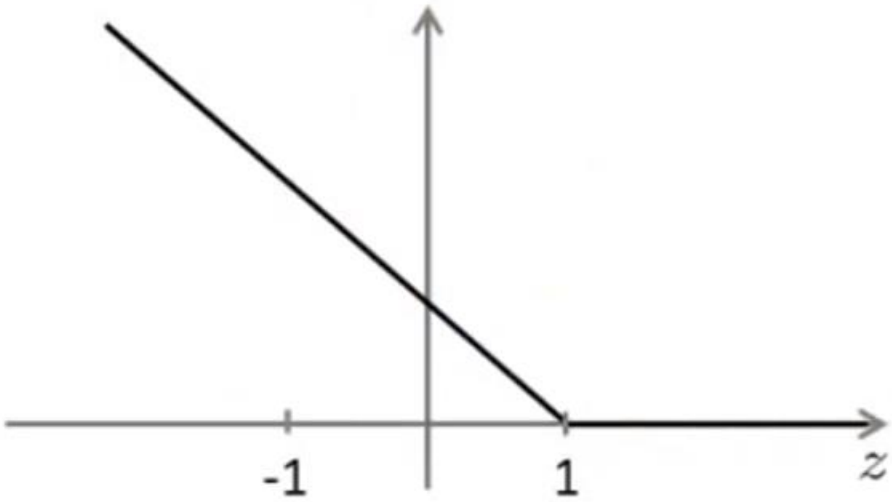
Large Margin Intuition

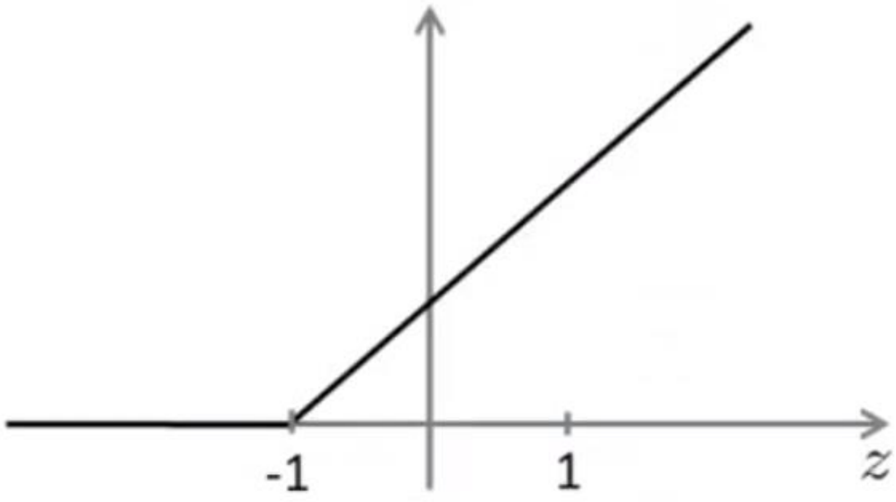
Large Margin Classification

Support Vector Machines

Support Vector Machine

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



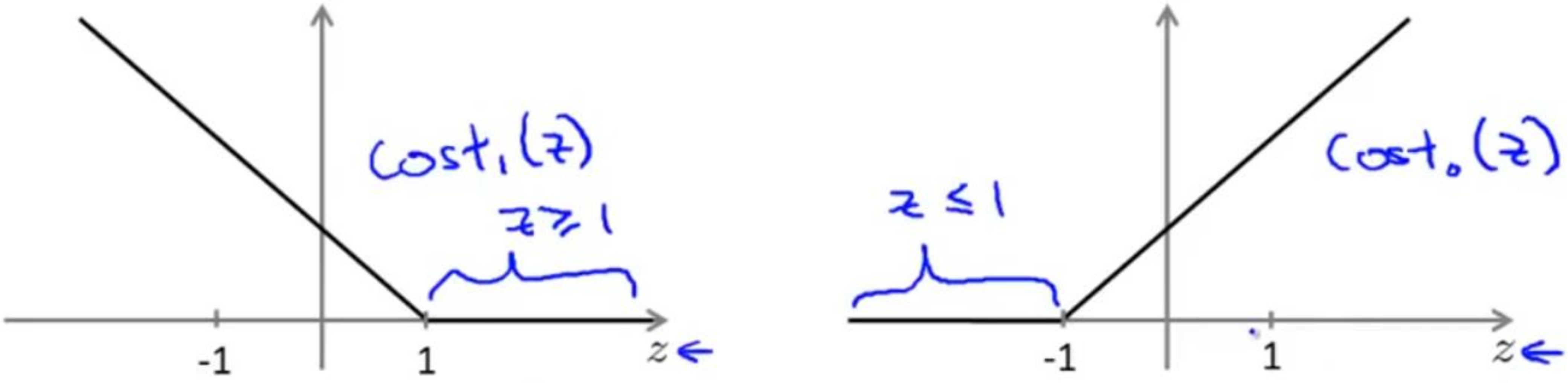


If $y = 1$, we want $\theta^T x \geq 1$ (not just ≥ 0)

