Operators and Matrices
Volume 1, Number 1 (2007), 113-119

UNIVERSAL ZERO PATTERNS FOR SIMULTANEOUS SIMILARITY OF SEVERAL MATRICES

DRAGOMIR Ž. ĐOKOVIĆ

(*communicated by Leiba Rodman*)

Abstract. Our field F is algebraically closed. Let M_n be the space of $n\times n$ matrices over F. If $(X_1,\dots,X_s)\in M_n^s$ and $X_k=\left[x_{ij}^{(k)}\right]$, we say that the triple (i,j,k) labels the entry $x_{ij}^{(k)}$. If P is a collection of labels (i,j,k), then M_P denotes the subspace of M_n^s consisting of $(X_1,\dots,X_s)\in M_n^s$ such that $x_{ij}^{(k)}=0$ for all $(i,j,k)\in P$. We present a method of proving that certain patterns P, with $|P|=n(n-1)/2$, are universal in the sense that for every s-tuple $(X_1,\dots,X_s)\in M_n^s$ there exists $S\in \mathrm{GL}_n(F)$ such that $(SX_1S^{-1},\dots,SX_sS^{-1})\in M_P$. We demonstrate the power of our method on several examples from recent literature.

1. Introduction

The many unresolved questions about simultaneous similarity of matrices and simultaneous unitary similarity of several complex matrices continue to attract the interest of researchers, see recent papers [2, 3, 7, 8] on these topics. Some other important references should be mentioned: Friedland's paper [4] on the problem of simultaneous similarity of matrix pairs, Shapiro's survey paper [10] on unitary similarity of complex matrices, and the recent book of Radjavi and Rosenthal [9] on simultaneous triangularization. The particular problem that we consider here has been raised in [7] (see also [3]), which served as a main motivation for this paper. Let us point out that the main objective of [7] was (and we quote) "to investigate the zero-patterns that can be created by unitary similarity in a given matrix, and the zero-patterns that can be created by simultaneous unitary similarity in a given sequence of matrices".

Throughout the paper n and s will denote fixed positive integers, F an algebraically closed field, and M_n the space of $n\times n$ matrices over F. An s-*matrix* is a sequence $A=(A_1,A_2,\dots,A_s)\in M_n^s$. We are interested in the diagonal conjugation action (also known as simultaneous similarity) of $G=\mathrm{GL}_n(F)$ on M_n^s:

$$SAS^{-1}=(SA_1S^{-1},SA_2S^{-1},\dots,SA_sS^{-1}).$$

Let Σ denote the set of ordered triples (i,j,k) of integers such that $i,j\in\{1,2,\dots,n\}$ and $k\in\{1,2,\dots,s\}$. If $i=j$, we say that the triple $(i,i,k)\in\Sigma$

Mathematics subject classification (2000): 15A21, 14L35.
Key words and phrases: Simultaneous similarity of matrices, unitary similarity, zero patterns.
The author was supported by the NSERC Grant A-5285.

Paper No. 01-07

is a *diagonal triple*. By a *zero pattern*, or just *pattern*, we mean a subset of Σ. For each pattern P we define its *complementary pattern* $P^c = \Sigma \setminus P$. Given a pattern P, we denote by M_P the subspace of M_n^s consisting of all s-matrices $A = (A_1, \dots, A_s)$ such that the (i,j) th entry of A_k is 0 if $(i,j,k) \in P$. For each pattern P, we have the direct decomposition $M_n^s = M_P \oplus M_{P^c}$. Let $\mathrm{pr}_P : M_n^s \to M_{P^c}$ be the corresponding projection with $\ker(\mathrm{pr}_P) = M_P$.

DEFINITION 1.1. For any pattern P we define the map $\varphi_P : G \times M_P \to M_n^s$ by $\varphi_P(S, Y) = SYS^{-1}$. We say that P is *universal* if φ_P is surjective, i.e., if for each $A \in M_n^s$ there exists $S \in G$ such that $SAS^{-1} \in M_P$.

Let us remark that if P contains a diagonal triple, then P is not universal.

The problem of deciding, in general, which patterns are universal is beyond our reach. We shall consider only some special patterns which were introduced in [7] under the name of "upper-form patterns". We say that a pattern P is an *upper pattern* if whenever $(i,j,k) \in P$ and $(i',j',k) \in \Sigma$ satisfy the inequalities $i' \geqslant i$ and $j' \leqslant j$, then $(i',j',k) \in P$.

