

딥러닝 세미나 2주차
김동협

3강 : 신경망
DSP Lab, Inha University
August , 2021



목차

- 활성화함수
- Step function을 쓰지 않는 이유
- Sigmoid, ReLU
- How



-활성화함수

- 정의 #

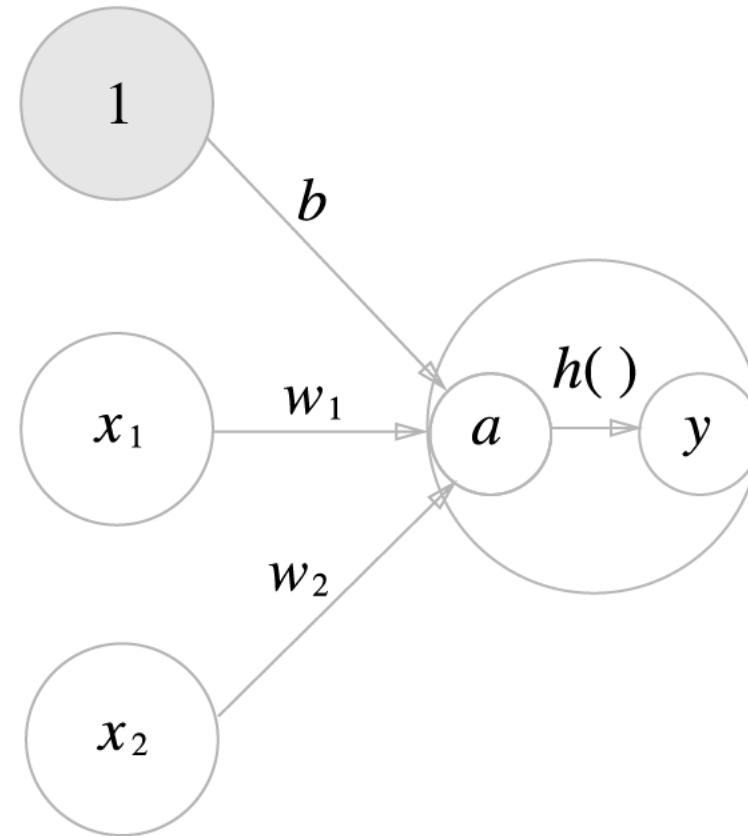
- 활성화 함수(Activation function)이란 입력된 데이터의 **가중 합**을 **출력 신호로 변환**하는 함수이다. 인공 신경망에서 이전 레이어에 대한 가중 합의 **크기에 따라 활성 여부가 결정**된다. 신경망의 목적에 따라, 혹은 레이어의 역할에 따라 선택적으로 적용한다.



-How

$$a = b + w_1x_1 + w_2x_2$$

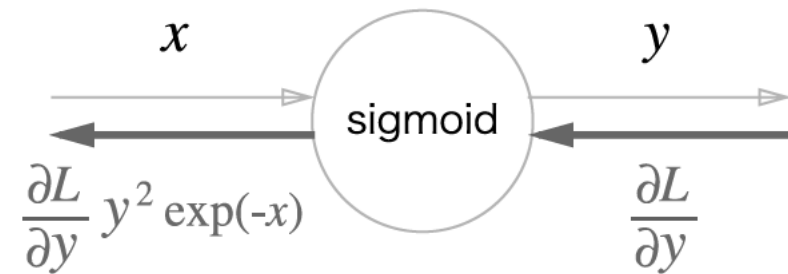
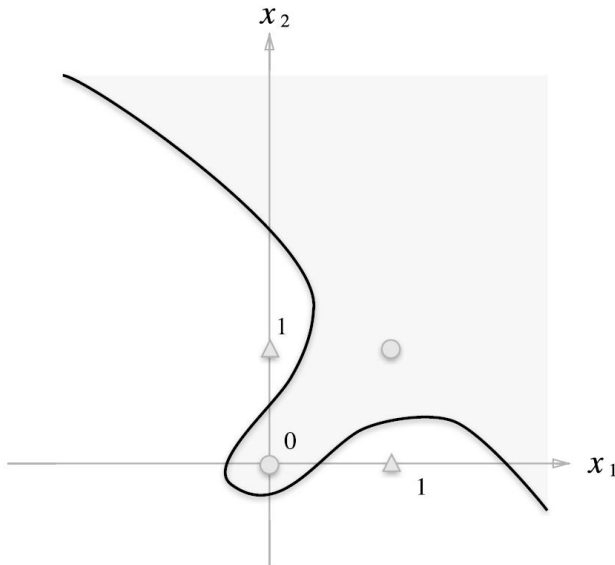
$$y = h(a)$$



