

11주차(3/3)

# 로지스틱 회귀 2

파이썬으로 배우는 기계학습

한동대학교  
김영섭 교수

# 로지스틱 회귀

---

- 학습 목표
  - 로지스틱 함수를 이해한다.
  - 로지스틱 회귀 비용함수(교차 엔트로피)를 이해한다.
  - 로지스틱 회귀 신경망의 역전파를 계산한다.
- 학습 내용
  - 로지스틱 함수 이해
  - 로지스틱 회귀 비용함수(교차 엔트로피) 미분하기
  - 로지스틱 회귀 신경망의 역전파 계산

# 1. 로지스틱 함수: 모델 이해하기

---

- 로지스틱 함수
  - 시그모이드 함수

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

# 1. 로지스틱 함수: 모델 이해하기

---

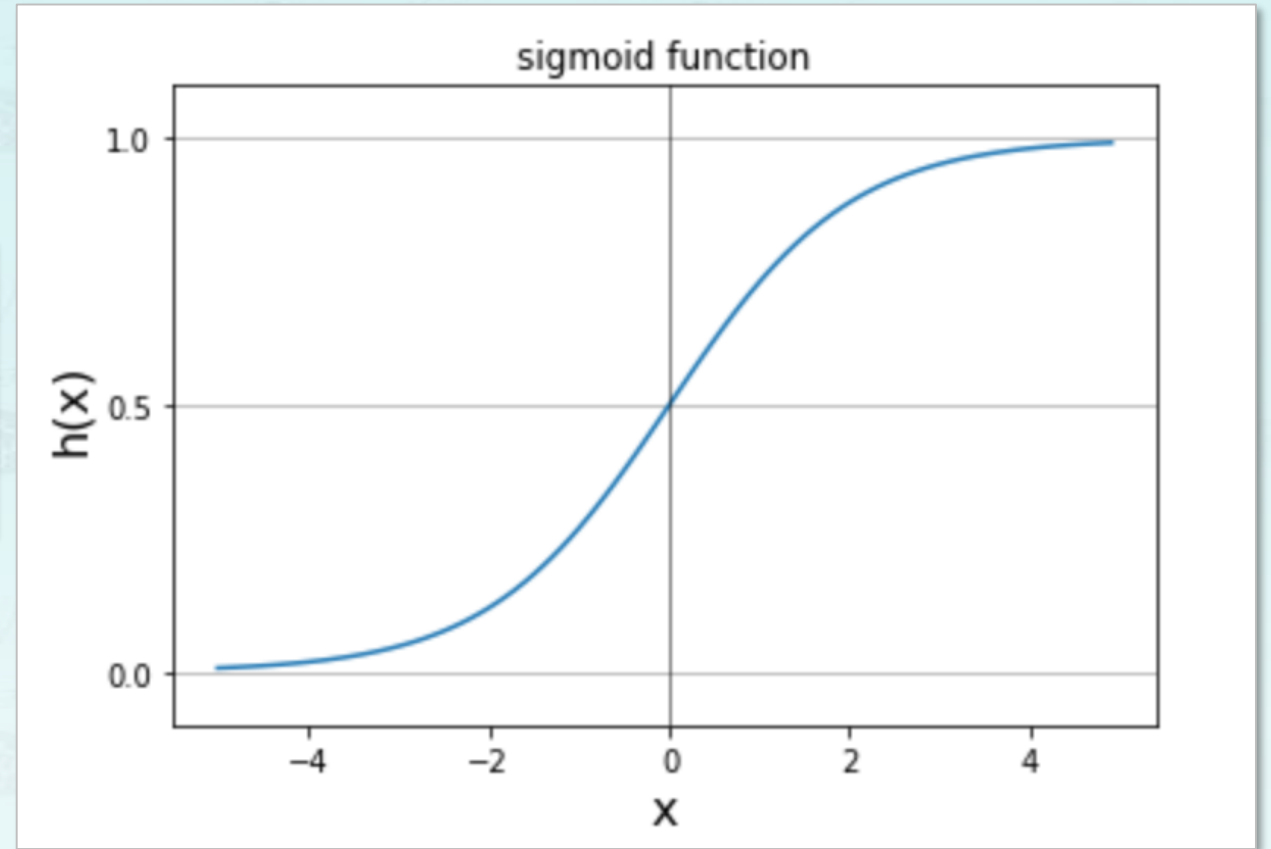
- 로지스틱 함수
  - 시그모이드 함수

$$\sigma(x) = \frac{1}{1 + e^{-x}} \begin{cases} \sigma(x) \rightarrow 1 & \text{if } x \rightarrow +\infty \\ \sigma(x) \rightarrow \frac{1}{2} & \text{if } x \rightarrow 0 \\ \sigma(x) \rightarrow 0 & \text{if } x \rightarrow -\infty \end{cases}$$

# 1. 로지스틱 함수: 모델 이해하기

- 로지스틱 함수
  - 시그모이드 함수

$$\sigma(x) = \frac{1}{1 + e^{-x}} \begin{cases} \sigma(x) \rightarrow 1 & \text{if } x \rightarrow +\infty \\ \sigma(x) \rightarrow \frac{1}{2} & \text{if } x \rightarrow 0 \\ \sigma(x) \rightarrow 0 & \text{if } x \rightarrow -\infty \end{cases}$$



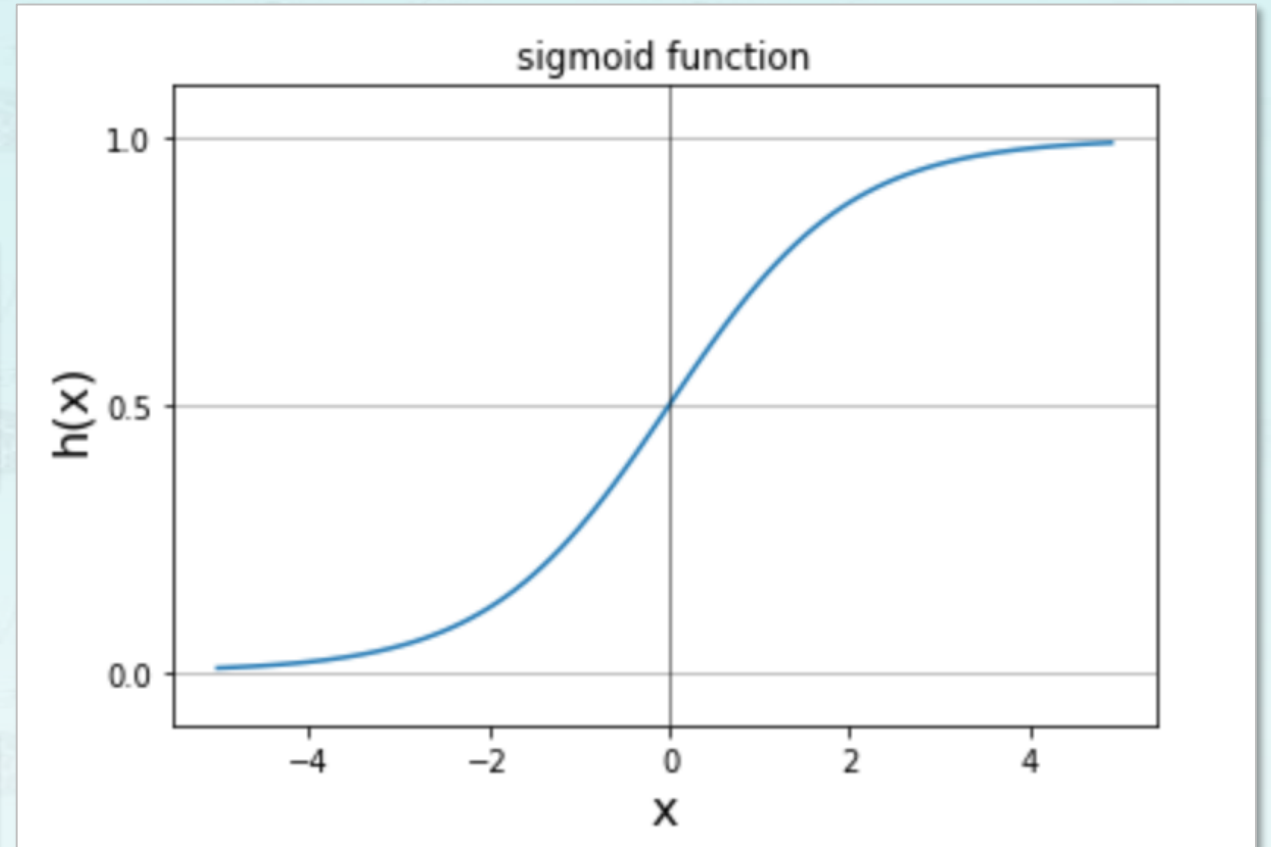
# 1. 로지스틱 함수: 모델 이해하기

- 로지스틱 함수
  - 시그모이드 함수

$$\text{sigmoid}(x) = \sigma(x) = \frac{1}{1 + e^{-x}}$$

- $h(z) = WX$

$$h(z) = \frac{1}{1 + e^{-z}} = \frac{1}{1 + e^{-W \cdot X}}$$



# 1. 로지스틱 함수: 모델 이해하기

- 로지스틱 함수
  - $y \in \{1,0\}$

*Classifier:*

$$y = h(z) = \frac{1}{1 + e^{-WX}} \begin{cases} y \rightarrow 1 \text{ if } WX \rightarrow \infty \\ y = \frac{1}{2} \text{ if } WX \rightarrow 0 \\ y \rightarrow 0 \text{ if } WX \rightarrow -\infty \end{cases}$$



- 시그모이드 함수

$$\sigma(x) = \frac{1}{1 + e^{-x}} \begin{cases} \sigma(x) \rightarrow 1 & \text{if } x \rightarrow +\infty \\ \sigma(x) \rightarrow \frac{1}{2} & \text{if } x \rightarrow 0 \\ \sigma(x) \rightarrow 0 & \text{if } x \rightarrow -\infty \end{cases}$$

