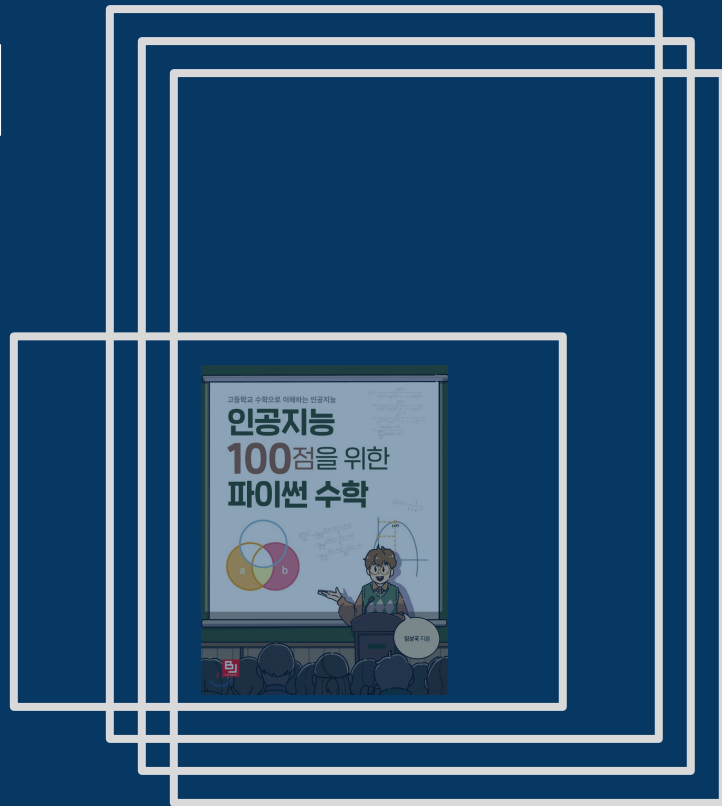


# 08. MNIST와 신경망의 입력과 출력

인공지능 100점을 위한 파이썬 수학



# Contents

1. 행렬식의 이해
2. 행렬의 계산과 선형대수
3. 행렬 계산식의 수학적 이해
4. 파이썬 코드
5. 신경망 데이터의 행렬 특징
6. 무작위 데이터로 신경망함수 만들기

# 1. 행렬식의 이해

# 01. 행렬식의 이해

## ● 행렬곱

$$A \quad \cdot \quad B \quad = \quad C$$

$$\begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \cdot \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix} = \begin{pmatrix} c_1 & c_2 \\ c_3 & c_4 \end{pmatrix}$$

$$A \cdot B = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix} = \begin{pmatrix} a_1b_1 + a_2b_3 & a_1b_2 + a_2b_4 \\ a_3b_1 + a_4b_3 & a_3b_2 + a_4b_4 \end{pmatrix}$$

## 2. 행렬의 계산과 선형대수

## 02. 행렬의 계산과 선형대수

### ● 연습문제 8-1. 행렬을 이용한 연산

포도, 귤, 사과, 배를 각각 1개를 구입하고 2,350원을 지불했습니다. 다음날 순서대로 5개, 2개, 4개, 7개를 구입하고 10,600원을 지불했습니다. 다시 다음날 귤 1개, 사과 1개, 배 2개를 구입하고 2,450원을 지불했습니다. 다음날 5개, 2개, 2개, 3개를 구입하고 6,800원을 지불했습니다. 받은 영수증에는 포도, 귤, 사과, 배의 개별 가격이 기록되지 않은 채 총금액만 적혀 있었습니다. 그러면 과일들의 개당 가격은 얼마일까요?

