

# What's this?

This document collects basic results on Reproducing Kernel Hilbert Space that are useful in the context of machine learning, aimed at readers with a basic knowledge of functional analysis. The content relies heavily on the existing literature listed in the bibliography, in particular [1] [2]. [3] [4] [5]

## Overview

## Contents

<b>1</b>	<b>Basic Property of RKHS</b>	<b>1</b>
1.1	Properties of Reproducing Kernel . . . . .	1
1.2	RKHS of a kernel . . . . .	3
1.3	Basic Properties of Kernel . . . . .	11
<b>2</b>	<b>Reconstruction of RKHS</b>	<b>11</b>
2.1	Reconstruction via Restriction . . . . .	11
2.2	Reconstruction via Operator . . . . .	11
2.3	Reconstruction via Sum and Product . . . . .	11
<b>3</b>	<b>Inheritance from Kernel to RKHS</b>	<b>11</b>
3.1	Measurability of RKHS . . . . .	11
3.2	Separability of RKHS . . . . .	11
<b>4</b>	<b>Mercer Representation</b>	<b>11</b>

## 1 Basic Property of RKHS

### 1.1 Properties of Reproducing Kernel

**Definition 1.1** (Reproducing kernel). *Let  $E$  be a nonempty set. A function  $K$  defined by*

$$K : E \times E \ni (x, y) \mapsto K(x, y) \in \mathbb{F}$$

*is called a reproducing kernel of a Hilbert space  $H$  of functions on  $E$  if it satisfies the following conditions:*

- (a)  $K(\cdot, x) \in H$  for every  $x \in E$
- (b)  $\langle f, K(\cdot, x) \rangle_H = f(x)$  for every  $x \in E$  and every  $f \in H$ .

Such a Hilbert space, associated with its reproducing kernel, is called a reproducing kernel hilbert space, and is denoted by  $(H, K)$ .

**Theorem 1.2** (Characterization of RKHS). *A Hilbert space  $H$  of functions on a nonempty set  $E$  admits a reproducing kernel  $K$  if and only if all evaluation functionals  $\{\text{ev}_x\}_{x \in E}$  are continuous on  $H$ .*

*Proof.* Suppose  $(H, K)$  is a RKHS. For  $x \in E$  and for  $f \in H$  we have

$$|\text{ev}_x(f)| = |f(x)| = |\langle f, K(\cdot, x) \rangle| \leq \|f\| \|K(\cdot, x)\| \leq \|f\| K(x, x)^{1/2} \rightarrow 0$$

as  $\|f\| \rightarrow 0$ . Thus,  $\text{ev}_x$  is continuous linear functional (with norm  $K(x, x)^{1/2}$ ).

Conversely, if  $\text{ev}_x : H \ni f \mapsto f(x) \in \mathbb{F}$  is continuous, then, by Riesz's representation theorem, there exists  $r_x \in H$  such that

$$\langle f, r_x \rangle = f(x)$$

for every  $f \in H$ . If this happens for every  $x \in E$ , then  $K(x, y) := r_x(y)$  is a reproducing kernel of  $H$ .  $\square$

**Corollary 1.3.** *Every convergent sequence in RKHS converges pointwise to the same limit.*

*Proof.*  $|f_n(x) - f(x)| = |\text{ev}_x(f_n - f)| \rightarrow 0$  when  $f_n \rightarrow f$  in norm by continuity of evaluation functional.  $\square$

**Proposition 1.4.** *(Uniqueness of  $H$  and  $K$ )*

- (a) *The subspace  $H_0$  spanned by  $\{K(\cdot, x)\}_{x \in E}$  is dense in  $H$ .*
- (b) *A Hilbert space admits at most one reproducing kernel.*
- (c) *A function  $K : E \times E \rightarrow \mathbb{F}$  is a reproducing kernel for at most one Hilbert space.*

