

PyTorch로 딥러닝 제대로 배우기

- 중급편 -

Part3-1. 심층 신경망

강사: 김 동 희



목차

I. 인공신경망의 기본 원리

- 1) 신경망 모델
- 2) 딥러닝 모델 구조
- 3) 역전파

II. Activation Functions

- 1) 딥러닝 모델 구조
- 2) 활성화 함수

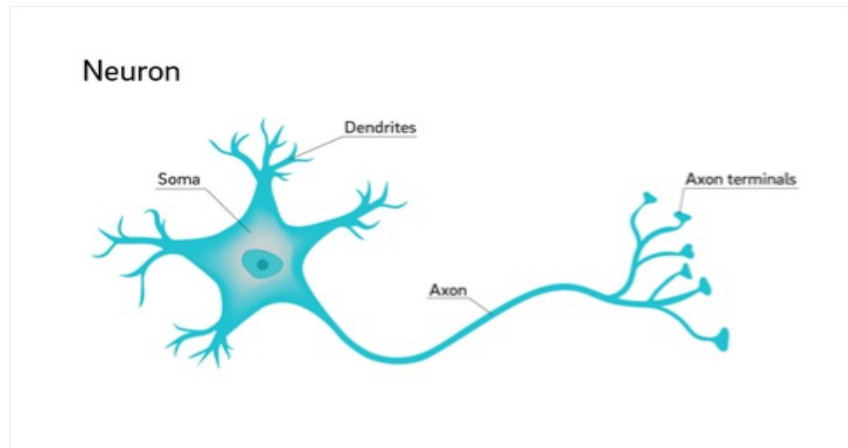


I. 인공신경망의 기본 원리

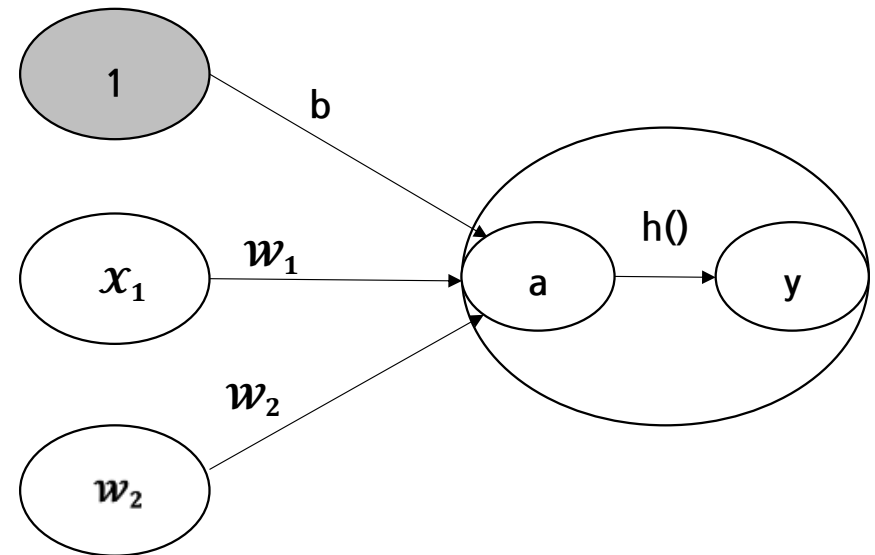
1. 신경망 모델(Model)

□ 신경망 모델 학습 원리

- 인간의 뇌 구조를 모방하여 만든 모델
- 정보가 입력되었을 때, 정보를 얼마나 어떻게 전달할지 결정
- 여러 개의 뉴런들이 상호 연결되어 동작
- 입력 값, 활성화 함수, 출력 값으로 구성
- 가중치와 편향에 따라 활성화 여부 결정



생물에서의 Neuron

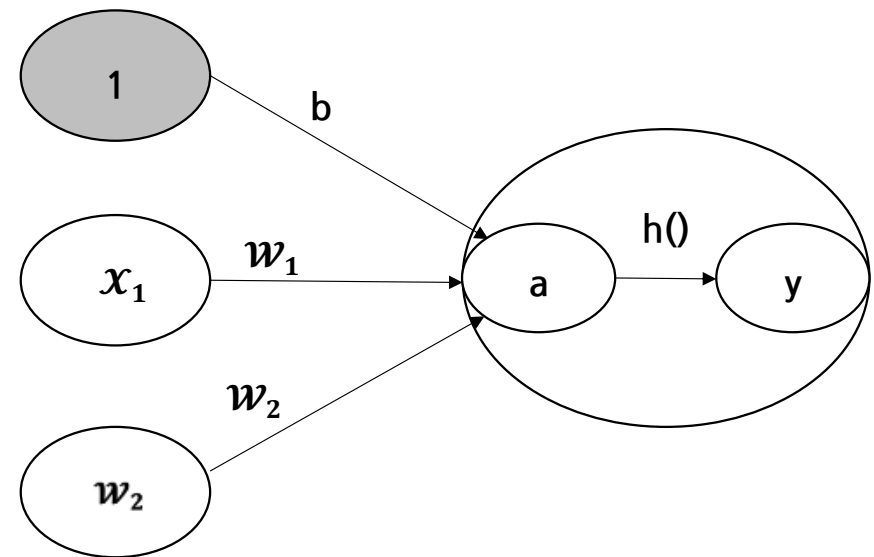


인공신경망에서 Neuron

1. 신경망 모델(Model)

