SUPERDIAGONALIZATION AND REAL NORMAL OPERATORS

Autor(en): **Geissinger, Ladnor**

Objekttyp: **Article**

Zeitschrift: **L'Enseignement Mathématique**

Band (Jahr): **22 (1976)**

Heft 1-2: **L'ENSEIGNEMENT MATHÉMATIQUE**

PDF erstellt am: **21.09.2024**

Persistenter Link: https://doi.org/10.5169/seals-48179

SUPERDIAGONALIZATION AND REAL NORMAL OPERATORS

by Ladnor GEISSINGER

We shall give short direct proofs of all the usual simultaneous superdiagonalization and diagonalization theorems found in linear algebra textbooks, as well as determine the structure of real normal operators. The proofs are all variations on the technique of superdiagonalization, a method well-known to be very useful not just for elementary linear algebra but also in the theory of finite dimensional Lie (and also associative) algebras [Flanders].

Henceforth E will be a finite dimensional vector space over a field K and S will be a set of commuting operators on E. A subspace W of E is *S-stable* if $AW \subseteq W$ for all A in S. Our main tool is the following strong commutative version of Schur's lemma.

PROPOSITION 1. *If* $\dim_K E > 1$, *then either there is a proper S-stable subspace or the subalgebra* S' *generated by* S *over the scalars* KI *is a field and* $\dim_K S' = \dim_K E$.

Proof. Suppose first that some operator A in S (or S') has a minimal polynomial which is not irreducible over K; so $m_A = pq$ where p and q are not constants. Since $p(A) \neq 0$ there is a vector v such that $w = p(A)v \neq 0$. Then $q(A)w = 0$ but $q(A) \neq 0$ so the kernel W of $q(A)$ is a proper subspace. Moreover, W is S-stable since if B is in S then $q(A)B(W) = Bq(A)(W) = 0$. Of course if all the operators in S are scalars (i.e. $S \subseteq KI$) every K-subspace is S-stable. Thus the only remaining case is when for each non-scalar A in S, m_A is irreducible over K of degree bigger than 1. Then the subalgebra K' generated by one such A over KI is a field and E is a vector space over K' of dimension less than $\dim_K E$ (in fact, $(\dim_{K'} E)(\dim_K K') = \dim_K E$). Moreover, the operators in S which are not in K' are K'-linear since they commute with A. Thus the proof can be completed by induction on the dimension of E.

We really only need the following special cases of Proposition 1 which result from successive restriction of S to S-stable subspaces.

Corollary. *If each of the operators in* S *has all of its characteristic roots in* K, *then a minimal* S*-stable subspace has dimension* 1 *over* K. *If* $K = \mathbf{R}$, *then a minimal* S*-stable subspace has dimension* 1 *or* 2 *over* $\mathbf{R}$.

Note that the first statement uses only the first half of the proof of Proposition 1. Also, when $K = \mathbf{R}$, the second half of the proof may be more transparent if E is considered to be a vector space over $\mathbf{C}$ by the isomorphism $s + ti \to sI + t\dfrac{A-aI}{b}$ where $m_A(X) = (X-a)^2 + b^2$ and then the proof is completed by appeal to the result already proved for $\mathbf{C}$ in the first statement.

A chain of subspaces $0 = W_0 < W_1 < W_2 < \ldots < W_r = E$ will be called a *flag*, or a *fan* if $\dim W_k/W_{k-1} = 1$ for all k, and a basis $x_1, \ldots, x_n$ will be called a *flag (fan) basis* if $x_1, x_2, \ldots, x_{d(k)}$ is a basis for W_k for each k, where $d(k) = \dim W_k$. A flag is S-stable if each W_k is S-stable.

Proposition 2. *If each operator in* S *has all of its characteristic roots in* K, *then there is an* S*-stable fan. If* $K = \mathbf{R}$ *there is an* S*-stable flag* $0 = W_0 < W_1 < \ldots < W_r = E$ *in which* $\dim W_k/W_{k-1}$ *is* 1 *or* 2 *for each* k.

Proof. Apply Proposition 1 or its Corollary to get an S-stable subspace W; then apply it again to S restricted to W and/or to the operators induced by S on the quotient space E/W.

