values of the error shown in the tables correspond to these (approximately) optimal values of μ and γ. These computations were performed on the CDC 6600 at the National Center for Atmospheric Research.

National Center for Atmospheric Research
Boulder, Colorado

1. J. Douglas, Jr. & C. M. Pearcy, "On convergence of alternating direction procedures in the presence of singular operators," *Numer. Math.*, v. 5, 1963, pp. 175–184. MR **27** #4384.
2. H. B. Keller, "On some iterative methods for solivng elliptic difference equations," *Quart. Appl. Math.*, v. 16, 1958, pp. 209–226. MR **22** #8667.
3. H. B. Keller, "On the solution of singular and semidefinite linear systems by iteration," *SIAM J. Numer. Anal. Series* B, v. 2, 1965, pp. 281–290.
4. R. B. Kellogg & J. Spanier, "On optimal alternating direction parameters for singular matrices," *Math. Comp.*, v. 19, 1965, pp. 448–452. MR **32** #1914.
5. R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, N. J., 1962, p. 226. MR **28** #1725.
6. J. H. Giese, "On the truncation error in a numerical solution of the Neumann problem for a rectangle," *J. Math. Phys.*, v. 37, 1958, pp. 169–177. MR **20** #2845.

Canonical Decomposition of Hessenberg Matrices†

By Beresford Parlett

1. Introduction. A square matrix A is said to be in (upper) Hessenberg form if $a_{ij} = 0$ for $i > j + 1$. Such matrices occur frequently in connection with the eigenvalue problem. In practical work it is an important fact that any square matrix may be transformed in a stable manner into a similar Hessenberg matrix, see [5]. Apart from possible economies in computing the eigenvalues we may ask whether a preliminary reduction of a full matrix to this form offers any other advantages.

We show here that this reduction replaces an arbitrary independent set of eigenvectors by one which has some useful theoretical properties. In other words if J is the (lower) Jordan canonical form of A, say

$$A = Y^{-1}JY, \tag{1.1}$$

then the rows of Y are the row eigenvectors of A. When A is defective we must interpret eigenvectors in the generalized sense (as principal vectors). For general A we can say nothing about Y other than det $(Y) \neq 0$. If A is a Hessenberg matrix then Y has the properties summarized in Theorem 1.

We should remark here that our results are fairly straightforward deductions from Lemma 1 which is well known, but not in the form used here. The purpose of this note is just to extract the properties which are latent in that lemma: essentially the triangular factorization of Vandermonde matrices.

As we show in [7] the existence of this factorization helps explain the remarkable convergence properties of the QR algorithm of J. G. F. Francis [1]. The result is also useful in discussing other problems involving Hessenberg matrices. We note that "the" Jordan form is unique only to within the order of the submatrices of which it is a direct sum. The factors in our decomposition depend on this order and here we

Received May 5, 1966. Revised July 6, 1966.

† This work was begun while the author was a summer visitor at the Argonne National Laboratory, Illinois. It was completed under contract NONR 3656(23) with the University of California.

prefer to speak of a Jordan form $J = J_1 \oplus \cdots \oplus J_r$ where each J_i is a simple Jordan submatrix and $\oplus$ indicates direct sum.

THEOREM 1. *Let J be any lower Jordan form of the upper Hessenberg matrix H. Then there exists a nonsingular Y such that*

(i) $H = Y^{-1}JY$,

(ii) *Y permits the triangular decomposition $L_Y U_Y$, L_Y unit lower triangular, U_Y upper triangular,*

(iii) *if L_Y is partitioned conformably with J as $L_Y = (L^{ij})$ then L^{ii} commutes with J_i, $i = 1, \cdots, r$.*

In Lemma 1 we prove that for $k = 1, \cdots, n$ certain $k \times k$ minors in the first k columns of the $n \times n$ matrix L_Y cannot vanish. In Lemma 4 we obtain explicit expressions for the elements of L_Y in terms of polynomials whose zeros are the eigenvalues of H. Thus L_Y depends only on the spectrum of H. The effect of the eigenvectors is concentrated in U_Y.

2. Previous Results. We should point out here that there is no loss of generality in restricting attention to Hessenberg matrices with nonzero subdiagonal elements. Any Hessenberg matrix is block triangular, the diagonal blocks being of the above type. If each block permits an LU factorization then the L-factor of the whole matrix is just their direct sum.

Definition. UHM = UHM$_n$ = $\{H : H \text{ Hessenberg}, h_{i+1,i} \neq 0, i = 1, \cdots, n-1\}$ is the set of unreduced $n \times n$ Hessenberg matrices. Such matrices may be reducible.

