Lecture 6-1

Softmax classification: Multinomial classification

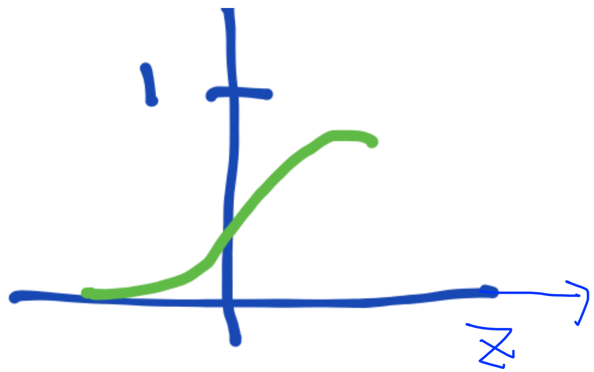
Sung Kim <hunkim+mr@gmail.com>

Linear 하기 때문에 0이나 1로 고르는 것에 적합하지 않았어

Logistic regression

$$H_L(x) = \underline{WX} - \left\{ \begin{matrix} 100 \\ 200 \\ -10 \end{matrix} \right\}$$

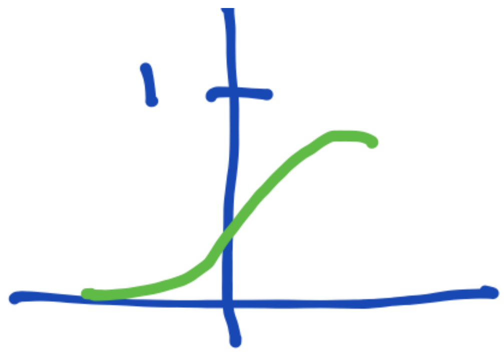
$$z = H_L(x), \quad \underbrace{g(z)} = \left\{ \begin{matrix} 0 \\ 1 \end{matrix} \right.$$