*Proof.* For density of  $H_0$ , observe  $f \in H$  fulfills  $f \perp H_0$  if and only if

$$\langle f, K(\cdot, x) \rangle_H = f(x) = 0$$

for every  $x \in E$ , which is the case precisely when  $f \equiv 0$ . To check (b), suppose  $K_1$  and  $K_2$  qualify as a reproducing kernel of  $H$ . By definition,

$$f(x) = \langle f, K_1(\cdot, x) \rangle_H = \langle f, K_2(\cdot, x) \rangle_H$$

for every  $x \in E$ , and hence

$$\langle f, K_1(\cdot, x) - K_2(\cdot, x) \rangle_H = 0$$

for every  $f \in H$  and  $x \in E$ . From this we conclude  $K_1 = K_2$ . Finally suppose that  $(H_1, K)$  and  $(H_2, K)$  are two RKHSs. Pick  $f \in H_1$ . By (a), there is  $\{f_n\} \subset H_0 \subset H_1 \cap H_2$  such that  $f_n \rightarrow f$  in  $H_1$ -norm. Since  $\{f_n\}$  is also an Cauchy sequence in  $H_2$ , it admits a limit  $g \in H_2$ . But Corollary 1.3 implies  $f = g$ , and hence  $f \in H_2$ . We then have

$$\|f\|_{H_1} = \lim_{n \rightarrow \infty} \|f_n\|_{H_1} = \lim_{n \rightarrow \infty} \|f_n\|_{H_0} = \lim_{n \rightarrow \infty} \|f_n\|_{H_2} = \|f\|_{H_2}.$$

Therefore,  $H_1$  is isometrically included in  $H_2$ . Symmetry thus shows that both Hilbert spaces coincide.  $\square$

**Proposition 1.5.** *(Representation of RK in terms of evaluation functional) In arbitrary RKHS  $(H, K)$ , the reproducing kernel  $K : E \times E \rightarrow \mathbb{F}$  always fulfills the identity*

$$K(x, y) = \langle \text{ev}_y, \text{ev}_x \rangle_{H^*}$$

for all  $x, y \in E$ , where  $H^*$  is the dual space of  $H$ .

*Proof.* It suffices to show that a function  $K$  defined by the above equation is also a reproducing kernel. Let a mapping  $I : H^* \rightarrow H$  be the isometric anti-linear surjection, guaranteed by Riesz's Representation Theorem, that assigns to every functional in  $H^*$  the corresponding representor in  $H$ , i.e.,  $g^*(f) = \langle f, Ig^* \rangle_H$  for all  $f \in H$  and  $g \in H^*$ , where  $g^*$  is the adjoint of  $g$ . Then we have

$$K(x, y) = \langle \text{ev}_y, \text{ev}_x \rangle_{H^*} = \langle I\text{ev}_y, I\text{ev}_x \rangle_H = \text{ev}_x(I\text{ev}_y) = (I\text{ev}_y)(x),$$

for all  $x, y \in E$ , and hence  $K(\cdot, y) = I\text{ev}_y \in H$ . From this it follows that

$$f(y) = \text{ev}_y(f) = \langle f, I\text{ev}_y \rangle_H = \langle f, K(\cdot, y) \rangle$$

for all  $y \in E$ . Thus,  $K$  is a reproducing kernel.  $\square$

## 1.2 RKHS of a kernel

**Definition 1.6.** *(Kernel, Feature Space, Feature Map) A function  $K : E \times E \rightarrow \mathbb{F}$  is called a kernel if there is a  $\mathbb{F}$ -Hilbert space  $H$  and a mapping  $\varphi : E \rightarrow H$  such that*

$$K(x, y) = \langle \varphi(y), \varphi(x) \rangle_H$$

for all  $x, y \in E$ . Such a  $\varphi$  is called a feature map, and  $H$  a feature space.

