

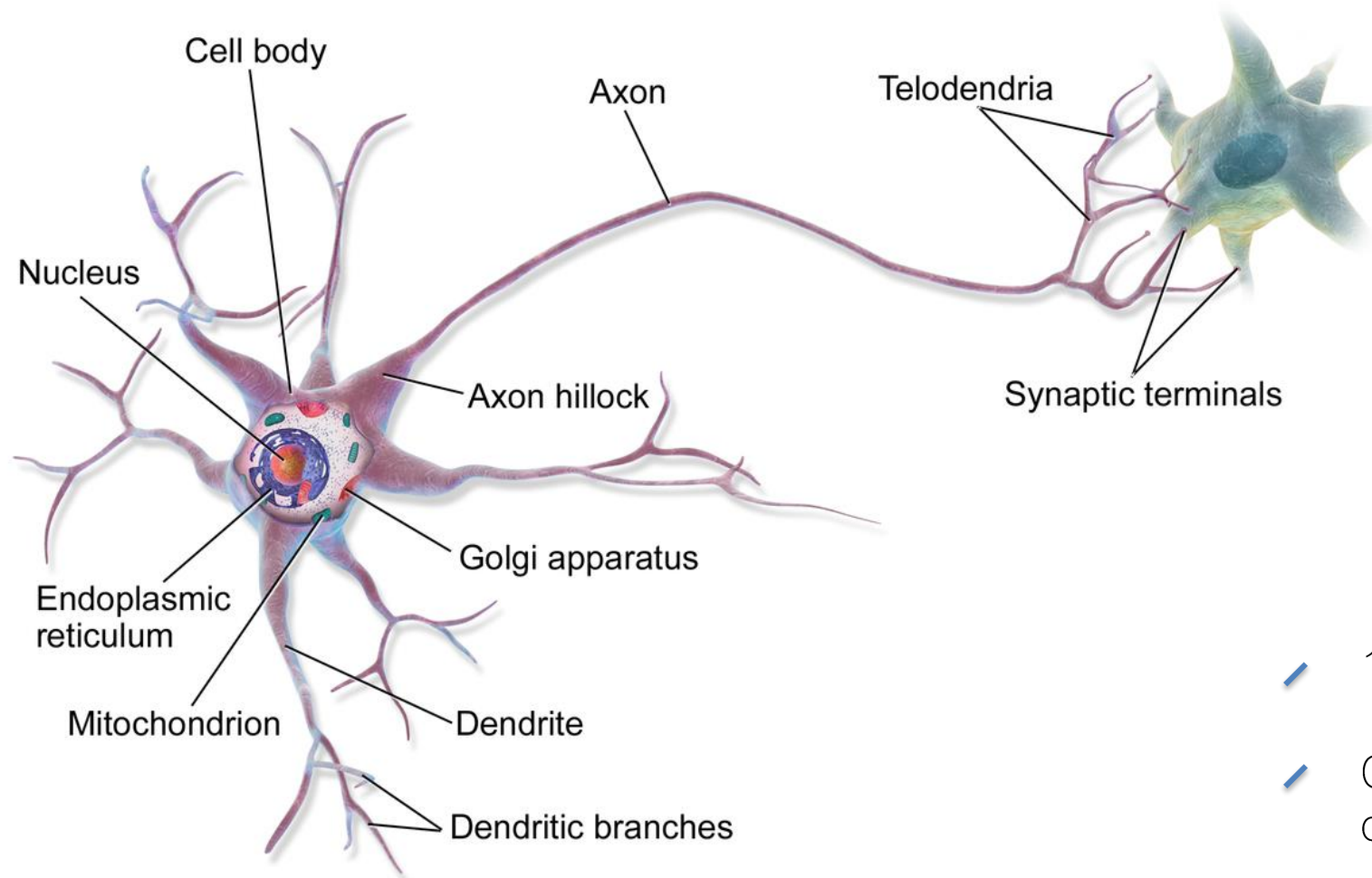
A blue circle with a textured, hand-drawn appearance, containing the letters 'ML' in a bold, black, sans-serif font.

