

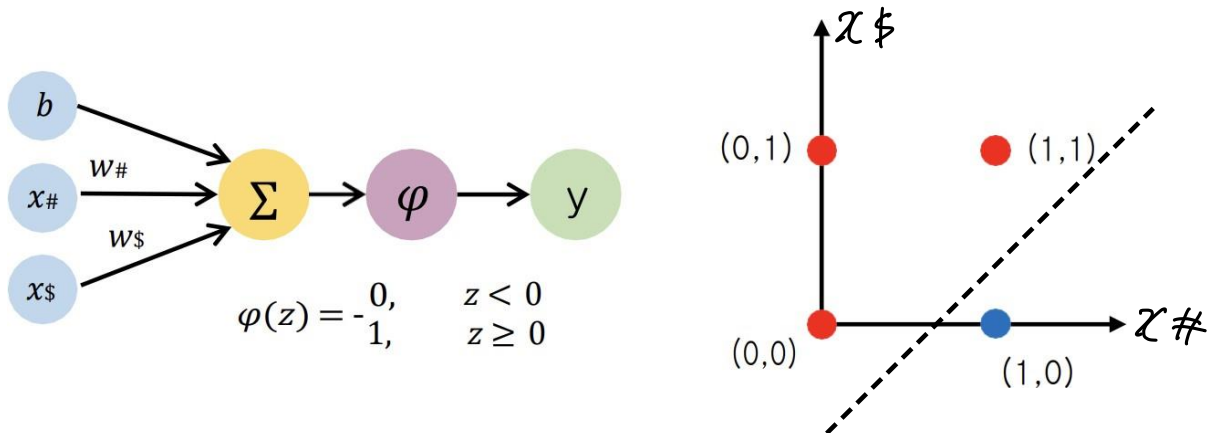
# Neural Network Basic Assignment

이름: 이혁준

1. Sigmoid Function을  $z$ 에 대해 미분하세요.

$$\begin{aligned}\sigma'(z) &= \frac{e^{-z}}{(1+e^{-z})^2} & \sigma(z) &= \frac{1}{1+e^{-z}} \\ &= \frac{1}{1+e^{-z}} \cdot \frac{e^{-z}}{1+e^{-z}} \\ &= \frac{1}{1+e^{-z}} \cdot \left(1 - \frac{1}{1+e^{-z}}\right) = \underline{\underline{\sigma(z)(1-\sigma(z))}}\end{aligned}$$

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



2-1. ●, ●을 분류하는 임의의  $b, w$ 를 선정하고 분류해보세요.

$x\#$	$x\$\$	$y$
0	0	1
0	1	1
1	0	0
1	1	1

$\begin{cases} w\# = -1.0 \\ w\$ = 1.0 \\ b = 1.0 \end{cases}$

$x\#$	$x\$\$	$s$	$0$
0	0	1	1
0	1	2	1
1	0	0	1
1	1	1	1

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

$$w\# = -1.0, w\$ = 1.0, b = 1.0$$

$$\eta = 0.05$$

$x\#$	$x\$\$	$s$	$0$	$y$
0	0	1	1	1
0	1	2	1	1
1	0	0	1	0
1	1	1	1	1

①

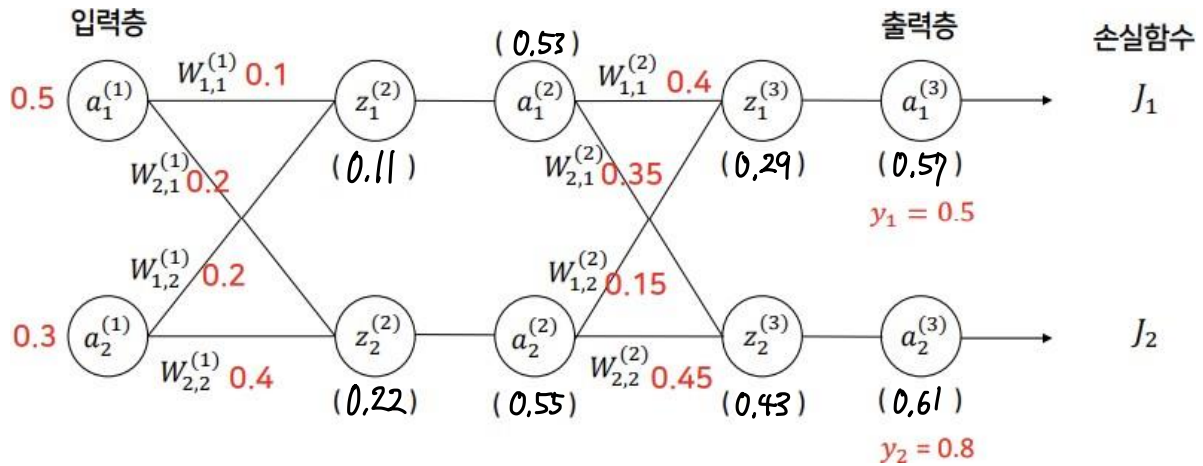
$$\textcircled{1} \quad b \leftarrow b + 0.05(0-1) \times 1 \quad b \leftarrow 1.0 + 0.05(0-1) \times 1 = 0.95$$

