ELSEVIER Linear Algebra and its Applications 358 (2003) 139–144
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

Inner deflation for symmetric tridiagonal matrices

I.S. Dhillon [a], A.N. Malyshev [b,*]

[a] Department of Computer Sciences, University of Texas, Austin, TX 78712-1188, USA
[b] Department of Informatics, University of Bergen, 5020 Bergen, Norway

Received 20 March 2001; accepted 23 August 2001

Submitted by B.N. Parlett

Abstract

Suppose that one knows an accurate approximation to an eigenvalue of a real symmetric tridiagonal matrix. A variant of deflation by the Givens rotations is proposed in order to split off the approximated eigenvalue. Such a deflation can be used instead of inverse iteration to compute the corresponding eigenvector.

Keywords: Eigenvector; Symmetric tridiagonal matrix; Deflation; Inverse iteration

1. Introduction

Suppose that λ approximates an exact eigenvalue λ_* of a real symmetric tridiagonal matrix

$$T = \begin{bmatrix} a_1 & b_1 & & & \\ b_1 & a_2 & b_2 & & \\ & \cdot & \cdot & \cdot & \\ & & \cdot & \cdot & b_{n-1} \\ & & & b_{n-1} & a_n \end{bmatrix}.$$

If $b_i \neq 0$, $i = 1, \dots, n-1$, then it is very tempting to compute an eigenvector x of T corresponding to λ_* by solving the almost singular system $(T - \lambda I)x = 0$ downwards:

* Corresponding author.

0024-3795/02/$ - see front matter
PII: S0024-3795(01)00479-7

$$x_1 = 1, \quad x_2 = -(a_1 - \lambda)/b_1,$$
$$x_i = -\big[b_{i-2}x_{i-2} + (a_{i-1} - \lambda)x_{i-1}\big]/b_{i-1}, \quad i = 3, \dots, n.$$

However, the above process is not always successful because the last equation $b_{n-1} x_{n-1} + (a_n - \lambda)x_n = 0$ may be strongly violated. The main reason is not the round-off error in computer arithmetic but the slight departure of λ from the true eigenvalue λ_*. Indeed the last equation may be violated even if all operations are done in exact arithmetic but with a slightly inaccurate λ. Wilkinson [6,7] tried to cure this defect but could not succeed. Only in 1983, Godunov [4] found an elegant solution to this problem. A detailed exposition of Godunov's solution is found, e.g., in [3]. In 1995, Fernando [2] independently proposed his own solution, which is applicable in more general situation and looks slightly simpler than Godunov's solution.

The present work was initially motivated by an attempt to combine Fernando's approach with the deflation techniques developed by Godunov and his collaborators, and Sorensen's implicitly restarted Lanczos [5] was considered among possible applications. In pursuit of this goal a new deflation procedure was discovered, which can be referred to as an "inner deflation". In the classical QR algorithm, deflation is restricted to one of the ends of the tridiagonal band. However, in the proposed inner deflation the Givens rotations start from both ends of the band and meet inside it. The deflated eigenvalue emerges at the meeting point on the main diagonal of the transformed tridiagonal matrix. The rest of the transformed matrix forms a tridiagonal band with a bulge near the meeting point that can be chased in any direction. The inner deflation is simple and robust and provides an alternative to inverse iteration for computing several eigenvectors of T. The inner deflation idea has previously been outlined in [1, Section 3.5].

2. The inner deflation

Let us denote the QR factorization of $T - \lambda I$ by Q^+R and QL factorization of $T - \lambda I$ by Q^-L. Then

$$Q^+ = G_1^+G_2^+ \cdots G_{n-1}^+, \quad Q^- = G_{n-1}^-G_{n-2}^- \cdots G_1^-,$$

where the Givens rotations

$$G_i^+ = \begin{pmatrix} I & & & \\ & c_i^+ & s_i^+ & \\ & -s_i^+ & c_i^+ & \\ & & & I \end{pmatrix}, \quad G_i^- = \begin{pmatrix} I & & & \\ & c_i^- & s_i^- & \\ & -s_i^- & c_i^- & \\ & & & I \end{pmatrix}$$

have the blocks

$$\begin{pmatrix} c_i^\pm & s_i^\pm \\ -s_i^\pm & c_i^\pm \end{pmatrix}$$

at the intersection of rows i, $i+1$ and columns i, $i+1$. The upper triangular matrix R is computed downwards:

$$
\begin{aligned}
&r_{11} = a_1 - \lambda, \quad r_{12} = b_1,\\
&\begin{bmatrix} c_i^+ & -s_i^+ \\ s_i^+ & c_i^+ \end{bmatrix}\begin{bmatrix} r_{ii} & r_{i,i+1} & 0 \\ b_i & a_{i+1}-\lambda & b_{i+1} \end{bmatrix} = \begin{bmatrix} R_{ii} & R_{i,i+1} & R_{i,i+2} \\ 0 & r_{i+1,i+1} & r_{i+1,i+2} \end{bmatrix},\\
&\quad i = 1, \ldots, n-2,\\
&\begin{bmatrix} c_{n-1}^+ & -s_{n-1}^+ \\ s_{n-1}^+ & c_{n-1}^+ \end{bmatrix}\begin{bmatrix} r_{n-1,n-1} & r_{n-1,n} \\ b_{n-1} & a_n-\lambda \end{bmatrix} = \begin{bmatrix} R_{n-1,n-1} & R_{n-1,n} \\ 0 & R_{nn} \end{bmatrix},
\end{aligned}
$$

while the lower triangular matrix L is computed upwards:

$$
\begin{aligned}
&l_{n,n-1} = b_{n-1}, \quad l_{nn} = a_n - \lambda,\\
&\begin{bmatrix} c_i^- & -s_i^- \\ s_i^- & c_i^- \end{bmatrix}\begin{bmatrix} b_{i-1} & a_i-\lambda & b_i \\ 0 & l_{i+1,i} & l_{i+1,i+1} \end{bmatrix} = \begin{bmatrix} l_{i,i-1} & l_{ii} & 0 \\ L_{i+1,i-1} & L_{i+1,i} & L_{i+1,i+1} \end{bmatrix},\\
&\quad i = n-1, \ldots, 2,\\
&\begin{bmatrix} c_1^- & -s_1^- \\ s_1^- & c_1^- \end{bmatrix}\begin{bmatrix} a_1-\lambda & b_1 \\ l_{21} & l_{22} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix}.
\end{aligned}
$$

Let us fix some k, $1 \leqslant k \leqslant n-1$, and apply the $k-1$ rotations $G_1^+, G_2^+, \ldots, G_{k-1}^+$ and $n-k-1$ rotations $G_{n-1}^-, G_{n-2}^-, \ldots, G_{k+1}^-$ to the matrix $T - \lambda I$ from the left. The result will be

$$
\begin{aligned}
&\left(G_1^+ G_2^+ \cdots G_{k-1}^+ G_{n-1}^- G_{n-2}^- \cdots G_{k+1}^-\right)^{\mathrm{T}}(T-\lambda I)\\
&= \begin{bmatrix}
R_{11} & R_{12} & R_{13} & & & & & \\
 & \cdot & \cdot & \cdot & & & & \\
 & & R_{k-1,k-1} & R_{k-1,k} & R_{k-1,k+1} & & & \\
 & & & r_{kk} & r_{k,k+1} & & & \\
 & & & l_{k+1,k} & l_{k+1,k+1} & & & \\
 & & & L_{k+2,k} & L_{k+2,k+1} & L_{k+2,k+2} & & \\
 & & & & \cdot & \cdot & \cdot & \\
 & & & & & L_{n,n-2} & L_{n,n-1} & L_{nn}
\end{bmatrix}.
\end{aligned}
$$

After this an additional Givens rotation

$$
G_k = \begin{pmatrix} I & & & \\ & c_k & s_k & \\ & -s_k & c_k & \\ & & & I \end{pmatrix}
$$

is applied in order to get a zero at entry $(k, k+1)$:

$$
\begin{bmatrix} c_k & -s_k \\ s_k & c_k \end{bmatrix}\begin{bmatrix} r_{kk} & r_{k,k+1} \\ l_{k+1,k} & l_{k+1,k+1} \end{bmatrix} = \begin{bmatrix} \check{L}_{kk} & 0 \\ \check{L}_{k+1,k} & \check{L}_{k+1,k+1} \end{bmatrix}.
$$

Note that we could similarly have applied a rotation to zero out entry $(k+1, k)$ to obtain a singleton in the $(k+1)$st row. For the sake of convenience we introduce the matrices $Q_n = Q^+$ and

$$Q_k = G_1^+ G_2^+ \cdots G_{k-1}^+ G_{n-1}^- G_{n-2}^- \cdots G_{k+1}^- G_k, \quad 1 \leqslant k \leqslant n-1.$$

