Re-iterative Methods For

Integral Equations

Paul Antony Burton

This dissertation is submitted in partial

fulfilment of the requirements of

the degree of Master of Science.

Department of Mathematics.

University of Reading.

September, 1993.

Abstract

A review is given of the Galerkin, iterated Galerkin and re-iterated Galerkin methods for finding approximate solutions to integral equations of the second kind. The process of re-iteration, first described by Porter Stirling, is applied to the Kantorovich method to produce two new re-iteration methods. The first, the modified iterated Kantorovich method, is a straightforward application of the re-iterated Galerkin method to a regularized Kantorovich equation. The second method differs by applying the Kantorovich regularization at each iteration. This second method is tested on some example integral equations and its convergence compared with that of the re-iterated Galerkin method. Finally it is shown how the presence of an eigenvalue can affect the convergence of these methods.

Acknowledgements

I wish to thank, first and foremost, my supervisor for this project, Dr.D.Porter, for his invaluable help and encouragement throughout. I would also like to thank my fellow MSc students and members of the department who have assisted me over the last year. Finally I wish to acknowledge the financial support of the Science and ngineering Research Council.

Contents

Table of Notation **iv**

1 Introduction **1**

1.1 Integral quations 1

1.2 Projection methods 3

2 Galerkin Methods **4**

2.1 The Galrkin Method 4

2.2 Degenerate Kernels 5

2.3 The Iterated Galerkin Method 7

2.4 The Re-iterated Galerkin Method 8

3 Kantorovich Methods **13**

3.1 The Kantorovich Method 13

3.2 The Iterated Kantorovich Method 15

3.3 The Modified Iterated Kantorovich Method 16

3.4 The Re-iterated Kantorovich Method 19

4 Numerical Considerations **26**

4.1 Numerical Integration 26

4.2 Choosing The Subspace 27

5 Examples and Results **29**

5.1 The Kernel 2max(x,t) 29

5.2 The $\frac{1}{2}\sin(\kappa_0|x-t|)$ Kernel 33

6 Convergence Regions and Eigenvalue Bounds **37**

6.1 The Kernel 2λmax(x,t) 38

6.1.1 Calculating $\rho(S)$ 39

6.2 The Kernel $\frac{\lambda}{2}sin(\kappa_0|x-t|)$ 44

7 Extending to 2-D **47**

8 Conclusions **51**

Table of Notation

$k(x,t)$	kernel
K	integral operator
S, SK	convergence operators
μ_1^+, μ_1^-	eigenvalues
$\tilde{\mu}_1$	approximate eigenvalue
(ϕ, ψ)	inner product
$\|.\|$	norm
H	Hilbert space
$L_2(0,1)$	space of square integrable functions
$R_K(\phi)$	Rayleigh quotient
I	identity operator
P_n	projection
$\rho(T)$	spectral radius of T
$\kappa_0, \alpha, \beta, \lambda$	parameters
$\phi(x)$	particular function
E_n	n-dimensional subspace
$\{\chi_1, ..., \chi_n\}$	basis functions
sup	supremum

In certain problems it is not uncommon to encounter equations in which the un-

b

a

b

a

1

0

We wish to find approximations to the solution of integral equations of the form

$$= \quad + \tag{1 3}$$

in some Hilbert space , such as $_2(0\ 1)$.

To do this we look for an approximation $_n$ to in some finite dimensional subspace $_n$ of , by solving the equation

$$_n = \ _n \ + \ _n \qquad _n \tag{1 4}$$

in $_n$, where $_n$ is an orthogonal projection of onto $_n$.

Ch pter 2

G lerkin Methods

2.1 The Galerkin Method

If our approximation p_n is chosen such that $Ap_n - f$ is orthogonal to the subspace E_n, then we have

$$(Ap_n - f, \chi_j) = 0, \quad j = 1, \ldots, n$$

$$\Rightarrow \quad (Ap_n, \chi_j) = (f, \chi_j), \quad j = 1, \ldots, n$$

where the χ_j's are the orthonormal basis functions of E_n.

Using our definition of p_n from equation (1.5) and $A = I - K$ (taking $\lambda = 1$ in equation (1.3) for convenience) we arrive at

$$\sum_{i=1}^{n} \alpha_i \{(\chi_i, \chi_j) - (K\chi_i, \chi_j)\} = (f, \chi_j), \quad j = 1, \ldots, n \tag{2.1}$$

This system of n equations can be solved simultaneously, provided the determinant of coefficients is non-zero, to obtain the α_i's which, via equation (1.5), provide us with our approximation p_n to the solution ϕ of our integral equation.

In terms of the projection P_n of H onto E_n,

$$P_n p_n = P_n f + P_n K p_n$$

$$_n = (\quad _n \quad)^{-1} \quad _n$$

since $\quad _n \ _n = \quad _n$.

n

n

$N \qquad N \quad N$

$N \qquad \overset{N}{\underset{i=1}{}} \quad i \qquad i$

N

N

Notice that this equation is identical to the expression for the 's in Galerkin's method, so the 's are the same for both methods. However, $_N = _n$. In fact

$$_N = + \sum_{n=1}^{N} _n \ _n = + _n \qquad (2\ 5)$$

n n -1 n

N

n n

n -1 n

n n -1

n n n -1

n -1

n n -1 -1

n -1 -1

n n

n -1

n n

n -1

n n

n n n -1

n n n n

n -1

n n

n

0 0

$$\frac{n}{n \quad 1}$$

n

0 0 0 1 0 1 0 1

0 0 1

n 1 n 0 1 0

n n 1 n n 1 0

$$\frac{1}{0}$$

n n 1

n n 0 n 0

in the initial (non-iterated) Galerkin approximations we find, (taking $n = 0$),

$$\begin{aligned} r_0 &= f + Kp_0 - p_0 \\ &= \hat{p_0} - p_0 \\ &= (I - KP)^{-1} f - (I - P_n K)^{-1} P_n K \\ &= (I - P_n)(I - KP_n)^{-1} f, \end{aligned}$$

which leads us to

$$\frac{\|\hat{r_n}\|}{\|r_n\|} \leq \|K\|, \tag{2.9}$$

ie. the ratio of the residual of the iterated Galerkin approximation, over the norm of the residual of the Galerkin approximation, provides a lower bound on the norm of the operator K. Whereas this quantity has no direct consequences for the convergence of the method, it is however valuable information to know about the kernel, and comes relatively free.

The re-iterated Galerkin method provides a means of calculating an approximation to the solution of an integral equation, providing certain conditions are met, to any desired accuracy. The number of re-iterations required to reach a specific accuracy depends upon the rate of convergence of the method, which, in turn, depends upon the size of the subspace used. An increase in the size of $_n$ will result in faster convergence, but requires more calculations.

Is there a way in which the re-iteration process can be made more efficient, increasing the rate of convergence whilst minimising the number of calculations?

quation(3.1) represents the application of the operator K to both sides of the original integral equation $\phi = f + K\phi$.

We get a clue that we may have improved matters by considering the Neumann series for our equation, given by

$$\phi = f + Kf + K^2 f + \ldots + K^n f + \ldots \tag{3.2}$$

The Kantorovich method gives $\phi = f + Kf + K\psi$, compared to $\phi = f + K\phi$ for Galerkin, and so contains an extra term of the Neumann series.

