An iterative method for Fredholm equations of the first kind

William Carl Peterson
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd
Part of the Mathematics Commons

Recommended Citation
https://lib.dr.iastate.edu/rtd/5436
INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in “sectioning” the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from “photographs” if essential to the understanding of the dissertation. Silver prints of “photographs” may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106
PETERSON, William Carl, 1939-
AN ITERATIVE METHOD FOR FREDHOLM EQUATIONS
OF THE FIRST KIND.

Iowa State University, Ph.D., 1975
Mathematics

Xerox University Microfilms, Ann Arbor, Michigan 48106

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED.
An iterative method for Fredholm equations of the first kind

by

William Carl Peterson

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Department: Mathematics

Major: Applied Mathematics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa
1975
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>INTRODUCTION AND REVIEW OF THE LITERATURE</td>
<td>1</td>
</tr>
<tr>
<td>1.1.</td>
<td>Statement of the Problem</td>
<td>1</td>
</tr>
<tr>
<td>1.2.</td>
<td>Iterative Methods</td>
<td>5</td>
</tr>
<tr>
<td>2.</td>
<td>DEVELOPMENT OF AN ITERATIVE METHOD FOR CONVERGENCE TO THE LEAST SQUARES SOLUTION OF MINIMUM NORM</td>
<td>7</td>
</tr>
<tr>
<td>2.1.</td>
<td>Introduction</td>
<td>7</td>
</tr>
<tr>
<td>2.2.</td>
<td>An Iterative Method</td>
<td>7</td>
</tr>
<tr>
<td>2.3.</td>
<td>Perturbations of $g$</td>
<td>21</td>
</tr>
<tr>
<td>2.4.</td>
<td>Some Error Bounds and Properties of $\lambda$</td>
<td>24</td>
</tr>
<tr>
<td>3.</td>
<td>IMPLEMENTING THE ALGORITHM</td>
<td>40</td>
</tr>
<tr>
<td>3.1.</td>
<td>Calculation of a Starting Element</td>
<td>40</td>
</tr>
<tr>
<td>3.2.</td>
<td>The Sequence ${f_{j,n}}_{n=0}^{\infty}$</td>
<td>45</td>
</tr>
<tr>
<td>3.3.</td>
<td>A Flow Diagram of the Basic Algorithm</td>
<td>47</td>
</tr>
<tr>
<td>4.</td>
<td>EXAMPLES</td>
<td>54</td>
</tr>
<tr>
<td>4.1.</td>
<td>Matrix Examples</td>
<td>54</td>
</tr>
<tr>
<td>4.2.</td>
<td>An Inner Product Space</td>
<td>60</td>
</tr>
<tr>
<td>4.3.</td>
<td>Discretizing Integral Equations</td>
<td>62</td>
</tr>
<tr>
<td>4.4.</td>
<td>Error Bounds</td>
<td>67</td>
</tr>
<tr>
<td>4.5.</td>
<td>Perturbations of $g$</td>
<td>72</td>
</tr>
</tbody>
</table>
4.6. Integral Equation Examples  

5. SUMMARY AND FUTURE RESEARCH  
   5.1. Summary  
   5.2. Future Research  

6. BIBLIOGRAPHY  

7. ACKNOWLEDGMENTS  

8. APPENDIX  
   8.1. Appendix A, Linear System  
   8.2. Appendix B, Subroutine BITER  

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.6. Integral Equation Examples</td>
<td>75</td>
</tr>
<tr>
<td>5. SUMMARY AND FUTURE RESEARCH</td>
<td>85</td>
</tr>
<tr>
<td>5.1. Summary</td>
<td>85</td>
</tr>
<tr>
<td>5.2. Future Research</td>
<td>86</td>
</tr>
<tr>
<td>6. BIBLIOGRAPHY</td>
<td>87</td>
</tr>
<tr>
<td>7. ACKNOWLEDGMENTS</td>
<td>90</td>
</tr>
<tr>
<td>8. APPENDIX</td>
<td>91</td>
</tr>
<tr>
<td>8.1. Appendix A, Linear System</td>
<td>91</td>
</tr>
<tr>
<td>8.2. Appendix B, Subroutine BITER</td>
<td>101</td>
</tr>
</tbody>
</table>
1. INTRODUCTION AND REVIEW OF THE LITERATURE

1.1. Statement of the Problem

Throughout this paper we will assume that \( H_1 \) is a real separable (or finite dimensional) Hilbert space, \( H_2 \) is a real Hilbert space and \( K \) is a compact (or completely continuous) operator from \( H_1 \) to \( H_2 \). The basic definitions used above can be found in Bachman and Narici [3, Chapter 6 and 17].

We will use the notation:

\[
R(K) = \text{range of } K, \\
N(K) = \{ f \in H_1 : Kf = 0 \}
\]

and

\[
M^\perp = \{ f \in H_1 : \langle f, h \rangle = 0 \text{ for all } h \in M \subset H_1 \}
\]

where \( \langle , \rangle \) denotes inner product in the Hilbert space.

From Strand [20, page 22] we have that for an operator \( K \)

\[
H_2 = \overline{R(K^*)} \oplus \overline{R(K)} \text{ where } \overline{R(K)} \text{ is the closure of } R(K), \\
H_1 = N(K) \oplus (N(K))^\perp,
\]

and

\[
\overline{R(K^*)} = [N(K)]^\perp.
\]
We are interested in finding an approximation to the least squares solution of minimum norm, \( f_0 \), of the operator equation \( Kf = g \) where \( g \in H_2 \), when such a solution exists. Kammerer and Nashed [11] give the following definitions.

**Definition 1.1.1** For \( f \in H_1 \), \( g \in H_2 \), an element \( \bar{f} \in H_1 \) is a least squares solution of \( Kf = g \) if and only if

\[
\|K\bar{f} - g\| = \inf\{\|Kf - g\| \mid f \in H_1\}.
\]

**Definition 1.1.2** For \( f \in H_1 \), \( g \in H_2 \), an element \( f_0 \in H_1 \) is a least squares solution of minimum norm of \( Kf = g \) if and only if it is a least squares solution of \( Kf = g \) and \( \|f_0\| \leq \|\bar{f}\| \) for all least squares solutions \( \bar{f} \) of \( Kf = g \).

Strand [20, pages 23-24] notes that the least squares solution of minimum norm to \( Kf = g \) exists (is unique and in \( N(K)^\perp \)) if and only if \( g \in H_2 \) can be written in the form

\[
g = g_1 + g_2
\]
where $g_1 \in \text{R}(K)$ and $g_2 \in \text{N}(K^*)$. For this paper we will assume $g$ is always in the form $g = g_1 + g_2$, with $g_1 \in \text{R}(K)$. From Strand [20] and Diaz and Metcalf [5], $f$ is a solution of $\|Kf - g\| = \inf\{\|Kh - g\| \mid h \in H_1\}$ if and only if $f$ is a solution to $Kf = g_1$. They also note that the least square solutions to $K^*Kf = K^*g$ and $Kf = g_1$ are the same, this will become important later on in this paper.

The least squares solution of minimum norm of $Kf = g$ can be written as a series in terms of $g$ and the eigenvalues and eigenvectors of $KK^*$ and $K^*K$, so is known exactly ([20, page 25], [24, page 143]). However, finding each eigenvector of $KK^*$ is as difficult a problem, numerically, as the original problem.


\begin{equation}
Q(f) = \|Kf - g\|^2 + \lambda\|f - p\|^2, \lambda > 0,
\end{equation}
where \( p \in H_1 \) is an estimate for the least squares solution of minimum norm of \( Kf = g \). He proves that (1.1.4) is minimized by

\[
(1.1.5) \quad f_\lambda = (K^*K + \lambda I)^{-1}(K^*g + \lambda p).
\]

Strand [20] gives a proof for the following theorem.

**Theorem 1.1.1.** \( \bar{f} = \lim_{\lambda \to 0^+} f_\lambda \) exists and \( \bar{f} \) is the unique solution of \( K^*Kf = K^*g \) for which \( \|f - p\|^2 \) is a minimum.

We will use this theorem in Chapter 2 for the case where \( p = 0 \), to obtain an approximation to the least squares solution of minimum norm of \( Kf = g \). In general a method for determining a good choice for \( \lambda \) in actual applications and construction of an inverse to represent \( (K^*K + \lambda I)^{-1} \) is difficult since \( (K^*K + \lambda I)^{-1} \) is generally ill-conditioned for a useful choice for \( \lambda \). In this dissertation, the problem of a choice for \( \lambda \) will be considered and finding \( (K^*K + \lambda I)^{-1} \) will be avoided.
1.2. Iterative Methods

Landweber [13], Diaz and Metcalf [5], Kammerer and Nashed [10,11,12] and Nashed [14,15] study iterative methods for finding the least squares solution of minimum norm of $Kf = g$, by constructing a sequence $\{f_i\}_{i=1}^{\infty}$ which converges to the least squares solution of minimum norm. The sequence $\{f_i\}_{i=1}^{\infty}$ is defined by

\begin{equation}
 f_{n+1} = f_n + \alpha_n W_n
\end{equation}

where $\alpha_n$ and $W_n$ are chosen to give the method of steepest descent, conjugate gradient or weak steepest descent method. These methods converge to the least squares solution of minimum norm of $Kf = g$ for various choices for a starting vector and restrictions on $g \in H_2$. For perturbations of $g$ by a function $\epsilon$ in $H_2$ we have the operator equation $Kf = g + \epsilon$ which, since $g + \epsilon$ might not be an element of $R(K) + N(K^*)$, may fail to have a least squares solution of minimum norm. Strand [20, page 47] shows that these methods are sensitive to small perturbations of $g$ and may fail to converge to a function near
the actual solution. This is a serious problem for a numerical method, since discretization error and roundoff error act like a perturbation of $g$.

The iterative method considered in this dissertation depend on operators of the form $(\lambda I + K^*K) : H_1 \to H_1$, $\lambda > 0$. Since $\lambda I + K^*K$ is bounded and positive definite the inverse $(\lambda I + K^*K)^{-1}$ exists and is defined on all of $H_1$ for each $\lambda > 0$. For $\lambda$ not too small the solutions of $(K^*K + \lambda I)f = K^*g$ will not be as sensitive to perturbation of $g$ as the equation $K^*Kf = K^*g$ (see section 2.3), but has a solution close to the least squares solution of minimum norm of $K^*Kf = K^*g$ (or $Kf = g$).

Some examples will be given, together with a discussion of some of the problems of discretization, of finding approximations to least square solutions of minimum norm of integral equations of the first kind.
2. DEVELOPMENT OF AN ITERATIVE METHOD FOR CONVERGENCE TO THE LEAST SQUARES SOLUTION OF MINIMUM NORM

2.1. Introduction

First, an iterative method is obtained for an approximate solution to the least squares solution of minimum norm of $Kf = g$, when such a solution exists.

Second, several theorems will be given which study the properties of changes in the solution to $Kf = g$ when $g$ is replaced with $g + \epsilon$, \( \epsilon \) an element of $H^2$, to give an indication of stability.

Third, theorems relating $\lambda \succ 0$ and $f_\lambda \in H^1$, where $f_\lambda$ is a solution to $(K^*K + \lambda I)f = K^*g$ will be given, as well as some theorems on error bounds.

2.2. An Iterative Method

The least squares solution of minimum norm, $f_0'$, is an element of $[N(K)]^\perp$. For any element $f \in N(K)$ we have $K(f_0' + f) = Kf_0' + Kf = Kf_0'$, so since $f_0' + f$ and $f_0'$ are mapped to the same element in the range of $K$ it is important for an iterative procedure to stay on $[N(K)]^\perp$. The goal is to construct a sequence of $f_\lambda$'s.
which satisfy the theorem below, converge to $f_0$ and stays in $N(K)^\perp$.

**Theorem 2.2.1.** If $f_\lambda \in H$ is the solution to $(K*K + \lambda I)f = K*g$, $\lambda > 0$, then $f_\lambda \in R(K^*) \subset N(K)^\perp$.

**Proof:** Since $(K*K + \lambda I)f_\lambda = K*g$ we have

$$\lambda f_\lambda = K*g - K*K f_\lambda$$

which in turn implies that

$$f_\lambda = \frac{K^*(g - K f_\lambda)}{\lambda},$$

thus $f_\lambda \in R(K^*) \subset [N(K)]^\perp$.

Now we develop an iterative method related to a steepest descent method for Tikhonov-Twomey regularization, which converges to the least squares solution of minimum norm, when such a solution exists. The method developed allows one to find a solution of $(K*K + \lambda I)f = K*g$ for a $\lambda > 0$. When the convergence fails or becomes slow then the advantage of this method is that a new starting vector can be obtained.

We really wish to implement the procedure in Step 1 and 2 following and in doing so the main result we use is given in the corollary on page 20 for Theorem 2.2.2 which follows.
Step 1: Choose a sequence of positive real numbers 
\( \{\lambda_i\}_{i=1}^{\infty} \) such that \( \lambda_i \to 0 \) as \( i \to \infty \), a starting element \( f_{1,0} \in H_1 \) for the sequence \( \{f_{1,n}\}_{n=0}^{\infty} \) converging to the solution, \( f_1 \), of \( (K^*K + \lambda_1 I)f = K^*g \), as required by Theorem 2.2.2.

Step 2: we now proceed inductively. For \( j = 2, 3, 4, \ldots \)
using \( f_{j,0} = f_{j-1} \) as a starting vector, obtain the sequence \( \{f_{j,n}\}_{n=0}^{\infty} \) converging to \( f_j \), where \( f_j \) is the solution to \( (K^*K + \lambda_j I)f = K^*g \). Continue until \( \lambda_j \) is less than some pre-assigned value.

Theorem 2.2.2. Let \( f_{1,n+1} = f_{1,n} + a_{1,n} \, W_{1,n} \), \( a_{1,n} \in \mathbb{R} \), for each integer \( n \geq 0 \) and \( f_{1,0} \in H_1 \) then the sequence \( \{f_{1,n}\}_{n=0}^{\infty} \) converges to the function \( f = f_1 \in F(K^*) \) that minimizes 
\[ Q_{\lambda_1} (f) = \|Kf - g\|^2 + \lambda_1 \|f\|^2 \]
where

\[ W_{1,n} = K^*K f_{1,n} + \lambda f_{1,n} - K^*g, \]

\[ a_{1,n} = -\frac{\langle W_{1,n}, W_{1,n} \rangle}{\langle KW_{1,n}, Kw_{1,n} \rangle + \lambda_1 \langle W_{1,n}, W_{1,n} \rangle} \quad \text{for} \quad W_{1,n} \neq 0, \]
and

if \( n_1 \) is the first \( n \) such that \( W_{1,n} = 0 \)
take \( \alpha_{1,n} = 0 \) and \( f_1 = f_{n,1} \) for all \( n \geq n_1 \).

