Deep Learning Basic

Jaewon Kim, Dankook Univ.

Chapter 1-1

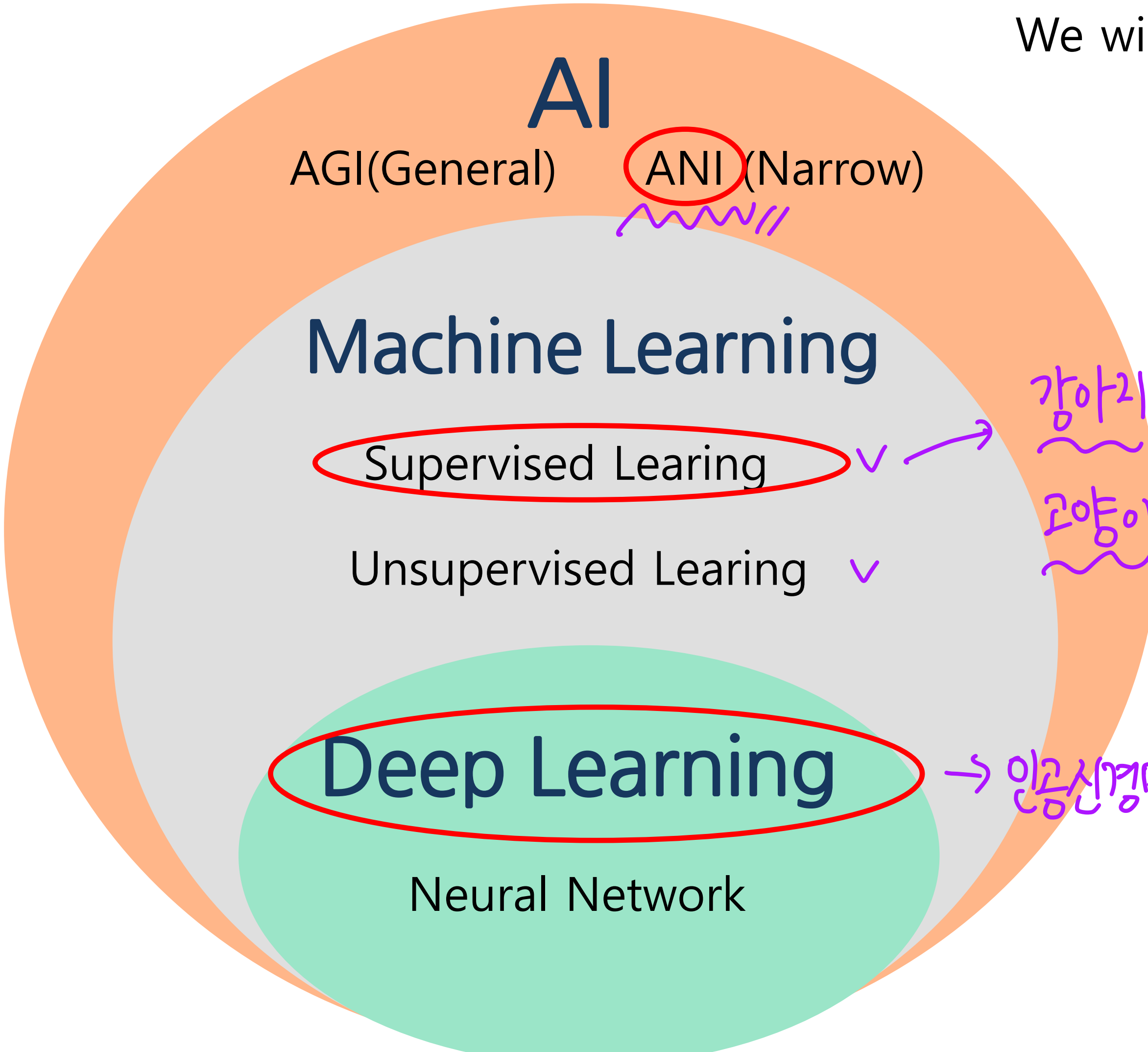
Contents

Part 1

About AI

AI vs Machine Learning vs Deep Learning

We will do



강아지
고양이

$\boxed{1}$ $\boxed{3}$ $\boxed{5}$ $\boxed{8}$

0
1
2
3
4

→ 인공신경망 (brain)

Machine Learning Categories

x, y, z

	Supervised Learing	Unsupervised Learing
Dicrete	**Classification**	**Clustering**
Continuous	**Regression**	**Dimensionality Reduction**

* Reinforcement Learning is excluded

Regression Ex.

출처 : https://finance.naver.com/item/fchart.nhn?code=020150

Classification Ex.

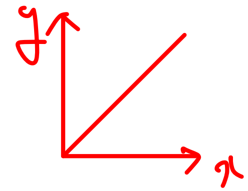
강아지
머핀

출처 : https://www.freecodecamp.org/news/chihuahua-or-muffin-my-search-for-the-best-computer-vision-api-cbda4d6b425d/

Part 2

Basic Regression

Hypothesis

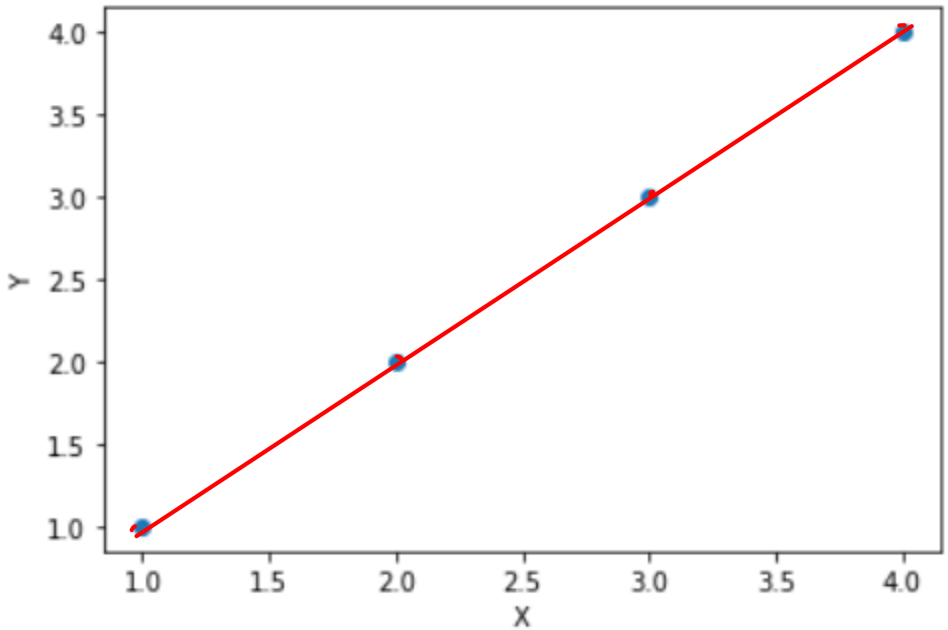


$f(x) = x$

If Data is

x	y
1	1
2	2
3	3
4	4

What is f(5) ?



