First, we check that K is p.d. Take $N \geq 1$, $a_1, \dots, a_N \in \mathbb{R}$, $x_1, \dots, x_N \in \mathcal{X}$

$$\sum_{i=1}^{N} \sum_{j=1}^{N} a_i a_j K(x_i, x_j) = \left\| \sum_{i=1}^{N} a_i \Psi(x_i) \right\|_{\mathcal{F}}^2 \geq 0.$$

Moreover, K is symmetric. Hence K is a p.d. kernel.

Define

$$F \begin{cases} \mathcal{F} \longrightarrow \mathbb{R}^{\mathcal{X}} \\ \omega \longmapsto \left(f_\omega : x \longmapsto \langle \Psi(x), \omega \rangle_{\mathcal{F}}\right) \end{cases}$$

We divide the answer into three parts:

1. We equip $\mathcal{H} := F(\mathcal{F})$ with the structure of a Hilbert space
2. We show that $\mathcal{H}$ is the RKHS of K
3. We explicit the norm $\|.\|_{\mathcal{H}}$

1. Note that $F : \mathcal{F} \longrightarrow \mathbb{R}^{\mathcal{X}}$ is a linear map (between vector space).

Consider its nullspace

$$\begin{aligned} N &:= F^{-1}(\{0\}) \\ &= \{\omega \in \mathcal{F} : \forall x \in \mathcal{X} \ \langle \omega, \phi(x) \rangle_{\mathcal{F}} = 0\} \\ &= \{\omega \in \mathcal{F} : \omega \in \phi(\mathcal{X})^{\perp}\} \\ &= \phi(\mathcal{X})^{\perp} \end{aligned}$$

We show that N is closed in $\mathcal{F}$.

Indeed, for any $x \in \mathcal{X}$, $\{\omega \in \mathcal{F} : \langle \omega, \phi(x) \rangle_{\mathcal{F}} = 0\}$ is closed as $\omega \longmapsto \langle \omega, \phi(x) \rangle_{\mathcal{F}}$ is a continuous linear form on $\mathcal{F}$.

Hence, N is closed since it is an intersection of closed sets.

Moreover, N is a linear subspace of $\mathcal{F}$.

Hence, N is a closed linear subspace of $\mathcal{F}$. Therefore

$$\mathcal{F} = N \oplus N^{\perp}$$

As a consequence, $F_{|N^{\perp}} : N^{\perp} \longrightarrow \mathbb{R}^{\mathcal{X}}$ is injective.

Define $\mathcal{H} := F(\mathcal{F}) = F(N^{\perp})$ linear subspace of $\mathbb{R}^{\mathcal{X}}$

Then $\tilde{F} := F_{|N^{\perp}} : N^{\perp} \longrightarrow \mathcal{H}$ is an isomorphism of vector spaces.

Moreover $N^{\perp} = \{\omega \in \mathcal{F} : \forall v \in N \ \langle v, \omega \rangle_{\mathcal{F}} = 0\}$ is closed (same reasoning as above).

Hence $N^\perp$ is a closed linear subspace of $\mathcal{F}$, so $N^\perp$ is a Hilbert space too

$\left(\text{with } \langle \cdot, \cdot \rangle_{N^\perp} = \langle \cdot, \cdot \rangle_{\mathcal{F} \mid N^\perp \times N^\perp} \right)$

Therefore, we define a structure of Hilbert space on $\mathcal{H}$ by defining

$$\forall f, g \in \mathcal{H} \qquad \langle f, g \rangle_{\mathcal{H}} = \langle \tilde{F}^{-1}(f), \tilde{F}^{-1}(g) \rangle_{\mathcal{F}}$$

2. We show that the Hilbert space $\mathcal{H}$ is the RKHS of K

- We show that, for any $x \in \mathcal{X}$, $K_x \in \mathcal{H}$

Indeed $\quad K_x = \langle \Psi(x), \Psi(\cdot) \rangle_{\mathcal{F}} = F(\Psi(x)) \in F(\mathcal{F}) = \mathcal{H}$

- Take $f \in \mathcal{H}$ and $x \in \mathcal{X}$; we show that $f(x) = \langle f, K_x \rangle_{\mathcal{H}}$

By construction of $\mathcal{H}$,

$$\langle f, K_x \rangle_{\mathcal{H}} = \langle \tilde{F}^{-1}(f), \tilde{F}^{-1}(K_x) \rangle_{\mathcal{F}}$$

But, both $\Psi(x)$ and $\tilde{F}^{-1}(K_x)$ are in $F^{-1}(\{K_x\})$

It must be that $\quad \Psi(x) - \tilde{F}^{-1}(K_x) \in N$

Since $\tilde{F}^{-1}(f) \in N^\perp$, this gives

$$\begin{aligned} \langle f, K_x \rangle_{\mathcal{H}} &= \langle \tilde{F}^{-1}(f), \Psi(x) \rangle_{\mathcal{F}} \\ &= F(\tilde{F}^{-1}(f))(x) \\ &= f(x) \end{aligned}$$

3. We show that

$$\forall f \in \mathcal{H} \qquad \|f\|^2_{\mathcal{H}} = \min \{ \|w\|^2_{\mathcal{F}} : w \in \mathcal{F},\ f_w = f \}$$

Indeed, take $f \in \mathcal{H}$ and define $w := \tilde{F}^{-1}(f) \in N^\perp$

Take some $w' \in \mathcal{F}$ s.t. $f_{w'} = f$ ie $w' \in F^{-1}(\{f\})$.

