90

The Combinatorial Structure of the Generalized Nullspace of a Block Triangular Matrix

Daniel Hershkowitz*[†]
Mathematics Department
Technion — Israel Institute of Technology
Haifa 32000, Israel

Uriel G. Rothblum
Faculty of Industrial Engineering and Management
Technion — Israel Institute of Technology
Haifa 32000, Israel

and

Hans Schneider[†‡]
Mathematics Department
University of Wisconsin — Madison
Madison, Wisconsin 53706

Submitted by Richard A. Brualdi

ABSTRACT

The combinatorial structure of the generalized nullspace of a block triangular matrix with entries in an arbitrary field is studied. Using an extension lemma, we prove the existence of a weakly preferred basis for the generalized nullspace. Independently, we study the height of generalized nullvectors. As a corollary we obtain the index theorem, which provides an upper bound for the index of a general matrix in terms of the indices of its diagonal blocks. We also investigate the case of equality in the index theorem.

*The research of this author was supported in part by the Technion VPR grant 100-708 (Japan TS Research Fund).

[†]The research of the first and the third author was supported also by their joint grant No. 85-00153 from the United States–Israel Binational Science Foundation (BSF), Jerusalem, Israel.

[‡]The research of this author was supported in part by NSF grant DMS-8521521.

655 Avenue of the Americas, New York, NY 10010 0024-3795/89/$3.50

INTRODUCTION

Beginning with Frobenius [2], many authors have investigated the combinatorial structure of a basis for the generalized eigenspace associated with the spectral radius of a (not necessarily irreducible) nonnegative matrix; see [10], [8], [7], [4], and the survey [11]. These results have been partially extended to other real eigenvalues of a nonnegative matrix, or equivalently to the real eigenvalues of a Z-matrix [12, 5]. In this paper we eliminate the restriction to nonnegativity and show that a somewhat weaker version of the combinatorial results on the structure of certain bases holds for matrices with entries in an arbitrary field. Thus, it is possible to apply, for example, our results to complex eigenvalues of real matrices. Our aim is to relate the structure of the generalized eigenspace of a matrix given in a block triangular form to the Jordan structure of its diagonal blocks and to the graph structure of the matrix. Formally our results are stated in terms of the eigenvalue 0 of a singular matrix, but this is a technicality, since a scalar matrix may always be added to the original matrix. Our principal results are the Extension Lemma (3.2), the theorem (4.9) on the existence of weakly preferred bases, and the Index Theorem (Corollary (5.8)) and the discussion of the equality case in the Index Theorem in Section 6.

The Extension Lemma (3.2) proved in Section 3 shows that every vector in the generalized nullspace of a diagonal block of a matrix has an extension to a vector in the generalized nullspace of the matrix which satisfies certain combinatorial properties. This lemma is a major tool used for subsequent results.

The existence of a weakly preferred basis for the generalized nullspace of a matrix is proved in Section 4; see Theorem (4.9). Such a basis is characterized by a very special combinatorial structure induced by the reduced graph of the matrix.

Section 5 is independent of Sections 3 and 4. It is an easy consequence of Theorem 2.1 in [6] that the index of a matrix given in a block triangular form is less than or equal to the sum of the indices of the diagonal blocks. In the Index Theorem (Corollary (5.8)) we improve this result. We show that the index of the matrix is less than or equal to the maximal sum of the indices of blocks along a path in the reduced graph. Another proof of the Index Theorem using an entirely different approach is given in [1]. A special case for nonnegative matrices for the eigenvalues that are possibly different from the spectral radius is contained in [9].

In Section 6 we discuss the equality case of the inequality in the Index Theorem. We give a necessary and sufficient condition when A is a 2×2 block matrix (see Theorem (6.8)), and we show that an analogous condition is necessary for the equality when the number of blocks is arbitrary (see

Theorem (6.13)). We show by means of examples that our condition is not sufficient.

This paper continues the series of papers [4, 5, 3]. The current paper is logically independent of these references.

2. NOTATION AND DEFINITIONS

In this paper we discuss $n \times n$ matrices A and vectors with n entries over an arbitrary field (which will not be mentioned explicitly in the sequel). The matrix A is always assumed to be in a (lower) block triangular form, with p diagonal blocks, all square. The diagonal blocks are *not* necessarily irreducible. The dimension of the jth block is n_j, $j \in \langle p \rangle$. Also, every vector b with n entries will be assumed to be partitioned into p vector components b_i conformably with A.

We follow the notation and definitions of [4], [5], and [3].

(2.1) NOTATION. For a positive integer n we denote by $\langle n \rangle$ the set $\{1,\dots,n\}$.

(2.2) NOTATION. Let b be a vector with n entries (partitioned as above). We denote

$$\operatorname{supp}(b) = \{ i \in \langle p \rangle : b_i \neq 0 \}.$$

(2.3) DEFINITION. The *reduced graph* $R(A)$ of the matrix A is defined to be the (directed) graph with vertices $1,\dots,p$ and where (i, j) is an arc if and only if $A_{ij} \neq 0$.

