Kernel methods for machine learning
Homework 1

Simon Queric simon.queric@telecom-paris.fr

February 6, 2024

Exercise 1

Let f be a function in the unit ball of the corresponding RKHS and x a point in $\mathbf{X}$. Then, the reproducting property writes :

$$\begin{aligned}
|f(x)| &= |\langle f, K_x \rangle| \\
&\leqslant \|f\|_{\mathcal{H}} K(x,x)^{1/2} && \text{by Cauchy-Schwarz's inequality.} \\
&\leqslant K(x,x)^{1/2} && \text{because } f \text{ belongs to the unit ball.} \\
&\leqslant b && \text{because the kernel is bounded.}
\end{aligned}$$

It holds for any $x \in \mathcal{X}$ so that $\|f\|_\infty = \sup_{x \in \mathcal{X}} |f(x)| \leqslant b$.

Exercise 2

$\Longrightarrow$ Let's suppose that K is positive definite.

By symmetry we have $K(x, x') = 1 \iff K(x', x) = 1$. To prove the second property, let's suppose by contradiction that $K(x, x') = K(x', x'') = 1$ but $K(x, x'') = 0$. Then, the Gram matrix of $\{x, x', x''\}$ writes $\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ whose determinant is equal to -1. which is impossible because K is positive definite. Hence $K(x, x') = K(x', x'') = 1 \Rightarrow K(x, x'') = 1$.

$\Longleftarrow$ Let K be a similarity measure which satisfies the properties given in the exercise 2.

Then, the binary relation defined by $x \sim x' \Leftrightarrow K(x, x') = 1$ is an equivalence relation on $\mathcal{X}$: it is reflexive, symmetric and transitive.

Let $n \geqslant 1$, $(x_1, \cdots, x_n) \in \mathcal{X}^n, (a_1, \cdots, a_n) \in \mathbb{R}^n$. Since $\sim$ is an equivalence relation on $\mathcal{X}$, there exists a partition $(C_k)_{1 \leqslant k \leqslant K}$ of $\{1, \cdots, n\}$ such that for $i, j \in C_k$ we have $x_i \sim x_j$ and for $i \in C_k, j \in C_\ell$ we have $K(x_i, x_j) = 0$ if $k \neq \ell$.

Then : $\displaystyle\sum_{i,j} a_i a_j K(x_i, x_j) = \sum_{k=1}^{K} \sum_{i,j \in C_k} a_i a_j = \sum_{k=1}^{K} \left(\sum_{i \in C_k} a_i \right)^2 \geqslant 0$ which proves K is positive definite kernel.

Exercise 3

1. Let's show that $K = \alpha K_1 + \beta K_2$ is a positive definite kernel. Let $x, x' \in \mathcal{X}$. It's symmetric because $\alpha K_1(x, x') + \beta K_2(x, x') = \alpha K_1(x', x) + \beta K_2(x', x)$ by symme-

try of K_1 and K_2. It's positive because for any $a_{1:n} \in \mathbb{R}^n, x_{1:n} \in \mathcal{X}^n$ we have $\sum_{i,j} a_i a_j (\alpha K_1(x_i, x_j) + \beta K_2(x_i, x_j)) = \alpha \sum_{i,j} a_i a_j K_1(x_i, x_j) + \beta \sum_{i,j} a_i a_j K_2(x_i, x_j) \geqslant 0$.

Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be the two RKHS corresponding to K_1 and K_2 respectively. Let's suppose that $\mathcal{H}_1 \cap \mathcal{H}_2 = \emptyset$. Then, we can endow $\mathcal{H}_1 \oplus \mathcal{H}_2$ with the following inner product : $\langle f + g, f' + g' \rangle = \langle f, f' \rangle_{\mathcal{H}_1}/\alpha + \langle g, g' \rangle_{\mathcal{H}_2}/\beta$ with $(f, g), (f', g') \in \mathcal{H}_1 \times \mathcal{H}_2$ so that it becomes an Hilbert space. Then if $(f, g) \in \mathcal{H}_1 \times \mathcal{H}_2$ and $x \in \mathcal{X}$ we have $\langle K_x, f + g \rangle = \langle \alpha K_{1,x}, f \rangle_{\mathcal{H}_1}/\alpha + \langle \beta K_{2,x}, g \rangle_{\mathcal{H}_2}/\beta = f(x) + g(x)$. According to Aronszajn's theorem $\mathcal{H}_1 \oplus \mathcal{H}_2$ is the RKHS associated to K.

