Forward Propagation

수업 목표

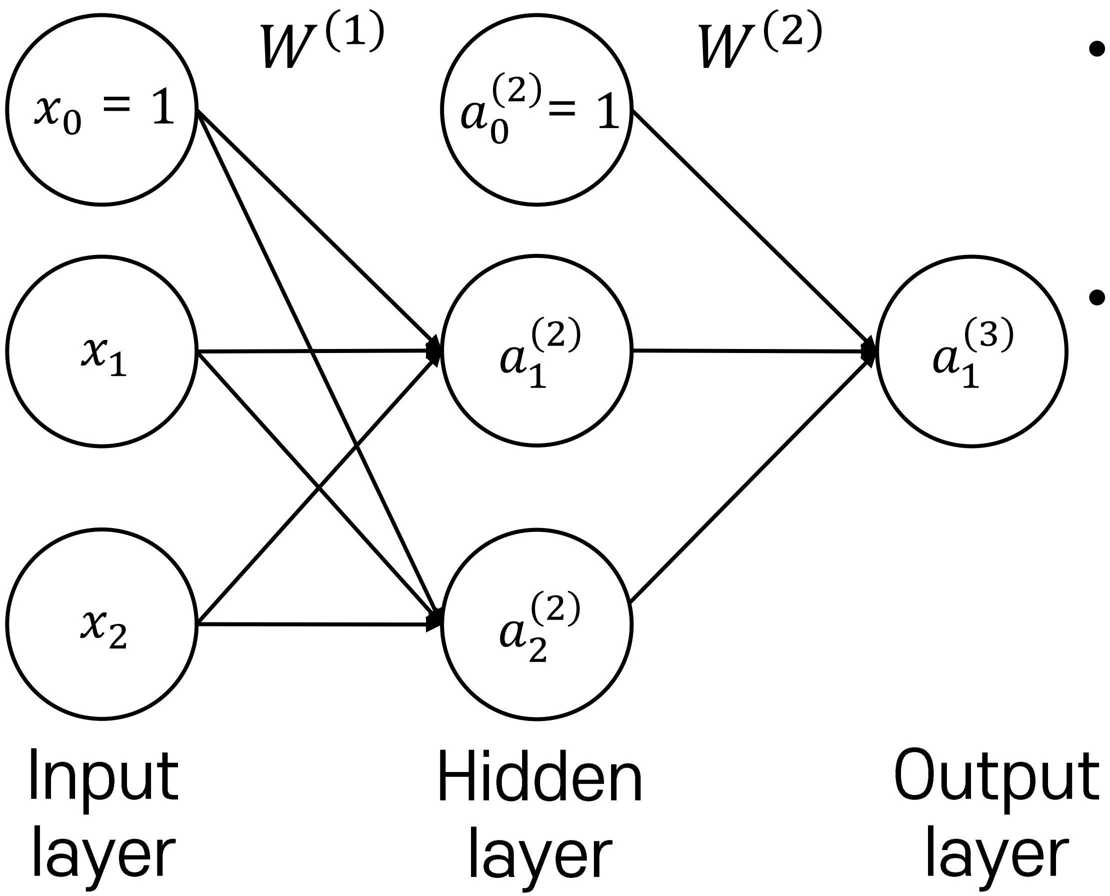
이번 수업의 핵심:

- Forward Propagation의 개념
- Neural Network에서의 Forward Propagation
- 행렬을 이용한 Forward Propagation 계산
- Linear Layer의 개념 및 특성

핵심 개념

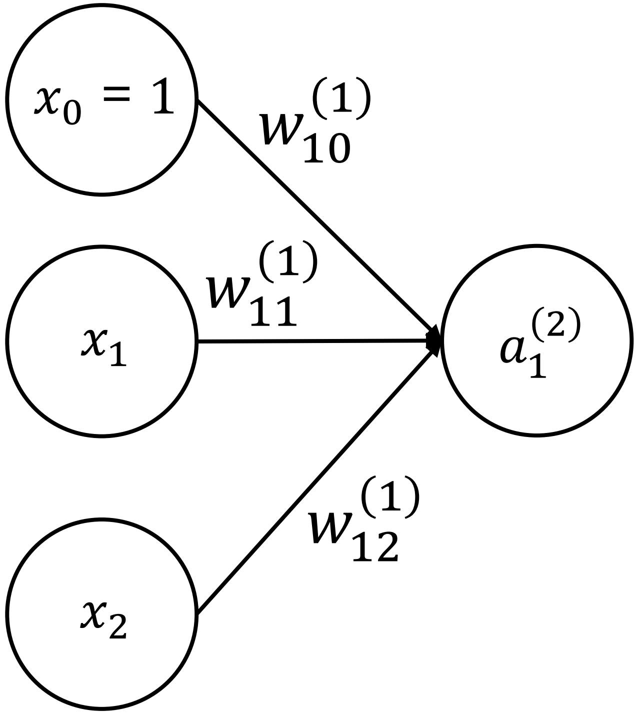
- Forward Propagation
- Activation, Weight Matrix
- Linear Layer

