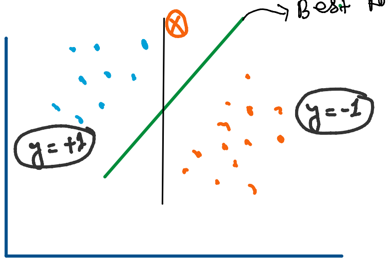
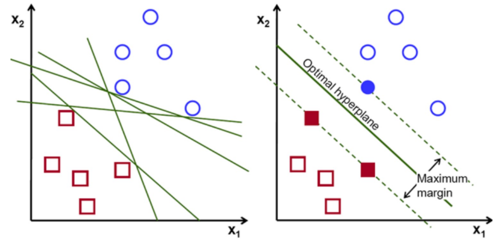
SVM

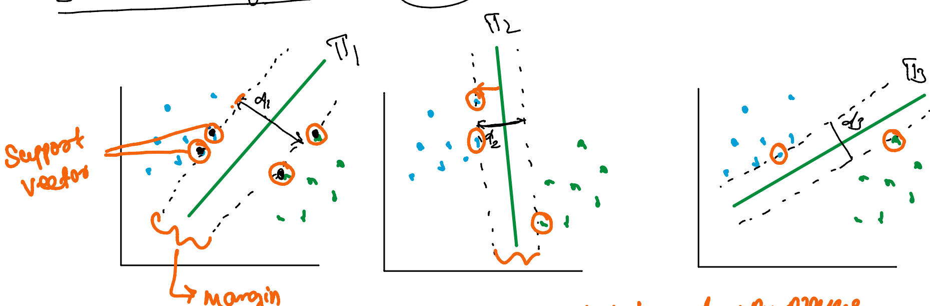
06 May 2022 06:38 PM

Support Vector Machine





SVM — Best hyper plane (How?)



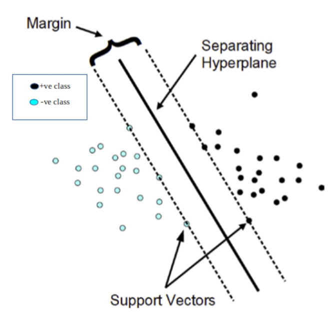
SVM ⇒ Margin maximizing hyperplane

Topic discuss

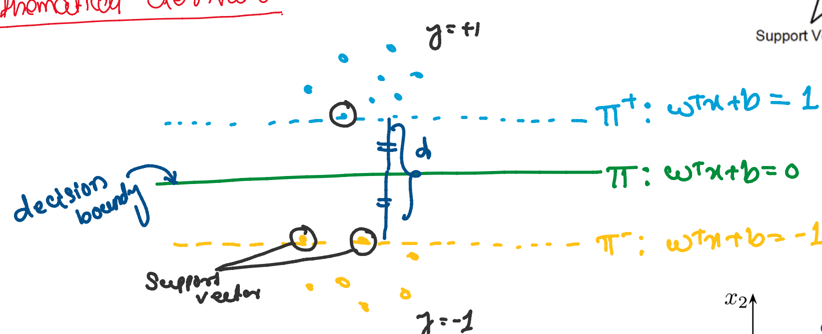
① Type
- hard SVM
- soft SVM → Robust to outliers

② Kernel tricks

③ SVM
- SVC → classification
- SVR → Regression

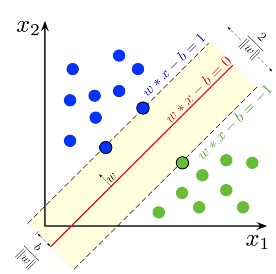


Mathematical derivation



Decision function

$$y = \begin{cases} +1, & \text{if } w^Tx+b \geq 0 \\ -1, & \text{if } w^Tx+b < 0 \end{cases}$$



Optimization function

↳ Maximize distance (d) given we have support vector

↳ we have π^+, π^-

⇓

for +ve point & π^+ ⇒ $w^T x + b \geq 1$, $y = +1$

for -ve point & π^- ⇒ $w^T x + b \leq -1$, $y = -1$

⇓

multiplying y_i and $w^T x_i + b$

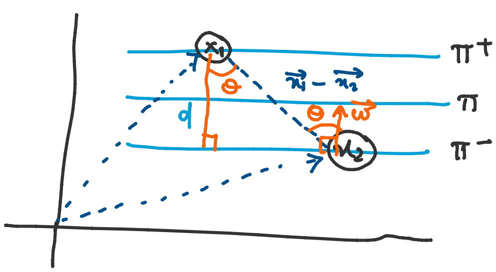
⇓

$$\boxed{y_i(w^T x + b) \geq 1}$$

∴ Optimization Equation

↳ $\boxed{\text{Maximize distance } (d) \text{ given } y_i(w^T x + b) \geq 1}$

How to find d?



