数理解析研究所講究録
第1980巻 2016年 95-102

An introduction of the book "Theory of Reproducing Kernels and Applications"

Yoshihiro Sawano (Tokyo Metropolitan University) and
Saburou Saitoh (Gunma University)

December 12, 2015

Abstract

In this paper, we introduce some fundamental theorems in the book "Theory of Reproducing Kernels and Applications" to explain what this book is oriented to. The book will be published from Springer. We pick up several theorems and explain what they are used for in our book. The detailed applications can be found in the book.

1 Definition of reproducing kernel Hilbert spaces

We start with the definition of reproducing kernel Hilbert spaces.

Definition 1. Let E be an arbitrary abstract nonempty set. Denote by $\mathcal{F}(E)$ the set of all complex-valued functions on E. *A reproducing kernel Hilbert space* (RKHS for short) on the set E is a Hilbert space $\mathcal{H} \subset \mathcal{F}(E)$ coming with a function $K : E\times E \to \mathcal{H}$, which is called *the reproducing kernel*, enjoying *the reproducing property* that

$$K_p \equiv K(\cdot, p) \in \mathcal{H} \tag{1.1}$$

for all $p \in E$ and that the representation

$$f(p) = \langle f, K_p\rangle_{\mathcal{H}} \tag{1.2}$$

holds for all $p \in E$ and all $f \in \mathcal{H}$. Denote by $(H_K =)H_K(E)$ the Hilbert space $\mathcal{H}$ whose corresponding reproducing kernel function is K.

The following two theorems show that the fundamental property of the space $H_K(E)$ and the correspondence $K \mapsto H_K(E)$.

Theorem 1.1. *Suppose that H is a Hilbert space consisting of functions on a set E.*

1. *The Hilbert space H is realized as a reproducing kernel Hilbert space $H_K(E)$ with a function $K : E \times E \to \mathbb{C}$ if and only if the embedding $H \subset \mathcal{F}(E)$ is continuous.*

2. *If a sequence* $\{f_j\}_{j=1}^{\infty}$ *in* $H_K(E)$ *converges to* f *in* $H_K(E)$*, then*

$$\lim_{j\to\infty} f_j(p) = f(p) \tag{1.3}$$

for all $p \in E$*. Furthermore, on any subset of* E *on which* $p \mapsto K(p,p)$ *is bounded, its convergence is uniform on there.*

In general, a complex-valued function $k : E \times E \to \mathbb{C}$ is called a *positive definite quadratic form function* on the set E, or shortly, *positive definite function*, when it satisfies the property that, for an arbitrary function $X : E \to \mathbb{C}$ and for any finite subset F of E,

$$\sum_{p,q\in F} \overline{X(p)}X(q)k(p,q) \geq 0.$$

Theorem 1.2. *For any positive definite quadratic form function* $K : E\times E \to \mathbb{C}$*, there exists a uniquely determined reproducing kernel Hilbert space* $H_K = H_K(E)$ *admitting the reproducing kernel* K *on* E*.*

2 Fundamental theorems

2.1 Basic operations on reproducing kernel Hilbert spaces

The next theorems are on fundamental operations of RKHS.

Theorem 2.1 (Restriction of RKHS). *Suppose that* $K : E\times E \to \mathbb{C}$ *is a positive definite quadratic form function on a set* E*. Let* E_0 *be a subset of* E*. Then the reproducing kernel Hilbert space that* $K|E_0 \times E_0 : E_0 \times E_0 \to \mathbb{C}$ *defines is given by*

$$H_{K|E_0\times E_0}(E_0) = \{f \in \mathcal{F}(E_0) \,:\, f = \tilde{f}|E_0 \text{ for some } \tilde{f} \in H_K(E)\}. \tag{2.1}$$

Furthermore, the norm is expressed in terms of the one of $H_K(E)$:

$$\|f\|_{H_{K|E_0\times E_0}(E_0)} = \min\{\|\tilde{f}\|_{H_K(E)} \,:\, \tilde{f} \in H_K(E),\, f = \tilde{f}|E_0\}. \tag{2.2}$$