Note that since A is in a block triangular form, $R(A)$ may contain loops but no other (directed, simple) cycles.

(2.4) DEFINITION. Let i and j be vertices in $R(A)$. We say that j accesses i if $i = j$ or there is a path in $R(A)$ from j to i. In this case we write that $i =< j$. We write $i -< j$ for $i =< j$ but $i \neq j$. We write $i \neq< j$ if $i =< j$ is false.

(2.5) DEFINITION. A set W of vertices in $R(A)$ is said to be *initial* if for every vertex j of $R(A)$ and every element i of W, $i =< j$ implies that $j \in W$.

(2.6) NOTATION. Let W be a set of vertices of $R(A)$. We denote

$$\text{below}(W) = \{\text{vertices } i \text{ of } R(A): \text{there exists } j \in W \text{ such that } j =< i\},$$

$$\text{top}(W) = \{i \in W: j \in W,\ j =< i \text{ imply that } i = j\}.$$

(2.7) DEFINITION. A vertex i of $R(A)$ is said to be *singular* if A_{ii} is singular. The set of all singular vertices of $R(A)$ is denoted by S.

(2.8) NOTATION. Let W be a set of vertices of $R(A)$. We denote by $A[W]$ the block matrix $(A_{ij})_{i,\,j \in W}$. Also, if b is an n-vector then we denote by $b[W]$ the block vector $(b_i)_{i \in W}$. Finally, if $W \neq \langle p \rangle$, then we denote by $A(W)$ and by $b(W)$ the block matrix $A[\langle p \rangle \backslash W]$ and the block vector $b[\langle p \rangle \backslash W]$ respectively.

Note that $A[W]$ is a principal submatrix of A.

(2.9) NOTATION. For an $n \times n$ matrix A we denote:

$m(A)$ = the algebraic multiplicity of 0 as an eigenvalue of A;
index(A) = the index of 0 as an eigenvalue of A, viz., the size of the largest Jordan block associated with 0;
$N(A)$ = the nullspace of A;
$E(A)$ = the generalized nullspace of A, viz. $N(A^n)$ [note that $m(A)$ is the dimension of $E(A)$];
range(A) = the range of A.

(2.10) DEFINITION. Let A be an $n \times n$ matrix. A sequence $(x^1, \ldots, x^m)$ of vectors is said to be a *chain* (*with respect to* A) if $Ax^i = x^{i+1}$, $i = 1, \ldots, m - 1$.

(2.11) DEFINITION. Let A be an $n \times n$ matrix. A chain $(x^1, \ldots, x^m)$ of vectors is said to be a *Jordan chain* (*with respect to* A) if $x^m \neq 0$ and $Ax^m = 0$.

(2.12) REMARK. As is well known, the generalized nullspace of a given square matrix has a basis which is a union of Jordan chains. Such a basis is called a *Jordan basis* for the generalized nullspace of the matrix.

(2.13) DEFINITION. Let A be a square matrix and let $x \in E(A)$. We define the *height* of x to be the minimal nonnegative integer k such that $A^k x = 0$. We denote it by height(x).

3. COMBINATORIAL EXTENSIONS OF GENERALIZED NULLVECTORS

(3.1) DEFINITION. Let A be a square matrix in a block triangular form, let x be a vector, and let i be a vertex in $R(A)$. The vector x is said to be a *weak i-combinatorial extension* of an n_i-vector y if $x_i = y$ and $x_j = 0$ whenever $i \neq< j$. The latter condition means that $\operatorname{supp}(x) \subseteq$ below(i).

The following Extension Lemma is a major tool in our results.

(3.2) LEMMA. *Let A be a square matrix in a block triangular form, and let u be a vector in $E(A_{ii})$ for some vertex i in $R(A)$. Then there exists a vector x in $E(A)$ such that x is a weak i-combinatorial extension of u.*

Proof. If $u = 0$, then $x = 0$ is the required vector. So assume that $u \neq 0$. Then A_{ii} is singular. Let $B = A[\text{below}(i)]$, and let $x^1, \ldots, x^{m(B)}$ be a basis for $E(B)$. Observe that $x_i^q \in E(A_{ii})$ for all $q \in \langle m(B) \rangle$. By performing elementary operations we may assume that $x_i^q \neq 0$ if and only if $q \in \langle t \rangle$, and that $x_i^1, \ldots, x_i^t$ are linearly independent vectors in $E(A_{ii})$. Since $x_i^q = 0$ for $t < q \leqslant m(B)$, it follows that the vectors $x^{t+1}(i), \ldots, x^{m(B)}(i)$ are linearly independent vectors in $E(B(i))$, where we recall that $x^j(i) = x^j[\text{below}(i) \setminus \{i\}]$ and $B(i) = B[\text{below}(i) \setminus \{i\}]$. Therefore, we have

