Asian Research Journal of Mathematics

18(11): 301-319, 2022; Article no.ARJOM.92844

ISSN: 2456-477X

On the Computation of the Minimum Polynomial and Applications

Nikolaos Halidias [a*]

[a] *Department of Statistics and Actuarial-Financial Mathematics, University of the Aegean, Karlovassi 83200 Samos, Greece.*

Author's contribution

The sole author designed, analysed, interpreted and prepared the manuscript.

Article Information

DOI: 10.9734/ARJOM/2022/v18i11603

Open Peer Review History:

This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/92844

Received: 08 August 2022

Accepted: 14 October 2022

Published: 21 October 2022

Review Article

Abstract

Aims/Objectives: In this review article we study the computation of the minimum polynomial of a matrix A and how we can use it for the computation of the matrix A^n. We also describe the form of the elements of the matrix A^{-n} and we will see that it is closely related with the computation of the Drazin generalized inverse of A. Next we study the computation of the exponential matrix and finally we give a simple proof of the Leverrier - Faddeev algorithm for the computation of the characteristic polynomial.

Keywords: Minimum polynomial; characteristic polynomial; exponential matrix; Drazin inverse; Leverrier - Faddeev algorithm.

2010 Mathematics Subject Classification: 15-01.

1 Introduction

We are interested here for the computation of the minimum polynomial of a matrix $A_{m\times m}$ (without first compute the characteristic polynomial) and its applications, see the paper [1] for this matter. Here we give a simpler proof than [1]. In particular we are interested for the computation of the matrix A^n for every $n \in \mathbb{Z}$ (if is invertible) and for the computation of the exponential matrix e^{tA},

**Corresponding author: E-mail: nikoshalidias@hotmail.com;*

one can see the book of [2] for this matter. It is well known how one can compute them in the case where the matrix A is diagonalizable but here we are interested for a general method. We will see that the Cayley-Hamilton theorem and the notion of the generalized Vandermonde determinant are very crucial in this direction. In order to compute the matrices A^n and e^{tA} we can use the characteristic polynomial or the minimum polynomial. In the case where the minimum polynomial is of much less degree in comparison with the characteristic then we will see that using the minimum polynomial one can compute the desired matrices by much less computational effort.

Moreover, we will prove that if the matrix A is invertible and we have computed the matrix A^n for $n \in \mathbb{N}$ then the matrix A^{-n} can be produced directly from the elements of the matrix A^n. If we denote these elements as $a_{ij}(n)$ then we shall prove that the elements of A^{-n} are the $a_{ij}(-n)$. Moreover, we will see that even in the case where the matrix is not invertible the elements $a_{ij}(-1)$ are well defined. So, what is the role of the matrix D that contains the elements $a_{ij}(-1)$? We shall see that the matrix D is the Drazin inverse (a notion of a generalized inverse of a matrix, see for example [3]). Therefore, one way to compute the Drazin generalized inverse of the matrix A is first to compute the matrix A^n and then substitute n by -1. Of course, if the matrix A is invertible then we arrive at the matrix A^{-1}.

Finally, we will study the Leverrier-Faddeev algorithm (see [4]) for the computation of the characteristic polynomial, giving a simple proof of the construction.

One can see the following books and articles that are closely related to this subject [5], [6], [7] and [8] - [14].

2 Minimum Polynomial

In this section we will restate the definition of the minimum polynomial and some well known results about it. By P_A we denote the set of all the polynomials that vanishes by the matrix A,

$$P_A = \{\text{all the polynomials } p(k) : p(A) = 0\}$$

Obviously P_A is nonempty because it contains the characteristic polynomial $\delta_A(k)$ of the matrix A. Let $q(k) = k^r + b_{r-1}k^{r-1} + \cdots + b_0$ the polynomial that belongs to P_A with less degree, that is $q(A) = 0$ and there is not another polynomial with less degree than $q(k)$ that vanishes by A. We can prove that this polynomial is unique and we call it the minimum polynomial.

Proposition 2.1. *If $p \in P_A$ then p is divided by the minimum polynomial $q(k)$. The minimum polynomial $q(k)$ is unique among the monic polynomials.*

Proposition 2.2. *The characteristic polynomial $\delta(k)$ of a matrix $A_{n\times n}$ and the minimum polynomial q have the same roots. Moreover, if the characteristic polynomial has n distinct roots then $\delta(k) = q(k)$.*

Theorem 2.1. *The matrix $A_{n\times n}$ is diagonalizable* $\mathtt{iff}$ *the minimum polynomial has the form*

$$q(k) = (k - k_1)\cdots(k - k_m)$$

where $k_1, \cdots, k_m$ the $m \leq n$ distinct eigenvalues of the matrix.

We will now describe a method in order to calculate the minimum polynomial directly, without compute first the characteristic polynomial.

Let a matrix $A_{n\times n}$ and $\delta(k) = k^n + b_{n-1}k^{n-1} + \cdots + b_0$ its characteristic polynomial. Obviously $\delta(A) = 0$.

Let the minimum polynomial has the following form

$$q(k) = k^r + a_{r-1}k^{r-1} + \cdots + a_0$$

with $r \leq n$.

Then

$$a_0\mathbb{I} + a_1A + \cdots + a_{r-1}A^{r-1} = -A^r$$

From the above matrix equality we can construct a linear system for the unknowns $a_0, \cdots, a_{r-1}$ and this linear system has a unique solution since the minimum polynomial is unique. What is the matrix of this linear system? The left hand side is a matrix and each element is equal with the corresponding element of the right hand side, i.e. with the corresponding element of the matrix $-A^r$. That is we have n^2 equations with r unknowns. The matrix of the above system is a matrix $C_{n^2\times r}$. In order to construct this matrix we will put the elements of the identity matrix (with any order) at the first column of C, then at the second column we will put the matrix A (with the same order as before), and final the matrix A^{r-1} at the r column of the matrix C. The row echelon form of this matrix will have r leading ones at the first r columns, side by side. If we construct now the augmented matrix $(C|X)$ where X is the column that contains the matrix $-A^r$ then the row echelon matrix will have again r (and not $r+1$) leading ones side by side at the first r columns. If it has $r+1$ leading ones then the above system will not have a solution and that contradicts with the fact that there exists a unique polynomial that satisfies the above matrix equality.

