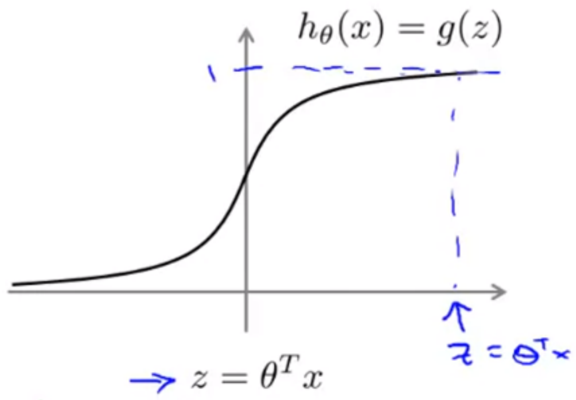
Support Vector Machine

Alternate view of Logistic Regression

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

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

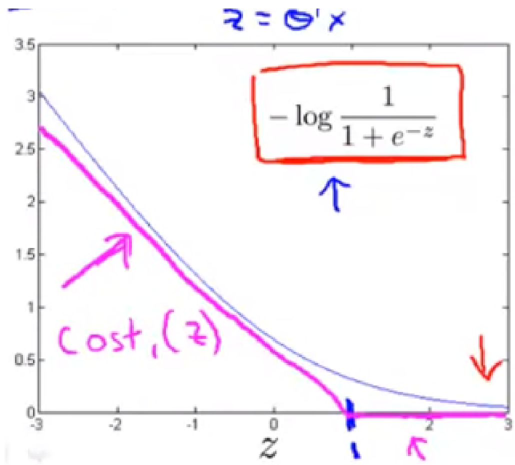
If $y=0$, we want $h_\theta(x) \approx 0$, which is $\theta^T x << 1$



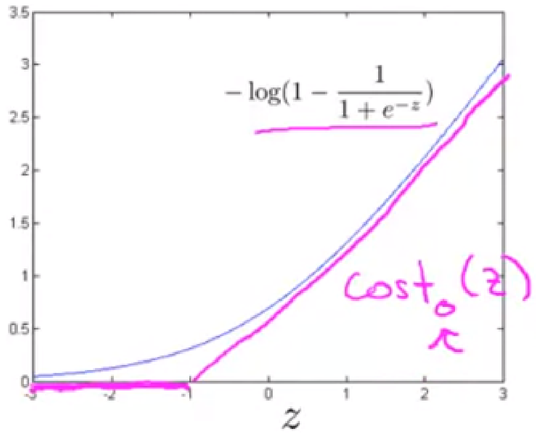
The term which a single training example contributes to the overall logistic regression:

$$\begin{aligned}\text{Cost of Example} &= -(y \log h_\theta(x) + (1-y)\log(1-h_\theta(x))) \\ &= -y \log \frac{1}{1+e^{-\theta^T x}} - (1-y)\log(1-\frac{1}{1+e^{-\theta^T x}})\end{aligned}$$

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



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



We had the cost function for Logistic Regression as:

$$\min_\theta \frac{1}{m}\left[\sum_{i=1}^{m} y^{(i)}\left(-\log_\theta(x^{(i)})\right) + (1-y^{(i)})\left(-log(1-h_\theta(x^{(i)}))\right)\right] + \frac{\lambda}{2m}\sum_{j=1}^{n}\theta_j^2$$

Support Vector Machine

Cost Function:

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

$$\text{Getting rid of } \frac{1}{m} \text{ terms}$$

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

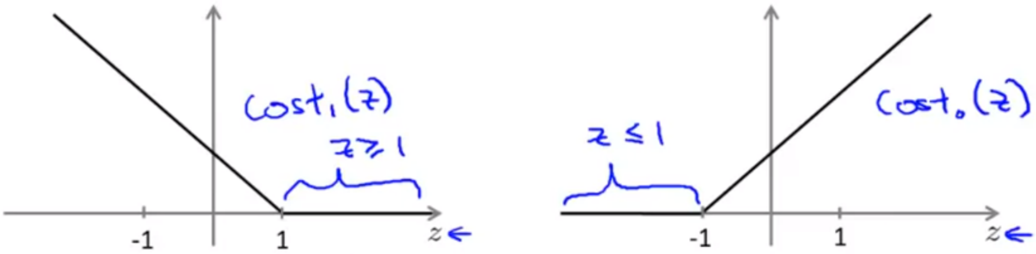
$$\text{Getting rid of } \lambda \text{ and adding new constant } C$$

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

Where $C \approx \frac{1}{\lambda}$

Hypothesis:

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



→ If $y = 1$, we want $\theta^T x \geq 1$ (not just ≥ 0)

→ If $y = 0$, we want $\theta^T x \leq -1$ (not just < 0)

SVM Decision Boundary

- Whenever $y^{(i)} = 1$

  $\theta^T x^{(i)} \geq 1$

- Whenever $y^{(i)} = 0$

  $\theta^T x^{(i)} \leq -1$