ILLINOIS JOURNAL OF MATHEMATICS
Volume 28, Number 3, Fall 1984

WEAKLY TRANSITIVE MATRICES

BY

JOSÉ BARRÍA AND P. R. HALMOS[1]

Introduction. A set of matrices is *transitive* if every number occurs as the $\langle 1, 1\rangle$ entry in at least one of them. (This is a tentative definition; in what follows it will be modified and quantified.) Question: is the conjugate class of a matrix transitive? (The conjugate class of a square matrix A is the set of all matrices SAS^{-1} similar to A.) Answer: not always. Example: $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, and, more generally, $\begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix}$ for each α. Sharper question: can the conjugate class of a matrix ever be transitive, and, if so, when? Special case: is the conjugate class of $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ transitive?

The purpose of this paper is to raise and answer a suitably general form of these questions. A set of matrices (square matrices, of size n, say, with entries in an arbitrary field) will be called *weakly k-transitive* ($1 \leqslant k \leqslant n$) if every square matrix of size k occurs as the top left corner in at least one of them. (Strong transitivity is something else and will be studied on another occasion; it has to do not with $k \times k$ squares in a corner but with $n \times k$ rectangles on a side.) The main problem is to determine when the conjugate class of a matrix is weakly k-transitive. Since weak transitivity is the only kind that will be considered, and since conjugate classes are the only sets for which the question will be raised, we propose to express the problem in the following abbreviated form: when is a matrix k-transitive? Explicitly: for which $n \times n$ matrices A is it true that corresponding to every $k \times k$ matrix X ($1 \leqslant k \leqslant n$) there exists an invertible $n \times n$ matrix S such that SAS^{-1} is of the form $\begin{pmatrix} X & * \\ * & * \end{pmatrix}$? It is sometimes convenient to express the relation between SAS^{-1} and X by saying that X is the *compression* (k-compression) of SAS^{-1}. In that language the question becomes: for which A (of size n) does every X (of size k) occur as the compression of some matrix similar to A?

The question is suggested by facts and problems about the weak density of certain sets of operators on infinite-dimensional spaces (see for instance [1]), but it seems to be interesting and will be studied here in its own right as a part of pure linear algebra.

---

[1] Research supported in part by a grant from the National Science Foundation.

Received December 18, 1981.

Simple transitivity. If $k = 1$, the answer is easy: *the only matrices that are not* 1*-transitive are the scalars.* (This answers the questions in the first paragraph of the introduction.)

The core of the proof is the observation that if A is not a scalar, then there exists a vector f such that f and Af are linearly independent. (We identify each matrix of size n with the linear transformation it induces on the naturally corresponding n-dimensional coordinate space.) The observation is an exercise. It is easy, but not automatic; a small idea is needed. Since a suitable generalization is basic in the sequel, we do not leave the exercise to the reader but proceed to offer two proofs.

A direct proof attacks the contrapositive head on. To say that f and Af are always linearly dependent means that $Af = \lambda(f)f$ for all f. It follows that

$$\lambda(\alpha f)\alpha f = A(\alpha f) = \alpha Af = \alpha\lambda(f)f,$$

and hence that $\lambda(\alpha f) = \lambda(f)$ whenever $\alpha f \neq 0$; similarly,

$$\lambda(f + g)(f + g) = A(f + g) = Af + Ag = \lambda(f)f + \lambda(g)g,$$

and hence

$$\lambda(f) = \lambda(f + g) = \lambda(g)$$

whenever f and g are linearly independent. For each non-zero f and g, either g is a scalar multiple of f, or f and g are linearly independent, and, in either case, $\lambda(f) = \lambda(g)$. That is: the function λ is a constant, and therefore A is a scalar.

The alternative proof is completely conceptual; the price is the use of a powerful theorem. Recall that a matrix A of size n (linear transformation on an n-dimensional space) is *cyclic* if there exists a vector f such that the vectors f, Af, ..., $A^{n-1}f$ span the whole space. (Equivalently: they are linearly independent.) Clearly for every cyclic matrix A of size 2 or greater there exists a free vector (i.e., a vector f such that f and Af are linearly independent.) The powerful theorem is the assertion (a consequence of the theory of the rational canonical form) that every matrix is the direct sum of cyclic ones [2, Section 6.7].

