

3.2 활성화 함수

활성화 함수

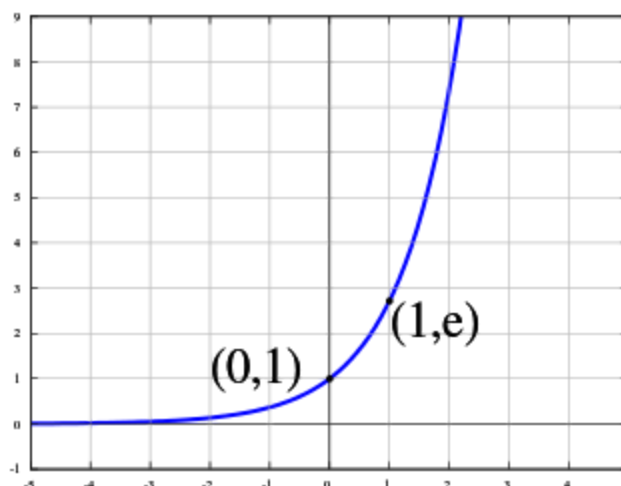
도입

- 노드에 들어오는 값들에 대해 바로 다음 레이어로 전달하지 않고 특정 함수를 통과시킨 후 전달한다.
- 이 함수를 Activation Function 이라고 부른다.
 - 선형함수
 - 비선형 함수
- 선형함수는 거의 사용하지 않으며 비 선형함수를 사용한다.
 - 선형함수는 의미 없기 때문
 - $f(x) = cx$ 가 레이어를 지날때마다 어떻게 변할까?

참고

지수함수(exponential function)

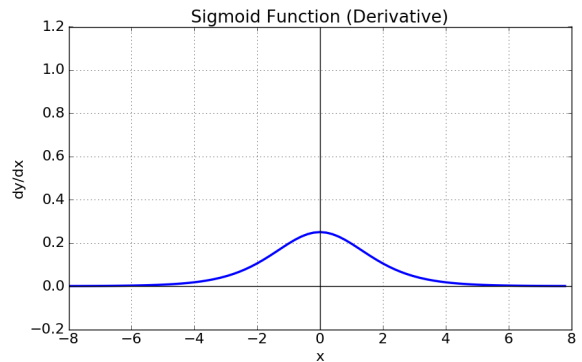
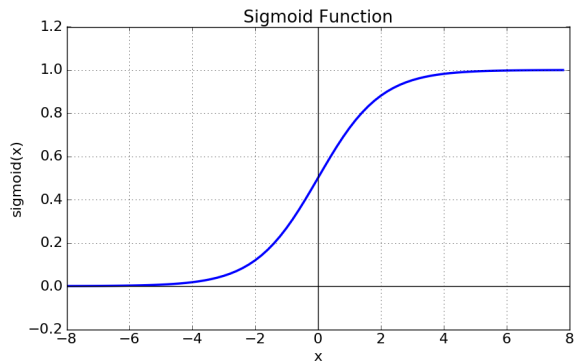
- 자연상수 $e = 2.718253\dots$
- $f(x) = e^x = \exp(x)$
-



Sigmoid

- Logistic Function 이라고도 불린다.

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

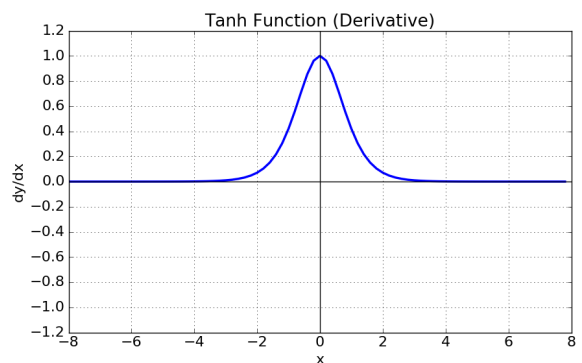
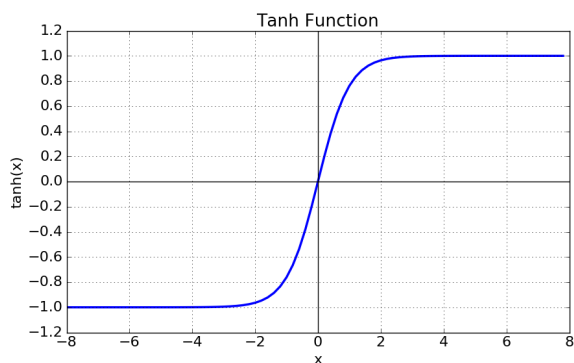


Sigmoid의 문제점

- 기울기 소실문제 발생
- 함수의 중심이 0이 아니다.
 - 학습이 느려질수 있음 (가중치가 Update 되는 과정에서 발생함)