ML

Neural Network

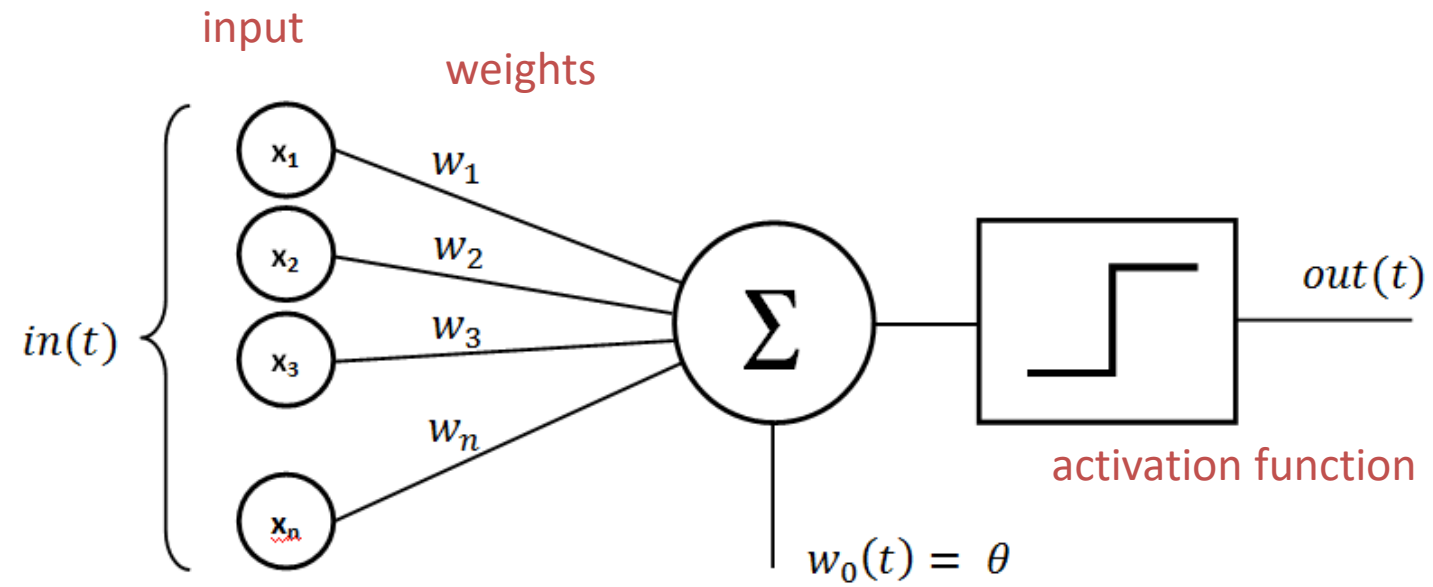
IDEA
LAB

Neuron



- 10¹¹ neurons in human brain
- One neuron is connected to 10³ other neurons

Perceptron



Decision boundary: $d(x) = W^T X + b$

일반적인 학습의 순서

1. 구조를 정의하고 분류 과정을 수학적식으로 표현. (매개변수 집합 θ 로 정의 됨) $\longrightarrow y = \tau(W^T X + b)$
2. 분류기의 품질을 측정할 수 있는 cost function $J(\theta)$ 를 정의 $\longrightarrow J(\theta) = \sum_{X_k \in Y} (-t_k)(W^T X_k + b)$
3. $J(\theta)$ 를 최소화 하는 θ 를 찾기 위한 알고리즘을 설계 \longrightarrow Gradient Descent method
$$\theta(h+1) = \theta(h) - \lambda \frac{\partial J(\theta)}{\partial \theta}$$

입력: 훈련 집합 $X = \{(\mathbf{x}_1, t_1), (\mathbf{x}_2, t_2), \dots, (\mathbf{x}_N, t_N)\}$, 학습률 ρ

출력: 퍼셉트론 가중치 \mathbf{w} , b

알고리즘:

1. \mathbf{w} 와 b 를 초기화한다.
2. **repeat** {
3. $Y = \emptyset$;
4. **for** ($i = 1$ **to** N) {
5. $y = \tau(\mathbf{w}^T \mathbf{x}_i + b)$; // (4.2)로 분류를 수행함
6. **if** ($y \neq t_i$) $Y = Y \cup \mathbf{x}_i$; // 오분류된 샘플 수집
7. }
8. $\mathbf{w} = \mathbf{w} + \rho \sum_{\mathbf{x}_k \in Y} t_k \mathbf{x}_k$; // (4.7)로 가중치 갱신
9. $b = b + \rho \sum_{\mathbf{x}_k \in Y} t_k$;
10. } **until** ($Y = \emptyset$);
11. \mathbf{w} 와 b 를 저장한다.