The system

$$a_0\mathbb{I} + a_1A + \cdots + a_{r-1}A^{r-1} = -a_rA^r$$

has infinitely many solutions with one free parameter. One of them is the coefficients of the minimum polynomial. Choosing $a_r = 1$ we get the coefficients of the minimum polynomial. Similarly the system

$$a_0\mathbb{I} + a_1A + \cdots a_{r-1}A^{r-1} = -a_rA^r - a_{r+1}A^{r+1} - \cdots - a_nA^n$$

will have infinitely many solutions. Choosing $a_r = 1$ and $a_{r+1} = \cdots = a_n = 0$ we will get the coefficients of the minimum polynomial. The r here is the number of the leading ones that are located side by side at the first r columns and the next column does not have a leading one.

Therefore, if $A_{n\times n}$ is a given matrix then we construct the matrix $B_{n^2\times(n+1)}$ which contains in its columns the matrices $I, A, \cdots, A^n$ with any order but the same for any column. We compute the reduced row echelon form $\hat{B}$. The number r of the leading ones located at the first columns side by side will give us the degree of the minimum polynomial. That is, in order to find the degree of the minimum polynomial one has to find the first column that has no leading one, i.e. find the minimum $r = 1, 2, \cdots$ such that $\hat{B}_{r+1,r+1} = 0$. Then the degree of the minimum polynomial equals r and we set $a_r = 1$ and $a_{r+1} = \cdots = a_n = 0$. Therefore the minimum polynomial will be the following

$$q(k) = k^r - \hat{B}_{r,r+1}k^{r-1} - \hat{B}_{r-1,r+1}k^{r-2} - \cdots - \hat{B}_{1,r+1}$$

Example 2.2. *Let the matrix*

$$A = \begin{pmatrix} -4 & 2 & 0 \\ -2 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

We will compute the minimum polynomial. The matrix B is

$$B = \begin{pmatrix} 1 & -4 & 12 & -28 \\ 0 & -2 & 10 & -34 \\ 0 & 0 & 0 & 0 \\ 0 & 2 & -10 & 34 \\ 1 & -1 & -3 & 23 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

and the reduced row echelon form is the following

$$\hat{B} = \begin{pmatrix} 1 & 0 & 0 & 8 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

That is the minimum polynomial is

$$q(k) = \delta(k) = k^3 + 4k^2 + 3k - 8$$

□

3 The Matrix A^n for $n \in \mathbb{Z}$

Let us recall first the Cayley-Hamilton theorem that is very crucial in our approach.

Theorem 3.1 (`Cayley-Hamilton`). *For every matrix $A_{m\times m}$ with characteristic polynomial*

$$\delta_A(k) = k^m + a_{m-1}k^{m-1} + \cdots + a_0,$$

it holds that

$$\delta_A(A) = 0.$$

Remark 3.1. By the fundamental theorem of algebra it is well known that the sum of the multiplicities of the roots of a polynomial of degree m equals m. Therefore any polynomial is such that the sum of the multiplicities is greater than m has to be the zero polynomial, i.e. all the coefficients are zero. □

Let a polynomial $p(k)$ of degree greater than or equal m and let the matrix $A_{m\times m}$. If we want to compute the matrix $p(A)$ then we can divide $p(k)$ by $\delta_A(k)$ arriving at $p(k) = \delta_A(k)\pi(k) + v(k)$ for every $k \in \mathbb{R}$. Therefore the polynomial $\lambda(k) = p(k) - \delta_A(k)\pi(k) - v(k)$ has all its coefficients equals to zero because vanishes for all $k \in \mathbb{R}$. So we see that $p(A) = \delta_A(A)\pi(A) + v(A) = v(A)$.

In particular, if we want to compute the n-th power of a matrix $A_{m\times m}$ where $n > m$, we divide (theoretically) the polynomial k^n by the characteristic polynomial arriving at $k^n = \delta_A(k)\pi(k)+v(k)$. Thus we have to compute the coefficients of the unknown polynomial $v(k)$. For this we can use the eigenvalues of the matrix A which satisfy the equation $k_i^n = \delta_A(k_i)\pi(k_i) + v(k_i) = v(k_i)$. In the case where some eigenvalue is of multiplicity greater than one then we construct more equations by

differentiating the equation $k^n = \delta_A(k)\pi(k) + v(k)$ and then substitute the eigenvalue. We end up by the following linear system of equations

$$\begin{aligned} v^{(i)}(z_1) &= (k^n)^{(i)}_{k=z_1}, \quad i = 0, \cdots, l_1 - 1 \\ &\vdots \\ v^{(i)}(z_t) &= (k^n)^{(i)}_{k=z_t}, \quad i = 0, \cdots, l_t - 1 \end{aligned}$$

where by $v^{(i)}$ and by $(k^n)^{(i)}$ we mean the i-th derivative of the function. This system has always a unique solution because the determinant is a generalized Vandermonde determinant (see [15]). To see that, one first observes that this system has at least one solution because the number of equations equals the number of unknowns. Supposing that there exists two solutions of the system we see that their difference satisfy the homogeneous system. Note that $v(k)$ is a polynomial of degree $m-1$ so the sum of the multiplicity of its roots equals to $m-1$. But we also see that the same polynomial has the z_i as its roots and the sum of their multiplicity are m. Therefore it has to vanish everywhere so there is only one solution of the system.

Once we have computed the polynomial $v(k)$, then from the relation

$$k^n = \delta_A(k)\pi(k) + v(k)$$

substituting A instead of k it follows that $A^n = v(A)$ because by the Cayley-Hamilton theorem we have that $\delta_A(A) = 0$. If the matrix A is real then one can easily prove that the matrix $v(A) = A^n$ is real.