Note that in the first case, if $x_1, \ldots, x_n$ is a fan basis, then the matrix of each A in S relative to this basis is *superdiagonal* (upper triangular). In the second case, if $x_1, \ldots, x_n$ is a flag basis then the matrix of each A in S relative to this basis is *nearly superdiagonal*, that is, there are 1×1 and 2×2 blocks down the diagonal and zeros below. Translating these statements entirely into the laguage of matrices gives the following result.

Corollary. *Let* S *be a commuting set of matrices over* K. *If each matrix in* S *has all its characteristic roots in* K, *there is an invertible matrix* P *with entries in* K *such that* $P^{-1}AP$ *is superdiagonal for all* A *in* S. *If* $K = \mathbf{R}$, *there is an invertible real matrix* P *such that* $P^{-1}AP$ *is nearly superdiagonal for all* A *in* S.

Recall that a polynomial with coefficients in K is *separable* if it factors into a product of distinct irreducible factors over K (and over every field extension of K if characteristic $(K) > 0$), that is, has no multiple roots in any field extension of K. Call an operator *separable* if its minimal polynomial is separable. If the minimal polynomial of a separable operator A in S factors as $m_A = qp$ over K, there are polynomials r and s such that $rp + sq = 1$. Then $I = r(A)p(A) + s(A)q(A)$ so E is the direct sum of $W = \text{Ker}\, q(A)$ and $V = \text{Ker}\, p(A)$.

As in the proof of Proposition 1 it follows that W and V are both S-stable. This gives a strengthened form of Proposition 1.

PROPOSITION 3. *If* $\dim_K E > 1$ *and if all the operators in* S *are separable, then either there are proper supplementary* S*-stable subspaces or the subalgebra* S' *generated by* S *over the scalars* KI *is a field (separable over* K*) and* $\dim_K S' = \dim_K E$.

COROLLARY. *If each operator in* S *is separable and has all of its characteristic roots in* K, *then* E *is the direct sum of* 1 *dimensional* S*-stable subspaces. If* $K = \mathbf{R}$, *then* E *is the direct sum of* 1 *or* 2 *dimensional* S*-stable subspaces, provided every operator in* S *is separable.*

In the first case, if $E = W_1 \oplus W_2 \oplus \ldots \oplus W_n$ with the W_i being S-stable and if x_i spans W_i, then each x_i is a common eigenvector for all A in S (or S') and the matrix of each A relative to this basis is diagonal. In the second case, $E = W_1 \oplus \ldots \oplus W_k \oplus V_1 \oplus \ldots \oplus V_r$ where $\dim W_i = 1$ and $\dim V_i = 2$ and the W_i and V_j are minimal S-stable subspaces. If x_i spans W_i and y_j, z_j span V_j, then the matrix of each A relative to the basis $x_1, \ldots, x_k, y_1, z_1, y_2, \ldots y_r, z_r$ is block diagonal with k eigenvalues followed by r 2×2-blocks down the diagonal. This gives the usual simultaneous diagonalization theorem for commuting (diagonalizable) matrices, as well as the criterion for diagonalizability of a single operator.

Next suppose K is $\mathbf{C}$ or $\mathbf{R}$ and E has an inner product $\langle\, , \rangle$, and S is a commuting set of normal operators on E. Recall that A is *normal* if it commutes with its adjoint A^*. Now for a normal A, if $m_A = pq$ and if $W = \text{Ker}\, q(A)$ then W is A^*-stable and so the orthogonal complement W^1 is A-stable since $\langle W, AW^1 \rangle = \langle A^*W, W^1 \rangle = 0$. But $p(A)(W^1) \subseteq \text{Ker}\, q(A) = W$ so $W^1 \subseteq \text{Ker}\, p(A)$ and hence l.c.m. $(p, q) = m_A$, from which we conclude that p and q are relatively prime and $W^1 = \text{Ker}\, p(A)$. Thus W and W^1 are S-stable, and every normal operator is separable. A *root space* for S is a maximal proper S-stable subspace W such that each A in S acts as a scalar on W, that is, W is the intersection of a collection of eigenspaces, one for each A in S.

PROPOSITION 4. *If each operator in S is normal and has all of its characteristic roots in K, then E is the orthogonal direct sum of root spaces for S. If $K = \mathbf{R}$ and each operator in S is normal, then E is the orthogonal direct sum of root spaces for S and S-stable subspaces V on which each A in S acts as a scalar or has irreducible (quadratic) minimal polynomial.*

Upon choosing an orthonormal basis for each root space, the first statement of Proposition 4 yields the simultaneous unitary diagonalization of commuting normal operators (or matrices) when $K = \mathbf{C}$. To see that it also yields the simultaneous real orthogonal diagonalization of commuting real symmetric operators (or matrices), it is only necessary to note that in this case all the characteristic roots are real.

