Support Vector Machines (SVM)

One way to write the equation for a line in a plane is

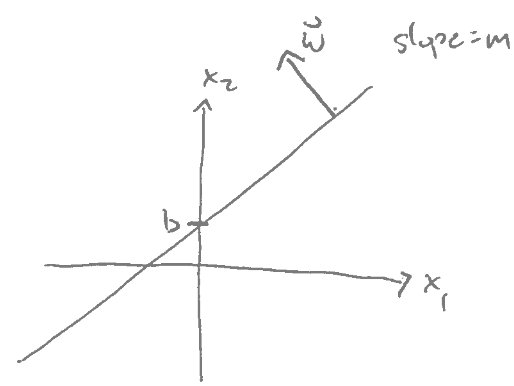
$$x_2 = m x_1 + b$$

In higher dimensions, it is more convenient to write

$$\begin{pmatrix} -m \\ 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = b$$

$$\downarrow$$

$$\vec{w} \cdot \vec{x} = b$$

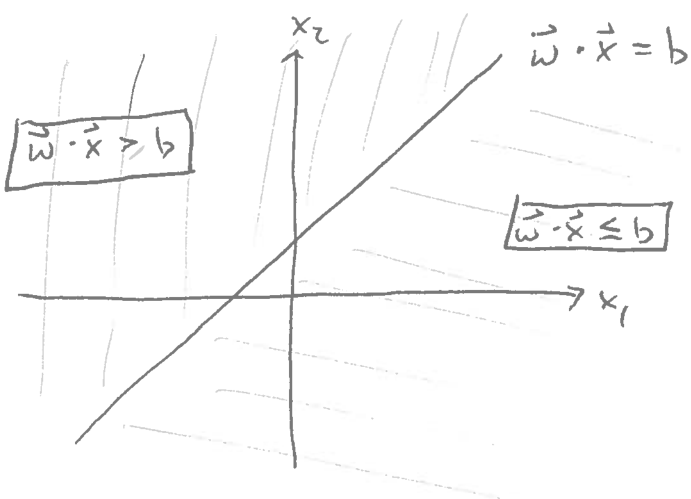


Note: $\vec{w}$ is perpendicular or normal to the line.

In 2-dimensions, the set of pts. satisfying $\vec{w} \cdot \vec{x} = b$ is a <u>line</u>. In 3-dim, its a <u>plane</u>. The vector $\vec{w}$ is normal to the plane. In n-dim, its an n-1 dim <u>hyperplane</u>.

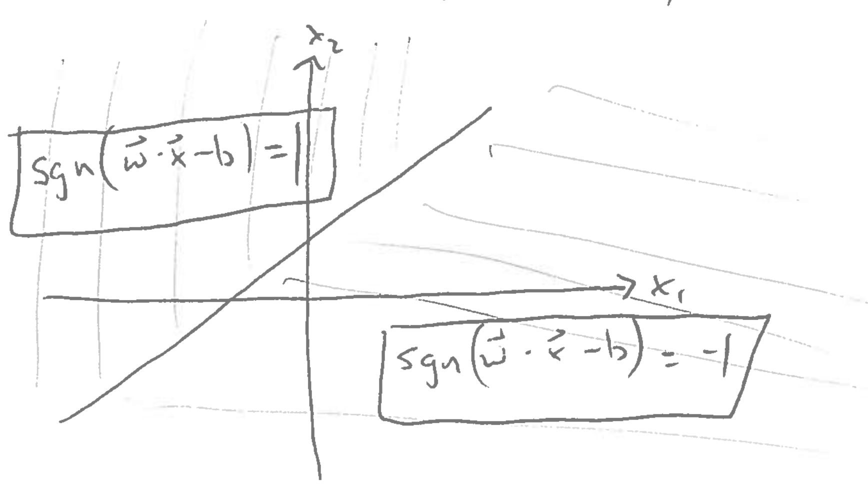
How do we describe the set of pts. above the line, $\vec{w}\cdot\vec{x} = b$, in 2-dim? It is the half-space of pts. $\vec{x}$ satisfying