$$m(B) \leqslant t + m(B(i)) \leqslant m(A_{ii}) + m(B(i)) = m(B),$$

the last inequality following from the fact that in a block triangular matrix the algebraic multiplicity of 0 as an eigenvalue equals the sum of the algebraic multiplicities of 0 as an eigenvalue of the diagonal blocks. Hence, we must have $t = m(A_{ii})$, and so $x_i^1, \ldots, x_i^t$ form a basis for $E(A_{ii})$. Thus, we have

$$v = \sum_{j=1}^{t} \alpha_j x_i^j$$

for some scalars $\alpha_1, \ldots, \alpha_t$. Since below$(i)$ is an initial set, it follows that by adjoining zero components to $x^1, \ldots, x^t$ we obtain vectors $y^1, \ldots, y^t$ in $E(A)$

that are weak i-combinatorial extensions of $x_i^1, \ldots, x_i^t$ respectively. It now follows that the vector x in $E(A)$, defined by

$$x = \sum_{j=1}^{t} \alpha_j y^j,$$

satisfies $x = u$ and $\mathrm{supp}(x) \subseteq \bigcup_{j=1}^{t} \mathrm{supp}(y^j) \subseteq \mathrm{below}(i)$. ∎

We remark that in general one cannot replace the generalized nullspaces in Lemma (3.2) by the nullspaces. In other words, if u is a vector in $N(A_{ii})$ for some vertex i in $R(A)$, then there does not necessarily exist a vector x in $N(A)$ such that x is a weak i-combinatorial extension of u. Note that in the proof of Lemma (3.2) we used the fact that the algebraic multiplicity of 0 as an eigenvalue of a block triangular matrix equals the sum of the algebraic multiplicities of 0 as an eigenvalue of the diagonal blocks. Clearly, this property does not hold in general for the geometric multiplicity.

The following elementary lemma is essentially known; see Lemma 3.1 in [5].

(3.3) Lemma. *Let A be a square matrix in block triangular form, and let x be a vector. Then* $\mathrm{supp}(Ax) \subseteq \mathrm{below}(\mathrm{supp}(x))$.

(3.4) Proposition. *Let A be a square matrix in block triangular form, let j be a vertex in $R(A)$, and let $(y^1, \ldots, y^m)$ be a chain of n_j-vectors with respect to A_{jj}. Let x^1 be a weak j-combinatorial extension of y^1. Then the vector $x^i = A^{i-1}x^1$ is a weak j-combinatorial extension of y^i, $i \in \langle m \rangle$.*

Proof. By Lemma (3.3) we have

$$\mathrm{supp}(x^2) \subseteq \mathrm{below}(\mathrm{supp}(x^1)) \subseteq \mathrm{below}(\mathrm{below}(j)) = \mathrm{below}(j).$$

Also, since $\mathrm{supp}(x^1) \subseteq \mathrm{below}(j)$ we have

$$x_j^2 = (Ax^1)_j = A_{jj}x_j^1 = A_{jj}y^1 = y^2.$$

Hence, x^2 is a weak j = combinatorial extension of y^2. An inductive argument completes the proof. ∎

Motivated by Proposition (3.4), we now define

(3.5) DEFINITION. Let A be a square matrix in block triangular form, let j be a vertex in $R(A)$, and let $\beta=(y^1,\dots,y^m)$ be a chain of n_j-vectors with respect to A_{jj}. Let x^1 be a weak j-combinatorial extension of y^1, and let $\alpha=(x^1,\dots,x^m)$ be a chain with respect to A, defined by $x^i=A^{i-1}x^1$, $i\in\langle m\rangle$. Then α is said to be a *weak j-combinatorial chain extension* of β. If β is a Jordan chain with respect to A_{jj} and $x^1\in E(A)$, then we call α a *weak j-combinatorial Jordan chain extension* of β.

(3.6) LEMMA. *Let A be a square matrix in block triangular form, let j be a vertex in $R(A)$, and let $\beta=(y^1,\dots,y^m)$ be a Jordan chain of n_j-vectors with respect to A_{jj}. Then there exists a weak j-combinatorial Jordan chain extension α of β.*

Proof. By Lemma (3.2), there exists $x^1\in E(A)$ which is a weak j-combinatorial extension of y^1. Therefore, the chain $\alpha=(x^1,\dots,x^m)$ is the required one. ∎

(3.7) DEFINITION. Let A be a square matrix in block triangular form, let j be a vertex in $R(A)$, and let γ be a Jordan basis of $E(A_{jj})$. A set that consists of weak j-combinatorial Jordan chain extensions of the Jordan chains in γ is said to be a *weak j-combinatorial extension of the Jordan basis γ for $E(A_{jj})$*.

