MVA: Kernel methods in machine learning (2022/2023)

Assignment 1

Solution by Matias Etcheverry January 30, 2023

1 Exercise 1

Let K_1 and K_2 be two positive definite kernels on a set $\mathcal{X}$ with corresponding RKHS's $\mathcal{H}_1$ and $\mathcal{H}_2$, and α, β two positive scalars.

1. Show that $\alpha K_1 + \beta K_2$ is positive definite.

   Let $n \in \mathbb{N}$, $x_1, \cdots, x_n \in \mathcal{X}$, $a_1, \cdots, a_n \in \mathbb{R}$:

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j K(x_i, x_j) = \alpha \underbrace{\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j K_1(x_i, x_j)}_{\geq 0} + \beta \underbrace{\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j K_2(x_i, x_j)}_{\geq 0}$$

   because K_1 and K_2 are positive definite. Moreover, K is symmetric. In the end, K is a positive definite kernel.

2. Express the norm of the RKHS $\mathcal{H}$ in terms of the norms of both RKHS $\mathcal{H}_1$ and $\mathcal{H}_2$ and describe $\mathcal{H}$ in terms of elements in $\mathcal{H}_1$ and $\mathcal{H}_2$.

   This answer is inspired from Aronszajn [1950].

   Let's consider $\mathcal{A}$ the Hilbert space made of elements (g', g'') with $g' \in \mathcal{H}_1$ and $g'' \in \mathcal{H}_2$. We define the following norm on $\mathcal{A}$:

$$\|(g', g'')\|_{\mathcal{A}} = \frac{1}{\alpha}\|g'\|_1^2 + \frac{1}{\beta}\|g''\|_2^2$$

   We consider $\mathcal{H}_0$ the class of functions belonging to $\mathcal{H}_1$ and to $\mathcal{H}_2$. We denote $\mathcal{A}_0$ the space of couples $(f, -f), f \in \mathcal{H}_0$.

   $\mathcal{A}_0$ is closed. Indeed, let $(f_n, -f_n) \to (f', -f'')$, with $f_n \in \mathcal{H}_0$. Then $f_n \to f'$ and $f_n \to f''$. The convergence in $\mathcal{H}_0$ implies a pointwise convergence, resulting in $f' = f''$, i.e. $f' \in \mathcal{H}_0$. Then, $(f', -f'') \in \mathcal{A}_0$.

   Then, we can construct $\mathcal{A}'$ such that $\mathcal{A} = \mathcal{A}_0 \oplus \mathcal{A}'$. We also construct $\boxed{\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2}$. $\mathcal{H}$ is a Hilbert space. Finally, we also construct a correspondence between $\mathcal{A}$ and $\mathcal{H}$: for any $(f_1, f_2) \in \mathcal{A}$, $f(x) = \alpha f_1(x) + \beta f_2(x)$. The elements of $\mathcal{A}$ transformed into the 0 function are elements of $\mathcal{A}_0$. Consequently, this correspondence transforms $\mathcal{A}'$ in a one-to-one way into $\mathcal{H}$. The inverse correspondence transforms every function $f \in \mathcal{H}$ into an element $(g'(f), g''(f)) \in \mathcal{A}'$. We define the inner product and the norm on $\mathcal{H}$:

$$\langle s \mid t \rangle_{\mathcal{H}} = \frac{1}{\alpha}\langle g'(s) \mid g'(t) \rangle_1 + \frac{1}{\beta}\langle g''(s) \mid g''(t) \rangle_2$$

$$\boxed{\|f\|_{\mathcal{H}}^2 = \|(g'(f), g''(f))\|_{\mathcal{H}}^2 = \frac{1}{\alpha}\|g'(f)\|_1^2 + \frac{1}{\beta}\|g''(f)\|_2^2}$$

We notice that:

(a) for $x \in \mathcal{X}$ fixed

$$K_x : y \longmapsto K(x, y) = \alpha K_1(x, y) + \beta K_2(x, y)$$

belongs to $\mathcal{H}$

(b) Reproducing property: for $y \in \mathcal{X}$, let

$$\begin{aligned} f' &= g'(f) \\ K'(x,y) &= g'(K(x,y) \end{aligned} \qquad\qquad \begin{aligned} f'' &= g''(f) \\ K''(x,y) &= g''(K(x,y)) \end{aligned}$$

i.e.

$$\begin{aligned} f(y) &= f'(y) + f''(y) \\ K(x,y) &= K'(x,y) + K''(x,y) = \alpha K_1(x,y) + \beta K_2(x,y) \end{aligned}$$

Thus, $K''(x,y) - \beta K_2(x,y) = -\left[K'(x,y) - \alpha K_1(x,y)\right]$, which means that $(K''(x,y) - \beta K_2(x,y), -[K'(x,y) - \alpha K_1(x,y)]) \in \mathcal{A}_0$. Finally,

$$\begin{aligned} f(y) &= f'(y) + f''(y) \\ &= \langle f' \mid K_1(x,y) \rangle_1 + \langle f'' \mid K_2(x,y) \rangle_2 \\ &= \langle (f', f'') \mid (\alpha K_1(x,y), \beta K_2(x,y)) \rangle_{\mathcal{H}} \\ &= \langle (f', f'') \mid (K'(x,y), K''(x,y)) \rangle_{\mathcal{H}} + \langle \underbrace{(f', f'')}_{\in \mathcal{A}'} \mid \underbrace{(\alpha K_1(x,y) - K'(x,y), \beta K_2(x,y) - K''(x,y))}_{\in \mathcal{A}_0} \rangle_{\mathcal{H}} \\ &= \langle (f', f'') \mid (K'(x,y), K''(x,y)) \rangle_{\mathcal{H}} \end{aligned}$$