Remark 3.2. The following linear system

$$\begin{aligned} v^{(i)}(k_1) &= (k^n)^{(i)}_{k=k_1}, \quad i = 0, \cdots, r_1 - 1 \\ &\vdots \\ v^{(i)}(k_l) &= (k^n)^{(i)}_{k=k_l}, \quad i = 0, \cdots, r_l - 1 \end{aligned}$$

have a unique solution $(a_{m-1}(n), \cdots, a_0(n))$ which is such that $v(A) = a_{m-1}(n)A^{m-1} + a_{m-2}(n)A^{m-2} + \cdots + a_0(n)I_{m\times m} = A^n$. This holds even in the case where $n \leq m$. The unique solution is $a_n(n) = 1$ and $a_i(n) = 0$ for $i = 0, 1, \cdots, n-1, n+1, \cdots, m$ that is

$$(0, 0, \cdots, \underbrace{1}_{\text{nth element}}, 0 \cdots, 0)$$

therefore it also holds that $v(A) = A^n$ where $n \leq m$.

Remark 3.3. We can use the minimum polynomial instead of the characteristic in order to compute the nth power of the matrix. Indeed, denoting by $m_A(k)$ the minimum polynomial we can write $k^n = m_A(k)l(k) + \hat{v}(k)$. In order to compute the polynomial $\hat{v}(k)$ we can use the roots of the minimum polynomial and construct a linear system as above. This system admits also a unique solution and therefore we have computed the matrix A^n which equals the matrix $\hat{v}(A)$. In this approach we see that the linear system that we construct in order to compute the coefficients of the polynomial $\hat{v}(k)$ has l equations when the degree of the minimum polynomial is $l-1$. Comparing with the system using the characteristic polynomial we see that the computational effort is much less in the case where l is much less than m. Note also that we can use any polynomial that vanishes by the matrix A in order to compute the matrix A^n. □

Let the elements of A^n are $a_{ij}(n)$. We shall prove that when A is invertible then the elements of A^{-1} are the $a_{ij}(-1)$ and generally the elements of A^{-n} are the $a_{ij}(-n)$.

Theorem 3.2 (`The elements of` A^{-n})**.** *Let the matrix $A_{m\times m}$ is invertible. If the (i,j) element of A^n is $a_{ij}(n)$ then the (i,j) element of A^{-n} is $a_{ij}(-n)$. In particular the elements of A^{-1} are the $a_{ij}(-1)$.*

Proof. Let

$$\delta_A(k) = a_m k^m + \cdots + a_0$$

the characteristic polynomial of $A_{m\times m}$. By the Cayley-Hamilton theorem we know that $\delta_A(A) = 0$. If the matrix A is invertible then from the equality $\delta_A(A) = 0$ it follows that

$$A^{-1} = b_{m-1}A^{m-1} + \cdots + b_0 I_{m\times m}, \quad b_k = -\frac{a_{k+1}}{a_0}, \quad k = 0, 1, \cdots, m-1$$

We can compute the matrix A^n by using the Cayley-Hamilton theorem as we have described. To do this, we consider a polynomial $v(k) = b_{m-1}A^{m-1} + \cdots + b_0$ and we solve the system

$$\begin{aligned} v^{(i)}(k_1) &= (k^n)^{(i)}_{k=k_1}, \quad i = 0, \cdots, r_1 - 1 \\ &\vdots \\ v^{(i)}(k_l) &= (k^n)^{(i)}_{k=k_l}, \quad i = 0, \cdots, r_l - 1 \end{aligned} \tag{3.1}$$

when A has eigenvalues $k_1, \cdots, k_l$ of multiplicity $r_1, \cdots, r_l$. We will prove that if we solve the above system for $n = -1$ we will get the inverse of A. This linear system has a unique solution for every $n \in \mathbb{Z}$ because its determinant is a generalized Vandermonde determinant. In the case where $n = -1$ we will see that this unique solution is for $b_k = -\frac{a_{k+1}}{a_0}$ for $k = 0, \cdots, m-1$.

Indeed, the first equation of this linear system is

$$b_{m-1}k_1^{m-1} + \cdots + b_0 = \frac{1}{k_1}$$

Substituting the b_k we arrive at $\delta_A(k_1) = 0$ because k_1 is an eigenvalue.

The second equation is

$$(m-1)b_{m-1}k_1^{m-2} + \cdots + b_1 = -\frac{1}{k_1^2}$$

Substituting the $b_k = -\frac{a_{k+1}}{a_0}$ we arrive at

$$(m-1)\left(-\frac{a_m}{a_0}\right)k_1^{m-2} + \cdots + \left(-\frac{a_2}{a_0}\right) = -\frac{1}{k_1^2}$$

Therefore

$$(m-1)a_m k_1^{m-1} + \cdots + a_2 k_1 = \frac{a_0}{k_1}$$

We write the left hand side in the form

$$m a_m k_1^{m-1} + \cdots + 2a_2 k_1 + a_1 - (a_m k_1^{m-1} + a_{m-1}k_1^{m-2} + \cdots + a_2 k_1 + a_1)$$

The quantity $m a_m k_1^{m-1} + \cdots + 2a_2 k_1 + a_1$ equals $\delta_A^{'}(k_1)$ which is equal to zero because the eigenvalue k_1 is of multiplicity r_1 therefore it holds that $\delta_A^{(l)}(k_1) = 0$ for $l = 0, \cdots, r_1 - 1$.

We have proved that solving the linear system 3.1 for $n = -1$ we take the polynomial $v(k)$ which is such that $v(A) = A^{-1}$.

Suppose now that we want to compute the matrix A^{-k} for $k \in \mathbb{N}$. Set $B = A^k$ and we compute the powers of B which are such that $B^n = (A^k)^n$. The elements of B^n are of the form $b_{ij}(n)$ which are equals $a_{ij}(nk)$ since $(A^k)^n = A^{kn}$. Therefore the elements of B^{-1} are $b_{ij}(-1)$. Because $(A^k)^{-1} = A^{-k}$ it follows that $b_{ij}(-1) = a_{ij}(-k)$. Therefore the elements of the matrix A^{-k} are $a_{ij}(-k)$ when the elements of A^k are the $a_{ij}(k)$. □

Remark 3.4 (Generalized inverse of Drazin). As we can see, even in the case where the matrix A is not invertible, there always exist the matrix B with elements $a_{ij}(-1)$. In this case the matrix B coincides with the generalized inverse of Drazin (see the proof of theorem 3.2 and theorem 7.5.2 of [16]). Therefore, one way to compute the generalized inverse of Drazin is to calculate the matrix A^n and then get the matrix B with elements the $a_{ij}(-1)$. □