4. WEAKLY PREFERRED BASES FOR THE GENERALIZED NULLSPACE OF A MATRIX

(4.1) DEFINITION. Let A be a square matrix in block triangular form, and let $m_i=m(A_{ii})$. Let H be a set of singular vertices in $R(A)$ (that is, H is a subset of S). A set of vectors x^{ij}, $j=1,\dots,m_i$, $i\in H$, is said to be a *weakly H-preferred set* (for A) if

(4.2) $\{(x^{ij})_i : j\in\langle m_i\rangle\}$ forms a Jordan basis for $E(A_{ii})$, $i\in H$,

(4.3) $\operatorname{supp}(x^{ij})\subseteq \operatorname{below}(i)$ for all $j\in\langle m_i\rangle$, $i\in H$ (i.e. x^{ij} is a weak i-combinatorial extension of $(x^{ij})_i$), and

$$(4.4) \qquad Ax^{ij} = \sum_{h \in H} \sum_{k=1}^{m_h} c_{hk}^{ij} x^{hk}, \qquad j \in \langle m_i \rangle, \quad i \in H,$$

where the c_{hk}^{ij}'s satisfy

$$(4.5) \qquad c_{hk}^{ij} = 0 \qquad \text{whenever} \quad i \not< h, \quad i, h \in H.$$

(4.6) REMARK. Observe that a weakly H-preferred set $\mathscr{S}$ is a set of linearly independent vectors which span an A-invariant subspace V of $E(A)$. The dimension of V equals $\sum_{i \in H} m_i$. Therefore, since $m(A) = \sum_{i \in S} m_i$, a weakly S-preferred set forms a basis for $E(A)$.

In view of Remark (4.6) we define

(4.7) DEFINITION. Let A be a square matrix in block triangular form, and let H be a set of singular vertices in $R(A)$. A weakly H-preferred set $\mathscr{S}$ is said to be a *weakly H-preferred basis* for span($\mathscr{S}$).

(4.8) REMARK. The notion of a weakly preferred set generalizes the notion of a preferred set defined in [4].

(4.9) THEOREM. *Let A be a square matrix in block triangular form. Then there exists a weakly preferred basis for E(A).*

Proof. Let $\gamma_i = \{ y^{ij} : j \in \langle m_i \rangle \}$ be a Jordan basis for $E(A_{ii})$, $i \in S$. We now choose the set $\mathscr{S}$ of n-vectors $\{ x^{ij} : j \in \langle m_i \rangle, i \in S \}$ such that $\{ x^{ij} : j \in \langle m_i \rangle \}$ is a weak i-combinatorial extension of the Jordan basis γ_i for $E(A_{ii})$. All we have to show is that $\mathscr{S}$ is a weakly S-preferred set. Note that (4.2) is given, that we have (4.3) by Proposition (3.4) and Definition (3.7), and that obviously (4.4) holds for appropriate c_{hk}^{ij}'s. We now establish (4.5). Let $i \in S$ and $j \in \langle m_i \rangle$ be given. Define the set $V = \{ h : c_{hk}^{ij} \neq 0 \text{ for some } k \in \langle m_h \rangle \}$. Then (4.5) asserts that $V \subseteq$ below(i). Evidently, the latter is equivalent to the assertion that top(V) $\subseteq$ below(i). So let $r \in$ top(V). We next argue that $(Ax^{ij})_r \neq 0$. Assume to the contrary that $(Ax^{ij})_r = 0$. By (4.4) we have

$$(4.10) \qquad 0 = (Ax^{ij})_r = \sum_{h \in S} \sum_{k=1}^{m_h} c_{hk}^{ij} (x^{hk})_r .$$

Since $r \in \text{top}(V)$, we have $c_{hk}^{ij} = 0$ whenever $h -< r$. Since, by (4.3), $\text{supp}(x^{hk}) \subseteq \text{below}(h)$, it follows that $(x^{hk})_r = 0$ whenever $h \neq< r$. Therefore, (4.10) becomes

$$(4.11) \qquad 0 = (Ax^{ij})_r = \sum_{k=1}^{m_r} c_{rk}^{ij}(x^{rk})_r = \sum_{k=1}^{m_r} c_{rk}^{ij} y^{rk}.$$

The linear independence of y^{rk}, $k \in \langle m_r \rangle$, yields from (4.11) that $c_{rk}^{ij} = 0$, $k \in \langle m_r \rangle$, in contradiction to $r \in V$. Therefore, we have $(Ax^{ij})_r \neq 0$, that is $r \in \text{supp}(Ax^{ij})$. Since, by Lemma (3.3) and by (4.3), we have $\text{supp}(Ax^{ij}) \subseteq \text{below}(\text{supp}(x^{ij})) \subseteq \text{below}(\text{below}(i)) = \text{below}(i)$, it now follows that $r \in \text{below}(i)$. ∎

5. THE HEIGHT OF GENERALIZED NULLVECTORS

(5.1) NOTATION. Let i be a vertex in $R(A)$. We denote by s_i the maximal sum of indices of diagonal blocks of A along a path in $\text{below}(i) \backslash \{i\}$.

(5.2) NOTATION. For a vector x in $E(A)$ we denote

$$q(x) = \max\{s_i + \text{height}(x_i) : i \in \text{top}(\text{supp}(x))\}.$$

Observe that for $x \in E(A)$, $q(x) = 0$ if and only $x = 0$.