**Proof:** From Theorem 2.2.1 we have \( f_1 \in R(K^*) \). To choose
\( \alpha_{1,n} \) consider \( f_{1,n+1} = f_{1,n} + \alpha W_{1,n} \) and

\[
Q_{\lambda_1}(f_{1,n+1}) = \|K f_1,n+1 - g\|^2 + \lambda_1 \|f_{1,n+1}\|^2
\]

\[
= \langle K^* K f_{1,n+1}, f_{1,n+1} \rangle - 2 \langle K^* g, f_{1,n+1} \rangle + \langle g, g \rangle + \lambda_1 \langle f_{1,n+1}, f_{1,n+1} \rangle
\]

\hspace{1cm} \text{(2.2.1)}

\[
+ \langle g, g \rangle + \lambda_1 \langle f_{1,n+1}, f_{1,n+1} \rangle
\]

\[
= \langle K^* K f_{1,n} + \alpha K^* W_{1,n}, f_{1,n} + \alpha W_{1,n} \rangle
\]

\[
- 2 \langle K^* g, f_{1,n} + \alpha W_{1,n} \rangle + \langle g, g \rangle + \lambda_1 \langle f_{1,n} + \alpha W_{1,n}, f_{1,n} + \alpha W_{1,n} \rangle
\]

\hspace{1cm} \text{(2.2.2)}

\[
+ \lambda_1 \langle f_{1,n} + \alpha W_{1,n}, f_{1,n} + \alpha W_{1,n} \rangle
\]

\[
= \langle K^* K f_{1,n}, f_{1,n} \rangle + 2 \alpha \langle K^* K f_{1,n}, W_{1,n} \rangle + \langle g, g \rangle + \lambda_1 \langle f_{1,n}, f_{1,n} \rangle
\]

\hspace{1cm} \text{(2.2.2)}

\[
+ \alpha^2 \langle K^* K W_{1,n}, W_{1,n} \rangle - 2 \langle K^* g, f_{1,n} \rangle
\]
We now take the derivative of $Q_{\lambda_1}(f_{l,n+1})$ with respect to $\alpha$ and set the derivative equal to zero:

\[
\frac{d}{d\alpha} Q_{\lambda_1}(f_{l,n+1}) = 2\langle K^* K f_{l,n}, w_{l,n} \rangle + 2\alpha \langle K^* W_{l,n}, w_{l,n} \rangle - 2\langle K^* g, w_{l,n} \rangle + 2\lambda_1 \langle f_{l,n}, w_{l,n} \rangle + 2\lambda_1 \alpha \langle w_{l,n}, w_{l,n} \rangle = 0.
\]

Note also

\[
\frac{d^2}{d\alpha^2} Q_{\lambda_1}(f_{l,n+1}) = 2(\langle K^* W_{l,n}, w_{l,n} \rangle + \lambda_1 \langle w_{l,n}, w_{l,n} \rangle) > 0
\]

for $w_{l,n} \neq 0$. Solving $\frac{d}{d\alpha} Q_{\lambda_1}(f_{l,n+1}) = 0$ for $\alpha$ gives
\[\alpha\left(\langle K^*Kw_1,n,\delta_1,n \rangle + \lambda \langle \delta_1,n,\delta_1,n \rangle\right)\]

\[= \langle K^*g,\delta_1,n \rangle - \langle K^*f_1,n,\delta_1,n \rangle - \lambda \langle f_1,n,\delta_1,n \rangle\]

\[= - \langle K^*f_1,n + \lambda f_1,n - K^*g,\delta_1,n \rangle\]

where \(\delta_1,n \neq 0\), yielding the result:

\[(2.2.3) \quad \alpha_{1,n} = \alpha = -\frac{\langle K^*f_1,n + \lambda f_1,n - K^*g,\delta_1,n \rangle}{\langle K^*Kw_1,n,\delta_1,n \rangle + \lambda \langle \delta_1,n,\delta_1,n \rangle} .\]

Now we write \(f_{1,n+1}\) as

\[f_{1,n+1} = f_{1,n} - \frac{\langle K^*f_1,n + \lambda f_1,n - K^*g,\delta_1,n \rangle}{\langle K^*Kw_1,n,\delta_1,n \rangle + \lambda \langle \delta_1,n,\delta_1,n \rangle} \cdot \delta_1,n\]

for \(\delta_1,n \neq 0\).

Now we show that \(\{Q_1(f_{1,n})\}_{n=0}^{\infty}\) is a decreasing sequence of positive real numbers. From the expression (2.2.2), replacing \(n+1\) with \(n\) in (2.2.1) and \(\alpha\) with \(\alpha_{1,n}\) in both (2.2.2) and (2.2.1) we get the following expression.
\[ Q_{\lambda_1} (f_{l, n+1}) - Q_{\lambda_1} (f_{l, n}) = \langle K^* K f_{l, n}, f_{l, n} \rangle \]

\[ + 2 \alpha_{l, n} \langle K^* K^\dagger f_{l, n}, W_{l, n} \rangle + \alpha_{l, n}^2 \langle K^* K W_{l, n}, W_{l, n} \rangle \]

\[ - 2 \langle K^* g, f_{l, n} \rangle - 2 \alpha_{l, n} \langle K^* g, W_{l, n} \rangle + \langle g, g \rangle + \lambda_{l} \langle f_{l, n}, f_{l, n} \rangle \]

\[ + 2 \lambda_{l, n} \langle f_{l, n}, W_{l, n} \rangle + \lambda_{l, n}^2 \langle W_{l, n}, W_{l, n} \rangle \]

\[ = 2 \alpha_{l, n} \langle K^* K^\dagger f_{l, n}, W_{l, n} \rangle + \alpha_{l, n}^2 \langle K^* K W_{l, n}, W_{l, n} \rangle \]

\[ - 2 \alpha_{l, n} \langle K^* g, W_{l, n} \rangle + 2 \lambda_{l, n} \langle f_{l, n}, W_{l, n} \rangle \]

\[ + \lambda_{l, n}^2 \langle W_{l, n}, W_{l, n} \rangle \]

\[ = 2 \alpha_{l, n} \langle (K^* K f_{l, n} - K^* g, W_{l, n}) + \lambda_{l} \langle f_{l, n}, W_{l, n} \rangle \rangle \]

\[ + \alpha_{l, n}^2 \langle (K^* K W_{l, n}, W_{l, n}) + \lambda_{l} \langle W_{l, n}, W_{l, n} \rangle \rangle . \]
So we have from the expression for $\alpha_{1,n}$

$$Q_{\lambda_1}(f_{1,n+1}) - Q_{\lambda_1}(f_{1,n}) = -\frac{2\langle K^*Kf_{1,n} + \lambda_1 f_{1,n} - K^*g, W_{1,n} \rangle^2}{\langle K^*KW_{1,n}, W_{1,n} \rangle + \lambda_1 \langle W_{1,n}, W_{1,n} \rangle}$$

$$+ \frac{\langle K^*Kf_{1,n} + \lambda_1 f_{1,n} - K^*g, W_{1,n} \rangle^2}{\langle K^*KW_{1,n}, W_{1,n} \rangle + \lambda_1 \langle W_{1,n}, W_{1,n} \rangle}$$

$$= -\frac{\langle K^*Kf_{1,n} + \lambda_1 f_{1,n} - K^*g, W_{1,n} \rangle^2}{\langle K^*KW_{1,n}, W_{1,n} \rangle + \lambda_1 \langle W_{1,n}, W_{1,n} \rangle}.$$

Therefore, we have $Q_{\lambda_1}(f_{1,n+1}) \leq Q_{\lambda_1}(f_{1,n})$ and equality occurs when $\langle K^*Kf_{1,n} + \lambda_1 f_{1,n} - K^*g, W_{1,n} \rangle = 0$ for some $n$.

Note that if $K^*Kf_{1,n} + \lambda_1 f_{1,n} - K^*g = 0$ for some $n$, say $n_1$, then $f_{1,n_1} = (K^*K + \lambda_1 I)^{-1}K^*g$ from (1.1.5), where $p = 0$, is the minimizing value for $Q_{\lambda_1}(f)$. Thus if $K^*Kf_{1,n} + \lambda_1 f_{1,n} - K^*g \neq 0$ for $n > n_1$ and $K^*Kf_{1,n_1} + \lambda_1 f_{1,n_1} - K^*g = 0$ for $n = n_1$ take $\alpha_{1,n} = 0$ for $n > n_1$ and $f_{1,n} = f_{1,n_1}$ for $n \geq n_1$. In general take $W_{1,n} = K^*Kf_{1,n} + \lambda_1 f_{1,n} - K^*g$. We have discussed the
case \( w_{1,n} \neq 0 \) for all \( n \), then (2.2.1) becomes

\[
\alpha_{1,n} = -\frac{\langle w_{1,n}, w_{1,n} \rangle}{\langle K\ast kW_{1,n}, w_{1,n} \rangle + \lambda_1 \langle w_{1,n}, w_{1,n} \rangle}
\]

and

(2.2.4)

\[
Q_{\lambda_1}(f_{1,n+1}) - Q_{\lambda_1}(f_{1,n}) = -\frac{\langle w_{1,n}, w_{1,n} \rangle^2}{\langle K\ast kW_{1,n}, w_{1,n} \rangle + \lambda_1 \langle w_{1,n}, w_{1,n} \rangle}
\]

So \( Q_{\lambda_1}(f_{1,n+1}) < Q_{\lambda_1}(f_{1,n}) \) for all \( n \). We now have

\[
Q_{\lambda_1}(f_{1,n+1}) = Q_{\lambda_1}(f_{1,n}) - \frac{\langle w_{1,n}, w_{1,n} \rangle^2}{\langle K\ast kW_{1,n}, w_{1,n} \rangle + \lambda_1 \langle w_{1,n}, w_{1,n} \rangle}
\]

\[
= Q_{\lambda_1}(f_{1,0}) - \sum_{k=0}^{n} \frac{\langle w_{1,k}, w_{1,k} \rangle^2}{\langle K\ast kW_{1,k}, w_{1,k} \rangle + \lambda_1 \langle w_{1,k}, w_{1,k} \rangle}.
\]

So \( \left\{ \sum_{k=0}^{n} \frac{\langle w_{1,k}, w_{1,k} \rangle^2}{\langle K\ast kW_{1,k}, w_{1,k} \rangle + \lambda_1 \langle w_{1,k}, w_{1,k} \rangle} \right\}_{n=0}^{\infty} \) is an increasing
sequence of positive real numbers bounded above by 

\[ Q_{\lambda_1} \langle f_1, 0 \rangle \] and so converges. Now since

\[ \langle K^{*}KW_1, k', W_1, k \rangle = \langle KW_1, k', KW_1, k \rangle \]

\[ \leq \|K\|^2 \langle W_1, k', W_1, k \rangle \]

we have

\[ \langle K^{*}KW_1, k', W_1, k \rangle + \lambda_1 \langle W_1, k', W_1, k \rangle \leq (\|K\|^2 + \lambda_1) \langle W_1, k', W_1, k \rangle \]

and

\[ (2.2.5) \]

\[ \frac{1}{\|K\|^2 + \lambda_1} \langle W_1, k', W_1, k \rangle \leq \frac{1}{\langle K^{*}KW_1, k', W_1, k \rangle + \lambda_1 \langle W_1, k', W_1, k \rangle} . \]

Multiply each side of (2.2.5) by \( \langle W_1, k', W_1, k \rangle^2 \) and sum over \( k \), yielding
\[
\frac{1}{(\|K\|^2 + \lambda_1)} \sum_{k=0}^{\infty} \langle W_{1,k}, W_{1,k} \rangle \\
\leq \sum_{k=0}^{\infty} \frac{\langle W_{1,k}, W_{1,k} \rangle^2}{\langle (K^*KW_{1,k}) W_{1,k} \rangle + \lambda_1 \langle W_{1,k} W_{1,k} \rangle} < \infty.
\]

So \[\sum_{k=0}^{\infty} \langle W_{1,k}, W_{1,k} \rangle\] converges, thus

(2.2.6) \[
\langle W_{1,k}, W_{1,k} \rangle \to 0 \text{ as } k \to \infty.
\]

We now show \[\{f_{1,n}\}_{n=0}^{\infty}\] converges to

\[f_1 = (K^*K + \lambda_1 I)^{-1}K^*g.\]

Once this is done the proof will be complete, since this \(f_1\) minimizes \(Q_{\lambda_1}(f)\). Since

\((K^*K + \lambda_1 I)\) is invertible, there exists a positive real number \(M_{\lambda_1} = \|(K^*K + \lambda_1 I)^{-1}\|^{-1}\) such that for all \(n\)
Therefore we have

\[(2.2.7) \quad \|f_{1,n} - f_1\| \leq \frac{\|W_{1,n}\|}{M_{\lambda_1}} \to 0 \text{ as } n \to \infty\]

from (2.2.6).

So we have shown that \(f_{1,n}\) converges to \(f_1\) in norm as \(n \to \infty\). This completes the proof of the theorem.

Now take \(\lambda_2\) from the sequence \(\{\lambda_j\}_{j=1}^\infty\) and use \(f_1 \in R(K^*)\), from Theorem 2.2.1, as a starting element, i.e. take \(f_{2,0} = f_1\), and as in the case for \(\lambda_1\) construct the sequence \(\{f_{2,n}\}_{n=0}^\infty\) which converges to \(f_2 \in R(K^*)\) such that \(f_{2,n+1} = f_{2,n} + \alpha_{1,n}W_{2,n}\), where...
\[ \alpha_{2,n} = -\frac{\langle W_{2,n}, W_{2,n} \rangle}{\langle K^* W_{2,n}, W_{2,n} \rangle + \lambda_2 \langle W_{2,n}, W_{2,n} \rangle}, \]

and

\[ W_{2,n} = K^* f_{2,n} + \lambda_2 f_{2,n} - K^* g \]

for

\[ K^* f_{2,n} + \lambda_2 f_{2,n} - K^* g \neq 0 \]

and \( \alpha_{2,n} = 0 \) whenever

\[ K^* f_{2,n} + \lambda_2 f_{2,n} - K^* g = 0. \]

In general, use \( f_{j-1} \) as a starting element in \( R(K^*) \), take \( f_{j,0} = f_{j-1} \), \( j = 2, 3, 4, \ldots \) and construct the sequence

\[ (2.2.8) \quad \{f_{j,n}\}_{n=0}^{\infty} \quad \text{which converges to } f_j \]

such that \( f_{j,n+1} = f_{j,n} + \alpha_{j,n} W_{j,n} \).
where \( \alpha_{j,n} = - \frac{\langle K^*W_j,n, W_j,n \rangle}{\langle K^*Kf_j,n, W_j,n \rangle + \lambda_j \langle W_j,n, W_j,n \rangle} \),

\[
W_j,n = K^*Kf_j,n + \lambda_j f_j,n - K^*g
\]

for

\[
K^*Kf_j,n + \lambda_j f_j,n - K^*g \neq 0
\]

and \( \alpha_{j,n} = 0 \) whenever \( K^*Kf_j,n + \lambda_j f_j,n - K^*g = 0 \).

From the remarks at the beginning of section 2.2, we need that the sequence defined by (2.2.8) stay in \( [N(K)]^\perp \). Since for \( j > 1 \) and \( f_{j,0} = f_{j-1} \in R(K^*) \), we have the following corollary to Theorem 2.2.2.

**Corollary 2.2.3.** If \( j > 1 \) then the sequence defined by (2.2.8), \( \{f_{j,n}\}_{n=0}^{\infty} \subset R(K^*) \subset [N(K)]^\perp \).

**Theorem 2.2.4.** Let \( f_{0} \) be the least squares solution of minimum norm of \( Kf = g \), \( f_{1,0} \in H_1 \), \( \{\{f_{j,n}\}_{n=0}^{\infty}\}_{j=1}^{\infty} \) constructed as in (2.2.8) and let \( \epsilon > 0 \), then there exists positive integers \( n \) and \( j \) such that
\[ \| f_{j,n} - f_0 \| < \epsilon. \]

**Proof:** Follows from \[ \| f_{j,n} - f_0 \| \leq \| f_{j,n} - f_j \| + \| f_j - f_0 \|. \]

### 2.3. Perturbations of \( g \)

We have that \( R(K) + N(K^*) \subseteq H_2 \) is dense in \( H_2 \). For a perturbation of \( g \) by an element \( \epsilon \) belonging to \( H_2 \), it is possible that \( g + \epsilon \) will fail to be an element of \( R(K) + N(K^*) \), in which case the least squares solution of minimum norm if \( Kf = g + \epsilon \) would fail to exist. The goal is to determine relationships of the solutions of

\[
(K^*K + \lambda I)f = K^*g,
\]

\[
(K^*K + \lambda I)f = K^*(g + \epsilon)
\]

and the least squares solution of minimum norm of \( Kf = g \). Then an approximate solution of minimum norm of \( Kf = g \) can be found, provided the perturbation of \( g \) is not too large.

