King Fahd University of Petroleum & Minerals

DEPARTMENT OF MATHEMATICAL SCIENCES

Technical Report Series

TR 416

Jun 2010

A NOTE ON THE JORDAN CANONICAL FORM

H. Azad

DHAHRAN 31261 ● SAUDI ARABIA ● www.kfupm.edu.sa/math/ ● E-mail: c-math@kfupm.edu.sa

A NOTE ON THE JORDAN CANONICAL FORM

H. Azad
Department of Mathematics and Statistics
King Fahd University of Petroleum & Minerals
Dhahran, Saudi Arabia
hassanaz@kfupm.edu.sa

Abstract

A proof of the Jordan canonical form, suitable for a first course in linear algebra, is given. The proof includes the uniqueness of the number and sizes of the Jordan blocks. The value of the customary procedure for finding the block generators is also questioned.

2000 MSC: 15A21.

The Jordan form of linear transformations is an exceeding useful result in all theoretical considerations regarding conjugacy classes of matrices, nilpotent orbits and the Jacobson-Morozov theorem. Classical references for this topic are the books of Smirnov [9,p.245-254] and Horn [7, p.121-131] .There is a very well known proof due to Fillipov [3], which is also given in Strang's book [10, p.422-425]. The American Mathematical Monthly has published at least seven proofs of the Jordan form over the years: [1, 2, 4, 5, 6, 8,11]. In particular, Shapiro's paper [8] discusses in detail the Weyr characteristics, which are essentially the number of blocks of a nilpotent transformation of various sizes. For an inexplicable reason, the Weyr characteristics are not named as such in all the modern books. The recent book of Weintraub [12] is devoted entirely to the topic of the Jordan form and its related computations.

The justification for approaching the subject yet another time can only be the clarity, brevity, arrangement of concepts, and a different viewpoint. Indeed, as soon as it has been established that the number of blocks for a nilpotent transformation is the dimension of its nullspace, the Weir characteristics can be determined more or less immediately, as is done here in the corollary to Proposition 3.

The arguments in this note have the added advantage that the most important parts can be taught in a first course on linear algebra, as soon as basic ideas have been introduced and the invariance of dimensions has been established. This note is thus also a contribution to the teaching of these ideas.

Finally, an efficient algorithm for finding cyclic bases is given, but not in any great detail- for that would have raised the level of exposition and defeated the very purpose of the note.

As is well known, the main technical step in establishing the Jordan canonical form is to prove its existence and uniqueness for nilpotent transformations. We will return to the general case towards the end of this note.

Let A be a nilpotent transformation on a finite dimensional vector space V, let v be a nonzero vector in V and n the smallest integer such that $A^n v = 0$.

Proposition 1 *The vectors $\{A^i v : 0 \leq i < n\}$ are linearly independent.*

Proof. Take an expression

$$\sum_{i=0}^{n-1} c_i A^i v = 0, \qquad (*)$$

in which the number of non-zero coefficients is as small as possible. If the coefficient c_j is the non-zero coefficient of largest index j, then multiplying by A^{n-j}, we obtain an expression like (*) of smaller length. So in (*) every c_i with $i < j$ is 0. Therefore $c_j A^j v = 0$ and therefore $A^j v = 0$, with $j \leq n-1$, which contradicts the choice of n. This proves the claim, ■

Proposition 2 *Let $R(A)$ be the range space of A and $N(A)$ be the null space of A. Let $\{A(v_i) : i = 1, \ldots, r\}$ be a basis of the range space. Let $\{n_j : j = 1, \ldots, s\}$ be a basis of the null space of A. Then $\{v_i : i = 1, \ldots, r, n_j : j = 1, \ldots, s\}$ is a basis of the vector space V.*

