

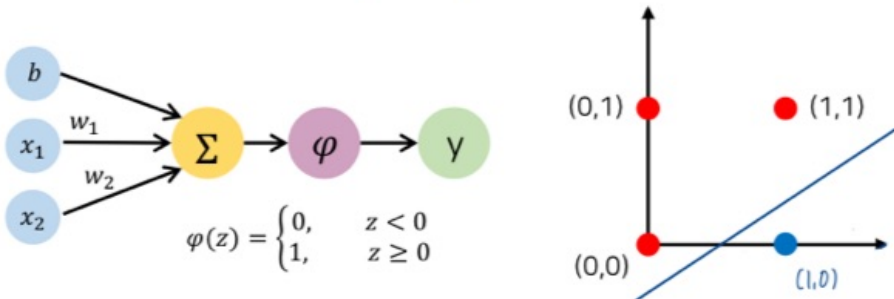
# Neural Network Basic Assignment 1

이름: 이소영

1. Sigmoid Function을  $z$ 에 대해 미분하세요.

$$\begin{aligned} \frac{d}{dz} \sigma(z) &= -(1+e^{-z})^{-2} \cdot (-e^{-z}) \\ &= (1+e^{-z})^{-2} \cdot e^{-z} \\ &= \frac{e^{-z}}{(1+e^{-z})^2} = \frac{1}{1+e^{-z}} \cdot \frac{e^{-z}}{1+e^{-z}} \\ &= \sigma(z)(1-\sigma(z)) \end{aligned} \quad \sigma(z) = \frac{1}{1+e^{-z}} = (1+e^{-z})^{-1}$$

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



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

$$b=1, w_1=1, w_2=-1$$

$x_1$	$x_2$	$z$	$y$
0	0	1	1
1	0	2	1
0	1	0	1
1	1	1	1

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

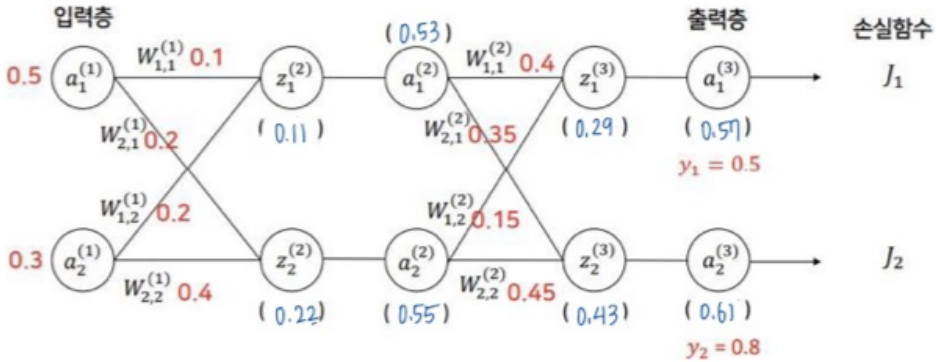
$$\text{학습률: } 0.05$$

$$b = 1 + 0.05(0-1) \times 1 = 0.95$$

$$w_1 = 1 + 0.05(0-1) \times 1 = 0.95$$

$$w_2 = -1 + 0.05(0-1) \times 0 = -1$$

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



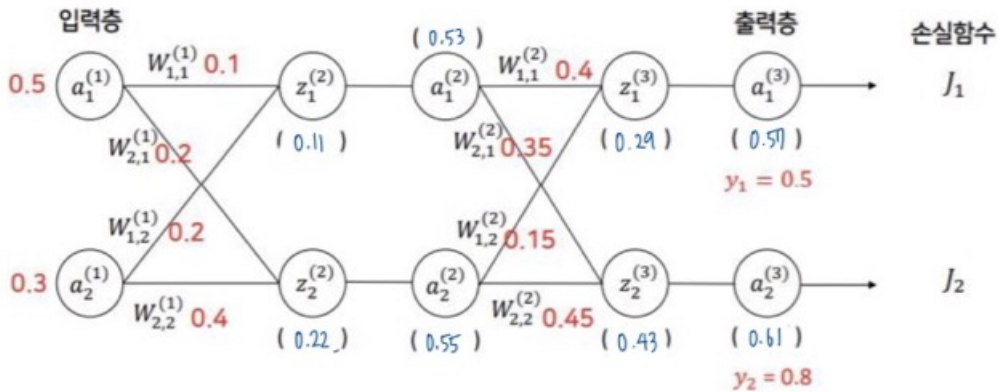
- 3-1. FeedForward가 일어날 때, 각 노드가 갖는 값을 빈칸에 써주세요. 단, 활성화함수는 sigmoid 함수입니다. (모든 계산의 결과는 소수점 셋째자리에서 반올림하여 둘째자리까지만 써주세요.)

