An Extremal Sparsity Property of the Jordan Canonical Form

Richard A. Brualdi[a], Pei Pei[b], Xingzhi Zhan[b1]

[a] Department of Mathematics, University of Wisconsin
Madison, WI 53706, USA

[b] Department of Mathematics, East China Normal University
Shanghai 200062, China

e-mail addresses: brualdi@math.wisc.edu (R.A. Brualdi),
math1303@hotmail.com (P. Pei), zhan@math.ecnu.edu.cn (X. Zhan)

To Richard Varga on the occasion of his 80th birthday

Abstract. We prove that among all the matrices that are similar to a given square complex matrix, the Jordan canonical form has the largest number of off-diagonal zero entries. We also characterize those matrices that attain this largest number.

AMS subject classifications. 15A21, 15A99

Key words. Jordan canonical form, similarity, sparsity, zero pattern

One of the central results in linear algebra (see e.g. [2] for a classical proof, and [1] for a largely combinatorial proof) is the Jordan canonical form theorem which states that every square complex matrix A is similar to a Jordan matrix

$$J(A) = \mathrm{diag}(J_{n_1}(\lambda_1), \ldots, J_{n_k}(\lambda_k))$$

[1]This author's research was supported by the NSFC grant 10571060.

where

$$J_{n_i}(\lambda_i) = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda_i & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_i \end{bmatrix}$$

is a Jordan block of order n_i, $i = 1, \ldots, k$. This Jordan canonical form $J(A)$ is unique up to permutations of the diagonal Jordan blocks (see [2]). We will use $J(A)$ to mean any of the Jordan canonical forms of A. The Jordan canonical form is a similarity invariant; indeed, two matrices are similar if and only if they have a common Jordan canonical form. Denote by $\mathcal{S}(A)$ the set of all complex matrices that are similar to A.

It seems that $J(A)$ has the simplest form among the matrices in $\mathcal{S}(A)$. But in what sense? Structure? Sparsity? One of the authors asked whether for any matrix A, $J(A)$ has the largest number of zero entries among all the matrices in $\mathcal{S}(A)$. The answer is no, as shown by the following example [3]:

$$A = \begin{bmatrix} 0 & 2 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad J(A) = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

A has 11 zero entries while $J(A)$ has 10 zero entries.

In this note we show that for any complex matrix A, $J(A)$ has the largest number of *off-diagonal* zero entries among all the matrices in $\mathcal{S}(A)$, and we characterize the matrices in $\mathcal{S}(A)$ that attain this largest number.

Throughout we use $\sigma(A)$ to denote the number of off-diagonal nonzero entries of a matrix A. The following purely combinatorial lemma is the basis of our analysis.

Lemma 1 *Let n, k be positive integers with $1 \le k \le n$. If a matrix A of order n satisfies $\sigma(A) \le n-k$, then there exists a permutation matrix P such that*

$$P^T A P = \operatorname{diag}(A_1, A_2, \ldots, A_k)$$

where A_j is square and non-void for $j = 1, \ldots, k$.

Proof. Let $A = [a_{ij}]$. With A we associate a graph G with vertex set $V = \{1, 2, \ldots, n\}$ where there is an edge between vertices i and j if and only if $i \neq j$, and $a_{ij} \neq 0$ or $a_{ji} \neq 0$. This graph G has p edges where $p \le \sigma(A) \le n-k$. We list the edges of G in some order $e_1, e_2, \ldots, e_p$. Let G_i be the graph with vertex set V and edges $\{e_1, e_2, \ldots, e_i\}$, $i = 0, 1, \ldots, p$. Thus G_0 has no edges, $G_p = G$, and G_i is obtained from G_{i-1} by including the new edge e_i, $i = 1, 2, \ldots, p$. It follows that G_i has at most one fewer connected component than G_{i-1}, since one edge can at best join together two connected components. Since G_0 has n (trivial) connected components, G_p has at least $n - p \ge n - (n-k) = k$ connected components $C_1, C_2, \ldots, C_{n-p}$. Let C_k' be the union of the connected components C_j with $j \ge k$. The principal submatrices $A_1, \ldots, A_{k-1}, A_k$ of A corresponding to $C_1, \ldots, C_{k-1}, C_k'$ satisfy the conclusion of the lemma. □

The following fact will be needed later.

