CS 56

QD-METHOD WITH NEWTON SHIFT

BY

F. L. BAUER

TECHNICAL REPORT NO. 56
MARCH 1, 1967

Supported in part by a grant from
the National Science Foundation.

COMPUTER SCIENCE DEPARTMENT
School of Humanities and Sciences
STANFORD UNIVERSITY

QD-METHOD WITH NEWTON SHIFT

BY

F. L. Bauer*

*Mathematisches Institut der Technischen Hochschule, München and Computer Science Department, Stanford University.

This work was in part supported by NSF.

1.

For determining the eigenvalues of a symmetric, tridiagonal matrix, various techniques have proven to be satisfactory; in particular the bisection method and &R-transformation with shifts determined by the last diagonal element or by the last 2X2 principal minor. Bisection, based upon a Sturm sequence, allows one to concentrate on the determination of any prescribed set of roots, prescribed by intervals or by ordering numbers. The &R-transformation is faster,; however, it gives the eigenvalues in a non-predictable ordering and is therefore mainly advocated for the determination of all roots.

Theoretically, for symmetric matrices, a &R-step is equivalent to two successive LR-steps, and the LR-transformation for a tridiagonal matrix is, apart from organizational details, identical with the qd-method. For non-positive definite matrices, however, the LR-transformation cannot be guaranteed to be numerically stable unless pivotal interchanges are made. This has led to preference for the &R-transformation, which is always numerically stable.

If, however, some of the smallest or some of the largest eigenvalues are wanted, then the &R-transformation will not necessarily give only -these, and bisection might seem too slow with its fixed convergence rate of $1/2$. In this situation, Newton's method would be fine if the Newton correction can be computed sufficiently simply, since it will always tend monotonically to the nearest root starting from a point outside the spectrum. Consequently, if one always worked with positive (or negative) definite matrices, there would be no objection to using the now stable qd-algorithm. In particular, for the determination of some of the smallest roots of a matrix known to be positive definite -- this problem arises frequently in connection with finite difference or Ritz approximations to analytical eigenvalue problems -- the starting value zero would be usually a quite good initial approximation.

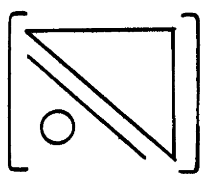
We shall show that for a qd-algorithm, the Newton correction can very easily be calculated, and accordingly a shift which avoids undershooting, or a lower bound. Since the last diagonal element gives an upper bound, the situation is quite satisfactory with respect to bounds.

2.

Let $\psi(\lambda) = (A - \lambda I)^{-1}$ be the resolvent of a matrix A . Then

$$\frac{d}{d\lambda}\psi(\lambda) = (A - \lambda I)^{-2} .$$

Assume that A is of Hessenberg form of order $n \geq 2$, viz,

A= 

with +1 's in the lower off-diagonal. Then since $A - \lambda I$ is again of Hessenberg form, the co-factor of the (1,n) element is ± 1 , and therefore the (n,1) -element of $(A - \lambda I)^{-1}$ is $-1/f(h)$, where $f(h)$ is the characteristic polynomial of A . From the result above we conclude that the (n,1) element of $(A - \lambda I)^{-2}$ is $f'(\lambda)/f^2(\lambda)$, and therefore $\delta(\lambda) = -f(\lambda)/f'(\lambda) = e_n^T(A - \lambda I)^{-1}e_1/e_n^T(A - \lambda I)^{-2}e_1$ is the Newton correction, $\lambda + \delta(\lambda)$ being the next approximation.

3.

The calculation of $e_n^T(A - \lambda I)^{-1}e_1$ and $e_n^T(A - \lambda I)^{-2}e_1$ can be based upon the solution of $(A - \lambda I)x = \alpha_1 e_1$ and $\alpha_2 y^T(A - \lambda I) = e_n^T$ by backward substitution starting with $(x)_n = 1$ and $(y)_1 = 1$. Then $\alpha_1 = \alpha_2$ $(= \alpha)$, $1/\alpha = e_n^T(A - \lambda I)^{-1}e_1$, and $y^T x/\alpha^2 = e_n^T(A - \lambda I)^{-2}e_1$; or $f(h) = \alpha$, $f'(h) = y^T x$, $6(h) = \alpha/y^T x$. The final result holds even when λ is an eigenvalue, α then being zero and $y^T x \neq 0$ unless A is defective. While for the approximation of eigenvectors this back-substitution, known as Hyman's technique, cannot be generally advocated, it offers a simple way to the calculation of $f'(\lambda) = y^T x$ and of the Newton correction. It can be used when Newton's method can safely be used; e.g., when the roots of the Hessenberg matrix A are known to be all real, in connection with some deflation technique.

4.

