KERNEL METHODS IN M.L.

Homework 2 - by DAVID SOTO

Exercise 1: Let K_1 and K_2 be two positive definite kernels on a set $\mathcal{X}$ with corresponding RKHSs $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, and two positive scalars α and β.

1. Let's prove that $\alpha K_1 + \beta K_2$ is a positive definite kernel.

* First of all, by definition of K_1 and K_2 being p.d. kernels, we have that K_1 and K_2 are symmetric.
Therefore we have: $\forall (x,y) \in \mathcal{X}^2$:

$$(\alpha K_1 + \beta K_2)(x,y) = \alpha K_1(x,y) + \beta K_2(x,y)$$
$$= \alpha K_1(y,x) + \beta K_2(y,x)$$

$$(\alpha K_1 + \beta K_2)(x,y) = (\alpha K_1 + \beta K_2)(y,x)$$

$\Longrightarrow$ $\alpha K_1 + \beta K_2$ is symmetric.

* By definition of K_1 and K_2 being 2 p.d. kernels, we have that $\forall N \in \mathbb{N}$, $\forall (x_1, \dots, x_N) \in \mathcal{X}^N$, $\forall (a_1, \dots, a_N) \in \mathbb{R}^N$:

$$\begin{cases} \sum_{i=1}^{N} \sum_{j=1}^{N} a_i a_j K_1(x_i, x_j) \geq 0 \\ \sum_{i=1}^{N} \sum_{j=1}^{N} a_i a_j K_2(x_i, x_j) \geq 0 \end{cases}$$

Let us recall that $\alpha \geq 0$ and $\beta \geq 0$.

$\rightarrow$ Let $N \in \mathbb{N}$, let $(x_1, \dots, x_N) \in \mathcal{X}^N$ and let $(a_1, \dots, a_N) \in \mathbb{R}^N$: we have:

$$A = \sum_{i=1}^{N} \sum_{j=1}^{N} a_i a_j (\alpha K_1 + \beta K_2)(x_i, x_j)$$

$$A = \sum_{i=1}^{N} \sum_{j=1}^{N} a_i a_j \left[\alpha K_1(x_i, x_j) + \beta K_2(x_i, x_j) \right]$$

$$A = \sum_{i=1}^{N}\sum_{j=1}^{N}\left[\alpha\, a_i a_j K_1(x_i,x_j) + \beta\, a_i a_j K_2(x_i,x_j)\right]$$

$$A = \sum_{i=1}^{N}\sum_{j=1}^{N} \alpha\, a_i a_j K_1(x_i,x_j) + \sum_{i=1}^{N}\sum_{j=1}^{N} \beta\, a_i a_j K_2(x_i,x_j)$$

$$A = \alpha \underbrace{\sum_{i=1}^{N}\sum_{j=1}^{N} a_i a_j K_1(x_i,x_j)}_{\substack{\geqslant 0 \\ \text{by definition of } K_1}} + \beta \underbrace{\sum_{i=1}^{N}\sum_{j=1}^{N} a_i a_j K_2(x_i,x_j)}_{\substack{\geqslant 0 \\ \text{by definition of } K_2}}$$

And since $\alpha \geqslant 0$ and $\beta \geqslant 0$, we have

$A \geqslant 0$, that is :

$$A = \sum_{i=1}^{N}\sum_{j=1}^{N} a_i a_j \, (\alpha K_1 + \beta K_2)(x_i, x_j) \geqslant 0$$

Hence, we can conclude that $\alpha K_1 + \beta K_2$ is a positive definite kernel

* Now, we are going to describe the RKHS of $\alpha K_1 + \beta K_2$.

Since K_1 and K_2 are 2 p.d. kernels with corresponding RKHSs $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, and since $\alpha \geq 0$ and $\beta \geq 0$, we can deduce from the "combining kernels" theorem of the course that αK_1 and βK_2 are p.d. kernels.