**Remark 1.7.** Proposition 1.5 tells us that every RK is indeed a kernel, and that the map  $E \ni x \mapsto \text{ev}_x \in H^*$  is a feature map with a feature space  $H^*$ . Every RKHS  $(H, K)$  also admits a more simple feature map  $\varphi_K$ , called a canonical feature map, given by

$$\varphi_K : E \ni x \mapsto K(\cdot, x) \in H.$$

This clearly shows that, given a kernel, neither feature space nor feature map are uniquely determined.  $\triangleleft$

**Definition 1.8.** (*Positive definite function*) Let  $E$  be a nonempty set. A function  $K : E \times E \rightarrow \mathbb{C}$  is called positive definite if for any  $n \in \mathbb{N}$  and for any  $a \in \mathbb{C}^n$  and  $x \in E^n$  there holds

$$\sum_{i=1}^n \sum_{j=1}^n a_i \bar{a}_j K(x_i, x_j) \geq 0, \quad (1)$$

where  $\bar{c}$  is the complex conjugate of  $c$ .

**Proposition 1.9.** (*Characterization of real positive definite function*) A real function  $K : E \times E \rightarrow \mathbb{R}$  is positive definite if and only if it has the following properties:

- (a)  $K$  is symmetric.
- (b) The defining inequality (1) holds for any  $\alpha \in \mathbb{R}^n$  instead of  $\mathbb{C}^n$ .

**Corollary 1.10.** (*Kernel is positive definite*) A kernel is positive definite.

*Proof.* For the case  $\mathbb{F} = \mathbb{C}$ , we have

$$\sum_{i=1}^n \sum_{j=1}^n a_i \bar{a}_j \overline{K(x_i, x_j)} = \sum_{i=1}^n \sum_{j=1}^n a_i \bar{a}_j \langle \varphi(x_i), \varphi(x_j) \rangle = \left\| \sum_{i=1}^n a_i \varphi(x_i) \right\|^2 \geq 0,$$

and hence  $\bar{K}$  as well as  $K$  are positive definite.  $\square$

**Proposition 1.11.** Every positive definite function  $K : E \times E \rightarrow \mathbb{C}$  satisfies

- (a)  $K(x, x) \geq 0$  for every  $x \in E$
- (b)  $K(x, y) = \overline{K(y, x)}$  for every  $x, y \in E$
- (c)  $\bar{K}$  is also positive definite, and conversely
- (d)  $|K(x, y)| \leq \sqrt{K(x, x)K(y, y)}$  for every  $x, y \in E$ .

*Proof.* (a) and (c) clearly hold. For  $\alpha, \beta \in \mathbb{C}$  and  $x, y \in E$ , we have

$$g(\alpha, \beta) := |\alpha|^2 K(x, x) + \alpha \bar{\beta} K(x, y) + \bar{\alpha} \beta K(y, x) + |\beta|^2 K(y, y) \geq 0.$$

Choose  $\alpha = \beta = 1$  and  $\alpha = i, \beta = 1$  to get

$$\begin{aligned} K(x, y) + K(y, x) &= g(1, 1) - K(x, x) - K(y, y) =: A \in \mathbb{R} \\ iK(x, y) - iK(y, x) &= g(i, 1) - K(x, x) - K(y, y) =: B \in \mathbb{R}. \end{aligned}$$

Therefore,

$$\begin{aligned} 2K(y, x) &= A + iB \\ 2K(x, y) &= A - iB, \end{aligned}$$

which proves (b). Finally, for  $x, y \in E$  with  $K(x, y) \neq 0$  and for  $r \in \mathbb{R}$ , (b) gives

$$0 \geq g(r, K(x, y)) = r^2 K(x, x) + 2r |K(x, y)|^2 + |K(x, y)|^2 K(y, y).$$

As RHS is quadratic in  $r$ , it must satisfy

$$|K(x, x)|^4 - |K(x, y)|^2 K(x, x) K(y, y) \leq 0,$$

from which (d) follows.  $\square$

**Theorem 1.12.** (*RKHS generated by inner product space*) Let  $H_0$  be the subspace of  $\mathbb{F}^E$ , equipped with an inner product  $\langle \cdot, \cdot \rangle_{H_0}$  with norm  $\|\cdot\|_{H_0}$ . Then there exists unique RKHS  $(H, K)$  that extends  $H_0$  in the sense that

