A Divide-and-Conquer Method for the Tridiagonal Generalized Eigenvalue Problem

L. Elsner, A. Fasse and E. Langmann *
Fakultät für Mathematik
Universität Bielefeld
Postfach 100131
33501 Bielefeld
Germany

December 17, 1996

Abstract

A divide-and-conquer method is developed for solving the generalized eigenvalue problem $Ax = \lambda Bx$, where A, B are real symmetric tridiagonal and B is positive definite. It is a generalization of Cuppens method for the standard eigenvalue problem, where $B = I$, being based on rank one modifications. This method is complementing a method developed by Borges-Gragg using bordering modifications.

1 Introduction

In this note we discuss a numerical method for solving the tridiagonal eigenvalue problem

$$Ax = \lambda Bx, \tag{1}$$

where A, B are real symmetric tridiagonal matrices and B is in addition positive definite.

In [3] Cuppen introduced a divide-and-conquer method for the case $B = I$. It is based on rank one modifications.

*Dedicated to W.B. Gragg at the occasion of his 60. birthday

Borges and Gragg, using modifications by bordering, developed later on in [1] a method for the general case (1). Independently Gu studied the bordering approach in [5] for the standard eigenvalue problem $B = I$.

Somewhat earlier, around 1988, the first author had sketched a generalization of Cuppens approach, namely using rank one modifications, to the case that B too is tridiagonal. In the Diplomarbeit [7] an algorithm based on these notes was developed and tested. As the numerical results seemed not to be very promising, they were never published outside [7]. Inspired by the publication of [1] the study of this topic was resumed in the Diplomarbeit [8]. By introducing a better solver for the secular equation numerically satisfying results were achieved. It turned out that further improvements were possible by using an updating for the eigenvectors developed by Gu-Eisenstat, see [6].

Encouraged by some people in the audience of a talk at the Householder meeting in Pontresina 1996, and in order to complete the picture we dare to publish these results belatedly.

2 The algorithm

In the sequel we develop an algorithm for (1). It consists of two phases, a *divide* or splitting phase, and a *conquer*

or construction phase.

Let

$$A-\lambda B = \begin{pmatrix} a_1 & \alpha_1 & & & 0 \\ \alpha_1 & a_2 & \alpha_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \alpha_{n-2} & a_{n-1} & \alpha_{n-1} \\ 0 & & & \alpha_{n-1} & a_n \end{pmatrix} - \lambda \begin{pmatrix} b_1 & \beta_1 & & & 0 \\ \beta_1 & b_2 & \beta_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \beta_{n-2} & b_{n-1} & \beta_{n-1} \\ 0 & & & \beta_{n-1} & b_n \end{pmatrix} \tag{2}$$

In the divide phase we find a split index s, which is typically near $n/2$ and a vector v such that

$$B = \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix} - vv^T, \tag{3}$$

where B_1 is s -by - s. Here v is given by

$$v_s = \sqrt{(|\beta_s|)}, \quad v_{s+1} = -sgn(\beta_s)v_s, \quad v_j = 0 \quad else \tag{4}$$

Observe that B_1 and B_2 are positive definite too and that $A + \epsilon vv^T$ splits up in a direct sum of tridiagonal matrices, for $\epsilon = \alpha_s/\beta_s$, so we have

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} - \epsilon vv^T. \tag{5}$$

Here again A_1 is s-by-s.

This requires $\beta_s \neq 0$. If $\beta_s = 0$, then B splits up anyhow and we can chose the vector v with

$$v_s = \sqrt{|\alpha_s|}, \quad v_{s+1} = sgn(\alpha_s)v_s, \quad v_j = 0 \quad else, \tag{6}$$