Perceptron | Training - example

$$a = (0,0)^T, t_a = -1$$

$$b = (1,0)^T, t_b = 1$$

$$c = (0,1)^T, t_c = 1$$

$$d = (1,1)^T, t_d = 1$$

$$\text{Initial } w(0) = (-0.5, 0.75)^T, b(0) = 0.375$$

$$\text{Learning rate} = 0.4$$

Multi-layer Perceptron

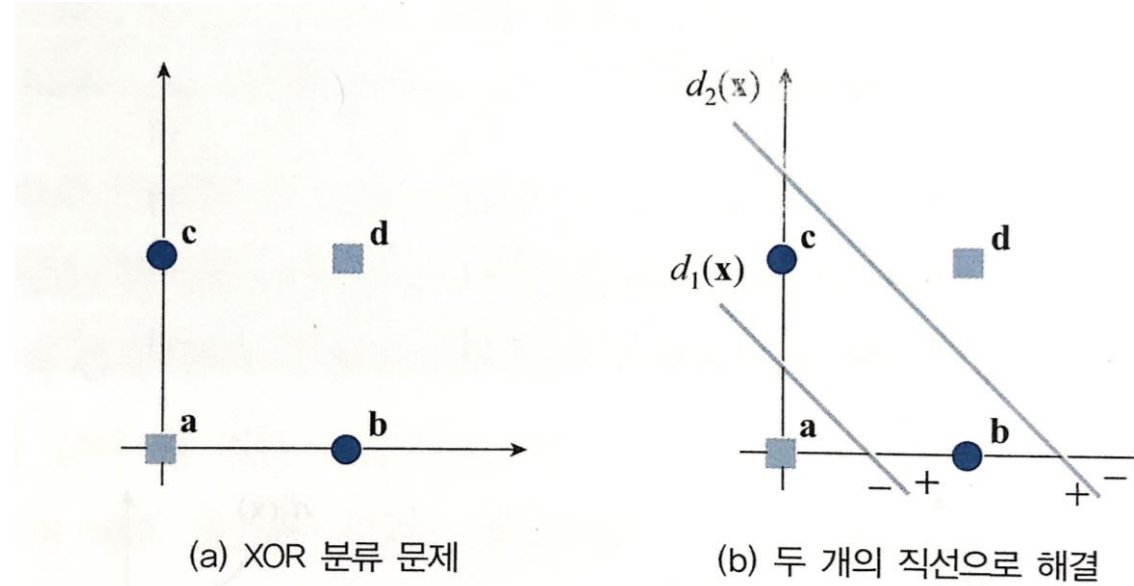


그림 4.6 XOR 분류 문제의 해결

$W_1^T X + b_1 > 0$ 이고, $W_2^T X + b_2 > 0$ 이면, $X \in \omega_1$

$W_1^T X + b_1 < 0$ 이거나, $W_2^T X + b_2 < 0$ 이면, $X \in \omega_2$

샘플	특징 벡터 (x)		첫 번째 단계	
	x_1	x_2	퍼셉트론1	퍼셉트론2
a	0	0	-1	+1
b	1	0	+1	+1
c	0	1	+1	+1
d	1	1	+1	-1

