Numer. Math. 32, 139–146 (1979)

Numerische Mathematik

Error Bounds and Estimates for Eigenvalues of Integral Equations. II

Alastair Spence
School of Mathematics, University of Bath, Claverton Down, Bath, BA2 7AY, England

Summary. In a previous paper computable error bounds and dominant error terms are derived for the approximation of simple eigenvalues of non-symmetric integral equations. In this note an alternative analysis is presented leading to equivalent dominant error terms with error bounds which are quicker to calculate than those derived previously.

Subject Classifications. AMS (MOS): 65R05; CR: 5.18.

1. Introduction

In earlier papers the author discussed the numerical solution of the integral equation eigenvalue problem using the Nyström method. The equation considered is

$$\lambda x(s) = \int_a^b k(s,t)\, x(t)\, dt \tag{1.1}$$

or, in operator notation,

$$\lambda x = K x \tag{1.2}$$

and the matrix equation derived by the Nyström method is

$$\nu^{(n)} \mathbf{u} = \mathbf{K}_n \mathbf{u}. \tag{1.3}$$

See [2] and [3] for details. Briefly, equation (1.2) is shown to be equivalent to

$$\lambda \mathbf{x} = [\mathbf{K}_n + \mathbf{B}_n(\lambda)]\, \mathbf{x} \tag{1.4}$$

where $\mathbf{B}_n(\lambda)$ is given by (5)–(7) of [3]. $\mathbf{B}_n(\lambda)$ is regarded as a perturbation of $\mathbf{K}_n$ and the solutions of (1.4) and (1.3) are compared using perturbation theory, modified to deal with the nonlinear λ-term in $\mathbf{B}_n(\lambda)$. In [3] computable error

0029-599X/79/0032/0139/$01.60

bounds ((25) and (27)) and dominant error terms ((30) and (33)) were derived for the approximation of simple eigenvalues and eigenvectors. The bounds and dominant error estimates were numerically very good but were expensive to calculate because of the need to know all the eigenvalues and left and right eigenvectors of $\mathbf{K}_n$ in (1.3). The purpose of this note is to present an alternative theory which gives rise to simpler expressions for the error bounds and estimates which are easier to calculate than those in [3].

The approach of this note is as follows. Assume that $\nu^{(n)}$ and $\mathbf{u}$, solutions of (1.2) have been calculated, and that $\nu^{(n)}$ is simple. For large enough n, $\nu^{(n)} \approx \lambda$ and $\mathbf{u} \approx \mathbf{x}$ (assuming a suitable normalization of $\mathbf{u}$ and $\mathbf{x}$). We are interested in estimating the quantitites $|\lambda - \nu^{(n)}|$ and $\|\mathbf{x} - \mathbf{u}\|$ for a given n without the calculation of any other eigenvalues or eigenvectors of (1.2). In §2 we derive theoretical results and illustrate them in §3 with reference to the trapezoidal rule method. In §4 we comment on the time taken to calculate the bounds and estimates.

2. Theoretical Error Bounds

Assume that ν and $\mathbf{u}$ satisfy (1.3) with ν a simple eigenvalue. (For convenience we shall drop the superscript on $\nu^{(n)}$ in this section.) Normalise $\mathbf{u}$ by putting

$$\|\mathbf{u}\|_\infty = 1 \tag{2.1}$$

and, for convenience only, assume that the largest component is the first and is real i.e.

$$\mathbf{u}^T = [1, \boldsymbol{\xi}_{n-1}^T]. \tag{2.2}$$

(Note that since $k(s,t)$ is not necessarily symmetric we must allow for complex eigenvalues and eigenvectors.) Since ν and $\mathbf{u}$ are approximations to λ and $\mathbf{x}$ we write

$$\lambda = \nu + \mu \tag{2.3}$$

and

$$\mathbf{x} = \mathbf{u} + \mathbf{z} \tag{2.4}$$

where

$$\mathbf{z}^T = [0, \boldsymbol{\eta}_{n-1}^T]. \tag{2.5}$$

Our aim is to bound and estimate μ and $\boldsymbol{\eta}_{n-1}$. Equation (1.4) becomes

$$[\mathbf{K}_n + \mathbf{B}_n(\lambda)](\mathbf{u}+\mathbf{z}) = (\nu+\mu)(\mathbf{u}+\mathbf{z})$$
$$\therefore\ \mu\mathbf{u} - [\mathbf{K}_n - \nu\mathbf{I}_n]\mathbf{z} = \mathbf{B}_n(\nu+\mu)(\mathbf{u}+\mathbf{z}) - \mu\mathbf{z}. \tag{2.6}$$