-Step function을 쓰지 않는 이유

- Step ,Sigmoid, ReLU

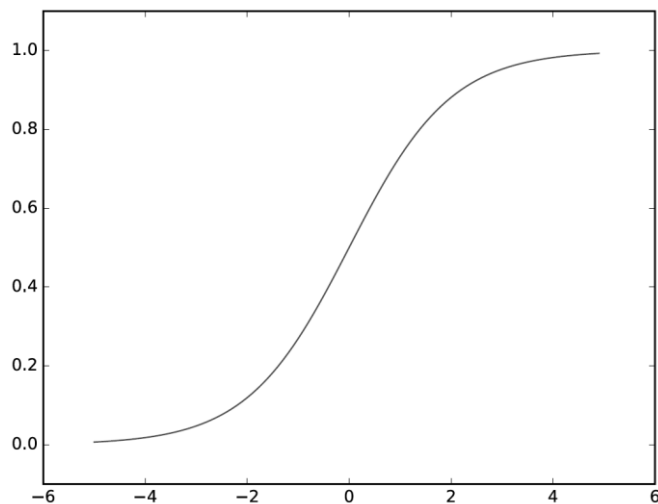
Step Function의 한계 : XOR 해결 불가, 미분 문제 -> D/L 의미X



-Sigmoid, ReLU

- Sigmoid

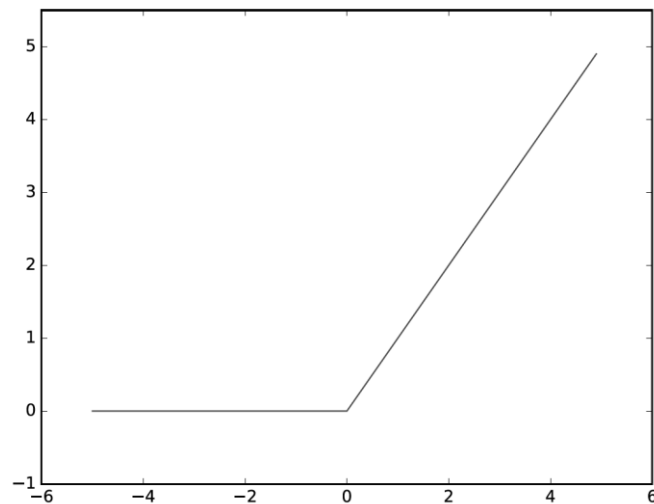
$$h(x) = \frac{1}{1 + \exp(-x)}$$



지수 함수적 특성을 이용해
계단 함수의 단점은 극복하고 장점은 살린다

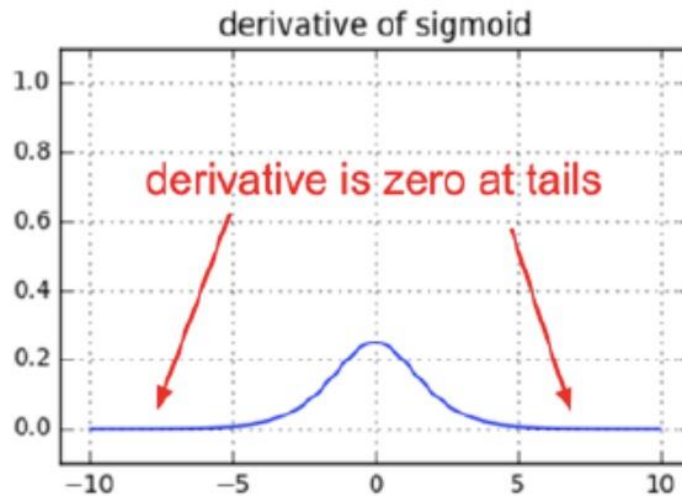
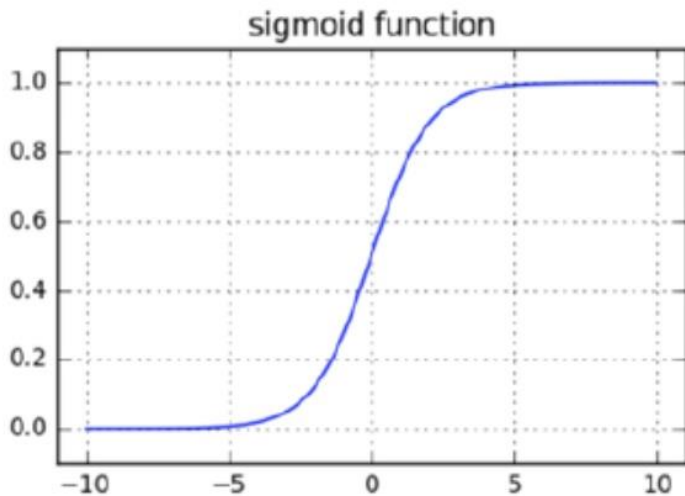
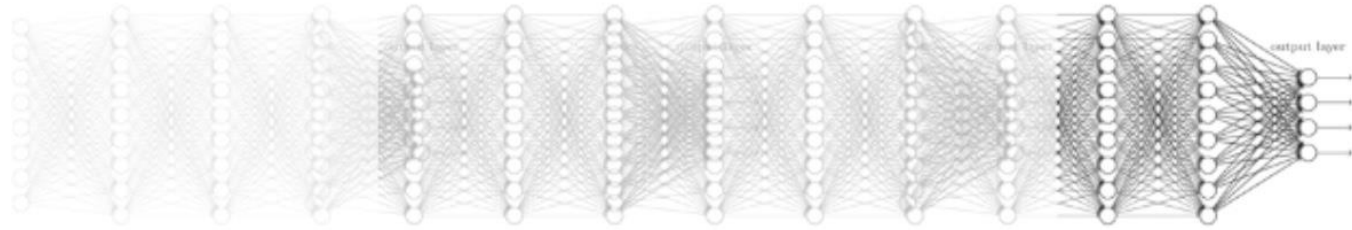
- ReLU

