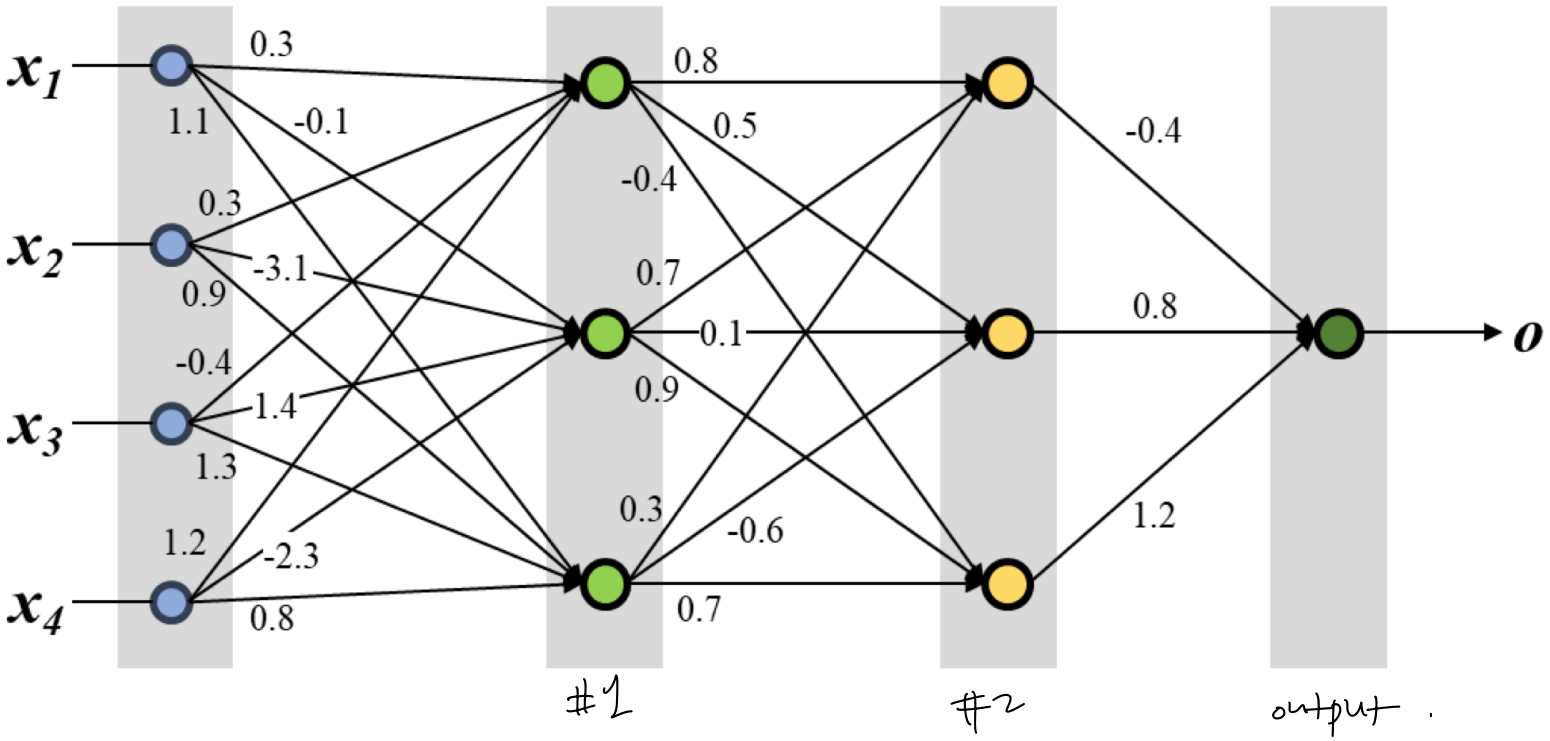
20181106 22 강지호 HW#1.

Problem 1. (2points)

Consider a MLP below that consist of one input layer, two hidden layers, and one output layer. Assume that the first and second hidden layers use a ReLU activation function and the output layer has a hyperbolic tangent activation function. Numbers represent weights associated with the corresponding perceptrons.



