S. SAITOH
KODAI MATH. J.
28 (2005), 359–367

BEST APPROXIMATION, TIKHONOV REGULARIZATION AND REPRODUCING KERNELS

Saburou Saitoh

In the memory of Professor Nobuyuki Suita

Abstract

We shall give a definite and good application of the theory of reproducing kernels to the Tikhonov regularization that is powerful in best approximation problems in numerical analysis.

1. Introduction

At first, we recall a fundamental theorem for the best approximation by the functions in a reproducing kernel Hilbert space (RKHS) based on [1, 3].

Let E be an arbitrary set, and let H_K be the RKHS admitting a reproducing kernel $K(p,q)$ on E. For any Hilbert space $\mathscr{H}$ we consider a bounded linear operator L from H_K into $\mathscr{H}$. Then, we shall consider the best approximate problem

$$\inf_{f \in H_K} \|Lf - \mathbf{d}\|_{\mathscr{H}} \tag{1.1}$$

for a member $\mathbf{d}$ of $\mathscr{H}$. Then, we have

Proposition 1.1. *For a member* $\mathbf{d}$ *of* $\mathscr{H}$, *there exists a function* $\tilde{f}$ *in* H_K *such that*

$$\inf_{f \in H_K} \|Lf - \mathbf{d}\|_{\mathscr{H}} = \|L\tilde{f} - \mathbf{d}\|_{\mathscr{H}} \tag{1.2}$$

if and only if, for the RKHS H_k *defined by*

$$k(p,q) = (L^*LK(\cdot,q), L^*LK(\cdot,p))_{H_K}, \tag{1.3}$$

2000 *Mathematics Subject Classification.* Primary 44A15; 35K05; 30C40.

Keywords. Approximation of functions, best approximation, reproducing kernel, Tikhonov regularization, Sobolev space, generalized inverse, Moore-Penrose generalized inverse, approximate inverse.

Received September 30, 2003; revised March 16, 2004.

L^* is the adjoint of L,

$$L^*\mathbf{d} \in H_k. \tag{1.4}$$

Furthermore, if the existence of the best approximation $\tilde{f}$ satisfying (1.2) *is ensured, then there exists a unique extremal function $f_{\mathbf{d}}^*$ with the minimum norm in H_K, and the function $f_{\mathbf{d}}^*$ is expressible in the form*

$$f_{\mathbf{d}}^*(p) = (L^*\mathbf{d}, L^*LK(\cdot\,,p))_{H_k} \quad \text{on } E. \tag{1.5}$$

In Proposition 1.1, note that

$$(L^*\mathbf{d})(p) = (L^*\mathbf{d}, K(\cdot\,,p))_{H_K} = (\mathbf{d}, LK(\cdot\,,p))_{\mathscr{H}}; \tag{1.6}$$

that is, $L^*\mathbf{d}$ is expressible in terms of the known $\mathbf{d}$, L, $K(p,q)$ and $\mathscr{H}$.

In Proposition 1.1, when $L^*\mathbf{d}$ does not belong to H_k, the function

$$f_{\mathbf{d}}^{**}(p) = (\mathbf{d}, LL^*LK(\cdot\,,p))_{\mathscr{H}} \tag{1.7}$$

is still well-defined and the function is the extremal function in the best approximate problem

$$\inf_{f \in H_K} \|L^*Lf - L^*\mathbf{d}\|_{H_K}, \tag{1.8}$$

as we see from Proposition 1.1, directly.

Let P be the projection map of $\mathscr{H}$ to $\overline{\mathscr{R}(L)}$ (closure). Then, there exists $\tilde{f}$ in H_K satisfying (1.2) if and only if $P\mathbf{d} \in \mathscr{R}(L)$. This condition is equivalent to

$$\mathbf{d} = P\mathbf{d} + (I - P)\mathbf{d} \in \mathscr{R}(L) + \mathscr{R}(L)^{\perp}.$$

Further, this condition is equivalent to

$$Lf - \mathbf{d} \in \mathscr{R}(L)^{\perp} = \mathscr{N}(L^*)$$

for some $f \in H_K$; that is, for some $f \in H_K$,

$$L^*Lf = L^*\mathbf{d}.$$

$f_{\mathbf{d}}^*$ in (1.5) is the Moore-Penrose generalized inverse of the equation

$$Lf = \mathbf{d}.$$

In particular, if the Moore-Penrose generalized inverse $f_{\mathbf{d}}^*$ exists, it coincides with $f_{\mathbf{d}}^{**}$ in (1.7).