Multi-layer Perceptron

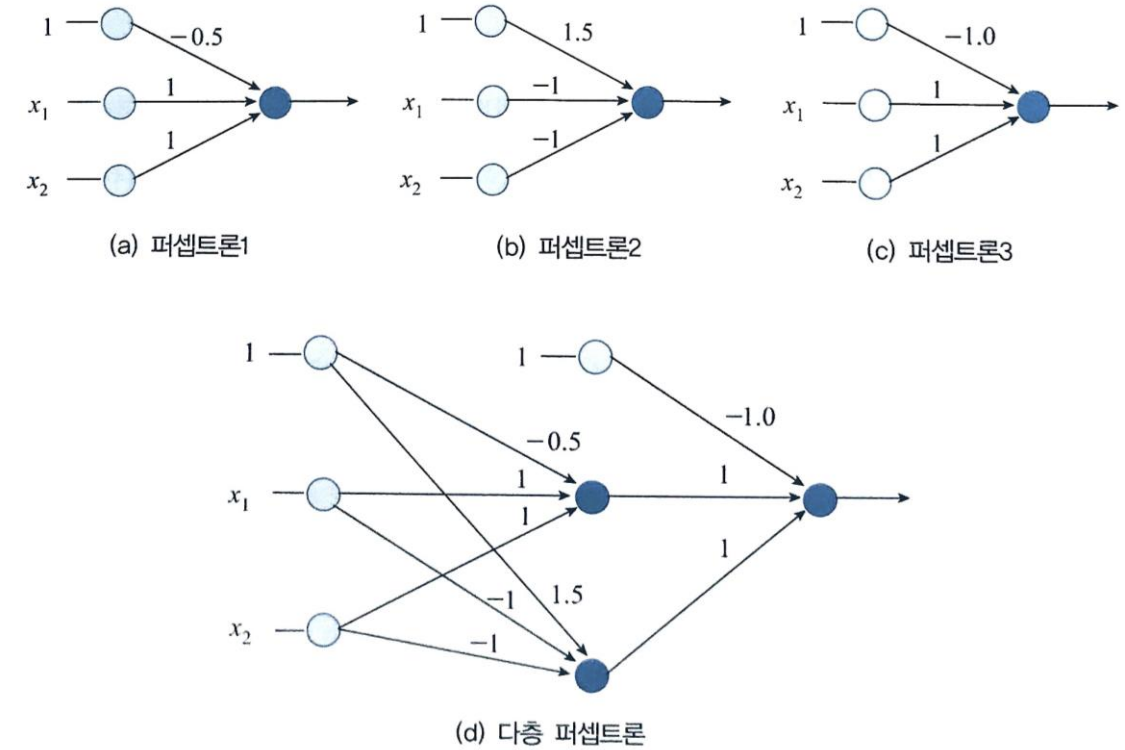
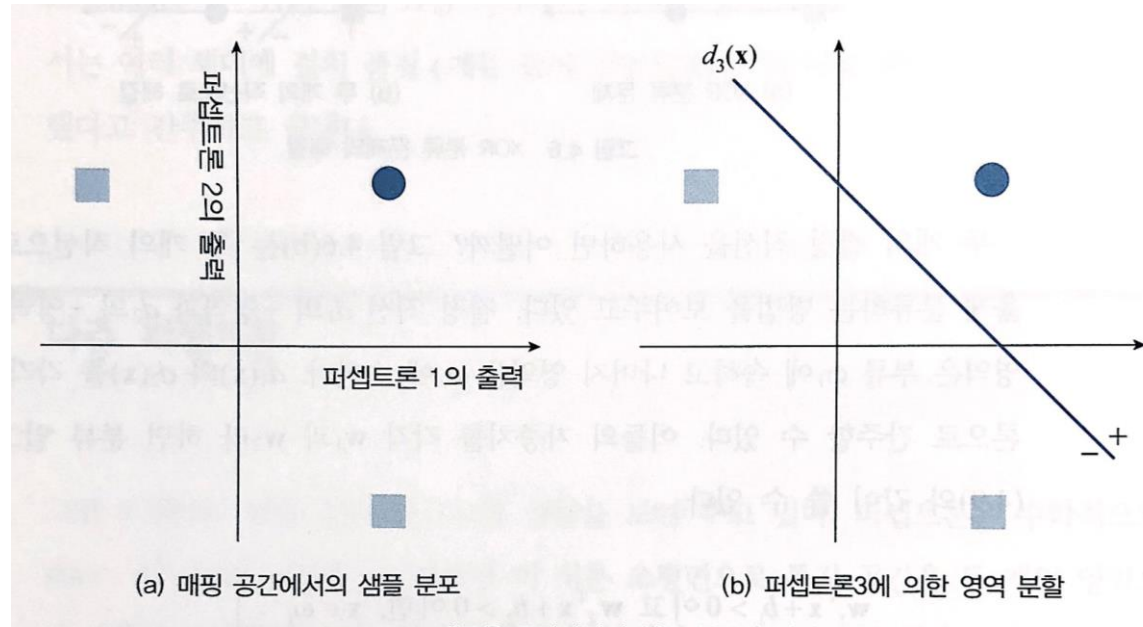


그림 4.8 세 개의 퍼셉트론과 이들을 연결하여 만든 다층 퍼셉트론

Chain rule

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

$$z = (x + y)^2 \quad \Rightarrow \quad \begin{aligned} z &= t^2 \\ t &= x + y \end{aligned}$$

z의 x에 대한
미분의 연쇄법
칙 표현

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial t} \frac{\partial t}{\partial x} \quad \Rightarrow \quad \frac{\partial z}{\partial x} = \frac{\partial z}{\cancel{\partial t}} \frac{\cancel{\partial t}}{\partial x}$$