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

Support Vector Machine

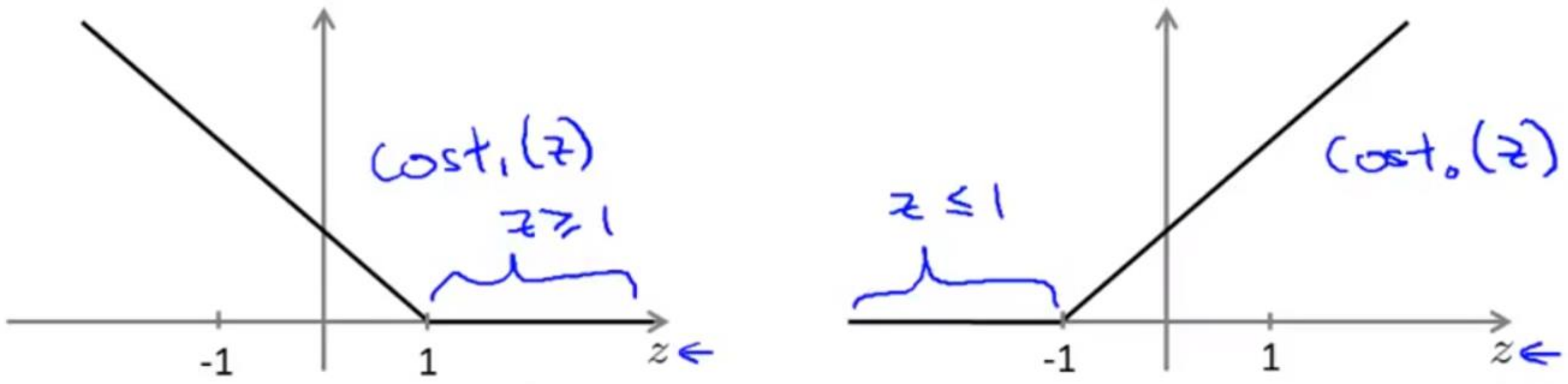
$$\min_\theta C \sum_{i=1}^{m} \left[y^{(i)} \underline{cost_1(\theta^T x^{(i)})} + (1-y^{(i)}) \underline{cost_0(\theta^T x^{(i)})} \right] + \frac{1}{2}\sum_{i=1}^{n} \theta_j^2$$



- If $\underline{y=1}$, we want $\underline{\theta^T x \geq 1}$ (not just ≥ 0)
- If $y=0$, we want $\theta^T x \leq -1$ (not just < 0)