$$w\# \leftarrow w\# + 0.05(0-1) \times 1 \quad w\# \leftarrow -1.0 + 0.05(0-1) \times 1 = -1.05$$

$$w\$ \leftarrow w\$ + 0.05(0-1) \times 0 \quad w\$ \leftarrow 1.0 + 0.05(0-1) \times 0 = 1.0$$

$$\langle w\# = 0.95, w\$ = -1.05, b = 1.0 \rangle$$

3. 다음과 같이 입력과 가중치가 주어진 퍼셉트론이 있을 때, 아래의 물음에 답해주세요. 모든 문제는 풀이과정을 자세하게 적어주세요! (3-3까지 있습니다.)



- 3-1. FeedForward가 일어날 때, 각 노드가 갖는 값을 빈칸에 써주세요. 단, 활성화함수는 sigmoid 함수입니다. (모든 계산의 결과는 소수점 셋째자리에서 반올림하여 둘째자리까지만 써주세요.)  $\phi(z) = \frac{1}{1+e^{-z}}$

입력층 $\rightarrow$ 1층	1층 활성화 함수	1층 $\rightarrow$ 2층	2층 활성화 함수
$z_1^{(2)} = W_{11}^{(1)} \cdot a_1^{(1)} + W_{12}^{(1)} \cdot a_2^{(1)}$ $= 0.11$	$a_1^{(2)} = \phi(z_1^{(2)})$ $= 0.53$	$z_1^{(3)} = 0.29$	$a_1^{(3)} = \phi(z_1^{(3)})$ $= 0.57$
$z_2^{(2)} = W_{21}^{(1)} \cdot a_1^{(1)} + W_{22}^{(1)} \cdot a_2^{(1)}$ $= 0.22$	$a_2^{(2)} = \phi(z_2^{(2)})$ $= 0.55$	$z_2^{(3)} = 0.43$	$a_2^{(3)} = \phi(z_2^{(3)})$ $= 0.61$

- 3-2. 3-1에서 구한 값을 이용하여 손실함수  $J_1$ 과  $J_2$ 의 값을 구해주세요. ( $J_1$ 과  $J_2$ 는 반올림하지 말고 써주세요.)

손실함수 : 
$$MSE = \frac{1}{2N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

$$J_1 = \frac{1}{2} (a_1^{(3)} - y_1)^2$$

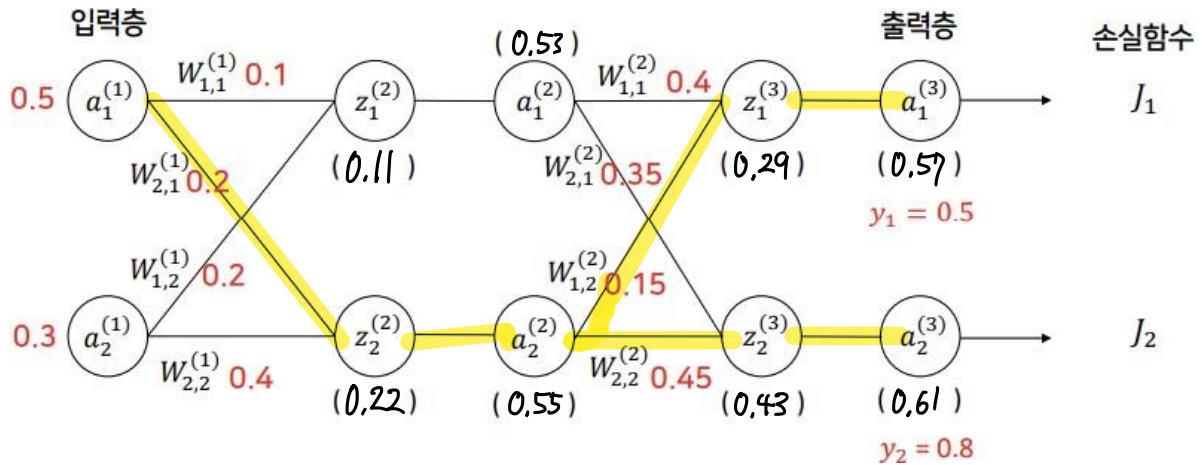
$$= \frac{(0.57 - 0.5)^2}{2}$$

$$= \underline{0.00245}$$

$$J_2 = \frac{1}{2} (a_2^{(3)} - y_2)^2$$

$$= \frac{(0.61 - 0.8)^2}{2}$$

$$= \underline{0.01805}$$



- 3-3. 위에서 구한 값을 토대로, BackPropagation이 일어날 때  $w_{2,2}^{(2)}$ 과  $w_{2,1}^{(1)}$ 의 조정된 값을 구해주세요.  
단, learning rate는 0.1입니다. (계산 과정에서 소수점 넷째자리에서 반올림하여 셋째자리까지만 써주시고, 마지막 결과인  $w_{2,1}^{(1)}$ 과  $w_{2,2}^{(2)}$ 의 값만 반올림하지 말고 써주세요.)

