투빅스 16기 정규세션 6주차

Neural Network Basic Assignment

이름: 김유호

1. Sigmoid Function을 z에 대해 미분하세요.

$1+e^{-z}=f(z)$ 로 두자. (즉, sigmoid function $=\frac{1}{f(z)}$)

$$\frac{\partial\sigma(z)}{\partial z}=\frac{\partial}{\partial z}\left(\frac{1}{f(z)}\right)=\frac{-f'(z)}{\{f(z)\}^2},\qquad \sigma(z)=\frac{1}{1+e^{-z}}$$

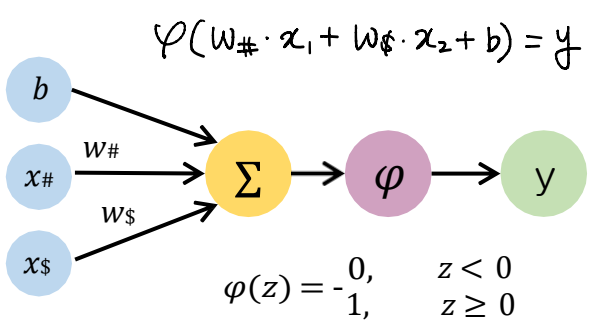
$f'(z)=-e^{-z}$

$$\therefore \frac{-f'(z)}{\{f(z)\}^2}=\frac{e^{-z}}{(1+e^{-z})^2}=\frac{1+e^{-z}-1}{(1+e^{-z})^2}=\frac{1}{1+e^{-z}}-\frac{1}{(1+e^{-z})^2}=\frac{1}{1+e^{-z}}\left(1-\frac{1}{1+e^{-z}}\right)$$

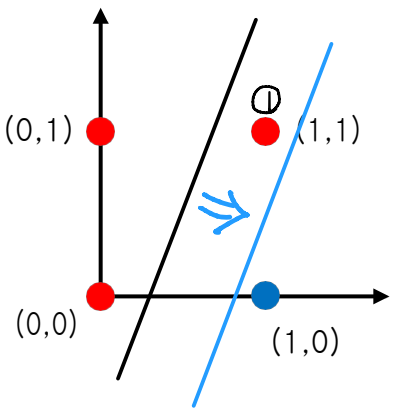
(시그모이드 함수)

$$=\sigma(z)(1-\sigma(z))$$

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.

$b=0.5$, $w_\#=-1.2$, $w_\$=0.6$

(x_1, x_2)

$(0,0)$: $\varphi(0.5)=1$

$(0,1)$: $\varphi(1.1)=1$

$(1,0)$: $\varphi(-0.7)=0$

$(1,1)$: $\varphi(-0.1)=0$

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

* 학습률 $=0.05$

① $b=0.5+0.05(1-0)\cdot 1=0.55$

$w_\#=-1.2+0.05(1-0)\cdot 1=-1.15$

$w_\$=0.6+0.05(1-0)\cdot 1=0.65$

(업데이트 후, 분류)

$(0,0)$: $\varphi(0.55)=1$

$(0,1)$: $\varphi(1.2)=1$

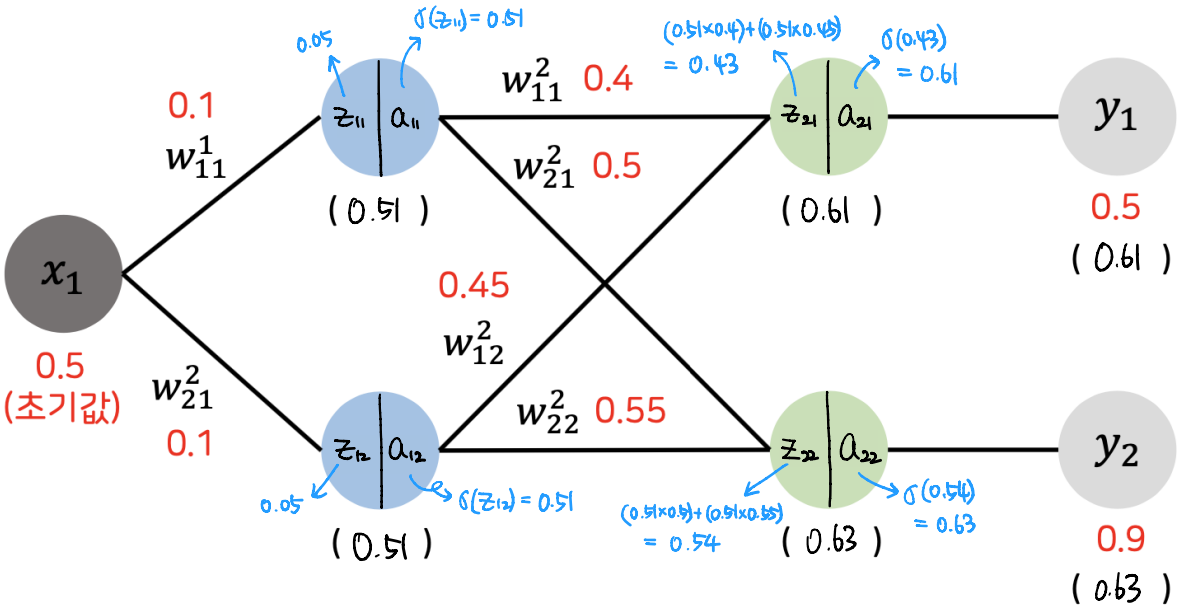
$(1,0)$: $\varphi(-0.6)=0$

$(1,1)$: $\varphi(0.05)=1$

⇒ $b=0.55$, $w_\#=-1.15$, $w_\$=0.65$ 로 가중치와 bias가 업데이트 되어 올바르게 분류된다.

