JORDAN CANONICAL FORMS OF MATRICES AB AND BA (*)

by SIMON J. BERNAU (in El Paso) and ALEXANDER ABIAN (in Ames) (**)

SOMMARIO. - *Siano A e B matrici rettangolari tali che esistano AB e BA. Si dimostra che nelle forme canoniche di Jordan di AB e BA ad ogni blocco di Jordan invertibile di AB corrisponde un identico blocco di BA, e viceversa. Si dimostra anche che i blocchi di Jordan nilpotenti di AB e BA si possono accoppiare in modo che gli indici di nilpotenza dei blocchi in ogni coppia differiscano al più per* 1, *e che ogni altro eventuale blocco nilpotente ha indice* 1.

SUMMARY. - *Let A and B be rectangular matrices such that AB and BA exist. It is shown that in the Jordan canonical forms of AB and BA to every invertible Jordan block of AB there corresponds an identical block of BA and vice versa. Also it is shown that the nilpotent Jordan blocks of AB and BA can be paired off in such a way that the nilpotency indices of the blocks in each pair differ at most by* 1, *and a nilpotent block which is left unpaired (if any) is of index* 1.

In what follows matrices are over an algebraically closed field, say, the field of complex numbers.

It is well known that the Jordan canonical form of a square matrix is determined by its diagonal Jordan blocks, uniquely up to the order in which these blocks occur [2, p. 151].

Also it can be readily verified that in the Jordan canonical form of a square matrix M it is the case that:

(1) The number of Jordan blocks of M is equal to the number of the linearly independent eigenvectors of M. Thus, there exists a one-to-one correspondence between the set of all Jordan blocks and any maximal set of linearly independent eigenvectors of M.

(*) Pervenuto in Redazione il 6 agosto 1988.

(**) Indirizzi degli Autori:
S. J. Bernau - Department of Mathematics - University of Texas - El Paso - Texas - 79968 (U.S.A.);
A. Abian - Department of Mathematics - Iowa State University - Ames - Iowa 50011 (U.S.A).

(2) The Jordan block of M corresponding to *any* eigenvector v of M which belongs to the eingenvalue e of M is an m by m matrix J such that e appears on every diagonal entry of J and if $m \geq 2$ then 1 appears on every entry of J which is immediately above an e, otherwise, 0 appears on every remaining entry of J. Moreover, m is equal to the maximum number of linearly independent vectors $w_2, ..., w_m$ which are obtained by starting with the eingevector $w_1 = v$ and proceeding cyclicly as follows:

(3) $(M - eI)\, w_1 = 0, \;\; (M - eI)\, w_2 = w_1, \;\; ..., \;\; (M - eI)\, w_m = w_{m\text{-}1}$

and where

(4) $(M - eI)\, x = w_m$ *has no solution*

as shown by the Lemma below.

LEMMA 1. *Let H be an m by m matrix and u_1 a nonzero m by 1 vector. Then the set $\{u_1, ..., u_k\}$ of vectors u_i given by*

(5) $Hu_1 = 0, \;\; Hu_2 = u_1, \;\; ..., \;\; Hu_k = u_{k\text{-}1}$

is a maximal set of linearly independent vectors obtained by the above cyclic process iff

(6) $Hx = u_k$ *has no solution.*

Proof. Since u_1 is a nonzero vector, from (5) it follows that u_i is a nonzero vector for every $i = 1, ..., k$. Also from (5) it follows that $H^i u_i = 0$ for every i. Next, we show that the vectors $u_1, ..., u_k$ are linearly independent. Indeed, let for some complex numbers c_i it be the case that $\Sigma c_i u_i = 0$ with $i = 1, ..., k$. But then $\Sigma c_i H^{k\text{-}1} u_i = 0$ which implies $c_k u_1 = 0$ and therefore $c_k = 0$. Consequently, $\Sigma c_i H^{k\text{-}2} u_i = 0$ which implies $c_{k\text{-}1} u_1 = 0$ and therefore $c_{k\text{-}1} = 0$.

In a similar manner we obtain $c_i = 0$ for every $i = 1, ..., k$. Thus, u_i's are linearly independent.

