INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer. .

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI®

Sc

A thesis submitted

by

C.Y. Chan

to

the Faculty of Pure and Applied Science

of the University of Ottawa

in partial fulfillment of the requirements

for the degree of

Master of Science

in the subject of

Mathematics

Université d'Ottawa
BIBLIOTHÈQUES
LIBRARIES
University of Ottawa

UMI Number: EC52077

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI®

UMI Microform EC52077
Copyright 2007 by ProQuest LLC
All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Application of

Lanczos' Minimized Iterations

To

Non-homogeneous Linear Integral Equations

With

Weak Singularities

Acknowledgement

I would like to thank Dr. J.L. Howland for suggesting the topic of the thesis to me, and for his important criticism and assistance. I must also thank Dr. G.F.D. Duff for his valuable help and encouragement.

Abstract

In this thesis, Lanczos' Method of Minimized Iterations is applied in a Hilbert space framework to solve a non-homogeneous linear integral equation of the second kind. The kernel of the integral equation is real, non-symmetric and has a weak singularity of a type frequently occurring in Potential Theory.

In Chapter 1, the given operator is symmetrized. Theorem 2 shows that this symmetrization process does not affect the solution, and Theorem 5 shows that the symmetrized operator is completely continuous and self-adjoint.

In Chapter 2, we use the concept of invariant (closed) subspace and Lanczos' Method of Minimized Iterations to find the approximate solution. Theorem 9 shows that this converges to the required solution faster than any geometrical progression.

Contents

Chapter 1

§ 1. Introduction

Many problems of mechanics, mathematical physics and technology lead to the consideration of non-homogeneous linear integral equations of the second kind with weak singularities with the solutions as elements of the Hilbert space, $L_2(\Omega)$, the space of all equivalence classes of (Lebesgue) square integrable real-valued functions in a bounded open connected set Ω of the n-dimensional Euclidean space [9; p.80]. Fredholm formulae are too complicated for numerical calculation. In 1950, C. Lanczos [8] published his Method of Minimized Iterations for numerical solutions. For a symmetric operator, this method gives rise to a three term recurrence formula in the construction of orthogonal elements; this leads to a formula of simple algebraic form (15). C. Lanczos did not prove the three term recurrence formula; a proof is given here. W. Karush [6; p. 848] and Y.V. Vorobyev [16; p.77] showed that for a completely continuous self-adjoint operator on a Hilbert space, the approximate solution, which is the solution of the corresponding finite-dimensional problem, tends to the original solution of the given non-homogeneous linear integral equation of the second kind faster than any geometrical progression. This makes the method particularly useful for numerical solutions, because the solution of the corresponding problem in a space of finite low dimension may differ from that of the original problem by a very small quantity. These observations motivate the idea of finding a technique of symmetrizing the given non-symmetric operator with a weak singularity such that (i) the symmetrization process does not affect the original solution, and (ii) the resulting operator is completely continuous and self-adjoint.

In this thesis, a special technique due to Erhard Schmidt is used to symmetrize the given non-symmetric operator. It is then shown that the requirements (i) and (ii) are achieved. References for proving the properties of the symmetrized operator can be found in [3; p.123],[9; pp.82-85, p.114], but in proving that the symmetrized operator is completely continuous and self-adjoint, a shorter and entirely independent approach has been used. The concept of invariant (closed) subspace is employed so that Lanczos' Method can be used to construct the approximate solution. The method used here is genuinely constructive as we begin our construction by taking the known element

$$f \equiv g - \mu L^* g$$

as the starting vector (see §5). Y.V. Vorobyev's Method [16; p.77] is then used to prove that the rate of convergence of the approximate solution to the original solution is faster than any geometrical progression.

§ 2. Real Hilbert Space

A space in contemporary mathematics is understood to be a set of arbitrary objects between which there exist relations analogous to some of the spatial relations in three-dimensional space. D.Hilbert took the decisive step by investigating infinite-dimensional space. It was, however, not before 1929 that J. von Neumann [15] gave the first axiomatic treatment of the theory of Hilbert space. Later important contributions are due to M.H. Stone [13].

Definition 1. A real linear space X is called an inner product space if there is defined on XxX a real-valued function, $< x,y >$, called the inner product of x and y with the following properties:

(a) $< x + y,z > = < x,z > + < y,z >$

(b) $< x,y > = < y,x >$

(c) $< \alpha x,y > = \alpha < x,y >$

(d) $< x,x > \geq 0$, and $< x,x > \neq 0$ if $x \neq 0$.

Definition 2. An inner product space X which is infinite-dimensional and complete in the metric determined by the norm

$$||x|| = < x,x >^{\frac{1}{2}}$$

is called a Hilbert space, H.

§ 3. Symmetrization

We shall consider the Hilbert space, $L_2(\Omega)$. Let p and q be two points in Ω, and $y, g \in L_2(\Omega)$. Let us consider the equation

(1) $$y(p) - \mu \int_\Omega L(p,q)\, y(q)\, dq = g(p)$$

where y is the unknown and μ is a given real constant.

Definition 3. (1) is called a non-homogeneous linear integral equation of the second kind.

Let the distance between p and q be denoted by r.

Definition 4. An integral operator of the form

$$\int_\Omega \frac{F(p,q)}{r^\alpha}\, y(q)\, dq$$

where the function $F(p,q)$ is bounded in Ω, and α is a constant such that $o \leq \alpha < n$, is called an integral operator with a weak singularity.

