C*環を用いたカーネル法の拡張

2. C*環を用いたカーネル法の理論的側面

Yuka Hashimoto

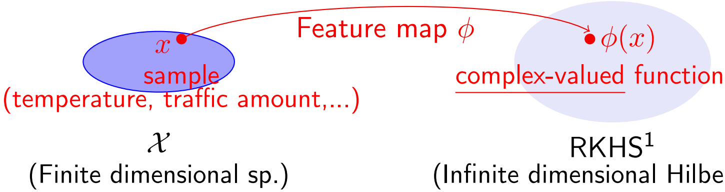
NTT / RIKEN AIP

January 9th, 2026

- Y. Hashimoto, I. Ishikawa, M. Ikeda, F. Komura, T. Katsura, and Y. Kawahara, JMLR, 22(267):1–56. (updated version : arXiv:2101.11410v2)
- Y. Hashimoto, F. Komura, and M. Ikeda, Matrix and Operator Equations, pp. 1–27.

Contents

Background: Kernel methods



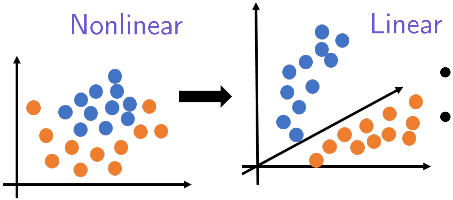
x Feature map ϕ $\phi(x)$

sample
(temperature, traffic amount,...)

complex-valued function

$\mathcal{X}$
(Finite dimensional sp.)

RKHS[1]
(Infinite dimensional Hilbert sp.)



- Kernel PCA, Kernel SVM
- Learning complex-valued functions

Advantages of RKHS

- Nonlinearity in the original space is transformed into a linear one.
- We can compute inner products in RKHS exactly by computers.

---

[1] Schölkopf and Smola, MIT Press, Cambridge, 2001

Background: Reproducing kernel Hilbert space (RKHS)

Let $\mathcal{X}$ be a set. A map $k : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$ is called a positive definite kernel if it satisfies:

1. $k(x, y) = \overline{k(y, x)}$ for $x, y \in \mathcal{X}$ and
2. $\sum_{t,s=1}^{n} \overline{c_t} k(x_t, x_s) c_s \geq 0$ for $n \in \mathbb{N}$, $c_1, \ldots, c_n \in \mathbb{C}$, $x_1, \ldots, x_n \in \mathcal{X}$.

$\phi(x) := k(\cdot, x)$ ($\phi : \mathcal{X} \to \mathbb{C}^{\mathcal{X}}$: feature map associated with k),

$$\mathcal{H}_{k,0} := \left\{ \sum_{t=1}^{n} \phi(x_t) c_t \,\middle|\, n \in \mathbb{N},\ c_t \in \mathbb{C},\ x_t \in \mathcal{X} \right\}. \tag{1}$$

We can define an inner product $\langle \cdot, \cdot \rangle_k : \mathcal{H}_{k,0} \times \mathcal{H}_{k,0} \to \mathbb{C}$ as

$$\left\langle \sum_{s=1}^{n} \phi(x_s) c_s, \sum_{t=1}^{l} \phi(y_t) d_t \right\rangle_k := \sum_{s=1}^{n} \sum_{t=1}^{l} \overline{c_s} k(x_s, y_t) d_t. \tag{2}$$

Reproducing property: $\langle \phi(x), v \rangle_k = v(x)$ for $v \in \mathcal{H}_k$ and $x \in \mathcal{X}$

RKHS $\mathcal{H}_k$: completion of $\mathcal{H}_{k,0}$

Background: Representer theorem in RKHSs

The representer theorem guarantees that solutions of a minimization problem are represented only with given samples[2].