LEMMA. *All characteristic roots of a real symmetric operator are real.*

Proof. If A is symmetric and $p(X) = (X-a)^2 + b$ is a factor of the minimal polynomial of A there is a vector $v \neq 0$ such that $p(A)(v) = 0$. Then $0 \leqslant \langle (A-aI)(v), (A-aI)(v) \rangle = \langle (A-aI)^2(v), v \rangle = (-b) \langle v, v \rangle$ so $b \leqslant 0$ and p has real roots.

For real normal operators we need only consider a subspace V of the kind in the second statement of Proposition 4. If A is normal with minimal polynomal $(X-a)^2 + b^2 = p(X)$ on V, there is by the Corollary of Proposition 1 a minimal 2-dimensional subspace W of V stable under the commuting operators $\frac{A + A^*}{2}$ (symmetric) and $\frac{A - A^*}{2}$ (skewsymmetric). Relative to an orthonormal basis y_1, y_2 of W, the matrix of $\frac{A + A^*}{2}$ must be $\begin{pmatrix} a' & 0 \\ 0 & a' \end{pmatrix}$, the matrix of $\frac{A - A^*}{2}$ must be $\begin{pmatrix} 0 & -b' \\ b' & 0 \end{pmatrix}$, and so the matrix of A must be $\begin{pmatrix} a' & -b' \\ b' & a' \end{pmatrix}$. Then $p(X) = X^2 - 2a'X + (a')^2 + (b')^2$ so that $a = a'$ and $\pm b = b'$. Since W^1 is also stable under the symmetric operator $\frac{A + A^*}{2}$ and the skew-symmetric operator $\frac{A - A^*}{2}$, both W and W^1 are A-stable. It follows that V is the orthogonal sum of such subspaces W, that $\frac{A + A^*}{2} = aI$ on V, that $\left(\frac{A - A^*}{2}\right)^2 = -b^2 I$ on V, and that $P = (a^2 + b^2)^{-\frac{1}{2}} A$ is orthogonal on V. Since $P^* = P^{-1}$ is a polynomial in P, A^* is a polynomial in A and hence for every A-stable subspace W of V, W is A^*-stable and W^1 is A-stable. In particular, the Corollary of Proposition 1 gives a

2-dimensional S-stable subspace W_1 of V, and by the preceding argument $W_1^{\perp}$ is also S-stable. Thus V is the orthogonal sum of 2-dimensional S-stable subspaces $W_1, \ldots, W_r$. Let y_i, z_i be an orthonormal basis for W_i, then relative to the basis $y_1, z_1, \ldots, y_r, z_r$ the matrix of each A in S is block diagonal and the blocks are positive multiples of 2×2 rotation matrices. Thus we have determined the structure of real normal operators.

Finally, return to Proposition 2 and suppose that $y_1, \ldots, y_n$ is a fan (flag) basis for an S-stable fan (flag) and that E has an inner product. The Gram-Schmidt process applied to this basis yields an orthonormal fan (flag) basis $z_1, \ldots, z_n$ for the same S-stable fan (flag). Thus the matrix P in the Corollary can be taken to be unitary or real orthogonal. Moreover, if S contains A and A^* for some A (hence A is normal) we get directly that the matrices of A and A^* relative to the basis $z_1, \ldots, z_n$ are both (nearly) super-diagonal, and since one is the adjoint of the other, they are both diagonal (block diagonal with blocks at most 2×2 in size). This observation could be used to give another proof for Proposition 4 and the structure of real normal operators. In either case, the argument can be simplified a bit if S is assumed to be $*$-closed.

REFERENCES

Flanders, H. Methods of Proof in Linear Algebra. *American Mathematical Monthly*, January 1956, pp. 1-15.

Hoffman, K. and R. Kunze. *Linear Algebra.* Second edition, Prentice-Hall, 1971.

(Reçu le 16 octobre 1975)

Ladnor Geissinger

Mathematics Department
University of North Carolina
Chapel Hill, North Carolina 27514