ERROR BOUNDS FOR ITERATIVE SOLUTIONS OF
FREDHOLM INTEGRAL EQUATION

Louis Baker Rall
ERROR BOUNDS FOR ITERATIVE SOLUTIONS OF FREDHOLM INTEGRAL EQUATIONS

L. B. RALL

1. Introduction. A number of iterative procedures for obtaining the solution \( x(s) \) of the integral equation of Fredholm type and second kind,

\[
y(s) = x(s) - \lambda \int_a^b K(s, t)x(t)dt,
\]

have been developed, notably by G. Wiarda [10, pp. 119–128], Hans Bückner [2, pp. 68–71], Carl Wagner [8], and P.A. Samuelson [7]. These methods are generalizations of the one due to Neumann [3, pp. 119–120] in the sense that they converge where the Neumann process fails, or else offer the possibility of more rapid convergence. The purpose of this paper is to obtain estimates for the error resulting from the use of a finite number of steps of these iterative processes in forms suitable for numerical computation.

The author wishes to thank Professor A.T. Lonseth for many enlightening discussions concerning the material presented here, and the Reviewer for his helpful remarks.

2. The solution of linear equations. Methods for the approximate solution of Fredholm integral equations such as (1.1) and error estimates for these methods may be obtained directly from known results concerning the solution of linear equations in certain abstract spaces; it will be convenient to summarize some of these results here.

A set \( X = \{x\} \) of elements is called a linear space if \( x \in X \) implies \( (\theta x) \in X \), where \( \theta \) is any real number, and a binary operation \( + \) is defined in \( X \), with respect to which \( X \) is an Abelian group. The identity element of \( X \) for the operation \( + \) will be denoted by \( 0 \). In order to discuss convergence and error estimation, with each \( x \in X \) associate a finite, non-negative real number \( \| x \| \), called the norm of \( x \), which satisfies the following conditions:

1°. \( \| x \| > 0 \) if \( x \neq 0 \), \( \| 0 \| = 0 \);

2°. \( \| \theta x \| = |\theta| \| x \| \) for any real number \( \theta \);

3°. \( \| x + y \| \leq \| x \| + \| y \| \) for all \( x, y \in X \).

The space $X$ is now said to be a \textit{normed} linear space, and all spaces considered subsequently will be of this type.

A sequence $\{x_n\}$ in $X$ is said to \textit{converge} to the element $x \in X$, in symbols, $x_n \to x$ as $n \to \infty$, if $\|x-x_n\| \to 0$ as $n \to \infty$. A normed linear space $X$ is called \textit{complete} if for every sequence $\{x_n\}$ in $X$ such that $\|x_n-x_{n+p}\| \to 0$ as $n \to \infty$ for all positive integers $p$, there exists an $x \in X$ such that $x_n \to x$ as $n \to \infty$.

A transformation which carries each $x \in X$ into a $y \in X$ is symbolized by $Tx=y$, where $T$ is called an \textit{operator} in $X$. $T$ is \textit{additive} if $T(x+y) = Tx + Ty$ for all $x, y \in X$, and \textit{continuous} if $x_n \to x$ as $n \to \infty$ implies that $Tx_n \to Tx$ as $n \to \infty$. An additive and continuous $T$ is said to be \textit{linear}; for such a $T$, the nonnegative real numbers

\begin{align}
(2.1) \quad M(T) &= \text{l.u.b. } \langle \|Tx\|/\|x\| \rangle, \quad \|x\| 
eq 0, \\
(2.2) \quad m(T) &= \text{g.l.b. } \langle \|Tx\|/\|x\| \rangle, \quad \|x\| 
eq 0,
\end{align}

exist and are finite [1, p. 54]. A linear $T$ is \textit{homogeneous}, that is $T(\theta x) = \theta (Tx)$ for any real $\theta$ [1, p. 36]. The \textit{sum} $T+U$ and \textit{product} $TU$ of two linear operators $T$ and $U$ in $X$ are defined respectively by the relations $(T+U)x = Tx + Ux$ and $(TU)x = T(Ux)$ for all $x \in X$. Furthermore,

\begin{align}
(2.3) \quad M(T+U) &\leq M(T) + M(U), \\
(2.4) \quad M(TU) &\leq M(T)M(U),
\end{align}

