Numerical solution of inverse Sturm–Liouville problems

Alan L. Andrew*

(Received 8 August 2003, revised 4 February 2004)

Abstract

A new algorithm is proposed for solving the inverse Sturm–Liouville problem of reconstructing a symmetric potential from eigenvalues. It uses Numerov's method instead of the second order method of the related algorithm of Fabiano, Knobel and Lowe. An extension by Andrew and Paine of the asymptotic correction technique of Paine, de Hoog and Anderssen is the key to the success of the new algorithm. Numerical results show that it can produce good accuracy even from very limited data.

Contents

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

See http://anziamj.austms.org.au/V45/CTAC2003/Andr for this article, Published June 1 2004, amended June 7, 2004. ISSN 1446-8735

1 Introduction

Numerical methods for computing eigenvalues, λ, and eigenfunctions, y, of the Sturm–Liouville problem

$$-y'' + qy = \lambda y\,, \tag{1}$$
$$y(0) = y(\pi) = 0 \tag{2}$$

(the "direct" problem) have reached a high level of sophistication [31]. (Herein y and q denote functions of position, whereas $y(x)$ and $q(x)$ denote their values at a position x.) This paper is concerned with the more difficult *inverse* Sturm–Liouville problem of computing q from a knowledge of the eigenvalues. It is well known that the eigenvalues of (1–2) do not define q uniquely, since, if y and q are replaced by $\widehat{y}$ and $\widehat{q}$, where $\widehat{y}(x) = y(\pi - x)$ and $\widehat{q}(x) = q(\pi - x)\,$, then the eigenvalues are unchanged. Some additional conditions ensuring uniqueness, and some conditions ensuring existence, are discussed in [4, 26] and the references given there. We consider a case first studied in [15]—that in which q is required to be continuous and also symmetric:

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

Inverse eigenvalue problems have been more intensively studied in the finite dimensional case [16, 17], and the idea of using finite difference methods to replace the inverse Sturm–Liouville problem by a matrix inverse eigenvalue problem is often given as motivation for work on the matrix inverse eigenvalue problem [27]. It is now known [28] that the naïve approach suggested in some early papers gives answers that are not even the correct order of magnitude, as the finite difference eigenvalues do not have the same asymptotic

behaviour as those of the continuous problem. The difficulty is obscured in the numerical example of [27] which uses as input, not the eigenvalues of the differential operator, but those of its finite difference approximation, which would not be known in real applications. That example tests the effectiveness of the methods of [27] for solving the *matrix* inverse eigenvalue problem, but provides no evidence of the usefulness of the methods of [27] for the more difficult inverse eigenvalue problem for differential equations.

Two methods of overcoming this difficulty have been found. The first [23] uses a completely different (and non-sparse) matrix eigenvalue problem. The second, first developed by Paine [28], combines the classical finite difference equations with a simple "asymptotic correction" technique [5, 10], introduced in [29] for the direct problem. The technique takes advantage of the following property of the asymptotic expansion (as $i \to \infty$) of the error in the approximation to the ith eigenvalue obtained by many finite difference and finite element methods. At least for sufficiently smooth q, *the leading term in this expansion is independent of* q. Moreover, in the case of constant q, this error is often known in closed form. The idea is that, before the matrix inverse eigenvalue problem is solved, a correction is added to the eigenvalues. This correction is calculated so that, in the case of constant q, the corrected eigenvalues would be those of the discrete problem rather than the continuous problem [4].

