

인공지능개론

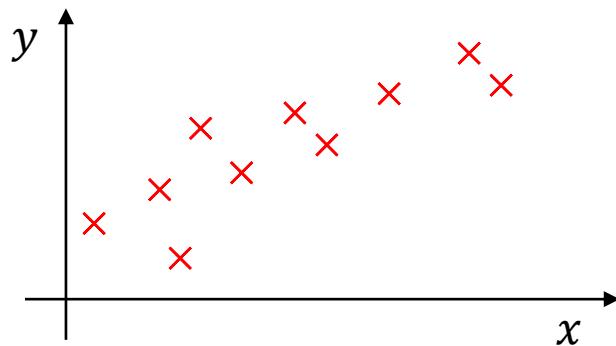
기계학습

Logistic regression

Prediction

- 인공지능 투자 알고리즘
- 날씨 예측 알고리즘
- 수요 공급 예측 알고리즘

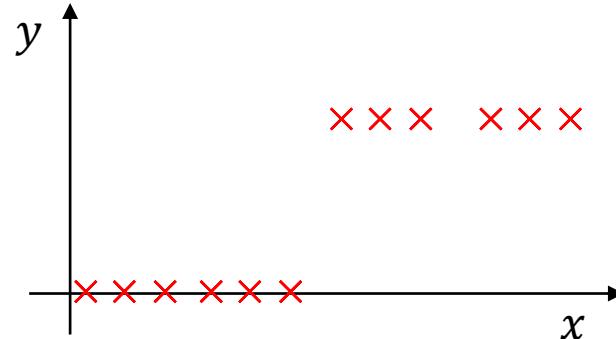
$$y \in \mathbb{R}^n$$



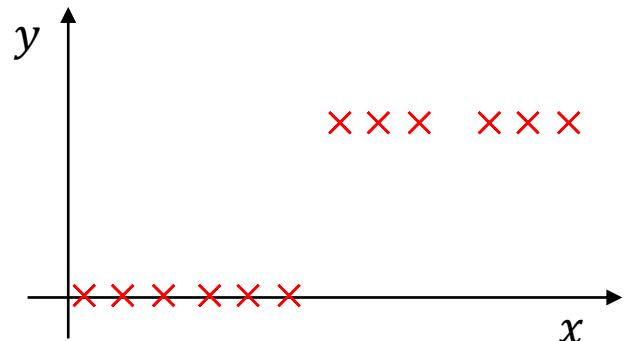
Classification

- 병 진단 기술
- 불법 게임 소프트웨어 판단
- 사진 분류 기술

$$y \in \{0,1\}$$



Logistic regression



$$h_{\theta}(x) = \theta^T x, \quad \theta = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_n \end{bmatrix}, x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

Threshold classifier: $\rho = 0.5$

if $h_{\theta}(x) < \rho, \ y = 0$
if $h_{\theta}(x) \geq \rho, \ y = 1$

How about additional data set??

Logistic regression

$h_{\theta}(x)$ can be > 1 or < 0



$0 \leq h_{\theta}(x) \leq 1$

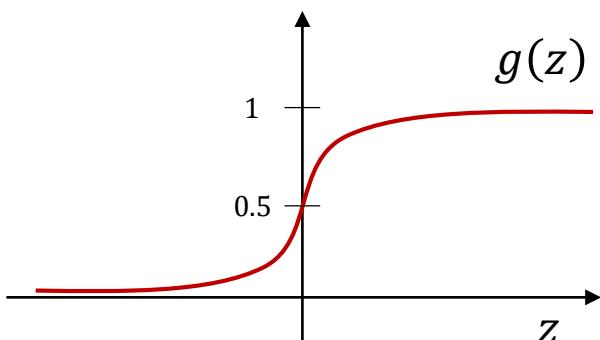
Logistic regression

Logistic regression

We want $0 \leq h_\theta(x) \leq 1$ for all x

How to design the model (function) $h_\theta(x)$??