$$h(x) = \begin{cases} x & (x > 0) \\ 0 & (x \leq 0) \end{cases}$$



-Sigmoid의 한계점

- 1. Gradient Vanishing



$$\frac{1}{1+e^{-\alpha}} \left[1 - \frac{1}{1+e^{-\alpha}} \right]$$



Sigmoid의 한계점

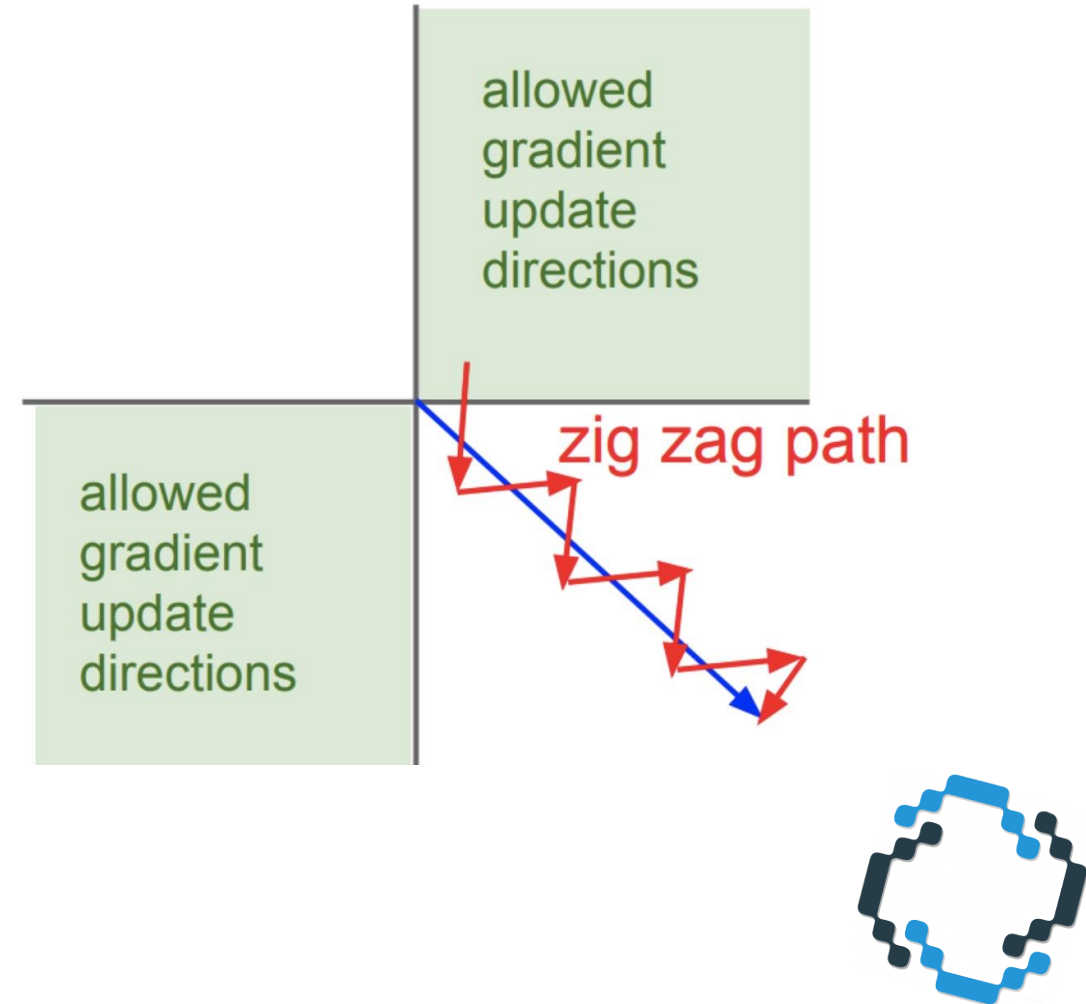
2. 중심값이 0이 아님

$$f = \sum w_i x_i + b$$

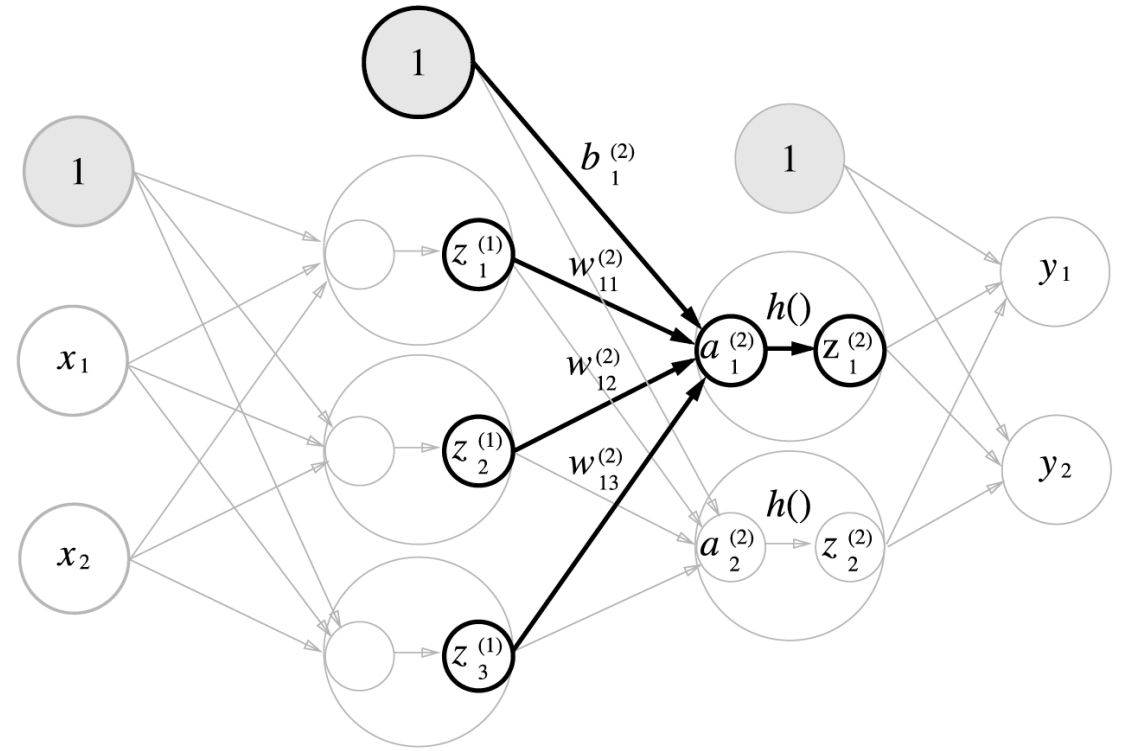
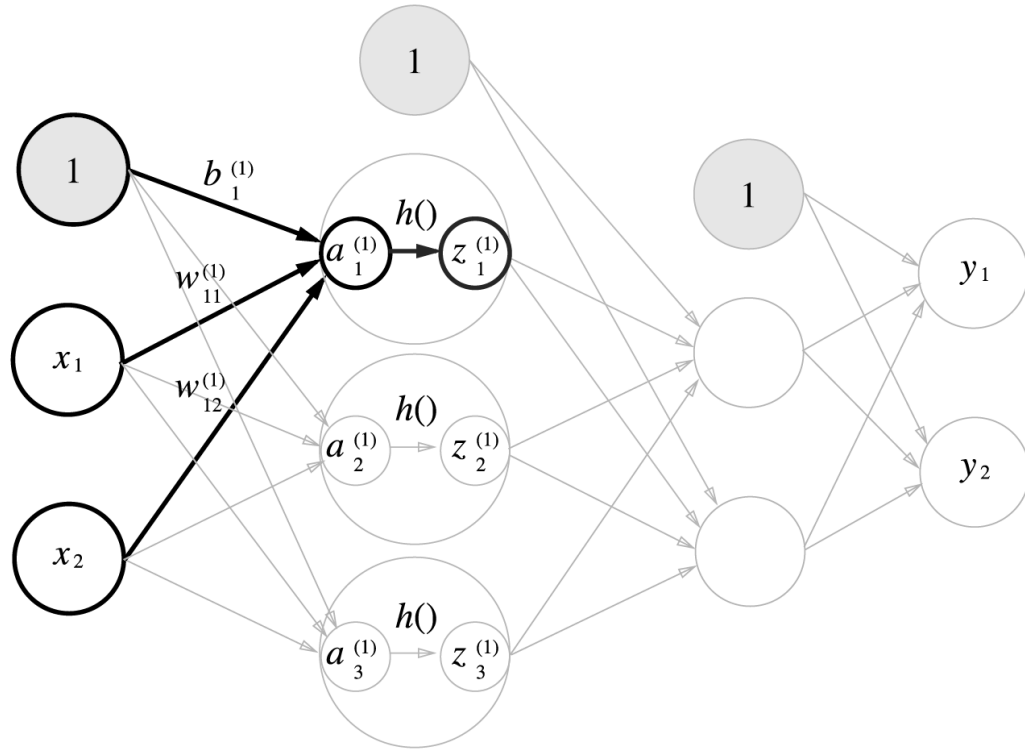
