TECHNICAL NOTES AND SHORT PAPERS

The Reduction of an Arbitrary Real Square Matrix to Tri-Diagonal Form Using Similarity Transformations

By C. Donald La Budde

Abstract. In this paper a new algorithm for reducing an arbitrary real square matrix to tri-diagonal form using real similarity transformations is described. The method is essentially a generalization of a method due to A. S. Householder for accomplishing the same reduction in the case where the matrix is real and symmetric.

1. Introduction. In the calculation of the eigenvalues and eigenvectors of real symmetric matrices, several authors [1, 3, 7] have considered methods based upon an initial reduction of the matrix to tri-diagonal form using orthogonal transformations. (A tri-diagonal matrix is a matrix having zeroes everywhere except on the three principal diagonals.) Two of the advantages of this initial reduction are: the storage requirements decrease from $\frac{1}{2}n(n + 1)$ to $2n - 1$ locations for the matrix during the computation of the eigenvalues and eigenvectors of the reduced matrix; and the tri-diagonal form readily yields a Sturm sequence of polynomials terminating with the characteristic polynomial of the matrix. Then the eigenvalues may be calculated rapidly and accurately using an algorithm based upon this Sturm sequence.

This initial reduction to tri-diagonal form is so attractive that it seems desirable to devise algorithms that would apply in the case where the matrix is real and non-symmetric. One such algorithm [5] involves first reducing the matrix to Hessenberg form and thereafter completing the reduction using elementary similarity transformations. The difficulty with this method is that it requires division by elements of the matrix which may have become zero during the earlier part of the computation. Hence the method may not be applicable to a given matrix. Another algorithm [4] involves the construction of a biorthogonal set of vectors x_j, y_j, $j = 1, \cdots, n$ in that order given the initial vectors x_1, y_1 which may be chosen "almost arbitrarily." The difficulty here is that if $x_j y_j = 0$ and $x_j \neq 0$; $y_j \neq 0$, $2 \leqq j \leqq n$, the construction cannot be continued and the matrix cannot be reduced to tri-diagonal form. The improper choice of the vectors x_1, y_1 could cause this to happen. Then it would be necessary to begin all over again with different choices of x_1 and y_1.

The algorithm described in this paper is a generalization of one of the methods [3] used to reduce a real symmetric matrix to tri-diagonal form. The matrices employed for this reduction are no longer orthogonal matrices, but the computation remains in the real domain. Throughout this paper we shall be working with a fixed non-symmetric real n by n matrix A.

Received February 16, 1962, revised March 18, 1963. This work was sponsored in part by the National Aeronautics and Space Administration, Washington 25, D. C.

2. Definitions and Notation. We shall define three special types of transformation matrices as follows: (1) for $j \neq k$,

$$T_{jk}(a) = I_n + aE_{jk} \tag{2.1}$$

where I_n is the n by n identity matrix, $a \neq 0$, and E_{jk} is an n by n matrix with all of its entries zero except in the jth row, kth column where the entry is 1; (2) for $j \neq k$,

$$P_{jk} = I_n \tag{2.2}$$

except for the elements in the jth row jth column, jth row kth column, kth row jth column, kth row kth column which instead are 0, 1, 1, 0 respectively; and (3) matrices of the type

$$V_j(T) = \left(\begin{array}{c|c} I_j & 0 \\ \hline 0 & T \end{array}\right) \begin{array}{l} j \text{ rows} \\ (n-j) \text{ rows} \end{array} \tag{2.3}$$

$$j \text{ columns} \quad (n-j) \text{ columns}$$

which will subsequently be referred to as matrices of type j. It is then obvious that the following equations hold:

$$\begin{aligned} T_{jk}(a)T_{jk}(-a) &= I_n \\ P_{jk}^2 &= I_n \\ V_j(T)V_j(T^{-1}) &= I_n \end{aligned} \tag{2.4}$$

the latter equation being true if T is a non-singular $n - j$ by $n - j$ matrix. Associated with an arbitrary n by n real matrix A with entries a_{jk} are the quantities S_j ($j = 1, \cdots, n - 2$) defined as follows:

$$S_j = \sum_{k=j+1}^{n} a_{jk}a_{kj} . \tag{2.5}$$

We now state two theorems which we will use in the subsequent development of the algorithm for reducing an arbitrary real square matrix to tri-diagonal form by similarity transformations.

THEOREM 1. *Let a matrix A be of the form*

$$A = \left(\begin{array}{cc|c} \multicolumn{2}{c|}{A_j} & \begin{array}{c} 0 \\ \hline B_j^t \end{array} \\ \hline 0 & C_j & D_j \end{array}\right) \tag{2.6}$$

where A_j is a j by j tri-diagonal matrix, B_j^t is an $n - j$ dimensional row vector, C_j is an $n - j$ dimensional column vector, and D_j is an $n - j$ by $n - j$ matrix. If $V_j(T)$ is a matrix of type j with T an $n - j$ by $n - j$ non-singular matrix, then $A' = V_j(T)AV_j(T^{-1})$ will have the same form as A with

$$A' = \left(\begin{array}{cc|c} \multicolumn{2}{c|}{A_j'} & \begin{array}{c} 0 \\ \hline B_j'^t \end{array} \\ \hline 0 & C_j' & D_j' \end{array}\right) \tag{2.7}$$

where A_j' is a j by j tri-diagonal matrix, $B_j'^t$ is an $n - j$ dimensional row vector, C_j' is an $n - j$ dimensional column vector, and D_j' is an $n - j$ by $n - j$ matrix.