Logistic regression

$$H_L(x) = Wx$$

$$z = H_L(x), \quad g(z)$$



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

logistic 또는 sigmod라고 부름

$$H_R(x) = g(H_L(x))$$

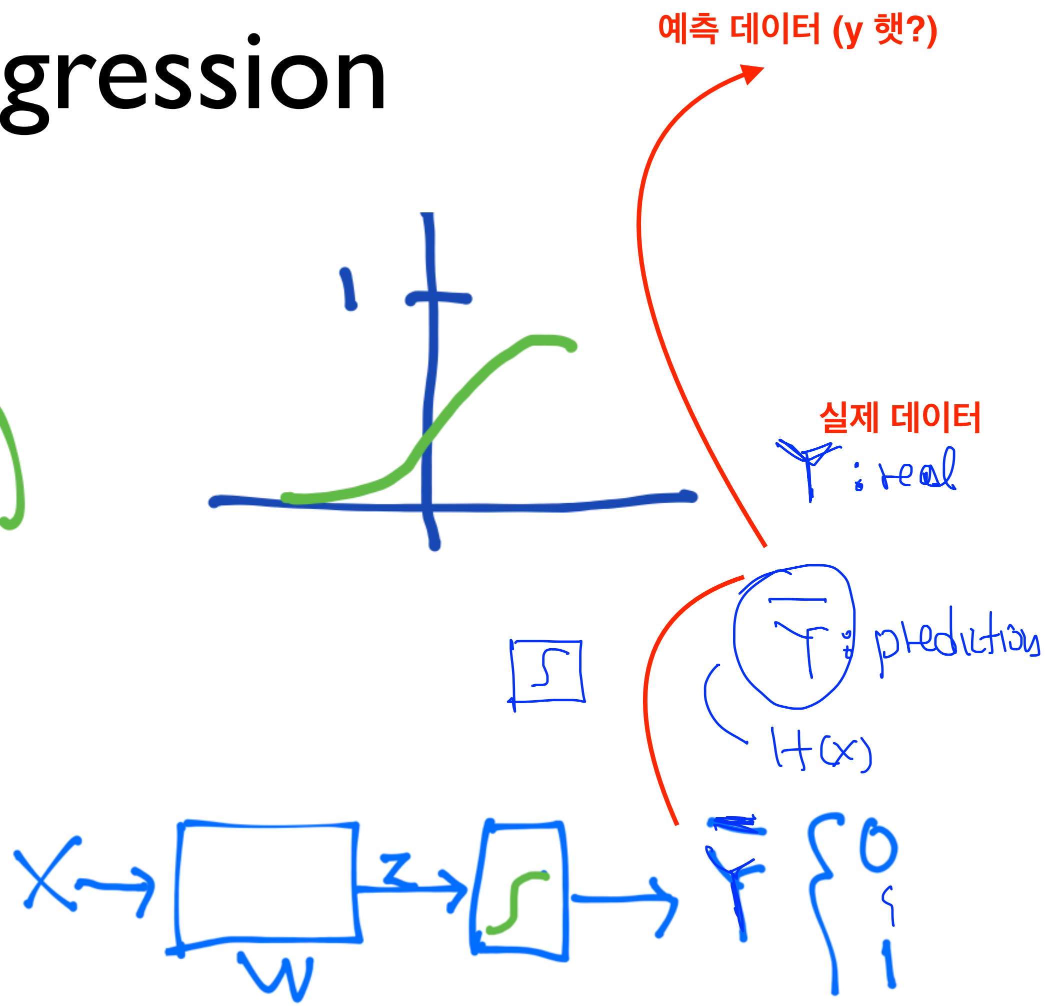
Logistic regression

$$H_L(X) = WX$$

$$z = H_L(X), \quad g(z)$$

