JOURNAL OF THE
CHUNGCHEONG MATHEMATICAL SOCIETY
Volume 7, June 1994

Numerical Computation for the Least Squares Solution of Minimum Norm to the First Kind Integral Equations

JAE HEON YUN

ABSTRACT. This paper proposes a numerical method approximating the least squares solution of minimum norm (LSSMN) to the first kind integral equation $Kf = g$ with a special kernel, and then some numerical experiments for this method are provided.

1. Introduction

Throughout this paper, it is assumed that $K : L_2[a,b] \to L_2[c,d]$ is an integral operator defined by $(Kf)(x) = \int_a^b k(x,y)f(y)\,dy$ with the kernel function $k(x,y) \in L_2([c,d] \times [a,b])$, and all functions are real-valued for simplicity of exposition. Then, it is clear that K is a compact operator. Some notations used in this paper are described below. The symbol $(\cdot,\cdot)$ and $\|\cdot\|$ denote the inner product and the associated norm in a given Hilbert space, respectively. If M is a subset of a Hilbert space, $\vee M$ denotes the *closed linear span* of M and $M^{\perp}$ denotes the *orthogonal complement* of M. If T is a bounded linear operator from a Hilbert space H_1 into a Hilbert space H_2, then T^* denotes the *adjoint operator* of T, $N(T)$ denotes the *null* space of T, $R(T)$ denotes the *range* space of T, and $T^{\dagger} : R(T) + R(T)^{\perp} \to H_1$ denotes the *Moore-Penrose generalized inverse* of T [3].

A numerical method for approximating the minimum norm solution to the first kind integral equation $(Kf)(x) = g(x)$ for all $x \in [c,d]$ with $g \in R(K)$ was proposed in [6]. Notice that the condition $g \in R(K)$

Received by the editors on June 27, 1994.
1980 *Mathematics subject classifications*: Primary 65R20, 45L10.

is required for this method to work properly. Since $g \in R(K)$ is a somewhat strong condition, we need to consider the case where $g \in R(K)+R(K)^{\perp}$. The purpose of this paper is to propose a numerical method for approximating the least squares solution of minimum norm (LSSMN) to $Kf = g$ with $g \in R(K) + R(K)^{\perp}$. Section 2 provides some theoretical results for a numerical method to be proposed, Section 3 contains some numerical results for this method, and some conclusions are drawn in Section 4.

2. Rationale of Numerical Method

Since $K : L_2[a,b] \to L_2[c,d]$ is assumed to be an integral operator defined by $(Kf)(x) = \int_a^b k(x,y)f(y)\,dy$ with the kernel function $k(x,y) \in L_2([c,d] \times [a,b])$, it is easy to show that $K^* : L_2[c,d] \to L_2[a,b]$ is an integral operator defined by $(K^*h)(y) = \int_c^d k^*(y,x)h(x)\,dx$ for all $h \in L_2[c,d]$, where $k^*(y,x) = k(x,y)$. It can be also shown that $K^*K : L_2[a,b] \to L_2[a,b]$ is an integral operator defined by $(K^*Kf)(y) = \int_a^b (k^*k)(y,u)f(u)\,du$ for all $f \in L_2[a,b]$, where $(k^*k)(y,u) = \int_c^d k^*(y,x)k(x,u)\,dx = \int_c^d k(x,y)k(x,u)\,dx$. Since $k(x,y) \in L_2([c,d] \times [a,b])$, it is clear that $k^*(y,x) \in L_2([a,b] \times [c,d])$ and $(k^*k)(y,u) \in L_2([a,b] \times [a,b])$.

Since the kernel $k(x,y)$ is assumed to be in $L_2([c,d] \times [a,b])$, by Fubini's theorem $k_x \in L_2[a,b]$ for almost all $x \in [c,d]$ and $k^*_y \in L_2[c,d]$ for almost all $y \in [a,b]$, where $k_x(y) = k(x,y)$ and $k^*_y(x) = k^*(y,x) = k(x,y)$. However, we can not guarantee that $k_x \in L_2[a,b]$ for every $x \in [c,d]$ or $k^*_y \in L_2[c,d]$ for every $y \in [a,b]$. For this reason, two kinds of integral operators K are defined below. Definition 2.1 in the below was already mentioned in [6].