$\mathcal{H}_k$: RKHS
$\mathbb{R}_+ := \{a \in \mathbb{R} \mid a \geq 0\}$

Theorem 1 Representer theorem in RKHSs

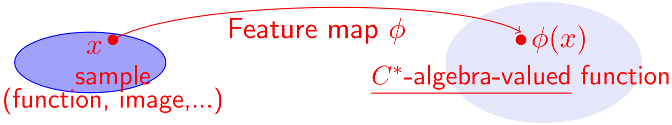
Let $x_1, \ldots, x_n \in \mathcal{X}$ and $a_1, \ldots, a_n \in \mathbb{C}$. Let $h : \mathcal{X} \times \mathbb{C}^2 \to \mathbb{R}_+$ be an error function and $g : \mathbb{R}_+ \to \mathbb{R}_+$ satisfy $g(c) < g(d)$ for $c < d$. Then, any $u \in \mathcal{H}_k$ minimizing $\sum_{i=1}^n h(x_i, a_i, u(x_i)) + g(\|u\|_k)$ admits a representation of the form $\sum_{i=1}^n \phi(x_i) c_i$ for some $c_1, \ldots, c_n \in \mathbb{C}$.

The result can be applied to supervised problems.

---

[2] Schölkopf et al., COLT 2001.

Goal: Generalization of data analysis in RKHS to RKHM



$\mathcal{X}$
(Structured data sp.)

RKHM
(Infinite dimensional Hilbert C^*-module)

Nonlinear

Linear + C^*-algebra-valued inner product

Advantages of RKHM:

- C^*-algebra-valued inner products extract information of structures.

We constructed a framework of data analysis with RKHM.

- We can reconstruct existing RKHSs by using RKHMs.
- We have shown fundamental properties for data analysis in RKHMs (e.g. representer theorem, kernel mean embedding).

C^*-algebra and von Neumann-algebra

C^*-algebra : Banach space equipped with a product & an involution * + C^*-property

e.g.

- $C(\mathcal{Z})$ for a compact space $\mathcal{Z}$
  Norm : sup norm, **Product** : pointwise product,
  Involution : pointwise complex conjugate
- $\mathcal{K}(\mathcal{H}) = \{$compact operators on a Hilbert space $\mathcal{H}\}$
  Norm : operator norm, **Product** : composition, **Involution** : adjoint

Von Neumann-algebra : C^*-algebra that is closed in the strong operator topology

e.g.

- $L^\infty(\mathcal{Z})$ for a measure space $\mathcal{Z}$
- $\mathcal{B}(\mathcal{H}) = \{$bounded linear operators on a Hilbert space $\mathcal{H}\}$

Positivity and order in C^*-algebras

For optimization, we need the notion of "positive" and order.

$\mathcal{A}$: C^*-algebra

Definition 1 Positive

Let $a \in \mathcal{A}$. If $a = b^*b$ for some $b \in \mathcal{A}$, then a is called positive. We put $\mathcal{A}_+ = \{a \in \mathcal{A} \mid a \text{ is positive}\}$.

We can define a (partial) order $\leq_{\mathcal{A}}$ in $\mathcal{A}$ by
"$a \leq_{\mathcal{A}} b$ if and only if $b - a$ is positive".

We denote $a <_{\mathcal{A}} b$ if $b - a$ is positive and not zero.

We consider supremum, maximum, infimum, and minimum in $\mathcal{A}$ with respect to the order $\leq_{\mathcal{A}}$.

Hilbert C^*-module

$\mathcal{A}$: C^*-algebra
$\mathcal{M}$: right $\mathcal{A}$-module ($u \in \mathcal{M}$, $c \in \mathcal{A} \to uc \in \mathcal{M}$)

Definition 2 $\mathcal{A}$-valued inner product

A map $\langle \cdot, \cdot \rangle : \mathcal{M} \times \mathcal{M} \to \mathcal{A}$ is called an $\mathcal{A}$-valued inner product if it satisfies the following properties for $u, v, w \in \mathcal{M}$ and $c, d \in \mathcal{A}$:

1. $\langle u, vc + wd \rangle = \langle u, v \rangle c + \langle u, w \rangle d$,
2. $\langle v, u \rangle = \langle u, v \rangle^*$,
3. $\langle u, u \rangle \geq 0$ (positive) and if $\langle u, u \rangle = 0$ then $u = 0$.

