The Perceptron Algorithm 1/2

Conventions

- We will use vector notation: a row of numbers $(a_1, a_2, ..., a_N)$ can be written concisely as: $\vec{a}$.
- In the literature one often finds the dot product: given two rows of numbers, one can combine them in a weighted sum: e.g. for $(x_1, x_2, ..., x_N)$ with weights $(w_1, w_2, ..., w_N)$, the weighted sum is: $\sum_i w_i x_i$. This is often written more concisely with the dot product:

$$\vec{x} \cdot \vec{w} \equiv \sum_i x_i w_i$$

- The length of an N-dimesnional vector $\vec{x}$ is defined as:

$$\mid \vec{x} \mid \equiv \sqrt{\sum_i x_i^2}$$

- Note that $\mid x \mid \equiv \sqrt{\vec{x} \cdot \vec{x}}$

The Perceptron Algorithm (continued) 2/2

Suppose that we have sets of data points. The data points are labeled: some of them belong to a class F^+, others belong to a class F^-. No data points belongs to two classes.

START:
Generate random weights $\vec{w}$
TEST: Choose any $\vec{x}$ from ether F^+ or F^-
if $\vec{x} \epsilon F^+$ and $\sum w_i x_i > 0$ goto TEST:
if $\vec{x} \epsilon F^+$ and $\sum w_i x_i \leq 0$ goto ADD:
if $\vec{x} \epsilon F^-$ and $\sum w_i x_i \leq 0$ goto TEST:
if $\vec{x} \epsilon F^-$ and $\sum w_i x_i > 0$ goto SUB:
ADD:
$\vec{w} \rightarrow \vec{w} + \vec{x}$
goto TEST:
SUB:
$\vec{w} \rightarrow \vec{w} - \vec{x}$
goto TEST:

In this algorithm it is assumed that the threshold θ has been converted to one of the weights! For a perceptron with weights w_1, w_2 and bias θ, the classification of point $\vec{x}$ amounts to checking if amounts to checking whether:

$$w_1 x_1 + w_2 x_2 < \theta \tag{1}$$

By introducting a $w_3 = -\theta$ and an $x_3 \equiv 1$, this is equivalent to checking if:

$$w_1 x_1 + w_2 x_2 + w_3 x_3 < 0 \tag{2}$$