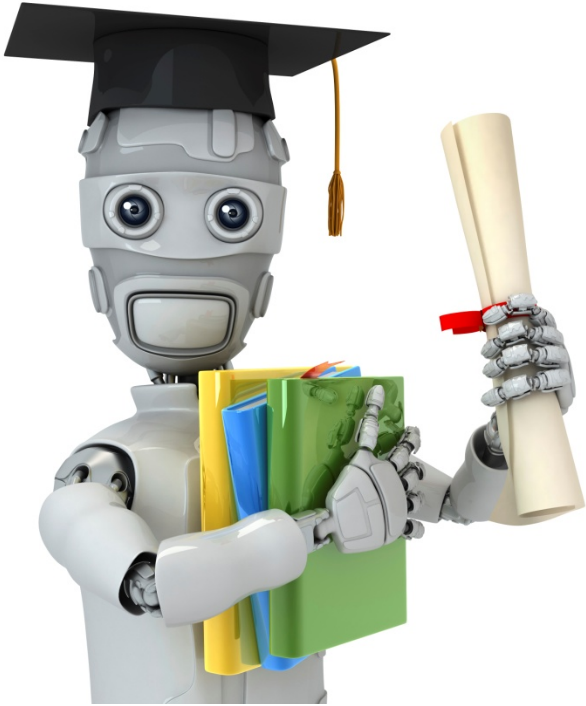


Machine Learning

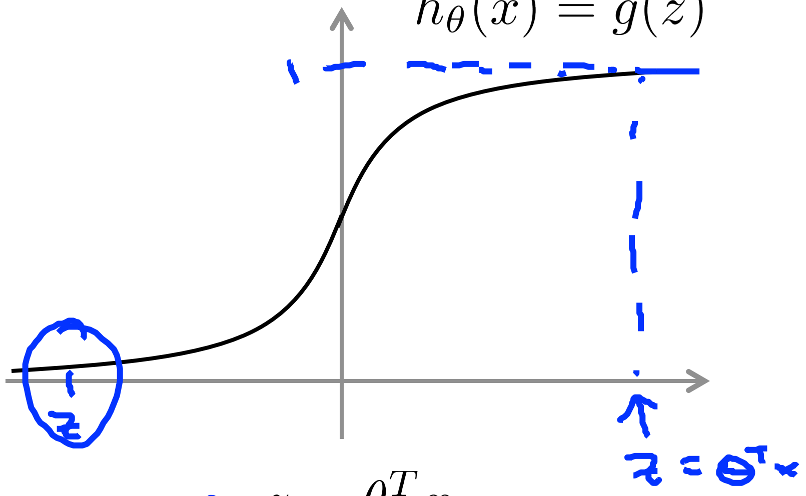
Support Vector Machines

Optimization objective

Alternative view of logistic regression

sigmoid

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



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

$$z = \theta^T x$$

muito maior que ZERO

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

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

Alternative view of logistic regression

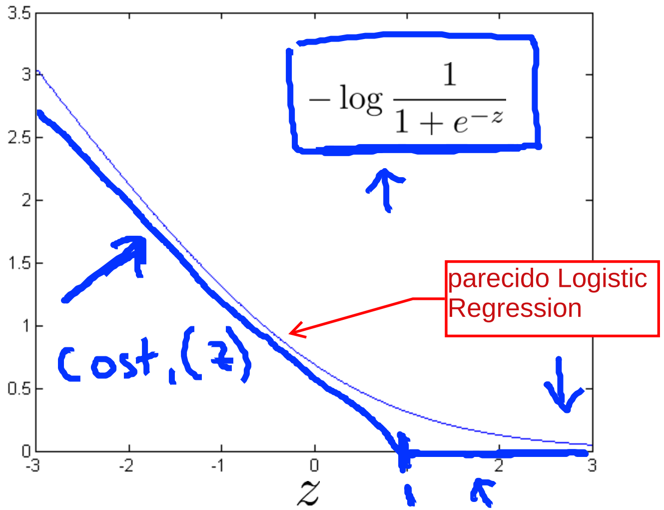
(x, y)

Cost of example: $-(y\log h_\theta(x) + (1-y)\log(1-h_\theta(x)))$

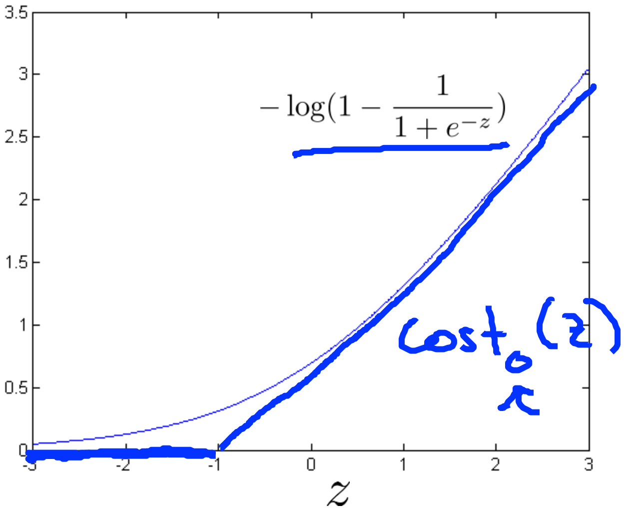
$$= \boxed{-y\log\frac{1}{1+e^{-\theta^T x}}} \boxed{-(1-y)\log\left(1-\frac{1}{1+e^{-\theta^T x}}\right)}$$

If $y = 1$ (want $\theta^T x \gg 0$):

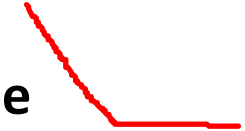
$z = \theta^T x$

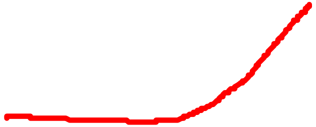


If $y = 0$ (want $\theta^T x \ll 0$):



Support vector machine





A

Logistic regression:

B

$$\overbrace{\min_{\theta}\frac{1}{m}\left[\sum_{i=1}^{m} y^{(i)}\left(\underbrace{-\log h_\theta(x^{(i)})}_{\text{cost1}}\right) + \underbrace{(1-y^{(i)})\left((-\log(1-h_\theta(x^{(i)}))\right)}_{\text{cost0}}\right]}^{A} + \overbrace{\frac{\lambda}{2m}\sum_{j=1}^{n}\theta_j^2}^{B}$$

Support vector machine:

$$\min \cancel{\frac{1}{m}} \sum_{i=1}^{m} y^{(i)} cost_1(\theta^T x^{(i)}) + (1-y^{i}) cost_0(\theta^T x^{i}) + \frac{\lambda}{2m}\sum_{j=0}^{m}\theta_j^2$$

$\min\ (u-5)^2+1 \to u=5$

$A + \lambda B \quad \Big\}\ C = \frac{1}{\lambda}$

$CA + B$

$$\min_{\theta} C\sum_{i=1}^{m}\left[y^{(i)}cost_1(\theta^T x^{(i)}) + (1-y^{(i)})cost_0(\theta^T x^{(i)})\right] + \frac{1}{2}\sum_{i=1}^{n}\theta_j^2$$

SVM hypothesis

$$\min_{\theta} C \sum_{i=1}^{m} \left[y^{(i)} cost_1(\theta^T x^{(i)}) + (1 - y^{(i)}) cost_0(\theta^T x^{(i)}) \right] + \frac{1}{2} \sum_{i=1}^{n} \theta_j^2$$

Hypothesis:

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