We approximate ψ by finding the Galerkin solution of the equation $\psi = Kf + K\psi$, ie.

$$\psi \quad q_n = \sum_{i=1}^{n} \beta_i \chi_i,$$

where $q_n \quad E_n$, a subspace spanned by the orthonormal basis $\chi_1, \ldots, \chi_n$.

We want to choose q_n so that $A\psi \quad Kf$ is orthogonal to E_n, where $A = I \quad K$. Therefore

$$\begin{aligned} (Aq_n \quad Kf, \chi_j) &= 0, \qquad j = 1, \ldots, n \\ (Aq_n, \chi_j) &= (Kf, \chi_j), \quad j = 1, \ldots, n \\ \sum_{i=1}^{n} \beta_i \ (\chi_i, \chi_j) \quad (K\chi_i, \chi_j) &= (Kf, \chi_j), \quad j = 1, \ldots, n \end{aligned}$$

Notice that the $_i$'s given by the system of equations above differ from the $_i$'s given by Galerkin's method because we have a different free term in the integral equation, instead of . The solution to our original problem is therefore approximated by $\tilde{}_n = + \ _n$, which is different from the $_n$ of Galerkin's method.

The benefit of making the substitution in the Kantorovich method becomes clear when we compare the differences between the approximations and =

$(I-K)^{-1}f$. If we denote by P_n the orthogonal projection onto E_n, we see that

$$p_n - \phi = (I - P_nK)^{-1}(P_n - I)(I-K)^{-1}f, \tag{3.3}$$

and

$$\tilde{p_n} - \phi = (I - P_nK)^{-1}(P_nK - K)(I-K)^{-1}f. \tag{3.4}$$

The difference between the two errors is that for the first, P_n has to be chosen such that $(P_n - I)(I-K)^{-1}f$ is small, so the choice depends on f, whereas, for the Kantorovich method, it is only necessary to choose P_n such that $\|P_nK - K\|$ is small, independently of f, which is more easily satisfied (see Sloan [6]).

Another advantage of the Kantorovich method is that it removes f from the calculations, replacing it by Kf. This is desirable if f is not a particularly smooth function. The application of K to f has the effect of smoothing the free term. Kf is then used in the calculations to approximate ψ, with f being added on afterwards to give $\tilde{p}$, the approximation to ϕ.

3.2 The Iterated antorovich Method

Sloan [6] demonstrated that the principle of iteration, which he first applied to Galerkin's method, could also be brought to bear on Kantorovich's method.

We wish to find the solution $\phi = f + \psi$, where ψ satisfies the integral equation $\psi = Kf + K\psi$. The Galerkin approximation to ψ gives

$$\psi \approx q_0 = (I - P_nK)^{-1}P_nKf \quad \Rightarrow \quad \phi \approx \tilde{p_0} = f + q_0.$$

If we apply the iteration process to q_0, by replacing ψ in the equation, we arrive at

$$\hat{q_0} = Kf + Kq_0$$

ie.

$$\hat{q}_0 = (I - KP_n)^{-1}Kf$$

$$\Rightarrow \qquad \phi \approx \hat{\tilde{p}}_0 = f + \hat{q}_0.$$

The error in the approximation of ψ is given by

$$\begin{aligned}
\hat{q}_0 - \psi &= (I - KP_n)^{-1}Kf - (I - K)^{-1}Kf \\
&= (I - KP_n)^{-1}(K - KP_n)(q_0 - \psi) \\
&= S(q_0 - \psi), \qquad \text{where } S = (I - KP_n)^{-1}(K - KP_n),
\end{aligned}$$

which in turn implies

$$\hat{\tilde{p}}_0 - \phi = S(p_0 - \phi).$$

The operator S is the same as that encountered in the iterated Galerkin method , which is as one might expect considering that the two methods differ only in the choice of free term. As with the Galerkin method, iteration will improve the Kantorovich approximation if the norm of the operator S is less than one, which can be guaranteed by a suitable choice of subspace. The advantage of the iterated Kantorovich method over the iterated Galerkin method lies in the possibility that $\psi = \phi - f$ will be smoother than ϕ, and thus easier to approximate.

3.3 The Modified Iterated antorovich Method

Although it has not previously been demonstrated, it seems fair to suppose that, since Kantorovich's method benefits from iteration, it might equally benefit from the process of re-iteration.

The iterated Kantorovich method is obtained by applying the iterated Galerkin method to the equation $\psi = Kf + K\psi$, where $\psi = \phi - f$. The iterated approximation $\hat{q}_0$ exhibits an error which can be compensated for by adding a correction term, ψ_1 say, so that

$$\psi = \hat{q}_0 + \psi_1.$$

By substituting ψ into the modified integral equation we see

$$\begin{aligned} \psi_1 &= Kf + K\hat{q}_0 - \hat{q}_0 + K\psi_1 \\ &= \hat{r}_0 + K\psi_1, \end{aligned}$$

where $\hat{r}_0$ is the residual error incurred in approximating ψ by $\hat{q}_0$.

We now have an integral equation to solve for ψ_1, with free term

$$\begin{aligned} \hat{r}_0 &= Kf - (I - K)\hat{q}_0 \\ &= K(Kf - (I - K)q_0) \\ &= Kr_0. \end{aligned}$$

As before, the ratio of the residual norms of the iterated and non-iterated approximations estimate $\|K\|$.

Once again we can use iterated Galerkin to find an approximation $\hat{q}_1$ to ϕ_1, and hence our new approximated solution to the modified equation becomes

$$\psi \approx \hat{q}_0 + \hat{q}_1,$$

where $\hat{q}_1 = (I - KP_n)^{-1}\hat{r}_0$, which has an error

$$\begin{aligned} \hat{q}_0 + \hat{q}_1 - \psi &= (I - KP_n)^{-1}\hat{r}_0 + (\hat{q}_0 - \psi) \\ &= -(I - KP_n)^{-1}(I - K)(\hat{q}_0 - \psi) + (\hat{q}_0 - \psi) \end{aligned}$$

$$= \quad (\qquad\quad _n)^{-1}(\qquad\quad _n)(\hat{}_0 \qquad)$$

$$= \quad (\hat{}_0 \qquad)$$

$$= \quad ^2(\ _0 \qquad)$$

ie. $\hat{\tilde{}}_1 \quad = \quad ^2(\ _0 \qquad)$, where $\qquad \tilde{}_1 = \quad + \hat{}_0 + \hat{}_1$. So $\hat{}_0 + \hat{}_1$ is an improved approximation to $\quad$ if $\qquad 1$, which we have previously assumed.

As before, this process can be repeated by noting that

$$= \hat{}_0 + \hat{}_1 + \ _2$$

where $\ _2$ is a correction term which satisfies the equation

$$_2 = \hat{}_1 + \qquad _2$$

where $\hat{}_1 = \hat{}_0 \quad (\qquad)\hat{}_1 = (\qquad\quad _n)(\qquad\quad _n)^{-1}\hat{}_0$, the residual error in $\hat{}_0 + \hat{}_1$. As with the re-iterated Galerkin method, the ratio of successive residual norms, $\frac{\|\hat{r}_n\|}{\|\hat{r}_{n-1}\|}$, gives an underestimate for the norm of $\quad$.

In general, then, we have $\qquad \hat{}^n_{kant} = \quad + \hat{}_0 + \ldots + \hat{}_n$, where each correction term $\ _n$ is determined by using iterated Galerkin on $\ _n = \hat{}_{n-1} + \qquad _n$, giving $\hat{}_n = (\qquad\quad _n)^{-1}\hat{}_{n-1}$, where $\hat{}_{n-1} = \qquad (\qquad)\hat{}^{n-1}_{kant}$. The error in the current approximation being given by $\hat{}^n_{kant} \qquad = \quad (\hat{}^{n-1}_{kant} \qquad) = \quad ^{n+1}(\ _0 \qquad)$.