Paine [28] used a method of [18] to retrieve a tridiagonal matrix from the corrected eigenvalues and, following [22], used a numerical implementation of the Liouville transformation to extract information from the computed off-diagonal elements. A simplification of this approach was given in [19, 20]. Marti [25] used the fact that the off-diagonal elements are already known, and solved for the diagonal elements only, using Newton's method. This approach was further developed in [30]. The work of [25] used a (second order) finite element method [14], but [30] used the (second order) finite difference scheme of [28, 19, 20]. The big advantage of the approach of [25, 30] is that, since the off-diagonal elements are no longer treated as unknowns, (3)

ensures that, when n eigenvalues are available, we may choose the mesh length $\pi/(2n+1)$, so that the matrix dimension is $2n$. As well as refining the mesh that can be used with a given number of eigenvalues, this ensures that only the first half of the matrix eigenvalues are required for the calculation. It is precisely these eigenvalues for which asymptotic correction is known to be most successful [10, 13, 29]. Fabiano, Knobel and Lowe [21] extended the work of [25, 30], using the same second order finite difference scheme as in [28, 19, 20, 30] and using the modified Newton's method [2] instead of Newton's method. They considered both the case (3) and a case in which certain other spectral data are used with more general q.

This paper refines the algorithm of [21] by using Numerov's method [13] instead of the second order method. As in [13, 19, 20, 25, 28, 30], only the boundary conditions (2) are considered, but asymptotic correction has also been successfully used with Numerov's method for other boundary conditions [9, 11, 12]. The use of Numerov's method for inverse problems with more general boundary conditions will be considered elsewhere.

2 New algorithm

Given numbers $\lambda_1 < \cdots < \lambda_n$, we seek a vector $\mathbf{q} = (q_1, \ldots, q_n)^T$ where, for $i = 1, \ldots, n$, q_i is a good approximation of $q(ih)$, where $h = \pi/(2n+1)$ and q is a function satisfying (3) such that the first n eigenvalues of (1–2) are $\lambda_1, \ldots, \lambda_n$. Numerov's method for the direct problem (1–2) approximates $\lambda_1, \ldots, \lambda_n$ by the eigenvalues $\Lambda_1 < \cdots < \Lambda_n$ of

$$-A\mathbf{u} + BQ\mathbf{u} = \Lambda B\mathbf{u}\,, \tag{4}$$

where $A = (a_{ij})$ is a symmetric tridiagonal $2n \times 2n$ matrix with $a_{ii} = -2/h^2$ and $a_{i,i+1} = 1/h^2$, $B = I + h^2A/12$, and Q is the centrosymmetric [8] diagonal $2n \times 2n$ matrix $\mathrm{diag}(q(h), q(2h), \ldots, q(nh), q(nh), \ldots, q(h))$. The results of [13] show, by the argument used in [4], that the appropriate corrected

eigenvalues are

$$\widetilde{\lambda}_i = \lambda_i + \frac{12\sin^2(ih/2)}{h^2[3-\sin^2(ih/2)]} - i^2\,, \quad i = 1,\ldots,n\,. \tag{5}$$

Some methods for solving the inverse eigenvalue problem for the matrix equation (4) are given in [24], but, following [21], we used the modified Newton method. More precisely, we used the recurrence relation

$$\mathbf{q}(k+1) = \mathbf{q}(k) - [\Lambda'(0)]^{-1}\mathbf{f}(k)\,, \tag{6}$$

where $\mathbf{f}(k)$ and $\mathbf{q}(k)$ are n-vectors with ith components $\Lambda_i(k) - \widetilde{\lambda}_i$ and $q_i(k)$ respectively; $q_i(k)$ is the kth approximation of q_i; $\Lambda_i(k)$ is the ith eigenvalue of (4) when Q is replaced by $\mathrm{diag}(q_1(k),\ldots,q_n(k),q_n(k),\ldots,q_1(k))$; Λ' is the $n\times n$ Jacobian matrix whose ijth element, $\Lambda_{i,j}$, is the partial derivative of Λ_i in (4) with respect to the jth diagonal element of Q; and $\Lambda'(0)$ is the value of Λ' when $Q = \mathrm{diag}(q_1(0),\ldots,q_n(0),q_n(0),\ldots,q_1(0))$.