Hypothesis

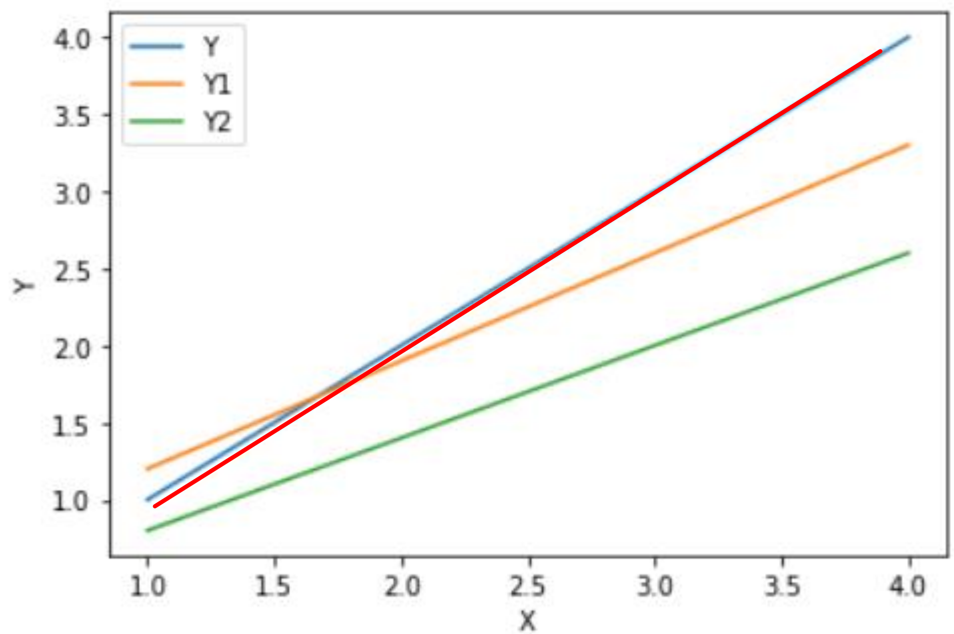
$$H(x) = wx + b$$

일차함수 기본식,,

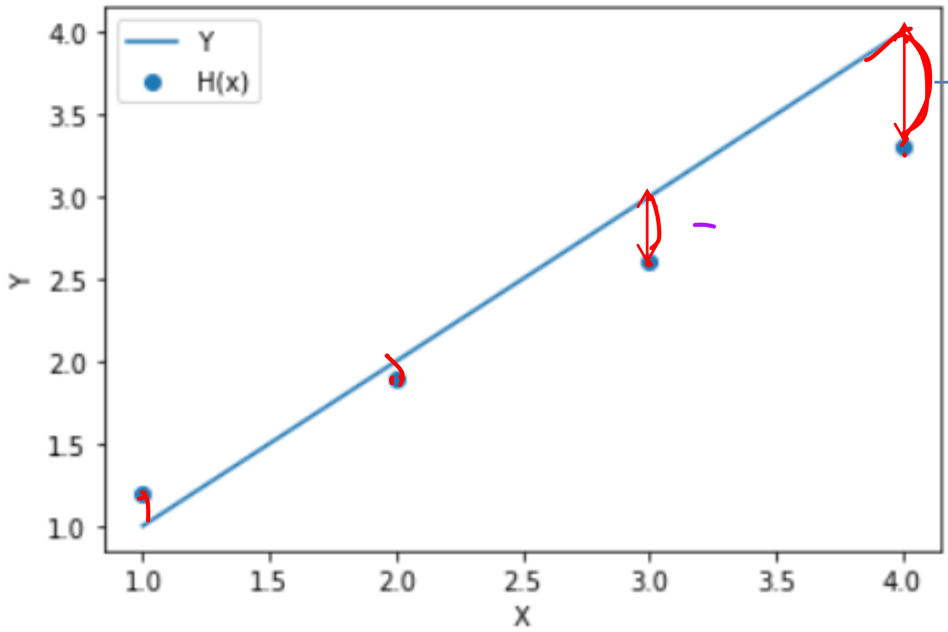
Approximate !!

w = weight,,

b = bias,,



Hypothesis



Cost :
H(x) - Y

Cost = H(x) - true_Y , H(x) = wx + b

To reduce cost, Find proper w,b

Cost Function (MSE)

부호 날리고 크기만 제곱 ,,

Mean Square Error

$$\frac{(H(x^{(1)}) - y^{(1)})^2 + (H(x^{(2)}) - y^{(2)})^2 + (H(x^{(3)}) - y^{(3)})^2}{3}$$

$$cost = \frac{1}{m} \sum_{i=1}^{m} (H(x^{(i)}) - y^{(i)})^2$$

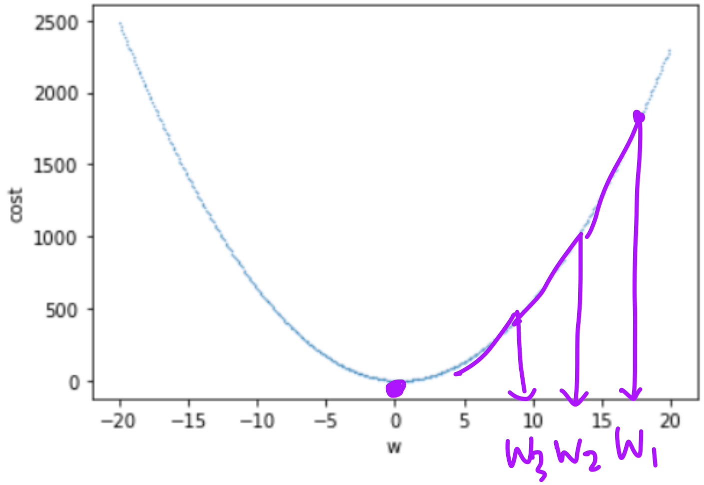
$$\underset{W,b}{\text{minimize}}\ \boxed{cost(W, b)}$$

Now, You may have an quetion !