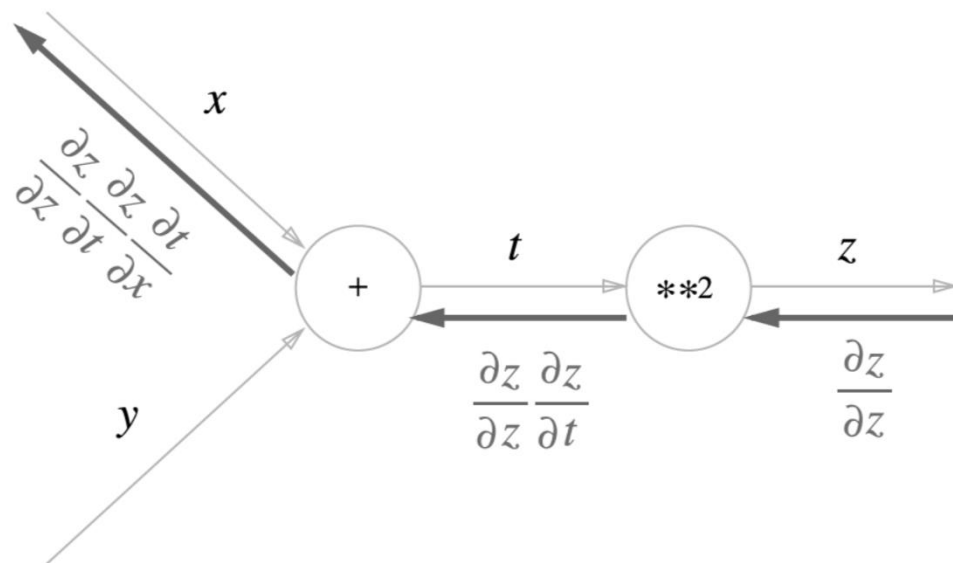
$$z = (x + y)^2 \quad \Rightarrow \quad \begin{aligned} z &= t^2 \\ t &= x + y \end{aligned}$$

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial t} \frac{\partial t}{\partial x} \quad \Rightarrow \quad \begin{aligned} \frac{\partial z}{\partial t} &= 2t \\ \frac{\partial t}{\partial x} &= 1 \end{aligned}$$

결과 $\frac{\partial z}{\partial x} = \frac{\partial z}{\partial t} \frac{\partial t}{\partial x} = 2t \cdot 1 = 2(x + y)$

- 연쇄법칙과 계산 그래프
역전파 과정

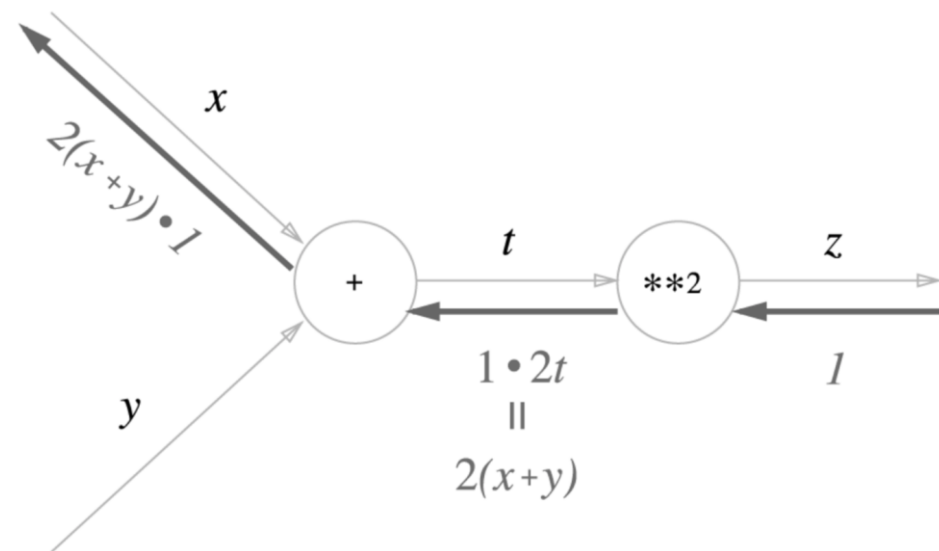
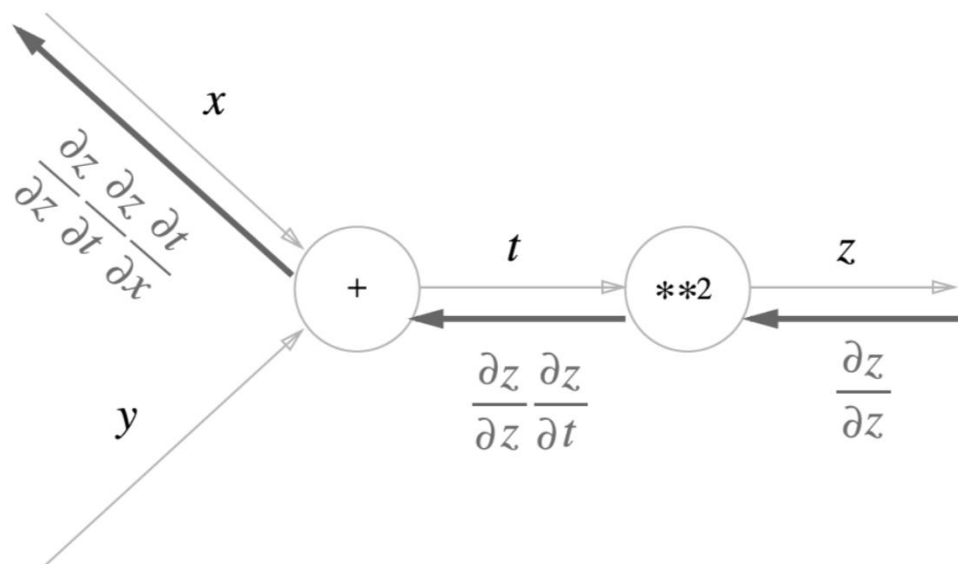
$$z = (x + y)^2 \Rightarrow \begin{aligned} z &= t^2 \\ t &= x + y \end{aligned}$$