**Theorem 2.3.1.** If \( f_{\lambda} \in H_1 \) is the solution to \( (K^*K + \lambda I)f = K^*g \) and \( f_{\lambda, \epsilon} \in H_1 \) is the solution to \( (K^*K + \lambda I)f = K^*(g + \epsilon) \) where \( \epsilon \) belongs to \( H_2 \) then
\[ \| f_{\lambda, \varepsilon} - f_{\lambda} \| \leq \frac{\| K \|}{\| (K K + \lambda I)^{-1} \|^{-1}} \| \varepsilon \| \]

**Proof:** We have that

\[ (K K + \lambda I) f_{\lambda} = K g, \]
\[ (K K + \lambda I) f_{\lambda, \varepsilon} = K (g + \varepsilon). \]

Subtracting the above equations, we get

\[ (K K + \lambda I) f_{\lambda, \varepsilon} - (K K + \lambda I) f_{\lambda} = K (g + \varepsilon) - K g, \]
\[ (K K + \lambda I) (f_{\lambda, \varepsilon} - f_{\lambda}) = K \varepsilon. \]

Now since \( K K + \lambda I \) has a bounded inverse, there exists an \( M_{\lambda} = \| (K K + \lambda I)^{-1} \|^{-1} \) such that

\[ M_{\lambda} \| f_{\lambda, \varepsilon} - f_{\lambda} \| \leq \| (K K + \lambda I) (f_{\lambda, \varepsilon} - f_{\lambda}) \| \]
\[ = \| K \varepsilon \| \]
\[ \leq \| K \| \| \varepsilon \|. \]
Remark: With reference to Theorem 2.3.1, if we let $\mu$ be the minimum (or inf of the) non-negative eigenvalues of $K^*K$ then

\[ \| (K^*K + \lambda I)^{-1} \|^{-1} = \frac{\|K\|}{M_\lambda} \| \epsilon \|. \]

Theorem 2.3.2. If given $\delta > 0$ and

1. $f_0$ is the least squares solution of minimum norm of $Kf = g$,
2. $f_j, \epsilon$ is the solution of $(K^*K + \lambda I)f = K^*(g + \epsilon)$ for each $j$, $j = 1, 2, \ldots$,
3. $f_j, n, \epsilon$ defined by (2.2.8),
4. $\epsilon$ is an element of $H_2$ then

\[ \| f_{\tilde{j}, \tilde{n}, \epsilon} - f_0 \| \leq \frac{\|K\| \| \epsilon \|}{\| (K^*K + \lambda I)^{-1} \|^{-1}} + \delta \] for some $\tilde{j}$ and $\tilde{n}$.
Proof: \[ \| f_{j,n} - f_0 \| = \| f_{j,n} - f_j + f_j - f_0 \| \]

\[ \leq \| f_{j,n} - f_j, \epsilon \| + \| f_j, \epsilon - f_j \| + \| f_j - f_0 \| \]

\[ \leq \frac{\| K \| \| \epsilon \|}{\| (K^*K + \lambda I)^{-1} \|^{-1}} + \| f_{j,n}, \epsilon - f_{j}, \epsilon \| + \| f_{j} - f_{o} \|. \]

Now there exists a \( j \) and \( n \) such that

\[ \| f_{j,n}, \epsilon - f_{j}, \epsilon \| < \frac{\delta}{2} \quad \text{and} \quad \| f_{j} - f_{o} \| < \frac{\delta}{2}, \quad \text{say for} \quad j = j \]

\( n = n \) we have

\[ \| f_{j,n}, \epsilon - f_{o} \| \leq \frac{\| K \| \| \epsilon \|}{\| (K^*K + \lambda I)^{-1} \|^{-1}} + \delta. \]

2.4. **Some Error Bounds and Properties of \( \lambda \)**

Now we will prove some theorems which will yield some bounds on \( \| f_{\lambda} - f_0 \| \) in terms of \( \lambda > 0 \) where \( f_{\lambda} \in H_1 \) is a solution to \((K^*K + \lambda I)f = K^*g\) and \( f_0 \) is the least squares solution of minimum norm of \( Kf = g \). The bound on \( \| f_{\lambda} - f_0 \| \) will depend on knowledge of the singular system.
(Uₙ, Vₙ; ωₙ) for the compact operator K which is described below. The notation and results follow that of Strand [20, page 19], [21] and Tricomi [24]. Define

σ₀(K*K) = σ₀(KK*) = {γₙ : γₙ is an eigenvalue of K*K, γₙ > 0, n ∈ N_k} where N_k = {1,2,...,k}, (k may be 31).

Let

U = {Uₙ : KK*Uₙ = γₙ Uₙ, n ∈ N_k}

and

V = {Vₙ : K*KVₙ = γₙ Vₙ, n ∈ N_k}.

Since K*K is compact, symmetric and non-negative definite, assume that the γₙ are ordered such that

γ₁ ≥ γ₂ ≥ ... ≥ γₙ ≥ ... > 0. Define μₙ = γₙ⁻¹/₂, n ∈ N_k.

Then we have 0 < μ₁ ≤ μ₂ ≤ ... ,

Uₙ = μₙ KVₙ

and
\[ V_n = \mu_n K^* U_n. \]

Now in this notation we can state Picard's Theorem, a proof of which is given in Strand [20, page 25].

**Theorem 2.4.1. (Picard).** Let \((U_n, V_n; \mu_n)\) be a singular system for the compact operator \(K: H_1 \to H_2\) and \(\bar{g} \in H_2\).

Then the equation \(Kf = \bar{g}\) has a solution \(f_0 \in H_1\) if

and only if \(\sum_{n \in \mathbb{N}_m} \mu_n^2 |\langle \bar{g}, U_n \rangle|^2 < \infty\) and \(\bar{g} \in R(K)\). Also

for \(\bar{g} \in R(K)\) we have \(f_0 = \sum_{n \in \mathbb{N}_m} \langle \bar{g}, U_n \rangle \mu_n V_n\).

Strand [20, page 26] notes that for

\(g = g_1 + g_2, g_1 \in R(K)\) and \(g_2 \in N(K^*)\) then

\[ f_0 = \sum_{n \in \mathbb{N}_m} <g, U_n> \mu_n V_n\] is the least squares solution of minimum norm of the equation \(Kf = g\).
Theorem 2.4.2. Let

1. $K^*K$ have finite non-zero spectrum $\sigma_K, N_K = m,$
2. $f_\lambda \in H_\perp$ be the least squares solution of minimum norm of $Kf = g$ then

$$\|f_\lambda - f_0\| \leq \frac{\lambda}{\lambda + \gamma_m} \sqrt{\sum_{i=1}^{m} |\langle g, u_i \rangle| \mu_i^2} .$$

Proof:

$$\|f_\lambda - f_0\|^2 = \| (K^*K + \lambda I)^{-1}K^*g - f_0 \|^2$$

$$= \| (K^*K + \lambda I)^{-1}K^*f_0 - f_0 \|^2$$

$$= \| (K^*K + \lambda I)^{-1}(K^*K - (K^*K + \lambda I))f_0 \|^2$$

$$= \lambda^2 \| (K^*K + \lambda I)^{-1}f_0 \|^2$$

$$= \lambda^2 \| (K^*K + \lambda I)^{-1} \sum_{n=1}^{m} \langle g, u_n \rangle \mu_n v_n \|^2$$

$$= \lambda^2 \sum_{n=1}^{m} \langle g, u_n \rangle \mu_n (K^*K + \lambda I)^{-1}v_n \|^2 ,$$
and since \((\mathbf{K}\mathbf{K} + \lambda \mathbf{I})^{-1} \mathbf{v}_n = \frac{\mathbf{v}_n}{\gamma_n + \lambda}\) we have

\[
\| \mathbf{e}_\lambda - \mathbf{e}_0 \|^2 = \lambda^2 \sum_{n=1}^{m} (\langle \mathbf{g}, \mathbf{u}_n \rangle \mu_n \frac{\mathbf{v}_n}{\gamma_n + \lambda})^2
\]

\[
= \lambda^2 \sum_{n=1}^{m} \frac{|\langle \mathbf{g}, \mathbf{u}_n \rangle|^2 \mu_n^2}{(\gamma_n + \lambda)^2}
\]

\[
\leq \left( \frac{\lambda}{\gamma_m + \lambda} \right)^2 \sum_{n=1}^{m} |\langle \mathbf{g}, \mathbf{u}_n \rangle|^2 \mu_n^2.
\]

Now we take the square root of each side of the above inequality, giving

\[
\| \mathbf{e}_\lambda - \mathbf{e}_0 \| \leq \frac{\lambda}{\gamma_m + \lambda} \sqrt{\sum_{n=1}^{m} |\langle \mathbf{g}, \mathbf{u}_n \rangle|^2 \mu_n^2}.
\]

This completes the proof of the theorem.

Theorem 2.4.2 can be extended to the case where \(\mathbf{K}\mathbf{K}\) does not have a finite non-zero spectrum.
Theorem 2.4.3. If under the same hypotheses as Theorem 2.4.2, except that $K^*K$ has non-finite spectrum and an integer $N > 0$, a real number $\delta > 0$ are known such that

$$\sum_{n=N+1}^{\infty} |\langle g, u_n \rangle|^2 \mu_n^2 \leq \delta$$

then

$$\|f_\lambda - f_o\| \leq \sqrt{\left(\frac{\lambda}{\lambda + \gamma_N}\right)^2 \sum_{n=1}^{N} |\langle g, u_n \rangle|^2 \mu_n^2 + \delta}.$$

Proof: Similar to the proof of Theorem 2.4.2 we have

$$\|f_\lambda - f_o\|^2 = \lambda^2 \sum_{n=1}^{\infty} |\langle g, u_n \rangle| \frac{\mu_n^2}{\gamma_n + \lambda}$$

$$= \lambda^2 \sum_{n=1}^{\infty} |\langle g, u_n \rangle|^2 \frac{\mu_n^2}{(\gamma_n + \lambda)^2}.$$
\[ = \lambda^2 \sum_{n=1}^{N} |\langle g, U_n \rangle|^2 \frac{\mu_n^2}{(\gamma_n + \lambda)^2} \]

\[ + \sum_{n=N+1}^{\infty} |\langle g, U_n \rangle|^2 \mu_n \left(\frac{\lambda}{\gamma_n + \lambda}\right)^2 \]

\[ \leq \left(\frac{\lambda}{\gamma_N + \lambda}\right)^2 \sum_{n=1}^{N} |\langle g, U_n \rangle|^2 \mu_n^2 + \sum_{n=N+1}^{\infty} |\langle g, U_n \rangle|^2 \mu_n^2 \]

\[ \leq \left(\frac{\lambda}{\gamma_N + \lambda}\right)^2 \sum_{n=1}^{N} |\langle g, U_n \rangle|^2 \mu_n^2 + \delta. \]

Now we take the square root of each side of the above inequality, giving

\[ \| f_{\lambda} - f_0 \| \leq \sqrt{\left(\frac{\lambda}{\gamma_N + \lambda}\right)^2 \sum_{n=1}^{N} |\langle g, U_n \rangle|^2 \mu_n^2 + \delta.} \]
Remark: Let
\[
A_\lambda = \begin{cases} 
\frac{\lambda}{\gamma_n + \lambda} \sum_{n=1}^{m} |\langle g, u_n \rangle|^2 \mu_n^2 & \text{if } m < \infty \\
\sqrt{\left(\frac{\lambda}{\gamma_n + \lambda}\right)^2 \sum_{n=1}^{N} |\langle g, u_n \rangle|^2 + \delta} & \text{otherwise.}
\end{cases}
\]

From (2.2.7) and (2.2.8) we have

\[
(2.4.1) \quad \|f_{j,n} - f_o\| \leq \|f_{j,n} - f_j\| + \|f_j - f_o\| \\
\leq \frac{\|w_{j,n}\|}{\| (K^*K + \lambda_n I)^{-1} \|^{-1}} + A_{\lambda j}.
\]

Remark: Since the eigenvectors of $K^*K$ associated with its non-zero eigenvalues span $N(K)^\perp$ and $f_o, f_j, f_j,n$ belong to $N(K)^\perp$ for $j > 1$, we have

\[
\gamma \|f_j - f_o\| \leq \|K^*K(f_j - f_o)\|
\]
and

\[ (\gamma + \lambda_j) \| f_j, n - f_j \| \leq \| (K^*K + \lambda I)(f_j, n - f_j) \| \]

where \( \gamma = \inf_{\gamma_n} \gamma_n \) (see Helmberg [7, page 225]).

Since \( K^*K(f_j - f_\infty) = \lambda_j f_j \) and

\[ (K^*K + \lambda I)(f_j, n - f_j) = w_j, n \]

we have

\[ \| f_j - f_\infty \| \leq \frac{\lambda_j \| f_j \|}{\gamma} \]

and

\[ \| f_j, n - f_j \| \leq \frac{\| w_j, n \|}{\gamma + \lambda_j} \].

Thus (2.4.1) can be rewritten as

\[ (2.4.2) \quad \| f_j, n - f_\infty \| \leq \frac{\| w_j, n \|}{\gamma + \lambda_j} + \frac{\lambda_j \| f_j \|}{\gamma} \]
giving non-a priori error bounds when $\gamma \neq 0$.

Using the same techniques as in the above remark, we can replace $\| (K^*K + I\lambda)^{-1} \|^{-1}$ with $\gamma + \lambda$ in Theorems 2.3.1 and 2.3.2.

**Theorem 2.4.4.** If $f_\lambda$ is the solution to

$$(K^*K + \lambda I)f = K^*g$$

then $\| f_\lambda \|$ is strictly decreasing in $\lambda > 0$.

**Proof:** Since $K^*K : H_1 \rightarrow H_1$ is a symmetric, non-negative definite and compact linear operator, there is an orthonormal bases $\{ e_i \}$ for $H_1$ such that

$$(K^*K + \lambda I)e_i = (\gamma_i + \lambda)e_i$$

for all $i$. Since $K^*K$ is compact and non-negative we may assume $\gamma_i \geq \gamma_{i+1} \geq 0$ for all $i$. So we have that $K^*K + \lambda I = \text{diag}(\gamma_1 + \lambda, \gamma_2 + \lambda, \ldots)$ relative to $\{ e_i \}$. Thus

$$(K^*K + \lambda I)^{-1} = \text{diag}\left(\frac{1}{\gamma_1 + \lambda}, \frac{1}{\gamma_2 + \lambda}, \ldots\right).$$

Since $f_\lambda = (K^*K + \lambda I)^{-1}K^*g$, it follows that

$$f_\lambda = \sum_i \frac{\langle K^*g, e_i \rangle}{\gamma_i + \lambda} e_i.$$
The theorem follows immediately.

We now prove that $\|Kf_\lambda\|$ is increasing as $\lambda > 0$ is decreasing. The theorem following is of interest, in its own right, when compared to Theorem 2.4.4.

Lemma 2.4.5. If $\alpha > 0$, $\beta > 0$, $f_\alpha \in \mathcal{R}(K^*)$ is the solution to $(K^*K + \alpha I)f = K^*g$ and $f_\beta \in \mathcal{R}(K^*)$ is the solution to $(K^*K + \beta I)f = K^*g$ then

$$\langle f_\alpha, f_\beta \rangle \geq \frac{\alpha \langle f_\alpha, f_\alpha \rangle + \beta \langle f_\beta, f_\beta \rangle}{\alpha + \beta}$$

and

$$\langle Kf_\alpha, Kf_\beta \rangle \geq \frac{\alpha \langle Kf_\alpha, Kf_\alpha \rangle + \beta \langle Kf_\beta, Kf_\beta \rangle}{\alpha + \beta}$$
Proof: We have the result that

\[ K^*Kf_\alpha + \alpha f_\alpha = K^*Kf_\beta + \beta f_\beta \]

which implies

\[(2.4.3) \quad K^*K(f_\alpha - f_\beta) = \beta f_\beta - \alpha f_\alpha.\]

Now taking the inner product of each side of (2.4.3) with \( f_\alpha - f_\beta \) we obtain the following result.

\[ 0 \leq \langle K^*K(f_\alpha - f_\beta), f_\alpha - f_\beta \rangle \]

\[ = \langle \beta f_\beta - \alpha f_\alpha, f_\alpha - f_\beta \rangle \]

\[ = \beta \langle f_\beta, f_\alpha - f_\beta \rangle - \alpha \langle f_\alpha, f_\alpha - f_\beta \rangle \]

\[ = \beta \langle f_\beta, f_\alpha \rangle - \beta \langle f_\beta, f_\alpha \rangle - \alpha \langle f_\alpha, f_\alpha \rangle + \alpha \langle f_\alpha, f_\beta \rangle \]

\[ = (\alpha + \beta) \langle f_\alpha, f_\beta \rangle - (\beta \langle f_\beta, f_\beta \rangle + \alpha \langle f_\alpha, f_\alpha \rangle). \]

So we have
\[ \frac{\beta \langle f_\beta, f_\beta \rangle + \alpha \langle f_\alpha, f_\alpha \rangle}{\alpha + \beta} \leq \langle f_\beta, f_\alpha \rangle. \]

The result

\[ \frac{\beta \langle Kf_\beta, Kf_\beta \rangle + \alpha \langle Kf_\alpha, Kf_\alpha \rangle}{\alpha + \beta} \leq \langle Kf_\alpha, Kf_\beta \rangle \]

has a similar proof, starting with

\[ K^*K(K^*K(f_\alpha - f_\beta)) = K^*K(\beta f_\beta - \alpha f_\alpha) \]

and then taking the inner product of each side with \( f_\alpha - f_\beta \).

