Reducing a Matrix to Hessenberg Form

By P. A. Businger

Abstract. It has been an open problem whether the reduction of a matrix to Hessenberg (almost triangular) form by Gaussian similarity transformations is numerically stable [2, p. 364]. We settle this question by exhibiting a class of matrices for which this process is unstable.

As a major step towards the numerical solution of the non-Hermitian algebraic eigenvalue problem, a matrix is usually first reduced to Hessenberg (almost triangular) form either by a sequence of Householder similarity transformations, [2, p. 347] or else by some form of Gaussian elimination [2, p. 353]. In practice, the latter process is favored because it requires fewer arithmetic operations. However, it is known that Householder's reduction is numerically stable [2, p. 350], while it has been an open problem whether Gaussian elimination is stable [2, p. 364]. We settle this question by exhibiting a class of matrices for which Gaussian elimination is unstable.

The column-by-column reduction of an n by n matrix $A = (a_{ij})$ to upper Hessenberg form by Gaussian elimination produces a sequence of similarity transforms

$$A_k = (G_k^{-1}P_k^{-1})A_{k-1}(P_kG_k)\ , \quad k = 1, 2, \cdots, n-2\ ,$$

of $A_0 = A$. $G_k^{-1} = (\bar{g}_{ij}^{(k)})$ differs from the identity matrix only in the elements $\bar{g}_{i,k+1}^{(k)}$, $i = k+2, k+3, \cdots, n$, which are chosen such that A_k has zeros in positions $(k+2, k)$, $(k+3, k)$, $\cdots$, (n, k). The permutation matrices P_k are chosen such that

$$|\bar{g}_{i,k+1}^{(k)}| \leqq 1, \qquad (i = k+2, k+3, \cdots, n).$$

Wilkinson [2, p. 364] points to the danger of potential worsening of the condition of the eigenvalues caused by large elements in the matrices

$$F_k^{-1} = (\bar{f}_{ij}^{(k)}) = (G_k^{-1}P_k^{-1})(G_{k-1}^{-1}P_{k-1}^{-1}) \cdots (G_1^{-1}P_1^{-1})\ .$$

It is implicit in [2] that the largest element of F_k^{-1} is bounded in magnitude by 2^{k-1}; the theorem below shows this bound to be sharp and the example illustrates the consequences of this fact. The proof of the theorem uses a construction analogous to that of [2, p. 212].

Theorem. *There exist matrices for which*

$$\max_{i,j} |\bar{f}_{ij}^{(k)}| = 2^{k-1}\ .$$

Proof. Let

Received April 25, 1969.

$$A = \begin{bmatrix} a_1 & a_2 & \cdot & \cdot & \cdot & a_{n-1} & a_n \\ 1 & & & & & & \\ -1 & 1 & & 0 & & & \\ \cdot & & \cdot & & & & \\ \cdot & 0 & & \cdot & & & \\ \cdot & & & & \cdot & & \\ -1 & & & & & 1 & \end{bmatrix}$$

where $a_1, a_2, \cdots, a_n$ are arbitrary. Then, for $P_k = I$, $k = 1, 2, \cdots, n-2$, $\bar{g}_{k+2,k+1}^{(k)} = \bar{g}_{k+3,k+1}^{(k)} = \cdots = \bar{g}_{n,k+1}^{(k)} = 1$ and $\max |\bar{f}_{ij}^{(k)}| = 2^{k-1}$.

While the following example does not quite achieve the bound 2^{k-1}, it illustrates the effect of large elements of F_k^{-1}.

Example. Consider the 6 by 6 matrix

$$B = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & -1 \\ -1 & 1 & 0 & 0 & 0 & -1 \\ -1 & 0 & 1 & 0 & 0 & -1 \\ -1 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$

whose eigenvalues $\lambda_i(B)$ are given to five figures in the table. (The calculations were performed on a GE645 computer by programs given in [1].) We also give approximate values of the condition numbers

$$s_i(B) = \frac{\|y_i^H\| \, \|x_i\|}{|y_i^H x_i|},$$

where y_i^H and x_i are row- and column-eigenvectors of B. Since no s_i is large, all eigenvalues of B are insensitive to small perturbations in the elements of B. After column-reducing B to upper Hessenberg form we obtain

$$H = B_4 = \begin{bmatrix} 0 & -2 & -1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 0 & 2 & -2 \\ 0 & 0 & 1 & 0 & 4 & -4 \\ 0 & 0 & 0 & 1 & 8 & -8 \\ 0 & 0 & 0 & 0 & 8\frac{1}{2} & -8 \end{bmatrix}.$$

Since H is similar to B, its eigenvalues $\lambda_i(H)$ agree with those of B. However, the transformed condition numbers

$$s_i(H) = \frac{\|y_i^H F_4\| \, \|F_4^{-1} x_i\|}{|y_i^H x_i|}$$

indicate considerably greater sensitivity of the eigenvalues to perturbations in the elements of H.

TABLE. *Numerical Results for the Example.*

$\lambda_i(B) = \lambda_i(H)$	$s_i(B)$	$s_i(H)$	$\lambda_i(H')$
1.0000	1.2	7.1	2.2725
−1.1869	1.4	3.8	−1.8652
.47473 ± 1.4373i	1.9	2.6	.31126 ± 1.4433i
− .38127 ± 1.2286i	1.7	4.0	− .51492 ± .77502i

Suppose now that the reduction to Hessenberg form was carried out in truncated 4-bit arithmetic. (This example can be generalized to n dimensions and $(n - 2)$-bit arithmetic.) The reduced matrix becomes

$$H' = \begin{bmatrix} 0 & -2 & -1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 0 & 2 & -2 \\ 0 & 0 & 1 & 0 & 4 & -4 \\ 0 & 0 & 0 & 1 & 8 & -8 \\ 0 & 0 & 0 & 0 & 8 & -8 \end{bmatrix}$$

whose (6, 5)-element differs slightly from the corresponding element of H. Due to sensitivity to this difference, the eigenvalues $\lambda_i(H')$ show little resemblance to those of B.

We have shown that there exist matrices which cannot be stably reduced to Hessenberg form by means of Gaussian elimination in finite precision arithmetic. Householder transformations, however, provide unconditional stability.

Bell Telephone Laboratories, Incorporated
Murray Hill, New Jersey 07974

1. J. M. VARAH, *The Computation of Bounds for the Invariant Subspaces of a General Matrix Operator*, Tech. Report No. CS66, Stanford University, Stanford, Calif., 1967.

2. J. H. WILKINSON, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965. MR **32** #1894.