$$\vec{w} \cdot (\vec{x_1} - \vec{x_2}) = |\vec{w}| \, |\vec{x_1} - \vec{x_2}| \cos\theta$$

$$= |w| \, \cancel{k} \cdot \frac{d}{\cancel{k}}$$

$$\Rightarrow d = \frac{\vec{w}(\vec{x_1} - \vec{x_2})}{|w|}$$

$$\Rightarrow d = \frac{\vec{w} \cdot \vec{x_1} - \vec{w}\,\vec{x_2}}{|w|}$$

$$= \frac{1 - b + 1 + b}{|w|}$$

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

$$\Rightarrow \boxed{d = \frac{2}{|w|}}$$

$y_i(w^T x + b) \geq 1$

for $x_1 \rightarrow y = +1$

$+1(w^T x_1 + b) = 1$

$\Rightarrow \boxed{w^T x_1 = 1 - b}$ —— (1)

for $x_2 \rightarrow y = -1$

$-1(w^T x_2 + b) = 1$

$\Rightarrow \boxed{w^T x_2 = -1 - b}$ —— (2)

Optimization function

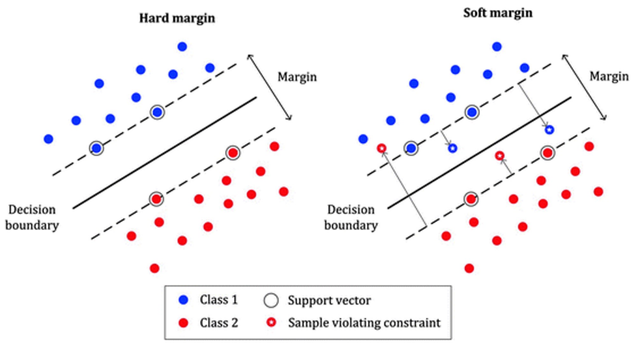
$$\underset{(w^*, b^*)}{\text{argMax}} \left(\frac{2}{\|w\|} \right) \quad \text{such that } y_i(w^T x + b) \geq 1$$

or

$$\underset{(w^*, b^*)}{\text{argMin}} \left(\frac{\|w\|}{2} \right) \quad \text{such that } y_i(w^T x + b) \geq 1$$

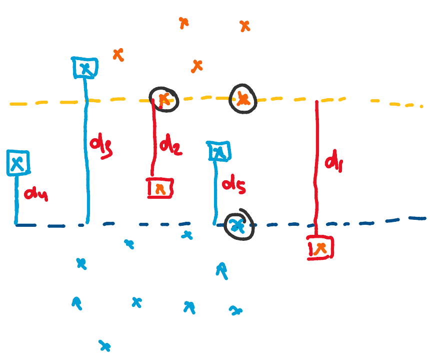
$\because \max(f(x)) = \min\left(\frac{1}{f(x)}\right)$

$\hookrightarrow$ **Hard Margin SVM**



Soft Margin SVM

$\hookrightarrow$ Allow some Error and Try to maximize d



$$\text{Error} = d_1 + d_2 + d_3 + \dots$$

$$= \sum_{i=1}^{n} \xi_i \quad \hookrightarrow \text{zeta}$$

$\xi_i = 0$ for correctly classified x

$\xi_i = d_i$ for incorrectly " "

Our Target $\Rightarrow$ $\min(\text{loss}) + \max(d)$

$$= \min\left(\sum \xi_i\right) + \max\left(\frac{2}{\|w\|}\right)$$

$$= \min\left(\sum \xi_i\right) + \min\left(\frac{\|w\|}{2}\right)$$

Optimization function

Optimization function

$$\boxed{\underset{(w^*, b^*)}{\arg\min}\left[\underbrace{C}_{\text{Hyper parameter}} \cdot \underbrace{\sum_{i=1}^{n} \xi_i}_{\substack{\text{Classification Error} \\ \text{(hinge loss)}}} + \underbrace{\frac{\|w\|}{2}}_{\substack{\text{Margin Error} \\ \text{(Regularization)}}}\right]} \rightarrow \text{Soft margin SVM}$$

$$\boxed{\begin{aligned} &\text{SVM} = C \times \text{hinge loss} + \text{Regularization} \\ &\text{logistic Reg} = \lambda \times \text{logistic loss} + \text{Regularization} \end{aligned}}$$

where $\boxed{C = \frac{1}{\lambda}}$