It is quite patent that the RKHS of αK_1 is $(\mathcal{H}_1, \frac{1}{\alpha}\langle .,.\rangle_{\mathcal{H}_1})$ and that the RKHS of βK_2 is $(\mathcal{H}_2, \frac{1}{\beta}\langle .,.\rangle_{\mathcal{H}_2})$.

↳ Indeed, we can easily prove that:

① $\forall x \in \mathcal{X}, \quad \alpha K_{1x} \in (\mathcal{H}_1, \frac{1}{\alpha}\langle .,.\rangle_{\mathcal{H}_1})$

② $\forall x \in \mathcal{X}, \ \forall f \in \mathcal{H}_1, \quad \frac{1}{\alpha}\langle f, \alpha K_{1x}\rangle_{\mathcal{H}_1} = \langle f, K_{1x}\rangle_{\mathcal{H}_1} = f(x)$

i.e. $f(x) = \frac{1}{\alpha}\langle f, \alpha K_{1x}\rangle_{\mathcal{H}_1}$

(Same proof for βK_2)

Hence, we have that αK_1 is a p.d. kernel with $(\mathcal{H}_1, \frac{1}{\alpha}\langle .,.\rangle_{\mathcal{H}_1})$ as RKHS, and βK_2 is a p.d. kernel with $(\mathcal{H}_2, \frac{1}{\beta}\langle .,.\rangle_{\mathcal{H}_2})$ as RKHS.

* Let $\mathcal{F}$ be the Hilbert space such that $\mathcal{F} := \mathcal{H}_1 \times \mathcal{H}_2$, endowed with the inner product: $\forall (f_1, f_2), (g_1, g_2) \in (\mathcal{H}_1 \times \mathcal{H}_2)^2$

$$\langle (f_1, f_2), (g_1, g_2)\rangle_{\mathcal{F}} := \frac{1}{\alpha}\langle f_1, g_1\rangle_{\mathcal{H}_1} + \frac{1}{\beta}\langle f_2, g_2\rangle_{\mathcal{H}_2}.$$

Let $u : \mathcal{F} \longrightarrow \mathcal{H}_1 + \mathcal{H}_2$, $(f_1, f_2) \longmapsto f_1 + f_2$. u is clearly linear and surjective. Moreover, $\ker(u)$ is clearly closed.

Therefore, we can deduce that $\ker(u)^\perp$ is in direct sum with $\ker(u)$.

, i.e $\mathcal{F} = \ker(u) \oplus \ker(u)^{\perp}$.

We define $v := u|_{\ker(u)^{\perp}}$. A standard result of linear algebra shows that v is bijective.

$\Rightarrow$ v^{-1} is bijective

$\Rightarrow$ v^{-1} is a linear isomorphism

(N.B. $\forall f \in \mathcal{H}_1 + \mathcal{H}_2$, $\exists! (f_1, f_2) \in \mathcal{H}_1 \times \mathcal{H}_2$ such that $v^{-1}(f) = (f_1, f_2)$.)

If we define the following inner product on $\mathcal{H}_1 + \mathcal{H}_2$:

$\forall f, g \in (\mathcal{H}_1 + \mathcal{H}_2)^2$, $\langle f, g \rangle_{\mathcal{H}_1 + \mathcal{H}_2} := \langle v^{-1}(f), v^{-1}(g) \rangle_{\mathcal{F}}$

We have that, since v^{-1} is an isomorphisme,

$(\mathcal{H}_1 + \mathcal{H}_2, \langle .,. \rangle_{\mathcal{H}_1 + \mathcal{H}_2})$ is a Hilbert space.

* Let us prove that $(\mathcal{H}_1 + \mathcal{H}_2, \langle .,. \rangle_{\mathcal{H}_1 + \mathcal{H}_2})$ is the RKHS of $\alpha K_1 + \beta K_2$.