**Theorem 2.4.5.** If \( 0 < \beta < \alpha \), \( f_\alpha \in R(K^*) \) is the solution to \( (K^*K + \alpha I)f = K^*g \) and \( f_\beta \in R(K^*) \) is the solution to \( (K^*K + \beta I)f = K^*g \) then \( \|Kf_\beta\| \geq \|Kf_\alpha\| \).

**Proof:** Since we have

\[ (K^*K + \alpha I)f_\alpha = (K^*K + \beta I)f_\beta, \]
take the inner product of each side with $f_a$ yielding the following result.

$$\langle K*K f_a + \alpha f_a, f_a \rangle = \langle K*K f_\beta + \beta f_\beta, f_a \rangle,$$

$$\langle K*K f_a, f_a \rangle + \alpha \langle f_a, f_a \rangle = \langle K*K f_\beta, f_a \rangle + \beta \langle f_\beta, f_a \rangle,$$

$$\langle K f_a, K f_\beta \rangle + \alpha \langle f_a, f_\beta \rangle = \langle K f_\beta, K f_\alpha \rangle + \beta \langle f_\beta, f_\alpha \rangle.$$

In a similar way, taking the inner product with $f_\beta$ yields:

$$\langle K f_\alpha, K f_\beta \rangle + \alpha \langle f_\alpha, f_\beta \rangle = \langle K f_\beta, K f_\beta \rangle + \beta \langle f_\beta, f_\beta \rangle.$$

Now using the previous lemma we have

$$\langle K f_\alpha, K f_\alpha \rangle + \alpha \langle f_\alpha, f_\alpha \rangle = \langle K f_\beta, K f_\beta \rangle + \beta \langle f_\beta, f_\alpha \rangle$$

$$\geq \frac{\beta \langle K f_\beta, K f_\beta \rangle + \alpha \langle K f_\alpha, K f_\alpha \rangle}{\alpha + \beta} + \beta \frac{\langle f_\beta, f_\beta \rangle + \alpha \langle f_\alpha, f_\alpha \rangle}{\alpha + \beta}$$

When we multiply each side of the previous result by $\alpha + \beta$ and simplify the algebraic expression, we obtain
(2.4.4) \[
\beta \langle \mathbf{K}_\alpha \mathbf{f}_\alpha, \mathbf{K}_\alpha \mathbf{f}_\alpha \rangle + \alpha^2 \langle \mathbf{f}_\alpha, \mathbf{f}_\alpha \rangle \geq \beta \langle \mathbf{K}_\beta \mathbf{f}_\beta, \mathbf{K}_\beta \mathbf{f}_\beta \rangle + \beta^2 \langle \mathbf{f}_\beta, \mathbf{f}_\beta \rangle.
\]

In a similar way from

\[
\langle \mathbf{K}_\alpha \mathbf{f}_\alpha, \mathbf{K}_\beta \mathbf{f}_\beta \rangle + \alpha \langle \mathbf{f}_\alpha, \mathbf{f}_\beta \rangle = \langle \mathbf{K}_\beta \mathbf{f}_\beta, \mathbf{K}_\alpha \mathbf{f}_\alpha \rangle + \beta \langle \mathbf{f}_\beta, \mathbf{f}_\alpha \rangle
\]

we obtain:

(2.4.5) \[
\alpha \langle \mathbf{K}_\beta \mathbf{f}_\beta, \mathbf{K}_\beta \mathbf{f}_\beta \rangle + \beta^2 \langle \mathbf{f}_\beta, \mathbf{f}_\beta \rangle \geq \alpha \langle \mathbf{K}_\alpha \mathbf{f}_\alpha, \mathbf{K}_\alpha \mathbf{f}_\alpha \rangle + \alpha^2 \langle \mathbf{f}_\alpha, \mathbf{f}_\alpha \rangle.
\]

Now add equations (2.4.4) and (2.4.5), obtaining the inequality:

(2.4.6) \[
\beta \langle \mathbf{K}_\alpha \mathbf{f}_\alpha, \mathbf{K}_\alpha \mathbf{f}_\alpha \rangle + \alpha \langle \mathbf{K}_\beta \mathbf{f}_\beta, \mathbf{K}_\beta \mathbf{f}_\beta \rangle \geq

\beta \langle \mathbf{K}_\beta \mathbf{f}_\beta, \mathbf{K}_\beta \mathbf{f}_\beta \rangle + \alpha \langle \mathbf{K}_\alpha \mathbf{f}_\alpha, \mathbf{K}_\alpha \mathbf{f}_\alpha \rangle.
\]

So (2.4.6) yields the result:
\[(\alpha - \beta) \langle \text{Kf}_\beta, \text{Kf}_\beta \rangle \geq (\alpha - \beta) \langle \text{Kf}_\alpha, \text{Kf}_\alpha \rangle.\]

Since \(0 < \beta < \alpha\) we have

\[\langle \text{Kf}_\beta, \text{Kf}_\beta \rangle \geq \langle \text{Kf}_\alpha, \text{Kf}_\alpha \rangle.\]

Thus we have \(\|\text{Kf}_\beta\| \geq \|\text{Kf}_\alpha\|\) and the theorem.
3. IMPLEMENTING THE ALGORITHM

3.1. Calculation of a Starting Element

For the basic algorithm described in section 3.2 we require an initial choice for \( \lambda, \lambda_1 \) and any starting element \( f_{1,0} \in H_1 \).

In practice we will choose any initial starting element \( f \in H_1 \) such that \( f \neq 0 \) and \( \|K^*Kf\| 
eq 0 \). Then we determine an actual starting vector \( f_{1,0} \in H_1 \) and a initial \( \lambda, \lambda_1 \), such that \( f_{1,0} \) is, in some sense, close to a solution of \( K^*Kf + \lambda_1 f = K^*g \).

Now to choose the actual \( f_{1,0} \), consider the element

\[
f_s = \frac{K^*Kf}{\|K^*Kf\|} \in R(K^*).
\]

The element \( f_s \) is chosen normalized so as to control the size of the norm of the actual starting element \( f_{1,0} \), which will be either \( f_s \) or \(-f_s\).

Choose \( \lambda_1 \geq 0 \) so that

\[
\hat{Q}_\lambda(f_s) = \| (K^*K + \lambda I)f_s - K^*g \|^2
\]
is minimized. Note that if \( \hat{\Omega}_\lambda (f_s) = 0 \) for some choice for \( \lambda > 0 \) then we would have obtained an equation

\[
(K*K + \lambda I)f = K*g
\]

which has \( f_s \) as its solution.

To minimize \( \hat{\Omega}_\lambda (f_s) \) with respect to \( \lambda \), we write

\[
\hat{\Omega}_\lambda (f_s) = \|K*Kf_s - K*g\|^2
\]

\[+ 2\lambda \langle f_s, K*Kf_s - K*g \rangle + \lambda^2 \|f_s\|^2.\]

Taking the derivative \( \hat{\Omega}_\lambda (f_s) \) with respect to \( \lambda \) and setting \( \frac{d}{d\lambda} \hat{\Omega}_\lambda (f_s) = 0 \) we get the result

\[
\frac{d}{d\lambda} \hat{\Omega}_\lambda (f_s) = 2\langle f_s, K*Kf_s - K*g \rangle + 2\lambda \|f_s\|^2
\]

\[= 2\langle f_s, K*Kf_s - K*g \rangle + 2\lambda = 0.\]
Note also that \( \frac{d^2 \hat{Q}(f_s)}{d\lambda^2} = 2\langle f_s, f_s \rangle > 0 \), so that the solution to \( \frac{d \hat{Q}(f_s)}{d\lambda} = 0 \) yields a minimum.

By solving \( \frac{d \hat{Q}_\lambda(f_s)}{d\lambda} = 0 \) for \( \lambda = \lambda^* \), the minimum, we have the result:

\[
(3.1.1) \quad \lambda^* = -\langle f_s, K^* K f_s - K^* g \rangle
\]
\[
= \langle g - K f_s, K f_s \rangle
\]
\[
= \langle K f_s, g \rangle - \langle K f_s, K f_s \rangle.
\]

By substituting \( \lambda = \lambda^* \) in \( \hat{Q}_\lambda(f_s) \) we obtain

\[
\hat{Q}_\lambda(f_s) = \| K^* K f_s - K^* g \|^2 - \langle K f_s, g - K f_s \rangle^2.
\]

For the choice for \( \lambda_1 \) we need \( \lambda_1 \geq 0 \), so consider the following three cases:

**Case 1.** If \( \lambda^* = 0 \), take \( \lambda_1 = \lambda^* \) and we have computed an exact solution, \( f_s \in R(K^*) \subset N(K)^{\perp} \), to \( K^* K f = K^* g \).
Since $K^*$ is one-to-one on the range of $K$ we have that

$$K^*(Kf_s - g_1) = 0,$$

so $Kf_s = g_1$. Thus we have computed the least squares solution of minimum norm, $f_s = f_0$, of $Kf = g$.

**Case 2.** If $\bar{\lambda} > 0$ then take $\lambda_1 = \bar{\lambda}$ and $f_{1,0} = f_s$ for a starting element.

**Case 3.** If $\bar{\lambda} < 0$ then replace $f_s$ with $-f_s$ in (3.1.1), since then the new value of $\bar{\lambda}$ is

$$\bar{\lambda} = \langle K(-f_s), g - K(-f_s) \rangle$$

$$= \langle -Kf_s, g + Kf_s \rangle$$

$$= -\langle Kf_s, g \rangle - \langle Kf_s, Kf_s \rangle,$$

as compared with equation (3.1.1). So if $\langle Kf, g - Kf \rangle$ is negative for $f_s$ and positive for $-f_s$ take $f_{1,0} = -f_s$ and $\lambda_1 = \bar{\lambda} = -\langle Kf_s, g \rangle - \langle Kf_s, Kf_s \rangle$.

If $\langle Kf_s, g \rangle = 0$ then a new starting element $f$ must be chosen.
Now from here on suppose that \( \langle Kf_s, g \rangle \neq 0 \).

If \( \langle Kf, g - Kf \rangle \) is negative for both \( f_s \) and \( -f_s \), then from (3.1.1) and (3.1.2) there exists a real number \( a > 1 \) such that for the scaled equation

\[
Kf = ag
\]

we have \( \bar{\lambda} = \langle Kf, ag - Kf \rangle \geq 0 \) for \( f \) one of the elements \( f_s \) or \( -f_s \). In the above computation it was critical that \( \langle Kf_s, g \rangle \neq 0 \). So we take \( \lambda_1 = \bar{\lambda} \) and \( f_{1,0} \) to be the element \( f_s \) or \( -f_s \) which makes \( \lambda_1 \geq 0 \). The new equation \( Kf = ag \) is then solved for \( af_{1,0} \), the least squares solution of minimum norm of \( Kf = ag \), and scaled back to the least squares solution of minimum norm, \( f_{1}' \), of the original equation \( Kf = g \) at the end of the algorithm.

In summary, for each \( f \neq 0 \) where \( \|K^*Kf\| \neq 0 \) and \( f \in H_1 \) a starting element \( f_{1,0} \in R(K^*) \) and a \( \lambda_1 \geq 0 \) can be chosen so that \( \| (K^*K + \lambda I)f_s - aK^*g \| \) is a minimum (here \( a \geq 1 \)). Thus we start as close as possible (in norm) to a solution of \( (K^*K + \lambda I)f - aK^*g = 0 \).
Once $\lambda_1$ is chosen, a sequence $\{\lambda_j\}_{j=1}^\infty$ can be constructed such that $\lambda_j > 0$, $\lambda_j \to 0$ as $j \to \infty$, by taking $\lambda_{j+1} = x\lambda_j$ where $0 < x < 1$. In practice the sequence $\{\lambda_j\}_{j=1}^\infty$ is terminated when $\lambda_j$ is less than some pre-assigned value $\mu$. For the first $\lambda_j$ such that $\lambda_j \leq \mu$ take $\lambda_j = \mu$.

$x$ is determined by numerical experimentation.

3.2. The Sequence $\{f_{j,n}\}_{n=0}^\infty$

The sequence $\{f_{j,n}\}_{n=0}^\infty$, for $j$ fixed, is constructed by taking

$$f_{j,n+1} = f_{j,n} + \alpha_{j,n}w_{j,n}$$

where $\alpha_{j,n}$ and $w_{j,n}$ are given by (2.2.8). The construction of the sequence $\{f_{j,n}\}_{n=0}^\infty$ is terminated when a measure of convergence fails, due to rounding errors in computation, to give an approximate solution $f_j$ to the equation $(K^*K + \lambda_j I)f = K^*g$ (possibly scaled). The measure of convergence used is to note that from (2.2.4),
\[ Q_{\lambda_j}(f_{j,n}) - Q_{\lambda_j}(f_{j,n+1}) = \frac{\langle W_{j,n}, W_{j,n} \rangle^2}{\langle W_{j,n}, K W_{j,n}, W_{j,n} \rangle + \lambda_j \langle W_{j,n}, W_{j,n} \rangle} \]

for \( W_{j,n} \neq 0 \), i.e. \( Q_{\lambda_j}(f_{j,n+1}) < Q_{\lambda_j}(f_{j,n}) \), an approximation to \( f_j \) is chosen to be the least value of \( n \) for which \( Q_{\lambda_j}(f_{j,n+1}) < Q(f_j,n) \) fails in the computations, or \( \|W_{j,n}\| \) is less than some pre-assigned value; which ever occurs first.

We use \( \|W_{j,n}\| \) as a measure of \( \|f_{j,n} - f_o\| \), since

\[ \|f_{j,n} - f_o\| \leq \frac{\|W_{j,n}\|}{\|K^{*}K + \lambda I\|^{-1}} + \|f_j - f_o\| \]

(see 2.4.1). Ideally we first choose \( \lambda \) so that \( \|f_j - f_o\| \) is small and then make \( \|W_{j,n}\| \) small. In practice it is difficult to make \( \|f_j - f_o\| \) small with out some knowledge of the solution.
3.3. A Flow Diagram of the Basic Algorithm

The following diagram, given in Figure 3.3.1, is a basic description of the algorithm to find the least squares solution of minimum norm of the operator equation $Kf = g$. The flow diagram describes the program in Appendix A for matrix equations and the subroutine BITER for discretized integral equations in Appendix B.
48

Input:
Operator K
(Operator K* input or constructed)
Lambda multiplier
Terminal lambda
$g$
Control on $|\mathbf{w}_{\lambda,n}|^2$.
Initial starting vector $f \neq 0$
$M$, total count of iterations at each step

Initialize:
$LL=0$, total count of iterations

Compute:
$K^*$ (or omit)
$K^*g$
$||K^*g||$

Output:
$||K^*g||$

If $||K^*g|| > 10^{-40}$

Write:
The Solution is the Zero Function

Figure 3.3.1. Flow Chart of the Basic Algorithm
Compute:
\[ K^*K \]
\[ f_s = \frac{K^*K_f}{\|K^*K_f\|} \]
\[ K_f \]
Initialize:
CR=1., scaling factor
U=2., scaling multiplier
j=1

Compute:
\[ \lambda_1 = \langle s-K_f s, f_s \rangle \]
Starting lambda

If \( \lambda_1 \geq 0 \)

yes

\[ g = U \cdot g \]
\[ j = j + 1 \]
\[ CR = U \cdot CR \]
\[ K_f s = K_f s \]

no

If \( j = 2 \)

yes

no

Figure 3.3.1. (Continued)
Comment:
\( f_s = f_{1,0} \)

Compute:
\( Kf_s^{-g} \)
\( <Kf_s^{-g}, Kf_s^{-g}> \)
\( K^*(Kf_s^{-g}) \)
\( <K^*(Kf_s^{-g}), K^*(Kf_s^{-g})> - \lambda_1^2 \) (minimum)
\( <Kf_s^{-g}, Kf_s^{-g}> + \lambda_1 \) (initial \( Q_1(f_1) \))
\( W_{1,0} = K^*(Kf_s^{-g}) + \lambda_1 f_s \)

Output:
Initial lambda \( \lambda_1 \)
Starting vector \( f_s \)
Initial \( W_{1,0} \)
Scaling factor CR
\( <K^*(Kf_s^{-g}), K^*(Kf_s^{-g})> - \lambda_1^2 \)

Figure 3.3.1. (Continued)
Figure 3.3.1. (Continued)
Write:
At $\lambda = \ldots$ the maximum value of $M$ has been attained.