which leads to

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} + vv^T. \tag{7}$$

Suppose now that we have solved the subproblems

$$A_i x = \lambda B_i x, \qquad i = 1, 2, \tag{8}$$

i.e. we have determined matrices $Q_i, i = 1, 2$ satisfying

$$Q_1^T (A_1 - \lambda B_1) Q_1 = D_1 - \lambda I_s \tag{9}$$

$$Q_2^T (A_2 - \lambda B_2) Q_2 = D_2 - \lambda I_{n-s} \tag{10}$$

with diagonal matrices $D_i, i = 1, 2$. Then defining $Q = Q_1 \oplus Q_2, D = D_1 \oplus D_2$ and $w = Q^T v$ we get

$$Q^T (A - \lambda B) Q = (D - \epsilon ww^T) - \lambda(I - ww^T) \tag{11}$$

in the case $\beta_s \neq 0$, while for $\beta_s = 0$ we get

$$Q^T (A - \lambda B) Q = (D + ww^T) - \lambda I. \tag{12}$$

Observe that we have $w^T w < 1$ in (11), as B and hence $I - ww^T$ is positive definit. In the conquer phase we have to solve the problems (8) and then solve (11) or (12). While the solution of (12) is well established, see e.g. [6], [2], we have to develop a solution for (11).

This is done in the next three sections. First we reduce the solution of (11) to the problem of finding roots of a certain rational function and determine intervals in which these roots are located. Then we develop a solver in each of the intervals using rational interpolation. This gives the approximate eigenvalues. Then we discuss the calculation of the eigenvectors which are used in the recursive solution of the problem, namely to find the matrices Q_i in (9) and (10).

Though this approach parallels that of Cuppen [3], the appearance of $B \neq I$ complicates things to a certain extent. In particular the secular equation, whose roots are the eigenvalues, has an additional term which destroys monotonicity. This is in contrast to the methods in [1], where for $B \neq I$ the form of the secular equation does not change.

3 Solving $(D - \epsilon ww^T)x = \lambda(I - ww^T)x$

Here we discuss the solution of

$$(D - \epsilon ww^T)x = \lambda(I - ww^T)x. \tag{13}$$

We first simplify the problem by deflation. Deflation for the problem (13) is possible in the following cases:

Case (i): $w_i = 0$, then $\lambda_i = d_i$ is an eigenvalue with corresponding eigenvector $x = e_i$, where e_i is the i-th unit vector.

Case (ii): $d_i = d_j$ for some $i \neq j$. Then a suitable Givens rotation in the $i - j$-plane leaves D invariant, but produces a zero entry in position j of w. So we have case (i) and can deflate.

Case (iii): $\epsilon = d_i$ for some i. Also here it is easy to check that $x = e_i$ is eigenvector to the eigenvalue $\lambda_i = \epsilon = d_i$.

Hence we can assume in the following that

$$d_1 < d_2 < \ldots < d_n,$$

and that all w_i are nonzero. We do not yet rule out Case (iii).

Theorem 1 *For the generalized eigenvalue problem*

$$(D - \epsilon ww^T)x = \lambda(I - ww^T)x$$

assume that

$$w^T w < 1$$
$$w_i \neq 0, i = 1, \ldots, n$$
$$d_1 < d_2 < \ldots < d_n.$$

If $\epsilon \neq d_i$ for all i then the eigenvalues are given by the roots of the rational function

$$f(\lambda) = \sum_{i=1}^{n} \frac{w_i^2(d_i - \epsilon)}{\lambda - d_i} - 1 + w^T w.$$

There is exactly one eigenvalue in each of the intervals $I_r = (d_r, d_{r+1}), r = 0, \ldots, n$ except that one containing ϵ. Here it is understood that $d_0 = -\infty$ and $d_{n+1} = \infty$.

If $\epsilon = d_i$ for some i, then $\lambda = \epsilon$ is an eigenvalue with eigenvector $x = e_i$, while the other eigenvalues are the n-1 roots of $f(\lambda)$ in $(d_r, d_{r+1}), r \neq i-1, i$.

