Exercise Sheet 8.

Hand in due to: **Friday, 12/05/2023, 14:00**

Exercise 1 (Orthonormal basis of H_X | 4 points).

Let $\Omega \in \mathbb{R}^d$ be a nonempty open set and $H_K(\Omega)$ be a RKHS with the reproducing kernel $K : \Omega \times \Omega \to \mathbb{R}$. Let $\emptyset \neq X_0 \subset X_n = \{x_1, \dots, x_n\} \subset X_N \subset \Omega$ be sets of pairwise distinct points. Consider $\{v_i(x)\}_{i=1}^N$ as the Gram-Schmidt orthonormalization of $\{K(\cdot, x_i)\}_{i=1}^N$, i.e.,

$$v_1(x) := \frac{K(x, x_1)}{\|K(\cdot, x_1)\|_{H_K(\Omega)}},$$

and for all $1 < n \leq N$

$$v_n(x) = \frac{\widetilde{v}_n(x)}{\|\widetilde{v}_n\|_{H_K(\Omega)}}, \quad \text{where } \widetilde{v}_n(x) := K(x, x_n) - \sum_{i=1}^{n-1} \big(K(\cdot, x_n), v_i\big)_{H_K(\Omega)} v_i(x).$$

Show that the power function is

$$P_{H_{X_{n-1}}}(x_n) = v_n(x_n).$$

Hint. *Show the identity* $P_{H_{X_n}}(x)^2 = K(x, x) - \sum_{i=1}^n v_i(x)^2$.

Exercise 2 (Iterative interpolation | 4 points).

Let $\Omega \in \mathbb{R}^d$ be a nonempty, open set and $H_K(\Omega)$ be a RKHS with the reproducing kernel $K : \Omega \times \Omega \to \mathbb{R}$. Let $\emptyset \neq X_0 \subset X_n = \{x_1, \dots, x_n\} \subset X_N \subset \Omega$ sets of pairwise distinct points and $\{v_i(x)\}_{i=1}^N$ be defined as in Exercise 1. For $f \in H_K(\Omega)$, we set $r_{X_0} := f$ and $r_{X_n} := f - f_{X_n}$. Show that for all $n \geq 1$:

(a) The interpolant can be computed as $f_{X_N}(x) = \sum_{i=1}^N (f, v_i)_{H_K(\Omega)} v_i(x)$ and there holds

$$\begin{aligned} f_{X_n}(x) &= f_{X_{n-1}}(x) + (f, v_n)_{H_K(\Omega)} v_n(x), \\ P^2_{H_{X_n}}(x) &= P^2_{H_{X_{n-1}}}(x) - v_n^2(x), \\ r_{X_n}(x) &= r_{X_{n-1}}(x) - (f, v_n)_{H_K(\Omega)} v_n(x). \end{aligned}$$

(b) For the coefficients of the basis, there holds $(f, v_n)_{H_K(\Omega)} = r_{X_{n-1}}(x_n)/P_{H_{X_{n-1}}}(x_n)$.

Exercise 3 (Kernel characterization by Fourier transform | 4 points).

Let $0 \not\equiv \Phi \in L^1(\mathbb{R}^d) \cap C(\mathbb{R}^d)$ be nonnegative, i.e., $\widehat{\Phi}(x) \geq 0$ for all $x \in \mathbb{R}^d$. Show that $K(x,y) := \Phi(x-y)$ is a reproducing kernel by verifying the following steps:

(1) First, for $g \in L^1(\mathbb{R}^d)$, show that

$$\int_{\mathbb{R}^d} \Phi(x)\widehat{g}(x)dx = \int_{\mathbb{R}^d} \widehat{\Phi}(x)g(x)dx.$$

(2) By using the first step, show that

$$\Phi(0) = \lim_{m\to\infty} \int_{\mathbb{R}^d} \widehat{\Phi}(y)\widehat{g}_m(y)dy,$$

where g_m is defined in Exercise 2 from Sheet 7.

(3) Next, show $\widehat{\Phi} \in L^1(\mathbb{R}^d)$ by applying the Monotone Convergence Theorem.

(4) By previous step and fact that $\widehat{\Phi} \in L^1(\mathbb{R}^D)$ implies $\Phi = F^{-1}(\widehat{\Phi})$ conclude that $K(x,y)$ is positive definite.

Exercise 4 (Representer Theorem | 4 points).

Let $\Omega \in \mathbb{R}^d$ be a nonempty, open set and $H_K(\Omega)$ be a RKHS with a reproducing kernel $K : \Omega \times \Omega \to \mathbb{R}$. Let $X_N = \{x_1, \dots, x_n\} \subset \Omega$ be a set of pairwise distinct points and $\{f_i = f(x_i)\}_{i=1}^N \subset \mathbb{R}$ the associated point evaluations of a given function $f \in H_K(\Omega)$. For some regularization parameter $\lambda > 0$, let us define the regularized interpolant f_X^λ as

$$f_X^\lambda := \underset{g\in H_K(\Omega)}{\operatorname{argmin}} J_\lambda(g),$$

where

$$J_\lambda(g) = \sum_{i=1}^N \left(f_i - g(x_i)\right)^2 + \lambda \|g\|^2_{H_K(\Omega)}.$$

Show that for every $f \in H_K(\Omega)$ there exists a regularized interpolant

$$f_X^\lambda(\cdot) = \sum_{i=1}^N c_i K(x_i, \cdot),$$

where the vector of coefficients $c \in \mathbb{R}^N$ is the solution of the linear system

$$(A + \lambda I)c = b, \quad A = [K(x_i, x_j)]_{i,j=1}^N, \quad b = [f(x_1), \dots, f(x_N)]^\top.$$

Moreover, if K is strictly positive definite, this is the unique solution of the minimization problem.

Hint. *Prove first that for every $h \in H_K(\Omega)$ there exists $g \in H_X := span\{K(\cdot, x_1), \dots, K(\cdot, x_N)\}$ such that $J_\lambda(h) \leq J_\lambda(g)$. Then, restrict the minimization over H_X, i.e., consider only functions of the form*

$$g := \sum_{i=1}^N c_i K(\cdot, x_i).$$