CUBO *A Mathematical Journal*
Vol. 9, Nº 3, (65 - 74). December 2007.

Tikhonov regularization and the theory of reproducing kernels

S. Saitoh [1]
Department of Mathematics, Faculty of Enigeering, Gunma University
Kiryu 376-8515 Japan
ssaitoh@math.sci.gunma-u.ac.jp

ABSTRACT

In this paper, some definite applications of the theory of reproducing kernels to the Tikhonov regularization representing the extremal functions in the regularization are established.

RESUMEN

En este articulo se establecen algunas aplicaciones definidas de la teoria de núcleos reproductores a la regularización de Tikhonov que representan las funciones extremales en la regularización.

Key words and phrases: *Tikhonov Regularization, Reproducing Kernels, Spectral Theory*
Math. Subj. Class.: *41A50, 46B22*

[1] Supported by the Grant-in-Aid for the Scientific Research (C)(2) (No. 16540137) from the Japan Society for the Promotion Science and by the Mitsubishi Foundation, the 36th, Natural Sciences, No. 20 (2005-2006).
The author thanks the referee for his/her carefully reading the manuscript and valuable comments on the paper.

1 Introduction

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

$$\inf_{f\in H_K} \|Lf - \mathbf{d}\|_{\mathcal{H}} \tag{1}$$

for a vector $\mathbf{d}$ in $\mathcal{H}$. Then, we have

Proposition 1.1 *([1,9]) For a vector* $\mathbf{d}$ *in* $\mathcal{H}$*, there exists a function* $\tilde{f}$ *in* H_K *such that*

$$\inf_{f\in H_K} \|Lf - \mathbf{d}\|_{\mathcal{H}} = \|\mathbf{L}\tilde{\mathbf{f}} - \mathbf{d}\|_{\mathcal{H}} \tag{2}$$

if and only if, for the RKHS H_k *admitting the reproducing kernel defined by*

$$k(p,q) = (L^*LK(\cdot,q), L^*LK(\cdot,p))_{H_K}, \tag{3}$$

$$L^*\mathbf{d} \in \mathbf{H_k}. \tag{4}$$

Furthermore, if the best approximation $\tilde{f}$ *satisfying (2) exists, 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*

$$f_{\mathbf{d}}(p) = (L^*\mathbf{d}, \mathbf{L^*LK}(\cdot,\mathbf{p}))_{\mathbf{H_k}} \quad on \quad \mathbf{E}. \tag{5}$$

In Proposition 1.1, note that

$$(L^*\mathbf{d})(\mathbf{p}) = (\mathbf{L^*d}, \mathbf{K}(\cdot,\mathbf{p}))_{\mathbf{H_K}} = (\mathbf{d}, \mathbf{LK}(\cdot,\mathbf{p}))_{\mathcal{H}}; \tag{6}$$

that is, $L^*\mathbf{d}$ is expressible in terms of the known $\mathbf{d}, \mathbf{L}, \mathbf{K}(\mathbf{p},\mathbf{q})$ and $\mathcal{H}$. $f_{\mathbf{d}}$ in (5) is the Moore-Penrose generalized inverse solution $L^{\dagger}\mathbf{d}$ of the equation $Lf = \mathbf{d}$. Therefore, Proposition 1.1 gives a necessary and sufficient condition for the existence of the Moore-Penrose generalized inverse. See [3,13] for the details. Proposition 1.1 is rigid and is not practical in practical applications, because, practical data contain noises or errors and the criteria (4) is not suitable. So, we shall consider the Tikhonov regularization and we shall establish a good relation between the Tikhonov regularization and the theory of reproducing kernels. For the Tikhonov regularization, see, for example, [2,3].

In this paper, we, in particular, establish the important error estimates Theorem 3.1 and Theorem 5.1 and an important general discretization Theorem 6.1 with the related error estimate. The author now thinks that the application of the theory of reproducing kernels to the Tikhonov regularization is completed, in a sense, in a general theory.

2 Spectral theory

In order to discuss operator equations for general bounded linear operators L, following [2] we shall fix the well-established theory among spectral theory, the Moore-Penrose generalized inverse and the Tikhonov regularization. See [3] for the corresponding results for compact operators L.

