

Convergence Rates of Approximate Least Squares Solutions of Linear Integral and Operator Equations of the First Kind*

By M. Z. Nashed and Grace Wahba

Abstract. We consider approximations $\{x_n\}$ obtained by moment discretization to (i) the minimal \mathcal{L}_2 -norm solution of $\mathcal{K}x = y$ where \mathcal{K} is a Hilbert-Schmidt integral operator on \mathcal{L}_2 , and to (ii) the least squares solution of minimal \mathcal{L}_2 -norm of the same equation when y is not in the range $\mathcal{R}(\mathcal{K})$ of \mathcal{K} . In case (i), if $y \in \mathcal{R}(\mathcal{K})$, then $x_n \rightarrow \mathcal{K}^\dagger y$, where \mathcal{K}^\dagger is the generalized inverse of \mathcal{K} , and $\|x_n\| \rightarrow \infty$ otherwise. Rates of convergence are given in this case if further $\mathcal{K}^\dagger y \in \mathcal{K}^*(\mathcal{L}_2)$, where \mathcal{K}^* is the adjoint of \mathcal{K} , and the Hilbert-Schmidt kernel of $\mathcal{K}\mathcal{K}^*$ satisfies certain smoothness conditions. In case (ii), if $y \in \mathcal{R}(\mathcal{K}) \oplus \mathcal{R}(\mathcal{K})^\perp$, then $x_n \rightarrow \mathcal{K}^\dagger y$, and $\|x_n\| \rightarrow \infty$ otherwise. If further $\mathcal{K}^\dagger y \in \mathcal{K}^*\mathcal{K}(\mathcal{L}_2)$, then rates of convergence are given in terms of the smoothness properties of the Hilbert-Schmidt kernel of $(\mathcal{K}\mathcal{K}^*)^2$.

Some of these results are generalized to a class of linear operator equations on abstract Hilbert spaces.

1. Introduction. It is well known that the numerical solution of the integral equation of the first kind

$$(1.1) \quad \mathcal{K}x := \int_0^1 K(\cdot, t)x(t) dt = y$$

leads to certain difficulties which are not encountered in integral equations of the second kind

$$(1.2) \quad (I - \lambda\mathcal{K})x = y.$$

Considerable light can be shed on the problems (1.1) and (1.2) when viewed in the context of linear operator theory. The operator \mathcal{K} in (1.1) under mild conditions is compact on $\mathcal{L}_2[0, 1]$ into itself; hence, its range $\mathcal{R}(\mathcal{K})$ is always a nonclosed subspace of $\mathcal{L}_2[0, 1]$ (unless it is a degenerate operator, i.e., $\mathcal{R}(\mathcal{K})$ is finite dimensional). In contrast, $\mathcal{R}(I - \lambda\mathcal{K})$ associated with (1.2) is always a closed subspace of $\mathcal{L}_2[0, 1]$ for any λ . Consequently, the generalized inverse of \mathcal{K} (and in particular the inverse of \mathcal{K} if it exists) is unbounded and densely defined, whereas the generalized inverse of $I - \lambda\mathcal{K}$ is bounded and everywhere defined (see [5] or [6] for amplification of these remarks if necessary). Also, since $\mathcal{R}(\mathcal{K})$ is only dense in $\mathcal{N}(\mathcal{K}^*)^\perp$, the orthogonal complement of the null space of \mathcal{K}^* , the Fredholm alternative theorem for (1.2) fails

Received September 18, 1972.

AMS (MOS) subject classifications (1970). Primary 45B05, 45C05, 45L10, 65F20, 65J05, 65R05; Secondary 45N05, 46N05, 47B10, 47B30.

Key words and phrases. First kind integral equations, linear operator equations, moment discretization, convergence rates, least squares solutions, generalized inverses, reproducing kernel Hilbert spaces, Picard criterion.

* Sponsored by the United States Army under Contract DA-31-124-ARO-D-462 at the Mathematics Research Center, University of Wisconsin, Madison.

Copyright © 1974, American Mathematical Society

to apply to (1.1) and one needs a condition (Picard's criterion [12]) under which an element

$$y \in \mathfrak{R}(\mathfrak{K}^*)^\perp = \overline{\mathfrak{R}(\mathfrak{K})}$$

belongs to $\mathfrak{R}(\mathfrak{K})$.

From the numerical and approximation aspects of (1.1) and (1.2), the preceding operator-theoretic facts are reflected in the following ways.

Firstly, the problem (1.1) is *ill*-posed: The solution (or generalized inverse solution) of (1.1) does not depend continuously on the data y in the \mathfrak{L}_2 -setting. This inherent instability also carries over to the solution of the algebraic system arising from discretization of the integral equation, whenever one seeks greater accuracy. The numerical manifestation of this instability and various approaches to approximate numerical solutions of (1.1) have been described by many authors. We cite the recent works of Hanson [3] and Strand [13], where the reader may backtrack the literature on this subject from the bibliography cited therein; see also the comprehensive bibliography on integral equations prepared recently by Noble [8].

Secondly, iterative methods for the solution of (1.1) exhibit convergence patterns which are markedly different from their counterpart for the solution of (1.2). Hellinger and Toeplitz remarked in [4] that a method of solution of (1.1) by iteration is not available. Since then several authors have analyzed the problem of convergence of successive approximation for (1.1). See in particular the recent paper of Diaz and Metcalf [2], the earlier work of Landweber, Fridman, and Bialy cited in [2] and [6, pp. 344–345], and [5], [13]. The method of successive approximation (with a parameter), steepest descent and conjugate gradient methods for least squares solutions of (1.1), all exhibit slow convergence patterns (as $1/n$); see [5], [6], and [11]. In contrast, the rate of convergence of these methods to least squares solutions of (1.2) is at least geometric (see [5], [6], [9]).

The purpose of this paper is to establish the convergence of approximation schemes based on moment discretization to the least squares solution of minimal norm of (1.1) (or to the minimal norm solution if $y \in \mathfrak{R}(\mathfrak{K})$), and to provide sharp convergence rates under some mild smoothness assumptions on the function y and the kernel $K(s, t)$. Furthermore we show that if the minimal norm solution is smooth, then the domain of \mathfrak{K} may be viewed as a reproducing kernel Hilbert space (RKHS) and thus we obtain pointwise convergence in this case.