$$\frac{df}{dw_i} = x_i$$

$$\frac{dL}{dw_i} = \frac{dL}{df} \frac{df}{dw_i} = \frac{dL}{df} x_i$$

1. x_i (바로 앞의 sigmoid 출력) 는 무조건 양수
2. w 파라미터 변화량 대한 손실 함수의 변화량 (dL/dw_i) 가 무조건 dL/df 의 부호와 같음.
3. 따라서 f 와 w 의 증,감 관계가 4사분면을 봤을 때 오른쪽 그림과 같음



-실제 신경망에선?



```
a1 = np.dot(x, W1) + b1  
z1 = sigmoid(a1)  
a2 = np.dot(z1, W2) + b2  
y = softmax(a2)
```



참고자료, 링크

- <http://www.incodom.kr/>활성화함수
- <https://github.com/WegraLee/deep-learning-from-scratch-3/blob/2ac26c51a0329ae6dbef55abc646bb4a97e45ef7/%EB%B0%91%EB%B0%94%EB%8B%A53%20%EA%B7%B8%EB%A6%BC%EA%B3%BC%20%EC%88%98%EC%8B%9D.zip>
- 밑바닥부터 시작하는 딥러닝 1
- <https://astralworld58.tistory.com/62>
- <https://brunch.co.kr/@chris-song/39>