## 02. 행렬의 계산과 선형대수

### ● 행렬을 이용한 연산

포도, 귤, 사과, 배를 각각  $W$ ,  $X$ ,  $Y$ ,  $Z$ 라고 놓으면,

$$W + X + Y + Z = 2350 \text{ --- (1)}$$

$$5W + 2X + 4Y + 7Z = 10600 \text{ --- (2)}$$

$$X + Y + 2Z = 2450 \text{ --- (3)}$$

$$5W + 2X + 2Y + 3Z = 6800 \text{ --- (4)}$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 5 & 2 & 4 & 7 \\ 0 & 1 & 1 & 2 \\ 5 & 2 & 2 & 3 \end{bmatrix} \begin{bmatrix} W \\ X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 2350 \\ 10600 \\ 2450 \\ 6800 \end{bmatrix}$$

# 3. 행렬 계산식의 수학적 이해



## 03. 행렬 계산식의 수학적 이해

### 행렬을 이용한 연산

귤 하나, 사과 하나의 가격이 1, 100원이고, 귤 둘, 사과 셋의 가격이 2, 800원이라면

$$\begin{array}{lcl} X + Y = 1100 & \text{---} & (1) \\ 2X + 3Y = 2800 & \text{---} & (2) \end{array} \quad \begin{pmatrix} 1 \\ 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} 1100 \\ 2800 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}, \quad X = \begin{pmatrix} X \\ Y \end{pmatrix}, \quad B = \begin{pmatrix} 1100 \\ 2800 \end{pmatrix}$$

### 03. 행렬 계산식의 수학적 이해

#### ● 행렬을 이용한 연산

The diagram illustrates the matrix multiplication  $A(2,2) \cdot X(2,1) = B(2,1)$ . Red arrows and text explain the compatibility of dimensions:

- A red arrow from the second dimension of A (2) to the first dimension of X (2) is labeled "A 행의 크기 = B 행의 크기".
- A red arrow from the second dimension of X (1) to the second dimension of B (1) is labeled "B 열의 크기 = B 열의 크기".
- A bracket under the first two dimensions of A (2,2) is labeled "A 열의 크기 = X 행의 크기".

$A(2,2) \cdot X(2,1) = B(2,1)$

## 03. 행렬 계산식의 수학적 이해

### ● 행렬을 이용한 연산

$$\begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}^{-1} \begin{pmatrix} 1100 \\ 2800 \end{pmatrix} \quad \begin{matrix} \mathbf{AX} = \mathbf{B} \\ \mathbf{X} = \mathbf{A}^{-1}\mathbf{B} \end{matrix}$$

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}^{-1} \begin{pmatrix} 1100 \\ 2800 \end{pmatrix}$$

## 4. 파이썬 코드

## 04. 파이썬 코드

### ● 행렬계산식

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}, \quad X = \begin{pmatrix} X \\ Y \end{pmatrix}, \quad B = \begin{pmatrix} 1100 \\ 2800 \end{pmatrix} \text{ 일 때}$$

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}^{-1} \begin{pmatrix} 1100 \\ 2800 \end{pmatrix} \text{ 이므로}$$

$$X = A^{-1} B$$

## 04. 파이썬 코드

### 1.2.1

CODE

$$X = A^{-1} B$$

```
import numpy as np
from numpy.linalg import inv

A = np.array([[1,1],
              [2,3]])
B = np.array([[1100],
              [2800]])

invA = inv(A)
print(np.dot(invA , B))
```

## 04. 파이썬 코드

### 🕒 연습문제 8-1 을 파이썬코드로 풀어보세요

포도, 귤, 사과, 배를 각각 1개를 구입하고 2,350원을 지불했습니다. 다음날 순서대로 5개, 2개, 4개, 7개를 구입하고 10,600원을 지불했습니다. 다시 다음날 귤 1개, 사과 1개, 배 2개를 구입하고 2,450원을 지불했습니다. 다음날 5개, 2개, 2개, 3개를 구입하고 6,800원을 지불했습니다. 받은 영수증에는 포도, 귤, 사과, 배의 개별 가격이 기록되지 않은 채 총금액만 적혀 있었습니다. 그러면 과일들의 개당 가격은 얼마일까요?