1. First of all, since $\forall x \in \mathcal{X}$, $\begin{cases} \alpha K_{1x} \in \mathcal{H}_1 \\ \beta K_{2x} \in \mathcal{H}_2 \end{cases}$

$\Rightarrow$ $(\alpha K_1 + \beta K_2)_x = \alpha K_{1x} + \beta K_{2x} \in \mathcal{H}_1 + \mathcal{H}_2$

2. Let $x \in \mathcal{X}$, let $f \in \mathcal{H}_1 + \mathcal{H}_2$ $\Rightarrow$ $\exists f_1 \in \mathcal{H}_1$ and $\exists f_2 \in \mathcal{H}_2$ such that $f = f_1 + f_2$. We have:

$$\langle f, (\alpha K_1 + \beta K_2)_x \rangle_{\mathcal{H}_1 + \mathcal{H}_2} = \langle f_1 + f_2, \alpha K_{1x} + \beta K_{2x} \rangle_{\mathcal{H}_1 + \mathcal{H}_2}$$

$$= \langle v^{-1}(f_1 + f_2), v^{-1}(\alpha K_{1x} + \beta K_{2x}) \rangle_{\mathcal{F}}$$

$$= \langle (f_1, f_2), (\alpha K_{1x}, \beta K_{2x}) \rangle_{\mathcal{F}}$$

(1)

$$\langle f, (\alpha K_1 + \beta K_2)_x \rangle_{\mathcal{H}_1 + \mathcal{H}_2} = \frac{1}{\alpha} \langle f_1, \alpha K_{1x} \rangle_{\mathcal{H}_1} + \frac{1}{\beta} \langle f_2, \beta K_{2x} \rangle_{\mathcal{H}_2}$$

$$= \langle f_1, K_{1x} \rangle_{\mathcal{H}_1} + \langle f_2, K_{2x} \rangle_{\mathcal{H}_2}$$

$$= f_1(x) + f_2(x)$$

$$= (f_1 + f_2)(x)$$

$$\langle f, (\alpha K_1 + \beta K_2)_x \rangle_{\mathcal{H}_1 + \mathcal{H}_2} = f(x)$$ The reproducing property holds!

Hence, we can conclude that $(\mathcal{H}_1 + \mathcal{H}_2, \langle \cdot, \cdot \rangle_{\mathcal{H}_1 + \mathcal{H}_2})$ is the RKHS of $\alpha K_1 + \beta K_2$.

2. We have shown that $\mathcal{H} = (\mathcal{H}_1 + \mathcal{H}_2, \langle .,. \rangle_{\mathcal{H}_1+\mathcal{H}_2})$ is the RKHS of $\alpha K_1 + \beta K_2$.

Now let us express the norm of the RKHS $(\mathcal{H}_1 + \mathcal{H}_2, \langle .,. \rangle_{\mathcal{H}_1+\mathcal{H}_2})$ in terms of the norms of both RKHS $(\mathcal{H}_1, \langle .,. \rangle_{\mathcal{H}_1})$ and $(\mathcal{H}_2, \langle .,. \rangle_{\mathcal{H}_2})$

Let $f \in \mathcal{H}_1 + \mathcal{H}_2$, then $\exists (f_1, f_2) \in \mathcal{H}_1 \times \mathcal{H}_2$ such that $f = f_1 + f_2$.

Using the same inner products and functions introduced in question 1, we have:

$$
\begin{aligned}
\|f\|^2_{\mathcal{H}_1+\mathcal{H}_2} &= \langle f, f \rangle_{\mathcal{H}_1+\mathcal{H}_2} \\
&= \langle f_1 + f_2, f_1 + f_2 \rangle_{\mathcal{H}_1+\mathcal{H}_2} \\
&= \langle v^{-1}(f_1+f_2), v^{-1}(f_1+f_2) \rangle_{\mathcal{F}} \\
&= \langle (f_1, f_2), (f_1, f_2) \rangle_{\mathcal{F}} \\
&= \frac{1}{\alpha} \langle f_1, f_1 \rangle_{\mathcal{H}_1} + \frac{1}{\beta} \langle f_2, f_2 \rangle_{\mathcal{H}_2}
\end{aligned}
$$

$$\|f\|^2_{\mathcal{H}_1+\mathcal{H}_2} = \frac{1}{\alpha} \|f_1\|^2_{\mathcal{H}_1} + \frac{1}{\beta} \|f_2\|^2_{\mathcal{H}_2}$$

