

Asymptotic correction and inverse eigenvalue problems: an overview

Alan L. Andrew*

(Received 20 October 2004, revised 27 January 2005)

Abstract

Asymptotic correction was first used by Paine, de Hoog and Anderssen to improve the accuracy of finite difference approximations of higher Sturm–Liouville eigenvalues. Later it was used to develop an important class of methods for numerical solution of inverse Sturm–Liouville problems. It also shows promise as a method for the solution of more general inverse eigenvalue problems, including some involving partial differential equations and higher order operators. We critically review the literature on this subject and discuss some important open questions.

* Mathematics Department, La Trobe University, Victoria 3086, AUSTRALIA.
<mailto:a.andrew@latrobe.edu.au>

See <http://anziamj.austms.org.au/V46/CTAC2004/Andr> for this article, © Austral. Mathematical Soc. 2005. Published 7 March 2005, amended March 22, 2005. ISSN 1446-8735

Contents

1	Introduction	C2
2	Symmetric Sturm–Liouville problems	C4
3	Convergence questions	C7
4	Other problems	C9
	References	C11

1 Introduction

Most work on numerical methods involving eigenvalue problems for differential operators has concerned the *direct* problem of computing eigenvalues and eigenfunctions of a known operator. This paper concerns the *inverse* problem of using information on eigenvalues and eigenfunctions to obtain information about the differential operator. Such problems arise, for example, in the search for cracks in aircraft and other structures, and in the use of seismic data to infer properties of the earth’s crust [17].

Most work on numerical solution of inverse eigenvalue problems for differential operators has concerned Sturm–Liouville problems, especially the problem of using the eigenvalues of

$$-y'' + qy = \lambda y, \quad (1)$$

$$y(0) = y(\pi) = 0, \quad (2)$$

to find the (unique) $q \in L_2(0, \pi)$ satisfying (1), (2) and

$$q(x) = q(\pi - x), \quad \text{for almost all } x \in [0, \pi]. \quad (3)$$

Some of the best numerical methods for inverse Sturm–Liouville problems, such as those of [27], depend on specific properties of these problems, and it is not obvious how they may be adapted to other differential operators. Although Sturm–Liouville problems, especially the problem (1), (2), (3), are also emphasized here, the methods considered are applicable to a wide class of inverse eigenvalue problems, including many involving partial differential equations and higher order operators.

The key to the methods considered here is the technique of asymptotic correction [4] (sometimes called “algebraic correction” [14] or the “AAAdHP correction” [26]), which was introduced for the direct problem (1), (2) in [22]. This enables the successful combination of finite difference and finite element methods for the direct problem with known methods [11] for matrix inverse eigenvalue problems, to produce a powerful battery of methods for inverse eigenvalue problems for differential equations. The importance of asymptotic correction lies in the fact [25] that, for all $q \in L_2(0, \pi)$, the i th eigenvalue of (1), (2) is $i^2 + \int_0^\pi q(x)dx/\pi + \alpha_i(q)$, where $\{\alpha_i(q)\}_{i=1}^\infty \in \ell^2$. The first term, i^2 , gives no information on q , and the second gives only the mean value. All information on the variation of q is contained in the small term $\alpha_i(q)$, which, without asymptotic correction, would be swamped by discretization errors in all except the smallest eigenvalues [7]. Asymptotic correction overcomes this difficulty by adding a suitable correction to the eigenvalues before solving the discrete inverse eigenvalue problem.

The aim of this paper is to encourage further work on the use of asymptotic correction for the numerical solution of inverse eigenvalue problems, by giving a critical review of the existing literature on the subject and drawing attention to several open questions on which I believe significant progress should not be too hard to achieve. Section 2 discusses the problem of computing q satisfying (3) from eigenvalues of (1) with various boundary conditions. Section 3 discusses convergence properties of methods for this problem and Section 4 discusses the application of asymptotic correction to other inverse eigenvalue problems.