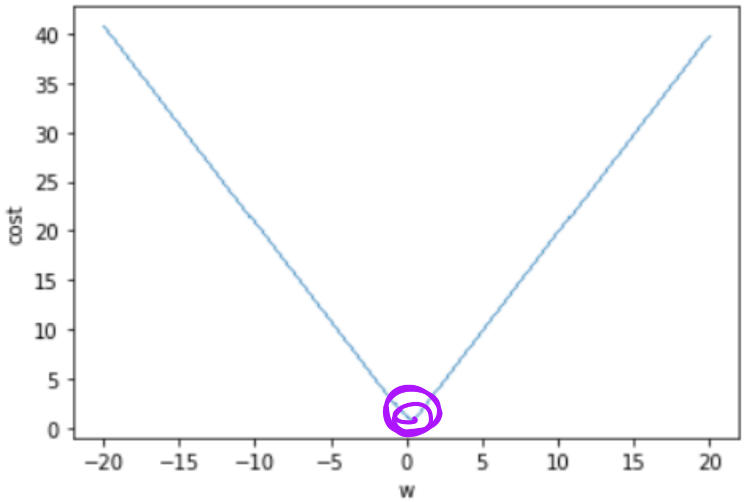
MSE vs MAE

V//

Mean Square Error //



Mean Absolute Error



Hypothesis

$$cost = \frac{1}{m}\sum_{i=1}^{m}(H(x^{(i)}) - y^{(i)})^2$$

Ex) H(x) = 13x + 1

x	2	3	4	5
Pred	25	50	42	61
True	27	40	53	66
Cost	-2	10	-9	-5

$$cost = \sum_{i=1}^{n}(H(x^i) - y(x))^2$$

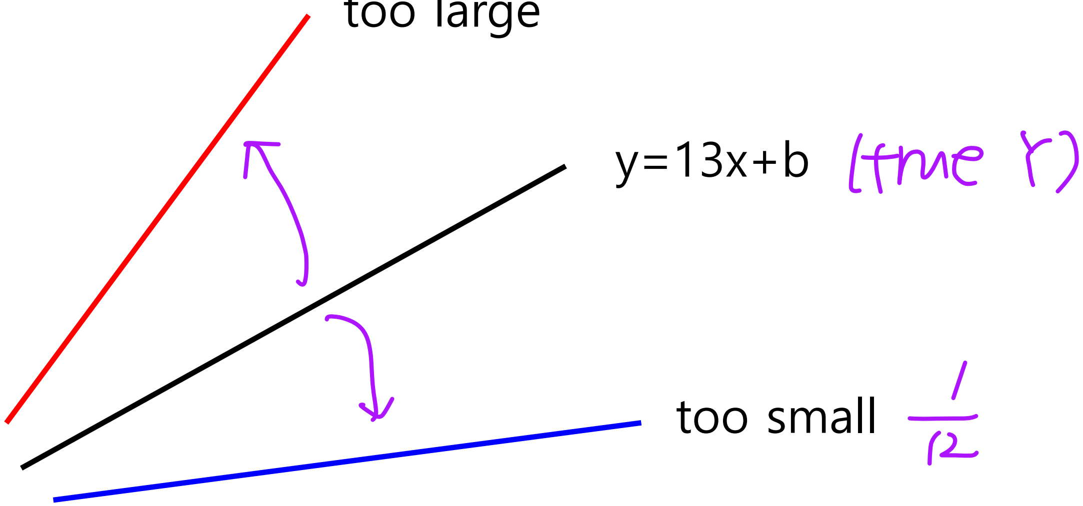
합 $= (-2)^2 + (10)^2 + (-9)^2 + (-5)^2 = 210$

$\therefore cost = 210/4 = 52.5$

Hypothesis

Optimization (Gradient Decent)

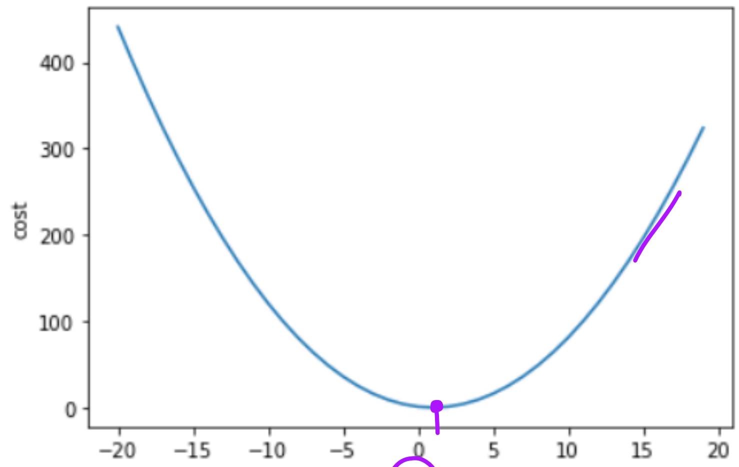
$H(x) = Wx + b$



$$\textbf{Gradient} \propto \boldsymbol{Cost}$$

Optimization (Gradient Decent)

Ex) H(x) = wx



$\frac{d}{dw}$ cost

$$Cost(w) = \frac{1}{n}\sum_{n=i}^{n}(H(x^{(i)}) - y^{(i)})^2$$

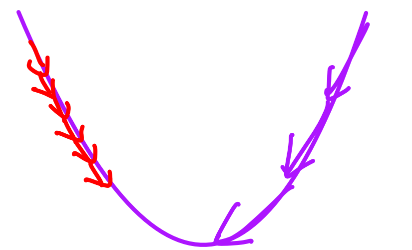
$$H(x^{(i)}) = w(x^{(i)})$$

Optimization (Gradient Decent)

$$cost(w) = \frac{1}{n}\sum_{n=i}^{n}\frac{1}{2}\left(w(x^i) - y^i\right)^2$$

$$\frac{\partial}{\partial w} cost(w) = \frac{1}{n}\sum_{n=i}^{n}\frac{1}{\cancel{2}}\left(w(x^i) - y^i\right)^2 \cdot \cancel{2} \cdot x^i$$

$$\therefore \frac{\partial}{\partial w} cost(w) = \frac{1}{n}\sum_{n=i}^{n}\left(w(x^i) - y^i\right)^2 \cdot x^i$$



