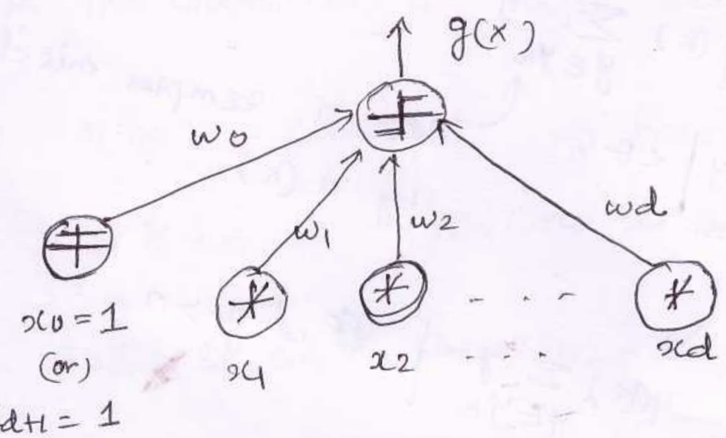
15/4/2019

Perceptron - Example:

The two-category case:

d-dimensional feature vector.



Algorithm: Batch Perceptron

1. begin initialize a, $\eta(\cdot)$, $k \leftarrow 0$
2. do $k \leftarrow k+1$
3. $a \leftarrow a + \eta(k) \sum_{y \in \mathcal{Y}_k} y$
4. until $\left| \eta(k) \sum_{y \in \mathcal{Y}_k} y \right| < \theta$

   set of samples misclassified by $a(k)$.

5. return a
6. end

$\eta(k) \sum_{y \in \mathcal{Y}_k} y \;\Big|\; \cancel{} w^t y > 0$.

Batch perceptron:

The next weight vector is obtained by adding some multiple of the sum of the misclassified samples

$\eta(k) * \sum y$ to the present weight vector $a(k)$.

$$\boxed{a \leftarrow a + \eta(k) * \sum_{y \in \mathcal{Y}_k} y}$$

(or)

$$\boxed{\begin{array}{l} a(0) = \text{Initial weight vector, arbitrary} \\ a(k+1) = a(k) + \eta(k) \cdot \sum y \quad \forall y \text{ misclassified.} \end{array}}$$

(~~??~~)

different variant exits a that is easier to analyze.

* we shall consider the samples in a sequence and shall modify the weight vector whenever it misclassifies a single sample.

* $\eta(k)$ - constant $\Rightarrow$ fixed-increment case.

$\eta(k) = 1$ with no loss in generality.

* The second simplification, when the samples are consided sequentially, some will be misclassified.

$$\overset{\downarrow}{y_1}, y_2, \overset{\downarrow}{y_3}, \overset{\downarrow}{y_1}, \overset{\downarrow}{y_2}, y_3, y_1, y_2$$

$$\boxed{\begin{aligned} &a(0) \text{ - arbirary} \\ &a(k+1) = a(k) + \underbrace{1}_{\eta(k)=1} \cdot y^k \quad k \geq 1 \end{aligned}}$$

Algorithm (Fixed-Increment Single-Sample Perceptron)

1. begin Initialize a, $k \leftarrow 0$
2. do $k \leftarrow (k+1) \bmod n$ $\quad a(k) \quad a(k+1) \leftarrow a(k) + y^k$
3. If y^k is misclassified by a then $a \leftarrow a + y^k$
4. until all samples properly classified
5. return a
6. end

Description of the Patterns:- Example: Perceptron learning Algorithm.