2 Symmetric Sturm–Liouville problems

The classical second order finite difference method for the direct problem approximates the first $2n$ eigenvalues, $\lambda_1(q) < \dots < \lambda_{2n}(q)$, of (1), (2), (3) by the eigenvalues, $\mu_1(q) < \dots < \mu_{2n}(q)$, of the $2n \times 2n$ symmetric tridiagonal matrix $h^{-2}A(n) + Q((\mathbf{q})_h)$. Here and *throughout this paper* we define the $2n \times 2n$ matrix $A(n) := \text{tridiag}(-1, 2, -1)$, $(\mathbf{q})_h := (q(h), q(2h), \dots, q(nh))^T$, where $h := \pi/(2n + 1)$, and, for all $\mathbf{p} := (p_1, \dots, p_n)^T \in \mathbb{R}^n$, $Q(\mathbf{p}) := \text{diag}(p_1, p_2, \dots, p_n, p_n, \dots, p_1)$. Although possible application to the numerical solution of inverse Sturm–Liouville problems is often given as motivation for work on matrix inverse eigenvalue problems, most authors ignore the problem of how to deal with the different asymptotic behaviour of the eigenvalues of the discrete and continuous problems. Instead, they appear to endorse methods which do not even give solutions of the correct order of magnitude. This is discussed more fully in [1, 6, 7, 14, 19, 20, 21].

The first successful treatment of this problem using asymptotic correction is due to Paine [21], who considered the problem (1), (2), (3). Asymptotic correction uses the fact that, at least for sufficiently smooth q , the leading term in the asymptotic expansion (as $i \rightarrow \infty$) of the error in the i th computed eigenvalue is independent of q . Moreover, this error is often known in closed form in the case when q is constant. In particular, for the second order finite difference method, $\lambda_i(0) - \mu_i(0) = \varepsilon_1(i, h)$ where, as in [3], we use the notation

$$\varepsilon_r(i, h) := i^2 - \frac{12 \sin^2(ih/2)}{h^2[3 + (1 - r) \sin^2(ih/2)]}. \quad (4)$$

Hence, Paine sought a $2n \times 2n$ matrix $h^{-2}A(n) + Q(\mathbf{p})$ whose i th eigenvalue is $\lambda_i(q) - \varepsilon_1(i, h)$, where $\lambda_i(q)$ is the given i th eigenvalue of (1), (2), (3).

There has been some work [18] on methods for computing a diagonal matrix D such that $A(n) + D$ has specified eigenvalues, but there has been more on recovering tridiagonal matrices from eigenvalues. Perhaps for this reason, Paine’s method computes a symmetric centrosymmetric [2] tridiago-

nal matrix with *all* its eigenvalues specified and then proceeds to refine the diagonal elements using the errors in the computed off-diagonal elements. A simpler method for the latter task is given in [13]. A different approach was used by Pirovino [23, 24], who used the same finite difference scheme as in [21, 13]. Following Marti [19] (who used the finite element method of [9]), he used the fact that the non-diagonal elements are already known. Given $\lambda_1(q), \dots, \lambda_n(q)$, he computed an n -vector $\mathbf{q}(n)$ such that the first n eigenvalues of $h^{-2}A(n) + Q(\mathbf{q}(n))$ are $\lambda_1(q) - \varepsilon_1(1, h), \dots, \lambda_n(q) - \varepsilon_1(n, h)$. Then the i th component q_i of $\mathbf{q}(n)$ is the accepted value of $q(ih)$. The same approach was used in [15], the main difference being that, whereas Pirovino, following Marti, used Newton’s method to compute $\mathbf{q}(n)$, [15] used the modified Newton’s method. An advantage of the approach of [15, 23], which uses only the first half of the matrix eigenvalues, is that the error in the corrected eigenvalue estimate is never bigger than $\mathcal{O}(h)$ whereas the error in the largest eigenvalue used in the method of [21, 13] is $\mathcal{O}(1)$. Numerical results of [14] show that, for a given *mesh length*, neither the method of [21, 13] nor the method of [23] is consistently more accurate than the other. However, in applications the limiting factor is normally the number of *eigenvalues* known with sufficient accuracy. For a given mesh length, the method of [21, 13] requires twice as many eigenvalues as the methods of [19, 23, 24, 15], and the important question is usually, “Which method extracts most information from a given set of data?” A theoretical analysis comparing the errors of the two approaches would be of interest.