Here then is the alternative proof. Write A as a direct sum of cyclic matrices. If any of the direct summands is of size 2 or greater, a free vector exists. If they are all of size 1, then A is diagonal. Since it is not a scalar, it has eigenvectors g and h corresponding to distinct eigenvalues; in that case the sum $f = g + h$ is free.

Granted now that there exists a free vector f for A, form a basis whose first two elements are f and Af. With respect to such a basis the first column of A becomes

$$\langle 0, 1, 0, \ldots, 0\rangle^t.$$

(The superscript stands for "transpose"; it is used for typographical convenience, to avoid the display of a column. Observe that if A is not a scalar, then the size of A is at least 2.) In other words A is similar to a matrix with the indicated first column, and there is, therefore, no loss of generality in assuming that A itself has that first column. Once that is assumed, let x be an arbitrary element of the underlying field and write

$$S = \begin{pmatrix} 1 & x & 0 & & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & 0 \\ & \vdots & & & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix},$$

note that S is invertible, with

$$S^{-1} = \begin{pmatrix} 1 & -x & 0 & & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & 0 \\ & \vdots & & & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix},$$

and vertify that the $\langle 1, 1 \rangle$ entry of SAS^{-1} is x.

Non-transitivity. It is not obvious how to generalize either the statement or the proof of the result of the preceding section to k-transitivity when $k > 1$. The assertion nearest the surface is a negative one: *if $n < 2k$, then, over an infinite field, no matrix of size n is k-transitive.*

Approach the proof by contradiction; assume that $n < 2k$ and A is a k-transitive matrix of size n. The k-transitivity of A implies in particular that A is similar to $\begin{pmatrix} 0 & * \\ * & * \end{pmatrix}$, where the size of the top left 0 is k. We proceed to derive two statements that turn out to be contradictory: (1) a property of all matrices of the form just indicated, and (2) a property of all k-transitive matrices.

(1) If $n < 2k$ and if a matrix A of size n has a zero compression of size k, then A is singular. Reason: A maps a subspace of dimension k into one of dimension $n - k$, and $n - k < k$. (Alternative proof: every term in the expansion of the determinant of A must be 0.)

(2) Every scalar translate of a k-transitive matrix is k-transitive. Proof: if A is k-transitive and X is an arbitrary matrix of size k, then, for each scalar λ, some conjugate of A has the compression $X + \lambda$, and, consequently, some conjugate of $A - \lambda$ has the compression X.

The statements (1) and (2) imply that under the present assumptions $\det(A - \lambda) = 0$ for all λ, which is impossible (since the underlying field is infinite).

Main theorem. The condition on a matrix A that ensures that every matrix of size 1 occurs as the compression of some conjugate of A, namely that A be non-scalar, can be expressed this way:

$$\operatorname{rank}(A - \lambda) \geqslant 1 \quad \text{for all } \lambda.$$

This suggests consideration of the condition

$$\operatorname{rank}(A - \lambda) \geqslant k \quad \text{for all } \lambda.$$

In geometric terms the latter condition says that, for each λ, the geometric multiplicity of λ as an eigenvalue of A is not more than $n - k$ (where n is, as always, the size of A). If k is 1, only the most "concentrated" matrices (the scalars) can fail to satisfy it; to increase k tends to have the effect of making A less like a scalar, more scattered. Proposed technical term: A is *k-scattered.*

Our principal result about transitivity is the following necessary and sufficient condition.

THEOREM. *A matrix A of size n over an infinite field is k-transitive if and only if $n \geqslant 2k$ and A is k-scattered.*

The condition is usable. Sample non-obvious application: a truncated shift

$$\begin{pmatrix} 0 & 0 & 0 & & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & & 0 & 0 \\ & \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & & 1 & 0 \end{pmatrix}$$

of size n is k-transitive if and only if $n \geqslant 2k$.