Proposition 1.1 is rigid and is not practical in practical applications, because, practical data contain noises or errors and the criteria (1.4) is not suitable.

Meanwhile, the representation (1.7) is convenient in these senses. However, the function $f_{\mathbf{d}}^{**}(p)$ is, in general, not suitable for the problem (1.1). Indeed, we can see an estimate of $\|Lf_{\mathbf{d}}^{**} - \mathbf{d}\|_{\mathscr{H}}$ as follows:

For the best approximate function $f_{\mathbf{d}}^{**}(p)$, we have

$$\begin{aligned} f_{\mathbf{d}}^{**}(p) &= (L^*\mathbf{d}, L^*LK(\cdot\,,p))_{H_K} \\ &= (L^*LL^*\mathbf{d})(p) \end{aligned}$$

and so for the image of $f_{\mathbf{d}}^{**}(p)$, we thus obtain the estimate

$$\|Lf_{\mathbf{d}}^{**} - \mathbf{d}\|_{\mathscr{H}} \le \|LL^*LL^* - I\| \, \|\mathbf{d}\|_{\mathscr{H}}.$$

We shall establish good relationship between the Tikhonov regularization and the theory of reproducing kernels. For the Tikhonov regularization, see, for example, [2].

2. Tikhonov regularization

We shall introduce the Tikhonov regularization in the framework of the theory of reproducing kernels based on ([1], [3, pp. 50–53]). However, from the viewpoint of Tikhonov regularization we shall give a further result constructing the associated reproducing kernels and a new viewpoint for the previous results.

Let L be a bounded linear operator from a reproducing kernel Hilbert space H_K admitting a reproducing kernel $K(p,q)$ on a set E into a Hilbert space $\mathscr{H}$. Then, by introducing the inner product, for any fixed positive $\lambda > 0$

$$(f,g)_{H_K(L;\lambda)} = \lambda(f,g)_{H_K} + (Lf,Lg)_{\mathscr{H}}, \tag{2.9}$$

we shall construct the Hilbert space $H_K(L;\lambda)$ comprising functions of H_K. This space, of course, admits a reproducing kernel and we shall denote it by $K_L(p,q;\lambda)$. Then, we first have the elementary properties:

Lemma 2.1. *For the linear mapping from $\mathscr{H}$ into the RKHS H_K defined by*

$$f(p) = (\mathbf{d}, LK_p)_{\mathscr{H}} \quad \text{for } \mathbf{d} \in \mathscr{H}; \tag{2.10}$$

$$K_p := K(\cdot\,, p),$$

we have

$$f = L^*\mathbf{d} \tag{2.11}$$

and

$$\|f\|_{H_K}^2 = (\mathbf{d}, L(\mathbf{d}, LK_p)_{\mathscr{H}})_{\mathscr{H}}. \tag{2.12}$$

Lemma 2.2. *The following items are equivalent*:

$$\text{(i)} \quad K(p,q) \gg (LK_q, LK_p)_{\mathscr{H}};$$

that is, $K(p,q) - (LK_q, LK_p)_{\mathscr{H}}$ is a positive matrix on E,

$$\text{(ii)} \quad \|L\| \le 1,$$

and

$$\text{(iii)} \quad \|f\|_{H_K} \le \|\mathbf{d}\|_{\mathscr{H}} \quad \text{in (2.10)}.$$

Lemma 2.3. *The reproducing kernel $K_L(p,q;\lambda)$ is determined as the unique solution $\tilde{K}(p,q;\lambda)$ of the equation*:

$$\tilde{K}(p,q;\lambda) + \frac{1}{\lambda}(L\tilde{K}_q, LK_p)_{\mathscr{H}} = \frac{1}{\lambda}K(p,q) \tag{2.13}$$

with

$$\tilde{K}_q = \tilde{K}(\cdot\,, q; \lambda) \in H_K \quad \text{for } q \in E. \tag{2.14}$$

Note here, in general, that the norm of the RKHS $H_{\lambda K}$ admitting the reproducing kernel $\lambda K(p,q)$ $(\lambda > 0)$ is given by

$$\|f\|^2_{H_{\lambda K}} = \frac{1}{\lambda}\|f\|^2_{H_K} \tag{2.15}$$

and the members of functions of $H_{\lambda K}$ are the same of those of H_K.

