Support Vector Machine (SVM)

May 2, 2020
Guorui Shen
guorui233@outlook.com

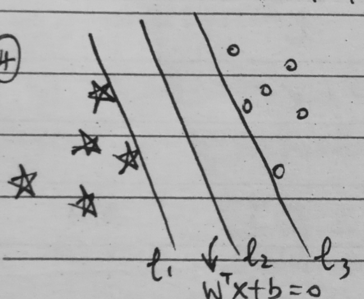
① $\{(x_i, y_i)\}_{i=1}^{N}$, $x_i \in R^n$, $y_i = 1$ or -1

② An affine hyperplane is defined as $w^T x + b = 0$, or $w_1 x_1 + \cdots + w_n x_n + b = 0$, here $\boxed{x_i \in R}$ is different from $x_i \in R^n$.

$x = [x_1, \cdots, x_n]^T \in R^n$ is a n dimensional variable vector.

③ The task of SVM is to find an affine hyperplane $w^T x + b = 0$ such that $w^T x_i + b > 0$ if $y_i = 1$, or $w^T x_i + b < 0$ if $y_i = -1$.

④



$l_1: W^T x + b = k$
$l_2: W^T x + b = 0$
$l_3: W^T x + b = -k$

$\Longleftrightarrow$

$l_1: W_{new}^T x + b_{new} = 1$
$l_2: W_{new}^T x + b_{new} = 0$
$l_3: W_{new}^T x + b_{new} = -1$

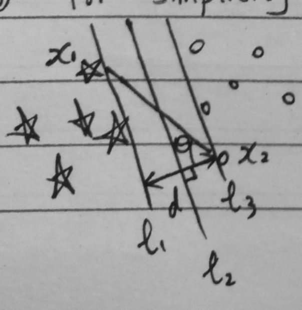
$\boxed{\text{Therefore}}$: Finding $w^T x + b = 0$ is equivalent to finding $W_{new}^T x + b_{new} = 0$

Once we know W_{new} & b_{new}, we can compute $W_{new}^T x_i + b_{new}$ to decide $y_i = 1$ or $y_i = -1$.

⑤ For simplicity, let ~~≡~~ $l_1: W^T x + b = 1$

$l_2: W^T x + b = 0$

$l_3: W^T x + b = -1$



The goal is to maximize the distance between l_1 & l_2 while satisfying $W^T x_i + b \geq 1$ if $y_i = 1$ & $W^T x_i + b \leq -1$ if $y_i = -1$

⑥ In the figure of step ⑤, $\begin{cases} x_1 \text{ is in } \ell_1, \text{ then } w^T x_1 + b = 1 \\ x_2 \text{ is in } \ell_3, \text{ then } w^T x_2 + b = -1 \end{cases}$

hence $w^T(x_1 - x_2) = 2$. Since w & $x_1 - x_2$ represents two n-dimensional vectors, we have $w^T(x_1 - x_2) = \|w\| \cdot \|x_1 - x_2\| \cdot \cos\theta$

therefore $d = \|x_1 - x_2\| \cdot \cos\theta = \dfrac{w^T(x_1 - x_2)}{\|w\|} = \dfrac{2}{\|w\|}$

convex problem:

$$\min \tfrac{1}{2} w^T w \quad \text{s.t. } -y_i(w^T x_i + b) + 1 \le 0, \quad i = 1, \cdots, N$$

⑦ Finally, we have:

$$\max \frac{2}{\|w\|} \quad \text{s.t. } \begin{cases} w^T x_i + b \ge 1 & \text{if } y_i = +1 \\ w^T x_i + b \le -1 & \text{if } y_i = -1 \end{cases}$$

$\Updownarrow$

$$\max_{w,b} \frac{2}{\|w\|} \quad \text{s.t. } y_i(w^T x_i + b) \ge 1, \quad i = 1, 2, \cdots, N$$

$\Longleftrightarrow$

$$\cancel{\max}\ \min \tfrac{1}{2} w^T w \quad \text{s.t. } y_i(w^T x_i + b) \ge 1, \quad i = 1, \cdots, N$$

$\Updownarrow$ (convex problem)

⑧ Let $L(w, b, \alpha) = \frac{1}{2} w^T w - \sum_{i=1}^{N} \alpha_i [y_i(w^T x_i + b) - 1]$, $\alpha_i \ge 0$, $i = 1, \cdots, N$

~~Analysis: If $\exists i$, $y_i(w^T x_i + b) \ge 1$ is violated, then let $\alpha_i = +\infty \Rightarrow L = +\infty$.~~

~~else for all $i = 1, \cdots, N$, $y_i(w^T x_i + b) \ge 1$ holds, then let $\alpha_i = 0$, $\Rightarrow L = \frac{1}{2} w^T w$.~~

~~Conclusion: Solving~~

$$\frac{\partial L}{\partial w} = w - \sum_{i=1}^{N} \alpha_i y_i x_i = 0 \Rightarrow w = \sum_{i=1}^{N} \alpha_i y_i x_i$$

$$\frac{\partial L}{\partial b} = -\sum_{i=1}^{N} \alpha_i y_i = 0 \Rightarrow \sum_{i=1}^{N} \alpha_i y_i = 0$$

Therefore, the dual function of the primal convex problem is

$$g(\alpha) = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j x_i^T x_j - \sum_{i=1}^{N} \alpha_i$$

$\boxed{\max g(\alpha) \quad \sum_{i=1}^{N} \alpha_i y_i = 0, \quad \alpha_i \ge 0, \ i = 1, \cdots, N}$ is a ~~unconstrain~~ quadratic programming, $\Rightarrow \hat{\alpha}$, then

$\hat{w} = \sum_{i=1}^{N} \hat{\alpha}_i y_i x_i$. $\exists j$, such that $y_j(x_j^T \hat{w} + \hat{b}) = 1$

$\Rightarrow \hat{b} = y_j - \hat{w} x_j$