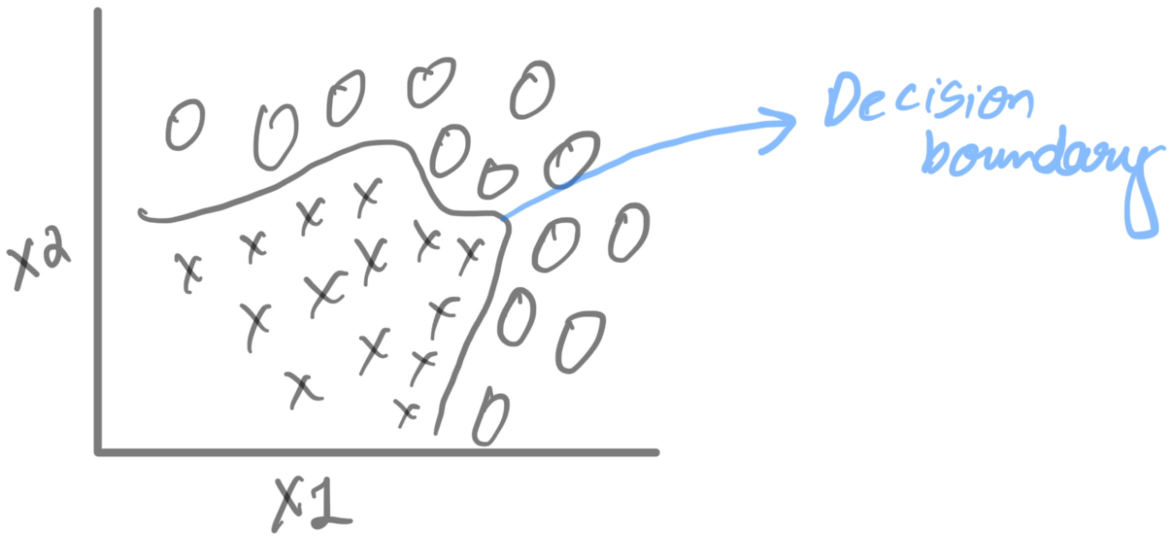
Stanford CS229 Lecture-5

Support Vector machines :- (Classification)

↳ for non linear classification

Ex:-

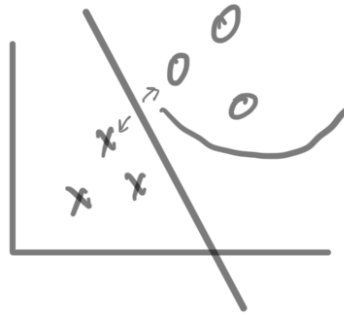


- SVM takes x_1 & x_2 and maps it to higher dimensions

ex:- $\begin{bmatrix} x_1 \\ x_2 \\ x_1^2 \\ x_2^2 \end{bmatrix}$ → so that we can obtain a non linear boundary

Linear S.V.M

↳



We have to maximize this distance (Geometric margin)

Formula :-

$$h_{w,b}(x) = g(w^T x + b) \rightarrow \mathbb{R}$$

$w \rightarrow \mathbb{R}^n$

g ⇓

$$\sum_{i=1}^{i=n} w_i x_i + b$$

Functional margin of hyperplane is defined by (w, b) w.r.t (x^i, y^i) → some point

$$\hat{\gamma}^{(i)} = y^{(i)}\left(w^T x^i + b\right)$$ } Functional margin :- the output

If $y^{(i)} = 1$, we want $w^T x^i + b >> 0$

if $y^{(i)} = -1$, we want $w^T x^i + b << 0$

Since,
SVM classify either 1 or -1,
if $w^T x^i + b >> 0$ & $y^i = 1$ then margin is high

if $w^T x^i + b << 0$ & $y^i = -1$ then margin is high here too.

→ essentially $\hat{\gamma}^{(i)}$ should either be $>> 0$ or $<< 0$,,

Functional margin w.r.t to the entire training set:-

$$\hat{\gamma} = \min_i \hat{\gamma}^{(i)} \qquad i = 1 \dots m$$

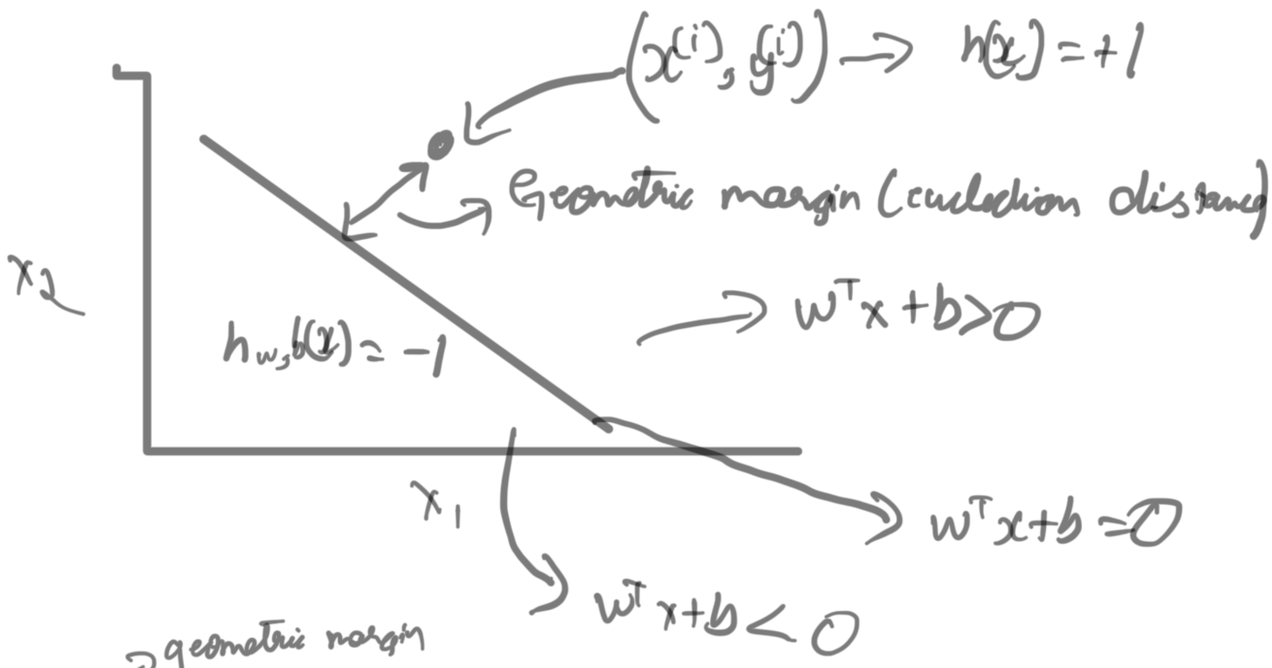
* Caveat

increasing w & b increases $w^T x_i + b$ which increases margin but does nothing to algorithm (Cheating)

→ So normalise w & b

$$(w, b) \rightarrow \left(\frac{w}{\|w\|}, \frac{b}{\|w\|}\right)$$

Geometric margin w.r.t single point



geometric margin

$$\gamma^{(i)} = \frac{w^T x^{(i)} + b}{\|w\|} \cdot y^{(i)}$$

$$\gamma^{(i)} = \frac{\hat{\gamma}^{(i)}}{\|w\|} \rightarrow \text{functional margin}$$

GOAL :-

→ choose w, b to maximize γ (Geometric margin)