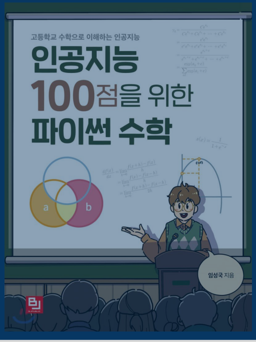
13. 역전파

인공지능 100점을 위한 파이썬 수학



Contents

1. 계산그래프

01. 계산그래프

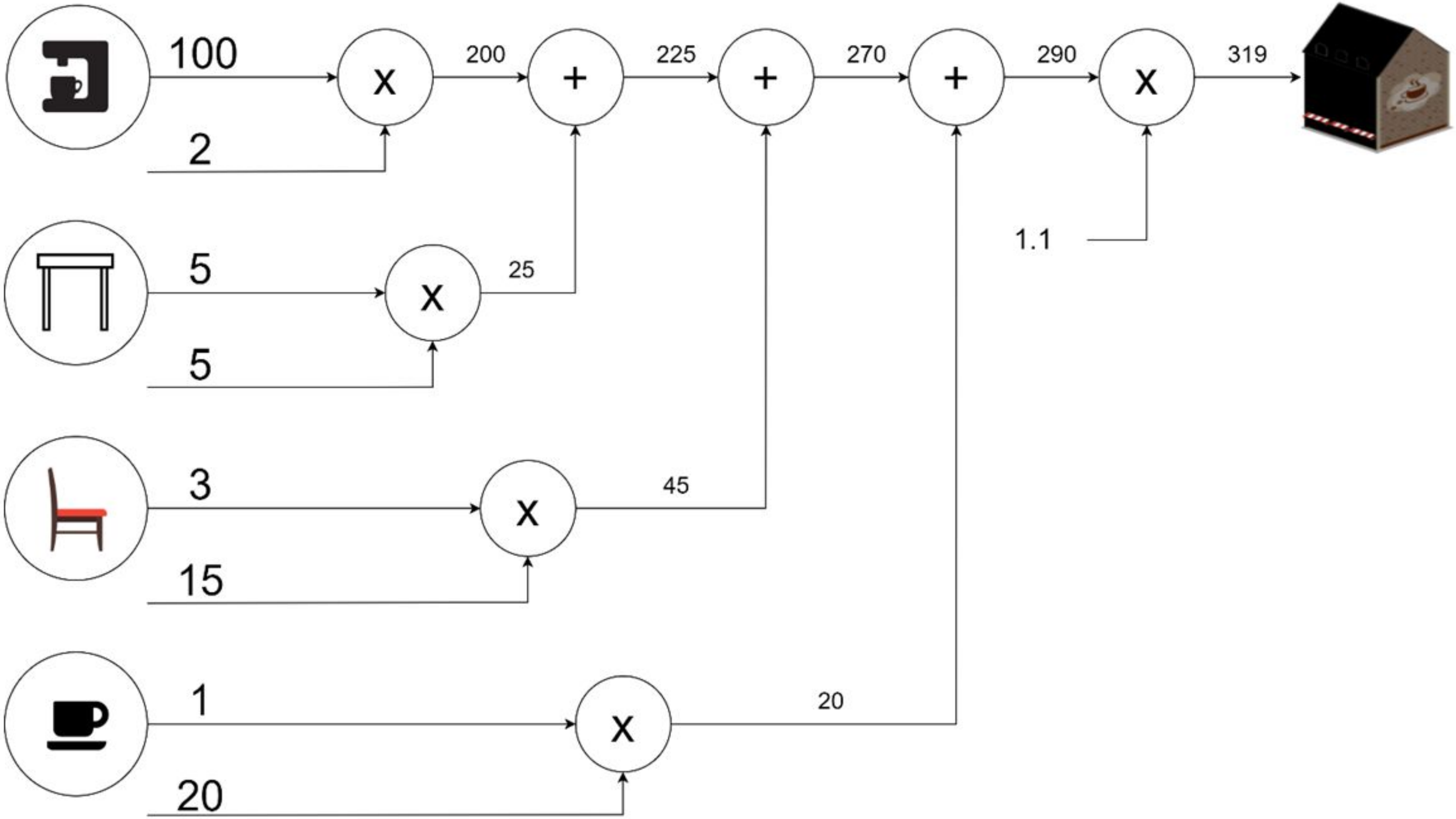
계산 그래프

계산그래프는 계산 과정을 그래프로 나타낸 것으로 원으로 표시되는 정점 (node), 정점과 정점을 연결하는 간선(edge)으로 구성됩니다. 보통 정점이라는 말 대신 노드를 그냥 사용하는 경우가 많아서 앞으로는 노드와 간선이라는 말을 사용하겠습니다. 그래프를 사용하는 이유로 빠른 이해가 있습니다. 그래프를 이용하여 표기하면 보다 직관적으로 이해할 수 있습니다. 간단한 문제를 풀어보겠습니다.

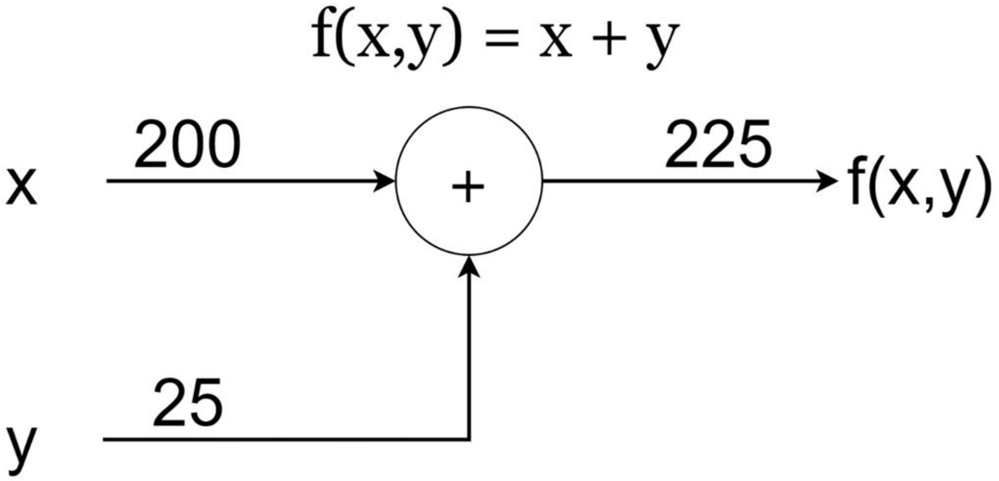
계산 그래프

길동은 카페 창업을 하기 위해 100만 원짜리 커피머신 2개, 5만 원 테이블 5개, 3만 원 의자 15개, 1만 원 컵 20개를 구입했습니다. 세금 10%를 포함해서 창업에 든 비용을 계산해 봅시다. 이 문제를 계산그래프로 표현하면 다음과 같습니다.

01. 계산그래프



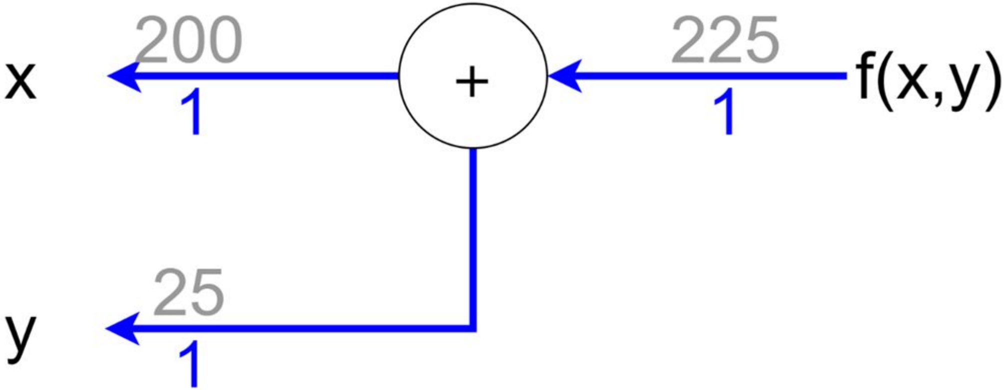
덧셈노드 순전파



$$f(x,y) = x + y \quad \Rightarrow \quad \frac{\partial f}{\partial x} = 1, \quad \frac{\partial f}{\partial y} = 1$$

