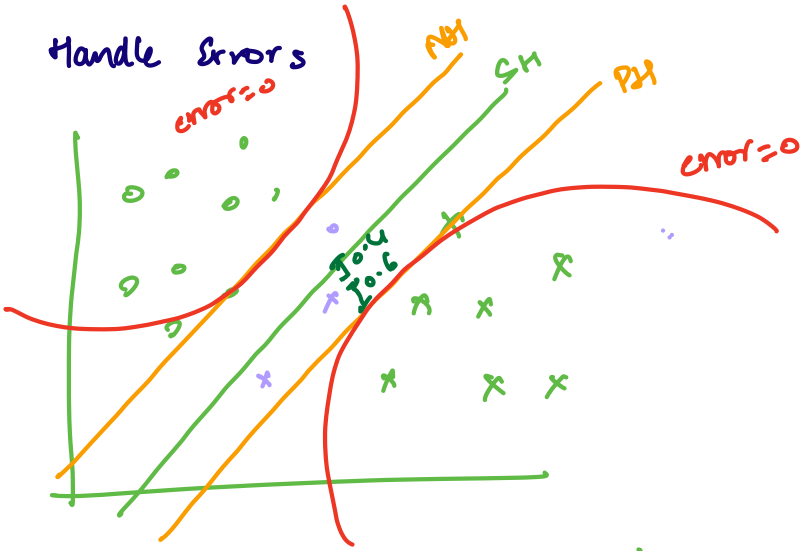
SVM objective:

$$\begin{cases} \min \dfrac{\|w\|^2}{2} \\ \text{such that } \; y^{(i)}\left(w^T x^{(i)} + b\right) \geq 1 \end{cases}$$

Handle Errors



$\varepsilon^{(i)}$ denotes the distance of ith point from true hyperplane

$$y^{(i)}\left(w^T x^{(i)} + b\right) \geq 1$$

$$y^{(i)}\left(w^T x^{(i)} + b\right) \geq \underbrace{1 - \underbrace{\varepsilon^{(i)}}_{0.6}}_{0.4}$$

SVM objective:

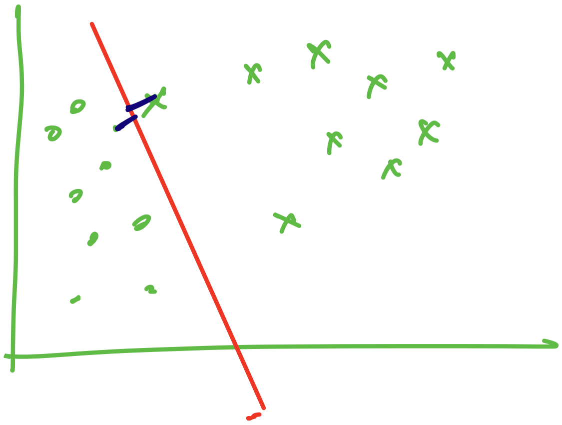
$$\begin{cases} \min \dfrac{\|w\|^2}{2} \xrightarrow{} w^T w \\ \text{such that } \; y^{(i)}\left(w^T x^{(i)} + b\right) \geq 1 \end{cases}$$

Loss:

$$\left[\begin{array}{l} \min \left(\dfrac{w^T w}{2} + C \displaystyle\sum_{i=1}^{m} \varepsilon^{(i)} \right) \\ \text{such that } \; y^{(i)}\left(w^T x^{(i)} + b\right) \geq 1 - \varepsilon^{(i)} \end{array} \right.$$

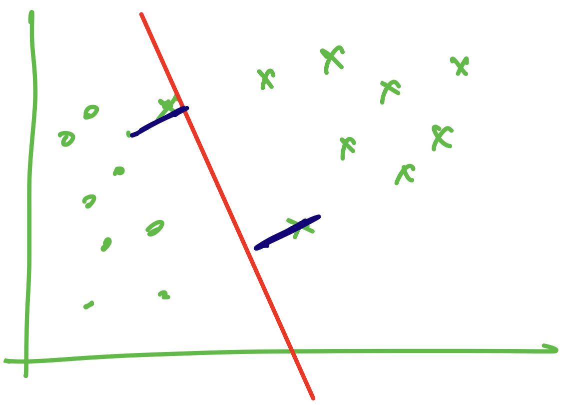
C = hyperparameter

$C = \infty$ No error

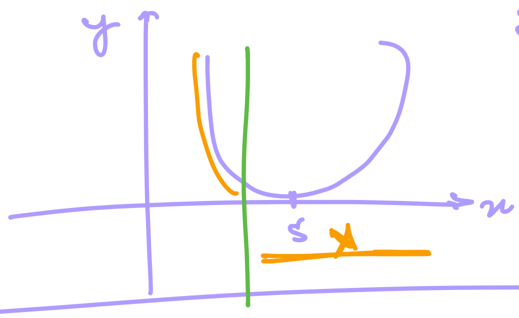


$\min\left(\_\_\_\_ + 1000 \cdot \_\_\_\_ \right)$

$C = 1$ afford some errors, hyperplane maximum margin



$y = (x-5)^2$



value of x for which y is min?