In Section 2, we state some preliminaries about RKHS and recast Picard's criterion in this setting. The main convergence results for integral equations of the first kind are stated in Sections 3 and 4. Finally, in Section 5, some of these results are generalized to a class of linear operator equations on Hilbert space.

2. Picard's Criterion in Reproducing Kernel Spaces and the Generalized Inverse of \mathfrak{K} . A Hilbert space \mathfrak{H} of (real-valued) functions f on the interval T is said to be a reproducing kernel Hilbert space (RKHS) if all the evaluation functionals $\mathfrak{E}_t(f) = f(t)$, $f \in \mathfrak{H}$, for each fixed $t \in T$ are continuous. Then, by Riesz's theorem for each $t \in T$, there exists a unique element, call it Q_t , in \mathfrak{H} such that

$$(2.1) \quad \langle f, Q_t \rangle = f(t), \quad f \in \mathfrak{H}.$$

Let

$$(2.2) \quad Q(t, t') := \langle Q_t, Q_{t'} \rangle, \quad t, t' \in T.$$

Note that $Q_t(t') = Q(t, t')$ by applying (2.1) to (2.2); for this reason, $Q(t, t')$ is called the reproducing kernel (RK). Let \mathcal{H}_Q denote the RKHS with RK Q and inner product $\langle \cdot, \cdot \rangle_Q$. Clearly, $Q(t, t')$ is a nonnegative definite symmetric kernel. Conversely, by the Aronszajn-Moore theorem (see [1]), every nonnegative definite symmetric function Q on $T \times T$ determines a unique Hilbert function space \mathcal{H}_Q for which Q is the reproducing kernel. In an RKHS, the element representing a given bounded linear functional ϕ can be expressed by means of the RK; more precisely, $\phi(f) = \langle f, h \rangle_Q$ for all $f \in \mathcal{H}_Q$, where $h(t) = \phi(Q_t)$, and similarly for a bounded linear operator L on \mathcal{H}_Q to \mathcal{H}_Q :

$$(Lf)(t) = \langle Lf, Q_t \rangle_Q = \langle f, L^*Q_t \rangle_Q.$$

Note that $\mathcal{L}_2[0, 1]$ is not an RKHS. If $Q(t, t')$ is continuous on $T \times T$ (the only case we will consider here), then \mathcal{H}_Q is a space of continuous functions. To see this, note that

$$|f(t) - f(t')| = |\langle f, Q_t - Q_{t'} \rangle_Q| \leq \|f\|_Q \|Q_t - Q_{t'}\|_Q,$$

and

$$\|Q_t - Q_{t'}\|_Q^2 = Q(t, t) - 2Q(t, t') + Q(t', t').$$

One reason for using RKHS in approximation theory and numerical analysis is that strong convergence in \mathcal{H}_Q implies pointwise convergence, viz.,

$$|f(t) - f_n(t)| = |\langle f - f_n, Q_t \rangle_Q| \leq \|f - f_n\|_Q \sqrt{Q(t, t)}.$$

For properties of RKHS see Aronszajn [1] and Shapiro [10].

For the purpose of this paper, we show that the range of a Hilbert-Schmidt operator applied to $\mathcal{L}_2(S)$ is an RKHS. We shall use (\cdot, \cdot) to denote inner product in \mathcal{L}_2 , reserving the notation $\langle \cdot, \cdot \rangle_Q$ for inner product in RKHS, and omitting the subscript when there is no confusion. The \mathcal{L}_2 -norm is denoted by $\|\cdot\|$, whereas the RKHS-norm is denoted by $\|\cdot\|_Q$.

PROPOSITION 2.1. *Let S, T be intervals, and let $G(t, s)$ be defined on $T \times S$ with the property that, for each fixed $t \in T$, $G(t, s) \in \mathcal{L}_2(S)$. Consider the family of functions f defined by*

$$(2.3) \quad f(t) = \int_S G(t, s)p(s) ds =: (\mathcal{G}p)(t)$$

for $p \in \mathcal{L}_2(S)$. This collection of f 's is an RKHS $\mathcal{H}_Q = \mathcal{G}[\mathcal{L}_2(S)]$ with RK

$$(2.4) \quad Q(t, t') = \int_S G(t, s)G(t', s) ds$$

where the inner product in \mathcal{H}_Q is given by

$$(2.5) \quad \langle f_1, f_2 \rangle_Q = \int_S p_1(s)p_2(s) ds$$

where $p_i(s)$, $i = 1, 2$, is the element of minimal $\mathcal{L}_2(S)$ -norm which satisfies (2.3).

Proof. To see that \mathcal{H}_Q is a Hilbert space with the given inner product, let V be the smallest closed subspace in $\mathcal{L}_2(S)$ containing the functions $G(t, \cdot)$, $t \in T$. Then

for each f of the form (2.3), there is a unique $p \in V$ which satisfies (2.3), and this p is that element of minimal $\mathcal{L}_2(S)$ -norm for which (2.3) holds. This correspondence between \mathcal{H}_Q and V is linear and, in fact, is a metric isomorphism if we adopt the inner product (2.5). To see that $Q(t, t')$ is the RK, we must show that $\langle f, Q_t \rangle_Q = f(t)$, for $f \in \mathcal{H}_Q$, $t \in T$, where $Q_t(t') = Q(t, t')$. Let $f(t) = \int_S G(t, s)p(s) ds$ with $p \in V$. Note that $Q_{t'}(t) = \int_S G(t, s)G(t', s) ds$. Letting $f_1 = f$, $p_1 = p$, $f_2 = Q_{t'}$, $p_2 = G(t', \cdot)$, we get, using (2.5),

$$\begin{aligned} \langle f_1, f_2 \rangle_Q &= \int_S p_1(s)p_2(s) ds = \langle f, Q_{t'} \rangle_Q \\ &= \int_S p(s)G(t', s) ds = f(t') \end{aligned}$$

so that Q is the RK for \mathcal{H}_Q .