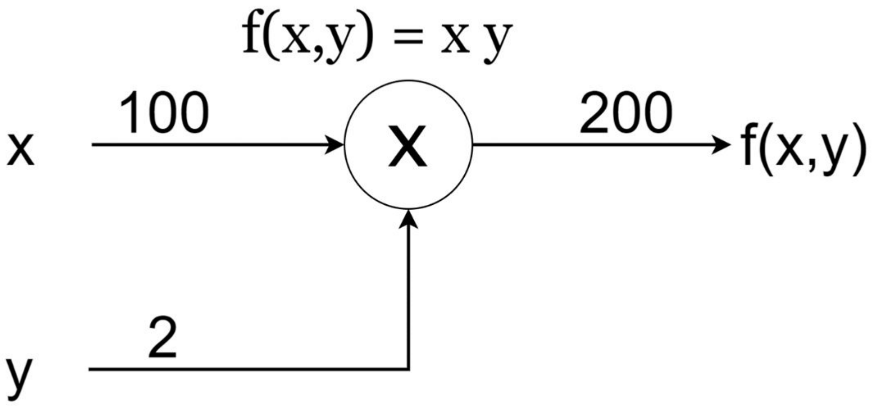
01. 계산그래프

덧셈노드 역전파



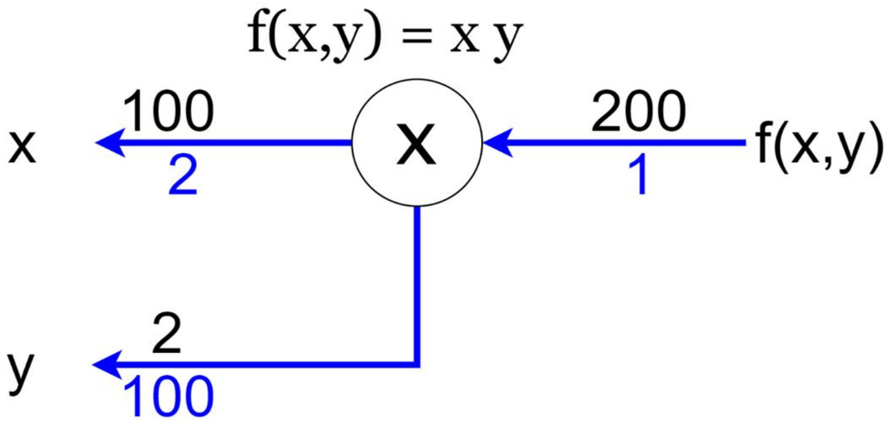
$$f(x,y) = x + y \quad \Rightarrow \quad \frac{\partial f}{\partial x} = 1, \quad \frac{\partial f}{\partial y} = 1$$

곱셈노드 순전파



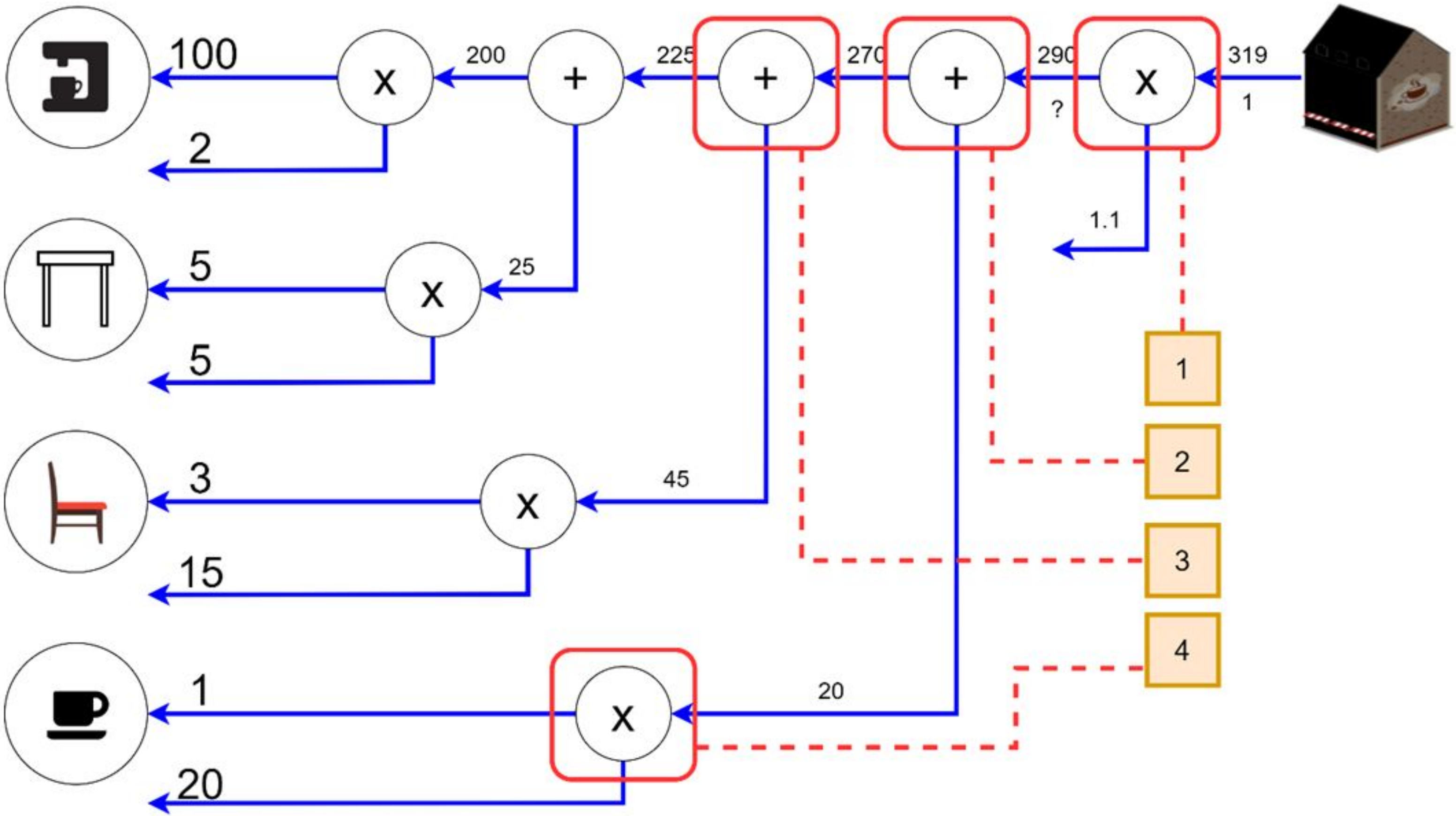
$$f(x,y) = xy \quad \Rightarrow \quad \frac{\partial f}{\partial x} = y, \quad \frac{\partial f}{\partial y} = x$$

곱셈노드 역전파



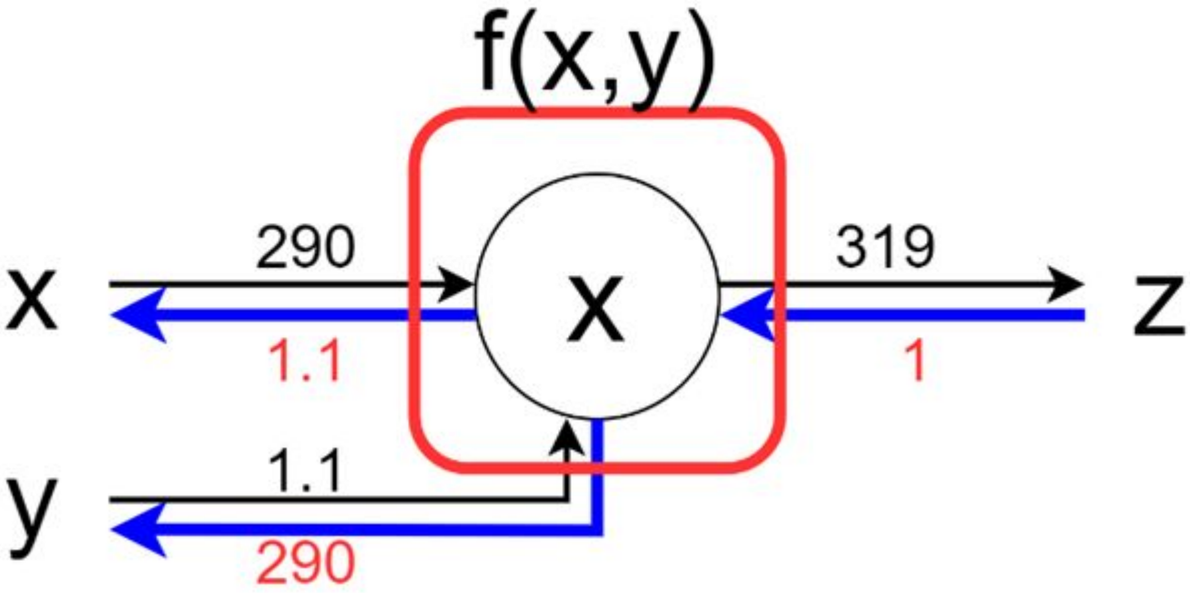
$$f(x,y) = xy \quad \Rightarrow \quad \frac{\partial f}{\partial x} = y, \quad \frac{\partial f}{\partial y} = x$$