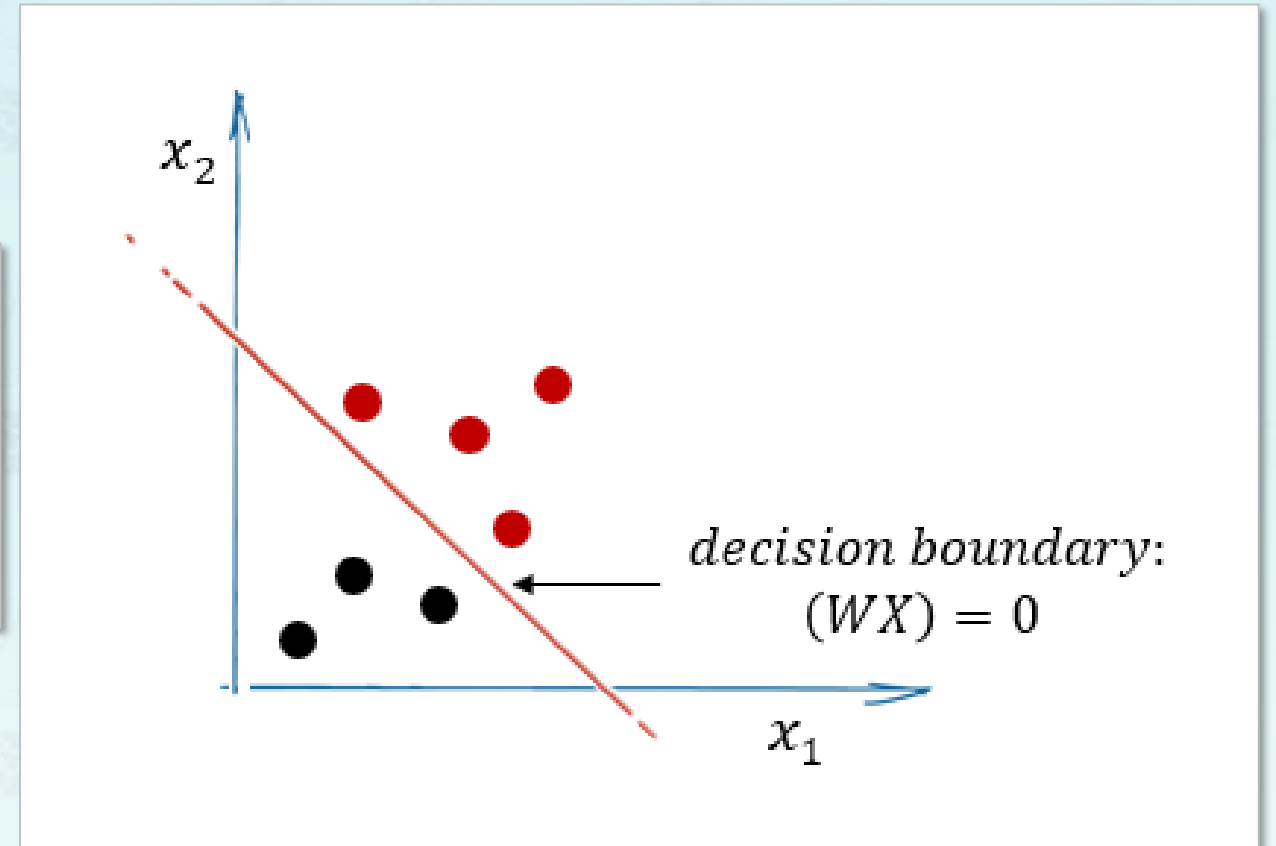
# 1. 로지스틱 함수: 모델 이해하기

- 로지스틱 함수
  - $y \in \{1,0\}$

*Classifier:*

$$y = h(z) = \frac{1}{1 + e^{-WX}} \begin{cases} y \rightarrow 1 \text{ if } WX \rightarrow \infty \\ y = \frac{1}{2} \text{ if } WX \rightarrow 0 \\ y \rightarrow 0 \text{ if } WX \rightarrow -\infty \end{cases}$$

- 결정경계선
  - hyperplane





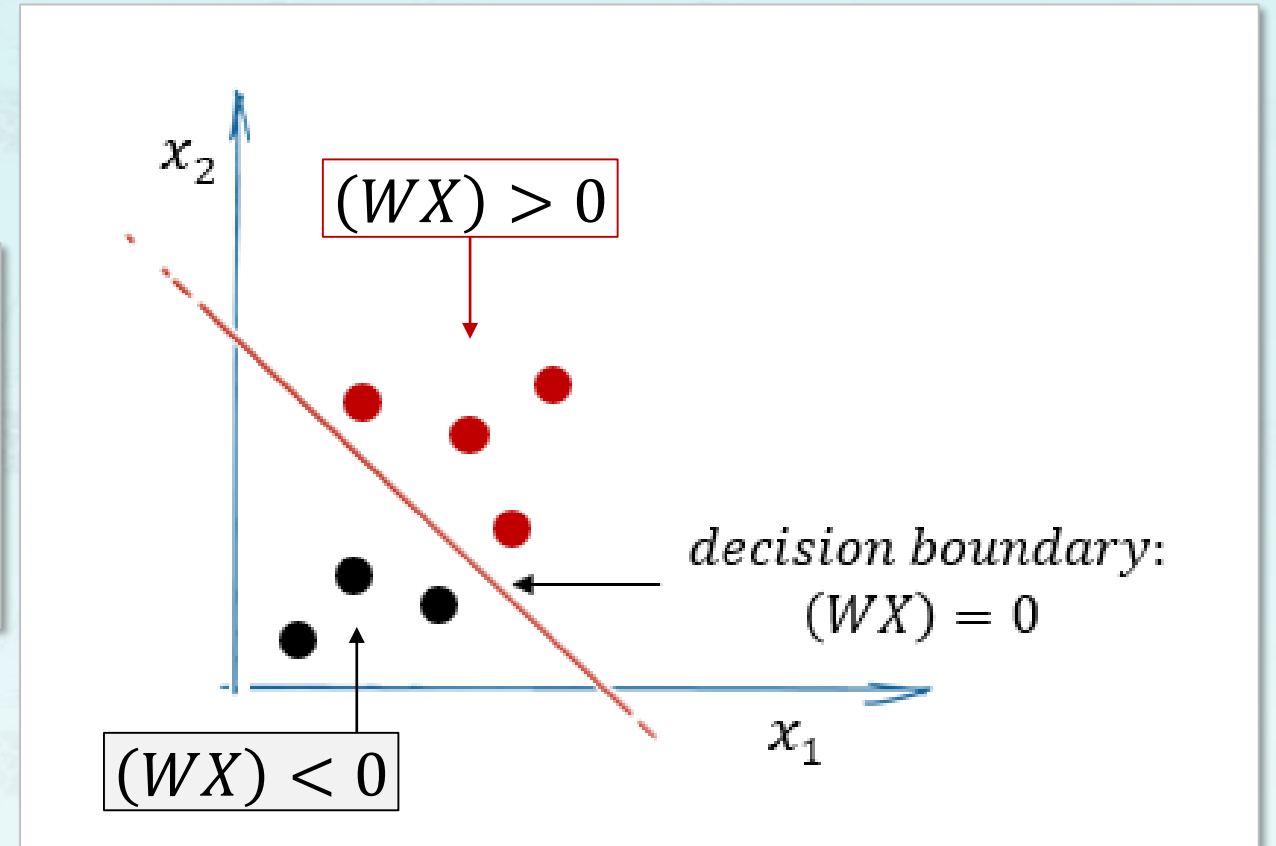
# 1. 로지스틱 함수: 모델 이해하기

- 로지스틱 함수
  - $y \in \{1,0\}$

*Classifier:*

$$y = h(z) = \frac{1}{1 + e^{-WX}} \begin{cases} y \rightarrow 1 \text{ if } WX \rightarrow \infty \\ y = \frac{1}{2} \text{ if } WX \rightarrow 0 \\ y \rightarrow 0 \text{ if } WX \rightarrow -\infty \end{cases}$$

- 결정경계선
  - **hyperplane**
  - 예시:  $W = [-3, 1, 1]$



# 1. 로지스틱 함수: 모델 이해하기

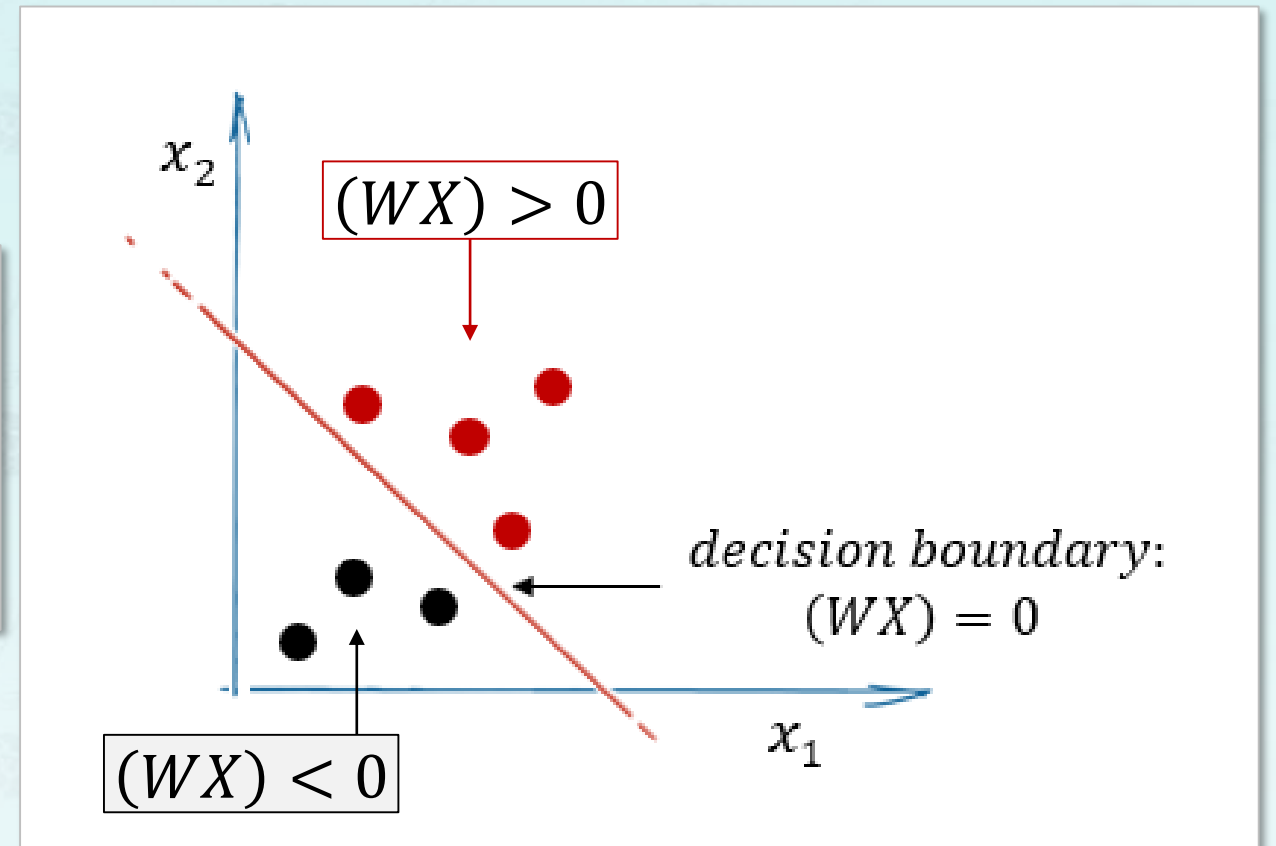
- 로지스틱 함수
  - $y \in \{1,0\}$

*Classifier:*

$$y = h(z) = \frac{1}{1 + e^{-WX}} \begin{cases} y \rightarrow 1 \text{ if } WX \rightarrow \infty \\ y = \frac{1}{2} \text{ if } WX \rightarrow 0 \\ y \rightarrow 0 \text{ if } WX \rightarrow -\infty \end{cases}$$

- 결정경계선
  - **hyperplane**
  - 예시:  $W = [-3, 1, 1]$

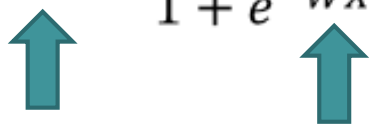
$$\begin{aligned} h(z) &= h(w_0 + w_1x_1 + w_2x_2) \\ &= -3 + x_1 + x_2 \end{aligned}$$



# 1. 로지스틱 함수: 모델 이해하기

- 로지스틱 함수
  - $y \in \{1,0\}$

*Classifier:*

$$y = h(z) = \frac{1}{1 + e^{-WX}} \begin{cases} y \rightarrow 1 \text{ if } WX \rightarrow \infty \\ y = \frac{1}{2} \text{ if } WX \rightarrow 0 \\ y \rightarrow 0 \text{ if } WX \rightarrow -\infty \end{cases}$$
The diagram shows the sigmoid function formula. Two teal arrows point upwards to the variables in the formula: one points to the 'z' in 'h(z)' and the other points to the 'W' in 'WX'.

