

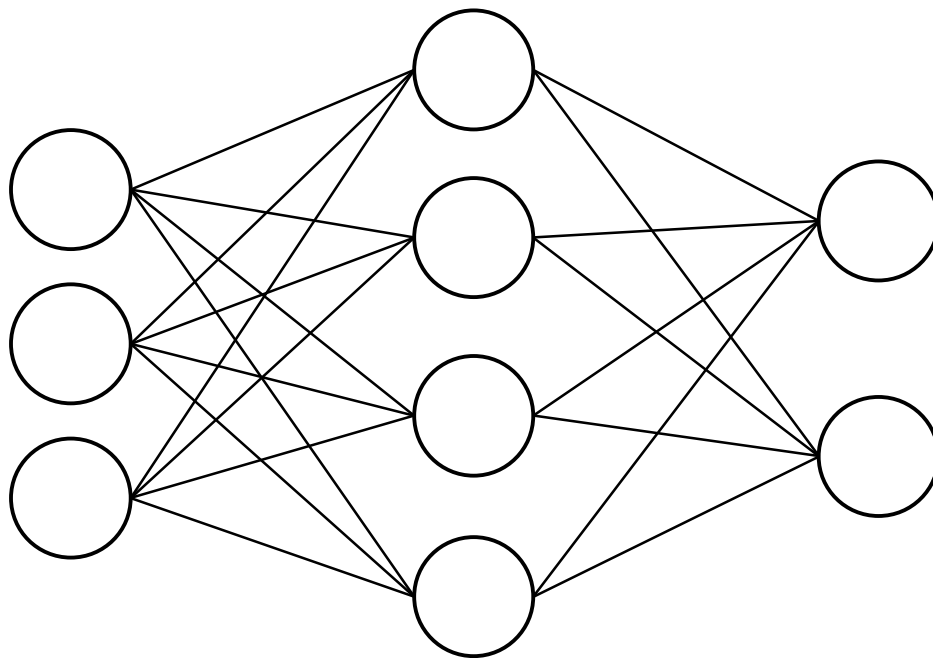
Chapter 02. 가장 단순한 신경망을 통해 작동 원리 이해하기

# STEP2.

## 얕은 신경망의 수식적 이해

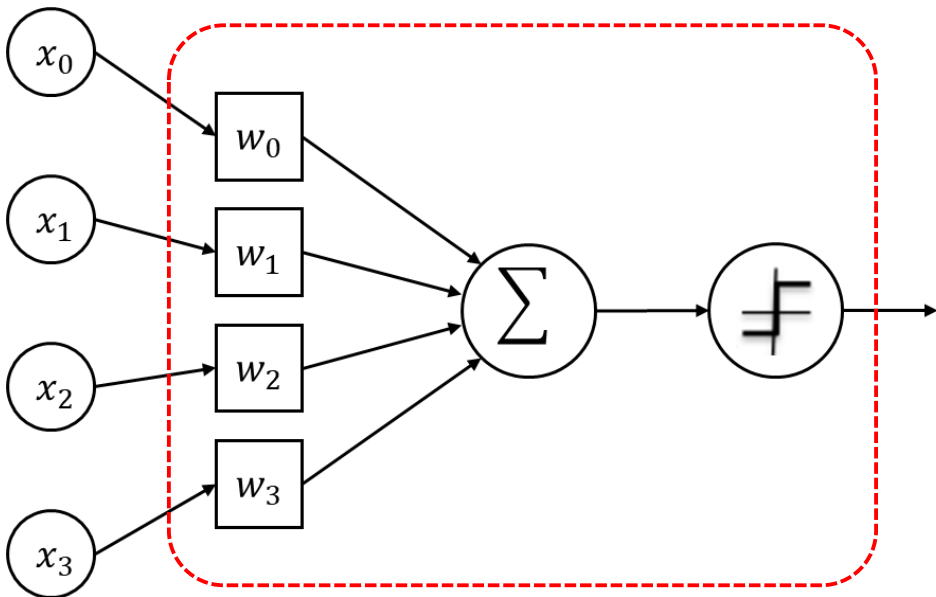
# 얕은 신경망을 깊게 이해해 보자

입력 계층 (Input layer)      은닉 계층 (Hidden layer)      출력 계층 (Output layer)



얕은 신경망에 대한 깊은 이해가 깊은 신경망(심층 신경망)에 대한 이해를 도울 것이다.

