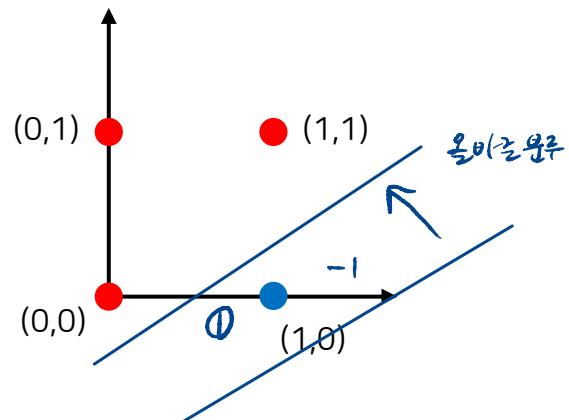
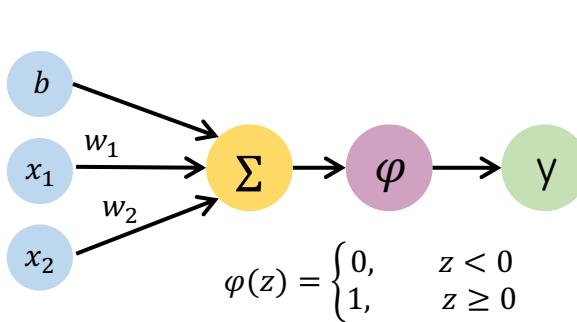


Neural Network Basic Assignment 1

이름: 강지우

1. Sigmoid Function을 z 에 대해 미분하세요.

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

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.

$$\begin{aligned}
 b &= 0.4, w_1 = -0.3, w_2 = 0.4 \\
 (0,1) : \varphi(0.8) &= 1 \\
 (0,0) : \varphi(0.4) &= 1 \\
 (1,1) : \varphi(0.5) &= 1 \\
 (1,0) : \varphi(0.1) &= 1
 \end{aligned}$$

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

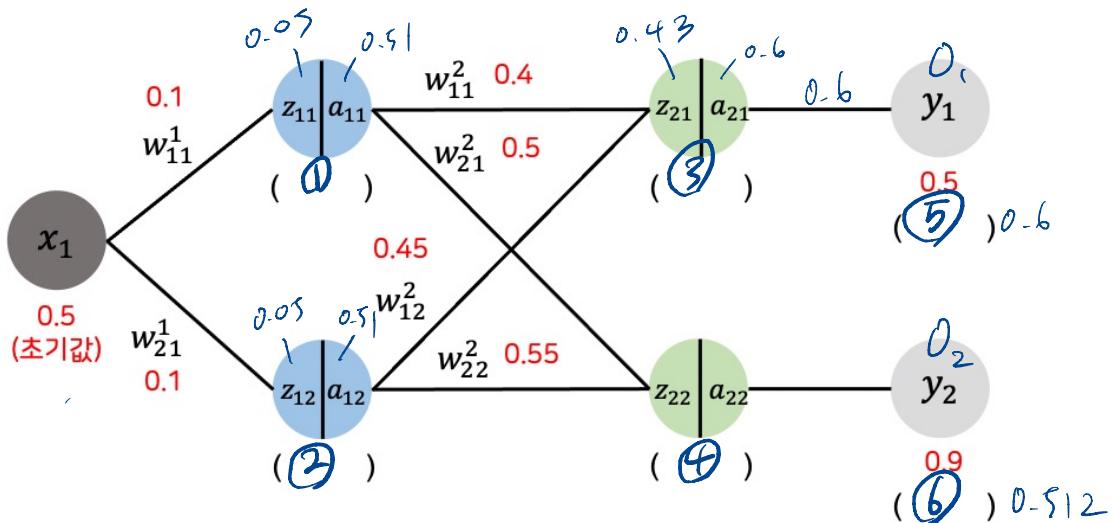
학습률 = 0.05

$$\begin{aligned}
 ① \quad b &= 0.4 + 0.05(-1-1) = 0.3 \\
 w_1 &= -0.3 + 0.05(-1-1) \times 1 = -0.7 \\
 w_2 &= 0.4 + 0.05(-1-1) \times 0 = 0.4
 \end{aligned}$$

$$b = 0.3, w_1 = -0.7, w_2 = 0.4 \quad ?$$

업데이트되면, ①이 올바르게 분류될 수 있다.

3. 다음과 같은 구조와 초기값을 가진 Multilayer Perceptron이 있습니다. (수정된 사진 적용)



3-1. Forward Propagation이 일어날 때, 각 노드는 어떤 값을 갖게 되는지 빈 칸을 채워주세요.

$$\text{(Sigmoid Function 사용)} \quad ① : z_{11} = 0.05, a_{11} = \frac{1}{1+e^{-0.05}} = 0.51 \quad ② z_{21} = 0.45 \times 0.51 = 0.43, a_{21} = 0.6$$

$$③ z_{12} = 0.05, a_{12} = \frac{1}{1+e^{-0.05}} = 0.51 \quad ④ z_{22} = 0.1 \times 0.51 = 0.051, a_{22} = 0.512$$

3-2. output layer에 있는 노드들의 Mean Squared Error를 구해주세요.

$$MSE_1 = \frac{1}{2} (0.5 - 0.6)^2 = 0.02$$

$$MSE_2 = \frac{1}{2} (0.9 - 0.512)^2 = 0.388$$

3-3. 3-2에서 구한 답을 토대로, Back Propagation이 일어날 때 가중치 w_{11}^1 과 w_{11}^2 의 조정된 값을 구해주세요. (학습률 $\eta = 0.5$)

두장 참고해주세요 !

