

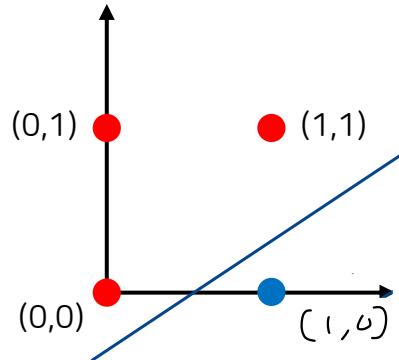
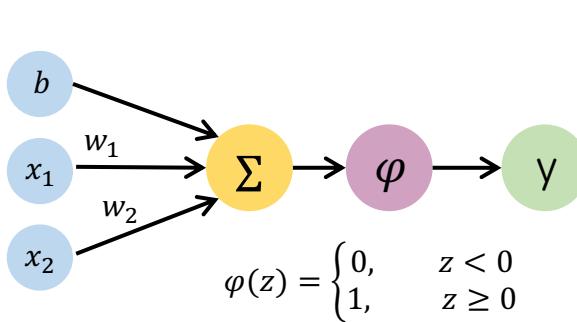
Neural Network Basic Assignment 1

이름: 양승현

1. Sigmoid Function을 z 에 대해 미분하세요.

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

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

2. 다음과 같은 구조의 Perceptron과 $\bullet (=1)$, $\circ (=0)$ 을 평면좌표상에 나타낸 그림이 있습니다.2-1. \bullet , \circ 를 분류하는 임의의 b, w 를 선정하고 분류해보세요.

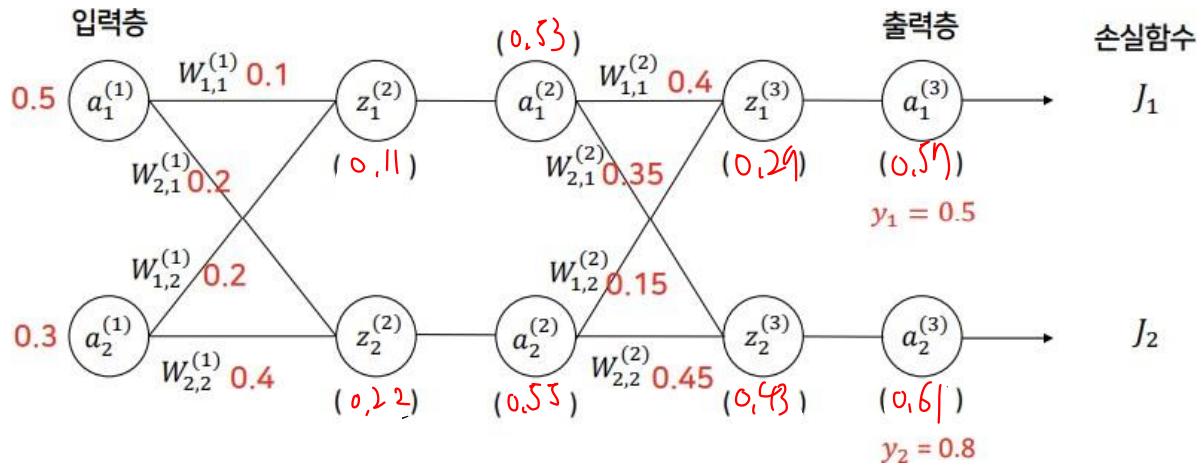
$w_1x_1 + w_2x_2 + b = 2 \quad \ell(b) = 1, \quad \ell(w_1+b) = 1, \quad \ell(w_1+w_2+b) = 1, \quad \ell(w_2+b) = 0$

$b \geq 0, \quad w_2+b \geq 0, \quad w_1+w_2+b \geq 0, \quad w_1+b < 0,$

$w_1 = -1, \quad w_2 = 1, \quad b = 0.5$

2-2. Perceptron 학습 규칙에 따라 임의의 학습률을 정하고 b, w 를 1회 업데이트 해주세요.오늘은 허락해요 $\times (.,?)$