# 뉴런의 수학적 표현



$$\sum_{i=0}^{N-1} w_i x_i = \mathbf{w} \cdot \mathbf{x} = \mathbf{w}^T \mathbf{x}$$

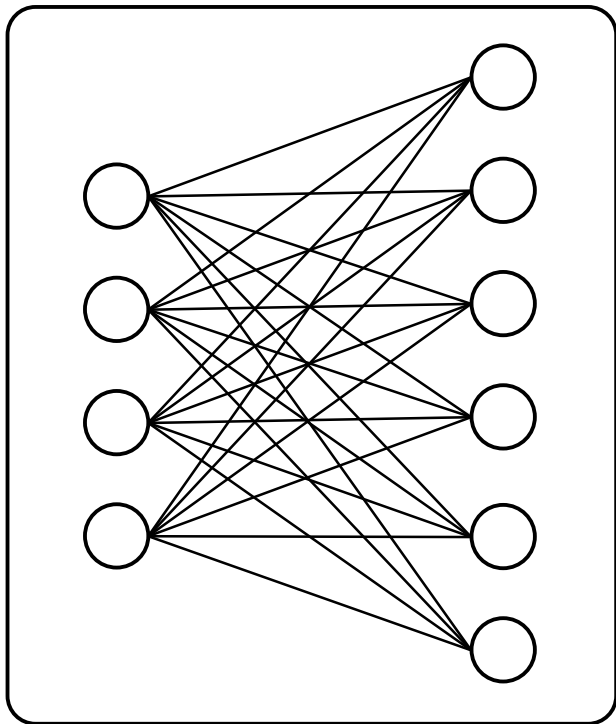
$$y = a \left( \underbrace{\sum_{i=0}^{N-1} w_i x_i}_{\text{두 벡터의 내적}} + b \right) \leftarrow \text{편향(Bias)}$$

활성함수

$$y = a(\mathbf{w}^T \mathbf{x} + b)$$

뉴런은 수학적으로 **두 벡터의 내적**으로 쉽게 표현할 수 있다.

# 전결합 계층의 수학적 표현



$$W = [\mathbf{w}_0, \mathbf{w}_1, \dots, \mathbf{w}_{M-1}]^T$$

$$\mathbf{b} = [b_0, b_1, \dots, b_{M-1}]^T$$

$$y_0 = a(\mathbf{w}_0^T \mathbf{x} + b_0)$$

$$y_1 = a(\mathbf{w}_1^T \mathbf{x} + b_1)$$

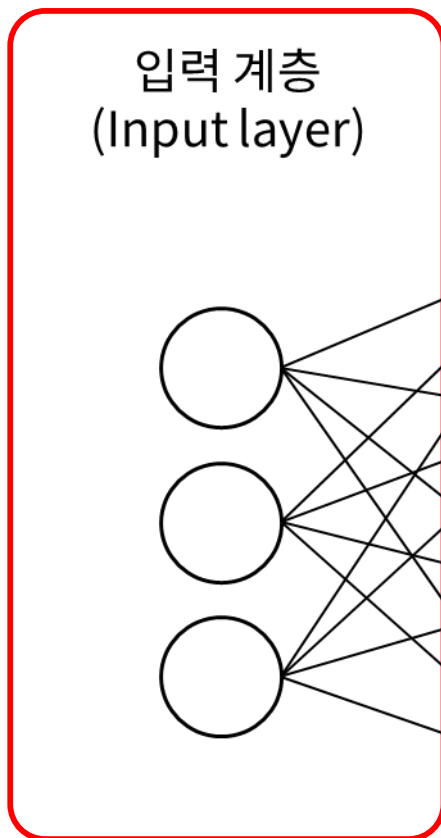
$$\vdots$$

$$y_{M-1} = a(\mathbf{w}_{M-1}^T \mathbf{x} + b_{M-1})$$

$$\mathbf{y} = a(W\mathbf{x} + \mathbf{b})$$

FC 계층은 여러 개의 뉴런을 한 곳에 모아둔 것으로, **Matrix 곱셈 연산**으로 표현된다.