01. 계산그래프

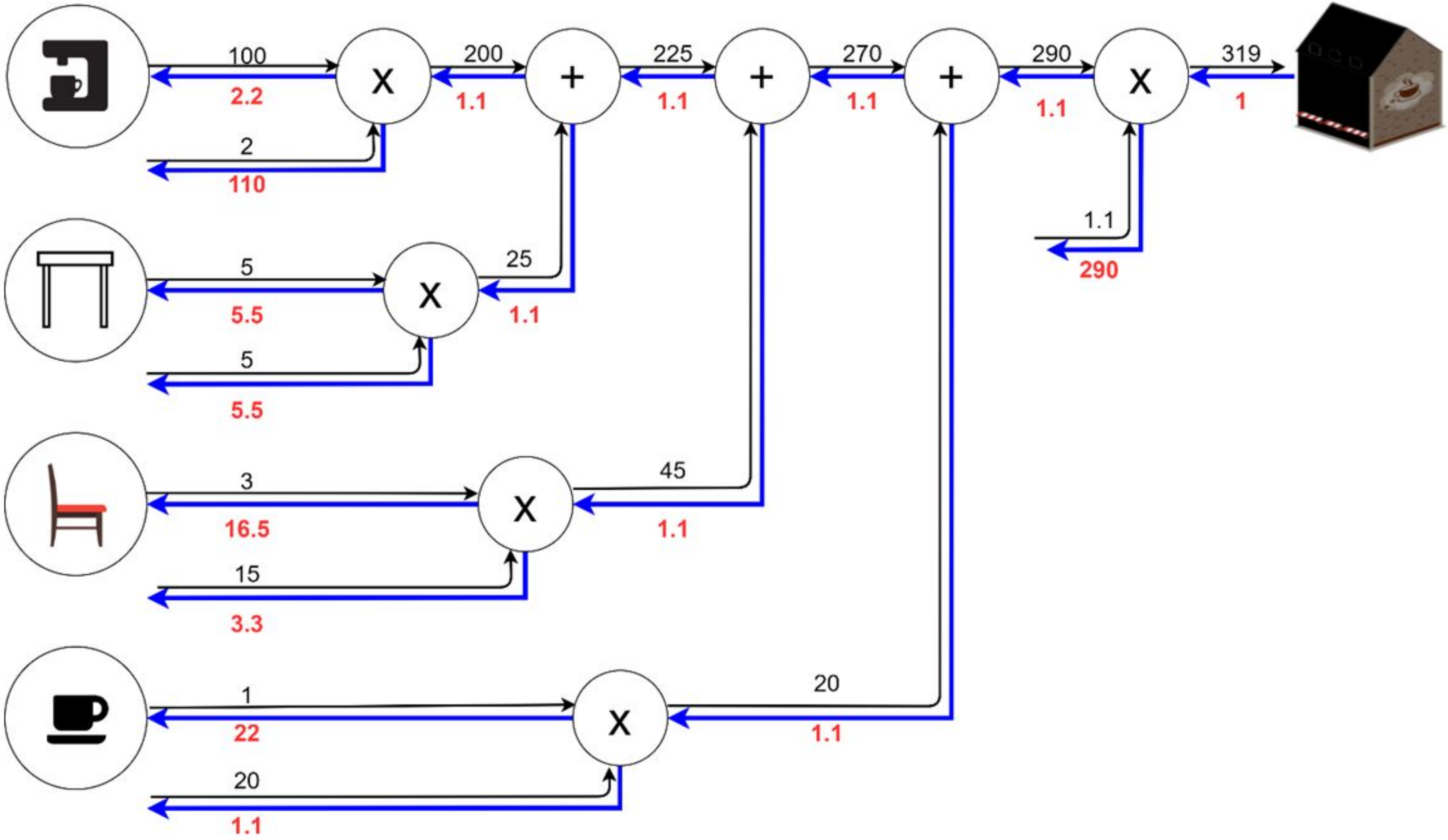


01. 계산그래프

카페창업 #1 - 곱셈노드 & 역전파



01. 계산그래프



100
2.2
2
110
X
200
1.1
+
225
1.1
+
270
1.1
+
290
1.1
X
319
1
1.1
290
5
5.5
5
5.5
X
25
1.1
3
16.5
15
3.3
X
45
1.1
1
22
20
1.1
X
20
1.1

01. 계산그래프

체인룰(Chain Rule)

덧셈, 곱셈 등과 위의 미분함수들은 개별적으로는 그다지 어렵지 않습니다. 하지만 서로 복합적으로 연결되면 복잡하게 보이게 됩니다. 이런 복잡한 함수를 단순화시키기 위해서 체인룰을 사용합니다.