Proof. Let $v \in V$. So $A(v) = \sum_{i=1}^{r} c_i A(v_i)$. Therefore $v - \sum_{i=1}^{r} c_i v_i$ belongs to the null space of A, hence it is a linear combination of the $\{v_i\}$ and $\{n_j\}$. To see that these vectors are linearly independent, suppose $\sum_{i=1}^{r} c_i v_i + \sum_{j=1}^{s} d_j n_j = 0$. This gives $\sum_{i=1}^{r} cA(v_i) = 0$ and by linear independence of the vectors $A(v_i)$, we get $c_i = 0,\ i = 1, \ldots, r$. The linear independence of $\{n_j\}$ then shows that $d_j = 0,\ j = 1, \ldots, s$. ■

Proposition 3 *V is a direct sum of cyclic subspaces.*

Proof. We prove this, as in the standard proofs [9, 10], by induction on dimension. The null space of A is a non-zero subspace and therefore the range space of A is a proper subspace of V. If this is the zero subspace, then a basis of V gives the decomposition into cyclic subspaces. So suppose that $R(A)$ is a nonzero subspace. It is an A invariant subspace. By induction on dimensions, it is a direct sum of cyclic subspaces, with generators $v_i,\ i = 1, \ldots, k$, and basis $A^j v_i,\ 0 \le j \le n_i$, and $A^{n_i+1} v_i = 0$. Let $v_i = Aw_i$. So $A^j v_i = AA^j w_i$ shows, using Proposition 1, that the vectors $A^j w_i,\ 0 \le j \le n_i$ are linearly independent. Also $A^{n_i+1} v_i = A^{n_i+2} w_i = 0$, so $A^{n_i+1} w_i = A^{n_i} v_i$ belong to the null space of A.

By Proposition 2, if we enlarge $A^{n_i} v_i,\ i = 1, \ldots, k$, to a basis of the null space of A by adjoining independent vectors $n_1, \ldots, n_l$ in the null space of A, then $A^j w_i, 0 \le j \le n_i, 0 \le i \le k,\ A^{n_i} v_i, i = 1, \ldots, k,\ n_1, \ldots, n_l$ form a basis of V.

Therefore, the cyclic subspaces generated by $w_i,\ i = 1, \ldots, k$ and the one-dimensional subspaces generated by $n_r,\ 1 \le r \le l$ give a direct sum decomposition of V into cyclic subspaces. ■

From this description, it is clear that in each summand only $A^{n_i+1} w_i = A^{n_i} v_i$ contributes to the null space of A in that summand and therefore the number of summands in the above given decomposition is the dimension of the null space of A.

Corollary Let $d_i = dim\,(N(A|R(A^i)),\ i = 0, 1, \ldots, n)$, where n is the smallest positive integer so that $A^n = 0$. The differences $d_0 - d_1, d_1 - d_2, \ldots, d_{n-1} - d_n$ give the number of Jordan blocks of sizes $1, 2, \ldots, n$.

Proof. As shown in the proof of Proposition 3, the number of summands in the Jordan decomposition is the dimension of the null space of A. Therefore the number of blocks of size ≥ 1 is $\dim(N(A))$. Applying A removes all blocks, if any, of size 1, and so the number of blocks of size ≥ 2 is $\dim(N(A|R(A))) = d_1$. Continuing, we get that d_i is the number of blocks of size $\ge i + 1,\ i = 1, \ldots, n$. Therefore the difference $d_{i-1} - d_i$

gives the number of blocks of size i, for $i = 1, \ldots, n$. ■

Examples

1. Let A be any nilpotent upper triangular matrix whose entries to the right of the main diagonal give a non-singular matrix. Then the null space of A is 1 dimensional and therefore the canonical form of A consists of only one block.

   In particular, the matrices

$$\begin{bmatrix} 0 & 2 & & & \\ & 0 & 1 & & \\ & & 0 & -1 & \\ & & & 0 & -2 \\ & & & & 0 \end{bmatrix}$$

   and

$$\begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ & 0 & 7 & 6 & 5 \\ & & 0 & 8 & 9 \\ & & & 0 & 10 \\ & & & & 0 \end{bmatrix}$$

   are conjugate matrices as they are conjugate to

$$\begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & 1 & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}.$$

2. If

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

   then N(A)works out to be 1 dimensional, so there is only 1 Jordan block.