In all these cases the eigenvector belonging to λ is given by

$$x = (D - \lambda I)^{-1} w.$$

Proof A simple calculation gives

$$T = (I - ww^T)^{-1}(D - \epsilon ww^T) = D + (1 - w^T w)^{-1} ww^T (D - \epsilon I).$$

We remark that T is a *nonsymmetric* rank one modification of D having n real eigenvalues. Assume that λ is an eigenvalue of T with eigenvector x and $\lambda \neq d_i \quad \forall i$. Then from

$$Dx + (1 - w^T w)^{-1} ww^T (D - \lambda I)x = \lambda x \tag{14}$$

we get

$$x = (D - \lambda I)^{-1} w \tag{15}$$

and

$$(1 - w^T w)^{-1} w^T (D - \epsilon I)(\lambda I - D)^{-1} w = 1,$$

whence $f(\lambda) = 0$. On the other hand, if $\lambda \neq d_i$ is a number satisfying $f(\lambda) = 0$, and x is given by (15), then $Tx = \lambda x$.

If $\lambda = d_i$ is an eigenvalue then (14) and $w_i \neq 0$ implies $w^T(D - \epsilon I)x = 0$. But then again (14) implies $Dx = \lambda x$, hence $x = e_i$ and $\lambda = \epsilon = d_i$, a contradiction.

Now for $d_i < \epsilon$ we have

$$\lim_{s \to +0} f(d_i + s) = -\infty$$

and

$$\lim_{s \to +0} f(d_i - s) = +\infty,$$

while for $d_i > \epsilon$ we have the opposite limiting behaviour at the pole d_i of $f(\lambda)$. This shows that in the cases $\epsilon < d_i < d_{i+1}, \quad i < n$ there is a root of f in (d_i, d_{i+1}). Also $\lim_{\lambda \to \infty} f(\lambda) < 0$ so there is an additional root in (d_n, d_{n+1}), if we are not in case that $\epsilon > d_n$. A similar reasoning gives the inclusion for intervals left of ϵ. As f has only n roots, we get the uniqueness of the roots in the intervals.

If $\epsilon = d_i$ for some i, then it is obvious that $\lambda = \epsilon$ is an eigenvalue with eigenvector e_i. For the other eigenvalues we have that they are roots of f. As above we see that there are roots in the intervals $(d_j, d_{j+1}) j \neq i, i-1$.

□

4 Solving the secular equation

Instead of solving $f(\lambda) = 0$ we look for zeros of the function

$$g(\lambda) = (\epsilon - \lambda)^{-1} f(\lambda) = \sum_{i=1}^{n} \frac{w_i^2}{d_i - \lambda} - \frac{1}{\epsilon - \lambda}. \tag{16}$$

We call g the secular equation. There are several possibilities to construct solvers for $g(\lambda) = 0$ in a fixed interval. Here we describe that solver which showed the best numerical behaviour. The mathematical analysis is not very complete.

We first describe the solver for *interior* (finite) intervals.

Let $1 \leq j \leq n-1$ and assume that $I_j = (d_j, d_{j+1})$ contains one and hence exactly one root , say μ, of the secular equation. Hence $\epsilon \neq I_j$. Let $x_i \in I_j$ be an approximation to μ, which we want to improve. For this, we replace g in I_j by

$$\Phi_i(x) = p + \frac{r}{d_j - x} + \frac{s}{d_{j+1} - x} - \frac{1}{\epsilon - x}$$

and determine the parameters p, r, s such that

$$\Phi_i^{(\nu)}(x_i) = g^{(\nu)}(x_i), \nu = 0, 1, 2. \tag{17}$$

The solution of this linear system is given by

$$s = w_{j+1}^2 + \frac{(d_{j+1} - x_i)^3}{d_{j+1} - d_j} {\sum_k}' \frac{w_k^2 (d_k - d_j)}{(d_k - x_i)^3},$$

$$r = w_j^2 - \frac{(d_j - x_i)^3}{d_{j+1} - d_j} \sum_k{}' \frac{w_k^2 (d_k - d_{j+1})}{d_k - x_i)^3},$$

$$p = \sum_k{}' \frac{w_k^2 (d_s - d_{j+1})(d_k - d_j)}{(d_k - x_i)^3}.$$

Here $\sum'$ indicates that in the sum the indices j and $j+1$ are left out.