Example 3.3. *We will compute the n-th power of the matrix P where*

$$P = \begin{pmatrix} 3 & 0 \\ 1 & 1 \end{pmatrix}$$

by using the Cayley-Hamilton theorem. We compute the eigenvalues of the matrix which are $k_1 = 1$ and $k_2 = 3$. Next we write

$$k^n = t(k)\delta_P(k) + v(k)$$

By the Cayley-Hamilton theorem we have that $P^n = v(P)$. We have to compute the coefficients of the polynomial $v(k)$. In order to do this we construct the following linear system

$$\begin{aligned} 1^n &= a + b \\ 3^n &= 3a + b \end{aligned}$$

which have the solution $a = \frac{3^n-1}{2}$ and $b = \frac{3}{2}(1 - 3^{n-1})$. Thus

$$P^n = aP + bI_{2\times 2} = \begin{pmatrix} 3^n & 0 \\ \frac{3^n-1}{2} & 1 \end{pmatrix}$$

By using theorem 3.2 we can now compute the inverse of the matrix P which is

$$P^{-1} = \begin{pmatrix} 3^{-1} & 0 \\ \frac{3^{-1}-1}{2} & 1 \end{pmatrix}$$

□

Example 3.4. *We will compute the n-th power of the matrix*

$$P = \begin{pmatrix} -3 & 6 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

The eigenvalues are $k_1 = 3$ (double) and $k_2 = -5$.

Dividing k^n by the characteristic polynomial we obtain $k^n = t(k)\delta_P(k) + v(k)$ where $v(k) = ak^2 + bk + c$. In order to compute the unknown coefficients of $v(k)$ we will solve the following linear system

$$\begin{aligned} 3^n &= 9a + 3b + c \\ 3^{n-1}n &= 6a + b \\ (-1)^n 5 &= 25a - 5b + c \end{aligned}$$

Finally

$$P^n = aP^2 + bP + c = \begin{pmatrix} 21a - 3a + c & -12a + 6b & 0 \\ -4a + 2b & 13a + b + c & 0 \\ 0 & 0 & 9a + 3b + c \end{pmatrix}$$

□

Example 3.5. *Let the matrix*

$$A = \begin{pmatrix} k & 1 & 0 \\ 0 & k & 1 \\ 0 & 0 & k \end{pmatrix}$$

where $k \neq 0$. As before we have to compute the unknown coefficients of the polynomial $v(k) = b_2k^2 + b_1k + b_0$ in order to compute the matrix A^n. So we have to solve the following system

$$\begin{aligned} b_2k^2 + b_1k + b_0 &= k^n \\ 2b_2k + b_1 &= nk^{n-1} \\ 2b_2 &= n(n-1)k^{n-2} \end{aligned}$$

The solution is

$$\begin{aligned} b_0(n) &= k^n \frac{n^2 - 3n + 2}{2} \\ b_1(n) &= nk^{n-1}(2 - n) \\ b_2(n) &= \frac{n(n-1)}{2} k^{n-2} \end{aligned}$$

Therefore the matrix A^n has the form

$$A^n = b_2(n)A^2 + b_1(n)A + b_0(n)I_{3\times 3} = \begin{pmatrix} k^n & nk^{n-1} & \frac{1}{2}n(n-1)k^{n-2} \\ 0 & k^n & nk^{n-1} \\ 0 & 0 & k^n \end{pmatrix}$$

Using theorem 3.2 we can immediately compute the inverse of the matrix A which is

$$A^{-1} = \begin{pmatrix} \frac{1}{k} & -\frac{1}{k^2} & \frac{1}{k^3} \\ 0 & \frac{1}{k} & -\frac{1}{k^2} \\ 0 & 0 & \frac{1}{k} \end{pmatrix}$$

Similarly, we can have the matrix A^{-n} for every $n \in \mathbb{N}$. □

4 The Exponential Matrix

In this section we will study the function of a matrix and in particular the exponential matrix. We will give three definitions for the function of a matrix. The first one is by the diagonalization of the matrix, the second via the Taylor series of the function and the third one via the Cayley-Hamilton theorem. We shall prove that all the above definitions are equivalent. More about this topic can be found in [2].

• Let $A_{n\times n}$ a matrix which is diagonalizable, that is there exist an invertible matrix P such that $A = PDP^{-1}$ where the diagonal matrix D contains the eigenvalue. Let a function $f : \mathbb{C} \to \mathbb{C}$ such that it is well possed at the eigenvalues of the matrix A. Then by $f(A)$ we mean the matrix

$$f^D(A) := P \cdot \begin{pmatrix} f(k_1) & 0 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ \vdots & 0 & f(k_i) & 0 \\ 0 & \cdots & 0 & f(k_n) \end{pmatrix} \cdot P^{-1}$$

where $k_1, \cdots, k_n$ are the eigenvalues of the matrix A not necessarily distinguished.

• Let the function $f :\to$ which is such that all the derivatives are well defined in all of and that are absolutely bounded in every bounded interval (for example the function e^x). We define the matrix $f(A)$ as

$$f^{Taylor}(A) := \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} A^k$$

This matrix is well defined because the infinite series converges. Indeed, define by $\Delta = \max_{i,j} |A_{ij}|$ we have that

$$|(A^k)_{ij}| \leq n^{k-1}\Delta^k$$

We will prove it by induction. For $k = 1$ is obvious that $|(A^1)_{ij}| \leq n^0\Delta$. We assume that it holds for some k, that is $|(A^k)_{ij}| \leq n^{k-1}\Delta^k$ and we will prove that it holds for $k + 1$. We have that

$$|(A^{k+1})_{ij}| = |\sum_{l=1}^{n} (A^k)_{il} A_{lj}| \leq \sum_{l=1}^{n} n^{k-1}\Delta^k A_{lj} \leq n^k \Delta^{k+1}$$

Therefore

$$\left|\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} (A^k)_{ij}\right| \leq 1 + \left|\sum_{k=1}^{\infty} \frac{f^{(k)}(0)}{k!} (A^k)_{ij}\right| \leq 1 + \sum_{k=1}^{\infty} \frac{|f^{(k)}(0)|}{k!} n^{k-1}\Delta^k$$

That means that $\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} (A^k)_{ij}$ converges absolutely and therefore the matrix $f(A)$ is well defined.