[6]. The operator $I$ such that $Ix = x$ for all $x \in X$ is defined to be the \textit{identity operator} in $X$. The \textit{nth power} $T^n$ of an operator $T$ in $X$ is defined by $T^n = TT^{n-1}$ for all positive integers $n$, with $T^0 = I$ by definition. The \textit{inverse} of an operator $T$ in $X$ is the operator $T^{-1}$ such that $T^{-1}T = TT^{-1} = I$ if such exists. If $T$ is linear and $T^{-1}$ exists, $T^{-1}$ is likewise linear; moreover, if $m(T) > 0$,

\begin{equation}
(2.5) \quad m(T)M(T^{-1}) = 1
\end{equation}

[6]. If $T$ is a linear operator in a complete space $X$ and $M(T) < 1$,

\begin{equation}
(2.6) \quad (I-T)^{-1} = \sum_{j=0}^{\infty} T^j,
\end{equation}

[6, 9]. This result in combination with (2.3, 4, 5) gives

\begin{equation}
(2.7) \quad 1 - M(T) \leq m(I-T) \leq M(I-T) \leq 1 + M(T)
\end{equation}

for $M(T) < 1$. An operator $T$ in a normed linear space $X$ (not necessarily complete) is called \textit{completely continuous} if for every bounded set $B = \{x: \|x\| \leq \theta\}$ for $\theta$ finite, in the set $TB = \{Tx: x \in B\}$ every infinite
sequence converges to an element of \( X \). In a general normed linear space \( X \), (2.6) and (2.7) hold with the additional assumption that \( T \) is completely continuous [6]. These results furnish the following theorems:

**Theorem 1.** If \( F \) is a given linear operator in a complete normed linear space \( X \), then the linear equation

\[
Fz = y
\]

has a unique solution \( z \in X \) for every \( y \in X \) if and only if there exists a linear operator \( P \) in \( X \) such that \( P^{-1} \) exists, and

\[
M(I-PF)<1.
\]

The solution \( z \) of (1) in this case is given by

\[
z = \sum_{j=0}^{\infty} (I-PF)^j Py.
\]

**Proof:** To prove the sufficiency of Theorem 1, assume that a linear operator \( P \) having the desired properties exists. The series thus converges to an element, say \( z \), of \( X \); furthermore, \((I-PF)z = z - Py\), so \( Pfz = Py \). The application of \( P^{-1} \) yields \( Fz = y \), and thus \( z \) satisfies (1). If \( Fz_1 = y \) and \( Fz_2 = y \), then \( F(z_1 - z_2) = 0 \), so that \((I-PF)(z_1 - z_2) = z_1 - z_2 \), and if \( z_1 \neq z_2 \), \( M(I-PF) \geq 1 \), contrary to assumption; hence \( x = z \) is the unique solution of (1), and is given by (3). The necessity of Theorem 1 results from the fact that if there is a unique solution \( x \) of (1) for every \( y \in X \), \( F^{-1} \) exists. Taking \( P = F^{-1} \), \( P^{-1} \) exists and \( M(I-PF) = M(I-F^{-1}F) = M(0) = 0 < 1 \), which completes the proof of the theorem.

**Corollary 1.** Subject to the conditions of Theorem 1, \( F \) has the unique inverse

\[
F^{-1} = \sum_{j=0}^{\infty} (I-PF)^j P.
\]

These results hold in a general normed linear space \( X \) subject only to the additional condition that \((I-PF)\) be completely continuous.

**Theorem 2.** If a unique solution \( x \in X \) of (1) exists for every \( y \in X \), \( X \) a normed linear space, and an operator \( P \) on \( X \) exists such that
(2) is satisfied, then the iterative process

\[ x_n = (I - PF)x_{n-1} + Py \]

is totally convergent (Bückner) to the solution \( x \) of (1), that is, for all \( x_0 \in X \), \( x_n \to x \) as \( n \to \infty \), and its error is bounded by

\[ \| x - x_n \| \leq [M(I - PF)]^n \| x - x_0 \| \]

and

\[ \| x - x_n \| \leq \frac{M(I - PF)}{m(PF)} \| x_n - x_{n-1} \|. \]