The benefits of re-iteration for the Kantorovich method are the same as for the

generally smaller than that for the re-iterated Galerkin method, $_0$. All subsequent approximations, however, improve by the same factor, , for both methods. So the rates of convergence are virtually the same.

The improvement in the initial approximation, brought about by the Kantorovich "regularization" of the original equation, leads us to wonder whether subsequent approximations might also benefit from similar regularization.

Suppose, once again, that we wish to solve

$$= + \tag{3 5}$$

in a Hilbert space , where is a compact linear map on , is given, and is sought. We first perform a Kantorovich "regularization" on the above equation, whereby we write the solution as

$$= + \tag{3 6}$$

where = satisfies the equation

$$= + \tag{3 7}$$

We seek a Galerkin approximation $\tilde{}_0$ to in a subspace $_n$ of , so that $\tilde{}_0 = (\quad _n \quad)^{-1} \ _n$, where $_n$ is the orthogonal projection of onto $_n$. Our Kantorovich approximation, $\tilde{}_0$, is therefore

$$\tilde{}_0 = \tilde{}_0 +$$

Next we form the iterated Galerkin solution of equation (3.7), ie.

$$\hat{}_0 \qquad\qquad _0 \qquad\qquad _n{}^{-1}$$

giving us the iterated Kantorovich approximation

$$\phi \approx \hat{\tilde{p}}_0 = \hat{\tilde{q}}_0 + f,$$

where

$$\hat{\tilde{p}}_0 - \phi = S(\tilde{p}_0 - \phi).$$

As we have seen before, iteration improves the initial approximation by a factor $\|S\|$ for both the Kantorovich and Galerkin methods, the Kantorovich method having the possible advantage that $\tilde{p}_0$ has better convergence properties than p_0, since it relies upon $\|P_n K - K\| \to 0$ as $n \to \infty$, rather than $(P_n - I)(I - K)^{-1} \to 0$ as $n \to \infty$.

We can utilise this convergence property in a re-iterative method by performing a Kantorovich regularization at each iteration. Instead of finding a correction term ψ_1 to $\hat{\tilde{q}}_0$, to form $\psi = \hat{\tilde{q}}_0 + \psi_1$, as we did for the modified iterated Kantorovich method, we now look for a correction term for ϕ.

We know $\phi \approx \hat{\tilde{p}}_0 = f + \hat{\tilde{q}}_0$, so this implies that

$$\phi = \hat{\tilde{p}}_0 + \phi_1, \tag{3.8}$$

which in turn implies that

$$\phi_1 = \hat{r}_0 + K\phi_1, \tag{3.9}$$

where $\hat{r}_0 = f - (I - K)\hat{\tilde{p}}_0 = Kr_0$.

This method differs from the modified iterated Kantorovich method in that we now perform a second Kantorovich regularization, this time to equation (3.9).

Let $K\phi_1 = \psi_1$, then $\phi_1 = \hat{r}_0 + \psi_1$, where ψ_1 satisfies

$$\psi_1 = K\hat{r}_0 + K\psi_1.$$

The Galerkin solution to this is

$$1 \quad \tilde{}_1 = (\quad _n \quad)^{-1} \; _n \quad \hat{}_0$$

which we can iterate to obtain

$$1 \quad \hat{\tilde{}}_1 = \quad \hat{}_0 + \quad \tilde{}_1 = (\quad _n)^{-1} \quad \hat{}_0$$

The correction term $_1$ is therefore given by

$$1 \quad \hat{\tilde{}}_1 = \hat{}_0 + \hat{\tilde{}}_1$$

1

0 1

0 1 0 1 0

0 $_n^{-1}$ 0 $^{-1}$ 0

$_n^{-1}$ $_n$ 2 $^{-1}$ 0

$_n^{-1}$ $_n$ 0

0 0

0 1 0 0

0 1 2

where ϕ_2 is a correction term which satisfies the equation

$$\phi_2 = \hat{r}_1 + K\phi_2, \tag{3.11}$$

where $\hat{}_1 = \quad (\quad)(\hat{\tilde{}}_0 + \hat{\tilde{}}_1) = \quad \hat{}_1$. Performing a Kantorovich regularization on the above equation gives us

$$_2 = \hat{}_1 + \quad _2$$

where $\quad _2$ satisfies

$$_2 = \quad \hat{}_1 + \quad \quad _2$$

Applying Galerkin's method gives us $\quad _2 \quad \tilde{}_2 = (\quad \quad _n \quad)^{-1} \quad _n \quad \hat{}_1$, and upon iteration, $\quad _2 \quad \hat{\tilde{}}_2 = \quad \hat{}_1 + \quad \tilde{}_2 = (\quad \quad _n)^{-1} \quad \hat{}_1$.

So now

$$_2 = \hat{\tilde{}}_2 = \hat{}_1 + \hat{\tilde{}}_2$$

which implies that

$$\hat{\tilde{}}_0 + \hat{\tilde{}}_1 + \hat{\tilde{}}_2$$

The error in this approximation can be shown to be

$$\begin{aligned} \hat{\tilde{}}_0 + \hat{\tilde{}}_1 + \hat{\tilde{}}_2 \quad &= \quad (\quad \quad _n)^{-1}(\quad \quad _n) \quad (\hat{\tilde{}}_0 + \hat{\tilde{}}_1 \quad \quad) \\ &= \quad \quad (\hat{\tilde{}}_0 + \hat{\tilde{}}_1 \quad \quad) \end{aligned}$$

so that

$$\hat{\tilde{}}_0 + \hat{\tilde{}}_1 + \hat{\tilde{}}_2 \qquad\qquad \hat{\tilde{}}_0 + \hat{\tilde{}}_1$$

provided, once again, that $\quad \quad 1$.

In general, if $\quad \hat{}^n_{rekant} = \hat{\tilde{}}_0 + \ldots + \hat{\tilde{}}_n$, and $\hat{}_n = \quad \quad (\quad \quad)\hat{}^n_{rekant}$ denotes the residual after the nth iteration, then the correction term $\quad _n = \hat{}_{n-1} + \quad _n$, is

approximated by $\hat{\tilde{}}_n = \hat{}_{n-1} + \hat{\tilde{}}_n$, where $\hat{\tilde{}}_n = (\quad\quad_n)^{-1}\quad \hat{}_{n-1}$ is the iterated solution of the regularized equation $\quad_n = \quad \hat{}_{n-1} + \quad\quad_n$.

This gives

$$\hat{}^{n+1}_{rekant} \qquad\qquad \hat{}^{n}_{rekant}$$

$$= \qquad\qquad {}^{n+1}\ \hat{\tilde{}}_0$$

$$\qquad\qquad {}^{n+1}\qquad\quad \tilde{}_0$$

Note that the operator controlling the convergence of this method is $\quad$, rather than $\quad$ produced by the re-iterated Galerkin and modified iterated Kantorovich methods. This gives us a new condition on the size of the subspace $\quad_n$. We now require that $\quad_n$ is chosen such that $\quad\quad 1$.

