Approximate SVM

labeled -1, 1

linearly separable

S: a set of pts in $\mathbb{R}^d$

γ^*: maximum margin of all separation planes

Goal: Find a separation plane with margin $\geq \gamma^*/4$

Margin Perceptron

We don't know

γ_{guess}: a value $\leq \gamma^*$

Returns a separation plane with margin $\geq \frac{\gamma_{guess}}{2}$ ← Goal

d-dim. describes a linear classifier

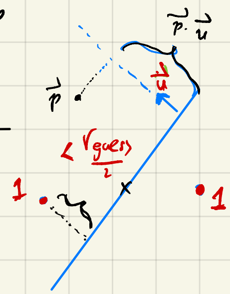
$h_{\vec w}(x) = \begin{cases} 1 & \text{if } \vec w \cdot \vec x \geq 0 \\ -1 & \text{otherwise} \end{cases}$

$\vec w = \vec 0$

while $\exists$ violation pt $\vec p \in S$

- if label($\vec p$) = 1 $\Rightarrow \vec w \leftarrow \vec w + \vec p$
- else $\vec w \leftarrow \vec w - \vec p$

return $\vec w$



Violation:

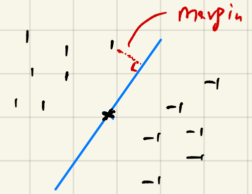
$\vec p$: 1

$\vec p \cdot \frac{\vec w}{|\vec w|} < \frac{\gamma_{guess}}{2}$

a unit vector of the current plane $\vec w \cdot \vec x = 0$

distance from p to the plane

signed



Thm: If $\gamma_{guess} \leq \gamma^*$, then Margin Perceptron finishes after $12 \frac{R^2}{\gamma^{*2}}$ corrections

maximum distance of the pts in S to the origin.

Corollary: If margin Perceptron has already made $1 + 12\frac{R^2}{\gamma_{guess}^2}$ corrections then we must have: $\gamma_{guess} > \gamma^*$.

Proof: By contradiction. Suppose $\gamma_{guess} \leq \gamma^*$. By the thm, margin Perceptron should perform at most $12\frac{R^2}{\gamma^{*2}} \leq 12\frac{R^2}{\gamma_{guess}^2}$

$\therefore$ Contradiction. □

A Progressive Strategy

Start with a very large γ_{guess}

- if normal termination $\Rightarrow$ happy!
- otherwise $\Rightarrow$ Reduce γ_{guess}

In the beginning, set $\gamma_{guess} \leftarrow R$

ROUND: Run Margin Perceptron with γ_{guess}

- if normal termination then return the $\vec w$ found
- otherwise $\gamma_{guess} \leftarrow \gamma_{guess}/2$

geometric series

$n + \frac{n}{2} + \frac{n}{4} + \frac{n}{8} + \cdots \leq 2n$

$n + cn + c^2 n + c^3 n + \cdots = \frac{n}{1-c} = O(n)$, $c<1$

		# corrections
Round 1	$\gamma_{guess} = R$	$1 + 12\frac{R^2}{R^2} = 13 = O(1)$
Round 2	$\gamma_{guess} = R/2$	$1 + 12\frac{R^2}{R^2/4} = 49 = 4 \cdot O(1)$
Round 3	$\gamma_{guess} = R/4$	$16 \cdot O(1)$

$O(1)\cdot(1 + 4 + 4^2 + 4^3 + \cdots + 4^h)$

$= O(1) \cdot O(4^h)$ → # corrections in the last round

$\frac{R^2}{\gamma_{guess}^2} < \frac{4R^2}{\gamma^{*2}} = O(\frac{R^2}{\gamma^{*2}})$

$\gamma_{guess} > \frac{\gamma^*}{2}$

Lmm: $\gamma^* \leq R$

unit normal

Proof: $\vec u$ opt sep plane

Consider any pt $\vec p \in S$

distance from $\vec p$ to the plane

$= \vec u \cdot \vec p = |\vec u| \cdot |\vec p| \cos\theta$

$\leq |\vec p| \leq R$. □

Lmm: At the end of the progressive strategy, we must have $\gamma_{guess} > \gamma^*/2$.

Proof: By contradiction. Set d = the final γ_{guess}. Suppose our final $d \leq \gamma^*/2$ $\Rightarrow$ the γ_{guess} in the prev. round must be exactly $2d \leq \gamma^*$. In that case the strategy should have finished at the prev round. □

Corollary: the final sep. plane obtained has margin $\geq \gamma^*/4$.

