Assignement#1

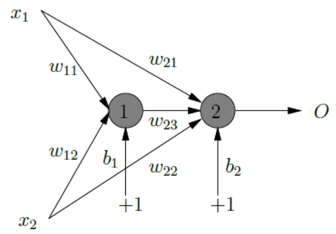
Ana Araújo, Nº 59457

Pedro Almeida, Nº58844

Backpropagation exercises

Problem 1 (adapted from Haykin, 2009 - problem 4.1)

Consider the following NN. Assume a McCulloch-Pitts model with a step function and inputs and output in $\{0, 1\}$.



Suppose weights are:

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

a) Determine the separation planes.

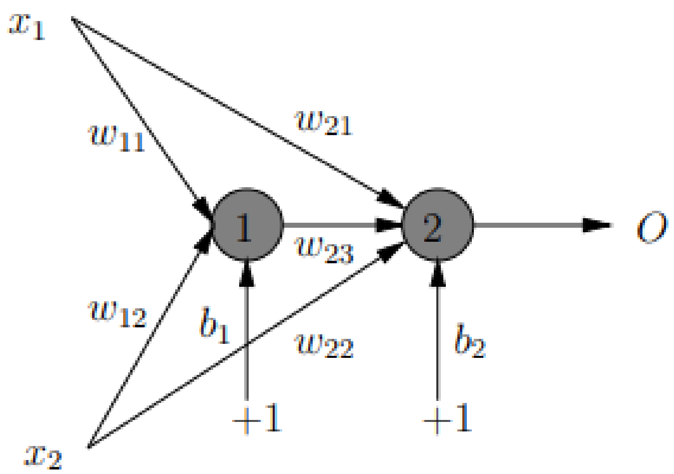
b) Obtain a truth table for the NN.

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

Problem 1 Resolution:

The McCulloch-Pitts model is a simple artificial neuron. The inputs and the outputs can be either a zero or a one. And the output as a zero or a one.

a)



From the figure above, we can observe that:

- x1 and x2 = i-th input of neuron j
- w11, w12, w21, w22, w23 = synapse weight: output of i to input of j
    - if negative: the input is inhibitory
    - if positive: the input is excitatory
- b1 and b2 = bias

Where the suppose weights are:

w11 = w12 = w21 = w22 = +1 (excitatory)

w23 = −2 (inhibitory)

b1 = −1.5

b2 = −0.5

To determine the separation planes, we need to find the decision boundaries between the two classes, where the values can be either 0 or 1.

From figure above, there are two neurons: 1 and 2.

For neuron 1:

$$y1 = \varphi\ (x_1 w_{11} + x_2 w_{12} + b_1)$$

For neuron 2:

$$y2 = \varphi\ (x_1 w_{21} + x_2 w_{22} + b_2 + w_{23} * y1)$$

Thus, the expression for the output:

$$y = \varphi\ \left(x_1 w_{21} + x_2 w_{22} + b_2 + w_{23} \times \varphi\ (x_1 w_{11} + x_2 w_{12} + b_1)\right) \Leftrightarrow$$

$$y = \varphi\ \left(x_1 + x_2 - 0.5 - 2 \times \varphi\ (x_1 + x_2 - 1{,}5)\right) \qquad (1)$$

Considering the input value of 0 for x1, we have:

$$y = \varphi\ \left(x_1 + x_2 - 0.5 - 2 \times \varphi\ (x_1 + x_2 - 1{,}5)\right) \Leftrightarrow$$

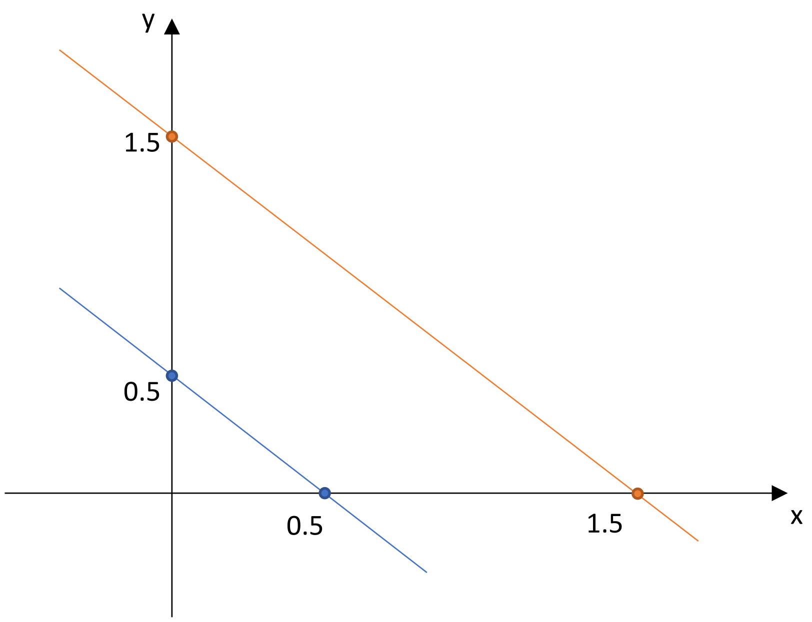
$$\Leftrightarrow y = \varphi\ \left(x_2 - 0.5 - 2\varphi\ (x_2 - 1.5)\right)$$

