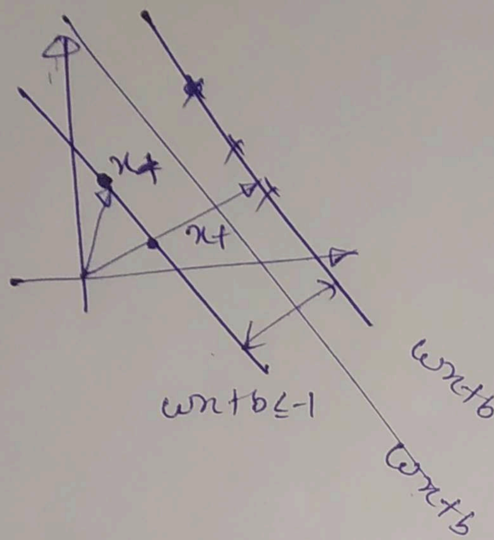
SVM Binary classification

Naveneet Kumar Singh

2020201064



$x \cdot w + b \geq 1$ for $y_i = +1$

$wx + b \leq -1$ for $y_i = -1$

$y_i(wx_i + b) - 1 \geq 0 \;\; \forall i$

width of margin $= (x_+ - x_-) \dfrac{w}{||w||}$

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

$\left(\dfrac{2}{||w||}\right)$ maximize $\Rightarrow$ $||w||$ minimize

for our convenience computation $\Rightarrow$ $\frac{1}{2}||w||^2$

min $\frac{1}{2}||w||^2$ such that $y_i(x_i w + b) - 1 \geq 0 \;\; \forall i$

$\Downarrow$ Lagrangian form

$L = \frac{1}{2}||w||^2 - \sum^n \alpha_i [y_i(wx_i + b) - 1]$ — (1)

$L = \frac{1}{2}||w||^2 - \sum_{i=1}^{n} \alpha_i [y_i(wx_i + b) - 1]$

$L = \frac{1}{2}||w||^2 - \sum_{i=1}^{n} \alpha_i y_i (wx_i + b) + \sum_{i=1}^{n} \alpha_i$

$\dfrac{\partial L_P}{\partial w} = 0 \Rightarrow \frac{1}{\cancel{2}} \cancel{2} ||w|| - \sum_{i=1}^{n} \alpha_i y_i x_i = 0$

$$\boxed{w = \sum_{i=1}^{n} \alpha_i y_i x_i}$$

$$\frac{\partial L_P}{\partial b} = 0 \Rightarrow \quad 0 - 0 - \sum_{i=1}^{n} \alpha_i y_i = 0$$

$$\boxed{\sum \alpha_i y_i = 0}$$

substituting value of w in eqn ①

$$L = \frac{1}{2}\left(\sum \alpha_i y_i x_i\right)\left(\sum \alpha_i y_i x_i\right) - \sum \alpha_i y_i x_i \left(\sum \alpha_j y_j x_j\right) + \sum \alpha_i$$

$$L = \sum \alpha_i - \sum \alpha_i \alpha_j y_i y_j \underbrace{(x_i \cdot x_j)}_{\text{kernel}} \quad (\underline{\text{Linear}})$$

$$L = \sum \alpha_i - \sum \alpha_i \alpha_j y_i y_j (\phi(x_i) \cdot \phi(x_j)) \quad \{\phi = \text{linear, poly, rbf}\}$$

$\Downarrow$

$$L = -\sum \alpha_i \alpha_j y_i y_j x_i x_j + \sum \alpha_i$$

Quadratic Program

minimize

$(1/2) x^T P x + q^T x$

Subject to $Gx \leq h$

$Ax = b$

Now

$K = x_i \cdot x_j$

$P = y_i y_j \cdot K$

so, $P = y_i y_j x_i x_j$

so,

$$L = -\sum \alpha_i P \alpha_j + \sum \alpha_i$$

$\Downarrow$

$$L = (1/2) \alpha^T P \alpha + 1^T \alpha_i$$

$\alpha_i \geq 0$ and $\sum \alpha y = 0$

$\Downarrow$ $\quad Ax = b$

$-\alpha_i \leq 0 \quad A = y$

So here $G = -1$ identity meta

Here $\Rightarrow$ $P = \alpha_i y_j x_i x_j \ (n \times n)$

$q = 1^T \ (n \times 1)$

$G = -1 \ \underline{(n \times n)}$

$$\boxed{\begin{array}{l} \min \left(\frac{1}{2}\alpha^T P \alpha + 1^T \alpha\right) \\ -\alpha \leq 0 \\ \sum \alpha y = 0 \end{array}}$$

$\xrightarrow{\text{Opsolve}}$ we get α.

after that we calculate w

$w = \sum \alpha_i y_i x_i$

then

$a_i > 0$ {from the condition}

if $0 < \alpha \leq C$ {for slack variables}

then compute b -

$b = \frac{1}{N} \sum_{s \in S} (y_s - \sum \alpha_a y_a x_a \cdot x_b)$

Last step from w, b we check the sign of

X_{test} $\quad$ $y_{test} = (w \cdot x\_test + b)$

$y_{pred} = sign(y_{test})$

- y pred contain = {+, -}
  ↳ Binary class

Summary ——— x ———

① Create H matrix, where $H_{ij} = y_i y_j k$
where $k = \langle x_i \cdot x_j \rangle$ {kernel}

② find α. $\sum_{i=1}^{n} \alpha_i - \frac{1}{2} \alpha^T H \alpha$ {maximize}

⇕

$\sum \frac{1}{2} \alpha^T H \alpha - \sum \alpha_i$ {minimize}

$\alpha_i \geq 0$ and $\sum \alpha_i y_i = 0$

this quadratic problem solve by the QP solver
ex. cvxpy (——, ——)

(3) $w = \sum_{i=}^{n} \alpha_i y_i x_i$

(4) from w now calculate b where be is

(5) And determine α_i which is $\alpha_i > 0$

now $b = \frac{1}{N_s} \sum_{s \in S} (y_s - \sum_{t \in S} \alpha_t y_t x_t \cdot x_s)$

where S = Number of support vector.

(6) After computation of w and b we ~~can~~ than classify new test data that are done by sign checking of y test

y_test = sign($w \cdot x' + b$)

y_pos_class = y_test > 0
y_neg_class = y_test < 0

Here we see that both class are computed for test data.

This is for Hard margin

for soft margin only ~~the~~ refine the α value that is

$$\boxed{\begin{array}{l} \alpha_i > 0 \text{ for Hard Margin} \\ C \geq \alpha_i > 0 \text{ for Soft Margin.} \end{array}}$$

for the soft margin the slack variable is constant value after derivation of $\frac{\partial L}{\partial w}$ and $\frac{\partial L}{\partial b}$ and $\frac{\partial L_P}{\partial \xi}$ we get

$$-\sum_{i=1}^{L} \frac{1}{2} \alpha^T H \alpha + \sum \alpha_i \text{ such that } 0 \leq \alpha_i \leq C \ \forall i$$

after that same procedure like Hard margin.