Forward Propagation



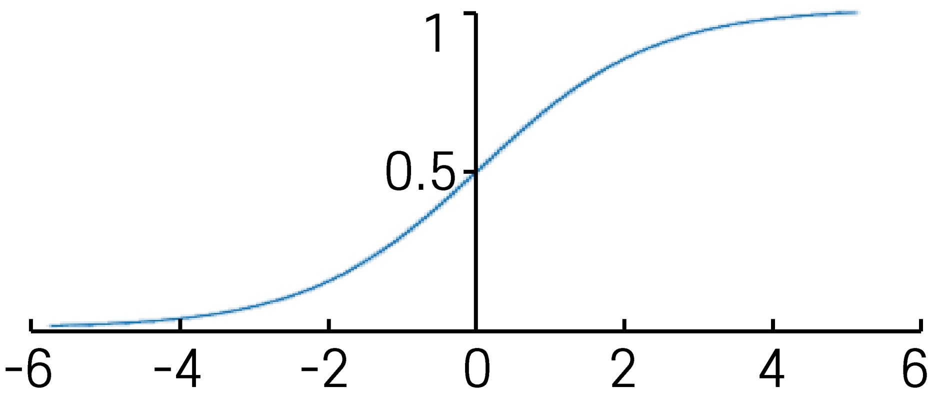
- $a_j^{(i)}$: i번째 층 j번째 유닛의 "Activation (활성화도)"
- $W^{(i)}$: i번째 층과 $(i+1)$번째 층을 잇는 "Weight Matrix (가중치 행렬)"

Forward Propagation



$$z_1^{(2)} = w_{10}^{(1)} x_0 + w_{11}^{(1)} x_1 + w_{12}^{(1)} x_2$$
$$= \begin{bmatrix} w_{10}^{(1)} & w_{11}^{(1)} & w_{12}^{(1)} \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix}$$

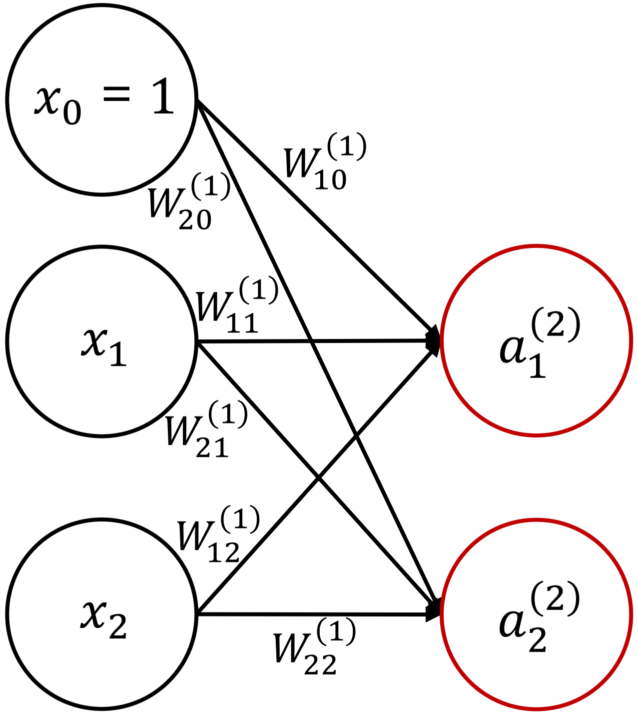
$$a_1^{(2)} = g(z_1^{(2)})$$



Logistic function
(Sigmoid function)