We shall consider that the reproducing kernel $K(p,q)$ is known and we wish to construct the reproducing kernel $K_L(p,q;\lambda)$. For this construction we can obtain a very effective method by using the Neumann series. We define the bounded linear operator $\tilde{L}$ from H_K into H_K defined by

$$(\tilde{L}f)(p) = (Lf, LK_p)_{\mathscr{H}} = (L^*Lf)(p).$$

Then, from (2.13) we obtain directly

THEOREM 2.4. *If* $\|\tilde{L}\| < \lambda$, *then* $K_L(p,q;\lambda)$ *is expressible in terms of* $K(p,q)$ *by the Neumann series*:

$$K_L(p,q;\lambda) = \left(I + \frac{\tilde{L}}{\lambda}\right)^{-1}\frac{1}{\lambda}K(p,q) = \sum_{n=0}^{\infty}\left(-\frac{\tilde{L}}{\lambda}\right)^{n}\frac{1}{\lambda}K(p,q), \tag{2.16}$$

where $\left(I + \frac{\tilde{L}}{\lambda}\right)^{-1}$ *is a bounded linear operator from* H_K *into* H_K *satisfying*

$$\left\|\frac{1}{I + \frac{\tilde{L}}{\lambda}}\right\| \le \frac{1}{1 - \left\|\frac{\tilde{L}}{\lambda}\right\|}.$$

Of course, if the operator $\tilde{L}$ is compact, then we can apply the spectral theory to the equation (2.13) without the restriction $\|\tilde{L}\| < \lambda$. In particular, then $\left(I + \frac{\tilde{L}}{\lambda}\right)^{-1}$ is a bounded linear operator and

$$K_L(p,q;\lambda) = \left(I + \frac{\tilde{L}}{\lambda}\right)^{-1}\frac{1}{\lambda}K(p,q).$$

We shall consider the best approximation problem, for any given $f_0 \in H_K$ and $\mathbf{d} \in \mathscr{H}$:

$$\inf_{f \in H_K}\{\lambda\|f_0 - f\|^2_{H_K} + \|\mathbf{d} - Lf\|^2_{\mathscr{H}}\}, \tag{2.17}$$

in connection with the Tikhonov regularization for the equation $Lf = \mathbf{d}$.

We shall introduce the direct sum space

$$\hat{\mathscr{H}}_\lambda = \sqrt{\lambda} H_K \oplus \mathscr{H} \tag{2.18}$$

and the bounded linear operator $\hat{L}_\lambda$ from $H_K(L;\lambda)$ into $\hat{\mathscr{H}}_\lambda$ defined by

$$\hat{L}_\lambda f = \sqrt{\lambda} f \oplus Lf \in \hat{\mathscr{H}}_\lambda \tag{2.19}$$

and the adjoint operator $\hat{L}_\lambda^*$ of $\hat{L}_\lambda$ defined by, for $f \in H_K$ and for $\mathbf{d} \in \mathscr{H}$

$$\hat{L}_\lambda^*\{f, \mathbf{d}\} = \sqrt{\lambda} f + L^* \mathbf{d} \in H_K(L;\lambda). \tag{2.20}$$

Then, by Proposition 1.1 we see that in (2.17) the best approximation $f^* \in H_K(L;\lambda) = H_K$ (as sets of functions) in the sense

$$\inf_{f \in H_K} \{\lambda \|f_0 - f\|_{H_K}^2 + \|\mathbf{d} - Lf\|_{\mathscr{H}}^2\} = \lambda \|f_0 - f^*\|_{H_K}^2 + \|\mathbf{d} - Lf^*\|_{\mathscr{H}}^2 \tag{2.21}$$

exists if and only if, for $f_0 \in H_K$ and $\mathbf{d} \in \mathscr{H}$

$$\hat{L}_\lambda^*\{f_0, \mathbf{d}\} \in H_{k_\lambda}, \tag{2.22}$$

where H_{k_λ} is the RKHS admitting the reproducing kernel

$$k_\lambda(p, q) = (\hat{L}_\lambda^* \hat{L}_\lambda K_L(\cdot, q; \lambda), \hat{L}_\lambda^* \hat{L}_\lambda K_L(\cdot, p; \lambda))_{H_K(L;\lambda)}.$$