Let $\{E_\lambda\}$ be a spectral family for the self-adjoint operator L^*L. If L^*L is continuously invertible, then

$$(L^*L)^{-1} = \int \frac{1}{\lambda} dE_\lambda.$$

In this case, the Moore-Penrose generalized inverse (5) can be represented by the Gaussian normal equation

$$f_{\mathrm{d}}(p) = \int \frac{1}{\lambda} dE_\lambda L^* \mathrm{d}. \tag{7}$$

If $\mathcal{R}(L)$ is non-closed and $\mathrm{d} \notin \mathcal{D}(L^\dagger)$, i.e. if the equation $Lf = \mathrm{d}$ is ill-posed, then the integral in (7) does not exist. Then, we shall define

$$f_{\mathrm{d},\alpha}(p) = \int \frac{1}{\lambda + \alpha} dE_\lambda L^* \mathrm{d}. \tag{8}$$

By construction, the operator on the right-hand side of (8) acting on d is continuous, so that, for noisy data d^δ with $\|\mathrm{d} - \mathrm{d}^\delta\|_{\mathcal{H}} \leq \delta$, we can bound the error between $f_{\mathrm{d},\alpha}$ and

$$f^\delta_{\mathrm{d},\alpha}(p) = \int \frac{1}{\lambda + \alpha} dE_\lambda L^* \mathrm{d}^\delta \tag{9}$$

as follows:

Proposition 2.1 *([2], pages 71-73) For any* $\mathrm{d} \in \mathcal{D}(L^\dagger)$,

$$\lim_{\alpha \to 0} \frac{1}{L^*L + \alpha I} L^* \mathrm{d} = \lim_{\alpha \to 0} f_{\mathrm{d},\alpha} = f_{\mathrm{d}}. \tag{10}$$

Furthermore,

$$\|Lf_{\mathrm{d},\alpha} - Lf^\delta_{\mathrm{d},\alpha}\|_{\mathcal{H}} \leq \delta \tag{11}$$

and

$$\|f_{\mathrm{d},\alpha} - f^\delta_{\mathrm{d},\alpha}\|_{H_K} \leq \frac{\delta}{\sqrt{\alpha}}. \tag{12}$$

Proposition 2.2 *([2], pages 117-118) For any* $\mathrm{d} \in \mathcal{D}(L^\dagger)$ *with* $\|\mathrm{d} - \mathrm{d}^\delta\|_{\mathcal{H}} \leq \delta$, *the function* $f^\delta_{\mathrm{d},\alpha}$ *defined by (9) is the unique minimizer of the Tikhonov functional*

$$\inf_{f \in H_K} \{\alpha \|f\|^2_{H_K} + \|\mathrm{d}^\delta - Lf\|^2_{\mathcal{H}}\}. \tag{13}$$

If $\alpha = \alpha(\delta)$ *is such that*

$$\lim_{\delta \to 0} \alpha(\delta) = 0$$

and

$$\lim_{\delta \to 0} \frac{\delta^2}{\alpha(\delta)} = 0,$$

then

$$\lim_{\delta \to 0} f_{\mathbf{d},\alpha}^{\delta} = f_{\mathbf{d}} = L^{\dagger}(\mathbf{d}). \tag{14}$$

3 Representation of the extremal functions in Tikhonov regularization

Our main purpose here is to give an effective representation of the extremal functions $f_{\mathbf{d},\alpha}$ or $f_{\mathbf{d},\alpha}^{\delta}$ in the Tikhonov regularization, since the representation by spectral theory is abstract, in many practical problems.

We set

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

Then, by introducing the inner product, for any fixed positive $\alpha > 0$

$$(f, g)_{H_K(L;\alpha)} = \alpha(f, g)_{H_K} + (Lf, Lg)_{\mathcal{H}}, \tag{15}$$

we shall construct the Hilbert space $H_K(L; \alpha)$ comprising functions of H_K. This space, of course, admits a reproducing kernel. Furthermore, we obtain, directly