- 결정경계선
  - **hyperplane**
  - 예시:  $W = [-3, 1, 1]$

# 1. 로지스틱 함수: 모델 이해하기

- 로지스틱 함수
  - $y \in \{1,0\}$

*Classifier:*

$$y = h(z) = \frac{1}{1 + e^{-WX}} \begin{cases} y \rightarrow 1 \text{ if } WX \rightarrow \infty \\ y = \frac{1}{2} \text{ if } WX \rightarrow 0 \\ y \rightarrow 0 \text{ if } WX \rightarrow -\infty \end{cases}$$

- 결정경계선
  - **hyperplane**
  - 예시:  $W = [-3,1,1]$



$$z = w_0 + w_1x_1^2 + w_2x_2^2$$

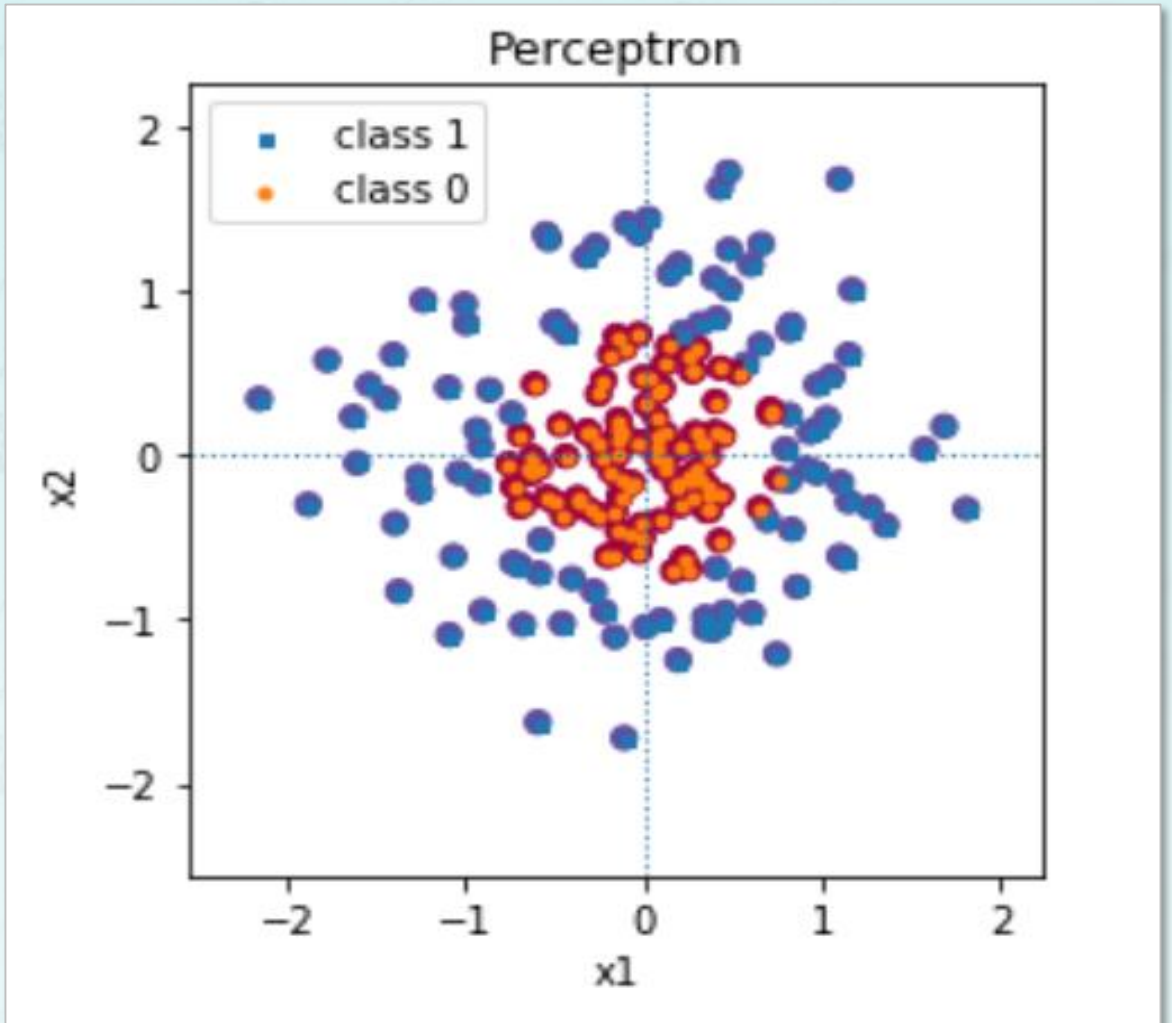
# 1. 로지스틱 함수: 모델 이해하기

- 로지스틱 함수
  - $y \in \{1,0\}$

*Classifier:*

$$y = h(z) = \frac{1}{1 + e^{-WX}} \begin{cases} y \rightarrow 1 \text{ if } WX \rightarrow \infty \\ y = \frac{1}{2} \text{ if } WX \rightarrow 0 \\ y \rightarrow 0 \text{ if } WX \rightarrow -\infty \end{cases}$$

- 결정경계선
  - **hyperplane**
  - 예시:  $W = [-3, 1, 1]$



$$z = w_0 + w_1 x_1^2 + w_2 x_2^2$$

## 2. 로지스틱 회귀: 비용함수

---

- 오차함수  $E$

## 2. 로지스틱 회귀: 비용함수

---

- 오차함수  $E$  - 오차 제곱의 합, 평균
  - **SSE: Sum of Squared Error**
  - **MSE: Mean of Squared Error**

$$SSE = \sum_{i=1}^m \frac{1}{2} (y^{(i)} - \hat{y}^{(i)})^2$$

$$MSE = \frac{1}{m} \sum_{i=1}^m \frac{1}{2} (y^{(i)} - \hat{y}^{(i)})^2$$

## 2. 로지스틱 회귀: 비용함수

- 오차함수  $E$  - 오차 제곱의 합, 평균
  - **SSE: Sum of Squared Error**
  - **MSE: Mean of Squared Error**

$$SSE = \sum_{i=1}^m \frac{1}{2} (y^{(i)} - \hat{y}^{(i)})^2$$

$$MSE = \frac{1}{m} \sum_{i=1}^m \frac{1}{2} (y^{(i)} - \hat{y}^{(i)})^2$$

- 선형 회귀
  - $y = WX$
- 로지스틱 회귀
  - $y = \frac{1}{1 + e^{-WX}}$



## 2. 로지스틱 회귀: 비용함수

---

- 비용함수  $J$

$$J = \begin{cases} -\log(h(z)) & \text{if } y = 1. \\ -\log(1 - h(z)) & \text{if } y = 0. \end{cases}$$

- 로지스틱 함수

$$h(z) = \frac{1}{1 + e^{-W \cdot X}}$$

## 2. 로지스틱 회귀: 비용함수

- 비용함수  $J$

$$J = \begin{cases} -\log(h(z)) & \text{if } y = 1. \\ -\log(1 - h(z)) & \text{if } y = 0. \end{cases}$$

- $y$ : 클래스 레이블
- $h(z)$ : 신경망의 출력

- 로지스틱 함수

$$h(z) = \frac{1}{1 + e^{-W \cdot X}}$$

## 2. 로지스틱 회귀: 비용함수

- 비용함수  $J$

$$J = \begin{cases} -\log(h(z)) & \text{if } y = 1. \\ -\log(1 - h(z)) & \text{if } y = 0. \end{cases}$$

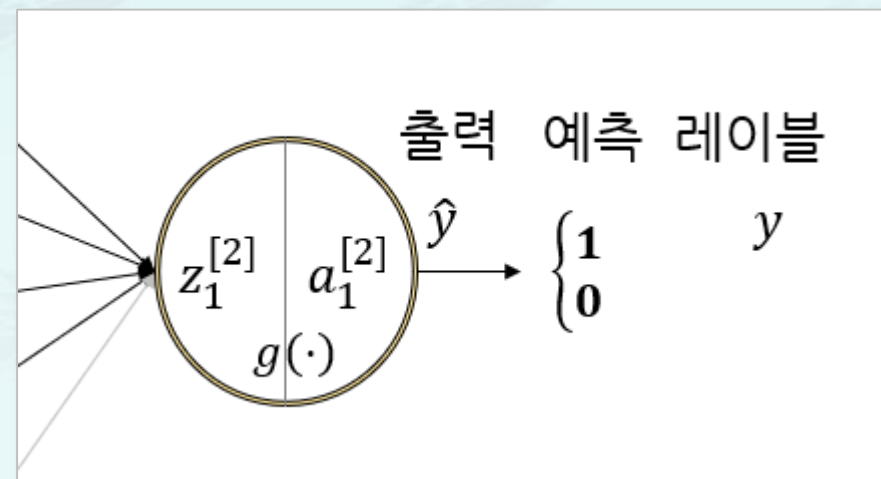
$$e^{(i)} = \begin{cases} -\log(a^{[2](i)}) & \text{if } y = 1. \\ -\log(1 - a^{[2](i)}) & \text{if } y = 0. \end{cases}$$

