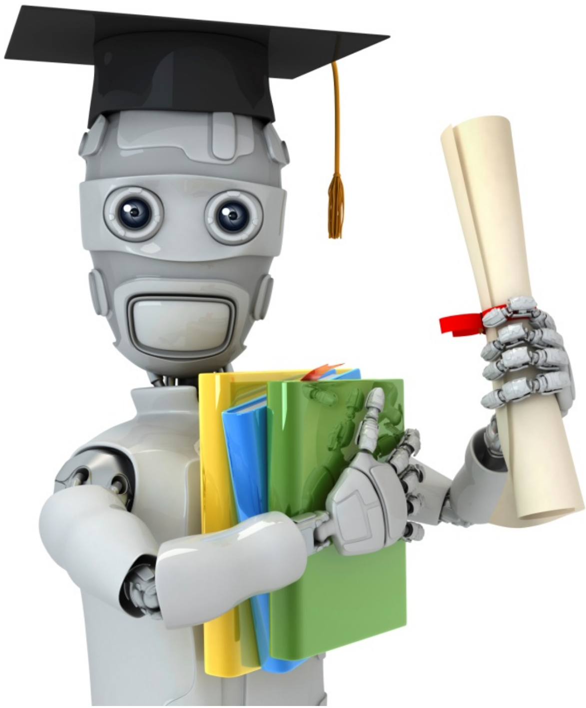
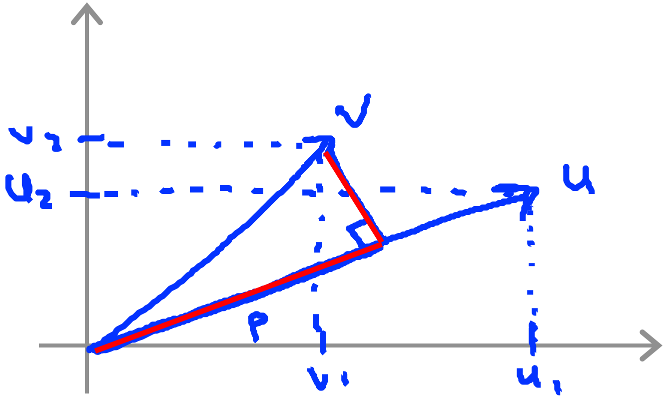


Machine Learning

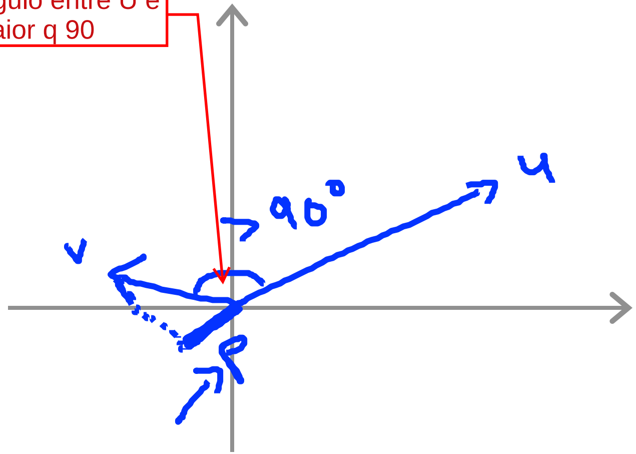
Support Vector Machines

The mathematics behind large margin classification (optional)

Vector Inner Product



se o angulo entre U e V for maior q 90



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

$u^T v = ?$ $\quad [u_1 \ \ u_2] \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$

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

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

p = length of projection of v onto u.

Signed

$u^T v = \underline{p} \cdot \underline{\|u\|} \leftarrow \quad = v^T u$

$= u_1 v_1 + u_2 v_2 \leftarrow \quad p \in \mathbb{R}$

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

$p < 0$

se o angulo for menor entao p épositivo

Andrew Ng

SVM Decision Boundary

$\omega = (\sqrt{\omega})^2$

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

$$\sqrt{\theta_1^2 + \theta_2^2} = \|\theta\|$$

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

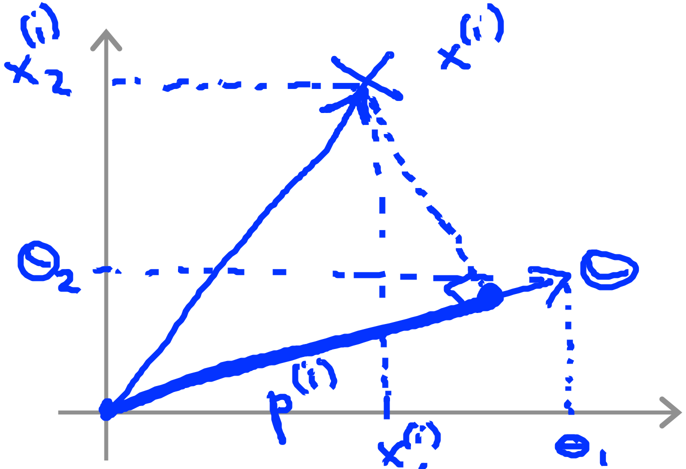
$\theta^T x^{(i)} \leq -1$ if $y^{(i)} = 0$