3-3에서 의도하신 문제가 시그모이드 함수 계산이 2개의 층에 모두 있는 경우인 것 같아 그것 수정했습니다.

3. 다음과 같은 구조와 초기값을 가진 Multilayer Perceptron이 있습니다. (소수 둘째자리까지 반올림)



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.

$$MSE_t = \frac{1}{2}\left\{(0.5-0.61)^2+(0.9-0.63)^2\right\}$$

$$= \frac{1}{2}(0.0121+0.0729) = 0.0425$$

$(MSE_1 = 0.00605,\ MSE_2 = 0.03645)$

※ $MSE = \frac{1}{n}\sum\frac{1}{2}(y_i-\hat{y}_i)^2$ (본문제의 input은 하나이므로 여기서 $n=1$, $\frac{1}{2}$: 편의상 시, 사용자 편의를 위한 $\frac{1}{2}$)

3-3. 3-2에서 구한 답을 토대로, Back Propagation이 일어날 때 가중치 w_{11}^1과 w_{11}^2의 조정된 값을 구해주세요. (learning rate : 0.4)

① $\frac{\partial E_t}{\partial w_{11}^1} = \frac{\partial E_1}{\partial w_{11}^1} + \frac{\partial E_2}{\partial w_{11}^1} = \left(\frac{\partial E_1}{\partial a_{21}}\cdot\frac{\partial a_{21}}{\partial z_{21}}\cdot\frac{\partial z_{21}}{\partial a_{11}}\cdot\frac{\partial a_{11}}{\partial z_{11}}\cdot\frac{\partial z_{11}}{\partial w_{11}^1}\right) + \left(\frac{\partial E_2}{\partial a_{22}}\cdot\frac{\partial a_{22}}{\partial z_{22}}\cdot\frac{\partial z_{22}}{\partial a_{11}}\cdot\frac{\partial a_{11}}{\partial z_{11}}\cdot\frac{\partial z_{11}}{\partial w_{11}^1}\right)$

② $\frac{\partial E_1}{\partial w_{11}^2} = \frac{\partial E_1}{\partial a_{21}}\cdot\frac{\partial a_{21}}{\partial z_{21}}\cdot\frac{\partial z_{21}}{\partial w_{11}^2}$ (뒷장에서 계속)

수고하셨습니다.

① $\frac{\partial E_t}{\partial w^1_{11}} = \frac{\partial E_1}{\partial w^1_{11}} + \frac{\partial E_2}{\partial w^1_{11}} = \left(\frac{\partial E_1}{\partial a_{21}} \cdot \frac{\partial a_{21}}{\partial z_{21}} \cdot \frac{\partial z_{21}}{\partial a_{11}} \cdot \frac{\partial a_{11}}{\partial z_{11}} \cdot \frac{\partial z_{11}}{\partial w^1_{11}}\right) + \left(\frac{\partial E_2}{\partial a_{22}} \cdot \frac{\partial a_{22}}{\partial z_{22}} \cdot \frac{\partial z_{22}}{\partial a_{11}} \cdot \frac{\partial a_{11}}{\partial z_{11}} \cdot \frac{\partial z_{11}}{\partial w^1_{11}}\right)$

㉠ $\frac{\partial}{\partial a_{21}}\left(\frac{1}{2}(y_1 - a_{21})^2 + \frac{1}{2}(y_2 - a_{22})^2\right) = a_{21} - y_1 = 0.11$

㉡ 시그모이드함수 미분 ⇒ $\sigma(z_{21}) \cdot (1 - \sigma(z_{21})) = 0.61 \times 0.39 = 0.24$

㉢ 0.4

㉣ $(0.51 \times 0.49) \times 0.5 = 0.12$

㉠′ : $a_{22} - y_2 = -0.27$

㉡′ : $0.63 \times 0.37 = 0.23$

㉢′ : 0.5

㉣ : 0.12

$\therefore \frac{\partial E_t}{\partial w^1_{11}} = (0.11 \times 0.24 \times 0.4 \times 0.12) + (-0.27 \times 0.23 \times 0.5 \times 0.12) = -0.0025$

$\therefore W^1_{11}(\text{update}) = 0.1 - 0.4(-0.0025) = \underline{0.101}$

② $\frac{\partial E_1}{\partial w^2_{11}} = \frac{\partial E_1}{\partial a_{21}} \cdot \frac{\partial a_{21}}{\partial z_{21}} \cdot \frac{\partial z_{21}}{\partial w^2_{11}}$

$= 0.11 \times 0.24 \times 0.51$

$= 0.013464$

$\therefore W^2_{11}(\text{update}) = 0.4 - 0.4(0.013464)$

$= \underline{0.3946}$