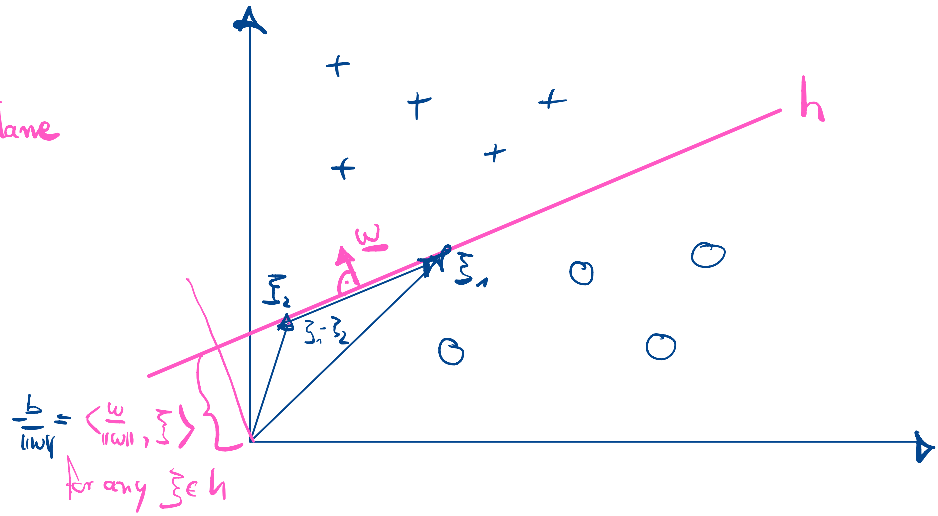
LARGE MARGIN CLASSIFICATION

Hyperplane $h_{\underline{w},b}$

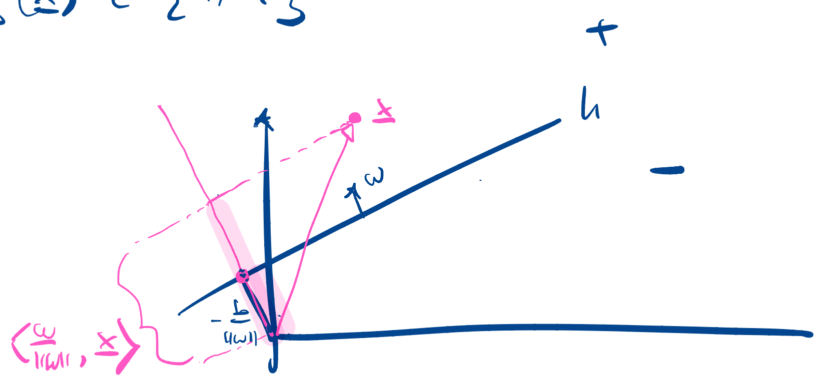


$$\langle \underline{w}, \underline{\xi}_1 - \underline{\xi}_2 \rangle \overset{!}{=} 0 \iff \langle \underline{w}, \underline{\xi}_1 \rangle = \underline{\langle \underline{w}, \underline{\xi}_2 \rangle =: -b}$$

$$\Rightarrow h_{\underline{w},b} := \{ \underline{\xi} \mid \langle \underline{w}, \underline{\xi} \rangle + b = 0 \}$$

Decision function

$$f_{\underline{w},b}(\underline{x}) \in \{1, -1\}$$



$$\Rightarrow \left\langle \frac{w}{\|w\|}, \underline{x} \right\rangle + \frac{b}{\|w\|} > 0 \quad \Rightarrow \quad f_{w,b}(\underline{x}) = +1$$

$$\text{———} \;\; " \;\; \text{———} < 0 \quad \Rightarrow \quad f_{w,b}(\underline{x}) = -1$$

$$\Rightarrow \quad f_{w,b}(\underline{x}) := \operatorname{sgn}\left(\underline{\langle w, \underline{x} \rangle + b} \right)$$

$$\Downarrow$$

find w, b s.th. the training points satisfy $f_{w,b}(x_i) = y_i \quad \forall i = 1, \dots, N$ (*)

how?