As commented on previously, these norms give a weaker bound on the con-

$$\frac{\quad_n}{\quad_{n-1}} \qquad\qquad \quad_n \qquad\quad \quad_n^{-1}$$

$$\quad^i_{gal} \qquad 0 \qquad\quad i$$

$$\quad^i_{kant} \qquad\quad 0 \qquad\quad i$$

$$\quad^i_{rekant} \qquad 0 \qquad\quad i$$

Then, for i=0,

$$\begin{aligned}
\hat{\sigma}^0_{gal} &= f + Kp_0, \\
\hat{\sigma}^0_{kant} &= f + Kf + Kq_0, \\
\hat{\sigma}^0_{rekant} &= f + Kf + K\tilde{q}_0,
\end{aligned}$$

and for i=1,

$$\begin{aligned}
\hat{\sigma}^1_{gal} &= f + Kf + K^2p_0 + Kp_1, \\
\hat{\sigma}^1_{kant} &= f + Kf + K^2f + K^2q_0 + Kq_1, \\
\hat{\sigma}^1_{rekant} &= f + Kf + K^2f + K^3f + K^3\tilde{q}_0 + K\tilde{q}_1.
\end{aligned}$$

In general, for i=n,

$$\begin{aligned}
\hat{\sigma}^n_{gal} &= \sum_{j=0}^{n} K^j f + \sum_{j=0}^{n} K_{j+1} p_{n-j}, \\
\hat{\sigma}^n_{kant} &= \sum_{j=0}^{n+1} K^j f + \sum_{j=0}^{n} K_{j+1} q_{n-j}, \\
\hat{\sigma}^n_{rekant} &= \sum_{j=0}^{2n+1} K^j f + \sum_{j=0}^{n} K^{2j+1} \tilde{q}_{n-j}.
\end{aligned}$$

From this we can see that after n iterations, the re-iterated Galerkin approximation contains the first n+1 terms of the Neumann series for $\phi = f + K\phi$, the modified iterated Kantorovich approximation contains the first n+2 terms, and the re-iterated Kantorovich approximation contains the first 2n+2 terms of the Neumann series, twice as many as for the re-iterated Galerkin method. This property appears to suggest that, under suitable conditions, the re-iterated Kantorovich method will have a much faster rate of convergence than the other two methods. In reality, if $\|K\|$ is greater than one, the Neumann series diverges, and it is the other terms which have to compensate. The Neumann series itself cannot therefore be responsible for the rate of convergence.

In the following chapters we will be comparing the efficiency and accuracy of the re-iterated Galerkin method with the new re-iterated Kantotovich method. The modified iterated Kantorovich method has not been considered further due to the belief that any improvement over the re-iterated Galerkin method will be minimal.

Ch pter

Numeric l Consider tions

4.1 Numerical Integration

The process of calculating successive corrections to previous approximations, a feature of all the re-iterative methods, ensures that the accuracy of the numerical solution can be increased until it is of the same order as the computational accuracy of the method. Once the residual $\hat{r}_n$ reaches this accuracy there is no point in continuing the iteration, as no further improvement can be made. The main factor governing the computational accuracy is the method of numerical integration used. In all the cases described here modified Gauss-Legendre quadrature was employed, using 10 points and 20 subintervals. In this way, the value of every function used in the method is known at each of the two hundred Gauss points in the interval.

Numerical integration routines tend to encounter difficulties when dealing with kernels containing slope discontinuities, such as

$$(Kf)(x) = \int_0^1 max(x,t)f(t)dt, \qquad 0 \leq x \leq 1,$$

which has such a discontinuity at $= .$

The problem may be remedied by noting that

$$(\quad)(\ ) = \int_0^1 (\quad)\ (\ )$$
$$= \int_0^1 (\quad)(\ (\ ) \quad (\ )) + (\ ) \int_0^1 (\quad)$$

This reformulated equation has a first term with a continuous first derivative, and a second term that may be evaluated exactly. (See Chamberlain [1], who used this device in the case where $(\quad) = \frac{1}{2\alpha} \quad (\qquad)\ (\ ))$.

In order to keep calculations as simple as possible, attention is restricted to finding approximations in subspaces of minimal dimension. Provided that the subspace is of sufficient size to ensure the conditions $(\ ) \quad 1$ and $(\quad) \quad 1$ are satisfied,

$$\sum_{n=1}^{\infty} \quad n-1$$

$$n \qquad 2 \qquad n-1$$

would be a good approximation to . For this reason we conclude that we ought to be looking for approximations belonging to the n-dimensional subspace spanned by ^{2} ... $^{n-1}$. Indeed, it has been shown (Porter Stirling [3]) that this constitutes the optimal choice of subspace.

For our one-dimensional case, therefore, we are looking for Galerkin approximations of the form = $_1$ $_1$, where $_1$ is chosen to be the free term of the integral equation. For the re-iterated Galerkin method this is obviously $_1$ = . What is perhaps not so obvious is what we mean by the free term for the Kantorovich method.

Whilst it is true we are trying to find satisfying = + , which would suggest taking $_1$ = , the Kantorovich regularization effectively removes from the approximation, so that we are instead seeking to approximate , where satisfies = + . Here the free term is not , but , so we should perhaps take our trial function to be = instead.

Since the aim of this project is to give a comparison of the different methods, it seems sensible that they should use the same subspaces. On the other hand, each

	0	1	2	3	4	$\frac{\|\hat{r}_n\|}{\|\hat{r}_{n-1}\|}$	$\frac{\|\hat{r}_n\|}{\|r_n\|}$
$\hat{}^0_{rekant}$	2 500001	2 627627	3 032719	3 785712	4 996088		0 356
$\hat{}^2_{rekant}$	1 813364	1 927510	2 283396	2 926562	3 932512	0 179	0 354
$\hat{}^4_{rekant}$	1 791510	1 905215	2 259509	2 899160	3 898587	0 179	0 354
$\hat{}^6_{rekant}$	1 790809	1 904501	2 258744	2 898282	3 897500	0 179	0 354
$\hat{}^8_{rekant}$	1 790787	1 904478	2 258719	2 898254	3 897465	0 179	0 354
$\hat{}^9_{rekant}$	1 790786	1 904477	2 258718	2 898253	3 897464	0 179	0 354
( )	1 790786	1 904477	2 258718	2 898253	3 897464		

Table 5.2: Re-iterated Kantorovich approximations (=)

$\frac{\|\hat{r}_n\|}{\|\hat{r}_{n-1}\|}$

$\frac{\|\hat{r}_n\|}{\|r_n\|}$

	0	1	2	3	4	$\frac{\|\hat{r}_n\|}{\|\hat{r}_{n-1}\|}$	$\frac{\|\hat{r}_n\|}{\|r_n\|}$
0_{rekant}							
2_{rekant}							
4_{rekant}							
6_{rekant}							

the projection P_n, and thus the choice of subspace, is approximated as 0·058. This is smaller still than is obtained by taking $\chi = f$. As we might expect, therefore, the approximations to ϕ, given by $\hat{\sigma}^n_{rekant} = \hat{\hat{p}}_0 + \cdots + \hat{\hat{p}}_n$, converge to an accuracy of six decimal places after only seven iterations. (Naturally the rate of convergence for all three methods would be improved by taking a higher dimensional subspace.)

The value given here for $\|K\|$ is noticeably different from that given in the two previous tables. It isn't immediately obvious why this should be; K hasn't changed, so the reason must be linked to the choice of subspace. In order to see what is happening we need to look at the Rayleigh quotient, denoted $R(\phi)$.