Proposition 3.1 *([13,14]) The extremal function* $f_{\mathbf{d},\alpha}(p)$ *in the Tikhonov regularization*

$$\inf_{f \in H_K} \{\alpha \|f\|_{H_K}^2 + \|\mathbf{d} - Lf\|_{\mathcal{H}}^2\} \tag{16}$$

is represented in terms of the kernel $K_L(p, q; \alpha)$ *as follows:*

$$f_{\mathbf{d},\alpha}(p) = (\mathbf{d}, LK_L(\cdot, p; \alpha))_{\mathcal{H}} \tag{17}$$

where the kernel $K_L(p, q; \alpha)$ *is the reproducing kernel for the Hilbert space* $H_K(L; \alpha)$ *and it is determined as the unique solution* $\tilde{K}(p, q; \alpha)$ *of the equation:*

$$\tilde{K}(p, q; \alpha) + \frac{1}{\alpha}(L\tilde{K}_q, LK_p)_{\mathcal{H}} = \frac{1}{\alpha} K(p, q) \tag{18}$$

with

$$\tilde{K}_q = \tilde{K}(\cdot, q; \alpha) \in H_K \quad \text{for} \quad q \in E, \tag{19}$$

and

$$K_p = K(\cdot, p) \in H_K \quad \text{for} \quad p \in E.$$

In (17), when d contains errors or noises, we need its error estimate. For this, we can obtain the general result:

Theorem 3.1 *In* (17), *we obtain the estimate*

$$|f_{\mathrm{d},\alpha}(p)| \leq \frac{1}{\sqrt{\alpha}}\sqrt{K(p,p)}\|\mathrm{d}\|_{\mathcal{H}}.$$

Proof.
From (15), we have

$$\begin{aligned} |f_{\mathrm{d},\alpha}(p)| &= |(\frac{1}{\alpha I + L^*L}L^*\mathrm{d}, K(.,p))_{H_K}| \\ &\leq \|\frac{1}{\alpha I + L^*L}L^*\mathrm{d}\|_{H_K}\sqrt{K(p,p)}. \end{aligned}$$

As we see from the spectoral theory, since

$$\frac{1}{\alpha I + L^*L}L^* = L^*\frac{1}{\alpha I + LL^*},$$

we have the inequality

$$\begin{aligned} \|\frac{1}{\alpha I + L^*L}L^*\mathrm{d}\|^2_{H_K} &= \left(L^*\frac{1}{\alpha I + LL^*}\mathrm{d}, L^*\frac{1}{\alpha I + LL^*}\mathrm{d}\right)_{H_K} \\ &= \left(LL^*\frac{1}{\alpha I + LL^*}\mathrm{d}, \frac{1}{\alpha I + LL^*}\mathrm{d}\right)_{\mathcal{H}} \\ &\leq \left\|LL^*\frac{1}{\alpha I + LL^*}\mathrm{d}\right\|_{\mathcal{H}}\left\|\frac{1}{\alpha I + LL^*}\mathrm{d}\right\|_{\mathcal{H}} \\ &\leq \|\mathrm{d}\|_{\mathcal{H}}\frac{1}{\alpha}\|\mathrm{d}\|_{\mathcal{H}}, \end{aligned}$$

([3], pp. 71-73), and so, we obtain the desired result. ∎

For many concrete applications of these general theorems, see, for example, [4-8,10-12].

4 New algorithm

In several concrete examples, we consider as the reproducing kernel Hilbert space H_K the Sobolev Hilbert spaces on the whole spaces which admit concrete reproducing kernels and as the Hilbert space $\mathcal{H}$ the Hilbert spaces L_2 on the whole spaces. Then the related reproducing kernels $K_L(p,q;\alpha)$ and the extremal functions $f_{\mathrm{d},\alpha}$ can be determined concretely in terms of the Fourier integrals from the general equation (18). See, [4-8,10-12]. Here, we shall propose a new algorithm to solve numerically the equation (18) which is, in general, an integral equation of Fredholm of the second kind. Our algorithm will give a new type discretization whose effectivity was proved by examples ([8]), since to solve the equation (18) is decisively important to obtain the concrete representation (17).

We take a complete orthonormal system $\{e_j\}_{j=1}^{\infty}$ of the Hilbert space $\mathcal{H}$.