출력층 → 2층

$$\begin{aligned} w_{1,1}^{(2)} &= w_{1,1}^{(2)} - \frac{\partial J_{total}}{\partial w_{1,1}^{(2)}} \\ &= w_{1,1}^{(2)} - \frac{\partial J_1}{\partial a_1^{(3)}} \cdot \frac{\partial a_1^{(3)}}{\partial z_1^{(2)}} \cdot \frac{\partial z_1^{(2)}}{\partial w_{1,1}^{(2)}} \\ &= w_{1,1}^{(2)} - \left\{ (a_1^{(3)} - y_1) (a_1^{(3)} (1 - a_1^{(3)})) a_1^{(2)} \right\} \\ &= w_{1,1}^{(2)} - \left\{ \delta_1^{(3)} a_1^{(2)} \right\} \end{aligned}$$

$$\begin{cases} \frac{\partial J_1}{\partial a_1^{(3)}} = \frac{1}{2} \frac{\partial}{\partial a_1^{(3)}} (a_1^{(3)} - y_1)^2 = a_1^{(3)} - y_1 \\ \frac{\partial a_1^{(3)}}{\partial z_1^{(2)}} = \frac{2}{2z_1^{(2)}} \left( \frac{1}{1 + e^{-2z_1^{(2)}}} \right) = a_1^{(3)} (1 - a_1^{(3)}) \\ \frac{\partial z_1^{(2)}}{\partial w_{1,1}^{(2)}} = \frac{\partial}{\partial w_{1,1}^{(2)}} (w_{1,1}^{(2)} a_1^{(2)} + w_{1,2}^{(2)} a_2^{(2)}) = a_1^{(2)} \end{cases}$$

learning rate →  $w_{2,2}^{(2)} = w_{2,2}^{(2)} - \eta \frac{\partial J_2}{\partial w_{2,2}^{(2)}} = w_{2,2}^{(2)} - \eta \left\{ (a_2^{(3)} - y_2) (a_2^{(3)} (1 - a_2^{(3)})) a_2^{(2)} \right\}$

$$= 0.45 - (0.1) \left\{ (0.61 - 0.8) (0.61 (1 - 0.61)) (0.61) \right\}$$

$$\approx 0.447$$

2층 → 1층 (J<sub>total</sub> = J<sub>1</sub> + J<sub>2</sub>)

learning rate →  $w_{2,1}^{(1)} = w_{2,1}^{(1)} - \eta \frac{\partial J_{total}}{\partial w_{2,1}^{(1)}} = w_{2,1}^{(1)} - \eta \left( \frac{\partial J_{total}}{\partial a_2^{(3)}} \cdot \frac{\partial a_2^{(3)}}{\partial z_2^{(1)}} \cdot \frac{\partial z_2^{(1)}}{\partial w_{2,1}^{(1)}} \right)$

$$= w_{2,1}^{(1)} - \eta \left( \frac{\partial J_1 + \partial J_2}{\partial a_2^{(3)}} \right) (a_2^{(3)} (1 - a_2^{(3)}) \cdot a_2^{(1)})$$

$$= 0.2 - (0.1) \left( (0.07 \cdot 0.57 \cdot 0.43) \cdot 0.4 + (-0.19) (0.61 \cdot 0.39 \cdot 0.45) \right)$$

$$\times (0.55 \cdot 0.45 \cdot 0.5)$$

$$= 0.20016 \dots$$

$$\approx 0.200$$

\*  $\frac{\partial J_{total}}{\partial a_2^{(3)}} = \frac{\partial J_1 + \partial J_2}{\partial a_2^{(3)}}$

$$\Rightarrow \frac{\partial J_1}{\partial a_2^{(3)}} = \frac{\partial J_1}{\partial z_1^{(2)}} \times \frac{\partial z_1^{(2)}}{\partial a_2^{(3)}} = \delta_1^{(2)} \cdot w_{1,2}^{(2)}$$

$$= (a_1^{(2)} - y_1) (a_1^{(2)} (1 - a_1^{(2)}))$$

$$\frac{\partial J_2}{\partial a_2^{(3)}} = \frac{\partial J_2}{\partial z_2^{(1)}} \times \frac{\partial z_2^{(1)}}{\partial a_2^{(3)}} = \delta_2^{(1)} \cdot w_{2,2}^{(1)}$$

$$= (a_2^{(3)} - y_2) (a_2^{(3)} (1 - a_2^{(3)}))$$