We observe that $r > 0$ and $s > 0$. Hence $\Phi_i(x) = 0$ has at least one solution in I_j. Unfortunately we were not able to prove that in all cases there is only one such root in I_j, though this was never observed numerically. We take this root (or in general the root in I_j nearest to x_i) as new approximation x_{i+1} to μ.

For *outer* (infinite) intervals we interpolate $g(x)$ by

$$\Phi_i(x) = p + rx + \frac{s}{d_1 - x} - \frac{1}{\epsilon - x}$$

in the case $I_0 = (-\infty, d_1)$, determining the parameters p, r, s by requiring (17). This gives

$$s = w_1^2 + \sum_{j=2}^{n} \frac{w_j^2 (d_1 - x_i)}{(d_j - x_i)^3},$$

$$r = \sum_{j=2}^{n} \frac{w_j^2 (d_j - d_1)}{(d_j - x_i)^3},$$

and

$$p = \sum_{j=2}^{n} \frac{w_j^2 (d_j - d_1)(d_j + d_1 - 3x_i)}{(d_j - x_i)^3}.$$

As in the finite case we see that $r > 0$ and $s > 0$. Hence there is at least one solution of $\Phi_i(x) = 0$ in I_0. Again we were not able to show that there is only one such root in I_0. Similar formulas hold for I_n.

In all cases determining this root amounts to the solution of a cubic equation. In our experiments we used the MATLAB routine *roots*. Cubic convergence, as expected by (17), was observed throughout.

5 Calculating the eigenvectors

After having determined approximate eigenvalues $\tilde{\lambda}_i$, for the conquer phase we have to calculate the associated eigenvectors $x^{(i)}, i = 1, \ldots, n$. According to Theorem 1 we take

$$x_r^{(i)} = w_r(d_r - \tilde{\lambda}_i)^{-1}, i, r = 1, \ldots, n.$$

While this leads often to satisfactory results, an improvement could be achieved by the following modification. Due to the errors in the eigenvalue these vectors are not orthogonal with respect to the inner product given by $I - ww^T$. Following an idea in [6] and also applied in [1] we replace the vector w by a vector $\tilde{w}$ in such a way that the $\tilde{\lambda}_i$ are the exact eigenvalues of the modified problem

$$(D - \epsilon \tilde{w}\tilde{w}^T)x = \lambda(I - \tilde{w}\tilde{w}^T)x.$$

Then we take

$$x_r^{(i)} = \tilde{w}_r(d_r - \tilde{\lambda}_i)^{-1}, i, r = 1, \ldots, n$$

as approximate eigenvectors. We show in the following theorem that this inverse eigenvalue problem is always solvable. The solution is given in the course of the proof.

Theorem 2 *Let*

$$-\infty = d_0 < d_1 < d_2 < \ldots < d_n < d_{n+1} = \infty$$

and $\epsilon \neq d_i, i = 1, \ldots, n$. *Let* $I_j = (d_j, d_{j+1}), j = 0, \ldots, n$ *and* $\epsilon \in I_s$. *Let* n *numbers* ξ_i *be given, such that there is one* ξ_i *in each of the intervals* $I_j, j \neq s$. *Then there exists a real vector* $v \in R^n$ *with* $v^T v < 1$ *so that the eigenvalue problem*

$$(D - \epsilon vv^T)x = \lambda(I - vv^T)x$$

has the eigenvalues $\xi_i, i = 1, \ldots, n$.