Let B denote the Borel subgroup of G consisting of the invertible upper triangular matrices, and $\mathfrak{b}$ its Lie algebra consisting of all upper triangular matrices. We denote by $\mathfrak{n} \subseteq M_n$ the subspace of strictly lower triangular matrices. Thus, $M_n = \mathfrak{b} \oplus \mathfrak{n}$. It is easy to verify that if P is an upper pattern, then the subspace M_P is B-stable, i.e., $SM_PS^{-1} = M_P$ for all $S \in B$.

Note that if P is an upper pattern, with $|P| > n(n-1)/2$, then P is not universal for dimension reason. Indeed, let us consider the space M_P as a left B-module and form the associated fibre bundle $G \times_B M_P$ (see [1, Section 5]). This fibre bundle is the geometric quotient of $G \times M_P$ for the B-action $b \cdot (S, X) = (Sb^{-1}, bXb^{-1})$. The map $\varphi_P : G \times M_P \to M_n^s$ induces a morphism $\psi_P : G \times_B M_P \to M_n^s$, and we have

$$\dim(G \times_B M_P) = \dim(\mathfrak{n}) + sn^2 - |P| < sn^2.$$

Consequently, ψ_P (and also φ_P) is not surjective, i.e., P is not universal.

Let us say that an upper pattern P is *feasible* if it contains no diagonal triples and $|P| = n(n-1)/2$. Our main result, Theorem 2.4, gives a simple characterization of the universal feasible patterns. This is proved in the next section. In Section 3. we demonstrate the power of our main result by giving quick uniform proofs for most of the results proved in the recent paper [7]. In particular, Example 3.2 provides a simple proof of a theorem of Pati [8] which asserts that every 4×4 complex matrix is unitarily similar to a tridiagonal matrix, see [7] for details.

Let I_n denote the identity matrix in M_n. For $X, Y \in M_n$ let $[X, Y] = XY - YX$. All topological notions refer to the Zariski topologies of M_n^s and G.

We thank John Holbrook for sending us the preprint [7]. We also thank the referee for his suggestion to include more details for the convenience of the reader (in particular, on the short proof of Pati's theorem).

2. Main result and proof

From now on P will denote a fixed feasible pattern and, to simplify notation, we write φ instead of φ_P. We denote by $\mathrm{d}\varphi_{(S_0,X)}$ the differential of φ at the point $(S_0,X) \in G \times M_P$, where $X = (X_1,\dots,X_s) \in M_P$. The tangent space of $G \times M_P$ at the point (I_n,X) can be identified with the product $M_n \times M_P$. We also identify the tangent space of M_n^s at the point $\varphi(I_n,X) = X$ with M_n^s. With these identifications in force, this differential is a linear map

$$\mathrm{d}\varphi_{(I_n,X)} : M_n \times M_P \to M_n^s. \tag{2.1}$$

The image of the tangent vector $(Z,U) \in M_n \times M_P$ is given by

$$\mathrm{d}\varphi_{(I_n,X)}(Z,U) = ([Z,X_1],\dots,[Z,X_s]) + U \in M_n^s. \tag{2.2}$$

Since $([Z,X_1],\dots,[Z,X_s]) \in M_P$ for $Z \in \mathfrak{b}$, it follows that

$$\mathrm{d}\varphi_{(I_n,X)}(\mathfrak{b} \times M_P) = M_P. \tag{2.3}$$

The following two lemmas are needed for the proof of our main result.

LEMMA 2.1. *If the differential $\mathrm{d}\varphi_{(S_0,X)}$ is surjective for some point $(S_0,X) \in G \times M_P$, then the differential $\mathrm{d}\varphi_{(I_n,X)}$ is also surjective.*

Proof. Define the invertible morphisms $\alpha : G\times M_P \to G\times M_P$ and $\beta : M_n^s \to M_n^s$ by $\alpha(S,Z) = (S_0^{-1}S,Z)$ and $\beta(Y) = S_0YS_0^{-1}$, respectively. Since $\beta\circ\varphi\circ\alpha = \varphi$, we have

$$\mathrm{d}\beta_X \circ \mathrm{d}\varphi_{(I_n,X)} \circ \mathrm{d}\alpha_{(S_0,X)} = \mathrm{d}\varphi_{(S_0,X)}.$$

Now the assertion of the lemma follows from the fact that α and β (and their differentials) are isomorphisms. □

The second lemma is due to R. Steinberg [11, Lemma 2, p. 68] (see also [5, Corollary 3.2.12 (a)]).