Now, let $\{u_1, ..., u_k\}$ be a maximal set of linearly independent vectors obtained by the cyclic process indicated in (5). We show that (6) holds. Indeed, if $Hx = u_k$ had a solution $x = u_{k+1}$, then our proof above would imply that $\{u_1, ..., u_k, u_{k+1}\}$ is a linearly independent set of vectors, contradicting the maximality of $\{u_1, ..., u_k\}$. Conversely, if (6) holds then the cyclic process in (5) cannot produce more then k vectors.

The following lemmas are needed.

LEMMA 2. *Let A and B be rectangular matrices such that AB and BA exist. Then $e \neq 0$ is an eigenvalue of AB iff $e \neq 0$ is an eigenvalue of BA.*

Proof. Let $e \neq 0$ be an eigenvalue of AB and v a corresponding eigenvector. Thus,

(7) $(AB)\,v = ev$ *where* $Bv \neq 0$ *since* $(AB)\,v \neq 0$.

But then (7) implies $BA\,(Bv) = e\,(Bv)$ and since $Bv \neq 0$ it follows that e is an eingenvalue of BA. The converse is proved in a similar way.

LEMMA 3. *Let A, B, e be as in Lemma 1. Then n is the maximum number of linearly independent eigenvectors of AB belonging to e iff n is the maximum number of linearly independent eigenvectors of BA belonging to e.*

Proof. Let $v_1, ..., v_n$ be the n eigenvectors of AB referred to in the above. Thus, for complex numbers c_i it is the case that

(8) $\Sigma c_i v_i = 0$ *implies* $c_i = 0$ *for* $i = 1, ..., n$.

From the proof of Lemma 1 it follows that $Bv_1, ..., Bv_n$ are eigenvectors of BA belonging to the eingenvalue e of BA. We show that these Bv_i's are linearly independent. Indeed, let for some complex numbers c_i' it be the case that $\Sigma c_i'(Bv_i) = 0$ with $i = 1, ..., n$. But then $\Sigma c_i'(AB)\,v_i = 0$ which by (4) implies $\Sigma c'ev_i = 0$ and therefore $e\Sigma c_i'v_i = 0$, and, since $e \neq 0$ it follows that $\Sigma c_i'v_i = 0$. But the latter in view of (8) implies $c_i' = 0$ for $i = 1, ..., n$. Thus, BA has at least n linearly independent eigenvectors belonging to e. Clearly, the proof of the converse establishes the Lemma.

Let us call a Jordan block J, mentioned in (2), *invertible* iff the corresponding eigenvalue e is nonzero. But then from (1), in view of Lemmas 2 and 3, we have immediately:

COROLLARY 1. *Let A and B be rectangular matrices such that AB and BA exist. Then AB and BA have the same number of invertible Jordan blocks in their Jordan canonical forms.*

LEMMA 4. *Let A, B, e be as in Lemma 1. Then J is a Jordan block of AB corresponding to an eigenvector v of AB belonging to e iff J is a Jordan block of BA corresponding to the eigenvector Bv of BA belonging to e.*

Proof. Let J be an m by m Jordan block of AB corresponding to an eigenvector v of AB belonging to e. Thus, (3) and (4) hold when in them

M is replaced by AB.

Hence, $w_1 = v$ and $\{w_1, ..., w_m\}$ is the maximal set of linearly independent vectors obtained by the cyclic process

(9) $(AB - eI)\, w_1 = 0,\ (AB - eI)\, w_2 = w_1,\ ...,\ (AB - eI)\, w_m = w_{m-1}$

and where

(10) $(AB - eI)\, x = w_m$ *has no solution.*

Next, the left multiplication by B of both sides of each equality in (9) yields

(11) $(BA - eI)\, Bw_1 = 0,\ (BA - eI)\, Bw_2 = Bw_1,\ ...,\ (BA = eI)\, Bw_m = Bw_{m-1}.$

We show that

(12) $(BA - eI)\, x = Bw_m$ *has no solution.*

Let us assume to the contrary that $x = s$ is a solution of (12). Hence, $(BA - eI)\, s = Bw_n$. Multiplying both sides of the latter equality on the left by A we obtain $(AB - eI)\, As = ABw_n$. However, from the last equality in (9) we have $ABw_m = ew_m + w_{m-1}$. Consequently, $(AB - eI)\, As - w_{m-1} = ew_m$ which again in view of the last equality in (9) implies $(AB - eI)\, As - (AB - eI)\, w_m = ew_m$. Therefore, $(AB - eI)\, (As - w_m) - ew_m$ and since $e \neq 0$, we have $(AB - eI)\, (As - w_m)\, e^{-1} = w_m$ contradicting (10). Thus, our assumption is false and (12) is established. But then, from (11), (12) and (2), it follows that J is a Jordan block of BA corresponding to the eigenvector $Bw_1 = Bv$ of BA belonging to the eigenvalue e of BA.