□ 가중치와 편향

- $y = \begin{cases} 0 & (b + w_1x_1 + w_2x_2) \leq 0 \\ 1 & (b + w_1x_1 + w_2x_2) > 0 \end{cases}$
- 가중치: 신호 전달의 정보를 나타냄
- 편향: '뉴런이 얼마나 쉽게 활성화 되느냐'를 결정

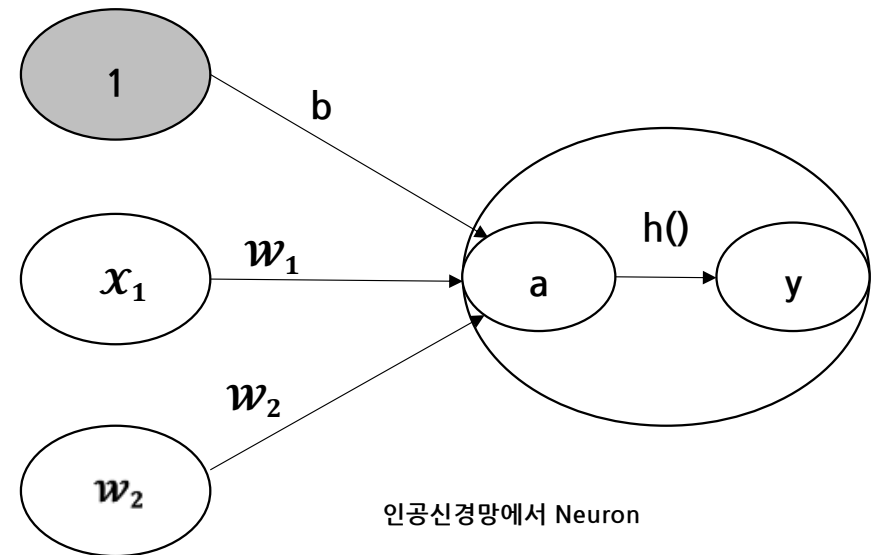


인공신경망에서 Neuron

1. 신경망 모델(Model)

□ 활성화 함수

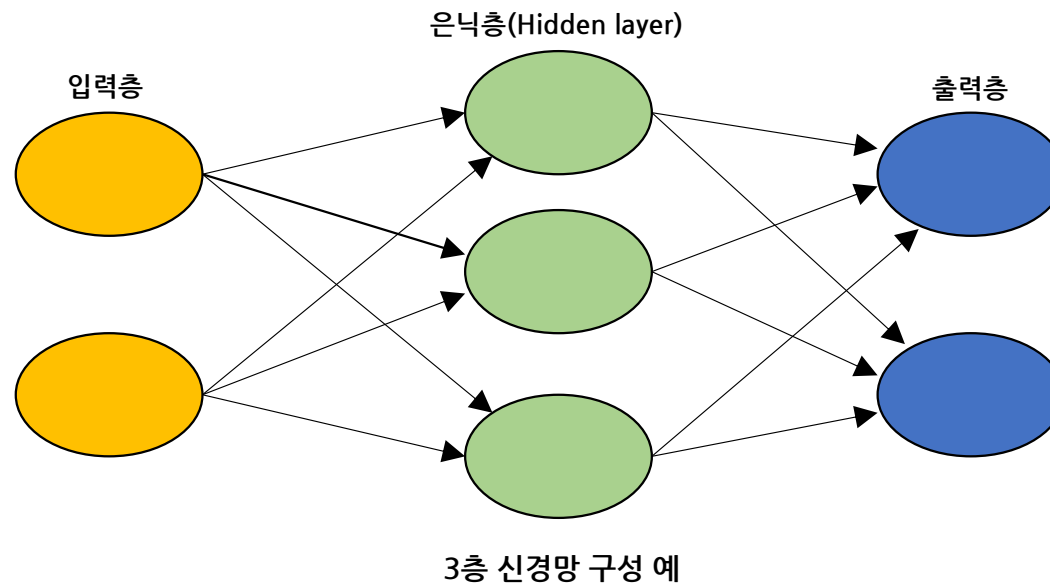
- $y = h(b + w_1x_1 + w_2x_2)$
- $h(x) = \begin{cases} 0(x \leq 0) \\ 1(x > 0) \end{cases}$
- 활성화 함수: 입력신호의 총합을 출력 신호로 변환하는 함수
- 임계 값을 경계로 출력하는 특징이 있음



