Hyman's Method Applied To the General Eigenvalue Problem

By John Gary

This note is concerned with the problem of finding the roots of the determinantal equation $| A + \lambda B | = 0$. If B is nonsingular, then the problem is equivalent to finding the roots of the equation $| B^{-1}A + \lambda I | = 0$. The latter problem can be solved by reduction of $B^{-1}A$ to Hessenberg form followed by application of Hyman's method coupled with the Laguerre algorithm to locate the eigenvalues [1]. The purpose of this note is to show that the reduction to Hessenberg form can be carried out without first reducing the problem to the standard form $| A + \lambda I | = 0$. It is possible to define elementary row and column transformations with corresponding matrices P and Q such that $\hat{A} = PAQ$ is an upper Hessenberg matrix and $\hat{B} = PBQ$ is an upper triangular matrix. Since $| P | = | Q | = \pm 1$, the eigenvalue problem $| \hat{A} + \lambda \hat{B} | = 0$ is equivalent to the original problem. The eigenvalues of the transformed equation may be computed efficiently using the method devised by Parlett [1].

The referee for this paper called the author's attention to the fact that the general problem $| A + \lambda B | = 0$ can be reduced to the standard problem $| A + \lambda I | = 0$ even if B is singular. This apparently involves a determination of the rank of B and also a check of the linear independence of certain rows of a transformation of A. The method described herein requires almost twice the computing time that Hyman's method does when applied to the standard problem. Hyman's method is probably not as fast as the Q-R algorithm. Therefore the method described herein is probably inferior to one which reduces the problem to standard form. The reduction to standard form might introduce greater error than Hyman's method applied directly to $A + \lambda B$, but we have not made any comparisons.

We will now describe this reduction to Hessenberg form. Gaussian elimination with interchanges applied to the rows of B can be used to reduce B to upper triangular form. This enables us to define a matrix $P^{(1)}$ such that $B^{(1)} = P^{(1)}B$ is upper triangular (that is, $b_{ij}^{(1)} = 0$ if $j < i$). Also we have $| P^{(1)} | = \pm 1$ and $| p_{ij}^{(1)} | \leqq 1$. We use the notation $A^{(1)} = P^{(1)}A$ and denote the order of A and B by N.

Next we interchange the Nth and $(N - 1)$st rows of $A^{(1)}$ (if necessary) so that $| a_{N-1,1}^{(1)} | \geqq | a_{N,1}^{(1)} |$. Then we add a multiple of the $(N - 1)$st row to the Nth row so that the element $a_{N,1}^{(1)}$ is replaced by zero. This defines a matrix $P^{(2)}$. Note that the matrix $B^{(2)} = P^{(2)}B^{(1)}$ is no longer triangular. It may have a nonzero element $b_{N,N-1}^{(2)}$. By Gaussian elimination with interchanges on the last two columns of $B^{(2)}$ we may replace $b_{N,N-1}^{(2)}$ by zero. This defines a matrix $Q^{(3)}$ such that $| Q^{(3)} | = \pm 1$ and $B^{(2)}Q^{(3)}$ is triangular. Note that these column operations have no effect on the first column of $A^{(2)}$, thus $a_{N,1}^{(3)} = 0$, where $A^{(3)} = A^{(2)}Q^{(3)}$. We may continue until we have reduced the first column of A to Hessenberg form. The same method will reduce the remaining columns of A to Hessenberg form leaving B in triangular form.

The diagrams below indicate the reduction procedure after B is in triangular form. The lines indicate the rows and columns in which operations were performed.

$$A^{(1)}$$

X	X	X	X
X	X	X	X
X	X	X	X
X	X	X	X

$$B^{(1)}$$

X	X	X	X
0	X	X	X
0	0	X	X
0	0	0	X

$\rightarrow$

$$A^{(2)}$$

X	X	X	X
X	X	X	X
X	X	X	X
0	X	X	X

$$B^{(2)}$$

X	X	X	X
0	X	X	X
0	0	X	X
0	0	X	X

$\rightarrow$

$$A^{(3)}$$

X	X	X	X
X	X	X	X
X	X	X	X
0	X	X	X

$$B^{(3)}$$

X	X	X	X
0	X	X	X
0	0	X	X
0	0	0	X

$\rightarrow$

Thus we have reduced the equation $|\, A + \lambda B \,| = 0$ to $|\, \hat{A} + \lambda \hat{B} \,| = 0$, where $\hat{A}$ is in Hessenberg form and $\hat{B}$ is triangular. The reduction will require $13N^3/6 + O(N^2)$ multiplications and additions. This estimate ignores the work required to perform the interchanges.

The method of Parlett uses the Laguerre root finder to compute the eigenvalues. This requires evaluation of $P(\lambda)$, $dP(\lambda)/d\lambda$ and $d^2P(\lambda)/d\lambda^2$, where $P(\lambda) = |\, \hat{A} + \lambda \hat{B} \,|$. These are obtained by Hyman's method.