Lemma 2 *Let*

$$A = \begin{bmatrix} a & x^T \\ 0 & B \end{bmatrix}$$

be a complex matrix of order n where B has order $n-1$. If $J(A)$ has only one Jordan block, then $J(B)$ has only one Jordan block.

Proof. The unique Jordan block of A must be $J_n(a)$. To the contrary, suppose $J(B)$ has at least two Jordan blocks. Then there exists a nonsingular matrix W such that $W^{-1}BW = \mathrm{diag}(B_1, B_2)$ where B_1 and B_2 are square and are of orders r and s respectively with $1 \le r, s \le n-2$, $r + s = n - 1$. Setting $E = \mathrm{diag}(1, W)$ we get

$$E^{-1}AE = \begin{bmatrix} a & y_1^T & y_2^T \\ 0 & B_1 & O \\ 0 & O & B_2 \end{bmatrix} = H \tag{1}$$

where $(y_1^T, y_2^T) = x^T W$ and $y_1 \in \mathbb{C}^r$, $y_2 \in \mathbb{C}^s$. Since H is similar to A, $J(H) = J_n(a)$. Denote by I an identity matrix whose order is taken from the context. From (1) we get

$$(H - aI)^{n-1} = \begin{bmatrix} 0 & y_1^T(B_1 - aI)^{n-2} & y_2^T(B_2 - aI)^{n-2} \\ 0 & (B_1 - aI)^{n-1} & O \\ 0 & O & (B_2 - aI)^{n-1} \end{bmatrix}. \tag{2}$$

This is a contradiction, because $(J_n(a) - aI)^{n-1} \ne O$ and H similar to $J_n(a)$ imply that $(H - aI)^{n-1} \ne O$ while, on the other hand, the matrix on the right side of (2) is the zero matrix. In fact, as B_1 and B_2 have all eigenvalues equal to a and the orders of B_1 and B_2 are at most $n-2$, $(B_1 - aI)^{n-2} = O$, $(B_2 - aI)^{n-2} = O$. Therefore, $J(B)$ has only one Jordan block. $\square$

A square complex matrix is called a *monomial matrix* if it has exactly one nonzero entry in each row and each column. Let Γ_n be the set of monomial matrices of order n. Then $M \in \Gamma_n$ if and only if there exist a permutation matrix P and a nonsingular diagonal matrix D such that $M = PD$ if and only if there exist a permutation matrix Q and a nonsingular diagonal matrix E such that $M = EQ$. Obviously, Γ_n is a multiplicative group. We now show that $J(A)$ has the greatest off-diagonal sparsity of all matrices similar to A.

Theorem 3 *Let A be a square complex matrix and $B \in \mathcal{S}(A)$. Then*

$$\sigma(B) \geq \sigma(J(A)). \tag{3}$$

Equality in (3) holds if and only if there exists a monomial matrix M such that

$$M^{-1}BM = J(A). \tag{4}$$

Proof. Suppose A and B are of order n and $J(A)$ has exactly k Jordan blocks so that $\sigma(J(A)) = n - k$. We denote by Π_n the set of permutation matrices of order n. To the contrary, suppose that $\sigma(B) < n - k$, i.e., $\sigma(B) \leq n - (k + 1)$. Note that $2 \leq k + 1 \leq n$. By Lemma 1, there exists a $P \in \Pi_n$ such that $P^T BP = \text{diag}(B_1, \ldots, B_{k+1})$ where each B_j is square and non-void. This implies that $J(B)$ has at least $k + 1$ Jordan blocks. This is a contradiction, since $B \in \mathcal{S}(A)$ implies that $J(B)$ and $J(A)$ have the same Jordan blocks, in particular, the same number k of Jordan blocks. Therefore, $\sigma(B) \geq n - k$.