- 연쇄법칙과 계산 그래프

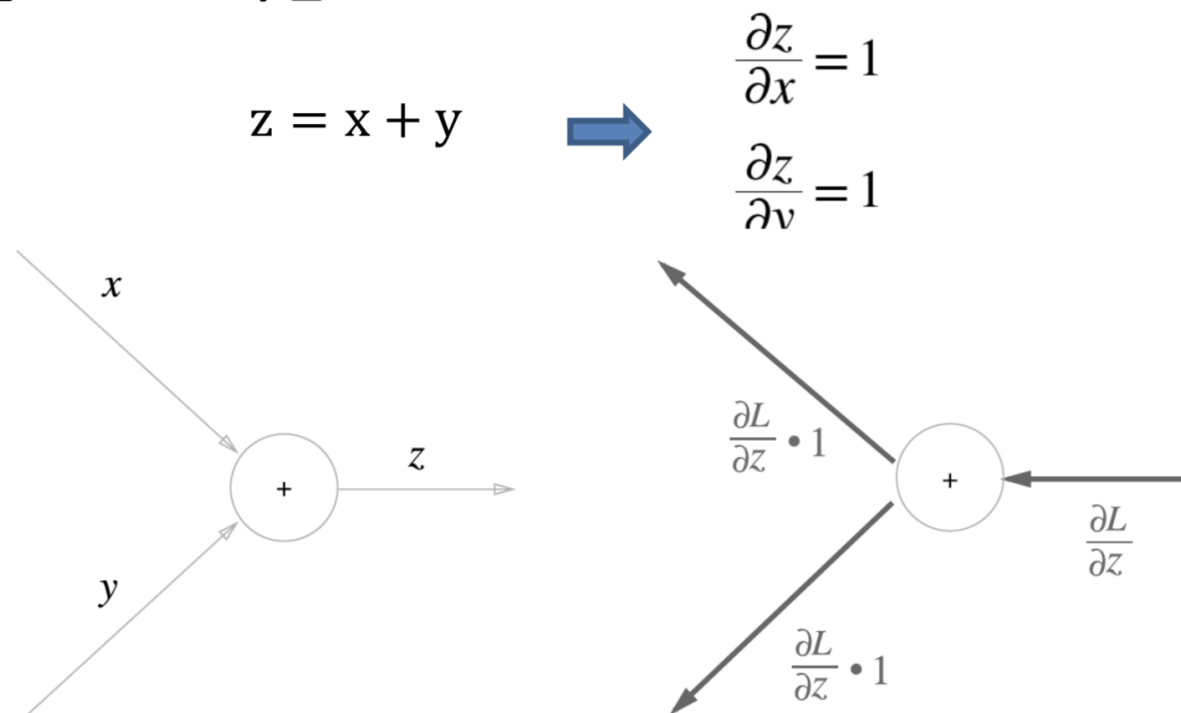
역전파 과정

$$z = (x + y)^2 \Rightarrow \begin{aligned} z &= t^2 \\ t &= x + y \end{aligned}$$