Simplification: $\theta_0 = 0$ $n=2$

$\begin{bmatrix} \cancel{\theta_0} \\ \theta_1 \\ \theta_2 \end{bmatrix}$ $\theta_0 = 0$

$\theta^T x^{(i)} = ?$

$u^T v$



$\theta^T x^{(i)} = \boxed{p^{(i)} \cdot \|\theta\|}$

$= \theta_1 x_1^{(i)} + \theta_2 x_2^{(i)}$

Andrew Ng

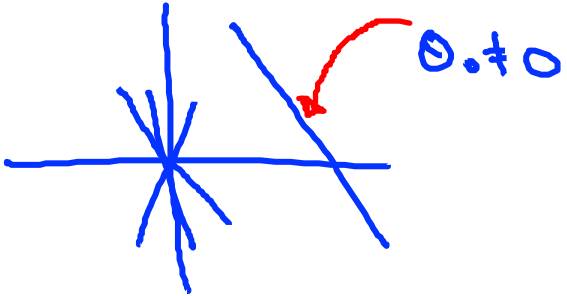
SVM Decision Boundary

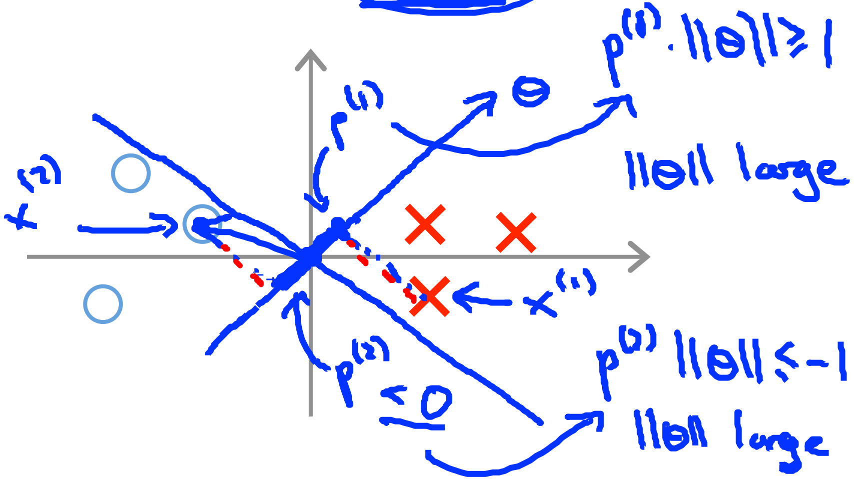
$$\min_\theta \frac{1}{2}\sum_{j=1}^{n}\theta_j^2 = \frac{1}{2}\|\theta\|^2$$

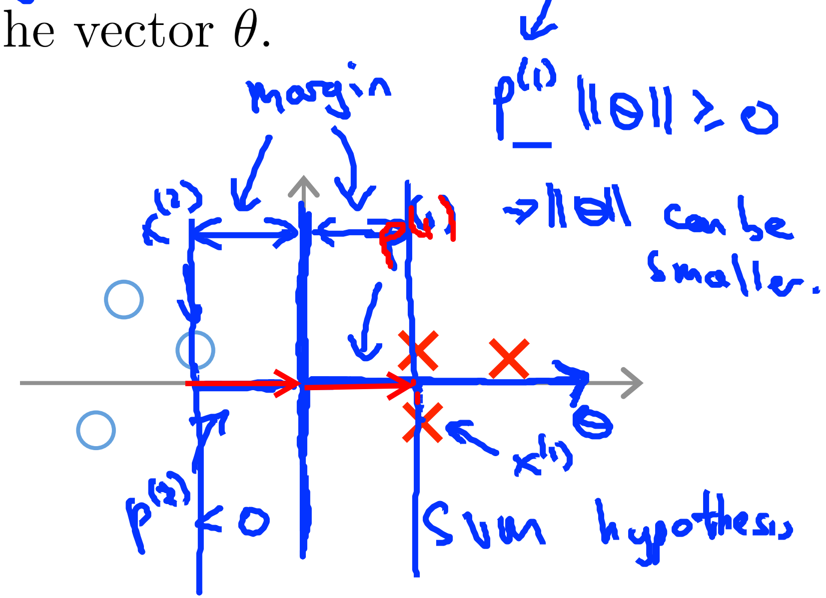
$$\left.\begin{array}{ll} \text{s.t.} \quad \boxed{p^{(i)}\cdot\|\theta\| \geq 1} & \text{if } y^{(i)} = 1 \\ \quad\quad\; p^{(i)}\cdot\|\theta\| \leq -1 & \text{if } y^{(i)} = 1 \end{array}\right\} \text{C very large}$$

where $p^{(i)}$ is the projection of $x^{(i)}$ onto the vector θ.

Simplification: $\theta_0 = 0$







Andrew Ng