1) Write the weights matrices for the two hidden layers and the output layer.

hidden layer #1. $W_1 = \begin{bmatrix} 0.3 & 0.3 & -0.4 & 1.2 \\ -0.1 & -3.1 & 1.4 & -2.3 \\ 1.1 & 0.9 & 1.3 & 0.8 \end{bmatrix}$

hidden layer #2 $W_2 = \begin{bmatrix} 0.8 & 0.7 & 0.3 \\ 0.5 & 0.1 & -0.6 \\ -0.4 & 0.9 & 0.7 \end{bmatrix}$

output layer $W_0 = \begin{bmatrix} -0.4 & 0.8 & 1.2 \end{bmatrix}$

2) Given 3 inputs **X**:

$$\mathbf{X} = \begin{bmatrix} 1.2 & 0.3 & 0.9 \\ 0.7 & 0.1 & 0.2 \\ 1.5 & 2.9 & 1.4 \\ 2.3 & 1.4 & 0.9 \end{bmatrix},$$

Compute the output ***o***. Show the intermediate results.

$$\vec{O} = \tau(W_0 \sigma(W_2 \sigma(W_1 \vec{X})))$$

$$W_1\vec{X} = \begin{bmatrix} 0.3 & 0.3 & -0.4 & 1.2 \\ -0.1 & -3.1 & 1.4 & -2.3 \\ 1.1 & 0.9 & 1.3 & 0.8 \end{bmatrix} \cdot \begin{bmatrix} 1.2 & 0.3 & 0.9 \\ 0.7 & 0.1 & 0.2 \\ 1.5 & 2.9 & 1.4 \\ 2.3 & 1.4 & 0.9 \end{bmatrix} = \begin{bmatrix} 2.173 & 0.64 & 0.85 \\ -5.48 & 0.5 & -0.82 \\ 5.54 & 5.31 & 3.71 \end{bmatrix}$$

$$\therefore \sigma W_1\vec{X} = \begin{bmatrix} 2.173 & 0.64 & 0.85 \\ 0 & 0.5 & 0 \\ 5.54 & 5.31 & 3.71 \end{bmatrix}$$

$$W_2\sigma(W_1\vec{X}) = \begin{bmatrix} 0.8 & 0.7 & 0.3 \\ 0.5 & 0.1 & -0.6 \\ -0.4 & 0.9 & 0.7 \end{bmatrix} \begin{bmatrix} 2.173 & 0.64 & 0.85 \\ 0 & 0.5 & 0 \\ 5.54 & 5.31 & 3.71 \end{bmatrix} = \begin{bmatrix} 3.906 & 2.455 & 1.793 \\ -2.079 & -2.816 & -1.801 \\ 2.926 & 3.911 & 2.257 \end{bmatrix}$$

$$\therefore \sigma(W_2\sigma(W_1\vec{X})) = \begin{bmatrix} 3.906 & 2.455 & 1.793 \\ 0 & 0 & 0 \\ 2.926 & 3.911 & 2.257 \end{bmatrix}$$

$$W_0\sigma(W_2\sigma(W_1\vec{X})) = \begin{bmatrix} -0.4 & 0.8 & 1.2 \end{bmatrix} \begin{bmatrix} 3.906 & 2.455 & 1.793 \\ 0 & 0 & 0 \\ 2.926 & 3.911 & 2.257 \end{bmatrix}$$

$$= \begin{bmatrix} 1.9488 & 3.9112 & 1.9912 \end{bmatrix}$$

$$\therefore \vec{O} = \tau(\begin{bmatrix} 1.9488 & 3.9112 & 1.9912 \end{bmatrix}) = \begin{bmatrix} 0.9662 & 0.9988 & 0.9634 \end{bmatrix}$$

3) Replace the ReLU activation function in the first hidden layer with the following ELU (exponential linear unit) activation function:

$$\mathrm{ELU}(z) = \begin{cases} z & \geq 0 \\ e^z - 1 & < 0 \end{cases}.$$

Repeat 2). Show the intermediate results.

$$\vec{O} = T(w_0 \Gamma(w_2 E(w_1 \vec{X})))$$

$$E(w_1\vec{X}) = E\left(\begin{bmatrix} 2.73 & 0.64 & 0.85 \\ -5.48 & 0.5 & -0.82 \\ 5.94 & 5.31 & 3.71 \end{bmatrix}\right) = \begin{bmatrix} 2.73 & 0.64 & 0.85 \\ -0.9958 & 0.5 & -0.5596 \\ 5.94 & 5.31 & 3.71 \end{bmatrix}$$

$$\Gamma(w_2 E(w_1\vec{X})) = \Gamma\left(\begin{bmatrix} 3.2089 & 2.455 & 1.0013 \\ -2.1786 & -2.816 & -1.8570 \\ 2.0298 & 3.911 & 1.3534 \end{bmatrix}\right) = \begin{bmatrix} 3.2089 & 2.455 & 1.0013 \\ 0 & 0 & 0 \\ 2.0298 & 3.911 & 1.3534 \end{bmatrix}$$

$$\vec{O} = T([1.1522 \quad 3.9112 \quad 1.5436]) = [0.8185 \quad 0.9988 \quad 0.9129].$$