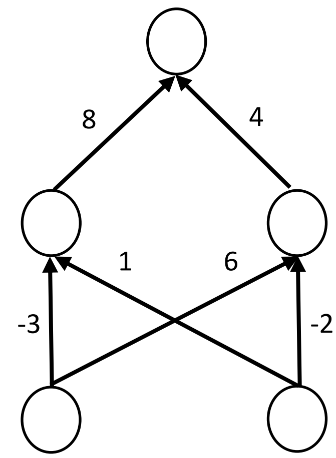
Given the neural network depicted in the image below, compute the approximate value of the weights after running the backpropagation algorithm in one iteration, using the input (2,6) with the target = 0.

Details about the neural network configuration:

- All activation functions are logistic (sigmoid),
- The Cost Function is Cross-entropy.
- The neurons don't have bias
- The learning rate is : $\eta = 0.5$
- Approximate the output of the sigmoid function using the graphic below.



Usefull formulas:

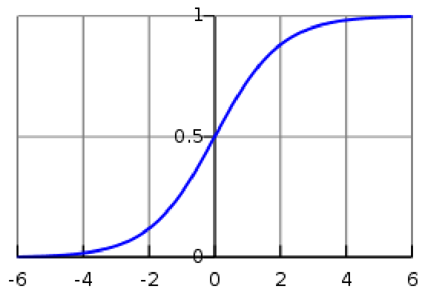
$$\delta_i^L = y_i^L - t_i^L$$

$$\delta_i^l = y_i^l(1 - y_i^l)\sum_k \delta_k^{l+1} \cdot w_{ik}^{l+1}$$

$$\frac{\partial C}{\partial w_{ij}^l} = \delta_j^l y_i^{l-1}$$

$$w_{ij}^l = w_{ij}^l - \eta * \frac{\partial C}{\partial w_{ij}^l}$$

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$



---

1. *Forward Pass:*

The neurons from the input layer do not have an activation function!

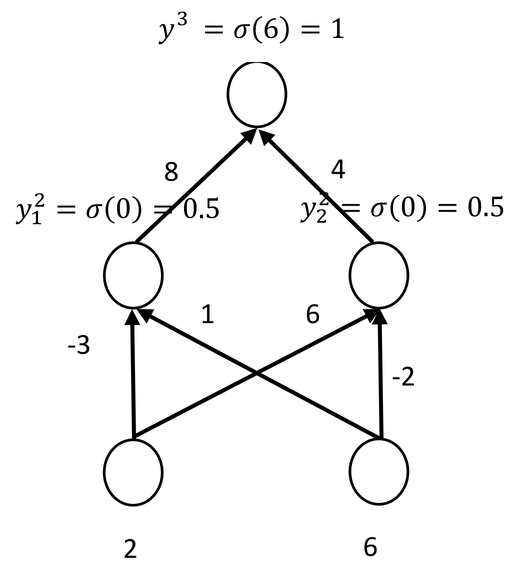
$z_1^2 = 2 * (-3) + 6 * 1 = 0;\ \ y_1^2 = \sigma(0) = 0.5$

$z_2^2 = 2 * (6) + 6 * (-2) = 0;\ \ y_2^2 = \sigma(0) = 0.5$

$z^3 = 8 * (0.5) + 4 * (0.5) = 6; y^3 = 1$

Compute the error in the output layer:

$\delta^3 = y^3 - t = 1 - 0 = 1$



2a. Backpropagate the error to the neurons in the hidden layer

$$\delta_1^2 = y_1^2 * (1 - y_1^2) * (\delta^3 * w_{11}^3) = 0.5 * (1 - 0.5) * (1 * 8) = 2$$

$$\delta_2^2 = y_2^2 * (1 - y_2^2) * (\delta^3 * w_{21}^3) = 0.5 * (1 - 0.5) * (1 * 4) = 1$$

2b. Adust the weights between the hidden and the output layer

$\frac{\partial C}{\partial w_{11}^3} = \delta^3 * y_1^2 = 1 * 0.5$=0.5. $\quad w_{11}^3 = 8 - 0.5 * 0.5 = 7.75$

$\frac{\partial C}{\partial w_{21}^3} = \delta^3 * y_2^2 = 1 * 0.5 = 0.5.\ w_{21}^3 = 4 - 0.5 * 0.5 = 3.75$

2c. Compute the gradients for the weihts between between the input and the hidden layer based on the previously backpropagated error (from 2a).

$$\frac{\partial C}{\partial w_{11}^2} = \delta_1^2 * y_1^1 = 2 * (2) = 4$$

$$\frac{\partial C}{\partial w_{12}^2} = \delta_1^2 * y_1^1 = 1 * (2) = 2$$

$$\frac{\partial C}{\partial w_{21}^2} = \delta_1^2 * y_2^1 = 2 * (6) = 12$$

$$\frac{\partial C}{\partial w_{22}^2} = \delta_2^2 * y_2^1 = 1 * (6) = 6$$

Based on the above computed gradients we compute the new weights

$w_{11}^2 = -3 - 0.5 * 4 = -5$ $\quad$ $w_{12}^2 = 6 - 2 * 0.5 = 5$

$w_{21}^2 = 1 - 0.5 * 12 = -5$ $\quad$ $w_{22}^2 = -2 - 0.5 * 6 = -5$