• We will now describe the definition of the matrix $f(A)$ via the Cayley-Hamilton theorem.

Let $A_{n\times n}$ a matrix with characteristic polynomial $\delta_A(k)$.

In the case where the function $f : \mathbb{C} \to \mathbb{C}$ is not a polynomial we can approximate it with a polynomial $p(k)$ and then set $f(A) := p(A)$. We will describe next how this can be done.

Definition 4.1. Let $A_{n\times n}$ a matrix which has eigenvalues $k_1, \cdots, k_r$ of multiplicity $l_1, \cdots, l_r$ respectively. We say that a function $\phi : U \subseteq\to$ is well posed at the eigenvalues of the matrix A when there exist the derivatives $\phi^{(d)}(k_i)$ for $0 \leq d \leq l_1 - 1$. Let the minimum polynomial has roots the $k_1, \cdots, k_r$ of multiplicity $l_1^{'}, \cdots, l_r^{'}$ respectively. We say that the function ϕ is well posed on the roots of the minimum polynomial when there exist the derivatives $\phi^{(d)}(k_i)$ for $0 \leq d \leq l_1^{'} - 1$.

Let a function f which is well posed at the eigenvalues of the matrix A, which are $k_1, \cdots, k_m$ of multiplicity $r_1, \cdots, r_m$ respectively. By $P_{\sigma(A)}$ we denote all the polynomials $p(k)$ which have the property that $p^{(l_i)}(k_i) = f^{(l_i)}(k_i)$ for every $0 \leq l_i \leq r_i - 1$ where $l_i \in \mathbb{N}$ and for every eigenvalue k_i of the matrix A. Note that the set $P_{\sigma(A)}$ is nonempty (Hermite interpolation) and moreover it contains infinite many polynomials because in this set there exist polynomials that approximate the function f and in other points excluding the eigenvalues of A. Here by $f^{(l_i)}$ we mean the l_i-th derivative of the function.

The polynomials that belong in this set have a common property. Dividing any polynomial $p(k) \in P_{\sigma(A)}$ by the characteristic of A we get that

$$p(k) = \pi(k)\delta_A(k) + v(k) \tag{4.1}$$

where $v(k)$ is a polynomial of at most $n-1$ degree and is independent of $p(k) \in P_{\sigma(A)}$, which is very crucial in our point of view. Setting $k = A$ and using the fact that $\delta_A(A) = 0$ we arrive at $p(A) = v(A)$.

We can compute the polynomial $v(k)$ constructing n equations using the eigenvalues (and their multiplicity) of A. For example substituting k by the eigenvalue k_1 we obtain the equation

$$v(k_1) = p(k_1) = f(k_1)$$

If this eigenvalue has multiplicity two or more then we construct more equations by differentiation and so we have

$$v^{'}(k_1) = f^{'}(k_1)$$

supposing that the derivatives of f are well defined on the eigenvalues of A.

Therefore we construct the following linear system for the unknown polynomial $v(k) = b_{m-1}k^{m-1} + \cdots + b_0$,

$$\begin{aligned} v^{(i)}(k_1) &= f^{(i)}(k_1), \quad i = 0, \cdots, r_1 - 1 \\ &\vdots \\ v^{(i)}(k_m) &= f^{(i)}(k_m), \quad i = 0, \cdots, r_m - 1 \end{aligned}$$

where by $v^{(i)}$ and $f^{(i)}$ we mean the i-th derivative of the function. For $i = 0$ we mean the function itself. The above linear system has a unique solution because the determinant of the corresponding matrix is a generalized Vandermonde determinant.

Solving the above system we will compute the polynomial $v(k)$ and we define as the matrix $f(A)$ the matrix $v(A)$, that is

$$f^{Cayley}(A) := v(A)$$

Note that the matrix $v(A)$ depends only on the function f and the eigenvalues of A. Moreover, if the matrix A is a real matrix with eigenvalues $k_1, \cdots, k_r$ of multiplicity $l_1, \cdots, l_r$ and $f : \mathbb{C} \to \mathbb{C}$ is such that $f^{(q_i)}(\overline{k_i}) = \overline{f^{(q_i)}(k_i)}$ (as for example the function $f(z) = e^z$) for every $q_i \in \mathbb{N}$ such that $0 \leq q_i \leq l_i - 1$ and for every $i = 1, \cdots, r$ then the matrix $f(A)$ is real.

Example 4.1. *Let the matrix*

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

We will compute the exponential matrix e^{tA} for $t \in$. The eigenvalues of the matrix are $k_1 = i$ and $k_2 = -i$. We will compute the polynomial $v(k) = ak + b$ which is such that

$$\begin{aligned} ai + b &= e^{it} \\ a(-i) + b &= e^{-it} \end{aligned}$$

Then we can write

$$a = \frac{e^{it} - e^{-it}}{2i} = \sin t$$

Thus

$$b = \frac{e^{it} + e^{-it}}{2} = \cos t$$

So the exponential matrix is the matrix $e^{tA} = aA + bI = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}$. □

5 Equivalence of the Definitions of the Function Matrix

In this section we will prove that all the above definitions are equivalent. That is, we will prove that

$$f^D(A) = f^{Taylor}(A) = f^{Cayley}(A)$$

Suppose that the matrix $A_{n\times n}$ is diagonalizable with eigenvalues $k_1, \cdots, k_n$ not necessarily distinguished.

