투빅스 15기 정규세션 6주차

Neural Network Basic Assignment 1

이름: 조준혁

1. Sigmoid Function을 z에 대해 미분하세요.

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

$$\frac{d}{dz}\sigma(z) = \frac{d}{dz}\left(\frac{1}{1+e^{-z}}\right) = \frac{d}{dz}(1+e^{-z})^{-1}$$

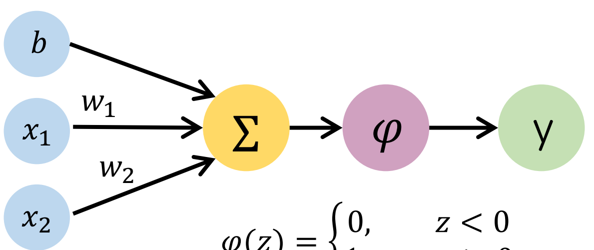
$$= -1\cdot(1+e^{-z})^{-2}\cdot\frac{d}{dz}(1+e^{-z})$$

$$= -(1+e^{-z})^{-2}\cdot(-e^{-z})$$

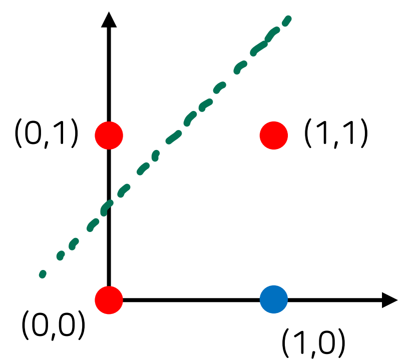
$$= \frac{e^{-z}}{(1+e^{-z})^2} = \frac{1+e^{-z}-1}{(1+e^{-z})^2} = \frac{1+e^{-z}}{(1+e^{-z})^2} - \frac{1}{(1+e^{-z})^2} = \frac{1}{1+e^{-z}} - \frac{1}{(1+e^{-z})^2} = \underbrace{\frac{1}{1+e^{-z}}}_{=\sigma(z)}\left(1-\frac{1}{1+e^{-z}}\right)$$

$$= \boxed{\sigma(z)(1-\sigma(z))}$$

2. 다음과 같은 구조의 Perceptron과 ●(=1), ●(=0)을 평면좌표상에 나타낸 그림이 있습니다.



$$\varphi(z) = \begin{cases} 0, & z < 0 \\ 1, & z \geq 0 \end{cases}$$



2-1. ●, ●를 분류하는 임의의 b, w를 선정하고 분류해보세요.

$b=0.5$, $w_1=2.0$, $w_2=-1.0$

	x_1	x_2	S	y
①	0	0	0.5	1
②	0	1	-0.5	0
③	1	0	2.5	1
④	1	1	1.5	1

① $x_1=0,\ x_2=0 \rightarrow (2.0\times0)+(-1.0\times0)+0.5=0.5$, $\varphi(0.5)=1$

② $x_1=0,\ x_2=1 \rightarrow (2.0\times0)+(-1.0\times1)+0.5=-0.5$, $\varphi(-0.5)=0$

③ $x_1=1,\ x_2=0 \rightarrow (2.0\times1)+(-1.0\times0)+0.5=2.5$, $\varphi(2.5)=1$

④ $x_1=1,\ x_2=1 \rightarrow (2.0\times1)+(-1.0\times1)+0.5=1.5$, $\varphi(1.5)=1$

2-2. Perceptron 학습 규칙에 따라 임의의 학습률을 정하고 b, w를 1회 업데이트 해주세요.

$lr=0.05$

(오분류 된 ②,③에 대하여 업데이트)