2. 딥러닝 모델 구조

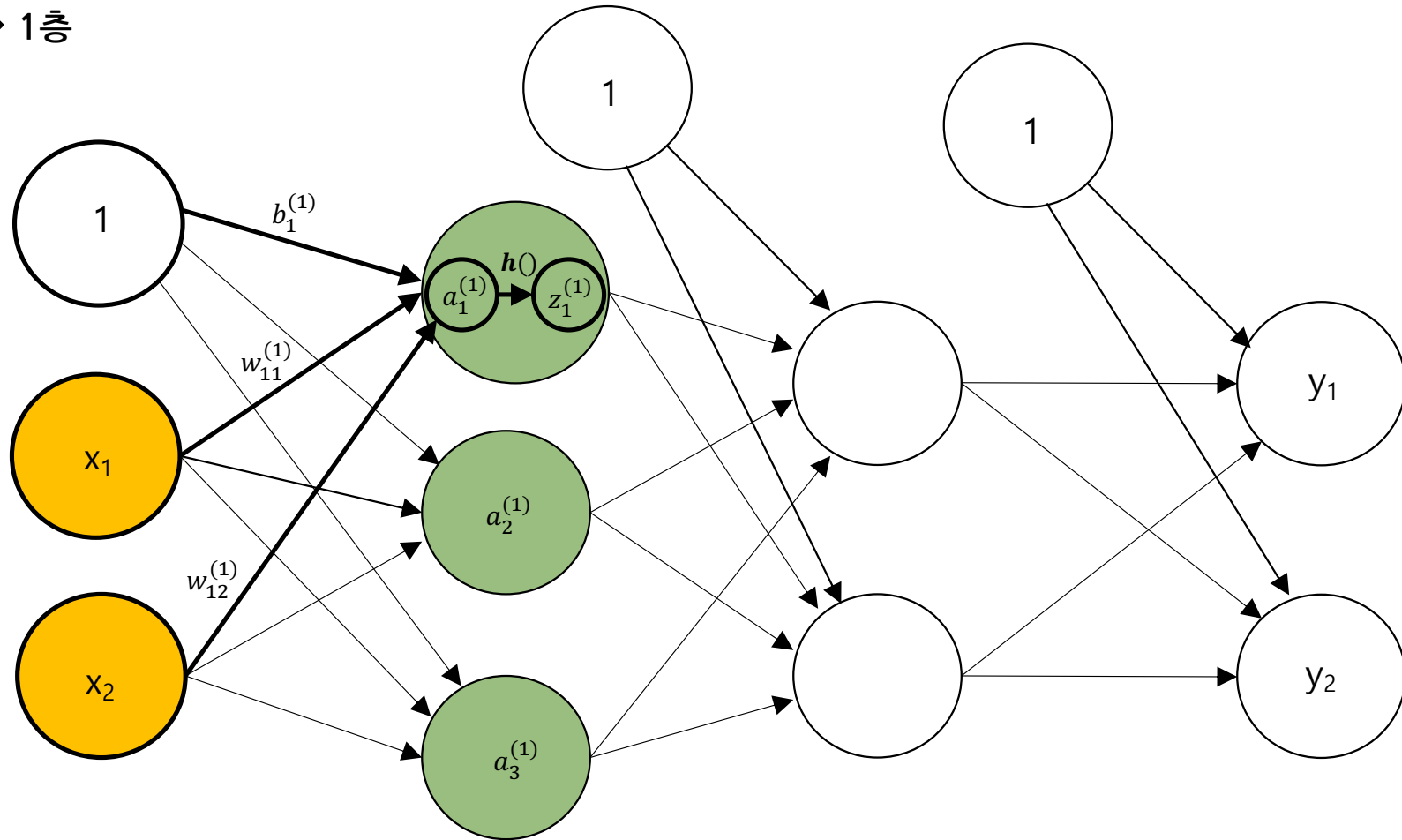
□ 다층 Perceptron

- 입력층: 데이터를 입력 받는 계층
- 은닉층: 학습을 진행하는 계층, 은닉층은 N개 이상으로 구성되며 정보를 담아냄
- 출력층: 입력된 데이터에 따라 결과를 출력하는 계층. Task에 따라 다르게 구성 가능