$$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}$$

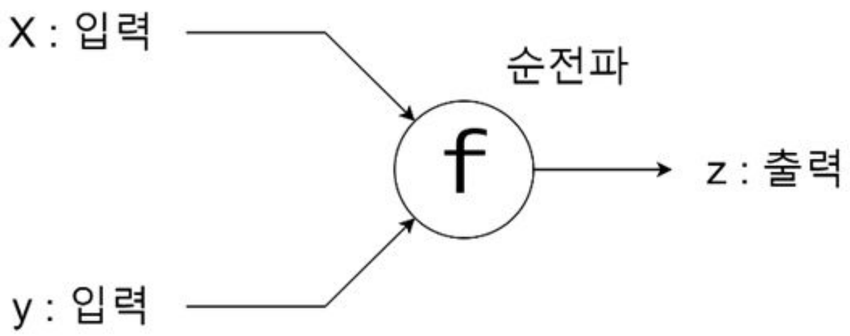
$$F'(x) = f'(g(x))\, g'(x)$$

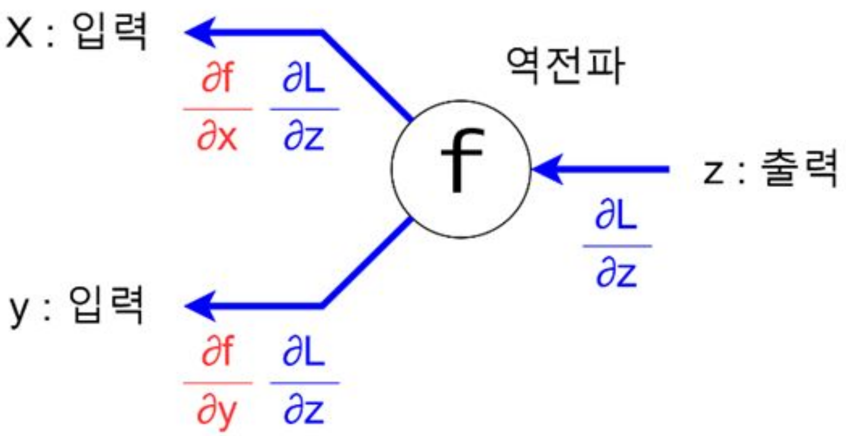
Differentiate outer function

Differentiate inner function

01. 계산그래프

함수의 기울기와 계산 그래프

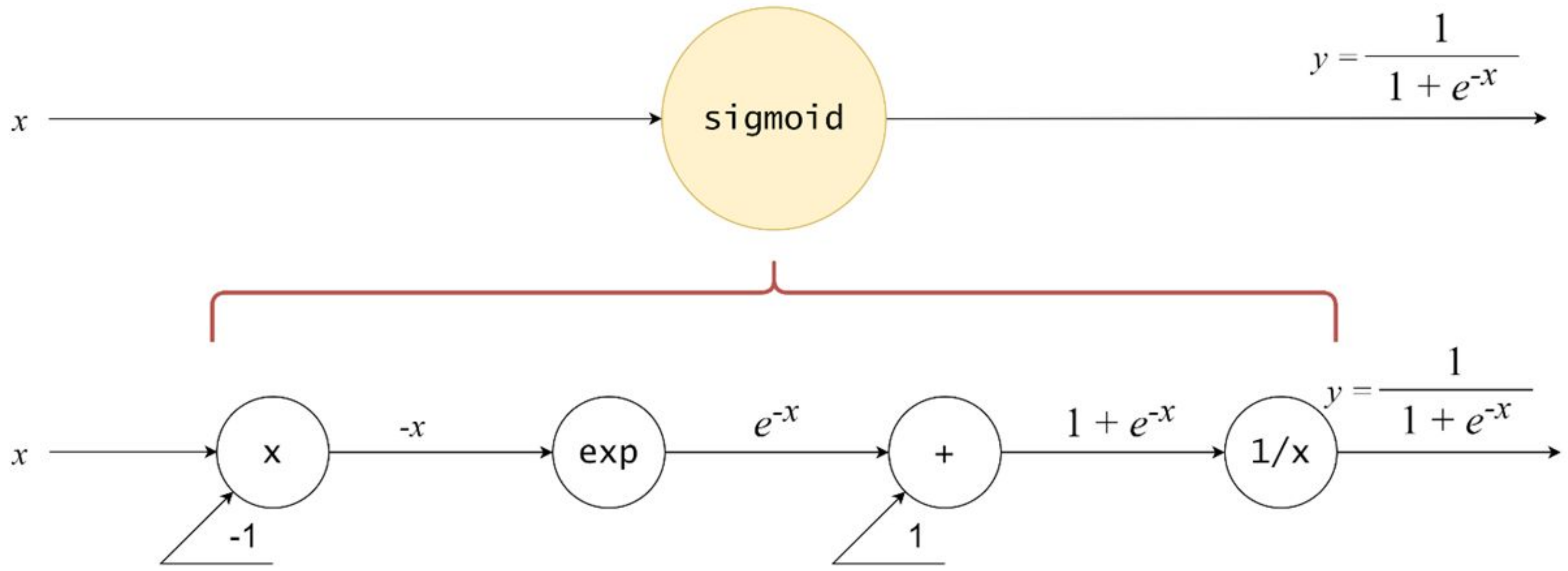




2. 시그모이드 기울기

02. 시그모이드 기울기

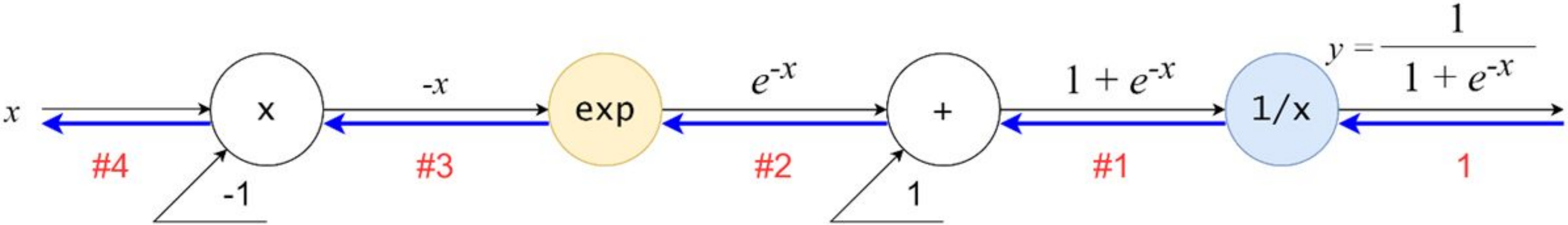
시그모이드



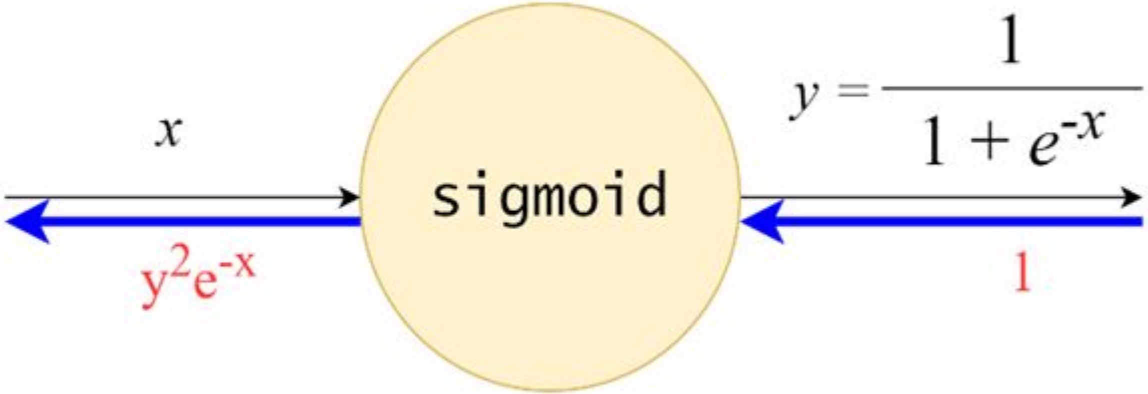
02. 시그모이드 기울기

$f(x) = e^x$ .... $f`(x) = e^x$

시그모이드



$$\#1 = -y^2$$
$$\#2 = -y^2$$
$$\#3 = -y^2 \times e^{-x}$$
$$\#4 = y^2 \times e^{-x}$$



02. 시그모이드 기울기

시그모이드

$$\begin{aligned} y^2 e^{-x} &= \frac{1}{(1+e^{-x})^2} e^{-x} \\ &= \frac{1}{1+e^{-x}} \frac{e^{-x}}{1+e^{-x}} \\ &= \frac{1}{1+e^{-x}} \left(\frac{1+e^{-x}}{1+e^{-x}} - \frac{1}{1+e^{-x}} \right) \\ &= y(1-y) \end{aligned}$$