Let a function $f(x)$ which can be expanded in a Taylor series around zero and converges absolutely in every closed interval of the form $[-M, M]$, that is

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k, \text{ for every } x \in [-M, M]$$

Because the matrix is diagonalizable it holds that

$$\begin{aligned} f^D(A) &= P \cdot \begin{pmatrix} \sum_{l=0}^{\infty} \frac{f^{(l)}(0)}{l!} k_1^l & 0 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ \vdots & 0 & \sum_{l=0}^{\infty} \frac{f^{(l)}(0)}{l!} k_i^l & 0 \\ 0 & \cdots & 0 & \sum_{l=0}^{\infty} \frac{f^{(l)}(0)}{l!} k_n^l \end{pmatrix} \cdot P^{-1} \\ &= P \cdot \left(\sum_{l=0}^{\infty} \frac{f^{(l)}(0)}{l!} D^l \right) \cdot P^{-1} \\ &= \sum_{l=0}^{\infty} \frac{f^{(l)}(0)}{l!} A^l \qquad (5.1) \\ &= f^{Taylor}(A) \end{aligned}$$

On the other hand, in order to compute the matrix A^l by using the Cayley - Hamilton theorem we should compute the unique solution of the system

$$\begin{aligned} a_{n-1}(l)k_1^{n-1} + \cdots + a_0(l) &= k_1^l \\ &\vdots \\ a_{n-1}(l)k_n^{n-1} + \cdots + a_0(l) &= k_n^l \end{aligned} \qquad (5.2)$$

Therefore the form of A^l is

$$A^l = a_{n-1}(l)A^{n-1} + \cdots + a_0(l)I_{n\times n}, \quad \text{for } l = 0, 1, \cdots,$$

Remark 5.1. Note that when $l < n$ then the unique solution of the system is $a_l(l) = 1$ and all the other coefficients are all equal to zero.

Substituting on 5.1 we conclude that the matrix $f(A)$ via diagonalization or by the Taylor series is

$$f^D(A) = A^{n-1}\underbrace{\sum_{l=0}^{\infty}\frac{f^{(l)}(0)}{l!}a_{n-1}(l)}_{b_{n-1}} + \cdots + I\underbrace{\sum_{l=0}^{\infty}\frac{f^{(l)}(0)}{l!}a_0(l)}_{b_0} \tag{5.3}$$

We are allowed to rearrange the terms because we have proved the absolute convergence.

In order to prove that we obtain the same form of the matrix $f(A)$ via the Cayley-Hamilton theorem we should multiply every equation of the system 5.2 by $\frac{f^{(l)}(0)}{l!}$. Summing every equation for $l = 0$ to m we arrive at

$$\begin{aligned} k_1^{n-1}\underbrace{\sum_{l=0}^{m}\frac{f^{(l)}(0)}{l!}a_{n-1}(l)}_{b_{n-1}^m} + \cdots + \underbrace{\sum_{l=0}^{m}\frac{f^{(l)}(0)}{l!}a_0(l)}_{b_0^m} &= \sum_{l=0}^{m}\frac{f^{(l)}(0)}{l!}k_1^l \\ &\vdots \\ k_n^{n-1}\underbrace{\sum_{l=0}^{m}\frac{f^{(l)}(0)}{l!}a_{n-1}(l)}_{b_{n-1}^m} + \cdots + \underbrace{\sum_{l=0}^{m}\frac{f^{(l)}(0)}{l!}a_0(l)}_{b_0^m} &= \sum_{l=0}^{m}\frac{f^{(l)}(0)}{l!}k_1^l \end{aligned} \tag{5.4}$$

The coefficient b_i^m are obvious finite for every m. This linear system has a unique solution because the determinant of the matrix of the system is a generalized Vandermonde determinant (see [15]). Moreover, the matrix of the system is invertible and independent of m. Therefore it follows that

$$b_i^m = (K^{-1}\cdot F)_i \tag{5.5}$$

where K is the matrix of this linear system and F the right hand side of the system. Because $\sum_{l=0}^{m}\frac{f^{(l)}(0)}{l!}k_i^l$ converges to $f(k_i)$ as $m \to \infty$ it easily follows from the relation 5.5 that $b_i^m \to b_i = \sum_{l=0}^{\infty}\frac{f^{(l)}(0)}{l!}a_i(l) < \infty$ as $m \to \infty$.

Setting $m = \infty$ in the system 5.4 we compute the unique solution of the system that is $b_0, \cdots, b_{n-1}$.

Therefore the matrix $f(A)$ has the following form via the Cayley-Hamilton theorem

$$f^{Cayley}(A) = b_{n-1}A^{n-1} + \cdots + b_0 I \tag{5.6}$$

Noting the relation 5.3 we observe that the above system drive us to the matrix $f^D(A)$.

In the case where there exist some eigenvalues with multiplicity two or more then at the system 5.2 will appear equalities via the differentiation. Let the eigenvalue k_i is of multiplicity r. Then the system 5.2 contains the following equations,

$$\begin{aligned} a_{n-1}(l)k_i^{n-1} + \cdots + a_0(l) &= k_i^l \\ a_{n-1}(l)(n-1)k_i^{n-2} + \cdots + a_1(l) &= lk_i^{l-1} \\ &\vdots \\ a_{n-1}(l)(n-r)k_i^{n-1-r} + \cdots + a_r(l) &= l\cdot(l-1)\cdots(l-r+1)k_i^{l-r} \end{aligned}$$

We multiply these equations by $\frac{f^{(l)}(0)}{l!}$ and then we sum from $l=0$ to ∞. Therefore the system 5.4 will contain the following equations

$$\begin{aligned}
k_i^{n-1}\underbrace{\sum_{l=0}^{\infty} a_{n-1}(l)\frac{f^{(l)}(0)}{l!}}_{b_{n-1}}+\cdots+\underbrace{\sum_{l=0}^{\infty} a_0(l)\frac{f^{(l)}(0)}{l!}}_{b_0} &= \sum_{l=0}^{\infty} k_i^l\frac{f^{(l)}(0)}{l!}\\
(n-1)k_i^{n-1}\underbrace{\sum_{l=0}^{\infty} a_{n-1}(l)\frac{f^{(l)}(0)}{l!}}_{b_{n-1}}+\cdots+\underbrace{\sum_{l=0}^{\infty} a_1(l)\frac{f^{(l)}(0)}{l!}}_{b_1} &= \sum_{l=0}^{\infty} lk_i^l\frac{f^{(l)}(0)}{l!}\\
&\vdots\\
(n-r)k_i^{n-1-r}\underbrace{\sum_{l=0}^{\infty} a_{n-1}(l)\frac{f^{(l)}(0)}{l!}}_{b_{n-1}}+\cdots+\underbrace{\sum_{l=0}^{\infty} a_r(l)\frac{f^{(l)}(0)}{l!}}_{b_r} &= \\
\sum_{l=0}^{\infty} l\cdot(l-1)\cdots(l-r+1)k_i^{l-r}\frac{f^{(l)}(0)}{l!}
\end{aligned}$$

We observe that the right hand side of the above equations equals to $f(k_i), f^{'}(k_i), \cdots, f^{(r)}(k_i)$.

