Far East Journal of Mathematical Education
Volume 4, Number 2, 2010, Pages 141-149
Published Online: June 17, 2010
This paper is available online at http://pphmj.com/journals/fjme.htm

A NUMERICAL CONSTRUCTION OF A NATURAL INVERSE OF ANY MATRIX BY USING THE THEORY OF REPRODUCING KERNELS WITH THE TIKHONOV REGULARIZATION

K. IWAMURA, T. MATSUURA and S. SAITOH

Department of Mathematics
Josai University
Sakado-shi 350-0295, Keyakidai, Japan
e-mail: kiwamura@math.josai.ac.jp

Department of Mechanical System Engineering
Graduate School of Engineering
Gunma University
Kiryu 376-8515, Japan
e-mail: matsuura@gunma-u.ac.jp

Department of Mathematics
University of Aveiro
3810-193 Aveiro, Portugal
e-mail: saburou.saitoh@gmail.com

Abstract

In this paper, by a new concept and method, combining the two theories of the Tikhonov regularization and reproducing kernels, we give a practical and numerical "inverse" for any matrix and show its numerical experiments by using computers.

2010 Mathematics Subject Classification: Primary 15A09, 15A06, 30C40.

Keywords and phrases: linear simultaneous equation, matrix, Moore-Penrose generalized inverse, reproducing kernel, Tikhonov regularization, generalized inverse, approximate inverse, noise, error.

Received March 18, 2010

1. Introduction

We have, of course, many methods and applications for linear simultaneous equations. We shall introduce a new method which gives simple and natural approximate solutions for the linear simultaneous equations: for any matrix $A_{m\times n}$ of type (m, n):

$$A\mathbf{x} = \mathbf{y}. \tag{1.1}$$

For simplicity, we shall consider all on the real number field. In order to solve the equation (1.1), following the idea of the Tikhonov regularization, we shall consider the extremal problem: for any fixed $\lambda > 0$ and for any $\mathbf{y} \in \mathbf{R}^m$,

$$\inf_{\mathbf{x}\in\mathbf{R}^n} (\lambda \| \mathbf{x} \|^2_{\mathbf{R}^n} + \| A\mathbf{x} - \mathbf{y} \|^2_{\mathbf{R}^m}) \tag{1.2}$$

and we shall represent the extremal vector $\mathbf{x}^*_\lambda(\mathbf{y})$ in the form

$$\mathbf{x}^*_\lambda(\mathbf{y}) = B_\lambda \mathbf{y}. \tag{1.3}$$

Then, by letting $\lambda \to 0$, we will be able to obtain an inverse of A and the natural solution of (1.1). Of course, as well-known, B_λ is given by

$$B_\lambda = ({}^t AA + \lambda I)^{-1}{}^t A.$$

We are interested in some practical and effective construction of B_λ. On the construction of $\mathbf{x}^*_\lambda$ in the representation (1.3), following the idea of the theory of reproducing kernels, we shall give a constructive method by iteration.

2. Reproducing Kernels and the Tikhonov Regularization

The good application of reproducing kernels to the Tikhonov regularization is given by the following two general theorems:

Theorem 1 ([2, 5]). *Let H_K be a Hilbert space admitting the reproducing kernel $K(p, q)$ on a set E. Let $L : H_K \to \mathcal{H}$ be a bounded linear operator on H_K into a Hilbert space $\mathcal{H}$. For $\lambda > 0$ introduce the inner product in H_K and call it H_{K_λ} as*

$$\langle f_1, f_2 \rangle_{H_{K_\lambda}} = \lambda \langle f_1, f_2 \rangle_{H_K} + \langle Lf_1, Lf_2 \rangle_{\mathcal{H}}, \tag{2.4}$$

then H_{K_λ} *is the Hilbert space with the reproducing kernel* $K_\lambda(p, q)$ *on E and satisfying the equation*

$$K(\cdot, q) = (\lambda I + L^* L) K_\lambda(\cdot, q), \tag{2.5}$$

where L^* *is the adjoint of* $L : H_K \to \mathcal{H}$.

Theorem 2 ([5]). *Let* H_K, L, $\mathcal{H}$, E *and* K_λ *be as in Theorem* 1. *Then, for any* $\lambda > 0$ *and for any* $g \in \mathcal{H}$, *the extremal function in*

$$\inf_{f \in H_K} (\lambda \| f \|_{H_K}^2 + \| Lf - g \|_{\mathcal{H}}^2) \tag{2.6}$$

exists uniquely and the extremal function is represented by

$$f_{\lambda, g}^*(p) = \langle g, LK_\lambda(\cdot, p) \rangle_{\mathcal{H}} \tag{2.7}$$

which is the member of H_K *attaining the infimum in* (2.6).

