Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Linear Algebra and its Applications 421 (2007) 63–68
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

Unitarily achievable zero patterns and traces of words in A and $A^{\star}$

Dragomir Ž. Đoković *, Charles R. Johnson

Department of Pure Mathematics, University of Waterloo, Waterloo, Ont., Canada N2L 3G1
Department of Mathematics, College of William and Mary, Williamsburg, VA 23187-8795, USA

Received 10 January 2006; accepted 3 March 2006
Available online 27 April 2006
Submitted by R. Guralnick

Abstract

Our primary objective is to identify a natural and substantial problem about unitary similarity on arbitrary complex matrices: which 0-patterns may be achieved for any given n-by-n complex matrix via some unitary similarity of it. To this end, certain restrictions on "achievable" 0-patterns are mentioned, both positional and, more important, on the maximum number of achievable 0's. Prior results fitting this general question are mentioned, as well as the "first" unresolved pattern (for 3-by-3 matrices!). In the process a recent question is answered.

A closely related additional objective is to mention the best known bound for the maximum length of words necessary for the application of Specht's theorem about which pairs of complex matrices are unitarily similar, which seems not widely known to matrix theorists. In the process, we mention the number of words necessary for small size matrices.

AMS classification: Primary 15A21, 15A57; Secondary 15A72

Keywords: Unitary similarity; Schur's triangularization theorem; Specht's criterion for unitary similarity

☆ The first author was supported in part by the NSERC Grant A-5285.
* Corresponding author. Address: Department of Pure Mathematics, University of Waterloo, Waterloo, Ont., Canada N2L 3G1.

E-mail addresses: djokovic@uwaterloo.ca, djokovic@hypatia.uwaterloo.ca (D.Ž. Đoković), crjohnso@math.wm.edu (C.R. Johnson).

0024-3795/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2006.03.002

1. Unitarily achievable zero patterns

We say that a collection of positions $\mathscr{C} = \{(i_1, j_1), \ldots, (i_k, j_k)\}$ in an n-by-n matrix is a 0-*pattern achievable by unitary similarity* (*achievable* for short) if for any n-by-n complex matrix A, there is a unitary matrix U such that $B = U^*AU = (b_{ij})$ satisfies

$$b_{i_t, j_t} = 0, \quad t = 1, \ldots, k.$$

Schur's theorem on unitary triangularization [8] says that both triangular 0-patterns

$$\{(i, j) : 1 \leqslant i < j \leqslant n\} \quad \text{and} \quad \{(i, j) : 1 \leqslant j < i \leqslant n\}$$

are achievable. Since a permutation matrix is unitary, if $\mathscr{C}$ is achievable and σ is a permutation then

$$\mathscr{C}_\sigma = \big\{(\sigma(i_1), \sigma(j_1)), \ldots, (\sigma(i_k), \sigma(j_k))\big\}$$

is also achievable. Thus any 0-pattern that is a subset of one permutable to a triangular pattern is achievable.

There are some clear restrictions on achievable patterns. Since the identity matrix is preserved by unitary similarity, no diagonal position is contained in an achievable set. If A is of size n but has an eigenvalue of algebraic multiplicity n and geometric multiplicity 1, then A cannot be similar (let alone unitarily similar) to a nontrivial direct sum. Thus, no pattern of the form

$$\mathscr{C}_S = (S \times S^c) \cup (S^c \times S)$$

can be achievable for any nonempty proper subset S of $\{1, 2, \ldots, n\}$.

Let M_n denote the space of n-by-n complex matrices. For any 0-pattern $\mathscr{C}$, let $V_{\mathscr{C}}$ denote the subspace of M_n consisting of all matrices having zero entries at the positions indicated by $\mathscr{C}$. Let us write $u \cdot x$ for uxu^{-1}, where $u \in U(n)$ and $x \in M_n$.

Our primary purpose in this section lay in

(1) Popularizing the general question of which patterns are achievable (of course, the maximal ones suffice), and
(2) Making a quantitative observation that limits the cardinalities of achievable patterns (see Theorem 1.4).

Remark 1.1. There are achievable patterns of cardinality $n(n-1)/2$ whose achievability does not follow from Schur's theorem. For example, for $n = 4$, it has recently been shown [14] (see also [2] for a shorter proof) that the pattern corresponding to a tridiagonal matrix is achievable. This result has been strengthened recently in [9] where the authors show, by modifying a proof from [2], that given any two 4-by-4 complex matrices x and y, there exists a unitary matrix u such that both matrices uxu^{-1} and uyu^{-1} are in upper Hessenberg form, i.e., have zero entries in the positions (3, 1), (4, 1) and (4, 2).