Support Vector Machine

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

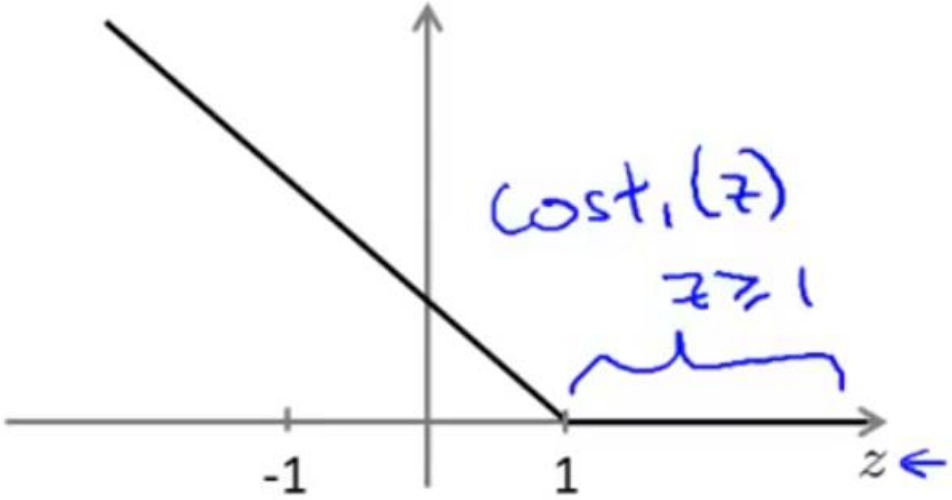


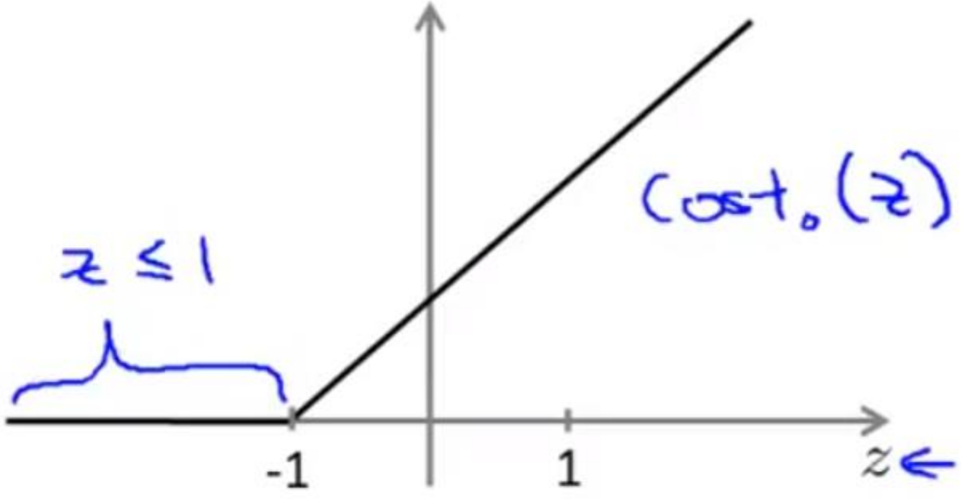
If $y = 1$, we want $\theta^T x \geq 1$ (not just ≥ 0) $\qquad \theta^T x \geq \cancel{0}\ 1$

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

Support Vector Machine

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





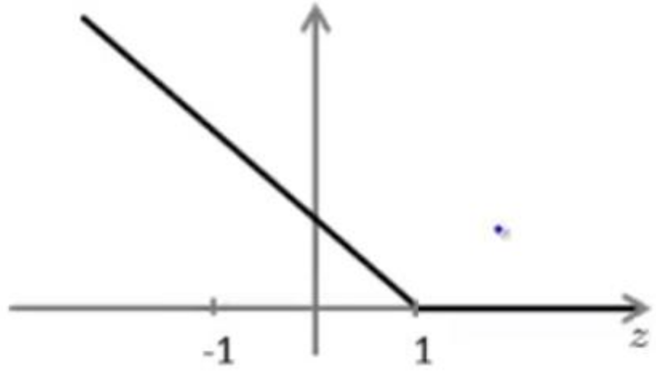
- If $y = 1$, we want $\theta^T x \geq 1$ (not just ≥ 0) $\theta^T x \geq \cancel{0}\ 1$
- If $y = 0$, we want $\theta^T x \leq -1$ (not just < 0) $\theta^T x \leq \cancel{0}\ -1$

$C = 100{,}000$

SVM Decision Boundary

$$\min_\theta C \underbrace{\sum_{i=1}^{m} \left[y^{(i)} cost_1(\theta^T x^{(i)}) + (1 - y^{(i)}) cost_0(\theta^T x^{(i)}) \right]}_{=0} + \frac{1}{2} \sum_{i=1}^{n} \theta_j^2$$

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

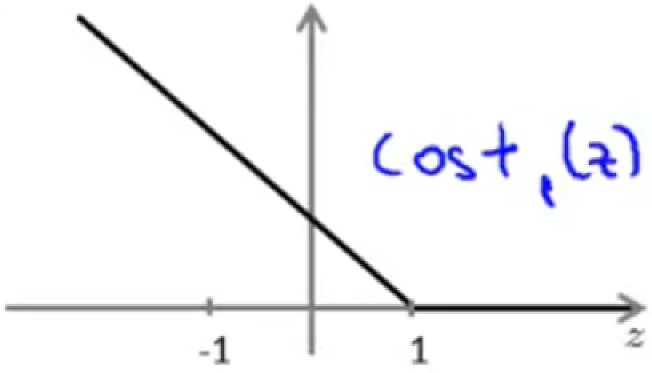


SVM Decision Boundary

$$\min_\theta C \underbrace{\sum_{i=1}^{m}\left[y^{(i)}cost_1(\theta^T x^{(i)}) + (1-y^{(i)})cost_0(\theta^T x^{(i)})\right]}_{=0} + \frac{1}{2}\sum_{i=1}^{n}\theta_j^2$$