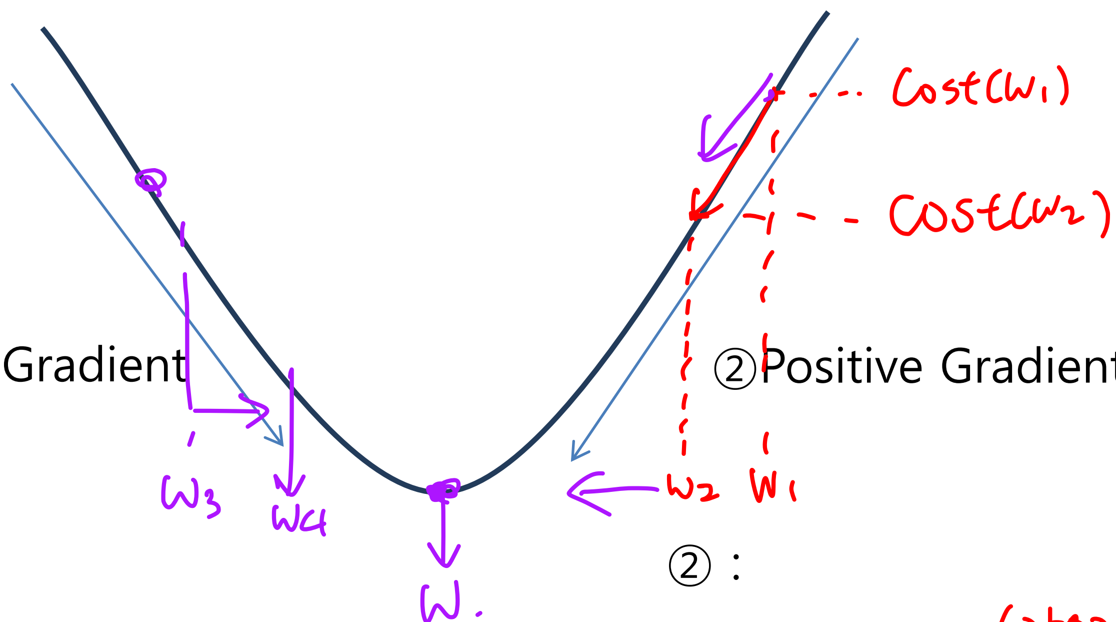
$\alpha \uparrow \downarrow$

$$W := W - \alpha\frac{1}{m}\sum_{i=1}^{m}(Wx^i - y^i)x^i \quad \alpha : Learning\ Rate$$

Optimization (Gradient Decent)

$$W := W - \alpha \frac{\partial}{\partial W} cost(W)$$

$$\frac{cost(W_1) - cost(W_2)}{W_1 - W_2}$$



①Negative Gradient

②Positive Gradient

① :

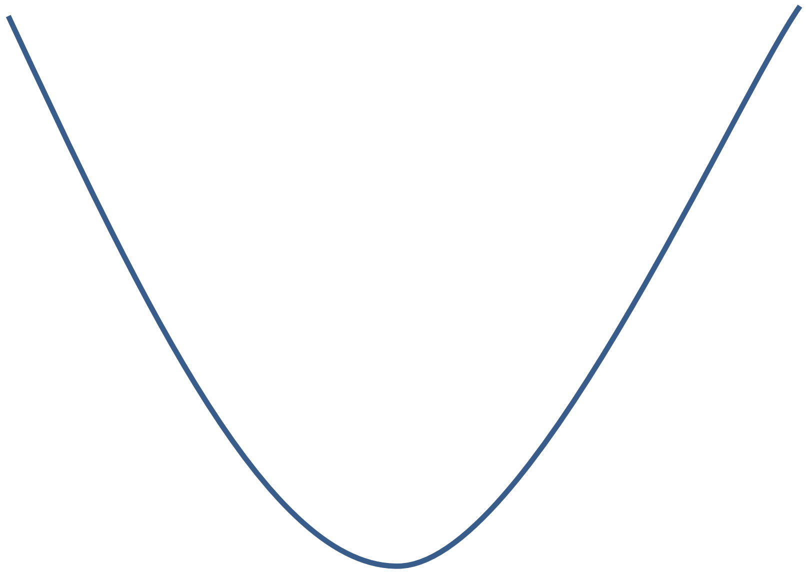
$W := W + \alpha$ (음수기울기)

② :

$W := W - \alpha$ (양수기울기)

Optimization (Gradient Decent) - Learing Rate

If Learning Rate is too large or small ?



Multivariable Linear Regression

x = RGB (3 channel) , y is 4 label.

$$H(x) = Wx + b$$

$$H(x_1, x_2, x_3) = w_1x_1 + w_2x_2 + w_3x_3 + b$$

$$cost(W, b) = \frac{1}{m}\sum_{I=1}^{m}(H(x_1^{(i)}, x_2^{(i)}, x_3^{(i)}) - y^{(i)})^2$$

$$w_1x_1 + w_2x_2 + w_3x_3$$

$$(x_1, x_2, x_3) \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}$$

Matrix

Perceptron

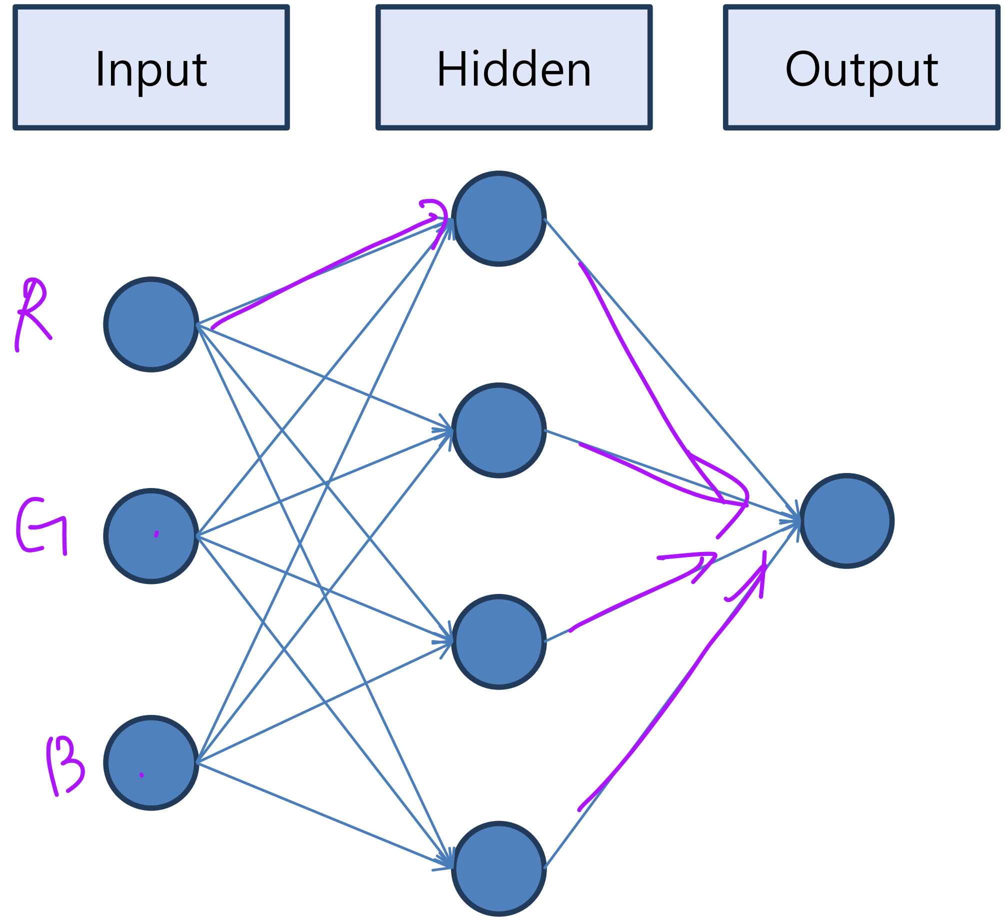
Multivariable Linear Regression - Matrix