$\Longrightarrow \quad \forall f = f_1 + f_2 \in \mathcal{H}_1 + \mathcal{H}_2 :$

$$\boxed{\|f\|_{\mathcal{H}_1+\mathcal{H}_2} = \left(\frac{1}{\alpha} \|f_1\|^2_{\mathcal{H}_1} + \frac{1}{\beta} \|f_2\|^2_{\mathcal{H}_2} \right)^{1/2}}$$

(2)

Exercise 2: Let $\mathcal{X}$ be a set and $\mathcal{F}$ be a Hilbert space.
Let $\Psi : \mathcal{X} \longrightarrow \mathcal{F}$, and $K : \mathcal{X} \times \mathcal{X} \longrightarrow \mathbb{R}$ be:
$\forall (x,y) \in \mathcal{X}^2, \quad K(x,y) = \langle \Psi(x), \Psi(y) \rangle_{\mathcal{F}}$

1. Let us prove that K is a positive definite kernel on $\mathcal{X}$.

* By definition, the inner-product in $\mathcal{F}$, $\langle .,. \rangle_{\mathcal{F}}$, is symmetric, therefore we have: $\forall (x,y) \in \mathcal{X}^2$

$$K(x,y) = \langle \Psi(x), \Psi(y) \rangle_{\mathcal{F}}$$
$$= \langle \Psi(y), \Psi(x) \rangle_{\mathcal{F}}$$
$$K(x,y) = K(y,x)$$

$\Longrightarrow$ K is symmetric

* Let $N \in \mathbb{N}$, let $(x_1, \dots, x_N) \in \mathcal{X}^N$, and let $(a_1, \dots, a_N) \in \mathbb{R}^N$: we have:

$$A = \sum_{i=1}^{N} \sum_{j=1}^{N} a_i a_j \, K(x_i, x_j)$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{N} a_i a_j \langle \Psi(x_i), \Psi(x_j) \rangle_{\mathcal{F}}$$, using the bilinearity of $\mathcal{F}$ we have:

$$= \sum_{i=1}^{N} \sum_{j=1}^{N} \langle a_i \Psi(x_i), a_j \Psi(x_j) \rangle_{\mathcal{F}}$$

$$= \left\langle \sum_{i=1}^{N} a_i \Psi(x_i), \sum_{j=1}^{N} a_j \Psi(x_j) \right\rangle_{\mathcal{F}}$$

$$A = \left\| \sum_{i=1}^{N} a_i \Psi(x_i) \right\|_{\mathcal{F}}^2 \geqslant 0$$ (where $\|.\|_{\mathcal{F}}$ is the norm of $\mathcal{F}$)

That is:

$$A = \sum_{i=1}^{N} \sum_{j=1}^{N} a_i a_j \, K(x_i, x_j) \geqslant 0$$

Hence, we can conclude that K is a positive definite kernel on X.

* Now let's describe the RKHS $\mathcal{H}$ of the p.d. kernel K.

Let $E = \{ \sum_{i=1}^{N} a_i \Psi(x_i),\ N \in \mathbb{N}^*, (x_1, \dots, x_N) \in \mathcal{X}^N, (a_1, \dots, a_N) \in \mathbb{R}^N \}$.

↳ E is a linear subspace of $\mathcal{F}$!