We can also construct a norm on $\mathcal{H}$ without going through the g' and g'' functions. We can define directly:

$$\boxed{\|f\|_{\mathcal{H}}^2 = \min_{f_1 \in \mathcal{H}_1, f_2 \in \mathcal{H}_2} (\alpha \|f_1\|_1^2 + \beta \|f_2\|_2^2)}$$

In order to prove that this new definition corresponds to the previous one, we notice that:

$$\beta f_2 - g''(f) = -[\alpha f_1 - g'(f)]$$

Thus, $((\beta f_2 - g''(f)), -(\alpha f_1 - g'(f))) \in \mathcal{A}_0$. We then have

$$\begin{aligned} \|f\|_{\mathcal{H}}^2 &= \|(\alpha f_1, \beta f_2)\|_{\mathcal{H}}^2 \\ &= \|(g'(f), g''(f))\|_{\mathcal{H}}^2 + \|(\alpha f_1 - g'(f), \beta f_2 - g''(f))\|_{\mathcal{H}}^2 \end{aligned}$$

The very last term is minimum when $g'(f) = \alpha f_1$ and $g''(f) = \beta f_2$.

2 Exercise 2

Let $\mathcal{X}$ be a set and $\mathcal{F}$ be a Hilbert space. Let $\Psi : \mathcal{X} \to \mathcal{F}$, and $\mathcal{K} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ be:

$$\forall x, x' \in \mathcal{X}, K(x, x') = \langle \Psi(x) \mid \Psi(x') \rangle_{\mathcal{F}}$$

1. Show that $\mathcal{K}$ is a positive definite kernel on $\mathcal{X}$. K is symmetric.

   Let $n \in \mathbb{N}$, $x_1, \cdots, x_n \in \mathcal{X}$, $a_1, \cdots, a_n \in \mathbb{R}$:

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

   Thus, K is a positive definite kernel.

2. Express the norm of the RKHS $\mathcal{H}$ in terms of the norm in $\mathcal{F}$ and describe $\mathcal{H}$ in terms of elements in $\mathcal{F}$.

   Let's define $g : z \in \mathcal{F} \mapsto \langle z \mid \Psi \rangle_{\mathcal{F}}$. We also define $\mathcal{H} = \left\{ g(z) | z \in ker(g)^{\perp} \right\}$. $\mathcal{H}$ is a space of function. By definition, g is injective and surjective on $\mathcal{H}$ (unless Ψ is the zero function. We assume it is not). $\mathcal{H}$ is a Hilbert space:

   - it is linear.
   - for $f, f' \in \mathcal{H}$, we give the inner product:

     $$\langle f \mid f' \rangle_{\mathcal{H}} = \left\langle g^{-1}(f) \;\middle|\; g^{-1}(f') \right\rangle_{\mathcal{F}}$$

   - $\mathcal{H}$ is complete for the above inner product. Let $(f_n)_{n \geq 0}$ be a Cauchy sequence in $\mathcal{H}$. Then, for $n \in \mathbb{N}, m \geq n$, $\|f_n - f_m\|_{\mathcal{H}} = \left\|g^{-1}(f_n) - g^{-1}(f_m)\right\|_{\mathcal{F}}$. Thus $(g^{-1}(f_n))_{n \geq 0}$ is a Cauchy sequence in $\mathcal{F}$. $\mathcal{F}$ being an Hilbert space, the Cauchy sequence $(g^{-1}(f_n))_{n \geq 0}$ converges to z in $\mathcal{F}$. Thus, the Cauchy sequence $(f_n)_{n \geq 0}$ converges to $g(z) \in \mathcal{H}$.

   We finally need to show that $\mathcal{H}$ is a RKHS of K:

   - for $x \in \mathcal{X}$ fixed

     $$K_x : y \longmapsto K(x, y) = \langle \Psi(x) \mid \Psi(y) \rangle_{\mathcal{F}} = g(\Psi(x))(y)$$

     belongs to $\mathcal{H}$
   - Reproducing property: for $y \in \mathcal{X}$ and $f \in \mathcal{H}$,

     $$\begin{aligned} f(y) &= \left\langle g^{-1}(f) \;\middle|\; \Psi(y) \right\rangle_{\mathcal{F}} \\ &= \langle f \mid g(\Psi(y)) \rangle_{\mathcal{H}} \\ &= \langle f \mid \langle \Psi(\cdot) \mid \Psi(y) \rangle_{\mathcal{F}} \rangle_{\mathcal{H}} \\ &= \langle f \mid K(\cdot, y) \rangle_{\mathcal{H}} \end{aligned}$$

References

N. Aronszajn. Theory of reproducing kernels. *Transactions of the American Mathematical Society*, 68 (3):337–404, 1950. ISSN 00029947. URL `http://www.jstor.org/stable/1990404`.