COMPLEX SYMMETRIC MATRICES

B. D. CRAVEN

1. Introduction

It is well known that a real symmetric matrix can be diagonalised by an orthogonal transformation. This statement is not true, in general, for a symmetric matrix of complex elements. Such complex symmetric matrices arise naturally in the study of damped vibrations of linear systems. It is shown in this paper that a complex symmetric matrix can be diagonalised by a (complex) orthogonal transformation, when and only when each eigenspace of the matrix has an orthonormal basis; this implies that no eigenvectors of zero Euclidean length need be included in the basis. If the matrix cannot be diagonalised, then it has at least one invariant subspace which consists entirely of vectors of zero Euclidean length.

A symmetric normal form has been obtained for the non-diagonalisable case, and its application shown to the solution of equations of damped linear vibrations. This normal form is different from the (non-unique) symmetric normal form which Wellstein [1] (quoted in Gantmacher [2], pages 9—12) obtained for a complex symmetric matrix, by transforming the Jordan normal form. Although Wellstein's form contains fewer non-zero elements, his result and method do not exhibit the essential role of vectors of zero Euclidean length, in the non-diagonalisable case. This is done in the present paper.

The following notation and terms will be used. A matrix G, of real or complex elements, is *orthogonal* if its transpose equals its inverse, $G' = G^{-1}$. The $n \times n$ matrices A and B are *similar* if $B = T^{-1}AT$ for some non-singular matrix T, and *orthogonally similar* if $B = G'AG$, where G is orthogonal. The matrix A is *complex symmetric* if $A' = A$, but the elements of A are not necessarily real numbers. Vectors u, v, in complex n-space C_n will be considered, in matrix notation, as column vectors, though usually written, for brevity, in row form as $u = \{u^1, u^2, \cdots, u^n\}$. The *inner product* and the resulting squared Euclidean (quasi-) norm are defined respectively by

$$(1) \qquad u'v = \sum_{k=1}^{n} u^k v^k$$

$$(2) \qquad ||u||^2 = u'u.$$

$||u||$ is not defined uniquely, but only $||u||^2$ is required later. The vector u is *quasi-null* (q.n.) if $||u||^2 = 0$ but $u \neq 0$, (e.g. the vector $\{1, i\}$ is q.n.) A finite set of vectors $u_1, u_2, \cdots, u_r$ is *orthogonal* if $u_j' u_k = 0$ ($j \neq k$; $j, k = 1, \cdots, r$); the set is *orthonormal* (o.n.) if, in addition, $||u_j||^2 = 1$ ($j = 1, \cdots, r$). A basis for a subspace is *quasi-null-free* (q.n.f.) if no vector in the basis is q.n. Denote also the vectors

$$e_1 = \{1, 0, \cdots, 0\}, \, e_2 = \{0, 1, 0, \cdots, 0\}, \cdots, e_n = \{0, 0, \cdots, 0, 1\}.$$

Let I denote a unit matrix.

THEOREM 1. (*Gantmacher* [2], *page* 8, *Theorem* 4.) *If two complex symmetric matrices are similar, then they are orthogonally similar.*

It follows that a complex symmetric matrix is diagonalisable by a similarity transformation when and only when it is diagonalisable by a (complex) orthogonal transformation. For this reason, orthogonal matrices and the Euclidean norm (2) are relevant to the problem, and not unitary matrices and the Hermitian norm. (A complex symmetric matrix is Hermitian only if it is real.) Arguments based on linear independence are unaffected by the choice of norm.

Most of the usual diagonalisation proof for real symmetric matrices applies also to complex symmetric matrices, but the proof assumes at one stage, in constructing an orthogonal set of vectors, that any non-zero vector can be normalised. This is not always so in the complex case, since q.n. vectors can occur. A preliminary investigation of the orthogonalisation procedure is therefore required.

2. Orthogonalisation

A number of preliminary results are stated as lemmas.

LEMMA 1. *Orthogonality does not imply linear independence.*

PROOF. In C_2, the vectors $\{1, i\}$ and $\{\alpha, \beta\}$ are orthogonal if $\alpha + i\beta = 0$, i.e. if $\{\alpha, \beta\} = -i\beta\{1, i\}$. So the only vector orthogonal to the quasi null vector $\{1, i\}$ is $\{1, i\}$ itself, apart from constant multipliers. No orthogonal basis of C_2 contains $\{1, i\}$.