Theorem 2.2. *Let* $K_1, K_2 : E \times E \to \mathbb{C}$ *be positive definite. Set* $K \equiv K_1 + K_2$*.*

1. *We have*

$$H_K(E) = \{f_1 + f_2 \in \mathcal{F}(E) \,:\, f_1 \in H_{K_1}(E), f_2 \in H_{K_2}(E)\}$$

as a set. So as a linear space, we have $H_{K_1+K_2}(E) = H_{K_1}(E) + H_{K_2}(E)$*.*

2. *The norm of* $H_K(E)$ *has another expression in terms of those of* $H_{K_1}(E)$ *and* $H_{K_2}(E)$:

$$\|f\|_{H_K} = \min_{\substack{f_1\in H_{K_1}(E),\, f_2\in H_{K_2}(E),\\ f=f_1+f_2}} \sqrt{\|f_1\|^2_{H_{K_1}(E)} + \|f_2\|^2_{H_{K_2}(E)}}. \tag{2.3}$$

Theorem 2.3. *Let $K_0, K : E \times E \to \mathbb{C}$ be positive definite quadratic form functions. Then the following are equivalent:*

1. *The Hilbert space $H_{K_0}(E)$ is a subset of $H_K(E)$;*
2. *there exists $\gamma > 0$ such that $K_0 \ll \gamma^2 K$.*

If these conditions hold, then the embedding $H_{K_0}(E) \subset H_K(E)$ is actually continuous and its norm is given by $M = \inf\{\gamma > 0 \,:\, K_0 \ll \gamma^2 K\}$.

Theorem 2.4. *Let $\{E_1, E_2\}$ be a partition of a set E. Suppose that we are given a reproducing kernel K on E. Denote by K_1, K_2 the restrictions of K to $E_1 \times E_1$ and $E_2 \times E_2$ respectively. Then the following are equivalent:*

(1) $K|E_1 \times E_2 \equiv 0$;

(2) $f \in H_K(E) \mapsto (f|E_1, f|E_2) \in H_{K_1}(E_1) \oplus H_{K_2}(E_2)$ *is an isomorphism.*

If one of these conditions is fulfilled, then we have

$$K(x,y) = \begin{cases} K_1(x,y) & x, y \in E_1, \\ K_2(x,y) & x, y \in E_2, \\ 0 & \textit{otherwise}. \end{cases} \tag{2.4}$$

Theorem 2.5. *Let $K_1 : E_1 \times E_1 \to \mathbb{C}$ and $K_2 : E_2 \times E_2 \to \mathbb{C}$ be positive definite quadratic form functions. Then $K_1 \otimes K_2 : E_1 \times E_2 \times E_1 \times E_2 \to \mathbb{C}$ is a positive definite quadratic form function and*

$$H_{K_1}(E_1) \otimes H_{K_2}(E_2) = H_{K_1 \otimes K_2}(E_1 \times E_2). \tag{2.5}$$

Theorem 2.6. *Suppose that $K_1, K_2 : E \times E \to \mathbb{C}$ are positive definite quadratic form functions. Then so is the pointwise product $K \equiv K_1 \cdot K_2 : E \times E \to \mathbb{C}$.*

Theorem 2.7. *Let n be a natural number and $H_K(E)$ be a reproducing kernel Hilbert space on E. Then the function $\wedge^n K$, given by*

$$\wedge^n K(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n) \equiv \frac{1}{n!} \det\{K(x_i, y_j)\}_{i,j=1,2,\ldots,n}$$

for $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n \in E$, is positive definite and a reproducing kernel of the space $\wedge^n H_K(E)$.