→ *$\mathcal{A}$-valued absolute value* $|u| := \langle u, u \rangle^{1/2}$ → Norm $\|u\| := \| \langle u, u \rangle \|_{\mathcal{A}}^{1/2}$

Hilbert C^*-module $\mathcal{M}$[3]: complete $\mathcal{A}$-module equipped with an $\mathcal{A}$-valued inner-product

---

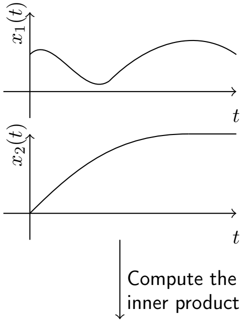
[3] Lance, Cambridge University Press, 1995.

Advantages of RKHM (functional data)

Algorithms in RKHS

x_1,x_2 : Functional data

$x_1, x_2 \in \mathcal{H}$

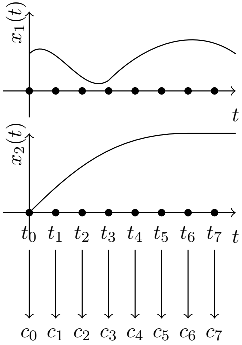


Compute the inner product

$\langle x_1, x_2 \rangle_{\mathcal{H}} \in \mathbb{C}$

Degenerates information along t

$x_1(t), x_2(t) \in \mathbb{C}$

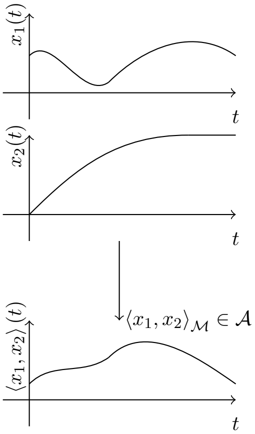


$c_i = \langle x_1(t_i), x_2(t_i) \rangle_{\mathcal{X}} \in \mathbb{C}$

Fails to capture continuous behavior (derivatives, total variation, frequency components,...)

Algorithms in RKHM

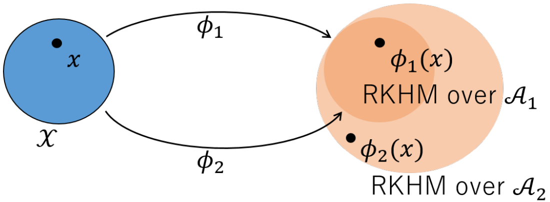
$x_1, x_2 \in \mathcal{M}$



Capture and control continuous behavior

Advantages of RKHM

- Enlarge representation spaces using C^*-algebras (e.g. use the C^*-algebra of continuous functions for functional data).



$\mathcal{A}_1, \mathcal{A}_2$: C^*-algebra

- Make use of the product structure.
  e.g. polynomial kernel $k(x, y) = x^* y + x^* x^* y y$ $(x, y \in \mathcal{A}_1$ or $\mathcal{A}_2)$
- Use the operator norm to alleviate the dependency of the error on data dimension.

Review of reproducing kernel Hilbert C^*-module

$\mathcal{A}$: C^*-algebra

RKHS ($\mathcal{H}_k$):

- $\mathbb{C}$-valued positive definite kernel k
- $\mathbb{C}$-valued functions
- $\mathbb{C}$-valued inner product

RKHM over $\mathcal{A}$ ($\mathcal{M}_k$):

- $\mathcal{A}$-valued positive definite kernel k
- $\mathcal{A}$-valued functions
- $\mathcal{A}$-valued inner product

Reproducing kernel Hilbert C^*-module (RKHM)

Let $\mathcal{X}$ be a set. A map $k : \mathcal{X} \times \mathcal{X} \to \mathcal{A}$ is called an $\mathcal{A}$-valued positive definite kernel if it satisfies:

1. $k(x, y) = k(y, x)^*$ for $x, y \in \mathcal{X}$ and
2. $\sum_{t,s=1}^{n} c_t^* k(x_t, x_s) c_s \geq 0$ for $n \in \mathbb{N}$, $c_1, \ldots, c_n \in \mathcal{A}$, $x_1, \ldots, x_n \in \mathcal{X}$.