DEFINITION 2.1. The integral operator K with kernel function $k(x,y)$ is said to satisfy *Property (c)* if

(a) $k_x \in L_2[a,b]$ for every $x \in [c,d]$, and

(b) if $(Kf)(x) = 0$ a.e. for a $f \in L_2[a,b]$, then $(Kf)(x) = 0$ for all $x \in [c,d]$.

DEFINITION 2.2. The integral operator K with kernel function $k(x,y)$ is said to satisfy *Property* (c^*) if

(i) $k_y^* \in L_2[c,d]$ for every $y \in [a,b]$, and

(ii) if $(K^*h)(y) = 0$ a.e. for a $h \in L_2[c,d]$, then $(K^*h)(y) = 0$ for all $y \in [a,b]$.

If K satisfies Property (c^*), then from the condition (i) and Schwarz inequality, $(K^*h)(y)$ exists for all $y \in [a,b]$ whenever $h \in L_2[c,d]$. If $k(x,y)$ is a continuous function on $[c,d] \times [a,b]$, then it is obvious that K satisfies Property (c^*). The following lemma describes a result for the integral operator K satisfying Property (c), see [6] for the proof.

LEMMA 2.3. *If the integral operator K satisfies Property (c), then*

$$N(K)^{\perp} = \vee\{\, k_x \; : \; x \in [c,d] \,\}.$$

Now an important result for the integral operator K satisfying Property (c^*) is described.

THEOREM 2.4. *If K satisfies Property (c^*), then the integral operator K^*K satisfies Property (c).*

PROOF. First, we will show that $(k^*k)_y \in L_2[a,b]$ for all $y \in [a,b]$. Let y be a fixed point in $[a,b]$. Using the hypothesis (i) and Schwarz inequality, we obtain

$$
\begin{aligned}
\| (k^*k)_y \|^2 &= \int_a^b |(k^*k)_y(u)|^2\, du \\
&= \int_a^b \Big| \int_c^d k^*(y,x)k(x,u)\, dx \Big|^2 du \\
&= \int_a^b \Big| \int_c^d k_y^*(x)k_u^*(x)\, dx \Big|^2 du \\
&= \int_a^b |(k_y^*, k_u^*)|^2\, du \qquad (1) \\
&\le \int_a^b \| k_y^* \|^2 \| k_u^* \|^2\, du \\
&= \| k_y^* \|^2 \int_a^b \Big(\int_c^d |k^*(u,x)|^2\, dx \Big)\, du \\
&= \| k_y^* \|^2 \Big(\int_a^b \int_c^d |k(x,u)|^2\, dx\, du \Big).
\end{aligned}
$$

Since $k_y^* \in L_2[c,d]$ for every $y \in [a,b]$ and $k(x,y) \in L_2([c,d] \times [a,b])$, from the inequality (1) $(k^*k)_y \in L_2[a,b]$ for every $y \in [a,b]$.

To show that the operator K^*K satisfies the condition (b), suppose that $(K^*Kf)(y) = 0$ a.e. for a $f \in L_2[a,b]$. Since $Kf \in L_2[c,d]$, from the hypothesis (ii) $(K^*Kf)(y) = \big(K^*(Kf)\big)(y) = 0$ for all $y \in [a,b]$. Therefore, K^*K satisfies Property (c).

LEMMA 2.5. $N(K^*K) = N(K)$.

PROOF. Let $f \in N(K)$. Then, $Kf = 0$ and so $K^*Kf = 0$. Hence, $N(K) \subset N(K^*K)$. To show the converse relation, suppose that $f \in N(K^*K)$. Then, $K^*Kf = 0$ and thus $(K^*Kf, g) = 0$ for all $g \in L_2[a,b]$. Since $(K^*Kf, g) = (Kf, Kg)$ and $f \in L_2[a,b]$, $(Kf, Kf) = 0$. It follows that $Kf = 0$, i.e., $f \in N(K)$. Hence, $N(K^*K) \subset N(K)$.

Observe that from Lemma 2.5, $N(K^*K)^\perp = N(K)^\perp$. Using this fact, Lemma 2.3, and Theorem 2.4, the following theorem is immediately obtained.