LEMMA 2. *Any two-dimensional subspace S of C_n contains a q.n. vector.*

PROOF. By a suitable rotation (i.e. orthogonal transformation), consider S as spanned by $u = \{1, 0, 0, \cdots, 0\}$ and $v = \{p, q, 0, \cdots, 0\}$, where $q \neq 0$. If $p^2 + q^2 = 0$, the lemma is proved. If not, then $p^2 + q^2 = 1$ may be assumed. Then $\alpha u + \beta v$ is q.n. when

$$(\alpha + \beta p)^2 + (\beta q)^2 = 0,$$

i.e. when
$$\beta = \alpha\left(-p \pm \sqrt{(p^2 - 1)}\right).$$

LEMMA 3. *In C_n $(n > 2)$ there exists a two-dimensional subspace S, which has a basis containing no q.n. vectors, but which does not have a basis of orthonormal vectors.*

PROOF. Consider the subspace of C_3 spanned by $e_1 = \{1, 0, 0\}$ and $z = \{1, 1, i\}$. The vectors $\alpha e_1 + \beta z = \{\alpha + \beta, \beta, i\beta\}$ and $\gamma e_1 + \delta z$ are not q.n. only if $||\alpha e_1 + \beta z||^2 = (\alpha + \beta)^2 \neq 0$, and likewise $(\gamma + \delta)^2 \neq 0$. But then their inner product $(\alpha + \beta)(\gamma + \delta) \neq 0$.

LEMMA 4. *If the vectors $v_1, v_2, \cdots, v_r$ $(r \leq n)$ are orthogonal and not q.n., then they are linearly independent.* (The usual proof applies.)

LEMMA 5. *Let $v_1, v_2, \cdots, v_r$ $(r < n)$ be orthonormal vectors in C_n. Then there exists a vector v_{r+1}, such that $v_1, v_2, \cdots, v_r, v_{r+1}$ are orthonormal.*

REMARK. The usual orthogonalisation procedure constructs a vector v_{r+1} orthogonal to $v_1, \cdots, v_r$, but then v_{r+1} is *not* necessarily quasi-null. (Compare Lemma 3.)

PROOF. At most r of the vectors e_j can be linearly dependent on $v_1, \cdots, v_r$; by renumbering the e_j, suppose that $e_1, \cdots, e_s$ $(0 \leq s \leq r)$ are linearly dependent on $v_1, \cdots, v_r$. For $s < k < n$, set

$$(3) \qquad u_k = e_k - \sum_{i=1}^{r} \alpha_i v_i,$$

where

$$\alpha_i = v_i' e_k = v_i^k \qquad (i = 1, 2, \cdots, r.)$$

Then u_k is orthogonal to $v_1, \cdots, v_r$, and

$$||u_k||^2 = (e_k - \sum_{i=1}^{r} \alpha_i v_i)'(e_k - \sum_{i=1}^{r} \alpha_i v_i)$$

$$(4) \qquad = 1 - 2\sum_{i=1}^{r} \alpha_i v_i^k + \sum_{i=1}^{r} \alpha_i^2 \quad \text{since } v_1, \cdots, v_r \text{ are orthogonal}$$

$$= 1 - \sum_{i=1}^{r} (v_i^k)^2.$$

If there is k $(s < k \leq n)$ such that $||u_k||^2 \neq 0$, then u_k may be normalised to unit norm, and e_k is linearly dependent on $v_1, \cdots, v_r, v_{r+1} = u_k$. A contrary assumption leads to a contradiction. Since $e_1, \cdots, e_s$ are linearly dependent on $v_1, \cdots, v_r$, there are constants β_j^m such that

$$(5) \qquad e_m = \sum_{j=1}^{r} \beta_j^m v_j \qquad (1 \leq m \leq s).$$

Then

$$v_j^m = v_j' e_m = \beta_j^m,$$

so that

(6) $$1 = ||e_m||^2 = \sum_{j=1}^{r} (\beta_j^m)^2 = \sum_{j=1}^{r} (v_j^m)^2.$$

If also (6) holds for $s+1 \leq m \leq n$, i.e. $||u_k||^2 = 0$ for all k, then

$$n = \sum_{m=1}^{n} \left[\sum_{j=1}^{r} (v_j^m)^2 \right] = \sum_{j=1}^{r} \left[\sum_{m=1}^{n} (v_j^m)^2 \right] = r,$$

contradicting $r < n$.

From this follows the orthogonalisation theorem required for the diagonalisation process.

THEOREM 2. *If* $v_1, v_2, \cdots, v_r$ *are orthonormal vectors in* C_n, *then there exist vectors* $w_{r+1}, \cdots, w_n$ *such that* $v_1, \cdots, v_r, w_{r+1}, \cdots, w_n$ *are orthonormal.*

PROOF. By induction, using Lemma 5.