Since A and B are symmetric commuting invertible matrices, it follows that $AB^{-1} = B^{-1}A = (B^{-1}A)^T$ and hence by (4)

$$-\mathbf{u}_i^T B^{-1}A + \mathbf{u}_i^T Q = \Lambda_i u_i^T\,, \quad i = 1,\ldots,2n\,. \tag{7}$$

One of several ways [6, 7] to compute $\Lambda_{i,j}$ is to differentiate (4) with respect to q_j. Since A and B are independent of Q, it follows that, provided all the required partial derivatives exist,

$$-B^{-1}A\mathbf{u}_{i,j} + Q_{,j}\mathbf{u}_i + Q\mathbf{u}_{i,j} = \Lambda_{i,j}\mathbf{u}_i + \Lambda_i\mathbf{u}_{i,j}\,, \tag{8}$$

where the subscript "$,j$" denotes the partial derivative with respect to q_j, so that the only nonzero elements of $Q_{,j}$ are its jth and $(2n+1-j)$th diagonal elements, both of which are one. Premultiplying (8) by $\mathbf{u}_i^T$ and using (7) gives

$$\Lambda_{i,j} = \mathbf{u}_i^T Q_{,j}\mathbf{u}_i / \mathbf{u}_i^T\mathbf{u}_i\,. \tag{9}$$

(This result can also be derived without the requirement that partial derivatives of the eigenvectors exist [6].) In Algorithm 1 below, $\mathbf{q}(0)$ is chosen so that all its components are equal. In this case, the jth component of $\mathbf{u}_i$ is $\sin(ijh)$ and hence, since $\sum_{j=1}^{2n} \sin^2(ijh) = (2n+1)/2$, it follows from (9) and the definition of $Q_{,j}$ that

$$\Lambda_{i,j}(0) = \frac{4}{2n+1} \sin^2(ijh)\,. \tag{10}$$

Algorithm 1:

Input: n, eigenvalues $\lambda_1, \ldots, \lambda_n$ and parameters ε, N.

1. Set $k = 0$ and $\delta = \varepsilon + 1$. For $i = 1, \ldots, n$ set

$$q_i(0) = \frac{1}{n} \sum_{i=1}^{n} \lambda_i - \frac{1}{6}(n+1)(2n+1)\,. \tag{11}$$

2. Compute the corrected eigenvalues $\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_n$ using (5).
3. Compute the n^2 elements, $\Lambda_{i,j}(0)$, of $\Lambda'(0)$ using (10).
4. **Repeat until** $\delta < \varepsilon$ or $k > N$.
   - 4.1. Compute the eigenvalues $\Lambda_i(k)$, $i = 1, \ldots, n$ of (4).
   - 4.2. Compute $\mathbf{f}(k)$ and hence compute $\mathbf{q}(k+1)$ using (6).
   - 4.3. Compute $\delta = \|\mathbf{q}(k+1) - \mathbf{q}(k)\|$.
   - 4.4. $k \leftarrow k+1$.
5. **Output** k, δ, $q_1(k), \ldots, q_n(k)$.

In the special case when q in (1) is constant, (11) ensures that $q_i(k) = q(ih)$ for all k (since $\lambda_i = i^2 + q$ in that case). This choice of $\mathbf{q}(0)$ allows

the condition "q is sufficiently small" imposed in [21] (which takes $\mathbf{q}(0) = \mathbf{0}$) to be weakened to "$q$ is sufficiently close to a constant". The output values of k and δ indicate the number of iterations used in Step 4 and how well $q_i(k)$ approximates $q_i = \lim_{k\to\infty} q_i(k)$. The number of iterations is controlled by the input parameters ε and N. The amount of time required by additional iterations is generally very small, especially when the methods of [1] are used to take advantage of the tridiagonal centrosymmetric structure of the coefficient matrices. However, our goal is to find $q(ih)$, and how well q_i approximates $q(ih)$ depends on how well (4) approximates (1). For all n we tested with $q(x) = \sin(x)$, we found that $|q_i(k) - q_i| < |q(ih) - q_i|$ for all $k > 4$ and all i, so that using very large k produced only slightly better estimates of $q(ih)$ than $k = 5$. As with all numerical algorithms, attainable accuracy is limited by the quantity and quality of the available data, that is by the number of accurately known eigenvalues. It is n, not k, that is critical in determining how well (4) approximates (1).