#### CODE

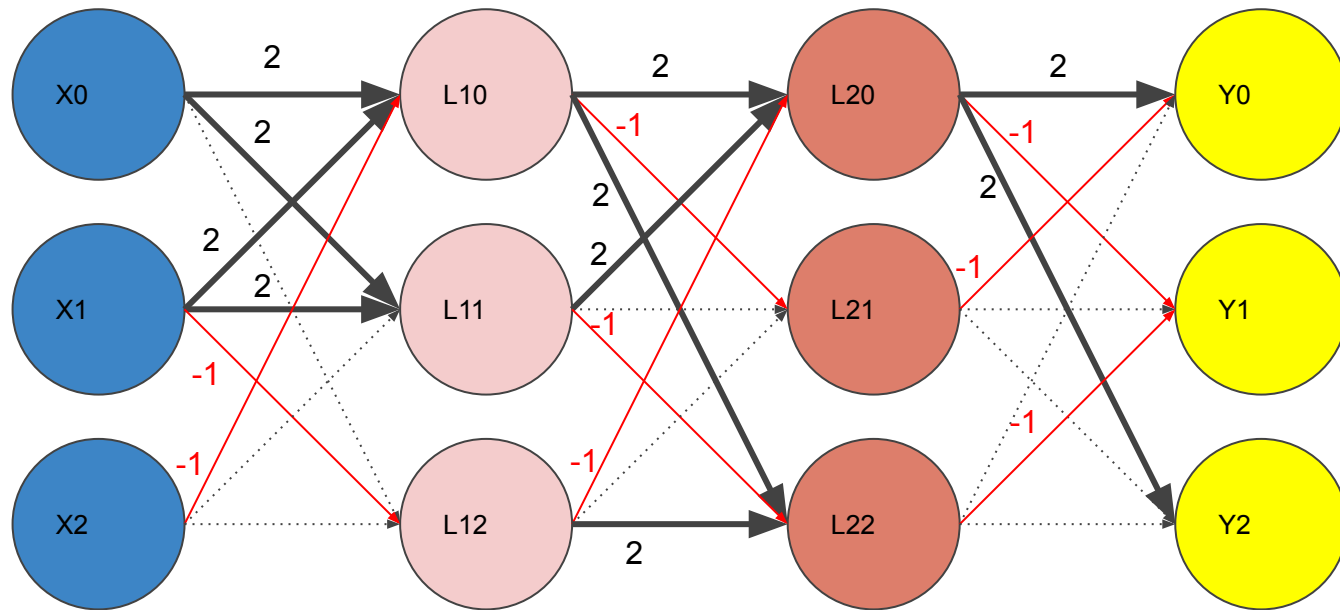
```
# ml_ex_08_01.py
import numpy as np
from numpy.linalg import inv

.....
```