As demonstrated above the only possible nonzero entry in the kth row of $N_k = Q_k^{\mathrm{T}}(T - \lambda I)$ is $\check{L}_{kk}$ for $1 \leqslant k \leqslant n-1$ and R_{nn} for $k = n$. At the same time the matrix $Q_k^{\mathrm{T}}(T - \lambda I) Q_k$ has the following structure:

$$Q_k^{\mathrm{T}}(T - \lambda I) Q_k = \begin{bmatrix} \times & \times & & & & & & & & & \\ \times & \times & \times & & & & & & & & \\ & \cdot & \cdot & \cdot & & & & & & & \\ & & \times & \times & \times & & & & & & \\ & & & \times & \times & \times & \times & \times & & & \\ & & & & \times & \times & \times & \times & & & \\ & & & & \times & \times & \times & \times & & & \\ & & & & \times & \times & \times & \times & \times & & \\ & & & & & & & \times & \times & \times & \\ & & & & & & & & \cdot & \cdot & \cdot \\ & & & & & & & & & \times & \times \end{bmatrix},$$

which is a real symmetric tridiagonal matrix plus a bulge at entries $(k+2, k-1)$ and $(k-1, k+2)$. Since the 2-norms of the kth rows of $Q_k^{\mathrm{T}}(T - \lambda I) Q_k$ and $Q_k^{\mathrm{T}}(T - \lambda I)$ are equal, the norm of the kth row of $Q_k^{\mathrm{T}}(T - \lambda I) Q_k$ is Γ_k, where

$$\Gamma = \left[\Gamma_1, \Gamma_2, \ldots, \Gamma_n\right] = \left[|\check{L}_{11}|, |\check{L}_{22}|, \ldots, |\check{L}_{n-1,n-1}|, |R_{nn}|\right].$$

Let us choose k such that $\Gamma_k = \min_i \Gamma_i$. It is shown in the next section that $\min_i \Gamma_i \leqslant \sqrt{n}\,|\lambda_* - \lambda|$. Therefore, if $\sqrt{n}\,|\lambda_* - \lambda|$ is negligibly small, then the kth row and column of $Q_k^{\mathrm{T}}(T - \lambda I) Q_k$ are approximately null, i.e.,

$$Q_k^{\mathrm{T}}(T - \lambda I) Q_k \approx \begin{bmatrix} \times & \times & & & & & & & & & \\ \times & \times & \times & & & & & & & & \\ & \cdot & \cdot & \cdot & & & & & & & \\ & & \times & \times & \times & & & & & & \\ & & & \times & \times & 0 & \times & \times & & & \\ & & & & 0 & 0 & 0 & 0 & & & \\ & & & & \times & 0 & \times & \times & & & \\ & & & & \times & 0 & \times & \times & \times & & \\ & & & & & & & \times & \times & \times & \\ & & & & & & & & \cdot & \cdot & \cdot \\ & & & & & & & & & \times & \times \end{bmatrix}. \qquad (1)$$

In order to recognize the latter structure we use the permutation matrix

$$\Pi_k = \begin{pmatrix} I_{k-1} & 0 & 0 \\ 0 & 0 & 1 \\ 0 & I_{n-k} & 0 \end{pmatrix},$$

which is obtained from I by moving its kth column to the end. Then

$$\Pi_k^{\mathrm{T}} Q_k^{\mathrm{T}} (T - \lambda I) Q_k \Pi_k \approx \begin{pmatrix} \widehat{T} & 0 \\ 0 & 0 \end{pmatrix},$$

where $\widehat{T}$ is symmetric tridiagonal with a bulge at entries $(k-1, k+1)$ and $(k+1, k-1)$:

$$\begin{pmatrix} \widehat{T} & 0 \\ 0 & 0 \end{pmatrix} = \begin{bmatrix} \times & \times & & & & & & & & & & \\ \times & \times & \times & & & & & & & & & \\ & \cdot & \cdot & \cdot & & & & & & & & \\ & & \times & \times & \times & & & & & & & \\ & & & \times & \times & \times & \times & & & & & \\ & & & & \times & \times & \times & & & & & \\ & & & & \times & \times & \times & \times & & & & \\ & & & & & & \times & \times & \times & & & \\ & & & & & & & \cdot & \cdot & \cdot & & \\ & & & & & & & & \times & \times & \times & \\ & & & & & & & & & \times & \times & 0 \\ & & & & & & & & & & 0 & 0 \end{bmatrix}.$$

The bulge can be easily chased down or up by suitable Givens rotations.

It is easy to show that one inner deflation costs $\Theta(n)$ arithmetic operations. Note that in (1), the kth column of Q_k, i.e., $Q_k e_k$, is an approximate eigenvector. Thus m eigenvectors can be recovered from m successive inner deflations in $\Theta(m^2 n)$ operations.