- (a)  $H_0 \subset H \subset \mathbb{F}^E$  and the subspace topology of  $H_0$  in  $H$  coincides with the topology of  $(H_0, \|\cdot\|_{H_0})$

if and only if  $H_0$  satisfies the following conditions:

- (b) every evaluation functional  $\text{ev}_x$  is continuous in  $(H_0, \|\cdot\|_{H_0})$   
(c) any Cauchy sequence  $\{f_n\} \subset H_0$  converging pointwise to 0 converges to 0 also in  $H_0$ -norm.

*Proof.* Suppose such an extension  $H$  exists.  $H$  satisfies (b) by Theorem 1.2. Since  $H$  is complete, a Cauchy sequence  $\{f_n\} \subset H_0$  tends to some  $f$ , for which we have

$$f(x) = \text{ev}_x(f) = \lim_{n \rightarrow \infty} \text{ev}_x(f_n) = \lim_{n \rightarrow \infty} f_n(x) = 0$$

for every  $x \in E$ . Therefore,  $f$  is identically 0.

Conversely, suppose (b)(c) hold. Let  $H$  be the set of all functions  $f \in \mathbb{F}^E$  for which there exists a Cauchy sequence  $\{f_n\} \subset H_0$  converging pointwise to  $f$ . Clearly,  $H_0 \subset H \subset \mathbb{F}^E$ . The rest of proof consists of the following Lemmas.  $\square$

**Lemma 1.13.** *Let  $f, g \in H$  and let  $\{f_n\}$  and  $\{g_n\}$  be two Cauchy sequences in  $H_0$  that converge pointwise to  $f$  and  $g$  respectively.*

- (A) *The sequence  $\langle f_n, g_n \rangle_{H_0}$  is convergent.*
- (B) *The limit  $\lim_{n \rightarrow \infty} \langle f_n, g_n \rangle_{H_0}$  depends solely on  $f$  and  $g$ , independent of the choice of  $f_n$  and  $g_n$ .*
- (C)  *$\langle f, g \rangle_H := \lim_{n \rightarrow \infty} \langle f_n, g_n \rangle_{H_0}$  is an inner product on  $H$ .*

*Proof.* It follows from the definition of  $f_n$  and  $g_n$  that

$$\begin{aligned} |\langle f_n, g_n \rangle_{H_0} - \langle f_m, g_m \rangle_{H_0}| &= |\langle f_n - f_m, g_n \rangle - \langle f_m, g_n - g_m \rangle| \\ &\geq \|g_n\| \|f_n - f_m\| + \|f_m\| \|g_n - g_m\| \rightarrow 0, \end{aligned}$$

which proves (A). In order to verify (B), suppose  $\{f'_n\}$  and  $\{g'_n\}$  are also such approximating sequences. We then similarly deduce that

$$|\langle f_n, g_n \rangle - \langle f'_n, g'_n \rangle| \leq \|g_n\| \|f_n - f'_n\| + \|f'_n\| \|g_n - g'_n\|.$$

$\{f_n - f'_n\}$  and  $\{g_n - g'_n\}$  are Cauchy sequences tending pointwise to 0. Thus, assumption (c) gives  $\|f_n - f'_n\| \rightarrow 0$  and  $\|g_n - g'_n\| \rightarrow 0$ . So, (A) and (B) show that  $\langle f, g \rangle_H$  is well-defined. Note that if  $\langle f, f \rangle_H = 0$ , then for every  $x \in E$

$$f(x) = \text{ev}_x(f) = \lim_{n \rightarrow \infty} \text{ev}_x(f_n) = \lim_{n \rightarrow \infty} f_n(x) = 0,$$

and hence  $f \equiv 0$ . As the symmetry, positivity, linearity are quite obvious, we conclude that (C) is true.  $\square$