PROPOSITION 2.2. *Let $Q(t, t')$ be a continuous reproducing kernel. Then*

$$(a) \quad Q(t, t') = \sum_{\nu=1}^{\infty} \lambda_{\nu} \phi_{\nu}(t) \phi_{\nu}(t')$$

converges pointwise, where $\{\lambda_{\nu}\}$, $\{\phi_{\nu}\}$ are the eigenvalues and orthonormalized eigenfunctions of Q , and $\sum \lambda_{\nu}^2 = \iint Q^2(t, t') dt dt'$.

$$(b) \quad \mathcal{H}_Q = \left\{ f : \sum_{\nu=1}^{\infty} (f, \phi_{\nu})^2 / \lambda_{\nu} < \infty \right\},$$

where if $\lambda_{\nu} = 0$, we must have $(f, \phi_{\nu}) = 0$, and the inner product in \mathcal{H}_Q is given by

$$\langle f, g \rangle_Q = \sum_{\nu=1}^{\infty} \frac{1}{\lambda_{\nu}} (f, \phi_{\nu})(g, \phi_{\nu}) = (Q^{-1/2}f, Q^{-1/2}g)$$

for $f, g \in \mathcal{H}_Q$.

Proof. Part (a) is simply a restatement of Mercer's theorem (see for instance [12]) since $Q(t, t')$ is a nonnegative kernel.

To prove (b) note that $(Q_t, \phi_{\nu}) = \lambda_{\nu} \phi_{\nu}(t)$, and

$$\sum_{\nu=1}^{\infty} \frac{1}{\lambda_{\nu}} (f, \phi_{\nu})(Q_t, \phi_{\nu}) = \sum_{\nu=1}^{\infty} (f, \phi_{\nu}) \phi_{\nu}(t) = f(t)$$

and thus $\langle Q_t, f \rangle_Q = f(t)$.

We now consider the integral equation

$$(2.6) \quad (\mathcal{K}x)(t) := \int_S K(t, s)x(s) ds = y(t), \quad t \in T.$$

We assume that the functions $k_t := K(t, \cdot) \in \mathcal{L}_2(S)$ for $t \in T$, and $k_{s^*} := K(\cdot, s) \in \mathcal{L}_2(T)$ for $s \in S$, and the kernel $Q(t, t')$ defined by

$$(2.7) \quad Q(t, t') = (k_t, k_{t'}) = \int_S K(t, s)K(t', s) ds$$

is continuous on $T \times T$. Then $\int_T \int_S K^2(t, s) dt ds < \infty$, and the integral operator \mathcal{K} maps $\mathcal{L}_2(S)$ into $\mathcal{L}_2(T)$. It is well known that, for $y \in \mathcal{L}_2(T)$, the necessary and sufficient condition that $y \in \mathcal{R}(\mathcal{K})$ belongs to $\mathcal{R}(\mathcal{K})$ is that

$$(2.8) \quad \sum_{\nu=1}^{\infty} (y, \phi_{\nu})^2 / \lambda_{\nu} < \infty,$$

where $\{\lambda_{\nu}\}$ and $\{\phi_{\nu}\}$ are the eigenvalues and orthonormalized eigenfunctions of the nonnegative definite symmetric Hilbert-Schmidt operator $\mathcal{Q} = \mathcal{K}\mathcal{K}^*$ with kernel $Q(t, t')$. If ϕ_{ν} is an eigenfunction with $\lambda_{\nu} = 0$, then we must have $(y, \phi_{\nu}) = 0$ for $y \in \mathcal{R}(\mathcal{K})$. Condition (2.8) is often known as Picard's criterion. Equivalent manifestations of this criterion have been given recently by Diaz and Metcalf [2]. Now by Proposition 1.1, $\mathcal{R}(\mathcal{K})$ is an RKHS and, in view of Proposition 2.2(b), condition (2.8) is equivalent to

$$(2.9) \quad y \in \mathcal{H}_Q$$

where \mathcal{H}_Q is the RKHS with RK $Q(t, t')$ given by (2.7), and where the norm in \mathcal{H}_Q is given by

$$\|y\|_Q^2 = \sum_{\nu=1}^{\infty} (y, \phi_{\nu})^2 / \lambda_{\nu}.$$

It should be noted that the introduction of RHKS and the casting of Picard's criterion (2.8) in the form (2.9) is not a mere formality. For one thing, $\mathcal{R}(\mathcal{K})$ is non-closed in $\mathcal{L}_2(T)$, unless $\mathcal{R}(\mathcal{K})$ is finite dimensional; whereas $\mathcal{R}(\mathcal{K})$ being an RKHS is a closed subspace in the RKHS topology. The usefulness of RKHS in the context of linear integral equations is highlighted when we consider approximate methods in the next three sections; for then we are able to establish convergence and sharp rates.

The approximations that we study will converge to the generalized inverse of \mathcal{K} , and, in case $y \in \mathcal{R}(\mathcal{K})$, to the minimal norm solution of (2.6). We now describe the generalized inverse for our setting. Suppose first that $y \in \mathcal{R}(\mathcal{K})$. Then there exists a unique element of minimal norm in $\mathcal{L}_2(S)$ which satisfies $\mathcal{K}x = y$. This element may be obtained as follows. Let V be the closure of the span of $\{k_t : t \in T\}$ in $\mathcal{L}_2(S)$ and P_V be the projection operator onto V . Note that $\mathcal{R}(\mathcal{K}) = V^{\perp}$ since $v \in \mathcal{R}(\mathcal{K})$ if and only if $(v, k_t) = 0, t \in T$. Thus if x is any element in $\mathcal{L}_2(S)$ satisfying (2.6), $P_V x$ is the unique solution of minimal norm. Denote $P_V x$ by $\mathcal{K}^{\dagger}y$. The domain of \mathcal{K}^{\dagger} may be extended to $\mathcal{R}(\mathcal{K}) \oplus \mathcal{R}(\mathcal{K})^{\perp}$ by defining $\mathcal{K}^{\dagger}y = \mathcal{K}^{\dagger}y_0$ where $y = y_0 + y_1, y_0 \in \mathcal{R}(\mathcal{K})$ and $y_1 \in \mathcal{R}(\mathcal{K})^{\perp}$. Thus $x = \mathcal{K}^{\dagger}y$ is the unique element in $\mathcal{L}_2(S)$ of minimal \mathcal{L}_2 -norm for which $\|y - \mathcal{K}x\|$ attains its infimum.