Pattern no	1	2	Class	
x_1	0.5	3.0	X, 1	ω_2
x_2	1	3.0	X, 1	
x_3.	0.5	2.5	X, 1	
x_4.	1	2.5	X, 1	
x_5.	1.5	2.5	X, 1	
x_6.	4.5	1	O, 2	ω_1
x_7.	5	1	O, 2	
x_8.	4.5	0.5	O, 2	
x_9.	5.5	0.5	O, 2	

Pattern no	1	2	3	
x_1	-0.5	-3.0	-1	ω_2
x_2	-1	-3.0	-1	
x_3	-0.5	-2.5	-1	
x_4	-1	-2.5	-1	
x_5	-1.5	-2.5	-1	
x_6	4.5	1	1	ω_1
x_7	5	1	1	
x_8	4.5	0.5	1	
x_9	5.5	0.5	1	

augment the vector and negate it $\begin{bmatrix} -x_1 \\ -x_2 \\ -1 \end{bmatrix}$

augment the vector ~~and negate it~~

1. $w_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ and $x_1 \begin{pmatrix} -0.5 \\ -3.0 \\ -1 \end{pmatrix}$

here $w_1^t x_1 = 0$ so $w_2 = w_1 + x_1$ which is represented by

$$\boxed{a(k+1) = a(k) + \eta(k) \cdot \sum_{\forall y \text{ misclassified}} y}$$

$$w_2 = w_1 + x_1$$

$$= \begin{pmatrix} -0.5 \\ -3 \\ -1 \end{pmatrix}$$

2. next we consider pattern x_2. $w_2^t x_2$

$$\begin{pmatrix} -0.5 & -3 & -1 \end{pmatrix} \begin{pmatrix} -1 \\ -3 \\ -1 \end{pmatrix} \leftarrow x_2 = 10.5 > 0$$

x_3, x_4 & x_5 are also properly classified.

$$\begin{pmatrix} -0.5 & -3 & -1 \end{pmatrix} \begin{pmatrix} -0.5 \\ -2.5 \\ -1 \end{pmatrix} \leftarrow x_3 = 8.75$$

$$\begin{pmatrix} -0.5 & -3 & -1 \end{pmatrix} \begin{pmatrix} -1 \\ -2.5 \\ -1 \end{pmatrix} = 9$$

$$\begin{pmatrix} 0.5 & -3 & -1 \end{pmatrix} \begin{pmatrix} -1.5 \\ -2.5 \\ -1 \end{pmatrix} = 9.25$$

3. However, $w_2^t x_6$

$$\begin{pmatrix}-0.5 & -3 & -1\end{pmatrix}\begin{pmatrix}4.5\\1\\1\end{pmatrix} = \begin{matrix}<0\\-6.25\end{matrix}$$

so update weight vector

$$w_3 = w_2 + x_6$$

$$= \begin{pmatrix}-0.5\\-3\\-1\end{pmatrix}\begin{pmatrix}4.5\\1\\1\end{pmatrix} = \begin{pmatrix}4\\-2\\0\end{pmatrix}$$

note that w_3 classifies patterns x_7, x_8, x_9, and in the next iteration x_1, x_2, x_3 & x_4 correctly.

$$w_3^t x_7 = \begin{pmatrix}4 & -2 & 0\end{pmatrix}\begin{pmatrix}5\\1\\1\end{pmatrix} = 18$$

$$w_3^t x_8 = \begin{pmatrix}4 & -2 & 0\end{pmatrix}\begin{pmatrix}4.5\\0.5\\1\end{pmatrix} = 17$$

$$w_3^t x_9 = \begin{pmatrix}4 & -2 & 0\end{pmatrix}\begin{pmatrix}5.5\\0.5\\1\end{pmatrix} = 21$$

$$w_3^t x_1 = \begin{pmatrix}4 & -2 & 0\end{pmatrix}\begin{pmatrix}-0.5\\-3.0\\-1\end{pmatrix} = 4$$

$$w_3^t x_2 = \begin{pmatrix}4 & -2 & 0\end{pmatrix}\begin{pmatrix}-1\\-3\\-1\end{pmatrix} = 2$$

$$w_3^t x_3 = \begin{pmatrix} 4 & -2 & 0 \end{pmatrix} \begin{pmatrix} -0.5 \\ -2.5 \\ -1 \end{pmatrix} = 3$$

$$w_3^t x_4 = \begin{pmatrix} 4 & -2 & 0 \end{pmatrix} \begin{pmatrix} -1 \\ -2.5 \\ -1 \end{pmatrix} = 1$$