Convergence of (6), though not a sufficient condition to ensure that $|q_i - q(ih)|$ is small, is nevertheless necessary for the success of Algorithm 1. This is where (5) is essential. When the eigenvalues, $\lambda_1, \ldots, \lambda_n$, of the differential equation were used instead of the corrected eigenvalues $\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_n$ in (6), that is when the ith element of $\mathbf{f}(k)$ was taken as $\Lambda_i(k) - \lambda_i$ instead of $\Lambda_i(k) - \widetilde{\lambda}_i$, the iteration did not even converge. Even in the trivial case of constant q (when Algorithm 1 gives $q_i(k) = q(ih)$ for all k and i), if the uncorrected eigenvalues λ_i are used instead of the corrected eigenvalues $\widetilde{\lambda}_i$, then the iteration of Step 4 diverges. As noted in [25] and [30], the second order methods studied there also have this property of "no convergence without asymptotic correction".

Table 1 shows the error (computed q − true q) in the results given by Algorithm 1 in the case $q(x) = \sin(x)$, with q_i approximated by $q_i(12)$. "Exact" eigenvalues for this table were computed using the formula $C_k(175, 125)$ of [11, equation (28)]. The three values of n chosen (5, 16 and 49) give mesh lengths $\pi/11$, $\pi/33$ and $\pi/99$ respectively, so that, if convergence is $\mathcal{O}(h^p)$

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

i	$n = 5$	$n = 16$	$n = 49$	ratio1	ratio2
1	–6.25E–3	–3.59E–4	–1.46E–5	17	25
2	2.15E–3	9.32E–5	3.53E–6	23	26
3	–9.60E–4	–3.79E–5	–1.41E–6	25	27
4	4.51E–4	1.69E–5	6.20E–7	27	27
5	–1.32E–4	–4.89E–6	–1.80E–7	27	27

as $h \to 0$ (that is as $n \to \infty$), the ratios of the errors in consecutive columns should be approximately 3^p. In Table 1, `ratio1` gives the ratio of the error with $h = \pi/11$ to that with $h = \pi/33$ while `ratio2` gives the ratio of the error with $h = \pi/33$ to that with $h = \pi/99$. These appear to approach $27 = 3^3$, indicating a convergence rate of $\mathcal{O}(h^3)$ as $h \to 0$. The rate of convergence is not uniform, being initially slower near the boundary, but even near the boundary, convergence was better than $\mathcal{O}(h^2)$. Indeed, convergence was sufficiently regular to suggest that $\mathcal{O}(h^3)$ extrapolation may be useful, even quite close to the boundary. For the direct problem [3, 5], extrapolation has been found to perform better with the corrected Numerov method than with the corrected second order method. It would be interesting to see whether similar results could be established for the inverse problem. However, even without extrapolation and even with quite small n, the results obtained by our method are already so accurate that if they were displayed by the (graphical) method used to present results in [21], the error would be too small to detect. Since the original version of this paper was written, the author has obtained further numerical results comparing Algorithm 1 with the method of [21] for a number of different potentials, q. These results, which he hopes to discuss more fully elsewhere, suggest that Algorithm 1 gives more information about q from a given set of eigenvalues than does the method of [21].

References

[1] A. L. Andrew. Eigenvectors of certain matrices. *Linear Algebra Appl.*, 7:151–162, 1973. C332

[2] A. L. Andrew. Error bounds for the modified Newton's method. *Bull. Austral. Math. Soc.*, 14:427–433, 1976. C329

[3] A. L. Andrew. Efficient computation of higher Sturm-Liouville eigenvalues. *Internat. Schriftenreihe Numer. Math.*, Birkhäuser, 86:1–9, 1988. C333

[4] 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. C327, C328, C329

[5] A. L. Andrew. Asymptotic correction of computed eigenvalues of differential equations. *Ann. Numer. Math.*, 1:41–51, 1994. C328, C333