Let $L(p,q)$ be a kernel of weak singularity, i.e. $L(p,q)$ is of the form

$$\frac{F(p,q)}{r^\alpha} .$$

Theorem 1 [9; p.82]. If the function $\phi(p)$ is square integrable in Ω, then the integral

$$\psi(p) \equiv \int_\Omega L(p,q)\, \phi(q)\, dq$$

exists for almost all $p \in \Omega$, and represents a square integrable function in Ω.

Proof: We shall divide the proof in four parts:

(a) The integral

$$\int_\Omega \frac{dp}{r^\alpha}$$

is bounded for any position of the point q inside or on the boundary of Ω. Let $d\omega$ be an element of surface area of the hypersphere S with unit radius. Now let us denote the diameter of Ω by h. Then,

$$\int_\Omega \frac{dp}{r^\alpha} \leq \int_{r<h} \frac{dp}{r^\alpha}$$

$$= \int_S d\omega \int_0^h r^{n-1-\alpha}\, dr$$

$$= \frac{\omega h^{n-\alpha}}{n-\alpha}$$

where

$$\omega = \frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})} .$$

(b) There exists the "double" integral

$$\int_\Omega \int_\Omega \frac{|\phi(q)|^2}{r^\alpha}\, dp\, dq = \int_{\Omega} |\phi(q)|^2\, dq \int_\Omega \frac{dp}{r^\alpha}$$

$$\leq \frac{\omega h^{n-\alpha}}{n-\alpha} ||\phi||^2 .$$

According to the Fubini theorem and a partial converse to it [2:pp.142-143] , there exists a function of the point p almost everywhere in Ω and integrable in Ω given by

$$\int_\Omega \frac{|\phi(q)|^2}{r^\alpha}\, dq \quad .$$

(c) The function $L(p,q)\ \phi(q)$ is integrable for almost all $p \in \Omega$. In fact, if $|F(p,q)| \leq C$, then

$$|L(p,q)\ \phi(q)| \leq C \frac{|\phi(q)|}{r^{\frac{\alpha}{2}}} \frac{1}{r^{\frac{\alpha}{2}}}$$

$$\leq \frac{C}{2}\left\{\frac{|\phi(q)|^2}{r^\alpha} + \frac{1}{r^\alpha}\right\} \quad .$$

By what has been proved in (a) and (b), the product $L(p,q)\ \phi(q)$ is integrable for almost all $p \in \Omega$. Thus, we have proved that $\psi(p)$ exists for almost all $p \in \Omega$.

(d) We observe that

$$|\psi(p)|^2 \leq C^2 \left\{\int_\Omega \frac{|\phi(q)|}{r^{\frac{\alpha}{2}}} \cdot \frac{1}{r^{\frac{\alpha}{2}}}\, dq\right\}^2$$

$$\leq \frac{C^2\ \omega\ h^{n-\alpha}}{n-\alpha} \int_\Omega \frac{|\phi(q)|^2}{r^\alpha}\, dq \ .$$

According to what has been proved in (b), the right hand side of the above inequality, and so also the function $|\psi(p)|^2$ is integrable in Ω, where

(2) $$||\psi|| = \{ \int_\Omega |\psi(p)|^2 \, dp\}^{\frac{1}{2}} \leq \frac{C\omega \; h^{n-\alpha}}{n-\alpha} \, ||\phi|| \quad .$$

Thus, $\psi(p)$ represents a square integrable function in Ω .

<u>Definition 5.</u> Let X and Y be two linear spaces over the same field. Let A be a function with domain X and range contained in Y. Then A is called a <u>linear operator</u> on X into Y if the following two conditions are satisfied:

$$A(x+y) = Ax + Ay$$
$$A(\alpha x) = \alpha Ax$$

where α is an arbitrary scalar, and x,y are arbitrary vectors from X .

In what follows, we shall assume that the kernel is real and non-symmetric. We shall use the notation

$$Ly = \int_\Omega L(p,q) \; y(q) \; dq$$

where L is the operator corresponding to the kernel $L(p,q)$.

(1) can be written into the form

(3) $$y - \mu Ly = g \; .$$

As we are interested in finding the approximate solution for (3), it is natural to assume that (3) has a unique solution, i.e. by using Fredholm alternative, the corresponding equation

$$Ly = \frac{1}{\mu} \, y$$

has no eigenvalue. Let

$$L^*(p,q) = L(q,p),$$

and let L^* be the operator corresponding to $L^*(p,q)$. From (3), we get

(4) $$L^*y - \mu L^*Ly = L^*g .$$

On multiplying (4) by $-\mu$ and adding to (3), we get

$$y - \mu Ly - \mu L^*y + \mu^2 L^*Ly = g - \mu L^*g ,$$

i.e.

(5) $$y - \mu(L + L^* - \mu L^*L)y = g - \mu L^*g .$$

The kernel of (5) is now symmetric. A symmetrizing procedure of this kind has been used, for example, in [5; p.1381] and is also described in [12; pp.459-461].

Theorem 2. (3) and (5) have the same solution.

Proof: Since the unique solution of (3) satisfies (5), it is sufficient to show that (5) has a unique solution. Let

$$z = y - \mu Ly .$$

Then, (5) becomes

(6) $$(I - \mu L^*)z = (I - \mu L^*)g.$$

Thus, if z is a unique solution of (6), it follows, from the assumption that (3) has a unique solution, that y is the unique solution of (5), so it is sufficient to prove that (6) has a unique solution. We shall prove this by contradiction.