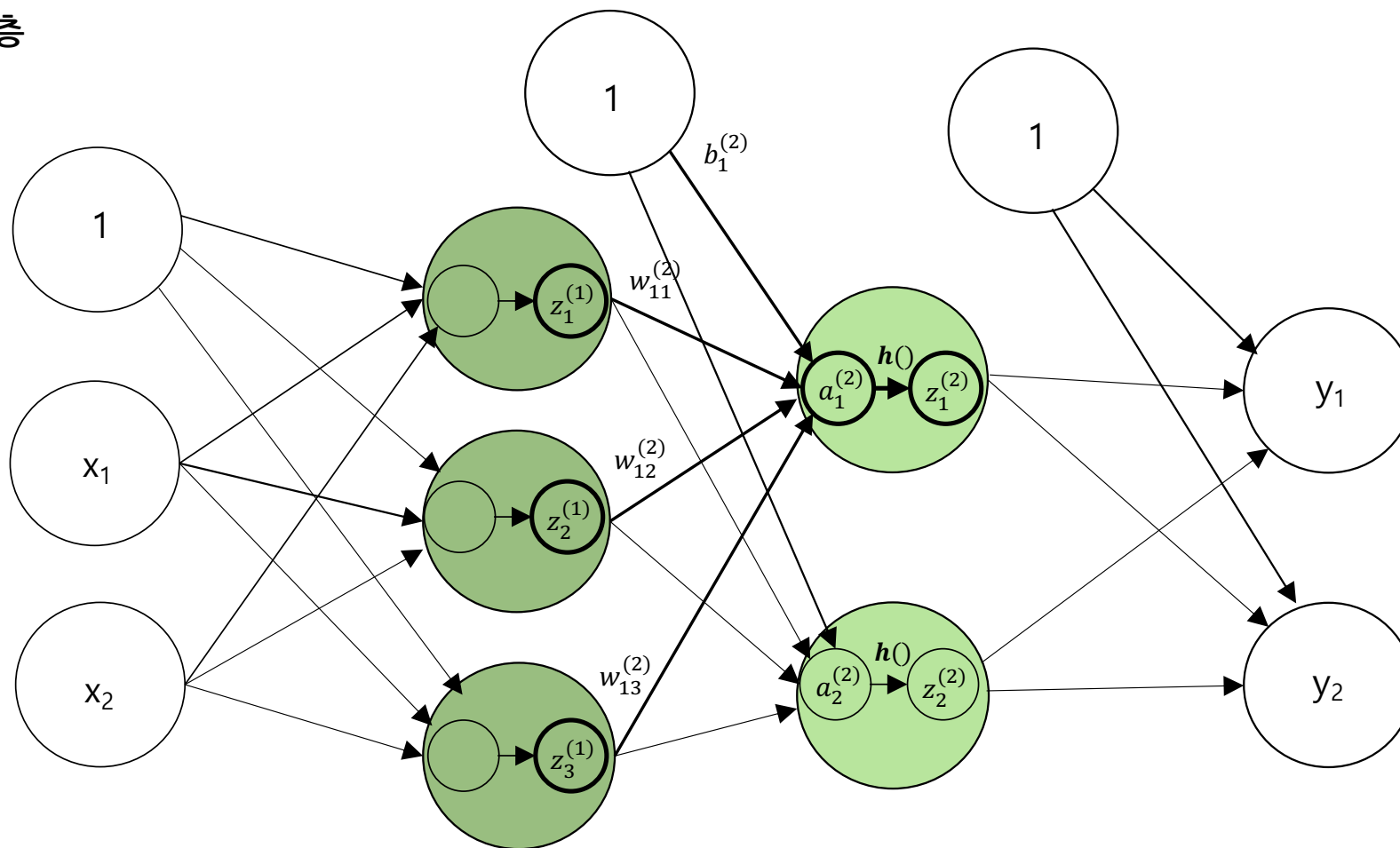
2. 딥러닝 모델 구조

□ 입력층 → 1층



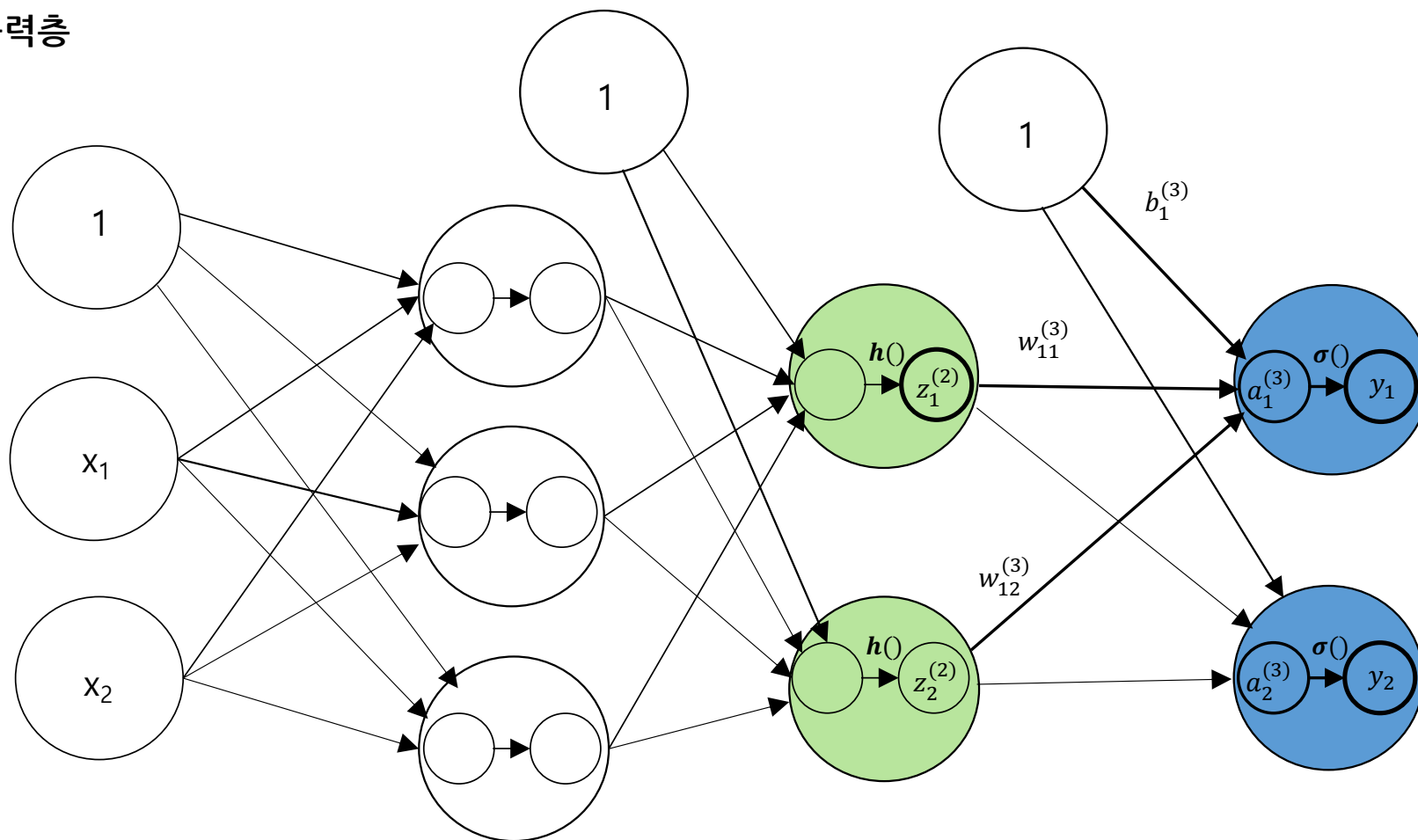
2. 딥러닝 모델 구조

□ 1층 → 2층



2. 딥러닝 모델 구조

□ 2층 → 출력층



