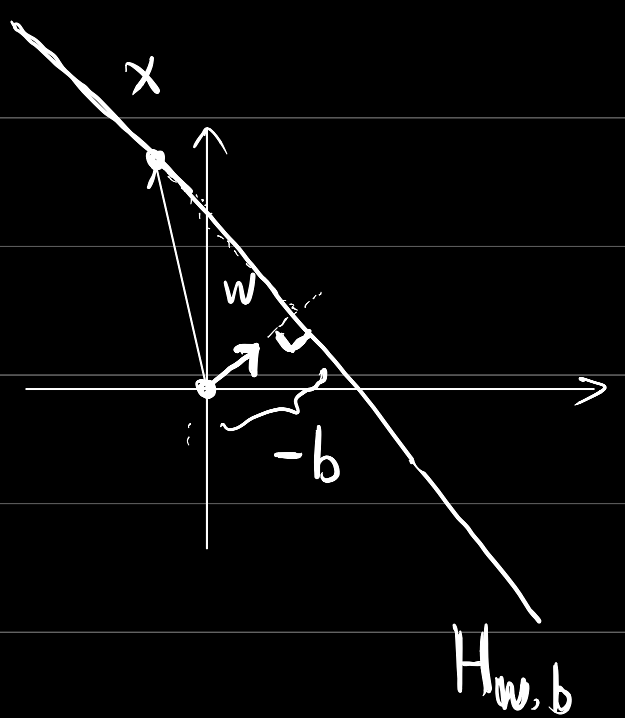
4.2.2. Theory

Recall that given a unit vector $w \in \mathbb{R}^p$ with $\|w\| = 1$ and $b \in \mathbb{R}$. Then

$$H_{w,b} \overset{\text{def}}{=} \{ x \in \mathbb{R}^p : x \cdot w + b = 0 \}$$

$$= \{ x \in \mathbb{R}^p : \underbrace{x \cdot w}_{\text{magnitude of proj of } x \text{ onto } w.} = -b \}$$

represents a hyperplane that is orthogonal to w and of distance $|b|$ to the origin.



Remark:

(1) For any const $c \neq 0$,

$$H_{cw, cb} = \{ x \in \mathbb{R}^p : cw \cdot x + cb = 0 \}$$

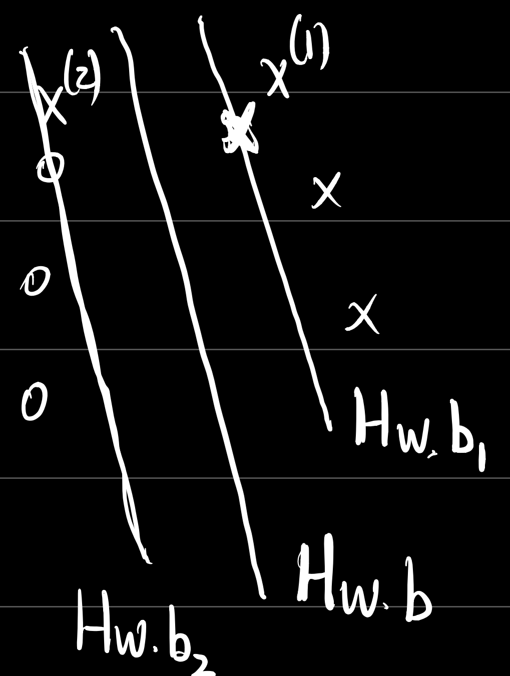
$$= \{ x \in \mathbb{R}^p : w \cdot x + b = 0 \} = H_{w,b}$$

In particular, $H_{-w,-b} = H_{w,b}$.

(2) By varying b, we get a series of parallel hyperplanes. By varying w, we get hyperplanes with different orientations.

(3) $H_{w,b}$ correctly classifies the sample pts

$\Leftrightarrow$ for each i,



$$\begin{cases} x^{(i)} \cdot w + b > 0 & \text{if } Y_i = 1 \\ x^{(i)} \cdot w + b < 0 & \text{if } Y_i = -1 \end{cases}$$

Without loss of generality, suppose

$$x^{(1)} \cdot w + b = \min_{\{i : y_i = 1\}} x^{(i)} \cdot w + b$$

i.e. $x^{(1)}$ is the "x" closest to $H_{w,b}$.

$$x^{(2)} \cdot w + b = \max_{\{i : y_i = -1\}} x^{(i)} \cdot w + b$$

i.e. $x^{(2)}$ is the "o" closest to $H_{w,b}$.

By varying b, $\exists\, b_1$ such that $x^{(1)} \in H_{w,b_1}$

(i.e. $x^{(1)} \cdot w + b_1 = 0$)

and $x^{(i)}$'s with $y_i = 1$ are "above" H_{w,b_1}.

$\exists\, b_2$ such that $x^{(2)} \in H_{w,b_2}$

(i.e. $x^{(2)} \cdot w + b_2 = 0$)

and $x^{(i)}$'s with $y_i = -1$ are "below" H_{w,b_2}.

Note in the picture $-b_1 > -b_2$, i.e. $b_1 < b_2$

Consider the hyperplane $H_{w, \frac{b_1+b_2}{2}} = H_{\tilde{w}, \tilde{b}}$

where $\tilde{w} = \frac{2}{b_2-b_1} w$, $\tilde{b} = \frac{2}{b_2-b_1} \cdot \frac{b_1+b_2}{2} = \frac{b_1+b_2}{b_2-b_1}$

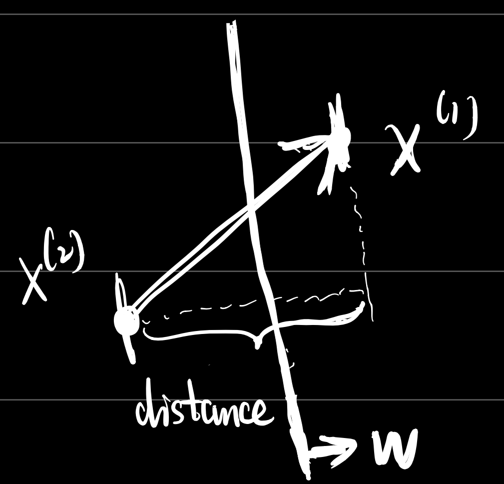
Then (1) it is most stable with respect to perturbations of inputs.

(2) it is equally distant from H_{w,b_1} and H_{w,b_2}.

(3) The distance between H_{w,b_1} and H_{w,b_2} ("margin")

$$= \left| \left(x^{(1)} - x^{(2)}\right) \cdot w \right|$$

$$\overset{\text{eqns of } H_{w,b_1} \text{ and } H_{w,b_2}}{=\!=} \left| -b_1 + b_2 \right| = b_2 - b_1$$



$$\|\tilde{w}\| = \left\| \frac{2}{b_2 - b_1} w \right\| = \frac{2}{b_2 - b_1}$$

$$\frac{2}{\|\tilde{w}\|}$$