such that $x \le 3$

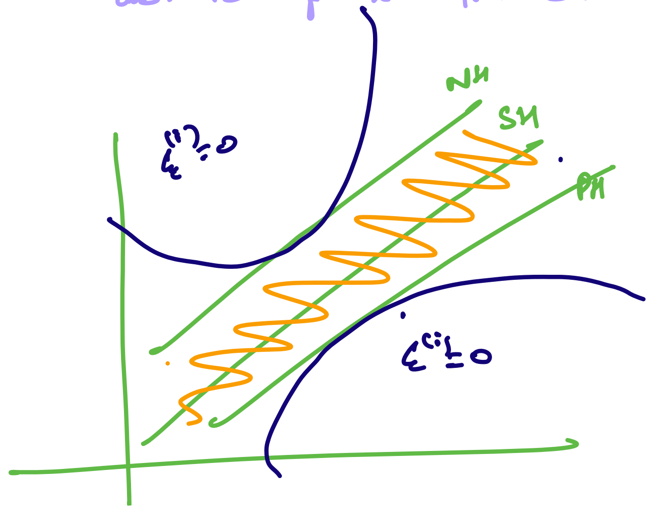
Remove Constraint

$$\min \left(\frac{w^T w}{2} + C \sum_{i=1}^{m} \varepsilon^{(i)} \right)$$

such that $y^{(i)}\left(w^T x^{(i)} + b\right) \ge 1 - \varepsilon^{(i)}$

$$\varepsilon^{(i)} \geq 1 - \underbrace{y^{(i)}\left(w^T x^{(i)} + b\right)}_{t^{(i)}}$$

$t^{(i)}$: unnormalized absolute distance of $x^{(i)}$ from SH



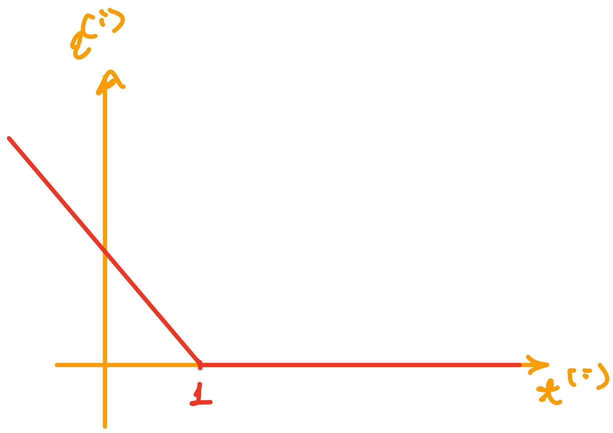
$$\varepsilon^{(i)} \geq 1 - t^{(i)}$$

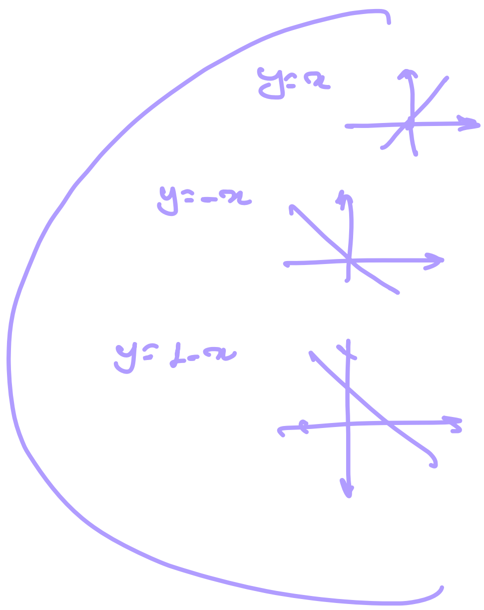
if $t^{(i)} \geq 1$: $\varepsilon^{(i)} = 0$