LEMMA 2.2. *Let G be an affine algebraic group acting on a variety V. Let H be a closed subgroup of G and $U \subseteq V$ be a closed subset of V, invariant under the action of H. Assume that the homogeneous space G/H is a complete variety. Then $G\cdot U$ is closed.*

For $u,v \in \{1,2,\dots,n\}$ let $E_{uv} \in M_n$ be the matrix with (u,v)th entry 1 and all other entries 0. The matrices E_{uv} for all $u,v \in \{1,\dots,n\}$ form a basis of M_n to which we refer as the *standard basis* of M_n. One defines similarly the standard basis of M_n^s.

For each $X = (X_1,\dots,X_s) \in M_P$, denote by W_X the subspace of M_{P^c} spanned by the s-matrices

$$\mathrm{pr}_P\left([E_{uv},X_1],\dots,[E_{uv},X_s]\right), \quad n \geqslant u > v \geqslant 1.$$

By setting $U = 0$ and by taking $Z \in \mathfrak{n}$ in (2.2), we obtain that

$$W_X = \mathrm{pr}_P(\mathrm{d}\varphi_{(I_n,X)}(\mathfrak{n})).$$

Let us define the linear map $f_{P,X} : \mathfrak{n} \to M_{P^c}$ by

$$f_{P,X}(Z) = \mathrm{pr}_P([Z,X_1],\dots,[Z,X_s]).$$

Clearly, W_X is just the image of $f_{P,X}$.

DEFINITION 2.3. We say that a feasible pattern P is *regular* if $W_X = M_{P^c}$ for some $X \in M_P$.

We remark that the condition $W_X = M_{P^c}$ is equivalent to the statement that the differential (2.1) is surjective.

Our main result is the following theorem.

THEOREM 2.4. *Let P be a feasible pattern. If P is regular then it is universal. If the characteristic of F is 0, then the converse holds.*

Proof. To prove the first assertion, assume that P is regular. Since P is feasible, the subspace $M_P \subseteq M_n^s$ is B-stable. As G/B is a projective variety (and hence complete), we can apply Lemma 2.2 to deduce that the set

$$G \cdot M_P = \left\{SXS^{-1} : S \in G,\, X \in M_P\right\}$$

is closed.

As P is regular, we may assume that $X = (X_1, \dots, X_s) \in M_P$ is chosen so that $W_X = M_{P^c}$. We claim that the differential $\mathrm{d}\varphi = \mathrm{d}\varphi_{(I_n,X)}$ is surjective. Since M_P is contained in the image of $\mathrm{d}\varphi$ and

$$\mathrm{d}\varphi(Z,0) = ([Z,X_1], \dots, [Z,X_s])\,, \tag{2.4}$$

the vector

$$\mathrm{pr}_P\left([Z,X_1], \dots, [Z,X_s]\right)$$

belongs to the image of $\mathrm{d}\varphi$. As $W_X = M_{P^c}$, we conclude that the image of $\mathrm{d}\varphi$ contains M_{P^c}. Since it also contains M_P, our claim is proved.

By the differential criterion for dominance (see e.g. [5, Proposition 1.4.15]), φ is a dominant map. Since its image is also closed, φ must be surjective. This means that the pattern P is universal.

To prove the second assertion, assume that the characteristic of F is 0 and that P is universal. Consequently, the map φ is surjective. By [6, Proposition 14.4], there exists a point $(S_0, X) \in G \times M_P$ such that the differential $\mathrm{d}\varphi_{(S_0,X)}$ is surjective. By Lemma 2.1, the differential $\mathrm{d}\varphi_{(I_n,X)}$ is also surjective. From (2.3) and $M_n = \mathfrak{b} \oplus \mathfrak{n}$ we infer that the vectors (2.4) for $Z \in \mathfrak{n}$ span a complement of M_P in M_n^s. Hence, $W_X = M_{P^c}$ which means that the pattern P is regular. $\square$

3. An application and examples

In this section we apply our main result to unitary similarity and discuss several examples.