   Also

$$A^3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

   so

$$N(A^3) = \begin{bmatrix} x \\ y \\ z \\ 0 \end{bmatrix}$$

and, as $A^4 = 0$, a basis of $N(A^4)/N(A^3)$ is

$$\nu = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Therefore, this must be a generator of the block.

3. Let

$$A = \begin{bmatrix} 2 & 0 & 2 & 1 \\ 0 & 2 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 4 \end{bmatrix}.$$

The eigenvalue 2 is of multiplicity 3, so the generalized eigenspace $V_{(2)}$ is $3-$ dimensional, whose basis works out to be $(1,0,0,0), (0,1,0,0), (0,0,1,0)$ and the matrix of $A|V_{(2)}$ is therefore

$$\begin{bmatrix} 2 & 0 & 2 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}.$$

We have

$$(A - 2I)|V_{(2)} = \begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Let $\tilde{A} = (A - 2I)|V_{(2)}$. This gives $d_0 = \dim(N(\tilde{A}) = 2$, $d_1 = \dim(N(\tilde{A}|R(\tilde{A})) = 1$, $d_2 = \dim(N(\tilde{A}|R(\tilde{A}^2)) = 0$.

Therefore $\tilde{A}$ has $d_0 - d_1 = 1$ block of size 1 and $d_1 - d_2 = 1$ block of size 2.

The Jordan form of $\tilde{A}$ is therefore

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

and of $A|V_{(2)}$ is

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

The eigenspace for eigenvalue 4 is one-dimensional. Therefore, the Jordan form of A is

$$\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}.$$

In view of such examples, it is not clear to us why an algorithmic procedure is needed to find the precise generators of the various blocks, because all one needs to find the form of the Jordan blocks is to compute the invariants d_i. Nevertheless, for the sake of completeness, we outline such a procedure- at the expense of increase in level of exposition.

Step 1:

Find all eigenvalues. For an eigenvalue λ, compute the generalized eigenspace corresponding to λ. Although, one needs to compute only all vectors annihilated by $(A - \lambda I)^{\dim V}$, it is algorithmically better to compute the vectors annihilated by $(A - \lambda I)^n = 0$, where n is the multiplicity of the eigenvalue λ in the characteristic polynomial of A. So, by working in the generalized eigenspace for λ, and replacing $(A - \lambda I)$ by A, we may assume that A is a nilpotent transformation of index $\leq n$.

From now on, we assume that A is a nilpotent transformation defined on a vector space V

Step 2:

Find the number and sizes of blocks of this nilpotent transformation according to the algorithm given below: it is a restatement of the *Corollary* given on p.3. This is the most important step, which is needed to complete the next step efficiently.

Algorithm for finding the Jordan Form

For a nilpotent transformation A on a finite dimensional vector space V, let N be the smallest integer such that $A^N = 0$. Let $d_i = \dim(N(A|R(A^i)),\ i = 0, 1, \ldots, N)$.

The differences

$$d_0 - d_1, d_1 - d_2, \ldots, d_{N-1} - d_N$$

give the number of Jordan blocks of sizes $1, 2, \ldots, N$.

Step 3: *Algorithm for finding the block generators*

Call a nonzero vector v is of height n if n is the smallest integer so that $A^n(v) = 0$. The vector space spanned by $v, Av, \ldots, A^{n-1}v$ is n-dimensional. A block of size n is an A-invariant subspace generated by a vector of height n.

Let n be the size of the largest block. Choose a basis of $N(A^n)/N(A^{n-1})$. This is a non-zero space, because there exist blocks of size n. The smallest A-invariant subspace of the preimages gives a direct sum of blocks, each of size n. Call this space W_1.

Let m be the size of the block immediately below n. Consider $N(A^m)/N(A^{m-1})$. Find a basis of $N(A^m|W_1)/N(A^{m-1}|W_1)$.
Let $w_1, \ldots, w_r$ be the preimages of these basis elements; they are all of height m.

Extend this basis of $N(A^m|W_1)/N(A^{m-1}|W_1)$ to a basis of $N(A^m)/N(A^{m-1})$ by adjoining independent elements with preimages $v_1, \ldots, v_s$.