The necessity of the condition $n \geqslant 2k$ was proved in the preceding section. The necessity of the scattering condition is also easy to prove. Indeed: if A is k-transitive, then some conjugate of A has a compression that is invertible (e.g., the identity matrix of size k), and therefore rank $A \geqslant k$. Since both k-transitivity and k-scattering are invariant under scalar translation (for k-transitivity this was proved in the preceding section and for k-scattering it is obvious), the proof of necessity is complete.

Observe that the two parts of the condition (which between them are necessary and sufficient for k-transitivity) are almost independent of each other. Indeed: if $k = 1$ and

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

then $n = 2k$, but A is not 1-scattered; if $k = 2$ and

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

then A is 2-scattered, but $n < 2k$. The only implication along these lines is a trivial one: if A is 1-scattered, then its size must be at least 2.

The remainder of this paper is devoted to the proof that the condition stated in the theorem is sufficient as well as necessary; that is the part that takes the most work.

Free sets. With hindsight we can say that the sufficiency proof could have been discovered as a generalization of the case $k = 1$. The idea is, assuming that A is k-scattered, to think of A as a block matrix consisting of blocks of size k as far as possible, allowing some smaller pieces at the bottom and at the right. The major difficulty is to show that A is similar to a matrix whose first column (in the block sense) is

$$\langle 0, 1, 0, \ldots, 0\rangle'.$$

From there the last step is trivial. The similarity of a matrix with the indicated first column to a matrix with arbitrarily prescribed compression is formally the same as for $k = 1$: use the same transforming matrix S, but interpret the entries as blocks (matrices) rather than as elements of the field.

The problem of showing that A is similar to a matrix with "first" column $\langle 0, 1, 0, \ldots, 0\rangle'$ (in reality the symbol indicates the first k columns) is the same as the problem of showing that there exists a set

$$\{f_1, \ldots, f_k\}$$

of vectors such that the set

$$\{f_1, \ldots, f_k, Af_1, \ldots, Af_k\}$$

is linearly independent. Indeed, given such a set, extend it to a basis, and then express A as a matrix with respect to that basis. We propose to call such sets *free* (or, more precisely, A-free). If $k = 1$, then the singleton $\{f_1\}$ is free in this sense exactly when the vector f_1 is free in the sense used above. The rest of the proof assumes that $n \geqslant 2k$ and A is k-scattered, and shows that under these conditions a free set of size k always exists; that is exactly what the proof for $k = 1$ did.

If the matrix A happens to be cyclic, the conclusion is near at hand. Indeed, in that case (by definition) there exists a vector f such that the vectors

$$f, Af, \ldots, A^{n-1}f$$

form a basis; the desired free set consists of the first k vectors of the sequence

$$f, A^2f, A^4f, \ldots .$$

To see that, note that

$$\{f, A^2f, \ldots, A^{2(k-1)}f\} \subset \{f, Af, \ldots, A^{n-2}f\},$$

and hence that the union of the set and its image under A is included in a basis.

The reasoning in the cyclic case contains the germ of the idea in every case. The point is that since every matrix is the direct sum of cyclic ones, a free set of the appropriate size can be obtained by stitching together smaller free sets contributed by the direct summands. ("Stitching" will turn out to mean something different from just forming unions.) The cyclic argument above did not use the scattering assumption; that will be needed for a part of the stitching only.

Eigenvalues. The proof will be achieved by successive reductions to more and more simple cases; the first reduction embeds most of the visible eigenvalues into larger cyclic pieces.

Write A as a direct sum of cyclic matrices, and consider the ones of size 1; they are given by eigenvectors, which, from the point of view of the construction of free sets, are obstacles. How many distinct eigenvalues do the direct summands of size 1 correspond to? If the answer is more than one, choose an eigenvector corresponding to each eigenvalue and form the span of the chosen eigenvectors; the result is a cyclic direct summand of A of size 2 or greater. (It is cyclic because all its eigenvalues have multiplicity 1.) If the remaining direct summands of size 1 still correspond to more than one eigenvalue, the steps just described can be repeated. The process so begun can be continued so long as there remain direct summands of size 1 corresponding to more than one eigenvalue. When the process terminates, A is exhibited as the direct sum of a scalar multiple of the identity matrix and cyclic summands of size 2 or greater. If, moreover, the scalar in question is λ, we replace A by $A - \lambda$ and thus assume that the scalar summand is 0. The effect of the latter reduction is a small but helpful simplification of notation.