# 행렬식과 신경망



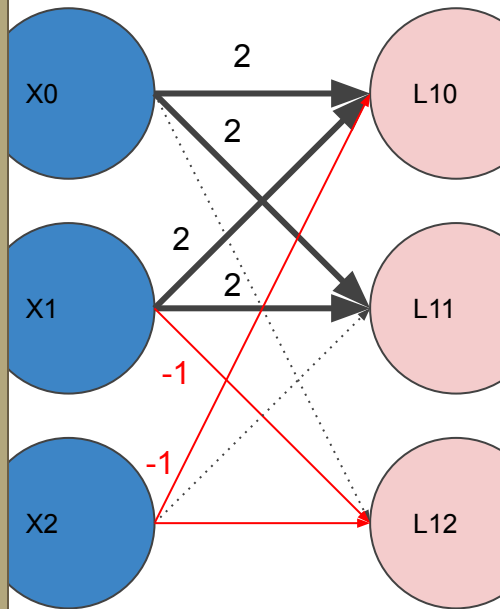
# 행렬식과 신경망



# 행렬식과 신경망

입력 :  
 $X_0, X_1, X_2$

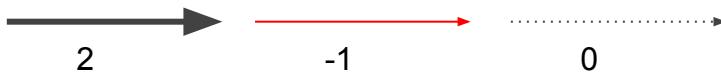
출력 :  
 $L_{10}, L_{11}, L_{12}$



$$L_{10} = 2X_0 + 2X_1 - X_2$$

$$L_{11} = 2X_0 + 2X_1$$

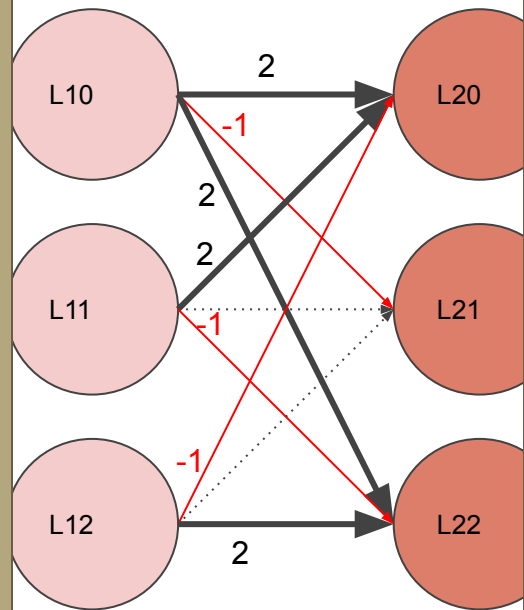
$$L_{12} = -X_1 - X_2$$



# 행렬식과 신경망

입력 : L10, L11, L12

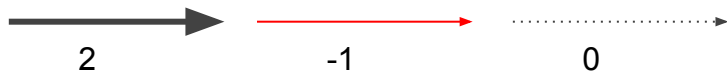
출력 : L20, L21, L22



$$L20 = 2L10 + 2L11 - L12$$

$$L21 = -L10$$

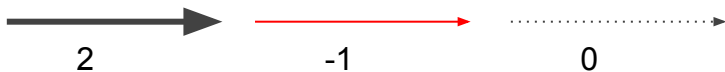
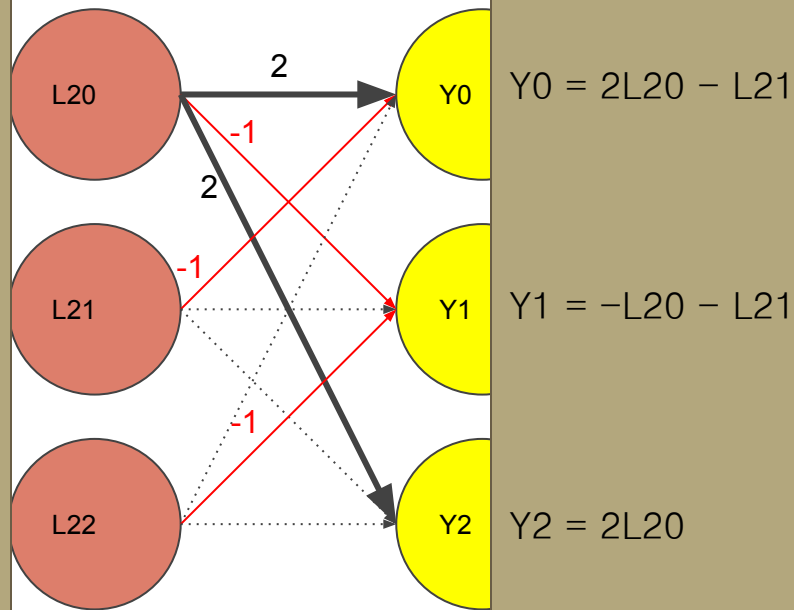
$$L22 = 2L10 - L11 + 2L12$$