Figure 3.3.1. (Continued)
Figure 3.3.1. (Continued)
4. EXAMPLES

4.1. Matrix Examples

Let the real Hilbert spaces $H_1 = E^n$, $H_2 = E^m$ where

$$E^n = \{(x_1, \ldots, x_n)^T : x_i \text{ real}\}$$

be given. In this finite dimensional case each linear transformation $A : H_1 \to H_2$ has a unique matrix representation

$$A = (a_{ij})_{i=1}^m_{j=1}^n$$

with respect to each bases $\{v_1, v_2, \ldots, v_n\}$ in $H_1$ and $\{w_1, w_2, \ldots, w_m\}$ in $H_2$. We shall always take our bases vectors to be the canonical bases. Since $H_1$ and $H_2$ are finite dimensional, all linear transformations are bounded and compact. Since $A$ is of finite rank we have $R(A) = \overline{R(A)}$, so the algorithm (see section 2.2) will
converge to the least squares solution of minimum norm, \( x_0 \), of \( Ax = g \) for all \( g \in \mathbb{H}_2 \).

Let

\[
A = \begin{bmatrix}
1 & 1 & 1 & 0 & 0 & 0 \\
0 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 \\
1 & 2 & 2 & 1 & 0 & 0 \\
3 & 3 & 3 & 1 & 1 & 1 \\
1 & 2 & 2 & 2 & 1 & 1 \\
\end{bmatrix}.
\]

Then \( A \) has rank three.

Consider \( Ax = g \) for the following two cases.

**Case 1.** \( g_1 = (10,12,13,22,43,35)^T \in R(A) \).

**Case 2.**

\[
g_2 = g_1 + (-5,-2,-2,1,1,1)^T
\]

\[
= (5,10,11,23,24,36)^T
\]

where \((-5,-2,-2,1,1,1)^T \in N(A)\).

In both Case 1 and Case 2 the least squares solution of minimum norm, \( x_0 \), of \( Ax = g_i \), \( i = 1,2 \) is
\[ \overrightarrow{x}_o = (17/6, 43/12, 43/12, 29/6, 49/12, 49/12) \in N(A). \]

\[ \overrightarrow{x}_o \] was constructed by finding a basis

\[
\begin{bmatrix}
-1 \\
1 \\
0 \\
-1 \\
1 \\
0 \\
\end{bmatrix}
\begin{bmatrix}
-1 \\
0 \\
1 \\
-1 \\
1 \\
0 \\
\end{bmatrix}
\begin{bmatrix}
0 \\
0 \\
0 \\
0 \\
0 \\
1 \\
\end{bmatrix}
\]

for \( N(A) \). Then a basis

\[
\begin{bmatrix}
1 \\
0 \\
-1 \\
0 \\
0 \\
0 \\
\end{bmatrix}
\begin{bmatrix}
0 \\
1 \\
1 \\
0 \\
0 \\
0 \\
\end{bmatrix}
\begin{bmatrix}
0 \\
0 \\
1 \\
\end{bmatrix}
\]

is found for \( N(A)^\perp \).

Since \( \overrightarrow{x}_o \) is contained in \( N(A)^\perp \), \( \overrightarrow{x}_o \) must be a linear combination of the above three vectors. Therefore to solve

\[ Ax = q_i \]

we need only solve
The algorithm in Appendix A was used to find the approximate least squares solution of minimum norm for Case 1 and Case 2 with the results given in Table 1. Comparisons with the actual solution were computed separately.

For Run 1 Case 1 we will satisfy the error bounds of Statement (2.4.1) so that \( \| f_{j,n} - f_o \| < 10^{-6} \). We will assume that we know the eigenvalues of \( A^T A \), which were computed to be 56.96, 4.72, 2.32, 0, 0, 0. Also we will assume that we have an approximation for \( \| f_o \| \), the norm of the least squares solution of minimum norm, say 9.5. Now from (2.4.1) we have

\[
\| f_{j,n} - f_o \| \leq \frac{\| W_{\lambda_j,n} \|}{\lambda_j + 2.32} + \frac{9.5\lambda_j}{\lambda_j + 2.32}
\]
We will choose terminal lambda, \( \mu = \lambda_{j} \) and \( ||w_{j,n}|| \) so that

\[
\frac{9.5\mu}{\mu + 2.32} < \frac{10^{-6}}{2}
\]

and

\[
\frac{\|w_{j,n}\|}{\mu + 2.32} < \frac{10^{-6}}{2}
\]

We actually used \( \mu = 1 \times 10^{-7} \) and made \( \|w_{j,n}\|^2 \leq 10^{-12} \).

The results of several runs of Case 1 and Case 2 are given in Table 1. For each run we used the initial starting vector \( F = (1, 1, 1, 0, 0, 0)^T \). We computed \( \|A^T q_{i}\| = 470.14 \).
Table 1. Matrix Example

<table>
<thead>
<tr>
<th>Run 1</th>
<th>Run 2</th>
<th>Run 3</th>
<th>Run 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>g</td>
<td>g₀₁</td>
<td>g₁</td>
<td>g₁</td>
</tr>
<tr>
<td>M</td>
<td>400</td>
<td>400</td>
<td>400</td>
</tr>
<tr>
<td>TL</td>
<td>1.(-7)</td>
<td>1.(-16)</td>
<td>1.(-16)</td>
</tr>
<tr>
<td>W₁</td>
<td>1.(-12)</td>
<td>1.(-10)</td>
<td>1.(-20)</td>
</tr>
<tr>
<td>λm</td>
<td>1.(-5)</td>
<td>1.(-5)</td>
<td>1.(-5)</td>
</tr>
<tr>
<td>TL</td>
<td>411.68</td>
<td>411.68</td>
<td>411.68</td>
</tr>
<tr>
<td>EN</td>
<td>4.2(-7)</td>
<td>3.3(-6)</td>
<td>3.3(-11)</td>
</tr>
<tr>
<td>E₁</td>
<td>1.(-6)</td>
<td>4.4(-6)</td>
<td>4.4(-11)</td>
</tr>
<tr>
<td>E₂</td>
<td>1.(-6)</td>
<td>4.4(-6)</td>
<td>4.4(-11)</td>
</tr>
<tr>
<td>E₃</td>
<td>2.3(-6)</td>
<td>1.9(-6)</td>
<td>1.8(-11)</td>
</tr>
<tr>
<td>W₂</td>
<td>8.8(-7)</td>
<td>8.6(-6)</td>
<td>8.7(-11)</td>
</tr>
<tr>
<td>Re</td>
<td>6.5(-7)</td>
<td>5.(-6)</td>
<td>5.1(-11)</td>
</tr>
<tr>
<td>TI</td>
<td>132</td>
<td>104</td>
<td>293</td>
</tr>
<tr>
<td>Ex</td>
<td>3.19</td>
<td>2.60</td>
<td>6.93</td>
</tr>
</tbody>
</table>

M = maximum number of allowable iterations for a given step.

TL = terminal Lambda.

W₁ = terminal choice for \[ ||W_{j,n}||^2 \].

λm = Lambda multiplier.

TL = computed initial \( \lambda_1 \), approximately.

EN = computed \( ||f_{T\lambda,n} - f_o|| \).

E₁ = error bound computed from (2.4.1).

E₂ = error bound computed from (2.4.2).

E₃ = computed \( \max|f_{T\lambda,n} - f_o| \).

W₂ = computed \( ||(A^TA + \lambda I)f_{T\lambda,n} - A^Tg|| \).

Re = computed \( ||(A^Tf_{T\lambda,n} - g)|| \).

TI = total iterations.

Ex = execution time, seconds, WATFIVE.
4.2. An Inner Product Space

We will consider the real \( n \)-dimensional space \( \mathbb{R}^n \) with scalar field the real numbers. Define a mapping from \( \mathbb{R}^n \times \mathbb{R}^n \) to \( \mathbb{R} \) by

\[
(u, v) \rightarrow \langle u, v \rangle = \sum_{j=1}^{n} T_j u_j v_j
\]

where \( u = (u_1, u_2, \ldots, u_n)^T \), \( v = (v_1, v_2, \ldots, v_n)^T \) and \( T_j > 0, T_j \in \mathbb{R} \) for all \( j = 1, \ldots, n \). Thus we have a Hilbert space, which will be denoted by \( H^n_T \). Let \( \{e_1^n, e_2^n, \ldots, e_n^n\} \) be the orthogonal (not necessarily normalized) bases for \( H^n_T \) where \( e_j^n = (0, 0, \ldots, 1_j, \ldots, 0)^T \), \( j = 1, n \).

If \( A : H^n_T \rightarrow H^m_S \) then \( A^* : H^m_S \rightarrow H^n_T \) will denote the adjoint.

**Theorem 4.2.1.** If \( A : H^n_T \rightarrow H^n_S \) is \( (T_j a_{i,j})_{i=1,n} \) \( j=1,m \)

\[
A^* = (S_i a_{i,j})_{j=1,m}^{i=1,n}
\]
Proof: We will show that the above $A^*$ is the adjoint of $A$.

\[
\langle A e^m_k, e^n_{\ell} \rangle = \langle \sum_{n} T_{k} a_{i,k} e^n_{i}, e^n_{\ell} \rangle
\]

\[
= T_{k} a_{\ell,k} \langle e^n_{\ell}, e^n_{\ell} \rangle
\]

\[
= T_{k} S_{\ell,k} a_{\ell,k}
\]

and

\[
\langle e^m_k, A^* e^n_{\ell} \rangle = \langle e^m_k, \sum_{m} S_{k} a_{j,k} e^m_{j} \rangle
\]

\[
= S_{k} a_{\ell,k} \langle e^m_{\ell}, e^m_{\ell} \rangle
\]

\[
= S_{k} e_{\ell,k} a_{\ell,k}
\]

\[
= \langle A e^m_k, e^n_{\ell} \rangle
\]

for all $k$ and $\ell$, therefore $A^*$ is the adjoint of $A$. 
4.3. **Discretizing Integral Equations**

We will obtain a discretized form of Fredholm integral equations of the first kind,

\[(4.3.1) \quad Kf(y) = \int_0^1 k(y,x)f(x)dx = g(y), \quad 0 \leq y \leq 1\]

where \( g \in \text{R}(K) + N(K^*) \subset L_2[0,1], f \in L_2[0,1] \) and \( k(y,x) \) is square integrable with respect to the product Libesque measure on \([0,1] \times [0,1]\).

Akhiezer and Glazman [1] prove that if
\[
\int_0^1 \int_0^1 |k(y,x)|^2 dy \, dx < \infty \quad \text{then } K \text{ is compact.}
\]

For \( Kf(s) = \int_0^1 k(s,x)f(x)dx \) Bachman and Narici [3, page 403] note that the adjoint to \( K \) is \( K^* \), given by

\[
K^*f(s) = \int_0^1 k(x,s)f(x)dx.
\]

Of particular interest is the equation
(4.3.2) 
\[
\int_0^1 k(y,s) \int_0^1 k(y,x) f(x) \, dx \, dy + \lambda f(s) = \int_0^1 k(y,s) g(y) \, dy
\]
for \( \lambda > 0 \).

For numerical purposes, we will restrict the discussion to the functions \( k(y,x), g \) and \( f_\lambda \) where we have Riemann integration.

The methods of discretization of the integrals will initially parallel that found in Anselone [2, page 13] and Isaacson and Keller [9].

We have that

(4.3.3) 
\[
\int_0^1 k(y,s) g(y) \, dy = \sum_{j=1}^{m} S_j k(y_j,s) g(y_j) + R_m(s)
\]
where \( R_m(s) \) is the discretization error. Here \( S_j > 0 \)

\[
\sum_{j=1}^{m} S_j = 1 \quad \text{are the quadrature weights for}
\]

\( 0 = y_1 < y_2 < \ldots < y_m = 1 \), a partition of the interval [0,1], chosen equally spaced for the Newton-Cotes closed
integration formulae.

For \( \int_0^1 k(y,s) \int_0^1 k(y,x) f(x) \, dx \, dy \) we have

\[
\int_0^1 k(y,s) \int_0^1 k(y,x) f(x) \, dx \, dy = \sum_{j=1}^m S_j k(y_j,s) \sum_{i=1}^n T_i k(y_i,x_i) f(x_i) + R(s)
\]

where \( R(s) \) is the discretization error. Here \( T_i > 0 \) for \( i = 1, n, \sum_{i=1}^n T_i = 1 \) are the quadrature weights for \( 0 = x_1 < x_2 < \ldots < x_n = 1 \), a partition of the interval [0,1] chosen equally spaced for the Newton-Cotes closed integration formulae. The partition and weights for the \( y_j \)'s are chosen as for (4.3.3).

Thus from (4.3.3) and (4.3.4), (4.3.2) becomes

\[
\sum_{j=1}^m S_j k(y_j,s) \sum_{i=1}^n T_i k(y_j,x_i) f(x_i) + \lambda f(s) + R(s)
\]

\[
= \sum_{j=1}^m S_j k(y_j,s) g(y_j) + R_m(s).
\]
Now, (4.3.5) can be written in the equivalent form

\[(S_j k(y_j, s))_{j=1, m} (T_i k(y_j, x_i))_{j=1, m} (f(x_i))_{i=1, n} + \lambda f(s) + \bar{R}(s)_{i=1, n}\]

\[= \sum_{j=1}^{m} S_j k(s, y_j) g(y_j) + R_m(s). \text{ Where } (v_i)_{i=1, \ell} \text{ denotes an } \ell\text{-dimensional column vector.}\]

Now partition the interval \([0,1]\) for \(s\) such that \(s_i = x_i\) for \(i = 1, n\). Thus we obtain the linear system

(4.3.6)

\[(S_j k(y_j, x_i))_{j=1, n} \cdot (T_i k(y_j, x_i))_{j=1, m} (f(x_i))_{i=1, n}\]

\[+ \lambda (f(x_i))_{i=1, n} = (S_j k(Y_j, x_i))_{j=1, m} (g(Y_j))_{j=1, m}\]

\[+ (R_m(x_i) - \bar{R}(x_i))_{i=1, n}.

Let
\[ K = \left( T_i k(y_j, x_i) \right)_{j=1,m, \atop i=1,n}; \quad \mathbb{K} : \mathbb{H}_T^n \rightarrow \mathbb{H}_S^m; \]

\[ \mathbf{\chi}^* = \left( S_j k(y_j, x_i) \right)_{i=1,n, \atop j=1,m}; \quad \mathbf{\chi}^* : \mathbb{H}_S^m \rightarrow \mathbb{H}_T^n; \]

\[ \bar{f} = (f(x_i))_{i=1,n}, \quad \bar{f} \in \mathbb{H}_T^n; \]

\[ \bar{g} = (g(x_j))_{j=1,m}, \quad \bar{g} \in \mathbb{H}_S^m \]

and

\[ \bar{R}_l = (R_n(x_i) - \bar{R}(x_i))_{i=1,n}, \quad \bar{R} \in \mathbb{H}_T^n. \]

Thus (4.3.6) can be written in the form

(4.3.7) \[ \mathbf{\chi}^* \bar{K} \bar{f} + \lambda \bar{f} = \mathbf{\chi}^* \bar{g} + \bar{R}_l. \]

Similar to (4.3.5) we have the integral equation

(4.3.8) \[ \int_0^1 k(y_s) \int_0^1 k(y, x) f(x) \, \mathrm{d}x \, \mathrm{d}y = \int_0^1 k(y, s) g(y) \, \mathrm{d}y \]
and its discretized form

\begin{equation}
K^*K\vec{f} = K^*\vec{g} + \vec{R}_1,
\end{equation}

where \( \vec{R}_1 \) is the discretization error vector. To solve equation (4.3.8) numerically we consider the system

\begin{equation}
K^*K\vec{f} + \lambda \vec{f} = K^*\vec{g}
\end{equation}

for \( \lambda > 0 \) and small.