수고하셨습니다.

W_{11}^1

$$\frac{\partial E_{\text{total}}}{\partial W_{11}^1} = \frac{\partial E_{\text{total}}}{\partial a_{11}} \cdot \frac{\partial a_{11}}{\partial z_{11}} \cdot \frac{\partial z_{11}}{\partial w_{11}^1}$$

$$\frac{\partial E_{\text{total}}}{\partial a_{11}} = \frac{\partial E_1}{\partial a_{11}} + \frac{\partial E_2}{\partial a_{11}}$$

$$\frac{\partial E_1}{\partial a_{11}} = \frac{\partial E_1}{\partial a_{21}} \cdot \frac{\partial a_{21}}{\partial z_{21}} \cdot \frac{\partial z_{21}}{\partial a_{11}} = -(0.5 - 0.6) \cdot 0.6 \cdot (1 - 0.6) \cdot 0.4 = 0.0096$$

$$\frac{\partial E_2}{\partial a_{11}} = \frac{\partial E_2}{\partial a_{22}} \cdot \frac{\partial a_{22}}{\partial z_{22}} \cdot \frac{\partial z_{22}}{\partial a_{11}} = -(0.9 - 0.512) \cdot 0.512 \cdot (1 - 0.512) \cdot 0.5 = -0.048$$

$$\frac{\partial a_{11}}{\partial z_{11}} = a_{11} \cdot (1 - a_{11}) = 0.51 \cdot (1 - 0.51) = 0.249$$

$$\frac{\partial z_{11}}{\partial w_{11}^1} = x_1 = 0.5$$

$$\therefore W_{11}^1 = -0.004 \quad , \quad W_{11}^{1+} = 0.1 - 0.5(-0.004) \\ = 0.02$$

W_{11}^2

$$\frac{\partial E_{\text{total}}}{\partial W_{11}^2} = \frac{\partial E_{\text{total}}}{\partial a_{21}} \cdot \frac{\partial a_{21}}{\partial z_{21}} \cdot \frac{\partial z_{21}}{\partial w_{11}^2}$$

$$= -(0.5 - 0.6) \cdot 0.6 \cdot (1 - 0.6) \cdot 0.5 / \\ = 0.012$$

* 참고자료

오차역전파

: 인공적인 오차가 나왔다면 이를 퍼지아웃으로 표기해보자
그 값을 반영해서 업데이트를 한다. 그런데 이렇게 단계별로 업데이트를
계산량은 너무 크다

→ 오차역전파는 미분의 chain rule에 근거하여 미분값을 두번에서
앞으로 전파시키면서 부등한 분자와 denominator의 곱의 차분률을 구한다.

$\frac{\partial L}{\partial y}$: 끝에 대한 Loss의 변화량,
 $\frac{\partial y}{\partial x}$: Loss로부터 출력들은 그때마다

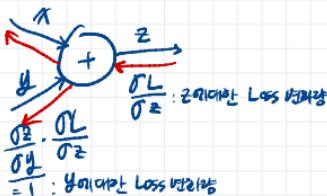
$$\frac{\partial L}{\partial x} = \frac{\partial L}{\partial y} \cdot \frac{\partial y}{\partial x}$$

↳ 2단계 미분

① 덧셈노드

$$z = x + y$$

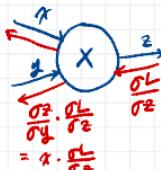
$$\frac{\partial z}{\partial x} = 1, \quad \frac{\partial z}{\partial y} = 1$$



② 곱셈노드

$$z = xy$$

$$\frac{\partial z}{\partial x} = y, \quad \frac{\partial z}{\partial y} = x$$

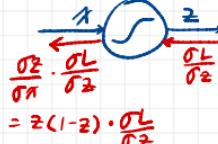


∴ 곱셈노드에서는 위치를 두어주어서 서로의 값을 공유하면 된다.

③ Sigmoid 노드

$$z = \frac{1}{1 + e^{-x}}$$

$$z' = z(1-z)$$

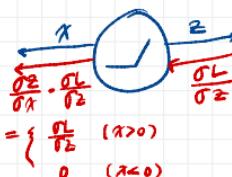


∴ 시그모이드 노드에서는 시그모이드 함수의 미분값을 공유해준다.

④ ReLU 노드

$$z = \begin{cases} x & x > 0 \\ 0 & x \leq 0 \end{cases}$$

$$\frac{\partial z}{\partial x} = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0 \end{cases}$$



∴ ReLU 노드에서도 ReLU 함수의 미분값을 공유해준다.

∴ 덧셈노드에서는 그대로 앞으로 전해주면 된다.