Machine Learning from Data -IDC
HW5–Theory+ SVM

227367455 and 323081950

June 19, 2021

Question 1)

a)

Let K, L be two kernels (operating on the same space V) and let α, β be two positive scalars

Prove that $\alpha K + \beta L$ is a kernel.

Let $K(x, x') = \langle \phi_1(x), \phi_1(x') \rangle$
Let $L(x, x') = \langle \phi_2(x), \phi_2(x') \rangle$
Let $M(x, x') = \langle \phi(x), \phi(x') \rangle$

Define the mapping $\phi(x) = (\sqrt{\alpha} \cdot \phi_1(x), \sqrt{\beta} \cdot \phi_2(x'))$
$M(x, x') = \langle (\sqrt{\alpha} \cdot \phi_1(x), \sqrt{\beta} \cdot \phi_2(x')), (\sqrt{\alpha} \cdot \phi_1(x), \sqrt{\beta} \cdot \phi_2(x')) \rangle$
$= \langle \sqrt{\alpha} \cdot \phi_1(x), \sqrt{\alpha} \cdot \phi_1(x') \rangle + \langle \sqrt{\beta} \cdot \phi_2(x), \sqrt{\beta} \cdot \phi_2(x') \rangle$
$= \alpha \cdot \langle \cdot \phi_1(x), \cdot \phi_1(x') \rangle + \beta \cdot \langle \cdot \phi_2(x), \cdot \phi_2(x') \rangle$

$$= \alpha \cdot K(x, x') + \beta \cdot L(x, x')$$

b)

Provide (two different) examples of non-zero kernels K, L (operating on the same space), so that:

i. $K - L$ is a kernel

Let $K(x,y) = \varphi_1(x) \cdot \varphi_1(y)$ and $L(x,y) = \varphi_2(x) \cdot \varphi_2(y)$

Define $\varphi_1(x) = (2x, -4x^2)$ and $\varphi_2(x) = (x, x^2)$ Notice that $K(x,y) - L(x,y) = \varphi_1(x) \cdot \varphi_1(y) - \varphi_2(x) \cdot \varphi_2(y)$

$= (2x, -4x^2) \cdot (2y, -4y^2) - (x, x^2) \cdot (y, y^2)$

$= 2x \cdot 2y + (-4x^2) \cdot (-4y^2) - x \cdot y - x^2 \cdot y^2$

$= 4(x \cdot y) + 16(x^2 \cdot y^2) - x \cdot y - x^2 \cdot y^2$

$= 3(x \cdot y) + 15(x^2 \cdot y^2)$

$= 3x \cdot 3y + 15x^2 \cdot 15y^2$

This is a kernel with the mapping $\varphi(x) = (3x, 15x)$.

ii. $K - L$ is not a kernel

Let $K(x,y) = \varphi_1(x) \cdot \varphi_1(y)$ and $L(x,y) = \varphi_2(x) \cdot \varphi_2(y)$

Define $\varphi_1(x) = (x^4, x^2)$ and $\varphi_2(x) = (2x^4, -4x^2)$

Notice that $K(x,y) - L(x,y) = \varphi_1(x) \cdot \varphi_1(y) - \varphi_2(x) \cdot \varphi_2(y)$

$$= (x^4, x^2) \cdot (y^4, y^2) - (2x^4, -4x^2) \cdot (2y^4, -4y^2)$$

$= x^4 \cdot y^4 + x^2 \cdot y^2 - 2x^4 \cdot 2y^4 - (-4x^2) \cdot (-4y^2)$

$= x^4 \cdot y^4 + x^2 \cdot y^2 - 2(x^4 \cdot y^4) - 4(x^2 \cdot y^2)$

$= -(x^4 \cdot y^4) - 3(x^2 \cdot y^2)$

Notice that this is not a kernel. Proof:

ATC that this is a kernel. That means there exists a mapping $\varphi(x)$ s.t. $-(x^4 \cdot y^4) - 3(x^2 \cdot y^2) = \varphi(x) \cdot \varphi(y)$

Notice that the left side is always negative.

However, by the basic properties of an inner product, then $\varphi(x) \cdot \varphi(x) \geq 0$ for any x.

This means we have a norm that is negative, which is a contradiction!

Question 2)

Use Lagrange Multipliers to find the maximum and minimum values of the function subject to the given constraints:

Function: $f(x, y, z,) = x^2 + y^2 + z^2$.

Constraint: $g(x, y, z) = \frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} + \frac{z^2}{\beta^2} = 1$, where $\alpha > \beta > 0$

$$\bigtriangledown f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{bmatrix} = 2 \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$\bigtriangledown g = \begin{bmatrix} \frac{\partial g}{\partial x} \\ \frac{\partial g}{\partial y} \\ \frac{\partial g}{\partial z} \end{bmatrix} = 2 \begin{bmatrix} \frac{x}{\alpha^2} \\ \frac{y}{\beta^2} \\ \frac{z}{\beta^2} \end{bmatrix}$$

$$\bigtriangledown f = \lambda \bigtriangledown g$$

$$\iff \begin{bmatrix} x \\ y \\ z \end{bmatrix} - \lambda \begin{bmatrix} \frac{x}{\alpha^2} \\ \frac{y}{\beta^2} \\ \frac{z}{\beta^2} \end{bmatrix} = 0$$

$$\iff \begin{bmatrix} x(1 - \frac{\lambda}{\alpha^2}) \\ y(1 - \frac{\lambda}{\beta^2}) \\ z(1 - \frac{\lambda}{\beta^2}) \end{bmatrix} = 0$$

Notice that if $x \neq 0 \Rightarrow \lambda = \alpha^2$ and if $y \neq 0 \bigvee z \neq 0 \Rightarrow \lambda = \beta^2$. Therefore: $\alpha = \beta$ contradicting our constraint.

Consider $x \neq 0$:

$$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} + \frac{z^2}{\beta^2} = \frac{x^2}{\alpha^2} = 1 \iff x = \pm\alpha$$

Consider $x = 0$:

$$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} + \frac{z^2}{\beta^2} = \frac{y^2}{\beta^2} + \frac{z^2}{\beta^2} = 1$$

$$y^2 + z^2 = \beta^2$$

Notice: $\alpha > \beta > 0 \Rightarrow \alpha^2 > \beta^2$

Therefore we can define the maxima $f(\pm\alpha, 0, 0) = \alpha^2$

And minima $f(0, y, z) = y^2 + z^2 = \beta^2$
(Must satisfy the set of solution $y^2 + z^2 = \beta^2$ eg. $x = y = 0, z = \pm\beta$)

Question 3)

Let $X = \mathbb{R}^2$. Let $C = H = \{h(a,b,c) = \{(x,y,z) \mid |x| \leq a, |y| \leq b, |z| \leq c\} \mid a,b,c \in \mathbb{R}^+\}$ the set of all origin centered boxes.

Describe the polynomial sample complexity algorithm L that learns C using H. State the time complexity and sample complexity of your suggested algorithm. Prove all your steps.

Let there be m points in our training set and define all the points as such:

$$p_i = (x_i, y_i, z_i), \forall 0 \geq i \geq m$$

Define

$$\pi(\{D \in X^m \mid Err(L(D), c) > \varepsilon\}) \leq \sum_{i=1}^{6} (\pi(X - B_i))^m$$

We get by union bound:

$$\leq 6(1 - \frac{\varepsilon}{6})$$

$$\leq 6e^{\frac{m\varepsilon}{6}}$$