For fixed $\{\lambda_j\}_{j=1}^{\infty}(\lambda_j > 0)$, we consider the general extremal problem for (16)

$$\inf_{f\in H_K} \left\{ \alpha\|f\|_{H_K}^2 + \sum_{j=1}^{\infty} \lambda_j |(\mathbf{d} - Lf, e_j)_{\mathcal{H}}|^2 \right\}. \tag{20}$$

That is,

$$\|\mathbf{d} - Lf\|_{\mathcal{H}}^2$$

is replaced by

$$\sum_{j=1}^{\infty} \lambda_j |(\mathbf{d}, e_j)_{\mathcal{H}} - (Lf, e_j)_{\mathcal{H}}|^2.$$

Then, we shall give an algorithm constructing the reproducing kernel $K_{\alpha,\lambda_j}(p,q)$ of the Hilbert space $H_{K_{\alpha,\lambda_j}}$ with the norm square

$$\alpha\|f\|_{H_K}^2 + \sum_{j=1}^{\infty} \lambda_j |(Lf, e_j)_{\mathcal{H}}|^2. \tag{21}$$

Here, of course, we assume that (21) converges for $\{\lambda_j\}_{j=1}^{\infty}(\lambda_j > 0)$. However, in a practical application, of course, we consider only finite terms in (21) and by finite terms we can give a good approximation of (21).

We shall start with the first step. The reproducing kernel $K^{(1)}(p,q)$ of the Hilbert space with the norm square

$$\alpha\|f\|_{H_K}^2 + \sum_{j=1}^{1} \lambda_j |(Lf, e_j)_{\mathcal{H}}|^2 \tag{22}$$

is given by

$$K^{(1)}(p,q) = K^{(0)}(p,q) - \frac{\lambda_1 (e_1, LK_p^{(0)})_{\mathcal{H}} (LK_q^{(0)}, e_1)_{\mathcal{H}}}{1 + \lambda_1 (L(e_1, LK_q^{(0)})_{\mathcal{H}}, e_1)_{\mathcal{H}}}, \tag{23}$$

for

$$K^{(0)}(p,q) = \frac{1}{\alpha} K(p,q).$$

For the second step, the reproducing kernel $K^{(2)}(p,q)$ of the Hilbert space with the norm square

$$\alpha\|f\|^2_{H_K} + \sum_{j=1}^{2} \lambda_j |(Lf, e_j)_{\mathcal{H}}|^2 \tag{24}$$

is given by

$$K^{(2)}(p,q) = K^{(1)}(p,q) - \frac{\lambda_2 (e_2, LK_p^{(1)})_{\mathcal{H}} (LK_q^{(1)}, e_2)_{\mathcal{H}}}{1 + \lambda_2 (L(e_2, LK_q^{(1)})_{\mathcal{H}}, e_2)_{\mathcal{H}}}, \tag{25}$$

by using the reproducing kernel $K^{(1)}(p,q)$. In this way, we can obtain the desired representation of $K_{\alpha,\lambda_j}(p,q) = K^{(\infty)}(p,q)$. Then, we obtain

Proposition 4.1 *For any* $\mathrm{d} \in \mathcal{H}$, *the extremal function* $f_{\alpha,\lambda,\mathrm{d}}$ *in the extremal problem (20) is given by*

$$f_{\alpha,\lambda,\mathrm{d}}(p) = \sum_{j=1}^{\infty} \lambda_j (\mathrm{d}, e_j)_{\mathcal{H}} (e_j, LK_{\alpha,\lambda_j}(\cdot,p))_{\mathcal{H}}, \tag{26}$$

where we assume that (21) converges on E.

We consider a general extremal problem in (20) by considering a general weight $\{\lambda_j\}$. This means that for a larger λ_{j_0}, the speed of the convergence

$$(Lf, e_{j_0})_{\mathcal{H}} \to (\mathrm{d}, e_{j_0})_{\mathcal{H}}$$

is higher. This technique is a very important for practical applications. For examples, see [6,8].