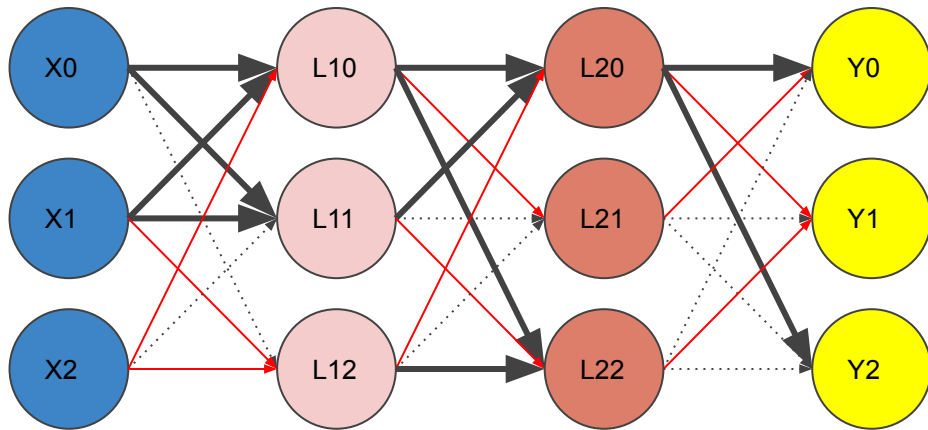


# 행렬식과 신경망

입력 : L20, L21, L22

출력 : Y0, Y1, Y2





$$\begin{aligned} L10 &= 2X0 + 2X1 - X2 \\ L11 &= 2X0 + 2X1 \\ L12 &= -X1 - X2 \end{aligned}$$

$$\begin{aligned} L20 &= 2L10 + 2L11 - L12 \\ L21 &= -L10 \\ L22 &= 2L10 - L11 + 2L12 \end{aligned}$$

$$\begin{aligned} Y0 &= 2L20 - L21 \\ Y1 &= -L20 - L21 \\ Y2 &= 2L20 \end{aligned}$$

$$\begin{aligned} L10 &= 2X0 + 2X1 - X2 \\ L11 &= 2X0 + 2X1 \\ L12 &= -X1 - X2 \end{aligned}$$

$$\begin{aligned} L20 &= 2L10 + 2L11 - L12 \\ L21 &= -L10 \\ L22 &= 2L10 - L11 + 2L12 \end{aligned}$$

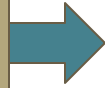
$$\begin{aligned} Y0 &= 2L20 - L21 \\ Y1 &= -L20 - L21 \\ Y2 &= 2L20 \end{aligned}$$



$$\begin{bmatrix} L10 \\ L11 \\ L12 \end{bmatrix} = \begin{bmatrix} 2 & 2 & -1 \\ 2 & 2 & 0 \\ 0 & -1 & -1 \end{bmatrix} \begin{bmatrix} X0 \\ X1 \\ X2 \end{bmatrix}$$



$$\begin{bmatrix} L20 \\ L21 \\ L22 \end{bmatrix} = \begin{bmatrix} 2 & 2 & -1 \\ -1 & 0 & 0 \\ 2 & -1 & 2 \end{bmatrix} \begin{bmatrix} L10 \\ L11 \\ L12 \end{bmatrix}$$



$$\begin{bmatrix} Y0 \\ Y1 \\ Y2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & -1 & 0 \\ 2 & 0 & 0 \end{bmatrix} \begin{bmatrix} L20 \\ L21 \\ L22 \end{bmatrix}$$

출력

세번째 계산

두번째 계산

첫번째 계산

입력

$$\begin{bmatrix} Y_0 \\ Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & -1 & 0 \\ 2 & 0 & 0 \end{bmatrix} \begin{bmatrix} L_{20} \\ L_{21} \\ L_{22} \end{bmatrix} = \begin{bmatrix} 2 & 2 & -1 \\ -1 & 0 & 0 \\ 2 & -1 & 2 \end{bmatrix} \begin{bmatrix} L_{10} \\ L_{11} \\ L_{12} \end{bmatrix} = \begin{bmatrix} 2 & 2 & -1 \\ 2 & 2 & 0 \\ 0 & -1 & -1 \end{bmatrix} \begin{bmatrix} X_0 \\ X_1 \\ X_2 \end{bmatrix}$$

$$\begin{aligned} L_{10} &= 2X_0 + 2X_1 - X_2 \\ L_{11} &= 2X_0 + 2X_1 \\ L_{12} &= -X_1 - X_2 \end{aligned}$$

$$\begin{aligned} L_{20} &= 2L_{10} + 2L_{11} - L_{12} \\ L_{21} &= -L_{10} \\ L_{22} &= 2L_{10} - L_{11} + 2L_{12} \end{aligned}$$

$$\begin{aligned} Y_0 &= 2L_{20} - L_{21} \\ Y_1 &= -L_{20} - L_{21} \\ Y_2 &= 2L_{20} \end{aligned}$$