For the properties and error estimates for the limit

$$\lim_{\lambda \to 0} f_{\lambda, g}^*(p),$$

see [7, 8]. In particular, when there exists the Moore-Penrose generalized solution for the operator equation

$$Lf = g$$

in (2.6), the limit converges uniformly to the Moore-Penrose generalized solution on any subset of *E* such that $K(p, p)$ is bounded.

For many concrete applications of these general theorems, see, for example, [1, 3, 5, 6].

3. Construction of Approximate Solutions – by Iteration

Following the idea and method in Section 2, we shall consider the extremal problem:

$$\inf_{\mathbf{x} \in \mathbf{R}^n} (\lambda \| \mathbf{x} \|_{\mathbf{R}^n}^2 + \| A\mathbf{x} - \mathbf{y} \|_{\mathbf{R}^m}^2). \tag{3.8}$$

We wish to construct the reproducing kernel for the inner product space $H_{A,\lambda}$ with the norm square

$$\lambda \| \mathbf{x} \|_{\mathbf{R}^n}^2 + \| A\mathbf{x} \|_{\mathbf{R}^m}^2. \tag{3.9}$$

This reproducing kernel $K_\lambda(i, j)$ is calculated by Theorem 1 as the solution of the equation:

$$K_\lambda^{(0)}(i, j) := \frac{1}{\lambda}\delta(i, j) = K_\lambda(i, j) + \frac{1}{\lambda}(AK_\lambda(\cdot, j), A\delta(\cdot, i))_{\mathbf{R}^m}. \tag{3.10}$$

Note here that $\delta(i, j) = \delta_{ij}$, the Kronecker's delta, is the reproducing kernel of the usual inner product space $\mathbf{R}^n$.

Then, the extremal function in (3.8) is given by, for each i component of the vector $\mathbf{x}_\lambda^*(\mathbf{y})$,

$$\mathbf{x}_\lambda^*(\mathbf{y})(i) = (AK_\lambda(\cdot, i), \mathbf{y})_{\mathbf{R}^m}. \tag{3.11}$$

We can solve the equation (3.10) directly ([4, p. 53]), however, in order to avoid the inverse of a large size matrix of $n \times n$ and in order to obtain a constructive method, we shall introduce a new algorithm based on an iterative method.

We shall write (3.9) as follows: For ${}^t\mathbf{a}_j = (a_{j1}, a_{j2}, \ldots, a_{jn})$,

$$\lambda \| \mathbf{x} \|_{\mathbf{R}^n}^2 + \sum_{j=1}^{m} | (\mathbf{a}_j, \mathbf{x})_{\mathbf{R}^n} |^2. \tag{3.12}$$

We shall start with one step of

$$\lambda \| \mathbf{x} \|_{\mathbf{R}^n}^2 + \sum_{j=1}^{1} | (\mathbf{a}_j, \mathbf{x})_{\mathbf{R}^n} |^2. \tag{3.13}$$

Then, the reproducing kernel $K_\lambda^{(1)}(i, j)$ for the space with the norm square (3.13) is given by

$$K_\lambda^{(1)}(i, j) = \frac{1}{\lambda}\delta_{ij} - \frac{a_{1i}a_{1j}}{\lambda(1 + \lambda \| \mathbf{a}_1 \|_{\mathbf{R}^n}^2)}, \tag{3.14}$$

([4, p. 81]). For the second step, for the space with the norm square

$$\lambda \| \mathbf{x} \|_{\mathbf{R}^n}^2 + \sum_{j=1}^{2} | (\mathbf{a}_j, \mathbf{x})_{\mathbf{R}^n} |^2, \tag{3.15}$$

its reproducing kernel $K_\lambda^{(2)}(i, j)$ is constructed from $K_\lambda^{(1)}(i, j)$ similarly as follows:

$$K_\lambda^{(2)}(i, j) = K_\lambda^{(1)}(i, j) - \frac{(\mathbf{a}_2, K_\lambda^{(1)}(\cdot, i))_{\mathbf{R}^n} (\mathbf{a}_2, K_\lambda^{(1)}(\cdot, j))_{\mathbf{R}^n}}{1 + (\mathbf{a}_2, (\mathbf{a}_2, K_\lambda^{(1)}(\cdot, \cdot))_{\mathbf{R}^n})_{\mathbf{R}^n}}. \tag{3.16}$$