3. Convergence of Approximations to Minimal-Norm Solution of $\mathcal{K}x = y$.

We consider first the case when $y \in \mathcal{R}(\mathcal{K})$, but \mathcal{K} is not necessarily one-to-one. Suppose that $y(t)$ is known on the set $T_n = \{t_1, t_2, \dots, t_n\} \subset T$ and consider the set of linear equations obtained by moment discretization on T_n of (2.6):

$$(3.1) \quad \int_0^1 K(t_i, s)x(s) ds = y(t_i), \quad i = 1, \dots, n.$$

We introduce the notation $K_i = K(t_i, \cdot) = k_{t_i}(\cdot), y_i = y(t_i)$ and rewrite (3.1) in the form

$$(K_i, x) = y_i, \quad i = 1, \dots, n.$$

Define the operator $\mathcal{C}_n : \mathcal{L}_2(S) \rightarrow R^n$ by

$$\mathcal{C}_n x = \begin{pmatrix} (K_1, x) \\ \vdots \\ (K_n, x) \end{pmatrix} \quad \text{and} \quad Y_n = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

$$(3.2) \quad \mathcal{C}_n x = Y_n.$$

Since $\mathcal{R}(\mathcal{C}_n)$ is finite dimensional, the generalized inverse \mathcal{C}_n^\dagger exists on R^n and the set of all least squares solutions of (3.2) is given by

$$S_n = \mathcal{C}_n^\dagger Y_n \oplus \mathfrak{N}(\mathcal{C}_n).$$

In particular, $x_n = \mathcal{C}_n^\dagger Y_n$ is the element of minimal $\mathcal{L}_2(S)$ -norm which minimizes $\|\mathcal{C}_n x - Y_n\|_{R^n}$. We consider convergence of $\{x_n\}$ to $\mathfrak{K}^\dagger y$ for $y \in \mathcal{D}(\mathfrak{K}^\dagger) = \mathcal{R}(\mathfrak{K}) \oplus \mathcal{R}(\mathfrak{K})^\perp$.

The solution x_n is given explicitly by

$$(3.3) \quad x_n(\cdot) = (y_1, y_2, \dots, y_n) Q_n^\dagger (k_{t_1}, k_{t_2}, \dots, k_{t_n})',$$

where Q_n is the $n \times n$ matrix whose ij th element is given by $Q(t_i, t_j)$ of (2.7) and the prime denotes transpose. We shall assume that the set $\{k_t : t \in T_n\}$ is linearly independent for every finite n . This is not a serious restriction, for if there exist constants c_1, \dots, c_n , not all zero, for which $\sum_{i=1}^n c_i k_{t_i} = 0$, then every $y \in \mathcal{R}(\mathfrak{K})$ would satisfy $\sum_{i=1}^n c_i y(t_i) = 0$. Under this assumption, Q_n^\dagger becomes Q_n^{-1} . Let

$$(3.4) \quad \Delta_n := \sup_{t \in T} \left(\inf_{t_i \in T_n} |t - t_i| \right).$$

Note that if $\lim_{n \rightarrow \infty} \Delta_n = 0$, then $\bigcup_n T_n$ is dense in T . We now state the first convergence theorem.

THEOREM 3.1. (a) *If $Q(t, t')$ is continuous on $T \times T$ and if $y \in \mathcal{R}(\mathfrak{K})$, then*

$$(3.5) \quad \lim_{\Delta_n \rightarrow 0} \|x_n - \mathfrak{K}^\dagger y\| = 0,$$

where Δ_n is defined as in (3.4).

(b) *If $y \in \mathcal{R}(\mathfrak{K})$, $\mathfrak{K}^\dagger y \in \mathfrak{K}^*(\mathcal{L}_2(T))$, or equivalently $y \in \mathcal{Q}(\mathcal{L}_2(T))$, and $Q(t, t')$ satisfies*

(i) *$(\partial^l / \partial t^l) Q(t, t')$ exists and is continuous on $T \times T$ for $t \neq t', l = 0, 1, 2, \dots, 2m$, $(\partial^l / \partial t^l) Q(t, t')$ exists and is continuous on $T \times T$ for $l = 0, 1, 2, \dots, 2m - 2$; and*

$$(ii) \quad \lim_{t \nearrow t'} \frac{\partial^{2m-1}}{\partial t^{2m-1}} Q(t, t') \quad \text{and} \quad \lim_{t \searrow t'} \frac{\partial^{2m-1}}{\partial t^{2m-1}} Q(t, t')$$

exist and are bounded for all $t' \in T$, then

$$(3.6) \quad \|x_n - \mathfrak{K}^\dagger y\| = O(\Delta_n^m).$$

(c) *If y does not coincide with some element $y_0 \in \mathcal{R}(\mathfrak{K})$ on $\bigcup_n T_n$, then $\|x_n\| \rightarrow \infty$.*

Proof. (a) Let $Q_t(\cdot)$ be that function on T defined by $Q_t(\cdot) = Q(t, \cdot)$. By properties of RKHS, we have $Q_t \in \mathfrak{K}_Q$ for all $t \in T$, and $\langle Q_t, h \rangle_Q = h(t)$ for $h \in \mathfrak{K}_Q, t \in T$. This shows that the set $\{Q_t : t \in T\}$ spans \mathfrak{K}_Q , and $\langle Q_t, Q_{t'} \rangle_Q = Q(t, t')$. Thus there exists an isometric isomorphism between \mathfrak{K}_Q and V generated by the correspondence

$$(3.7) \quad Q_t \in \mathfrak{K}_Q \sim k_t \in V.$$

To see this, note that $\langle Q_t, Q_{t'} \rangle_Q = Q(t, t') = (k_t, k_{t'})$ for $t, t' \in T$. Further note that $Q_t = \mathcal{K}k_t$, so that $\mathcal{K}^\dagger Q_t = k_t$, since $k_t \in V$, and $y \in \mathcal{K}Q \sim x \in V$ by the correspondence (3.7) if and only if $y = \mathcal{K}x$, $x = \mathcal{K}^\dagger y$.