2. The symmetry of K is derived from the symmetry of $\langle \cdot, \cdot \rangle_{\mathcal{F}}$. Let $a_{1:n} \in \mathbb{R}^n, x_{1:n} \in \mathcal{X}^n$, then

$$
\begin{aligned}
\sum_{i,j} a_i a_j K(x_i, x_j) &= \sum_{i,j} a_i a_j \langle \psi(x_i), \psi(x_j) \rangle_{\mathcal{F}} \\
&= \left\langle \sum_i a_i \psi(x_i), \sum_j a_j \psi(x_j) \right\rangle_{\mathcal{F}} \\
&= \left\| \sum_{i=1}^n a_i \psi(x_i) \right\|_{\mathcal{F}}^2 \geqslant 0
\end{aligned}
$$

Hence K is a positive definite kernel on $\mathcal{X}$.

Let $\mathcal{H} = \left\{ f_u : \mathcal{X} \ni x \mapsto \mathbb{R} \in \langle \psi(x), u \rangle_{\mathcal{F}} \mid u \in \overline{\text{Vect}\{\psi(x), x \in \mathcal{X}\}} \right\} \subseteq \mathbb{R}^{\mathcal{X}}$.

We can endow $\mathcal{H}$ with the following inner product $\langle f_u, f_v \rangle_{\mathcal{H}} = \langle u, v \rangle_{\mathcal{F}}$ which does not depend of the choice of u and v. Indeed, we can take u' stuch that $f_u = f_{u'}$ and then $u - u' \perp \overline{\text{Vect}\{\psi(x), x \in \mathcal{X}\}}$ and $u - u' \in \overline{\text{Vect}\{\psi(x), x \in \mathcal{X}\}}$ which implies that $u = u'$ and then $\langle f_u, f_v \rangle = \langle f_{u'}, f_v \rangle$. With this inner product, $\mathcal{H}$ is an Hilbert space.

We have for $f_u \in \mathcal{H}$, $\langle K_x, f_u \rangle = \langle \psi(x), u \rangle = f_u(x)$. Hence, according to Aronszajn's theorem, $\mathcal{H}$ is the reproducing Hilbert Space corresponding to K.

3. $\Longrightarrow$ The case $f = 0$ is obvious. Let's assume that $f \neq 0$. We write $K'(x, x') = K(x, x') - \lambda f(x) f(x')$. By Cauchy-Schwarz's inequality, we have $f(x)^2 \leqslant \|f\|_{\mathcal{H}}^2 K(x, x)$ so that we can think about taking $\lambda = 1/\|f\|_{\mathcal{H}}^2 > 0$. Now, we can show that it indeed works. Let $a_{1:n} \in \mathbb{R}^n, x_{1:n} \in \mathcal{X}^n$. We have :

$$
\begin{aligned}
\sum_{i,j} a_i a_j f(x_i) f(x_j) &= \left(\sum_{i=1}^n a_i f(x_i) \right)^2 \\
&= \left(\sum_{i=1}^n a_i \langle K_{x_i}, f \rangle_{\mathcal{H}} \right)^2 \\
&= \left\langle \sum_{i=1}^n a_i K_{x_i}, f \right\rangle_{\mathcal{H}}^2 \quad \text{by bilinearity of the inner product} \\
&\leqslant \left\| \sum_{i=1}^n a_i K_{x_i} \right\|_{\mathcal{H}}^2 \|f\|_{\mathcal{H}}^2 \\
&= \sum_{i,j} a_i a_j K(x_i, x_j) \|f\|_{\mathcal{H}}^2
\end{aligned}
$$

We've shown that $K'(x,x') = K(x,x') - f(x)f(x')/\|f\|_{\mathcal{H}}^2$ is indeed a positive definite kernel.

$\Longleftarrow$ Let's assume that $K_1(x,x') = K(x,x') - \lambda f(x)f(x')$ is p.d.

We can decompose K into $K(x,x') = (K(x,x') - \lambda f(x)f(x')) + \lambda f(x)f(x')$. According to question 2, using that $(\mathbb{R}, \times)$ is an Hilbert space, $K_2(x,x') = f(x)f(x') = f(x) \times f(x')$ is positive definite kernel. Then, f belong to the RKHS of K_2 because $f(x) = 1 \times f(x)$ and $f(x) = 0 + f(x)$ belongs to the RKHS of K according to the question 1.