In the present case, $\hat{L}_\lambda$ is an isometric operator from $H_K(L;\lambda)$ into $\hat{\mathscr{H}}_\lambda$ and so, $\hat{L}_\lambda^* \hat{L}_\lambda$ is the identity on $H_K(L;\lambda)$. Therefore, we have the simple identity

$$k_\lambda(p, q) = K_L(p, q; \lambda). \tag{2.23}$$

Therefore, the best approximation f^* exists always in (2.21). Furthermore, since $\mathscr{N}(\hat{L}_\lambda^* \hat{L}_\lambda) = \{0\}$, it is uniquely determined. Again, note that as members of functions

$$H_K(L;\lambda) = H_K.$$

We thus obtain, from Proposition 1.1:

THEOREM 2.5. *In our situation, the generalized solution f^* of the equations*

$$f_0 = f \quad \text{in } H_K$$

and

$$\mathbf{d} = Lf \quad \text{in } \mathscr{H}$$

in the sense (2.21) *exists uniquely and it is represented by*

$$f^*(p) = \lambda(f_0(\cdot), K_L(\cdot, p; \lambda))_{H_K} + (\mathbf{d}, LK_L(\cdot, p; \lambda))_{\mathscr{H}}. \tag{2.24}$$

In Theorem 2.5, in particular, we shall consider the best approximate function, for $f_0 = 0$

$$f_{\lambda,\mathbf{d}}^*(p) = (\mathbf{d}, LK_L(\cdot, p; \lambda))_{\mathscr{H}}, \tag{2.25}$$

which is the extremal function in the Tikhonov regularization (2.21) for $f_0 = 0$.

In general, in the Tikhonov regularization, the operator L is compact and the extremal functions are represented by using the singular values and singular functions of the selfadjoint operator L^*L. So, the representations are, in a sense, abstract. And the behaviour of the extremal functions as λ tends to zero is an important problem, because the limit function may be expected as a solution of the equation $Lf = \mathbf{d}$ as in the Moore-Penrose generalized inverse.

From many examples in our situation ([4, 5, 6]), however we see that

$$\lim_{\lambda\to 0} K_L(p,q;\lambda) \tag{2.26}$$

and

$$\lim_{\lambda\to 0}(\mathbf{d}, LK_L(p,q;\lambda))_{\mathscr{H}} \tag{2.27}$$

do, in general, not exist.

3. Main results

We now give our main results in this paper:

THEOREM 3.1. *For any* $\mathbf{d}\in\mathscr{H}$ *and for the two best approximate functions* $f^*_{\lambda,\mathbf{d}}(p)$ *in* (2.25) *satisfying*

$$\inf_{f\in H_K}\{\lambda\|f\|^2_{H_K} + \|\mathbf{d} - Lf\|^2_{\mathscr{H}}\} = \lambda\|f^*_{\lambda,\mathbf{d}}\|^2_{H_K} + \|\mathbf{d} - Lf^*_{\lambda,\mathbf{d}}\|^2_{\mathscr{H}}$$

and $f^{**}_{\mathbf{d}}(p)$ *in* (1.7) *satisfying*

$$\inf_{f\in H_K}\|L^*Lf - L^*\mathbf{d}\|_{H_K} = \|L^*Lf^{**}_{\mathbf{d}} - L^*\mathbf{d}\|_{H_K},$$

we have the estimate

$$|f^*_{\lambda,\mathbf{d}}(p) - f^{**}_{\mathbf{d}}(p)| \le \left(\lambda\|L\| + \|LL^*LL^* - I\|\frac{1}{\sqrt{2\lambda}}\right)\sqrt{K(p,p)}\|\mathbf{d}\|_{\mathscr{H}}. \tag{3.28}$$