We define three vectors

$$\{x_i\}, \quad \{x_i'\}, \quad \{x_i''\}, \qquad 0 \leqq i \leqq N,$$

by the relations

$$x_N = 1, \qquad x_N' = 0, \qquad x_N'' = 0,$$

$$-\hat{a}_{i,i-1}x_{i-1} = \sum_{k=i}^{N} x_k \hat{a}_{i,k} + \lambda \sum_{k=i}^{N} x_k \hat{b}_{i,k},$$

$$-\hat{a}_{i,i-1}x_{i-1}' = \sum_{k=i}^{N} x_k' \hat{a}_{i,k} + \sum_{k=i}^{N} x_k \hat{b}_{i,k} + \lambda \sum_{k=i}^{N} x_k' \hat{b}_{i,k},$$

$$-\hat{a}_{i,i-1}x_{i-1}'' = \sum_{k=i}^{N} x_k'' \hat{a}_{i,k} + 2 \sum_{k=i}^{N} x_k' \hat{b}_{i,k} + \lambda \sum_{k=i}^{N} x_k'' \hat{b}_{i,k},$$

where $\hat{a}_{i,0} = -1$. Then we have $P(\lambda) = Cx_0$, $P'(\lambda) = Cx_0'$, and $P''(\lambda) = Cx_0''$, where

$$C = (-1)^{N-1} \prod_{i=1}^{N-1} A_{i+1,i}.$$

This requires $6N^2 + O(N)$ multiplications and additions. If $\hat{B}$ is diagonal, as it is in Parlett's case, then $3N^2 + O(N)$ operations are required.

The degree of the polynomial $P(\lambda)$ is a parameter in the Laguerre algorithm. If we overestimate the degree, the root finder will usually still converge, therefore we may assume the degree is N. We computed the roots in one test case with $N = 98$ and the degree of $P(\lambda)$ equal to 47. With the degree set equal to 98 rather than 47 the Laguerre algorithm required about 13% more iterations for convergence.

We could have used plane rotations instead of similarity transformations to compute $\hat{A}$ and $\hat{B}$. This might increase the stability of the reduction. In fact, the method of Wilkinson would doubtless yield a proof of stability in this case [2]. By stability we mean that the computed values of $\hat{A}$ and $\hat{B}$ should be close to those obtained from a computation without roundoff error. In the test case mentioned above we knew exact values for 23 of the 47 eigenvalues. The maximum error for the 23 computed roots was two in the tenth digit. The machine carries approximately 11 digits.

In general, the matrices $\hat{A}$ and $\hat{B}$ will be full Hessenberg and triangular matrices, even if the original matrices A and B were banded (that is $a_{ij} = b_{ij} = 0$ if $|i - j| > t$). If the band width t is sufficiently small it may be better to use Gaussian elimination on the original matrix $A + \lambda B$ rather than reduction to Hessenberg form.

An exercise in Householder's book shows how the polynomial eigenvalue problem $|\lambda^s I + \lambda^{s-1} P_1 + \cdots + \lambda P_{s-1} + P_s| = 0$ may be reduced to the standard eigenvalue problem $|A - \lambda I| = 0$ [3]. The same trick may be used to reduce the eigenvalue problem $|\lambda^s P_0 + \lambda^{s-1} P_1 + \cdots + P_s| = 0$ to $|A + \lambda B| = 0$. We define:

$$A = \begin{pmatrix} 0 & 0 & \cdots & & & -P_s \\ I & 0 & & & & -P_{s-1} \\ \cdots & & & & & \cdot \\ \cdots & & & & & \cdot \\ \cdots & & & & & \cdot \\ \cdots & & I & 0 & & -P_2 \\ \cdots & & 0 & I & & -P_2 \end{pmatrix}, \qquad B = \begin{pmatrix} -I & 0 & \cdots & & & \\ 0 & -I & \cdots & & & \\ \cdot & \cdot & & & & \\ \cdot & \cdot & & & & \\ \cdot & \cdot & & & & \\ & & & & -I & 0 \\ & & & & 0 & -P_0 \end{pmatrix}.$$

If the order of $P_0, \cdots, P_s$ is N, then the order of A and B is sN.

This method has been tested successfully on a few matrix problems which arose from differential eigenvalue problems [4]. In one case with $N = 98$ and the degree of $P(\lambda)$ equal to 47, approximately five minutes were required to compute the 47 eigenvalues. The program was written in FORTRAN and run on a CDC 3600.

National Center for Atmospheric Research
Boulder, Colorado

1. B. Parlett, "Laguerre's method applied to the matrix eigenvalue problem," *Math. Comp.*, v. 18, 1964, pp. 464–484.
2. J. Wilkinson, "Plane rotations in floating point arithmetic," Proc. Sympos. Appl. Math., Vol. 15, Amer. Math. Soc., Providence, R. I., 1963, pp. 185–198.
3. A. S. Householder, *The Theory of Matrices in Numerical Analysis*, Blaisdell, New York, 1964, p. 173.
4. J. Gary, "Computing eigenvalues of ordinary differential equations by finite differences," submitted to *Math. Comp.*