$$\begin{bmatrix} L_{10} \\ L_{11} \\ L_{12} \end{bmatrix} = \begin{bmatrix} 2 & 2 & -1 \\ 2 & 2 & 0 \\ 0 & -1 & -1 \end{bmatrix} \begin{bmatrix} X_0 \\ X_1 \\ X_2 \end{bmatrix}$$

$$\begin{bmatrix} L_{20} \\ L_{21} \\ L_{22} \end{bmatrix} = \begin{bmatrix} 2 & 2 & -1 \\ -1 & 0 & 0 \\ 2 & -1 & 2 \end{bmatrix} \begin{bmatrix} L_{10} \\ L_{11} \\ L_{12} \end{bmatrix}$$

$$\begin{bmatrix} Y_0 \\ Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & -1 & 0 \\ 2 & 0 & 0 \end{bmatrix} \begin{bmatrix} L_{20} \\ L_{21} \\ L_{22} \end{bmatrix}$$

`np.array([])`

출력

세번째 계산

두번째 계산

첫번째 계산

입력

$$\begin{bmatrix} Y0 \\ Y1 \\ Y2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & -1 & 0 \\ 2 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 2 & -1 \\ -1 & 0 & 0 \\ 2 & -1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 2 & -1 \\ 2 & 2 & 0 \\ 0 & -1 & -1 \end{bmatrix} \begin{bmatrix} X0 \\ X1 \\ X2 \end{bmatrix}$$

`np.dot(A,B)`

`X = np.array([0,0,1])`

`W1 = np.array([[2,2,-1],[2,2,0],[0,-1,-1]])`

`W2 = np.array([[2,2,-1],[-1,0,0],[2,-1,2]])`

`W3 = np.array([[2,-1,0],[-1,-1,0],[2,0,0]])`

`Y = np.dot(np.dot(np.dot(W3, W2), W1), X)`

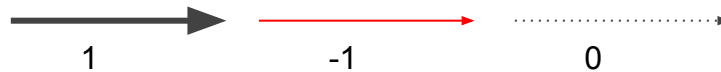
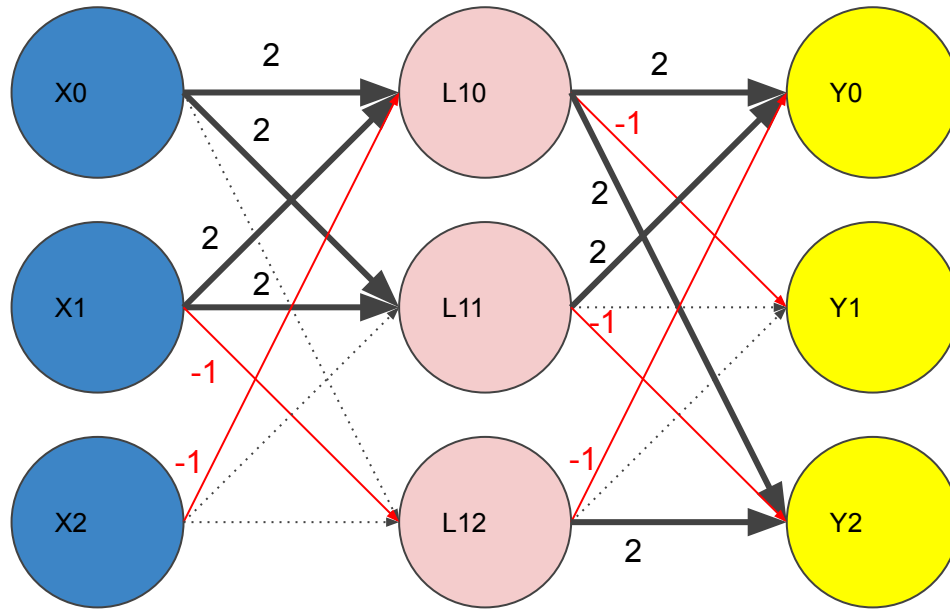
`Y = np.dot(W3, np.dot(W2, np.dot(W1, X)))`