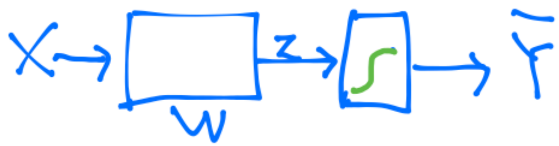
$$g(z) = \frac{1}{1+e^{-z}}$$

$$H_R(X) = g(H_L(X))$$

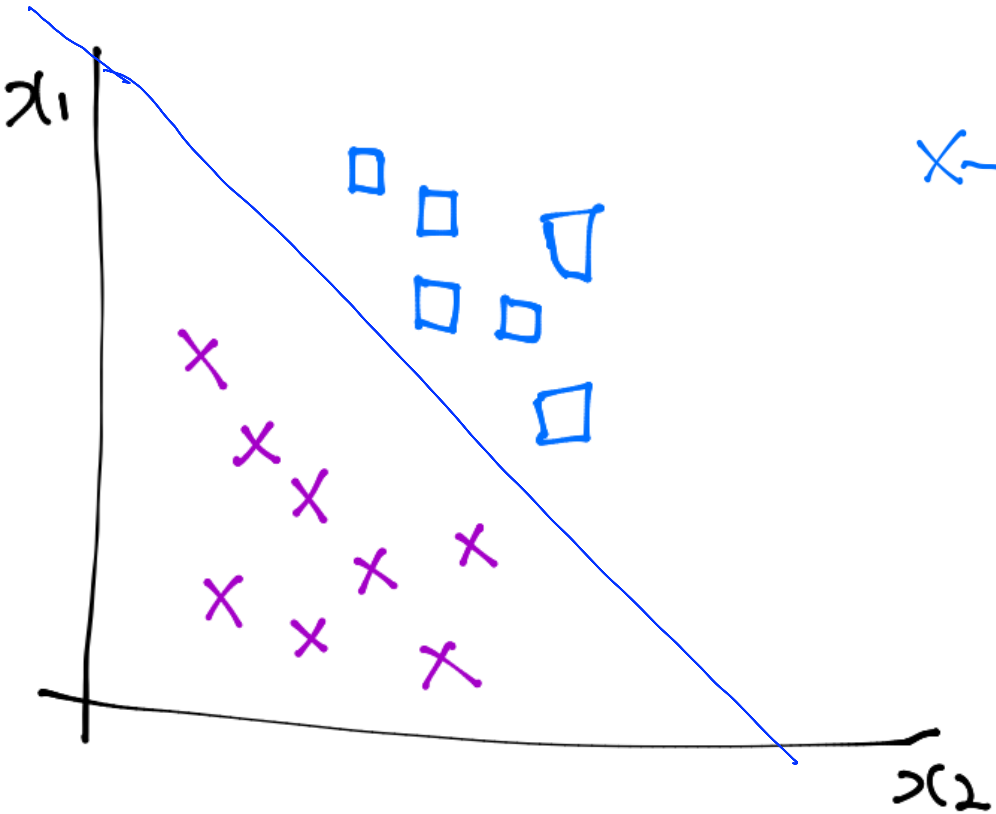


Logistic regression

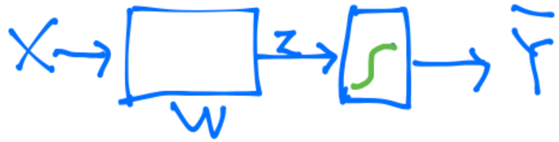
$$g(z) = \frac{1}{1+e^{-z}} \qquad H_R(x) = g(H_L(x))$$



