American Journal of Mathematics and Statistics 2016, 6(6): 238-241
DOI: 10.5923/j.ajms.20160606.03

Elementary Proofs of the Jordan Decomposition Theorem for Niloptent Matrices

Mohamed Kobeissi[*], Bilal Chebaro

Larifa Laboratory, Faculty of Sciences, Lebanese University of Beirut, Lebanon

Abstract In this paper we use elementary concepts of linear algebra to show that a nilpotent matrix is similar to a Jordan matrix.

Keywords Jordan decomposition, Nilpotent matrices

1. Introduction

The Jordan decomposition theorem for nilpotent matrices is treated in simple way. While the result is known, the interest of our proofs lies in their simplicity. Note that the usual proofs are mostly based on module theory and/or quotient spaces.

Definition 1 A nilpotent Jordan block of size n, denoted J_{n_i}, is a square matrix of the form:

$$J_n = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ \vdots & & \ddots & 1 & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \dots & \dots & 0 & 0 \end{pmatrix}$$

Definition 2 *A nilpotent Jordan matrix is a block diagonal matrix of the form:*

$$J = \begin{pmatrix} J_{n_1} & 0 & \dots & 0 \\ 0 & J_{n_2} & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & J_{n_r} \end{pmatrix}$$

where each J_{n_i} *is a nilpotent Jordan block.*

Theorem 1 *Let E be a vector space over a field K, of finite dimension n, and f a linear operator on E, nilpotent of index p. There exists some basis B, in which the matrix representing f in B is a Jordan matrix.*

Theorem 1 can be expressed in matrix form as follows:

Theorem 2 *Every nilpotent n×n matrix N of index p is similar to an n×n Jordan matrix J in which the size of the largest Jordan block is k×k, where k is the rank of N.*

A. Galperin and Z. Waksman [1] used elementary concepts to show that "λ-Jordan matrix" is similar to Jordan matrix. Gohberg and Goldberg [2] gave an algorithm that builds Jordan form of an operator *A* on an *n*-dimensional space if the Jordan form restricted to an *n*–1 dimensional invariant subspace is known. In what follows, we give two proofs of theorem 2. The first by using elementary operations on matrices, and the second by using a decomposition of *E* into direct sums of subspaces.

2. Method 1 - Elementary Operations

We shall prove theorem 2 by induction. The following two lemmas are first proved:

Lemma 1 *There exists a matrix A representing f having the following form:*

$$\begin{pmatrix} J_p & \vdots & B \\ \dots & \dots & \dots \\ 0 & \vdots & C \end{pmatrix}$$

Proof 1 *As* f *is nilpotent of index* p*, there exists* $x \epsilon E$ *such that* $f^{p-1}(x) \neq 0$*. The family* $\{x, f(x), \dots, f^{p-1}(x)\}$ *is then linearly independent (and therefore* $p \leq n$*). Suppose the contrary, then there exist constants* $\lambda_0, \lambda_1, \dots, \lambda_{p-1} \epsilon K$ *not all zero such that*

$$\sum_{i=0}^{p-1} \lambda_i f^i(x) = 0$$

Now let

$$i_0 = \min\{i \epsilon \{0, 1, 2, \dots, p-1\};\ \lambda_k \neq 0\}$$

therefore

$$\sum_{i=i_0}^{p-1} \lambda_i f^i(x) = 0$$

and

$$f^{p-i_0-1} \sum_{i=0}^{p-1} \lambda_i f^i(x) = 0$$

* Corresponding author:
mohamed.kobeissi@ul.edu.lb (Mohamed Kobeissi)
Published online at http://journal.sapub.org/ajms

i.e. $\lambda_{i_0} f^{p-1}(x) = 0$ *and* $\lambda_{i_0} = 0$, *which is a contradiction*

By the incomplete basis theorem, there exist vectors $e_1, e_2, \dots e_{n-p}$ *in E such that the family*

$$\tau = \{x, f(x), \dots, f^{p-1}(x), e_1, e_2, \dots e_{n-p}\}$$

spans E. The matrix $A^{'} = M_\tau(f)$ *representing f in that basis has the needed property.*

The lemma 2 below is the key to prove our theorem. We shall prove (again in two ways!) that the bloc matrix *B* found in lemma 1 is in fact the zero matrix.