Remark 1.2. There are considerable differences between the problem that we have raised (complex field) and the corresponding one over the real field (real orthogonally achievable patterns for real matrices). In the case of Schur's theorem, a triangular pattern is not real achievable (i.e., by similarity with a real invertible matrix) because there exist real matrices having a nonreal eigenvalue. For $n = 4$, the tridiagonal pattern is not real orthogonally achievable [3]. Even for

$n = 3$, the pattern $\mathscr{C} = \{(1, 2), (2, 3), (3, 1)\}$ is *not* real orthogonally achievable (consider a skew-symmetric matrix of rank 2), while its achievability is still unresolved in the complex case. It is an easy exercise to show that this particular pattern is real achievable iff the coefficients of the characteristic polynomial $\lambda^3 - c_1\lambda^2 + c_2\lambda - c_3$ of the given 3-by-3 real matrix A satisfy the inequality $3c_2 \leqslant c_1^2$. This inequality can be rewritten as $\operatorname{tr}(A)^2 \leqslant 3\operatorname{tr}(A^2)$.

Remark 1.3. For any given 0-pattern $\mathscr{C}$, there is, of course, a well-defined subset, $U(n) \cdot V_{\mathscr{C}} = \{uV_{\mathscr{C}}u^{-1} : u \in U(n)\}$, from which $\mathscr{C}$ may be achieved by some unitary similarity. Of course, this subset, which may be a proper one, is also an interesting one to study. A yet more general problem is the one of 0-patterns among k matrices that are achievable by simultaneous unitary similarity. For which lists of k 0-patterns $\mathscr{P}_1, \ldots, \mathscr{P}_k$, and any k n-by-n complex matrices $A_1, \ldots, A_k$, is there a unitary matrix U such that U^*A_iU has the pattern P_i, $i = 1, \ldots, k$?

Our primary result shows that the number of positions in an achievable pattern is a maximum in the case of Schur's theorem mentioned earlier. This is proved below by using a simple dimension argument. The same argument was used previously, in a special case, in the recent paper [2].

Theorem 1.4. *If $\mathscr{C}$ is an achievable 0-pattern for n-by-n matrices, then*

$$|\mathscr{C}| \leqslant \frac{1}{2}n(n-1).$$

Moreover, there are achievable patterns of this cardinality.

Proof. Let us denote by L the submanifold of the unitary group $U(n)$ consisting of matrices having real first row. The projection from the unitary group to the unit sphere $U(n) \to S^{2n-1}$, which sends a unitary matrix to its first column is a fibration with base S^{2n-1} and fibre $U(n-1)$. Our manifold L is the pullback of this bundle by the inclusion map of the unit sphere $S^{n-1} \subseteq \mathbf{R}^n$ into the sphere $S^{2n-1} \subseteq \mathbf{C}^n$. As the dimension of $U(n)$ is n^2, it follows that the dimension of L is $n^2 - n$. (All dimensions in this proof are real dimensions.)

Moreover, we have $U(n) = LT$ (product in the unitary group), where T is the maximal torus of $U(n)$ consisting of the diagonal matrices. Note also that the dimension of T is n.

Let $d = n^2 - |\mathscr{C}|$. The real dimension of $V_{\mathscr{C}}$ is $2d$. Since the 0-pattern $\mathscr{C}$ is achievable (by hypothesis), we have

$$U(n) \cdot V_{\mathscr{C}} = M_n.$$

Since T preserves $V_{\mathscr{C}}$, we have

$$\begin{aligned} 2n^2 &= \dim M_n = \dim(U(n) \cdot V_{\mathscr{C}}) \\ &= \dim(LT \cdot V_{\mathscr{C}}) = \dim(L \cdot V_{\mathscr{C}}) \\ &\leqslant \dim L + \dim V_{\mathscr{C}} = (n^2 - n) + 2d. \end{aligned}$$

Hence, $|\mathscr{C}| = n^2 - d \leqslant n(n-1)/2$. □

This theorem allows us to answer a related question.

Corollary 1.5. *There exist n-by-n complex matrices for which no unitarily similar matrix has more than $n(n-1)/2$ entries equal to* 0.

Proof. There are only finitely many 0-patterns $\mathscr{C}$ with $|\mathscr{C}| = 1 + n(n-1)/2$. For each of these patterns, the set $U(n) \cdot V_{\mathscr{C}}$ is a closed subset of the matrix space M_n of real dimension $<2n^2$. Hence, this finite family of subsets of M_n cannot cover M_n. □

Remark 1.6. The above corollary answers a question that has been raised by others, e.g. by X. Zhan at the 5th China Matrix Theory Conference, Shanghai, China, in August 2002 and again mentioned by him at the 12th ILAS Conference, Regina, Canada, in June 2005 [24]. It would be of interest to find, for each $n > 2$, a concrete n-by-n matrix having the property mentioned in the corollary.