Lmm: Our progressive strategy performs in total $O(\frac{R^2}{\gamma^{*2}})$ corrections.

Thm: If $\gamma_{guess} \le \gamma^*$, then Margin Perceptron finishes after $12 \frac{R^2}{\gamma^{*2}}$ corrections

$\vec{w}_0 = \vec{0}$

$\vec{w}_i$ = the $\vec{w}$ vector after i corrections

Claim 1: $|\vec{w}_k| \ge k \cdot \gamma^*$ holds for any k.

Claim 2: $|\vec{w}_k| \le \frac{2R^2}{\gamma^*} + \frac{3k}{4}\gamma^* + R$

holds for any k

If we have claims 1 and 2, then we have the thm because

$$k \cdot \gamma^* \le |\vec{w}_k| \le \frac{2R^2}{\gamma^*} + \frac{3k}{4}\gamma^* + R$$

$$\Rightarrow \frac{k \cdot \gamma^*}{4} \le \frac{2R^2}{\gamma^*} + R$$

$$\Rightarrow k \le \frac{8R^2}{\gamma^{*2}} + 4 \cdot \frac{R}{\gamma^*}$$

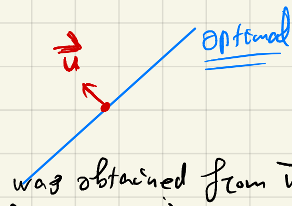
≥ 1

$$\le \frac{8R^2}{\gamma^{*2}} + 4 \cdot \frac{R^2}{\gamma^{*2}}$$

$$\le 12 \frac{R^2}{\gamma^{*2}}. \quad \square$$

Claim 1: $|\vec{w}_k| \ge k \cdot \gamma^*$ holds for any k.

Proof:



$\vec{w}_k$ was obtained from $\vec{w}_{k-1}$ using a violation of $\vec{p}$.

Consider $label(\vec{p}) = 1 \Rightarrow$

$\vec{w}_k = \vec{w}_{k-1} + \vec{p} \Rightarrow$

$$\vec{w}_k \cdot \vec{u} = \vec{w}_{k-1} \cdot \vec{u} + \vec{p} \cdot \vec{u}$$

dist. of $\vec{p}$ to the sep! plane $\ge \gamma^*$

$$\ge \vec{w}_{k-1} \cdot \vec{u} + \gamma^*$$

$$\ge (\vec{w}_{k-2} \cdot \vec{u} + \gamma^*) + \gamma^*$$

$$= \vec{w}_{k-2} \cdot \vec{u} + 2\gamma^*$$

$$\ge \dots$$

$$\ge \vec{w}_0 \cdot \vec{u} + k \cdot \gamma^* = k \cdot \gamma^*$$

$$|\vec{w}_k| \ge \underbrace{|\vec{w}_k| \cdot |\vec{u}|}_{=1} \cdot \underbrace{\cos\theta}_{\le 1} = \vec{w}_k \cdot \vec{u} \ge k \cdot \gamma^*. \quad \square$$

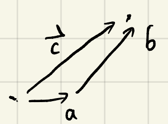
Claim 3: $|\vec{w}_k| \le |\vec{w}_{k-1}| + R$

Proof: Let $\vec{p}$ be the violation pt used to obtain $\vec{w}_k$ from $\vec{w}_{k-1}$

Consider $label(\vec{p}) = 1 \Rightarrow$

$\vec{w}_k = \vec{w}_{k-1} + \vec{p} \Rightarrow$

$|\vec{w}_k| \le |\vec{w}_{k-1}| + \underbrace{|\vec{p}|}_{\le R} \quad \square$



- $\vec{c} = \vec{a} + \vec{b}$
- $|\vec{c}| \le |\vec{a}| + |\vec{b}|$

Claim 4: When $|\vec{w}_{k-1}| \ge \frac{2R^2}{\gamma^*}$,

$$|\vec{w}_k| \le |\vec{w}_{k-1}| + \frac{3}{4}\gamma^*.$$

Proof of Claim 2

$> \frac{2R^2}{\gamma^*}$

$$|\vec{w}_0|, |\vec{w}_1|, |\vec{w}_2|, \dots, |\vec{w}_j| \; | \; |\vec{w}_{j+1}| \; |\vec{w}_{j+2}| \; |\vec{w}_{j+3}| \quad \dots \quad |\vec{w}_k|$$

