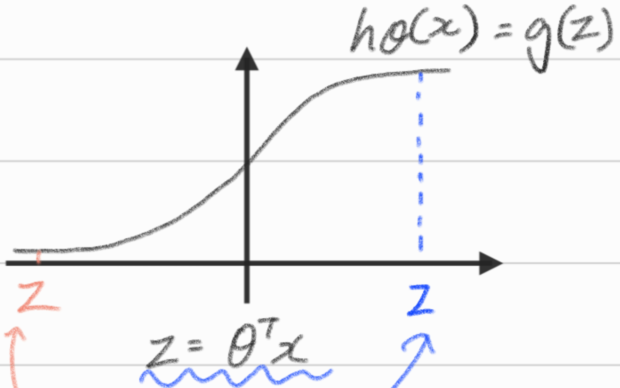
Optimization Objective

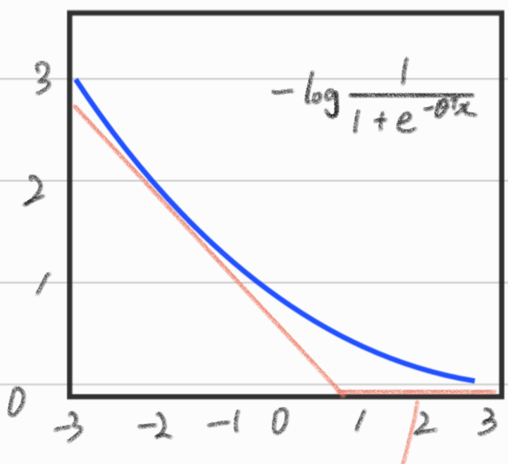
$$h_\theta(x) = \frac{1}{1+e^{-\theta^T x}}$$

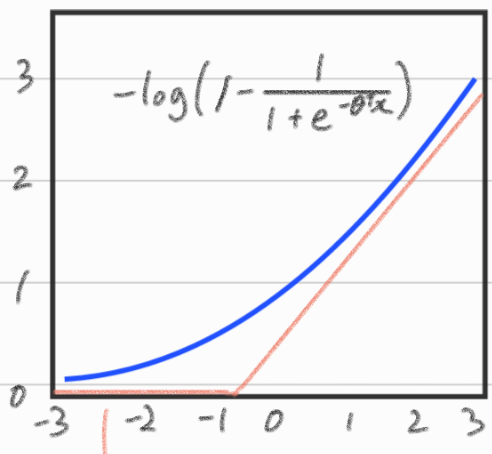


If $y=1$, we want $h_\theta(x) \approx 1$, $\theta^T x \gg 0$

$y=0$ ≈ 0 $\ll 0$

Cost: $-y \log \frac{1}{1+e^{-\theta^T x}} - (1-y) \log \left(1 - \frac{1}{1+e^{-\theta^T x}}\right)$





$\text{Cost}_1(z)$ $\text{Cost}_0(z)$

for SVM

Support Vector Machine (SVM)

$$\min_\theta C \sum_{i=1}^{m} \left[y^{(i)} \text{cost}_1(\theta^T x^{(i)}) + (1-y^{(i)}) \text{cost}_0(\theta^T x^{(i)}) \right] + \frac{1}{2} \sum_{i=1}^{n} \theta_j^2$$

규제

$\Rightarrow$ parameters (θ) learned by SVM

$A + \lambda B$ 람다가 크면 규제가 큼

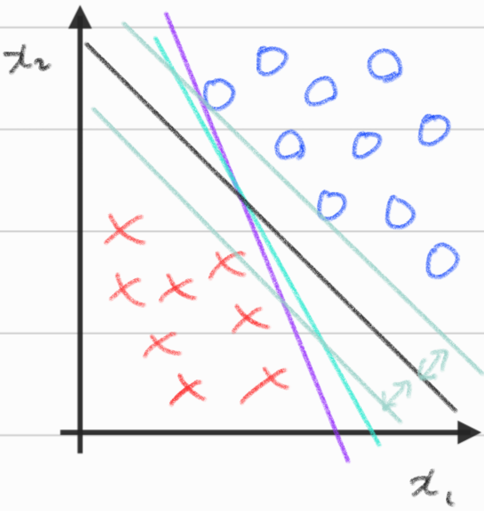
$C A + B$ C가 작으면 규제가 큼 $(C = \frac{1}{\lambda})$

$$h_\theta(x) \begin{cases} 1 & \text{if } \theta^T x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Large Margin Intuition

Whenever $y^{(i)} = 1$: $\theta^T x^{(i)} \geq 1$

Whenever $y^{(i)} = 0$: $\theta^T x^{(i)} \leq -1$



모두 완벽히 분류하지만,

가 가장 큰 margin 가짐.

Consequence of Optimization

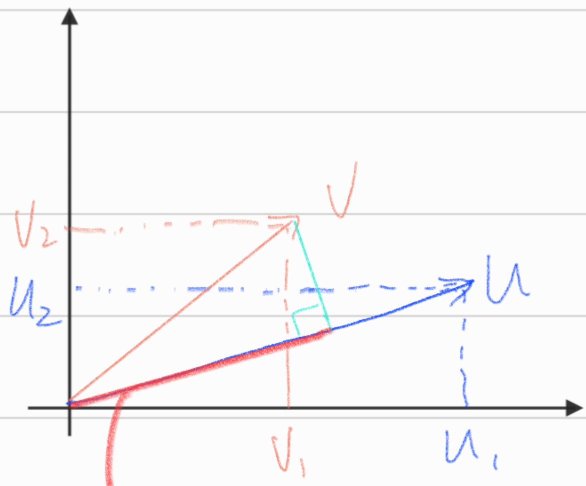
Separates positive and negative ex. with as big a margin as possible

Thus, Sensitive to outlier,

Mathematics Behind Large Margin Classification

Vector Inner Product

$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$



$\|u\|$ = length of vector u

$= \sqrt{u_1^2 + u_2^2} \in \mathbb{R}$ (피타고라스 정리)

p = length of projection of v onto u.

$u^T v = p \cdot \|u\|$

$= u_1 v_1 + u_2 v_2$