2.2 Transforms of reproducing kernel Hilbert spaces

The next theorem is used to describe the inverse function of a general mapping φ even for those whose inverse does not exit; that is, the inverse is a multipy-valued case, – indeed, we can consider transforms for arbitrary mappings. Positive definite quadratic forms are preserved under arbitrary mappings, and so, we can consider the transform of a reproducing kernel HIlbert space by any mapping. We use the following theorem:

Theorem 2.8 (Pullback of RKHS). *Set*

$$\mathcal{H}(E) \equiv \bigcap_{p \in F} \ker(\mathrm{ev}_{\varphi(p)}) \subset H_K(E).$$

Denote by $\mathcal{H}^{\perp}(E)$ the orthogonal complement of $\mathcal{H}(E)$ in $H_K(E)$ and by P the projection from $H_K(E)$ to $\mathcal{H}^{\perp}(E) \subset H_K(E)$. Then the pullback $H_{\varphi^* K}(F)$ is described as follows:

$$H_{\varphi^* K}(F) = \{f \circ \varphi \,:\, f \in H_K(E)\} \tag{2.6}$$

as a set and the inner product is given by

$$\langle f \circ \varphi, g \circ \varphi \rangle_{H_{\varphi^* K}(F)} = \langle Pf, P, g \rangle_{H_K(E)} \tag{2.7}$$

for all $f, g \in H_K(E)$.

2.3 Approximations and reproducing kernel Hilbert spaces

The next theorem is used to control the speed of convergence of the approximate solutions.

Theorem 2.9 (Highlight of several points). *Suppose that we are given a finite number of points $\Theta = \{\theta_j\}_{j=1}^N \subset E$ and a positive sequence $\{\lambda_j\}_{j=1}^N$. If we set*

$$A_\Theta \equiv \{A_{\Theta,j,j'}\}_{j,j'=1}^N \equiv \left({}^t\{\delta_{j,j'} + \lambda_j\, K(\theta_{j'}, \theta_j)\}_{j,j'=1}^N\right)^{-1}$$

$$K_\Theta(p,q) \equiv K(p,q) - \sum_{j,j'=1}^N \lambda_j\, K(p,\theta_j) A_{\Theta,j,j'} K(\theta_{j'}, q)$$

for $p, q \in E$. Then $H_{K_\Theta}(E) = H_K(E)$ as a set and the inner product of $H_{K_\Theta}(E)$ is given by

$$\langle f, g \rangle_{H_{K_\Theta}(E)} = \langle f, g \rangle_{H_K(E)} + \sum_{j=1}^N \lambda_j\, f(\theta_j)\overline{g(\theta_j)} \quad f, g \in H_K(E). \tag{2.8}$$

2.4 Dirac delta and reproducing kernels

We can approximate Hilbert spaces by using reproducing kernel Hilbert spaces as follows:

Theorem 2.10. *Suppose that we are given an decreasing sequence $\{K_t\}_{t>0}$ of positive definition quadratic form functions satisfying*

$$\langle f, g\rangle_{H_{K_{t_1}}} = \langle f, g\rangle_{H_{K_{t_2}}} \tag{2.9}$$

for all $t_2 > t_1 > 0$ and $f, g \in H_{K_{t_2}}(E)$.

1. *Let $f \in H_0$. If we define*

$$f_t^*(x) = \langle f, K_t(\cdot, x)\rangle_{H_0} \quad (x \in E),$$

   then $f_t^* \in H_{K_t}(E)$ for all $t > 0$, and as $t \downarrow 0$, $f_t^* \to f$ in the topology of H_0.

2. *The space $H_{K_t}(E)$ is a closed subspace of H_0.*

2.5 The structure of separable reproducing kernel Hilbert spaces

The following theorem is a simple consequence of the definition but this theorem is useful when we calculate reproducing kernels.

Theorem 2.11. *Let $\{v_j\}_{j=1}^{\infty}$ be a complete orthonormal basis in $H_K(E)$. Then we have*

$$K(p, q) = \sum_{j=1}^{\infty} v_j(p)\overline{v_j(q)} \quad (p, q \in E). \tag{2.10}$$

The next theorem is used to create some algorithms.

Theorem 2.12. *Let $H_K(E)$ be a separable reproducing kernel Hilbert space. Choose an orthonormal basis $\{e_j\}_{j=1}^{\infty} \subset H_K(E)$. If $\ell : H_K(E) \to \mathbb{C}$ is a bounded linear operator, then the expression*

$$\ell \bowtie K \equiv \sum_{j=1}^{\infty} \overline{\ell(e_j)} e_j \tag{2.11}$$

converges in $H_K(E)$ and it does not depend on the choice of $\{e_j\}_{j=1}^{\infty}$.