(5.3) LEMMA. *Let x be a nonzero vector in $E(A)$. Then $q(Ax) < q(x)$.*

Proof. Let $y = Ax$. If $y = 0$, then the result is obvious. So assume that $y \neq 0$ and let $i \in \text{top}(\text{supp}(y))$. We distinguish between two cases:

I. $i \in \text{top}(\text{supp}(x))$. In this case $\text{height}(y_i) < \text{height}(x_i)$, and hence

$$(5.4) \qquad s_i + \text{height}(y_i) < s_i + \text{height}(x_i).$$

II. $i \notin \text{top}(\text{supp}(x))$. Since, by Lemma (3.3), $\text{supp}(y) \subseteq \text{below}(\text{supp}(x))$, there exists $k \in \text{top}(\text{supp}(x))$ with $k -< i$. Then

$$(5.5) \qquad s_i + \text{height}(y_i) \leqslant s_i + \mu_i \leqslant s_k < s_k + \text{height}(x_k),$$

where $\mu_i = \text{index}(A_{ii})$.

In either case we have a $k \in \text{top}(\text{supp}(x))$ for which the extreme inequalities hold in (5.5), and hence, in view of Notation (5.2), we have $q(Ax) < q(x)$. ∎

We now have three corollaries.

(5.6) COROLLARY. *Let A be a square matrix in block triangular form, and let $x \in E(A)$. Then* $\text{height}(x) \leqslant q(x)$.

Proof. Let $\text{height}(x) = k$. If $k = 0$ then $\text{height}(x) = q(x) = 0$. If $k > 0$ then, by Lemma (5.3), we have

$$q(x) > q(Ax) > \cdots > q(A^{k-1}x) > 0,$$

which yields that $q(x) \geqslant k$. ∎

(5.7) COROLLARY. *Let A be a square matrix in block triangular form, and let $x \in E(A)$. Then* $\text{height}(x)$ *is less than or equal to the maximal sum of indices of diagonal blocks of A along a path in* $\text{below}(\text{supp}(x))$.

Proof. In view of Notation (5.2), the assertion follows immediately from Corollary (5.6), observing that $\text{height}(x_i) \leqslant \mu_i$. ∎

(5.8) COROLLARY. *Let A be a square matrix in block triangular form. Then the maximal height of a vector in $E(A)$ is less than or equal to the maximal sum of indices of diagonal blocks of A along a path in $R(A)$.*

It is an immediate consequence of a lemma in [6] that the index of a matrix in block triangular form is less than or equal to the sum of the indices of the diagonal blocks. We improve this result in the following Index Theorem for general matrices, which is equivalent to Corollary (5.8). A different proof for the Index Theorem may be found in [1], and a special case was proved in [9].

(5.9) THEOREM (The Index Theorem). *Let A be a square matrix in block triangular form. Then the index of A is less than or equal to the maximal sum of indices of diagonal blocks of A along a path in $R(A)$.*

Corollary (5.8) and Theorem (5.9) raise the natural question of when the index of A is *equal* to the maximal sum of indices of diagonal blocks of A along a path in $R(A)$. One equality case, that is for M-matrices, is well known ([8]; see also [11] and [4]). In the following section we shall investigate this problem in general.

(5.10) REMARK. Corollary (5.7) (and thus also Corollary (5.8) and Theorem (5.9)) can also be derived from Theorem (4.9). However, the proof is more complicated, and requires an analysis of the coefficients c_{hk}^{ij} in (4.4).

6. EQUALITY CASES IN THE INDEX THEOREM

(6.1) LEMMA. *Let A be a singular matrix and let* $\operatorname{index}(A)=\mu$. *Then for every $x\in N(A^{\mu})\setminus N(A^{\mu-1})$ we have $x\notin \operatorname{range}(A)+N(A^{\mu-1})$.*

Proof. Let $x\in\operatorname{range}(A)+N(A^{\mu-1})$, and suppose that $x\in N(A^{\mu})=E(A)$. We have $x=Aw+v$ for some vector w, and where $v\in N(A^{\mu-1})$. Thus, $A^{\mu-1}x=A^{\mu}w$. Since $x\in E(A)$, it follows that $w\in E(A)$ and hence $A^{\mu-1}x=A^{\mu}w=0$. It now follows that $x\in N(A^{\mu-1})$. Therefore, $x\in N(A^{\mu})\setminus N(A^{\mu-1})$ implies that $x\notin\operatorname{range}(A)+N(A^{\mu-1})$. ■

(6.2) PROPOSITION. *Let*

$$A=\begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22}\end{pmatrix},$$

where A_{11} and A_{22} are singular, and let η_1 and η_2 be positive integers such that

$$A_{21}\left[\operatorname{range}\left(A_{11}^{\eta_1-1}\right)\cap N(A_{11})\right]\not\subseteq\operatorname{range}(A_{22})+N\left(A_{22}^{\eta_2-1}\right). \tag{6.3}$$