Lemma 2 *There exists a matrix representing f having the form:*

$$\begin{pmatrix} J_p & \vdots & 0 \\ \dots & \dots & \dots \\ 0 & \vdots & C \end{pmatrix}$$

Proof 2 *method 1: The first proof is based on elementary matrix calculations. For this let the triangular matrix* T_X *defined as follows:*

$$T_X = \begin{pmatrix} I_p & \vdots & X \\ \dots & \dots & \dots \\ 0 & \vdots & I_{n-p} \end{pmatrix}$$

We can easily check that

$$T_X^{-1} = \begin{pmatrix} I_p & \vdots & -X \\ \dots & \dots & \dots \\ 0 & \vdots & I_{n-p} \end{pmatrix}$$

Define the matrix A' by $A^{'} = T_X A T_X^{-1}$, *clearly:*

$$A' = \begin{pmatrix} J_p & -J_p X + B + XC \\ & \\ 0 & I_{n-p} \end{pmatrix}$$

Let X_i *and* B_i *be the i-th rows of X and B respectively, then*

$$-J_p X + B + XC = \begin{pmatrix} -X_2 + B_1 + X_1 C \\ -X_3 + B_2 + X_2 C \\ \vdots \\ -X_p + B_{p-1} + X_{p-1} C \\ B_p + X_p C \end{pmatrix}$$

Now choose $X_1 = 0$, *and for* $1 \le i \le p-1$, $X_{i+1} = B_i + X_i C$, *we obtain a matrix A' , similar to A , of the form*

$$A^{'} = \begin{pmatrix} J_p & \vdots & 0 \\ 0 & \vdots & L \\ \dots & \dots & \dots \\ 0 & \vdots & C \end{pmatrix}, \mathit{with}\ L = \mathcal{M}_{1\times(n-p)}(\mathbb{R})$$

A simple calculation yields: $\forall\, k \in \{1,2,\dots,p\}$

$$(A^{'})^k = \begin{pmatrix} (J_p)^k & \sum_{i=0}^{k-1} (J_p)^{k-1-i} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ L \end{pmatrix} C^i \\ \dots & \dots \\ 0 & C^k \end{pmatrix}$$

As $(A^{'})^p = 0$, *then*

$$\sum_{i=0}^{p-1} (J_p)^{p-1-i} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ L \end{pmatrix} C^i = \sum_{i=0}^{p-1} (J_p)^{p-1-i} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ LC^i \end{pmatrix}$$

$$= \begin{pmatrix} L \\ LC \\ \vdots \\ LC^{p-1} \end{pmatrix} = 0$$

Therefore L=0, and hence

$$A^{'} = \begin{pmatrix} J_p & \vdots & 0 \\ \dots & \dots & \dots \\ 0 & \vdots & C \end{pmatrix}$$

method 2: Let $x \in E$ *such that the family* $\{x, f(x), \dots, f^{p-1}(x)\}$ *is linearly independent. Complete the basis of E by vectors* $e_p, \dots, e_n \in E$

i.e. the family

$$\{\, x, f(x), \dots, f^{p-1}(x)\ , e_p, \dots, e_n \,\}$$

is basis of E. Let x be in E, then

$$x = \sum_{i=0}^{p-1} \lambda_i f^i(x) + \sum_{i=p}^{n} \lambda_i e_i$$

And define the linear operator:

$$l : E \to K \text{ , by } l(x) = \lambda_{p-1}$$

We now state and prove the following three properties about the linear form l:

• *property 1: The family* $\{l, l\circ f, \dots, l\circ f^{p-1}\}$ *is linearly independent in* $E^* = L(E,K)$, *the dual of E*

Suppose the contrary, then there exist scalars $\lambda_0, \lambda_1, \dots, \lambda_{p-1} \in K$, *not all zeros, with*

$$\sum_{i=0}^{p-1} \lambda_i l \circ f^i = 0$$

Denote

$$q = \max\{i \in \{0,1,\dots,p-1\};\ \lambda_i \neq 0\}$$

then

$$\sum_{i=0}^{q} \lambda_i l \circ f^i = 0$$

and

$$\sum_{i=0}^{q} \lambda_i (l \circ f^i) \circ f^{p-q-1}(x) = 0$$

implies that $\lambda_q = 0$, *which contradicts our hypothesis*

Denote now

$$F = vect(x, f(x), \dots, f^{p-1}(x))$$
$$H = vect(l, l\circ f, \dots, l \circ f^{p-1})$$