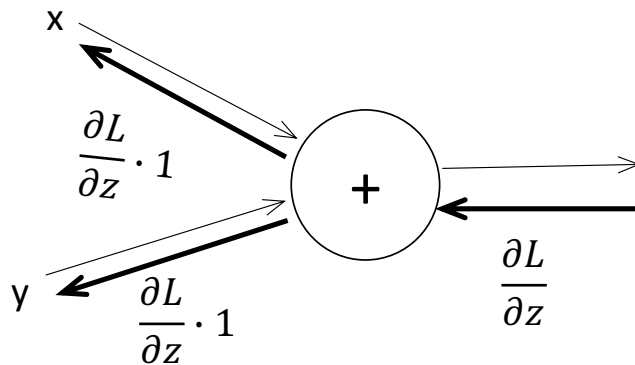
3. 역전파(Backpropagation)

□ 덧셈 노드의 역전파

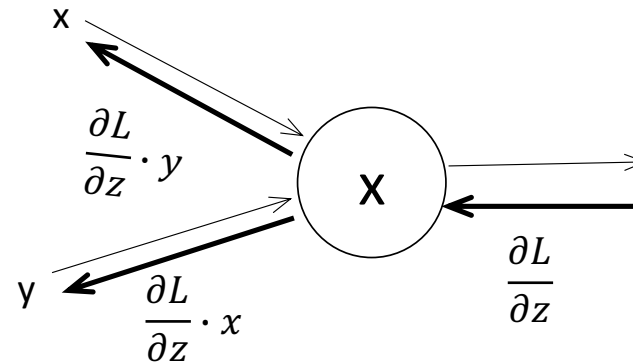
- 덧셈 노드의 역전파는 입력 값을 그대로 흘려보낸다.