- $a^{[2](i)}$

- 출력층 (i)번째 노드의 출력

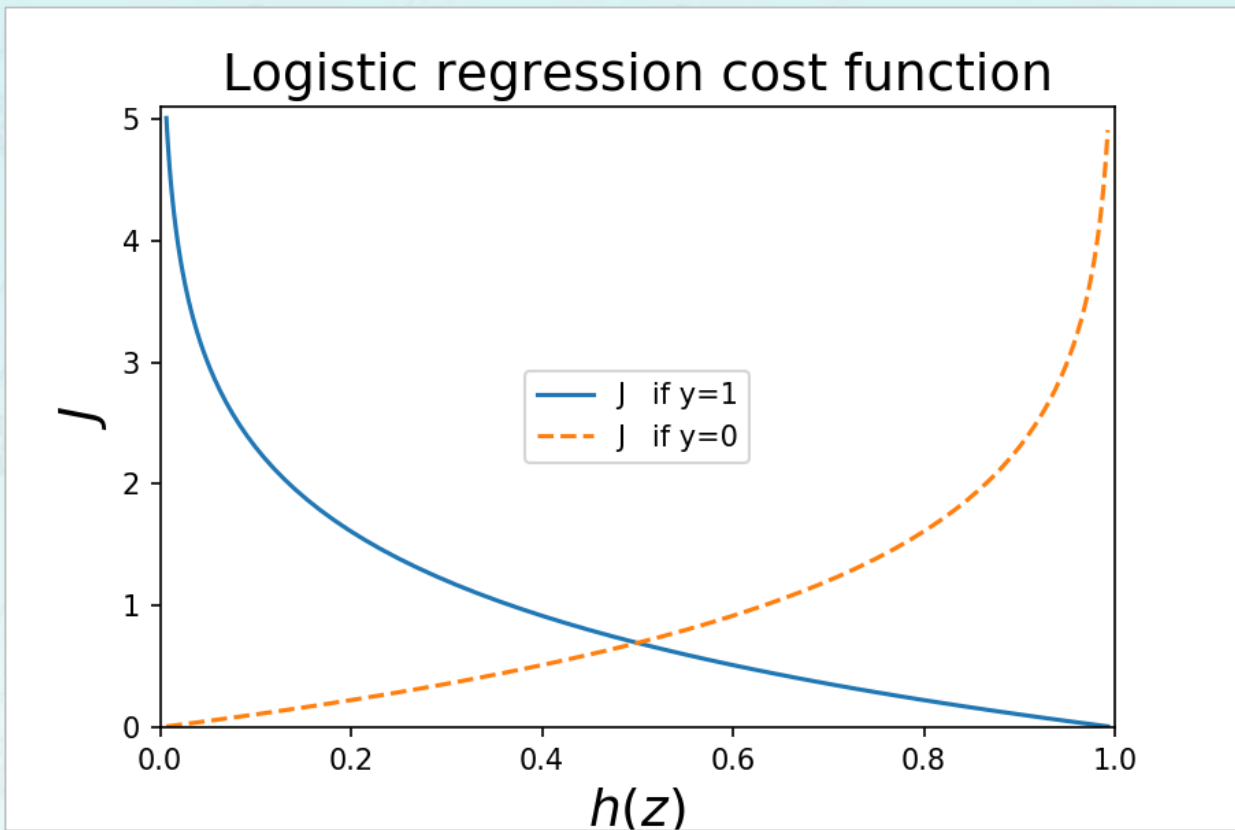
- 로지스틱 함수

$$h(z) = \frac{1}{1 + e^{-W \cdot X}}$$



## 2. 로지스틱 회귀: 비용함수

- 비용함수  $J$



$$e^{(i)} = \begin{cases} -\log(a^{[2](i)}) & \text{if } y = 1. \\ -\log(1 - a^{[2](i)}) & \text{if } y = 0. \end{cases}$$

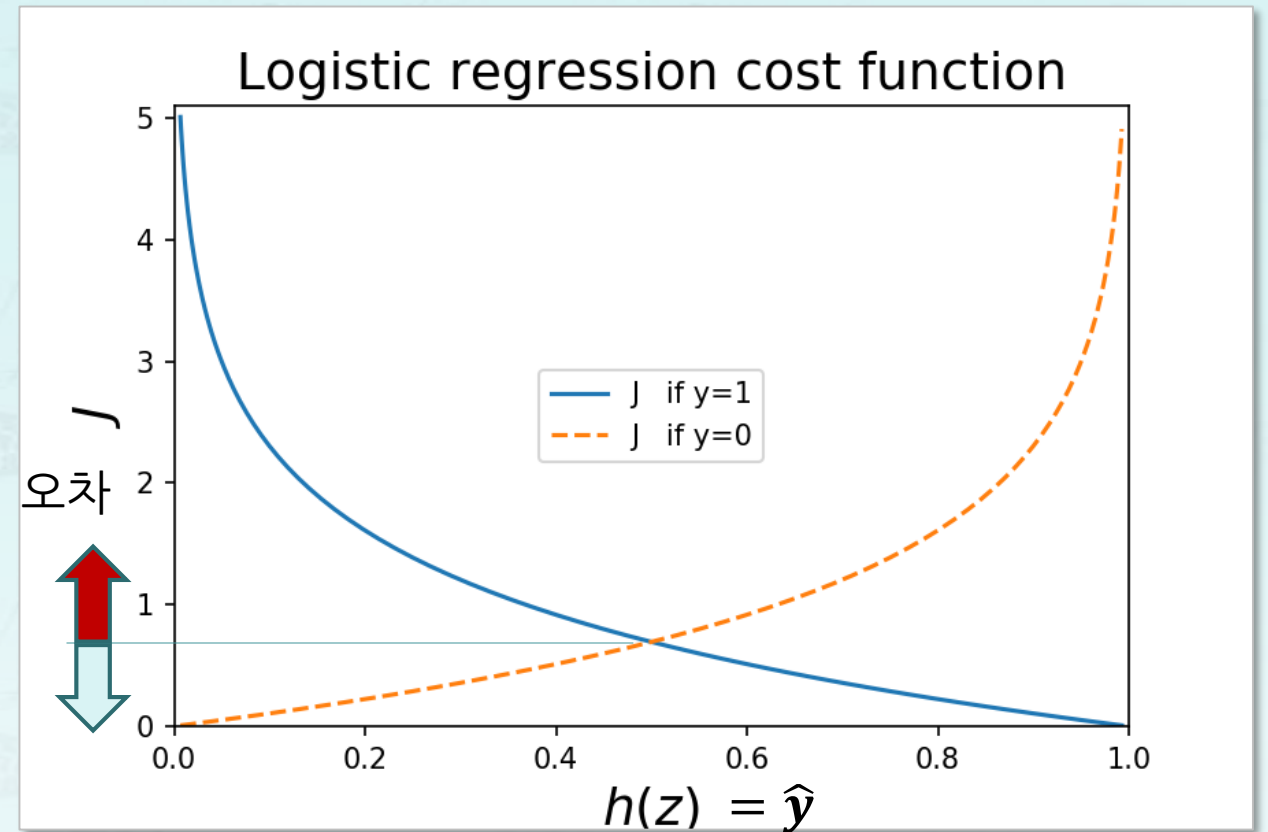
## 2. 로지스틱 회귀: 비용함수

- 비용함수  $J$

(1)  $y = 1$ ,  $\hat{y} = 0.2$  일 때 오차는?

(a) 작다 (b) 크다

(a) 실선 (b) 점선



## 2. 로지스틱 회귀: 비용함수

### ■ 비용함수 $J$

(1)  $y = 1$ ,  $\hat{y} = 0.2$  일 때 오차는?

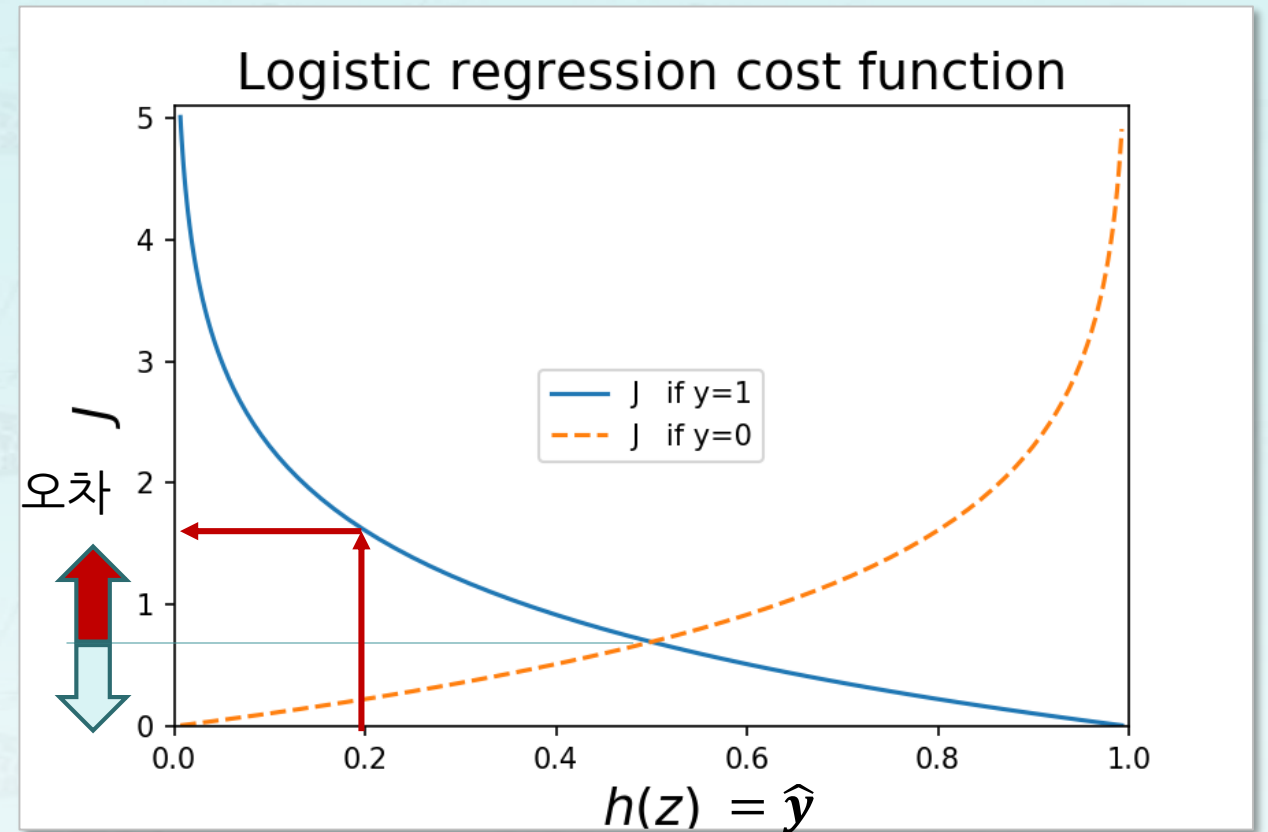
(a) 작다 (b) 크다

(a) 실선 (b) 점선

(2)  $y = 1$ ,  $\hat{y} = 0.8$  일 때 오차는?

(a) 작다 (b) 크다

(a) 실선 (b) 점선



## 2. 로지스틱 회귀: 비용함수

### ■ 비용함수 $J$

(1)  $y = 1$ ,  $\hat{y} = 0.2$  일 때 오차는?

(a) 작다 (b) 크다

(a) 실선 (b) 점선

(2)  $y = 1$ ,  $\hat{y} = 0.8$  일 때 오차는?