and let G be the set of all $y \in E$, *such that*

$$l \circ f^i(y) = 0, \forall i \in \{0, 1, \dots, p-1\}$$

• *property 2: The subspace G of E is stable by f*

- *for all* $\lambda \in K$ *and all* $y_1, y_2 \in G$ *, and* $\forall i \in \{0, 1, \dots, p-1\}$ *; we have* $l \circ f^i(\lambda y_1 + y_2) = \lambda l \circ f^i(y_1) + l \circ f^i(y_2) = 0$ *, then* $\lambda y_1 + y_2 \in G$ *. G is therefore a subspace of E*
- *If* $y \in G$*, then* $\forall i \in \{0, 1, \dots, p-1$*:* $l \circ f^i\big(f(y)\big) = l \circ f^{i+1}(y) = 0$*, and* $l \circ f^{p-1}\big(f(y)\big) = l\big(f^p(y)\big) = l(0) = 0$. *Therefore,* $f(y) \in G$.

And finally,

• *property 3:* $E = F \oplus G$

- *If* $\in \{G \cap F\} \setminus \{0\}$ *, then there exist* $\lambda_0, \lambda_1, \dots, \lambda_{p-1} \in K$ *, not all zeros, such that* $y = \sum_{i=0}^{p-1} \lambda_i f$ *; by letting*

  $$q = max\{i \in \{0,1, \dots, p-1\};\ \lambda_i \neq 0\}$$

  then $y = \sum_{i=0}^{q} \lambda_i f^i(x)$*, and* $0 = l \circ f^{p-1-q}(y) = \lambda_q$*, which contradicts the hypothesis*
- *Let* $\psi_{p+1}, \dots, \psi_n \in E^*$ *such that*

  $$\xi^* = \{l, l \circ f, \dots, l \circ f^{p-1}, \dots, \qquad \psi_{p+1}, \dots, \psi_n\}$$

  is a basis of E^**, and let* $\xi = \{\epsilon_{p+1}, \dots, \epsilon_n\}$ *be a basis of E whose dual basis is* ξ^*.
  $\forall\, i \in \{0,1, \dots, p-1\}$ *, and* $\forall\, k \in \{p+1, \dots, n\}$ *;* $l \circ f^i(\epsilon_k)$*=0, thus* $\epsilon_k \in G$
- *If* $z = \sum_{j=1}^{n} \lambda_j \epsilon_j \in G$*, then* $\forall\, i \in \{0,1, \dots, p-1\}$*;*
  $0 = l \circ f^i(\sum_{j=1}^{n} \lambda_j \epsilon_j) = \sum_{j=1}^{n} \lambda_j\, l \circ f^i(\epsilon_j) = \lambda_i = 0$ *,*

hence

$$z = \sum_{j=p+1}^{n} \lambda_j \epsilon_j$$

Thus

$$G = vect(\epsilon_{p+1}, \dots, \epsilon_n)$$

hence $dim(G) = n - p$*, and Consequently* $E = F \oplus G$

Is it now simple to see that the matrix representing f in the basis

$$\left\{f^{p-1}(x), \dots, f(x), x, \epsilon_{p+1}, \dots, \epsilon_n\right\}$$

of E *is of the form:*

$$\begin{pmatrix} J_p & \vdots & 0 \\ \dots & \dots & \dots \\ 0 & \vdots & C \end{pmatrix}$$

The first proof of our theorem can now be completed. For n=1 and n=2, the result is obvious; Assume the result holds up to n–1.

By lemma 2, there exists an invertible matrix Q such that

$$Q^{-1}CQ = \begin{pmatrix} J_p & \vdots & 0 \\ \dots & \dots & \dots \\ 0 & \vdots & C \end{pmatrix}$$

By the induction hypothesis, the exists an invertible matrix P such that

$$P^{-1}CP = \begin{pmatrix} J_{n_2} & 0 & \dots & 0 \\ 0 & J_{n_3} & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & J_{n_r} \end{pmatrix}$$

Let n_1=p, and

$$R = Q \begin{pmatrix} I_p & \vdots & 0 \\ \dots & \dots & \dots \\ 0 & \vdots & P \end{pmatrix}$$

then

$$R^{-1}NR = \begin{pmatrix} J_{n_1} & 0 & \dots & 0 \\ 0 & J_{n_2} & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & J_{n_r} \end{pmatrix}$$

this completes the proof.

