

Part. 02

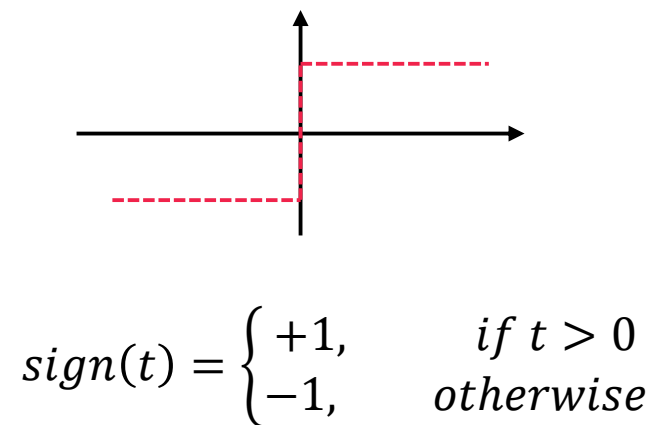
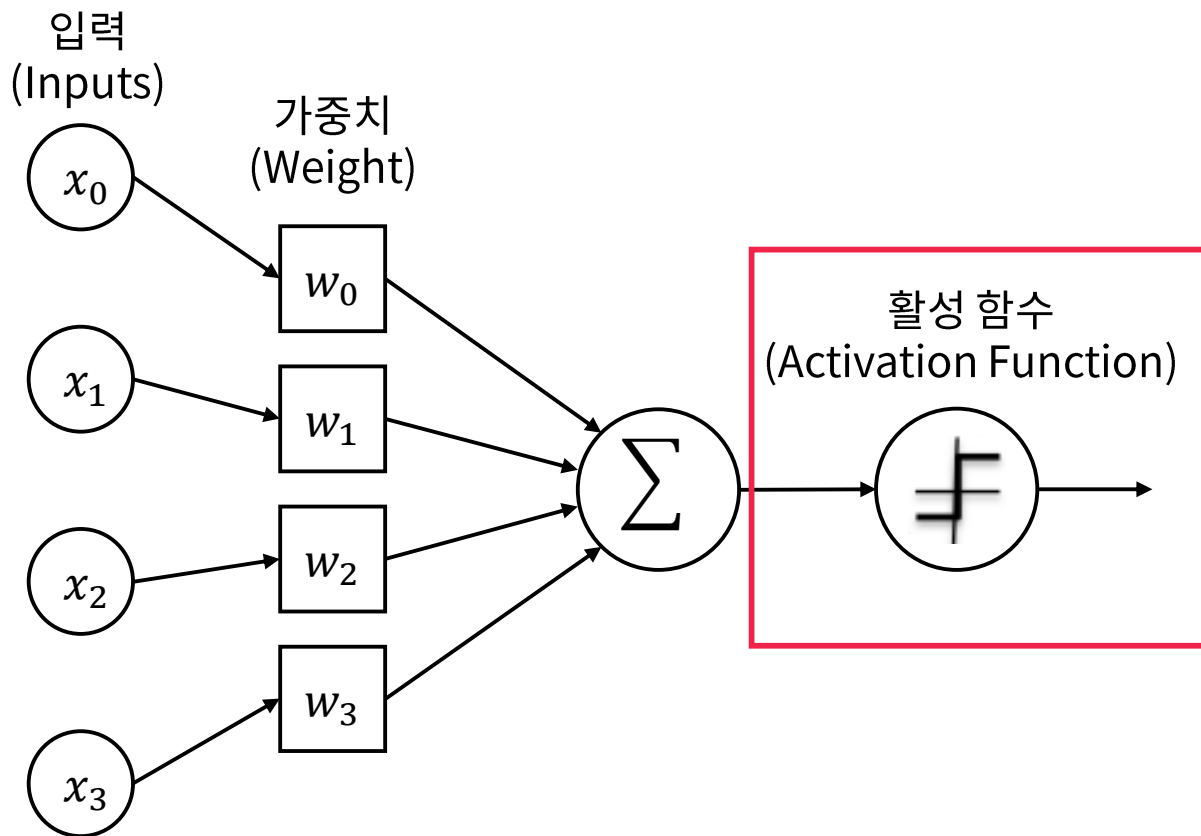
Neural Network Basics

