For the second neuron, $N_2 = 0$, we have the following expression:

$$N_2 = \varphi(x_1 w_{21} + x_2 w_{22} + N_1 w_{23} + b_2) \iff$$

$$\iff N_2 = \varphi(x_1 + x_2 - 2N_1 - 0.5)$$

Following the previous approach, and depending on the output of N_1, we can obtain the following planes:

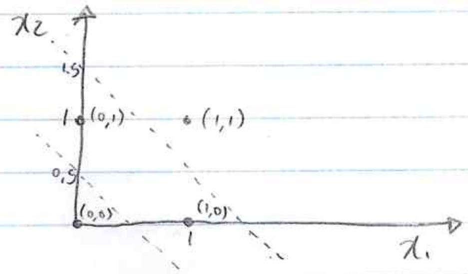
when $N_1 = 0$:

$$v_2 = 0 \iff$$
$$\iff x_1 + x_2 - 0.5 = 0 \iff$$
$$\iff x_2 = -x_1 + 0.5$$

when $N_1 = 1$:

$$v_2 \quad 0 \iff$$
$$\iff x_1 + x_2 - 2(\times 1) - 0.5 = 0 \iff$$
$$\iff x_1 + x_2 - 2.5 = 0 \iff$$
$$\iff x_2 = -x_1 + 2.5$$

However, since $x_1 + x_2$ can get a maximum value of 2 (when $x_1 = 1 \vee x_2 = 1$), the neuron can never be switched, meaning $x_1 + x_2 - 2.5 < 0$ for every scenario. For this reason, the representation of this last plane is redundant, so we will add only the first one to our plot:



We see that the lines separate the plane in 3 segments, with the first containing the (x_1, x_2) point $(0,0)$, the second containing $(0,1)$ and $(1,0)$, and the third containing $(1,1)$. To understand their classification, we will obtain the truth table in the next exercise

1 b) Obtain the truth table for the NN

We can obtain the truth table by replacing the x_1 and x_2 values by the different combinations of $\{0, 1\}$ and calculating the $\varphi(v_i)$ for each neuron

$$N_1 = \varphi(x_1 + x_2 - 1.5)$$
$$N_2 = \varphi(x_1 + x_2 - 2N - 0.5)$$

x_1	x_2	V_1	N_1	V_2	N_2
0	0	-1.5	0	-0.5	0
0	1	-0.5	0	0.5	1
1	0	-0.5	0	0.5	1
1	1	1	1	-0.5	0

The truth table is, therefore, a XOR:

x_1	x_2	O
0	0	0
0	1	1
1	0	1
1	1	0

1 c) Consider the activation function is a sigmoid with parameter $a = 1$. Calculate 1 iteration of the error backpropagation algorithm.

$$\text{sigmoid} = \frac{1}{1 + e^{(-a v_j)}}$$

since $a = 1$, $\varphi(v_j) = \dfrac{1}{1 + e^{-v_j}}$

We will calculate the error for when $x_1 = 0 \vee x_2 = 0$. First, we need to input these values in the N_2 expression:

$$N_2 = \varphi(x_1 + x_2 - 2\varphi(x_1 + x_2 - 1.5) - 0.5) \Leftrightarrow$$

$$\Leftrightarrow N_2 = \varphi(-2\varphi(-1.5) - 0.5) \Leftrightarrow$$

$$\Leftrightarrow N_2 = \varphi(-2 \times 0.1824 - 0.5) \Leftrightarrow$$

$$\Leftrightarrow N_2 = \varphi(-0.8648) \Leftrightarrow$$

$$\Leftrightarrow N_2 = 0.2963$$

Since $e_j(n) = d_j(n) - y_j(n)$

and, for this iteration n, $d_j(n) = 0$ and $y_j(n) = 0.2963$,

$$e_j(n) = 0 - 0.2963 = \boxed{-0.2963}$$

For the backpropagation step of w_{23}, we will recalculate the weight for an output neuron, with $d_2(n) = 0$

$$\Delta w_{23} = -\eta \frac{\varepsilon(n)}{w_{23}(n} = \eta\, \delta_2(n)\, y_1(n)$$

For an output neuron $y_j(n) = O_j(n)$

$$\delta_2(n) = a \times O_2(n)\,[1 - O_2(n)]\,[d_2(n) - O_2(n) \Leftrightarrow$$

$$\Leftrightarrow \delta_2(n) = 1 \times 0.2963(1 - 0.2963)(0 - 0.2963) \Leftrightarrow$$

$$\Leftrightarrow \delta_2(n) = -0.0618$$

since $y_1(n) = N_1 = 0.1824$, and assuming $\eta = 0.3$

$$w_{23}^{+} = w_{23} + \Delta w_{23} = -2 + (\eta\, \delta_2(n)\, y_1(n)) =$$

$$w_{23}^{+} = -2 + 0.3 \times (-0.0618) \times 0.1824 = -2.0034$$

where w_{23}^{+} is the recalculated weight.

w_{11}, w_{12}, w_{21}, and w_{22} do not need to be recalculated since, according to the formula,

$$\Delta w_{ji}(n) = \eta\, \delta_j(n)\, y_i(n)$$

and, for each of these weights, $y_i(n) = 0$ (since both x_1 and x_2 are zero), meaning $\Delta w_{ji}(n) = 0$

As for the biases, b_2 is in the same situation as w_{23}, meaning

$$\Delta b_2 = \eta\, \delta_2\, y_i(n)$$

since $\eta = 0.3$ and $y_i(n) = +1$,

$$b_2^+ = b_2 + \Delta b_2 = -0.5 + 0.3 \times (-0.0618) \times 1 = \cdot 0.5185$$

b_1 is directed to N_1, which is a hidden layer neuron, so

$$\delta_1(n) = a\, y_1(n)\,[1 - y_1(n)] \sum_K \delta_K(n)\, w_{Kj}(n)$$

since N_1 is only connected forward to N_2:

$$\delta_1(n) = 1 \times 0.1824\,(1 - 0.1824) \times \delta_2\, w_{23} =$$

$$= 0.1824\,(1 - 0.1824) \times (-0.0618) \times (-2) =$$

$$= 0.0184$$

$$b_1^+ = b_1 + \Delta b_1 = b_1 + \eta\, \delta_1\, y_i(n) = -1.5 + 0.3 \times 0.0184 \times 1 = -1.4945$$