Then

$$\operatorname{index}(A)\geqslant\eta_1+\eta_2.$$

Proof. Suppose that (6.3) holds. Let

$$y_1\in\operatorname{range}\left(A_{11}^{\eta_1-1}\right)\cap N(A_{11}) \tag{6.4}$$

be such that

$$(6.5)\qquad A_{21}y_1 \notin \text{range}(A_{22}) + N\left(A_{22}^{\eta_2 - 1}\right).$$

Obviously, $y_1 \neq 0$. It now follows from (6.4) that there exists $x_1 \in E(A_{11})$ such that

$$A_{11}^{\eta_1 - 1}x_1 = y_1,$$

and hence the height of x_1 with respect to A_{11} is η_1. By Lemma (3.2), there is an $x \in E(A)$ which is a weak 1-combinatorial extension of x_1. Observe that

$$(A^{\eta_1}x)_1 = 0,$$

and hence

$$(6.6)\qquad (A^{\eta_1}x)_2 \in E(A_{22}).$$

Let

$$(A^{\eta_1}x)_2 = A_{21}y_1 + A_{22}z_2,$$

where z is some vector. Since, by (6.5), $A_{21}y_1 + A_{22}z_2$ is not in $N(A_{22}^{\eta_2 - 1})$, it follows from (6.6) that

$$\text{height}((A^{\eta_1}x)_2) \geqslant \eta_2.$$

Therefore, we have $\text{height}(x) \geqslant \eta_1 + \eta_2$, and consequently $\text{index}(A) \geqslant \eta_1 + \eta_2$. ∎

The converse of Proposition 6.2 is not true in general, as is demonstrated by the following example.

(6.7) Example. Let A be the 2×2 block matrix

$$A = \left(\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{array}\right).$$

It is easy to verify that $\text{index}(A_{11}) = \text{index}(A_{22}) = 2$, and $\text{index}(A) = 3$. Now, observe that $A_{21}N(A_{11})$ consists of the zero vector only. Therefore, if we choose $\eta_1 = 1$ and $\eta_2 = 2$, then index$(A) \geqslant \eta_1 + \eta_2$, but (6.3) does not hold.

However, the converse of Proposition (6.2) does hold in the following important case.

(6.8) THEOREM. *Let*

$$A = \begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix},$$

where A_{11} and A_{22} are singular, and let μ_1 and μ_2 be the indices of A_{11} and A_{22} respectively. Then

$$\text{index}(A) = \mu_1 + \mu_2 \tag{6.9}$$

if and only if

$$A_{21}\left[\text{range}\left(A_{11}^{\mu_1 - 1}\right) \cap N(A_{11})\right] \not\subseteq \text{range}(A_{22}) + N\left(A_{22}^{\mu_2 - 1}\right). \tag{6.10}$$

Proof. Suppose that (6.10) holds. By Proposition (6.2) we have index$(A) \geqslant \mu_1 + \mu_2$. Since by Theorem (5.9) we have index$(A) \leqslant \mu_1 + \mu_2$, (6.9) now follows.

Conversely, suppose that (6.9) holds. Let $x \in E(A)$ be such that height$(x) = \mu_1 + \mu_2$. Observe that $(A^{\mu_1}x)_1 = A_{11}^{\mu_1}x_1 = 0$. Let $y_1 = A_{11}^{\mu_1 - 1}x_1$ and $z_2 = (A^{\mu_1 - 1}x)_2$. Then

$$y_1 = A_{11}^{\mu_1 - 1}x_1 \in \text{range}\left(A_{11}^{\mu_1 - 1}\right) \cap N(A_{11}), \tag{6.11}$$

and, since the height of $(A^{\mu_1}x)_2$ equals μ_2,

$$A_{21}y_1 + A_{22}z_2 = (A^{\mu_1}x)_2 \in N\left(A_{22}^{\mu_2}\right) \setminus N\left(A_{22}^{\mu_2 - 1}\right). \tag{6.12}$$

By Lemma 6.1, it follows from (6.12) that $A_{21}y_1 \notin \text{range}(A_{22}) + N(A_{22}^{\mu_2 - 1})$. In view of (6.11), (6.10) holds. ■

We comment that if A_{11} or A_{22} is nonsingular, then clearly index$(A) = \mu_1 + \mu_2$. This case is not covered by Theorem (6.8).

We now generalize one direction of Theorem (6.8) to the case that the number of diagonal blocks is greater than or equal to 2. We first prove a lemma.

(6.13) LEMMA. *Let A be a matrix in lower block triangular form with p square diagonal blocks. Let* $\operatorname{index}(A_{ii}) = \mu_i$, $i \in \langle p \rangle$, *and let μ be the maximal index sum along a chain in $R(A)$. Assume that* $\operatorname{index}(A) = \mu$, *and let x be a vector in $E(A)$ with* $\operatorname{height}(x) = \mu$. *Then there exists* $i \in \operatorname{top}(\operatorname{supp}(x))$ *such that* $\operatorname{height}(x_i) = \mu_i$ *and* $\operatorname{index}(A[\operatorname{below}(i)]) = \mu$.