For tridiagonal matrices, however, the LR-transformation or the qd-algorithm gives the Newton correction as a by-product. In the qd version of the LR-transformation we perform first for a certain value of λ the triangular decomposition of $A - \lambda I$,

$$A - \lambda I = \begin{bmatrix} q_1 & & & & 0 \\ 1 & q_2 & & & \\ & 1 & q_3 & & \\ \bigcirc & & \ddots & \ddots & \\ & & & 1 & q_n \end{bmatrix} \begin{bmatrix} 1 & e_1 & & \bigcirc \\ & 1 & e_2 & \\ & & 1 & \ddots \\ & & & \ddots & e_{n-1} \\ & & & & 1 \end{bmatrix}$$

multiply the factors conversely and decompose again

$$\begin{bmatrix} 1 & e_1 & & & \bigcirc \\ & 1 & e_2 & & \\ & & \ddots & \ddots & \\ 0 & & & 1 & e_{n-1} \\ & & & & 1 \end{bmatrix} \begin{bmatrix} q_1 & & & & \bigcirc \\ 1 & q_2 & & & \\ & 1 & q_2 & & \\ & & 1 & q_3 & \\ 0 & & & \ddots & \ddots \\ & & & & 1 \quad q_n \end{bmatrix} = A' - \lambda I =$$

$$1 \begin{bmatrix} q_2 & & & & 0 \\ 1 & q'_2 & & & \\ & 1 & q'_3 & & \\ & & \ddots & \ddots & \\ \bigcirc & & & 1 & q'_n \end{bmatrix} \begin{bmatrix} 1 & e'_1 & & & 0 \\ & 1 & e'_2 & & \\ & & \ddots & \ddots & \\ 0 & & & & e'_{n-1} \\ & & & & 1 \end{bmatrix}$$

The transformed matrix $A' - \lambda I$ has the same characteristic polynomial and may serve as well to calculate the New-ton correction. However, the solution of $(A' - \lambda I)x = \alpha_1 e_1$ is immediately given by

$$x = \begin{bmatrix} (-1)^{n+1} q_n \times q_{n-1} \times \ldots \times q_3 \times q_2 \\ (-1)^n q_n \times q_{n-1} \times \ldots \times q_3 \\ \\ \\ -q_n \\ 1 \end{bmatrix}, \quad \alpha_1 = \prod_{i=1}^{n} q_i \cdot (-1)^{n+1}$$

and likewise for the solution of $y^T (A' - \lambda I) = \alpha_2 e_n^T$ by

$y^T = (1, -q'_1, q'_1 \times q'_2, \ldots, (-1)^{n+1} q'_1 \times q'_2 \times$ [illegible] $, \quad \alpha_2 = \prod_{i=1}^{n} q'_i \cdot (-1)^{n+1}$

Note that $\alpha_1 = \alpha_2$ -det$(\lambda I - A)$ is the determinantal invariant of the qd-algorithm. Hence the relation

$$\frac{1}{\delta(\lambda)} = \frac{q_2 q_3 \cdots q_n}{q_1' q_2' \cdots q'n} + \frac{q_3 \cdots q_n}{q_2' \cdots q_n'} + \frac{q_n}{q_{n-1}' q_n'} + \frac{1}{q_n'}$$

or rather

$$\delta(\lambda) = q_n'/(\ldots((\frac{q_2}{q_1'} + 1) \frac{q_3}{q_2'} + 1) \ldots \frac{q_n}{q_{n-1}'} + 1) \quad .$$

The quotients q_{i+1}/q_i' in the nested product, however, are calculated as a matter of course in the LR-step with the quotient rule

$$e_i' := (q_{i+1}/q_i')e_i \quad .$$

The extra work amounts therefore to n-1 multiplications, n-1 additions of 1 , and one division.

The shift by' $\delta(\lambda)$ is now preferably made after the next inter+ mediate matrix A: $- \lambda I$ is formed and is done, as usual, implicitly in the difference rule. Thus, a shift is made every second qd-step. As mentioned in the introduction, numerical stability requires A in the beginning to be essentially symmetric and positive definite; i.e., $e_\mu > 0$ and $q_\mu > 0$. This property will then be preserved.

5.

The quantities q_μ and q_μ' can be calculated also by a continued fraction recurrence directly from

$$A' - \lambda I = \begin{vmatrix} a_1'-\lambda & b_1' & & & \\ 1 & a_2'-\lambda & b_2' & & 0 \\ & \ddots & \ddots & \ddots & \\ & & & & b_{n-1}' \\ 0 & & & 1 & a_n'-\lambda \end{vmatrix}$$

i.e., $q_n = a'_n - \lambda$, $q_i = a'_i - \lambda - b'_i/q_{i+1}$ $(i = n-1,\ldots,1)$

$q'_1 = a'_2 - \lambda$, $q'_i = a'_i - \lambda - b'_{i-1}/q'_{i-2}$ $(i = 2,3,\ldots,n)$.

For an essentially symmetric matrix; i.e., $b_\mu > 0$, the components of x and y^H together with $-\alpha$ form a Sturm sequence. Correspondingly, if all the $q_\mu \neq 0$ and $q'_\mu \neq 0$, the number of positive elements in the q-sequence counts the number of positive eigenvalues. This use of the continued fraction recurrence has some merits for the bisection method. The qd-transformation would not allow one to calculate the Sturm sequence in a stable way, apart from the trivial case where all $q_\mu > 0$.

Literature:

For the methods, concepts, and results used here, see

J.H. Wilkinson, The Algebraic Eigenvalue Problem. Oxford, 1965.

A. J. Householder, The Theory of Matrices in Numerical Analysis, New York, 1964.