**Proof.** Following [9], note that, from (1) and (5),

\[ x - x_n = (I - PF)^n(x - x_0), \]

from which (6) follows at once from (2.4). Condition (2) evidently insures that \( x_n \to x \) as \( n \to \infty \), whatever \( x_0 \). From (5),

\[ x_n - x_{n-1} = PF(x - x_{n-1}), \]

and

\[ x - x_n = (I - PF)(x - x_{n-1}), \]

from which (7) is obtained by (2.1) and (2.2). Condition (2) insures that \( m(PF) > 0 \), for, if \( PFz = 0 \) for any \( z \neq 0 \), then \( (I - PF)z = z \), and thus \( M(I - PF) \geq 1 \), contrary to assumption.

For the purposes of practical computation, it may prove expedient to calculate only one of the bounds \( M(I - PF) \), \( m(PF) \). By (2.7), the quantities \( m(PF) \) and \( 1 - M(I - PF) \) may be interchanged in (6) and (7); in what follows, the symbol \( \mu \) will be used to denote either of these quantities. These results have been obtained on the assumption that all operations have been carried out exactly, which is frequently not possible in practice. Set \( z_0 = x_0 \), and let \( z_n \) denote the results obtained from (5) by the use of some method of approximate evaluation. If \( \Delta_j \) is the difference of the exact and the approximate evaluation of \( (I - PF)z_{j-1} + Py \), from (5),

\[ x - z_n = x - x_0 + \sum_{j=0}^{n-1} (I - PF)^j \Delta_{n-j}. \]

Thus,

\[ \| x - z_n \| \leq \| x - x_n \| + \sum_{j=0}^{n-1} (1 - \mu)^j \| \Delta_{n-j} \|, \]

and as \( 0 < 1 - \mu < 1 \), for \( \delta = \max \| \Delta_{n-j} \|, (j = 0, \cdots, n-1) \),

\[ \| x - z_n \| \leq \| x - x_n \| + \delta/\mu, \]
where the estimate for \( \| x - x_n \| \) is obtained from the error bounds previously derived.

3. Application to integral equations. The space \( C \) of functions \( x = x(s) \) which are real, single-valued, and continuous on the interval \( a \leq s \leq b \) is an example of a linear space. For the purpose of error estimation, useful definitions of the norm of an element \( x \in C \) are:

\[
\begin{align*}
(i) & \quad \| x \| = \max_{[a, b]} |x(s)|, & (iii) & \quad \| x \| = \left[ \int_a^b x^2(s) \, ds \right]^{1/2}, \\
(ii) & \quad \| x \| = \int_a^b |x(s)| \, ds, & (iv) & \quad \| x \| = \left[ \int_a^b |x(s)|^\rho \, ds \right]^{1/\rho}, \quad \rho \geq 1;
\end{align*}
\]

all of these definitions are obtainable from (iv), (i) being the limit of (iv) as \( \rho \to \infty \), [5, pp. 134–150]. The inner product \( (x, y) \) of two elements \( x, y \in C \) is the real number

\[
(3.1) \quad (x, y) = \int_a^b x(s)y(s) \, ds.
\]

An operator \( Q \) in \( C \) is said to be positive definite if \( (Qx, x) > 0 \) for all \( x \neq 0 \) in \( C \), and to be positive semi-definite if \( (Qx, x) \geq 0 \) for all \( x \in C \). If \( Q \) is positive definite and \( M(Q) < 1 \), then \( M(I - Q) < 1 \) [11, p. 213], a fact which will be useful in establishing the convergence of iterative processes of the form (5). If \( K(s, t) \) is real, single-valued, and continuous on the square \( a \leq s, t \leq b \), the integral transform \( K \) defined by

\[
(3.2) \quad Kx = \int_a^b K(s, t) x(t) \, dt
\]

is a completely continuous linear operator in \( C \), so that the results of § 2 apply at once to the equation (1.1) with \( F = (I - \lambda K) \). A number \( \lambda \) is called a characteristic value of an integral transform \( K \) if \( m(I - \lambda K) = 0 \); Fredholm’s general theorem [4] states that (1.1) has a unique solution \( x(s) \) in \( C \) for every \( y(s) \) in \( C \) provided that \( \lambda \) is not a characteristic value of \( K \). If \( \lambda \) is a characteristic value of \( K \), it follows at once from Theorem 1 that (1.1) cannot have a unique solution, and thus it will be assumed throughout that \( \lambda \) is not a characteristic value of \( K \), unless the contrary is explicitly stated.