Hyperbolic tangent

- $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \tanh(x) = 2\sigma(2x) - 1$
- $\tanh'(x) = 1 - \tanh^2(x)$



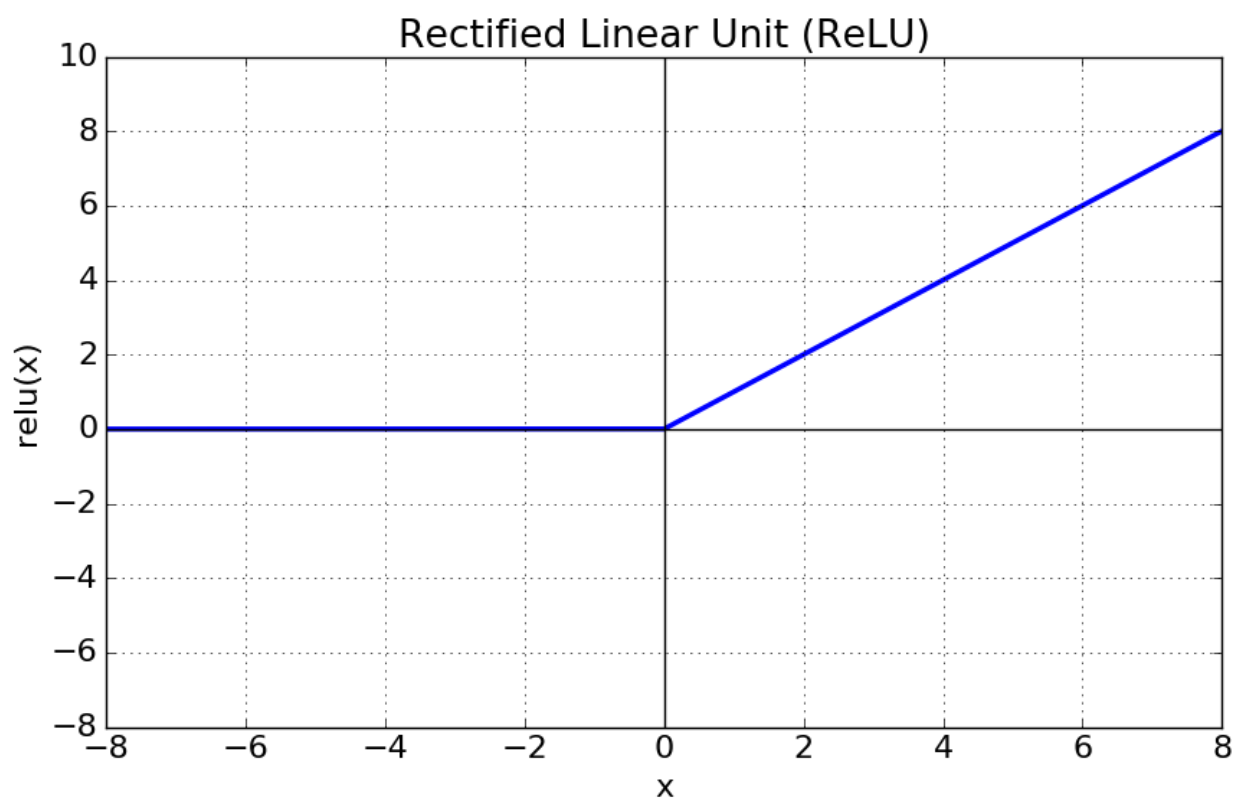
- 함수의 중심 0 → 학습속도 개선

tanh 의 문제점

- 기울기 소실문제 해결x

ReLU 함수 (Rectified Linear Unit)

- $f(x) = \max(0, x)$
- $f'(x) = 0 \ (x < 0), 1 \ (0 < x)$



- 학습의 속도가 빠르다.
- 연산속도가 빠르다.






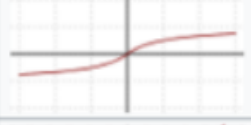

ReLU의 문제점

- $x < 0$ 에 대해서 기울기가 0이므로 학습시 뉴런이 죽어버린다.

Leakly ReLU

- $f(x) = \max(0.01x, x)$
- $f'(x) = 0.01, (x < 0), 1 \ (0 < x)$
- ReLU의 Dying Nuron 문제를 해결

정리

Identity		$f(x) = x$
Binary step		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Logistic (a.k.a. Soft step)		$f(x) = \frac{1}{1 + e^{-x}}$
TanH		$f(x) = \tanh(x) = \frac{2}{1 + e^{-2x}} - 1$
ArcTan		$f(x) = \tan^{-1}(x)$
Softsign ^{[7][8]}		$f(x) = \frac{x}{1 + x }$
Rectified linear unit (ReLU) ^[9]		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$