It is known that Fredholm theorems hold for linear integral equations with weak singularities [9; p.86]. Assume that (6) has more than one solution. Then by Fredholm alternative, the homogeneous equation,

$$z - \mu L^* z = 0$$

has a non-trivial solution. But this implies, by Fredholm theorem that its transpose

$$z - \mu L z = 0 ,$$

which can be written as

$$y - \mu L y = 0$$

as z is an unknown, has the same number of non-trivial solutions. This contradicts the original assumption that (3) has a unique solution (by Fredholm alternative). Therefore, (3) and (5) have the same unique solution.

§ 4. Properties of the symmetrized operator

In many physical problems, we ~~often~~ encounter kernels of the form

$$\frac{\partial}{\partial n_q} \frac{1}{r^\alpha} .$$

where n_q is the unit normal to the smooth, finite and closed surface [7; p.286], w = 0 , at q, where w is (at least) twice differentiable: it is therefore worthwhile to see whether such differentiation increases the order of the singularity as r tends to 0 . Here p and q are on the surface.

Theorem 3 [3; p.123]. The differentiation along the normal to the surface does not increase the order of the singularity.

Proof: Let $x^1, x^2, \ldots, x^n$ be a system of coordinates with reference to q. r is assumed to be very small. The Taylor expansion for w about q is given by

$$w(p) = w(q) + \sum_i x^i \left(\frac{\partial w}{\partial x^i}\right)_q + \frac{1}{2} \sum_{i,j} x^i x^j \left(\frac{\partial^2 w}{\partial x^i \partial x^j}\right)$$

where the second derivatives are evaluated at a suitable intermediate point between p and q . Since p and q lie on the surface,

$$w = 0 ,$$

therefore,

$$w(p) = 0 = w(q) .$$

Thus,

$$(7) \qquad \sum_i x^i \left(\frac{\partial w}{\partial x^i}\right)_q = -\frac{1}{2} \sum_{i,j} x^i x^j \left(\frac{\partial^2 w}{\partial x^i \partial x^j}\right) .$$

Since

$$dx^i = \left(\frac{dx^i}{dr}\right)_q dr \;,$$

$$x^i = \left(\frac{dx^i}{dr}\right)_q r \;.$$

Thus,

$$r\left(\frac{dw}{dr}\right)_q = r\left(\sum_i \frac{\partial w}{\partial x^i}\frac{dx^i}{dr}\right)_q$$

$$= \sum_i x^i \left(\frac{\partial w}{\partial x^i}\right)_q$$

$$= -\frac{1}{2}\sum_{i,j} x^i x^j \left(\frac{\partial^2 w}{\partial x^i \partial x^j}\right) \qquad \text{by (7)}$$

$$= -\frac{1}{2} r^2 \sum_{i,j}\left(\frac{dx^i}{dr}\right)_q \left(\frac{dx^j}{dr}\right)_q \left(\frac{\partial^2 w}{\partial x^i \partial x^j}\right)$$

$$= O(r^2) \;.$$

Hence,

$$\left(\frac{dw}{dr}\right)_q = O(r) \;.$$

$$\frac{\partial}{\partial n_q}\frac{1}{r^\alpha} = -\frac{\alpha}{r^{\alpha+1}}\frac{\partial r}{\partial n_q}$$

$$= -\frac{\alpha}{r^{\alpha+1}}\left(\frac{\nabla r \cdot \nabla w}{\sqrt{\nabla w \cdot \nabla w}}\right)_q$$

$$= O\left(\frac{1}{r^{\alpha+1}}\right)\left(\frac{dw}{dr}\right)_q$$

$$= O\left(\frac{1}{r^{\alpha+1}}\right) O(r)$$

$$= O\left(\frac{1}{r^{\alpha}}\right) \quad .$$

Thus, the theorem is proved.

Let us consider the operator

(8) $$K \equiv \mu(L + L^* - \mu L^* L) \quad .$$

Since it contains the term $\mu^2 L^*L$, so we shall study the following fundamental theorem in the theory of linear integral equations with weak singularities:

Theorem 4 [9; p.83]. Let

$$|L(p,q)| \le \frac{C_1}{r^{\alpha}} \quad , \qquad |M(p,q)| \le \frac{C_2}{r^{\beta}}$$

where C_1 and C_2 are constants and $o \le \alpha < n$, $o \le \beta < n$. Then the kernel of LM,

$$N(p,q) = \int_{\Omega} L(p,t)\, M(t,q)\, dt$$

has the value

(9) $$|N(p,q)| \le \begin{cases} C & \text{for } \alpha + \beta < n \\ C\,|\log_e r| + B & \text{for } \alpha + \beta = n \\ \dfrac{C}{r^{\alpha+\beta-n}} & \text{for } \alpha + \beta > n \end{cases}$$

where C and B are some constants.