#1 = $-y^2$

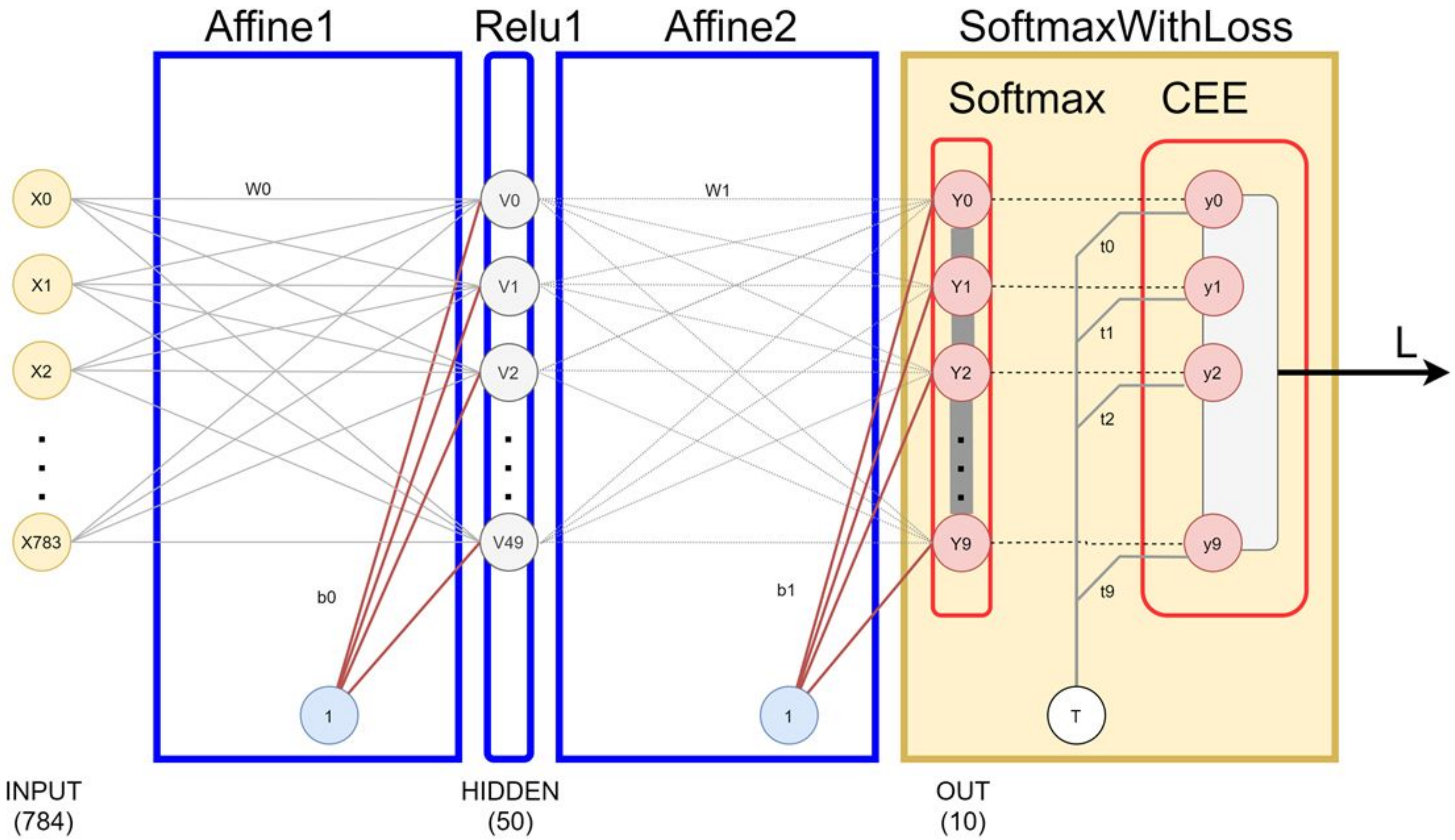
#2 = $-y^2$

#3 = $-y^2 \times e^{-x}$

#4 = $y^2 \times e^{-x}$

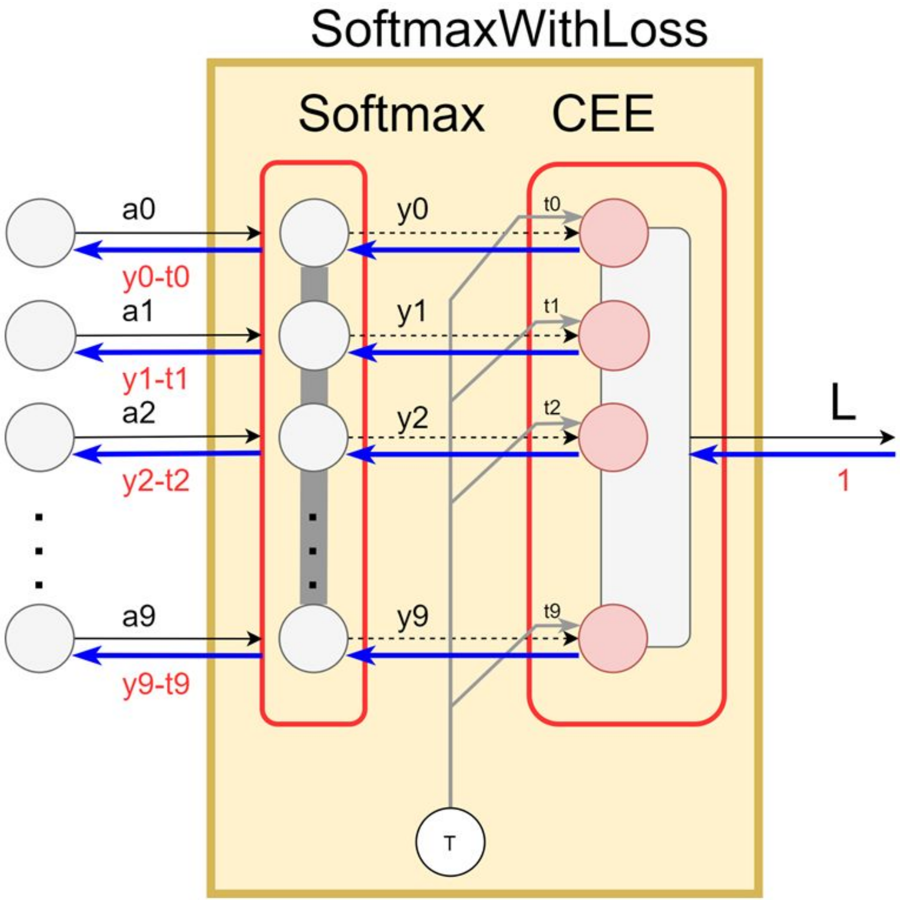
3. 소프트맥스와 CEE

03. 소프트맥스와 CEE



03. 소프트맥스와 CEE

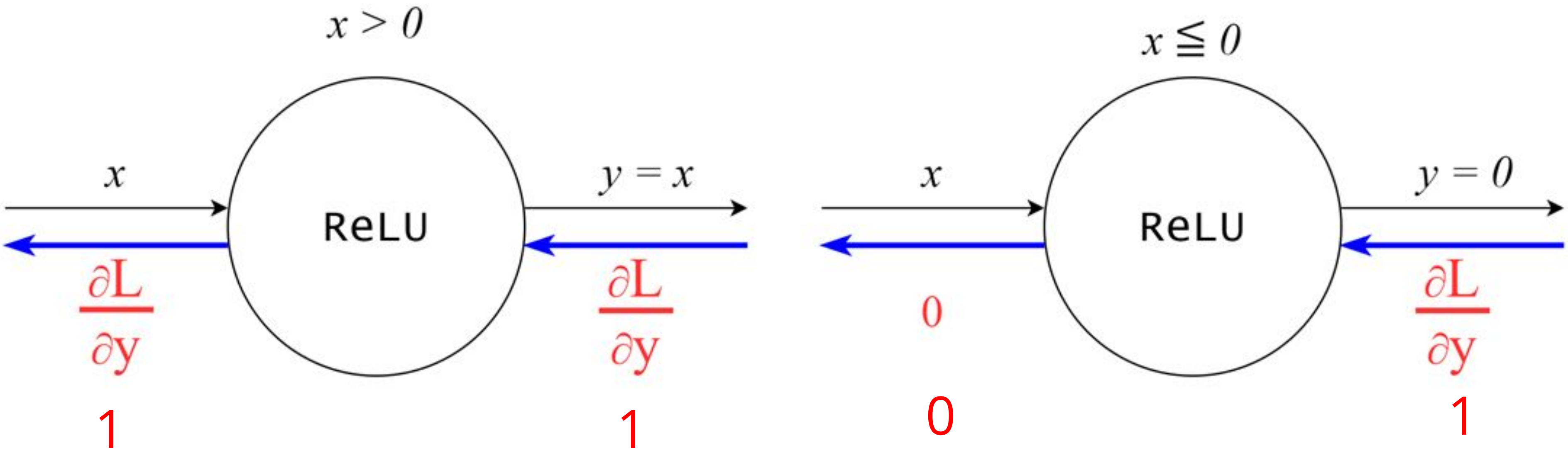
소프트맥스 역전파



4. 활성함수

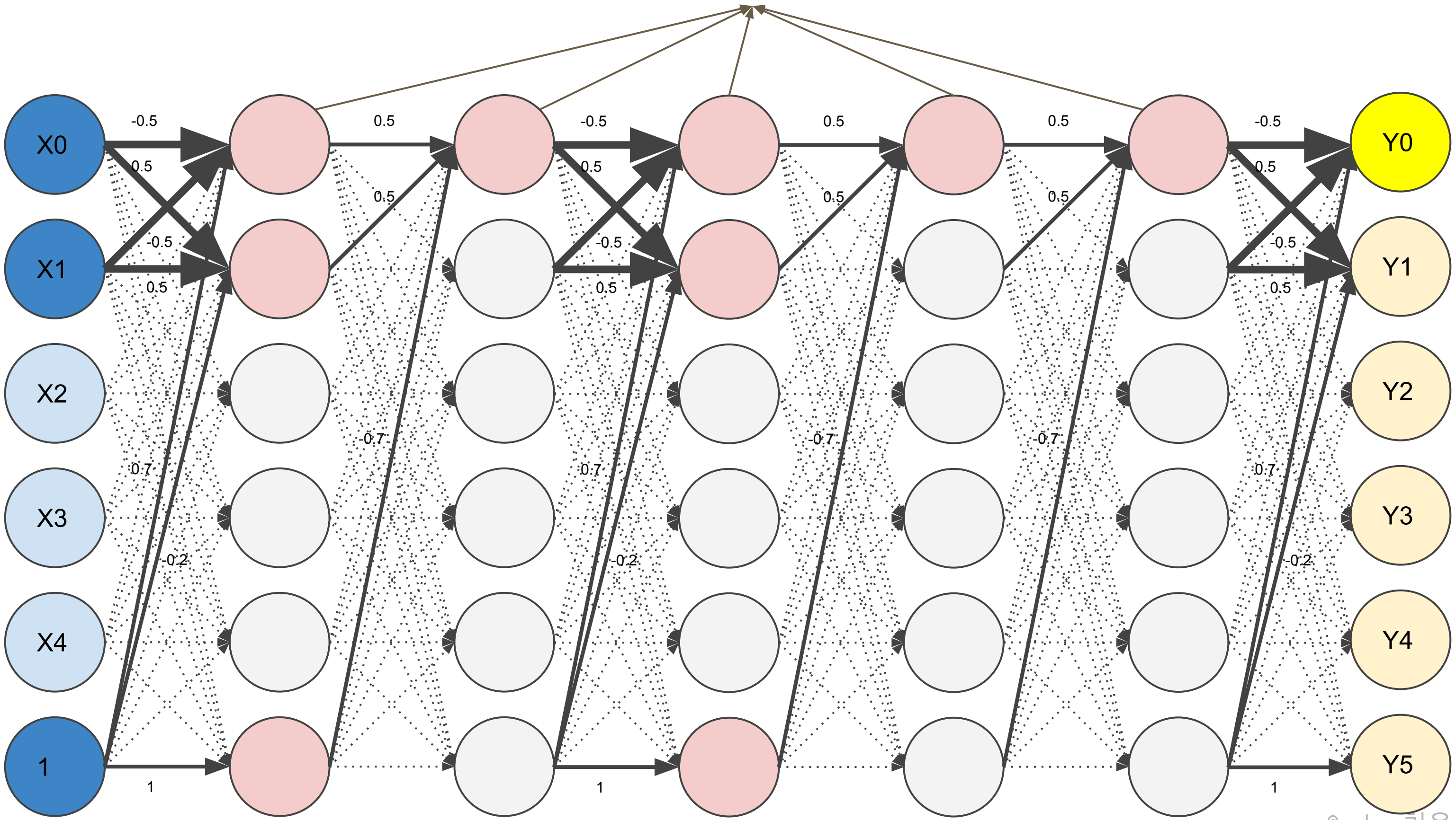
04. 활성함수

ReLu 역전파



5. Affine

Affine 계층/레이어



@echo 처음코딩 in YouTube

Affine 계층은 ...
행렬 내적을 기하학에서 어파인 변환(affine transformation)이라고 하고,
어파인 변환을 수행하는 처리를 Affine 계층이라는 이름으로 만든다.
즉, 이전 계층의 모든 뉴런과 연결되어 있어 행렬의 내적(np.dot())을
사용하여 계산하는 계층/레이어를 Affine 계층/레이어라 부른다.

```
>>>import numpy as np
>>>X = np.random.rand(2)
>>>W = np.random.rand(2,3)
>>>B = np.random.rand(3)
>>>Y = np.dot(X,W) + B
```

