Sapienza
Università Editrice

ANNALI DEL DIPARTIMENTO DI METODI E MODELLI PER L'ECONOMIA IL TERRITORIO E LA FINANZA

2019

Direttore Responsabile - Director
Alessandra De Rose

Direttore Scientifico - Editor in Chief
Roberta Gemmiti

Curatrici del numero - Managing Editors
Adriana Conti Puorger, Cristina Giudici

Comitato Scientifico - Editorial Board
Maria Giuseppina Bruno (Sapienza Università di Roma)
Adriana Conti Puorger (Sapienza Università di Roma)
Alessandra Faggian (The Ohio State University)
Francesca Gargiulo (Sapienza Università di Roma)
Roberta Gemmiti (Sapienza Università di Roma)
Cristina Giudici (Sapienza Università di Roma)
Ersilia Incelli (Sapienza Università di Roma)
Antonella Leoncini Bartoli (Sapienza Università di Roma)
Isabella Santini (Sapienza Università di Roma)
Marco Teodori (Sapienza Università di Roma)
Catherine Wihtol de Wenden (CERI-Sciences Po-CNRS Paris).

Copyright © 2019

Sapienza Università Editrice
Piazzale Aldo Moro 5 – 00185 Roma

www.editricesapienza.it
editrice.sapienza@uniroma1.it

Iscrizione Registro Operatori Comunicazione n. 11420

ISSN: 2385-0825

Pubblicato a dicembre 2019

Quest'opera è distribuita
con licenza Creative Commons 3.0
diffusa in modalità *open access.*

Roberto DE MARCHIS*, **Antonio GRANDE***, **Stefano PATRI'***, **Daniela SAITTA***

An Application of Jordan Canonical Form to the Proof of Cayley-Hamilton Theorem

Abstract. The statement of Cayley-Hamilton theorem is that every square matrix satisfies its own characteristic equation.

Cayley-Hamilton theorem holds both in a vector space over a field and in a module over a commutative ring.

The general proof of Cayley-Hamilton theorem is based on the concepts of *minimal polynomial* and *adjoint matrix* of a linear map (for the details of the general proof, see Lang (2002), page 561, or Liesen and Mehrmann (2011), page 96, or Shurman).

In the case of a diagonalizable matrix A over an algebraically closed field the proof becomes trivial because one can consider the diagonal form D of A and the relation for the k-th power matrix $A^k = CD^kC^{-1}$, where C is the matrix for the basis change to the basis of eigenvectors of A (for the details, see Sernesi (2000) or Lang (1987)).

The aim of this paper is to extend the simple proof for diagonalizable matrices to the case of non-diagonalizable ones over a generic field. First, we obtain a proof for non-diagonalizable matrices over an algebraically closed field and then, by virtue of the properties of field extensions, we show that this proof also holds in the case of a generic field.

Keywords: Jordan Canonical Form, Power of a Jordan Matrix

1. Short overview on spectral theory of matrices

The proof of Cayley-Hamilton theorem for non-diagonalizable matrices, as an extension of the diagonalizable case, is based on Jordan canonical form and the crucial key to this proof is in particular the lemma (1.1) about the structure of the N-th power matrix of a Jordan block. By virtue of the proof presented in this paper we shall have a similar proof of Cayley-Hamilton theorem in both cases of diagonalizable and non-diagonalizable matrix, and in this way we have then obtained a kind of unification of both cases over a field.

Untill the section (2.1.1), we consider an n dimensional vector space $V_{\bar{K}}$ over an algebraically closed field $\bar{K}$ and a square matrix $\bar{A}$ of size $n \times n$ whose elements belong to $\bar{K}$.

In order to begin, we recall some simple results about the spectral theory of matrices in linear algebra. For more details, see Lang (1987).

1.1 Block diagonal matrices

A square matrix $\bar{A}$ is called a *block diagonal matrix* if it has the form

$$\bar{A} = \begin{pmatrix} A_1 & 0 & 0 & \cdots & 0 \\ 0 & A_2 & 0 & \cdots & 0 \\ 0 & 0 & A_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & A_k \end{pmatrix}$$

where every block A_i is a square matrix of any size and every 0 represents a square zero matrix of such an order that the whole matrix $\bar{A}$ is complete.