3. Error analysis

If x is a unit eigenvector of T corresponding to λ_*, then $(T - \lambda I)x = (\lambda_* - \lambda)x$. Since the only possible nonzero element of the kth row of $Q_k^{\mathrm{T}}(T - \lambda I)$ lies on the main diagonal and its absolute value equals Γ_k, $\Gamma_k |x_k| = |e_k^{\mathrm{T}}[Q_k^{\mathrm{T}}(T - \lambda I)]x| = |\lambda_* - \lambda||e_k^{\mathrm{T}} Q_k^{\mathrm{T}} x|$. Let us choose k_0 such that $|x_{k_0}| = \max_i |x_i|$, whence $|x_{k_0}| \geqslant 1/\sqrt{n}$ owing to $\|x\|_2 = 1$. From the equality $\Gamma_k = |\lambda_* - \lambda||(Q_k e_k)^{\mathrm{T}} x|/|x_k|$ we derive the promised estimate

$$\min_i \Gamma_i \leqslant \Gamma_{k_0} \leqslant \sqrt{n}\, |\lambda_* - \lambda|.$$

Note that $Q_{k_0} e_{k_0}$ approximates the eigenvector x, so $|(Q_{k_0} e_{k_0})^{\mathrm{T}} x|$ is close to 1 when λ_* is an isolated eigenvalue.

Now the above analysis is modified in order to take rounding errors into account. Note that the effect of underflow is negligible in our case. It is well known that the computed value of $N_k = Q_k^{\mathrm{T}}(T - \lambda I)$ may be written as $\widetilde{N}_k = \widetilde{Q}_k^{\mathrm{T}}(T - \lambda I) + \Delta_k$, where $\widetilde{Q}_k$ and $\widetilde{N}_k$ are the computed values of Q_k and N_k, respectively, and $\|\Delta_k\|_2 \leqslant C(n)\epsilon \|T - \lambda I\|_2$, where $C(n)$ is a polynomial of small degree with coefficients of order $\mathrm{O}(1)$ and ϵ is the machine epsilon. Thus

$$\begin{aligned}\pm\widetilde{\Gamma}_k x_k &= e_k^{\mathrm{T}} \widetilde{N}_k x \\ &= e_k^{\mathrm{T}}\big[\widetilde{Q}_k^{\mathrm{T}}(T-\lambda I)\big]x + e_k^{\mathrm{T}}\Delta_k x \\ &= e_k^{\mathrm{T}}\widetilde{Q}_k^{\mathrm{T}}(\lambda_* - \lambda)x + e_k^{\mathrm{T}}\Delta_k x,\end{aligned}$$

where $\widetilde{\Gamma}_k = |\widetilde{N}_k(k,k)|$ is the computed value of Γ_k. Again by choosing k_0 such that $|x_{k_0}| = \max_i |x_i|$ we derive the final estimate

$$\min_i \widetilde{\Gamma}_i \leqslant \widetilde{\Gamma}_{k_0} \leqslant \sqrt{n}\,|\lambda_* - \lambda| + \sqrt{n}C(n)\epsilon\|T-\lambda I\|_2.$$

4. Numerical example

Consider the positive definite tridiagonal matrix from [4],

$$T = \begin{bmatrix} 2 & 1 & & & \\ 1 & 1+\rho & \rho & & \\ & \rho & 2\rho & \rho & \\ & & \rho & 1+\rho & 1 \\ & & & 1 & 2 \end{bmatrix},$$

where $\rho \ll 1$. The smallest eigenvalue of this matrix lies in the interval $(0, 2\rho)$. Setting ρ to machine precision ($\approx 2.22 \times 10^{-16}$ in IEEE double precision arithmetic) and $\lambda = 0$, we find the computed Γ values to be:

$$\tilde{\Gamma} = \big[.666667 \quad .408248 \quad .444089 \times 10^{-15} \quad .408248 \quad .666667\big].$$

Thus in this case $k_0 = 3$ and no other value of k allows deflation.

Acknowledgement

We would like to thank Beresford Parlett for his constructive suggestions.

References

[1] I.S. Dhillon, A new O(n^2) algorithm for the symmetric tridiagonal eigenvalue/-eigenvector problem, PhD Thesis, University of California, Berkeley, CA, 1997.
[2] K.V. Fernando, On computing an eigenvector of a tridiagonal matrix. Part I: Basic results, SIAM J. Matrix Anal. Appl. 18 (4) (1997) 1013–1034.
[3] S.K. Godunov, A.G. Antonov, O.P. Kiriljuk, V.I. Kostin, Guaranteed Accuracy in Numerical Linear Algebra, Kluwer Academic Publishers, Dordrecht, 1993.
[4] S.K. Godunov, V.I. Kostin, A.D. Mitchenko, Computation of an eigenvector of a symmetric tridiagonal matrix, Institute of Mathematics, Siberian Branch of the USSR Academy of Sciences. Novosibirsk, 1983 p. 58 (in Russian, Preprint 44).
[5] R.B. Lehoucq, D.C. Sorensen, C. Yang, ARPACK User's Guide. Solution of Large-scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods, SIAM, Philadelphia, PA, 1998.
[6] J.H. Wilkinson, The calculation of eigenvectors of codiagonal matrices, Comput. J. 1 (1958) 90–96.
[7] J.H. Wilkinson, The Algebraic Eigenvalue Problem, Clarendon Press, Oxford, 1965.