Proof. From (2.13), we have

$$\lambda K_L(p,q;\lambda) + L^*LK_L(p,q;\lambda) = K(p,q). \tag{3.29}$$

Hence,

$$\|\lambda K_L(p,q;\lambda)\|^2_{H_K} + \|L^*LK_L(p,q;\lambda)\|^2_{H_K} + 2\lambda\|LK_L(p,q;\lambda)\|^2_{\mathscr{H}} = K(q,q)$$

and so, in particular,

$$\|L^*LK_L(p,q;\lambda)\|^2_{H_K} \le K(q,q) \tag{3.30}$$

and

$$2\lambda\|LK_L(p,q;\lambda)\|^2_{\mathscr{H}} \le K(q,q). \tag{3.31}$$

Hence, from (3.29)

$$\begin{aligned}&LL^*LK(\cdot\,,p) - LK_L(\cdot\,,p;\lambda)\\&\qquad = \lambda LL^*LK_L(\cdot\,,p;\lambda) + LL^*LL^*LK_L(\cdot\,,p;\lambda) - LK_L(\cdot\,,p;\lambda).\end{aligned}$$

From (2.25) and (1.7), we have the desired result

$$\begin{aligned}(3.32)\quad &|f^*_{\lambda,\mathbf{d}}(p) - f^{**}_{\mathbf{d}}(p)|\\&\le \|\mathbf{d}\|_{\mathscr{H}}\|LL^*LK(\cdot\,,p) - LK_L(\cdot\,,p;\lambda)\|_{\mathscr{H}}\\&\le \lambda\|\mathbf{d}\|_{\mathscr{H}}\|LL^*LK_L(\cdot\,,p;\lambda)\|_{\mathscr{H}} + \|\mathbf{d}\|_{\mathscr{H}}\|LL^*LL^* - I\|\,\|LK_L(\cdot\,,p;\lambda)\|_{\mathscr{H}}\\&\le \lambda\|\mathbf{d}\|_{\mathscr{H}}\|L\|\,\|L^*LK_L(\cdot\,,p;\lambda)\|_{H_K} + \|\mathbf{d}\|_{\mathscr{H}}\|LL^*LL^* - I\|\frac{1}{\sqrt{2\lambda}}\sqrt{K(p,p)}\\&\le \lambda\|\mathbf{d}\|_{\mathscr{H}}\|L\|\sqrt{K(p,p)} + \|\mathbf{d}\|_{\mathscr{H}}\|LL^*LL^* - I\|\frac{1}{\sqrt{2\lambda}}\sqrt{K(p,p)}.\end{aligned}$$

COROLLARY 3.2. *If LL^* is unitary, then we have for the two best approximate functions $f^*_{\lambda,\mathbf{d}}(p)$ in* (2.25) *and $f^{**}_{\mathbf{d}}(p)$ in* (1.7) *we have the estimate*

$$(3.33)\qquad |f^*_{\lambda,\mathbf{d}}(p) - f^{**}_{\mathbf{d}}(p)| \le \lambda\|L\|\sqrt{K(p,p)}\|\mathbf{d}\|_{\mathscr{H}}$$

which shows that as λ tends to zero, $f^*_{\lambda,\mathbf{d}}(p)$ tends to $f^{**}_{\mathbf{d}}(p)$ with the order λ and the convergence is uniform on any subset E_0 of E satisfying $\sup_{p\in E_0} K(p,p) < \infty$.

THEOREM 3.3. *If L is a compact operator, then for the Moore-Penrose generalized inverse $f^*_{\mathbf{d}}$,*

$$(3.34)\qquad \lim_{\lambda\to 0} f^*_{\lambda,\mathbf{d}}(p) = f^*_{\mathbf{d}}(p),$$

uniformly on any subset E_0 of E satisfying $\sup_{p\in E_0} K(p,p) < \infty$.