When $F = \mathbf{C}$ (the field of complex numbers), we can apply the main theorem to the action of the complex unitary group $\mathrm{U}(n)$ on M_n^s by simultaneous similarity.

PROPOSITION 3.1. *Let $F = \mathbf{C}$ and let P be a regular pattern. Then for any $A \in M_n^s$ there exists $R \in \mathrm{U}(n)$ such that $RAR^{-1} \in M_P$.*

Proof. By Theorem 2.4 there exists $S \in \mathrm{GL}_n(\mathbf{C})$ such that $SAS^{-1} \in M_P$. Now it suffices to apply Remark 1 in [7]. For the reader's convenience we repeat this short argument. Recall that S can be factored as $S = TR$, where $R \in \mathrm{U}(n)$ and $T \in B$. It follows that $RAR^{-1} \in T^{-1}M_PT = M_P$. $\square$

We shall now give two examples of regular patterns. The first example gives another proof of Proposition 2 (and Theorem A.1) in [7].

EXAMPLE 3.2. Let $n = 4$ and $s = 2$. Then the (upper Hessenberg) pattern

$$P = \{(3,1,1),(4,1,1),(4,2,1),(3,1,2),(4,1,2),(4,2,2)\}$$

is regular. Indeed, if we choose $X = (X_1, X_2) \in M_P$, where $X_1 = E_{32} + E_{43}$ and $X_2 = E_{12} + E_{21}$, then an easy computation shows that $W_X = M_{P^c}$.

It is important to point out that this example indeed gives a proof of Pati's theorem mentioned in the introduction. Take $F = \mathbf{C}$, $n = 4$, $s = 2$, and let $A \in M_4$ be arbitrary. Since the pattern P from the example above is regular, Proposition 3.1 shows that there exists $R \in \mathrm{U}(4)$ such that $R(A, A^*)R^{-1} \in M_P$. This forces RAR^{-1} to be tridiagonal.

It is very easy to handle Example 3 from [7].

EXAMPLE 3.3. Let $n = 3$ and $s = 3$. Then the (upper Hessenberg) pattern

$$P = \{(3,1,1),(3,1,2),(3,1,3)\}$$

is regular. Indeed, one can easily verify that for $X = (E_{11}, E_{21}, E_{32}) \in M_P$, we have $W_X = M_{P^c}$.

The analogously defined pattern for $n = 4$ and $s = 6$ is not universal. This follows from the following proposition.

PROPOSITION 3.4. *Let F have characteristic 0 and let P be a feasible pattern. Assume that there exist integers u and v such that $n \geqslant u > v \geqslant 1$ and $u > i$ and $v > j$ for all triples $(i,j,k) \in P$. Then P is not universal.*

Proof. It suffices to observe that $f_{P,X}(E_{uv}) = 0$ for all $X \in M_P$. $\square$

Next we consider the non-universal pattern from Example 4 of [7].

EXAMPLE 3.5. Let F have characteristic 0 and let $n = 4$ and $s = 3$. Then the feasible pattern

$$P = \{(3,1,1),(4,1,1),(3,1,2),(4,1,2),(3,1,3),(4,1,3)\}$$

is not universal. Indeed, it is easy to check that for any $X = (X_1, X_2, X_3) \in M_P$,

$$f_{P,X}(E_{43}) = \mathrm{pr}_P\left([E_{43}, X_1], \ldots, [E_{43}, X_3]\right) = 0.$$

Consequently, W_X is a proper subspace of M_{P^c} for all $X \in M_P$, i.e., the pattern P is not regular. By Theorem 2.4, P is not universal.

Let us give yet another very simple example of a feasible pattern which is not universal.

EXAMPLE 3.6. Let $n = 4$ and $s = 2$. Then the feasible pattern

$$P = \{(2,1,1),(3,1,1),(4,1,1),(4,1,2),(4,2,2),(4,3,2)\}$$

is not universal.

Choose $A = (A_1, A_2) \in M_4^2$ such that both A_1 and A_2 have 4 distinct eigenvalues and no eigenvector of A_1 lies in any of the 4 three-dimensional invariant subspaces of A_2. It is immediate from these properties that there is no $S \in G$ such that $SAS^{-1} \in M_P$. Hence, P is not universal.

For an alternative proof (in case F has characteristic 0), observe that $f_{P,X}(E_{32}) = 0$ for all $X \in M_P$.