It follows that, for $y \in \mathcal{R}(\mathcal{K})$,

$$(3.8) \quad \|x_n - \mathcal{K}^\dagger y\| = \|y - P_{T_n} y\|_Q$$

where $P_{T_n} y$ is given by

$$(3.9) \quad P_{T_n} y = \mathcal{K}x_n = (y_1, \dots, y_n) Q_n^{-1} (Q_{t_1}, \dots, Q_{t_n})'$$

and is the projection (in $\mathcal{K}Q$) of y onto the subspace spanned by $\{Q_t : t \in T_n\}$. If Q is continuous and $\lim_{n \rightarrow \infty} \Delta_n = 0$, then $\text{span} \{Q_t : t \in \bigcup_n T_n\}$ is dense in $\mathcal{K}Q$; thus $\lim_{\Delta_n \rightarrow 0} \|y - P_{T_n} y\|_Q = 0$, and using (3.8) it follows that (3.5) holds.

(b) For $y \in \mathcal{R}(\mathcal{K})$ and $\mathcal{K}^\dagger y \in \mathcal{K}^*(\mathcal{L}_2(T))$, we may write $\mathcal{K}^\dagger y = \mathcal{K}^* \rho$, $y = \mathcal{K} \mathcal{K}^* \rho$ for some $\rho \in \mathcal{L}_2(T)$ and hence

$$(3.10) \quad y(t) = \int_T Q(t, t') \rho(t') dt'$$

For y of the form (3.10), the proof of Theorem 1 of [14] applies directly to give

$$(3.11) \quad \|y - P_{T_n} y\|_Q = O(\Delta_n^m).$$

Then (3.6) follows from (3.8) and (3.11).

To prove (c) we show that if $\lim_{\Delta_n \rightarrow 0} \|x_n\|$ is finite, then y must coincide with an element of $\mathcal{K}Q = \mathcal{R}(\mathcal{K})$ for $t \in \bigcup_{n=1}^\infty T_n$. From (3.3) and the fact $\langle Q_t, Q_{t'} \rangle_Q = Q(t, t') = (k_t, k_{t'})$, we get

$$\|x_n\|^2 = (y(t_1), \dots, y(t_n)) Q_n^{-1} (y(t_1), \dots, y(t_n))'.$$

Denote $\mathcal{K}x_n$ by w_n . Then $\|w_n\|_Q = \|x_n\|$. Note also that w_n is the element of minimal $\mathcal{K}Q$ -norm which satisfies $w_n(t) = y(t)$, $t \in T_n$. Without loss of generality, we may assume that $T_n \subset T_{n+m}$ and thus $P_{T_n} w_{n+m} = w_n$. Then $\|w_n\|_Q$, $n = 1, 2, \dots$, is a nondecreasing sequence, and $\|w_{n+m} - w_n\|_Q^2 = \|w_{n+m}\|_Q^2 - \|w_n\|_Q^2$. Therefore, if $\lim_{n \rightarrow \infty} \|w_n\|_Q$ is finite, then $\{w_n\}$ is a Cauchy sequence in $\mathcal{K}Q$ which has a limit \bar{w} in $\mathcal{K}Q$. Now, strong convergence in $\mathcal{K}Q$ implies pointwise convergence, so that we must have $y(t) = \bar{w}(t)$ for $t \in T_n$, $n = 1, 2, \dots$. Since $\|x_n\| = \|w_n\|_Q$, part (c) is proved.

The proof of Theorem 3.1 does not require $\mathcal{D}(\mathcal{K}) = \mathcal{L}_2$. (See Section 5, and [14].) As an application of Theorem 3.1, we consider the case when $u = \mathcal{K}^\dagger y$ is a continuously differentiable function on $[0, 1]$. Without loss of generality, we may take $u(0) = 0$ (since

$$\int_0^1 K(t, s)[u(s) - a] ds = y(t) - a \int_0^1 K(t, s) ds$$

and we may solve

$$\int_0^1 K(t, s) \tilde{u}(s) ds = \tilde{y}(t)$$

instead of the original equation).

We consider the problem: Find $u_n(s)$ to minimize $\int_0^1 [u'(s)]^2 ds$ subject to $u(0) = 0$ and

$$\int_0^1 K(t, s)u(s) ds = y(t), \quad t \in T_n = \{t_1, \dots, t_n\} \subset [0, 1].$$

We now show that this problem can be recast in the setting of RKHS and establish convergence of $\{u_n\}$ on the basis of Theorem 3.1.

Let \mathfrak{H}_R be the space of all absolutely continuous functions x with derivatives in $\mathcal{L}_2[0, 1]$, and $x(0) = 0$. The space \mathfrak{H}_R is an RKHS with inner product

$$\langle x_1, x_2 \rangle_R = \int_0^1 x_1'(s)x_2'(s) ds$$

and reproducing kernel

$$R(s, s') = \min(s, s').$$

Let $\eta_i(s) = \int_0^1 K(t, t')R_s(t') dt$. Then, for $x \in \mathfrak{H}_R$, we have

$$(3.12) \quad \int_0^1 K(t, s)x(s) ds = \langle \eta_t, x \rangle_R.$$

To show (3.12), we note that

$$\begin{aligned} \eta_t(s) &= \int_0^1 K(t, t') \min(s, t') dt' \\ &= \int_0^s t' K(t, t') dt' + s \int_s^1 K(t, t') dt' \end{aligned}$$

and $(d/ds)\eta_t(s) = \int_s^1 K(t, u) du$, which we shall denote by $\Gamma_t(s)$. Then

$$\begin{aligned} \langle \eta_t, x \rangle_R &= \int_0^1 \frac{d}{ds} \eta_t(s)x'(s) ds = \int_0^1 \Gamma_t(s)x'(s) ds \\ &= \int_0^1 K(t, s)x(s) ds. \end{aligned}$$