② $b \leftarrow b + 0.05(1-0)\times1$

$w_1 \leftarrow w_1 + 0.05(1-0)\times0$

$w_2 \leftarrow w_2 + 0.05(1-0)\times1$

$(b=0.55,\ w_1=2.0,\ w_2=-0.95)$

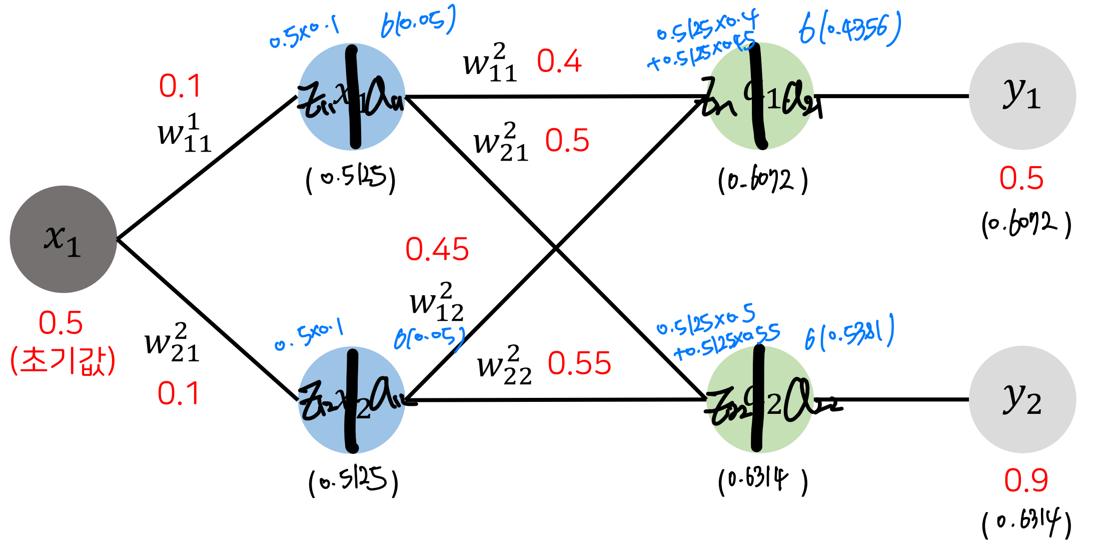
③ $b \leftarrow b + 0.05(0-1)\times1$

$w_1 \leftarrow w_1 + 0.05(0-1)\times1$

$w_2 \leftarrow w_2 + 0.05(0-1)\times0$

$(b=0.5,\ w_1=1.95,\ w_2=-0.95)$

3. 다음과 같은 구조와 초기값을 가진 Multilayer Perceptron이 있습니다.



3-1. Forward Propagation이 일어날 때, 각 노드는 어떤 값을 갖게 되는지 빈 칸을 채워주세요.
(Sigmoid Function 사용)

3-2. output layer에 있는 노드들의 Mean Squared Error을 구해주세요.

Cost Function : $\sum_{i}^{n} \frac{1}{2}(y_i - \hat{y}_i)^2$

n:2

$$\frac{1}{2}\left((0.5-0.6072)^2 + (0.9-0.6314)^2\right)$$

$$= \frac{1}{2}(0.01149184 + 0.07214596) \approx 0.0481$$

3-3. 3-2에서 구한 답을 토대로, Back Propagation이 일어날 때 가중치 w_{11}^1과 w_{11}^2의 조정된 값을 구해주세요. (학습률 $\eta = 0.5$)

뒷장 →

수고하셨습니다.

$$C = \sum \frac{1}{2}(y_i - \hat{y}_i)^2 = \frac{1}{2}(y_1 - a_{21})^2 + \frac{1}{2}(y_2 - a_{22})^2$$

$$= E_{y_1} + E_{y_2}$$

i) W_{11}^2

$$\rightarrow \frac{\partial C}{\partial W_{11}^2} = \frac{\partial C}{\partial a_{21}} \times \frac{\partial a_{21}}{\partial z_{21}} \times \frac{\partial z_{21}}{\partial W_{11}^2}$$

① $$\frac{\partial C}{\partial a_{21}} = \frac{1}{2} \times 2(y_1 - a_{21}) \times (-1) = -(y_1 - a_{21})$$

$$= -(0.5 - 0.6072) = 0.1072$$

② $$\frac{\partial a_{21}}{\partial z_{21}} = \left(\frac{1}{1+e^{-z_{21}}}\right)' = \underbrace{\frac{1}{1+e^{z_{21}}}}_{0.6072} \times \underbrace{\left(1 - \frac{1}{1+e^{-z_{21}}}\right)}_{0.3928}$$

$$= 0.2385$$

③ $$\frac{\partial z_{21}}{\partial W_{11}^2} = (a_{11} \times W_{11}^2 + a_{12} \times W_{12}^2)' = a_{11} = 0.5125$$

$$\therefore \frac{\partial C}{\partial W_{11}^2} = 0.1072 \times 0.2385 \times 0.5125 = 0.0131$$

$$W_{11}^2 = W_{11}^2 - 0.5 \times 0.0131$$

$$= 0.4 - 0.00655$$

$$= \boxed{0.39345}$$ 갱신

ii) W^1_{11}

$$\rightarrow \frac{\partial C}{\partial W^1_{11}} = \frac{\partial C}{\partial a_{11}} \times \frac{\partial a_{11}}{\partial z_{11}} \times \frac{\partial z_{11}}{\partial W^1_{11}}$$