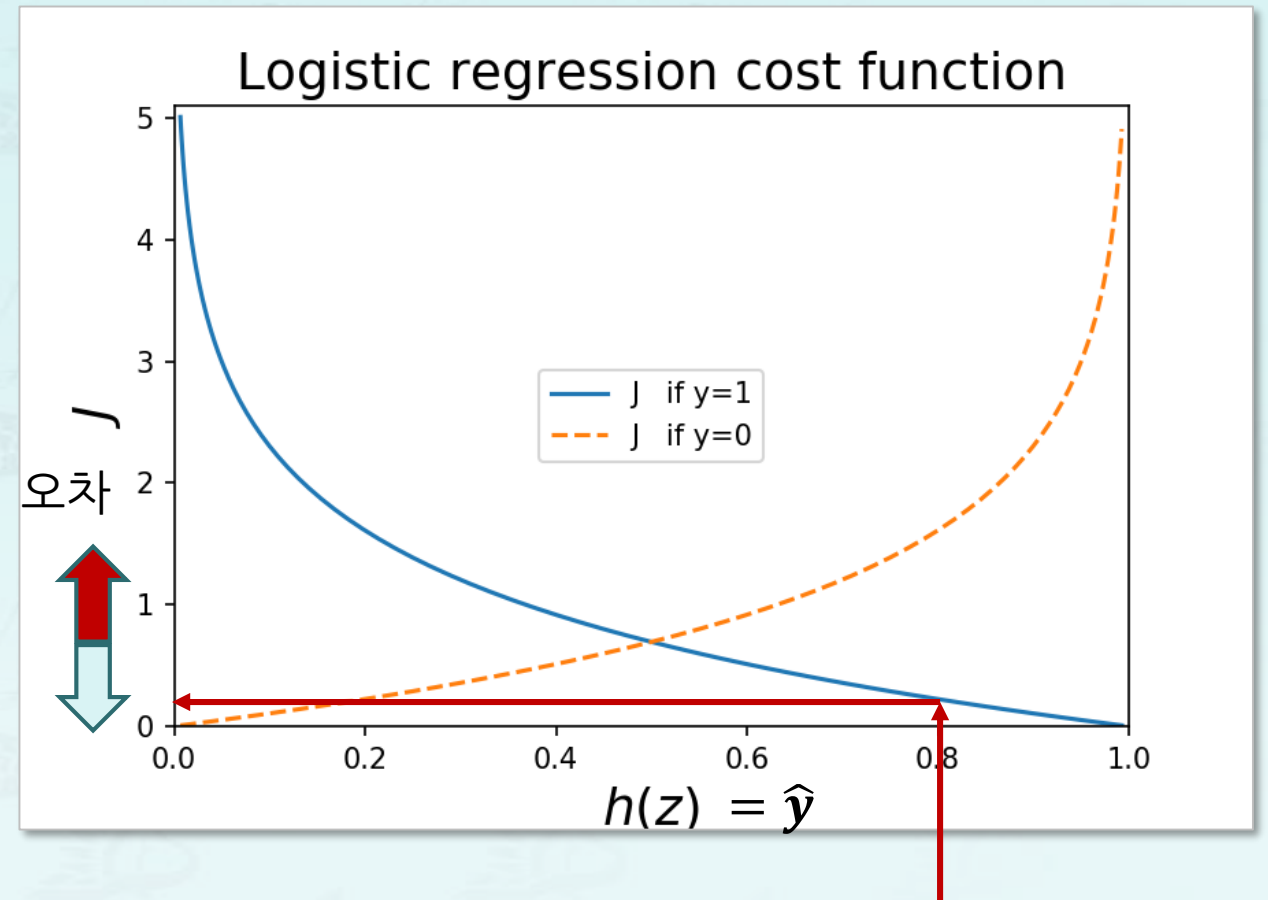
(a) 작다 (b) 크다

(a) 실선 (b) 점선

(3)  $y = 0$ ,  $\hat{y} = 0.9$  일 때 오차는?

(a) 작다 (b) 크다

(a) 실선 (b) 점선



## 2. 로지스틱 회귀: 비용함수

### ■ 비용함수 $J$

(1)  $y = 1$ ,  $\hat{y} = 0.2$  일 때 오차는?

(a) 작다 (b) 크다

(a) 실선 (b) 점선

(2)  $y = 1$ ,  $\hat{y} = 0.8$  일 때 오차는?

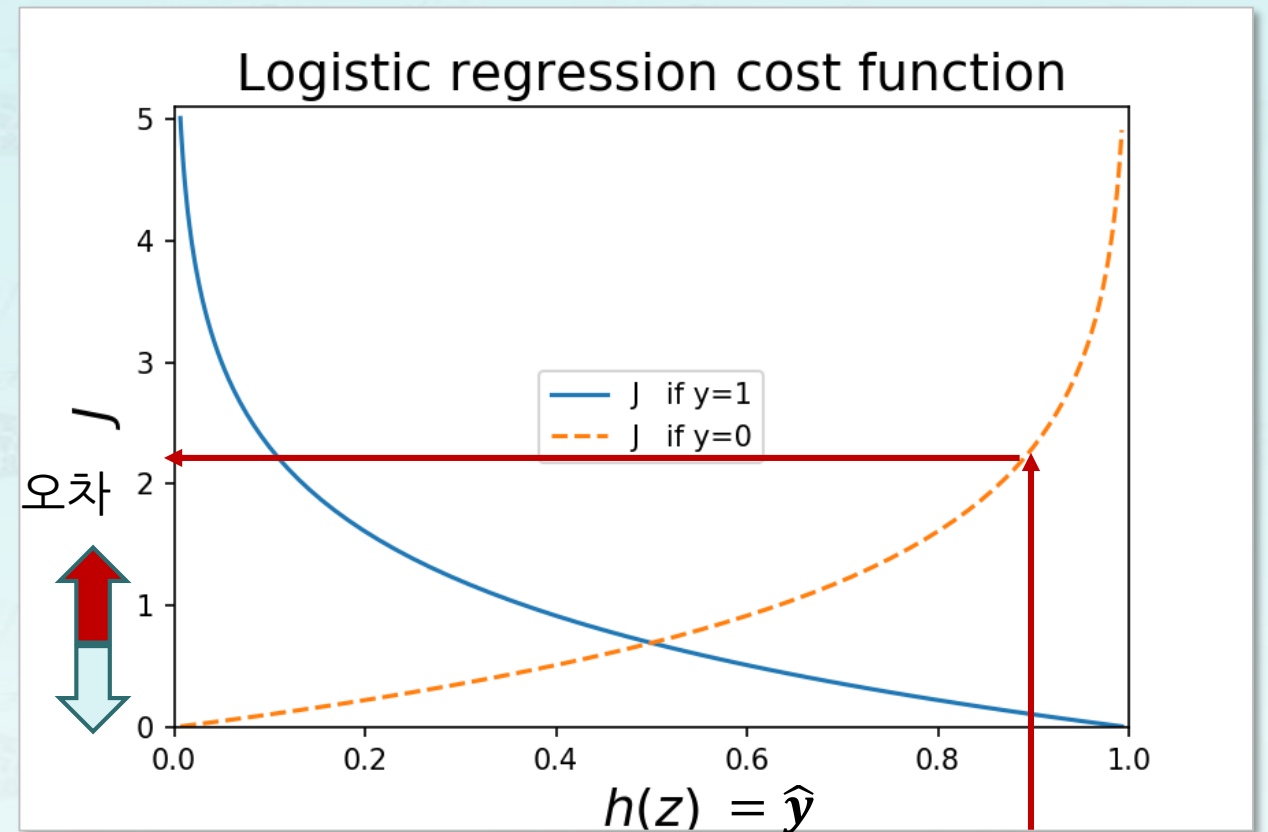
(a) 작다 (b) 크다

(a) 실선 (b) 점선

(3)  $y = 0$ ,  $\hat{y} = 0.9$  일 때 오차는?

(a) 작다 (b) 크다

(a) 실선 (b) 점선





## 2. 로지스틱 회귀: 비용함수

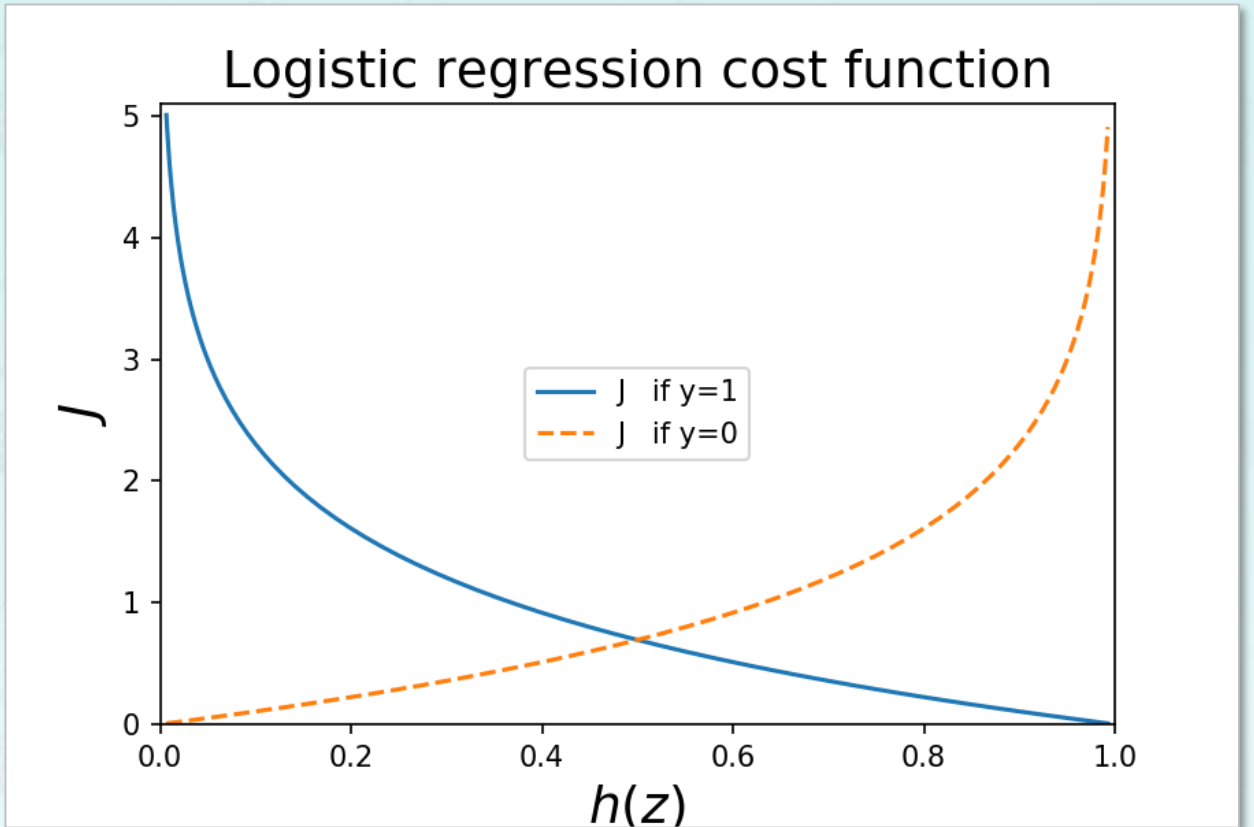
- 비용함수  $J$

$$J = \begin{cases} -\log(h(z)) & \text{if } y = 1. \\ -\log(1 - h(z)) & \text{if } y = 0. \end{cases}$$

$$e^{(i)} = \begin{cases} -\log(a^{[2](i)}) & \text{if } y = 1. \\ -\log(1 - a^{[2](i)}) & \text{if } y = 0. \end{cases}$$