**Lemma 1.14.** (A) *Let  $f \in H$  and let  $\{f_n\} \subset H_0$  be a Cauchy sequence converging pointwise to  $f$ . Then  $f_n \rightarrow f$  also in  $H$ -norm.*

(B)  *$H_0$  is dense in  $H$ .*

*Proof.* (A): Fix  $\epsilon > 0$ . Choose  $N \in \mathbb{N}$  large enough so that

$$\|f_n - f_m\|_{H_0} < \epsilon$$

for all  $n, m > N$ . For fixed  $n$ ,  $\{f_n - f_m\}_{m \in \mathbb{N}}$  is a Cauchy sequence converging pointwise to  $f_n - f$ . Therefore, by definition of  $\langle \cdot, \cdot \rangle_H$ ,

$$\|f - f_n\|_H = \lim_{m \rightarrow \infty} \|f_n - f_m\|_{H_0} \leq \epsilon.$$

(B) is obvious from (A).  $\square$

**Lemma 1.15.** *Every evaluation functional  $\text{ev}_x$  is continuous on  $H$ .*

*Proof.* Fix  $x \in E$ . As a linear functional  $\text{ev}_x$  is assumed to be continuous on  $H_0$ , it admits unique continuous extension  $T_x$  onto the closure of  $H_0$  in  $H$ , that is, onto whole  $H$ , where we use the assumption (a) and Lemma2(B).  $T_x$  is also the evaluation functional on  $H$ . Indeed, for  $f$  and  $f_n$  as in Lemma1.14, we have

$$T_x(f) = \lim_{n \rightarrow \infty} \text{ev}_x(f_n) = \lim_{n \rightarrow \infty} f_n(x) = f(x).$$

□

**Lemma 1.16.**  *$H$  is a RKHS satisfying (a) in Theorem1.12. Consequently,  $H$  is isomorphic to the completion of  $H_0$  and thus unique RKHS that meets the requirement.*

*Proof.* We first prove that  $H$  is actually a RKHS. In light of Theorem1.2 and Lemma1.15, it suffices to show that  $H$  is complete. Let  $\{f_n\}$  be a Cauchy sequence in  $H$ . Let  $x \in E$ .  $\{f_n(x)\}$  is also a Cauchy in  $\mathbb{F}$ , and hence converges to some  $f(x)$ . By Lemma1.14, for every  $n \in \mathbb{N}$ , there is  $g_n \in H_0$  such that  $\|f_n - g_n\|_H < n^{-1}$ . In view of the inequality

$$\|f - f_n\|_H \leq \|f - g_n\|_H + \|g_n - f_n\|_H,$$

it suffices to prove that  $\|f - g_n\|_H \rightarrow 0$ . To this end, we show that  $\{g_n\}$  is a Cauchy sequence converging pointwise to  $f$  (and then apply Lemma1.14).

For fixed  $x \in E$ , we have

$$\begin{aligned} |g_n(x) - f(x)| &\leq |g_n(x) - f_n(x)| + |f_n(x) - f(x)| \\ &= |\text{ev}_x(g_n - f_n)| + |f_n(x) - f(x)| \rightarrow 0 \end{aligned}$$

as  $n \rightarrow \infty$ , since  $\text{ev}_x$  is continuous and  $f_n(x) \rightarrow f(x)$  pointwise. Moreover,

$$\begin{aligned} \|g_n - g_m\|_{H_0} &= \|g_n - g_m\|_H \\ &\leq \|g_n - f_n\| + \|f_n - f_m\| + \|g_m - f_m\| \\ &= n^{-1} + \|f_n - f_m\| + m^{-1} \rightarrow 0 \end{aligned}$$

when  $n, m \rightarrow \infty$ , as required.  $H$  is isomorphic to the completion of  $H_0$  since  $H_0$  is dense in  $H$ . Finally, suppose  $H'$  is also an qualifying RKHS. As  $H$  is isomorphic to  $H'$  there exists an isometric isomorphism between them. Furthermore, by Lemma1.14 and Lemma1.15, they share the same evaluation functionals  $\{\text{ev}_x\}_{x \in E}$  and hence the same reproducing kernel by Proposition1.5. Thus, they must coincide by Proposition1.4. □

