

# What's this?

This document consists of a collection of basic results about Reproducing Kernel Hilbert Space useful in machine learning context. Nearly every statement and proof is borrowed mainly from [1] [2]. [3] [4] [5]

## Overview

### 1 Basic Property of RKHS

**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.  $\square$

**Definition 1.4.** (*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

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

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

**Proposition 1.5.** Suppose  $\varphi$  is a mapping of a set  $E$  into a Hilbert space  $H$ . Then the mapping  $K : E \times E \ni (x, y) \mapsto \langle \varphi(x), \varphi(y) \rangle \in \mathbb{C}$  is positive definite.

*Proof.* For  $a$  and  $x$  taken as in the definition, we have

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

$\square$

**Proposition 1.6.** Every reproducing kernel is positive definite.

*Proof.*  $\sum_{i=1}^n \sum_{j=1}^n a_i \bar{a}_j K(x_i, x_j) = \left\| \sum_{i=1}^n a_i K(\cdot, x_i) \right\|^2$ .  $\square$

**Proposition 1.7.** 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
- (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 (c) follows.  $\square$

**Theorem 1.8.** *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 1.9.** *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 1.10.** (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 1.11.** 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 Lemma1.10, we have

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

□

**Lemma 1.12.**  *$H$  is a RKHS.*

*Proof.* In light of Theorem1.2 and Lemma1.11, 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 Lemma1.10, 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 Lemma1.10).

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  $n \rightarrow \infty$ , as required. □

## 2 Reconstruction of RKHS

### 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.