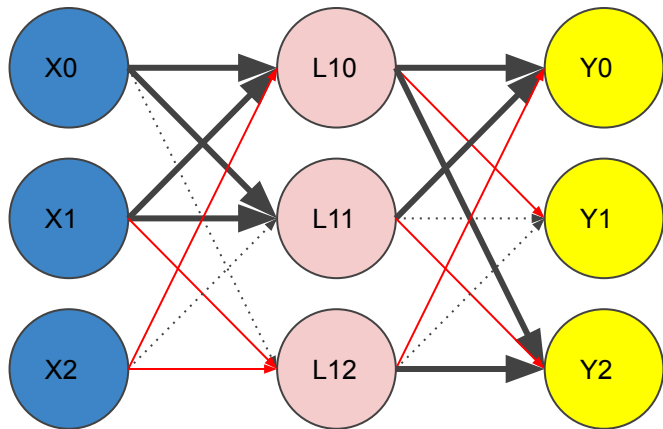
$((A \times B) \times C) \times D = A \times (B \times (C \times D)) \neq D \times C \times B \times A$





# 행렬식과 신경망





$$\begin{aligned} L10 &= 2X0 + 2X1 - X2 \dots + w783 X783 \\ L11 &= 2X0 + 2X1 \\ L12 &= -X1 - X2 \end{aligned}$$

$$\begin{aligned} L20 &= 2L10 + 2L11 - L12 \\ L21 &= -L10 \\ L22 &= 2L10 - L11 + 2L12 \end{aligned}$$

$$\begin{aligned} Y0 &= 2L20 - L21 \\ Y1 &= -L20 - L21 \\ Y2 &= 2L20 \end{aligned}$$

$$\begin{aligned} L10 &= 2X0 + 2X1 - X2 \\ L11 &= 2X0 + 2X1 \\ L12 &= -X1 - X2 \end{aligned}$$

$$\begin{aligned} L20 &= 2L10 + 2L11 - L12 \\ L21 &= -L10 \\ L22 &= 2L10 - L11 + 2L12 \end{aligned}$$

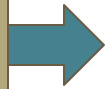
$$\begin{aligned} Y0 &= 2L20 - L21 \\ Y1 &= -L20 - L21 \\ Y2 &= 2L20 \end{aligned}$$



$$\begin{bmatrix} L10 \\ L11 \\ L12 \end{bmatrix} = \begin{bmatrix} 2 & 2 & -1 \\ 2 & 2 & 0 \\ 0 & -1 & -1 \end{bmatrix} \begin{bmatrix} X0 \\ X1 \\ X2 \end{bmatrix}$$



$$\begin{bmatrix} L20 \\ L21 \\ L22 \end{bmatrix} = \begin{bmatrix} 2 & 2 & -1 \\ -1 & 0 & 0 \\ 2 & -1 & 2 \end{bmatrix} \begin{bmatrix} L10 \\ L11 \\ L12 \end{bmatrix}$$



$$\begin{bmatrix} Y0 \\ Y1 \\ Y2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & -1 & 0 \\ 2 & 0 & 0 \end{bmatrix} \begin{bmatrix} L20 \\ L21 \\ L22 \end{bmatrix}$$

