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

National Center for Atmospheric Research
Boulder, Colorado


**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

$$A = Y^{-1}JY,$$

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
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) \( H = Y^{-1} J Y \),

(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 \( L_Y = \bigoplus_l L_{x_l} \) then \( L_{x_l} \) commutes with \( J_i, i = 1, \ldots, r \).

In Lemma 1 we prove that for \( k = 1, \ldots, n \) 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.** \( UHM = UHM_n = \{ H : H \text{ Hessenberg}, h_{i+1,i} \neq 0, i = 1, \ldots, n-1 \} \)

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 \( H \in UHM \).

(i) \( H = R^{-1} F R \), \( R \) upper triangular,

(ii) \( F = V^{-1} J V \), \( V \) 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, \( R^{-1} F R \) 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.

\[
J = \begin{pmatrix}
\lambda_1 & 1 & & \\
1 & \lambda_1 & & \\
& 1 & \lambda_2 & \\
& & \lambda_3 & \\
& & 1 & \lambda_3
\end{pmatrix}, \quad V_J = \begin{pmatrix}
1 & \lambda_1 & \lambda_1^2 & \lambda_1^3 & \lambda_1^4 & \lambda_1^5 \\
0 & 1 & 2\lambda_1 & 3\lambda_1^2 & 4\lambda_1^3 & 5\lambda_1^4 \\
0 & 0 & 1 & 3\lambda_1 & 6\lambda_1^2 & 10\lambda_1 \\
1 & \lambda_2 & \lambda_2^2 & \lambda_2^3 & \lambda_2^4 & \lambda_2^5 \\
1 & \lambda_3 & \lambda_3^2 & \lambda_3^3 & \lambda_3^4 & \lambda_3^5 \\
0 & 1 & 2\lambda_3 & 3\lambda_3^2 & 4\lambda_3^3 & 5\lambda_3^4
\end{pmatrix}
\]
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 \( H \in \text{UHM} \), its Jordan form

\[
J = J_1 \oplus \cdots \oplus J_r
\]

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

\[
\det (V) = \prod_{\alpha, \beta = 1}^{r} (\lambda_{\alpha} - \lambda_{\beta})^{m_{\alpha}m_{\beta}} \neq 0.
\]

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

\[
V = L_r U_r, \quad \text{diag} (L_r) = I.
\]

Since \( R \) is upper triangular we set \( Y = VR \) and by the uniqueness of the decomposition \( L_r = L_r, U_r = U_r R \). This establishes (ii) of Theorem 1.

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

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

\[
V = \begin{pmatrix} V_1 \\ \vdots \\ V_r \end{pmatrix}, \quad V_i \text{ is } m_i \times n, \quad i = 1, \cdots, r.
\]

Partition \( L = L_r \) 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 \( 0 \leq \mu_i \leq m_i \) and choose the top \( \mu_i \) rows of \( L_i \). This submatrix is determined by the vector \( \mu = (\mu_1, \cdots, \mu_r) \). It has \( \sum_{i=1}^{r} \mu_i = |\mu| \) rows and \( n = \sum_{i=1}^{r} m_i = |m| \) columns. We denote by \( L_{\mu} \) the submatrix obtained by taking only the first \( k \) columns, \( k \leq n \). Note that \( L = L_{|\mu|} \).

Lemma 2. Let \( V \) be the Vandermonde matrix associated with the nonderogatory Jordan form \( J = J_1 \oplus \cdots \oplus J_r \). Let \( V = V_{|\mu|} = LU, m = (m_1, \cdots, m_r), m_i \) the order of \( J_i \). Then for \( 0 \leq \mu_i \leq m_i, i = 1, \cdots, r \) the submatrices \( L_{\mu} \) are nonsingular.

Proof. In general \( V_{|\mu|}^{m_{|\mu|}} = L_{|\mu|} U_{|\mu|}^{m_{|\mu|}} \). Since we may write

\[
U_{|\mu|}^{m_{|\mu|}} = \begin{pmatrix} \tilde{U} \\ 0 \end{pmatrix},
\]

where \( \tilde{U} \) is \( |\mu| \times |\mu| \) and upper triangular, we have

\[
\det [V_{|\mu|}^{m_{|\mu|}}] = \det [L_{|\mu|}^{m_{|\mu|}}] \det [\tilde{U}].
\]
Now $V^a_{|\mu|}$ is the Vandermonde matrix determined by the $\lambda_i$ with multiplicity $\mu_i$, $i = 1, \ldots, r$. 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 $L_V = (L_V^{ij})$, $i \geq j$, where $L_V^{ij}$ is $m_i \times m_j$. We shall exhibit a typical element $l^{ij}_{\alpha \beta}$. We denote by $V_\delta$ the leading principal $\delta$-rowed submatrix of $V$ and by $V_{\gamma, \delta}$ the matrix obtained from $V_\delta$ by replacing row $\delta$ by the corresponding elements of row $\gamma$ of $V$. If $L_V = (l_{\alpha \beta})$ then by [3, Chapter 1]

\[(4.1) \quad l_{\gamma \delta} = \frac{\det [V_{\gamma, \delta}]}{\det [V_\delta]}, \quad \gamma \geq \delta.\]

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

By (3.2)

\[(4.2) \quad \det [V_{\delta}] = \prod_{\mu_\gamma, \mu_\delta = 1}^{\gamma-1} (\lambda_\mu - \lambda_\nu)^{m_\delta} \prod_{\nu_1 = 1}^{\gamma-1} (\lambda_\nu - \lambda_\gamma)^{m_\gamma} = \Gamma(p_j(\lambda_j))^{\beta},\]

since $\beta$ is the multiplicity of $\lambda_i$ in $V_\delta$.

\[\Gamma = \prod_{\mu_\gamma, \mu_\delta = 1}^{\gamma-1} (\lambda_\mu - \lambda_\nu)^{m_\delta} = \det [V_{\gamma-\delta}].\]

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

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

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

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

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 $n_i = \sum_{i=1}^{\gamma-1} m_i$ and $1 \leq \beta \leq \alpha \leq m_i$, $i \leq r$. Then

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

The result follows by induction on $\alpha$ 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

\[l^{ij}_{\alpha \beta} = p_i(\alpha - \beta)(\lambda_i)/(\alpha - \beta)!p_i(\lambda_i); \quad i = j, \quad \alpha \geq \beta,\]

\[(4.6) \quad \left(\frac{d}{d\lambda_i}\right)^{\alpha - 1} [p_j(\lambda_\gamma)(\lambda_i - \lambda_j)^{\beta - 1}]/(\alpha - 1)!p_j(\lambda_\gamma), \quad i > j.\]

**Corollary.** If $H$ has linear elementary divisors then

\[(4.7) \quad l_{\gamma \delta} = \frac{p_\delta(\lambda_\gamma)}{p_\delta(\lambda_\delta)}, \quad \gamma \geq \delta.\]

We are now in a position to establish (iii) of Theorem 1. We observe that $L^{ii}$ is
ELIMINATION METHOD FOR COMPUTING GENERALIZED INVERSE

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 $j$th 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

\begin{equation}
H = Q^{-1}A,
\end{equation}

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

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

\begin{equation}
H = \begin{bmatrix} B \\ 0 \end{bmatrix}
\end{equation}

where $B$ is $r \times n$.

Received July 13, 1966.

* Research supported by the National Science Foundation Grant GP-5230.