Proof: Let r_o and r_1 be the distances between p,t and q,t respectively. Then,

$$|N(p,q)| \leq C_1\ C_2 \int_\Omega \frac{dt}{r_o^\alpha\ r_1^\beta}$$

$$\leq C_1\ C_2 \int_{r_o<h} \frac{dt}{r_o^\alpha\ r_1^\beta} \quad .$$

Let p be the origin of the coordinates and let us choose x_1-axis through q so that the direction from p to q is positive. Then the points p and q have coordinates $(0,0,\ldots,0)$ and $(r,0,\ldots,0)$. Let $(x^1,x^2,\ldots,x^n)$ be the coordinates of t . Then

$$r_o^2 = \sum_i (x^i)^2 \quad \text{and} \quad r_1^2 = (x_1 - r)^2 + \sum_{i=2}^n (x^i)^2 \ .$$

Let $x^i = r\xi^i$ for $i = 1,2,\ldots,n,$ and $\rho = \sqrt{\sum_{i=1}^n (\xi^i)^2}$. Then,

$$|N(p,q)| \leq \frac{C_1\ C_2}{r^{\alpha+\beta-n}} \int_{\rho<\frac{h}{r}} \frac{d\xi^1 d\xi^2 \ldots d\xi^n}{\rho^\alpha(\rho^2-2\xi^1+1)^{\beta/2}} \quad .$$

Now,

(i) $$d\xi^1 d\xi^2 \ldots d\xi^n = \rho^{n-1} d\rho\, d\omega \quad ,$$

(ii) $$\rho^2 - 2\xi^1 + 1 \geq (\rho - 1)^2$$

and

(iii) $$(\rho - 1)^2 > \frac{1}{4}\rho^2$$

for $\rho > 2$ as may be seen by observing that

$$(\rho - 1)^2 - \frac{1}{4}\rho^2 = \frac{1}{4}(3\rho - 2)(\rho - 2) ,$$

which in term is positive for $\rho > 2$. Hence,

$$|N(p,q)| \le \frac{C_1 C_2}{r^{\alpha+\beta-n}} \{\int_{\rho\le 2} \frac{\rho^{n-1-\alpha} d\rho d\omega}{(\rho^2 - 2\xi^1 + 1)^{\beta/2}} + 2^\beta \int_{2<\rho<\frac{h}{r}} \rho^{n-1-\alpha-\beta} d\rho \, d\omega\}$$

The first integral in the brackets on the right is a constant quantity, which we shall denote by $a\,\omega$.

If $\alpha + \beta < n$, then

$$|N(p,q)| \le C_1 C_2 \omega \, (a r^{n-\alpha-\beta} + \frac{2^\beta h^{n-\alpha-\beta}}{n - \alpha - \beta})$$

$$\le C_1 C_2 h^{n-\alpha-\beta} \omega \, (a + \frac{2^\beta}{n - \alpha - \beta}) .$$

If $\alpha + \beta = n$, then

$$|N(p,q)| \le C_1 C_2 \omega \, (a + 2^\beta \log_e \frac{h}{2r}) .$$

Finally, if $\alpha + \beta > n$, then

$$|N(p,q)| \le \frac{C_1 C_2 \omega}{r^{\alpha+\beta-n}} (a + 2^\beta \int_2^{\frac{h}{r}} \frac{d\rho}{\rho^{\alpha+\beta+1-n}}$$

$$< \frac{C_1 \, C_2 \omega}{r^{\alpha+\beta-n}} (a + 2^\beta \int_2^\infty \frac{d\rho}{\rho^{\alpha+\beta+1-n}})$$

$$= \frac{C_1 \, C_2 \omega}{r^{\alpha+\beta-n}} (a + \frac{2^{n-\alpha}}{\alpha + \beta - n})$$

Thus, (9) is valid in all cases.

Corollary 1 [9; p.85]. If the kernel has a weak singularity, then all its iterated kernels, beginning at a certain one, are bounded.

Proof: Let the kernel having a weak singularity be denoted by $L(p,q)$. Then, from Theorem 4, its m-th iterated kernel, $L^m(p,q)$ satisfies the estimate:

$$|L^m(p,q)| \leq \begin{cases} \dfrac{C_m}{r^{m\alpha-(m-1)n}} & \text{for } m\alpha - (m-1)n > 0 \\ C_m & \text{for } m\alpha - (m-1)n < 0 \end{cases}$$

where C_m is some constant. Thus, $L^m(p,q)$ is bounded if

$$m > \frac{n}{n - \alpha} \quad .$$

Definition 6. A linear operator A given on a certain linear subspace D everywhere dense in the Hilbert space, is said to be symmetric if for any $x,y \in D$,

$$< Ax,y > = < x,Ay > \quad .$$

Definition 7[17; p.306]. A self-adjoint operator is a bounded symmetric operator on H.

Definition 8 [1; p.172]. An operator is said to be completely continuous if for any sequence of vectors $\{q_n\}$ such that $||q_n||$ is bounded, $\{Aq_n\}$ has a convergent subsequence.

Theorem 5. The operator K defined by (8) is completely continuous and self-adjoint.

Proof: Let L^*L be denoted by N. Then, by Theorem 4, $N, N^2, N^{2^2}, \ldots, N^{2^n}, \ldots$ are operators with weak singularities; by Theorem 1, the operators are defined on $L_2(\Omega)$, and by (2), they are bounded. By Corollary 1, N^{2^n} is a Fredholm operator [10; p. 460] for some n. A Fredholm operator is completely continuous in $L_2(\Omega)$ [10; p.460]. Now,

$$N^{2^n} = N^{2^{n-1}} N^{2^{n-1}} = (N^{2^{n-1}})^* N^{2^{n-1}}$$

and a bounded operator A is completely continuous if and only if A^*A is completely continuous. Therefore, $N^{2^{n-1}}$ is completely continuous; in the same way, $N^{2^{n-2}}, N^{2^{n-3}}, \ldots, N^2$, N and L are completely continuous. By repeating the above argument for LL^* as N, we obtain L^* is completely continuous. The scalar multiple of a completely continuous operator and the sum of completely continuous operators are completely continuous [1; p.173]. Thus, K is completely continuous.