This result shows that an orthonormal set of vectors $v_1, \cdots, v_r$ $(r < n)$ can always be extended to an orthonormal basis for C_n; therefore there is an orthogonal transformation, given by the matrix

$$P = [v_1, \cdots, v_r, w_{r+1}, \cdots, w_n]$$

which maps $v_1, \cdots, v_r$ on $e_1, \cdots, e_r$. But it is not always possible, as Lemma 3 shows, to construct an orthonormal basis for a given subspace.

LEMMA 6. *Let* S *be a subspace of* C_n. *Then the following three statements are equivalent*:
  (i) *every basis of* S *contains a q.n. vector*;
  (ii) *every basis of* S *consists of q.n. vectors*;
  (iii) *every element of* S *is q.n.*

PROOF. Obviously (iii) $\Rightarrow$ (i); and (iii) untrue $\Rightarrow$ (ii) untrue, by the Steinitz replacement theorem. To show that (ii) untrue $\Rightarrow$ (i) untrue, let S have a basis $w_1, \cdots, w_r$, in which $w_1, \cdots, w_s$ are q.n. and $w_{s+1}, \cdots, w_r$ are not q.n.; then the vectors $w_1 + c_1 w_r, \ w_2 + c_2 w_r, \ \cdots, \ w_s + c_s w_r,$ $w_{s+1}, \cdots, w_r$, where nonzero $c_j \neq -2w_j' w_r / ||w_r||^2$, form a q.n.f. basis for S.

There are therefore three types of subspaces of C_n, namely those possessing (a) an o.n. basis, (b) a q.n.f. basis but no o.n. basis, (c) q.n. elements only (as in Lemma 6.)

LEMMA 7. *A subspace* S *of* C_n *of type* (b) *is the direct sum of orthogonal subspaces of types* (a) *and* (c).

PROOF. Let S have a q.n.f. basis $w_1, \cdots, w_d$. Without loss of generality, $||w_1|| = 1$ can be assumed. By Theorem 2, there is an orthogonal transformation R of C_n which maps w_1 on e_1. Now the vectors e_1 and

$$x_j = Rw_j - (e_1' Rw_j)e_1 \qquad\qquad (2 \leq j \leq d)$$

form a basis for the subspace RS; so RS is the direct sum of the subspace A spanned by e_1 and the subspace B spanned by $x_2, \cdots, x_d$; since the first component of each x_j vanishes, A is orthogonal to B. Therefore S is the direct sum of the subspaces $R'A$, of type (a), and $R'B$.

If $R'B$ is of type (c), the lemma is proved. If it is of type (b), it may be expressed as a direct sum of a subspace of type (a), and another subspace; and so on. The process terminates in at most d steps.

COROLLARY. *Any subspace of C_n is a direct sum of orthogonal subspaces of types (a) and (c).*

LEMMA 8. *Let S be a subspace of C_n. If S has one orthonormal basis, then every orthonormal set in S can be extended to an orthonormal basis of S.*

PROOF. Let $b_1, \cdots, b_d$ be an o.n. basis for S. Let $w_1, \cdots, w_p$ be an o.n. set in S, which is not a basis; then $p < d$. There is an orthogonal transformation R of C_n which maps $b_1, \cdots, b_d$ on $e_1, \cdots, e_d$; then R maps $w_1, \cdots, w_p$ on o.n. vectors $w_1^*, \cdots, w_p^*$, say, in which the coordinates $d+1, \cdots, n$ vanish. Applying Theorem 2 to the d-space S^* defined by the first d coordinates, there is an orthonormal basis $w_1^*, \cdots, w_p^*, y_{p+1}, \cdots, y_d$ for S^*. Since $d > p$, the vectors

$$w_1, \cdots, w_p, w_{p+1} = R^{-1}y_{p+1}, \cdots, w_d = R^{-1}y_d$$

form a basis for S.

3. Diagonalisation

Let A be a $n \times n$ complex symmetric matrix. If λ is an eigenvalue of A, denote by S_λ the corresponding eigenspace, i.e. the subspace spanned by the eigenvectors corresponding to λ.

THEOREM 3. *Let A be a complex symmetric matrix. Then*
(a) A can be diagonalised by an orthogonal transformation if and only if
(b) every eigenspace S_λ of A possesses an orthonormal basis.

REMARKS. In (b), no hypothesis is made concerning the *dimension* of S_λ; in fact (b) implies that the dimension of S_λ equals the multiplicity of λ. An orthonormal basis is equivalent to an orthogonal basis containing no q.n. vectors. By Lemma 3, it is not sufficient to assume merely a basis for S_λ, containing no q.n. vectors. By Lemma 2, if S_λ has dimension greater than 1, then S_λ contains q.n. vectors; hypothesis (b) concerns a *basis* for S_λ, not all vectors in S_λ.