3 Nature of the book from its preface

The theory of reproducing kernels is starting from a paper in 1921 [4] and the one in 1922 [2] which dealt with typical reproducing kernels of Szegö and Bergman, and then the theory has been developed into a large and deep theory in complex analysis by many mathematicians. However, precisely, reproducing kernels were appeared previously during the first decade of 20th century by S. Zaremba [5] in his work on boundary value problems for harmonic and biharmonic functions. But he did not develop any further theory for the reproducing property. Furthermore, in fact, we knew many concrete reproducing kernels for spaces of polynomials and trigonometric functions in much older days, as we will see in this book. Meanwhile, the general theory of reproducing kernels was established in a complete form by N. Aronszajn [1] in 1950. Furthermore, L. Schwartz [3], who is Fields-Medalist and founded distribution theory, developed the general theory remarkably in 1964 with the paper of over 140 pages.

The general theory is certainly beautiful, it seems, however, that for a long time we have overlooked the importance of the general theory of reproducing kernels. We were not able to find an essential reason why the theory is important. Indeed, it was an abstract theory, and from the theory, we were not able to derive any definite results and any essential developments in mathematics. The theory by Schwartz is great, however its importance remained unnoticed for a long time: It is still ignored.

When we consider linear mappings in the framework of Hilbert spaces, we will encounter in a natural way the concept of reproducing kernels; then the general theory is not restricted to Bergman and Szegö kernels, but the general theory is as important as the concept of Hilbert spaces. It is a fundamental concept and important mathematics. The general theory of reproducing kernels is based on elementary theorems on Hilbert spaces. The theory of Hilbert spaces is the minimum core of functional analysis, however, when the general theory is combined with linear mappings on Hilbert spaces, it will have many relations in various fields, and its fruitful applications will spread over to differential equations, integral equations, generalizations of the Phytagorean theorem, inverse problems, sampling theory, nonlinear transforms in connection with linear mappings, various operators among Hilbert spaces and other many and broad fields. Furthermore, when we apply the general theory of reproducing kernels to the Tikhonov regularization, it produces approximate solutions for equations on Hilbert spaces which contain bounded linear operators. Looking from the viewpoint of computer users at numerical solutions, we will see that they are fundamental and have practical applications.

Concrete reproducing kernels like Bergman and Szegö kernels will produce many wide and broad results in complex analysis. They developed some deep theory and lead to profound results in complex analysis containing several complex variables. Meanwhile, the formal general theory by Aronszajn has also favorable connections with various fields like learning theory, support vector machines, stochastic theory and operator theory on Hilbert spaces.

In this book, we will concentrate on the general theory of reproducing kernels developed by Aronszajn while keeping in mind the theory combined with linear mappings and applications of the general theory to the Tikhonov regularization. We will present many concrete applications from the viewpoint of numerical solutions for computer use. These topics will be general and fundamental for many mathematical scientists beyond mathematicians as in calculus and linear algebra in the undergraduate course.

One of our strong motivations for writing this book is given by the historical success of numerical and real inversion formulas of the Laplace transform which is a famous ill-posed and difficult problem and, in fact, we will give their mathematical theory and formulas, as a clear evidence of definite power of the theory of reproducing kernels by combining the Tikhonov regularization. For the algorithm based on the theory, Hiroshi Fujiwara made the software and we can use it through his kind guide.

For these topics, we will need background materials like integration theory, fundamental Hilbert space theory, the Fourier transform and the Laplace transform. We describe the structure of this book.

In Chapter 1, we will give many concrete reproducing kernels first and in Chapter 2, we develop the general theory of reproducing kernels with general and broad applications by combining with linear mappings.

In Chapter 3, we will apply the general and global theory of reproducing kernels to the Tikhonov regularization in a lucid manner. We stand on the viewpoint of numerical solutions of bounded linear operator equations on Hilbert spaces for computer use in a definite and self-contained way.