Since K is completely continuous, it is continuous i.e. bounded. Obviously, K is symmetric. Thus, K is a bounded symmetric operator on $L_2(\Omega)$. Hence, K is self-adjoint.

Chapter 2

§ 5. Invariant (closed) subspace

(5) can be written in the form

$$y - Ky = f \tag{10}$$

where

$$f \equiv g - \mu L^* g,$$

which is known since μ , g, L^* are known. Thus, by Theorem 2, finding the solution of (3) is equivalent to finding the solution of (10). We shall make use of the fact that K is completely continuous and self-adjoint.

Assume that for $(0 \neq)$ x_o in $L_2(\Omega)$, a sequence of linearly independent elements x_o, Kx_o, K^2x_o, ... can be constructed. Let

$$\begin{aligned} z_o &= x_o \\ z_1 &= Kx_o = Kz_o \\ z_2 &= Kz_1 = K^2 z_o \\ &\vdots \qquad\qquad \vdots \\ z_n &= Kz_{n-1} = K^n z_o \\ &\vdots \qquad\qquad \vdots \end{aligned}$$

Let H_n denote the subspace spanned by $\{z_o, \ldots, z_{n-1}\}$; let L_z be the linear manifold generated by $\{z_o, z_1, \ldots, z_n, \ldots\}$.

Definition 9 [16: p.6]. Suppose a fundamental sequence has no limit in the inner product space; the element assigned to be the limit of the sequence is called an ideal element. If u and v are two ideal elements defined by the respective sequences $\{u_n\}$ and $\{v_n\}$, then u and v are assumed equal if

$$\lim_{n\to\infty} ||u_n - v_n|| = 0 .$$

The scalar product of ideal elements is defined by

$$< u,v > = \lim_{n\to\infty} < u_n, v_n >$$

and

$$||u|| = \lim_{n\to\infty} ||u_n|| .$$

Let H_z denote L_z and its ideal elements.

Definition 10. The (closed) subspace M is said to be invariant for the self-adjoint operator A if $Ax \in M$ for any $x \in M$.

Theorem 6. H_z is invariant for K .

Proof: Let $Q(\lambda)$ be an arbitrary polynomial. Then, for $q = Q(K)z_o$ in L_z ,

$$Kq = KQ(K)z_o \in L_z .$$

On the other hand, if q is an ideal element of L_z, then there exists a sequence $q_n = Q_n(K)z_o$ in L_z such that

$$||q - q_n|| \to 0.$$

In this case,

$$||Kq - Kq_n|| \leq ||K|| \; ||q - q_n|| \to 0.$$

Since $Kq_n \in L_z$, this implies that Kq, being the limit element of L_z, must belong to H_z. Thus H_z is invariant for K.

If the number k of linearly independent vectors is finite, then $H_z = H_k$, and the method of solution for the non-homogeneous equation (10) terminates with an exact solution at the kth step. As we are interested in the approximation of solutions, we shall assume that H_z is infinite dimensional.

Theorem 7. The orthogonal complement of H_z, $H_z^{\perp}$, is invariant for K.

Proof: Let $x \in H_z$, and y be an arbitrary element of $H_z^{\perp}$. Then,

$$< Ky,x > = < y, Kx > = 0 ,$$

because $Kx \in H_z$. Consequently, $Ky \in H_z^{\perp}$, and hence, $H_z^{\perp}$ is invariant for K.

This theorem shows that the notion of an invariant subspace has special interest for self-adjoint operators. Namely, if H_z is an invariant subspace of K, then K is decomposed into two independent parts: the operator K_{H_z} on the space H_z, and the operator $K_{H_z^\perp}$ on the space $H_z^\perp$. At the same time, the equation

$$< Kx,y > = < x,Ky >$$

for arbitrary x, y in $L_2(\Omega)$, characterizing self-adjoint operators and being fulfilled for K on $L_2(\Omega)$, is satisfied, in particular, on H_z and $H_z^\perp$. Consequently, the 'parts' of the operator K considered on each of the subspaces H_z and $H_z^\perp$ are also self-adjoint.

Since H_z is closed, it follows that H_z can be regarded as a Hilbert space. We can apply the above reasoning repeatedly. In this way, the study of K can be reduced to its study on the invariant subspaces.

§6. Lanczos' Method of Minimized Iterations [8; p.265].