* Department of *Methods and Models for Economics, Territory and Finance*. "Sapienza" University of Rome - Italy

For our aim, the main property of a block diagonal matrix $\bar{A}$ is the expression of its N-th power matrix

$$\left(\bar{A}\right)^N = \begin{pmatrix} (A_1)^N & 0 & 0 & \cdots & 0 \\ 0 & (A_2)^N & 0 & \cdots & 0 \\ 0 & 0 & (A_3)^N & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & (A_k)^N \end{pmatrix}.$$

A block diagonal matrix in which every block A_i is of order 1, is called a *diagonal matrix*.

1.2 Diagonalizable matrices

We recall that a square matrix $\bar{A}$ of size $n \times n$ is called *diagonalizable* if there exists a basis for $V_{\bar{K}}$ constitued by all eigenvectors of $\bar{A}$.

If a square matrix $\bar{A}$ is diagonalizable, we denote by C the matrix of the basis change to the basis of eigenvectors to obtain the diagonal matrix

$$D = C^{-1}\bar{A}C = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_k \end{pmatrix}$$

where $\lambda_1, \lambda_2, \dots, \lambda_k$ denote the eigenvalues which could not be all distinct.

By inverting the relation $D = C^{-1}\bar{A}C$, we get

$$\bar{A} = CDC^{-1} \tag{1.1}$$

from which the N-th power matrix

$$\left(\bar{A}\right)^N = \overbrace{(CDC^{-1})(CDC^{-1})(CDC^{-1})\cdots(CDC^{-1})}^{N \text{ times}} = CD^NC^{-1}$$

follows, where D^N is the diagonal matrix

$$D^N = \begin{pmatrix} (\lambda_1)^N & 0 & 0 & \cdots & 0 \\ 0 & (\lambda_2)^N & 0 & \cdots & 0 \\ 0 & 0 & (\lambda_3)^N & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & (\lambda_k)^N \end{pmatrix}.$$

1.3 Non-diagonalizable matrices

A square *non-diagonalizable* matrix $\bar{A}$ of size $n \times n$ is a matrix whose eigenvectors do not span $V_{\bar{K}}$.

We recall that a necessary and sufficient condition for a matrix $\bar{A}$ to be non-diagonalizable is that at least one of its eigenvalues has algebraic multiplicity greater than the geometric dimension (geometric multiplicity) of the corresponding eigenspace.

In this case we can always find a basis with respect to which the matrix $\bar{A}$ is represented by Jordan canonical form

$$J = C^{-1}\bar{A}C = \begin{pmatrix} J_1 & 0 & 0 & \cdots & 0 \\ 0 & J_2 & 0 & \cdots & 0 \\ 0 & 0 & J_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & J_k \end{pmatrix}$$

which is a block diagonal matrix. In this case we have

$$\bar{A} = CJC^{-1}\,. \tag{1.2}$$

If for a fixed eigenvalue λ_i, the condition $m_a = m_g = m$ holds, where m_a and m_g denote the algebraic and the geometric multiplicity of λ_i, respectively, then the corresponding block J_i is a diagonal matrix of order m in which all the diagonal elements are λ_i and the off-diagonal ones are zero.

If for a fixed eigenvalue λ_i, the condition $m_a = m > m_g$ holds, then the corresponding block J_i is a matrix of order m having the form

$$J_i = \begin{pmatrix} \lambda_i & g_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_i & g_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_i & g_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i & g_r \\ 0 & 0 & 0 & 0 & \cdots & \lambda_i \end{pmatrix} \tag{1.3}$$

where the elements $g_1, g_2, \ldots, g_r$ can only be 0 or 1.

It is clear that Jordan canonical form of every non-diagonalizable matrix has at least a block of the form (1.3).

If some g_j is equal to zero, then the block J_i is decomposable into *sub-blocks*: if, as an example but without loss of generality, we consider a case of order 5 with $g_3 = 0$, we have then the block J_i of the form

$$J_i = \begin{pmatrix} \lambda_i & 1 & 0 & 0 & 0 \\ 0 & \lambda_i & 1 & 0 & 0 \\ 0 & 0 & \lambda_i & 0 & 0 \\ 0 & 0 & 0 & \lambda_i & 1 \\ 0 & 0 & 0 & 0 & \lambda_i \end{pmatrix},$$

that is

$$J_i = \begin{pmatrix} J_{i1} & 0 \\ 0 & J_{i2} \end{pmatrix},$$

where

$$J_{i1} = \begin{pmatrix} \lambda_i & 1 & 0 \\ 0 & \lambda_i & 1 \\ 0 & 0 & \lambda_i \end{pmatrix} \qquad \text{e} \qquad J_{i2} = \begin{pmatrix} \lambda_i & 1 \\ 0 & \lambda_i \end{pmatrix}.$$

In this case we would have the N-th power matrix

$$(J_i)^N = \begin{pmatrix} (J_{i1})^N & 0 \\ 0 & (J_{i2})^N \end{pmatrix}$$

and it is then clear that we can always consider, without loss of generality, a block J_i as given in (1.3) where every g_j is equal to 1.