All the above methods used a second order method for solving the direct problem. An approach similar to that of [15], but using Numerov’s method, was described in [6]. In [7] the methods of [15] and [6] were both extended to deal with the boundary conditions

$$y'(0) = y'(\pi) = 0, \quad (5)$$

and the performance of the methods was compared numerically. Numerov’s method [3] for the direct problem with (5) uses some values of q outside the interval $[0, \pi]$ and, without some modification, this would introduce extra

unwanted unknowns into the inverse problem. Two strategies for dealing with this are suggested in [7]. The one which appears to give the best results is based on hypotheses suggested in [15] concerning the relationship between $\mathbf{q}(n)$ and truncated Fourier series for q . However, the question of precisely when that strategy will give better results than the other, or whether yet other methods will perform even better, awaits theoretical analysis.

The numerical results of [7] suggest the following conclusions, among others, though proofs are still lacking.

1. For discontinuous q , Numerov’s method and the method of [15] give very similar results, but for $q \in C^1$, and especially for $q \in C^4$, the results are strikingly different, as described below. (This result, if proved, could be used to determine the smoothness of q .)
2. In the case (5) with smooth q , Numerov’s results are significantly more accurate than those of the method of [15], though the regularity of convergence shows no consistent difference between the two methods.
3. For (2) and smooth q , the pointwise convergence of Numerov’s method is much more regular than that of the method of [15].

This regular convergence allows the Numerov results (but not the results of the method of [15]) to be significantly improved by extrapolation. In this case, results of the Numerov method of [6] are never significantly less accurate, and are sometimes significantly more accurate than those of the method of [15]. This is demonstrated in Table 1, which supplements results of [7] by comparing the errors $\mathbf{q}(n) - (\mathbf{q})_h$ obtained by the Numerov method of [6] with those of the method (FKL) of [15] for (1), (2), when $q(x) = \cos(2x)$. For other examples we found less difference in the accuracy of the two methods for (2) and more difference in the regularity of convergence. With (5) the difference in accuracy was greater [7]. The reason why the advantages of Numerov’s method are not the same for (5) as for (2) is another intriguing open question.

TABLE 1: Error in estimate of $q(i\pi/11)$ obtained from the first n eigenvalues

i	Numerov [6]			FKL [15]		
	$n = 5$	$n = 16$	$n = 49$	$n = 5$	$n = 16$	$n = 49$
1	2.8E-3	1.4E-4	5.6E-6	5.2E-3	2.8E-4	1.1E-5
2	-7.2E-4	-3.4E-5	-1.3E-6	1.3E-3	2.5E-4	3.2E-5
3	4.6E-4	1.5E-5	5.4E-7	4.8E-3	4.8E-4	5.1E-5
4	-1.0E-4	-5.4E-6	-2.2E-7	2.1E-3	2.5E-4	2.9E-5
5	4.9E-5	1.8E-6	6.7E-8	4.1E-4	4.0E-5	4.2E-6

3 Convergence questions

It was proved in [23, 15, 7] that, for all q “sufficiently close” in some norm to a constant, the various iterative methods used to solve matrix inverse eigenvalue problems in these papers always converge to $\mathbf{q}(n)$, even when $q \notin C$. The vital question of how close $\mathbf{q}(n)$ is to $(\mathbf{q})_h$ is more difficult. As a first step, if we define the norm $\|\mathbf{p}\|_h = (h \sum_{i=1}^n p_i^2)^{1/2}$, we would like to establish sufficient conditions to ensure that

$$\|\mathbf{q}(n) - (\mathbf{q})_h\|_h \rightarrow 0 \quad \text{as } n \rightarrow \infty. \quad (6)$$

Dun [13] proved a preliminary result along these lines using [16, Theorem 5], but his result requires stronger hypotheses than have yet been proved to hold for any non-constant q .

