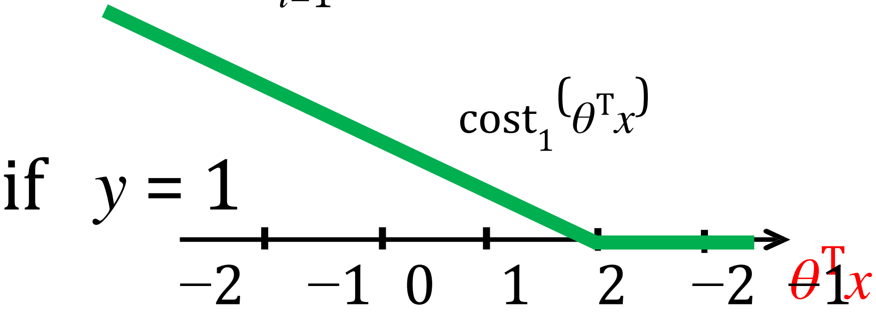
Large Margin Classifier

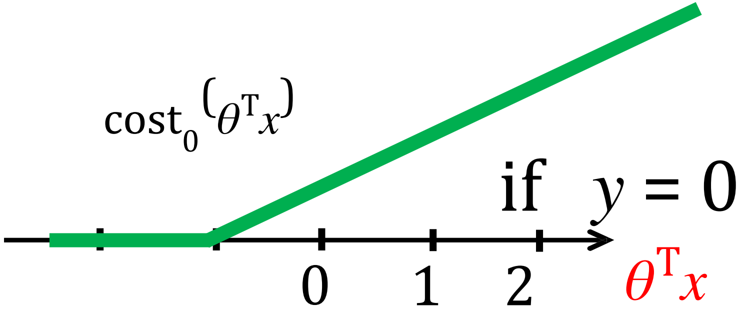
Name : Shashank Patel

Roll no. :220103027

Support vector machine

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





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)

SVM decision boundary

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

- Let's say we have a very large C...