**Theorem 1.17.** (Moore-Aronszajn) For arbitrary positive definite function  $K : E \times E \rightarrow \mathbb{F}$ , there exists unique RKHS  $H$  that has  $K$  as its reproducing kernel. Moreover, the subspace  $H_0$  spanned by  $\{K(\cdot, x)\}_{x \in E}$  is dense in  $H$ .

*Proof.* Define an inner product  $\langle \cdot, \cdot \rangle_{H_0}$  on  $H_0$  by setting

$$\langle f, g \rangle_{H_0} := \sum_{i=1}^n \sum_{j=1}^n \alpha_i \beta_j K(y_i, x_j),$$

where  $f = \sum_{i=1}^n \alpha_i K(\cdot, x_i)$  and  $g = \sum_{j=1}^n \beta_j K(\cdot, y_j)$ . Let us observe

$$\sum_{i=1}^n \sum_{j=1}^n \alpha_i \beta_j K(y_i, x_j) = \sum_{i=1}^n \alpha_i \overline{g(x_i)} = \sum_{j=1}^n \overline{\beta_j} f(y_j),$$

and therefore that the value  $\langle f, g \rangle_{H_0}$  is determined by solely by  $f$  and  $g$ , independent of the choice of representing linear combination. Choosing  $g = K(\cdot, x)$  yields

$$\langle f, K(\cdot, x) \rangle_{H_0} = \sum_{i=1}^n \alpha_i \overline{K(x, x_i)} = \sum_{i=1}^n \alpha_i K(x, x_i) = f(x).$$

So,  $K$  fulfills the reproducing identity under  $\langle \cdot, \cdot \rangle_{H_0}$ . In particular,

$$\|K(\cdot, x)\|_{H_0} = \langle K(\cdot, x), K(\cdot, x) \rangle_{H_0} = K(x, x) \geq 0.$$

From this definiteness of  $\langle \cdot, \cdot \rangle_{H_0}$  follows; indeed, if  $\langle f, f \rangle_{H_0} = 0$ , then we have

$$|f(x)| = |\langle f, K(\cdot, x) \rangle| \leq \langle f, f \rangle_{H_0}^{1/2} \|K(\cdot, x)\|_{H_0}^{1/2} = 0,$$

for every  $x \in E$ , implying  $f \equiv 0$ . We then conclude that  $\langle \cdot, \cdot \rangle_{H_0}$  is in fact an inner product on  $H_0$  as the other requirements are easy to check.

We now show that  $H_0$  fulfills the conditions of Theorem 1.12. First, each  $\text{ev}_x$  is continuous on  $H_0$ ; in fact, for  $f, g \in H_0$ ,

$$|\text{ev}_x(f) - \text{ev}_x(g)| = |\langle f - g, K(\cdot, x) \rangle_{H_0}| \leq \|f - g\|_{H_0} \|K(\cdot, x)\|_{H_0}^{1/2}$$

for every  $x \in E$ . To verify the other condition, let  $\{f_n\}$  be a Cauchy sequence in  $H_0$  converging pointwise to 0. Let  $B > 0$  be an upper bound for  $\|f_n\|_{H_0}$ . For  $\epsilon > 0$  and large  $N \in \mathbb{N}$  we have