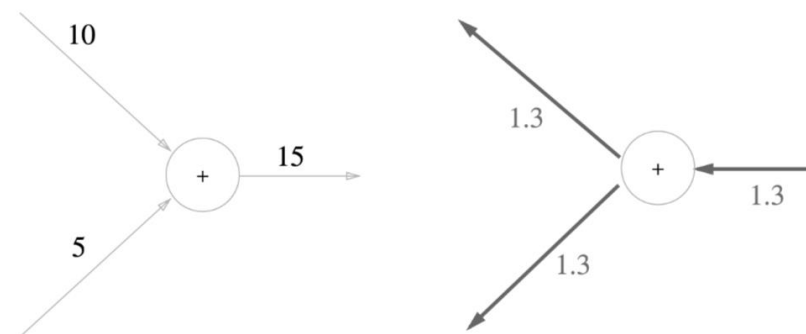


Computational Graph | Backpropagation

- 덧셈 노드의 역전파



예시

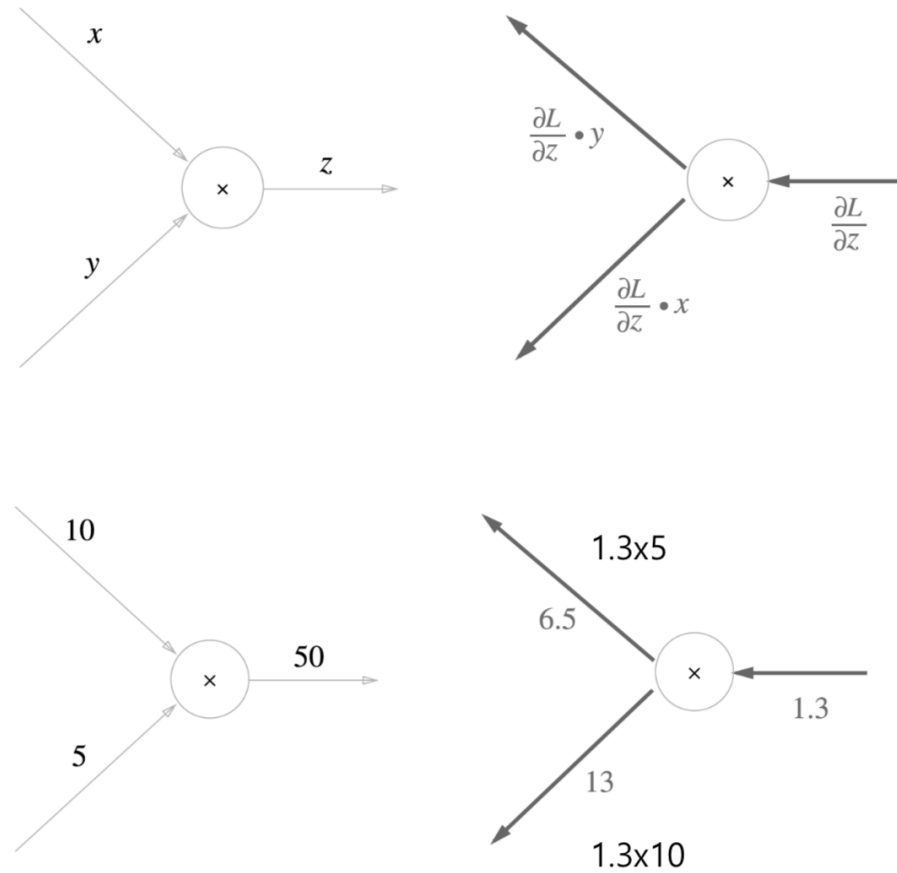


LYCA
LA13

- 곱셈 노드의 역전파

$$z = xy \Rightarrow \begin{aligned} \frac{\partial z}{\partial x} &= y \\ \frac{\partial z}{\partial y} &= x \end{aligned}$$