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

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



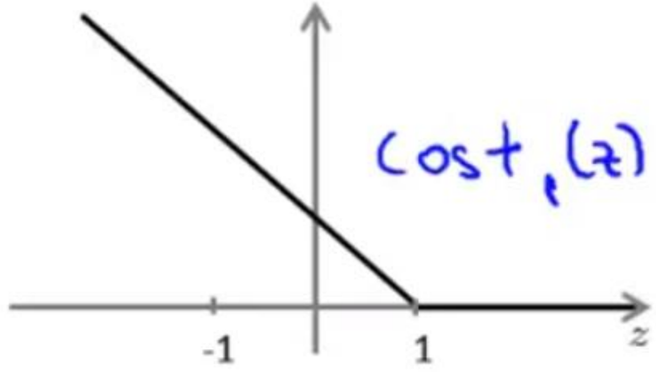
Cost is zero when z>1

SVM Decision Boundary

$$\min_\theta C \underbrace{\sum_{i=1}^{m} \left[y^{(i)} cost_1(\theta^T x^{(i)}) + (1 - y^{(i)}) cost_0(\theta^T x^{(i)}) \right]}_{=0} + \frac{1}{2} \sum_{i=1}^{n} \theta_j^2$$

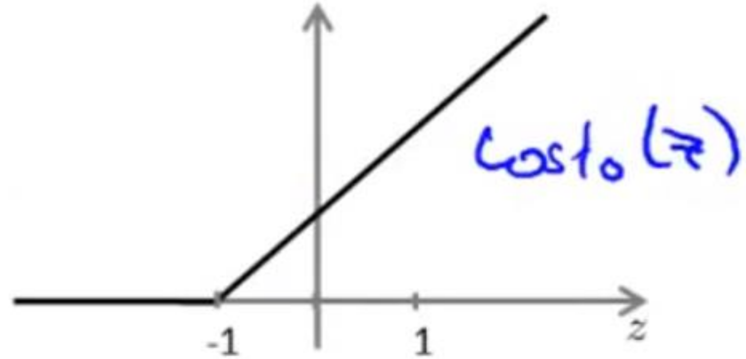
Whenever $y^{(i)} = 1$:

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



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

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



Cost is zero when z<-1

SVM Decision Boundary

$$\min_\theta C \underbrace{\boxed{\sum_{i=1}^{m} \left[y^{(i)} cost_1(\theta^T x^{(i)}) + (1 - y^{(i)}) cost_0(\theta^T x^{(i)}) \right]}}_{=0} + \frac{1}{2} \sum_{i=1}^{n} \theta_j^2$$

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

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

$$\min_\theta \cancel{C \times 0} + \frac{1}{2} \sum_{i=1}^{n} \theta_j^2$$

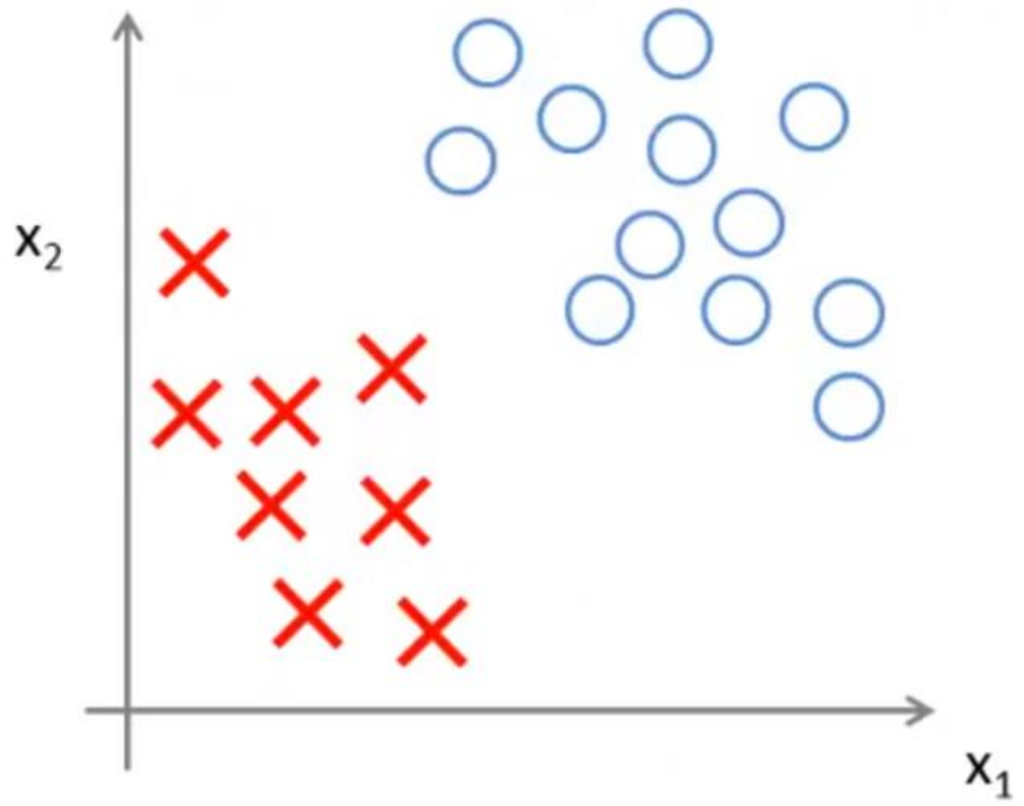
s.t. $\theta^T x^{(i)} \geq 1$ if $y^{(i)} = 1$