| 활성 함수

FASTCAMPUS
ONLINE

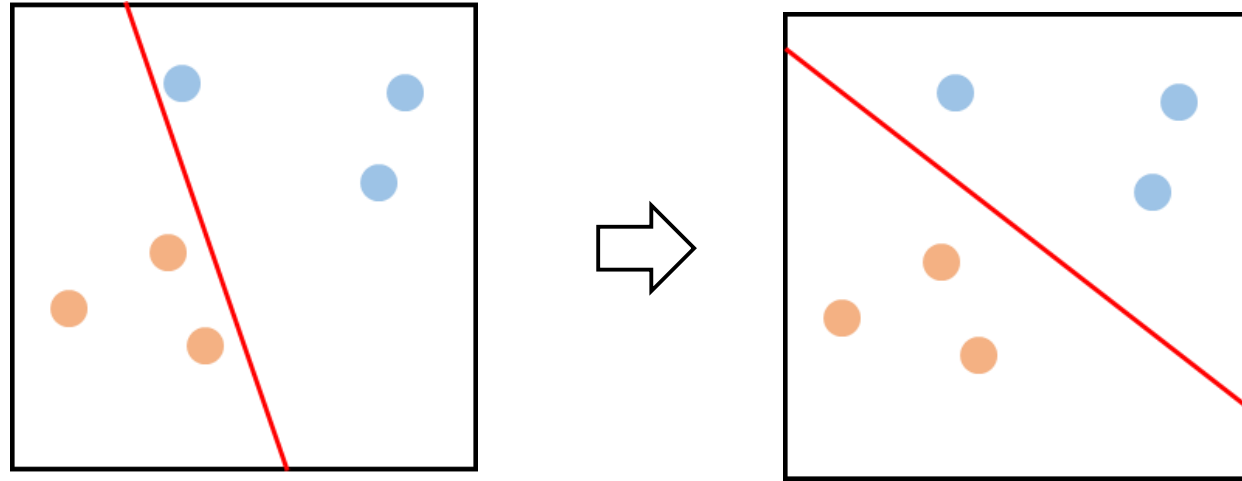
강사. 신제용

I 활성화 함수?



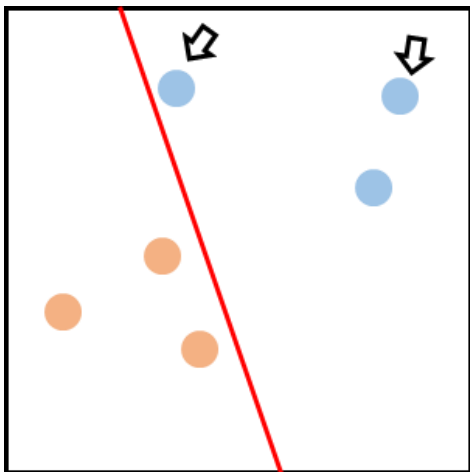
활성 함수는 퍼셉트론의 출력에 의미를 부여해 주며, 일반적으로 비선형 (Non-linear) 함수이다.

I 바람직한 결정 경계

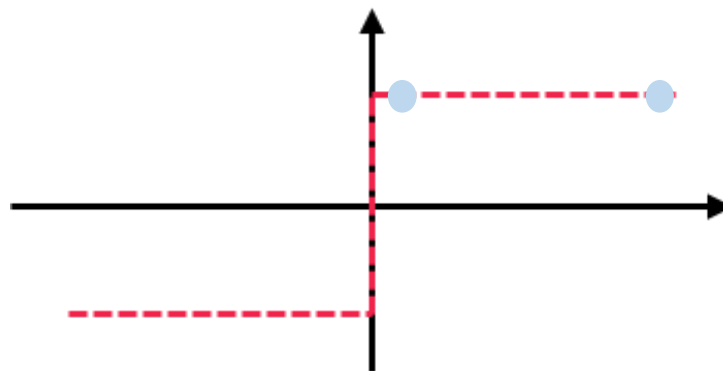


직관적으로 볼 때, 우측과 같은 결정 경계가 더 좋다고 볼 수 있다. 왜 그럴까?