Proof For any $v \in R^n$, $v^T v < 1$ we have

$$(I - vv^T)^{-1}(D - \epsilon vv^T) - \lambda I = (D - \lambda I)(I + (1 - v^T v)^{-1}(D - \lambda I)^{-1} vv^T (D - \epsilon I)).$$

We want to find v such that the determinant of the lefthand side is $\prod_{i=1}^{n}(\xi_i - \lambda)$. The determinant of the right hand side is given by

$$(1 + (1 - v^T v)^{-1} \sum_{j=1}^{n} \frac{v_j^2(d_j - \epsilon)}{d_j - \lambda}) \prod_{i=1}^{n}(d_i - \lambda).$$

Equating both terms for $\lambda = d_r, r = 1, \ldots, n$ gives after some manipulations

$$\frac{v_r^2}{1 - v^T v} = \frac{\prod_{i=1}^n (\xi_i - d_r)}{(d_r - \epsilon) \prod_{i \neq r} (d_i - d_r)} = h_r \quad r = 1, \ldots, n.$$

This leads to

$$v_r^2 = \frac{h_r}{1 + \sum_{i=1}^n h_i}, \quad r = 1, \ldots, n. \tag{18}$$

A somewhat tedious argument using the assumptions on the position of the ξ_i with respect to the d_i and ϵ gives

$$h_r > 0, r = 1, \ldots, n.$$

So there are real solutions of the problem and they are given by (18). $\Box$.

One chooses in (18) that root v_r which has the same sign as the original w_r. Observe that h_r can be calculated to high relative accuracy, as only products are involved.

6 Concluding remarks

In this paper it was our main intention to present an alternative approach to the numerical solution of the generalized tridiagonal eigenvalue problem. We have stressed the points in which we differ from the earlier papers [3] and [1]. Here we will touch upon the issue of numerical implementation and numerical experiments, but elaborating on this point is beyond the scope of these notes.

In [8] we have run many examples to compare the methods described above and the method of [7]. There a simpler solver was used. Also the modification of the eigenvector was not yet known then. In particular several choices of the tolerances used in the solver and in the deflation procedure were tested.

Our algorithm turned out to be stable and quite reliable. Due to the cubic convergence only very few steps were necessary to determine the eigenvalue in an interval.

In many cases using the modified version of the eigenvectors improved the accuracy of the result, though not always. This happened in particular, if $w - \tilde{w}$ was not very small.

A comparison with the method in [1] was not intended. We would expect that due to the faster size reduction and due to the fact that determining the zero of the solver in the latter method amounts to solving only a quadratic equation, while ours is cubic, the latter method might be faster.

References

[1] C.F. Borges and W.B. Gragg *A parallel Divide - and - Conquer Method for the Generalized Real Symmetric Definite Tridiagonal Eigenproblem,* Numerical Linear Algebra, Proceedings of the Conference in Numerical Linear Algebra and Scientific Computation, Kent(Ohio) 1992, Ed. L. Reichel, A. Ruttan and R.S. Varga, de Gruyter.

[2] J.R. Bunch, C.P. Nielsen and D.C. Sorensen *Rank one Modification of the Symmetric Eigenvalue problem,* Numerische Mathematik 31, 31-48,(1978)

[3] J.J. Cuppen *A Divide-and-Conquer Method for the Symmetric Tridiagonal Eigenproblem,* Numerische Mathematik 36, 177-195, (1981)

[4] J.J. Dongarra and D.C. Sorensen *A fully parallel algorithm for the Symmetric Eigenvaalue problem,* SIAM J. Sci. Stat. Comp., 8, 139-154, (1987)

[5] M. Gu *Studies in Numerical Linear Algebra,* Ph.D. Thesis, Department of Computer Science, Yale University (1993)

[6] M. Gu and S. Eisenstat *A Stable and Efficient Algorithm for the rank one Modification of the Symmetric Eigenproblem,* SIMAX 15, 1266-1276, (1994)

[7] E. Langmann *Eine Divide-and-Conquer Methode für das tridiagonale Eigenwertproblem* $Ax = \lambda Bx$, Diplomarbeit 1991, Fakultät für Mathematik, Universität Bielefeld.

[8] A. Fasse *Rekursive Divide-and-Conquer Verfahren für das tridiagonale Eigenwertproblem* $Ax = \lambda Bx$, Diplomarbeit 1996, Fakultät für Mathematik, Universität Bielefeld