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

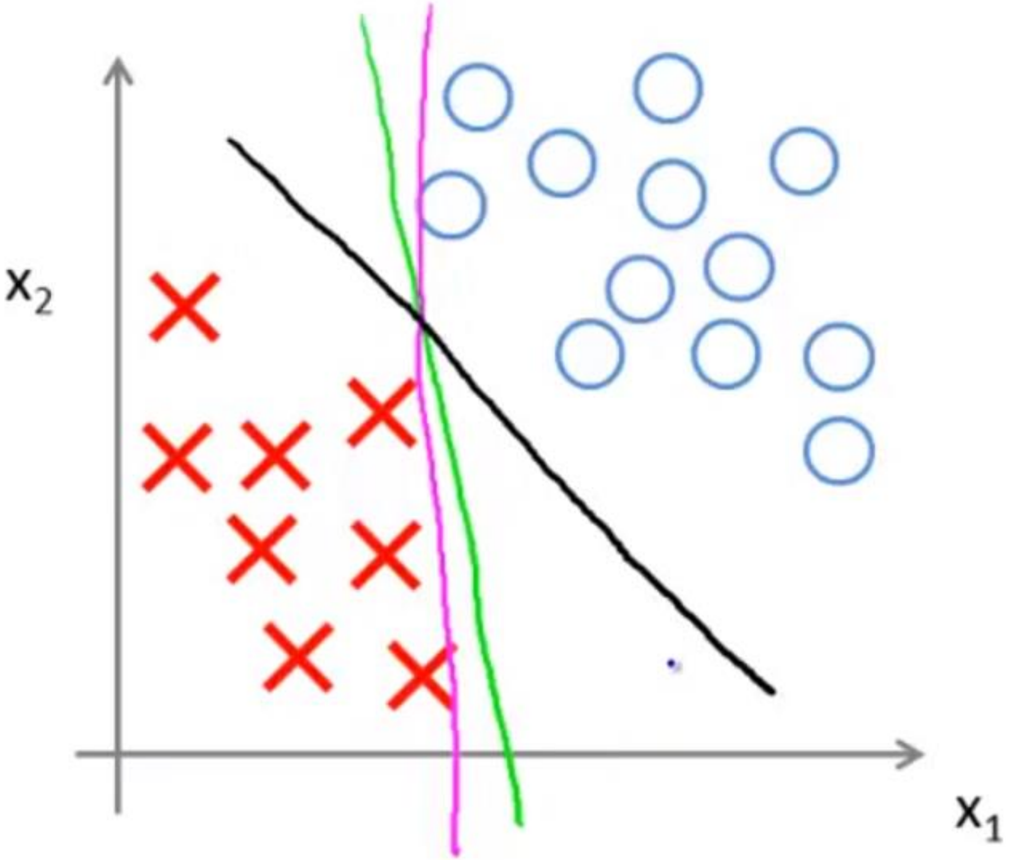
Whenever $y^{(i)} = 0$:

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

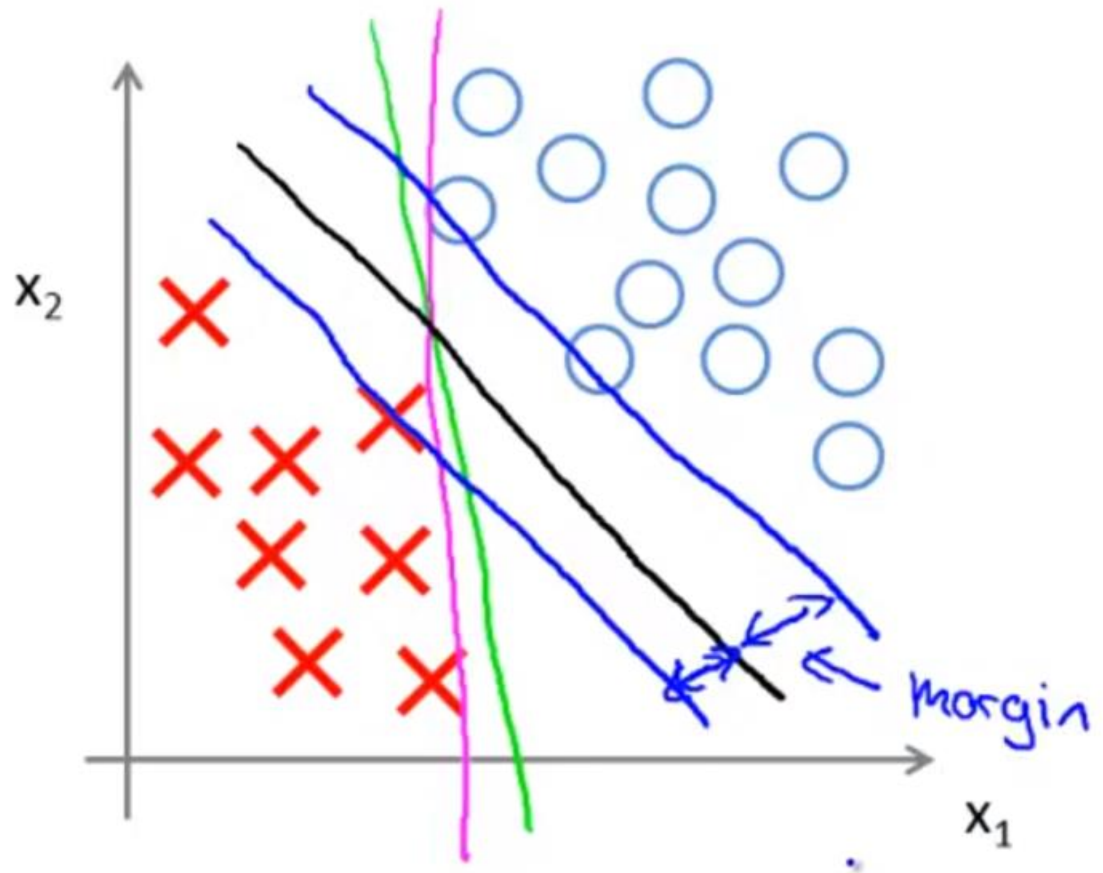
SVM Decision Boundary: Linearly separable case