# 입력 계층 (Input Layer)

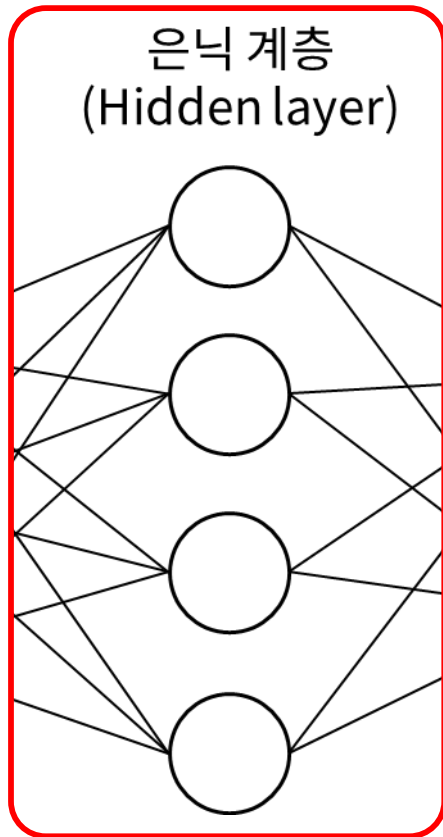


## 입력 계층 (Input Layer)

- 입력 계층은 아무런 연산도 일어나지 않는다.
- 신경망의 입력을 받아서 다음 계층으로 넘기는 역할
- 무엇을 입력으로 주어야 하는가? → 특징 추출 문제
- 계층의 크기 = Node의 개수 = 입력 Scalar의 수 = 입력 Vector의 길이

$$\mathbf{x} = [x_0, x_1, x_2, \dots, x_{N-1}]^T$$

# 은닉 계층 (Hidden Layer)

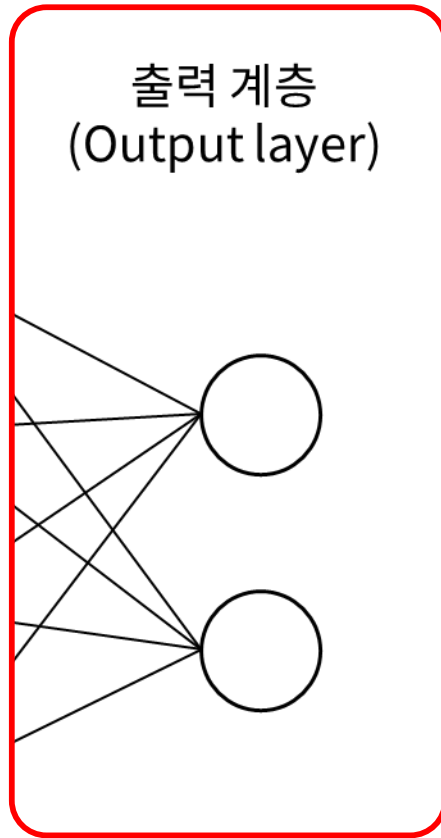


## 은닉 계층 (Hidden Layer)

- 은닉 계층은 입력 계층과 연결된 **전결합 계층**이다.
- 입출력 관점에서 볼 때 드러나지 않는다고 하여, 은닉 계층이라 한다.
- 복잡한 문제를 해결할 수 있게 하는 핵심적인 계층.
- 얕은 신경망에서는 1개의 은닉 계층만을 사용한다.

$$\mathbf{h} = a_h(W_h \mathbf{x} + \mathbf{b}_h)$$

# 출력 계층 (Output Layer)



## 출력 계층 (Output Layer)

- 출력 계층은 은닉 계층 다음에 오는 전결합 계층이다.
- 신경망의 외부로 출력 신호를 전달하는 데에 사용된다.
- 신경망의 **기능은 출력 계층의 활성화 함수에 의해 결정**된다.
- 출력 계층의 크기 = 출력의 Scalar 수 = 출력 벡터의 길이

$$\mathbf{y} = a_o(W_o\mathbf{h} + \mathbf{b}_o)$$