Logistic function / Sigmoid function



$$h_\theta(x) = \theta^T x$$

↓

$$h_\theta(x) = g(\theta^T x) = g(z)$$

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

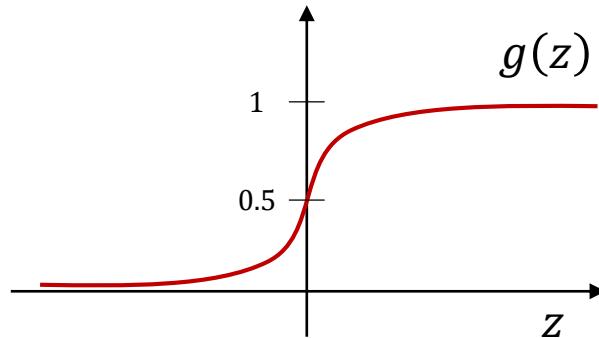
$h_\theta(x)$ can represent estimated probability

$$h_\theta(x) = P(y = 1|x; \theta)$$

$$P(y = 1|x; \theta) + P(y = 0|x; \theta) = 1$$

Logistic regression

Now, we can classify the dataset with logistic regression



$$h_{\theta}(x) = g(\theta^T x) = g(z)$$

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

Prediction result $y = 0$

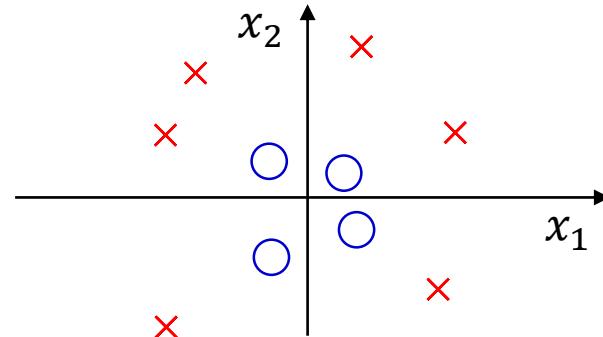
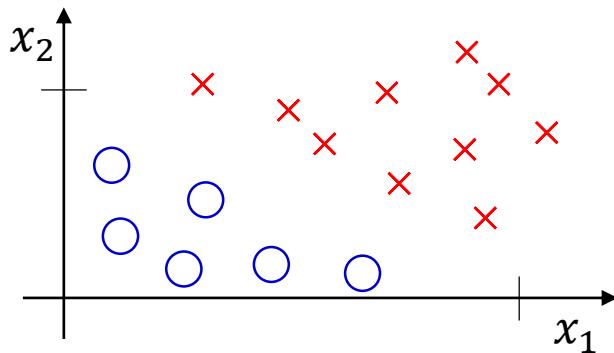
if $g(z) = h_{\theta}(x) < 0.5,$

Prediction result $y = 1$

if $g(z) = h_{\theta}(x) \geq 0.5,$

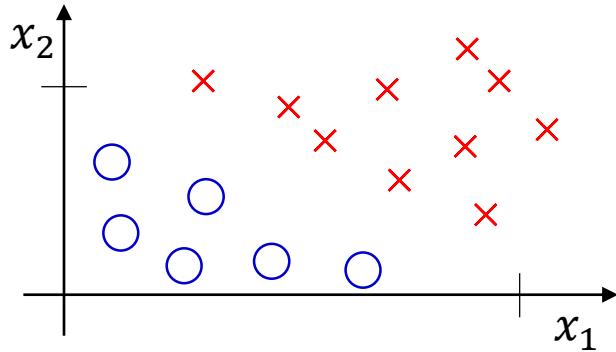
$\theta^T x = z$ is the only solution?

How can we handle multi variable?



Logistic regression

Define “Decision boundary”



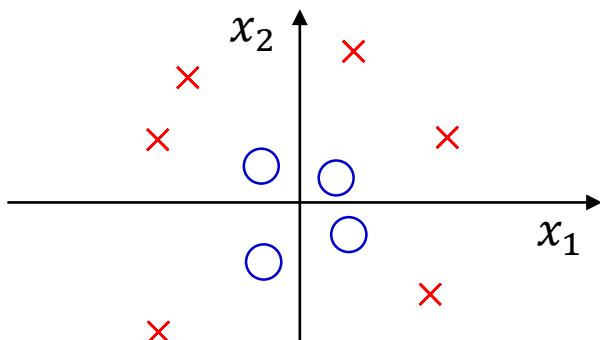
$$h_{\theta}(x) = g(\theta^T x) = g(z)$$



$$\begin{aligned} z &= \theta^T x \\ &= \theta_0 x_0 + \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_n x_n \end{aligned}$$



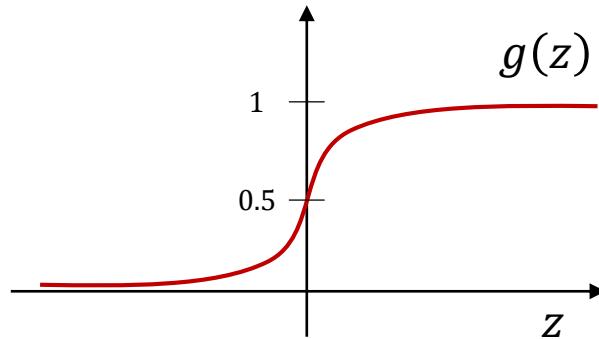
$$z = \theta_0 x_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_1^2 + \theta_4 x_2^2$$



$$\begin{aligned} z &= \theta_0 x_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_1^2 + \theta_4 x_2^2 \\ &\quad + \theta_5 x_1 x_2 + \theta_6 x_1 x_2^2 \dots \end{aligned}$$

Logistic regression

How to fine parameters θ ??