Now, the element of minimal \mathfrak{H}_R -norm which satisfies $\langle \eta_t, x \rangle_R = y(t)$, for $t \in T_n$, is

$$u_n(s) = (\eta_{t_1}(s), \dots, \eta_{t_n}(s))Q_n^{-1}(y(t_1), \dots, y(t_n))'$$

where Q_n is the $n \times n$ matrix whose ij th entry is given by

$$Q(t_i, t_j) = \langle \eta_{t_i}, \eta_{t_j} \rangle = \int_0^1 \Gamma_{t_i}(s)\Gamma_{t_j}(s) ds.$$

We assume that $Q(t, t')$ is continuous and that $\lim_{n \rightarrow \infty} \Delta_n = 0$ (cf. (3.4)). Then, as in Theorem 3.1, we have $\|u_n - \mathfrak{K}^t y\|_R \rightarrow 0$ as $n \rightarrow \infty$. and thus

$$\int_0^1 [u_n'(s) - u'(s)]^2 ds \rightarrow 0.$$

Furthermore, we also get pointwise convergence:

$$|u_n(s) - u(s)| \leq \|u_n - u\|_R \|R_s\|_R = \|u_n - u\|_R \sqrt{s} \rightarrow 0$$

since $\|R_s\|_R^2 = R(s, s) = s$.

4. **Convergence to $\mathcal{K}^\dagger y$ for $y \in \mathcal{R}(\mathcal{K}) \oplus \mathcal{R}(\mathcal{K})^\perp$.** We consider now the case $y \notin \mathcal{R}(\mathcal{K})$, and seek approximations to $\mathcal{K}^\dagger y$, the least squares solution of minimal norm of (2.6). Note that $\mathcal{K}^\dagger y$ is also the minimal norm solution of the equation

$$(4.1) \quad \mathcal{K}^* \mathcal{K} x = \mathcal{K}^* y$$

for $y \in \mathcal{D}(\mathcal{K}^\dagger)$. Furthermore, it is easy to show that $(\mathcal{K}^* \mathcal{K})^\dagger \mathcal{K}^* y = \mathcal{K}^\dagger y$ on $\mathcal{D}(\mathcal{K}^\dagger)$.

Let $\tilde{Q} = \mathcal{K}^* \mathcal{K}$, and $w = \mathcal{K}^* y$. The Hilbert-Schmidt operator \tilde{Q} has the kernel

$$\tilde{Q}(s, s') = (k_s^*, k_{s'}^*) = \int_T K(t, s) K(t, s') dt.$$

Let x_n be the element of minimal $\mathcal{L}_2(S)$ -norm which minimizes

$$\sum_{i=1}^n [(\tilde{Q}x)(s_i) - w(s_i)]^2,$$

where $S_n = \{s_1, \dots, s_n\} \subset S$.

By our previous assumptions, the function \tilde{q}_s , defined by $\tilde{q}_s(s') = \tilde{Q}(s, s')$, is in $\mathcal{L}_2(S)$ for each $s \in S$. If we assume that the set $\{\tilde{q}_s : s \in S_n\}$ is linearly independent for every finite n , then x_n is given by

$$(4.2) \quad x_n = (w_1, w_2, \dots, w_n) P_n^{-1} (\tilde{q}_{s_1}, \tilde{q}_{s_2}, \dots, \tilde{q}_{s_n})',$$

where P_n is the $n \times n$ matrix with ij th element given by $P(s_i, s_j)$,

$$P(s, s') = (\tilde{q}_s, \tilde{q}_{s'}) = \int_S \tilde{Q}(s, u) \tilde{Q}(s', u) du$$

and $w_i = (\mathcal{K}^* y)(s_i)$. Note that if $x \in \text{span}\{\tilde{q}_s : s \in S_n\}$ and $w = \mathcal{K}^* \mathcal{K} x$, then $x = x_n$. Let

$$\Delta_n := \sup_{s \in S} \inf_{s_i \in S_n} |s - s_i|,$$

and let \mathcal{O} be the Hilbert-Schmidt operator with kernel $P(s, s')$.

THEOREM 4.1. (a) *If $P(s, s')$ is continuous on $S \times S$ and $y \in \mathcal{R}(\mathcal{K}) \oplus \mathcal{R}(\mathcal{K})^\perp$, and x_n is given by (4.2), then*

$$(4.3) \quad \lim_{\Delta_n \rightarrow 0} \|x_n - \mathcal{K}^\dagger y\| = 0.$$

(b) *If $y \in \mathcal{R}(\mathcal{K}) \oplus \mathcal{R}(\mathcal{K})^\perp$, $\mathcal{K}^\dagger y \in \mathcal{K}^* \mathcal{K}(\mathcal{L}_2(S))$, or equivalently $w \in \mathcal{O}(\mathcal{L}_2(S))$ and $P(s, s')$ satisfies (i) and (ii) of Theorem 3.1 for some m , then*

$$(4.4) \quad \|x_n - \mathcal{K}^\dagger y\| = O(\Delta_n^m).$$

(c) *If $y \notin \mathcal{R}(\mathcal{K}) \oplus \mathcal{R}(\mathcal{K})^\perp$, then $\|x_n\| \rightarrow \infty$ as $\Delta_n \rightarrow 0$.*

Proof. This is a restatement of Theorem 3.1 when Q is replaced by P , $y \in \mathcal{H}_Q$ is replaced by $w \in \mathcal{H}_P$, where \mathcal{H}_P is the RKHS of functions on S with reproducing kernel P . Note that $w = \mathcal{K}^* y$ is in the space \mathcal{H}_P (which is the same as $\mathcal{R}(\mathcal{K}^* \mathcal{K})$) if and only if $y \in \mathcal{D}(\mathcal{K}^\dagger)$.

5. **Generalizations to Linear Operator Equations.** We note that some of the results of Sections 3 and 4 hold when $\mathcal{L}_2(S)$ is replaced by a general Hilbert space.