The proof that $(a) \Rightarrow (b)$ is trivial. The proof that $(b) \Rightarrow (a)$ adapts a standard proof for real symmetric matrices; (b), with Theorem 2, is just sufficient to prevent the occurrence of q.n. vectors in the construction.

PROOF *that* $(b) \Rightarrow (a)$. If the union of the orthonormal bases of the S_λ is *not* an orthonormal basis for C_n, then this union consists of orthonormal vectors $w_1, w_2, \cdots, w_r$, where $r < n$.

By Theorem 2, there are vectors $w_{r+1}, \cdots, w_n$ such that

$$P = [w_1, w_2, \cdots, w_n]$$

is an orthogonal matrix. Consequently

(7) $$P'AP = \begin{bmatrix} D & 0 \\ 0 & B \end{bmatrix},$$

where D is a diagonal matrix of eigenvalues of A, and $B = [w_i'Aw_j]$ $(i, j = r+1, \cdots, n)$ is also a complex symmetric matrix. The set of eigenvalues of B, with repetitions, coincides with the set of eigenvalues of A, with repetitions, minus those occurring as diagonal elements of D.

Since $r < n$, B has an eigenvalue λ, and a corresponding eigenvector g. Corresponding to λ and g, A has an eigenvalue λ, and an eigenvector Ph, where $h = \{0, g\}$, and Ph lies in the eigenspace S_λ of A. Now Ph is linearly independent of $w_1, \cdots, w_r$, for if

$$\sum_1^r \alpha_i w_i + \beta Ph = 0,$$

then, for $1 \leqq j \leqq r$,

$$\sum_{i=1}^r \alpha_i w_j'w_i + \beta(w_j'P)h = 0,$$

i.e. $\alpha_j + \beta \cdot 0 = 0$. So $\alpha_j = 0$, $\beta = 0$. Therefore S_λ is *not* spanned by a subset of $w_1, \cdots, w_r$. This contradiction shows that the union of the orthonormal bases of the S_λ is a basis for C_n. The orthogonal matrix whose columns are these basis vectors therefore diagonalises A.

4. Invariant subspaces

The diagonalisation process fails if there is an invariant subspace for which *every* basis contains a q.n. vector, or equivalently, by Lemma 6, if every basis consists of q.n. vectors. Some properties of such subspaces are as follows.

LEMMA 9. *If S is a subspace of C_q, and every basis of S consists entirely of q.n. vectors, then S has dimension $d \leqq \frac{1}{2}q$.*

PROOF. Let $u_1, \cdots, u_d$ be a basis for S, where, without loss of generality, u_1 may be assumed to be of the form $u_1 = \{1, \alpha, \beta, \cdots\}$. Since u_1 is q.n., $\alpha^2 + \beta^2 + \cdots = -1$. Since the $(q-1)$-vector $\{-i\alpha, -i\beta, \cdots\}$ constitutes, trivially, an orthonormal set, there exists, by Theorem 2, an orthonormal

basis for C_{q-1} which contains this vector. There exists, therefore, an orthogonal transformation R of C_q which leaves the first coordinate unchanged, and which maps u_1 on $v_1 = \{1, i, 0, \cdots\}$.

For $2 \leqq j \leqq n$, let $Ru_j = \gamma_j z_j$, where γ_j is constant, $z_j = \{a_j, b_j, c_j, \cdots\}$, and $a_j = 0$ or 1. In the basis for RS, $\gamma_j z_j$ may be replaced by

$$z_j - v_1 = \{a_j - 1, b_j - i, c_j, \cdots\}.$$

Since every vector in S is q.n., $z_j - v_1$ is q.n., and this holds if and only if

$$0 = (a_j - 1)^2 + (b_j - i)^2 + c_j^2 + \cdots;$$

since also z_j is q.n., this condition reduces to $b_j = ia_j$. So, for $2 \leqq j \leqq n$, if $a_j = 0$, then the first two coordinates of z_j vanish; if $a_j = 1$, the first two coordinates of $z_j - v_1$ vanish. Consequently, RS has a basis consisting of v_1 and $(n-1)$ vectors $v_2, \cdots, v_n$, for each of which the first two coordinates vanish.

This process can now be applied to $v_2, \cdots, v_n$, considering only coordinates $3, 4, \cdots, q$; and so on. The process cannot terminate with a single vector, since a 1-vector cannot be q.n., hence $d \leqq \frac{1}{2}q$.

