CS19BTECH11020

Assignment - 2

① Original margin boundaries be $w_1 x + b_1 = 1$, $w_1 x + b_1 = -1$

And replaced be $w_2 x + b_2 = \gamma$, $w_2 x + b_2 = -\gamma$

Mid hyperplane - $w_2 x + b_2 = 0$

L_P for replaced boundaries,

$$L_P = \frac{1}{2}\frac{\|w_2\|^2}{\gamma} - \sum \alpha_i [y_i (w_2 x_i + b_2) - \gamma]$$

$$\left(\frac{\partial L_P}{\partial w}\right)_{rep} = \frac{\|w_2\|}{\gamma} - \sum \alpha_i x_i y_i$$

for L_P to be min

$$\Rightarrow \boxed{\frac{\|w_2\|}{\gamma} = \sum \alpha_i x_i y_i} \longrightarrow ①$$

$$\left(\frac{\partial L_P}{\partial b}\right)_{rep} = \sum \alpha_i y_i \Rightarrow \boxed{\sum \alpha_i y_i = 0} - ②$$

As per original conditions $\left(\frac{\partial L_P}{\partial w}\right)_{org} = \|w_1\| - \sum \alpha_i x_i y_i$

$$\Rightarrow \boxed{\|w_1\| = \sum \alpha_i x_i y_i} \to ③ \quad \text{And} \quad \boxed{\sum \alpha_i y_i = 0}$$

Now divide ① & ③ $\Rightarrow \boxed{\|w_2\| = \gamma . \|w_1\|} \to ④$

Given hard margin SVM $\Rightarrow$ No slack variables for SV's

$$y_i (w_1 x + b_1) = 1$$

$$y_i (w_2 x + b_2) = \gamma$$

dividing $\Rightarrow \boxed{\frac{w_1 x + b_1}{w_2 x + b_2} = \frac{1}{\gamma}}$

from ② $\|w_2\| = \gamma \|w_1\|$

$$\Rightarrow \frac{w_1 x_1 + b_1}{\gamma w_1 x_1 + b_2} = \frac{1}{\gamma}$$

$$\Rightarrow \boxed{b_2 = \gamma b_1}$$

by substituting w_2, b_2 values in replaced margins

$$\Rightarrow \quad w_2 x + b_2 = \gamma \qquad w_2 x + b_2 = -\gamma$$

$$\gamma w_1 x + \gamma b_1 = \gamma \qquad \gamma w_1 x + \gamma b_1 = -\gamma$$

$$\Rightarrow \boxed{w_1 x + b_1 = 1 \quad , \quad w_1 x + b_1 = -1}$$

↳ Original margins

Both replaced margins are original margins.
Solution for Maximum margin hyperplane will be unchanged.

(2) Given. $\rho = \frac{1}{\|w\|}$, S.T $\frac{1}{\rho^2} = \sum_{i=1}^{w} \alpha_i$

$W = \sum \alpha_i x_i y_i$

$= \sum \underbrace{y_i \alpha_i}_{\text{scalar}} \underbrace{x_i}_{\text{vector}}$

$w \cdot w = \sum y_i \alpha_i (x_i \cdot w) \longrightarrow (1)$

And, $y_i (x_i \cdot w + b) = 1$

$x_i \cdot w = \frac{1}{y_i} - b \longrightarrow (2)$

from (1) & (2)

$w \cdot w = \sum y_i \alpha_i \left(\frac{1}{y_i} - b \right)$

$= \sum \alpha_i - b \sum y_i \alpha_i$

$w \cdot w = \sum \alpha_i \longleftarrow \boxed{\sum \alpha_i y_i = 0}$

ie $\frac{1}{\rho^2} = \sum_{i=1}^{N} \alpha_i$

Hence proved.

K_1, K_2 are valid kernals

(3) (a) $K(x,z) = K_1(x,z) + K_2(x,z)$

$\because K_1, K_2$ are valid $\Rightarrow K_1, K_2$ are symmetric positive definite kernals

so

$$K_1(x,z) = \phi_1(x) \cdot \phi_1(z)$$
$$K_2(x,z) = \phi_2(x) \cdot \phi_2(z)$$

$$\Rightarrow \phi(x) = (\phi_1(x) \ \phi_2(x))$$
$$\phi(z) = (\phi_1(z), \phi_2(z))$$

$$\phi(x) \cdot \phi(z) = (\phi_1(x) \ \phi_2(x)) \cdot (\phi_1(z) \ \phi_2(z))$$

$$= \phi_1(x) \cdot \phi_1(z) + \phi_2(x) \phi_2(z)$$

$$= K_1(x,z) + K_2(x,z)$$
$$= K(x,z)$$

$K(x,z)$ is expressed as dot pdt of $\phi(x) \cdot \phi(z)$

$\Rightarrow K(x,z) \rightarrow$ valid.

(b) $K(x,z) = K_1(x,z) \cdot K_2(x,z) = (\phi_1(x) \cdot \phi_1(z))(\phi_2(x) \cdot \phi_2(z))$

Using gram matrix K for $K(x,z)$, we can observe that each element of the matrix obtained by element pdt of K_1, K_2.

$\because K_1, K_2$ are SPDs $\Rightarrow K$ is also SPD.

$\Rightarrow K$ is valid kernal

(c) $k(x,z) = h(k_1(x,z))$

h is polynomial with +ve coeff.

$$\Rightarrow \quad k = a(k_1(x,z))^n + b(k_1(x,z))^{n-1} + \cdots$$

k canbe expressed as sum of pdts of diff kernals.

from (a) & (b)

We can say that k is valid kernal.

(d) $k(x,z) - \exp(k_1(x,z))$

from taylor series $e^x = 1 + x + \frac{x^2}{2!} + \cdots$

$$e^{k_1(x,z)} = 1 + k_1(x,z) + \frac{k_1(x,z)^2}{2!} + \frac{k_1(x,z)^3}{3!} + \cdots$$

from (a) & (b) & (c)

This is also summation of valid polynomial f^n kernal

$\therefore k(x,z)$ is valid

e) $k(x,z) = \exp\left(\frac{-\|x-z\|^2}{\sigma^2}\right) = \exp\left(\frac{-(|x|^2 + |z|^2 - 2|x||z|)}{\sigma^2}\right)$

$$= e^{\left(-\frac{|x|^2}{\sigma^2} + -\frac{|z|^2}{\sigma^2} + \frac{2|x||z|}{\sigma^2}\right)}$$

$$= e^{\frac{-|x|^2}{\sigma^2}} \cdot e^{\frac{-|z|^2}{\sigma^2}} \cdot e^{\frac{2|x||z|}{\sigma^2}}$$

from (d) both are valid

↳ from (b) (d) it's valid

$\Rightarrow k(x,z)$ is valid