3. Method 2 - Decomposition of E into Direct Sums

The second proof was suggested by Rached Mneimné [4] during my visit to the department of Mathematics at Université Diderot in April 2015. The following theorem is on the decomposition of E into direct.

Theorem 3 *Let E be a vector space over a field K, of finite dimension n, and f is a linear operator on E, nilpotent of index p. There exists* $s \in \mathbb{N}^*$ *and subspaces* $E_1, E_2, \dots, E_s$ *of* E *such that:*

a) $E = E_1 \oplus E_2 \oplus \dots \oplus E_s$
b) $\forall\, i \in \{1, 2, \dots, s\}$*,* E_i *is stable by f*
c) the operator f_i*, restriction of* f *over* E_i *, is nilpotent of index* $r_i = dim(E_i)$

Proof 3 *The proof is by induction. It is obviously true for* $n = 0$ *and* $n = 1$. *Now suppose that the result holds up to* $n - 1$.

As f is nilpotent, then $dim\big(Im(f)\big) \leq n - 1$*, and then by the induction hypothesis, there exists* $r \in \mathbb{N}^*$ *and subspaces* $F_1, F_2, \dots, F_r \in im(f)$ *, such that:*

a) $Im(f) = F_1 \oplus F_2 \oplus \dots \oplus F_r$
b) $\forall\, i \in \{1, 2, \dots, r\}$*,* F_i *is stable by f*
c) the operator f_i *, restriction of* f *over* F_i *, is nilpotent of index* $n_i = dim(F_i)$

As f_i *is nilpotent of order* n_i *, then by lemma 1, there exists* $y_i \in F_i$ *such that*

$$F_i = vect\{y_i, f(y_i), \dots, f^{n_i - 1}(y_i)\}$$

Moreover $y_i \in F_i \subseteq Im(f)$ *, then there exists* $x_i \in E$ *such that* $y_i \in f(x_i)$*. Denote*

$$G_i = vect\{x_i, f(x_i), \dots, f^{n_i}(x_i)\}$$

and $G = G_1 + G_2 + \dots + G_r$

Six properties for the subspaces $G_1, G_2, \dots, G_r$ *are stated and proved:*

1. $f(G_i) = F_i$

In fact, if $z = \sum_{j=0}^{n_i} \lambda_j\, f^j(x_i) \in G_i$; then $f(z) = \sum_{j=0}^{n_i - 1} \lambda_j\, f^j\big(y_i\big) \in F_i$ and $\forall\, t = \sum_{k=0}^{n_i} \alpha_k\, f^k(y_i)) \in F_i$; $t = f(\sum_{k=0}^{n_i} \alpha_k\, f^k(x_i)) \in f(G_i)$. Thus $f(G_i) = F_i$

2. $\{x_i, f(x_i), \dots, f^{n_i}(x_i)\}$ is a basis of G_i

If $\sum_{k=0}^{n_i} \alpha_k f^k(x_i) = 0$, then $f(\sum_{k=0}^{n_i} \alpha_k f^k(x_i)) = f(0) = 0$ i.e. $\sum_{k=0}^{n_i-1} \alpha_k f^k(y_i) = 0$, and as the family $\{y_i, f(y_i), \dots, f^{n_i-1}(y_i)\}$ is linearly independent, then $\alpha_0 = \alpha_1 = \cdots = \alpha_{n_i-1} = 0$ and $\alpha_{n_i} f^{n_i}(x_i) = 0$, and $\alpha_{n_i} = 0$, Therefore $\alpha_0 = \alpha_1 = \cdots = \alpha_{n_i} = 0$.

3. $G_i \cap \ker(f) = vect\{f^{n_i}(x_i)\}$

$f(f^{n_i}(x_i)) = f^{n_i}(y_i = 0)$, then $f^{n_i}(x_i) \in \ker(f)$, and $G_i \cap \ker(f) \subseteq vect\{f^{n_i}(x_i)\}$

if $z = \sum_{k=0}^{n_i} \alpha_k f^k(x_i) \in G_i \cap \ker(f)$, then $0 = f(z) = \sum_{k=0}^{n_i-1} \alpha_k f^k(y_i)$, and $\alpha_0 = \alpha_1 = \cdots = \alpha_{n_i-1} = 0$. Therefore $z = \lambda_{n_i} f^{n_i}(x_i) \in vect\{f^{n_i}(x_i)\}$

4. $G \cap \ker(f) = vect\{f^{n_1}(x_1), \dots, f^{n_i}(x_i)\}$

Let $z = z_1 + \cdots + z_r \in G \cap \ker(f)$ with $z_i \in G_i$ $(1 \le i \le r)$, then $0 = f(z) = f(z_1) + \cdots + f(z_r)$.