if $t^{(i)} < 1$: $\varepsilon^{(i)} = 1 - t^{(i)}$

combine

$$\varepsilon^{(i)} = \max\left(0, 1 - t^{(i)}\right)$$





for differentiating $\varepsilon^{(i)}$
concept of subgradient

$$\begin{cases} 0 & t^{(i)} \geq 1 \\ -1 & t^{(i)} < 1 \end{cases}$$

$$\min \left(\frac{w^T w}{2} + C \sum_{i=1}^{m} \varepsilon^{(i)} \right)$$

such that $y^{(i)}\left(w^T x^{(i)} + b\right) \geq 1 - \varepsilon^{(i)}$

$\downarrow$

$$L = \frac{1}{2} w^T \cdot w + c \sum_{i=1}^{m} \max(0, 1 - t^{(i)})$$

where $t^{(i)} = y^{(i)} (w^T x^{(i)} + b)$

SVM objective

Random value of w

How good w is ? $\longrightarrow$ loss

$w_1 x_1 + w_2 x_2 + b = 0$

update w

$$w = w - \eta \frac{\partial L}{\partial w} \quad ?$$

$$b = b - \eta \frac{\partial L}{\partial b}$$

$$\frac{1}{2} w^T \cdot w = \frac{1}{2} (w_1^2 + w_2^2 + w_3^2 - + w_j^2 + w_n^2)$$

$$\frac{\partial}{\partial w_j} \left(\frac{1}{2} w^T w\right) = \frac{1}{2} 2 w_j = w_j$$

$$\frac{\partial L}{\partial w_j} = w_j + \frac{\partial}{\partial w_j} \left(c \sum_{i=1}^{m} \max(0, 1 - t^{(i)}) \right)$$

$$\frac{\partial L}{\partial w_j} = w_j + c \sum_{i=1}^{m} \left(\frac{\partial}{\partial w_j} \underbrace{\max(0, 1 - t^{(i)})}_{f^{(i)}} \right)$$

$$\frac{\partial f}{\partial w_j} = \frac{\partial f}{\partial t} \cdot \frac{\partial t}{\partial w_j}$$

$$\frac{\partial L}{\partial w_j} = w_j + c \sum_{i=1}^{m} \frac{\partial f}{\partial t^{(i)}} \cdot \frac{\partial t^{(i)}}{\partial w_j}$$

$$\frac{\partial L}{\partial w_j} = w_j + c \sum_{i=1}^{m} \frac{\partial}{\partial t^{(i)}} (\max(0, 1 - t^{(i)})) \cdot \frac{\partial t^{(i)}}{\partial w_j}$$

$$\frac{\partial L}{\partial w_j} = w_j + c \sum_{i=1}^{m} \begin{bmatrix} 0 & t^{(i)} \geq 1 \\ -1 & t^{(i)} < 1 \end{bmatrix} \frac{\partial t^{(i)}}{\partial w_j}$$

$$t^{(i)} = y^{(i)}(w^T x^{(i)} + b)$$

$$t^{(i)} = y^{(i)}(w_1 x_1 + w_2 x_2 \cdots + b)$$

$$\frac{\partial t^{(i)}}{\partial w_j} = y^{(i)} x_j^{(i)}$$

$$\boxed{\frac{\partial L}{\partial w_j} = w_j + c \sum_{i=1}^{m} \begin{bmatrix} 0 & t^{(i)} \geq 1 \\ -1 & t^{(i)} < 1 \end{bmatrix} y^{(i)} x_j^{(i)}}$$

$$\frac{\partial L}{\partial b} = 0 + \frac{\partial}{\partial b}\left(c \sum_{i=1}^{m} \underbrace{\max(0, 1 - t^{(i)})}_{f^{(i)}}\right)$$

$$\frac{\partial L}{\partial b} = c \sum_{i=1}^{m} \frac{\partial f^{(i)}}{\partial t^{(i)}} \cdot \frac{\partial t^{(i)}}{\partial b}$$

$$\frac{\partial L}{\partial b} = c \sum_{i=1}^{m} \begin{bmatrix} 0 & \text{if } t^{(i)} \geq 1 \\ -1 & \text{if } t^{(i)} < 1 \end{bmatrix} \frac{\partial t^{(i)}}{\partial b}$$

$$\frac{\partial t^{(i)}}{\partial b} = \frac{\partial}{\partial b}\left(y^{(i)}(w^T x^{(i)} + b)\right)$$

$$= y^{(i)}$$

$$\frac{\partial L}{\partial b} = c \sum_{i=1}^{m} \begin{bmatrix} 0 & \text{if } t^{(i)} \geq 1 \\ -1 & \text{if } t^{(i)} < 1 \end{bmatrix} y^{(i)}$$

UPDATE RULE:

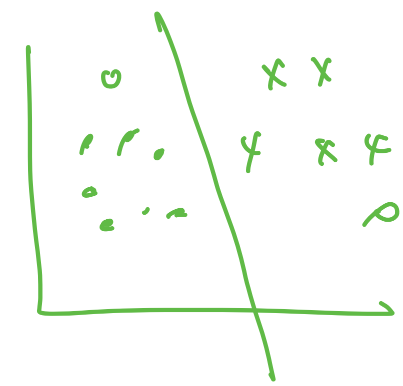
$$W = W - \eta\left(w_j + c\sum_{i=1}^{m}\begin{bmatrix}0 & t^{(i)}\geq 1\\ -1 & t^{(i)}<1\end{bmatrix} y^{(i)}x_j^{(i)}\right)$$

$$\boxed{w_j = w_j - \eta w_j + \sum_{i=1}^{m}\begin{bmatrix}0 & t^{(i)}\geq 1\\ \eta c y^{(i)} x_j^{(i)} & t^{(i)}<1\end{bmatrix}}$$