threshold neuron

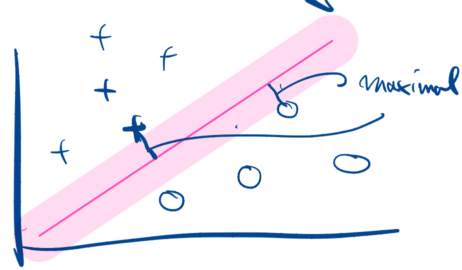
max. margin classifier $\Rightarrow$ SVMs

$$\min_{w,b} \; \text{loss}\left(y_i, f_{w,b}(x_i)\right)$$

$$\downarrow$$

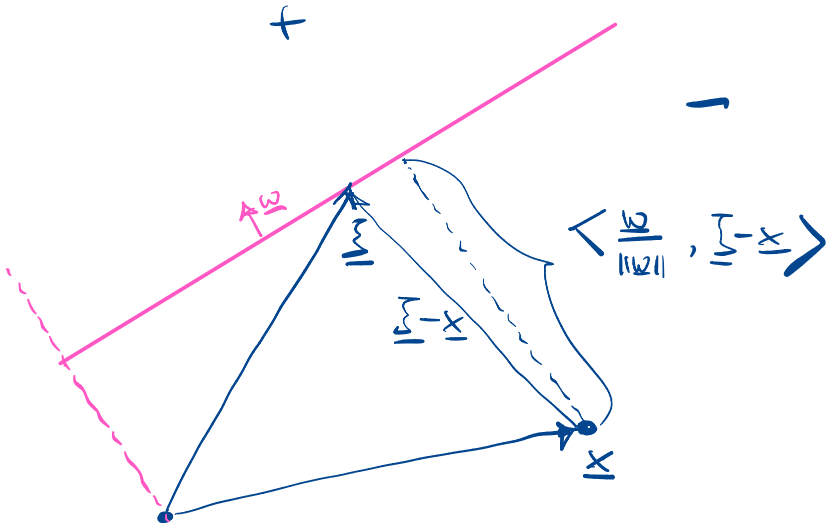
some h that solves (*)

maximize distance to the closest training points



margin

$m_{w,b}$



↓ $x^*_\pm$ are closest points with labels $\pm$ (respectively)

$$\underline{m_{w,b}} := \mathrm{dist}(h_{w,b}, x^*_\pm) = \left| \underbrace{\left\langle \frac{w}{\|w\|}, \xi \right\rangle}_{-\frac{b}{\|w\|}} - \left\langle \frac{w}{\|w\|}, x^*_\pm \right\rangle \right|$$

$$= \frac{1}{\|w\|} \left| \langle w, x^*_\pm \rangle + b \right|$$

$$= \pm \frac{1}{\|w\|} \cdot \left(\langle w, x^*_\pm \rangle + b \right)$$

Scaling invariance

$$h_{w,b} = \{\xi \mid \langle w, \xi \rangle + b = 0\}$$

$$h_{\lambda w, \lambda b} = \{\xi \mid \langle w, \lambda \xi \rangle + \lambda b = 0\} = h_{w,b}$$

where $\lambda \neq 0$

Scaling trick

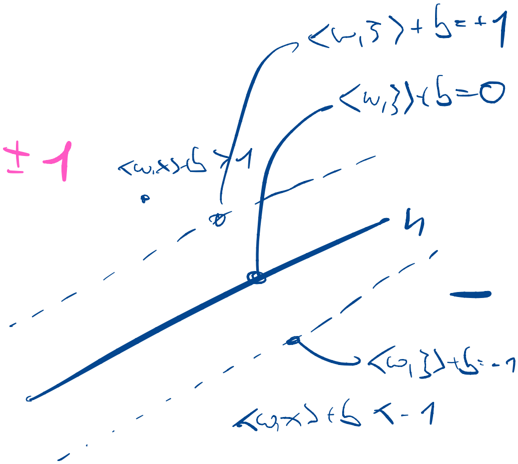
$$\underbrace{\|w\| \, m_{w,b}}_{=1} = \langle w, x_+^* \rangle + b$$

$$-\|w\| \, m_{w,b} = \langle w, x_-^* \rangle + b$$

$\Rightarrow$ pick λ s.th. $\|w\| = \frac{1}{m_{w,b}}$

$\Rightarrow$ $m_{w,b} = \frac{1}{\|w\|}$

$$\langle w, x_\pm^* \rangle + b = \pm 1$$



$$\Downarrow$$

linear SVM

"primal problem"

$$\min_{w,b} \|w\|^2 \quad \text{s.t.} \quad \underbrace{f_{w,b}(x_i) = y_i}_{} \quad \forall i$$

$$\Leftrightarrow \underline{y_i(\langle w, x_i \rangle + b) \geq 1 \ \forall i}$$

(constrained opt. problem)