LEMMA 2 (CLASSICAL). *Let J be a lower Jordan form and F the upper Frobenius form of $H \in$ UHM. Then*

(i) $H = R^{-1}FR$, *R upper triangular,*

(ii) $F = V^{-1}JV$, *V the (confluent) Vandermonde matrix determined by J.*

The results are implicit in [3, Chapter 6] and [5, Chapter 1] and we shall give a few observations instead of a proof.

Since H is nonderogatory, F is actually the companion matrix of the characteristic polynomial. The coefficients are in the last column. The triangle R may be associated, for example, with the Danilevski method. Note that as R ranges over all nonsingular upper triangles, $R^{-1}FR$ generates the equivalence class of members of UHM similar to F.

Conclusion (ii) relates the two forms. Note that if T is any matrix which commutes with J then $F = (TV)^{-1}J(TV)$ also relates the two forms and exhibits the basic freedom in choice of principal vectors. In general, TV is not a Vandermonde matrix.

The relation of V_J to J is best illustrated by an example.

$$J = \begin{pmatrix} \lambda_1 & & & & & \\ 1 & \lambda_1 & & & & \\ & 1 & \lambda_1 & & & \\ & & & \lambda_2 & & \\ & & & & \lambda_3 & \\ & & & & 1 & \lambda_3 \end{pmatrix}, \qquad V_J = \begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 & \lambda_1^3 & \lambda_1^4 & \lambda_1^5 \\ 0 & 1 & 2\lambda_1 & 3\lambda_1^2 & 4\lambda_1^3 & 5\lambda_1^4 \\ 0 & 0 & 1 & 3\lambda_1 & 6\lambda_1^2 & 10\lambda_1 \\ 1 & \lambda_2 & \lambda_2^2 & \lambda_2^3 & \lambda_2^4 & \lambda_2^5 \\ 1 & \lambda_3 & \lambda_3^2 & \lambda_3^3 & \lambda_3^4 & \lambda_3^5 \\ 0 & 1 & 2\lambda_3 & 3\lambda_3^2 & 4\lambda_3^3 & 5\lambda_3^4 \end{pmatrix}.$$

The fact that members of UHM are nonderogatory will be used heavily in what follows. It can be seen by noting that the nullity of $H - zI$ cannot exceed 1.

3. Decomposition of V. We consider $H \in$ UHM, its Jordan form

$$J = J_1 \oplus \cdots \oplus J_r \tag{3.1}$$

and the associated Vandermonde matrix V. Here $J_i = \mathrm{J}(\lambda_i)$ is the simple Jordan submatrix for the eigenvalue λ_i. Its order is the (algebraic) multiplicity of λ_i, say m_i. Since H is nonderogatory, different J_i correspond to distinct λ_i. A classical result, apparently first published in [6], is that

$$\det(V) = \prod_{\alpha,\beta=1;\alpha<\beta}^{r} (\lambda_\beta - \lambda_\alpha)^{m_\beta m_\alpha} \neq 0. \tag{3.2}$$

Now any leading principal submatrix of V is the Vandermonde matrix associated with the corresponding leading submatrix of J. Consequently, the leading principal minors of V are all different from zero. These are the necessary and sufficient conditions for the triangular decomposition, see [2] or [3]. Thus

$$V = L_V U_V, \qquad \operatorname{diag}(L_V) = I. \tag{3.3}$$

Since R is upper triangular we set $Y = VR$ and by the uniqueness of the decomposition $L_Y = L_V$, $U_Y = U_V R$. This establishes (ii) of Theorem 1.

It is a useful fact that L_V depends only on the eigenvalues of H and not on the eigenvectors. The next lemma gives information about L_V which plays an important role in [7].

Corresponding to the decomposition (3.1) we have a natural partitioning of V by rows as

$$V = \begin{pmatrix} V_1 \\ \vdots \\ V_r \end{pmatrix}, \qquad V_i \text{ is } m_i \times n, \qquad i = 1, \cdots, r. \tag{3.4}$$

Partition $L = L_V$ in the same way as V. We already know that the leading principal minors of L do not vanish (all are 1) and now we shall show that certain other minors of L cannot vanish either. First we want a notation for these submatrices.

