

Feb 14 2020

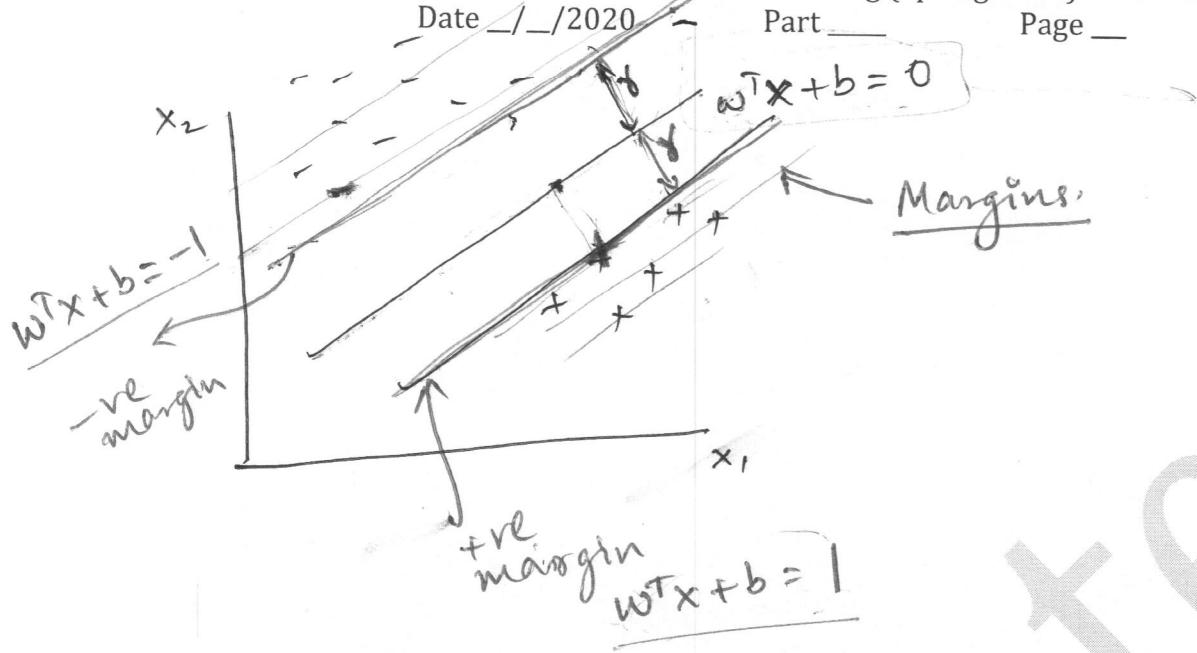
$\mathbf{w} \rightarrow \text{vector}$

$$[w_1 | w_2 | \dots | w_d]$$

$$\|\mathbf{w}\| = |w_1| + |w_2| + \dots + |w_d| \rightarrow l1 \text{ norm}$$

$$\|\mathbf{w}\| = \sqrt{w_1^2 + w_2^2 + \dots + w_d^2} \rightarrow l2 \text{ norm}$$

$$\|\mathbf{w}\|_p = (w_1^p + \dots + w_d^p)^{\frac{1}{p}} \rightarrow l_p \text{ norm}$$

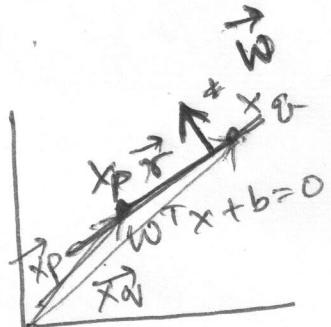


$$\mathbf{w} \rightarrow \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \quad \| \mathbf{w} \| = \left(\sum_{i=1}^2 w_i^2 \right)^{1/2} \quad \| \mathbf{w} \| = \sqrt{w_1^2 + w_2^2}$$

$$\mathbf{w}^T \mathbf{w} = \| \mathbf{w} \|^2$$

$$\gamma = \frac{1}{\| \mathbf{w} \|}$$

$$\text{Size of margin} = 2\gamma = \frac{2}{\| \mathbf{w} \|}$$



$$\begin{aligned} \vec{x}_q &= \vec{x}_p + \vec{\delta} \\ \vec{w}^T \vec{\delta} &= (\vec{x}_q - \vec{x}_p) \\ &= \vec{w}^T (\vec{x}_q - \vec{x}_p) \\ &= \vec{w}^T \vec{x}_q - \vec{w}^T \vec{x}_p \\ &= (\vec{w}^T \vec{x}_q + b) - (\vec{w}^T \vec{x}_p + b) \end{aligned}$$