분리

① $$\frac{\partial C}{\partial a_{11}} = \frac{\partial E_{y_1}}{\partial a_{11}} + \frac{\partial E_{y_2}}{\partial a_{11}} = \frac{\partial E_{y_1}}{\partial z_{21}} \times \frac{\partial z_{21}}{\partial a_{11}} + \frac{\partial E_{y_2}}{\partial z_{22}} \times \frac{\partial z_{22}}{\partial a_{11}}$$

$$= \left(\frac{\partial E_{y_1}}{\partial a_{21}} \times \frac{\partial a_{21}}{\partial z_{21}}\right) \times \frac{\partial z_{21}}{\partial a_{11}} + \left(\frac{\partial E_{y_2}}{\partial a_{22}} \times \frac{\partial a_{22}}{\partial z_{22}}\right) \times \frac{\partial z_{22}}{\partial a_{11}}$$

- $\frac{\partial E_{y_1}}{\partial a_{21}} = \frac{1}{2} \times 2\,(y_1 - a_{21}) \times (-1) = 0.1072$
- $\frac{\partial a_{21}}{\partial z_{21}} = \left(\frac{1}{1+e^{-z_{21}}}\right) \times \left(1 - \frac{1}{1+e^{-z_{21}}}\right) = 0.2385$
- $\frac{\partial z_{21}}{\partial a_{11}} = \frac{\partial}{\partial a_{11}}(a_{11} \times W^2_{11} + a_{12} \times W^2_{12}) = W^2_{11} = 0.4$
- $\frac{\partial E_{y_2}}{\partial a_{22}} = \frac{1}{2} \times 2 \times (y_2 - a_{22}) \times (-1) = -(0.9 - 0.6314) = -0.2686$
- $\frac{\partial a_{22}}{\partial z_{22}} = \underbrace{\frac{1}{1+e^{-z_{22}}}}_{0.6314} \times \left(1 - \underbrace{\frac{1}{1+e^{-z_{22}}}}_{0.3686}\right) = 0.2327$
- $\frac{\partial z_{22}}{\partial a_{11}} = \frac{\partial}{\partial a_{11}}(a_{11} \times W^2_{21} + a_{12} \times W^2_{22}) = W^2_{21} = 0.5$

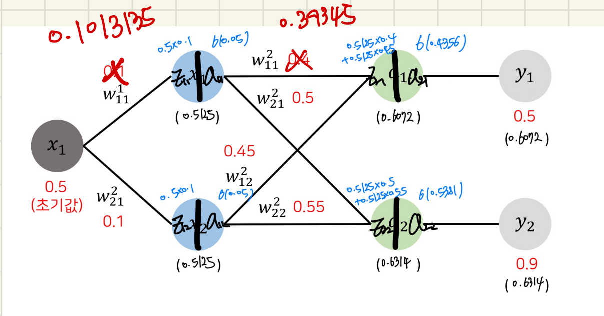
② $$\frac{\partial a_{11}}{\partial z_{11}} = \underbrace{\frac{1}{1+e^{-z_{11}}}}_{0.5125} \times \left(1 - \underbrace{\frac{1}{1+e^{-z_{11}}}}_{0.4875}\right) = 0.2499$$

③ $$\frac{\partial z_{11}}{\partial W^1_{11}} = \frac{\partial}{\partial W^1_{11}}(x_1 \times W^1_{11}) = x_1 = 0.5$$

$$\therefore \frac{\partial C}{\partial a_{11}} = [0.1072 \times 0.2385 \times 0.4 + (-0.2686) \times 0.2327 \times 0.5)] \times 0.2499 \times 0.5$$
$$= -0.002627$$

$$W^1_{11} = W^1_{11} - 0.5 \times (-0.002627)$$
$$= 0.1 + 0.0013135 = 0.1013135$$ 갱신

일단 업데이트된 2개의 가중치로 결괏값 산출 & y_1 값 비교



$$z_{11} = 0.5 \times 0.1013135 = 0.0507$$

$$a_{11} = \frac{1}{1+e^{-z_{11}}} = 0.5127$$

$$z_{21} = 0.5127 \times 0.39345 + 0.5125 \times 0.45 = 0.4323$$

$$a_{21} = \frac{1}{1+e^{-z_{21}}} = 0.6064$$

∴ 기존의 값인 0.6072에 비해 가중치 업데이트 후 결괏값인 0.6064가 실제값 0.5에 더 가까워 에러를 줄이는 방향으로 가중치가 업데이트 됐음을 알 수 있다.