Logistic regression

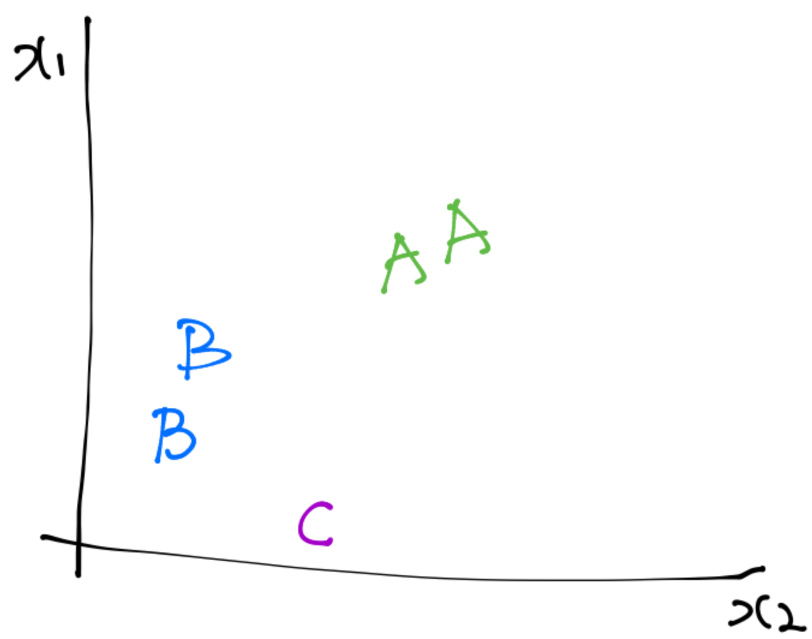


$$g(z) = \frac{1}{1+e^{-z}} \qquad H_R(x) = g(H_L(x))$$

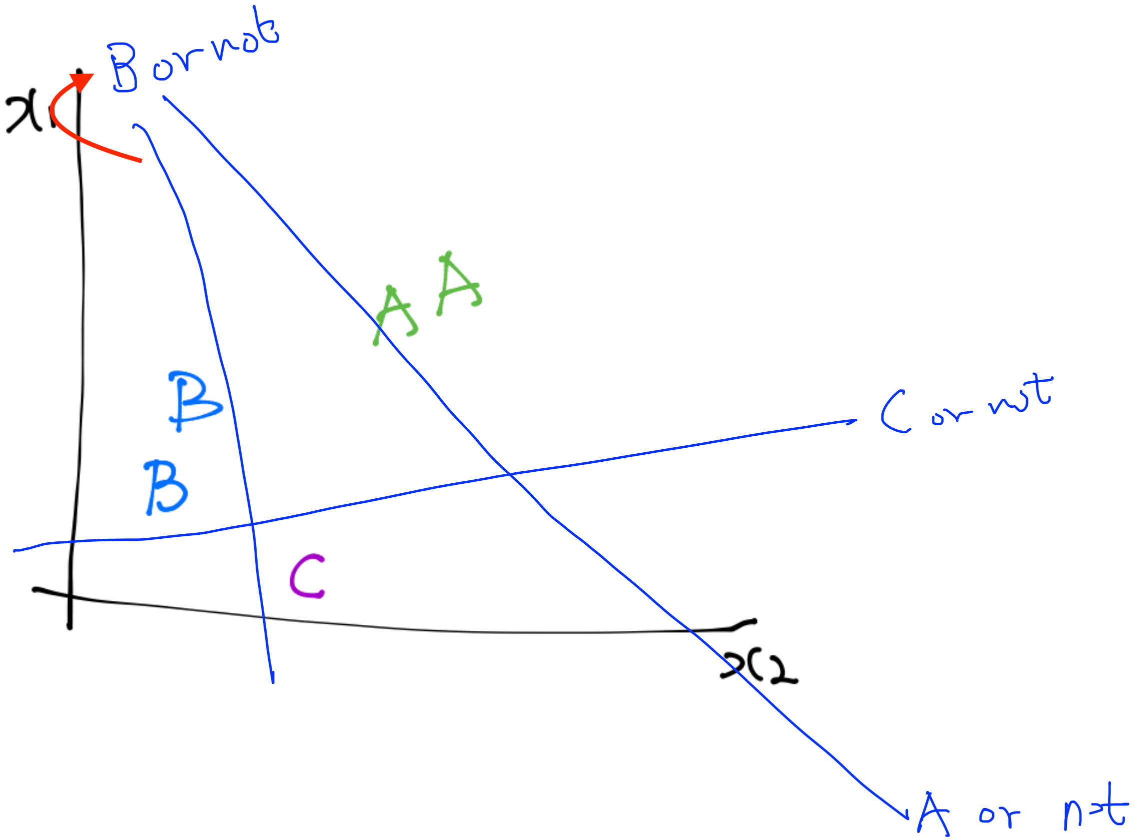


Multinomial classification

x1 (hours)	x2 (attendance)	y (grade)
10	5	A
9	5	A
3	2	B
2	4	B
11	1	C

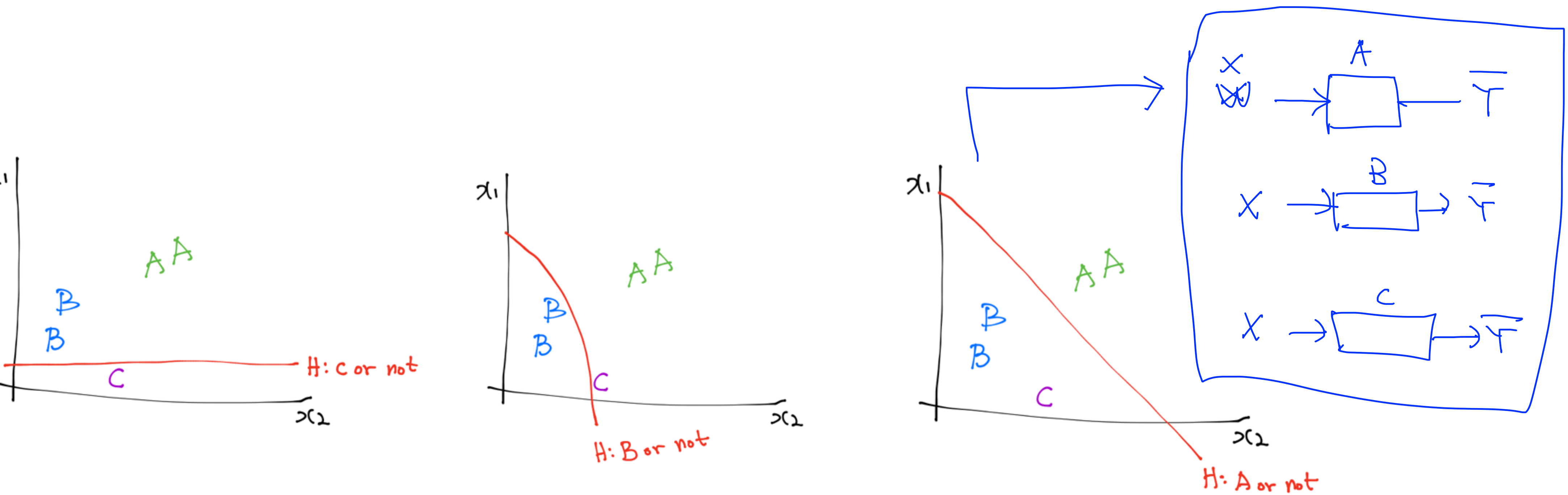


Multinomial classification

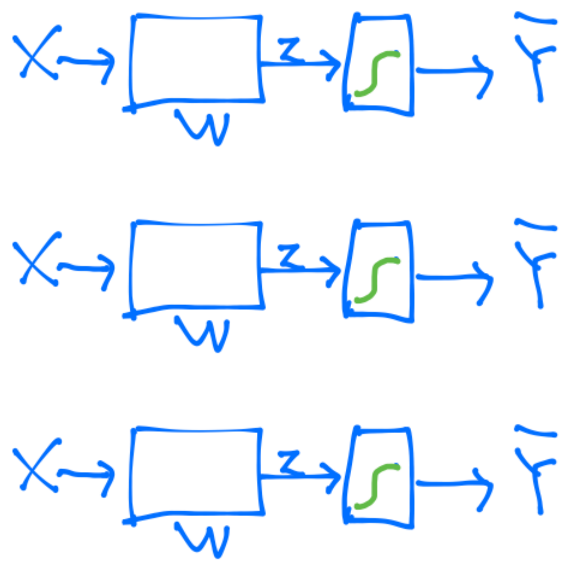


Multinomial classification