We define $\forall z \in \bar{E}$ the function $f_z : y \longmapsto \langle z, \Psi(y) \rangle_{\mathcal{F}}$ and the set $\mathcal{H} = \{ f_z,\ z \in \bar{E} \}$.

We define the function $Q : \begin{array}{l} \bar{E} \longrightarrow \mathcal{H} \\ z \longmapsto f_z \end{array}$. This function is a linear isomorphism. (it is clearly linear and surjective, I admit that Q is injective).

* We know that $\mathcal{F}$ is a Hilbert space. Since $\bar{E}$ is a closed subspace of $\mathcal{F}$, we can deduce that $\bar{E}$ is a Hilbert space (when endowed with the induced inner product), that is $(\bar{E}, \langle .,. \rangle_{\bar{E}})$ is a Hilbert space.
(where $\langle .,. \rangle_{\bar{E}} = \langle .,. \rangle_{\mathcal{F}}$)

Moreover, since Q is a linear isomorphism and since $\bar{E}$ is a Hilbert space, we can deduce that $\mathcal{H}$ endowed with the inner product $\langle f_z, f_{\bar{z}} \rangle_{\mathcal{H}} := \langle z, \bar{z} \rangle_{\bar{E}}$ is a Hilbert space.

Let us prove that $(\mathcal{H}, \langle .,. \rangle_{\mathcal{H}})$ is the RKHS of K.

① $\forall x \in \mathcal{X}$, $f_{\Psi(x)} : y \longmapsto \langle \Psi(x), \Psi(y) \rangle_{\mathcal{F}} = K(x,y)$
and since $f_{\Psi(x)} \in \mathcal{H}$ (by definition of $\mathcal{H}$),

We can deduce that $\forall x \in \mathcal{X}$, $K_x = f_{\Psi(x)} \in \mathcal{H}$.

(2) $\forall x \in \mathcal{X}$, $\forall z \in \bar{E}$ ($\Rightarrow \exists ! f_z \in \mathcal{H}$), we have:

$$\langle f_z , K_x \rangle_{\mathcal{H}} = \langle f_z , f_{\Psi(x)} \rangle_{\mathcal{H}}$$

$$= \langle z, \Psi(x) \rangle_{\bar{E}}$$

$$= \langle z, \Psi(x) \rangle_{\mathcal{F}}$$

$$= f_z(x)$$

i.e. $\langle f_z , K_x \rangle_{\mathcal{H}} = f_z(x)$ $\Rightarrow$ the reproducing property holds!

Thus, we can conclude that $(\mathcal{H}, \langle \cdot, \cdot \rangle_{\mathcal{H}})$ is the RKHS of K. (where $\mathcal{H} = \{ f_z , z \in \bar{E} \}$)

2. Now we will express the norm of RKHS $\mathcal{H}$ in terms of the norm in $\mathcal{F}$.

Let us recall that function $Q: \bar{E} \longrightarrow \mathcal{H}$, $z \longmapsto f_z$ is a linear isomorphism, so Q is bijective, which means that $\forall z \in \bar{E}$, $\exists! \, f_z \in \mathcal{H}$ such that $Q(z) = f_z$.

Therefore, in the following instead of writing "$\forall f \in \mathcal{H}$" I will write directly "$\forall f_z \in \mathcal{H}$".

* $\forall f_z \in \mathcal{H}$, we have:

$$\|f_z\|^2_{\mathcal{H}} = \langle f_z, f_z \rangle_{\mathcal{H}}$$

$$:= \langle z, z \rangle_{\bar{E}}$$

$$= \langle z, z \rangle_{\mathcal{F}}$$

$$= \|z\|^2_{\mathcal{F}}$$

i.e. $\|f_z\|^2_{\mathcal{H}} = \|z\|^2_{\mathcal{F}} \quad \forall f_z \in \mathcal{H}$.

$\Longrightarrow \forall f_z \in \mathcal{H}, \quad \|f_z\|_{\mathcal{H}} = \|z\|_{\mathcal{F}}$.