4.4. Error Bounds

Throughout this section we will assume that

\[ \gamma = \inf_{\vec{\gamma}_n} \gamma_n. \]

For \( \lambda > 0 \), first we will develop relationships between the solutions of the following equations

\[ K^*K\vec{f}_\lambda + \lambda \vec{f}_\lambda = K^*\vec{g} + \vec{R}_\lambda, \]

\[ K^*K\vec{f}_0 = K^*\vec{g} + \vec{R}_0 \]

and

\[ K^*K\vec{f}_\lambda + \lambda \vec{f}_\lambda = K^*\vec{g} \]
where $\vec{R}_\lambda$, $\vec{R}_0$ are the discretization error vector.

$\vec{f}_\lambda$, $\vec{f}_0$ and $\vec{f}_\lambda$ are the respective solutions to the above equations. These relationships will depend on an estimate for $\|\vec{R}_\lambda\|$, $\|\vec{R}_0\|$ and $\gamma$.

Assuming the above notation we have the following two theorems.

**Theorem 4.4.1.**

\[(4.4.1) \quad \|\vec{f}_\lambda - \vec{f}_\lambda\| \leq \frac{\|\vec{R}_\lambda\|}{\mu + \lambda}\]

where \(\mu = \inf_{\gamma_n \in \sigma(K^*K)} \gamma_n\).

**Proof:** Since \((K^*K + \lambda I)\vec{f}_\lambda = K^*\vec{g} + \vec{R}_\lambda\) and 
\((K^*K + \lambda I)\vec{f}_\lambda = K^*\vec{g}\) we have \((K^*K + \lambda I)(\vec{f}_\lambda - \vec{f}_\lambda) = \vec{R}_\lambda\). Thus

\[(\mu + \lambda)\|\vec{f}_\lambda - \vec{f}_\lambda\| = \| (K^*K + \lambda I)^{-1}\|^{-1}\|\vec{f}_\lambda - \vec{f}_\lambda\| \]

\[\leq \| (K^*K + \lambda I)(\vec{f}_\lambda - \vec{f}_\lambda)\| \]

\[= \|\vec{R}_\lambda\| .\]
Thus we have the theorem.

Remark: Since \( K^*Kf = K*g + R \), we have \( K^*(Kf - g) = R \).
Thus \( R \in R(K^*) \subset N(K)^\perp \). Therefore \( \tilde{f}_0 \in N(KK)^\perp \).

Theorem 4.4.2.

\[
(4.4.2) \quad \|K^*K(\tilde{f}_0 - \tilde{f}_\lambda)\| \leq \|R\| + \lambda \|\tilde{f}_\lambda\|, \\
(4.4.3) \quad \|\tilde{f}_0 - \tilde{f}_\lambda\| \leq \frac{\|R\| + \lambda \|\tilde{f}_\lambda\|}{\gamma} \quad \text{for} \quad \gamma \neq 0, \\
(4.4.4) \quad \|\tilde{f}_0 - \tilde{f}_\lambda\| \leq \frac{\|R\| + \lambda \|\tilde{f}_\lambda\|}{\lambda + \gamma}
\]

and

\[
(4.4.5) \quad \|\tilde{f}_0 - \tilde{f}_\lambda\| \leq \frac{(\lambda + \gamma)\|R\| + \lambda \|K^*g\|}{\gamma(\lambda + \gamma)} \quad \text{for} \quad \gamma \neq 0.
\]

Proof: For (4.4.2): Since \( K^*Kf = K^*g + R \) and
\( K^*Kf = K^*g + R \), we have \( K^*Kf = K^*g + R \).
Thus \( K^*K(\tilde{f}_0 - \tilde{f}_\lambda) = R \). Taking the norm of each side and using the triangular inequality we have
\[\|K^*K(\tilde{f}_o - \tilde{f}_\lambda)\| \leq \|R_o\| + \lambda \|\tilde{f}_\lambda\|.\]

**For (4.4.3):** Since \(\gamma\|\tilde{f}_o - \tilde{f}_\lambda\| \leq \|K^*K(\tilde{f}_o - \tilde{f}_\lambda)\|,\) the result follows from (4.4.2).

**For (4.4.4):** Since \(K^*K(\tilde{f}_o - \tilde{f}_\lambda) = R_o + \lambda \tilde{f}_\lambda\) we have \((K^*K + \lambda I)(\tilde{f}_o - \tilde{f}_\lambda) = R_o + \lambda \tilde{f}_o\). Taking norms of each side of the above and using the triangular inequality yields \(\|(K^*K + \lambda I)(\tilde{f}_o - \tilde{f}_\lambda)\| \leq \|R_o\| + \lambda \|\tilde{f}_o\|\). Since \((\lambda + \gamma)\|\tilde{f}_o - \tilde{f}_\lambda\| \leq \|(K^*K + \lambda I)(\tilde{f}_o - \tilde{f}_\lambda)\|\) (4.4.4) follows.

**For (4.4.5):** Since \(\|(K^*K + \lambda I)\tilde{f}_\lambda\| = \|K^*g\|\) and 
\[(\lambda + \gamma)\|\tilde{f}_\lambda\| \leq \|(K^*K + \lambda I)\tilde{f}_\lambda\|\) we have \(\|\tilde{f}_\lambda\| \leq \frac{\|K^*g\|}{\lambda + \gamma}\). Thus from (4.4.3) the result follows.

Now take \(\tilde{f}_{j,n}\) as a term in the sequence (2.2.8) converging to the least squares solution of minimum norm, \(\tilde{f}_o\), of \(K^*K\tilde{f} = K^*g + R_o\). The following error bounds on \(\|\tilde{f}_o - \tilde{f}_{j,n}\|\) can be obtained.
Theorem 4.4.3.

\begin{align}
\|\bar{x}_o - \bar{x}_{j,n}\| & \leq \frac{\|R_o\| + \|K*K\hat{\bar{x}}_{j,n} - K*g\|}{\gamma} \quad \text{if } \gamma \neq 0, \\
\|\bar{x}_o - \bar{x}_{j,n}\| & \leq \frac{\|W_{j,n}\|}{\lambda_j + \gamma} + \frac{\|R_o\| + \lambda_j \|\bar{x}_o\|}{\gamma} \quad \text{if } \gamma \neq 0,
\end{align}

(4.4.8)

\begin{align}
\|\bar{x}_o - \bar{x}_{j,n}\| & \leq \frac{\|W_{j,n}\|}{\lambda_j + \gamma} + \frac{\|R_o\| + \lambda_j \|\bar{x}_o\|}{\gamma(\lambda_j + \gamma)} \quad \text{if } \gamma \neq 0.
\end{align}

Proof: For (4.4.6): Since $K*K\bar{x}_o = K*g + R_o$, we have $K*K(\bar{x}_o - \bar{x}_{j,n}) = K*g - K*K\bar{x}_{j,n} + \bar{R}_o$. Take the norm of each side of the above result and use the triangle inequality, yielding $\|K*K(\bar{x}_o - \bar{x}_{j,n})\| \leq \|\bar{R}_o\| + \|K*g - K*K\bar{x}_{j,n}\|$. The result (4.4.6) follows from $\gamma\|\bar{x}_o - \bar{x}_{j,n}\| \leq \|K*K(\bar{x}_o - \bar{x}_{j,n})\|$. 
For (4.4.7), (4.4.8), (4.4.9): These bounds on \( \| \tilde{f}_o - \tilde{f}_j, n \| \)
follow from the triangle inequality:
\[
\| \tilde{f}_o - \tilde{f}_j, n \| \leq \| \tilde{f}_j, n - \tilde{f}_j \| + \| \tilde{f}_j - \tilde{f}_o \|,
\]
Theorem 4.4.2 and
\[
(\lambda_j + \gamma) \| \tilde{f}_j, n - \tilde{f}_j \| \leq \| (K^*K + \lambda_j I) (\tilde{f}_j, n - \tilde{f}_j) \| = \| w_{j, n} \|.
\]

### 4.5. Perturbations of \( g \)

In the case where \( g \) is not known exactly, we have the integral equation

\[
(4.5.1) \quad \int_0^1 k(y, x)f(x)dx = g(y) + \varepsilon(y),
\]
where \( \varepsilon(y) \in L_2[0,1] \). (4.4.10) may or may not have a solution (see 2.3). We follow the procedure and notation of section 4.3 and obtain the matrix representation for (4.5.1). Thus

\[
(4.5.2) \quad K^*K\tilde{f} + \lambda \tilde{f} = K^*\tilde{g} + K^*\tilde{\varepsilon} + \tilde{R}_2
\]
where \( \tilde{R}_2 \) is the discretization error vector. The system of equations actually solved is
\[ (4.5.3) \quad K^*K\vec{f} + \lambda\vec{f} = K^*\vec{g} + K^*\vec{e}. \]

Let \( \vec{f}_{\lambda,\varepsilon} \) be the solution to (4.5.3).

\[ \| \vec{f}_{\lambda,\varepsilon} - \vec{f}_{\lambda,\varepsilon} \| \leq \frac{\| K\| \| \vec{e} \|}{\gamma + \lambda} \quad \text{from Theorem 2.3.1 and statement (4.3.10).} \]

Let \( \vec{f}_{j,n,\varepsilon} \) be a term in the sequence defined by (2.2.8). If a bound for \( \| K^*\vec{e} \| \) or \( \| \vec{e} \| \) can be found, we can obtain bounds on \( \| \vec{f}_o - \vec{f}_{j,\varepsilon} \| \) and \( \| \vec{f}_o - \vec{f}_{j,n,\varepsilon} \| \).

Theorems 4.4.2 and 4.4.3 can be extended, with similar proof, to the following theorems.

**Theorem 4.5.1.**

\[ (4.5.3) \quad \| K^*K(\vec{f}_o - \vec{f}_{\lambda,\varepsilon}) \| \leq \| \vec{R}_o \| + \| K^*\vec{e} \| + \lambda \| \vec{f}_{\lambda,\varepsilon} \| \]

\[ (4.5.4) \quad \| \vec{f}_o - \vec{f}_{\lambda,\varepsilon} \| \leq \frac{\| \vec{R}_o \| + \| K^*\vec{e} \| + \lambda \| \vec{f}_{\lambda,\varepsilon} \|}{\gamma} \quad \text{if } \gamma \neq 0 \]

\[ (4.5.6) \quad \| \vec{f}_o - \vec{f}_{\lambda,\varepsilon} \| \leq \frac{\| \vec{R}_o \| + \| K^*\vec{e} \| + \lambda \| \vec{f}_o \|}{\lambda + \gamma} \]
Theorem 4.5.2.

\[ \left\| \tilde{x}_0 - \tilde{x}_{j,n,\epsilon} \right\| \leq \frac{\left(\lambda + \gamma\right)\left\| \tilde{R}_0 \right\| + \lambda\left\| K^* (\tilde{g} + \tilde{e}) \right\|}{\gamma (\lambda + \gamma)} \] if \( \gamma \neq 0 \).

It should be noted that the error bounds given by

(4.4.6) and (4.5.8) are numerically not as nice as they look. Even for \( \left\| \tilde{R}_0 \right\| = \left\| K^* \tilde{e} \right\| = 0 \) and \( \gamma \) fairly large,
making \( \|K^* K_{i,n}^* - K^* g\| \) small is not necessarily good measure of convergence for the following reason. The methods of numerical quadrature used to obtain the matrix representation of the integral equation are essentially piecewise polynomial approximations to the integrand. From numerical experimentation, making \( \|K^* K_{j,n}^* - K^* g\| \) as small as possible, will cause oscillations in the final result. This is due to computer convergence. An apparent measure of this effect can be noted in the output column AF-G of the algorithm in Appendix B.

4.6. Integral Equation Examples

For a first example we will consider the integral equation with symmetric kernel \( k(x,y) = x + y \),

\[(4.6.1) \quad Kf(y) = \int_0^1 (x+y)f(x)dx = \frac{1}{3} + \frac{y}{2} \in R(K), \quad y \in [0,1].\]

\( K^* K \) has non-zero eigenvalues

\[
\frac{7}{12} + \frac{\sqrt{3}}{3} = 1.160683 \quad \text{and} \quad \frac{7}{12} - \frac{\sqrt{3}}{3} = .005983.
\]
For (4.6.1) the least squares solution of minimum norm is
\[ f_0(x) = x. \]

The algorithm described in Chapter 2 and listed in Appendix B was used to obtain the following approximate solutions. Simpson's rule was used to determine the quadrature weights. In this example \( \| \tilde{R}_0 \| = \| \tilde{R}_1 \| = 0 \), the norm of the discretization error vectors. The interval \([0,1]\) was partitioned:

\[
\tilde{y} = \tilde{x} = \begin{bmatrix}
0.00 \\
0.25 \\
0.30 \\
0.75 \\
1.00
\end{bmatrix}.
\]

A 5 x 5 matrix representation for the integral equation (4.4.1) gives

\[
K = \begin{bmatrix}
0.0000 & 0.0833 & 0.0833 & 0.2500 & 0.0833 \\
0.0208 & 0.1667 & 0.1250 & 0.3333 & 0.1042 \\
0.0417 & 0.2500 & 0.1667 & 0.4167 & 0.1250 \\
0.0625 & 0.3333 & 0.2083 & 0.5000 & 0.1458 \\
0.0833 & 0.4167 & 0.2500 & 0.5833 & 0.1667
\end{bmatrix}
\]
The non-zero eigenvalues of $K^*K : H_T^5 \rightarrow H_T^5$ were computed to be 1.160677 and 0.005983. We assumed that an estimate for $\|\vec{f}_o\|$ is known, $\|\vec{f}_o\| = 0.577$ from (4.4.8).

$$\|\vec{f}_o - \vec{f}_{j,n}\| \leq \frac{\|W_{j,n}\|}{\lambda_j + 0.005983} + \frac{0.577\lambda_j}{\lambda_j + 0.005983}.$$

$\lambda_j = \text{terminal} \quad \lambda = 1 \times 10^{-9}$ and $\|W_{j,n}\|^2 = 10^{-8}$ are determined so that $\|\vec{f}_o - \vec{f}_{j,n}\| < 10^{-6}$. We used $M = 600$, maximum iterations per change in $\lambda$, $\lambda_{i+1} = 0.0001\lambda_i$, and initial starting vector $(f_i)_{i=1,5}' f_i = 1$, $i = 1, 5$. The algorithm computed an initial starting $\lambda, \lambda_1 = 0.133889238$. After 225 iterations, we found the approximate least squares solution of minimum norm to be $f_{j,n}$. 

Here is the matrix $K^*K$:

$$K^*K = \begin{bmatrix}
0.0278 & 0.1528 & 0.0972 & 0.2361 & 0.0694 \\
0.0382 & 0.2153 & 0.1389 & 0.3403 & 0.1007 \\
0.0486 & 0.2778 & 0.1806 & 0.4444 & 0.1319 \\
0.0590 & 0.3403 & 0.2222 & 0.5486 & 0.1632 \\
0.0694 & 0.4028 & 0.2639 & 0.6528 & 0.1944
\end{bmatrix}.$$
\[
\begin{bmatrix}
\bar{f}_j(0) \\
\bar{f}_j(.25) \\
\bar{f}_j(.50) \\
\bar{f}_j(.75) \\
\bar{f}_j(1.0)
\end{bmatrix} =
\begin{bmatrix}
0.000000176 \\
0.250000097 \\
0.500000023 \\
0.749999948 \\
0.999999873
\end{bmatrix}.
\]

Also the following information was computed:

\[
\|\bar{f}_{j,n} - \bar{f}_o\| < 9 \times 10^{-8} < 10^{-6},
\]

\[
\|\bar{f}_{j,n} - \bar{f}_o\| \leq \frac{\|K^*K\bar{f}_{j,n} - K^*g\|}{\gamma} < 10^{-6},
\]

and

\[
\max_{i=1,5} |\bar{f}_{j,n}(x_i) - f_o(x_i)| < 1.8 \times 10^{-7}.
\]

The same problem was run a second time using the same initial information, except to terminate when \(\lambda_j = 10^{-15}\) and \(\|W_{j,n}\|^2 < 10^{-28}\) obtaining, after 745 iterations,

\[
\|\bar{f}_{j,n} - \bar{f}_o\| < 8. \times 10^{-13}
\]
and

\[ \max_{i=1,5} \left| \frac{n}{n} f_i(x_i) - f_0(x_i) \right| < 1.6 \times 10^{-12}. \]

As a second example consider the integral equation

(4.6.2)

\[ Kf(y) = \int_0^1 (y-x)^2 f(x) \, dx = \frac{y^2}{2} - \frac{2y}{3} + \frac{1}{4} \in \mathcal{R}(K) \]

which Bellman et al. [4, page 159] used for testing purposes of an algorithm he developed for finding the least squares solution of minimum norm. (4.6.2) has least squares solution of minimum norm \( f_0(x) = x \). The quadrature method Bellman used was Simpson's rule, with 11 equally spaced points, as we will use.