$|\vec{w}_j|$: $\le \frac{2R^2}{\gamma^*}$; $|\vec{w}_{j+1}|$: $\le R$; $|\vec{w}_{j+2}|$, $|\vec{w}_{j+3}|$, …, $|\vec{w}_k|$: $\le \frac{3}{4}\gamma^*$ (Claim 4)

$$|\vec{w}_k| \le |\vec{w}_j| + R + \left(\frac{3}{4}\gamma^*\right) \cdot (k-j-1)$$

$$\le \frac{2R^2}{\gamma^*} + R + \left(\frac{3}{4}\gamma^*\right) k \quad \square$$

Claim 4: When $|\vec{w}_{k-1}| \geq \frac{2R^2}{\gamma^*}$, $|\vec{w}_k| \leq |\vec{w}_{k-1}| + \frac{3}{4}\gamma^*$.

$\gamma_{guess} \leq \gamma^*$.

$|\vec{w}|^2 = \vec{w} \cdot \vec{w}$

$|\vec{w}| = \sqrt{\sum_{i=1}^{d} \vec{w}[i]^2}$

$(a+b+c)^2 = a^2+b^2+c^2+2ab+2ac+2bc$.

Proof: Consider the violation pt $\vec{p}$ used to obtain $\vec{w}_k$ from $\vec{w}_{k-1}$

$label(\vec{p}) = 1 \Rightarrow \vec{w}_k = \vec{w}_{k-1} + \vec{p}$

$$|\vec{w}_k|^2 = \vec{w}_k \cdot \vec{w}_k = (\vec{w}_{k-1}+\vec{p})\cdot(\vec{w}_{k-1}+\vec{p})$$
$$= \vec{w}_{k-1}\cdot\vec{w}_{k-1} + \vec{p}\cdot\vec{p} + 2\cdot\vec{w}_{k-1}\cdot\vec{p}$$
$$= |\vec{w}_{k-1}|^2 + |\vec{p}|^2 + 2\,\vec{w}_{k-1}\cdot\vec{p}$$
$$\leq |\vec{w}_{k-1}|^2 + R^2 + 2\,\vec{w}_{k-1}\cdot\vec{p} \xrightarrow{\text{violation}} < |\vec{w}_{k-1}|^2 + R^2 + |\vec{w}_{k-1}|\gamma_{guess}$$

$\gamma_{guess} \leq \gamma^*$

$$\frac{\vec{w}_{k-1}\cdot\vec{p}}{|\vec{w}_{k-1}|} < \frac{\gamma_{guess}}{2} \qquad \leq |\vec{w}_{k-1}|^2 + R^2 + |\vec{w}_{k-1}|\gamma^*$$

$$\Rightarrow 2\,\vec{w}_{k-1}\cdot\vec{p} < |\vec{w}_{k-1}|\gamma_{guess}$$

Fact: $|\vec{w}_{k-1}|^2 + R^2 + |\vec{w}_{k-1}|\gamma^* \leq \left(|\vec{w}_{k-1}| + \frac{R^2}{2|\vec{w}_{k-1}|} + \frac{\gamma^*}{2}\right)^2$

Proof: Expand the right hand side

$$RHS = |\vec{w}_{k-1}|^2 + \frac{R^4}{4|\vec{w}_{k-1}|^2} + \frac{\gamma^{*2}}{4} + \cancel{2}\,\vec{w}_{k-1}\frac{R^2}{\cancel{2}|\vec{w}_{k-1}|} + \cancel{2}\,\vec{w}_{k-1}\frac{\gamma^*}{\cancel{2}} + \cancel{2}\frac{R^2}{\cancel{2}|\vec{w}_{k-1}|}\cdot\frac{\gamma^*}{2}$$ ✓ ✓ ✓ □

From fact
$$\leq \left(|\vec{w}_{k-1}| + \frac{R^2}{2|\vec{w}_{k-1}|} + \frac{\gamma^*}{2}\right)^2 \Rightarrow |\vec{w}_k| \leq |\vec{w}_{k-1}| + \frac{R^2}{2|\vec{w}_{k-1}|} + \frac{\gamma^*}{2} \leq |\vec{w}_{k-1}| + \frac{R^2}{2\cdot\frac{2R^2}{\gamma^*}} + \frac{\gamma^*}{2}$$

$|\vec{w}_{k-1}| \geq \frac{2R^2}{\gamma^*}$

$$= |\vec{w}_{k-1}| + \frac{\gamma^*}{4} + \frac{\gamma^*}{2}$$
$$= |\vec{w}_{k-1}| + \frac{3}{4}\gamma^*. \quad \square$$