THEOREM 2.6. *If the integral operator K satisfies Property (c^*), then*

$$N(K)^{\perp} = \vee\{\ (k^*k)_y \ :\ y \in [a,b]\ \},$$

where $(k^*k)_y(u) = (k^*k)(y,u) = (k_y^*, k_u^*)$ for all $u \in [a,b]$.

From now on, it is assumed that $g \in R(K)+R(K)^{\perp}$. Recall that f_0 is the least squares solution of minimum norm (LSSMN) to $Kf = g$ if and only if $f_0 \in N(K)^{\perp}$ and $Kf_0 = Pg$, where P is the orthogonal projection of $L_2[c,d]$ onto the closure of $R(K)$ [4]. Hence, Theorem 2.6 implies that the LSSMN f_0 to $Kf = g$ can be approximated as closely as desired by (finite) linear combinations of the $(k^*k)_y$'s if K satisfies Property (c^*).

LEMMA 2.7. *f_0 is the LSSMN to $Kf = g$ if and only if f_0 is the minimum norm solution to $K^*Kf = K^*g$, see [3] for the proof.*

For given n points $y_1, y_2, \dots, y_n$ in $[a,b]$, the next theorem shows how to choose constants $c_1, c_2, \dots, c_n$ so that $\sum_{j=1}^{n} c_j (k^*k)_{y_j}$ will best approximate the LSSMN f_0 to $Kf = g$ in the L_2-norm without knowing f_0, that is, only $g(x)$ and the kernel $k(x,y)$ need to be known to determine the c_j's. Let $\vec{c} = (c_1, c_2, \dots, c_n)^T$, $\vec{b} = \big((K^*g)(y_1), \dots, (K^*g)(y_n)\big)^T$, and let A be the $n \times n$ matrix whose ij-th component is $\big((k^*k)_{y_j}, (k^*k)_{y_i}\big)$. With these notations, the following theorem is obtained.

THEOREM 2.8. *Let f_0 be the LSSMN to the first kind integral equation $Kf = g$. If K satisfies Property (c^*), then the minimum of $\| f_0 - \sum_{j=1}^{n} c_j (k^*k)_{y_j} \|$ over all constants $c_1, c_2, \dots, c_n$ occurs when $\vec{c}$ is a solution of $A\vec{x} = \vec{b}$. Moreover, the value of $f_n = \sum_{j=1}^{n} c_j (k^*k)_{y_j}$ is independent of which solution $\vec{c}$ of $A\vec{x} = \vec{b}$ one takes.*

PROOF. Let f_0 be the LSSMN to $Kf = g$. Then, $f_0 \in N(K)^{\perp}$. Since K satisfies Property (c^*), $\vee\{(k^*k)_{y_1}, \dots, (k^*k)_{y_n}\}$ is a subspace of

$N(K)^{\perp}$. Hence, $\| f_0 - \sum_{j=1}^{n} c_j(k^*k)_{y_j} \|$ is minimized when the c_j's are chosen so that $\sum_{j=1}^{n} c_j(k^*k)_{y_j}$ is the orthogonal projection of f_0 onto $\vee\{(k^*k)_{y_1}, \ldots, (k^*k)_{y_n}\}$. Therefore, for each i,

$$\big(f_0 - \sum_{j=1}^{n} c_j(k^*k)_{y_j}, (k^*k)_{y_i}\big) = 0,$$

which implies that $\big(f_0, (k^*k)_{y_i}\big) = \sum_{j=1}^{n} \big((k^*k)_{y_j}, (k^*k)_{y_i}\big) c_j$. By Lemma 2.7, f_0 is the minimum norm solution to $K^*Kf = K^*g$ and thus $K^*Kf_0 = K^*g$. Using this fact, one obtains

$$\begin{aligned}\big(f_0, (k^*k)_{y_i}\big) &= \int_a^b f_0(u)(k^*k)_{y_i}(u)\, du \\ &= \int_a^b (k^*k)(y_i, u) f_0(u)\, du \\ &= (K^*Kf_0)(y_i) \;=\; (K^*g)(y_i).\end{aligned}$$

Hence, the first part of this theorem is accomplished.