□ 곱셈 노드의 역전파

- 곱셈 노드의 역전파는 입력 신호들을 서로 바꾼 값을 곱해서 하류로 보낸다.
- $z = xy, \frac{\partial z}{\partial x} = y, \frac{\partial z}{\partial y} = x$



덧셈 노드에서 역전파



곱셈 노드에서 역전파

3. 역전파(Backpropagation)

□ 연쇄법칙

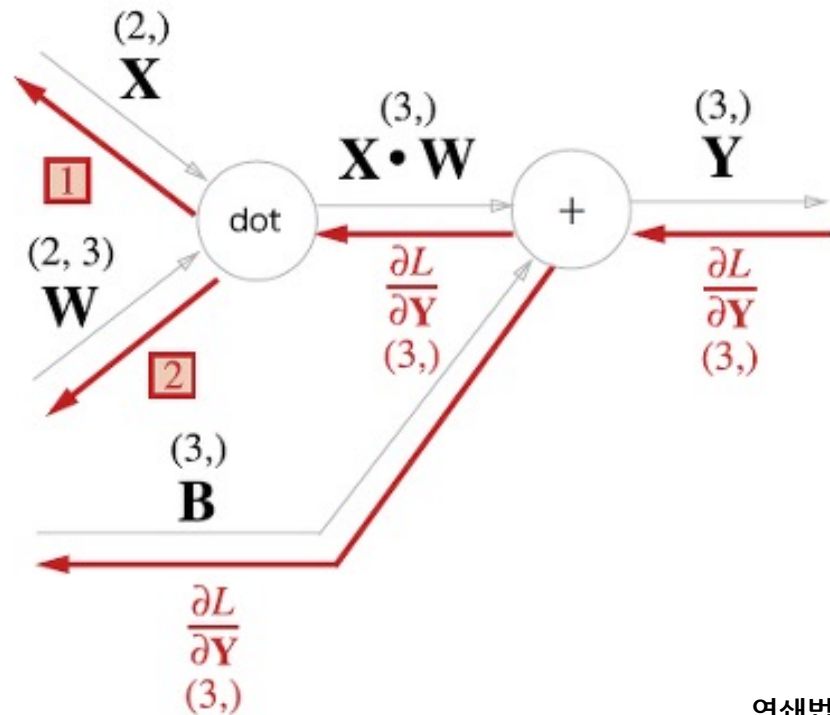
- 합성 함수의 미분은 합성 함수를 구성하는 각 함수의 미분의 곱으로 나타낼 수 있다.

$$\boxed{1} \quad \frac{\partial L}{\partial \mathbf{X}} = \frac{\partial L}{\partial \mathbf{Y}} \mathbf{W}^T$$

(2,) (3,) (3, 2)

$$\boxed{2} \quad \frac{\partial L}{\partial \mathbf{W}} = \mathbf{X}^T \frac{\partial L}{\partial \mathbf{Y}}$$

(2, 3) (2, 1) (1, 3)