Since w is also in $F^{-1}(\{f\})$, $w - w' \in N$

By Pythagore theorem,

$$\|w'\|^2_{\mathcal{F}} = \|w - w'\|^2_{\mathcal{F}} + \underbrace{\|w\|^2_{\mathcal{F}}}_{= \|f\|^2_{\mathcal{H}}}$$

$$\geq \|f\|^2_{\mathcal{H}} \quad \text{with equality iff } w = w',$$

which concludes the proof.

Exercise 2

($\Rightarrow$) We show that if $f \in \mathcal{H}$, then

$$\exists \lambda > 0 \qquad (x,x') \longmapsto k(x,x') - \lambda f(x) f(x') \quad \text{is p.d.}$$

Note that this kernel is symmetric.

Now, take $N \geq 1$, $a_1, \dots, a_N \in \mathbb{R}$, $x_1, \dots, x_N \in \mathcal{X}$

$$\sum_{i=1}^{N} \sum_{j=1}^{N} a_i a_j \left(k(x_i, x_j) - \lambda f(x_i) f(x_j) \right)$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{N} a_i a_j \langle k_{x_i}, k_{x_j} \rangle_{\mathcal{H}} - \lambda \sum_{i=1}^{N} \sum_{j=1}^{N} a_i a_j f(x_i) f(x_j) \quad \text{(Reproducing property)}$$

$$= \Big\| \sum_{i=1}^{N} a_i k_{x_i} \Big\|_{\mathcal{H}}^2 - \lambda \left(\sum_{i=1}^{N} a_i f(x_i) \right)^2$$

$$= \Big\| \sum_{i=1}^{N} a_i k_{x_i} \Big\|_{\mathcal{H}}^2 - \lambda \left(\sum_{i=1}^{N} a_i \langle f, k_{x_i} \rangle_{\mathcal{H}} \right)^2 \quad \text{(Reproducing property)}$$

$$= \Big\| \sum_{i=1}^{N} a_i k_{x_i} \Big\|_{\mathcal{H}}^2 - \lambda \langle f, \sum_{i=1}^{N} a_i k_{x_i} \rangle_{\mathcal{H}}^2$$

$$\geq \Big\| \sum_{i=1}^{N} a_i k_{x_i} \Big\|_{\mathcal{H}}^2 \left(1 - \lambda \| f \|_{\mathcal{H}}^2 \right) \quad \text{(Cauchy-Schwarz)} \quad (*)$$

Therefore, if we take any $\lambda > 0$ s.t. $\lambda \|f\|_{\mathcal{H}}^2 \leq 1$,

ie $\lambda \leq \dfrac{1}{\|f\|_{\mathcal{H}}^2}$ (with the convention $\frac{1}{0} = +\infty$)

then $(*)$ shows that,

$\forall N \geq 1$, $\forall a_1, \dots, a_N \in \mathbb{R}$, $\forall x_1, \dots, x_N \in \mathcal{X}$

$$\sum_{i=1}^{N} \sum_{j=1}^{N} a_i a_j \left(k(x_i, x_j) - \lambda f(x_i) f(x_j) \right) \geq 0$$

Hence, in this case, $(x,x') \longmapsto k(x,x') - \lambda f(x) f(x')$ is p.d.

($\Leftarrow$) We show that, if, for some $\lambda > 0$,

$$(x,x') \longmapsto k(x,x') - \lambda f(x) f(x') \quad \text{is p.d.}$$

then $f \in \mathcal{H}$.

First, note that if $f = 0$, f is necessarily in $\mathcal{H}$ so the result is trivial.

Throughout what follows, we assume that $f \neq 0$.

Define: $\quad K_1 : (x,x') \longmapsto k(x,x') - \lambda f(x) f(x')$

K_1 is p.d. by assumption. Denote by $\mathcal{H}_1$ its RKHS.

Define

$$K_2 : (x, x') \longmapsto \lambda f(x) f(x')$$

Let us check that K_2 is a p.d. kernel.

Indeed, by a lemma seen in class,

$$(x, x') \longmapsto f(x) f(x') = \langle f(x), f(x') \rangle_{\mathbb{R}}$$

is p.d.

Hence, as $\lambda > 0$, K_2 is a p.d. kernel.

Denote by $\mathcal{H}_2$ its RKHS.

We have seen in Exercise 5.1 that $K_1 + K_2$ is a p.d. kernel and that its RKHS $(\mathcal{H}_{12}, \langle \cdot, \cdot \rangle_{\mathcal{H}_{12}})$ satisfy

$$\mathcal{H}_{12} = \mathcal{H}_1 + \mathcal{H}_2$$

But $K = K_1 + K_2$, so, by uniqueness of the RKHS of K,

$$(\mathcal{H}_{12}, \langle \cdot, \cdot \rangle_{\mathcal{H}_{12}}) = (\mathcal{H}, \langle \cdot, \cdot \rangle_{\mathcal{H}})$$

ie $\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2$

As $0 \in \mathcal{H}_1$, $\mathcal{H}_2 \subset \mathcal{H}$.

Finally, we show that $f \in \mathcal{H}$ by showing that $f \in \mathcal{H}_2$.

Indeed, as $f \neq 0$, $\exists x_0 \in \mathcal{X}$ $f(x_0) \neq 0$.

By definition of the RKHS, $K_{2x_0} \in \mathcal{H}_2$

But, $K_{2x_0}(x) = \lambda f(x_0) f(x)$ for any $x \in \mathcal{X}$.

As $\lambda f(x_0) \neq 0$, and since $\mathcal{H}_2$ is a vector space,

$$f = \frac{K_{2x_0}}{\lambda f(x_0)} \in \mathcal{H}_2$$

Since $\mathcal{H}_2 \subset \mathcal{H}$, $f \in \mathcal{H}$