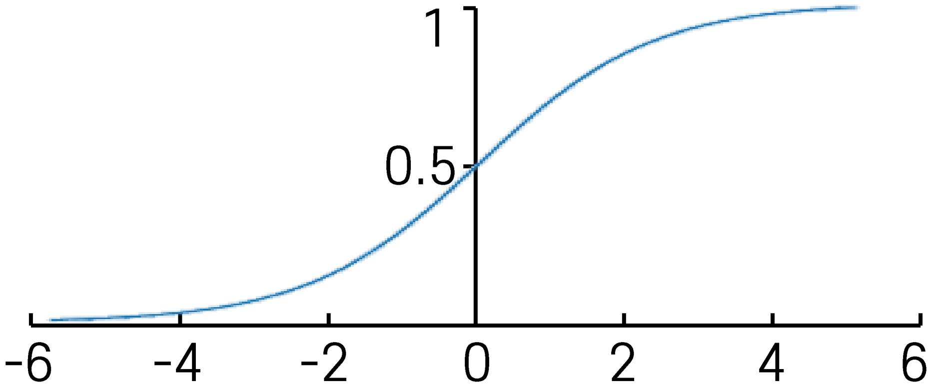
$$g(x) = \frac{1}{1+e^{-x}}$$

Forward Propagation



$$\begin{bmatrix} z_1^{(2)} \\ z_2^{(2)} \end{bmatrix} = \begin{bmatrix} w_{10}^{(1)} & w_{11}^{(1)} & w_{12}^{(1)} \\ w_{20}^{(1)} & w_{21}^{(1)} & w_{22}^{(1)} \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix}$$

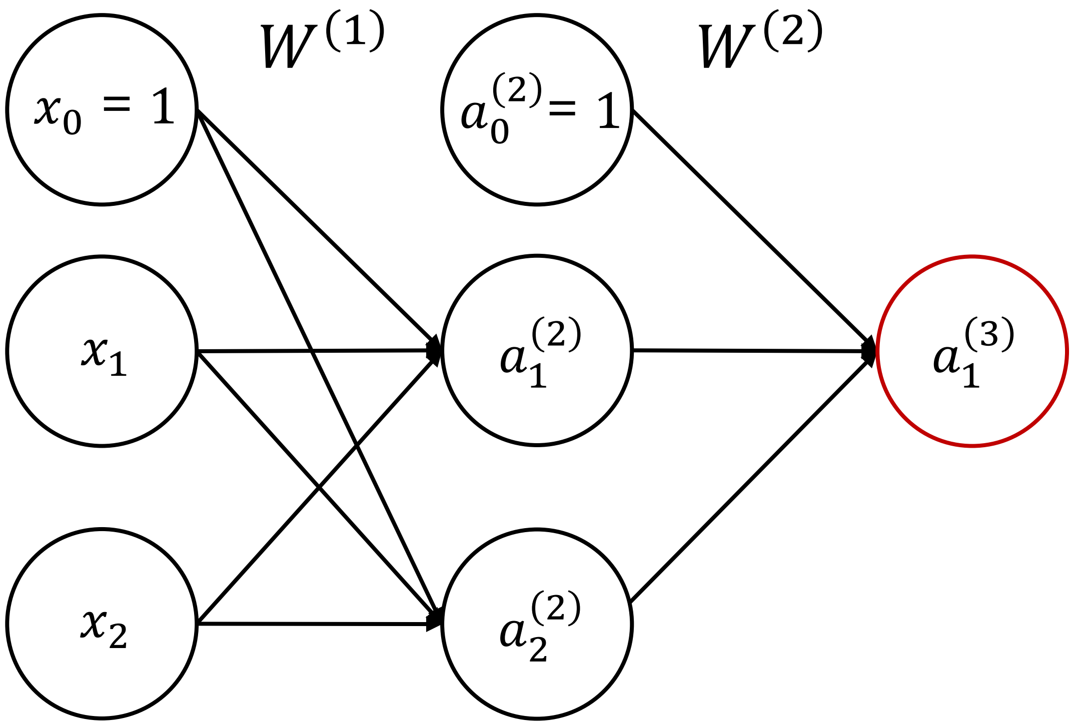
$$\boxed{\begin{bmatrix} a_1^{(2)} \\ a_2^{(2)} \end{bmatrix} = \begin{bmatrix} g\left(z_1^{(2)}\right) \\ g\left(z_2^{(2)}\right) \end{bmatrix}}$$



Logistic function
(Sigmoid function)