A bound for the norm of the quadrature error was found to be \( \|R_0\| = 0 \), since Simpson's rule integrates \( \int_0^1 (y-x)^2 f(x) \, dx \) for \( f(x) = x \) exactly and the errors for

\[ \int_0^1 (y-x)^2 \int_0^1 (y-x)^2 f(x) \, dx \, dy = \int_0^1 (y-x)^2 g(y) \, dy \]

\[ \int_0^1 (y-x)^2 \int_0^1 (y-x)^2 f(x) \, dx \, dy = \int_0^1 (y-x)^2 g(y) \, dy \]
subtract identically when \( f(x) = x \).

In Bellman's example, also based on a variation of regularization, he used \( \lambda = 10^{-7} \) and required a good initial approximation to the solution. To obtain somewhat comparable results, the algorithm in Appendix B with his \( \lambda \) was used.

The matrix operator \( K^*K : H^1_T \rightarrow H^1_T \) has computed non-zero eigenvalues \( .038101, .027778 \) and \( .000814 \). To compare the results of the algorithm in Appendix B with his we used \( \lambda_j = \text{Terminal Lambda} = 10^{-7} \), control on \( \|(K^*K + \lambda_j I)f - K^*g\|^2 = 10^{-16} \), lambda multiplier = \( 10^{-3} \), maximum steps for each \( \lambda_i = 300 \) and initial starting vector \( (f_i)^{l,11} \) where \( f_i = 0 \) for \( i = 1,10 \), \( f_{11} = 1 \).

The algorithm given in Appendix B choses a new starting vector and gave the results given in Table 2 for Run 1. The bound for \( \|\bar{f_j,n} - \bar{f}_o\| \) was computed using \( \gamma = .000814 \) and \( .577 \) as an estimate for \( \|\bar{f}_o\| \) in (4.4.8). The results were about the same as Bellman obtained.

Now Bellman also looked at the same example (4.6.2), but rounds the data representing \( \bar{g} \) correct to three places. We ran the same example with the same starting
information as for Run 1, to obtain two place accuracy compared to none for Bellman (see Table 2, Run 2). For Run 2, the bound for \( \| \tilde{f}_{j,n} - \tilde{f}_o \| \) was computed using \( \gamma = 0.000814, 0.577 \) as an estimate for \( \| \tilde{f}_o \| \) and

\[
\| K^* \| \| \tilde{e} \| \leq \| K^* \| \| e \| \approx \sqrt{0.038101 \cdot 0.0005} = 0.0001
\]

in (4.5.10).

For a third run of a discretized version of (4.6.2) the 11-point trapezoidal rule was used as the quadrature method. In this case, the matrix operator \( K^*K \) has computed non-zero eigenvalues \( 0.040182, 0.028900 \) and \( 0.000926 \). An upper bound on the norm of the discretization error vector was computed to be \( \| \tilde{e} \| \leq 0.056 \). To compare with Run 1, the same starting information was used and the computed error determined by using (4.4.8).

<table>
<thead>
<tr>
<th>Table 2. Integral Equation Example 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Run 1</strong></td>
</tr>
<tr>
<td>Total iterations \textbf{339}</td>
</tr>
<tr>
<td>( |	ilde{f}_{j,n} - \tilde{f}_o | ) \textbf{2.7(-5)}</td>
</tr>
<tr>
<td>( \max_i</td>
</tr>
<tr>
<td>Computed error \textbf{1.(-4)}</td>
</tr>
</tbody>
</table>

| **Run 2**                           |
| Total iterations \textbf{339}       |
| \( \|	ilde{f}_{j,n} - \tilde{f}_o \| \) \textbf{1.9(-3)} |
| \( \max_i | \tilde{f}_{j,n} - f_o | \) \textbf{3.(-3)} |
| Computed error \textbf{0.11}       |

| **Bellman**                         |
| Total iterations \textbf{341}       |
| \( \|	ilde{f}_{j,n} - \tilde{f}_o \| \) \textbf{1.2(-2)} |
| \( \max_i | \tilde{f}_{j,n} - f_o | \) \textbf{0.9}   |
| Computed error \textbf{61.}        |

| **Run 3**                           |
| Total iterations \textbf{341}       |
| \( \|	ilde{f}_{j,n} - \tilde{f}_o \| \) \textbf{1.2(-2)} |
| \( \max_i | \tilde{f}_{j,n} - f_o | \) \textbf{3.1(-2)} |
| Computed error \textbf{61.}        |
For a last example we will consider the integral equation of the first kind with kernel

\[
k(x,y) = \begin{cases} 
(1-y)x & \text{if } 0 \leq x \leq y \leq 1 \\
(1-x)y & \text{if } 0 \leq y \leq x \leq 1.
\end{cases}
\]

Tricomi [24, page 116] notes that the eigenvalues associated with

\[
K^*K(f)(x) = \int_0^1 k(x,y) \int_0^1 k(x,y)f(x) \, dx \, dy = \gamma f(x)
\]

are 
\[\gamma_1 = \frac{1}{\pi}, \quad \gamma_2 = \frac{1}{(2\pi)^2}, \ldots, \gamma_n = \frac{1}{(n\pi)^4}, \ldots.\]

Strand [20, page 71] shows that for 
\[g(y) = y(3-5y^2 + 3y^4 - y^5)/30 \in R(K),\]

\[Kf(y) = \int_0^1 k(x,y)f(x) \, dx = g(x)\]

has least squares solution of minimum norm \[f_0(x) = x - 2x^3 + x^4.\]

For the discretization of the integral equation

\[
\int_0^1 k(x,y) \int_0^1 k(x,y)f(x) \, dx \, dy = \int_0^1 k(x,y)g(y) \, dy
\]
the error for a 51 point Simpson's rule was estimated to be (upper bound) $1.4 \times 10^{-7}$.

The matrix representation for $K^*K$ has minimum positive non-zero eigenvalue $\gamma = 7.915 \times 10^{-9}$ (computed). Since $\gamma$ is small a choice for the terminal $\lambda_j$, from the error bounds is not practical (see Theorem 4.4.3).

Considering the expression

$$\|\tilde{x}_j - \tilde{x}_o\| \leq \frac{\lambda_j \|\tilde{x}_j\| + \|\tilde{R}_o\|}{\gamma},$$

numerical experimentation indicates that a reasonable choice for terminal $\lambda_j$ is for the order of magnitude of $\lambda_j \|\tilde{x}_j\|$ and $\|\tilde{R}_o\|$ be about the same. Assuming $\|\tilde{x}_j\| = .22 \approx \|\tilde{x}_o\|$ we get a choice for terminal $\lambda_j$, $\lambda_j = 6.4 \times 10^{-7}$. Using $\|K^*K\tilde{x}_j, n + \lambda\tilde{x}_j, n - K^*g\|^2 = 10^{-16}$ we get a computed error bound: $\|\tilde{x}_j, n - \tilde{x}_o\| < .46$ from (4.4.8). We allowed a maximum of 500 iterations per step and set the lambda multiplier = .001. The algorithm in Appendix B generated its own starting vector from the initial starting vector.
\[
\frac{\hat{f}(x_i)}{\hat{f}(x_i)} = \begin{cases} 
1 & \text{if } 17 \leq i \leq 35 \\
0 & \text{otherwise}
\end{cases}
\]

and initial \( \lambda_1 \approx .00795 \). After 52 iterations we obtained actual results:

\[
\|\frac{\hat{f}}{\hat{f}} (x_{j,n} - \hat{f}_0) \| < 1.332 \times 10^{-4}
\]

\[
\max_{i=1,52} |f_{j,n}(x_i) - f_0(x_i)| < 3.024 \times 10^{-4},
\]

and

\[
\|K^*(K_{j,n} - \hat{f}_0)\| < 7.121 \times 10^{-6}.
\]
5. SUMMARY AND FUTURE RESEARCH

5.1. Summary

The advantage of the iterative method developed in the paper is primarily that it avoids the calculation of $(K^*K + \lambda I)^{-1}$ directly and that the iterative process is stable. The unwanted oscillations in the final solution that often appear in solutions by other methods in the literature do not occur. In some applications the error bounds obtained become large numerically when an upper bound on $\|R^*_0\|$ or $\|\epsilon\|$ is large or when $\gamma$ is near zero. An upper bound for $\|R^*_0\|$ can be difficult to obtain unless some information about the solution is known.

When comparing with other iterative methods in the literature, we generally have at least two additional significate figures. The iterative method here converges for examples where other methods in the literature fail to converge to a solution.
5.2. Future Research

It would be desirable to modify the algorithms in Appendix A and B so that \( \gamma \) and the parameters for convergence are computed within the algorithm.

A method of discretization for integral equations with non-uniform mesh size should be developed that better describes the properties of a particular kernel and \( g \).

It would be nice to find a method of determining the weights for quadrature in representing the kernel, independent of any information about the form of the solution.
6. BIBLIOGRAPHY


I would like to express my thanks to Dr. R. J. Lambert for his patience, understanding and unfailing encouragement during my years at Iowa State as well as his guidance and assistance during this research work. I would like to thank Dr. G. R. Luecke for his guidance during this research and his perseverance during the final writing of this paper. I would like to thank my wife and growing family for their understanding during the many years of being a graduate student. I would like to extend my appreciation to Mrs. Madonna Blanshan for her excellent typing.
8. APPENDIX

8.1. Appendix A, Linear System

The following program is written to find the least squares solution of minimum norm of an \( NN \) by \( N \) system of equations, \( Ax = g \). The basic flow chart for the program is given in Figure 3.2.1.

The user must input the following information in the order given:

1. \( NN \) = the number of rows of the matrix \( A \). The format for reading \( NN \) is given in line \( SYST0280 \).

2. \( N \) = the number of columns of the matrix \( A \). The format for reading \( N \) is given in line \( SYST0280 \).

3. \( M \) = the maximum number of iterations per change in lambda. The format for reading \( M \) is given in line \( SYST0280 \).

4. \( BB \) = terminal lambda. The format for reading \( M \) is given in line \( SYST0320 \).

5. \( DD \) = control on \( \| A^T AF + \Lambda F - ATG \| ^2 \). The format for reading \( DD \) is given in line \( SYST0320 \).
6. The matrix A is read by rows. The format for reading A is given in line SYST0340.

7. Read the vector g. The format for reading g is given in line SYST0340.

8. Read the initial starting vector F. The format is given in line SYST0340.

9. XLAM = lambda multiplier. The format for reading XLAM is given in line SYST0340.

The output of the starting information begins at line SYST1440. The output of the final solution begins at line SYST2370.
THE FOLLOWING IS A PROGRAM TO SOLVE AN NN BY N SYSTEM OF
EQUATIONS, AF=G, FOR THE LEAST SQUARES SOLUTION OF MINIMUM
NORM. IF N OR NN IS GREATER THAN 10, THE DIMENSION
STATEMENT MUST BE CHANGED TO AT LEAST A(NN,N), G(NN), F(N),
AT(N,NN), ATA(N,N), R(NN), T(N), V(NN), W(N), S(NN), AND WW(N).

INPUT:

N= THE NUMBER OF COLUMNS OF THE MATRIX A.

NN= THE NUMBER OF ROWS OF THE MATRIX A.

W= THE MAXIMUM NUMBER OF ITERATIONS PER CHANGE IN
   LAMBDA.

BB= TERMINAL LAMBDA.

DD= CONTROL ON L2N (ATAF+LAMBDA*F-AT G)** 2.

XLAM= LAMBDA MULTIPLIER.

IMPLICIT REAL*(A-H,O-Z)
DIMENSION A(10,10), G(10), F(10), ATA(10,10), R(10)
DIMENSION T(10), V(10), W(10), S(10), WW(10)
READ(5,11) NN
READ(5,11) N
11 FORMAT(I5)
READ(5,11) M
READ(5,21) BB
READ(5,21) DD
21 FORMAT(D10,5)
31 FORMAT(8F10,5)
DO 10 I=1,NN
C
C READ MATRIX A BY ROWS.
C
10 READ(5,31) (A(I,J),J=1,N)

C

READ VECTOR G OF AF=G

C

READ(5,31) (G(I),I=1,NN)

C

READ INITIAL STARTING VECTOR F

C

READ(5,31) (F(I),I=1,N)

READ(5,31) XLAM

WRITE(6,12) NN,N,M

12 FORMAT(*0*",*NN=",I5,*5X,*N=",I5,*5X,*M=",I5)

WRITE(6,26) DD

26 FORMAT(*0"*CONTROL ON \| ATAF+LAMBDA*F-ATG \|**2=",D24,16)

WRITE(6,22) DD

22 FORMAT(*0"*CONTROL ON LAMBDA=",D24,16)

WRITE(6,24) XLAM

24 FORMAT(*0"*LAMBDA MULTIPLIER=",D24,16)

WRITE(6,32)

32 FORMAT(*0"*MATRIX A="/

DO 20 I=1,NN

20 WRITE(6,42) (A(I,J),J=1,N)

42 FORMAT(*0",10(F10.5,1X))

WRITE(6,52)

52 FORMAT(*0"*G="/

WRITE(6,42) (G(I),I=1,NN)

WRITE(6,62)

62 FORMAT(*0"*STARTING VECTOR F="/

WRITE(6,42) (F(I),I=1,N)

LL=0

DO 30 I=1,N

30 AT(I,J)=A(J,I)

DO 410 I=1,N

W(I)=0.00D0

DO 410 J=1,NN

410 W(I)=W(I)+AT(I,J)*G(J)
ZZ=0.000
DO 420 I=1,N
420 ZZ=ZZ+U(I)*U(I)
ZZ=DSQRT(ZZ)
WRITE(6,29) ZZ
29 FORMAT('0*E0D0 ATG | | =*E0D015) IF(ZZ*GE1.0D40) GO TO 430
WRITE(6,421)
421 FORMAT('0*6X' FORM AT '0*6X' SOLUTION IS THE ZERO VECTOR')
GO TO 600

C OBTAIN A STARTING VECTOR AND AN INITIAL LAMBDA
C

DO 40 I=1,N
DO 40 J=1,N
ATA(I,J)=0.000
DO 40 L=1,N
40 ATA(I,J)=ATA(I,J)+AT(I,L)*A(L,J)
DO 50 I=1,N
50 T(I)=0.000
DO 50 J=1,N
50 T(I)=T(I)+ATA(I,J)*F(J)
ZZ=0.000
DO 60 I=1,N
60 ZZ=ZZ+T(I)*T(I)
ZZ=DSQRT(ZZ)
DO 70 I=1,N
70 T(I)=T(I)/ZZ
DO 80 I=1,N
80 R(I)=0.000
DO 80 J=1,N
80 R(I)=R(I)+A(I,J)*T(J)
CR=1.000
L=2.000
J=1
90 X=0.000
DO 100 I=1,N
100 CR=1.000
L=2.000
J=1