$$\begin{aligned}
 z_1^{(2)} &= 0.5 \cdot 0.1 + 0.3 \cdot 0.2 = 0.11 & a_1^{(2)} &= (1 + e^{-0.11})^{-1} = 0.53 \\
 z_2^{(2)} &= 0.5 \cdot 0.2 + 0.3 \cdot 0.4 = 0.22 & a_2^{(2)} &= (1 + e^{-0.22})^{-1} = 0.55 \\
 z_1^{(3)} &= 0.53 \cdot 0.4 + 0.55 \cdot 0.15 = 0.29 & a_1^{(3)} &= (1 + e^{-0.29})^{-1} = 0.57 \\
 z_2^{(3)} &= 0.53 \cdot 0.35 + 0.55 \cdot 0.45 = 0.43 & a_2^{(3)} &= (1 + e^{-0.43})^{-1} = 0.61
 \end{aligned}$$

- 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.)

$$J_1 = \frac{1}{2} (0.5 - 0.57)^2 = 0.00245$$

$$J_2 = \frac{1}{2} (0.8 - 0.61)^2 = 0.01805$$



- 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 써 주시고, 마지막 결과인  $w_{2,1}^{(1)}$ 과  $w_{2,2}^{(2)}$ 의 값만 반올림하지 말고 써주세요.)

$$w_{2,2}^{(2)} = w_{2,2}^{(2)} - \eta \frac{\partial J_{\text{total}}}{\partial w_{2,2}^{(2)}}$$

$$\frac{\partial J_{\text{total}}}{\partial w_{2,2}^{(2)}} = \underbrace{\frac{\partial J_2}{\partial a_2^{(3)}}}_1 \times \underbrace{\frac{\partial a_2^{(3)}}{\partial z_2^{(3)}}}_2 \times \underbrace{\frac{\partial z_2^{(3)}}{\partial w_{2,2}^{(2)}}}_3$$

$$1 = \frac{1}{2} \frac{\partial}{\partial a_2^{(3)}} (a_2^{(3)} - y_2)^2 = a_2^{(3)} - y_2 = -0.19$$

$$2 = a_2^{(3)} (1 - a_2^{(3)}) = 0.61 \times 0.39 = 0.22$$

$$3 = a_2^{(2)} = 0.55$$

$$w_{2,2}^{(2)} = 0.45 - 0.1 (-0.19) \cdot (0.22) \cdot (0.55) = 0.452299$$

$$w_{2,1}^{(1)} = w_{2,1}^{(1)} - \eta \frac{\partial J_{\text{total}}}{\partial w_{2,1}^{(1)}}$$

$$\frac{\partial J_{\text{total}}}{\partial w_{2,1}^{(1)}} = \underbrace{\frac{\partial J_{\text{total}}}{\partial a_1^{(3)}}}_1 \cdot \underbrace{\frac{\partial a_1^{(3)}}{\partial z_1^{(3)}}}_2 \cdot \underbrace{\frac{\partial z_1^{(3)}}{\partial w_{2,1}^{(1)}}}_3$$

$$1 = \frac{\partial J_1}{\partial a_1^{(3)}} + \frac{\partial J_2}{\partial a_1^{(3)}} = \delta_1^{(3)} \cdot w_{1,1}^{(2)} + \delta_2^{(3)} \cdot w_{2,1}^{(2)}$$

$$2 = a_2^{(2)} (1 - a_2^{(2)}) = 0.55 \cdot 0.45 = 0.25$$

$$3 = a_2^{(1)} = 0.3$$