Here,

$$(\mathbf{a}_2, K_\lambda^{(1)}(\cdot, i))_{\mathbf{R}^n} = \sum_{\nu=1}^{n} a_{2\nu} K_\lambda^{(1)}(\nu, i)$$

and

$$(\mathbf{a}_2, (\mathbf{a}_2, K_\lambda^{(1)}(\cdot, \cdot))_{\mathbf{R}^n})_{\mathbf{R}^n} = \sum_{\nu, \mu=1}^{n} a_{2\mu} a_{2\nu} K_\lambda^{(1)}(\nu, \mu).$$

We obtain, repeatedly, the desired matrix

$$K_\lambda^{(m)}(i, j) = K_\lambda^{(m-1)}(i, j) - \frac{(\mathbf{a}_m, K_\lambda^{(m-1)}(\cdot, i))_{\mathbf{R}^n} (\mathbf{a}_m, K_\lambda^{(m-1)}(\cdot, j))_{\mathbf{R}^n}}{1 + (\mathbf{a}_m, (\mathbf{a}_m, K_\lambda^{(m-1)}(\cdot, \cdot))_{\mathbf{R}^n})_{\mathbf{R}^n}}. \tag{3.17}$$

Then, we obtain the desired representation

$$\mathbf{x}_\lambda^*(\mathbf{y}) = \| K_\lambda^{(m)}(i, j) \|_{n \times n} \, {}^t A \mathbf{y}. \tag{3.18}$$

4. Error Estimate

In (3.18), when the data $\mathbf{y}$ contains error or noises, we need the estimation of our solutions $\mathbf{x}_\lambda^*(\mathbf{y})$ in terms of $\mathbf{y}$. For this, we can obtain a good estimation in the form:

Theorem 3. *For each i component of the vector* $\mathbf{x}_\lambda^*(\mathbf{y})$,

$$| \mathbf{x}_\lambda^*(\mathbf{y})(i) | \le \frac{1}{\sqrt{\lambda}} \| \mathbf{y} \|_{\mathbf{R}^m}. \tag{4.19}$$

Proof. From (3.11), we have

$$| \mathbf{x}_\lambda^*(\mathbf{y})(i) |^2 \leq \| A(K_\lambda^{(m)}(\cdot, i))_{n\times n} \|_{\mathbf{R}^m}^2 \| \mathbf{y} \|_{\mathbf{R}^m}^2$$

$$= \left(\sum_{j=1}^{m} | (\mathbf{a}_j, K_\lambda^{(m)}(\cdot, i))_{\mathbf{R}^n} |^2 \right) \| \mathbf{y} \|_{\mathbf{R}^m}^2 .$$

Since $K_\lambda^{(m)}(i, j)$ is a reproducing kernel for the inner product space $H_{A,\lambda}$ in (3.9), we have, in particular,

$$\sum_{j=1}^{m} | (\mathbf{a}_j, K_\lambda^{(m)}(\cdot, i))_{\mathbf{R}^n} |^2 \leq \| K_\lambda^{(m)}(\cdot, i) \|_{H_{A,\lambda}}^2$$

$$= K_\lambda^{(m)}(i, i)$$

$$\leq K_\lambda^{(0)}(i, i) = \frac{1}{\lambda},$$

which implies the desired result.

5. Estimates of the Algorithm

In order to obtain (3.14) for all i, j, we need $n^2(n+4)$ times of multiplicity and $n^2(n+1)$ times of addition. Further, by (3.17), from $m = 2$ to $m = m$, in order to obtain $K_\lambda^{(m)}(i, j)$, we need $(m-1)n^2(n^2 + 3n + 2)$ times of multiplicity and $(m-1)n^2(n^2 + 2n + 1)$ times of addition. Therefore, we need the multiplicities and additions of the order

$$mn^4$$

in our algorithm to calculate $K_\lambda^{(m)}(i, j)$.

Note that, for matrices 30×40 and 40×30, their calculation numbers are quite different and the circumstances will be also looked by Table 1 for fixed λ. In matrices in Table 1 and Table 2, their elements are taken over random integers from –10 to +10. We used Mathematica 7.0 in Table 1 and Table 2.