I Sign 함수의 문제점

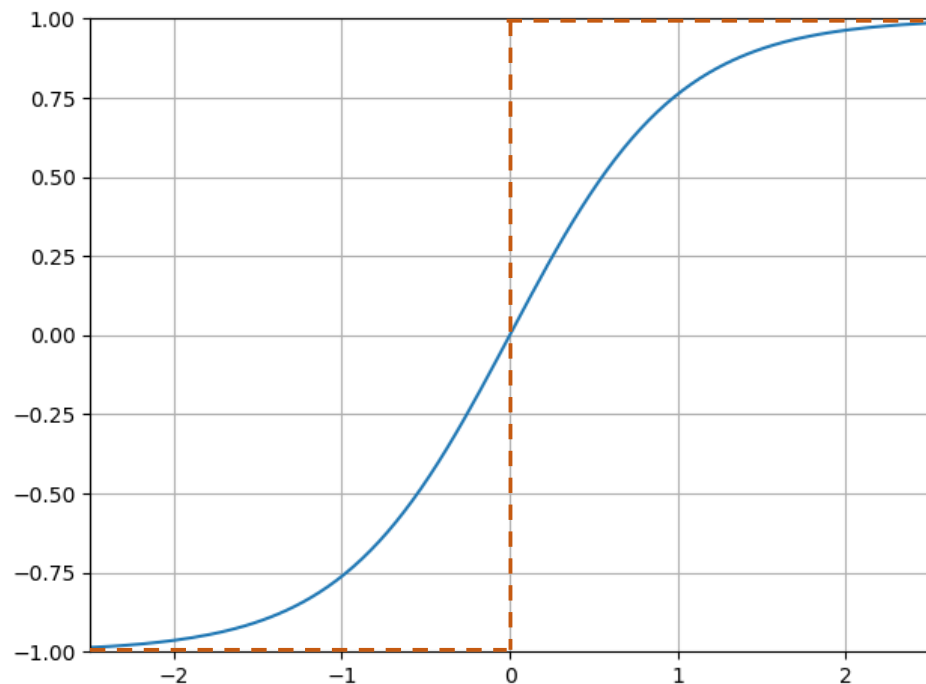


우측 점에 비해 좌측의 점은 지나치게
결정 경계 (Decision boundary)에 가깝다.



결정 경계로부터의 거리를 신경 쓰지 않은 탓!

I Tanh와 Soft decision

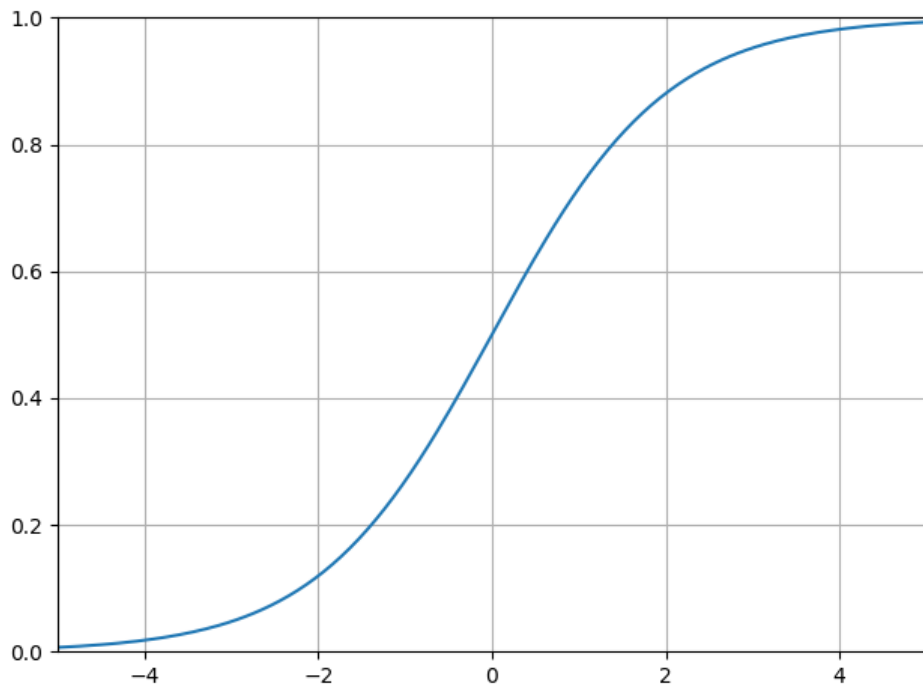


Tanh (Hyperbolic tangent) 함수

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

- 값이 작아질 수록 -1, 커질 수록 1에 수렴
- 모든 실수 입력 값에 대해 출력이 정의됨
- 출력이 실수값으로, Soft decision (cf. Hard decision)
- 입력 값이 0에 가까울 수록 출력이 빠르게 변함
- 모든 점에서 미분 가능