- $a^{[2](i)}$

- 출력층 (i)번째 노드의 출력



$$e^{(i)} = -y \log(a^{[2](i)}) - (1 - y) \log(1 - a^{[2](i)})$$

## 2. 로지스틱 회귀: 비용함수

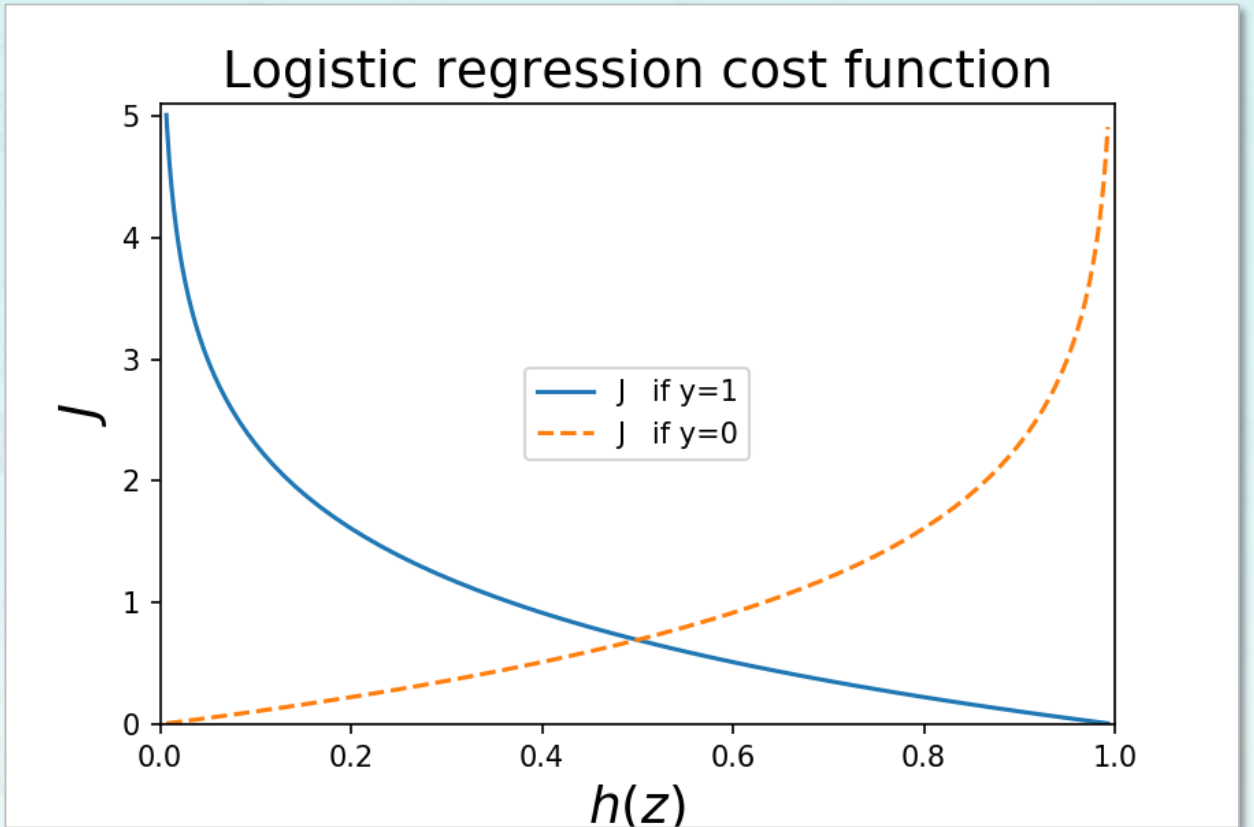
- 비용함수  $J$


$$J = \begin{cases} -\log(h(z)) & \text{if } y = 1. \\ -\log(1 - h(z)) & \text{if } y = 0. \end{cases}$$

$$e^{(i)} = \begin{cases} -\log(a^{[2](i)}) & \text{if } y = 1. \\ -\log(1 - a^{[2](i)}) & \text{if } y = 0. \end{cases}$$

- $a^{[2](i)}$

- 출력층 (i)번째 노드의 출력



$$e^{(i)} = -y \log(a^{[2](i)}) - (1 - y) \log(1 - a^{[2](i)})$$


## 2. 로지스틱 회귀: 비용함수

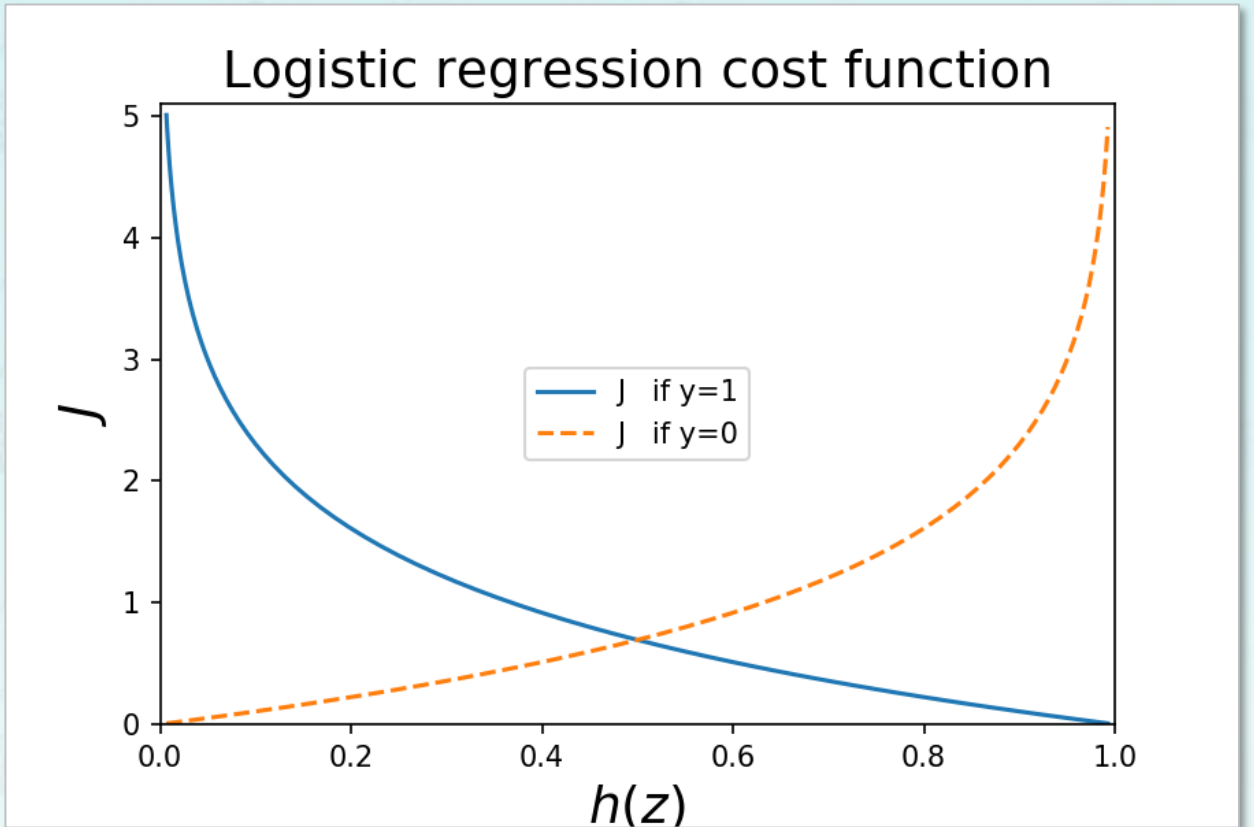
- 비용함수  $J$


$$J = \begin{cases} -\log(h(z)) & \text{if } y = 1. \\ -\log(1 - h(z)) & \text{if } y = 0. \end{cases}$$

$$e^{(i)} = \begin{cases} -\log(a^{[2](i)}) & \text{if } y = 1. \\ -\log(1 - a^{[2](i)}) & \text{if } y = 0. \end{cases}$$

- $a^{[2](i)}$

- 출력층 (i)번째 노드의 출력



$$e^{(i)} = -y \log(a^{[2](i)}) - (1 - y) \log(1 - a^{[2](i)})$$


## 2. 로지스틱 회귀: 비용함수

---

- 비용함수  $J$

$$e^{(i)} = -y \log(a^{[2](i)}) - (1 - y) \log(1 - a^{[2](i)})$$

## 2. 로지스틱 회귀: 비용함수

- 비용함수  $J$

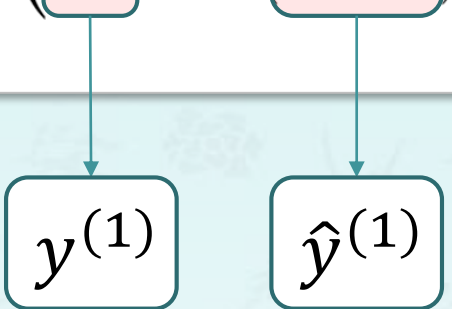
$$J(w) = -\frac{1}{m} \sum_{i=0}^m \left( y^{(i)} \log(a^{[2](i)}) + (1 - y^{(i)}) \log(1 - a^{[2](i)}) \right)$$



$$e^{(i)} = -y \log(a^{[2](i)}) - (1 - y) \log(1 - a^{[2](i)})$$

## 2. 로지스틱 회귀: 비용함수

- 비용함수  $J$

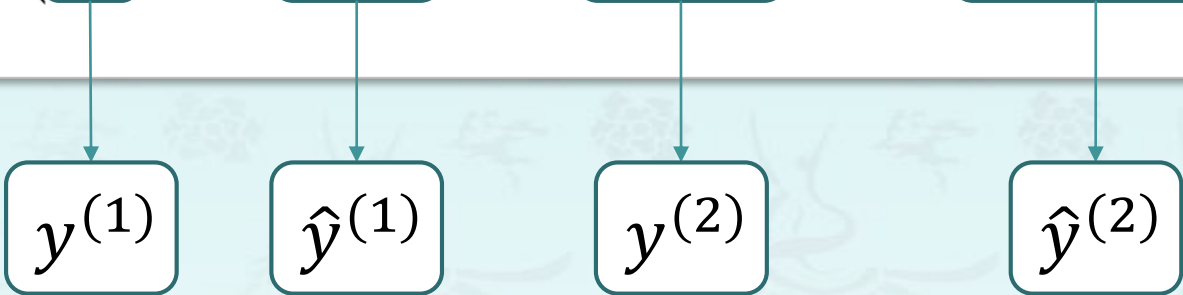
$$J(w) = -\frac{1}{m} \sum_{i=0}^m \left( y^{(i)} \log(a^{[2](i)}) + (1 - y^{(i)}) \log(1 - a^{[2](i)}) \right)$$


The diagram illustrates the mapping of terms in the cost function formula to specific variables for the first example. An arrow points from  $y^{(i)}$  in the formula to  $y^{(1)}$  in a box below. Another arrow points from  $a^{[2](i)}$  in the formula to  $\hat{y}^{(1)}$  in a box below.