REMARKS. In particular, if q is even, and $d = \frac{1}{2}q$, let G denote the subspace of S spanned by the vectors

$$g_1 = \{1, i, 0, 0, 0, 0, 0, \cdots\},$$
$$g_2 = \{0, 0, 1, i, 0, 0, 0, \cdots\},$$
$$g_3 = \{0, 0, 0, 0, 1, i, 0, \cdots\}, \text{ etc.}$$

Then G is the image of S by a suitable orthogonal transformation. The maximal subspace orthogonal to G is G itself (compare Lemma 1), for, if $w = \{a, b, c, d, \cdots\}$ is orthogonal to G, then $0 = a+bi = c+di = \cdots$, so $w = \{a, ai, b, bi, \cdots\} \in G$; and conversely.

Let M be a symmetric $q \times q$ matrix, all of whose eigenvalues are zero, and all of whose eigenvectors are q.n. Let S denote the eigenspace of M, i.e. the subspace spanned by its eigenvectors. By the construction of Lemma 7, there is an orthogonal transformation R of C_q which maps a basis for S, say $w_1, \cdots, w_d$ (where d is the dimension of S) onto the vectors $g_1, \cdots, g_d$ (which span a subspace, G_d say, of G). Since $g_j = Rw_j$ $(1 \leqq j \leqq d)$, and

$$w \in S \Leftrightarrow Mw = 0 \Leftrightarrow (RMR')Rw = 0,$$

the symmetric matrix RMR' has G_d as its eigenspace.

If, in particular, $d = \frac{1}{2}q$ (thus q is even), then the eigenspace of M is an orthogonal map of G, so that the subspace V spanned by the rows, or columns, of M is orthogonal to S. Since the maximal subspace orthogonal to S is S itself, $V \subset S$, so V consists entirely of q.n. vectors, and each column of M is a q.n. vector. Conversely, let M be a symmetric $q \times q$ matrix, all of

whose eigenvalues are zero, and such that the subspace V spanned by the columns of M contains only q.n. vectors, and has dimension $\frac{1}{2}q$. Then S, the eigenspace of M, is orthogonal to V, and since V is an orthogonal map of G, the maximal subspace orthogonal to V is V; hence $S \subset V$, and the eigenspace of M contains only q.n. vectors.

Lemma 9 shows that $d \leq \frac{1}{2}q$; the case $d < \frac{1}{2}q$ can occur, as is shown by the example

$$(9) \qquad\qquad M = \begin{bmatrix} 1 & i & C \\ i & -1 & iC \\ C & iC & 0 \end{bmatrix},$$

where $C \neq 0$. Here M has all eigenvalues zero, and all eigenvectors are multiples of $\{1, i, 0\}$.

LEMMA 10. *Given any subspace S of C_n, there exists a (complex) symmetric matrix A for which S is an eigenspace.*

PROOF. From Lemma 7 (Corollary), S is a direct sum of orthogonal subspaces with o.n. bases, and subspaces containing only q.n. vectors. Using the construction of Lemma 9, there is an orthogonal transformation R which reduces these subspaces to direct sums of the following subspaces (specified by their basis vectors):

$$(10) \qquad\qquad (a)\ \{1\}; \qquad (b)\ \{1, i\}.$$

Corresponding to these subspaces, suitable matrices are

$$(11) \qquad\qquad (a)\ [0]; \qquad (b)\ \begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix}.$$

Thus, for example, the subspace spanned by the vectors $\{1, i, 0, 0\}$, $\{0, 0, 1, i\}$ is an eigenspace of the matrix

$$(12) \qquad\qquad \begin{bmatrix} 1 & i & 0 & 0 \\ i & -1 & 0 & 0 \\ 0 & 0 & 1 & i \\ 0 & 0 & i & -1 \end{bmatrix},$$

which is of block diagonal form, the blocks being the relevant matrices from (11).

In general, the required matrix $A = R'BR$, where B is a block diagonal matrix of the required blocks from (11).

5. Symmetric normal form

Let A be any symmetric $n \times n$ matrix. The normal form given by Gantmacher [2] (page 9) shows that there is a (complex) orthogonal matrix Q such that $Q'AQ$ is of block diagonal form

$$(13) \qquad \begin{bmatrix} B_1 & 0 & 0 & \cdots \\ 0 & B_2 & 0 & \cdots \\ 0 & 0 & B_3 & \cdots \\ \cdots\cdots\cdots\cdots & & & \cdots \end{bmatrix},$$

where each block B is either (i) a 1×1 submatrix consisting of an eigenvalue of A, or (ii) $B = \lambda I + M$, where M is a $q \times q$ matrix, all of whose eigenvalues are zero, and (by a direct calculation from the Gantmacher form) all of whose eigenvectors are q.n., and λ is an eigenvalue of A. The same λ may correspond to several blocks. The remarks following Lemma 9 show that, by suitable choice of Q, the eigenspace of M is spanned by d vectors $\{1, i, 0, 0, 0, \cdots\}$, $\{0, 0, 1, i, 0, \cdots\}$, etc., where $d \leq \frac{1}{2}q$. The derivation of the Gantmacher form from the Jordan normal form shows that the blocks can be chosen to have each $d = 1$; but, for the present, d will not be so restricted.