Hence, we have proved that if J is a Jordan block of AB as describe in Lemma 4 then J is a Jordan block of BA as described in Lemma 4. The converse is proved in a similar way with AB replaced by BA.

Based on the above lemmas, we prove:

THEOREM 1. *Let A and B be rectangular matrices such that AB and BA exist. Then to every invertible Jordan block in the Jordan canonical form of AB there corresponds an identical Jordan block in the Jordan canonical form of BA and vice versa.*

Proof. By Corollary 1, AB and BA have the same number of invertible Jordan blocks in their Jordan canonical forms. On the other hand, by (1) and Lemma 4 we see that J is an invertible Jordan block of AB iff J is an invertible Jordan block of BA.

Thus, Theorem 1 is proved.

Next, we consider the relationship of the Jordan nilpotent blocks in the Jordan canonical forms of matrices AB and BA, i.e., the case of $e = 0$ in (2).

As usual, an m by m nilpotent Jordan block is referred to as a nilpotent Jordan block of *index* (or of *index of nilpotency*) m. Thus, by (2), a Jordan nilpotent block of a square matrix M, which corresponds to an eigenvector v of M which belongs to the eigenvalue 0 of M, is of index m iff m is equal to the maximum number of linearly independent vectors $w_1, \ldots, w_m$ which are obtained by starting with the eigenvector $w_1 = v$ and proceeding cyclicly as follows:

(13) $Mw_1 = 0, \; Mw_2 = w_1, \; \ldots, \; Mw_m = w_{m-1}$

and where

(14) $Mx = w_m$ *has no solution.*

Moreover, we call the eigenvector $v = w_1$ described above, a *nilpotent eigenvector of M of index m* and we denote this by:

(15) $N(v) = m$.

Based on the above, we prove the following lemmas.

LEMMA 5. *Let A and B be rectangular matrices such that AB and BA exist. Moreover, let v_1 be a nilpotent eigenvector of AB of index $m \geq 2$.*

(16) *If $Bv_1 \neq 0$ then Bv_1 is a nilpotent eigenvector of BA of index $\geq m$.*

(17) *If $Bv_1 = 0$ then Bv_2 is a nilpotent eigenvector of BA of index $\geq m$-1, where $(AB)\, v_2 = v_1$.*

Proof. By the hypothesis of the Lemma, in view of (13) and (14), we have:

(18) $(AB)\, v_1 = 0, \; (AB)\, v_2 = v_1, \; \ldots, \; (AB)\, v_m = v_{m-1}$ *with no solution for* $(AB)\, x = v_m$.

If $BV_1 \neq 0$ then multiplying both sides of each equality in (18) on the left be B we obtain

(19) $(BA)\,Bv_1 = 0,\ (BA)\,Bv_2 = Bv_1,\ \ldots,\ (BA)\,Bv_m = Bv_{m\text{-}1}$

which, by (13) and (14), implies (16).

If $Bv_1 = 0$ then, since from (18) it follows that $Bv_2 \neq 0$ we see that (19) implies

$$(BA)\,Bv_2 = 0,\ \ldots;\ (BA)\,Bv_m = Bv_{m\text{-}1}$$

which, in turn, by (13) and (14), implies (17).

LEMMA 6. *Let A and B be rectangular matrices such that AB and BA exist. Moreover, let $\{u_1, v_1, \ldots, w_1\}$ be a linearly independent set of nilpotent eigenvectors of AB of index $m \geq 2$ such that*

(20) $Bu_1 = Bv_1 = \ldots = Bw_1 = 0.$

Let

(21) $(AB)\,u_2 = u_1,\ (AB)\,v_2 = v_1,\ \ldots,\ (AB)\,w_2 = w_1.$

Then $\{Bu_2, Bv_2, \ldots, Bw_2\}$ is a linearly independent set of nilpotent eigenvectors of BA of index $\geq m$-1.