Let $x_o (\neq 0)$ be an arbitrarily chosen vector of $L_2(\Omega)$. The vector x_1 is constructed by

$$x_1 = Kx_o - \alpha_o x_o$$

satisfying

$$\begin{aligned} 0 &= < x_1, x_o > \\ &= < Kx_o - \alpha_o x_o, x_o > \\ &= < Kx_o, x_o > - \alpha_o < x_o, x_o > , \end{aligned}$$

i.e.

$$\alpha_o = \frac{< Kx_o, x_o >}{< x_o, x_o >} .$$

x_2 is constructed by

$$x_2 = Kx_1 - \alpha_1 x_1 - \beta_o x_o$$

satisfying

(a)

$$\begin{aligned} 0 &= < x_2, x_o > \\ &= < Kx_1 - \alpha_1 x_1 - \beta_o x_o, x_o > \\ &= < Kx_1, x_o > - \beta_o < x_o, x_o > , \end{aligned}$$

i.e.

$$\begin{aligned} \beta_o &= \frac{<Kx_1, x_o >}{< x_o, x_o >} \\ &= \frac{< x_1, Kx_o >}{< x_o, x_o >} \\ &= \frac{< x_1, x_1 + \alpha_o x_o >}{< x_o, x_o >} \\ &= \frac{< x_1, x_1 >}{< x_o, x_o >} \end{aligned}$$

and

(b)

$$\begin{aligned} 0 &= < x_2, x_1 > \\ &= < Kx_1 - \alpha_1 x_1 - \beta_o x_o, x_1 > \end{aligned}$$

$$= \langle Kx_1, x_1 \rangle - \alpha_1 \langle x_1, x_1 \rangle ,$$

i.e.

$$\alpha_1 = \frac{\langle Kx_1, x_1 \rangle}{\langle x_1, x_1 \rangle} .$$

In general, for x_{k+1} to be orthogonal to x_k and x_{k-1} in

(11) $$x_{k+1} = (K - \alpha_k I)x_k - \beta_{k-1}x_{k-1}$$

where I is the identity operator, one must choose

$$\alpha_k = \frac{\langle Kx_k, x_k \rangle}{\langle x_k, x_k \rangle}$$

and

$$\beta_{k-1} = \frac{\langle Kx_k, x_{k-1} \rangle}{\langle x_{k-1}, x_{k-1} \rangle}$$

$$= \frac{\langle x_k, x_k \rangle}{\langle x_{k-1}, x_{k-1} \rangle}$$

for $k = 0,1,2, \ldots,$ and $\beta_{-1} = 0$. The above method of constructing orthogonal elements is called Lanczos' Method of Minimized Iterations.

<u>Theorem 8.</u> <u>x_{k+1} constructed by Lanczos' Method of Minimized Iterations is orthogonal to $x_o, x_1, \ldots, x_{k-2}, x_{k-1}, x_k$.</u>

Proof: In order to prove this, the Principle of Mathematical Induction will be used. In (11), for $k = 0$, x_1 is clearly orthogonal to x_o. Suppose x_{k+1} is orthogonal to $x_o, x_1, \ldots, x_k$ for $k = 0, 1, 2, \ldots, s-1$. Then,

$$x_{s+1} = (K - \alpha_s I)x_s - \beta_{s-1}x_{s-1}$$

and x_{s+1}, by construction, is orthogonal to x_s and x_{s-1}. If $0 \leq t \leq s - 2$, then,

$$\begin{aligned} \langle x_t, x_{s+1} \rangle &= \langle x_t, (K - \alpha_s I)x_s - \beta_{s-1}x_{s-1} \rangle \\ &= \langle x_t, Kx_s \rangle \qquad \text{(by hypothesis)} \\ &= \langle Kx_t, x_s \rangle \\ &= \langle x_{t+1} + \alpha_t x_t + \beta_{t-1} x_{t-1}, x_s \rangle \\ &= 0 . \end{aligned}$$

Hence, x_{k+1} is, by the Principle of Mathematical Induction, orthogonal to $x_o, x_1, \ldots, x_{k-1}, x_k$.

From this theorem, it can be seen that the elements constructed by using Lanczos' Method can be obtained by the Gram-Schmidt Process. Thus, Lanczos' Method breaks down at constructing x_n if and only if $x_o, Kx_o, \ldots, K^n x_o$ are linearly dependent, and $x_o, Kx_o, \ldots, K^{n-1}x_o$ are linearly independent. When Lanczos' Method does not break down, it furnishes a basis for H_z.

§7. Approximate solution

Definition 11. Two elements x, y in H are said to be orthogonal if and only if

$$< x,y > = 0 .$$

If M is a subspace of $L_2(\Omega)$, then every element $x \in L_2(\Omega)$ can be represented uniquely in the form

$$x = u + v$$

where $u \in M$ and v is orthogonal to q for all $q \in M$.

Definition 12. The element u in the above representation is called the projection of x on the subspace M.

We shall construct a linear operator, K_n, defined on the subspace H_n by

$$(12) \qquad \begin{cases} z_k = K_n z_{k-1} & \text{for } k = 1, 2, \ldots, n-1, \\ E_n z_n = K_n z_{n-1}, & \end{cases}$$

where $E_n z_n$ is the projection of z_n on H_n. (12) defines K_n completely. In fact, let $x \in H_n$. Then,

$$x = \sum_{i=0}^{n-1} c_i z_i$$

where $c_o, c_1, \ldots, c_{n-1}$ are constants.

$$K_n x = \sum_{i=0}^{n-1} c_i K_n z_i$$

$$= c_o z_1 + \ldots + c_{n-1} E_n z_n \in H_n .$$

We shall use Lanczos' Method of Minimized Iterations to solve

(13) $$(I - K_n)y_n = f$$

by taking $z_o = f$. The existence of y_n is guaranteed by the unique solution of (3). Then, we shall show that y_n converges to the solution of (10) faster than any geometrical progression. By Theorem 2, y_n actually converges faster than any geometrical progression to the solution of our original problem, given by (3). Thus, we complete our aim.