In the following lemma, we recall the well-known recursive formula for the N-th power of a Jordan block which will be crucial for the proof of the theorem in the non-diagonalizable case.

Lemma 1.1 *If in the block J_i as given in* (1.3) *the equality $g_j = 1$ holds for all $j = 1, 2, \ldots, r$, then we have*

$$(J_i)^N = \begin{pmatrix} (\lambda_i)^N & a_1 & a_2 & a_3 & \cdots & a_{m-1} \\ 0 & (\lambda_i)^N & a_1 & a_2 & \cdots & a_{m-2} \\ 0 & 0 & (\lambda_i)^N & a_1 & \cdots & a_{m-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & (\lambda_i)^N & a_1 \\ 0 & 0 & 0 & 0 & \cdots & (\lambda_i)^N \end{pmatrix} \tag{1.4}$$

where the elements a_h, for $h = 1, 2, 3, \ldots, m-1$, are given by

$$a_h = \frac{1}{h!}\,\frac{d^h}{dx^h}\left(x^N\right)\Big|_{x=\lambda} = \begin{cases} \dbinom{N}{h}(\lambda_i)^{N-h}, & \textit{if} \quad h \le N \\[2ex] 0\,, & \textit{if} \quad h > N\,. \end{cases}$$

Proof 1.2 *By induction we have that for $N = 1$ the matrix in* (1.4) *becomes the matrix in* (1.3). *Now, if we suppose the equality* (1.4) *true for N, it follows for $N+1$*

$$(J_i)^{N+1} = \begin{pmatrix} (\lambda_i)^{N+1} & b_1 & b_2 & b_3 & \cdots & b_{m-1} \\ 0 & (\lambda_i)^{N+1} & b_1 & b_2 & \cdots & b_{m-2} \\ 0 & 0 & (\lambda_i)^{N+1} & b_1 & \cdots & b_{m-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & (\lambda_i)^{N+1} & b_1 \\ 0 & 0 & 0 & 0 & \cdots & (\lambda_i)^{N+1} \end{pmatrix}$$

where the elements b_j, for $j = 1, 2, 3, \dots, m-1$, are given by

$$b_j = a_{j-1} + \lambda_i a_j = \begin{cases} \binom{N+1}{j} (\lambda_i)^{(N+1)-j}, & \textit{if} \quad j \le N+1 \\ \\ 0\,, & \textit{if} \quad j > N+1\,. \end{cases}$$

Then, since the equality (1.4) *is equally true for $N+1$, we conclude that the* (1.4) *holds for every N and the proof is now complete.*

2. Cayley-Hamilton theorem

If we denote by $\tilde{0}$ the zero matrix and by $\mathcal{P}_{\bar{A}}(x)$ the characteristic polynomial corresponding to a square matrix $\bar{A}$ of size $n \times n$

$$\mathcal{P}_{\bar{A}}(x) = (-1)^n x^n + b_{n-1}x^{n-1} + b_{n-2}x^{n-2} + \cdots + b_2 x^2 + b_1 x + b_0\,,$$

the statement of Cayley-Hamilton theorem is:

$$for\ all\ square\ matrices\ \bar{A},\ we\ have\ \mathcal{P}_{\bar{A}}(\bar{A}) = \tilde{0}\,.$$

We briefly recall the proof of Cayley-Hamilton theorem in the case of diagonalizable matrix.

By using relation (1.1), we have for the diagonalizable matrix $\bar{A}$

$$\mathcal{P}_{\bar{A}}(\bar{A}) = (-1)^n \bar{A}^n + b_{n-1}\bar{A}^{n-1} + b_{n-2}\bar{A}^{n-2} + \cdots + b_2\bar{A}^2 + b_1\bar{A} + b_0 =$$

$$= (-1)^n C D^n C^{-1} + b_{n-1} C D^{n-1} C^{-1} + b_{n-2} C D^{n-2} C^{-1} + \dots$$

$$\cdots + b_2 C D^2 C^{-1} + b_1 C D C^{-1} + b_0 =$$

$$= C\,\Big[\,(-1)^n D^n + b_{n-1} D^{n-1} + b_{n-2} D^{n-2} + \cdots + b_2 D^2 + b_1 D + b_0\Big]\,C^{-1}$$