Proof. If the definition of $V_j(T)$ is used, it is trivial to show that $A_j' = A_j$, $D_j' = TD_jT^{-1}$, $B_j'^T = B_j^tT^{-1}$, $C_j' = TC_j$, and the zeroes in the upper right and lower left corners of A are preserved under the transformation

$$A' = V_j(T)AV_j(T^{-1}).$$

THEOREM 2. *Let A, A', and $V_j(T)$ be the matrices of theorem 1, and let S_j' be the quantity associated with A' defined in equation (2.5). Then $S_j' = S_j$.*

Proof. Let the entries of A' be denoted by a'_{pq}. Then

$$S_j' = \sum_{k=j+1}^{n} a'_{jk}a'_{kj} = B_j'^TC_j' = B_j^tT^{-1}TC_j = B_j^tC_j \tag{2.8}$$

$$= \sum_{k=j+1}^{n} a_{jk}a_{kj} = S_j.$$

3. Description of the Algorithm. The algorithm is comprised of $n - 2$ steps, the generic jth step consisting of at most two similarity transformations on the matrix A using matrices of type j. (The situation at step 1 may be handled in a slightly different way; we will discuss this situation a little later.) If, at the beginning of the jth step, A has the form assumed in theorem 1, then, at the end of the jth step, A' will also have the same form (theorem 1). If, for each j, the similarity transformations can be so chosen that

$$a'_{jk} = a'_{kj} = 0 \qquad (k = j + 2, \cdots, n) \tag{3.1}$$

then at the end of step $n - 2$, the matrix A will have been reduced to tri-diagonal form by similarity transformations.

We define special matrices of type j as follows: Let a, b be arbitrary non-zero numbers and let x, y be $n - j$ dimensional column vectors with entries x_p, y_p respectively ($p = j + 1, \cdots, n$). Then the special type j matrix $V_j(I_{n-j} + aXY^t)$ has the property that

$$V_j(I_{n-j} + aXY^t)V_j(I_{n-j} + bXY^t) = I_n \tag{3.2}$$

if

$$Y^tX = -(a + b)/ab. \tag{3.3}$$

Setting

$$A' = V_j(I_{n-j} + aXY^t)AV_j(I_{n-j} + bXY^t), \tag{3.4}$$

we see that equations (3.1) become

$$a_{jk} + by_k\sum_{p=j+1}^{n} x_pa_{jp} = a_{kj} + ax_k\sum_{p=j+1}^{n} y_pa_{pj} = 0 \qquad (k = j + 2, \cdots, n) \tag{3.5}$$

Equation (3.3) becomes

$$\sum_{k=j+1}^{n} x_ky_k = -(a + b)/ab. \tag{3.6}$$

We define two quantities c, d as follows:

$$c^{-1} = \sum_{k=j+1}^{n} x_ka_{jk}; \qquad d^{-1} = \sum_{k=j+1}^{n} y_ka_{kj}. \tag{3.7}$$

Solving these equations for c, d, x_k, y_k $(k = j + 1, \cdots, n)$ in that order we obtain

$$(3.8)\qquad \frac{1}{cd} = \frac{-(a+b)S_j \pm \sqrt{S_j^2(a+b)^2 + 4abS_j(a_{j,j+1}\, a_{j+1,j} - S_j)}}{2ab}$$

$$(3.9)\qquad x_{j+1} = \left[\frac{1}{cd} + \frac{S_j - a_{j,j+1}\, a_{j+1,j}}{a}\right] \cdot \frac{d}{a_{j,j+1}}$$

$$(3.10)\qquad y_{j+1} = \left[\frac{1}{cd} + \frac{S_j - a_{j,j+1}\, a_{j+1,j}}{b}\right] \cdot \frac{c}{a_{j+1,j}}$$

$$(3.11)\qquad x_k = -(d/a)a_{kj},\ y_k = -(c/b)a_{jk}, \qquad (k = j+2, \cdots, n).$$

We first observe that one of the variables c, d is arbitrary so for convenience we set $d = 1$. Then, to solve equations (3.8)–(3.11) for the x's and y's in the real domain it is sufficient to require that $S_j \neq 0$, $a_{j,j+1} \neq 0$, and $a_{j+1,j} \neq 0$. Since a, b are arbitrary non-zero numbers, their signs may be so chosen as to make the quantity under the radical in equation (3.8) positive. The sign of the radical may then be chosen so that c is finite. If these requirements are satisfied (and all the proper choices of signs are made) and X, Y are $n - j$ dimensional vectors defined by equations (3.8)–(3.11), then the similarity transformation of equation (3.4) will yield a matrix A' satisfying the conditions of equations (3.1).