$$w_1x_1 + w_2x_2 + w_3x_3$$

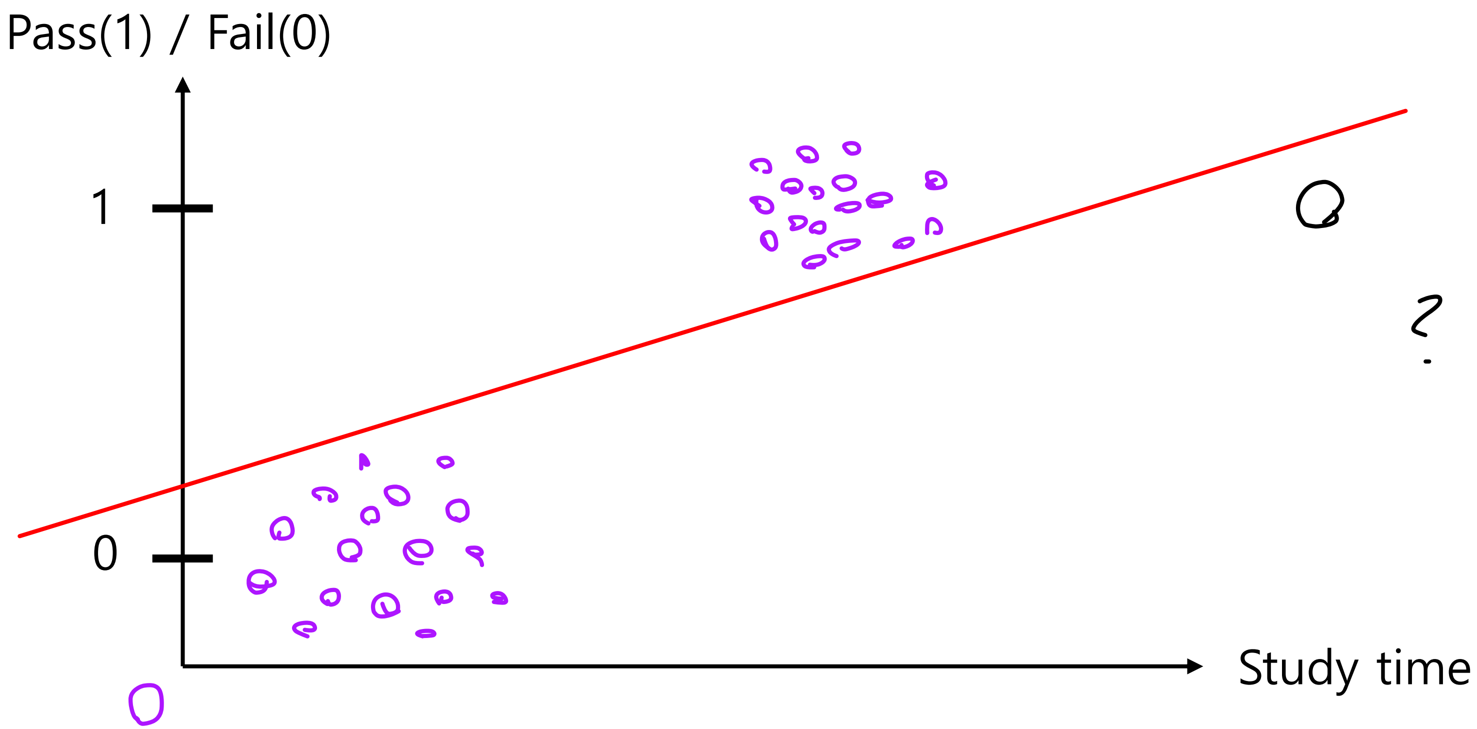
$$\boxed{\begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} x_1w_1 + x_2w_2 + x_3w_3 \end{pmatrix}}$$

H(X) = XW

Multivariable Linear Regression - Perceptron



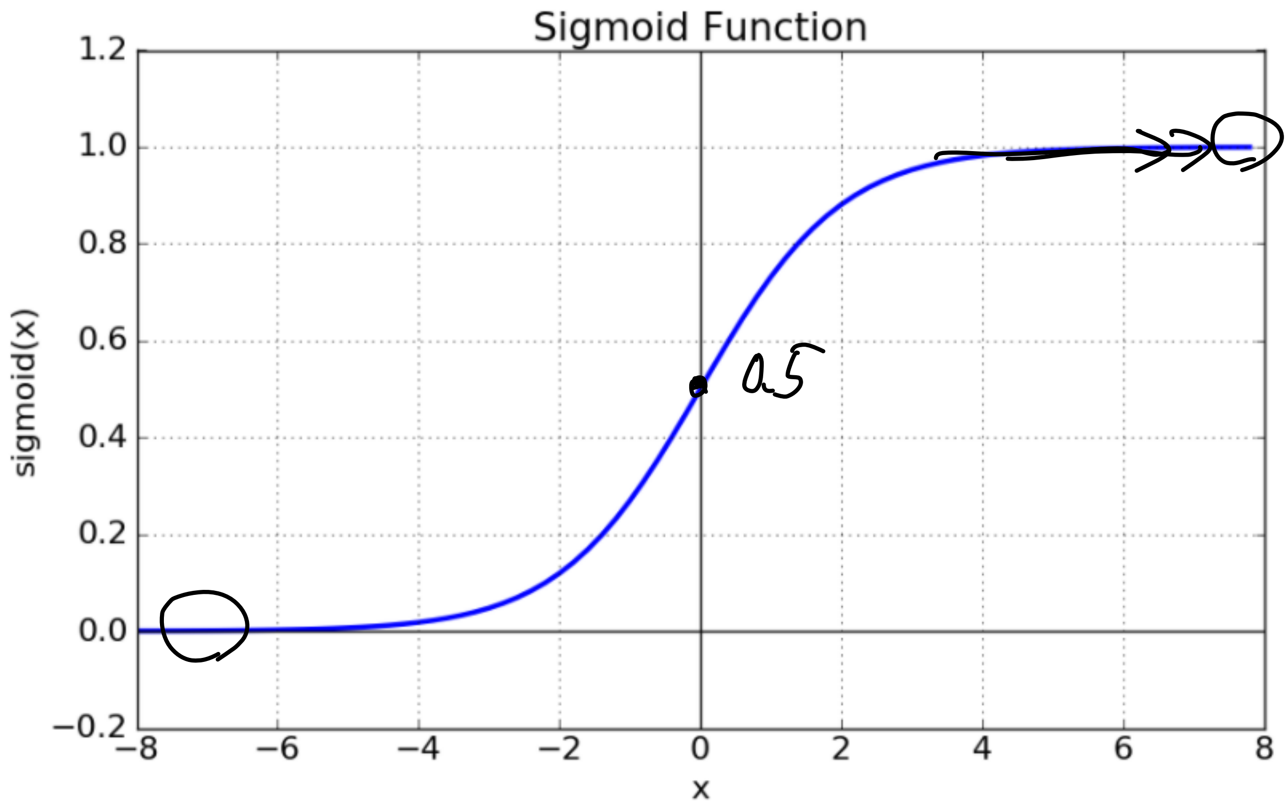
Binary Classification (이진분류) 둘중하나.