Next we consider the equality case:

$$\sigma(B) = \sigma(J(A)). \tag{5}$$

The condition (4) obviously implies the equality (5). Conversely, suppose B satisfies (5). We will prove the existence of M that satisfies (4). We first prove the assertion for the case when $J(A)$ has only one Jordan block: $J(A) = J_n(a)$ with $\sigma(J(A)) = n-1$. We use induction on n. The case $n = 1$ is trivial, and we now assume that $n \geq 2$. Suppose the conclusion for one Jordan block holds for all matrices of order $n-1$. Since $\sigma(B) = \sigma(J(A)) = n - 1$, B has a column containing no off-diagonal nonzero entries. Thus, there exists a $P \in \Pi_n$ such that

$$P^T BP = \begin{bmatrix} a & x^T \\ 0 & B_1 \end{bmatrix} \tag{6}$$

where $x \in \mathbb{C}^{n-1}$ and B_1 is of order $n-1$. Since $J(B)$ has only one Jordan block, $x \neq 0$. When $n = 2$, x has only one (nonzero) component. If x has more than one nonzero component when $n \geq 3$, then $\sigma(B_1) \leq n-3$, as $\sigma(B) = n-1$. By Lemma 1, there exists a $Q \in \Pi_{n-1}$ such that $Q^T B_1 Q = \text{diag}(C_1, C_2)$. Hence $J(B_1)$ has at least two Jordan blocks. This is impossible by Lemma 2, since the Jordan canonical form of the matrix in (6) has only one Jordan block. So x has exactly one nonzero component. Now $\sigma(B_1) = n-2$ and $J(B_1)$ has only one Jordan block by Lemma 2. By the induction hypothesis there exists an $M_1 \in \Gamma_{n-1}$ such that $M_1^{-1} B_1 M_1 = J_{n-1}(a)$. Setting $M_2 = \text{diag}(1, M_1)$, from (6) we get

$$M_2^{-1} P^T BPM_2 = \begin{bmatrix} a & y^T \\ 0 & J_{n-1}(a) \end{bmatrix} = H \tag{7}$$

where $y^T = x^T M_1$ has exactly one nonzero component. Since H is similar to B, $J(H) = J_n(a)$. Note that $J_{n-1}(0)^{n-1} = O$. Thus by (7) we have

$$O \neq (H - aI)^{n-1} = \begin{bmatrix} 0 & y^T J_{n-1}(0)^{n-2} \\ 0 & O \end{bmatrix}.$$

So $0 \neq y^T J_{n-1}(0)^{n-2}$. This implies that the first component of y is nonzero, since the only nonzero entry of $J_{n-1}(0)^{n-2}$ is 1 at the position $(1, n-1)$. Hence $y = (y_1, 0, \ldots, 0)^T$, $y_1 \neq 0$. Now setting $M = PM_2 \text{diag}(y_1, I_{n-1})$, from (7) we get $M^{-1}BM = J_n(a)$. This proves the case when $J(A)$ has only one Jordan block.

Next suppose that $J(A)$ has k Jordan blocks, $k \geq 2$:

$$J(A) = \text{diag}(J_{n_1}(a_1), \ldots, J_{n_k}(a_k)).$$

Then $\sigma(B) = \sigma(J(A)) = n - k$. By Lemma 1, there exists a $Q_1 \in \Pi_n$ such that

$$Q_1^T B Q_1 = \mathrm{diag}(B_1, \ldots, B_k)$$

where each B_j is square. Since B is similar to A, $J(Q_1^T B Q_1)$ and $J(A)$ have the same Jordan blocks $J_{n_i}(a_i)$, $i = 1, \ldots, k$. Hence each $J(B_i)$ has only one Jordan block, $1 \le i \le k$, and there is a permutation γ of $1, \ldots, k$ such that $J(B_{\gamma(i)}) = J_{n_i}(a_i)$, $1 \le i \le k$. Letting $H_i = B_{\gamma(i)}$ we see that there exists a $Q_2 \in \Pi_n$ such that

$$Q_2^T Q_1^T B Q_1 Q_2 = \mathrm{diag}(H_1, \ldots, H_k). \tag{8}$$

Since H_i is of order n_i and $J(H_i)$ has only one Jordan block, Lemma 1 implies that

