Machine Learning from Data -IDC
HW5–Theory+ SVM

227367455 and 323081950

June 11, 2021

Question 1)

a)

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

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

Let $\langle x, y \rangle_1 = \alpha K(x, y) + \beta L(x, y)$
We will show that $\langle x, y \rangle_1$ is a inner product by proving the four properties required to be an inner product operation.

Symmetry:
$\langle x, y \rangle_1 = \alpha K(x, y) + \beta L(x, y)$
$= \alpha K(y, x) + \beta L(y, x)$ by symmetry of inner products of K,L
$= \langle y, x \rangle_1$

Additivity:
$\langle u_1 + u_2, v \rangle_1 = \alpha K(u_1 + u_2, v) + \beta L(u_1 + u_2, v)$

$= \alpha(K(u_1, v) + K(u_2, v)) + \beta(L(u_1, v) + L(u_2, v))$ as K,L are inner products
$= (\alpha K(u_1, v) + \beta L(u_1, v)) + (\alpha K(u_2, v) + \beta L(u_2, v))$
$= \langle u_1, v\rangle_1 + \langle u_2, v\rangle_1$

Homogeneity

$\langle c \cdot u, v\rangle_1 = \alpha K(c \cdot u, v) + \beta L(c \cdot u, v)$
$= \alpha \cdot c \cdot K(u, v) + \beta \cdot c \cdot L(u, v)$ as K,L are inner products
$= c \cdot (\alpha K(u, v) + \beta L(u, v)) = c \cdot \langle u, v\rangle_1$

Positivity

$\langle u, u\rangle_1 = \alpha K(u, u) + \beta L(u, u)$
As K, L are inner products, then $K(u, u) \geq 0$ and $L(u, u) \geq 0$
As $\alpha, \beta > 0$ by the given assumption, then we can conclude that $\langle u, u\rangle_1 \geq 0$ as the right side is a sum and product of positive numbers

Now to show that $\langle u, u\rangle_1 = 0$ iff $u = 0$.
First assume that $\langle u, u\rangle_1 = 0$.
Then, as $\alpha, \beta > 0$, the only way this can occur are if $K(u, u) = L(u, u) = 0$.
As K, L are inner products, this means that $u = 0$.
Now assume that $u = 0$.
Notice that $\langle u, u\rangle_1 = \alpha K(u, u) + \beta L(u, u)$.
As K, L are inner products:
$\langle u, u\rangle_1 = \alpha \cdot 0 + \beta \cdot 0$
$= 0$ which proves the other direction.

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 $f(0, x, z) = y^2 + z^2 = \beta^2$

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.