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

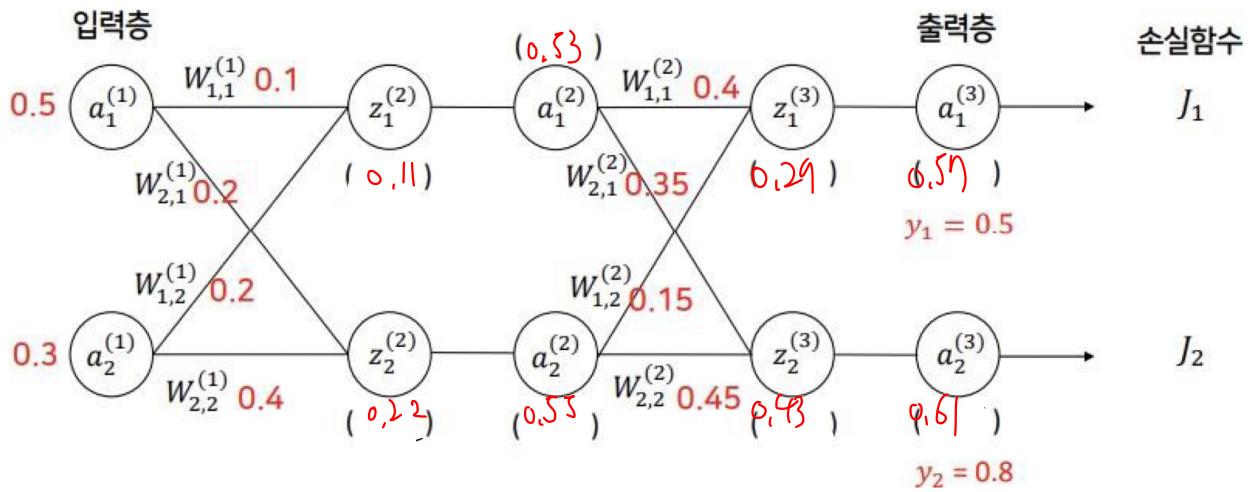


- 3-1. FeedForward가 일어날 때, 각 노드가 갖는 값을 빈칸에 써주세요. 단, 활성화함수는 sigmoid 함수입니다. (모든 계산의 결과는 소수점 셋째자리에서 반올림하여 둘째자리까지만 써주세요.)

- 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.)

$$J_1 = \frac{1}{2}(\alpha_1^3 - y_1)^2 = \frac{1}{2}(0.57 - 0.5)^2 = 0.00245$$

$$J_2 = \frac{1}{2}(\alpha_2^3 - y_2)^2 = \frac{1}{2}(0.61 - 0.8)^2 = 0.01805$$



3-3. 위에서 구한 값을 토대로, BackPropagation이 일어날 때 $W_{2,2}^{(1)}$ 과 $W_{2,1}^{(2)}$ 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.)

$$\begin{aligned}
 W_{2,2}^{(2)} &= W_{2,2}^{(2)} - \alpha \delta_2^{(3)} a_2^{(2)} \quad (\alpha = 0.1) \\
 &= W_{2,2}^{(2)} - \alpha \frac{\partial J_2}{\partial z_2^{(3)}} a_2^{(2)} = W_{2,2}^{(2)} - \alpha (a_2^{(3)} - y_2) \times a_2^{(2)} (1 - a_2^{(2)}) \times a_2^{(2)} \\
 &= 0.45 - \alpha (0.61 - 0.8) \times 0.61 \times (1 - 0.61) \times 0.55 \\
 &= 0.45 + 0.002486055 \\
 &= \boxed{0.452}
 \end{aligned}$$

$$\begin{aligned}
 W_{2,1}^{(1)} &= W_{2,1}^{(1)} - \alpha \frac{\partial J_{\text{total}}}{\partial a_2^{(1)}} \times \frac{\partial a_2^{(1)}}{\partial z_2^{(2)}} \times \frac{\partial z_2^{(2)}}{\partial w_{2,1}^{(1)}} \\
 &= W_{2,1}^{(1)} - \alpha \left(\frac{\partial J_1}{\partial a_2^{(1)}} + \frac{\partial J_2}{\partial a_2^{(1)}} \right) \times a_2^{(1)} (1 - a_2^{(1)}) \times a_2^{(1)} \\
 &= W_{2,1}^{(1)} - \alpha \left(f_1^{(1)} w_{1,2}^{(2)} + f_2^{(1)} w_{2,2}^{(2)} \right) \times a_2^{(1)} (1 - a_2^{(1)}) \times a_2^{(1)}
 \end{aligned}$$

$$f_1^{(1)} = (a_1^{(1)} - y_1) \times a_1^{(1)} (1 - a_1^{(1)}) = (0.5 - 0.5) \times (0.5) \times (1 - 0.5) = 0.017$$

$$f_2^{(1)} = (a_2^{(1)} - y_2) \times a_2^{(1)} (1 - a_2^{(1)}) = (0.3 - 0.8) \times (0.3) \times (1 - 0.3) = -0.028$$

$$W_{2,1}^{(1)} = W_{2,1}^{(1)} - \alpha \left(0.017 \times 0.15 - 0.028 \times 0.45 \right) \times 0.55 \times (1 - 0.55) \times 0.5$$

$$= 0.2 - 0.1 \times (-0.01005) \times 0.55 \times 0.48 \times 0.5$$

$$= 0.12601$$