Some convergence results are given in [23] (and summarized, with most proofs omitted, in [24]), but proofs depend on various quantities being “sufficiently small”. The key to this approach is Theorem 1 of [24], a discrete analogue of Theorem 3 of [17]. It gives a bound for the norm of the difference between two *centrosymmetric* diagonal perturbations of $A(n)$ in terms of the differences (assumed to be sufficiently small) between the corresponding

eigenvalues of the two perturbations. A corresponding result for centrosymmetric diagonal perturbations of the (real symmetric centrosymmetric) matrix $(2I - A(n))^{-1}A(n)$ would enable extension of Pirovino's analysis to Numerov's method. Pirovino used [24, Theorem 1] and the fact that

$$\mu_i(q) + \varepsilon_1(i, h) - \lambda_i(q) = \mathcal{O}(ih^2) \quad \text{for } i = 1, \dots, n, \quad (7)$$

to deduce that $\|\mathbf{q}(n) - (\mathbf{q})_h\|_2 = \mathcal{O}(h^{1/2})$ as $h \rightarrow 0$, provided $\sum_{i=1}^{\infty} (\lambda_i(q) - \lambda_i(0))^2$ is sufficiently small. Since results of [22] hold only for $ih < \alpha$, where α is some unknown number, they are not strong enough to prove (7), which needs $\alpha > nh$. However, it was shown in [8] that, for all $q \in C^2$, there is a number $c(q)$, independent of i and h , such that $|\mu_i(q) + \varepsilon_1(i, h) - \lambda_i(q)| < |c(q)i^2h^3/\sin(ih)|$. Since $ih/\sin(ih) < \pi/2$ for $i \leq n$, this implies (7). A different approach is needed for the method of [13, 21], as it uses all $2n$ eigenvalues rather than the first n , and $\mu_{2n}(q) + \varepsilon_1(2n, h) - \lambda_{2n}(q) = \mathcal{O}(1)$. Note that even an extension of Pirovino's results to Numerov's method would not explain the excellent convergence results observed in our numerical tests. Our results [7] suggest that, if $q \in C^4$ is monotonic in $(0, \alpha)$, then, for each $x \in (0, \alpha]$, there exists a number $\beta(q, x)$ such that the error in the Numerov solution to the inverse problem (1), (2), (3) is less than $\beta(q, x)h^3$. Even for step functions we obtained $\mathcal{O}(h)$ convergence for all $x \in (0, \pi)$. The convergence analysis of [13, 23, 24] considered only the norm of the error. Results of [7] suggest that this approach will always underestimate the rate of convergence at most points, since the first grid point is h and the coefficient $\beta(q, h) \rightarrow \infty$ as $h \rightarrow 0$.

As discussed in [7], a major source of error in solving inverse eigenvalue problems for differential operators is the fact that we are trying to use a finite set of data to calculate a quantity that requires an infinite set of data for its specification. This has nothing to do with the method used (and for this reason some of the analysis given for other methods such as [27] may be applicable to the methods considered here). At best we can hope to find an approximation to q which will be good for sufficiently "nice" q . For smooth q , numerical results do show a significant difference between the performance of

different methods, but more theoretical analysis is needed. Most theoretical work on inverse eigenvalue problems so far has concerned only the case when an infinite amount of data is available, although in applications the amount of available data is often quite small. There has also been some interesting work on artificially increasing the amount of available data by making certain assumptions [10], but more work is needed to determine the effectiveness of such methods.

4 Other problems

When q is not required to satisfy (3) the eigenvalues of (1), (2) do not determine q uniquely, but q can be determined uniquely if we have suitable additional data [1, 27]. One such set of data is the eigenvalues from a *suitable* second set of boundary conditions. I plan to consider the use of asymptotic correction in this case elsewhere. If the first boundary conditions are (2), then (5) will not be a suitable second set, as the eigenvalues of (1), (2) and those of (1), (5) are both unchanged if q is replaced by \hat{q} , where $\hat{q}(x) = q(\pi - x)$ for all x . However, for any $c \in \mathbb{R}$, $y(0) = y'(\pi) - cy(\pi) = 0$ is a suitable supplement to (2). Other sets of supplementary data, which are known to determine non-symmetric q uniquely, involve some information about the eigenvectors. Asymptotic correction has been used successfully in one such case [7, 15], although questions remain about the best way to implement Numerov's method for this problem [7].