We will now explain the fact that we can use the minimum polynomial instead of the characteristic.

Let the minimum polynomial is of the form

$$m_A(k) = (k-k_1)^{d_1}\cdots(k-k_m)^{d_m}$$

where $d_i \leq r_i$ for $i=1,\cdots,m$. We define the set of polynomials $\hat{P}_{\sigma(A)}$ which are such that $p^{(l_i)}(k_i) = f^{(l_i)}(k_i)$ for every $0 \leq l_i \leq d_i - 1$ where $l_i \in \mathbb{N}$ and for every eigenvalue k_i of the matrix A. It holds that $P_{\sigma(A)} \subseteq \hat{P}_{\sigma(A)}$. The polynomials that belongs to the set $\hat{P}_{\sigma(A)}$ has the following property. If we divide any them by the minimum polynomial then $p(k) = m_A(k)\pi(k) + \hat{v}(k)$. The polynomial $\hat{v}(k)$ is independent of $p(k)$.

Remark 5.2. Denoting by $v(k)$ the polynomial which is such that

$$p(k) = \pi(k)\delta_A(k) + v(k)$$

and by $\hat{v}(k)$ the polynomial which is such that

$$p(k) = \hat{\pi}(k)m_A(k) + \hat{v}(k)$$

where $\delta_A(k)$ is the characteristic polynomial of A, $p(k) \in P_{\sigma(A)}$ and $m_A(k)$ the minimum polynomial of A, it follows that $v(A) = \hat{v}(A)$. Therefore the matrix $f(A)$, where $f : \mathbb{C} \to \mathbb{C}$ can be defined uniquely either by the characteristic or by the minimum polynomial. □

Let us summarize below the steps that we have to do in order to compute the matrices A^n and e^{tA} given a matrix $A_{m\times m}$.

(*i*) We set $f(x) = x^n$.

(*ii*) We compute the characteristic polynomial (or even better the minimum polynomial) in the form $(k-k_1)^{r_1}(k-k_2)^{r_2}\cdots(k-k_l)^{r_l}$ where $k_1,\cdots,k_l$ the eigenvalues of the matrix A.

(*iii*) We consider a polynomial $v(k) = a_{m-1}k^{m-1} + \cdots + a_1 k + a_0$ in the case were we want to use the characteristic polynomial. If we have computed the minimum polynomial and $r_1 + \cdots + r_l = q \leq m$ then $v(k) = a_{q-1}k^{q-1} + \cdots + a_0$ where $r_1, \cdots, r_l$ are the multiplicities of the eigenvalues concerning the minimum polynomial. We construct the following linear system

$$\begin{aligned}
v(k_1) = a_{m-1}k_1^{m-1} + \cdots + a_1 k_1 + a_0 &= f(k_1) \\
v^{'}(k_1) = (m-1)a_{m-1}k_1^{m-2} + \cdots + a_1 &= f^{'}(k_1) \\
&\vdots \\
v^{(r_1-1)}(k_1) &= f^{(r_1-1)}(k_1) \\
v(k_2) = a_{m-1}k_2^{m-1} + \cdots + a_1 k_2 + a_0 &= f(k_2) \\
v^{'}(k_2) = (m-1)a_{m-1}k_2^{m-2} + \cdots + a_1 &= f^{'}(k_2) \\
&\vdots \\
v^{(r_2-1)}(k_2) &= f^{(r_2-1)}(k_2) \\
&\vdots \\
v(k_l) = a_{m-1}k_l^{m-1} + \cdots + a_1 k_l + a_0 &= f(k_l) \\
v^{'}(k_l) = (m-1)a_{m-1}k_l^{m-2} + \cdots + a_1 &= f^{'}(k_l) \\
&\vdots \\
v^{(r_l-1)}(k_l) &= f^{(r_l-1)}(k_l)
\end{aligned}$$

(*iv*) We solve for the unknown coefficients $a_0, \cdots, a_{m-1}$. Then

$$A^n = a_{m-1}A^{m-1} + \cdots + a_1 A + a_0 I_{m\times m}$$

In order to compute the exponential matrix e^{tA} we just set $f(x) = e^{tx}$ and we follow the same procedure. All the above remain true even in the case where the matrix A is complex. □

Example 5.1. *Let the matrix*

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0.4 & 0.6 & 0 \\ 0 & 0.4 & 0.6 \end{bmatrix}$$

We will compute the matrices A^n and e^{tA} using the minimum polynomial of A.

In order to compute the minimum polynomial we construct the matrix $B_{9\times 4}$ in which at the first column we place the matrix $\mathbb{I}$, at the second column the matrix A and so on. We arrive at the matrix

$$B = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0.4 & 0.64 & 0.784 \\ 0 & 0 & 0.16 & 0.352 \\ 0 & 0 & 0 & 0 \\ 1 & 0.6 & 0.36 & 0.216 \\ 0 & 0.4 & 0.48 & 0.432 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0.6 & 0.36 & 0.216 \end{bmatrix}$$

Next we compute the reduced row echelon form of the matrix B which is

$$\hat{B} = \begin{bmatrix} 1.0 & 0.0 & 0.0 & 0.36 \\ 0.0 & 1.0 & 0.0 & -1.56 \\ 0.0 & 0.0 & 1.0 & 2.2 \\ 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 \end{bmatrix}$$

We see that the number of the leading ones is three so the degree of the minimum polynomial is also three. That is the degree of the minimum polynomial equals to the degree of the characteristic polynomial. We find the coefficients of the minimum polynomial at the next column after the last leading one. That is, the minimum polynomial is the following

$$-0.36 + 1.56k - 2.2k^2 + k^3$$

Note that the matrix is not diagonalizable.