[6] A. L. Andrew. Finite and continuous perturbations of matrix eigenvalues. *Appl. Math. Lett.* 11(1):47–51, 1998. C330, C331

[7] A. L. Andrew. Computation of derivatives of eigenvalues and eigenvectors of parameter-dependent matrices. In B. J. Noye, M. D. Teubner and A. W. Gill, editors, *Computational Techniques and Applications: CTAC97*, pages 51–58. World Scientific, 1998. C330

[8] A. L. Andrew. Centrosymmetric matrices. *SIAM Rev.* 40:697–698, 1998. C329

[9] A. L. Andrew. Asymptotic correction of Numerov's eigenvalue estimates with natural boundary conditions. *J. Comput. Appl. Math.*, 125:359–366, 2000. C329

[10] 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. C328, C329

[11] A. L. Andrew. Asymptotic correction of Numerov's eigenvalue estimates with general boundary conditions. *ANZIAM J.*, 44(E):C1–C19, 2003. [Online] http://anziamj.austms.org.au/V44/CTAC2001/Andr. C329, C332

[12] A. L. Andrew. Asymptotic correction of more Sturm-Liouville eigenvalue estimates. *BIT*, 43:485–503, 2003. C329

[13] A. L. Andrew and J. W. Paine. Correction of Numerov's eigenvalue estimates. *Numer. Math.* 47:289–300, 1985. C329

[14] A. L. Andrew and J. W. Paine. Correction of finite element estimates for Sturm-Liouville eigenvalues. *Numer. Math.* 50:205–215, 1986. C328

[15] G. Borg. Eine Umkehrung der Sturm–Liouvilleschen Eigenwertaufgabe. *Acta Math.* 78:1–96, 1946. C327

[16] M. T. Chu. Inverse eigenvalue problems. *SIAM Rev.* 40:1–39, 1998. C327

[17] M. T. Chu and G. H. Golub. Structured inverse eigenvalue problems. *Acta Numer.* 11:1–71, 2002. C327

[18] C. De Boor and G. H. Golub. The numerically stable reconstruction of a Jacobi matrix from spectral data. *Linear Algebra Appl.* 21:245–260, 1978. C328

[19] C. R. Dun. *Algebraic correction methods for two-dimensional eigenvalue problems.* PhD thesis, Australian National University, Canberra, 1995. C328, C329

[20] 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. C328, C329

[21] 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. C329, C330, C332, C333

[22] O. H. Hald. *On discrete and numerical Sturm-Liouville problems.* PhD thesis, New York University, New York 1972. C328

[23] O. H. Hald. The inverse Sturm-Liouville problem and the Rayleigh-Ritz method. *Math. Comput.* 32:687–705, 1978. C328

[24] 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. C330

[25] J. T. Marti. Small potential corrections for the discrete eigenvalues of the Sturm-Liouville problem. *Numer. Math.* 57:51–62, 1990. C328, C329, C332

[26] J. R. McLaughlin. Analytical methods for recovering coefficients in differential equatiuons from spectral data. *SIAM Rev.* 28:53–72, 1986. C327

[27] P. Morel. Des algorithmes pour le problème inverse des valeurs propres. *Linear Algebra Appl.* 13:251–273, 1976. C327, C328

[28] J. Paine. A numerical method for the inverse Sturm-Liouville problem. *SIAM J. Sci. Stat. Comput.* 5:149–156, 1984. C327, C328, C329

[29] J. W. Paine, F. R. de Hoog and R. S. Anderssen. On the correction of finite difference eigenvalue approximations for Sturm–Liouville problems. *Computing* 26:123–139, 1981. C328, C329

[30] M. H. Pirovino. *Das Sturm-Liouville-Problem als direktes und inverses Eigenwertproblem und seine numerische Behandlung durch finite Differenzen.* Dissertation, ETH Zürich, 1992. C328, C329, C332

[31] J. D. Pryce. *Numerical solution of Sturm-Liouville problems.* Oxford University Press, 1993. C327