1.

1. Sigmoid Function을 Z에 대해 미분.

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

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

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

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

$$= \delta(1-\delta)$$

2.

2. 다음과 같은 구조의 Perceptron ①=1, ⑪=0 을 좌표평면상에
나타낸 그림이 있습니다.

x_2↑

(0,1) ○   ○ (1,1)

(b)
x_1 w_1 → (Σ) → (ψ) → (y)
x_2 w_2

(0,0) ⊕   ◉ (1,0) x_1

2-1. ①, ⓪을 분류하는 임의의 b, w를 선정하고 분류

$$b = 0.5 \quad w = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

x_1	x_2	S	y	Ground Truth	
0	0	0.5	1	1	
0	1	1.5	1	1	
1	0	-0.5	0	0	
1	1	0.5	1	1	

2-2. Perceptron 학습규칙에 따라 임의의 학습률을 설정하고 b, w를 1회 업데이트.

$$b_0 = -0.5 \quad w_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \eta = 0.5 \text{ 라 하자.}$$

○   ○

⊕   ◉

① $b_0 \leftarrow b_0 + 0.5 \times (1-0) \times 1$
  $w_1 \leftarrow w_1 + 0.5 \times (1-0) \times 0$
  $w_2 \leftarrow w_2 + 0.5 \times (1-0) \times 0.$ $b_1 = -0.5$

① ② 0시점 결정경계

$$w_1 = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix}$$

↓

② $b_0 \leftarrow b_0 + 0.5 \times (0-1) \times 1$
  $w_1 \leftarrow w_1 + 0.5 \times (0-1) \times 1$
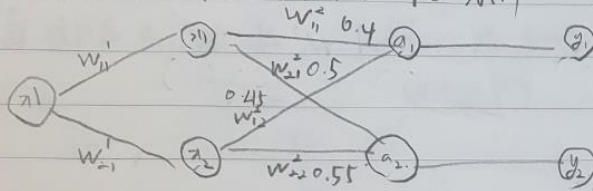 $w_2 \leftarrow w_2 + 0.5 \times (0-1) \times 0$

○   ○

⊕   ●

3.

3. 다음과 같은 구조와 초기값을 가진 MLP가 있습니다.

3-1. Forward Propagation이 일어날 때, 각 노드는 어떤 값을 갖게되는지 빈 칸을 채우세요.

$X = [0.5]$, $W_1 = \begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix}$ 이다.

$Z_1 = W_1 X = \begin{bmatrix} 0.05 \\ 0.05 \end{bmatrix}$, $a_1 = sigmoid(Z_1) = \begin{bmatrix} 0.5124 \\ 0.5124 \end{bmatrix}$

$W_2 = \begin{bmatrix} 0.4 & 0.45 \\ 0.5 & 0.55 \end{bmatrix}$ 이다.

$Z_2 = W_2 a_1 = \begin{bmatrix} 0.4356 \\ 0.5381 \end{bmatrix}$ $a_2 = sigmoid(Z_2) = \begin{bmatrix} 0.6072 \\ 0.6313 \end{bmatrix}$
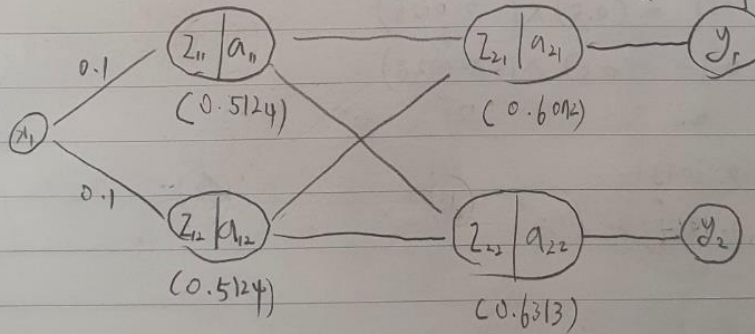
3-2. output layer에 있는 노드들의 MSE를 구하세요.

$$MSE = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{2} (y_i - \hat{y_i})^2 = \frac{1}{2} \left[\frac{1}{2} (0.5 - 0.6072)^2 + \frac{1}{2} (0.9 - 0.6313)^2 \right]$$

$$= 0.0209$$

3-3 Back propagation이 끝난 뒤 가중치 W_{11}' 과 W_{12}' 가 조정된 값

$$(\eta = 0.5)$$

$$E = \frac{1}{2}\sum_{k=1}^{m}(y_k - a_{2k})^2$$

＊ $\times$ 는 element wise product 이는 dot product 입니다.

$$\frac{\partial E}{\partial W_{2j}} = \frac{\partial E}{\partial a_{2k}} \cdot \frac{\partial a_{2k}}{\partial Z_{2k}} \cdot \frac{\partial Z_{2k}}{\partial a_{1j}} \cdot \frac{\partial a_{1j}}{\partial Z_{1j}} \cdot \frac{\partial Z_{1j}}{\partial W_{2j}}$$

$$= -(y - a_2)\times(a_2 \times (1-a_2)) \cdot W_2 \times (a_1 \times (1-a_1) \times \lambda)$$

$$= \begin{bmatrix} 0.1072 \\ -0.2686 \end{bmatrix} \times \begin{bmatrix} 0.2385 \\ 0.2327 \end{bmatrix} \cdot \begin{bmatrix} 0.4 & .45 \\ 0.5 & 0.55 \end{bmatrix} \times \begin{bmatrix} 0.2498 \\ 0.2498 \end{bmatrix} \times [0.5]$$

$$= \begin{bmatrix} -0.0026 \\ -0.0028 \end{bmatrix}$$

$W_{11}' \longleftarrow 0.1 - (0.5)\times(-0.0026)$

$W_{12}' \longleftarrow 0.1 - (0.5)\times(-0.0028)$

$W_{11}' = 0.10131\cdots$

$W_{12}' = 0.10142\cdots$