For the special case where all eigenvalues of A are zero, an alternative proof of (13) is outlined, showing how submatrices M with all eigenvectors q.n. occur. The proof is by induction; assume that, for some $r \leq n$, $Q'AQ = \begin{bmatrix} 0 & 0 \\ 0 & K \end{bmatrix}$, where Q is orthogonal, and K is $r \times r$ complex symmetric. If K has an eigenvalue (necessarily zero) and a corresponding eigenvector z which is not q.n., then, by Theorem 2, there is an orthogonal matrix P, containing z as one column, such that $P'KP = \begin{bmatrix} 0 & 0 \\ 0 & K* \end{bmatrix}$, where $K*$ is $(r-1) \times (r-1)$ symmetric. Then $Q*'AQ* = \begin{bmatrix} 0 & 0 \\ 0 & K* \end{bmatrix}$, where $Q* = Q\begin{bmatrix} 1 & 0 \\ 0 & P \end{bmatrix}$ is orthogonal. The induction begins with $K = A$, $Q = I$, and ends when the final $K*$ is either zero, or has an eigenspace containing only q.n. vectors.

Let F denote the matrix obtained from M by deleting all but the first $2d$ rows and columns. If now F is partitioned into 2×2 submatrices

$$(14) \qquad F = \begin{bmatrix} F_{11} & F_{12} & \cdots & F_{1d} \\ F_{21} & F_{22} & \cdots & F_{2d} \\ \cdots\cdots\cdots\cdots\cdots\cdots & & & \\ F_{d1} & F_{d2} & \cdots & F_{dd} \end{bmatrix},$$

then the requirement that $Mw = 0$ for each eigenvector w of M shows that $F_{rs}v = 0$ $(r, s = 1, \cdots, d)$, where $v = \{1, i\}$. Thus F_{rs} and F_{sr} have the respective forms

$$(15) \qquad F_{rs} = \begin{bmatrix} \alpha & i\alpha \\ i\gamma & -\gamma \end{bmatrix}, \qquad F_{sr} = \begin{bmatrix} \beta & i\beta \\ i\delta & -\delta \end{bmatrix}, \qquad (r \leq s)$$

where $\alpha, \beta, \gamma, \delta$ are constants. But since M is symmetric, $F'_{rs} = F_{sr}$, which implies that $\alpha = \beta = \gamma = \delta$. Therefore

$$(16) \qquad F_{rs} = F_{sr} = c_{rs}T,$$

where the c_{rs} form a symmetric $d \times d$ matrix C, and

(17)
$$T = \begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix}.$$

Since F has rank d, C is non-singular. The columns of F are q.n., and pairwise orthogonal. The matrix F may be written as a direct product of matrices: $F = T \times C$.

If $d < \frac{1}{2}q$, partition M in the form

(18)
$$M = \begin{bmatrix} F & J' \\ J & H \end{bmatrix}$$

where F is a $2d \times 2d$ matrix, H is a $q' \times q'$ matrix, where $q' = q - 2d$, and J is a $q' \times 2d$ matrix. Since $Mw = 0$ for each eigenvector w of M, the elements of J are related by $M_{r, 2s} = iM_{r, 2s-1} (2d < r \leq q, 1 \leq s \leq d)$, but it does not follow that $M_{r+1, 2s-1} = iM_{r, 2s-1}$ as in the case for $r \leq 2d$. Thus J is of the form

(19)
$$J = [\beta_1, i\beta_1, \beta_2, i\beta_2, \cdots, \beta_d, i\beta_d],$$

where each β_k is a q'-vector; but J does not decompose into submatrices T (compare (9)).

Denote by $\phi_t(M)$ (resp. $\phi_t(H)$) the sum of all principal $t \times t$ minors of M (resp. H). The determinants which contribute to $\phi_t(M)$, but not to $\phi_t(H)$, for $t \leq q'$, are of the forms:

(i) containing (part of) column $2s-1$ of M (and the corresponding row) for some s in $1 \leq 2s-1 < 2d$; but not column s;

(ii) as for (i), but with $2s$ and $2s-1$; interchanged;

(iii) containing both (part of) column $2s-1$ and column $2s$, for some s in $1 \leq 2s-1 < 2d$.