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

Forward Propagation



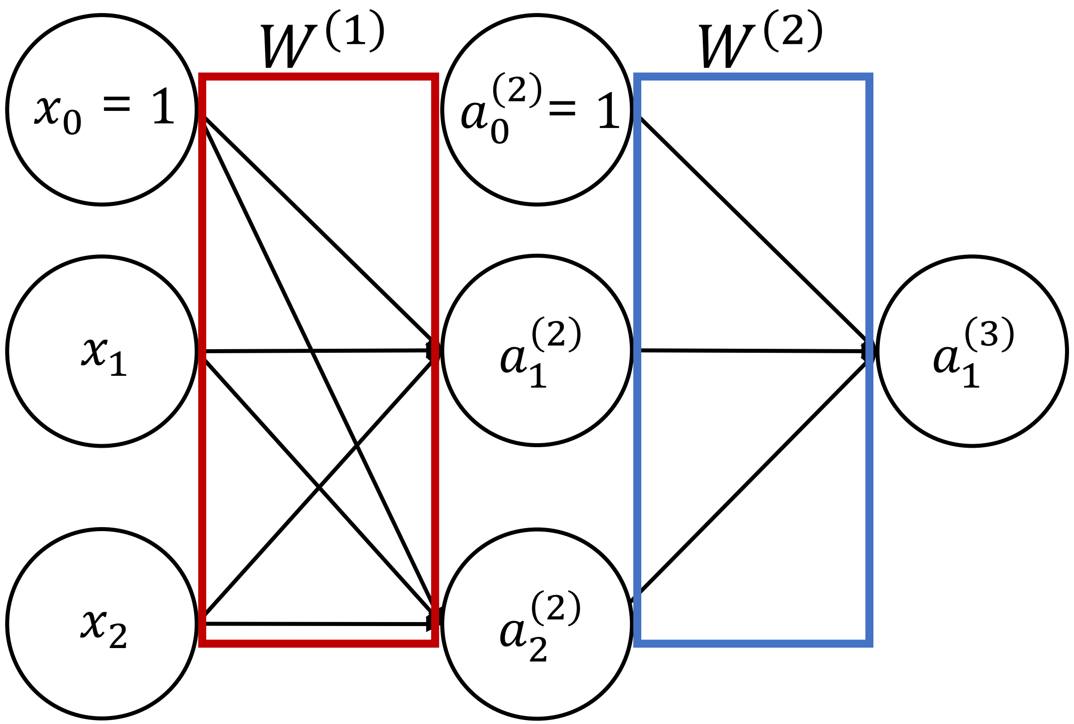
$$\begin{bmatrix} z_1^{(2)} \\ z_2^{(2)} \end{bmatrix} = \begin{bmatrix} w_{10}^{(1)} & w_{11}^{(1)} & w_{12}^{(1)} \\ w_{20}^{(1)} & w_{21}^{(1)} & w_{22}^{(1)} \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix}$$

$$\begin{bmatrix} a_1^{(2)} \\ a_2^{(2)} \end{bmatrix} = \begin{bmatrix} g\left(z_1^{(2)}\right) \\ g\left(z_2^{(2)}\right) \end{bmatrix}$$

$$z_1^{(3)} = \begin{bmatrix} w_{10}^{(2)} & w_{11}^{(2)} & w_{12}^{(2)} \end{bmatrix} \begin{bmatrix} a_0^{(2)} \\ a_1^{(2)} \\ a_2^{(2)} \end{bmatrix}$$