x가 주어졌을 때 각각 A, B, C 인지 아닌지 판단하는 독립된 classifier



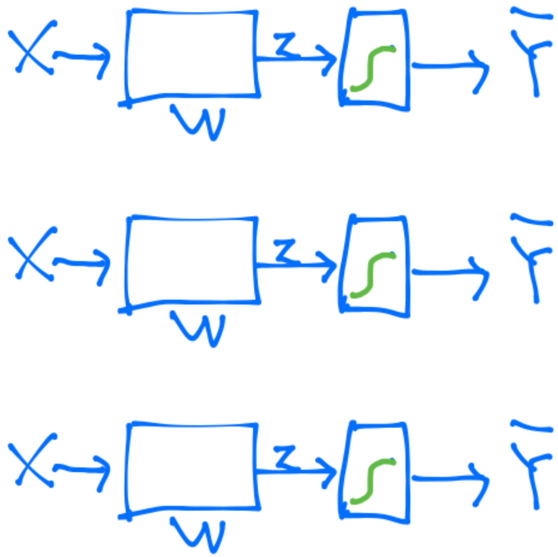
Multinomial classification



실제로 구현할 때는 행렬로

Multinomial classification

$$\underbrace{\begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix}}_{W} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{X} = \begin{bmatrix} \overbrace{w_1x_1 + w_2x_2 + w_3x_3}^{H(x)} \end{bmatrix}$$