Now determinant (iii) $= 0$, since (column $2s$) $= i$ (column $2s-1$); and to each determinant (i) there corresponds a (ii) with the same s, the other rows and columns being the same (and conversely); the extraction of a factor i from both column $2s$ and row $2s$ of (ii) shows that determinant (ii) $= i^2 \times$ determinant (i). Therefore $\phi_t(M) = \phi_t(H)$, for $t \leq q'$. Since the $\phi_t(M)$ (resp. $\phi_t(H)$) are, apart from sign, the coefficients of the characteristic equation of M (resp. H), and all eigenvalues of M are zero, it follows that all eigenvalues of H are zero. But it does *not* follow that all eigenvectors of H are q.n.; (9) is a counter example, where $H = [0]$ has $\{1\}$ as a non-q.n. eigenvector.

If H has a non-q.n. eigenvector, then the reduction (13) of A to block diagonal form may be applied to H. Thus there is a complex orthogonal transformation Q which reduces H to the form

$$Q'HQ = \begin{bmatrix} 0 & 0 \\ 0 & H_1 \end{bmatrix},$$

where H_1 has all its eigenvalues zero, and all its eigenvectors q.n. The orthogonal transformation

$$\begin{bmatrix} I & 0 \\ 0 & Q_1 \end{bmatrix},$$

applied to M, has this effect; this transformation also replaces J by $Q_1 J$, which is also of the form (19). The following theorem has therefore been proved:

THEOREM 4. *If A is a complex symmetric matrix, then there is a complex orthogonal matrix P such that $P'AP$ has the block diagonal form (13). Each block B is either a 1×1 submatrix consisting of an eigenvalue of A, or a $q \times q$ symmetric submatrix of the form $\lambda I + E$, where λ is an eigenvalue of A, E has all its eigenvalues zero, all its eigenvectors quasi-null, and its eigenspace of dimension $d \leq \frac{1}{2}q$. The matrix E has the partitioned form*

(20)
$$E = \left[\begin{array}{c|cc} F & \multicolumn{2}{c}{J'} \\ \hline & 0 & 0 \\ J & 0 & E_1 \end{array} \right]$$

where F has the form given by (14) and (16), J has the form (19), and E_1 has all its eigenvalues zero, and all its eigenvectors quasi-null.

It follows that a similar reduction may be applied to the submatrix E_1, leading recursively to a representation of M in terms of submatrices of the respective forms of F, J, and zero matrices. (The zero matrices do not always occur, but are required, e.g., in (9).)

REMARKS. As already noted, a block diagonal form (13) exists in which each $d = 1$; then each F is a 2×2 matrix cT, c being constant.

If $d = \frac{1}{2}q$, then M reduces to F. In this case, it is observed that the columns of F, and consequently those of M, are orthogonal, and either q.n. or null. Also, for $d = \frac{1}{2}q$, the decomposition into d 2×2 subblocks shows that E can be reduced to a block diagonal matrix, with each diagonal block of the form cT, the c being constants.

An example of the symmetric normal form is as follows (where $\alpha^2 + \beta^2 = -1$):

(21)
$$A = \begin{bmatrix} 1 & \alpha & \beta \\ \alpha & \alpha^2 & \alpha\beta \\ \beta & \alpha\beta & \beta^2 \end{bmatrix}; \qquad P = \begin{bmatrix} -i\alpha/\beta & i/\beta & 0 \\ i/\beta & i\alpha/\beta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$P'AP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\beta^2 & -i\beta^2 \\ 0 & -i\beta^2 & \beta^2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & -\beta^2 T \end{bmatrix}.$$

6. Application to damped oscillations

The small damped oscillations of a linear system, with a finite number of degrees of freedom, may be described by the matrix equation

(22) $N\ddot{x}+K\dot{x}+Lx = 0,$

where the $n \times n$ matrices N, K, L will be assumed real symmetric, with N and L positive definite, and x is a n-vector function of time t. The three matrices cannot, in general, be diagonalised by the same similarity transformation. Although the equation is readily solved by other means, it is of interest to note the relevance of complex symmetric matrices.