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

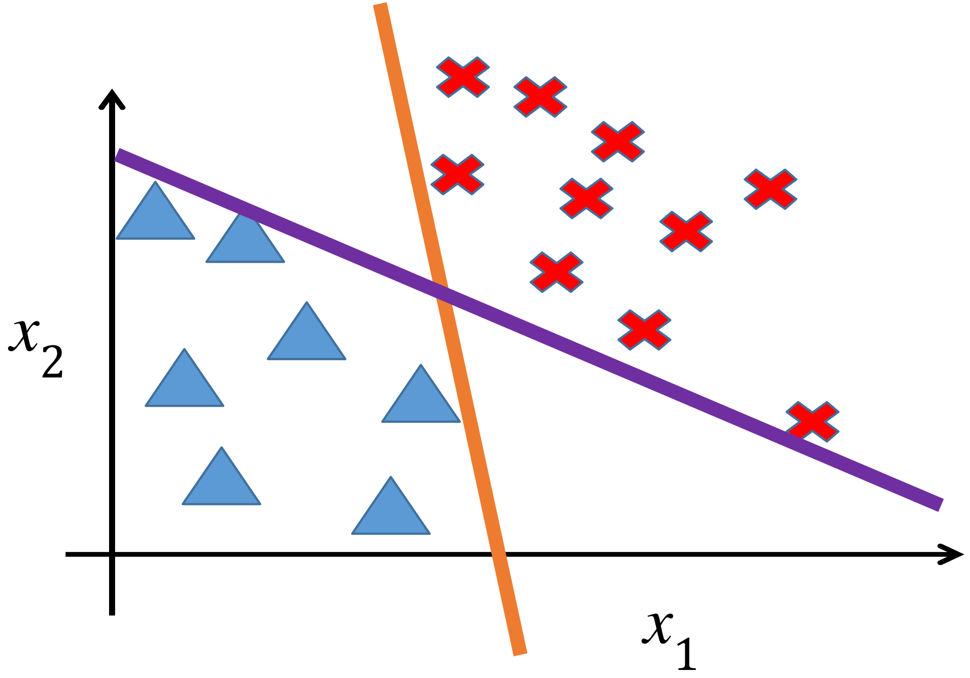
  $\theta^T x^{(i)} \geq 1$

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

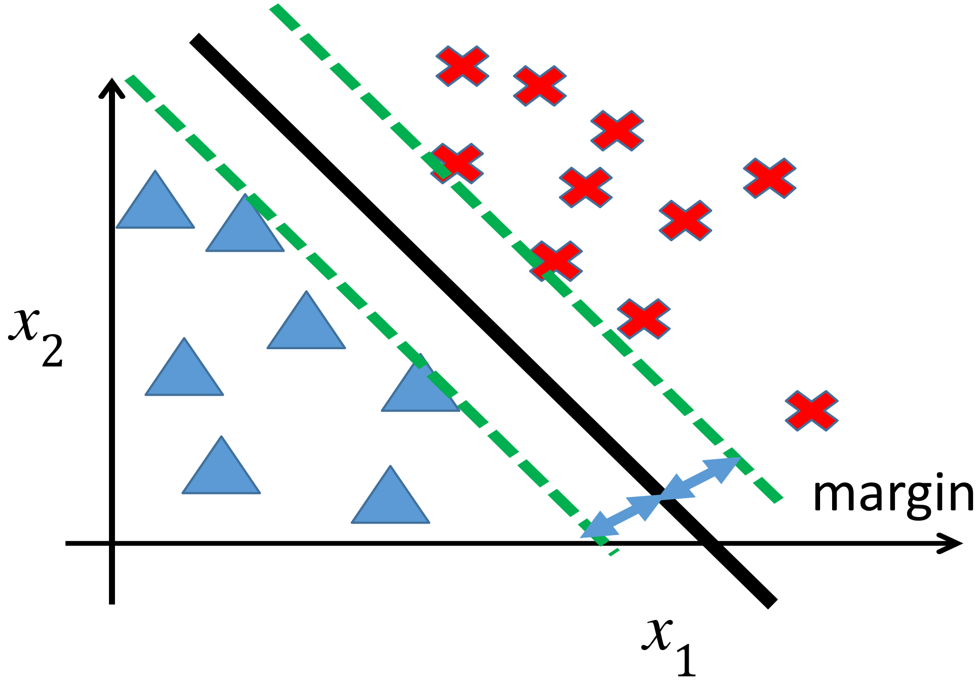
  $\theta^T x^{(i)} \leq -1$

$$\boxed{\begin{aligned} &\min_\theta \frac{1}{2}\sum_{j=1}^{n}\theta_j^2 \\ &\text{s.t.}\quad \theta^T x^{(i)} \geq 1 \quad \text{if } y^{(i)} = 1 \\ &\qquad\;\; \theta^T x^{(i)} \leq -1 \quad \text{if } y^{(i)} = 0 \end{aligned}}$$

SVM decision boundary: Linearly separable case



SVM decision boundary: Linearly separable case

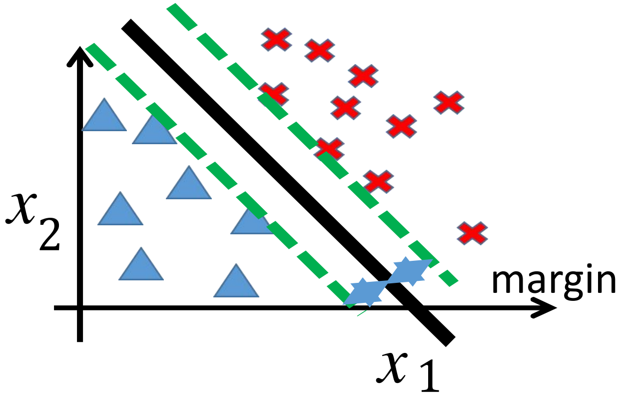


Why large margin classifiers?

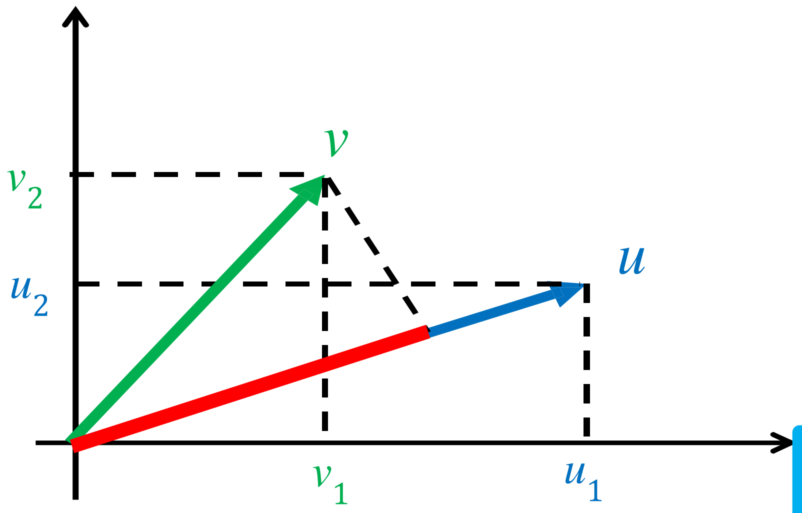
$$\min_{\theta} \frac{1}{2}\sum_{j=1}^{n} \theta_j^{\ 2}$$

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

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



Vector inner product



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

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

$= \sqrt{u_1^2 + u_2^2} \in \mathbb{R}$

$p =$ length of projection of v onto u

$$\boxed{\begin{aligned} u^{\mathrm{T}} v &= p \cdot \|u\| \\ &= \mathrm{u}_1 \mathrm{v}_1 + \mathrm{u}_2 \mathrm{v}_2 \end{aligned}}$$