$$\|f_n - f_N\| < \frac{\epsilon}{B}$$



for all  $n \geq N$ . We may write

$$f_N = \sum_{i=1}^k K(\cdot, x_i)$$

for some  $\alpha_i \in \mathbb{F}$  and  $x_i \in E$ , and for some fixed  $k$ . It then follows that

$$\|f_n\|_{H_0}^2 = \langle f_n - f_N, f_n \rangle_{H_0} + \langle f_N, f_n \rangle_{H_0} \leq \epsilon + \sum_{i=1}^k f(x_i)$$

for  $n \geq N$ , and hence  $\|f_n\| \rightarrow 0$  as  $n \rightarrow \infty$ . Therefore, there is a RKHS  $H$  that has  $H_0$  as a dense subspace. Furthermore, for each  $f \in H$  there is  $\{f_n\} \subset H_0$  such that  $f_n \rightarrow f$  pointwise as well as in  $H$ -norm, for which we have

$$f(x) = \lim_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} \langle f_n, K(\cdot, x) \rangle_{H_0} = \langle f, K(\cdot, x) \rangle_H,$$

for every  $x \in E$ . Thus,  $K$  is a reproducing kernel of  $H$ . Uniqueness follows from Proposition 1.4.  $\square$

**Theorem 1.18.** (*Characterization of positive definite function*) A function  $K : E \times E \rightarrow \mathbb{F}$  is positive definite (and thus a reproducing kernel of some RKHS) if and only if  $K$  is a kernel, that is, if and only if there exists some mapping  $\varphi$  of  $E$  into some  $\mathbb{F}$ -Hilbert space  $H$  such that

$$K(x, y) = \langle \varphi(y), \varphi(x) \rangle_H$$

for all  $x, y \in E$ .

*Proof.* If  $(H, K)$  is the RKHS generated by positive definite function  $K$ , then the canonical feature map  $\varphi_K : E \ni x \mapsto K(\cdot, x) \in H$  obviously qualifies. The converse follows from Corollary 1.10.  $\square$

**Remark 1.19.** Theorem 1.18 implies that RKHS  $(H, K)$  is a natural feature space.  $\triangleleft$

Theorem 1.18 is a powerful tool to construct a positive definite function as well as to prove a given function is a kernel *if we can find an appropriate feature space*, as the following example illustrates.

**Example 1.20.** Let us show that  $K(x, y) = \min(x, y)$ ,  $x, y \in \mathbb{R}_+$  is positive definite. Let  $H := L^2(\mathbb{R}_+, \mu)$  be the space of all square integrable functions on  $\mathbb{R}_+$  with respect to a  $\sigma$ -finite measure  $\mu$ . It is well-known that  $H$  is a Hilbert space with the inner product  $\langle f, g \rangle_H := \int_{\mathbb{R}_+} f \bar{g} d\mu$ . Then we have

$$K(x, y) = \int_{\mathbb{R}_+} 1_{[0, y]}(t) 1_{[0, x]}(t) d\mu(t) = \langle \varphi(y), \varphi(x) \rangle_H,$$

where  $\varphi : E \ni x \mapsto 1_{[0,x]}(\cdot) \in H$  is the feature map, and  $1_A(\cdot)$  is the indicator function of  $A$ . Therefore,  $K$  is positive definite.  $\triangleleft$

The next Theorem relates a feature map (therefore a kernel) and the RKHS the corresponding positive definite function generates.

**Theorem 1.21.** (*RKHS generated by feature map*) Let  $E \neq \emptyset$ . Suppose  $K$  is a positive definite kernel with a feature space  $H_0$  and a feature map  $\varphi_0 : E \rightarrow H_0$ . Then the Hilbert space

$$H := \{f : E \rightarrow \mathbb{F} \mid \exists w \in H_0 : f(x) = \langle w, \varphi_0(x) \rangle_{H_0} \forall x \in E\}$$

equipped with the norm

$$\|f\|_H := \inf\{\|w\|_{H_0} : w \in H_0, f = \langle w, \varphi_0(\cdot) \rangle_{H_0}\}$$

is the RKHS generated by  $K$ , and  $H$  and  $\|\cdot\|_H$  are determined independent of the choice of feature space  $H_0$  and feature map  $\varphi_0$ . Moreover, the function

$$V : H_0 \ni w \mapsto \langle w, \varphi_0(\cdot) \rangle_{H_0} \in H$$

is an isometric isomorphism on  $(\ker V)^\perp$ .