Proof. Let

$$T = \{ i \in \operatorname{top}(\operatorname{supp}(x)) : \operatorname{height}(x_i) = \mu_i \}.$$

Let $i \in T$. Since $i \in \operatorname{top}(\operatorname{supp}(x))$, we have $x_i \in E(A_{ii})$. By Lemma (3.2) we can find a vector y^i such that $y^i \in E(A)$, and y^i is a weak i-combinatorial extension of x_i. Let

$$z = x - \sum_{i \in T} y^i. \tag{6.14}$$

Since obviously $\operatorname{supp}(x) \subseteq \operatorname{below}(\operatorname{supp}(x))$, and since for all $i \in T$ we have $\operatorname{supp}(y^i) \subseteq \operatorname{below}(i) \subseteq \operatorname{below}(T) \subseteq \operatorname{below}(\operatorname{supp}(x))$, it follows from (6.14) that $\operatorname{supp}(z) \subseteq \operatorname{below}(\operatorname{supp}(x))$. Since

$$z_i = \begin{cases} 0, & i \in T, \\ x_i, & i \in \operatorname{top}(\operatorname{supp}(x)) \setminus T, \end{cases}$$

it follows that for all $i \in \operatorname{top}(\operatorname{supp}(x))$ we have $\operatorname{height}(z_i) < \mu_i$. Therefore, we have $q(z) < \mu$, and by Corollary (5.6) we have $\operatorname{height}(z) < \mu$. Since $\operatorname{height}(x) = \mu$, it now follows from (6.14) that for at least one $i \in T$ we have $\operatorname{height}(y^i) = \mu$. ∎

(6.15) THEOREM. *Let A be a matrix in lower block triangular form with p square diagonal blocks. Let* $\operatorname{index}(A_{ii}) = \mu_i$, $i \in \langle p \rangle$, *and let μ be the maximal index sum along a chain in $R(A)$. If*

$$\operatorname{index}(A) = \mu,$$

then there exists a singular chain $i_1 -\!< i_2 -\!< \cdots -\!< i_t$ with maximal index *sum, such that for every* r, $1 \leqslant r \leqslant t-1$, *we have*

$$(6.16)\qquad A_{kh}\left[\operatorname{range}\left(A_{hh}^{\mu_h-1}\right)\cap N(A_{hh})\right]+\sum_{\substack{l\\ h\prec l\prec k}}\operatorname{range}(A_{kl})$$

$$\not\subseteq \operatorname{range}(A_{kk})+N\left(A_{kk}^{\mu_k-1}\right),$$

where $h=i_r$ *and* $k=i_{r+1}$.

Proof. We prove our assertion by induction on p. If $p=1$ then there is nothing to prove. Assume that our claim holds for $p<m$, $m>1$, and let $p=m$. Let x be an element of a weakly S-preferred basis for $E(A)$ such that height$(x)=\mu$. By definition, there exists $j\in$ top(supp(x)) such that $j\in S$, and such that x is a weak j-combinatorial extension of x_j. Let s_j be the maximal sum of indices of diagonal blocks of A along a path in $T=$ below$(j)\backslash\{j\}$. Observe that $\mu=\mu_j+s_j$. Since, by Corollary (5.6), we have $\mu=$ height$(x)\leqslant q(x)=$ height$(x_j)+s_j$, it follows that

$$(6.17)\qquad \operatorname{height}(x_j)=\mu_j.$$

Also, we have height$(y)=\mu-\mu_j=s_j$, where $y=A^{\mu_j}x$. Let $T'=$ supp(y). By applying Lemma (6.13) to $A[\text{below}(j)\backslash\{j\}]$, we can find $i_2\in$ top(T') such that

$$(6.18)\qquad \operatorname{height}(y_{i_2})=\mu_{i_2},$$

and index$(A[\text{below}(i_2)])=s_j$. By the inductive assumption there exists a singular chain $i_2\prec i_3\prec\cdots\prec i_t$ with maximal index sum (in $R(A[\text{below}(i_2)])$), such that for every r, $2\leqslant r\leqslant t-1$, we have (6.16), where $h=i_r$ and $k=i_{r+1}$. Set $i_1=j$, and observe that the chain $i_1\prec i_2\prec\cdots\prec i_t$ has index sum μ. In order to complete the proof we have to show that (6.16) holds also for $h=i_1$ and $k=i_2$. So, let $h=i_1$ $(=j)$ and $k=i_2$. Let $w=A^{\mu_h-1}x$ (so that $y=Aw$). By (6.17) we have

$$(6.19)\qquad w_h=A_{hh}^{\mu_h-1}x_h\in\operatorname{range}\left(A_{hh}^{\mu_h-1}\right)\cap N(A_{hh}).$$