Affine 계층/레이어

행렬계산 복습 (순전파)

$$X \cdot W = O$$

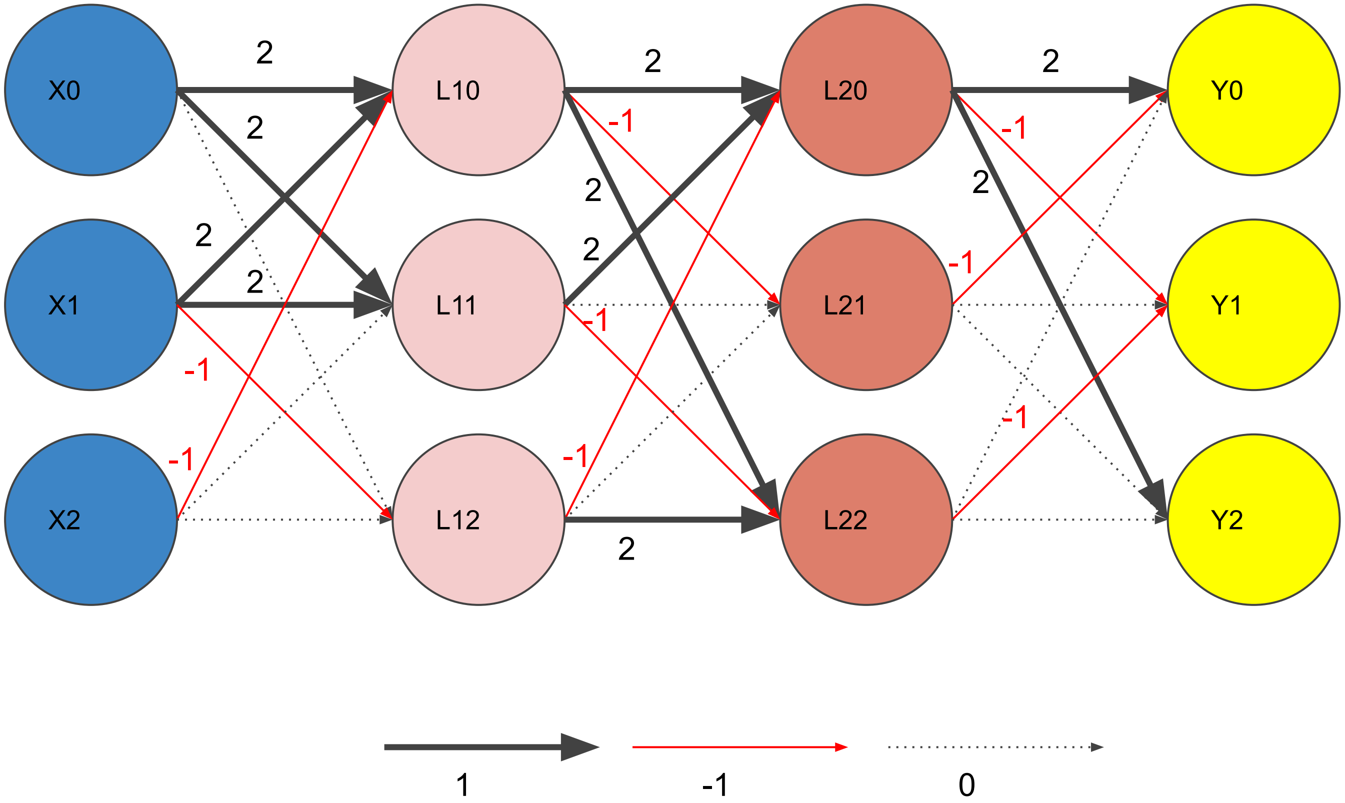
$$(2,) \qquad (2,3) \qquad (3,)$$

Affine 계층/레이어

행렬계산 복습 (순전파)

$$X \cdot W = O$$

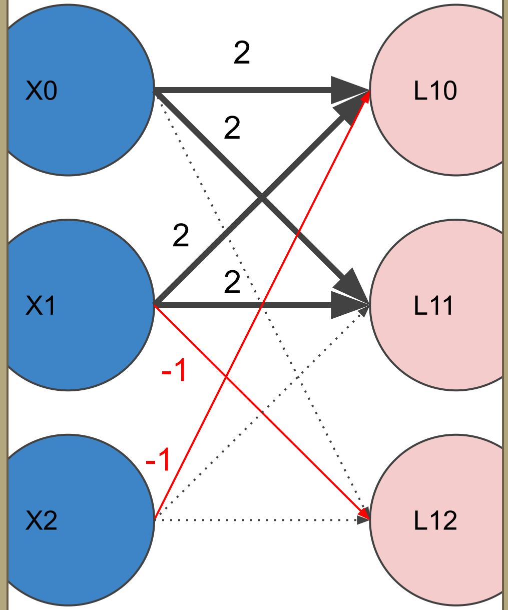
(2,)	(2,3)	(3,)
(1,2)	(2,3)	(1,3)