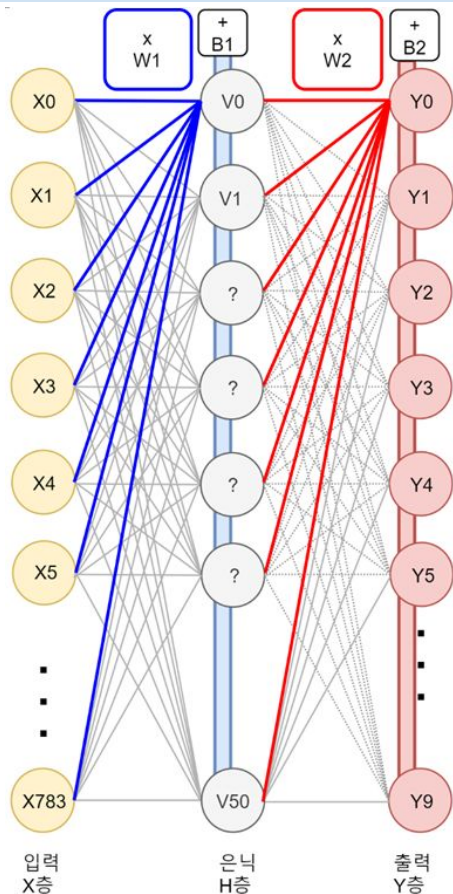
## 5. 신경망 데이터의 행렬 특징

## 05. 신경망 데이터의 행렬 특징

### ● 신경망

여기서는 신경망 전체의 코드 중 행렬을 주의해서 살펴봐야 하는 부분에 대해서 다루겠습니다.

784개의 입력과 10개의 출력이 있는 신경망에서 은닉층이 하나 있고 은닉층에 50개의 노드가 있는 간단한 신경망을 만들어보겠습니다.



## 05. 신경망 데이터의 행렬 특징

### ● 신경망

10행 1열 행렬로 2번째 값이 1이고 나머지가 0인 행렬 A

```
np.array([[0],[1],[0],[0],[0],[0],[0],[0],[0],[0]])
```

A 행렬의 행과 열을 바꿔서 1행 10열 행렬로 나타낸 A의 전치행렬

```
np.array([[0,1,0,0,0,0,0,0,0,0]])
```

$$\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}^T$$

$$= [0 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]$$

## 6. 무작위 데이터로 신경망함수 만들기

## 06. 무작위 데이터로 신경망함수 만들기

### ● 신경망함수의 파라미터들

지금까지 공부한 것으로 신경망함수를 만들 수 있습니다. 단, 학습되지 않은 신경망은 전혀 쓸모가 없습니다. 학습한다는 것은  $W1$ ,  $W2$ ,  $B1$ ,  $B2$  행렬에 담긴 값들을 변화시켜가면서 정답에 가까워지게 한다는 것입니다. 즉,  $W1$ ,  $W2$ ,  $B1$ ,  $B2$  행렬에 있는 모든 수치를 변화시켜야 한다는 뜻인데,  $50 \times 784 + 10 \times 50 + 50 + 10$ 인 총 39,760개의 수치를 변화시켜야 합니다. 즉, 39,760개의 수치들이 적절한 값을 가지면  $28 \times 28$  크기의 손글씨 영상인 784개의 픽셀 값을 받아서 출력인  $Y0$ 에서  $Y9$ 까지, 0에서 9 사이의 답에 1을 넣고 나머지에 0을 출력하게 된다는 것입니다.

## 06. 무작위 데이터로 신경망함수 만들기

### ● 신경망함수의 파라미터들

입력 784 개 (고정)

은닉 50 개 (고정) :  $w_{ij}$  784\*50,  $b_i$  50

$$v_0 = w_{0_0}x_0 + w_{0_1}x_1 + w_{0_2}x_2 + \dots + w_{0_{783}}x_{783} + b_0$$

$$v_1 = w_{1_0}x_0 + w_{1_1}x_1 + w_{1_2}x_2 + \dots + w_{1_{783}}x_{783} + b_1$$

...

$$v_{49} = w_{49_0}x_0 + w_{49_1}x_1 + w_{49_2}x_2 + \dots + w_{49_{783}}x_{783} + b_{49}$$

출력 10 개 (고정) :  $w_{ij}$  50\*10  $b_i$  10

$$Y_0 = w_{2_0_0}v_0 + w_{2_0_1}v_1 + w_{2_0_2}v_2 + \dots + w_{2_0_{49}}v_{49} + b_{2_0}$$

$$Y_1 = w_{2_1_0}v_0 + w_{2_1_1}v_1 + w_{2_1_2}v_2 + \dots + w_{2_1_{49}}v_{49} + b_{2_1}$$

...

$$Y_9 = w_{2_9_0}v_0 + w_{2_9_1}v_1 + w_{2_9_2}v_2 + \dots + w_{2_9_{49}}v_{49} + b_{2_9}$$