where the matrix

$$\mathfrak{M} = (-1)^n D^n + b_{n-1} D^{n-1} + b_{n-2} D^{n-2} + \cdots + b_2 D^2 + b_1 D + b_0$$

is given by

$$\mathfrak{M} = \begin{pmatrix} \mathcal{P}_{\bar{A}}(\lambda_1) & 0 & 0 & \cdots & 0 \\ 0 & \mathcal{P}_{\bar{A}}(\lambda_2) & 0 & \cdots & 0 \\ 0 & 0 & \mathcal{P}_{\bar{A}}(\lambda_3) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \mathcal{P}_{\bar{A}}(\lambda_k) \end{pmatrix}. \tag{2.1}$$

Since $\mathcal{P}_{\bar{A}}(\lambda_i) = 0$ holds for every eigenvalue (because the eigenvalues satisfy the characteristic equation $\mathcal{P}_{\bar{A}}(x) = 0$), we conclude $\mathcal{P}_{\bar{A}}(\bar{A}) = \tilde{0}$ and the proof is now complete.

We now extend this strategy to non-diagonalizable matrices over an algebraically closed field.

2.1 Proof for non-diagonalizable matrices

If a matrix $\bar{A}$ over an algebraically closed field is non-diagonalizable, we extend the strategy of the proof of the diagonalizable case.

Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be its distinct eigenvalues with algebraic multiplicities $m_1, m_2, \ldots, m_k$, respectively, such that

$$m_1 + m_2 + \cdots + m_k = n\,.$$

We notice that necessarily at least an eigenvalue has algebraic multiplicity greater than 1.

The characteristic polynomial is then of the form

$$\mathcal{P}_{\bar{A}}(x) = (-1)^n \cdot (x-\lambda_1)^{m_1} \cdot (x-\lambda_2)^{m_2} \cdot \cdots \cdot (x-\lambda_k)^{m_k}$$

with the obvious property (property of multiple zeros)

$$\frac{d^h}{dx^h}\,\mathcal{P}_{\bar{A}}(x)\bigg|_{x=\lambda_i} = 0\,, \quad \forall\ h = 0,1,2,3,\ldots,m_i-1\,. \tag{2.2}$$

In analogy with the diagonalizable case and with the unique difference that the matrix $\bar{A}$, under an appropriate basis change, transforms, as known, into a Jordan block diagonal matrix J, we have by virtue of the relation (1.2)

$$\begin{aligned}\mathcal{P}_{\bar{A}}(\bar{A}) &= (-1)^n \bar{A}^n + b_{n-1}\bar{A}^{n-1} + b_{n-2}\bar{A}^{n-2} + \cdots + b_2\bar{A}^2 + b_1\bar{A} + b_0 = \\ &= (-1)^n C J^n C^{-1} + b_{n-1} C J^{n-1} C^{-1} + b_{n-2} C J^{n-2} C^{-1} + \ldots \\ &\quad \cdots + b_2 C J^2 C^{-1} + b_1 C J C^{-1} + b_0 = \\ &= C\left[\,(-1)^n J^n + b_{n-1}J^{n-1} + b_{n-2}J^{n-2} + \cdots + b_2J^2 + b_1J + b_0\right] C^{-1}\,.\end{aligned}$$

Since Jordan matrix J is a diagonal block matrix, it follows that also the matrix

$$\mathcal{M} = (-1)^n J^n + b_{n-1}J^{n-1} + b_{n-2}J^{n-2} + \cdots + b_2J^2 + b_1J + b_0 \tag{2.3}$$

is a diagonal block matrix containing two kinds of blocks.

The first kind of block inside the matrix $\mathcal{M}$, denoted by $\mathcal{M}_1$, is the one corresponding to a diagonal block J_i and is of the form

$$\mathcal{M}_1 = \begin{pmatrix} \mathcal{P}_{\bar{A}}(\lambda_i) & 0 & 0 & \cdots & 0 \\ 0 & \mathcal{P}_{\bar{A}}(\lambda_i) & 0 & \cdots & 0 \\ 0 & 0 & \mathcal{P}_{\bar{A}}(\lambda_i) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \mathcal{P}_{\bar{A}}(\lambda_i) \end{pmatrix} = \tilde{0}$$

because the eigenvalue λ_i is a root of the characteristic polynomial.