The equation may be written in the form $G\dot{u}+Du = 0$, where

$$G = \begin{bmatrix} N & 0 \\ 0 & L \end{bmatrix}, \qquad D = \begin{bmatrix} K & -iL \\ -iL & 0 \end{bmatrix}, \qquad u = \begin{bmatrix} \dot{x} \\ ix \end{bmatrix}.$$

Further simplification gives $\dot{y}+Ay = 0$, where $y = G^{+\frac{1}{2}}u$, and $A = G^{-\frac{1}{2}}DG^{-\frac{1}{2}}$ is complex symmetric. If $P'AP$ is the symmetric normal form for A, given by Theorem 4, and $w = P'y$, then

(23) $\dot{w}+(P'AP)w = 0.$

Let $\lambda I+E$ be one of the block submatrices of $P'AP$; let z be the corresponding partitioned part of w. Then the solution w of (23) is obtained by combining, for each block, the solution z of $\dot{z}+(\lambda I+E)z = 0$, namely

$$z = e^{-\lambda t}(a+bt),$$

where a and b are constant vectors satisfying (for all t)

$$b+E(a+bt) = 0.$$

Thus $Eb = 0$, so that b is in the eigenspace of E; and $Ea = -b$.

If $E = 0$, then $b = 0$. If $E \neq 0$, then b is a quasi-null vector, of the form $b = \{\alpha_1 v, \alpha_2 v, \cdots, \alpha_d v, 0, \cdots 0\}$, where $v = \{1, i\}$, the α_k are constants, and $d \leqq \frac{1}{2}q$, where E is a $q \times q$ matrix. Denote (using (20))

$$E = \begin{bmatrix} F & J' \\ J & Y \end{bmatrix}, \qquad a = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \qquad b = \begin{bmatrix} b_1 \\ 0 \end{bmatrix}; \qquad \alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_d \end{bmatrix}.$$

Then

(24) $Fa_1+J'a_2 = -b_1$
                              $Ja_1+Ya_2 = 0.$

From (19), J is of the form $[\beta_1, i\beta_1, \cdots, \beta_d, i\beta_d]$, so that

$$J'a_2 = \{\beta_1' a_2 v, \cdots, \beta_d' a_2 v\}.$$

Therefore $-Fa_1 = J'a_2 + b_1$ is of the form $\{\gamma_1 v, \cdots, \gamma_d v\}$, where $\gamma_s = \alpha_s + \beta'_s a_2$. Let $a_1 = \{\lambda_1, \cdots, \lambda_d\}$, where each λ_s is a 2-vector. Since F has the form given by (14) and (16),

$$-\sum_{s=1}^{d} c_{rs}(T\lambda_s) = \gamma_r v \qquad\qquad (r = 1, \cdots, d)$$

so that $T\lambda_s = k_s v$, where $\{k_1, \cdots, k_d\} = -C^{-1}\{\gamma_1, \cdots, \gamma_d\}$. Substituting T from (17), this gives

(25) $$\lambda_s = k_s e + g_s v \qquad\qquad (s = 1, \cdots, d)$$

where $e = \{1, 0\}$, $v = \{1, i\}$, and the g_s are arbitrary constants.

Now, from (24),

$$-Ya_2 = Ja_1 = \sum_{1}^{d} \beta_s k_s \qquad\qquad \text{since } v'v = 0.$$

Denoting $\hat{J} = [\beta_1, \cdots, \beta_d]$ and $k = \{k_1, \cdots, k_d\}$, this becomes

$$-Ya_2 = \hat{J}k$$
$$= -\hat{J}C^{-1}(\alpha + \hat{J}'a_2)$$

using the previous expressions for k_s and γ_s. Therefore

(26) $$(Y - \hat{J}C^{-1}\hat{J}')a_2 = \hat{J}C^{-1}\alpha.$$

The solution to the differential equation for z must contain q arbitrary constants; of these, $2d$ are contributed by the two arbitrary d-vectors α and g. Consequently, *either* $2d = q$, so that J, Y and a_2 are eliminated, *or* an additional $q' = q - 2d$ arbitrary constants are contributed by the solution of (26) for the q'-vector a_2. The latter is only possible if the matrix coefficient of a_2 has zero rank, so that a_2 is an arbitrary q'-vector, and $\hat{J}C^{-1}$, and therefore also $\hat{J}$, is zero. In either case, a_1 is then determined by (25).

Then the solution to (22) is completely determined for the case where the complex symmetric matrix D cannot be diagonalised. This case corresponds exactly to the situation when the solution to (22) is not simply a sum of exponentials.

Acknowledgements

normal form of a complex symmetric matrix, given by Gantmacher, to my attention, for a simplification of my proof of Theorem **3**, for Lemma **6**, and for various details of presentation.

References

[1] J. Wellstein, 'Über symmetrische, alternierende und orthogonale Normalformen von Matrizen', *J. für reine und angew. Math.* **163** (1930), 166—182.
[2] F. Gantmacher, *The Theory of Matrices*, Vol. 2. (Chelsea, 1959/1964.)

University of Melbourne
Melbourne, Victoria