Chapter 4 intends an introduction to what Hiroshi Fujiwara did. In particular, Fujiwara solved linear simultaneous equations with 6000 unknowns by means of discretization of a Fredholm integral equation of the second kind. This integral equation of the second kind was derived by the Tikhonov regularization and the reproducing kernel method in the above real inversion formula. At this moment, theoretically we will use the whole data of the output – in fact, 6000 data. Fujiwara gave solutions in **600 digits precision** with the data of **10 GB** for solutions. This fact gave a great impact to the authors. Computer power and its algorithm will be improved year by year. Meanwhile, we can practically obtain a finite number of observation data, and so we expect to obtain solutions in terms of a finite number of data for various forward and inverse problems. Thanks to the power of computers, we will be able to realize more direct and simple algorithms and so, we had included results based on *a finite number of observation data*. This method will give a new discretization principle.

Chapter 5 deals with the applications to ODEs such as fundamental equations $y'' + \alpha y' + \beta y = 0$, where α and β can be general functions. Sometimes, we consider the case when the boundary condition comes into play.

As one main substance of new results, in Chapter 6 we present many concrete results for various fundamental PDEs. Here we take up the Poisson equation, the Laplace equation, the heat equation and the wave equation.

Similarly, in Chapter 7 we deal with integral equations. We will consider typical singular integral equations, convolution equations, convolution integral equations and integral equations with the mixed Toeplitz-Hankel kernel.

In Chapter 8, we refer to specially hot topics and important materials on reproducing kernels; namely, norm-inequalities, convolution inequalities, inversion of an arbitrary matrix, representation of inverse mappings, identification of nonlinear systems, sampling theory, statistical learning theory and membership problems – this will give a new method how to catch analyticity and smoothing properties of functions by computers. Furthermore, we will see basic relationships among eigenfunctions, initial value problems for linear partial differential equations, and reproducing kernels, and we will refer to a new type general sampling theory with numerical experiments. In the last two subsections, we added new fundamental results on generalized reproducing kernels, generalized delta functions, generalized reproducing kernel Hilbert spaces and general integral transform theory. In particular, any separable Hilbert space consisting of functions may be looked as generalized reproducing kernel Hilbert spaces and the general integral transform theory may be extended to a general framework.

Chapter 9 is an appendix of this book. In Section 9.1, we introduce the theory of Akira Yamada discussing equality problems in nonlinear norm-inequalities in reproducing kernel Hilbert spaces, indeed, we may be surprised at his general theory in the general theory of reproducing kernels. In Section 9.2, we introduce Yamada's unified and generalized inequalities for Opial's inequalities. Similar, but different generalizations were independently published by Nguyen Du Vi Nhan, Dinh Thanh Duc, and Vu Kim Tuan, in the same year. In Section 9.3, we introduce concrete integral representations of implicit functions. We rely upon the implicit function theory guaranteeing the existence of implicit functions. The fundamental result was obtained as a great development of a general abstract theory of reproducing kernels.

References

[1] N. Aronszajn, *Theory of reproducing kernels,* Trans. Amer. Math. Soc., **68**(1950) , 337–404.

[2] S. Bergman, *The kernel function and conformal mapping*, Amer. Math. Soc., Providence, R. I. (1950,1970).

[3] L. Schwartz, *Sous-espaces hilbertiens d'espaces vectoriels topologiques et noyaux associés (noyaux reproduisants),* J. Analyse Math., **13** (1964), 115–256.

[4] S. Szegö, *Über orthogonale Polynome, die zu einer gegebenen Kurve der Komplexen Ebene gehören*, Math. Z., **9** (1921), 218–270.

[5] S. Zaremba, *L'equation biharminique et une class remarquable de fonctions foundamentals harmoniques,* Bulletin International de l'Academie des Sciences de Cracovie, **39** (1907), 147–196.

Yoshihiro Sawano
Tokyo Metropolitan University
1-1 Minami-Ohsawa, Hachioji, 192-0397, Tokyo.

Saburou Saitoh
Institute of Reproducing Kernels
Kawauchi-cho, 5-1648-16, Kiryu 376-0041.