2. The length of words bound for Specht's theorem

The objective of this section is to bring to the attention of the linear algebra community some recent advances concerning the old problem of unitary similarity of complex matrices.

Let us recall Specht's theorem [22].

Theorem 2.1. *Two complex n-by-n matrices x and y are unitarily similar iff the equalities* $\mathrm{tr}(w(x, x^*)) = \mathrm{tr}(w(y, y^*))$ *hold for all (finite) words w in two letters.*

This theorem would be of little practical value without a concrete upper bound on the length of the words for which the equalities mentioned in the theorem have to be checked. The first such bound was produced by Pearcy [16]:

Theorem 2.2. *In Specht's theorem, it suffices to verify the equalities* $\mathrm{tr}(w(x, x^*)) = \mathrm{tr}(w(y, y^*))$ *for words w of length $\leqslant 2n^2$.*

Pearcy's bound is well known. For instance it is given in Kaplansky's book [10, p. 72], the book of Horn and Johnson [8, Theorem 2.2.8], as well as in the survey on unitary similarity of matrices by Shapiro [19].

There is a sharper bound which follows from Nagata–Higman theorem and its strengthening by Razmyslov [18]. For more information about this theorem we refer the reader to the recent book [4, Part A, Chapter 6]. Let us state the result of Razmyslov.

Theorem 2.3. *In any associative algebra over a field of zero characteristic in which the identity $y^n = 0$ holds, the identity*

$$x_1 x_2 \cdots x_{n^2} = 0$$

also holds.

It was observed by Procesi [17] that the bound in Nagata–Higman theorem can be used to obtain a bound on the degrees of a set of generators of simultaneous polynomial GL_n-invariants of two (or more) n-by-n matrices. For more details and the proof of this fact we refer the reader to one of the following references [6,17], or [4, Theorem 6.1.6, p. 78].

It is a well known fact that a complex polynomial, $p(x, y)$, in the entries of two complex matrices x and y, is (a simultaneous) $\mathrm{GL}_n(\mathbf{C})$-invariant iff the corresponding polynomial $p(z, z^*)$ in the real and imaginary parts of the entries of the complex matrix z is a $U(n)$-invariant. For instance this is proved in the book of Sibirskiǐ [21, Theorem 2.33, p. 193] as well as in Procesi's paper [17, Section 11].

Finally, since $U(n)$ is a compact group, the polynomial $U(n)$-invariants separate the unitary similarity classes in M_n.

Hence, from these remarks one obtains the following bound.

Theorem 2.4. *In Specht's theorem, it suffices to verify the equalities* $\operatorname{tr}(w(x, x^*)) = \operatorname{tr}(w(y, y^*))$ *for words* w *of length* $\leqslant n^2$.

Kuzmin [11] has constructed examples showing that $d(n) \geqslant n(n+1)/2$, where $d(n)$ denotes the best bound for the Nagata–Higman theorem. He has conjectured that in fact equality holds. This conjecture is true for $n \leqslant 4$, and also for $n = 5$ assuming that the algebra is generated by two elements [20].

Subsequently, by using a theorem of Paz [15], Laffey [12] lowered the bound n^2 in Theorem 2.4 to

$$\left\lceil \frac{2}{3}(n^2+2) \right\rceil,$$

where $\lceil x \rceil$ denotes the smallest integer $\geqslant x$. (This bound is cited incorrectly in the survey paper [19, p. 149].) A much better bound,

$$n \cdot \sqrt{\frac{2n^2}{n-1} + \frac{1}{4}} + \frac{n}{2} - 2$$

has been obtained recently by Pappacena [13] (see also [7]). Note that this bound is significantly lower than $n(n+1)/2$, i.e., the bound that one can expect to obtain from an improvement of Razmyslov's theorem. It has been conjectured [7] that Pappacena's bound can be replaced by one which is linear in n.

Bhattacharya and Mukherjea [1] have shown that there exist n^2+1 continuous $U(n)$-invariant functions $M_n \to \mathbf{C}$, which separate the unitary similarity classes in M_n. It is an interesting question to decide whether these functions can be chosen to be polynomials.

Let us also comment on the small values of n. For $n = 2$, in order to verify that two matrices are unitarily similar, it suffices to check the equality of the traces of only three words: x, x^2, xx^*. The following result of Pearcy is often quoted (see e.g. [8, Theorem 2.2.8, p. 76]): for $n = 3$ it suffices to check the equality of the traces of the following nine words:

$$x, x^2, xx^*, x^3, x^2x^*, x^2(x^*)^2, (xx^*)^2, x^2x^*xx^*, x^2(x^*)^2xx^*.$$

The fact that this list contains two redundant words, $(xx^*)^2$ and $x^2x^*xx^*$, is not so well known. This fact was proved by Sibirskiǐ, see his book [21, Theorem 3.45, Corollary 1.45, p. 260]. He also showed that the remaining set of seven words is minimal, i.e., it contains no redundant word.