Assume that $\mathbf{K}_n$ is partitioned in the form

$$\mathbf{K}_n = \left[\begin{array}{c|c} k_{11} & \mathbf{d}_{n-1}^T \\ \hline \mathbf{c}_{n-1} & \mathbf{A}_{n-1} \end{array}\right] \tag{2.7}$$

and, using (2.5) and (2.7), (2.6) becomes

$$\mathbf{T}_n \left[\begin{array}{c} \mu \\ \hline \boldsymbol{\eta}_{n-1} \end{array}\right] = \mathbf{B}_n(\nu+\mu) \left[\begin{array}{c} 1 \\ \hline \boldsymbol{\xi}_{n-1} + \boldsymbol{\eta}_{n-1} \end{array}\right] - \mu \left[\begin{array}{c} 0 \\ \hline \boldsymbol{\eta}_{n-1} \end{array}\right] \tag{2.8}$$

where

$$\mathbf{T}_n = \left[\begin{array}{c|c} 1 & -\mathbf{d}_{n-1}^T \\ \hline \boldsymbol{\xi}_{n-1} & \nu \mathbf{I}_{n-1} - \mathbf{A}_{n-1} \end{array}\right]. \tag{2.9}$$

Equation (2.8) is a system of n nonlinear equations for the n unknowns μ and $\boldsymbol{\eta}_{n-1}$. The solvability of (2.8) depends on the nonsingularity of $\mathbf{T}_n$ and it is easy to prove that "if ν is a simple eigenvalue of $\mathbf{K}_n$ then $\mathbf{T}_n$ is nonsingular".

Equation (2.8) can be written

$$\gamma = \mathbf{c}(\gamma) + \mathbf{H}(\gamma)\gamma + \mathbf{M}(\gamma)\gamma \tag{2.10}$$

where

$$\gamma^T = [\mu, \boldsymbol{\eta}_{n-1}^T] \qquad \mathbf{c}(\gamma) = \mathbf{T}_n^{-1}\mathbf{B}_n(\nu+\mu)\mathbf{u}$$
$$\mathbf{H}(\gamma)\gamma = \mathbf{T}_n^{-1}\mathbf{B}_n(\nu+\mu)\mathbf{z} \qquad \mathbf{M}(\gamma)\gamma = -\mu\mathbf{T}_n^{-1}\mathbf{z}. \tag{2.11}$$

The Lemma in §2 of [3] will now be used to bound the solution of (2.10). The constants C_0, C_1, H_0, H_1, M_1 needed by the lemma are

$$C_0 = \|\mathbf{c}(\mathbf{0})\| = \|\mathbf{T}_n^{-1}\mathbf{B}_n(\nu)\mathbf{u}\|, \tag{2.12}$$

$$C_1 \leqq \max_{|\mu|\in D} \|\mathbf{T}_n^{-1} r_n K(\nu+\mu-Q_n)^{-2} Q_n p_n \mathbf{u}\|, \tag{2.13}$$

$$H_0 = \|\mathbf{H}(\mathbf{0})\| = \|\mathbf{T}_n^{-1}\mathbf{B}_n(\nu)\|, \tag{2.14}$$

$$H_1 \leqq \max_{|\mu|\in D} \|\mathbf{T}_n^{-1} r_n K(\nu+\mu-Q_n)^{-2} Q_n p_n\| \tag{2.15}$$

and

$$M_1 \leqq \|\mathbf{T}_n^{-1}\|. \tag{2.16}$$

A direct application of the lemma gives the following bounds.

Theorem. Let ν be a simple eigenvalue of $\mathbf{K}_n$ with corresponding eigenvector $\mathbf{u}$. Define C_0, C_1, H_0, H_1 and M_1 as in (2.12) to (2.16) respectively, and let $w(g) = (1-\sqrt{1-2g})/g$.