100 X=X+(G(I)-R(I))*R(I)
IF(X\$GEo0:000) GO TO 1000
IF(J\$EQo2) GO TO 1010
DO 110 I=1,N
110 T(I)=-T(I)
DO 115 I=1,N
115 R(I)=-R(I)
GO TO 90
1010 DO 120 I=1,N
120 G(I)=U*G(I)
J=1
CR=CR*U
GO TO 90
1000 DO 130 I=1,N
130 R(I)=R(I)-G(I)
RR=0:000
DO 140 I=1,N
140 RR=RR+R(I)*R(I)
DO 150 I=1,N
W(I)=0:000
DO 150 J=1,N
150 W(I)=W(I)+AT(I,J)*R(J)
ZZ=0:000
DO 160 I=1,N
160 ZZ=ZZ+W(I)*W(I)
YY=ZZ-XX
AB=RR+XX
DO 170 I=1,N
170 W(I)=W(I)+XX*T(I)
C
C OUTPUT STARTING INFORMATION
C
WRITE(6,72) X
72 FORMAT(*0.***INITIAL LAMBDA=*:024.16)
WRITE(6,82)
82 FORMAT(*0.***STARTING VECTOR F*:10X,*ATAF+LAMBDA*F-ATG*:/)
180 WRITE(6,92) T(I),W(I)  
92 FORMAT(' ',3D24.16)  
96 FORMAT(' ',D24.16)  
WRITE(6,84)  
84 FORMAT(' ',I5,' AF=',D24.16)  
DO 132 I=1,NN  
182 WRITE(6,96) R(I)  
WRITE(6,94) CR  
94 FORMAT(' ',SCALING FACTOR=',D24.16)  
WRITE(6,102) YY  
102 FORMAT(' ',MINIMUM VALUE=',D24.16)  
XX=X  
C MAIN ITERATIONS OF THE ALGORITHM BEGIN  
C 1020 X=XX  
K=1  
1030 AA=AB  
YY=ZZ  
DO 190 I=1,NN  
WW(I)=W(I)  
W(I)=0.0D0  
DO 190 J=1,NN  
190 W(I)=W(I)+AT(I,J)*R(J)  
DO 200 I=1,NN  
200 W(I)=W(I)+X*T(I)  
ZZ=0.0D0  
DO 210 I=1,NN  
210 ZZ=ZZ+W(I)*W(I)  
IF(ZZ.LT.0.D0) GC TO 1040  
DO 220 I=1,NN  
220 V(I)=R(I)  
DO 225 I=1,NN  
225 F(I)=T(I)  
VV=RR  
DO 230 I=1,NN  
S(I)=0.0D0
DO 230 J=1,N
230 S(I)=S(I)+A(I,J)*W(J)
    Y=0.000
    DO 240 I=1,NN
    240 Y=Y+S(I)*S(I)
    Y=Y+X*ZZ
    E=ZZ/Y
    DO 250 I=1,N
250 T(I)=T(I)-S*W(I)
    DO 260 I=1,NN
    260 R(I)=R(I)+A(I,J)*T(J)
    DO 270 I=1,NN
270 R(I)=R(I)-G(I)
    RR=0.000
    DO 280 I=1,NN
    280 RR=RR+R(I)*R(I)
    AB=0.000
    DO 290 I=1,N
290 AB=AB+T(I)*T(I)
    AB=RR+X*AB
    IF(K0EQ01) GO TO 300
    IF(ABLT0AA) GO TO 300
    DO 310 I=1,N
310 T(I)=F(I)
    DO 295 I=1,NN
295 W(I)=W(I)
    DO 305 I=1,N
305 R(I)=V(I)
    RR=VV
    ZZ=YY
    K=K-1
    GO TO 330
300 IF(M-K) 600,330,340
340 K=K+1
    GO TO 1030
1040 K=K-1
330 LL=LL+K
   IF(K* LT*M) GO TO 440
   WRITE(6,112) X
   112 FORMAT('0°,3X,'AT LAMBDA='D24.16,3X,'THE MAXIMUM VALUE OF M HAS BEEN ATTAINED')
440 IF(XoEQoBB) GO TO 360
   XX=XLAM*X
   IF(XKoLToBB) GO TO 350
   GO TO 1020
350 XX=BB
   GO TO 1020
C
C RE-SCALE THE COMPUTATIONS AND OUTPUT
C
360 WRITE(6,122)
   122 FORMAT('0°,3X,'FINAL SOLUTION F*,10X,'ATAF+LAMBDA*F-ATG*/)
   DO 370 I=1,N
      T(I)=T(I)/CR
      W(I)=W(I)/CR
   370 WRITE(6,92) T(I),W(I)
   DO 375 I=1,NN
      R(I)=R(I)/CR
   375 WRITE(6,96) R(I)
   ZZ=ZZ/(CR*CR)
   RR=RR/(CR*CR)
   TT=DSQRT(ZZ)
   TT=DSQRT(TT)
   WRITE(6,132)
   132 FORMAT('0°,8X,'||F||,12X,'||ATAF+LAMBDA*F-ATG||,10X,'C
   C*AF-G||,9X)
   WRITE(6,92) TT,ZZ,RR
   DO 390 I=1,N
   390 TT=TT+T(I)*T(I)
   WRITE(6,132)
V(I) = 0.0D0
DO 390 J = 1, NN
390 V(I) = V(I) + AT(I, J) * R(J)
VVV = 0.0D0
DO 400 I = 1, N
400 VVV = VVV + V(I) * V(I)
VVV = DSGRT(VVV)
WRITE(6, 144) VVV
144 FORMAT('0', 8X, 'AT(ADF-G) ' = *D24.16)
WRITE(6, 142) LL
142 FORMAT('C', 8X, 'TOTAL ITERATIONS = ', I6)
C
C IF THE ACTUAL SOLUTION IS KNOWN, A PROGRAM TO COMPARE THE
C ACTUAL SOLUTION WITH THE COMPUTED SOLUTION CAN BE PLACED
C HERE. THE COMPUTED SOLUTION IS GIVEN BY THE VECTOR T. THE
C VECTOR F CAN BE ASSIGNED THE ACTUAL SOLUTION.
C
600 STCP
END
8.2. Appendix B, Subroutine BITER

Subroutine BITER is a program written to find a \( \lambda \)-approximate least squares solution of minimum norm of an integral equation of the first kind. The subroutine BITER assumes that the integral equation is of the form

\[
\int_0^1 k(y,x)f(x)\,dx = g(y), \quad y \in [0,1].
\]

A driver program must be written which obtains the following information:

1. Partitions the interval integrated over, \([0,1]\), into \(N-1\) equal subintervals for an \(N\)-point closed Newton-Cotes integration formula. The \(N\)-points constructed are stored in the vector \(XV\).

2. Partitions the domain of \(g\), \([0,1]\), into \(NN-1\) equal subintervals for an \(NN\)-point closed Newton-Cotes integration formula. The \(NN\)-points constructed are stored in the vector \(YV\).

3. Construct the vector \(GV(I) = g(YV(I))\) \(I = 1,NN\).
4. Construct the matrix representing the kernel

\[ k(x,y), A(k(XV(I),YV(j))\frac{1}{j=1,NN} \frac{1}{I=1,N} \]

5. Construct the vector \( WTK \) of weights for the kernel using a closed \( N \)-point Newton-Cotes formula.

6. Construct the vector \( WTKT \) of weights for integration of the adjoint of the kernel using a closed \( NN \)-point Newton-Cotes formula.

7. The driver program must be dimensioned as the dimension statement for the subroutine \( BITER \).

8. The driver program must supply the controls on convergence given by

\[ XIMDA = \text{Lambda multiplier}, \]
\[ BB = \text{Terminal lambda}, \]

and

\[ DD = \text{control on } \|ATAF + LAMBDA*F - ATG\|^2. \]

The output section of the starting vector and initial lambda begins at line \( BITR1080 \). The output section for the final results begins at line \( BITR1990 \).
CONSTRUCT THE ADJOINT=AT TO KERNEL=A*

TRANSPOSE THE UNWEIGHTED KERNEL

\[
\begin{align*}
\text{DO 510 } & I=1,NN \\
\text{DO 510 } & J=1,N \\
510 & \text{AT}(J,I)=A(I,J)
\end{align*}
\]

CONSTRUCT THE WEIGHTED KERNEL

\[
\begin{align*}
\text{DO 520 } & I=1,NN \\
\text{DO 520 } & J=1,N \\
520 & A(I,J)=A(I,J)*WTK(J)
\end{align*}
\]

CONSTRUCT THE WEIGHTED KERNEL TRANSPOSE

\[
\begin{align*}
\text{DO 530 } & I=1,N \\
\text{DO 530 } & J=1,NN \\
530 & \text{AT}(I,J)=\text{AT}(I,J)*WTKT(J)
\end{align*}
\]

TEST FOR A ZERO SOLUTION

\[
\begin{align*}
\text{DO 810 } & I=1,N \\
W(I)=0.000 \\
\text{DO 810 } & J=1,NN \\
810 & W(I)=W(I)+AT(I,J)*GV(J) \\
\text{ZZ}=0.000 \\
\text{DO 820 } & I=1,N \\
820 & \text{ZZ}=\text{ZZ}+WTK(I)*(W(I)*W(I))
\end{align*}
\]
ZZ=USCR(T(ZZ))
WRITE(6,29) ZZ
29 FORMAT(*0.9,|=ATG|=@D24:16)
IF(ZZ@GE=1D-40) GO TO 830
WRITE(6,821)
821 FORMAT(*0.9,6X.*SOLUTION IS THE ZERO FUNCTION*./)
GO TO 620
C
C CONSTRUCT THE WEIGHTED KERNEL TRANSPOSE*WEIGHTED KERNEL*
C
830 DO 540 I=1,N
 DO 540 J=1,N
 AT(A(I,J))=0.0D0
 DO 540 L=1,N
 540 AT(A(I,J))=AT(A(I,J))+AT(A(L,J))*A(L,J)
C
C OBTAIN AN ACTUAL STARTING VECTOR AND INITIAL LAMBDA*
C
 DO 20 I=1,N
 T(I)=0.0D0
 DO 20 J=1,N
 20 T(I)=T(I)+AT(A(I,J))*F(J)
 ZZ=0.0D0
 DO 30 I=1,N
 30 ZZ=ZZ+WTK(I)*(T(I)*T(I))
 ZZ=DSGRT(ZZ)
 DO 40 I=1,N
 40 T(I)=T(I)/ZZ
 DO 50 I=1,NN
 R(I)=0.0D0
 DO 50 J=1,N
 50 R(I)=R(I)+A(I,J)*T(J)
 CR=1.0D0
 U=2.0D0
 J=1
 52 X=0.0D0
 DO 60 I=1,NN
...
60 X=X+(GV(I)*R(I))(*(R(I)*WT(I))
   IF(X*GE*0*000) GO TO 1000
   IF(J*EQ*2) GO TO 1001
   J=J+1
   DO 70 I=1,N
   70 T(I)=T(I)
   DO 80 I=1,NN
   80 R(I)=R(I)
   GO TO 92
1001 DO 450 I=1,NN
450 GV(I)=U*GV(I)
   J=1
   CR=CR*U
   GO TO 92
1000 DO 90 I=1,NN
90 R(I)=R(I)-GV(I)
   RR=0.0000
   DO 100 I=1,NN
100 RR=RR+WT(I)*(R(I)*R(I))
   DO 110 I=1,N
   110 W(I)=W(I)+AT(I,J)*R(J)
   ZZ=0.0000
   DO 120 I=1,NN
120 ZZ=ZZ+WT(I)*(W(I)*W(I))
   YY=ZZ-X*X
   AB=RR+X
   DO 130 I=1,NN
130 W(I)=W(I)+X*T(I)

C
C OUTPUT SECTION FOR THE STARTING INFORMATION
C
WRITE(6,2) X
2 FORMAT("0", "INITIAL LAMBDA=", D24.16)
WRITE(6,3)
3 FORMAT("0", 12X, "X=", 18X, "STARTING=", 12X, "ATAF+LAMBDA=F-ATG="/)
DO 140 I=1,N
140 WRITE(6,4) XV(I),T(I),W(I)
   WRITE(6,5)
   WRITE(6,4)
   WRITE(6,5)
   WRITE(6,4)
   WRITE(6,5)
DO 150 I=1,N
150 WRITE(6,6) YV(I),R(I)
   WRITE(6,7) CR
   WRITE(6,7) CR
   WRITE(6,7) CR
   WRITE(6,7) CR
   DO 160 1=1,N
   DO 160 J=1,N
   DO 170 1=1,N
   DO 170 1=1,N
   DO 180 1=1,N
   DO 180 1=1,N
   DO 190 1=1,N
   DO 190 1=1,N
   DO 200 1=1,N
   DO 200 1=1,N
   DO 210 1=1,N
   DO 210 1=1,N

C THE MAIN ITERSATIONS OF THE ALGORITHM
C
171 X=XX
   K=1
41 AA=AB
   YY=ZZ
   DO 160 I=1,N
   DO 160 J=1,N
   DO 170 I=1,N
   DO 170 I=1,N
160 W(I)=W(I)+AT(I,J)*R(J)
   W(I)=0.000
   DO 160 J=1,N
170 W(I)=W(I)+X*T(I)
   ZZ=0.000
   DO 180 I=1,N
180 ZZ=ZZ+WTK(I)*W(I)
   IF(ZZ<LT.0)<GO TO 403
   DO 190 I=1,N
190 V(I)=R(I)
   DO 200 I=1,N
200 F(I)=T(I)
   VV=RR
   DO 210 I=1,N
$S(I) = 0, 0 \text{ DO 210 } J = 1, N$

$210 $S(I) = S(I) + A(I, J) * W(J)$

$Y = 0, 0 \text{ DO 220 } I = 1, \text{ NN}$

$220 Y = Y + WTKT(I) * (S(I) * S(I))$

$B = ZZ / Y$

$\text{ DO 230 } I = 1, N$

$230 T(I) = T(I) - B * W(I)$

$B = ZZ / Y$

$\text{ DO 240 } I = 1, \text{ NN}$

$240 R(I) = 0, 0 \text{ DO 240 } J = 1, N$

$240 R(I) = R(I) + A(I, J) * T(J)$

$\text{ DO 250 } I = 1, \text{ NN}$

$250 R(I) = R(I) - GV(I)$

$RR = 0, 0 \text{ DO 260 } I = 1, \text{ NN}$

$260 RR = RR + WTKT(I) * (R(I) * R(I))$

$AB = 0, 0 \text{ DO 270 } I = 1, N$

$270 AB = AB + WIK(I) * (T(I) * T(I))$

$AB = RR + X * AB$

$\text{ IF(KoE} \geq 1) \text{ GO TO 201}$

$\text{ IF(AB} \leq 0 \text{ AA) GO TO 201}$

$\text{ DO 280 } I = 1, \text{ NN}$

$280 T(I) = F(I)$

$W(I) = W(I)$

$\text{ DO 290 } I = 1, \text{ NN}$

$290 R(I) = W(I)$

$RR = VV$

$ZZ = YY$

$K = K - 1$

$\text{ GO TO 401}$

$201 \text{ IF(M} \leq K) 620, 401, 301$

$301 K = K + 1$

$\text{ GO TO 41}$
403 K=K-1
401 LL=LL+K
    IF(K<LT>M) GO TO 440
    WRITE(6,1) X
1  FORMAT('0',3X,'AT LAMBDA=',D24,16,3X,'THE MAXIMUM VALUE OF M HAS
CBEEN ATTAINED')
440 IF(X=EQ=BB) GO TO 600
    XX=XLMDA*X
    IF(XX<LT=BB) GO TO 610
    GO TO 171
610 XX=BB
    GO TO 171
C
C OUTPUT SECTION FOR THE FINAL RESULTS
C
600 WRITE(6,23)
23 FORMAT('0',12X,'X=',14X,'FINAL SOLUTION F=',10X,
     'ATAF+LAMBDA*F-ATG=',1X)
DO 320 I=1,N
    T(I)=T(I)/CR
    W(I)=W(I)/CR
320 WRITE(6,4) XV(I),T(I),W(I)
    WRITE(6,5)
    DO 330 I=1,NN
    R(I)=R(I)/CR
330 WRITE(6,6) YV(I),R(I)
    ZZ=ZZ/(CR*CR)
    RR=RR/(CR*CR)
    ZZ=DSCRT(ZZ)
    RR=DSCRT(RR)
    WRITE(6,26) ZZ,RR
26 FORMAT('0',6X,'ATAF+LAMBDA*F-ATG ||=0',D24,16,5X,'|| AF=G ||=0',
     D24,16)
    WRITE(6,28) LL
28 FORMAT('0',1X,'TOTAL ITERATIONS=',16)
    CC=0.000
    DD 340 I=1,N
340  CC = CC + wTK(I) * (T(I) * T(I))
     CC = DSQRT(CC)
     WRITE(6,9) CC
 9 FORMAT('0', ' NORM OF THE SOLUTION =', 'D24.16')
  DO 390 I = 1, N
     V(I) = 0.0D0
  DO 390 J = 1, NN
 390  V(I) = V(I) + AT(I,J) * R(J)
     VV = 0.0D0
  DO 400 I = 1, N
 400  VV = VV + wTK(I) * (V(I) * V(I))
     VV = DSQRT(VV)
     WRITE(6,144) VV
 144  FORMAT('0', '8X,' ' AT(AF-G) | | =', 'D24.16')
 620 RETURN
END