$\phi(x) := k(\cdot, x)$ ($\phi : \mathcal{X} \to \mathcal{A}^{\mathcal{X}}$: feature map associated with k),

$$\mathcal{M}_{k,0} := \left\{ \sum_{t=1}^{n} \phi(x_t) c_t \,\middle|\, n \in \mathbb{N},\ c_t \in \mathcal{A},\ x_t \in \mathcal{X} \right\}. \tag{3}$$

We can define an $\mathcal{A}$-valued inner product $\langle \cdot, \cdot \rangle_k : \mathcal{M}_{k,0} \times \mathcal{M}_{k,0} \to \mathcal{A}$ as

$$\left\langle \sum_{s=1}^{n} \phi(x_s) c_s, \sum_{t=1}^{l} \phi(y_t) d_t \right\rangle_k := \sum_{s=1}^{n} \sum_{t=1}^{l} c_s^* k(x_s, y_t) d_t. \tag{4}$$

Reproducing property: $\langle \phi(x), v \rangle_k = v(x)$ for $v \in \mathcal{M}_k$ and $x \in \mathcal{X}$

RKHM $\mathcal{M}_k$: completion of $\mathcal{M}_{k,0}$

Representer theorem in RKHMs

To generalize complex-valued supervised problems to $\mathcal{A}$-valued ones, we show a representer theorem.

$\mathcal{M}_k$: RKHM over $\mathcal{A}$, $|\cdot|_k$: absolute value in $\mathcal{M}_k$
$\mathcal{A}_+ := \{a \in \mathcal{A} \mid {}^\exists b \in \mathcal{A}$ such that $a = b^*b\}$

Theorem 2 Representer theorem in RKHMs

Let $\mathcal{A}$ be a unital C^*-algebra, $x_1, \ldots, x_n \in \mathcal{X}$ and $a_1, \ldots, a_n \in \mathcal{A}$. Let $h : \mathcal{X} \times \mathcal{A}^2 \to \mathcal{A}_+$ be an error function and $g : \mathcal{A}_+ \to \mathcal{A}_+$ satisfy $g(c) < g(d)$ for $c < d$. If $\mathrm{Span}_{\mathcal{A}}\{\phi(x_i)\}_{i=1}^n$ is closed, any $x \in \mathcal{M}_k$ minimizing $f(w) := \sum_{i=1}^n h(x_i, a_i, x(x_i)) + g(|x|_k)$ admits a representation of the form $\sum_{i=1}^n \phi(x_i)c_i$ for some $c_1, \ldots, c_n \in \mathcal{A}$.

Key point of the proof:
For a Hilbert C^*-module $\mathcal{M}$ over a unital C^*-algebra $\mathcal{A}$ and any finitely generated closed submodule $\mathcal{V}$ of $\mathcal{M}$, $x \in \mathcal{M}$ is decomposed into $x = x_1 + x_2$ where $x_1 \in \mathcal{V}$ and $x_2 \in \mathcal{V}^\perp$.

Approximate representer theorem in RKHMs

If $\mathcal{A}$ is a von Neumann algebra, we can show an approximate representer theorem under mild conditions.

Theorem 3 Approximate representer theorem in RKHMs

Let $\mathcal{A}$ be a von Neumann-algebra, $x_1, \ldots, x_n \in \mathcal{X}$ and $a_1, \ldots, a_n \in \mathcal{A}$. Let $h : \mathcal{X} \times \mathcal{A}^2 \to \mathcal{A}_+$ be a Lipschitz continuous error function with Lipschitz constant L and $g : \mathcal{A}_+ \to \mathcal{A}_+$ satisfy $g(c) < g(d)$ for $c < d$. Assume $f(x) := \sum_{i=1}^{n} h(x_i, a_i, x(x_i)) + g(|x|_k)$ has a minimizer x. Then, for any $\epsilon > 0$, there exists $v \in \mathcal{M}_k$ of the form $\sum_{i=1}^{n} \phi(x_i) c_i$ such that $\|f(v) - f(x)\|_{\mathcal{A}} \leq Ln\epsilon \|x\|_{\mathcal{A}}$.

Key point of the proof:

If $\mathcal{A}$ is a von Neumann-algebra, we can apply the Gram–Schmidt orthonormalization to construct a module approximating the module generated by $\{\phi(x_i)\}_{i=1}^{n}$.

Conclusion

- RKHM is a natural generalization of RKHS.
- We showed a representer theorem and an approximate representer theorem in RKHMs and in RKHMs.
- RKHMs are useful for analyzing image data and functional data.