Proof. From (20), by Lemma 5, $\{Bu_2, Bv_2, \ldots, Bw_2\}$ is a set of nilpotent eigenvectors of BA of index $\geq m$-1. To show that it is a linearly independent set, let $a(bu_2) + b(Bv_2) + \ldots + t(Bw_2) = 0$ for some complex numbers $a, b, \ldots, t$. But then multiplying the above equality on the left by A, in view of (21) we obtain $au_1 + bv_1 + \ldots + tw_1 = 0$ which by the hypothesis of the Lemma implies $a = b = \ldots = t = 0$. Consequently, $\{Bu_2, Bv_2, \ldots, Bw_2\}$ is linearly independent, as desired.

Based on Lemmas 5 and 6 and the notation (15), we prove:

THEOREM 2. *Let A and B be rectangular matrices such that AB and BA exist. Let the number of the nilpotent Jordan blocks in the Jordan canonical form of, say, AB be greater than or equal to the number of the nilpotent Jordan blocks in the Jordan canonical form of BA. Then the nilpotent blocks of AB and BA can be paired off (exhausting the nilpotent blocks of BA) in such a way that every nilpotent block of AB of index m is paired with a nilpotent block of BA of index m or m - 1 or $m + 1$ and a nilpotent block of AB which is left unpaired (if any) is of index* 1.

Proof. In view of (1), instead of considering the sets of nilpotent Jordan blocks of AB and BA, we may consider a maximal set $D = \{d_1, d_2, ..., d_k\}$ of linearly independent nilpotent eigenvectors of AB and a maximal set $G = \{g_1, g_2, ..., g_n\}$ of linearly independent nilpotent eigenvectors of BA. Moreover, we assume that D as well as G is well ordered according to the nonincreasing indices of nilpotency of its elements.

We pair off d_1 with g_1. Clearly, one and only one of the three cases $N(d_1) = N(g_1)$ or $N(d_1) > N(g_1)$ or $N(d_1) < N(g_1)$ must occur. If $N(d_1) > N(g_1)$ then because of the maximality of G, Lemma 6, in view of the well ordering of G implies $N(g_1) = N(d_1) - 1$. If $N(d_1) < N(G_1)$ then because of the maximality of D, Lemma 6 (where AB is interchanged with BA), in view of the well ordering of D implies $N(g_1) = N(d_1) + 1$. Thus, we see that d_1 and g_1 are paired off as described in the Theorem.

Next, we pair off d_2 with g_2 (of course if they exist) precisely as we proceeded for d_1 and g_1 and obtain the desired result.

Finally, since by the hypothesis $k \geq n$, let d_j be the first (in the well ordering of D) element (if any) of D which is left unpaired. Then $N(d_j) = 1$. This is because otherwise $N(d_j) \geq 2$ which, by the maximality of G, in view of Lemma 6 and the well ordering of G, would imply the existence of an element g_{n+1} of G, such that $N(g_{n+1}) \geq 1$, contradicting the fact that the last element of G is g_n. Thus, indeed $N(d_j) = 1$, as desired.

From Theorems 1 and 2 it follows that we have established:

THEOREM 3. *The Jordan canonical forms of matrices AB and BA have identical invertible blocks. On the other hand, the nilpotent blocks of AB and BA can be paired off in such a way that the nilpotency indices of the blocks in each pair differ at most by* 1, *and, a nilpotent block which is left unpaired (if any) is of index* 1.

REMARK. It is well known [2, p.193] that the *elementary divisors of* a square matrix M are uniquely determined by the Jordan blocks in the Jordan canonical form of M, and, conversely. Thus, if B is an m by m Jordan block of M belonging to an eigenvalue e of M then $(x - e)^m$ is the elementary divisor of M corresponding to B, and, conversely. We also recall [2, p.76] the definitions of the minimal polynomial and the characteristic polynomial, i.e. det $(M - xI)$ of M. But then from theorems 1 and 2 we readily obtain:

COROLLARY 2. *Let A be an m by n and B be an n by m matrix with $m \geq n$. Then the minimal polynomials of AB and BA differ at most with a factor of x. Moreover, the characteristic polynomial of AB is x^{m-n} times the characteristic polynomial of BA.*

It was recently brought to our attention that the subject matter of this paper is treated in [1] using a different method of proof.

REFERENCES

[1] H. FLANDERS, *Elementary Divisors of AB and BA.*, Proceedings of Amer. Math. Soc., *2* (1951), 871-874.

[2] F. R. GANTMACHER, *The Theory of Matrices*, vol. 1, Chelsea Pub. Co., New York, N.Y. 1960.