X0
X1
X2
L10
L11
L12
L20
L21
L22
Y0
Y1
Y2
2
-1
1
-1
0

행렬식과 신경망

입력 :
X0, X1, X2

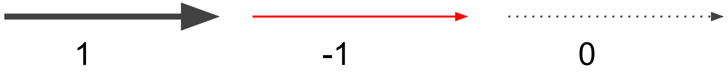
출력 :
L10, L11, L12



L10 = 2X0+2X1−X2

L11 = 2X0 + 2X1

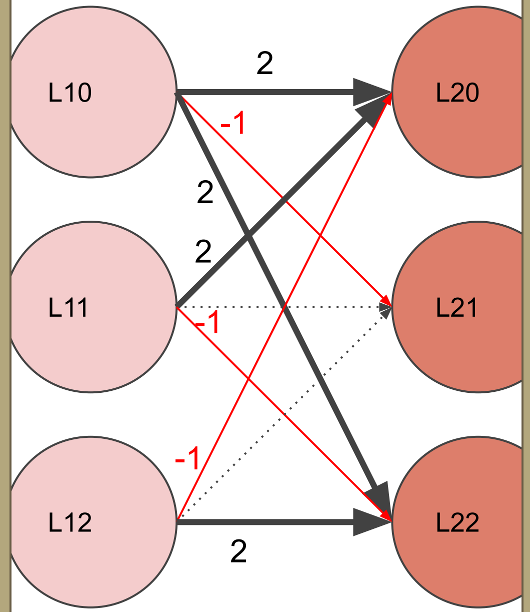
L12 = −X1 − X2



행렬식과 신경망

입력 : L10, L11, L12

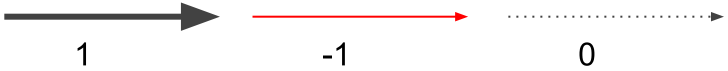
출력 : L20, L21, L22



$L20 = 2L10 + 2L11 - L12$

$L21 = -L10$

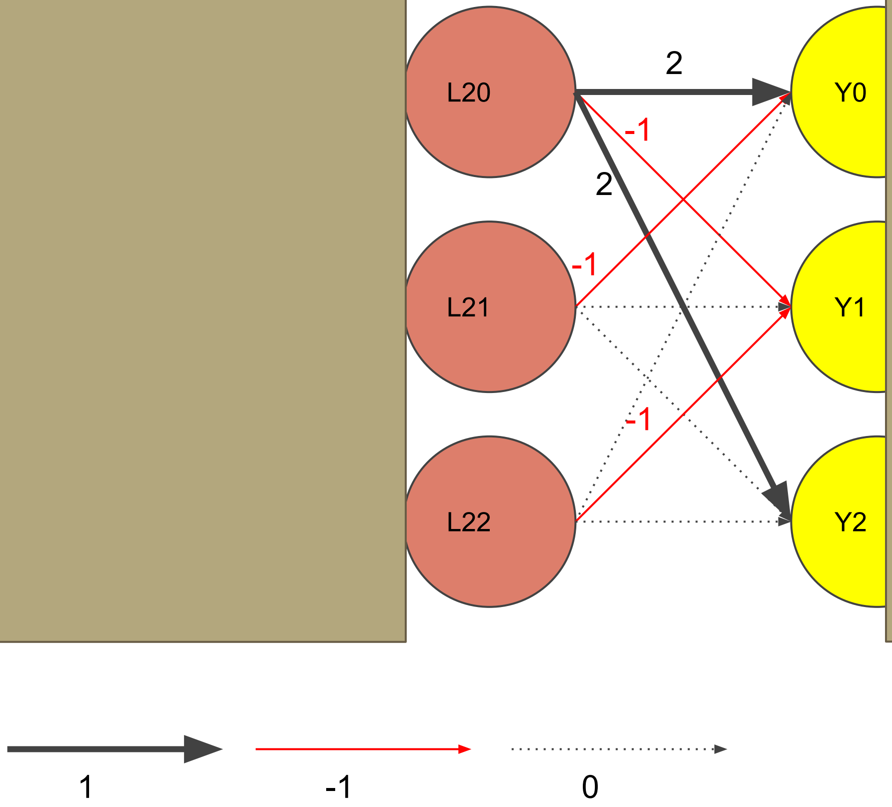
$L22 = 2L10 - L11 + 2L12$



행렬식과 신경망

입력 : L20, L21, L22

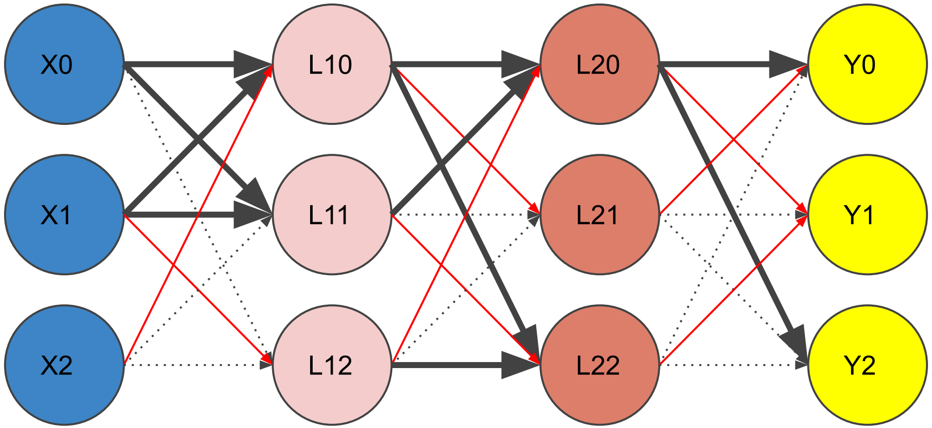
출력 : Y0, Y1, Y2



$Y0 = 2L20 - L21$

$Y1 = -L20 - L21$

$Y2 = 2L20$



L10 = 2X0+2X1-X2
L11 = 2X0 + 2X1
L12 = -X1 - X2

L20 = 2L10 + 2L11 - L12
L21 = -L10
L22 = 2L10 - L11 + 2L12

Y0 = 2L20 - L21
Y1 = -L20 - L21
Y2 = 2L20

L10 = 2X0+2X1-X2
L11 = 2X0 + 2X1
L12 = -X1 - X2

L20 = 2L10 + 2L11 - L12
L21 = -L10
L22 = 2L10 - L11 + 2L12

Y0 = 2L20 - L21
Y1 = -L20 - L21
Y2 = 2L20

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

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

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

출력 세번째 계산 두번째 계산 첫번째 계산 입력

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

```
L10 = 2X0+2X1-X2
L11 = 2X0 + 2X1
L12 = -X1 - X2

L20 = 2L10 + 2L11 - L12
L21 = -L10
L22 = 2L10 - L11 + 2L12

Y0 = 2L20 - L21
Y1 = -L20 - L21
Y2 = 2L20
```

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

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

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

```
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)
```

→

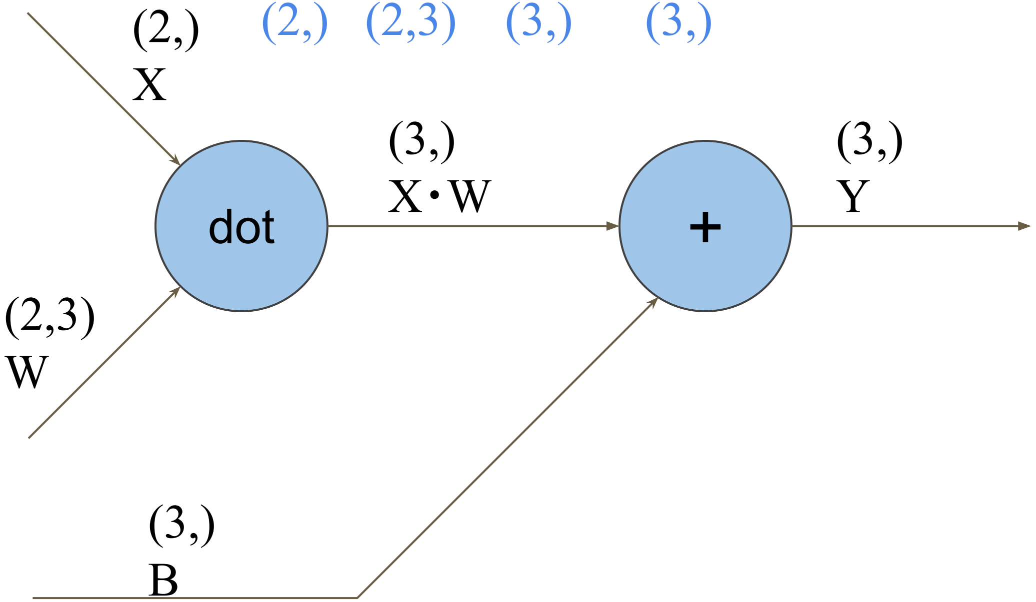
```
Y = np.dot(W3, np.dot(W2, np.dot(W1,X)))
```

←

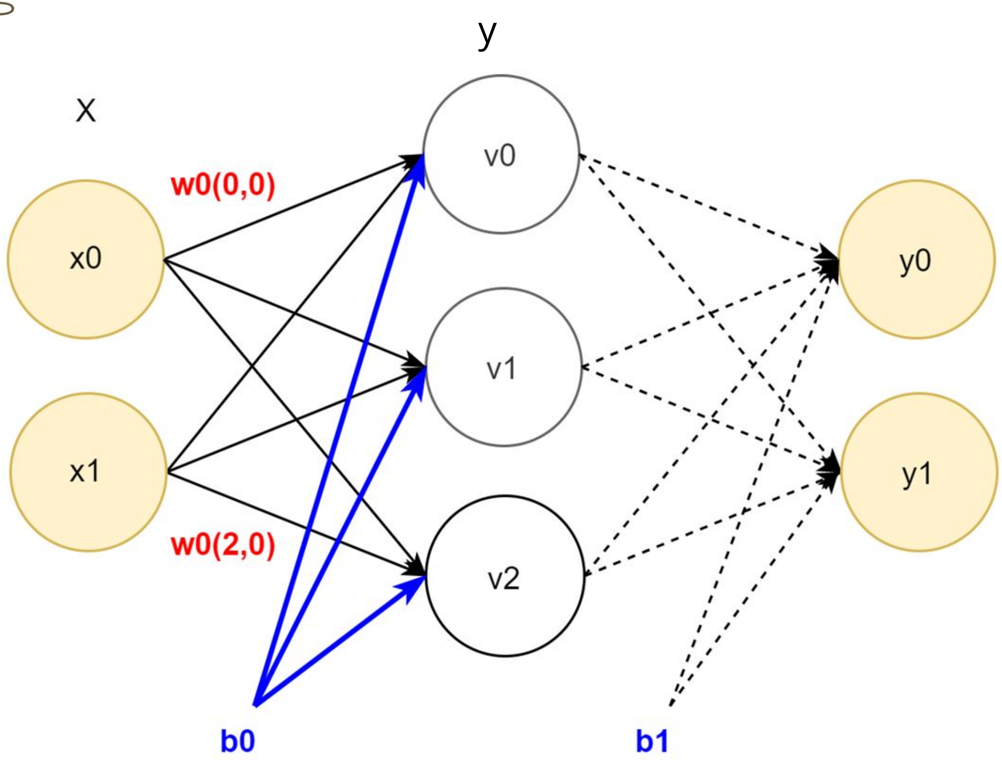
Affine 계층/레이어 (순전파)