Proof. Since L is compact, we have, from (2.13)

$$K_L(p,q;\lambda) = \frac{1}{\lambda I + L^*L}K(p,q).$$

Then,

$$\begin{aligned}f^*_{\lambda,\mathbf{d}}(p) &= (\mathbf{d}, LK_L(\cdot\,,p;\lambda))_{\mathscr{H}}\\&= (L^*\mathbf{d}, K_L(\cdot\,,p;\lambda))_{\mathscr{H}}\\&= \left(\frac{1}{\lambda I + L^*L}L^*\mathbf{d}, K(\cdot\,,p)\right)_{H_K}.\end{aligned}$$

As we see by using the singular value decomposition of L, for the Moore-Penrose generalized inverse $f^*_{\mathbf{d}}$, as $\lambda \to 0$,

$$\frac{1}{\lambda I + L^*L} L^* \mathbf{d} \to f_{\mathbf{d}}^*, \quad \text{in } H_K$$

(see Section 5.1 in [2]). Hence, from the identity

$$f_{\lambda,\mathbf{d}}^*(p) - f_{\mathbf{d}}^*(p) = \left(\frac{1}{\lambda I + L^*L} L^* \mathbf{d} - f_{\mathbf{d}}^*, K(\cdot\,, p) \right)_{H_K},$$

we have the desired result.

The result in Theorem 3.3 is very reasonable and good, because in the Tikhonov regularization, if the Moore-Penrose generalized inverse exits and if λ tends to zero, then (3.34) is valid. Of course, here we need to consider a reproducing kernel Hilbert space as the function space. However, the convergence rate in (3.34) or the error estimate in the form

$$|f_{\lambda,\mathbf{d}}^*(p) - f_{\mathbf{d}}^*(p)|$$

will be derived by case by case arguments depending on concrete problems. See, for example, [7].

In Corollary 3.2, the result (3.33) will be fresh, because $f_{\lambda,\mathbf{d}}^*$ and $f_{\mathbf{d}}^{**}$ are, in general, different, however, if λ tends to zero, then $f_{\lambda,\mathbf{d}}^*(p)$ tends to $f_{\mathbf{d}}^{**}(p)$ with the λ order, under the unitarieness of LL^*. However, at this moment, the author does not see the precise meanings of this result.

Theorem 3.1 is a much more delicate result in the general situation that the extremal functions $f_{\lambda,\mathbf{d}}^*$ and $f_{\mathbf{d}}^{**}$ exist, when we consider their difference in connection with the parameter λ, involving the unitarieness of LL^*. Indeed, the quantity $\|LL^*LL^* - I\|$ may be understood as a distance of the operator LL^* from being unitary.

4. Acknowledgements

The author wishes to express his deep thanks the referee for his/her careful readings the paper and valuable suggestions for the paper.

References

[1] D.-W. Byun and S. Saitoh, Best approximation in reproducing kernel Hilbert spaces, Proc. of the 2nd International Colloquium on Numerical Analysis, VSP-Holland, 1994, 55–61.

[2] C. W. Groetsch, Inverse Problems in the Mathematical Sciences, Vieweg & Sohn Verlagsgesellschaft mbH, Braunschweig/Wiesbaden, 1993.

[3] S. Saitoh, Integral Transforms, Reproducing Kernels and their Applications, Pitman Res. Notes in Math. Series **369**, Addison Wesley Longman Ltd., UK, 1997.

[4] S. Saitoh, Approximate Real Inversion Formulas of the Gaussian Convolution, Applicable Analysis **83** (2004), 727–733.

[5] S. Saitoh, T. Matsuura and M. Asaduzzaman, Operator Equations and Best Approximation Problems in Reproducing Kernel Hilbert Spaces, J. of Analysis and Applications **1** (2003), 131–142.

[6] S. SAITOH, Constructions by Reproducing Kernels of Approximate Solutions for Linear Differential Equations with L_2 Integrable Coefficients, International J. of Math. Sci. **2** (2003), 261–273.

[7] T. MATSUURA, S. SAITOH AND D. D. TRONG, Approximate and analytical inversion formulas in heat conduction on mutidimensional spaces, J. of Inverse and Ill-posed Problems (to appear).

DEPARTMENT OF MATHEMATICS
FACULTY OF ENGINEERING
GUNMA UNIVERSITY
KIRYU 376-8515
JAPAN
e-mail address: ssaitoh@math.sci.gunma-u.ac.jp