I Sigmoid function



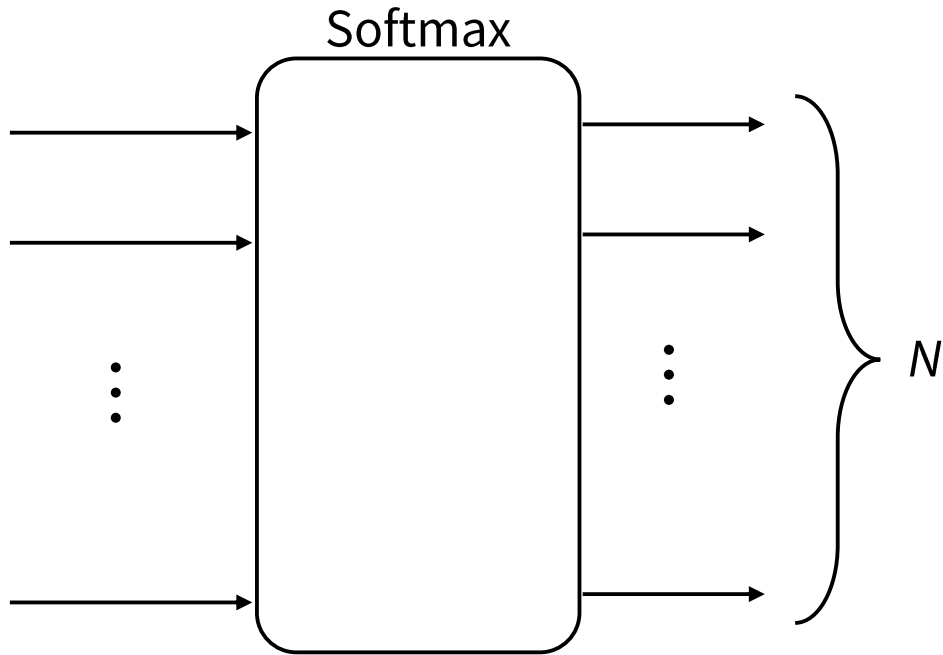
Sigmoid 함수

$$\text{sigmoid}(x) = \frac{1}{1 + e^{-x}}$$

- 값이 작아질 수록 0, 커질 수록 1에 수렴
- 모든 실수 입력 값에 대해 출력이 정의됨
- 출력이 0~1 사이로, '확률'을 표현할 수 있음
- 입력 값이 0에 가까울 수록 출력이 빠르게 변함
- 모든 점에서 미분 가능

Sigmoid 함수는 본래 S-Shape의 함수 모두를 가르키나, 딥러닝에서는 보통 위 함수로 통용된다.