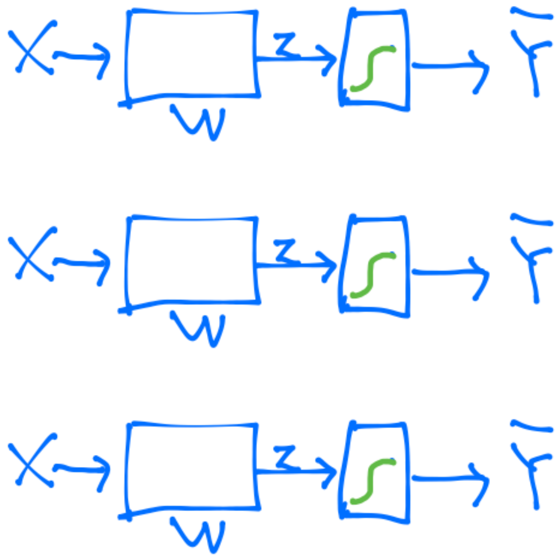


Multinomial classification

$$\begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} w_1x_1 + w_2x_2 + w_3x_3 \end{bmatrix}$$

$$\begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} w_1x_1 + w_2x_2 + w_3x_3 \end{bmatrix}$$

$$\begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} w_1x_1 + w_2x_2 + w_3x_3 \end{bmatrix}$$

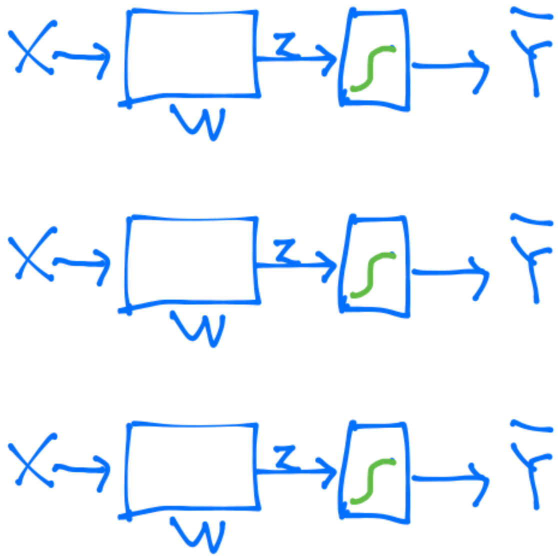


Multinomial classification

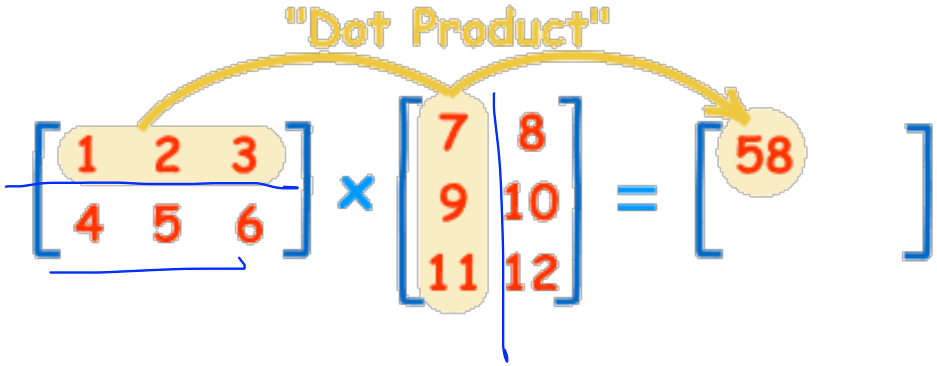
각각 구현하면 복잡해서 하나로 합친 후 구현

$$\begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} w_1x_1 + w_2x_2 + w_3x_3 \end{bmatrix}$$

$$\begin{bmatrix} w_{A1} & w_{A2} & w_{A3} \\ w_{B1} & w_{B2} & w_{B3} \\ w_{C1} & w_{C2} & w_{C3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} =$$