Binary Classification

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

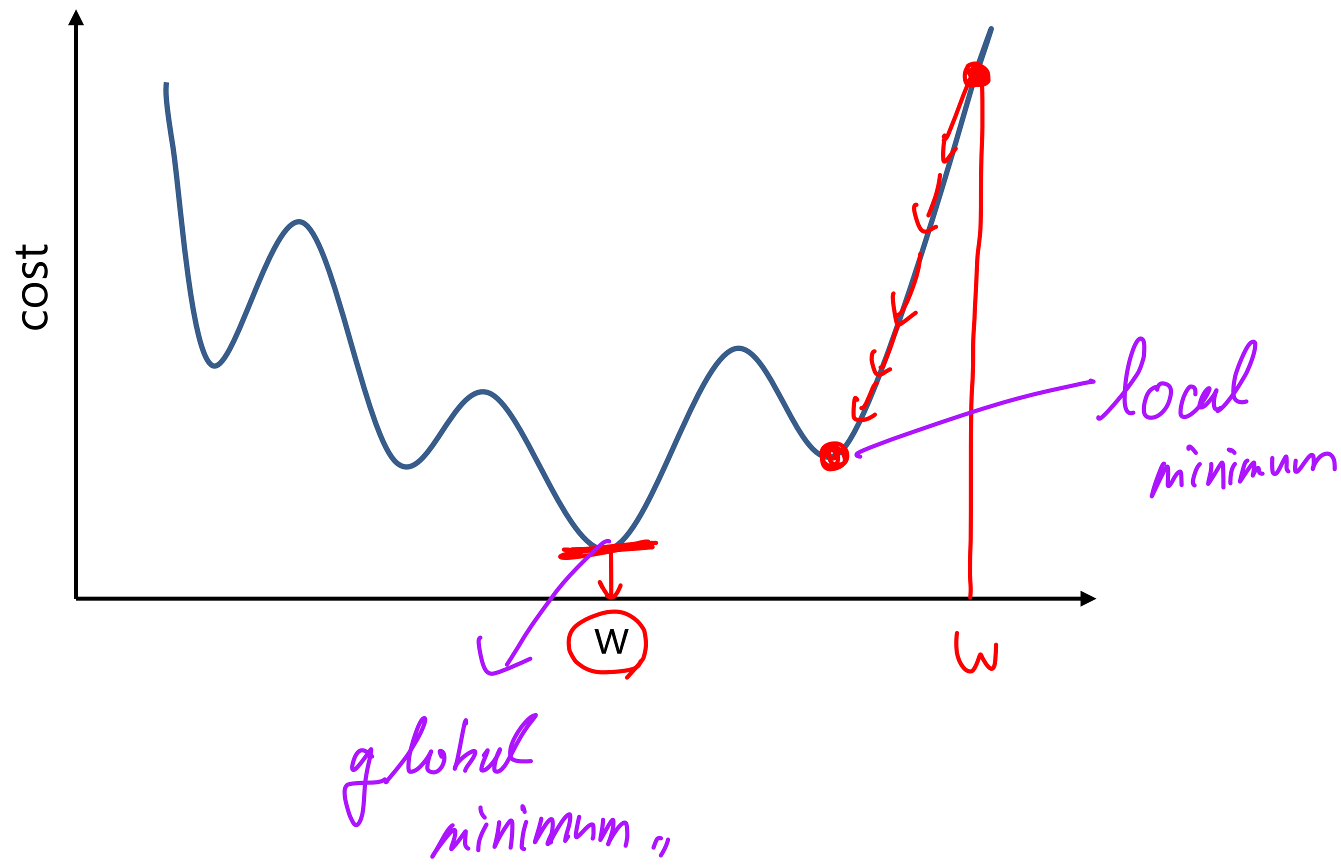
(0 , 1)



But! It has some problem.

Binary Classification

Problem (Cost Function)



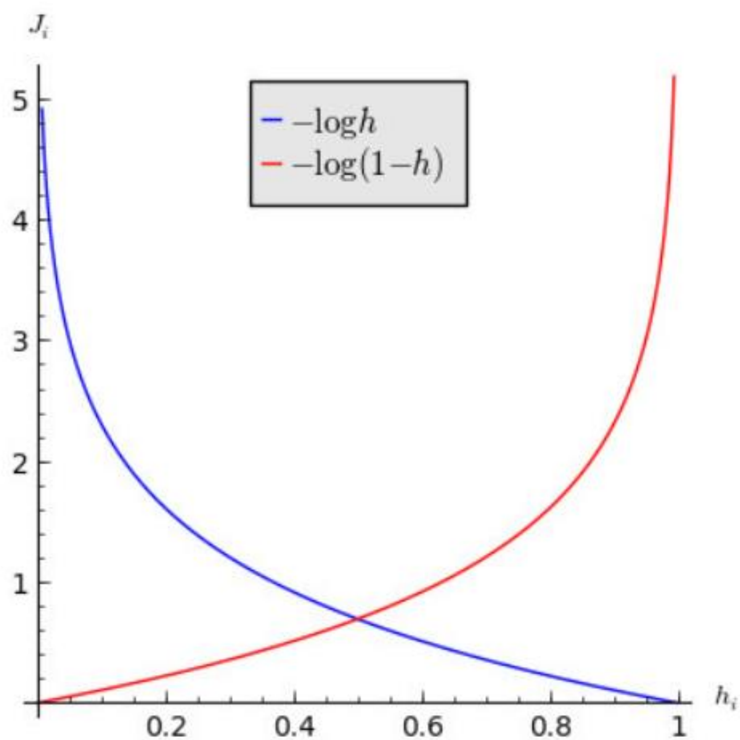
Binary Classification - Cost Function

- Cross Entropy

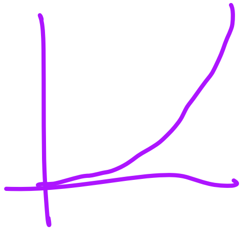
$$H(P,Q) = -\sum P(x)\log(Q(x))$$

$$cost(W) = \frac{1}{m}\sum c(H(x),y)$$

$$c(H(x),y) = \begin{cases} -log(H(x)) & : y=1 \\ -log(1-H(x)) & : y=0 \end{cases}$$