I Softmax function

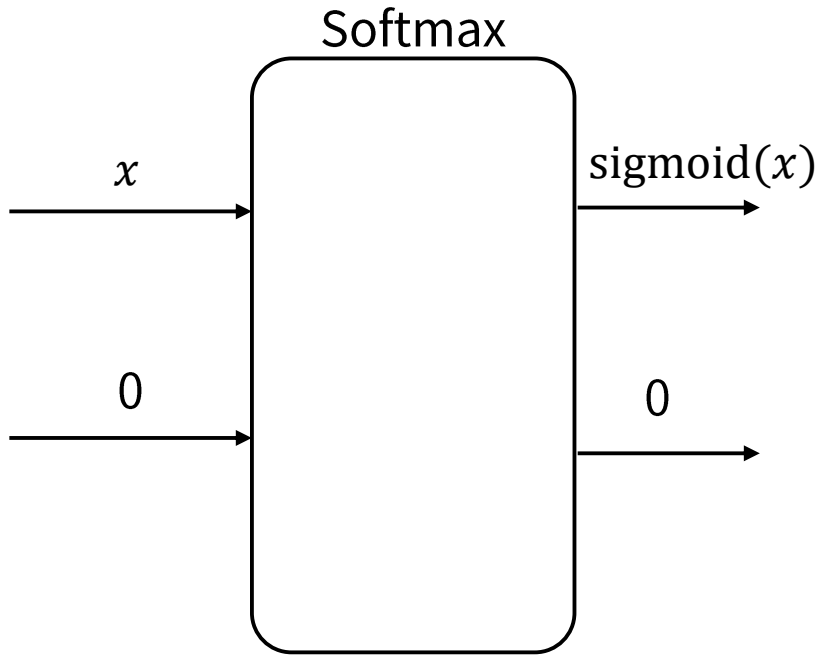


$$\text{softmax}(\mathbf{x})_i = \frac{e^{x_i}}{\sum_j e^{x_j}}$$

- 각 입력의 지수함수를 정규화한 것
- 각 출력은 0~1 사이의 값을 가짐
- 모든 출력의 합이 반드시 1이 됨
- 여러 경우의 수 중 한가지에 속할 ‘확률’을 표현

Softmax는 최종 출력 단에서 N 가지 범주로 분류하는 Multi-class classification에 쓰임

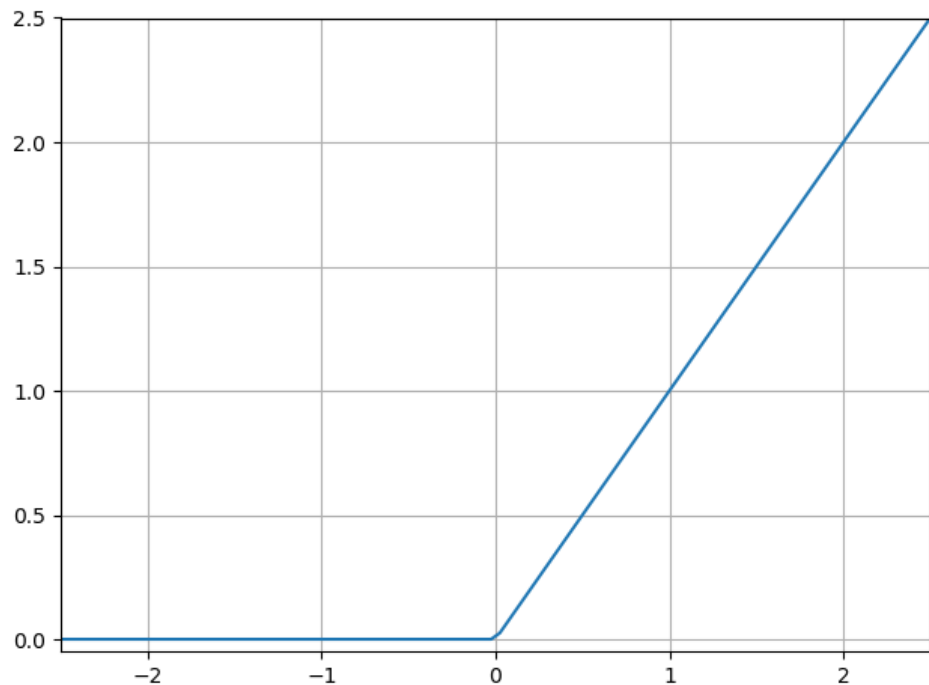
I Softmax vs. Sigmoid



$$\text{softmax}([x, 0])_0 = \frac{e^x}{e^x + e^0} = \frac{1}{1 + e^{-x}} \quad \text{sigmoid}(x)$$

Sigmoid는 하나의 입력을 0으로 강제한 2-Class Softmax 함수와 동일하다.
2가지 클래스를 구분하기 위해 1개의 입력을 받는다는 점에 주목.

I ReLU function



ReLU (Rectified linear unit) 함수

$$\text{ReLU}(x) = \max(0, x)$$

- 0보다 작은 값을 0으로 강제하는 함수
- 딥러닝에서 가장 많이 사용되는 활성화 함수
- 미분 값이 일정 (0 또는 1)해서 학습이 잘 되는 특성
- 단순한 구현으로 매우 빠른 연산이 가능

ReLU는 현재 쓰임을 자세히 알긴 어려우나, 지속적으로 다루어집니다.