Stitching. After the preceding reductions A is the direct sum of a zero matrix (possibly absent) and cyclic matrices of size 2 or greater. The next step is to show how nullvectors (eigenvectors of the zero summand) can be "absorbed" by the cyclic summands. The word, used here as an informal abbreviation, means "to use a non-zero cyclic summand together with some nullvectors to constuct a large free set". Absorption in this sense is a part of the "stitching" promised before.

To be specific, let a vector f be the generator of one of the cyclic direct summands of A, of size m say, so that the vectors

$$f,\ Af,\ \ldots,\ A^{m-1}f$$

form a basis for an m-dimensional subspace invariant under A. Suppose to begin with that m is odd (so that in particular $m \geqslant 3$) and that g is a nullvector; note that the set $\{f, Af, \ldots, A^{m-1}f, g\}$ is linearly independent. Assertion: the set

$$f,\ A^2f,\ \ldots,\ A^{m-3}f,\ A^{m-2}f + g$$

is free. (If $m = 3$, the set consists of just f and $Af + g$.) Reason: the image of the set under A is

$$Af, A^3f, \ldots, A^{m-2}f, A^{m-1}f.$$

The span of the union contains all A^jf ($j = 0, 1, \ldots, m - 2$) as well as $A^{m-1}f$ and $A^{m-2}f + g$, and, therefore, it contains each of the $m + 1$ vectors $f, Af, \ldots, A^{m-1}f, g$. Since the union consists of $m + 1$ vectors and spans an $(m + 1)$-dimensional space, it is necessarily linearly independent.

This shows how to absorb one nullvector (if m is odd); it is just as easy to absorb any odd number of them, up to and including $m - 2$. If, for instance, $m = 9$ and three nullvectors are to be absorbed, then the way to make a free set (of the maximal size $\frac{1}{2}(9 + 3)$) out of the linearly independent set

$$f, Af, A^2f, A^3f, A^4f, A^5f, A^6f, A^7f, A^8f, g_1, g_2, g_3$$

is to form

$$f, A^2f, A^4f, A^5f + g_1, A^6f + g_2, A^7f + g_3.$$

(Use even powers of A at the beginning and the sums of all powers of A with given nullvectors at the end.) The proof of freedom is the same as in the preceding paragraph, and, in fact, the construction and the argument remain the same in general.

If the restriction of A to the cyclic subspace generated by f happens to be invertible, then the number of nullvectors that can be absorbed can be raised to m. If, for a typical instance, $m = 5$, then a free set (of the maximal size 5) in the span of

$$f, Af, A^2f, A^3f, A^4f, g_1, g_2, g_3, g_4, g_5$$

is given by

$$f + g_1, Af + g_2, A^2f + g_3, A^3f + g_4, A^4f + g_5.$$

The argument that proves freedom has one extra feature this time. The image of the set to be proved free is

$$Af, A^2f, A^3f, A^4f, A^5f.$$

The invertibility of A on the span of $\{f, Af, A^2f, A^3f, A^4f\}$ implies that f itself is in the span of that image, and hence that the span of the union of the set with its image contains all A^if's ($i = 0, \ldots, 4$) and all g_j's ($j = 1, \ldots, 5$); since the union consists of 10 ($=2m$) vectors and spans a space of dimension 10, it is necessarily linearly independent.

There is no real difference between the even m's and the odd ones; the facts and the proofs are the same.

There is one more stitching step. Suppose that two of the cyclic direct summands (invertible or not) are of odd size, $2p + 1$ and $2q + 1$ say. To "stitch" them together means, again, "to put them together so as to construct a large free set (of size $p + q + 1$ to be precise)". The process goes like this. Given vectors f and g that generate two direct summands, so that

$$f, Af, \ldots, A^{2p}f \quad \text{and} \quad g, Ag, \ldots, A^{2q}g$$

span disjoint invariant subspaces of dimensions $2p + 1$ and $2q + 1$, form the set

$$f, A^2f, \ldots, A^{2p-2}f, A^{2p-1}f + g, Ag, A^3g, \ldots, A^{2q-1}g.$$

The image under A is

$$Af, A^3f, \ldots, A^{2p-1}f, A^{2p}f + Ag, A^2g, A^4g, \ldots, A^{2q}g.$$

Verify that the span of the union contains all A^if's and all A^jg's; it follows that the $2(p + q + 1)$ vectors span a space of that dimension and are, therefore, linearly independent.