The smallest A-invariant subspace spanned by $v_1, \ldots, v_s$ - call it W_2 has 0 intersection with W_1.

Let $W_1 \oplus W_2 = W_3$. Let l be the size of the blocks, if any, just below m. Extend a basis of $N(A^l|W_3)/N(A^{l-1}|W_3)$ to a basis of $N(A^l)/N(A^{l-1})$. As before, we will get the required number of blocks of size l complementary to $W_1 \oplus W_2 = W_3$. Continuing, this will give a Jordan decomposition.

Explanation

Step 3 is based on the following observations

1. If W is a direct sum of blocks and the size of the smallest block is n and $0 < j < n$, then the null-space of A^j in W is the range space of A^{n-j} in W.

2. If W is a direct sum of blocks of size n, generated by vectors $v_1, \ldots, v_k$ - all of height n, then these vectors are linearly independent in $N(A^n)/N(A^{n-1})$.

Conversely, if vectors $w_1, \ldots, w_l$ are in $N(A^n)$ and their images in the quotient $N(A^n)/N(A^{n-1})$ are linearly independent, then the smallest A-invariant subspace generated by $w_1, \ldots, w_l$ is a direct sum of blocks of size n, with generators $w_1, \ldots, w_l$.

Example:

Using the above algorithm, the reader can check that if

$$A = \begin{bmatrix} 0 & 1 & 4 & 5 & 6 & 7 & 8 \\ 0 & 0 & 1 & 6 & 7 & 8 & 9 \\ 0 & 0 & 0 & 0 & 7 & 8 & 9 \\ 0 & 0 & 0 & 0 & 0 & 10 & 11 \\ 0 & 0 & 0 & 0 & 0 & 11 & 12 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

then the smallest integer n so that $A^n = 0$ is 6. There are two blocks, of sizes 1 and 6 respectively, generated by

$$\begin{bmatrix} 0 \\ 19 \\ -6 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

A final remark on applications: A main application of the Jordan form in differential equations is in computation of matrix exponentials. However, it is computationally more efficient to calculate the matrix of A relative to a basis of generalized eigenvectors- not necessarily given by cyclic vectors -and compute its exponential relative to this basis; finally, conjugating by the change of basis matrix gives the exponential of A.

Acknowledgments The author thanks the King Fahd University of Petroleum and Minerals for its support and excellent research facilities.

References

[1] Richard. A. Brualdi, The Jordan canonical form: an old proof. Amer. Math. Monthly **94** no. 3 (1987), 257–267.

[2] S. Cater, An elementary development of the Jordan canonical form. Amer. Math. Monthly **69** no. 5 (1962), 391–393.

[3] A. F. Filippov, A short proof of the theorem on reduction of a matrix to Jordan form. Moscow Univ. Bull. **26** no. 2 (1971), 18–19. 8

[4] R. Fletcher and D. C. Sorensen, An algorithmic derivation of the Jordan canonical form. Amer. Math. Monthly **90** no. 1 (1983), 12–16.

[5] A. Galperin and Z. Waksman, An elementary approach to Jordan theory. Amer.Math. Monthly **87** no. 9 (1980), 728–732.

[6] Israel Gohberg and Seymour Goldberg, A simple proof of the Jordan decomposition theorem for matrices. Amer. Math. Monthly **103** no. 2 (1996), 157–159.

[7] Horn & Johnson, Matrix Analysis, (Cambridge University Press, 1985)

[8] H. Shapiro, The Weyr Characteristic, Amer. Math. Monthly **196** (1999), 919–929.

[9] V. I. Smirnov, Linear Algebra and Group Theory. (Dover Publications, New York, 1961).

[10] Gilbert Strang, Linear Algebra and its Applications. (4th edition, Thomsons, Brooks/Cole, 2006).

[11] H. Valiaho, An elementary approach to the Jordan form of a matrix. Amer. Math.Monthly **93** no. 9 (1986), 711–714.

[12] Steven Weintraub, Jordan Canonical Form: Theory and Practice. (Morgan and Claypool, 2010).