The second part follows immediately from the uniqueness of the orthogonal projection $f_n = \sum_{j=1}^{n} c_j(k^*k)_{y_j}$ of f_0 onto $\vee\{(k^*k)_{y_1}, \ldots, (k^*k)_{y_n}\}$.

In Theorem 2.8, if $(k^*k)_{y_j}$'s are linearly independent, then A is positive definite and hence $A\vec{x} = \vec{b}$ can be solved uniquely using the Cholesky factorization. But there are often some cases where $(k^*k)_{y_j}$'s are linearly dependent. In this case, Theorem 2.8 shows that any solution to $A\vec{x} = \vec{b}$ can be used to form an approximate solution $f_n = \sum_{j=1}^{n} c_j(k^*k)_{y_j}$ of f_0. Hence, the minimum norm solution, $\vec{c} = A^{\dagger}\vec{b}$, to $A\vec{x} = \vec{b}$ can be used to form the f_n, where $A^{\dagger}$ is the Moore-Penrose generalized inverse of A. Recall that if $A = U\Sigma V^T$ is the singular value decomposition of A, then $\vec{c} = A^{\dagger}\vec{b} = V\Sigma^{\dagger}U^T\vec{b}$ [5].

In what follows, let n be a fixed natural number. We now consider how an optimal set of n points $y_1, y_2, \ldots, y_n$ in $[a, b]$ can be chosen so that $\sum_{j=1}^{n} c_j(k^*k)_{y_j}$ can best approximate the LSSMN f_0 to $Kf = g$. Let $f_n = \sum_{j=1}^{n} c_j(k^*k)_{y_j}$ be the approximate solution obtained by Theorem 2.8 for the LSSMN f_0 to $Kf = g$. Using a property of the orthogonal projection,

$$(2) \quad \| f_0 - f_n\|^2 = \| f_0\|^2 - \| f_n\|^2 \text{ and } \| f_n\|^2 = \sum_{j=1}^{n} c_j(K^*g)(y_j).$$

By virtue of Theorem 2.8 and the equality (2), a *numerical method* for approximating the LSSMN f_0 to $Kf = g$ can be obtained in the following theorem.

THEOREM 2.9. *Let f_0 be the LSSMN to $Kf = g$ and let n be a fixed natural number. Suppose that K satisfies Property (c^*). Then, to minimize $\| f_0 - \sum_{j=1}^{n} c_j(k^*k)_{y_j} \|$ over all $c_1, c_2, \ldots, c_n$ and $y_1, y_2, \ldots, y_n$, one need only solve : minimize $-\sum_{j=1}^{n} c_j(K^*g)(y_j)$ over all $y_1, y_2, \ldots, y_n$, where $(c_1, c_2, \ldots, c_n)^T = A^{\dagger}\vec{b}$ and $\vec{b} = ((K^*g)(y_1), \ldots, (K^*g)(y_n))^T$.*

Notice that $\| f_0\|^2$ being unknown does not inhibit solving this minimization problem. Since $-\|f_n\|^2 = -\sum_{j=1}^{n} c_j(K^*g)(y_j)$ in Theorem 2.9 can be thought of as a function of n variables, $y_1, y_2, \ldots, y_n$, we have a problem of minimizing the nonlinear function $F(y_1, y_2, \ldots, y_n) = -\sum_{j=1}^{n} c_j(K^*g)(y_j)$ subject to $a \leq y_i \leq b$ $(i = 1, 2, \ldots, n)$. This minimization problem can be solved numerically by a subroutine E04JAF in the NAG (Numerical Algorithms Group) library. Once n is given, this numerical method provides an optimal set of n points and n basis functions to approximate the LSSMN f_0 of $Kf = g$.

3. Numerical Implementation

In this section, we consider the numerical implementation of the method developed in Section 2. All computer runs are done in double precision on the Hitachi Data Systems (HDS) AS/9160 using the IBM VS FORTRAN compiler and using the MVS operating system. Double precision arithmetic on this machine means about 15 decimal digits of accuracy.