As $Im(f) = F_1 \oplus F_2 \oplus \dots \oplus F_r$ and $f(z_i) \in F_i$, then $f(z_1) = \cdots = f(z_r) = 0$. By the previous property, there exists $\lambda_i \in K$ such that $z_i = \lambda_i f^{n_i}(x_i)$, and $z = \sum_{i=1}^{r} \lambda_i f^{n_i}(x_i) \in vect\{f^{n_1}(x_1), \dots, f^{n_i}(x_i)\}$. Therefore $G \cap \ker(f) = vect\{f^{n_1}(x_1), \dots, f^{n_i}(x_i)\}$

5. $G = G_1 \oplus G_2 \oplus \dots \oplus G_r$

If $z_1 + \cdots + z_r = 0$ with $z_i \in G_i$, then by the previous property $\forall\, i \in \{1, 2, \dots, r\}$ there exists $\lambda_i \in K$ such that $z_i = \lambda_i f^{n_i}(x_i) = \lambda_i f^{n_i-1}(y_i) \in F_i$. As the subspaces F_i are in direct sum, then $\lambda_i = 0$ and $z_i = 0$, therefore $G = G_1 \oplus G_2 \oplus \dots \oplus G_r$

6. $\dim(G) = \dim\big(Im(f)\big) + r$

$$
\begin{aligned}
\dim(G) &= \dim(G_1) + \cdots + \dim(G_r) \\
&= (n_1 + 1) + \cdots + (n_r + 1) \\
&= \dim(F) + \cdots + \dim(F_r) \\
&= \dim(F_1 \oplus F_2 \oplus \dots \oplus F_r) + r \\
&= \dim\big(Im(f)\big) + r
\end{aligned}
$$

The proof of theorem 3 can now be completed as follows: Let now H be a complement of $G \cap ker(f)$ in $ker(f)$, then:

$G \cap H = (G \cap \ker(f)) \cap H$, cause $H \cap \ker(f) = \{0\}$ (1)

Thus, G and H are in direct sum, and:

$$
\begin{aligned}
\dim(E) &= \dim(Im(f)) + \dim(\ker(f)) \\
&= \dim\big(Im(f)\big) + \dim\big(G \cap \ker(f)\big) + \dim(H) \\
&= \dim(Im(f)) + r + \dim(H) \\
&= \dim(G) + \dim(H)
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
E &= G \oplus H \\
&= G_1 \oplus G_2 \oplus \dots \oplus G_r + H
\end{aligned}
$$

And

$$H = vect(\epsilon_1, \dots, \epsilon_k)$$

with $\{\epsilon_1, \dots, \epsilon_k\}$ is a basis of H.

The second proof of theorem can be obtained. We can now check that the matrix representing f in the basis is:

$$
\begin{pmatrix}
J_{n_1} & 0 & \dots & 0 \\
0 & J_{n_2} & \ddots & 0 \\
\vdots & \ddots & \ddots & 0 \\
0 & \dots & 0 & J_{n_r}
\end{pmatrix}
$$

ACKNOWLEDGEMENTS

The authors are thankful to Salim Kobeissi, Professeur aggregé at Université Pierre Mendès France, Grenoble II, and to Rached Mneimné, Associate Professor at Université Diderot, Paris 7, for their valuable comments.

REFERENCES

[1] A. Galperin and Z. Waksman, An Elementary Approach to Jordan Theory, *The American Mathematical Monthly*, Vol. 87, No. 9 (Nov., 1980), pp. 728-732.

[2] Israel Gohberg and Seymour Goldberg, A Simple Proof of the Jordan Decomposition Theorem for Matrices, *the American Mathematical Monthly*, Vol. 103, No. 2 (Feb., 1996), pp. 157-159.

[3] N. Jacobson, Basic algebra I, *W. H. Freeman and Company, New York*, 1985.

[4] Rached Mneimne, Réduction des endomorphismes: Tableaux de Young, Cône nilpotent, Représentations des algèbres de Lie semi-simples, *Calvage & Mounet*, 2006.