Counting. The proof will now be concluded by a careful count of the free sets that absorption and stitching can produce.

Suppose that p is the size of the "invertible part" of A (that is the total size of the invertible direct summands put together), q is the size of the "singular part" of A (that is the total size of the non-zero singular direct summands), and r is the size of the "zero part" of A (that is the size of the zero direct summand). Suppose, moreover, that there are m non-zero singular direct summands, of sizes $q_1, \ldots, q_m$. (The case $m = 0$ is not excluded.) Note that

$$n = p + q + r \quad \text{and} \quad q = \sum_{j=1}^{m} q_j.$$

Since the nullity of A is $r + m$ (a contribution of r from the zero part and a contribution of 1 from each of the m singular summands), it follows that

$$\operatorname{rank} A = p + q - m.$$

Nullvectors are a source of difficulty; what is the maximum number of them that can be absorbed? The invertible part can absorb at most p of them. A singular direct summand of size q_j can absorb at most $q_j - 2$, and, therefore, the number of nullvectors that the entire singular part can absorb is at most $\Sigma_{j=1}^{m}(q_j - 2) = q - 2m$. The "absorption maximum" of A is $p + q - 2m$.

If the zero part is small, smaller than the absorption maximum,

$$r < p + q - 2m,$$

then all (or, in case of parity conflict, all but one) of the nullvectors can be absorbed. The following special case is almost completely typical. Suppose that

$$p + (q_1 - 2) + (q_2 - 2) < r < p + (q_1 - 2) + (q_2 - 2) + (q_3 - 2),$$

and suppose that $r - (p + (q_1 - 2) + (q_2 - 2))$ is odd and $q_3, \ldots, q_m$ are all even. Let the invertible part absorb p nullvectors, and absorb

$$(q_1 - 2) + (q_2 - 2)$$

in the first two direct summands of the singular part. Of the remaining

$$r - (p + (q_1 - 2) + (q_2 - 2))$$

nullvectors the third singular direct summand can absorb all but one. It follows that there exists a free set of size

$$p + (q_1 - 1) + (q_2 - 1) + \frac{1}{2}(r - (p + (q_1 - 2) + (q_2 - 2)) + q_3 - 1) + \frac{1}{2}(q_4 + \cdots + q_m) = \frac{1}{2}(p + q + r - 1) = \left[\frac{n}{2}\right] \geqslant k.$$

The only possibility that the preceding almost completely typical case does not indicate is the one in which there is no early parity conflict, the nullvectors are all absorbed, but some of untouched direct summands (invertible or singular) are of odd size. In that case use stitching; the result will always be a free set of size $n/2$ or $(n - 1)/2$, depending on parity.

If the zero part is large, large enough to make full use of the absorption maximum, the proof is easy. Assume, indeed, that $r \geqslant p + q - 2m$. Absorb p nullvectors in the invertible part and $q - 2m$ in the singular part; the result is a free set of size

$$\frac{1}{2}((p + p) + (q + (q - 2m))) = p + q - m = \operatorname{rank} A \geqslant k.$$

Note the essential use of the scattering assumption.

The proof is complete.

REFERENCES

1. D. W. HADWIN, E. A. NORDGREN, HEYDAR RADJAVI and PETER ROSENTHAL, *Most similarity orbits are strongly dense,* Proc. Amer. Math. Soc., vol. 76 (1979), pp. 250–252.

2. I. N. HERSTEIN, *Topics in algebra,* Xerox, Lexington, 1975.

INSTITUTO VENEZOLANO DE INVESTIGACIONES CIENTIFICAS
CARACAS, VENEZUELA

INDIANA UNIVERSITY
BLOOMINGTON, INDIANA