Multinormal Classification

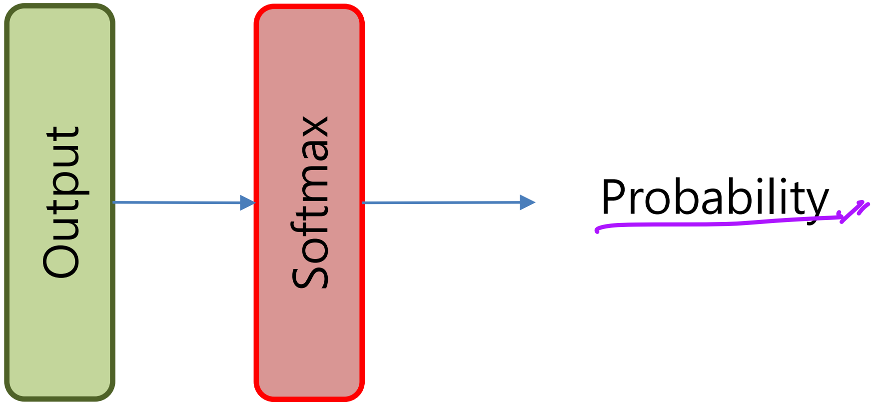


- Softmax Fuction

편차

$$P_j = \frac{e^{x_j}}{\sum_{k=1}^{k} e^{x_k}} = \frac{e^{x_j}}{e^{x_1} + e^{x_2} + e^{x_3} + \ \ldots + e^{x_k}} \quad for\ j = 1, \ldots, k$$

Sum is always 1



AND, OR, XOR (Prove)

AND

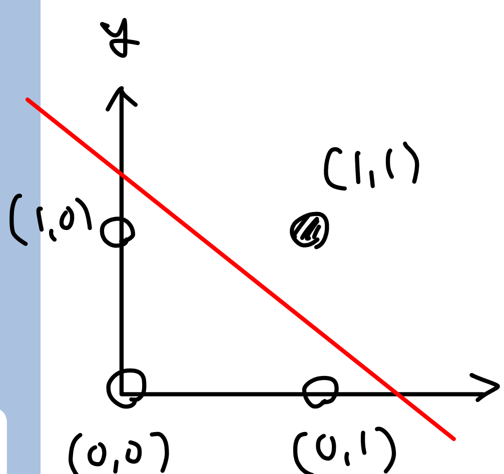
x1	x2	Y
0	0	0
0	1	0
1	0	0
1	1	1

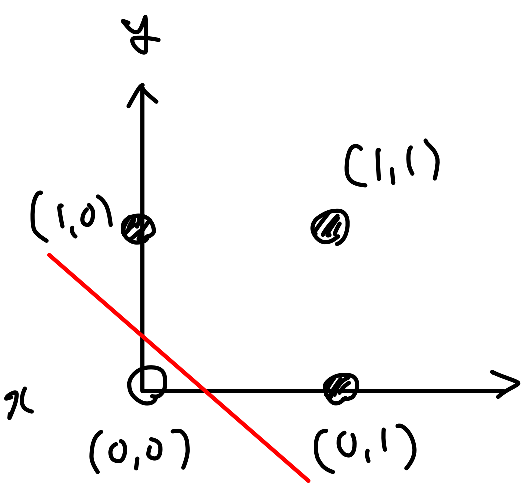
OR

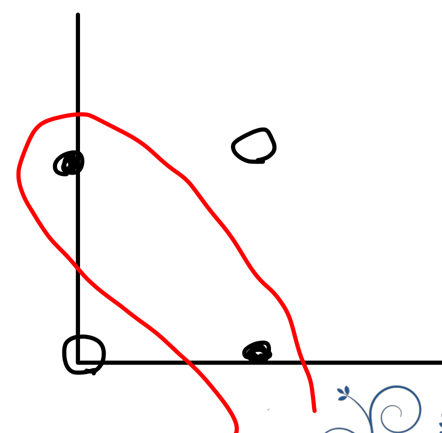
x1	x2	Y
0	0	0
0	1	1
1	0	1
1	1	1

XOR

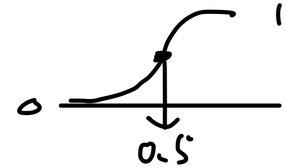
x1	x2	Y
0	0	0
0	1	1
1	0	1
1	1	0

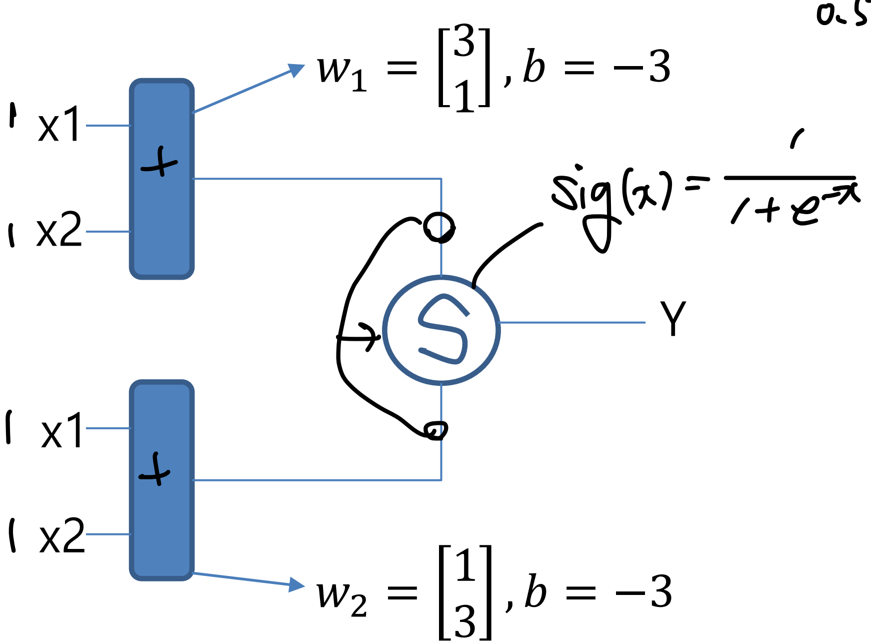






AND (Prove)