The domain of definition of $I - K_n$ is the intersection of the domain of definition of I and that of K_n . Thus, $y_n \in H_n$. Therefore, it can be written in the form

(14) $$y_n = \sum_{k=0}^{n-1} c_k x_k,$$

where $x_o, x_1, \ldots, x_{n-1}$ are constructed by Lanczos' Method and $c_o, c_1, \ldots, c_{n-1}$ are constants. By the construction of K_n,

$$K_n x_k = x_{k+1} + \alpha_k x_k + \beta_{k-1} x_{k-1} \quad \text{for} \quad k = 1,\ldots,n-2,$$

$$K_n x_o = x_1 + \alpha_o x_o$$

and $$K_n x_{n-1} = \alpha_{n-1} x_{n-1} + \beta_{n-2} x_{n-2} .$$

On substituting (14) into (13),

$$(I - K_n) \sum_{k=0}^{n-1} c_k x_k = f .$$

$$\sum_{k=0}^{n-1} c_k x_k - \sum_{k=0}^{n-1} c_k K_n x_k = f .$$

$$\sum_{k=0}^{n-1} c_k x_k - \sum_{k=1}^{n-2} c_k (x_{k+1} + \alpha_k x_k + \beta_{k-1} x_{k-1}) - c_o (x_1 + \alpha_o x_o)$$

$$- c_{n-1} (\alpha_{n-1} x_{n-1} + \beta_{n-2} x_{n-2}) = f .$$

By construction,

$$f = z_o = x_o .$$

Hence, on equating coefficients of like terms, c_k are determined by the system of equations:

$$(15) \quad \begin{cases} -c_{k-1} + (1 - \alpha_k) c_k - \beta_k c_{k+1} = 0 \quad \text{for} \quad k = 1,2,\dots,n-1, \\ c_o - c_1 \beta_o - c_o \alpha_o = 1 \\ c_n = 0 . \end{cases}$$

Theorem 9 [16: p.77]. y_n tends to the solution y in (10) faster than any geometrical progression.

Proof: By the construction,

$$z_o = f .$$

Since y_n is in H_n ,

$$y_n = Q_{n-1}(K_n) z_o ,$$

where $Q_{n-1}(\lambda)$ is a polynomial of $(n-1)$th degree. On substituting this into (13), we get

$$Q_{n-1}(K_n)z_o - K_nQ_{n-1}(K_n)z_o - f = 0 ,$$

i.e. $$[Q_{n-1}(K_n) - K_nQ_{n-1}(K_n) - I]z_o = 0 .$$

Since the characteristic polynomial $P_n(\lambda)$ of K_n is uniquely defined [16; p.17],

(16) $$Q_{n-1}(\lambda) - \lambda Q_{n-1}(\lambda) - 1 = CP_n(\lambda)$$

where C is a constant to be determined. Let $\lambda = 1$. Then,

(17) $$C = -\frac{1}{P_n(1)} .$$

Thus, from (16) and (17),

$$Q_{n-1}(K_n) = (I - K_n)^{-1}[I - \frac{P_n(K_n)}{P_n(1)}] .$$

Hence,

$$y_n = (I - K_n)^{-1} [I - \frac{P_n(K_n)}{P_n(1)}]f$$

$$= (I - K)^{-1} [I - \frac{P_n(K)}{P_n(1)}]f$$

since the right hand side involves an (n-1)th degree polynomial in K_n with a leading coefficient of $\frac{1}{P_n(1)}$. Since

$$(I - K_n)y_n - (I - K)y_n = [I - \frac{P_n(K_n)}{P_n(1)}]f - [I - \frac{P_n(K)}{P_n(1)}]f$$

and

$$\frac{P_n(K)f - P_n(K_n)f}{P_n(1)} = \frac{K^n f - K_n^n f}{P_n(1)} ,$$

hence,

(18) $$Ky_n - K_n y_n = \frac{K^n f - K_n^n f}{P_n(1)} .$$

Since $K_n^n f$ is the projection of $K^n f$ on H_n, it follows that

(19) $$||K^n f - K_n^n f|| \leq ||K^n f - Q_{n-1}(K)f||$$

holds for any arbitrary polynomial $Q_{n-1}(\lambda)$. Now, $Q_{n-1}(\lambda)$ is so chosen such that

$$\lambda^n - Q_{n-1}(\lambda) = (\lambda - \lambda_o)(\lambda - \lambda_1)(\lambda - \lambda_2) \ldots (\lambda - \lambda_{n-1})$$

where $\lambda_o, \ldots , \lambda_{n-1}$ are the eigenvalues of K arranged in the order of decreasing numerical value:

$$|\lambda_o| > |\lambda_1| > \ldots > |\lambda_n| > \ldots .$$

If f is expanded in terms of the eigenvectors u_k of K, that is,

$$f = \sum_{k=0}^{\infty} < f,u_k > u_k ,$$ then

(19) becomes

$$||K^n f - K_n^n f||^2 \leq \sum_{k=0}^{\infty} [\lambda_k{}^n - Q_{n-1}(\lambda_k)]^2 < f,u_k>^2$$