The problems considered so far have the advantage that the asymptotic correction is known in closed form. Even this class of problems is much wider than Sturm–Liouville problems. Closed form solutions for the errors in finite difference eigenvalues are also known for some higher order equations and some partial differential equations, and asymptotic correction has been found effective for the direct problem in these cases (see [5] for references and further information). The use of asymptotic correction for the corresponding

inverse problems is an obvious field for further investigation. Apart from some preliminary work on certain partial differential operators in [13] (which is discussed in [4]), this question seems to be unexplored.

Asymptotic correction can also be used for a wider class of problems. I believe it will be useful for all problems of the form

$$A_1 u + A_2 u = \lambda u, \quad (8)$$

where A_1 is an unbounded self-adjoint linear operator with compact inverse on some infinite dimensional Hilbert space and A_2 is a bounded symmetric linear operator on the same space, provided only that the exact eigenvalues of A_1 are known. It is *not* necessary that eigenvalues of the discretization of A_1 should also be known in closed form. Typically A_1 is a differential operator and A_2 may, for example, represent multiplication by another function in the same Hilbert space, as is the case with (1). To solve (8) using asymptotic correction when eigenvalues of the discretization of A_1 are not known in closed form, we must obtain finite difference or finite element approximations to the solutions of both (8) and $A_1 u = \lambda u$. For the method to be competitive for the direct problem, the cost of doubling the number of finite difference problems to be solved must be less than the cost of the mesh refinement required to produce the same increase in accuracy without asymptotic correction. Numerical results in [12] show that this is the case even for certain methods for solving Sturm–Liouville problems.

The potential value of asymptotic correction in cases where no closed form expression is available for the correction, is much greater for the inverse problem of estimating A_2 from spectral data than for the corresponding direct problem for two reasons:

1. for the inverse problem, indefinite mesh refinement is not an option, as we are limited by the number of known eigenvalues;
2. without asymptotic correction, finite difference and finite element methods still give good approximations for the lower eigenvalues of the direct

problem, but such methods fail completely for the inverse problem.

The reasons for this are discussed more fully for Sturm–Liouville problems in [7], but similar considerations hold for the more general problem (8). So far there has been almost no work done on the use of asymptotic correction for such problems. However, I suspect that they may eventually prove to be the most important use of asymptotic correction, since finite difference and finite element methods have been developed for such a wide variety of differential equations and no methods at all are yet available for many of the corresponding inverse problems.

References

- [1] A. L. Andrew. Some recent developments in inverse eigenvalue problems. In D. Stewart, H. Gardner and D. Singleton, editors, *Computational Techniques and Applications: CTAC93*, pages 94–102. World Scientific, 1994. C4, C9
- [2] A. L. Andrew. Centrosymmetric matrices. *SIAM Rev.*, 40:697–698, 1998. C4
- [3] A. L. Andrew. Asymptotic correction of Numerov’s eigenvalue estimates with natural boundary conditions. *J. Comput. Appl. Math.*, 125:359–366, 2000. C4, C5
- [4] A. L. Andrew. Twenty years of asymptotic correction for eigenvalue computation. *ANZIAM J.*, 42(E):C96–C116, 2000. [Online] <http://anziamj.austms.org.au/V42/CTAC99/Andr.> C3, C10
- [5] A. L. Andrew. Asymptotic correction of more Sturm–Liouville eigenvalue estimates. *BIT*, 43:485–503, 2003. C9