Let X be an arbitrary Hilbert space, with inner product $\langle \cdot, \cdot \rangle_X$, let \mathcal{Q} be a linear operator mapping X into the real-valued functions on T with the property that

$$(5.1) \quad |(\mathcal{Q}x)(t)| \leq M_t \|x\|_X, \quad t \in T, \quad x \in X.$$

Then there exists a family $\{a_t : t \in T\}$ of elements in X with the property that

$$(\mathcal{Q}x)(t) = \langle a_t, x \rangle_X, \quad t \in T, \quad x \in X.$$

Let $Q(t, t') = \langle a_t, a_{t'} \rangle_X$, $t, t' \in T$. Then Q is a nonnegative definite symmetric kernel on $T \times T$ and by the same argument as before, it is easy to see that $\mathcal{R}(X) = \mathcal{R}\mathcal{C}_Q$, when $\mathcal{R}\mathcal{C}_Q$ is the RKHS with RK Q . For $y \in \mathcal{R}\mathcal{C}_Q$, let x_n be that element in X of minimal norm which satisfies

$$(\mathcal{Q}x_n)(t) = y(t), \quad t \in T_n.$$

Then

$$x_n = (y_1, y_2, \dots, y_n)Q_n^{-1}(a_{t_1}, a_{t_2}, \dots, a_{t_n})'$$

where we are assuming that the set $\{a_t : t \in T_n\}$ is linearly independent; and Q_n is the $n \times n$ matrix with ij th entry given by $Q(t_i, t_j) = \langle a_{t_i}, a_{t_j} \rangle$. Letting $\hat{x} = \mathcal{Q}^\dagger y$ be the unique element of minimal X -norm for which $\mathcal{Q}x = y$, we easily see that Theorem 3.1 then holds with no change, except that (3.5) and (3.6) become, respectively,

$$\lim_{\Delta_n \rightarrow 0} \|x_n - \mathcal{Q}^\dagger y\|_X = 0, \quad \|x_n - \mathcal{Q}^\dagger y\|_X = O(\Delta_n^m).$$

If X possesses an RK, then the norm convergence in X implies pointwise convergence. For suitable choice of the RK, a large class of operators have the property (5.1) including differential, integral, and integro-differential operators (see [14], [15]).

Suppose now that X and Y are Hilbert spaces of real-valued functions on S and T , respectively. Let \mathcal{Q} be a linear operator on X into Y . Let \mathcal{Q}^\dagger be the generalized inverse of \mathcal{Q} defined on $\mathcal{R}(\mathcal{Q}) \oplus \mathcal{R}(\mathcal{Q})^\perp$.

If $\mathcal{Q}^* \mathcal{Q}x$ has the property that

$$|(\mathcal{Q}^* \mathcal{Q}x)(s)| \leq M_s \|x\|_X, \quad s \in S, \quad x \in X,$$

then there exists a family $\{\tilde{q}_s : s \in S\}$ in X with the property

$$(\mathcal{Q}^* \mathcal{Q}x)(s) = \langle \tilde{q}_s, x \rangle_X, \quad s \in S, \quad x \in X.$$

Let x_n be the element of minimal X -norm which satisfies

$$(\mathcal{Q}^* \mathcal{Q}x_n)(s) = w(s), \quad s = S_n,$$

where $w(s) = (\mathcal{Q}^* y)(s)$. Then x_n is given by

$$x_n = (w_1, w_2, \dots, w_n)P_n^{-1}(\tilde{q}_{s_1}, \tilde{q}_{s_2}, \dots, \tilde{q}_{s_n})'$$

where P_n is the $n \times n$ matrix with ij th element given by $P(s_i, s_j) = \langle \tilde{q}_{s_i}, \tilde{q}_{s_j} \rangle_X$. Theorem 4.1 then holds for this case with (4.3) and (4.4) replaced by

$$\lim_{\Delta_n \rightarrow 0} \|x_n - \mathcal{Q}^\dagger y\|_X = 0 \quad \text{and} \quad \|x_n - \mathcal{Q}^\dagger y\|_X = O(\Delta_n^m),$$

respectively.

6. Practical Considerations. Some comments on part (c) of Theorem 3.1 are in order. Both conceptually and practically, y can fail to be “in” $\mathcal{R}(\mathcal{K})$ in several ways. The notion of a generalized inverse $\mathcal{K}^\dagger y$ for $y \in \mathcal{R}(\mathcal{K}) \oplus \mathcal{R}(\mathcal{K})^\perp$ was considered in Section 4. In [7], the following situation is considered: It is assumed that $y = \mathcal{K}x + \xi$, where x is in an RKHS \mathcal{H}_R of real-valued functions on S , and $\xi \in \mathcal{H}_P$, where \mathcal{H}_P is an RKHS of real-valued functions on T which is not contained in $\mathcal{H}_Q \equiv \mathcal{K}(\mathcal{H}_R)$. The “desired” solution \bar{x} is taken as the (unique) solution to the minimization problem: Find $x \in \mathcal{H}_R$ to $\min J_y(x)$ where

$$J_y(x) = \|y - \mathcal{K}x\|_P^2 + \|x\|_R^2,$$

where $\|\cdot\|_P, \|\cdot\|_R$ are the norms in \mathcal{H}_P and \mathcal{H}_R and $\|g\|_P^2$ is interpreted as ∞ if $g \notin \mathcal{H}_P$. Approximations \bar{x}_n to \bar{x} are found and convergence rates obtained for the case $(Q + P)(t, t')$ is a smooth kernel.