SVM Decision Boundary: Linearly separable case

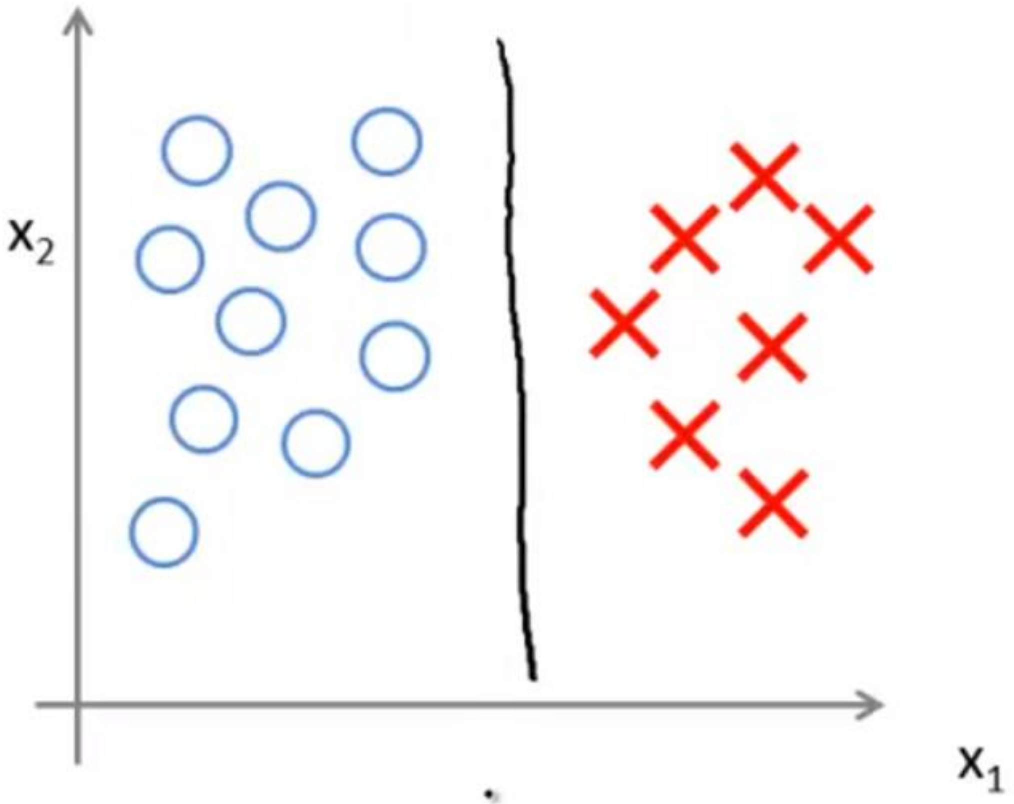


SVM Decision Boundary: Linearly separable case



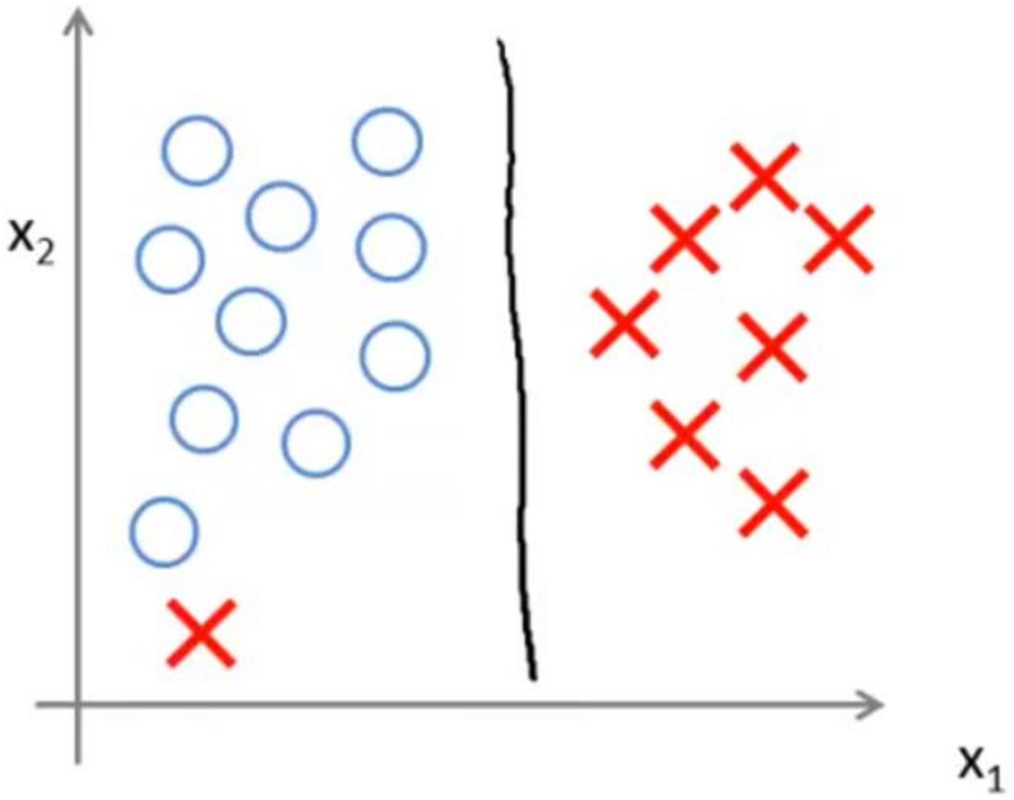
Large Margin Classifier

Large margin classifier in presence of outliers



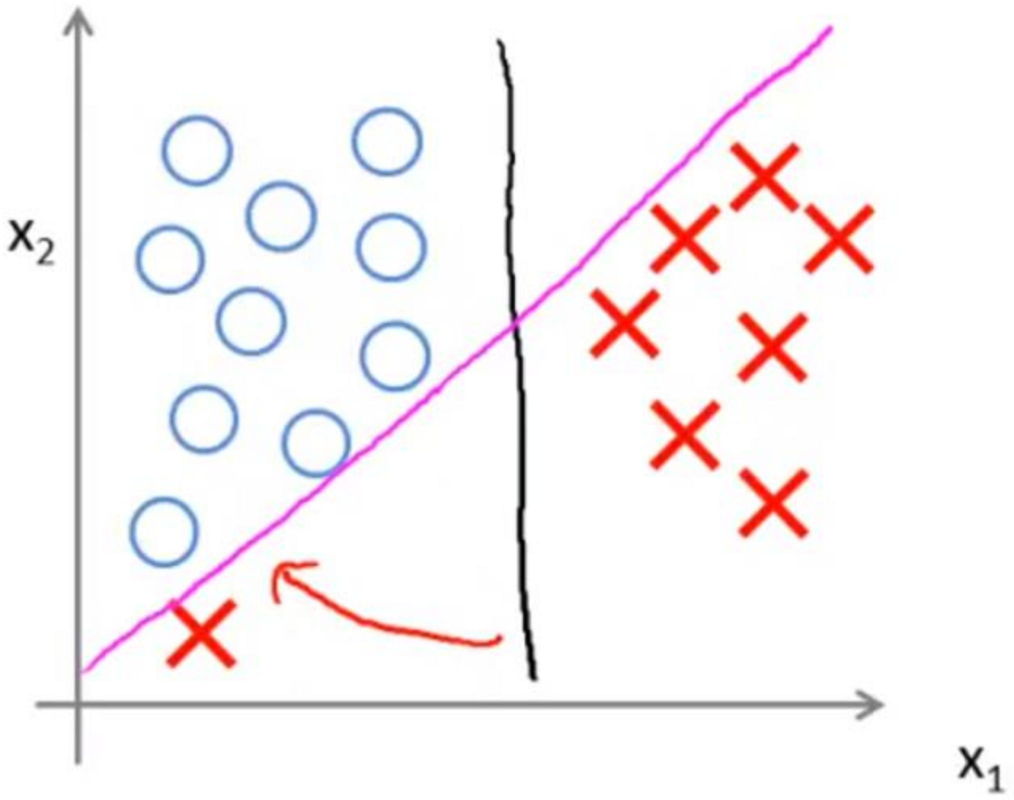