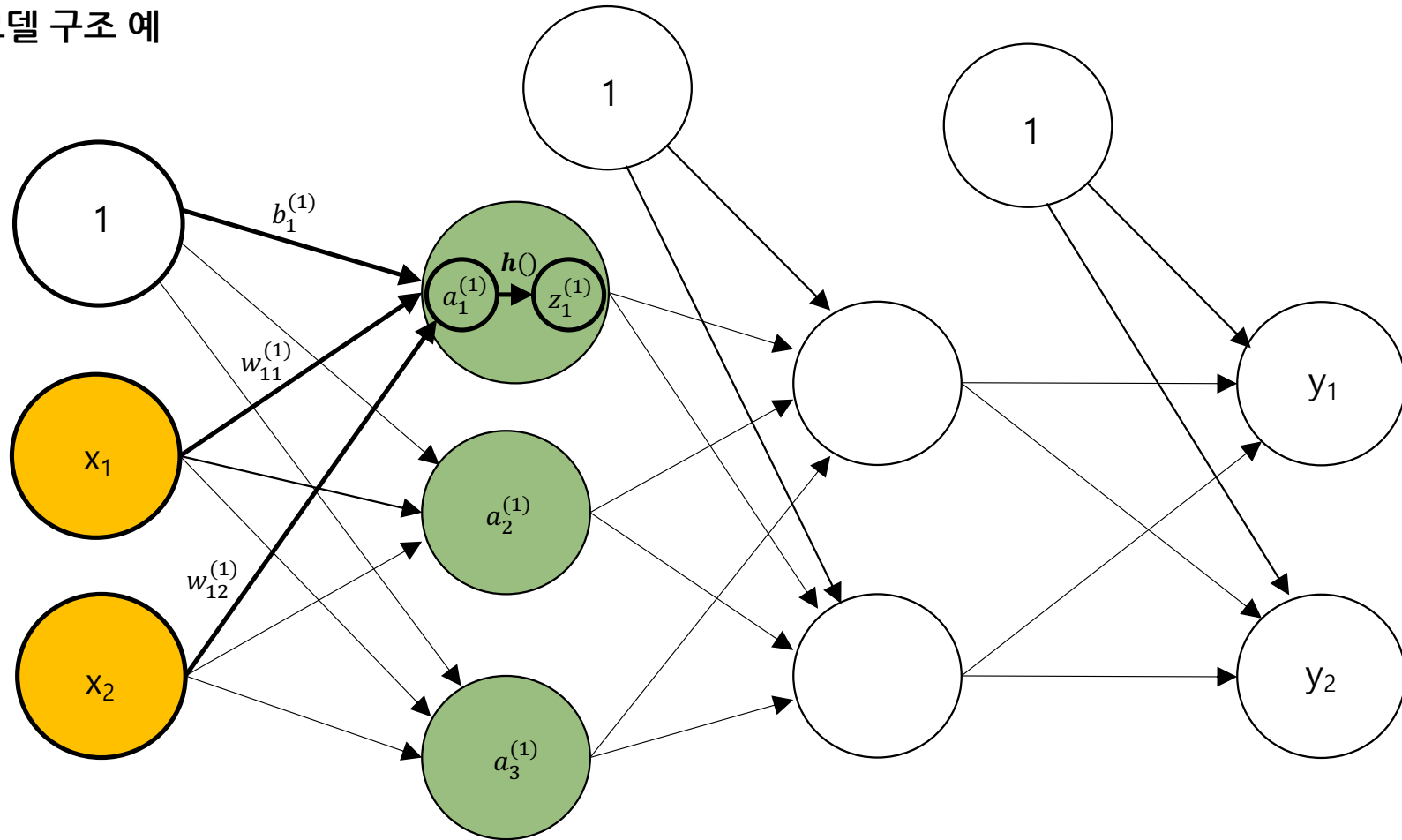
연쇄법칙



II. Activation Functions

1. 딥러닝 모델 구조

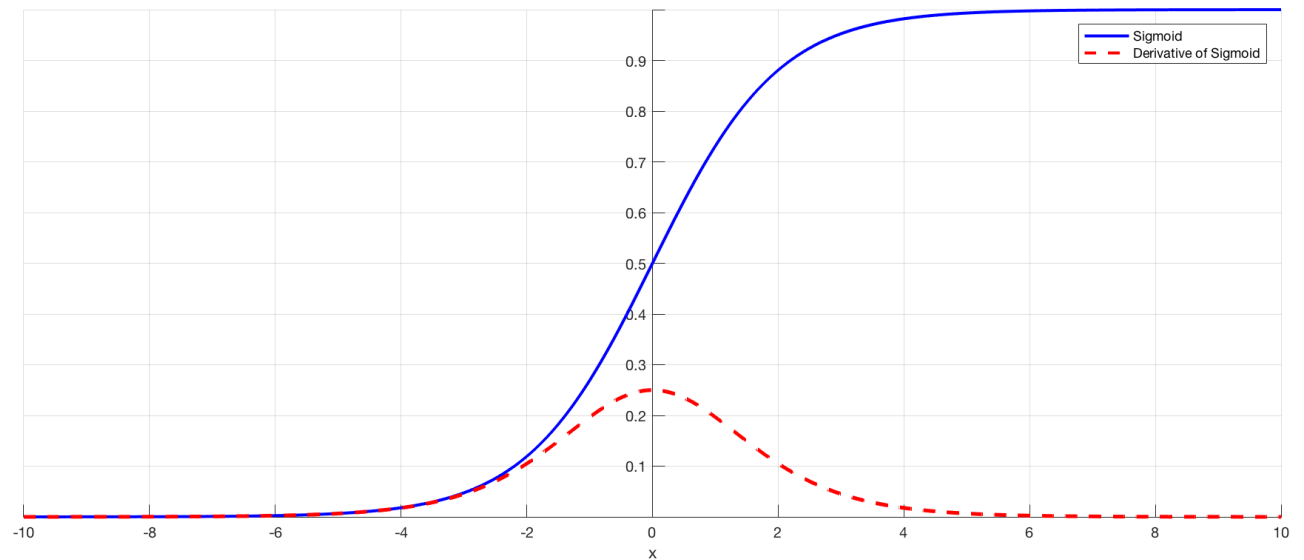
□ 딥러닝 모델 구조 예



2. 활성화 함수

□ Sigmoid 함수

- $h(x) = \frac{1}{1+\exp(-x)}$
- 자연 상수를 분모로하는 함수
- 0과 1사이 값을 가짐
- 미분한 값의 최대치가 0.25