- 교차 엔트로피(Cross Entropy)

## 2. 로지스틱 회귀: 비용함수

- 비용함수  $J$

$$J(w) = -\frac{1}{m} \sum_{i=0}^m \left( \boxed{y^{(i)}} \log(\boxed{a^{[2](i)}}) + \boxed{(1 - y^{(i)})} \log(\boxed{1 - a^{[2](i)}}) \right)$$


The diagram illustrates the components of the cost function for two data points. The first point's terms are  $y^{(1)}$  and  $\hat{y}^{(1)}$ , and the second point's terms are  $y^{(2)}$  and  $\hat{y}^{(2)}$ .

- 교차 엔트로피(Cross Entropy)

## 2. 로지스틱 회귀: 비용함수

- 비용함수  $J$

$$J(w) = -\frac{1}{m} \sum_{i=0}^m \left( y^{(i)} \log(a^{[2](i)}) + (1 - y^{(i)}) \log(1 - a^{[2](i)}) \right)$$

The diagram illustrates the components of the cost function for two data points. Arrows indicate the following mappings:

- $y^{(i)}$  maps to  $y^{(1)}$  and  $y^{(2)}$ .
- $a^{[2](i)}$  maps to  $\hat{y}^{(1)}$  and  $\hat{y}^{(2)}$ .
- $(1 - y^{(i)})$  maps to  $y^{(2)}$ .
- $(1 - a^{[2](i)})$  maps to  $\hat{y}^{(2)}$ .

- 교차 엔트로피(Cross Entropy)

$$J = - \sum_i y^{(i)} \log(\hat{y}^{(i)})$$



## 2. 로지스틱 회귀: 비용함수 미분

- 비용함수  $J$

$$J(w) = -\frac{1}{m} \sum_{i=0}^m \left( y^{(i)} \log(a^{[2](i)}) + (1 - y^{(i)}) \log(1 - a^{[2](i)}) \right)$$

$y$ : 클래스 레이블, 0 혹은 1  
 $a$ : 출력층의 출력,  $a^{[2](i)}$   
로지스틱 회귀 가설함수의 출력  
 $z: w x$

$$\frac{\partial J(w)}{\partial w_j} = \frac{\partial}{\partial w_j} \frac{-1}{m} \sum_{i=1}^m [y \log(a) + (1 - y) \log(1 - a)] \quad \because a = h(z) = \sigma(z)$$

## 2. 로지스틱 회귀: 비용함수 미분

### ■ 비용함수 $J$

$$J(w) = -\frac{1}{m} \sum_{i=0}^m \left( y^{(i)} \log(a^{[2](i)}) + (1 - y^{(i)}) \log(1 - a^{[2](i)}) \right)$$

$y$ : 클래스 레이블, 0 혹은 1

$a$ : 출력층의 출력,  $a^{[2](i)}$

로지스틱 회귀 가설함수의 출력

$z: w \cdot x$

$$\begin{aligned} \frac{\partial J(w)}{\partial w_j} &= \frac{\partial}{\partial w_j} \frac{-1}{m} \sum_{i=1}^m [y \log(a) + (1 - y) \log(1 - a)] && \because a = h(z) = \sigma(z) \\ &= \frac{\partial}{\partial w_j} \frac{-1}{m} \sum_{i=1}^m [y \log(\sigma(z)) + (1 - y) \log(1 - \sigma(z))] && \because \frac{d}{dx} \log(x) = \frac{1}{x} \\ &= \frac{-1}{m} \sum_{i=1}^m \left[ \frac{y}{\sigma(z)} \frac{\partial \sigma(z)}{\partial w_j} + \frac{(1 - y)}{1 - \sigma(z)} \frac{\partial (1 - \sigma(z))}{\partial w_j} \right] && \because \frac{d}{dx} f(g(x)) = f'(g(x)) g'(x) \end{aligned}$$

## 2. 로지스틱 회귀: 비용함수 미분

### ■ 비용함수 $J$

$$J(w) = -\frac{1}{m} \sum_{i=0}^m \left( y^{(i)} \log(a^{[2](i)}) + (1 - y^{(i)}) \log(1 - a^{[2](i)}) \right)$$

$y$ : 클래스 레이블, 0 혹은 1

$a$ : 출력층의 출력,  $a^{[2](i)}$

로지스틱 회귀 가설함수의 출력

$z$ :  $Wx$

$$\begin{aligned} \frac{\partial J(w)}{\partial w_j} &= \frac{\partial}{\partial w_j} \frac{-1}{m} \sum_{i=1}^m [y \log(a) + (1 - y) \log(1 - a)] && \because a = h(z) = \sigma(z) \\ &= \frac{\partial}{\partial w_j} \frac{-1}{m} \sum_{i=1}^m [y \log(\sigma(z)) + (1 - y) \log(1 - \sigma(z))] && \because \frac{d}{dx} \log(x) = \frac{1}{x} \\ &= \frac{-1}{m} \sum_{i=1}^m \left[ \frac{y}{\sigma(z)} \frac{\partial \sigma(z)}{\partial w_j} + \frac{(1 - y)}{1 - \sigma(z)} \frac{\partial (1 - \sigma(z))}{\partial w_j} \right] && \because \frac{d}{dx} f(g(x)) = f'(g(x))g'(x) \\ &= \frac{-1}{m} \sum_{i=1}^m \left[ \frac{y}{\sigma(z)} - \frac{(1 - y)}{1 - \sigma(z)} \right] \frac{\partial \sigma(z)}{\partial w_j} \end{aligned}$$

## 2. 로지스틱 회귀: 비용함수 미분

$$\begin{aligned}\frac{\partial J(w)}{\partial w_j} &= \frac{-1}{m} \sum_{i=1}^m \left[ \frac{y}{\sigma(z)} - \frac{(1-y)}{1-\sigma(z)} \right] \sigma'(z) x_j && \because z = w_j x_j \\&= \frac{-1}{m} \sum_{i=1}^m \left[ \frac{y(1-\sigma(z)) - \sigma(z)(1-y)}{\sigma(z)(1-\sigma(z))} \right] \sigma'(z) x_j \\&= \frac{-1}{m} \sum_{i=1}^m [y(1-\sigma(z)) - \sigma(z)(1-y)] x_j && \because \sigma'(z) = \sigma(z)(1-\sigma(z)) \\&= \frac{-1}{m} \sum_{i=1}^m [y - \sigma(z)] x_j \\&= -\frac{1}{m} \sum_{i=1}^m [y^{(i)} - \hat{y}^{(i)}] x_j^{(i)}\end{aligned}$$

## 2. 로지스틱 회귀: 비용함수 미분

$$\begin{aligned}\frac{\partial J(w)}{\partial w_j} &= \frac{-1}{m} \sum_{i=1}^m \left[ \frac{y}{\sigma(z)} - \frac{(1-y)}{1-\sigma(z)} \right] \sigma'(z) x_j && \because z = w_j x_j \\ &= \frac{-1}{m} \sum_{i=1}^m \left[ \frac{y(1-\sigma(z)) - \sigma(z)(1-y)}{\sigma(z)(1-\sigma(z))} \right] \sigma'(z) x_j \\ &= \frac{-1}{m} \sum_{i=1}^m [y(1-\sigma(z)) - \sigma(z)(1-y)] x_j && \because \sigma'(z) = \sigma(z)(1-\sigma(z)) \\ &= \frac{-1}{m} \sum_{i=1}^m [y - \sigma(z)] x_j \\ &= -\frac{1}{m} \sum_{i=1}^m [y^{(i)} - \hat{y}^{(i)}] x_j^{(i)}\end{aligned}$$

### 역전파 2: $W^{[2]}$ 의 오차함수 미분 - 4단계

$$\Delta w_{jk}^{[2]} = \frac{\partial E}{\partial w_{jk}} = \boxed{-(y_k - \hat{y}_k)} \cdot \boxed{g'(z_k)} \cdot \boxed{a_j}$$

$$\Delta W^{[2]} = \frac{\partial E}{\partial W^{[2]}} = \boxed{-E^{[2]}} \cdot g'(Z^{[2]}) \cdot A^{[1]T}$$

## 2. 로지스틱 회귀: 비용함수 미분

$$\begin{aligned}
 \frac{\partial J(w)}{\partial w_j} &= \frac{-1}{m} \sum_{i=1}^m \left[ \frac{y}{\sigma(z)} - \frac{(1-y)}{1-\sigma(z)} \right] \sigma'(z) x_j \\
 &= \frac{-1}{m} \sum_{i=1}^m \left[ \frac{y(1-\sigma(z)) - \sigma(z)(1-y)}{\sigma(z)(1-\sigma(z))} \right] \sigma'(z) x_j \\
 &= \frac{-1}{m} \sum_{i=1}^m [y(1-\sigma(z)) - \sigma(z)(1-y)] x_j \\
 &= \frac{-1}{m} \sum_{i=1}^m [y - \sigma(z)] x_j \\
 &= -\frac{1}{m} \sum_{i=1}^m [y^{(i)} - \hat{y}^{(i)}] x_j^{(i)}
 \end{aligned}$$

$$\because z = w_j x_j$$

$$\because \sigma'(z) = \sigma(z)(1-\sigma(z))$$

$$\Delta W^{[2]} = \frac{\partial J}{\partial W^{[2]}} = -\frac{1}{m} E^{[2]}(1) A^{[1]T}$$

### 역전파 2: $W^{[2]}$ 의 오차함수 미분 - 4단계

$$\Delta w_{jk}^{[2]} = \frac{\partial E}{\partial w_{jk}} = -(y_k - \hat{y}_k) \cdot g'(z_k) \cdot a_j$$

$$\Delta W^{[2]} = \frac{\partial E}{\partial W^{[2]}} = -E^{[2]} \cdot g'(Z^{[2]}) \cdot A^{[1]T}$$

## 2. 로지스틱 회귀: 비용함수 미분

$$\begin{aligned}
 \frac{\partial J(w)}{\partial w_j} &= \frac{-1}{m} \sum_{i=1}^m \left[ \frac{y}{\sigma(z)} - \frac{(1-y)}{1-\sigma(z)} \right] \sigma'(z) x_j && \because z = w_j x_j \\
 &= \frac{-1}{m} \sum_{i=1}^m \left[ \frac{y(1-\sigma(z)) - \sigma(z)(1-y)}{\sigma(z)(1-\sigma(z))} \right] \sigma'(z) x_j \\
 &= \frac{-1}{m} \sum_{i=1}^m [y(1-\sigma(z)) - \sigma(z)(1-y)] x_j && \because \sigma'(z) = \sigma(z)(1-\sigma(z)) \\
 &= \frac{-1}{m} \sum_{i=1}^m [y - \sigma(z)] x_j \\
 &= -\frac{1}{m} \sum_{i=1}^m [y^{(i)} - \hat{y}^{(i)}] x_j^{(i)}
 \end{aligned}$$

$$\Delta W^{[2]} = \frac{\partial J}{\partial W^{[2]}} = -\frac{1}{m} E^{[2]} (1) A^{[1]T}$$

