Cost function

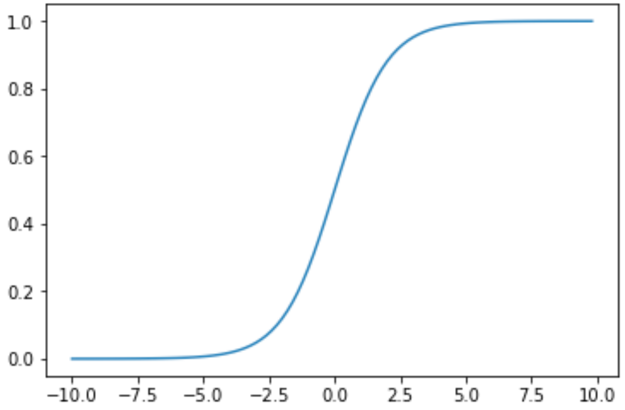
Logistic Regression Classifier

Director of TEAMLAB
Sungchul Choi

Logistic Regression에서 Weight 학습하기

가설 함수

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



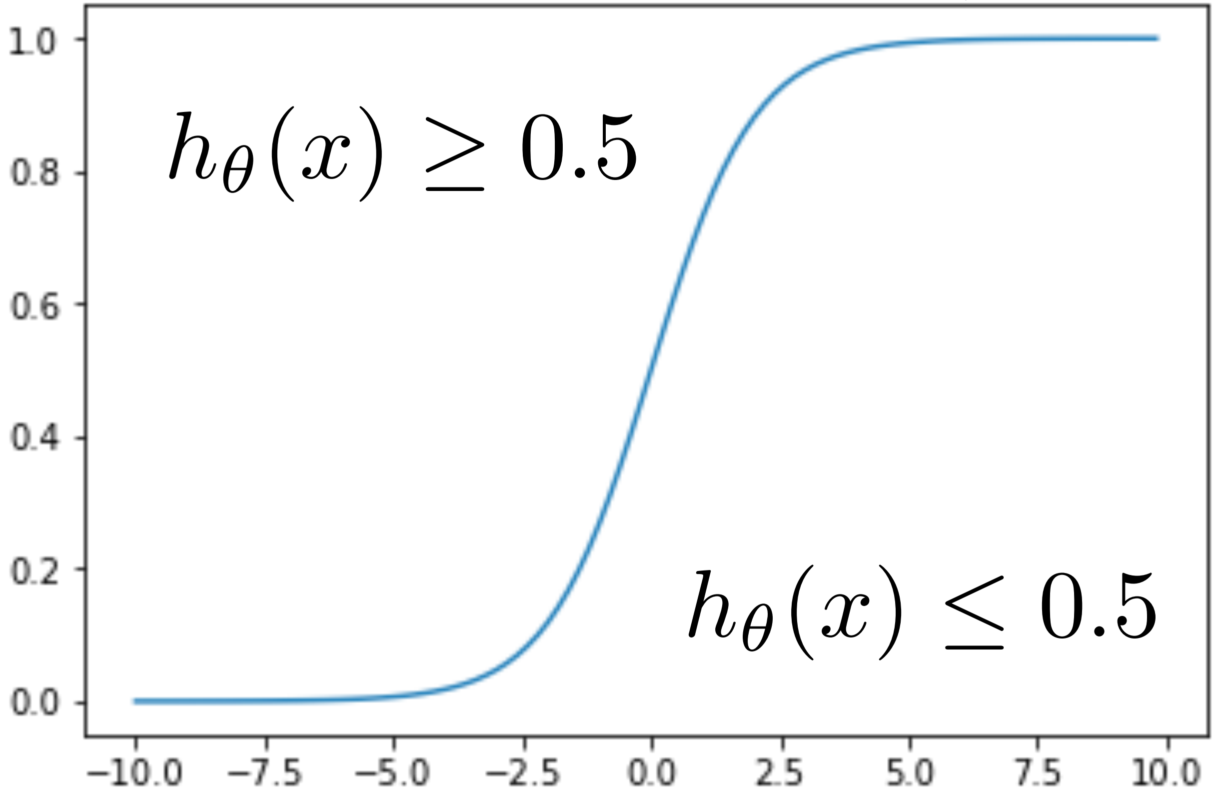
where:

$$\begin{aligned} z &= w_0x_0 + w_1x_1 + \cdots + w_nx_n \\ &= \theta^T\mathbf{x} \end{aligned}$$

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

가설 함수

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



Training θ

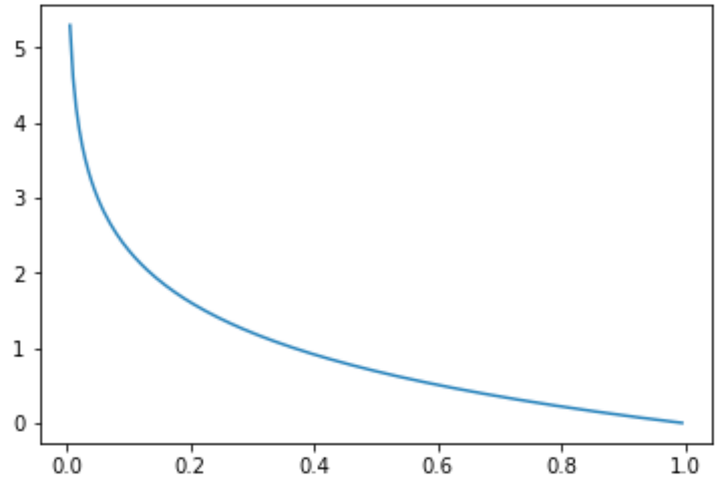
$$h_\theta(x) = \frac{1}{1 + e^{-\theta^T \mathbf{x}}}$$