Also, we have $y_k = A_{kh}w_h + z + A_{kk}w_k$, where

$$z = \sum_{\substack{l \\ h -< l -< k}} A_{kl}w_l. \tag{6.20}$$

It now follows from (6.18) that

$$A_{kh}w_h + z + A_{kk}w_k = y_k \in N(A_{kk}^{\mu_k}) \setminus N(A_{kk}^{\mu_k - 1}).$$

By Lemma (6.1) it follows that

$$A_{kh}w_h + z + A_{kk}w_k \notin \text{range}(A_{kk}) + N(A_{kk}^{\mu_k - 1}),$$

and so

$$A_{kh}w_h + z \notin \text{range}(A_{kk}) + N(A_{kk}^{\mu_k - 1}).$$

In view of (6.19) and (6.20) this proves (6.16). ■

(6.21) Remark. Observe that if a singular chain $i_1 -< i_2 -< \cdots -< i_t$ has a maximal index sum, then there do not exist $l \in \langle p \rangle$ and $r \in \langle t-1 \rangle$ such that A_{ll} is singular and $i_r -< l -< i_{r+1}$. Therefore, all the 1's in the second term on the left hand side of (6.16) correspond to nonsingular blocks A_{ll}. In particular, in the case that the diagonal blocks of A are all singular, (6.16) in Theorem (6.15) can be replaced by

$$A_{kh}\left[\text{range}(A_{hh}^{\mu_h - 1}) \cap N(A_{hh})\right] \not\subseteq \text{range}(A_{kk}) + N(A_{kk}^{\mu_k - 1}).$$

The converse of Theorem (6.15) does not hold in general, as demonstrated by the following examples.

(6.22) Example. We give three examples. First, let A be the 3×3 block matrix

$$\left(\begin{array}{c|cccc|c} 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right).$$

Here, $\mu_1 = \mu_3 = 1$ and $\mu_2 = 2$. Clearly, the chain $1 -\!< 2 -\!< 3$ has maximal index sum ($\mu = 4$). It is easy to verify that (6.16) is satisfied (note that $0^0 = 1$), yet index$(A) = 3$.

Another example, this time of a matrix in Frobenius normal form, is the following matrix, which is similar to A:

$$\left(\begin{array}{c|cccc|c} 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 3 & -1 & -2 & -2 & -2 & 0 \\ -1 & 1 & 1 & 1 & 1 & 0 \\ 0 & -1 & -1 & -1 & -1 & 0 \\ -1 & 1 & 2 & 2 & 1 & 0 \\ \hline 0 & 1 & 1 & 1 & 2 & 0 \end{array}\right).$$

We conclude with a matrix that has also a nonsingular diagonal block. The matrix

$$B = \left(\begin{array}{c|c|c} 0 & 0 & 0 \\ \hline 1 & 1 & 0 \\ \hline 1 & 1 & 0 \end{array}\right)$$

satisfies (6.16). However, index$(B) = 1$, while the maximal index sum along a chain in $R(B)$ is 2.

REFERENCES

1 S. Friedland and D. Hershkowitz, The rank of powers of matrices in a block triangular form, *Linear Algebra Appl.* 107:17–22 (1988).

2 G. F. Frobenius, Über Matrizen aus nicht negativen Elementen, *Sitzungsber. Kön. Preuss. Akad. Wiss. Berlin*, 1912, pp. 456–477, and *Ges. Abh.*, Springer, 1968, Vol. 3, pp. 546–567.

3 D. Hershkowitz, U. G. Rothblum, and H. Schneider, Characterizations and classifications of M-matrices using generalized nullspaces, *Linear Algebra Appl.* 109:59–69 (1988).

4 D. Hershkowitz and H. Schneider, On the generalized nullspace of M-matrices and Z-matrices, *Linear Algebra Appl.* 106:5–23 (1988).

5 D. Hershkowitz and H. Schneider, Solutions of Z-matrix equations, *Linear Algebra Appl.* 106:25–38 (1988).

6 C. D. Meyer and N. J. Rose, The index and the Drazin inverse of block triangular matrices, *SIAM J. Appl. Math.* 33:1–7 (1977).

7 D. J. Richman and H. Schneider, On the singular graph and the Weyr characteristic of an M-matrix, *Aequationes Math.* 17:208–234 (1978).

8 U. G. Rothblum, Algebraic eigenspaces of non-negative matrices, *Linear Algebra Appl.* 12:281–292 (1975).

9 U. G. Rothblum, Bounds on the indices of the spectral-circle eigenvalues of a nonnegative matrix, *Linear Algebra Appl.* 29:445–450 (1980).

10 H. Schneider, The elementary divisors associated with 0 of a singular *M*-matrix, *Proc. Edinburgh Math. Soc.* (2) 19:108–122 (1956).

11 H. Schneider, The influence of the marked reduced graph of a nonnegative matrix on the Jordan form and related properties: A survey, *Linear Algebra Appl.* 84:161–189 (1986).

12 H. D. Victory, On nonnegative solutions to matrix equations, *SIAM J. Algebraic Discrete Methods* 6:406–412 (1985).

Received 23 February 1988; final manuscript accepted 25 March 1988