Homework 1

Mathilde Bateson Thomas Debarre
Kernel methods in machine learning Homework 1
MVA 2016/2017

1 Kernel examples

Are the following kernels positive definite?

1.1

The kernel K_1 is defined as:

$$\forall (x,y) \in \mathbb{R}^2 \quad K_1(x,y) = 10^{xy}$$

The linear kernel K defined as $\forall (x,y) \in \mathbb{R}^2 \mapsto K(x,y) = \langle x,y \rangle_{\mathbb{R}} = xy$ is a positive definite kernel. Thus by multiplying by the constant $c = \ln 10$ and taking the exponential, the resulting kernel $K_1 = \exp(cK)$ is also positive definite.

The kernel K_2 is defined as:

$$\forall (x,y) \in \mathbb{R}^2 \quad K_2(x,y) = 10^{x+y}$$

With $\phi : \mathbb{R} \to \mathbb{R}$ defined as $\forall x \in \mathbb{R}, \phi(x) = 10^x$, $\forall (x,y) \in \mathbb{R}^2, K_2(x,y) = 10^x 10^y = \langle \phi(x), \phi(y) \rangle_{\mathbb{R}}$. Thus using Aronzsajn's theorem, K_2 is a positive definite kernel.

1.2

The kernel K_3 is defined as:

$$\forall (x,y) \in [0,1[^2 \quad K_3(x,y) = -\log(1-xy)$$

We have that:

$$\forall (x,y) \in [0,1[^2 \quad K_3(x,y) = \lim_{n\to\infty} K_n(x,y)$$

where

$$K_n(x,y) = \sum_{i=1}^{n} \frac{K(x,y)^i}{i}$$

with K the linear kernel. We have that K_n is a positive definite kernel as a sum of positive definite kernels, and it converges pointewisely to K_3. Thus K_3 is a positive definite kernel.

1.3

Given a set $\mathcal{X}$ and $f, g : \mathcal{X} \to \mathbb{R}_+$ two non-negative functions, the kernel K_4 is defined as:

$$\forall (x,y) \in \mathcal{X}^2 \quad K_4(x,y) = \min(f(x)g(y), f(y)g(x))$$

$\forall a \in \mathbb{R}^n, \forall (x_i)_{1\leq i \leq n} \in \mathcal{X}^n$, we have:

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j K_4(x_i,x_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \min(f(x_i)g(x_j), f(x_j)g(x_i))$$

We note that we can restrict ourselves to the case where $\forall i, g(x_i) > 0$. Thus we have:

$$\begin{aligned}
\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j K_4(x_i, x_j) &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j g(x_i) g(x_j) \min\left(\frac{f(x_j)}{g(x_j)}, \frac{f(x_i)}{g(x_i)}\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j g(x_i) g(x_j) \int_{t=0}^{+\infty} \mathbb{1}_{t \leq \frac{f(x_j)}{g(x_j)}} \mathbb{1}_{t \leq \frac{f(x_i)}{g(x_i)}} dt \\
&= \int_{t=0}^{+\infty} \left(\sum_{i=1}^{n} a_i g(x_i) \mathbb{1}_{t \leq \frac{f(x_i)}{g(x_i)}}\right) \left(\sum_{j=1}^{n} a_j g(x_j) \mathbb{1}_{t \leq \frac{f(x_j)}{g(x_j)}}\right) dt \\
&= \int_{t=0}^{+\infty} \left(\sum_{i=1}^{n} a_i g(x_i) \mathbb{1}_{t \leq \frac{f(x_i)}{g(x_i)}}\right)^2 dt \geq 0
\end{aligned}$$

Thus K_4 is a positive definite kernel.

2 Combining kernels

Find the Reproducing Kernel Hilbert Space of the following kernels.

2.1

- The Kernel K_1 is defined as:

$$\forall (x, y) \in \mathbb{R}^2 \quad K_1(x, y) = (xy + 1)^2 = x^2 y^2 + 2xy + 1 = \langle \phi(x), \phi(y) \rangle_{\mathbb{R}^3}$$

with the following ϕ:

$$\begin{aligned}
\phi : \quad & \mathbb{R} \to \mathbb{R}^3 \\
& x \mapsto \begin{pmatrix} x^2 \\ \sqrt{2}x \\ 1 \end{pmatrix}
\end{aligned}$$

We know that $\mathcal{H}_1$ contains all the functions

$$\begin{aligned}
f(x) &= \sum_i a_i K_1(x_i, x) \\
&= \sum_i a_i \langle \phi(x_i), \phi(x) \rangle_{\mathbb{R}^3} \\
&= \langle \sum_i a_i \phi(x_i), \phi(x) \rangle_{\mathbb{R}^3}
\end{aligned}$$

Let $G = Span(1, x, x^2)$. We will prove that G is the RKHS of K_1.

  - First, we have that $Span\{K_1(, z) \mid z \in \mathbb{R}^2\} \subset G$
  - Second, any polynomial of degree 2 belongs to $\mathcal{H}_1$, because for any (a_0, a_1, a_2), the following linear equation is solvable:

$$\begin{pmatrix} z_0^2 & z_0 & 1 \\ z_1^2 & z_1 & 1 \\ z_2^2 & z_2 & 1 \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_0 \\ \frac{a_1}{\sqrt{2}} \\ a_2 \end{pmatrix}$$

    Thus $\sum_{i=0}^{2} b_i k(x, z_i) = \sum_{i=0}^{2} a_i x^i$, and the reproducing property holds.
  - By remarking that G is a Hilbert space, we finally conclude that it is the RKHS.