ID	RPM	VIBRATION	STATUS
1	568	585	good
2	586	565	good
3	609	536	good
4	616	492	good
5	632	465	good
6	652	528	good
7	655	496	good
8	660	471	good
9	688	408	good
10	696	399	good

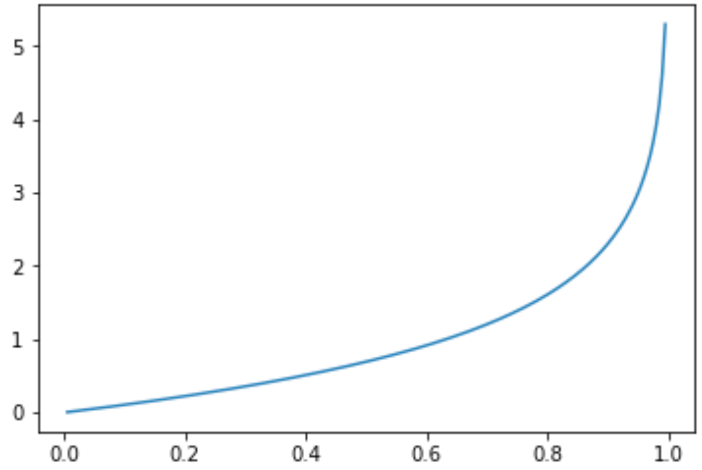
$$\theta^T \mathbf{x} = w_0 x_0 + w_1 x_1 + \cdots + w_n x_n$$

$$y = 0 \text{ or } 1$$

Cost Function



$$\mathrm{Cost}(h_\theta(x), y) = \begin{cases} -\log(h_\theta(x)) & \text{if } y = 1 \\ -\log(1 - h_\theta(x)) & \text{if } y = 0 \end{cases}$$



Cost Function

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m} \mathrm{Cost}\left(h_\theta(x^{(i)}), y^{(i)}\right)$$

$$= -\frac{1}{m}\sum_{i=1}^{m}[y^{(i)} \log h_\theta(x^{(i)}) + (1 - y^{(i)}) \log(1 - h_\theta(x^{(i)}))]$$

find θ, where $\min\limits_{\theta} J(\theta)$

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

Partial derivation of cost function

$$J(\theta) = -\frac{1}{m}\sum_{i=1}^{m}\left[-y^i(\log(1+e^{-\theta x^i})) + (1-y^i)(-\theta x^i - \log(1+e^{-\theta x^i}))\right]$$

$$J(\theta) = -\frac{1}{m}\sum_{i=1}^{m}\left[y_i\theta x^i - \theta x^i - \log(1+e^{-\theta x^i})\right]$$

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

$$= -\frac{1}{m}\sum_{i=1}^{m}\left[y_i\theta x^i - \log(1+e^{\theta x^i})\right]$$

$$\begin{aligned} -\theta x^i - \log(1+e^{-\theta x^i}) &= -\left[\log e^{\theta x^i} + \log(1+e^{-\theta x^i})\right] \\ &= -\log(1+e^{\theta x^i}). \end{aligned}$$

Partial derivation of cost function

$$-\frac{1}{m}\sum_{i=1}^{m}\left[y_i\theta x^i - \log(1+e^{\theta x^i})\right]$$

$$z = w_0x_0 + w_1x_1 + \cdots + w_nx_n$$
$$= \theta^T\mathbf{x}$$

$\boldsymbol{\theta}$ 에 관하여 미분하면

$$\frac{\partial}{\partial\theta_j}y_i\theta x^i = y_ix_j^i \qquad \frac{d}{dx}ln(2x) = \frac{2}{2x} = \frac{1}{x} \qquad \frac{d}{dx}e^{2x} = 2e^x$$

$$\frac{\partial}{\partial\theta_j}\log(1+e^{\theta x^i}) = \frac{x_j^ie^{\theta x^i}}{1+e^{\theta x^i}} = x_j^ih_\theta(x^i),$$

Partial derivation of cost function

$$\frac{\partial}{\partial \theta_j} \log(1 + e^{\theta x^i}) = \frac{x_j^i e^{\theta x^i}}{1 + e^{\theta x^i}} = x_j^i h_\theta(x^i),$$

$$\frac{x_j^i e^{\theta x^i}}{1 + e^{\theta x^i}} = \frac{x_j^i}{e^{-\theta x^i} * (1 + e^{\theta x^i})}$$

$$= \frac{x_j^i}{e^{-\theta x^i} + e^{-\theta x^i + \theta x^i}} = \frac{x_j^i}{e^{-\theta x^i} + e^0}$$

$$= \frac{x_j^i}{e^{-\theta x^i} + 1} = \frac{x_j^i}{1 + e^{-\theta x^i}}$$

$$h_\theta(x^i) = \frac{1}{1 + e^{\theta x^i}}$$

$$= x_j^i * h_\theta(x^i)$$

Partial derivation of cost function

$$-\frac{1}{m}\sum_{i=1}^{m}\left[y_i\theta x^i - \log(1+e^{\theta x^i})\right]$$

$$\frac{\partial}{\partial\theta_j}y_i\theta x^i = y_i x_j^i \qquad \frac{\partial}{\partial\theta_j}\log(1+e^{\theta x^i}) = \frac{x_j^i e^{\theta x^i}}{1+e^{\theta x^i}} = x_j^i h_\theta(x^i),$$

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

Weight update

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

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

모든 $\boldsymbol{\theta_j}$ 동시에 업데이트

$$:= \theta_j - \alpha \sum_{i=1}^{m} (h_\theta(x^i) - y^i) x_j^i$$

TEAMLAB

Human knowledge belongs to the world.