If T is a bounded, self-adjoint linear map from a Hilbert space, H, to itself, then the Rayleigh quotient

$$R_T(\phi) = \frac{(T\phi, \phi)}{(\phi, \phi)},$$

gives a lower bound on the largest positive eigenvalue of T, denoted $\mu_1^+(T)$. That is

$$\begin{aligned} \mu_1^+(T) &= \sup\{R_K(\phi) : \phi \neq 0\} \\ &= \max\{R_K(\phi) : \phi \neq 0\} : \mu_1^+(T) \neq 0. \end{aligned}$$

Also

$$\|T\| = \max\{|R_K(\phi)| : \phi \neq 0\}.$$

(See Porter Stirling [3] Lemma 5.1)

If we take $\phi = r_n$, and $T = K^2$, we see that

$$R_K(r_n) = \frac{(Kr_n, Kr_n)}{(r_n, r_n)} \leq \|K\|^2 = \mu_1^+(K^2),$$

where ${}^{+}_{1}$ denotes the largest positive eigenvalue of , with ${}^{+}_{1}$ being the corresponding eigenvector. This implies that

$$\frac{{}_{n}{}^{2}}{{}_{n}{}^{2}} = \frac{\hat{}_{n}{}^{2}}{{}_{n}{}^{2}} \qquad {}^{2}$$

${}_{n}$ ${}^{+}_{1}$

$\frac{||\hat{r}_n||}{||r_n||}$ ${}_{n}$ ${}_{0}$ ${}_{0}$

${}_{n}$

${}_{n}$ ${}^{+}_{1}$

$\frac{}{0}$

${}_{0}$ $\frac{1}{{}_{0}\;\;{}_{0}}$ ${}_{0}$

${}_{0}$

${}_{i}$ ${}_{i}$

where

$$i \qquad 0$$

$$0$$

$$1$$

$$0 \qquad \frac{}{0} \qquad 0$$

$$0$$

$$'' \qquad \begin{matrix}2\\0\end{matrix}$$

$$0 \qquad 0 \qquad 0 \qquad 0 \qquad ' \qquad 0 \qquad 0$$

$$'(1) \qquad 0 \qquad 0 \qquad 0 \qquad ' \qquad 0 \qquad 0$$

$$\frac{0}{} \qquad 0 \qquad 0 \qquad 0$$

$$\frac{0}{} \qquad 0 \qquad 0 \qquad 0$$

$$\frac{}{\begin{matrix}2\\0\end{matrix}} \qquad 0$$

$$\begin{matrix}2\\0\end{matrix} \qquad 2 \qquad 0 \qquad 0 \qquad 0$$

1

1

The results in the previous chapter lead us to wonder in what cases the Kantorovich based re-iterative methods are better, in the sense of convergence rates,

$$\frac{||\hat{r}_n||}{||\hat{r}_{n-1}||}$$

6.1 The ernel $2\lambda\max(x,t)$

The first equation we consider is

$$\phi(x) = 1 + \lambda \int_0^1 2\max(x,t)\phi(t)dt, \qquad 0 \leq x \leq 1. \tag{6.1}$$

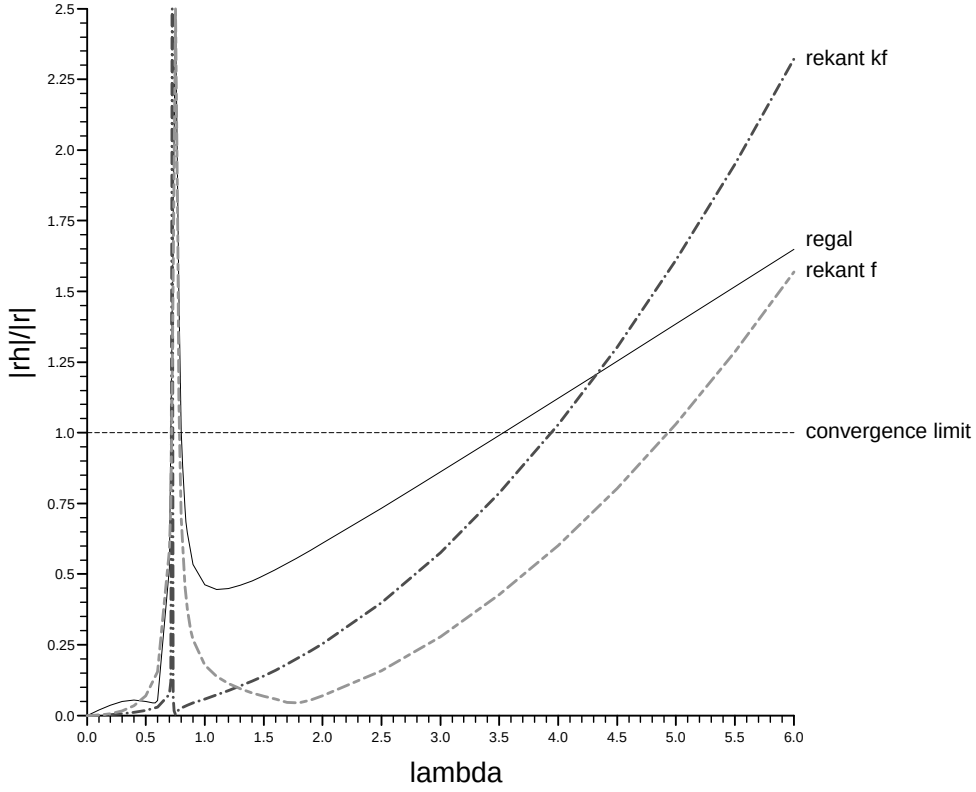


Figure 6.1: Convergence of methods for $k(x,t) = 2\lambda\max(x,t)$

Figure (6.1) illustrates the variation in $\frac{||\hat{r}_n||}{||\hat{r}_{n-1}||}$ (written as $|rh|/|r|$ in the Figure) for each of the methods, with λ taken in the range, $0 \leq \lambda \leq 6$. The most striking feature of the graph is the spike which occurs at around $\lambda = 0 \cdot 7$. It seems reasonable to suppose that this spike in some way represents the presence of an eigenvalue, since the ratio of the residual norms approximates the spectral radii of the operators S and SK.

There is a process by which it is occasionally possible to calculate, analytically,

Finally,

$$S = (I \quad KP_n)^{-1}(K \quad KP_n) = I \quad (I \quad KP_n)^{-1}(I \quad K),$$

so

$$\begin{aligned}
(Sf)(x) &= f(x) \quad (f(x) \quad (Kf)(x)) + \int_0^1 c\lambda(1+x^2)\ f(t) \quad (Kf)(t)\ dt \\
&= (Kf)(x) + c\lambda(1+x^2) \int_0^1 f(t)dt \quad \int_0^1 f(t)dt \int_0^1 2\lambda \max(s,t)ds \\
&= \int_0^1 2\lambda \max(x,t) + c\lambda(1+x^2) \quad c\lambda^2(1+x^2)(1+t^2)\ f(t)dt,
\end{aligned}$$

$$\text{ie. } (Sf)(x) = \int_0^1 2\lambda \max(x,t) + c\lambda(1+x^2)(1 \quad \lambda \quad \lambda t^2)\ f(t)dt.$$