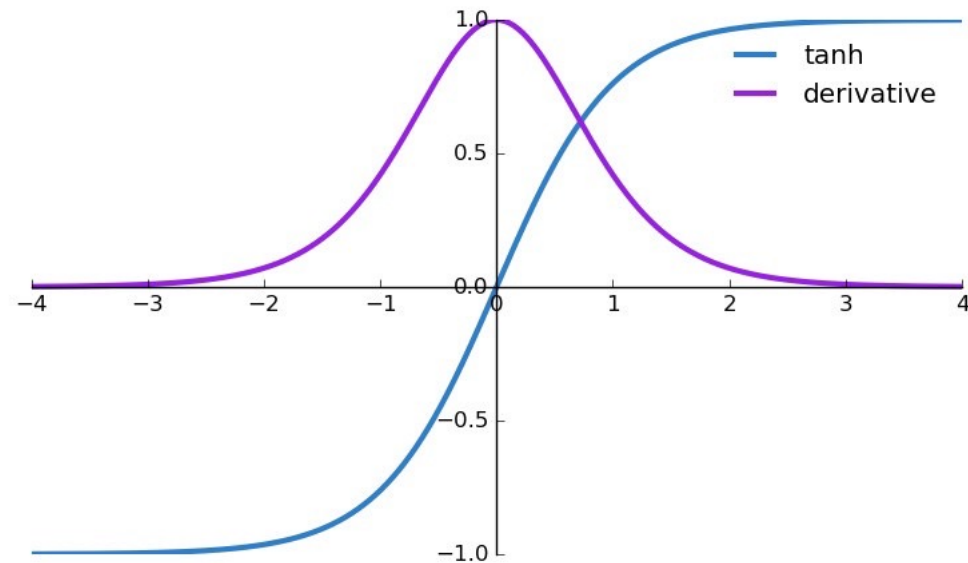


Sigmoid 함수와 미분값

2. 활성화 함수

□ tanh

- $\tanh(x) = \frac{\sinh(x)}{\cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}} = 2 \times \text{sigmoid}(2x) - 1$
- Sigmoid 함수를 rescale 한 함수
- 1과 -1사이 값을 가짐

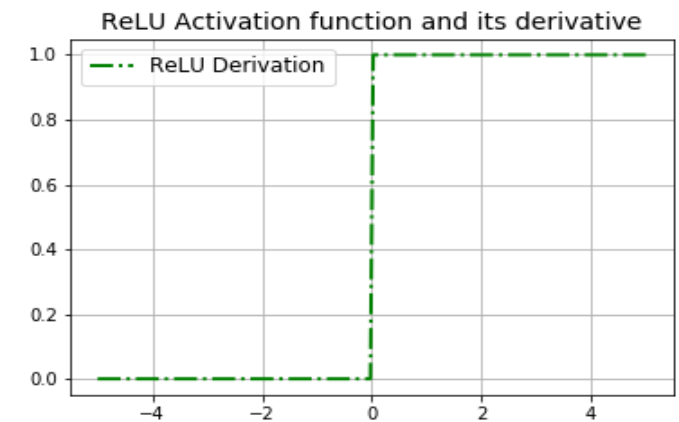
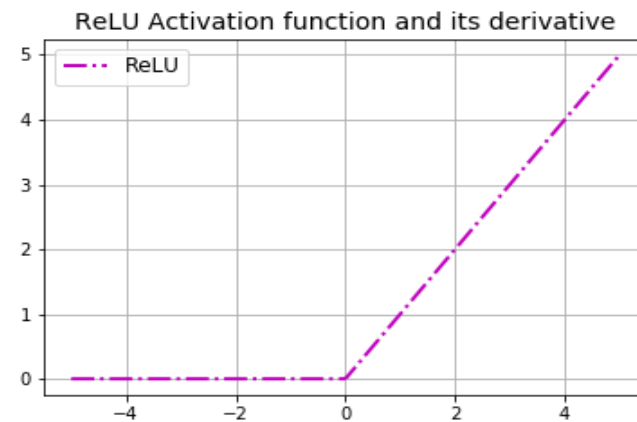


tanh 함수와 미분값

2. 활성화 함수

□ ReLU

- $ReLU(x) = \begin{cases} 0 & \text{if } (x < 0) \\ x & \text{if } (x \geq 0) \end{cases}$
- 극단적으로 값이 0보다 작으면 0, 0보다 크면 x값 그대로 출력
- 미분한 값의 최대치가 1또는 0
- 최근 딥러닝에서 가장 많이 사용



ReLU 함수와 미분값

감사합니다.