C very large

Large margin classifier in presence of outliers



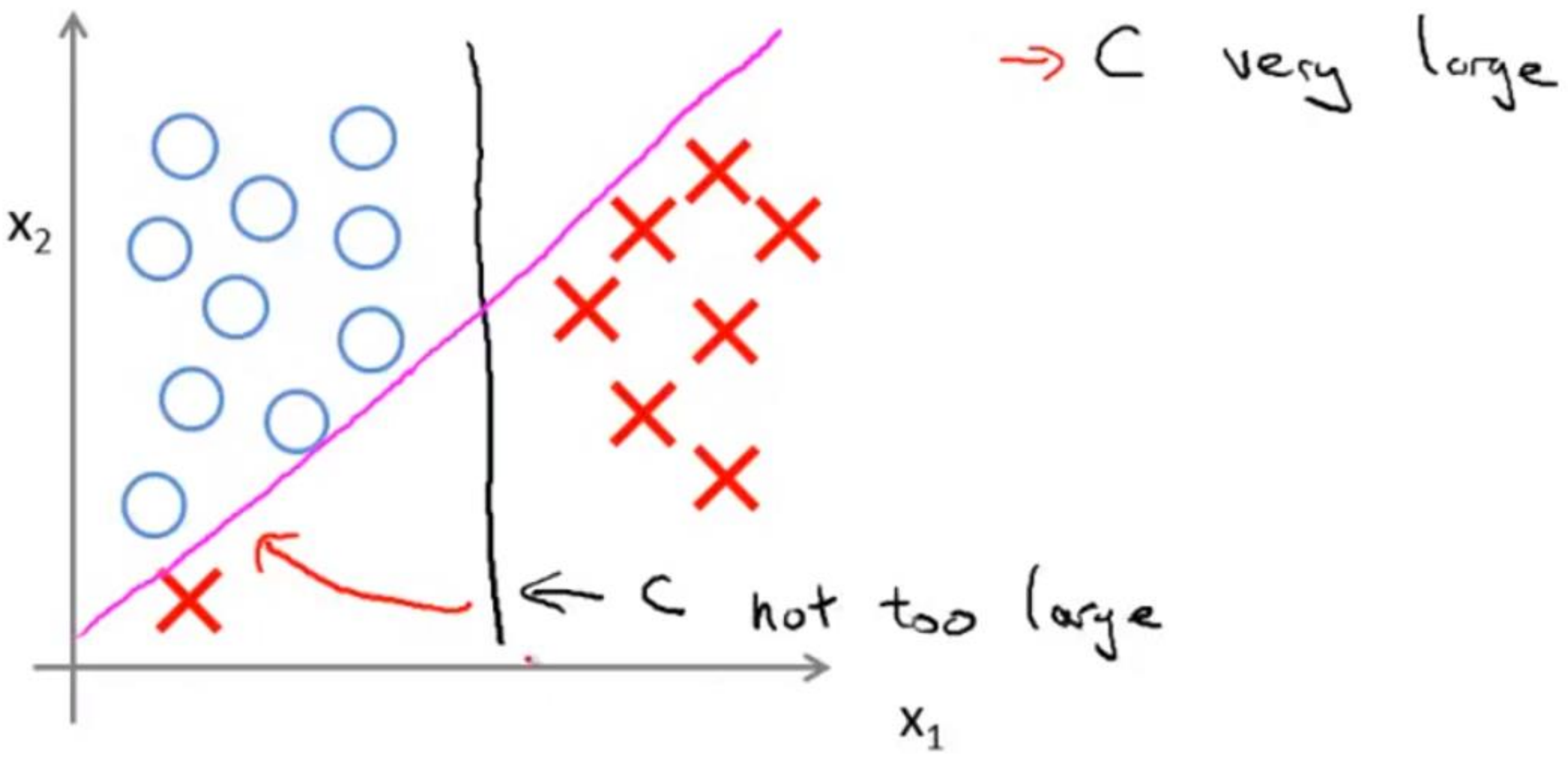
C very large

Large margin classifier in presence of outliers



C very large

Large margin classifier in presence of outliers



Andrew Ng

Large margin classifier in presence of outliers