5 Error estimate

In the representation of (26), when the data $(\mathrm{d}, e_j)_{\mathcal{H}}$ contain errors or noises, we need its error estimate. For this we obtain the good result, which is corresponding to Theorem 3.1:

Theorem 5.1 *In (26), we obtain the estimate*

$$|f_{\alpha,\lambda,\mathrm{d}}(p)|$$

$$\leq \frac{1}{\sqrt{\alpha}} \left(\sum_{j=1}^{\infty} (\lambda_j |(\mathrm{d}, e_j)_{\mathcal{H}}|^2) \right)^{1/2} \sqrt{K(p,p)}. \tag{27}$$

6 Discrete point data case

As a very general algorithm, we shall consider the discrete point data case such that: In (20), we shall consider the corresponding problem:

$$\inf_{f \in H_K} \left\{ \alpha \|f\|_{H_K}^2 + \sum_{j=1}^{\infty} \lambda_j |f(p_j) - b_j|^2 \right\}, \tag{28}$$

for fixed discrete points $\{p_j\}_j$ of the set E and for given values $\{b_j\}_j$. Then, the corresponding kernels for (23) and (25) are given similarly

$$K^{(1)}(p, q; \{p_1\}) = K^{(0)}(p, q) - \frac{\lambda_1 K^{(0)}(p, p_1) K^{(0)}(p_1, q)}{1 + \lambda_1 K^{(0)}(p_1, p_1)}, \tag{29}$$

and

$$K^{(2)}(p, q; \{p_1, p_2\}) = K^{(1)}(p, q; \{p_1\}) - \frac{\lambda_2 K^{(1)}(p, p_2; \{p_1\}) K^{(1)}(q, p_2; \{p_1\})}{1 + \lambda_2 K^{(1)}(p_2, p_2; \{p_1\})}. \tag{30}$$

In this way, we obtain the reproducing kernel $K_{\alpha,\lambda_j}(p, q; \{p_j\})$ and the corresponding results:

Theorem 6.1 *For any $\{b_j\}$, the extremal function $f_{\alpha,\lambda,\{b_j\}}$ in the extremal problem (28) is given by*

$$f_{\alpha,\lambda,\{b_j\}}(p) = \sum_{j=1}^{\infty} \lambda_j b_j K_{\alpha,\lambda_j}(\cdot, p; \{p_j\}), \tag{31}$$

where we assume that (31) converges on E. Furthermore, we obtain the estimate

$$|f_{\alpha,\lambda,\{b_j\}}(p)|$$

$$\leq \frac{1}{\sqrt{\alpha}} \left(\sum_{j=1}^{\infty} (\lambda_j |b_j|^2) \right)^{1/2} \sqrt{K(p,p)}. \tag{32}$$

The most prototype application of the general theory of this paper is a simple construction of the Moore-Penrose generalized inverse for any matrix:

A Construction of a Natural Inverse of Any Matrix by Using the Theory of Reproducing Kernels by K. Iwamura, T. Matsuura and S. Saitoh (PAJMS Vol. 1 no: 2 (December 2005)).

Received: Jan 2006 . Revised: May 2006.

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] H. W. Engl, M. Hanke and A. Neubauer, *Regularization of Inverse Problems*, Mathematics and Its Applications, **376**(2000), Kluwer Academic Publishers.

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

[4] T. Matsuura and S. Saitoh, *Analytical and Numerical Solutions of the Inhomogeneous Wave Equation*, Australian J. of Math. Anal. and Appl., **1**(2004), Volume 1, Issure 1, Article 7.

[5] T. Matsuura and S. Saitoh, *Numerical Inversion Formulas in the Wave Equation*, J. of Computational Mathematics and Optimization, **1**(2005), 1-19.

[6] T. Matsuura and S. Saitoh, *Analytical and Numerical Solutions of Linear Ordinary Differential Equations with Constant Coefficients*, J. of Analysis and Applications. **3**(2005), 1-17.

[7] T. Matsuura, S. Saitoh and D. D. Trong, *Numerical Solutions of the Poisson Equation*, Applicable Analysis, **83**(2004), 1037-1051.

[8] T. Matsuura and S. Saitoh, *Dirichlet's Principle Using Computers*, Applicable Analysis, **84**(2005), 989-1003.

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

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

[11] 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.

[12] 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.

[13] S. Saitoh, *Best approximation, Tikhonov regularization and reproducing kernels*, Kodai. Math. J., **28**(2005), 359-367.

[14] S. Saitoh, *Tikhonov regularization and the theory of reproducing kernels*, Finite or Infinite Dimensional Complex Analysis and Applications (Proceedings of the 12th ICFIDCAA), Kyushu University Press (2005), 291-298.