$$\boxed{a_1^{(3)} = g\left(z_1^{(3)}\right)}$$

Linear Layer



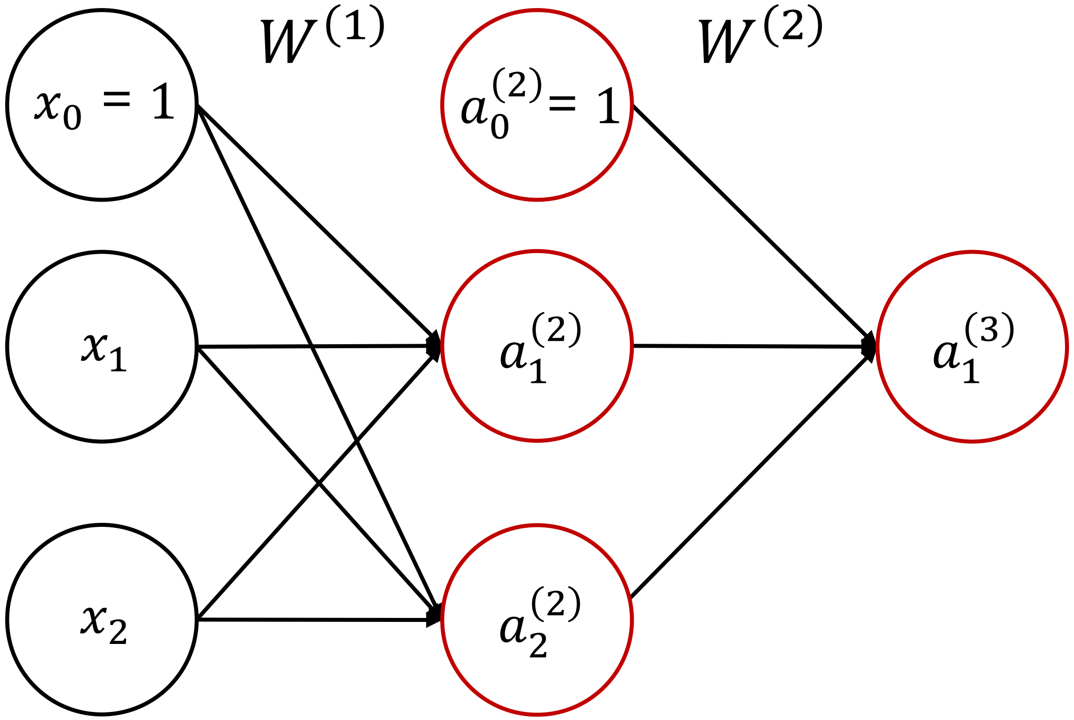
$$\begin{bmatrix} z_1^{(2)} \\ z_2^{(2)} \end{bmatrix} = \boxed{\begin{bmatrix} w_{10}^{(1)} & w_{11}^{(1)} & w_{12}^{(1)} \\ w_{20}^{(1)} & w_{21}^{(1)} & w_{22}^{(1)} \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix}}$$

$$z_1^{(3)} = \boxed{\begin{bmatrix} w_{10}^{(2)} & w_{11}^{(2)} & w_{12}^{(2)} \end{bmatrix} \begin{bmatrix} a_0^{(2)} \\ a_1^{(2)} \\ a_2^{(2)} \end{bmatrix}}$$

각 층은 선형 변환을 수행하기 때문에 Linear Layer (선형층)라고도 불림

Linear Layer = Fully-connected Layer

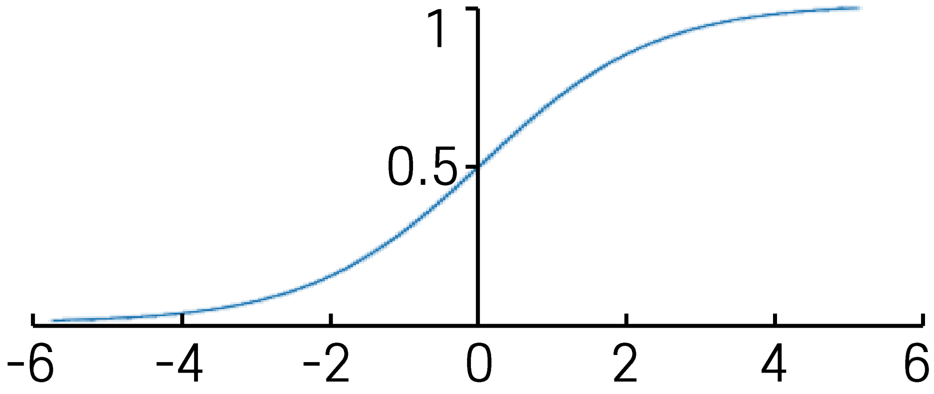
Activation Function의 필요성



$$\begin{bmatrix} a_1^{(2)} \\ a_2^{(2)} \end{bmatrix} = \begin{bmatrix} g\left(z_1^{(2)}\right) \\ g\left(z_2^{(2)}\right) \end{bmatrix}$$

$$a_1^{(3)} = g\left(z_1^{(3)}\right)$$

Activation function은 왜 있을까?



Logistic function
(Sigmoid function)

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

Activation Function의 필요성

일차 함수를 통한 Activation function의 필요성

- 선형 함수: $f_1(x) = 2x + 1,\ f_2(x) = -x + 3$
- Activation function: $g(x)$

$$f_2\left(g\left(f_1(x)\right)\right) = -g(2x + 1) + 3$$

Activation function이 만약 없다면?

$$f_2\left(f_1(x)\right) = -(2x + 1) + 3 = -2x + 2$$

→ 또다른 선형 함수

요약

- Neural Network에서 Forward Propagation 계산
- 행렬을 이용한 Forward Propagation
- Linear Layer와 Fully-Connected Layer의 연관성