The error bounds (6) and (7) for the iterative method (5) as applied to (1.1) may be put in the following convenient forms:

\[
\begin{align*}
(E1) & \quad \| x - x_n \| \leq (1 - \mu) \| x - x_0 \|; \\
(E2) & \quad \| x - x_n \| \leq \frac{1 - \mu}{\mu} \| x_n - x_{n-1} \|;
\end{align*}
\]
for \( k \) a nonnegative integer,

\[
\| x - x_{n+k} \| \leq \frac{(1 - \mu)^{k+1}}{\mu} \| x_n - x_{n-1} \| ;
\]

while for \( x_0 = y \),

\[
\| x - x_n \| \leq \frac{(1 - \mu)^n}{\mu} \| P(\lambda K) y \| .
\]

As before, \( \mu = m[P(I - \lambda K)] \) or \( \mu = 1 - M[I - P(I - \lambda K)] \). These bounds depend on the values of \( \mu \) and \( M(P) \). The operator \( P \) will now be specified to obtain several iterative methods of practical importance, for which explicit bounds for \( \mu \) and \( M(P) \) will be calculated.

**Method I (Neumann):**

\[
x_n = y + \lambda K x_{n-1} .
\]

This process is (5) with \( P = I \), and thus \( (I - PF) = \lambda K \). It follows from Theorem 2 that (3.3) is totally convergent provided that \( M(\lambda K) < 1 \). If this is the case, explicit error estimates are obtained from the general expressions by setting \( \mu = 1 - M(\lambda K) \) and noting that \( M(P) = M(I) = 1 \). Usually \( M(\lambda K) \) is not known exactly, but estimates for \( M(\lambda K) \) are obtainable for various definitions of \( \| x \| \) from known inequalities [5, loc. cit.; 6; 9].

**Method II (Wiarda):**

\[
x_n = (1 - \theta)x_{n-1} + \theta \lambda K x_{n-1} + \theta y , \quad 0 < \theta < 1 .
\]

This method is (5) with \( P = \theta I \). Sufficient conditions for (3.4) to be totally convergent are that \(-\lambda K\) is positive semi-definite and

\[
0 < \theta < 1 + M(\lambda K) .
\]

These conditions insure that \( PF = \theta(I - \lambda K) \) is positive definite and that \( M(PF) < 1 \); the total convergence of Method II is a consequence of Theorem 2 in this case. As \(-\lambda K\) is positive semi-definite, \( m(PF) = m[\theta(I - \lambda K)] \geq \theta \), and as \( 0 < \theta < 1 \), explicit error bounds for Method II may be obtained from the general expressions by the substitution \( \mu = \theta \), and noting that \( M(P) = \theta \).

**Method III (Bückner):**

\[
x_n = (1 + \theta)v_{n-1} - \theta \lambda K v_{n-1} - \theta y ,
\]
ITERATIVE SOLUTIONS OF FREDHOLM INTEGRAL EQUATIONS

where

\[(3.7) \quad v_{n-1} = (1 - \theta)x_{n-1} + \theta \lambda K x_{n-1} + \theta y . \]

This process is totally convergent provided that \( \theta \) satisfies (3.5) and the kernel \( K(s, t) \) of \( K \) is symmetric, that is, \( K(s, t) = K(t, s), a \leq s, t \leq b. \) From (3.6) and (3.7),

\[(3.8) \quad x_n = x_{n-1} - \theta(I - \lambda K)^2 x_{n-1} + \theta(I - \lambda K) y . \]

This is (5) with \( P = \theta(I - \lambda K). \) If the kernel \( K(s, t) \) of \( K \) is symmetric, direct calculation from (3.1) verifies that

\[(3.9) \quad \langle [I - \lambda K] x, x \rangle = \langle [I - \lambda K] x, [I - \lambda K] x \rangle , \]

which is positive for all \( x \neq 0 \) in \( C \) as \( \lambda \) is not a characteristic value of \( K. \) Thus \( PF = \theta(I - \lambda K) \) is positive definite, and if \( \theta \) satisfies (3.5), \( M(PF) < 1. \) By Theorem 2, Method III is totally convergent. If \( \{\lambda_m\} \) denotes the set of characteristic values of \( K, \) for the norm defined by (iii),

\[(3.10) \quad \mu = m(PF) = \theta \cdot \min_{\langle x, x \rangle} [1 - \lambda/I - \lambda K]^2, \]

[2, pp. 10–11; 3, pp. 112–113]. This, together with the fact that \( M(P) < \theta, \) as \( M[\theta(I - \lambda K)] < 1 \) from (3.5), allows the explicit evaluation of the general error estimates for Method III for the norm (iii).