If

$$1 - C_1 - H_0 > 0 \tag{2.17}$$

and

$$g=\frac{2(H_1+M_1)C_0}{(1-C_1-H_0)^2}<\frac{1}{2} \tag{2.18}$$

then

$$|\lambda-\nu|\leqq\frac{C_0}{1-C_1-H_0}w(g) \tag{2.19}$$

and

$$\|\mathbf{x}-\mathbf{u}\|\leqq\frac{C_0}{1-C_1-H_0}w(g). \tag{2.20}$$

Now if we assume that $\|\mathbf{T}_n^{-1}\|\leqq T$, independent of n, as $n\to\infty$ then (2.12) implies that

$$C_0=O(\|\mathbf{B}_n(\nu)\mathbf{u}\|),$$

and (2.19) and (2.20) give

$$|\lambda-\nu|=O(\|\mathbf{B}_n(\nu)\mathbf{u}\|),\quad \|\mathbf{x}-\mathbf{u}\|=O(\|\mathbf{B}_n(\nu)\mathbf{u}\|). \tag{2.21}$$

The assumption that $\|\mathbf{T}_n^{-1}\|\leqq T$ is equivalent to the assumption that ν tends to a simple eigenvalue of K. Thus the fact that $\|\mathbf{T}_n^{-1}\|$ is large only when ν is close to another eigenvalue of $\mathbf{K}_n$ implies that $\|\mathbf{x}-\mathbf{u}\|$ is not perturbed significantly by multiple eigenvalues well separated from ν. This result was not obvious from the approach of [3].

The bounds (2.19) and (2.20) can be sharpened in a straightforward manner as follows. From (2.10) we have

$$\begin{aligned}&\gamma-\mathbf{c}(\mathbf{0})=\mathbf{c}(\gamma)-\mathbf{c}(\mathbf{0})+\mathbf{H}(\gamma)\gamma+\mathbf{M}(\gamma)\gamma\\ &\therefore\|\gamma-\mathbf{c}(\mathbf{0})\|\leqq C_1\|\gamma\|+H_0\|\gamma\|+H_1\|\gamma\|^2+M_1\|\gamma\|^2\\ &\qquad=N(\gamma)\quad\text{say.}\end{aligned} \tag{2.22}$$

Write $\mathbf{c}(\mathbf{0})$ as

$$\mathbf{c}(\mathbf{0})^T=[c_1(\mathbf{0}),\mathbf{c}_{n-1}(\mathbf{0})^T] \tag{2.23}$$

and so, using (2.3)–(2.5), (2.11) and (2.22),

$$|\lambda-\nu|=|\mu|\leqq|c_1(\mathbf{0})|+|\mu-c_1(\mathbf{0})|\leqq|c_1(\mathbf{0})|+\|\gamma-\mathbf{c}(\mathbf{0})\|=|c_1(\mathbf{0})|+N(\gamma). \tag{2.24}$$

Similarly

$$\|\mathbf{x}-\mathbf{u}\|=\|\boldsymbol{\eta}_{n-1}\|\leqq\|\mathbf{c}_{n-1}(\mathbf{0})\|+N(\gamma). \tag{2.25}$$

The numerical performance of these bounds is given in Table 2 below.

Equation (2.22) does of course provide "dominant error terms" directly. Since $N(\gamma)=O(\|\mathbf{B}_n(\nu)\mathbf{u}\|^2)$ we obtain immediately

$$\lambda-(\nu+c_1(\mathbf{0}))=O(\|\mathbf{B}_n(\nu)\mathbf{u}\|^2) \tag{2.26}$$

and

$$\mathbf{x}^T-(\mathbf{u}^T+[0,\mathbf{c}_{n-1}(\mathbf{0})^T])=O(\|\mathbf{B}_n(\nu)\mathbf{u}\|^2). \tag{2.27}$$

Thus $c_1(\mathbf{0})$ and $\mathbf{c}_{n-1}(\mathbf{0})$ provide us with correction terms to improve our estimates ν and $\mathbf{u}$. (See Table 3 below.)

Equations (2.26) and (33) of [3] give different expressions for the correction of ν but it is straightforward to show their equivalence.

3. The Trapezoidal Rule Method

In this section we look at the numerical performance of the bounds (2.24) (2.25) and the expressions for the dominant error terms given by (2.26) and (2.27) when we use the trapezoidal rule method. The details of the techniques used are given in [3]. Briefly, if $k(s,t)\in C^2[a,b]\times[a,b]$ then

$$\|\mathbf{B}_n(\nu)\mathbf{u}\|_\infty=O(h^2) \tag{3.1}$$

providing $|\lambda|>\|Q_n\|_\infty$. Also, providing (2.17) and (2.18) are satisfied, C_0, C_1, H_0, H_1 and g are $O(h^2)$ (see Table 1 below). Hence (2.19) and (2.20) imply

$$|\lambda-\nu^{(n)}|=O(h^2),\qquad \|\mathbf{x}-\mathbf{u}\|_\infty=O(h^2) \tag{3.2}$$

(see Table 2). Similarly, (2.26) and (2.27) imply

$$\begin{aligned}&|\lambda-(\nu^{(n)}+c_1(\mathbf{0}))|=O(h^4)\\ &\|\mathbf{x}^T-\mathbf{u}^T-[0,\mathbf{c}_{n-1}(\mathbf{0})^T]\|_\infty=O(h^4)\end{aligned} \tag{3.3}$$

(see Table 3).