$$\vec{w}\cdot\vec{x} \geq b.$$

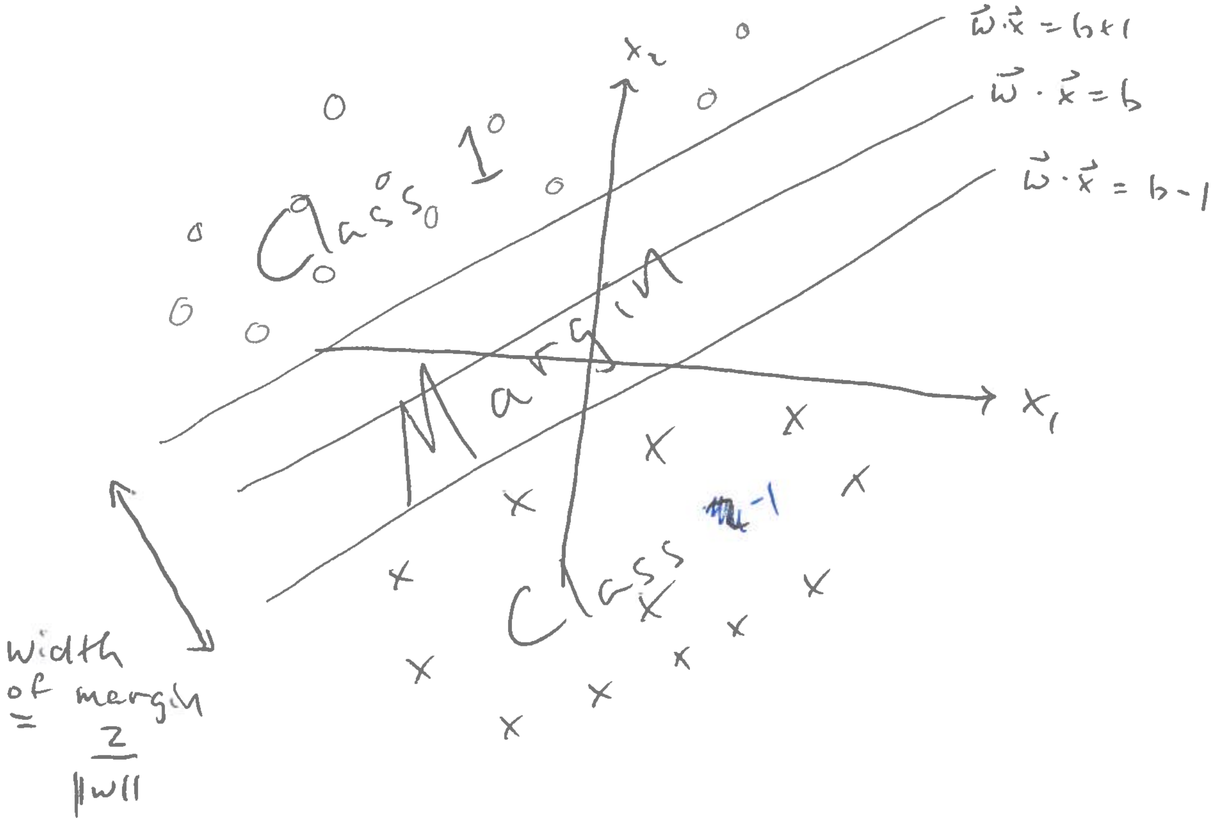


If I know $\vec{w}$ and b, I can define a classifier by

$$\vec{x} \longmapsto \operatorname{sgn}(\vec{w}\cdot\vec{x} - b)$$



If I take $b \mapsto b+1$ or $b \mapsto b-1$, I get 2 more lines parallel to the first.



Idea: Given data $\{x_i\}_i$ and labels $\{y_i\}_i$ $i=1,\dots,n$

$$\underset{w,b}{\text{maximize}} \quad \frac{1}{\|w\|} \quad \Leftarrow \text{width of margin}$$

s.t. $\vec{w}\cdot\vec{x}_i \geq b+1$ $\forall\, i$ belonging to class 1

constraints: $\nearrow$ $\vec{w}\cdot x_i \leq b-1$ $\forall\, i$ belonging to class ~~2~~ -1

each point lies on the correct side of the separating hyperplane.

It can be shown that this is equivalent to

$$\underset{w, b}{\text{minimize}} \quad \|w\|^2$$

$$\text{s.t.} \quad y_i (\vec{w} \cdot \vec{x}_i - b) \geq 1 \quad \forall i.$$

From the picture, the max-margin hyperplane will touch some of the x_i (otherwise, we could take $\|w\|$ to be smaller). These x_i are called <u>support vectors</u>.