In the next example we indicate how one can deduce [7, Theorem A.4] from our result.

EXAMPLE 3.7. Let $n > 4$ and $s = 2$. Then the feasible pattern

$$P = \{(i,j,1) : i > j+1\} \cup \{(i,1,2) : i > 2\} \cup \{(n,2,2)\}$$

is universal.

We shall only outline the proof. Let $X = (X_1, X_2) \in M_P$, where

$$X_1 = E_{11} + E_{32} + E_{43} + \cdots + E_{n-2,n-3} + E_{n,n-1} + cE_{n,n}, \quad c \neq 0,1;$$

and $X_2 = E_{21} + E_{n-1,2}$. It suffices to show that the map $f_{P,X} : \mathfrak{n} \to M_{P^c}$ is an isomorphism. Let us order the standard basis of $\mathfrak{n}$ as follows:

$$\begin{aligned}
&E_{n,1},\ E_{n-1,1},\ \ldots,\ E_{21},\ E_{n-1,2};\\
&E_{n,2},\ E_{n,3},\ldots,\ E_{n,n-1};\\
&E_{n-1,3},\ E_{n-1,4},\ldots,\ E_{n-1,n-2};\\
&E_{n-2,2},\ E_{n-2,3},\ldots,\ E_{n-2,n-3};\\
&\vdots\\
&E_{42},\ E_{43};\\
&E_{32}.
\end{aligned}$$

Then one can verify that, for each k, $f_{P,X}$ maps the subspace of $\mathfrak{n}$ spanned by the first k basis vectors onto a k-dimensional standard subspace of M_{P^c}. (By a standard subspace we mean a subspace spanned by a subset of the standard basis.) We omit this tedious verification.

Finally, we mention an interesting problem related to our main result. For simplicity, assume that F has characteristic 0. Let P be a universal feasible pattern. The varieties $G \times_B M_P$ and M_n^s have the same dimension and the morphism $\psi_P : G \times_B M_P \to M_n^s$ induced by φ_P is surjective. For such maps the degree is well defined, see e.g. [6, Lecture 7, p. 80].

PROBLEM. Compute the degree of the morphism ψ_P, for any universal feasible pattern P.

REFERENCES

[1] K. BONGARTZ, *Some geometric aspects of representation theory*, Algebras and Modules I, (I. Reiten, S.O. Smalø, Ø. Solberg, editors), CMS Conf. Proc. **23** (1998), 1–27.

[2] K. R. DAVIDSON AND D. Ž. ĐOKOVIĆ, *Tridiagonal forms in low dimensions*, Linear Alg. Appl. **407** (2005), 169–188.

[3] D. Ž. ĐOKOVIĆ AND C. R. JOHNSON, *Unitarily achievable zero patterns and traces of words in A and A^**, Linear Alg. Appl. **421** (2007), 63–68.

[4] S. FRIEDLAND, *Simultaneous similarity of matrices*, Advances in Math. **50** (1983), 189–265.

[5] M. GECK, *An Introduction to Algebraic Geometry and Algebraic Groups*, Oxford University Press, 2003.

[6] J. HARRIS, *Algebraic Geometry*, A First Course, Springer–Verlag, New York, 1992.

[7] J. HOLBROOK AND J.-P. SCHOCH, *Moving zeros among matrices*, with an appendix by T. Košir and B.Al. Sethuraman, Linear Alg. Appl. (to appear).

[8] V. PATI, *Unitary tridiagonalization in $M(4,\mathbf{C})$*, Proc. Indian Acad. Sci. (Math. Sci.) **111** (2001), 381–397.

[9] H. RADJAVI AND P. ROSENTHAL, *Simultaneous Triangularization*, Springer–Verlag, New York, 2000.

[10] H. SHAPIRO, *A survey of canonical forms and invariants for unitary similarity*, Linear Alg. Appl. **147** (1991), 101–167.

[11] R. STEINBERG, *Conjugacy Classes in Algebraic Groups*, Notes by Vinay V. Deodhar, Lect. Notes in Math. vol. 366, Springer–Verlag, Berlin–New York, 1974.

(Received June 20, 2006)

Department of Pure Mathematics
University of Waterloo
Waterloo
Ontario, N2L 3G1
Canada
e-mail: djokovic@uwaterloo.ca