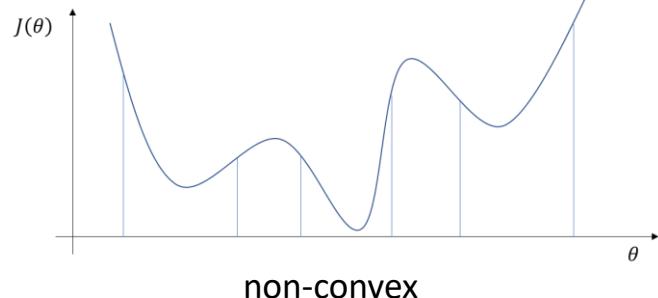


$$h_{\theta}(x) = g(\theta^T x) = g(z)$$

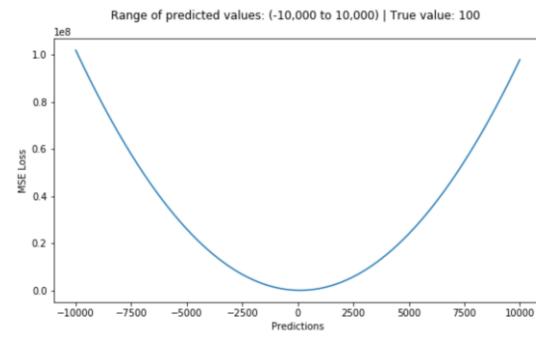
$$g(z) = \frac{1}{1 + e^{-z}}$$

remind linear regression

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^m (h(x^i) - y^i)^2$$



non-convex



convex

Logistic regression

How to fine parameters θ ??

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^m (h(x^i) - y^i)^2 \quad \xrightarrow{\hspace{1cm}} \quad J(\theta) = \text{Cost}(h_\theta(x), y)$$

Logistic regression

$$J(\theta) = \text{Cost}(h_\theta(x), y) \quad \left\{ \begin{array}{ll} -\log(h_\theta(x)) & \text{if } y = 1 \\ -\log(1 - h_\theta(x)) & \text{if } y = 0 \end{array} \right.$$

Logistic regression

Logistic regression cost function

$$\begin{aligned} J(\theta) &= \frac{1}{m} \sum_{i=1}^m \text{Cost}(h_\theta(x^{(i)}), y^{(i)}) \\ &= -\frac{1}{m} \left[\sum_{i=1}^m y^{(i)} \log h_\theta(x^{(i)}) + (1 - y^{(i)}) \log (1 - h_\theta(x^{(i)})) \right] \end{aligned}$$

Want $\min_{\theta} J(\theta)$:

Repeat {

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta)$$

}

(simultaneously update all θ_j)

$$h_\theta(x) = \frac{1}{1 + e^{-\theta^T x}} \quad \frac{\partial}{\partial \theta_j} J(\theta) = \sum_{i=1}^m (h_\theta(x^{(i)}) - y^{(i)}) x_j^{(i)}$$

Logistic regression

Summary

- Prediction vs Classification
- Linear vs Logistic
- Parameters
- Cost function
- Decision boundary : linear & non-linear

$$h_{\theta}(x) = \theta^T x \quad h_{\theta}(x) = \frac{1}{1 + e^{-\theta^T x}}$$

Logistic regression cost function

$$\begin{aligned} J(\theta) &= \frac{1}{m} \sum_{i=1}^m \text{Cost}(h_{\theta}(x^{(i)}), y^{(i)}) \\ &= -\frac{1}{m} \left[\sum_{i=1}^m y^{(i)} \log h_{\theta}(x^{(i)}) + (1 - y^{(i)}) \log (1 - h_{\theta}(x^{(i)})) \right] \end{aligned}$$

Want $\min_{\theta} J(\theta)$:

Repeat {

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta)$$

}

(simultaneously update all θ_j)

What's next?

- Classification : Support vector machine