We now consider in detail the computations in the jth step. We *assume* that in the beginning of the jth step one of the requirements, namely $S_j \neq 0$ is satisfied. This implies that at least one term $a_{pj}a_{jp} \neq 0$, for $j + 1 \leqq p \leqq n$. Then a similarity transformation on A using the matrix $P_{j+1,p}$ will interchange $a_{j+1,j}$ and a_{pj}, and $a_{j,j+1}$ and a_{jp}. Since $P_{j+1,j}$ is a matrix of type j, the matrix A after this similarity transformation will still have the property that $S_j \neq 0$ (theorem 2). Then, with all the proper choices of signs, the vectors X, Y may be determined, and the transformation of equation (3.4) on A may be carried out yielding a matrix A' satisfying equations (3.1).

In order to continue the algorithm into step $j + 1$, we must be certain that $S'_{j+1} \neq 0$. (In general, a similarity transformation on A using a matrix of type j will alter S_{j+1}.) Now a, b are arbitrary (except for sign) so we may theoretically choose $|a|$, $|b|$ so that $S'_{j+1} \neq 0$. At present, there doesn't seem to be any practical way of doing this. $|a|$, $|b|$ could be determined by trial and error starting, for example, with $a = b = -2$ (same sign required) or $a = +1$, $b = -1$ (opposite signs required). From these values of a, b trial vectors X, Y and hence S'_{j+1} could be computed before the similarity transformation of equation (3.4) is carried out. If $S'_{j+1} = 0$, the vectors X, Y could be effectively changed by either scaling or incrementing $|a|$, $|b|$ by some fixed positive constant. By scaling or incrementing as often as necessary, vectors X, Y would be obtained so that $S'_{j+1} \neq 0$. At step $n - 2$ it would not be necessary to alter a and b.

It only remains to discuss some of the special features of step 1. Since no step preceded it, it might happen that $S_1 = 0$. If $S_1 = 0$ and $a_{21} = 0$, then some off-diagonal element $a_{pq} \neq 0$ if the matrix A is not in diagonal form. Then a_{21}, a_{pq} can be interchanged by at most two similarity transformations on A using matrices of the kind P_{rs}. If S_1 is still zero, we may follow by a similarity transformation on A

using a matrix of the kind $T_{12}(a)$. Then the elements a_{1k} $(k = 2, \cdots, n)$ are altered as follows:

$$a_{12} \to (a_{12} + aa_{22}) - a(a_{11} + aa_{21}) \tag{3.12}$$

$$a_{1k} \to (a_{1k} + aa_{2k}) \qquad (k = 3, \cdots, n). \tag{3.13}$$

The elements a_{k1} $(k = 2, \cdots, n)$ remain fixed. It is easily seen that by making $|a|$ large enough, we can make $S_1 \neq 0$ and $((a_{12} + aa_{22}) - a(a_{11} + aa_{21})) \neq 0$ since $a_{21} \neq 0$. Then step 1 may be continued as described before.

In conclusion, we may point out some of the features (good and bad) of this algorithm. Each of the $n - 2$ steps (except possibly for the first) requires only two similarity transformations, the first involving no computations but only row and column permutations, and the second requiring only one square root. Another feature is the fact that at the end of step j we must have $S_j = a'_{j,j+1}a'_{j+1,j}$ (theorem 2 and equations (3.1)). This means that all of the elements $a'_{j,j+1}$, $a'_{j+1,j}$ will be non-zero and the characteristic polynomial of the matrix A will not factor under this algorithm. The most annoying feature of the algorithm is the determination of a, b by trial and error in each step to insure that the algorithm may be continued.

4. Acknowledgment. The author wishes to express his thanks to Dr. W. Givens and Dr. R. C. Sahni for helpful discussions in connection with this work.

Argonne National Laboratory and
Department of Chemistry
New York University

1. W. Givens, "Numerical computation of the characteristic values of a real symmetric matrix," Oak Ridge National Laboratory Report ORNL 1574, 1954.

2. W. Givens, "Computation of plane unitary rotations transforming a general matrix to triangular form," *J. Soc. Indust. Appl. Math.*, v. 6, 1958, p. 26–50.

3. A. S. Householder & F. L. Bauer, "On certain methods for expanding the characteristic polynomial," *Numer. Math.*, v. 1, 1959, p. 29–37.

4. C. Lanczos, "An iterative method for the solution of the eigenvalue problem of linear differential and integral operators," *J. Res. Nat. Bur. Standards*, v. 45, 1950, p. 255–282.

5. H. R. Schwarz, "Critere de stabilite pour des systemes d'equations a coefficients constants complexes," *C. R. Acad. Sci. Paris*, v. 242, 1956, p. 325–327.

6. J. H. Wilkinson, "Stability of the reduction of a matrix to almost triangular and triangular forms by elementary similarity transformations," *J. Assoc. Comput. Mach.*, v. 6, 1959, p. 336–359.

7. J. H. Wilkinson, "Householder's method for the solution of the algebraic eigenvalue problem," *Comput. J.*, v. 3, 1960, p. 23–27.