**Method IV (Wagner):**

\[(3.11) \quad x_n = x_{n-1} - (1/g)(I - \lambda K)x_{n-1} + (1/g)y , \]

where

\[(3.12) \quad g = g(s) = 1 - \lambda \int_a^b K(s, t) dt , \quad a \leq s \leq b ; \]

here it is assumed throughout that \( g(s) \equiv 0, a \leq s \leq b. \) If \( K(s, t) \) has a high maximum for \( s = t \) and is nearly zero elsewhere, then \( (I - \lambda K)x \approx gx \) for all \( x \in C. \) Define the function \( \phi(s; x) \) by

\[(3.13) \quad \phi(s; x) = (1/g)(I - \lambda K)x \]

for all \( x \in C. \) If

\[(3.14) \quad \omega = \max_{s \leq 0, x \in C} (1 - \phi(s; x)) < 1 , \]

then Method IV is totally convergent, as it is (5) with \( P = (1/g)I, \) and (3.14) gives \( M(I - PF) \leq \omega < 1. \) Explicit error bounds are obtained from
the general expressions by the substitution \( \mu = 1 - \omega \) and the fact that

\[
M(P) = M[(1/g)I] \leq [\min_{a \leq s \leq b} |g(s)|]^{-1}.
\]

For kernels of the type considered, it may be true that \( \omega \ll 1 \), in which case Method IV will converge rapidly.

**Method V (Samuelson):**

\[
x_n = x_{n-1} - (I + J)(I - \lambda K)x_{n-1} + (I + J)y
\]

is totally convergent, provided that

\[
M(G - J) \leq \frac{1}{1 + M(\lambda K)},
\]

where \( G \) is the resolvent operator for \( \lambda K \) which gives the solution \( x \) of (1.1) as

\[
x = (I + G)y,
\]

This follows at once from Theorem 2, as (3.16) is (5) with \( P = I + J \). Hence,

\[
PF = [(I + G) - (G - J)](I - \lambda K) = I - (G - J)(I - \lambda K),
\]

as \( I + G \) is the inverse of \( I - \lambda K \). Thus \( I - PF = (G - J)(I - \lambda K) \), and (3.17) insures that \( M(I - PF) < 1 \). Explicit error estimates for Method V are obtained by setting \( \mu = 1 - M(G - J)[1 + M(\lambda K)] \) and from \( M(P) = M(I + J) \leq 1 + M(J) \). In case that \( M(G - J) \) is very small, Method V converges rapidly.

### 4. Numerical example

To illustrate the application of some of the methods and error bounds given, an approximate solution of the integral equation

\[
s^2 = x(s) - \int_0^1 K(s, t)x(t) \, dt, \quad 0 \leq s \leq 1,
\]

where

\[
K(s, t) = \begin{cases} 
  s(1-t), & 0 \leq s \leq t \leq 1, \\
  t(1-s), & 0 \leq t \leq s \leq 1,
\end{cases}
\]

will be sought for various values of \( \lambda \). An approximation \( x_n(s) \) to \( x(s) \) will be considered to be satisfactory if \( \|x - x_n\| < 0.01 \) with the norm defined by (iii). The characteristic values of \( K \) are known to be \( \lambda_n \)
and $M(K) = 1/\pi^2$.

For $\lambda = -1$, $M(\lambda K) = 1/\pi^2 < 1$, and thus Method I will be used. For $y(s) = s^3$, as $\|s^2\| = 5^{-1/2}$, from (E4) the number of iterations required will not exceed one, so that

4.3) $x_1(s) = s^2 - s(1 - s^3)/12$

s a satisfactory approximation to $x(s)$ on $[0, 1]$.