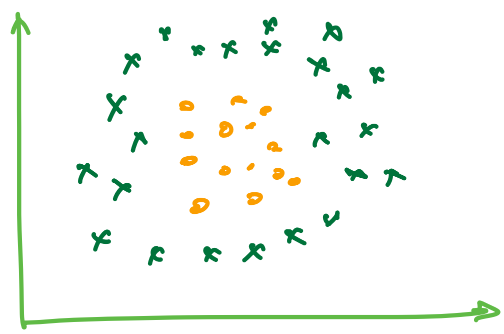
$$b = b - \eta\left(c\sum_{i=1}^{m}\begin{bmatrix}0 & \text{if } t^{(i)}\geq 1\\ -1 & \text{if } t^{(i)}<1\end{bmatrix} y^{(i)}\right)$$

$$\boxed{b = b + \sum_{i=1}^{m}\begin{bmatrix}0 & \text{if } t^{(i)}\geq 1\\ \eta c y^{(i)} & \text{if } t^{(i)}<1\end{bmatrix}}$$

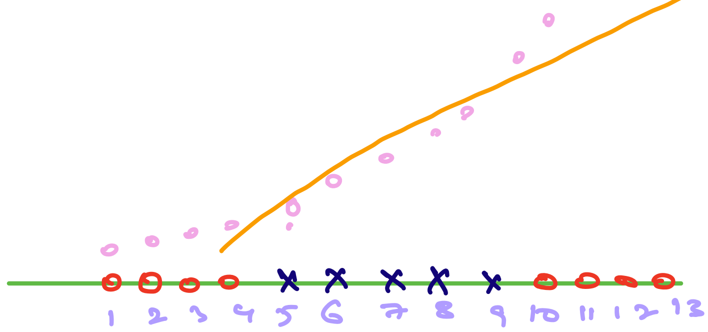
Linear SVM



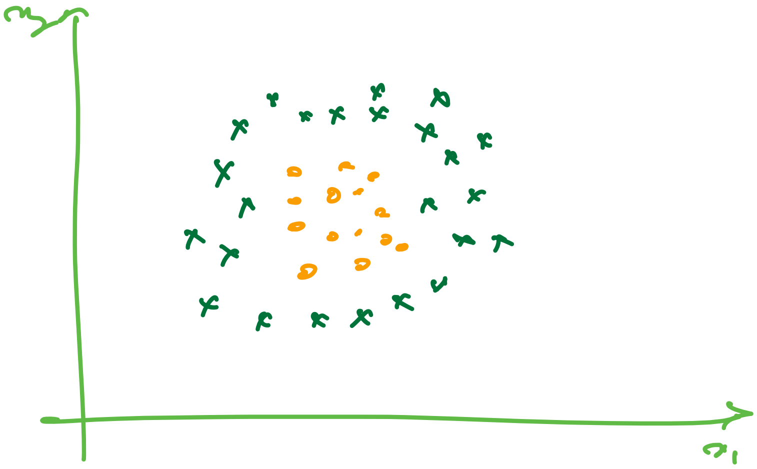
LSD ✓



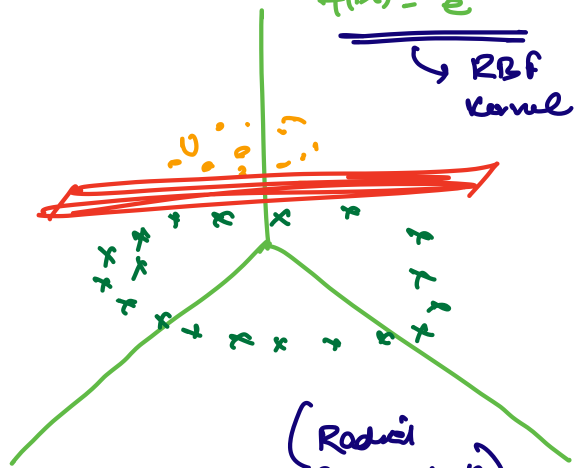
1D



$f(x) = x^2$ → Kernel



$f(x) = e^{-x^2}$ → RBF Kernel



(Radial Basis fxn)

Kernel:

- RBF
- Polynomial
- Sigmoid

$$L = \frac{1}{2} w^T \cdot w + C \sum_{i=1}^{m} \max(0, 1 - t^{(i)})$$

where $t^{(i)} = y^{(i)} (w^T x^{(i)} + b)$

$$x^{(i)} \longrightarrow f(x^{(i)})$$