The second kind of block inside the matrix $\mathcal{M}$, denoted by $\mathcal{M}_2$, is the one corresponding to a block J_i as given in (1.3) and, by virtue of the power matrix (1.4), is of the form

$$\mathcal{M}_2 = \begin{pmatrix} \mathcal{P}_{\bar{A}}(\lambda_i) & \mathcal{Q}_1(\lambda_i) & \mathcal{Q}_2(\lambda_i) & \mathcal{Q}_3(\lambda_i) & \cdots & \mathcal{Q}_{m_i-1}(\lambda_i) \\ 0 & \mathcal{P}_{\bar{A}}(\lambda_i) & \mathcal{Q}_1(\lambda_i) & \mathcal{Q}_2(\lambda_i) & \cdots & \mathcal{Q}_{m_i-2}(\lambda_i) \\ 0 & 0 & \mathcal{P}_{\bar{A}}(\lambda_i) & \mathcal{Q}_1(\lambda_i) & \cdots & \mathcal{Q}_{m_i-3}(\lambda_i) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \mathcal{P}_{\bar{A}}(\lambda_i) & \mathcal{Q}_1(\lambda_i) \\ 0 & 0 & 0 & 0 & \cdots & \mathcal{P}_{\bar{A}}(\lambda_i) \end{pmatrix},$$

where the elements above the diagonal are given by

$$\mathcal{Q}_h(\lambda_i) = \frac{1}{h!}\,\frac{d^h}{dx^h}\,\mathcal{P}_{\bar{A}}(x)\bigg|_{x=\lambda_i}\,, \quad \text{for } h = 1,2,3,\ldots,m_i-1\,.$$

The matrix $\mathcal{M}_2$ is the matrix $\tilde{0}$ because the eigenvalue λ_i is a root of the characteristic polynomial and, by virtue of (2.2), it yields

$$\mathcal{Q}_h(\lambda_i) = 0\,, \quad \forall\ h = 1,2,3,\ldots,m_i-1\,.$$

Since every block inside $\mathcal{M}$ is the matrix $\tilde{0}$, we have then obtained the proof of Cayley-Hamilton theorem for a non-diagonalizable matrix over an algebraically closed field.

2.1.1 Extension to a generic field

Let V_K be a vector space over a generic field K and $\bar{K}$ the algebraic closure[1] of K, such that we can extend the vector space V_K to the vector space $V_{\bar{K}}$ over $\bar{K}$ (that is $V_{\bar{K}} = V \otimes_K \bar{K}$). If A is an endomorphism of V_K, we can univocally extend A to an endomorphism $\bar{A}$ of $V_{\bar{K}}$ and both the endomorphisms have the same characteristic polynomial $\mathcal{P}_A(x) \equiv \mathcal{P}_{\bar{A}}(x)$. From the result $\mathcal{P}_{\bar{A}}(\bar{A}) = \tilde{0}$, the result $\mathcal{P}_A(A) = \tilde{0}$ follows.

3. Conclusions

Since we have used the matrix $\mathfrak{M}$ in the (2.1) for diagonalizable matrices and, by virtue of the lemma (1.1), the matrix $\mathcal{M}$ in the (2.3) for non-diagonalizable matrices, we conclude that over a field the proofs corresponding to both cases of diagonalizable and non-diagonalizable matrices are similar and we have then obtained a kind of unification of the proofs of Cayley-Hamilton theorem over a generic field.

Acknowledgments

We are very grateful to Professor Antonio Maschietti for his helpful and useful suggestions and comments.

References

APOSTOL T. M. (1985), *Calcolo. Vol. 2: Geometria*. Bollati Boringhieri.
ARTIN M. (1991), *Algebra*, Prentice-Hall.
JACOBSON N. (2009), *Basic Algebra 1*, Dover Publications, 2nd Ed.
LANG S. (1987), *Linear Algebra*, Springer-Verlag, 3rd Ed.
LANG S. (2002), *Algebra*, Springer-Verlag, revised 3rd Ed.
LIESEN J., MEHRMANN V. (2011), *Lineare Algebra*, Vieweg+Teubner Verlag (German Ed.).
SERNESI E. (2009), *Geometria 1*, Bollati Boringhieri.
SHURMAN J. (personal website: *http://people.reed.edu/ jerry/*), downloadable note from the website *http://people.reed.edu/ jerry/332/28ch.pdf*.

[1] By virtue of Zorn's lemma, every field K has an algebraic extension, called *algebraic closure*, which is algebraically closed and unique up to an isomorphism that fixes every member of K