SVM decision boundary

$$\min_{\theta} \frac{1}{2} \sum_{j=1}^{n} \theta_j^2$$

$$\frac{1}{2}\sum_{j=1}^{n} \theta_j^2 = \frac{1}{2}\left(\theta_1^2 + \theta_2^2\right) = \frac{1}{2}\left(\sqrt{\theta_1^2 + \theta_2^2}\right)^2 = \frac{1}{2}\|\theta\|^2$$

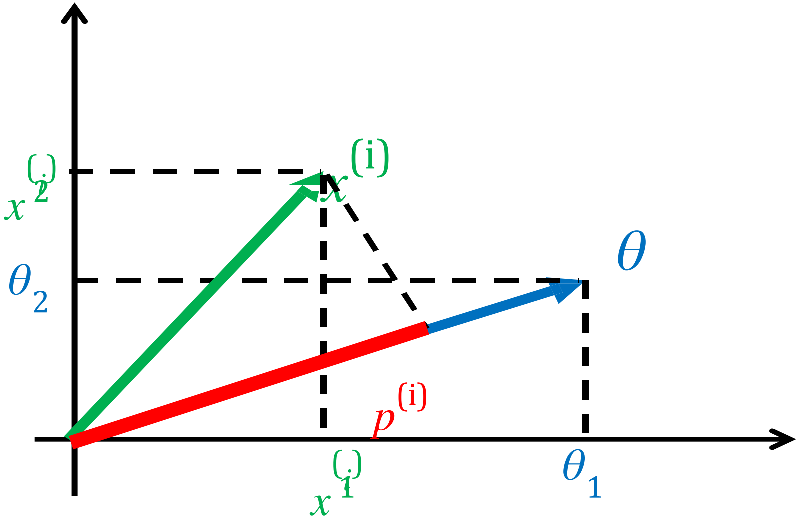
s. t. $\theta^{\mathrm{T}} x^{(i)} \geq 1$ if $y^{(i)} = 1$

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

Simplication: $\theta_0 = 0, n = 2$

What's $\theta^{\mathrm{T}} x^{(i)}$?

$$\theta^{\mathrm{T}} x^{(i)} = p^{(i)} \|\theta\|^2$$



SVM decision boundary

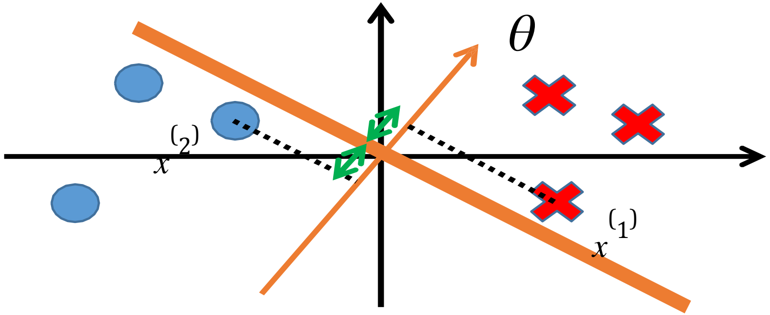
$$\min_{\theta} \frac{1}{2}\|\theta\|^2$$

$$\text{s.t. } p^{(i)}\|\theta\|_2 \geq 1 \quad \text{if} \quad y^{(i)} = 1$$

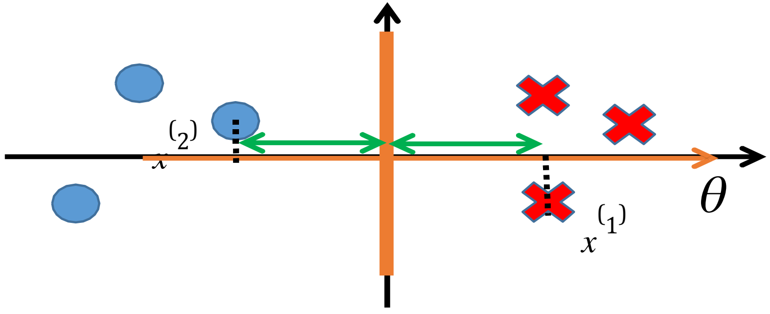
$$p^{(i)}\|\theta\|_2 \leq -1 \quad \text{if} \quad y^{(i)} = 0$$

Simplication: $\theta_0 = 0, n = 2$

$\boxed{p^{(1)}, p^{(2)} \text{ small} \rightarrow \|\theta\|_2 \text{ large}}$



$\boxed{p^{(1)}, p^{(2)} \text{ large} \rightarrow \|\theta\|_2 \text{ can be small}}$



Thank You

?