- The Kernel K_2 is defined as:

$$\forall (x,y) \in \mathbb{R}^2 \quad K_2(x,y) = (xy-1)^2 = x^2y^2 - 2xy + 1$$

We observe that when choosing $x_1 = 0$ and $x_2 = 1$ the similarity matrix $\mathcal{K}$ is:

$$\mathcal{K} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

The eigenvalues of this matrix are $\frac{1+\sqrt{5}}{2} > 0$ and $\frac{1-\sqrt{5}}{2} < 0$, so it is not positive semidefinite and thus K_2 isn't a reproducing kernel.

- The Kernel K is defined as: $K = K_1 + K_2$

$$\forall (x,y) \in \mathbb{R}^2 \quad K(x,y) = (xy+1)^2 + (xy-1)^2 = 2x^2y^2 + 2 = \langle \phi(x), \phi(y) \rangle_{\mathbb{R}^3}$$

with the following ϕ:

$$\begin{aligned} \phi: \quad & \mathbb{R} \to \mathbb{R}^3 \\ & x \mapsto \begin{pmatrix} \sqrt{2}x^2 \\ \sqrt{2} \end{pmatrix} \end{aligned}$$

We use the same arguments as for K_1 to find the RKHS $\mathcal{H}$ of K:$Span(1, x^2)$

2.2

Let K_1 and K_2 be two positive definite kernels on a set $\mathcal{X}$ and α and β two positive scalars. We want to show that $K = \alpha K_1 + \beta K_2$ is positive definite.

- $\forall (x,y) \in \mathcal{X}^2$: $K(x,y) = \alpha K_1(x,y) + \beta K_2(x,y) = \alpha K_1(y,x) + \beta K_2(y,x) = K(y,x)$, which proves that K is symmetric.

- Moreover: $\forall (x_i)_{1\leq i \leq N} \in \mathcal{X}^N$ and $\forall (a_i)_{1 \leq i \leq N} \in \mathbb{R}^N$:

$$\begin{aligned} \sum_{i=1}^N \sum_{j=1}^N a_i a_j K(x_i, x_j) &= \sum_{i=1}^N \sum_{j=1}^N a_i a_j \left(\alpha K_1(x_i,x_j) + \beta K_2(x_i,x_j)\right) \\ &= \alpha \underbrace{\left(\sum_{i=1}^N \sum_{j=1}^N a_i a_j K_1(x_i,x_j) \right)}_{\geq 0} + \beta \underbrace{\left(\sum_{i=1}^N \sum_{j=1}^N a_i a_j K_2(x_i,x_j) \right)}_{\geq 0} \geq 0 \end{aligned}$$

Which proves that K is positive. Finally, we have proven that K is positive definite.

3 Uniqueness of the RKHS

Exercise 3: Let $\mathcal{H}$ and $\mathcal{H}'$ be 2 RKHS of the same reproducing kernel function $\mathcal{K}$. We want to prove that $\mathcal{H} = \mathcal{H}'$, and that both inner products are the same.
Let $\mathcal{F}$ be the linear space spanned by the K_x functions:

$$\mathcal{F} = \{K_x, x \in \mathcal{X}\} \tag{1}$$

We will first show that $\mathcal{F}$ is a dense subset of both $\mathcal{H}$ and $\mathcal{H}'$. Let $h \in \mathcal{H}$ be a vector orthogonal to the set $\mathcal{F}$ (considered as a subset of $\mathcal{H}$). Using the reproducing property of an RKHS, for any $x \in \mathcal{X}$, we have:

$$\langle K_x, h \rangle_{\mathcal{H}} = h(x) = 0 \tag{2}$$

since $K_x \in \mathcal{H}$. We thus have $h = 0$; this proves that the only orthogonal vector to $\mathcal{F}$ in $\mathcal{H}$ is 0. Yet this property is equivalent to $\mathcal{F}$ being dense in $\mathcal{H}$. Similarly, $\mathcal{F}$ is also dense in $\mathcal{H}'$. Therefore, $\overline{\mathcal{F}} = \mathcal{H}$ and $\overline{\mathcal{F}} = \mathcal{H}'$. These closures are respectively defined for the norm of the inner products $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ and $\langle \cdot, \cdot \rangle_{\mathcal{H}'}$ restricted to $\mathcal{F} \times \mathcal{F}$. However, using the reproducing property, it is clear that the restrictions of both those inner products $\mathcal{F} \times \mathcal{F}$ are equal (since $\langle K_x, K_y \rangle = K(x, y)$). Both closures are therefore defined with respect to the same norm, which yields $\mathcal{H} = \mathcal{H}'$. Finally, let us show that inner products of $\mathcal{H}$ and $\mathcal{H}'$ are identical. Let $h, g \in \mathcal{H} = \mathcal{H}'$. Since $\mathcal{F}$ is dense in $\mathcal{H} = \mathcal{H}'$, there exists sequence $h_n, g_n \in \mathcal{F}^{\mathbb{N}}$ such that $h_n \to h$ and $g_n \to g$. As explained earlier, we have:

$$\langle g_n, h_n \rangle_{\mathcal{H}} = \langle g_n, h_n \rangle_{\mathcal{H}'} \tag{3}$$

Since the inner products are continuous, we have $\langle g_n, h_n \rangle_{\mathcal{H}} \to \langle g, h \rangle_{\mathcal{H}}$ and $\langle g_n, h_n \rangle_{\mathcal{H}'} \to \langle g, h \rangle_{\mathcal{H}'}$, and so:

$$\langle g, h \rangle_{\mathcal{H}} = \langle g, h \rangle_{\mathcal{H}'} \tag{4}$$

This concludes our proof that there is a unique RKHS for a given reproducing kernel K.