There are m_i rows in L_i. For each $i = 1, \cdots, r$ let $0 \leqq \mu_i \leqq m_i$ and choose the *top* μ_i rows of L_i. This submatrix is determined by the vector $\mu = (\mu_1, \cdots, \mu_r)$. It has $\sum_{i=1}^{n} \mu_i = |\mu|$ rows and $n = \sum_{i=1}^{r} m_i = |m|$ columns. We denote by $L_{k\mu}$ the submatrix obtained by taking only the first k columns, $k \leqq n$. Note that $L = L^m_{|m|}$.

LEMMA 2. *Let V be the Vandermonde matrix associated with the nonderogatory Jordan form $J = J_1 \oplus \cdots \oplus J_r$. Let $V = V^m_{|m|} = LU$, $m = (m_1, \cdots, m_r)$, m_i the order of J_i. Then for $0 \leqq \mu_i \leqq m_i$, $i = 1, \cdots, r$ the submatrices $L^\mu_{|\mu|}$ are nonsingular.*

Proof. In general $V^\mu_{|\mu|} = L^\mu_{|m|} U^m_{|\mu|}$. Since we may write

$$U^m_{|\mu|} = \begin{pmatrix} \bar{U} \\ 0 \end{pmatrix},$$

where $\bar{U}$ is $|\mu| \times |\mu|$ and upper triangular, we have

$$\det[V^\mu_{|\mu|}] = \det[L^\mu_{|\mu|}] \det[\bar{U}]. \tag{3.5}$$

Now $V^{\mu}_{|\mu|}$ is the Vandermonde matrix determined by the λ_i with multiplicity μ_i, $i = 1, \cdots, r$. As such it is nonsingular. Since $\bar{U}$ is the leading principal $|\mu| \times |\mu|$ submatrix of U it is nonsingular and the lemma follows.

4. The Elements of L_V. It is natural to partition L_V conformably with J as $L_V = (L^{ij})$, $i \geqq j$, where L^{ij} is $m_i \times m_j$. We shall exhibit a typical element $l^{ij}_{\alpha\beta}$. We denote by V_δ the leading principal δ-rowed submatrix of V and by $V_{\gamma,\delta}$ the matrix obtained from V_δ by replacing row δ by the corresponding elements of row γ of V. If $L_V = (l_{\gamma\delta})$ then by [3, Chapter 1]

$$l_{\gamma\delta} = \det [V_{\gamma,\delta}]/\det [V_\delta], \qquad \gamma \geqq \delta. \tag{4.1}$$

We shall express $l^{ij}_{\alpha\beta}$ in terms of the polynomials $p_1(z) = 1$, $p_k(z) = \prod_{\nu=1}^{k-1} (z - \lambda_\nu)^{m_\nu}$, $k = 2, \cdots, r$. Let $l_{\gamma\delta} = l^{ij}_{\alpha\beta}$; thus $\gamma = \sum_{\nu=1}^{i-1} m_\nu + \alpha$, $\delta = \sum_{\nu=1}^{j-1} m_\nu + \beta$, $\alpha \leqq m_i$, $\beta \leqq m_j$.

By (3.2)

$$\det [V_\delta] = \prod_{\mu,\nu=1;\mu<\nu}^{j-1} (\lambda_\nu - \lambda_\mu)^{m_\nu m_\mu} \prod_{\nu=1}^{j-1} (\lambda_j - \lambda_\nu)^{\beta m_\nu} = \Gamma(p_j(\lambda_j))^\beta, \tag{4.2}$$

since β is the multiplicity of λ_j in V_δ.

$$\Gamma = \prod_{\mu,\nu=1;\mu<\nu}^{j-1} (\lambda_\nu - \lambda_\mu)^{m_\nu m_\mu} = \det [V_{\delta-\beta}].$$

To find $\det [V_{\gamma,\delta}]$ we consider first the case when $\alpha = 1$, $i > j$. The last row of V_δ is $(1/(\beta - 1)!)(d/d\lambda_j)^{\beta-1} r(\lambda_j)$ where $r(\lambda) = (1, \lambda, \lambda^2, \cdots, \lambda^{\delta-1})$. It is replaced by $r(\lambda_i)$ in $V_{\gamma,\delta}$. Again (3.2) yields, for $\alpha = 1$,

$$\det [V_{\gamma,\delta}] = \Gamma(p_j(\lambda_j))^{\beta-1} p_j(\lambda_i)(\lambda_i - \lambda_j)^{\beta-1}. \tag{4.3}$$