Considering the input value of 0 for x2, we have:

$$y = \varphi\ \left(x_1 + x_2 - 0.5 - 2 \times \varphi\ (x_1 + x_2 - 1{,}5)\right) \Leftrightarrow$$

$$\Leftrightarrow y = \varphi\ \left(x_1 - 0.5 - 2\varphi\ (x_1 - 1.5)\right)$$

The separation planes are a straight line in the input space with slope -0.5 and y-intercept 1 for neuron 1 and with slope -1.5 and y-intercept 1.5 for neuron 2:



The blue separation plane is equal to:

$y1 = -x + 0.5$

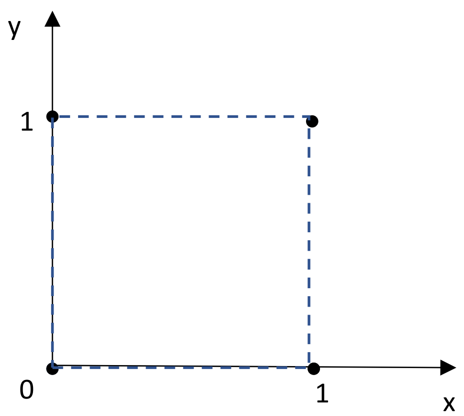
This means that if for the value of input x1 = 0, there is an increment of 0.5, the value of the input x1 changes from 0 to 1. Thus, if $y1 < -x + 0.5$, corresponds to output value 0 (y=0), thus classe 0, and if $y1 > -x + 0.5$, corresponds to output value 1 (y=1), thus classe 1.

The orange separation plane is equal to:

$y2 = -x + 1.5 \rightarrow$ this means that if for the value of input x2 = 0, there is an increment of 1.5, the value of the input x2 changes from 0 to 1. Thus, if $y2 < -x + 1.5$, corresponds to output value 0 (y=0), thus classe 0, and if $y2 > -x + 1.5$, corresponds to output value 1 (y=1), thus classe 1.

b)

The input values x_1 and x_2 can be either 0 or 1, as the out put value y can also be either 0 or 1, depending on the input values of x_1 and x_2:



Considering the possible input values for x_1 and x_2, we can obtain the following truth table:

- For x_1 = 0 and x_2 = 0:

  Neuron 1:

  $y1 = \varphi\ (x_1 w_{11} + x_2 w_{12} + b_1) \Leftrightarrow y1 = (0 \times 1 + 0\ \times 1 - 1.5) = 0$

  For neuron 2:

  $y2 = \varphi\ (x_1 w_{21} + x_2 w_{22} + b_2 + w_{23} * y1) \Leftrightarrow y2 = (0 \times 1 + 0\ \times 1 - 0.5 - 2\ \times 0) = 0$

  Thus the output for y is 0.

- For x_1 = 0 and x_2 = 1:

  Neuron 1:

  $y1 = \varphi\ (x_1 w_{11} + x_2 w_{12} + b_1) \Leftrightarrow y1 = (0 \times 1 + 1\ \times 1 - 1.5) = 0$

  For neuron 2:

  $y2 = \varphi\ (x_1 w_{21} + x_2 w_{22} + b_2 + w_{23} * y1) \Leftrightarrow y2 = (0 \times 1 + 1\ \times 1 - 0.5 - 2\ \times 0) = 1$

  Thus the output for y is 0.

- For x_1 = 1 and x_2 = 0:

  Neuron 1:

  $y1 = \varphi\ (x_1 w_{11} + x_2 w_{12} + b_1) \Leftrightarrow y1 = (1 \times 1 + 0\ \times 1 - 1.5) = 0$

  For neuron 2:

  $y2 = \varphi\ (x_1 w_{21} + x_2 w_{22} + b_2 + w_{23} * y1) \Leftrightarrow y2 = (1 \times 1 + 0\ \times 1 - 0.5 - 2\ \times 0) = 1$

  Thus the output for y is 1.