Example. The integral equation (1.1) with $k(s,t)$ given by

$$k(s,t)=\frac{s}{10}+\frac{7}{60}(8s^3-s)(5t^3-2)$$

with $a=0$, $b=1$ has solutions

$$\lambda_1=0.2\qquad x_1(s)=s^3.$$

The trapezoidal rule was used to provide numerical solutions. Table 1 shows the behaviour of the quantities C_0, C_1, H_0, H_1, M_1, and g for the example.

Table 1

n	C_0	C_1	H_0	H_1	M_1	g
5	0.106E-1	0.126E-1	0.118	0.148E 2	5.29	0.567E 1
9	0.281E-1	0.435E-2	0.331E-1	0.405E 1	8.29	0.750
17	0.722E-2	0.121E-2	0.875E-2	0.109E 1	9.51	0.156
33	0.182E-2	0.310E-3	0.224E-2	0.280	9.88	0.371E-1
65	0.455E-3	0.781E-4	0.565E-3	0.709E-1	9.97	0.916E-2

The quantities C_0, C_1, H_0, H_1 and g all show $O(h^2)$ convergence and M_1 $(=\|\mathbf{T}_n^{-1}\|_\infty)$ tends to a constant value. This is in agreement with the theoretical results (2.12)–(2.16) with (3.1).

The performance of the bounds (2.24) and (2.25) is shown in Table 2.

Table 2

n	$\lvert\lambda-\nu^{(n)}\rvert$	bound (2.24)	$\|\mathbf{x}-\mathbf{u}\|_\infty$	bound (2.25)
5	0.109	$g>\frac{1}{2}$	0.183 E–1	$g>\frac{1}{2}$
9	0.285 E–1	$g>\frac{1}{2}$	0.690 E–2	$g>\frac{1}{2}$
17	0.725 E–2	0.797 E–2	0.199 E–2	0.266 E–2
33	0.1820 E–2	0.1858 E–2	0.515 E–3	0.549 E–3
65	0.4556 E–3	0.4578 E–3	0.1301 E–3	0.1321 E–3

Note the $O(h^2)$ convergence of $|\lambda-\nu^{(n)}|$ and $\|\mathbf{x}-\mathbf{u}\|$ in agreement with (3.2) and the sharpness of the bounds with increasing n. These results are slightly better than those in Table 6 of [3]. (Note that the example in this paper is different from that in [3] by a factor of 10).

To calculate the dominant error term we approximate $\mathbf{B}_n(\nu)\mathbf{u}$ by

$$\mathbf{G}_n(\nu)\mathbf{u}=\left[r_n K p_n-\mathbf{K}_n+\frac{1}{\nu}r_n K Q_n p_n\right]\mathbf{u}. \tag{3.4}$$

This approximation is discussed in §5 of [3].

An alternative method of approximating the dominant error term is provided by the following approach. Define $\boldsymbol{\delta}$ by

$$\boldsymbol{\delta}=r_n K q_n\mathbf{u}-\nu\mathbf{u} \tag{3.5}$$

where $q_n\mathbf{u}=\frac{1}{\nu}\sum_{j=1}^{n} w_j k(s,t_j)u_j$, the Nyström extension. It is not difficult to show that, for the trapezoidal rule, $\|\boldsymbol{\delta}-\mathbf{B}_n(\nu)\mathbf{u}\|=O(h^4)$ provided $k(s,t)$ is smooth enough and so, recalling (2.11),

$$\|\mathbf{c}(\mathbf{0})-\mathbf{T}_n^{-1}\boldsymbol{\delta}\|=O(h^4). \tag{3.6}$$