For simplicity of exposition, $(\cdot,\cdot)_{[a,b]}$ and $(\cdot,\cdot)_{[c,d]}$ denote the inner products on $L_2[a,b]$ and $L_2[c,d]$, respectively. Let n denote the number of points to be used. Once n points $y_1, y_2, \dots, y_n$ in $[a,b]$ are given, the $n \times n$ linear system $A\vec{x} = \vec{b}$ is constructed and then its minimum norm solution is found using the Linpack subroutine DSVDC [2]. Owing to the finite precision of the computer's arithmetic, some computed singular values may be small instead of zero. Thus, computed singular values that give the smallest residual norm $\|\vec{b} - A\vec{x}\|_2$ have been used. All inner products required to form $A = (a_{ij})$ and $\vec{b} = (b_i)$ are computed using the repeated Gauss-Legendre 4-point rule. Specifically, the a_{ij} and b_i are computed using the following relations:

$$\begin{aligned} a_{ij} &= \left((k^*k)_{y_j}, (k^*k)_{y_i}\right)_{[a,b]} \\ &= \left((k^*_{y_j}, k^*_u)_{[c,d]}, (k^*_{y_i}, k^*_u)_{[c,d]}\right)_{[a,b]} \\ b_i &= (K^*g)(y_i) \;=\; (k^*_{y_i}, g)_{[c,d]}. \end{aligned}$$

Once an approximate solution $f_n = \sum_{j=1}^{n} c_j (k^*k)_{y_j}$ for f_0 has been calculated, three different kinds of errors - AE, RE, and RES - between f_0 and f_n are approximated on the set $S = \{a + ih \mid h = \frac{b-a}{100},\ i = 0, 1, \dots, 100\}$ to see how well our method works. They are computed as follows:

$$\begin{aligned} \text{AE} &= \max\{\ |f_0(y) - f_n(y)|\ :\ y \in S\ \} \\ \text{RE} &= \max\{\ \left|\frac{f_0(y) - f_n(y)}{f_0(y)}\right|\ :\ y \in S,\ f_0(y) \neq 0\ \} \\ \text{RES} &= \max\{\ |(K^*Kf_n)(y) - K^*g(y)|\ :\ y \in S\ \} \\ &= \max\{\ \left|\left((k^*k)_y, f_n\right)_{[a,b]} - (k^*_y, g)_{[c,d]}\right|\ :\ y \in S\ \}. \end{aligned}$$

All test examples given below are the first kind integral equations of the form $(Kf)(x) = \int_a^b k(x,y)f(y)\,dy = g(x)$, $c \le x \le d$, which satisfy Property (c^*). The n initial points $y_1, y_2, \dots, y_n$ in $[a,b]$ are chosen by the formula $y_i = a+i\left(\frac{b-a}{n}\right)$ or $y_i = a+i\left(\frac{b-a}{n+1}\right)$ for $i = 1, 2, \dots, n$. It is observed that there is a small difference between these two choices in some of test problems. For this case, the better results are reported. The Fortran code for this numerical method is available from the author upon request. The operator P in the following examples refers to the orthogonal projection of $L_2[c,d]$ onto the closure of $R(K)$.

EXAMPLE 3.1. Let $k(x,y) = x + y$ and $g(x) = x^2 + \frac{1}{6}$, where $x \in [0,1]$ and $y \in [0,1]$. Since $R(K) = \vee\{1, x\}$, $(Pg)(x) = x$. Since $N(K)^\perp = \vee\{1, y\}$, by a simple calculation we obtain $f_0(y) = 4 - 6y$.

EXAMPLE 3.2. Let $k(x,y) = (x-y)^2$ and $g(x) = x^3 - x^2 - \frac{x}{15} + \frac{1}{5}$, where $x \in [0,1]$ and $y \in [0,1]$. Since $R(K) = \vee\{1, x, x^2\}$, $(Pg)(x) = \frac{x^2}{2} - \frac{2}{3}x + \frac{1}{4}$. From the fact that $N(K)^\perp = \vee\{1, y, y^2\}$, $f_0(y) = y$ is easily obtained.

EXAMPLE 3.3. Let $k(x,y) = e^{xy}$ and $g(x) = \frac{e^{x+1}-1}{x+1}$, where $x \in [0,1]$ and $y \in [0,1]$. Observe that $N(K)^\perp = L_2[0,1]$ (see Example 4.1 in [6]). Using this fact, $f_0(y) = e^y$.