To find the eigenvalues of S, let ϕ now satisfy $\phi = \Lambda S\phi$,

$$\begin{aligned}
ie. \qquad \phi(x) &= 2\Lambda\lambda \int_0^x x\phi(t)dt + 2\Lambda\lambda \int_x^1 t\phi(t)dt \quad c\Lambda\lambda(1+x^2) \int_0^1 (\lambda t^2 + \lambda \quad 1)\phi(t)dt \\
\phi'(x) &= 2\Lambda\lambda \int_0^x \phi(t)dt + 2x\Lambda\lambda x\phi(x) \quad 2cx\Lambda\lambda \int_0^1 (\lambda t^2 + \lambda \quad 1)\phi(t)dt \\
\phi''(x) &= 2\Lambda\lambda\phi(x) \quad 2c\Lambda\lambda \int_0^1 (\lambda t^2 + \lambda \quad 1)\phi(t)dt
\end{aligned}$$

where $\phi'(0) = 0 \qquad \phi'(1) = \quad \phi(1).$

Now,

$$\begin{aligned}
\int_0^1 (\ ) \quad &= \Lambda \int_0^1 (\ ) \int_0^1 2 \ \max(\quad) \quad (\ ^2 + \quad 1)(1 + \ ^2) \\
&= \frac{1}{3}\ \Lambda \int_0^1 (1 + 3\ ^2)\ (\ )
\end{aligned}$$

Hence

$$(1 \quad \frac{1}{3}\ \Lambda\ ) \int_0^1 (\ ) \quad = \quad \Lambda \int_0^1 \ ^2 \ (\ )$$

and so,

$$''(\ ) \quad 2\Lambda \quad (\ ) = 2\ (1 \quad \frac{1}{3}\ \Lambda\ ) \int_0^1 (\ ) \quad 2\ \Lambda\ (\quad 1) \int_0^1 (\ ) \quad = 2$$

where $= (1 \quad \Lambda) \int_0^1 (\ )$.

Suppose $\Lambda \quad 0$, and write $2\Lambda = \quad ^2 (\quad \mathbb{R})$.

Then

$$( ) = \cos(\quad) + \sin(\quad) \frac{}{\Lambda}$$

$$= \cos(\quad) \frac{}{\Lambda}$$

(since $'(0) = 0$). Therefore

$$\int_0^1 ( ) = -\sin \frac{}{\Lambda}$$

$$= (1 \quad \Lambda)\frac{a}{\omega}\sin \quad (1 \quad \Lambda)\frac{\xi}{\Lambda\lambda}$$

Λ-Pxj8jCA8ff$\frac{\Omega}{\omega}$C

′

—

—

$\frac{1-\Lambda}{\omega}$ $\frac{1}{2\lambda}$ 2

— 2 2

2 2

n n

$_n^{-1}$ $\frac{1}{2\lambda}$ $_n^2$ n $\frac{-2\lambda}{\hat{\omega}_n^2}$ 1

2

2 2

Let the roots of this equation be $\check{}_n$ ($\check{}_n$ 0) for IN, and denote the smallest root by $\check{}_1$.

The spectral radius of the operator is therefore given by

$$( ) = \frac{2}{\frac{2}{1}} \tag{6 4}$$

where $_1 = \min(\hat{}_1 \ \check{}_1)$.

In order to determine the minimum of $\hat{}_1$ and $\check{}_1$ it is necessary to explore the behavior of the roots of equations (6.2) and (6.3). Such analysis is quite involved and so only the results are given here.

For 0 0 5 equation (6.3) has no non-trivial roots, unlike equation (6.2), and so we take $_1 = \hat{}_1$, giving ( ) $= \frac{2\lambda}{\hat{\omega}_1}$ which increases linearly as tends to 0 5.

For 0 5 0 75 equation (6.3) has one root which starts off being very large, but which tends to zero like $\frac{1}{\epsilon}$ (where) as 0 75. This results in ( ) as $\check{}_1$ 0, giving ( ) $= \frac{2\lambda}{\tilde{\omega}_1}$.

$\frac{2\lambda}{\hat{\omega}_1}$

$\frac{||\hat{r}_n||}{||\hat{r}_{n-1}||}$

eigenvalue of the operator K (ie. for $\lambda = \frac{1}{\mu_n}$, where μ_n is an eigenvalue of K), the numerical method would not be able to find an approximation and hence would not converge. This would result in the ratio of the residual norms, and therefore $\rho(S)$ and $\rho(SK)$, exceeding the convergence limit, and in fact, becoming infinite. If so, the spikes in Figure (6.1) would appear to mark the presence of an eigenvalue of the kernel operator K, given by $\mu = \frac{1}{\lambda}$. Since λ is positive in the examples, and K has only one positive eigenvalue, the graphs seem to provide a means of obtaining the maximum positive eigenvalue of K.

Unfortunately the graphs do not tend to infinity at the same point. The RIG and RIK($\chi = f$) methods both become infinite at $\lambda = 0 \cdot 75$, whilst the RIK($\chi = Kf$) method becomes infinite at a value closer to $\lambda = 0 \cdot 72$. The problem is that we are not actually solving $\phi = f + K\phi$, but rather an approximate equation lying in a reduced space. The eigenvalues we are locating, therefore, are not of the original equation, but belong instead to the approximate equation.

The approximate eigenvalues arise from the initial Galerkin method and are unaltered by iteration or re-iteration. Thus, in the one-dimensional case with $\chi = f$, the Galerkin equation is $\alpha\{\|f\|^2 - \lambda(Kf, f)\} = \|f\|^2$, which produces the approximate eigenvalue $\tilde{\mu}_1 = \frac{(Kf,f)}{\|f\|^2}$. Since K is self-adjoint we can deduce that $\mu_1^- \leq \tilde{\mu}_1 \leq \mu_1^+$ by considering the Rayleigh quotient. The eigenvalue apparent in the graphs is therefore an underestimate of the maximum positive eigenvalue of K. When we change the subspace, as for the RIK($\chi = Kf$) method, we also change the value of $\tilde{\mu}_1$, in this case to $\tilde{\mu}_1 = \frac{\|Kf\|^2}{(Kf,f)}$. Since $\frac{(Kf,f)}{\|f\|^2} \leq \frac{\|Kf\|^2}{(Kf,f)}$, this second $\tilde{\mu}_1$ will be no less than (and in practice greater than) the first $\tilde{\mu}_1$. Hence, taking $\chi = Kf$ produces a closer underestimate to the maximum positive

eigenvalue of K than can be achieved by taking $\chi = f$.

This result is echoed in the graphs, where the RIK($\chi = f$) and RIG (also with $\chi = f$) methods offer $\tilde{\mu}_1 = 1 \cdot 33$, corresponding to $\lambda = 0 \cdot 75$, whilst the RIK($\chi = Kf$) method gives $\tilde{\mu}_1 = 1 \cdot 39$, since $1 \cdot 38 \leq \mu_1^+ \leq 1 \cdot 42$ can be obtained by standard eigenvalue estimation methods, as mentioned earlier.

Away from the eigenvalue, both of the RIK methods exhibit better convergence properties than the RIG method, as illustrated by their lower $\frac{||\hat{r}_n||}{||\hat{r}_{n-1}||}$ values, which in turn lead to increased intervals of convergence. Although the RIK graphs intersect the RIG graph, they do not do so until all the methods have exceeded the limit for convergence. For this particular problem we can conclude that the RIK methods exhibit consistently better convergence properties than the RIG method, and that, away from the eigenvalue, taking $\chi = f$ constitutes a better choice of subspace than $\chi = Kf$.