Table 3

n	$\lvert\lambda-(\nu^{(n)}+c_1(\mathbf{0}))\rvert$	Ratio	$\|\mathbf{x}^T-\mathbf{u}^T-[0,\mathbf{c}_{n-1}(\mathbf{0})^T]\|_\infty$	Ratio	$\lvert\lambda-(\nu^{(n)}+(\mathbf{T}_n^{-1}\boldsymbol{\delta})_1)\rvert$	Ratio
5	0.294 E–2		0.669 E–2		0.201 E–1	
		7.4		6.7		19.1
9	0.399 E–3		0.985 E–3		0.106 E–2	
		12.3		11.9		19.2
17	0.325 E–4		0.830 E–4		0.553 E–4	
		14.8		14.8		19.3
33	0.219 E–5		0.562 E–5		0.287 E–5	
		15.6		15.7		18.4
65	0.140 E–6		0.359 E–6		0.156 E–6	

Such an approach is close to the method of "deferred correction" of Fox and Goodwin [1]. From Table 3 we see that the estimates given by (3.3) are superior to those given by (3.6) for small n but there is little to choose between them for large n. However, the second estimate is cheaper to calculate (see Table 4 below). The columns headed "Ratio" give the ratio of successive values of the entries in the previous column.

Clearly all errors behave as $O(h^4)$. The results for the dominant error term for the eigenvalue, $c_1(\mathbf{0})$, are identical with those given in Table 3 of [3]. The results for the corrections to the eigenvectors are better than those given in [3] but the rate of convergence is almost identical.

4. Practical Considerations

The time taken to calculate the bounds and estimates is, of course, important. As a comparison we quote the time taken to solve the eigenvalue problem (1.2) using a standard method. We make the following definitions.

T_1 = time to calculate all the eigenvalues and eigenvectors of $\mathbf{K}_n$ using reduction to upper Hessenberg form and then the QR algorithm (Wilkinson and Reinsch [4], pp. 339–358, pp. 372–395).

T_2 = time to calculate the matrices $r_n K p_n - \mathbf{K}_n$ and $r_n K Q_n p_n$ and their norms.

T_3 = time taken to calculate the bounds (2.24) (2.25) for one eigenvalue and eigenvector.

T_4 = time to calculate $\mathbf{G}_n(\nu)\mathbf{u}$ from (3.4).

T_5 = time to calculate $\boldsymbol{\delta}$ using (3.6).

T_6 = time to solve $\mathbf{T}_n^{-1}\mathbf{G}_n(\nu)\,\mathbf{u}$ or $\mathbf{T}_n^{-1}\boldsymbol{\delta}$.

Table 4

n	T_1	T_2	T_3	T_4	T_5	T_6
5	0.06	0.03	0.04	0.01	0.004	0.01
9	0.43	0.12	0.12	0.03	0.007	0.03
17	2.70	0.69	0.53	0.10	0.014	0.70
33	15.19	4.34	3.36	0.38	0.026	1.04
65	104.45	31.21	23.84	1.49	0.051	7.32

(The times are given in seconds on the ICL System 4-70 at University College, Cardiff, Wales.)

Note that if bounds were required for p eigenvalues and eigenvectors then the total time would be $T_2 + pT_3$. A major disadvantage is the need to calculate $\mathbf{T}_n^{-1}$ for the quantities H_0, H_1 and M_1 (see (2.14)–(2.16)). This requires $O(n^3)$ multiplications and takes over 90% of the time T_3 for large n. Even so the times given above compare favourably with those in Table 7 of [3].

Often we shall be content with estimates of the dominant error terms instead of bounds (2.24) (2.25). We noted the numerical performance of two such estimates in Table 3 and we give the times to calculate them in Table 4.

Acknowledgements. I would like to express my sincere appreciation to Professor Ben Noble for his guidance and encouragement throughout my research at Oxford. During that time I was supported by the Science Research Council of Great Britain. I would also like to thank Dr. Ken Thomas, for the many interesting discussions we have had.

References

1. Fox, L., Goodwin, E.T.: The numerical solution of non-singular linear integral equations. Phil. Trans. Roy. Soc. London Ser. A. **245**, 501–534 (1953)
2. Spence, A.: On the convergence of the Nyström method for the integral equation eigenvalue problem. Numer. Math. **25**, 57–66 (1975)
3. Spence, A.: Error bounds and estimates for eigenvalues of integral equations. Numer. Math. **29**, 133–147 (1978)
4. Wilkinson, J.H., Reinsch, C.: Handbook for automatic computation, Vol. II. In: Linear algebra. Berlin-Heidelberg-New York: Springer 1971

Received October 24, 1977