예시

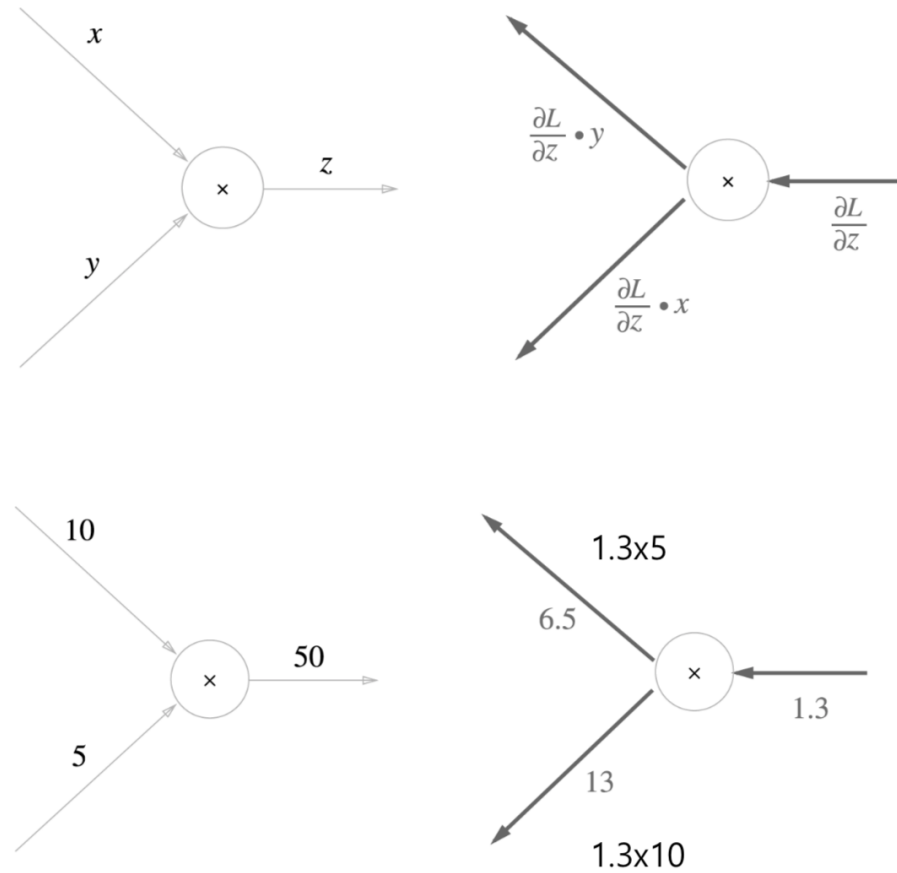


곱셈 노드의 역전파는 순방향 입력 신호가 필요함

- 곱셈 노드의 역전파

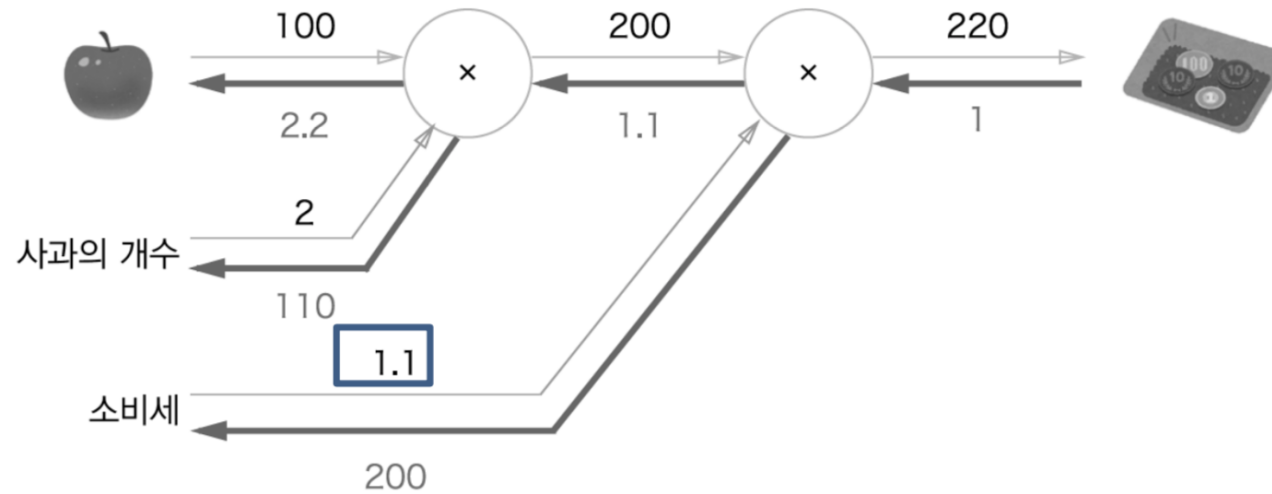
$$z = xy \Rightarrow \begin{aligned} \frac{\partial z}{\partial x} &= y \\ \frac{\partial z}{\partial y} &= x \end{aligned}$$

예시



곱셈 노드의 역전파는 순방향 입력 신호가 필요함

Computational Graph | Backpropagation



- 사과가격이 1원 오르면, 최종 가격 2.2 증가
 - $101(\text{사과 가격}) \times 2 \times 1.1 = 222.2$ 원
- 소비세가 100%(1.1+1 = 2.1)오르면 가격 200증가
 - $100 \times 2 \times 2.1(\text{소비세}) = 420$ 원

Computational Graph | Backpropagation