6.2 The ernel $\frac{\lambda}{2\kappa_0} sin(\kappa_0|x-t|)$

The second equation we consider is the general form of equation (5.2), ie.

$$\phi(x) = \cos(\kappa_0 x) + \lambda \int_0^1 \frac{1}{2\kappa_0} \sin(\kappa_0|x-t|)\phi(t)dt, \qquad 0 \leq x \leq 1 \tag{6.5}$$

where κ_0 is a parameter (taken to be $\kappa_0 = 1 \cdot 5$ in the following example).

Unlike equation (6.1), finding both $\rho(S)$ and$\rho(SK)$ analytically proves too difficult, so we have to rely entirely on the numerical results in order to gain information about the convergence properties of the methods for this kernel. Figure (6.2) shows the graphs of $\frac{||\hat{r}_n||}{||\hat{r}_{n-1}||}$ against λ for the three methods when

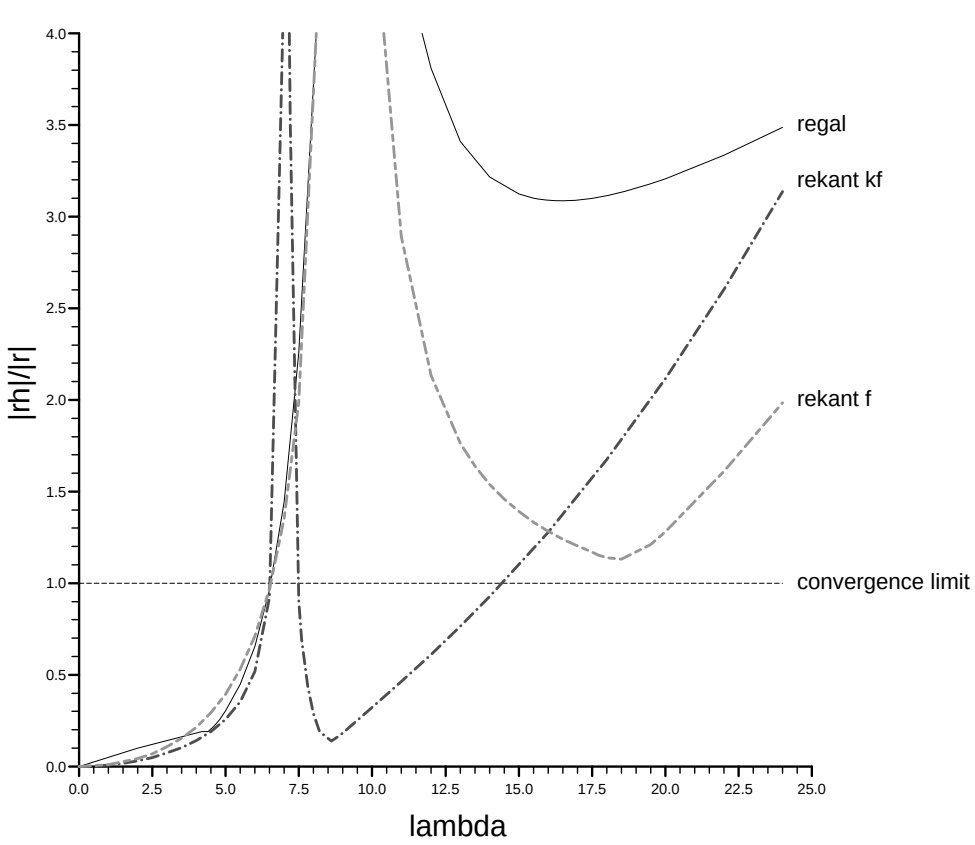


Figure 6.2: Convergence of methods for ( ) = $\frac{\lambda}{2\ _0}$ sin($_0$)

approximating equation (6.5). As before we notice discontinuities in the results, pointing to the presence of approximate eigenvalues. This time, however, the singularities occur at markedly different values for the methods with different subspaces. The = methods become infinite at around = 9 1, whilst the = method becomes infinite at approximately = 7 1. As in the previous example, we can show that the latter choice of subspace provides the closest underestimate to the maximum positive eigenvalue of , giving us $\tilde{}_1 = \frac{1}{\lambda}$ 0 14 $^+_1$.

gives

$$\begin{aligned} {}^{+}_{1} \quad & \min(0 \quad 173704 + \frac{{}^{2}_{0}}{24 \quad \overline{7}} \quad \frac{1}{2 \quad 2 \; {}^{2}_{0}} \quad \overline{{}^{2}_{0} \quad \sin^2 \quad {}_0}) \\ {}^{-}_{1} \quad & \max(\quad {}^{-2} \quad \frac{1}{2 \quad 2 \; {}^{2}_{0}} \quad \overline{{}^{2}_{0} \quad \sin^2 \quad {}_0}) \end{aligned}$$

ie. taking ${}_0 = 1 \quad 5$ we have

$$\begin{aligned} {}^{+}_{1} \quad & 0 \quad 17604 \\ {}^{-}_{1} \quad & 0 \quad 10132 \end{aligned}$$

Our approximation, $\check{}_1$, certainly lies below Chamberlain's upper bound, and since it is an underestimate, provides us with a lower bound on ${}^{+}_{1}$.

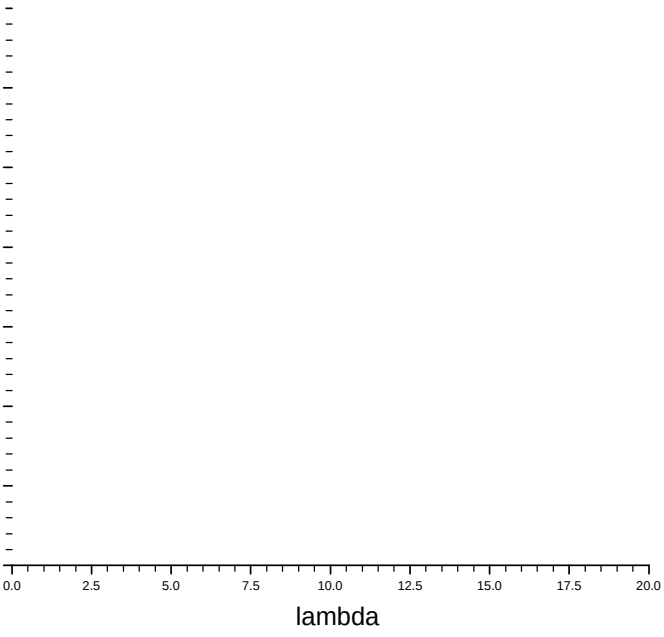
All three methods, in this example, demonstrate convergence up to $\quad 6 \quad 5$, where the effect of the approximate eigenvalue becomes too great. Neither the