By differentiating (4.3) with respect to λ_i we obtain for $1 \leqq \alpha \leqq m_i$

$$\det [V_{\gamma,\delta}] = (\Gamma/(\alpha - 1)!)(p_j(\lambda_j))^{\beta-1}(d/d\lambda_i)^{\alpha-1}[p_j(\lambda_i)(\lambda_i - \lambda_j)^{\beta-1}]. \tag{4.4}$$

There remains the case when $i = j$. The result does not seem to have appeared in literature (at least not in Sir Thomas Muir's history) and so we give it here as

LEMMA 3. *Let* $n_i = \sum_{\nu=1}^{i-1} m_\nu$ *and* $1 \leqq \beta \leqq \alpha \leqq m_i$, $i \leqq r$. *Then*

$$\det [V_{n_i+\alpha, n_i+\beta}] = \Gamma(p_i(\lambda_i))^{\beta-1} p_i^{(\alpha-\beta)}(\lambda_i)/(\alpha - \beta)! \tag{4.5}$$

The result follows by induction on α and uses Leibnitz' rule. The proof is left as an exercise.

Substituting (4.2)–(4.5) into (4.1) gives us

LEMMA 4. *With the notation developed above the elements of* L_V *are given by*

$$\begin{aligned} l^{ij}_{\alpha\beta} &= p_i^{(\alpha-\beta)}(\lambda_i)/(\alpha - \beta)!\, p_i(\lambda_i); \qquad i = j, \qquad \alpha \geqq \beta, \\ &= \left(\frac{d}{d\lambda_i}\right)^{\alpha-1} [p_j(\lambda_i)(\lambda_i - \lambda_j)^{\beta-1}]/(\alpha - 1)!\, p_j(\lambda_j), \qquad i > j. \end{aligned} \tag{4.6}$$

COROLLARY. *If H has linear elementary divisors then*

$$l_{\gamma\delta} = p_\delta(\lambda_\gamma)/p_\delta(\lambda_\delta), \qquad \gamma \geqq \delta. \tag{4.7}$$

We are now in a position to establish (iii) of Theorem 1. We observe that L^{ii} is of order m_i, depends only on λ_i and, by (4.6), its (α, β) element is a function of $\alpha - \beta$. If we define $N_i = (e_2, \cdots, e_{m_i}, 0)$, where $I = (e_1, \cdots, e_{m_i})$, then

$$L^{ii} = \sum_{\nu=0}^{m_i-1} \frac{p_i^{(\nu)}(\lambda_i)}{\nu!\, p_i(\lambda_i)} N_i^{\nu}, \tag{4.8}$$

and is a polynomial in N_i. Since J_i is also a polynomial in N_i it must commute with L^{ii}.

The above results were derived for $H \in$ UHM. However, properties (ii) and (iii) generalize immediately to all Hessenberg matrices by the remarks at the beginning of Section 2.

University of California
Berkeley, California

1. J. G. F. Francis, "The QR transformation. I, II," *Comput. J.*, v. 4, 1961/1962, pp. 265–271, 332–345.
2. V. N. Faddeeva, *Computational Methods of Linear Algebra*, Dover, New York, 1959, p. 20.
3. A. S. Householder, *The Theory of Matrices in Numerical Analysis*, Blaisdell, New York, 1964. MR **30** #5475.
4. M. Marcus & H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, Allyn and Bacon, Boston, Mass., 1964. MR **29** #112.
5. J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965. MR **32** #1894.
6. V. Zeipel, "On Determinates, hvars elementer aro Binomialkoefficenter," *Lunds Univ. Asskr.* II, 1865, pp. 1–68.
7. B. Parlett, "Convergence theory for the QR algorithm on a Hessenberg matrix," *Math. Comp.* (To appear.)

An Elimination Method for Computing the Generalized Inverse*

By Leopold B. Willner

0. Notations. We denote by

A an $m \times n$ complex matrix,
A^* the conjugate transpose of A,
$A_j, j = 1, \cdots, n$ the jth column of A,
A^+ the generalized inverse of A [7],
H the Hermite normal form of A, [6, pp. 34–36],
Q^{-1} the nonsingular matrix satisfying

$$H = Q^{-1}A, \tag{1}$$

$e_i, i = 1, \cdots m$ the ith unit vector $e_i = (\delta_{ij})$,
r the rank of A ($=$ rank H).

1. Method. The Hermite normal form of A is written as

$$H = \begin{bmatrix} B \\ 0 \end{bmatrix} \quad \text{where } B \text{ is } r \times n. \tag{2}$$

Received July 13, 1966.
* Research supported by the National Science Foundation Grant GP-5230.