	x1	x2	Y
①	0	0	0
②	0	1	0
③	1	0	0
④	1	1	1

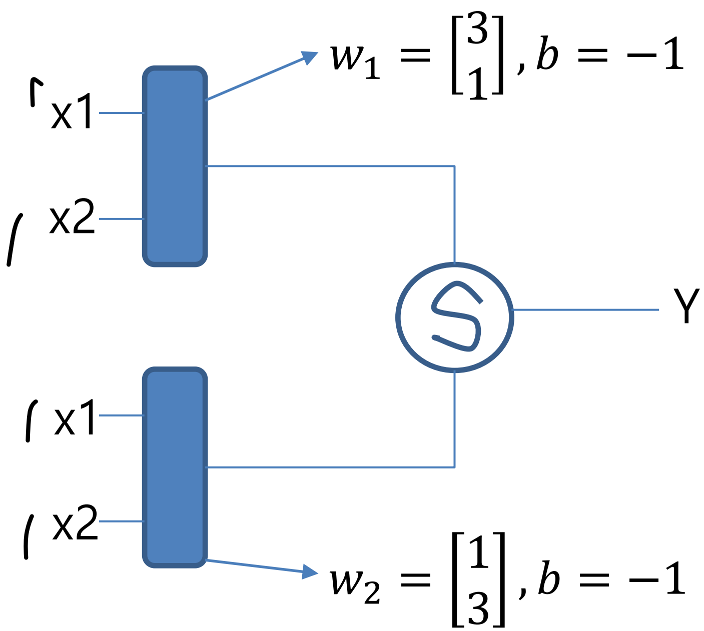
① $\left.\begin{array}{l} 3\times0 + 1\times0 - 3 = -3 \\ 1\times0 + 3\times0 - 3 = -3 \end{array}\right) \oplus\ -6 \rightarrow \text{sig} \rightarrow 0$

② $\left.\begin{array}{l} 0\times3 + 1\times1 - 3 = -2 \\ 0\times1 + 1\times3 - 3 = 0 \end{array}\right) \oplus\ -2 \rightarrow \text{sig} \rightarrow 0$

③ $\left.\begin{array}{l} 1\times3 + 0\times1 - 3 = 0 \\ 1\times1 + 0\times3 - 3 = -2 \end{array}\right) \oplus\ -2 \rightarrow \text{sig} \rightarrow 0$

④ $\left.\begin{array}{l} 3\times1 + 1\times1 - 3 = 1 \\ 1\times1 + 1\times3 - 3 = 1 \end{array}\right) \oplus\ 2 \rightarrow \text{sig} \rightarrow 1$

OR (Prove)



	x1	x2	Y
①	0	0	0
②	0	1	1
③	1	0	1
④	1	1	1

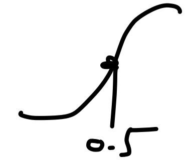
① $-1 + -1 = -2 \rightarrow 0$

② $0 + 2 = 2 \rightarrow 1$

③ $2 + 0 = 2 \rightarrow 1$

④ $3 + 3 = 6 \rightarrow 1$

XOR (Prove)



↳ 두 값이 (x_1, x_2) 같으면 0, 다르면 1

$(0,0)\ \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} + b' \rightarrow b' < 0.5$ ... ①

$(1,1)\ \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} + b' \rightarrow w_1 + w_2 + b' < 0.5$

↓

$w_1 + w_2 < 0.5 - b'$ ... ②

$(1,0)\ \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} + b'$

$(0,1)\ \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} + b'$

→ $w_1 + b' > 0.5$

$+\ \ w_2 + b' > 0.5$

$w_1 + w_2 + 2b' > 1$

$w_1 + w_2 > 1 - 2b'$ ... ③

② ③ 연결.

$1 - 2b' < w_1 + w_2 < 0.5 - b'$

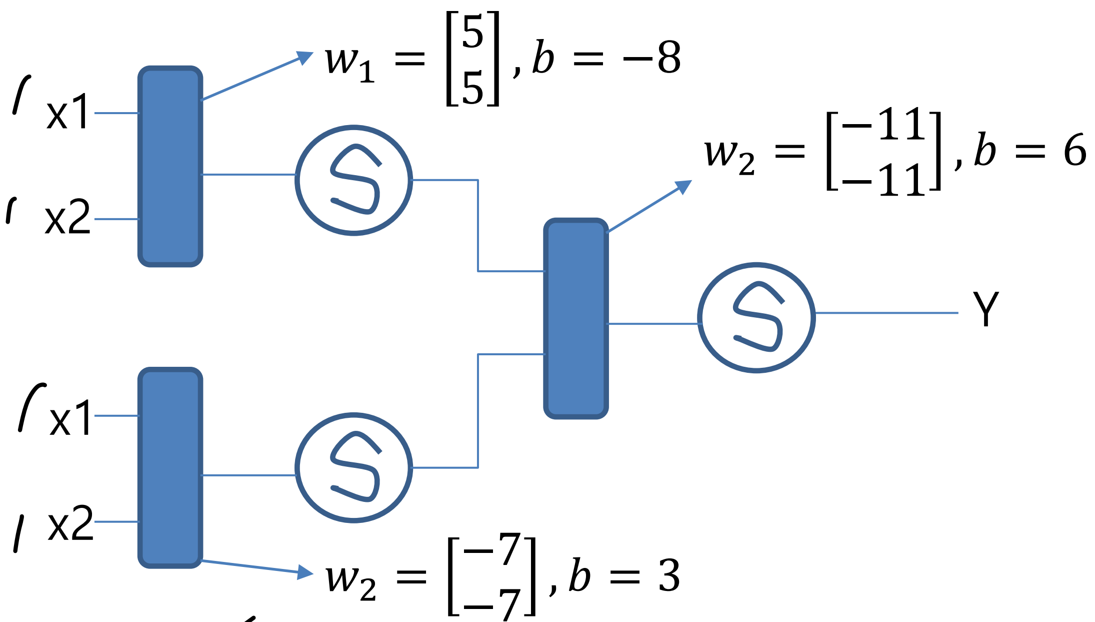
↓

$2(0.5 - b') < w_1 + w_2 < (0.5 - b')$ 모순 !!

linear f., ~~XOR~~ ↓ ~~classifi~~

XOR (Solution About Contradiction)

	x1	x2	Y
①	0	0	0
②	0	1	1
③	1	0	1
④	1	1	0



① $\cancel{-11 \times} 0 + -11 \times 1 + 6 = -5$

② $-3 \to 0$ $\quad 6 \to 1$

$-4 \to 0$

④ $10 - 8 = 2 \to 1$) $\quad -11 + 0 + 6 = -5 \to 0$

$-14 + 3 = -11 \to 0$

AND, OR, XOR (Prove)

AND, OR, XOR (Prove)

AND, OR, XOR (Prove)

AND, OR, XOR (Prove)

AND, OR, XOR (Prove)

Thank you...!!!

Next Time ?

- Live Coding in Jupyter Notebook (you prepare jupyter notebook, Colab, Anaconda or Vscode)
- Probability and Statistics
- Backpropagation

Assignment

- Diffrentiate The Sigmoid Function

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

Mail : wodnjs284@naver.com