$$X \cdot W + B = Y$$

$$(2,)\quad(2,3)\quad(3,)\quad(3,)$$

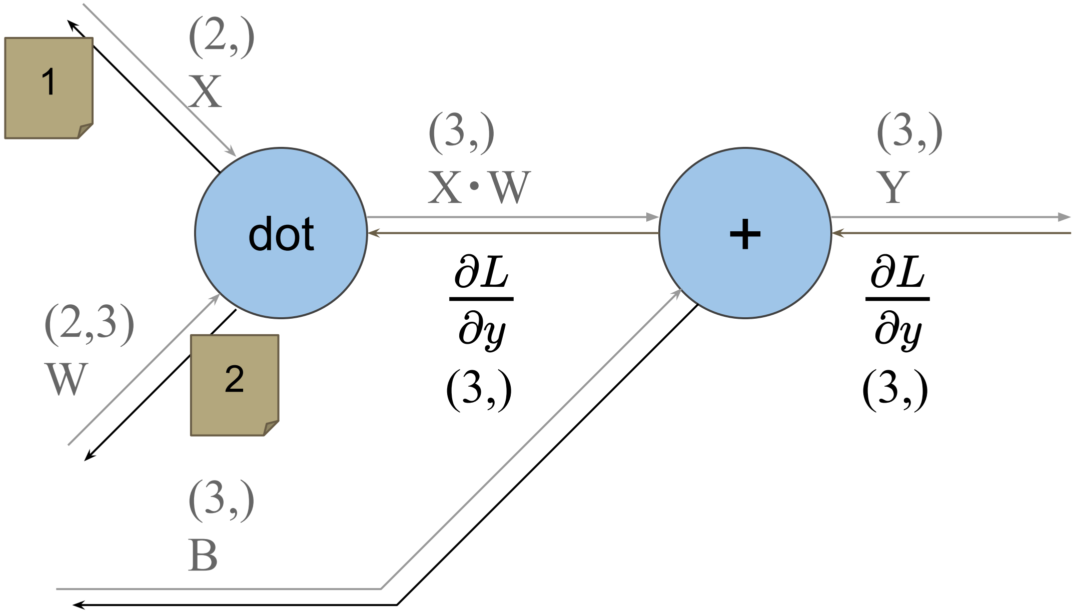


어파인 계층



Affine 계층/레이어

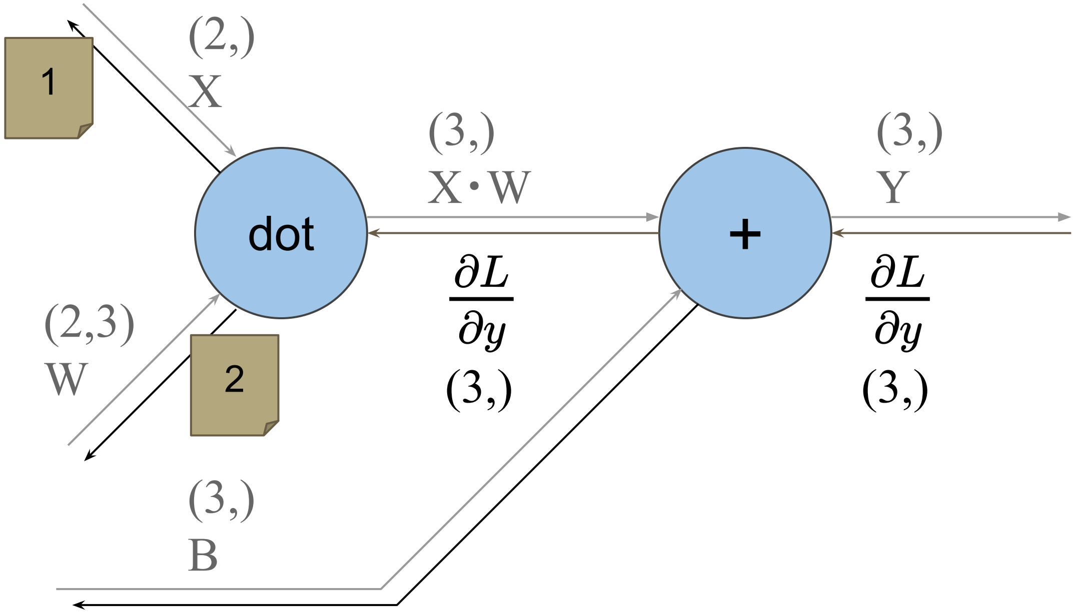
$$X \cdot W + B = Y$$
$$(2,) \quad (2,3) \quad (3,) \quad (3,)$$



@echo 처음코딩 in YouTube

Affine 계층/레이어

$$\underset{(2,)}{X} \cdot \underset{(2,3)}{W} + \underset{(3,)}{B} = \underset{(3,)}{Y}$$



1

$$\underset{(1,2)}{\frac{\partial L}{\partial X}} = \underset{(1,3)}{\frac{\partial L}{\partial Y}} \cdot \underset{(3,2)}{W^T}$$

2

$$\underset{(2,3)}{\frac{\partial L}{\partial W}} = \underset{(2,1)}{X^T} \cdot \underset{(1,3)}{\frac{\partial L}{\partial Y}}$$

전치행렬

$$W = \begin{pmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \end{pmatrix}$$

$$W^T = \begin{pmatrix} w_{11} & w_{21} \\ w_{12} & w_{22} \\ w_{13} & w_{23} \end{pmatrix}$$

전치행렬

$$\frac{\partial L}{\partial X} = \frac{\partial L}{\partial Y} \cdot W^T$$

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

$$\begin{pmatrix} lx_{11} & lx_{12} \end{pmatrix} = \begin{pmatrix} ly_{11} & ly_{12} & ly_{13} \end{pmatrix} \begin{pmatrix} w_{11} & w_{21} \\ w_{12} & w_{22} \\ w_{13} & w_{23} \end{pmatrix}$$

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

전치행렬

$$\frac{\partial L}{\partial X} = \frac{\partial L}{\partial Y} \cdot W^T$$

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

$$\begin{pmatrix} lx_{11} & lx_{12} \end{pmatrix} = \begin{pmatrix} ly_{11} & ly_{12} & ly_{13} \end{pmatrix} \begin{pmatrix} w_{11} & w_{21} \\ w_{12} & w_{22} \\ w_{13} & w_{23} \end{pmatrix}$$

$$= \begin{pmatrix} ly_{11}w_{11} + ly_{12}w_{12} + ly_{13}w_{13} & ly_{11}w_{21} + ly_{12}w_{22} + ly_{13}w_{23} \end{pmatrix}$$

전치행렬

$$\frac{\partial L}{\partial X} = \frac{\partial L}{\partial Y} \cdot W^T$$

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

$$\begin{pmatrix} lx_{11} & lx_{12} \end{pmatrix} = \begin{pmatrix} ly_{11} & ly_{12} & ly_{13} \end{pmatrix} \begin{pmatrix} w_{11} & w_{21} \\ w_{12} & w_{22} \\ w_{13} & w_{23} \end{pmatrix}$$

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

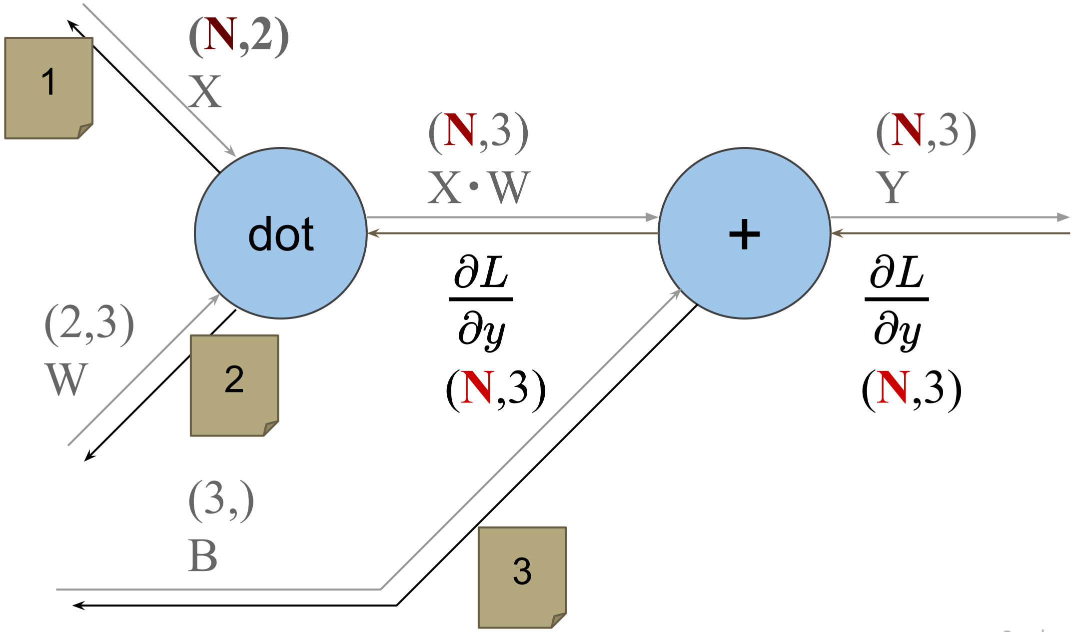
$$\frac{\partial L}{\partial W} = X^T \cdot \frac{\partial L}{\partial Y}$$

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

$$\begin{pmatrix} lw_{11} & lw_{12} & lw_{13} \\ lw_{21} & lw_{22} & lw_{23} \end{pmatrix} = \begin{pmatrix} x_{11} \\ x_{12} \end{pmatrix} \begin{pmatrix} ly_{11} & ly_{12} & ly_{13} \end{pmatrix}$$

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

N 개의 데이터를 묶어서 처리하는 (순전파) Affine 계층 - 배치 처리



6. 오차역전파

06. 오차역전파

과정별

13.6.1 [STEP1] 미분과 역전파 선택

13.6.2 [STEP2] MNIST 데이터 가져오기

13.6.3 [STEP3] 함수 정의 : 수치미분, 소프트맥스, CEE

13.6.4 [STEP4] 클래스 정의 : ReLU, Affine, SoftmaxWithLoss,

13.6.5 [STEP5] 클래스 정의 : SimpleNetwork

13.6.6 [SETP6] 학습을 위한 설정치 입력

13.6.7 [STEP7] 학습과 검증