4. However x_5 is misclassified by w_3. note that

$w_3^t x_5$ is -1

$$w_3^t x_5 = \begin{pmatrix} 4 & -2 & 0 \end{pmatrix} \begin{pmatrix} -1.5 \\ -2.5 \\ -1 \end{pmatrix} = -1 \quad < 0$$

So, update weight vector $w_4 = w_3 + x_5$.

$$w_4 = \begin{pmatrix} 4 \\ -2 \\ 0 \end{pmatrix} + \begin{pmatrix} -1.5 \\ -2.5 \\ -1 \end{pmatrix} = \begin{pmatrix} 2.5 \\ -4.5 \\ -1 \end{pmatrix}$$

w_4 classifies patterns $x_6, x_7, x_8, x_9, x_1, x_2, x_3, x_4$ & x_5 correctly.

$$w_4^t x_6 = \overset{2.5 \quad -4.5 \quad -1}{\begin{pmatrix} 4 & -2 & 0 \end{pmatrix}} \begin{pmatrix} 4.5 \\ 1 \\ 1 \end{pmatrix} = 5.75$$

$$w_4^t x_7 = \overset{2.5 \quad -4.5 \quad -1}{\begin{pmatrix} \cancel{4} & \cancel{-2} & \cancel{0} \end{pmatrix}} \begin{pmatrix} 5 \\ 1 \\ 1 \end{pmatrix} = 7$$

$$w_4^t x_8 = \begin{pmatrix} \overset{2.5}{\cancel{4}} & \overset{-4.5}{\cancel{-2}} & \overset{-1}{\cancel{0}} \end{pmatrix} \begin{pmatrix} 4.5 \\ 0.5 \\ 1 \end{pmatrix} = 8$$

$$w_4^t x_9 = \begin{pmatrix} \overset{2.5}{\cancel{4}} & \overset{-4.5}{\cancel{-2}} & \overset{-1}{\cancel{0}} \end{pmatrix} \begin{pmatrix} 5.5 \\ 0.5 \\ 1 \end{pmatrix} = 10.5$$

$$w_4^t x_1 = \begin{pmatrix} \overset{2.5}{\cancel{4}} & \overset{-4.5}{\cancel{-2}} & \overset{-1}{\cancel{0}} \end{pmatrix} \begin{pmatrix} -0.5 \\ -3.0 \\ -1 \end{pmatrix} = 13.25$$

$$w_4^t x_2 = \begin{pmatrix} \overset{2.5}{\cancel{4}} & \overset{-4.5}{\cancel{-2}} & \overset{-1}{\cancel{0}} \end{pmatrix} \begin{pmatrix} -1 \\ -3 \\ -1 \end{pmatrix} = 11.5$$

$$w_4^t x_3 = \begin{pmatrix} \overset{2.5}{\cancel{4}} & \overset{-4.5}{\cancel{-2}} & \overset{-1}{\cancel{0}} \end{pmatrix} \begin{pmatrix} -0.5 \\ -2.5 \\ -1 \end{pmatrix} = 11$$

$$w_4^t x_4 = \begin{pmatrix} \overset{2.5}{\cancel{4}} & \overset{-4.5}{\cancel{-2}} & \overset{-1}{\cancel{0}} \end{pmatrix} \begin{pmatrix} -1 \\ -2.5 \\ -1 \end{pmatrix} = 9.75$$

$$w_4^t x_5 = \begin{pmatrix} \overset{2.5}{\cancel{4}} & \overset{-4.5}{\cancel{-2}} & \overset{-1}{\cancel{0}} \end{pmatrix} \begin{pmatrix} -1.5 \\ -2.5 \\ -1 \end{pmatrix} = 8.5$$

So w_4 is the desired vector. (a) In other words

$\boxed{2.5 x_1 - 4.5 x_2 - 1 = 0}$ is the equation of the decision boundary. equivalently, the line separating the two classes is $\boxed{5x_1 - 9x_2 - 2 = 0.}$; $\boxed{w_1 = 5 \; ; \; w_2 = -9 \; ; \; w_0 = -2}$