EXAMPLE 3.4. Let $k(x,y) = \cos(x-y)$ and $g(x) = \frac{\pi}{2}\sin x + \sin 2x - \frac{8}{3\pi}\cos x$, where $x \in [0,\pi]$ and $y \in [0,\pi]$. Since $R(K) = \vee\{\sin x, \cos x\}$, $(Pg)(x) = \frac{\pi}{2}\sin x$ and $f_0(y) = \sin y$.

The numerical results for the above examples are listed in Table 1. All numbers have been rounded to 3 decimal places. Notice that RES can be thought of as an approximation for the residual norm $\|K^*Kf - K^*g\|$. As can be seen in Table 1, satisfactory results are obtained when $n = 3$ for Examples 3.1 and 3.2 and $n = 2$ for Example 3.4. For Example 3.3, we have tried numerical experiments for various values of $n > 4$, but the errors AE and RE decrease very slowly as

n increases. This may be because it is very difficult to approximate $f_0(y) = e^y$ by a finite linear combination of the $(k^*k)_{y_j}$'s. Notice also that for Example 3.3 the values of AE and RE are relatively much larger than those of RES. This may be because the problem is ill-posed.

Table 1 : Numerical results for test problems

Example	n	AE	RE	RES
3.1	2	2.61×10^{-11}	2.06×10^{-10}	2.21×10^{-13}
	3	8.72×10^{-12}	7.09×10^{-11}	1.05×10^{-13}
	4	1.39×10^{-12}	1.16×10^{-11}	1.28×10^{-13}
3.2	2	3.19×10^{-6}	3.09×10^{-4}	8.53×10^{-7}
	3	2.26×10^{-14}	2.13×10^{-12}	1.95×10^{-16}
	4	1.18×10^{-14}	2.45×10^{-13}	2.14×10^{-16}
3.3	2	2.12×10^{-2}	1.31×10^{-2}	4.50×10^{-5}
	3	5.67×10^{-4}	5.07×10^{-4}	1.14×10^{-10}
	4	5.70×10^{-4}	5.10×10^{-4}	8.38×10^{-12}
3.4	2	2.97×10^{-10}	9.43×10^{-9}	9.44×10^{-16}
	3	2.96×10^{-10}	9.44×10^{-9}	9.44×10^{-16}
	4	2.94×10^{-10}	9.40×10^{-9}	9.38×10^{-16}

4. Conclusions

An advantage of the above method is that for a fixed value of n, it automatically provides an optimal set of n basis functions that best approximate the LSSMN f_0 of the first kind integral equation $Kf = g$ satisfying Property (c^*). A second advantage of this method is that the coefficients of the $(k^*k)_{y_j}$'s are chosen to minimize the norm difference between the approximate solution f_n and the LSSMN f_0 instead of minimizing the residual norm. From numerical results in Table 1, it may be concluded that the accuracy of this method depends on how well one can approximate the LSSMN f_0 by a (finite) linear combination of the $(k^*k)_y$'s and how serious the ill-posedness of the problem is.

References

1. Philip J. Davis, *Interpolation and Approximation*, Dover Publications, Inc., New York, 1975.
2. J. J. Dongarra, J. R. Bunch, C. B. Moler, and G. W. Stewart, *LINPACK User's Guide*, SIAM Publications, Philadelphia, 1979.
3. C. W. Groetsch, *Generalized Inverses of Linear Operators*, Marcel Dekker, Inc., New York, 1977.
4. C. W. Groetsch, *The Theory of Tikhonov regularization for Fredholm Equations of the First Kind*, Pitman Publishing Limited, Boston, 1984.
5. Charles L. Lawson and Richard J. Hanson, *Solving Least Squares Problems*, Prentice-Hall, Inc., Englewood Cliffs, 1974.
6. Jae H. Yun, *A Numerical Method for the Minimum Norm Solution to the First Kind Integral Equations*, J. Chungcheong Math. Soc. **6** (1993), 25-43.

Department of Mathematics
College of Natural Sciences
Chungbuk National University
Cheongju, 360-763, Korea