The algebra of simultaneous polynomial $\mathrm{GL}_4(\mathbf{C})$-invariants of two complex matrices x and y has a minimal set of generators consisting of 32 polynomials (see [5,23]). All of these polynomials can be taken to be of the form $\operatorname{tr}(w(x, y))$, where w is a word in two letters. These 32 words have lengths $\leqslant$10 and only one of them has length 10. There is a subset consisting of only 20 words which suffices for testing the unitary similarity of 4-by-4 complex matrices. This will be elaborated upon in a separate paper (under preparation) by the first author.

Acknowledgments

The authors thank K.R. Davidson for his valuable comments on a preliminary version, and R. Guralnick for his comments on the second part of the note.

References

[1] R. Bhattacharya, K. Mukherjea, On unitary similarity of matrices, Linear Algebra Appl. 126 (1989) 95–105.
[2] K. Davidson, D.Ž. Đoković, Tridiagonal forms in low dimensions, Linear Algebra Appl. 407 (2005) 169–188.
[3] D.Ž. Đoković, M.L. MacDonald, Orthogonal invariants of a matrix of order four and applications, J. Pure Appl. Algebra 202 (2005) 259–283.
[4] V. Drensky, E. Formanek, Polynomial Identity Rings, Advanced Courses in Mathematics CRM Barcelona, Birkhäuser, Basel–Boston–Berlin, 2004.
[5] V. Drensky, L. Sadikova, Generators of invariants of two 4×4 matrices, arXiv:math.RA/0503146 v1 8 March 2005.
[6] E. Formanek, The Polynomial Identities and Invariants of $n \times n$ Matrices, Regional Conference Series in Mathematics No. 78, Amer. Math. Soc. 1991.
[7] A. Freedman, R.N. Gupta, R.M. Guralnick, Shirshov's theorem and representations of semigroups, Pacific J. Math. 18 (3) (1997) 159–176.
[8] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, 1985.
[9] J. Holbrook, J.-P. Schoch (with an appendix by T. Košir and B. Al Sethuraman), Moving zeros among matrices, preprint, 19pp.
[10] I. Kaplansky, Linear Algebra and Geometry, A second course, Allyn and Bacon, Boston, 1969.
[11] E.N. Kuzmin, On the Nagata–Higman theorem, in: Mathematical Structures, Computational Mathematics, Mathematical Modeling. Proc. Dedicated to the 60-th birthday of Acad. L. Iliev, Sofia, 1975, pp. 101–107.
[12] T.J. Laffey, Simultaneous reduction of sets of matrices under similarity, Linear Algebra Appl. 84 (1986) 123–138.
[13] C.J. Pappacena, An upper bound for the length of a finite-dimensional algebra, J. Algebra 197 (1997) 535–545.
[14] V. Pati, Unitary tridiagonalization in $M_4(\mathbf{C})$, Proc. Indian Acad. Sci. 111 (4) (2001) 381–397.
[15] A. Paz, An application of the Cayley–Hamilton theorem to matrix polynomials in several variables, Linear and Multilinear Algebra 15 (1984) 161–170.
[16] C. Pearcy, A complete set of unitary invariants for operators generating finite W^*—algebras of type I, Pacific J. Math. 12 (1962) 1405–1414.
[17] C. Procesi, The invariant theory of $n \times n$ matrices, Adv. Math. 19 (1976) 306–381.
[18] Yu.P. Razmyslov, Trace identities of full matrix algebras over fields of zero characteristic, Izv. Akad. Nauk SSSR, Ser. Math. 38 (1974) 723–756.
[19] H. Shapiro, A survey of canonical forms and invariants for unitary similarity, Linear Algebra Appl. 147 (1991) 101–167.
[20] I. Shestakov, N. Zhukavets, On associative algebras satisfying the identity $x^5 = 0$, Algebra Discrete Math. (1) (2004) 112–120.
[21] K.S. Sibirskiĭ, Algebraic Invariants of Differential Equations and Matrices (in Russian), Stiintsa, Kishinev, 1976.
[22] W. Specht, Zur Theorie der Matrixen II, Jahresber. Deutsche Math. Ver. 50 (1940) 19–23.
[23] Y. Teranishi, The ring of invariants of matrices, Nagoya Math. J. 104 (1986) 149–161.
[24] X. Zhan, private communication, July 2005.