- For x_1 = 1 and x_2 = 1:

  Neuron 1:

  $y1 = \varphi\ (x_1 w_{11} + x_2 w_{12} + b_1) \Leftrightarrow y1 = (1 \times 1 + 1\ \times 1 - 1.5) = 1$

  For neuron 2:

  $y2 = \varphi\ (x_1 w_{21} + x_2 w_{22} + b_2 + w_{23} * y1) \Leftrightarrow y2 = (1 \times 1 + 1\ \times 1 - 0.5 - 2\ \times 1) = 0$

  Thus the output for y is 0.

x_1	x_2	y
0	0	0
0	1	1
1	0	1
1	1	0

c)

The activation function is a sigmoid with parameter α =1:

Sigmoid function:

$y_j = \frac{1}{1+e^{(-\alpha v_i)}}$ with α =1

Where:

y_j = neuron j output

v_i = induced local field (weighted sum of all inputs)

In the NN presented in the exercise, there are two activation functions (1 and 2).

For x_1 = 1 and x_2 = 0

$$v_1 = (w_{11} \times x_1 + w_{12} \times x_2 - b_1) = 1 + 0 - 1.5 = -0.5$$

$$v_2 = (w_{21} \times x_1 + w_{22} \times x_2 - b_1) = 1 + 0 - 1.5 = -0.5$$

For activation function 1 (sigmoid function):

$$y_1 = \frac{1}{1+e^{(-\alpha v_1)}} \Leftrightarrow y_1 = \frac{1}{1+e^{(-1\times(-0.5))}} \Leftrightarrow y_1 = 0.3775$$

For activation function 2 (sigmoid function):

$$y_2 = \frac{1}{1+e^{(-\alpha v_2)}} \Leftrightarrow y_2 = \frac{1}{1+e^{(-1\times(-0.5))}} \Leftrightarrow y_2 = 0.3775$$

Thus, for the sum in the second output:

$$v_3 = (w_{23} \times y_1 + w_{23} \times y_2 - b_2) \Leftrightarrow v_3 = (-2 \times 0.3775 + (-2) \times 0.3775 - 0.5 = 2.01$$

$$y_3 = \frac{1}{1+e^{(-\alpha v_3)}} \Leftrightarrow y_3 = \frac{1}{1+e^{(-1\times(-2.01))}} \Leftrightarrow y_3 = 0.8818 \text{ (predicted value)}$$

Calculate the error:

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

Where:

$e_j(n)$ error of neuron j

$y_j(n)$ real output

$d_j(n)$) desired output

Since we consider x_1 = 1 and x_2 = 0, the desired output is 1.

Thus, the error:

$e_j(n) = d_j(n) - y_j(n) \iff e_2(1) = d_2(1) - y_2(1) = 1 - 0.8818 = 0.1182$

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

$\varphi = \dfrac{1}{1 + e^{(-v\ \ )}}$

$\varphi' = \varphi - (1 - \varphi)$

Ponto: $x_1 = 0;\ x_2 = 1;\ y = 1$

Forward propagation:

Compute the weighted sum and activation of the input layer:

$v_1 = w_1 * x + b_1 = 0 + 1 - 1.5 = -0.5$

$a_1 = \varphi(v_1) = -1.5415$

Compute the weighted sum and activation of the output layer:

$v_2 = w_{23} * a_1 + b_2 = 1 – 0.5 + 3.083 = 3.583$

$a_2 = \varphi(v_2) = 0.973$

Compute the error of the output layer:

$d_2 = (a_2 - y) * \varphi'(v_2) = (0.973 – 1) * (3.582 * (1 – 3.582)) = 0.25$

Backward propagation:

Compute the error of the input layer:

$d_1 = (w_{23} * d_2) * \varphi'(v_1) = (-2 * 0.25) * (-0.5 * (1 – (-0.5)) = 0.375$

Compute the gradient of the weights and biases:

$dw_{23} = d_2 * a_1 = 0.25 * (-1.542) = -0.3855$

$db_2 = d_2 = 0.25$

$dw_1 = d_1 * x = [d_1 * x_1, d_1 * x_2] = [0, 0.375]$ # quero rever uma coisa

$db_1 = d_1 = 0.375$

Let the learning rate be:

$\alpha = 0.1$

Update the weights and biases:

$w_{23} = w_{23} - \alpha * dw_{23} = 1 + 0.0386 = 1.0386$

$b_2 = b_2 - \alpha * db_2 = -0.5 – (0.1 * 0.25) = -0.525$

$w_1 = w_1 - \alpha * dw_1 = [w_{x1}, w_{x2}] = [1, 0.9625]$

$w_{11} = w_{21} = 1$

$w_{12} = w_{22} = 0.9625$

$b_1 = b_1 - \alpha * db_1 = -1.5 - (0.1 * 0.375) = -1.5375$