### 역전파 2: $W^{[2]}$ 의 오차함수 미분 - 4단계

$$\Delta w_{jk}^{[2]} = \frac{\partial E}{\partial w_{jk}^{[2]}} = -(y_k - \hat{y}_k) \cdot g'(z_k) \cdot a_j$$

$$\Delta W^{[2]} = \frac{\partial E}{\partial W^{[2]}} = -E^{[2]} \cdot g'(Z^{[2]}) \cdot A^{[1]T}$$



## 2. 로지스틱 회귀: 비용함수 미분

$$\begin{aligned}
 \frac{\partial J(w)}{\partial w_j} &= \frac{-1}{m} \sum_{i=1}^m \left[ \frac{y}{\sigma(z)} - \frac{(1-y)}{1-\sigma(z)} \right] \sigma'(z) x_j \\
 &= \frac{-1}{m} \sum_{i=1}^m \left[ \frac{y(1-\sigma(z)) - \sigma(z)(1-y)}{\sigma(z)(1-\sigma(z))} \right] \sigma'(z) x_j \\
 &= \frac{-1}{m} \sum_{i=1}^m [y(1-\sigma(z)) - \sigma(z)(1-y)] x_j \\
 &= \frac{-1}{m} \sum_{i=1}^m [y - \sigma(z)] x_j \\
 &= -\frac{1}{m} \sum_{i=1}^m [y^{(i)} - \hat{y}^{(i)}] x_j^{(i)}
 \end{aligned}$$

$$\because z = w_j x_j$$

$$\because \sigma'(z) = \sigma(z)(1-\sigma(z))$$

$$\Delta W^{[2]} = \frac{\partial J}{\partial W^{[2]}} = -\frac{1}{m} E^{[2]}(1) A^{[1]T}$$

### 역전파 2: $W^{[2]}$ 의 오차함수 미분 - 4단계

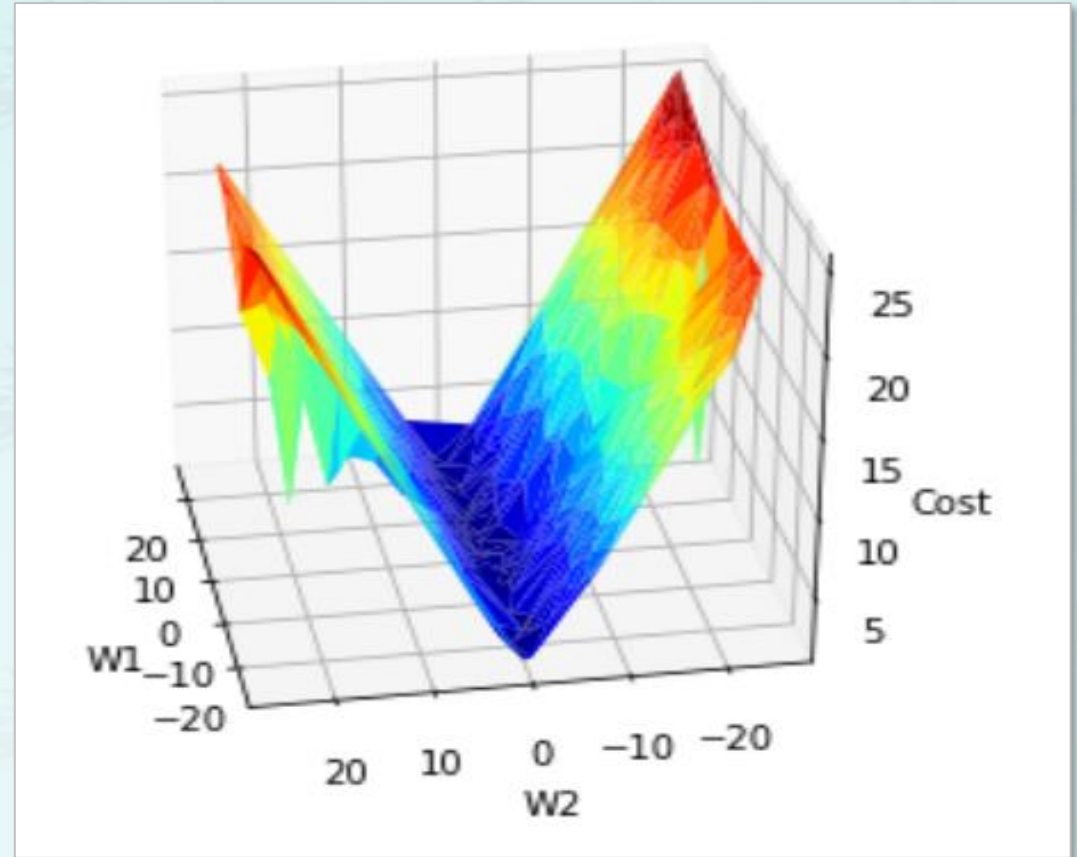
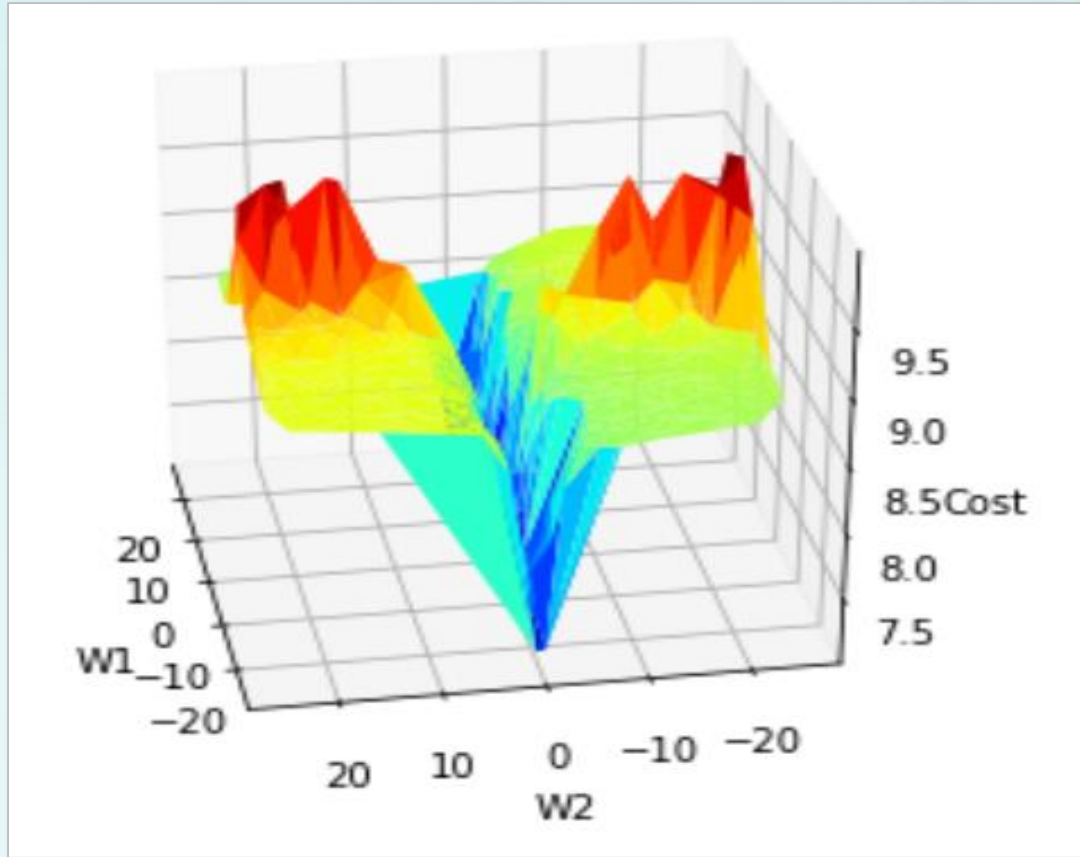
$$\Delta w_{jk}^{[2]} = \frac{\partial E}{\partial w_{jk}} = \boxed{-(y_k - \hat{y}_k)} \cdot \boxed{g'(z_k)} \cdot \boxed{a_j}$$

$$\Delta W^{[2]} = \frac{\partial E}{\partial W^{[2]}} = \boxed{-E^{[2]}} \cdot g'(Z^{[2]}) \cdot A^{[1]T}$$



## 2. 로지스틱 회귀: 비용함수 미분

- 제곱 합 오차(**SSE**)함수
- 교차 엔트로피(**Cross Entropy**)



# 로지스틱 회귀

---

- 학습 정리
  - 로지스틱 함수 이해하기
  - 로지스틱 회귀의 비용함수를 이해하기
  - 로지스틱 회귀 비용함수(교차 엔트로피)를 미분하기
- 차시 예고
  - **11-1 로지스틱 회귀 3**

11주차(3/3)

# 로지스틱 회귀 2

파이썬으로 배우는 기계학습

한동대학교  
김영섭 교수

여러분 곁에 항상 열려 있는 K-MOOC 강의실에서 만나 뵙기를 바랍니다.

## 2. 로지스틱 회귀: 비용함수 미분

- 제곱 합 오차(**SSE**)함수

$$\begin{aligned}\frac{\partial E}{\partial w_{jk}} &= -(y_k - \hat{y}_k) \cdot g'(z_k) \cdot a_j \\ &= -(y_k - \hat{y}_k) \cdot \sigma(z_k)(1 - \sigma(z_k)) \cdot a_j\end{aligned}$$

- 교차 엔트로피(**Cross Entropy**)

## 2. 로지스틱 회귀: 비용함수 미분

- 제곱 합 오차(**SSE**)함수

$$\begin{aligned}\frac{\partial E}{\partial w_{jk}} &= -(y_k - \hat{y}_k) \cdot g'(z_k) \cdot a_j \\ &= -(y_k - \hat{y}_k) \cdot \sigma(z_k)(1 - \sigma(z_k)) \cdot a_j\end{aligned}$$

- 교차 엔트로피(**Cross Entropy**)

$$\begin{aligned}\frac{\partial J}{\partial w_{jk}} &= -\frac{1}{m}(y_k - \hat{y}_k) \cdot g'(z_k) \cdot a_j \\ &= -\frac{1}{m}(y_k - \hat{y}_k) \cdot a_j\end{aligned}$$

## 2. 로지스틱 회귀: 비용함수 미분

- 제곱 합 오차(**SSE**)함수

$$\begin{aligned}\frac{\partial E}{\partial w_{jk}} &= -(y_k - \hat{y}_k) \cdot g'(z_k) \cdot a_j \\ &= -(y_k - \hat{y}_k) \cdot \sigma(z_k)(1 - \sigma(z_k)) \cdot a_j\end{aligned}$$

- 교차 엔트로피(**Cross Entropy**)

$$\begin{aligned}\frac{\partial J}{\partial w_{jk}} &= -\frac{1}{m}(y_k - \hat{y}_k) \cdot g'(z_k) \cdot a_j \\ &= -\frac{1}{m}(y_k - \hat{y}_k) \cdot a_j\end{aligned}$$

- 선형회귀 가설함수

- $y = h_w(x) = WX$

- 로지스틱 회귀 가설함수

- $y = h_w(x) = \frac{1}{1 + e^{-WX}}$