(Implies that \vec{w} is \perp to the line)

Size of the margin is $\frac{2}{\|w\|}$

$$\vec{\gamma} = \gamma \frac{\vec{w}}{\|\vec{w}\|}$$

$$\begin{aligned} \vec{x} &= \vec{x}' + \vec{\gamma} \\ &= \vec{x}' + \gamma \frac{\vec{w}}{\|\vec{w}\|} \end{aligned}$$

$$\vec{w}^T \vec{x} + b = 1$$

$$\vec{w}^T \vec{x}' + \frac{\vec{w}^T (\gamma \vec{w})}{\|\vec{w}\|} + b = 1$$

$$\vec{w}^T \vec{x}' + b = 0$$

$$\gamma \frac{\vec{w}^T \vec{w}}{\|\vec{w}\|} = 1$$

$$\gamma \|\vec{w}\| = 1$$

$$\gamma = \frac{1}{\|\vec{w}\|} \quad \leftarrow \text{Proof.}$$

$$\text{Size of Margin} = \frac{2}{\|\vec{w}\|}$$

SRM

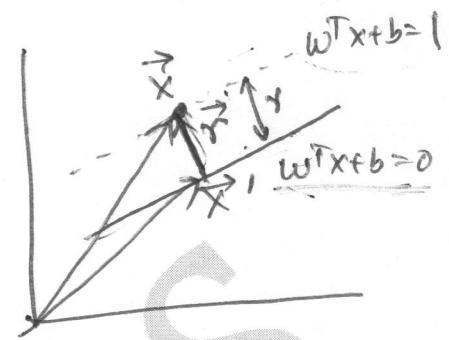
formulation:

$$\min_{w, b}$$

$$\frac{\|\vec{w}\|^2}{2}$$

s.t.

$$y_i (\vec{w}^T \vec{x}_i + b) \geq 1 \quad \forall i=1, \dots, N$$



$$f(w_1, w_2) = w_1^2 + 3w_1^3 + 2w_1 + 7$$

$$w_1 + w_2 \leq 1 \rightarrow w_1 + w_2 - 1 \leq 0$$

$$L(w, \alpha) = (w_1^2 + 3w_1^3 + 2w_1 + 7) + \alpha(w_1 + w_2 - 1)$$

$$\text{S.t } \underline{\alpha} \geq 0$$

Let $\delta = \log \alpha$ $\alpha = \exp(\delta)$
 "log-barrier"

$$L_p(w, b, \alpha) = \frac{\|w\|^2}{2} + \sum_{i=1}^N \alpha_i [1 - y_i (w^T x_i + b)]$$

We will first minimize this w.r.t w & b .

$$\frac{\partial L_p}{\partial w} = w + \sum_{i=1}^N (-x_i y_i x_i) \quad \frac{\partial \|w\|^2}{\partial w} = 2w$$

Setting to 0

$$\frac{\partial L_p}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^N \alpha_i y_i x_i$$

$$\frac{\partial}{\partial w} w^T w = 2w$$

$$\frac{\partial L_p}{\partial b} = - \sum_{i=1}^N \alpha_i y_i = 0$$

$$\sum_{i=1}^N \alpha_i y_i = 0$$

$$w = \sum_{i=1}^N \alpha_i y_i x_i$$

$$= \alpha_1 y_1 x_1 + \alpha_2 y_2 x_2 + \dots + \alpha_N y_N x_N$$

$$w = \sum_{i \text{ is a.s.v}} \alpha_i y_i x_i$$

for a new test point, x^*

$$\underline{w^T x^* + b}$$

$$\left(\sum_{x_i \text{ is a.s.v}} \alpha_i y_i x_i^T x^* \right) + b$$