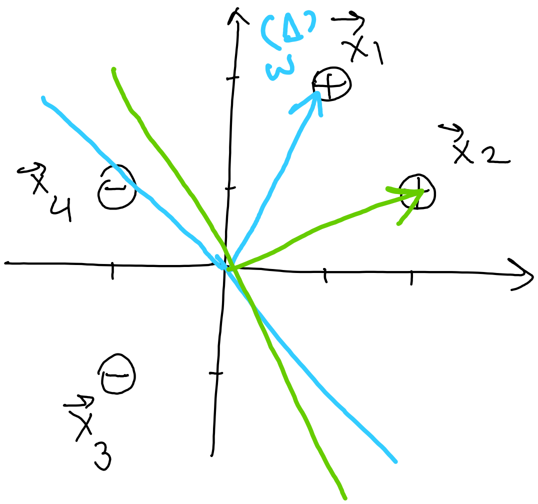
Lecture 9/19: Convergence of Perceptron

Thursday, September 19, 2019 2:42 PM



$\vec{x}$	y
$\vec{x}_1 = (1,2)$	$+1$
$\vec{x}_2 = (2,1)$	$+1$
$\vec{x}_3 = (-1,-1)$	-1
$\vec{x}_4 = (-1,1)$	-1

$$\vec{w}^{(0)} = (0,0)$$

$\hookrightarrow \vec{x}_1 = y_1(\vec{w}^{(0)} \cdot \vec{x}_1) = 0 \nless 0$

$\hookrightarrow \vec{w}^{(1)} = \vec{w}^{(0)} + y_1\vec{x}_1 = (1,2)$

$$\vec{w}^{(1)} = (1,2)$$

$\hookrightarrow x_1 : \; y_1(\vec{w}^{(1)} \cdot \vec{x}_1) = (1,2)(1,2) = 5 > 0$

$x_2 : \; y_2(\vec{w}^{(1)} \cdot \vec{x}_2) = (1,2)(2,1) = 4 > 0$

$x_3 = \; y_3(\vec{w}^{(1)} \cdot \vec{x}_3) = -1(1,2)(-1,-1) = 3 > 0$

$$x_4 = y_4\left(\vec{w}^{(1)} \cdot \vec{x}_4\right) = -1\,(1,2)\,(-1,+1) = -1 \times +1 < 0 \quad \times$$

$$\hookrightarrow \vec{w}^{(2)} = \vec{w}^{(1)} + y_4 \vec{x}_4 = (1,2) - (-1,1) = (2,1)$$

$$\vec{w}^{(2)} = (2,1)$$

$\hookrightarrow \vec{x}_1$

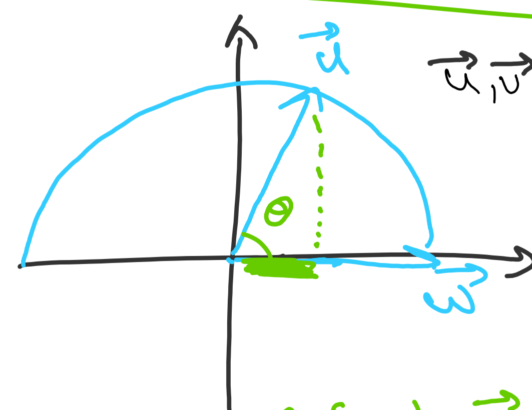
$\hookrightarrow \vec{x}_2$

$\hookrightarrow \vec{x}_3$

$\hookrightarrow \vec{x}_4$

✓

$$\cos\left(\theta(\vec{u},\vec{v})\right) = \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\|\,\|\vec{v}\|} \in [-1, 1]$$



$\vec{u}, \vec{v}$ normalized

$\cos(\theta) = \vec{u} \cdot \vec{w}$

Idea: Proof by Contradiction

if $t+1 > \frac{R^2}{\gamma^2} \implies \cos\left(\vec{w}^{(t+1)}, \vec{w}^*\right) > 1.$

---

(1) lower bound on $\vec{w}^{(t+1)} \cdot \vec{w}^*$

$$\vec{w}^* \cdot \vec{w}^{(t+1)} = \vec{w}^* \cdot \left(\vec{w}^{(t)} + y_i \vec{x}_i\right)$$

b/c Update on $(\vec{x}_i, y_i)$

$$= \vec{w}^* \cdot \vec{w}^{(t)} + \underbrace{(\vec{w}^* \cdot \vec{x}_i) y_i}_{\geq \gamma}$$

By Margin.

$$\geq \vec{w}^* \cdot \vec{w}^{(t)} + \gamma$$

$\vdots$ $t, t-1, \dots, 1$

$$\geq \underbrace{\vec{w}^* \cdot \vec{w}^{(0)}}_{0} + (t+1)\gamma$$

$\vec{w}^{(0)} = \vec{0}$

$$\geq (t+1)\gamma$$

---

(2) Upper bound $\|\vec{w}^{(t+1)}\|$

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

$$\|\vec{w}^{(t+1)}\|^2 = \vec{w}^{(t+1)} \cdot \vec{w}^{(t+1)}$$

$$= \left(\vec{w}^{(t)} + y_i \vec{x}_i\right) \circ \left(\vec{w}^{(t)} + y_i \vec{x}_i\right)$$

Updated. $(\vec{x}_i, y_i)$

$\vec{w}^{(t)}$ was wrong.

$\Downarrow$

$y_i(\vec{w}^{(t)} \circ \vec{x}) \leq 0$

$$= \vec{w}^{(t)} \circ \vec{w}^{(t)} + 2(\vec{x}_i \circ \vec{w}^{(t)}) y_i + (\vec{x}_i \circ \vec{x}_i) \underline{y_i^2}_{+1}$$

$$= \|\vec{w}^{(t)}\|^2 + \underbrace{2(\vec{x}_i \circ \vec{w}^{(t)}) y_i}_{\leq 0} + \underbrace{\|\vec{x}_i\|^2}_{R^2}$$

$$\leq \|\vec{w}^{(t)}\|^2 + \quad + R^2$$

$$\leq \|\vec{w}^{(t)}\|^2 + R^2 \leq \ldots\ldots$$ repeat for $t, t-1, \ldots, 1$

$$\leq \|\vec{w}^{(0)}\|^2 + (t+1)R^2 \leq (t+1)R^2$$

---

1. $\vec{w}^{(t+1)} \circ \vec{w}^* \geq (t+1)\gamma$

2. $\|\vec{w}^{t+1}\| \leq R\sqrt{t+1}$

3. $t+1 > \dfrac{R^2}{\gamma^2}$

$$\cos(\vec{w}^{(t+1)}, \vec{w}^*) = \frac{\vec{w}^{(t+1)} \circ \vec{w}^*}{\|\vec{w}^{(t+1)}\|}$$

unit vector.

(1), (2)

$$\geq \frac{(t+1)\gamma}{R\sqrt{t+1}} \overset{(3)}{>} \frac{R}{\gamma} \cdot \frac{\gamma}{R} > 1$$ Contradiction

$$\Longrightarrow t+1 < \frac{R^2}{\gamma^2}. \Longrightarrow \leq \frac{R^2}{\gamma^2} \text{ updates}$$

Remarks: Gaurantees independent of

\# features, order, $|S|$, ..., scaling, rotated

actual # of update.

$\longrightarrow$

Outcome $\longrightarrow$ order matters