(20) $$= \sum_{k=n}^{\infty} (\lambda_k - \lambda_o)^2 \ldots (\lambda_k - \lambda_{n-1})^2 <f,u_k>^2$$

$$\leq (|\lambda_o|+|\lambda_n|)^2 \ldots (|\lambda_{n-1}|+|\lambda_n|)^2 \sum_{k=n}^{\infty} < f,u_k>^2 .$$

Let η_n be given by

$$\eta_n = y - y_n .$$

Then,

$$\eta_n - K\eta_n = y - y_n - Ky + Ky_n$$

$$= Ky_n - y_n + f \qquad \text{by (10)}$$

$$= Ky_n - (K_n y_n + f) + f \qquad \text{by (13)}$$

$$= Ky_n - K_n y_n$$

$$= \frac{K^n f - K_n^n f}{P_n(1)} \qquad \text{by (18)} .$$

Thus,

$$\eta_n = (I - K)^{-1} \frac{K^n f - K_n^n f}{P_n(1)} .$$

By Theorem 2, (10) has a unique solution. Thus, $(I - K)^{-1}$ is bounded [4; p.903]. With the help of (20),

$$||y - y_n|| = ||\eta_n||$$

$$\leq ||(I - K)^{-1}|| \; (|\lambda_o|+|\lambda_n|) \; \ldots \; (|\lambda_{n-1}|+|\lambda_n|) \frac{\sqrt{\sum_{k=n}^{\infty} <f,u_k>^2}}{P_n(1)}$$

$$= ||(I - K)^{-1}|| \frac{|\lambda_o|+|\lambda_n|}{|1 - \lambda_o^{(n)}|} \ldots \frac{|\lambda_{n-1}|+|\lambda_n|}{|1 - \lambda_{n-1}^{(n)}|} \sqrt{\sum_{k=n}^{\infty} <f,u_k>^2}$$

where $\lambda_o^{(n)}, \ldots, \lambda_{n-1}^{(n)}$ are the eigenvalues of K_n. As $n \to \infty$,

$$\lambda_k^{(n)} \to \lambda_k \qquad \text{and} \qquad |\lambda_n| \to 0 .$$

Therefore, y_n tends to y faster than any geometrical progression.

§ 8. Conclusion

To recapitulate, in solving (3), (i) it is transformed to (10), (ii) the transformed operator is shown to be completely continuous and self-adjoint, and (iii) Lanczos' Method is used to find the approximate solution and Vorobyev's Method is used to show that the approximate solution converges to the solution of (10), and hence to the solution of (3), faster than any geometrical progression.

If the given kernel is symmetric, then no symmetrization is necessary. W. Karush [6] described a method equivalent to Lanczos' Method, and in our case, in the construction of the orthogonal elements, we have seen that Lanczos' Method coincides with the Gram-Schmidt Process. Instead of using Vorobyev's Method to prove the rate of convergence, the method by Karush [6; p.848] may be used.

Bibliography

[1] S.K. Berberian, Introduction To Hilbert Space, Oxford University Press, Inc., 1961.

[2] S.K. Berberian, Measure And Integration, The Macmillan Company, New York, 1965.

[3] G.F.D. Duff, Partial Differential Equations, University of Toronto Press, 1962.

[4] N. Dunford and J.T. Schwartz (With the assistance of W.G. Bade and R.G. Bartle), Linear Operators Part 1 and Part 2, Interscience Publisher, New York, 1958 and 1964.

[5] E. Hellinger und O. Toeplitz, Integralgleichungen Und Gleichungen Mit Unendlichvielen Unbekannten, Chelsea Publishing Company, New York, N.Y., 1953.

[6] W. Karush, Convergence Of A Method Of Solving Linear Problems, Proc. AMS 3, 1952, pp.839 - 851 .

[7] O.D. Kellogg, Foundations Of Potential Theory, New York Dover Publications, Inc., 1953.

[8] C. Lanczos, An Iterative Method For The Solution Of The Eigenvalue Problem Of Linear Differential and Integral Operators, Journal of Research of the National Bureau of Standards, Vol. 45, No. 4, October, 1950, Research Paper 2133, pp.255 - 282.

[9] S.G. Michlin, Vorlesungen Über Lineare Integralgleichungen, VEB Deutscher Verlag Der Wissenschaften, Berlin, 1962.

0] S.G. Mikhlin, Variational Methods In Mathematical Physics, Translated by T.Boddington, Pergamon Press, 1964.

1] F. Riesz and B.Sz-Nagy, Functional Analysis, Translated by L.F. Boron, Third Edition, Frederick Ungar Publishing Co., 1965.

E. Schmidt, Entwicklung Willkürlicher Funktionen Nach Systemen Vorgeschriebener, Diss. Göttingen 1905, 33 S.= Math. Ann. 63 (1907), pp.433 - 476.

M. Stone, Linear Transformations In Hilbert Space And Their Applications To Analysis, Amer. Math. Soc. Colloquium Publications, New York, Vol. 15, 1932.

A.E. Taylor, Introduction To Functional Analysis, New York, John Wiley and Sons, Inc., 1958.

J. von Neumann, Allgemeine Eigenwerttheorie Hermitescher Funktionaloperatoren, Math. Ann. 102, Vol. II, No. 1, 1929, pp.49 - 131.

V. Vorobyev, Method Of Moments In Applied Mathematics, Translated by B. Seckler, Gordon and Breach, 1964.

Z. Vulikh, Introduction To Functional Analysis For Scientists And Technologists, Translated by I.N. Sneddon, Pergamon Press, 1963.