FoML – Assignment 2

Naitik Malav

CS19BTECH11026

Ans 1. Let $w_1^T x + b_1 = 0$ be the maximum margin hyperplane.

so $w_1^T x + b_1 = Y$ and $w_1^T x + b_1 = -Y$ are the boundaries.

The margin is →

$$(+1)\frac{w_1^T x_+ + b_1}{||w_1||} + \frac{(-1)\, w^T x_- + b}{||w_1||}$$

$$= \frac{2Y}{||w_1||}$$

$\therefore \frac{2Y}{||w_1||}$ should be maximized and $\forall \{(x_i, y_i)\}$

→ $\therefore \frac{||w_1||}{2Y}$ is minimized.

constraint → $\left[\begin{array}{l} y_i(w_1^T x_i + b_1) \geq Y \\ \text{and } \frac{||w_1||^2}{2Y} \text{ is minimized} \end{array}\right]$

$\therefore$ Using Lagrange Multipliers we write:

Date:__.__.__

$$L(x) = \min_{\vec{w}_1, b_1} \max_{\vec{\alpha} \geq 0} \frac{\|w\|^2}{2Y} - \sum_j \alpha_j \left((\vec{w}_1 \cdot \vec{x}_j + b_1) y_i - Y \right)$$

or by using slator's cond^n.

$$D(x) = \max_{\vec{\alpha} \geq 0} \min_{\vec{w}_1, b_1} \frac{1}{2} \|\vec{w}\|^2 - \sum_j \alpha_j \left((\vec{w}_1 \cdot \vec{x}_j + b_1) y_j - Y \right)$$

for optimal w_1, b_1

$$\frac{\partial D}{\partial w_1} = \frac{w_1}{Y} - \sum_j \alpha_j y_j x_j = 0$$

$$\Rightarrow \boxed{w_1 = Y \cdot \sum \alpha_j y_j x_j} \quad -(1)$$

$$\frac{\partial D}{\partial b_1} = 0 \Rightarrow -\sum_j \alpha_j y_j = 0$$

$$\Rightarrow \boxed{\sum_j \alpha_j y_j = 0} \quad -(2)$$

apsara

Date:

Eq^n for Suppor Vectors :-

→ $y_i(x_i \cdot w_1 + b_1) - Y = 0$

$b_1 y_i = Y - y_i(x_i \cdot w_1)$

$\therefore \boxed{b_1 = \frac{Y - y_i(x_i \cdot w_1)}{y_i}}$ — (3)

and $\boxed{= \frac{Y - y_i(x_i \cdot Y \sum_j \alpha_j y_j x_j)}{y_i}}$

$\Rightarrow$

$\therefore \quad w_1^T x + b_1 = 0$

$Y w^T x + Y b = 0$

$\Rightarrow \quad \boxed{w^T x + b = 0}$

Hence, the new eq^n is same as the previous one, which mean that eq^n doesn't change.

apsara

Date: __.__.__

Ans 2. Given : $p = \frac{1}{\|w\|}$

$$\frac{1}{p^2} = \|w\|^2 \Rightarrow w \cdot w$$

also we know that

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

$$w \cdot w = \sum_{i=1}^{N} \alpha_i y_i (x_i \cdot w) \quad \text{--- (1)}$$

→ For support vectors → we know that →

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

$$x_i \cdot w = \frac{1}{y_i} - b$$

$$\Rightarrow w \cdot w = \sum_{i=1}^{N} \alpha_i y_i (x_i \cdot w)$$

$$\Rightarrow w \cdot w = \underbrace{\sum \alpha_i y_i (x_i \cdot w)}_{\text{Support vectors}} + \underbrace{\cancel{\sum \alpha_i y_i (x_i \cdot w)}}_{\text{Non-support vectors}}$$

Here $\alpha_i = 0$

apsara

Date:

$$w \cdot w = \sum_{\substack{\text{support} \\ \text{vec.}}} \alpha_i y_i \left(\frac{1}{y_i} - b \right)$$

$$= \sum_{s \cdot v} (\alpha_i - \alpha_i y_i b)$$

$$= \sum_{sv} \alpha_i - b \underbrace{\sum \alpha_i y_i}_{0}$$

Hence $\boxed{w \cdot w = \sum_{s \cdot vec} \alpha_i = \sum_{i=1}^{N} \alpha_i}$ as

Page No.:

Scanned with CamScanner

apsara

Date: __.__.__

Ans 3. (a) To prove →

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

Proof :- $K_1(x,z) = \langle \phi_1(x), \phi_1(z) \rangle$ — ①

and $K_2(x,z) = \langle \phi_2(x), \phi_2(z) \rangle$ — ②

putting eqⁿ ① & ② in the above given eqⁿ.

$$K(x,z) = \langle \phi_1(x), \phi_1(z) \rangle + \langle \phi_2(x), \phi_2(z) \rangle$$

$$= \langle [\phi_1(x)\ \phi_2(x)], [\phi_1(z)\ \phi_2(z)] \rangle \quad — ④$$

→ ~~Since~~ From eqⁿ ④ we can see that $K(x,z)$ can be ~~ep~~ expressed as an inner product.

apsara

Date: __.__.____

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

Proof → Gram matrix K for K is the element-by-element product (i.e. Hadamard Product) of K_1 & K_2.

Suppose that K_1 & K_2 are covariance matrices of $(X_1 \cdots X_n)$ and $(Y_1 \cdots Y_n)$.

Then K is simply the covariance matrix of $(X_1Y_1, \cdots X_nY_n)$ implying that it is symmetric and positive definite.

~~(c) $K(x,z) = h(K_1(x,z))$~~

~~Proof: Because Each polynomial term is a product of kernels with a +ve coeff. (The proof~~

(c) $K(x,z) = h(K_1(x,z))$

→ ∵ each polynomial term is a product of kernels with a positive coeff. ∴ we can easily prove it using part (a) and (b).

Hence Proved.

(d) $K(x,z) = \exp\left(K_1(x,z)\right)$

Proof :- we know that

$$\exp(x) = \lim_{n\to\infty}\left(1 + x + \cdots \frac{x^n}{n!}\right)$$

Hence →

$$K(x,z) = \lim_{n\to\infty}\left(K_n(x,z)\right)$$

(By using part (c) property)

(e) ~~Free Proof~~ ~~f can be transformed in the same domain~~

~~(f). $K(x,z) = E B$~~

apsara

Date:

(c) $K(x,z) = \exp\left(\frac{-\|x-z\|_2^2}{\sigma^2}\right)$

Proof :- $K(x,z) = \exp\left(\frac{-\|x-z\|^2}{\sigma^2}\right)$

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

$\Rightarrow \exp\left(\frac{-\|x\|^2}{\sigma^2}\right) \cdot \exp\left(\frac{-\|z\|^2}{\sigma^2}\right) \cdot \exp\left(\frac{2x'z}{\sigma^2}\right)$

$= f(x)\, f(z) \exp(K_1(x,z))$ — (1)

Now we can express the gram matrix K as the outer product of vector $r = [f(x_1) \ldots f(x_n)]'$.

Hence K is symmetric and +ve semi definite.

$\therefore$ $f(x)\, f(z)$ is a kernel

Page No.:

$\therefore$ in eqn 1 $\longrightarrow$

$$K(x, z) = \underbrace{f(x)\, f(z)}_{\text{kernel}} \quad \underbrace{\exp(k_1(x, z))}_{\text{kernel}}$$

$\therefore$ Product of 2 kernel is also a valid kernel.