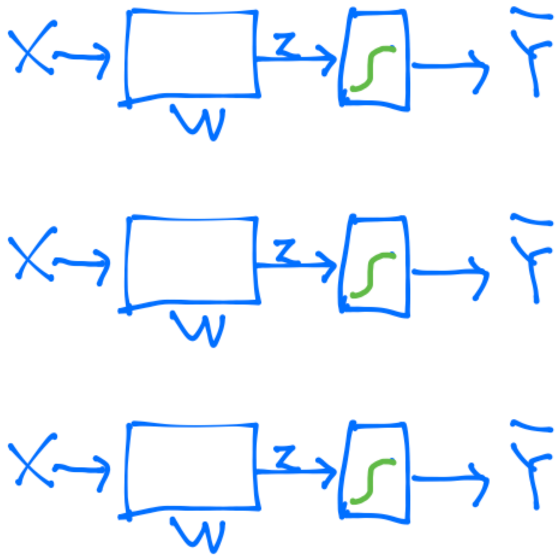
Matrix multiplication



https://www.mathsisfun.com/algebra/matrix-multiplying.html

Multinomial classification

$$\begin{bmatrix} w_{A1} & w_{A2} & w_{A3} \\ w_{B1} & w_{B2} & w_{B3} \\ w_{C1} & w_{C2} & w_{C3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} w_{A1}x_1 + w_{A2}x_2 + w_{A3}x_3 \\ w_{B1}x_1 + w_{B2}x_2 + w_{B3}x_3 \\ w_{C1}x_1 + w_{C2}x_2 + w_{C3}x_3 \end{bmatrix}$$



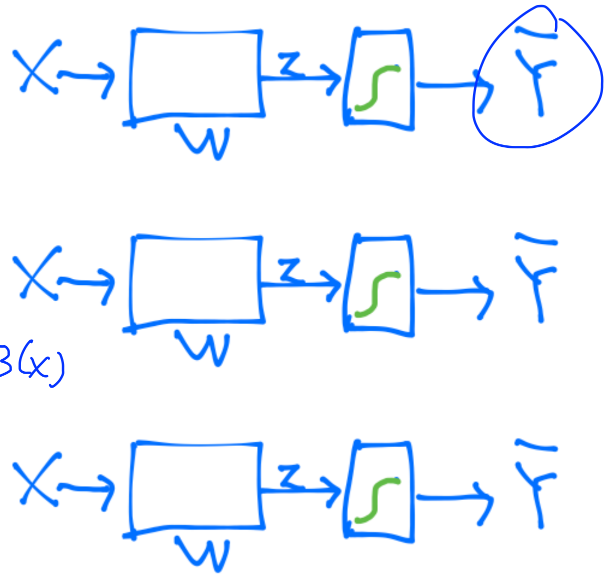
Multinomial classification

$$\begin{bmatrix} w_{A1} & w_{A2} & w_{A3} \\ w_{B1} & w_{B2} & w_{B3} \\ w_{C1} & w_{C2} & w_{C3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} w_{A1}x_1 + w_{A2}x_2 + w_{A3}x_3 \\ w_{B1}x_1 + w_{B2}x_2 + w_{B3}x_3 \\ w_{C1}x_1 + w_{C2}x_2 + w_{C3}x_3 \end{bmatrix} = \begin{bmatrix} \bar{y}_A \\ \bar{y}_B \\ \bar{y}_C \end{bmatrix}$$

$H_A(x)$

$H_B(x)$

$H_C(x)$



Where is sigmoid?

각각 sigmod를 적용하면 돼

$$\begin{bmatrix} w_{A1} & w_{A2} & w_{A3} \\ w_{B1} & w_{B2} & w_{B3} \\ w_{C1} & w_{C2} & w_{C3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} w_{A1}x_1 + w_{A2}x_2 + w_{A3}x_3 \\ w_{B1}x_1 + w_{B2}x_2 + w_{B3}x_3 \\ w_{C1}x_1 + w_{C2}x_2 + w_{C3}x_3 \end{bmatrix} = \begin{bmatrix} \bar{y}_A \\ \bar{y}_B \\ \bar{y}_C \end{bmatrix}$$