For many matrices appearing from finite element methods and difference equations, the components of the vectors $\{{}^t\mathbf{a}_j\}$ are zero except for around j; that is, except for around the diagonal part. So, in our algorithm, we can save calculations for such matrices. Indeed, we shall assume that $m = n$ and

$${}^t\mathbf{a}_j = (0, 0, ..., a_{j(j-r)}, ..., a_{jj}, ..., a_{j(j+r)}, 0, 0, ..., 0)$$

for a large n for fixed $r > 0$. Then, we can estimate the multiplicities of

$$\frac{n}{3} + \frac{2}{3}n^3 + \frac{1}{12}r + 3n^2 r + 2nr^2 - \frac{1}{12}r^3$$

times to obtain (3.18) in our algorithm; that is, as an order

$$\frac{2}{3}n^3$$

times.

Table 1

Size	Time (sec.)
5×7	1.35×10^{-17}
7×5	1.20×10^{-17}
5×10	0.04
10×5	0.016
10×10	0.015
20×15	0.109
20×20	0.203
30×30	0.985
30×40	1.812
40×30	1.547
40×40	3.079
40×50	4.953
50×40	4.500
50×50	7.546

Table 2

Size	λ	$\sum c_{ij}^2$	$\max\lvert c_{ij} \rvert$
4×6	1.0×10^{-0}	2.99×10^{-4}	6.51×10^{-3}
	1.0×10^{-2}	2.99×10^{-8}	6.51×10^{-5}
	1.0×10^{-4}	2.99×10^{-12}	6.51×10^{-7}
	1.0×10^{-8}	2.99×10^{-20}	6.51×10^{-11}
20×15	1.0×10^{-0}	2.06×10^{-3}	8.15×10^{-3}
	1.0×10^{-2}	2.07×10^{-7}	8.16×10^{-5}
	1.0×10^{-4}	2.07×10^{-11}	8.16×10^{-7}
	1.0×10^{-8}	2.07×10^{-19}	8.16×10^{-11}
50×50	1.0×10^{-0}	9.80×10^{-3}	7.91×10^{-3}
	1.0×10^{-2}	9.85×10^{-7}	7.93×10^{-5}
	1.0×10^{-4}	9.85×10^{-11}	7.93×10^{-7}
	1.0×10^{-8}	9.85×10^{-19}	7.93×10^{-11}

6. Numerical Experiments

We shall illustrate numerical experiments for our algorithm depending on the matrix sizes and λ. For this purpose, we shall estimate

$$C_\lambda := AA_\lambda^\dagger A - A$$

for

$$\| K_\lambda^{(m)}(i, j) \|_{n \times n} \, {}^t A = A_\lambda^\dagger,$$

since the Moore-Penrose generalized inverse $A^\dagger$ satisfies the identity

$$AA^\dagger A = A.$$

In order to look convergence of the C_λ to the zero matrix as λ tending to zero, we estimate the two values

$$\sum c_{ij}^2 \quad \text{and} \quad \max|\, c_{ij}\,|,$$

in Table 2.

Acknowledgements

S. Saitoh is supported in part by Research Unit Mathematics and Applications, University of Aveiro, Portugal, through FCT - Portuguese Foundation for Science and Technology. The author is also supported in part by the Grant-in-Aid for the Scientific Research (C)(2)(No. 21540111). T. Matsuura is supported in part by the Grant-in-Aid for the Scientific Research (C)(2)(No. 20540105) and also in part by Takahashi Industrial and Economic Research Foundation.

References

[1] M. Asaduzzaman, T. Matsuura and S. Saitoh, Constructions of approximate solutions for linear differential equations by reproducing kernels and inverse problems, Advances in Analysis, World Scientific Publishers, Hackensack, NJ, 2005, pp. 335-343.

[2] 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, pp. 55-61.

[3] T. Matsuura and S. Saitoh, Analytical and numerical solutions of linear ordinary differential equations with constant coefficients, J. Anal. Appl. 3 (2005), 1-17.

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

[5] S. Saitoh, Approximate real inversion formulas of the Gaussian convolution, Appl. Anal. 83 (2004), 727-733.

[6] S. Saitoh, Applications of reproducing kernels to best approximations, Tikhonov regularizations and inverse problems, Advances in Analysis, Proceedings of the 4th International ISAAC Congress, World Scientific, 2005, pp. 439-445.

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

[8] 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, pp. 291-298.