*Proof.* In light of Theorem 1.17, it suffices to prove that  $H$  is RKHS with reproducing kernel  $K$ . The property of  $V$  are automatically obtained in the process. It is easy to verify that  $\|\cdot\|_H$  is actually a norm on  $H$ . As  $\ker V$  is closed subspace of  $H_0$ , we get the orthogonal decomposition  $H_0 = \ker V \oplus (\ker V)^\perp$ . Let  $H_1 := (\ker V)^\perp$  and the restriction of  $V$  onto  $H_1$  be denoted by  $V_1$ . Since every  $f \in H$  can be written as  $f = V(w_0 + w_1) = V_1 w_1$ , with  $w_0 \in \ker V$ ,  $w_1 \in H_1$ , we see that  $V_1 : H_1 \rightarrow H$  is an isomorphism. Similarly, we have

$$\begin{aligned} \|f\|_H &= \inf\{\|w_0 + w_1\|_{H_0}^2 : w_0 \in \ker V, w_1 \in H_1, w_0 + w_1 \in V^{-1}(\{f\})\} \\ &= \inf\{\|w_0\|^2 + \|w_1\|^2 : w_0 \in \ker V, w_1 \in H_1, w_0 + w_1 \in V^{-1}(\{f\})\} \\ &= \|V_1^{-1}(f)\|_{H_1} \left( := \|V_1^{-1}(f)\|_{H_0} \right). \end{aligned}$$

From this equation, we conclude that  $V_1 : H_1 \rightarrow H$  is an isometric isomorphism, as required, and that  $H$  is a Hilbert space.

It remains to show that  $K$  qualifies as the reproducing kernel. Observe

$$K(\cdot, x) = \langle \varphi_0(x), \varphi_0(\cdot) \rangle_{H_0} = V\varphi_0(x) \in H.$$

Moreover, the fact  $\langle w, \varphi_0(x) \rangle_{H_0} = 0$  for all  $w \in \ker V$  implies

$$f(x) = \langle V_1^{-1}f, \varphi_0(x) \rangle_{H_0} = \langle f, V\varphi_0(x) \rangle_H = \langle f, K(\cdot, x) \rangle_H$$

for all  $f \in H$  and  $x \in E$ .  $\square$

### 1.3 Basic Properties of Kernel

## 2 Reconstruction of RKHS

### 2.1 Reconstruction via Restriction

### 2.2 Reconstruction via Operator

### 2.3 Reconstruction via Sum and Product

## 3 Inheritance from Kernel to RKHS

### 3.1 Measurability of RKHS

### 3.2 Separability of RKHS

## 4 Mercer Representation

## Further Readings

## References

- [1] Alain Berlinet and Christine Thomas-Agnan. *Reproducing Kernel Hilbert Spaces in Probability and Statistics*. Springer, 1st edition, 2004.
- [2] Andreas Christmann and Ingo Steinwart. *Support vector machines*. Springer, 1st edition, 2008.
- [3] Felipe Cucker and Ding Xuan Zhou. *Learning theory: An approximation theory viewpoint*. Cambridge University Press, 1st edition, 2007.
- [4] Vern I. Paulsen and Mrinal Raghupathi. *An Introduction to the Theory of Reproducing Kernel Hilbert Spaces*. Cambridge University Press, 1st edition, 2016.
- [5] Bernhard Schölkopf and Alexander J. Smola. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. The MIT Press, 1st edition, 2001.