A practical case is the situation where

$$(6.1) \quad y(t_i) = (\mathcal{K}x)(t_i) + \epsilon(t_i), \quad i = 1, 2, \dots, n,$$

where $\{\epsilon(t_i)\}_{i=1}^n$ are realizations of some “noise” random variables $\{\epsilon(t_i)\}_{i=1}^n$. This is an appropriate description if $(\mathcal{K}x)(t_i)$ is measured with some experimental error. Assume that $E\epsilon(t_i) = 0, i = 1, 2, \dots, n, E\epsilon(t_i)\epsilon(t_j) = 0, t_i \neq t_j$, and $E\epsilon^2(t_i) = \sigma^2$, where E is mathematical expectation. In this case, as n becomes large, the vector $(y(t_1), y(t_2), \dots, y(t_n))$ begins to look less and less like the restriction of an element of \mathcal{H}_Q . This model also describes roundoff error, if $(\mathcal{K}x)(t_{i+1}) - (\mathcal{K}x)(t_i)$ tends to be larger than, say, 3σ , and if further the roundoff contribution to the overall error is not negligible compared to the “projection error” given by (3.6). The situation (6.1) is considered in [16]. In [16] it is assumed that $x \in \mathcal{H}_R$, an RKHS of real-valued functions on S , and the approximate solution $x_{n,\lambda}$ is taken to be the solution to the problem: Find $x \in \mathcal{H}_R$ to $\min J_y^{(n)}(x)$ where

$$(6.2) \quad J_y^{(n)}(x) = \frac{1}{n} \sum_{i=1}^n (y(t_i) - (\mathcal{K}x)(t_i))^2 + \lambda \|x\|_R^2,$$

where λ is to be found.

λ is chosen to minimize a bound on $E\|x_{n,\lambda} - \mathcal{K}^\dagger g\|_R^2$. λ depends on n , and σ^2 as well as (generally unknown) properties of $g \equiv \mathcal{K}x$. If g is known to be very smooth, σ^2 is fixed, but λ is allowed to vary appropriately, then convergence will obtain. Rates are given. These results apply to the situation of Section 3 if we replace $\|x\|_R$ in (6.2) by $\|x\|$. Then the solution $x_{n,\lambda}$ to: Find $x \in \mathcal{L}_2$ to $\min J_y^{(n)}(x)$ where

$$J_y^{(n)}(x) = \frac{1}{n} \sum_{i=1}^n (y(t_i) - (\mathcal{K}x)(t_i))^2 + \lambda \|x\|^2$$

is

$$x_{n,\lambda}(\cdot) = (y_1, y_2, \dots, y_n)(Q_n + n\lambda I)^{-1}(k_{t_1}, k_{t_2}, \dots, k_{t_n}).$$

The expected mean square error $E\|x_{n,\lambda} - \mathcal{K}^\dagger g\|^2$ is of larger order than the error of (3.6). We must interpret the present results to apply to the situation where the contribution of roundoff and/or experimental errors are controlled so as to be small compared to the contribution of the projection error.

School of Mathematics
Georgia Institute of Technology
Atlanta, Georgia 30332

Department of Statistics
The University of Wisconsin
Madison, Wisconsin 53706

1. N. ARONSZAJN, "Theory of reproducing kernels," *Trans. Amer. Math. Soc.*, v. 68, 1950, pp. 337–404. MR 14, 479.
2. J. B. DIAZ & F. T. METCALF, "On iteration procedures for equations of the first kind, $Ax = y$, and Picard's criterion for the existence of a solution," *Math. Comp.*, v. 24, 1970, pp. 923–935. MR 43 #7094.
3. R. J. HANSON, "A numerical method for solving Fredholm integral equations of the first kind using singular values," *SIAM J. Numer. Anal.*, v. 8, 1971, pp. 616–622, MR 45 #2943.
4. E. HELLINGER & O. TOEPLITZ, "Integralgleichungen und Gleichungen mit unendlich vielen Unbekannten," *Enzyklopädie der Mathematischen Wissenschaften II C13*, 1928, p. 1349.
5. W. J. KAMMERER & M. Z. NASHED, "Iterative methods for best approximate solutions of linear integral equations of the first and second kinds," MRC Technical Summary Report #1117, Mathematics Research Center, The University of Wisconsin, 1971; *J. Math. Anal. Appl.*, v. 40, 1972, pp. 547–573.
6. M. Z. NASHED, "Generalized inverses, normal solvability, and iteration for singular operator equations," in *Nonlinear Functional Analysis and Applications* (Proc. Advanced Semi., Math. Res. Center, Univ. of Wisconsin, Madison, Wis., 1970), L. B. Rall (editor), Academic Press, New York, 1971, pp. 311–359. MR 43 #1003.
7. M. Z. NASHED & G. WAHBA, *Approximate Regularized Solutions to Linear Operator Equations When the Data-Vector is Not in the Range of the Operator*, MRC Technical Summary Report #1265, Mathematics Research Center, The University of Wisconsin, 1973.
8. B. NOBLE, *A Bibliography on: "Methods for Solving Integral Equations"*—Subject Listing, MRC Technical Summary Report #1177, Mathematics Research Center, The University of Wisconsin, 1971.
9. W. V. PETRYSHYN, "On generalized inverses and uniform convergence of $(I - \beta K)^n$ with applications to iterative methods," *J. Math. Anal. Appl.*, v. 18, 1967, pp. 417–439. MR 34 #8191.
10. H. L. SHAPIRO, *Topics in Approximation Theory*, Lecture Notes in Math., vol. 187, Springer-Verlag, Berlin, 1970.
11. D. W. SHOWALTER & A. BEN-ISRAEL, "Representation and computation of the generalized inverse of a bounded linear operator between two Hilbert spaces," *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur.*, (8), v. 48, 1970, pp. 184–194. MR 42 #8310.
12. F. SMITHIES, *Integral Equations*, Cambridge Tracts in Math and Math. Phys., no. 49, Cambridge Univ. Press, New York, 1958. MR 21 #3738.
13. O. N. STRAND, *Theory and Methods for Operator Equations of the First Kind*, Ph.D. Thesis, Colorado State University, Fort Collins, Colorado, 1972, 86 pp.
14. G. WAHBA, "Convergence rates for certain approximate solutions to Fredholm integral equations of the first kind," *J. Approximation Theory*, v. 7, 1973, pp. 167–185.
15. G. WAHBA, "A class of approximate solutions to linear operator equations," *J. Approximation Theory* (To appear.)
16. G. WAHBA, *Convergence Properties of the Method of Regularization for Noisy Linear Operator Equations*, MRC Technical Summary Report #1132, Mathematics Research Center, The University of Wisconsin, 1973.