For $\lambda = -10$, $M(\lambda K) = 10/\pi^2 > 1$, and the condition for the total convergence of Method I is not satisfied. However, $-\lambda K$ is positive definite, so that Method II is applicable. Take $x_0(s) = s^2$ and

4.4) $\theta = 0.49650 < 1/(1 + 10/\pi^2)$.

From (E4), the number of iterations will not exceed six. Successive iterations yield

4.5) $x_1(s) = s^2 - (0.41375)s(1 - s^3)$,

4.6) $x_2(s) = s^2 + (0.34238)s(1 - s^3) - (0.62207)s(1 - s^3) - (0.06848)s(1 - s^3)$,

with

4.7) $\|x_2 - x_1\| = 0.01140$,

and thus from (E3),

4.8) $\|x - x_1\| \leq 0.006$.

It follows that

4.9) $x_3(s) = s^2 + (0.46050)s(1 - s^3) - (0.72960)s(1 - s^3) + (0.08500)s(1 - s^3) - (0.13542)s(1 - s^3) - (0.00607)s(1 - s^3)$

is a satisfactory approximation to $x(s)$ on $[0, 1]$.

For $\lambda = 25$, $M(\lambda K) = 25/\pi^2 > 1$, so that Method I is not applicable. As $(-\lambda Ks, s) = -5/9$, $-\lambda K$ is not positive semi-definite, and Method II also fails. However, $K(s, t)$ is symmetric, and 25 is not a characteristic value of $K$, so Method III is totally convergent in this case. Choose

4.10) $\theta = 0.28394 < 1/(1 + 25/\pi^2)$

and $x_0(s) = s^2$. From (3.10),

4.11) $\mu = 0.01084$.

The upper bound for the number of iterations necessary is calculated from (E4) to be 727. The slowness of convergence in this case excludes manual methods of computation, but would be of little concern if a high-speed computing machine is available.
REFERENCES


OREGON STATE COLLEGE

*Added in proof:* Error bounds for Methods II and III are also contained in:

Nesmith Cornett Ankeny and Theodore Joseph Rivlin, *On a theorem of S. Bernstei* .................................................. 849
Louis Auslander, *The use of forms in variational calculation* ................................................................. 853
Paul Civin, *Abstract Riemann sum* .......................................................... 861
Paul Civin, *Some ergodic theorems involving two operator* ................................................................. 869
Eckford Cohen, *The number of solutions of certain cubic congruence* .................................................. 877
Richard M. Cohn, *Specializations over difference field* ................................................................. 887
Jean Dieudonné, *Pseudo-discriminant and Dickson invariant* .......................................................... 907
Ky Fan, *A comparison theorem for eigenvalues of normal matrix* .................................................. 911
Richard P. Gosselin, *On the convergence behaviour of trigonometric interpolating polynomial* .......................................................................................................................... 915
Peter K. Henrici, *On generating functions of the Jacobi polynomial* .................................................. 923
Meyer Jerison, *An algebra associated with a compact group* .......................................................... 933
Wilhelm Magnus, *Infinite determinants associated with Hill’s equation* .................................................. 941
G. Power and D. L. Scott-Hutton, *The slow steady motion of liquid past a semi-elliptical body* .......................................................................................................................... 953
Lyle E. Pursell, *An algebraic characterization of fixed ideals in certain function ring* .......................................................................................................................... 963
C. T. Rajagopal, *Additional note on some Tauberian theorems of O. Szás* .................................................. 971
Louis Baker Rall, *Error bounds for iterative solutions of Fredholm integral equation* .......................................................................................................................... 977
Shigeo Sasaki and Kentaro Yano, *Pseudo-analytic vectors on pseudo-Kählerian manifold* .......................................................................................................................... 987
Eugene Schenkman, *On the tower theorem for finite group* .......................................................... 995
P. Stein and John E. L. Peck, *On the numerical solution of Poisson’s equation over a rectangle* .......................................................................................................................... 999
Morgan Ward, *The mappings of the positive integers into themselves which preserve divisio* .......................................................................................................................... 1013
Seth Warner, *Weak locally multiplicatively-convex algebra* .................................................................. 1025
Louis Weisner, *Group-theoretic origin of certain generating function* .................................................. 1033