$$\frac{n}{n \quad 1}$$

$$\frac{n}{n \quad 1}$$

$$n$$

$$\frac{}{1}$$


0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
lambda

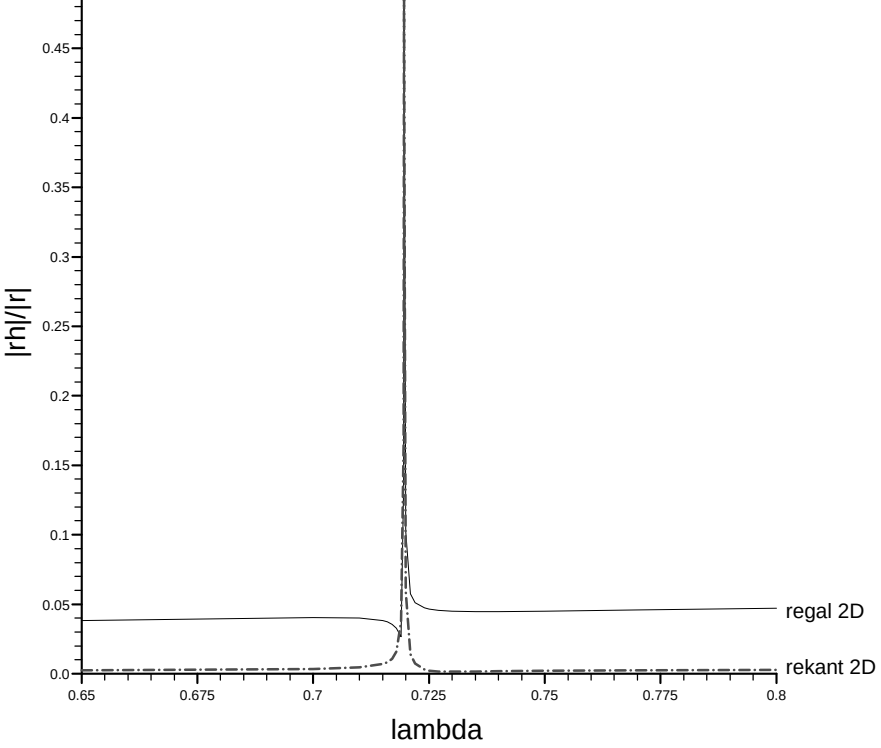
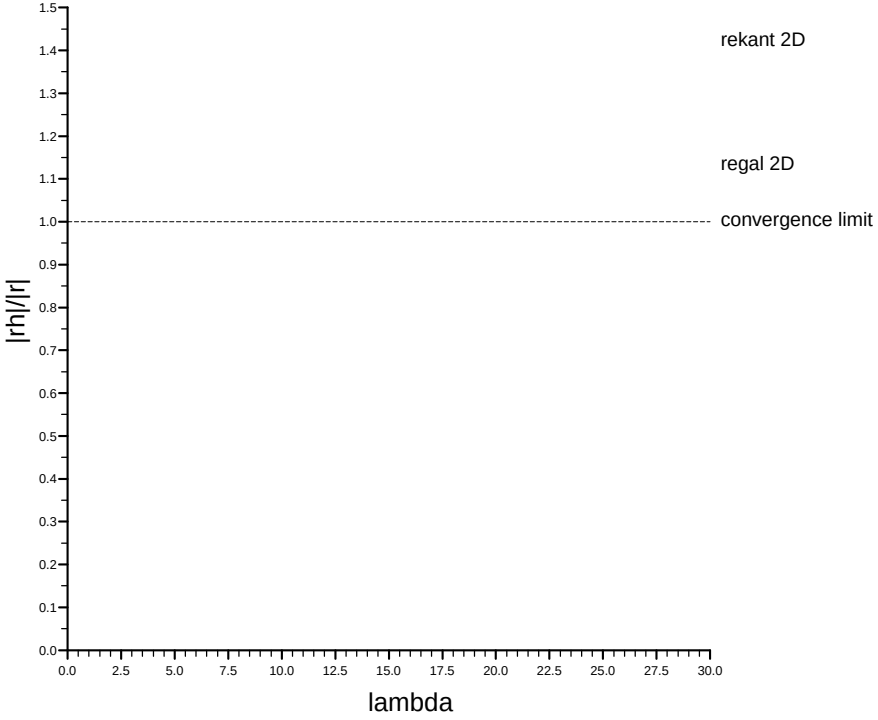


Figure 7.2: The eigenvalue of ( ) = 2 max( ) for a 2-D subspace

magnification in Figure (7.2), the value of at which the approximate eigenvalue occurs can be pinpointed at = 0 71967, corresponding to ˜ 1 1 38949. The disruption caused to the convergence of the methods by the presence of the eigenvector is greatly reduced, allowing us to see more clearly the linear nature of the RIG method, and the quadratic nature of the RIK method. It also demonstrates that the RIG method, which was the most disrupted in the one-dimensional case, may converge faster than the RIK method for certain values of , and hence for certain kernels. The point of cross-over in this example does not, however, occur until $\frac{\|\hat{r}_n\|}{\|\hat{r}_{n-1}\|}$ is significantly greater than 0 75, a value at which, in a practical situation, we would be looking to utilise an even higher dimensional subspace.

Finally, Figure (7.3) illustrates the effect of moving to a two-dimensional subspace for equationequCBjjF6-Figu2R8n-N8FC-er,xffC62rxffCrxffC8-casuCBjjF6-Figu2R8n-e-ix6FN

AuAiu


rekant 2D
regal 2D
convergence limit
|rh|/|r|
lambda
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
1.1
1.2
1.3
1.4
1.5
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
22.5
25.0
27.5
30.0

Ch pter 8

Conclusions

The results in the previous two Chapters suggest that the new re-iterated Kantorovich method will converge, and do so faster than the re-iterated Galerkin method, if

$$\rho(SK) < \rho(S) \leq 1.$$

When $\|K\|$ is large, ie. when λ is large in the previous graphs, the RIG method can display the better convergence properties. However, if the view is taken that values of $\frac{\|\hat{r}_n\|}{\|\hat{r}_{n-1}\|}$ greater than $0 \cdot 5$, corresponding to roughly forty iterations, signal the need for a higher dimensional subspace in which to seek a solution, then the RIK method represents the best practical method of approximation. Naturally, increasing the size of the subspace, and therefore introducing more degrees of freedom into the approximation, has the effect of increasing the rate of convergence, leading to increased regions of convergence.

The presence of approximate eigenvalues can cause a fair amount of disruption to the methods. This is especially true when using a subspace of low dimension, where the number of terms in the expansion of the approximation available to

compensate for the rapid growth of the residuals, is low. It has been shown that choosing a subspace spanned by $\quad = \quad$, rather than by $\quad = \quad$, for the one-dimensional case, not only gives a closer underestimate to the exact eigenvalue $\quad_1^+$ of the problem, but also reduces the effect of the approximate eigenvalue on convergence. The position of the approximate eigenvalue may be calculated beforehand and so be avoided. This is done by considering the initial Galerkin equation, given by (in 1-D case)

$$\quad_1 \quad \quad_1{}^2 \quad (\quad \quad_1 \quad \quad_1) \; = (\quad \quad_1)$$

hence the eigenvalue, $\tilde{\quad}_1$, is given by

$$\tilde{\quad}_1 = \frac{(\quad \quad_1 \quad \quad_1)}{\quad_1{}^2}$$

where $\quad_1$ is the choice of trial function. Also to be avoided, naturally, are the

Whereas the estimates of ( ) proved to be very accurate, a similar degree of accuracy could not be attributed to the underestimates of , given by $\frac{\|\hat{r}_n\|}{\|r_n\|}$ for both methods. These were shown to be very poor in comparison, depending upon the residual, $_n$, being a close approximation to a constant multiple of the eigenvector $^+_1$.

There is obviously a need for more exhaustive testing of the re-iterated Kantorovich method, as well as more in-depth analysis, before its strengths and weaknesses are fully understood. In the limited time available, it has only been possible to explore some basic ideas in relation to a few particular examples and many issues have arisen which require closer analysis and examination. Some obvious features that have not been explored in this report include: a closer study

$\frac{1}{2}_0$ 0 0

0

iterated Galerkin method could be exploited in variational principles. It would