- [6] A. L. Andrew. Numerical solution of inverse Sturm–Liouville problems. *ANZIAM J.*, 45(E):C326–C337, 2004. [Online] <http://anziamj.austms.org.au/V45/CTAC2003/Andr.> C4, C5, C6, C7
- [7] A. L. Andrew. Numerov’s method for inverse Sturm–Liouville problems. *Inverse Problems*, 21:223–238, 2005. [Online] <http://stacks.iop.org/0266-5611/21/223>. C3, C4, C5, C6, C7, C8, C9, C11
- [8] A. L. Andrew and J. W. Paine. Correction of Numerov’s eigenvalue estimates. *Numer. Math.*, 47:289–300, 1985. C8
- [9] A. L. Andrew and J. W. Paine. Correction of finite element estimates for Sturm–Liouville eigenvalues. *Numer. Math.*, 50:205–215, 1986. C5
- [10] Z.-J. Bai, R. H. Chan and B. Morini. An inexact Cayley transform method for inverse eigenvalue problems. *Inverse Problems*, 20:1675–1689, 2004. C9
- [11] M. T. Chu. Inverse eigenvalue problems. *SIAM Rev.*, 40:1–39, 1998. C3
- [12] D. C. Condon. *Asymptotic correction of Sturm–Liouville eigenvalue estimates*. PhD thesis, La Trobe University, Melbourne, 2002. C10
- [13] C. R. Dun. *Algebraic correction methods for two-dimensional eigenvalue problems*. PhD thesis, Australian National University, Canberra, 1995. C5, C7, C8, C10
- [14] C. R. Dun and R. S. Anderssen. Algebraic correction methods for inverse Sturm–Liouville problems. In D. Stewart, H. Gardner and D. Singleton, editors, *Computational Techniques and Applications: CTAC93*, pages 202–210. World Scientific, 1994. C3, C4, C5

- [15] R. H. Fabiano, R. Knobel and B. D. Lowe. A finite difference algorithm for an inverse Sturm-Liouville problem. *IMA J. Numer. Anal.*, 15:75–88, 1995. C5, C6, C7, C9
- [16] O. H. Hald. Inverse eigenvalue problems for Jacobi matrices. *Linear Algebra Appl.*, 14:63–85, 1976. C7
- [17] O. H. Hald. The inverse Sturm-Liouville problem and the Rayleigh-Ritz method. *Math. Comput.*, 32:687–705, 1978. C2, C7
- [18] M. Hegland and J. T. Marti. Algorithms for the reconstruction of special Jacobi matrices from their eigenvalues. *SIAM J. Matrix Anal. Appl.*, 10:219–228, 1989. C4
- [19] J. T. Marti. Small potential corrections for the discrete eigenvalues of the Sturm-Liouville problem. *Numer. Math.*, 57:51–62, 1990. C4, C5
- [20] F. Natterer. A Lanczos type algorithm for inverse Sturm-Liouville problems. *Proc. Centre Math. Appl. Austral. Nat. Univ.*, 57:82–88, 1992. C4
- [21] J. Paine. A numerical method for the inverse Sturm-Liouville problem. *SIAM J. Sci. Stat. Comput.*, 5:149–156, 1984. C4, C5, C8
- [22] J. W. Paine, F. R. de Hoog and R. S. Anderssen. On the correction of finite difference eigenvalue approximations. *Computing*, 26:123–139, 1981. C3, C8
- [23] M. H. Pirovino. *Das Sturm-Liouville-Problem als direktes und inverses Eigenwertproblem und seine numerische Behandlung durch finite Differenzen*. Dissertation 9683, ETH Zürich, 1992. C5, C7, C8
- [24] M. H. Pirovino. *The inverse Sturm-Liouville problem and finite differences* Res. Rep. 93-04, Sem. Angew. Math. ETH Zürich, 1993. C5, C7, C8

- [25] J. Pöschel and E. Trubowitz. Inverse spectral theory. Academic Press, 1987. [C3](#)
- [26] J. D. Pryce. *Numerical solution of Sturm-Liouville problems*. Oxford University Press, 1993. [C3](#)
- [27] W. Rundell and P. E. Sacks. Reconstruction techniques for classical Sturm–Liouville problems. *Math. Comp.*, 58:161–183, 1992. [C3](#), [C8](#), [C9](#)