$$\sigma(H_i) \ge n_i - 1, \quad i = 1, \ldots, k.$$

We have

$$n - k = \sum_{i=1}^{k} \sigma(H_i) \ge \sum_{i=1}^{k} (n_i - 1) = n - k.$$

So $\sigma(H_i) = n_i - 1 = \sigma(J(H_i))$ for each $i = 1, \ldots, k$. Applying the theorem for the established case of matrices whose Jordan canonical forms have only one Jordan block to H_i, we deduce that there exists an $M_i \in \Gamma_{n_i}$ such that

$$M_i^{-1} H_i M_i = J_{n_i}(a_i), \quad i = 1, \ldots, k. \tag{9}$$

Let $M = Q_1 Q_2 \mathrm{diag}(M_1, \ldots, M_k)$. Then $M \in \Gamma_n$. By (8) and (9) we have

$$M^{-1} B M = \mathrm{diag}(J_{n_1}(a_1), \ldots, J_{n_k}(a_k)) = J(A).$$

This completes the proof. $\square$

Theorem 3 shows that up to permutation similarity, $J(A)$ is the unique zero-nonzero pattern among matrices in $\mathcal{S}(A)$ that attains the largest number of off-diagonal zero entries.

A referee asked whether a conclusion similar to that in Theorem 3 holds for the real Jordan canonical form (see e.g. [2]), that is, if A is a real, square matrix, does the real Jordan canonical form of A have the minimum number of off-diagonal nonzero entries among all *real* matrices in $\mathcal{S}(A)$? That this is not so is illustrated by the following example.

Let

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ -1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix}.$$

Then A is a matrix in real Jordan canonical form with eigenvalues $1 \pm i$, each of multiplicity 2. Let

$$T = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ where } T^{-1} = \begin{bmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Then

$$T^{-1}AT = \begin{bmatrix} 1 & 1 & 2 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix}.$$

We have

$$\sigma(A) = 6 > 5 = \sigma(T^{-1}AT).$$

Now suppose that A is a real, diagonalizable matrix of order n. Then the Jordan canonical form of A is a diagonal matrix. There is a nonnegative integer k such that $2k \leq n$, where A has $2k$ non-real eigenvalues in conjugate pairs and $n - 2k$ real eigenvalues. Each pair $a \pm bi, b \neq 0$ of conjugate, non-real eigenvalues contributes 2 off-diagonal nonzero entries to the real Jordan canonical form of A via the matrix

$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix},$$

and thus we deduce that the real Jordan canonical form of A has exactly $2k$ off-diagonal nonzero entries. Assume that B is a real matrix similar to A with $\sigma(B) \leq 2k-1$. Let B have irreducible components of orders $n_1, \ldots, n_p, n_{p+1}, \ldots, n_q$ where $n_1 + n_2 + \cdots + n_q = n$, and $n_1, \ldots, n_p \geq 2$ and $n_{p+1} = \cdots = n_q = 1$. Each irreducible component of B of order $n_i \geq 2$ contributes at least n_i to $\sigma(B)$. Hence $n_1 + \cdots + n_p \leq \sigma(B) \leq 2k-1$. Therefore B has at least $n-2k+1$ irreducible components of order 1. Since B is a real matrix, B and hence A have at least $n-2k+1$ real eigenvalues, a contradiction. Thus $\sigma(B) \geq 2k$, and *the real Jordan canonical form of a diagonalizable real matrix contains the largest number of off-diagonal zeros among all real matrices similar to A.* In contrast to the Jordan canonical form, there are cases of equality other than the type given in Theorem 3. For example, let

$$A = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Then the characteristic polynomial of A is $\lambda^4 + 1$, and it has four distinct, non-real eigenvalues. Thus A is diagonalizable over the complex number field, and the real Jordan canonical form B of A has four off-diagonal nonzeros, the same number as A. But clearly $M^{-1}AM \neq B$ for any monomial matrix M.

References

[1] R.A. Brualdi, The Jordan canonical form: an old proof, *Amer. Math. Monthly*, 94 (1987), 257–267.

[2] R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge, 1985, 121–127 and 150–153.

[3] C.K. Li, private communication, 2005.