Set now $f(x) = x^n$ and $g(x) = e^{tx}$. We construct the following linear system

$$\begin{aligned} a + b + c &= f(1) \quad (or)\ g(1) \\ 0.36a + 0.6b + c &= f(0.6) \quad (or)\ g(0.6) \\ 1.2a + b &= f^{'}(0.6) \quad (or)\ g^{'}(0.6) \end{aligned}$$

The matrix of this system is

$$F = \begin{bmatrix} 1 & 1 & 1 \\ 0.36 & 0.6 & 1 \\ 1.2 & 1 & 0 \end{bmatrix}$$

This matrix is invertible and the inverse is

$$F^{-1} = \begin{bmatrix} 6.25 & -6.25 & -2.5 \\ -7.5 & 7.5 & 4 \\ 2.25 & -1.25 & -1.5 \end{bmatrix}$$

Therefore (a, b, c) *will be equal to* $F^{-1} \cdot \begin{bmatrix} f(1) \\ f(0.6) \\ f^{'}(0.6) \end{bmatrix}$ *for the matrix* A^n *or* $F^{-1} \cdot \begin{bmatrix} g(1) \\ g(0.6) \\ g^{'}(0.6) \end{bmatrix}$ *for the matrix* e^{tA}.
For the matrix A^n *it follows that*

$$\begin{aligned} a &= 6.25 - 6.25 \cdot 0.6^n - 2.5 \cdot n \cdot 0.6^{n-1} \\ b &= -7.5 + 7.5 \cdot 0.6^n + 4 \cdot n \cdot 0.6^{n-1} \\ c &= 2.25 - 1.25 \cdot 0.6^n - 1.5 \cdot n \cdot 0.6^{n-1} \end{aligned}$$

and therefore $A^n = aA^2 + bA + c\mathbb{I}$, *that is*

$$A^n = \begin{bmatrix} 1.0 & 0.0 & 0.0 \\ 1.0 - 0.6^n & 0.6^n & 0.0 \\ 1. - 0.6^n - 0.4 \cdot n \cdot 0.6^{n-1} & 0.4 \cdot n \cdot 0.6^n & 0.6^n \end{bmatrix}$$

In order to check the validity of the above one can use induction. Moreover, substituting n *by* -1 *we take the inverse of the matrix* A *(see theorem 3.2).*

For the exponential matrix we have

$$\begin{aligned} a &= 6.25 \cdot e^t - 6.25 \cdot e^{0.6t} - 2.5 \cdot t \cdot e^{0.6t} \\ b &= -7.5 \cdot e^t + 7.5 \cdot e^{0.6t} + 4te^{0.6t} \\ c &= 2.25 \cdot e^t - 1.25 \cdot e^{0.6t} - 1.5 \cdot t \cdot e^{0.6t} \end{aligned}$$

Therefore $e^{tA} = aA^2 + bA + c\mathbb{I}$, *that is*

$$e^{tA} = \begin{bmatrix} e^t & 0.0 & 0.0 \\ e^t - e^{0.6t} & e^{0.6t} & 0.0 \\ e^t - e^{0.6t} - 0.4te^{0.6t} & 0.4\,t\mathrm{e}^{0.6\,t} & e^{0.6\,t} \end{bmatrix}$$

In order to check the validity of the above we can use the fact that the exponential matrix is the unique matrix $P(t)$ *that satisfies the following differential equation*

$$\begin{aligned} P^{'}(t) &= AP(t) \\ P(0) &= \mathbb{I} \end{aligned}$$

□

6 Leverrier - Faddeev Algorithm

In this section we will study the Leverrier - Faddeev algorithm for the computation of the characteristic polynomial of a matrix A. The construction of this algorithm that we are going to present comes from the paper [4] giving a simpler form.

Let the matrix $A_{m\times m}$ and $\delta_A(k) = k^m + a_1k^{m-1} + \cdots + a_m$ its characteristic polynomial. Let the matrices $N_1, \cdots, N_m$ (that we are going to compute them below) such that

$$(k\mathbb{I} - A)N(k) = \mathbb{I}\delta_A(k) \tag{6.1}$$

where $N(k) = N_1k^{m-1} + \cdots + N_m$. We consider the family of matrices $B(s) = s\mathbb{I} - A$. Note that if the eigenvalues of A are $k_1, \cdots, k_m$ (not necessarily distinguished) then the eigenvalues of $B(s)$ are the $\lambda_i = s - k_i$. Indeed, because

$$|\lambda\mathbb{I} - B(s)| = |\lambda\mathbb{I} - sI + A| = (-1)^m|(s-\lambda)\mathbb{I} - A|$$

then we easily get the desired result.

The matrix $B(s)$ is invertible if and if $\lambda_i = s - k_i \neq 0$ for $i = 1, \cdots, m$. In this case (that is $s \neq k_i$ for $i = 1, \cdots, m$) the inverse is $(s\mathbb{I} - A)^{-1}$ with eigenvalues $\frac{1}{s-k_i}$.

Then it holds that

$$\frac{N(s)}{\delta_A(s)} = (s\mathbb{I} - A)^{-1}, \quad \text{when } s \neq k_i \tag{6.2}$$

where k_i for $i = 1, \cdots, m$ the eigenvalues of the matrix A.

Moreover

$$trace(s\mathbb{I} - A)^{-1} = \frac{1}{s-k_1} + \cdots + \frac{1}{s-k_m} = \frac{\delta_A^{'}(s)}{\delta_A(s)} \tag{6.3}$$

But

$$A = s\mathbb{I} - (s\mathbb{I} - A)$$

If $s \neq k_i$ we multiply by $(s\mathbb{I} - A)^{-1}$ and we have that

$$s(s\mathbb{I} - A)^{-1} - \mathbb{I} = A(s\mathbb{I} - A)^{-1}$$

Then

$$s\frac{\delta_A^{'}(s)}{\delta_A(s)} - m = trace\frac{AN(s)}{\delta_A(s)}, \quad s \neq k_i$$

from which it follows that

$$\begin{aligned}&-a_1s^{m-1} - 2a_2s^{m-2} - \cdots - (m-1)a_{m-1}s - ma_m = \\ &(traceAN_1)s^{m-1} + (traceAN_2)s^{m-2} + \cdots + (traceAN_{m-1})s + traceAN_m\end{aligned}$$

Thus

$$a_i = -\frac{1}{i}traceAN_i, \quad i = 1, 2, \cdots, m$$

In order to compute the matrix N_i we will use the relation

$$(k\mathbb{I} - A)N(k) = \mathbb{I}\delta_A(k)$$

From this relation it follows that

$$N_1k^m + N_2k^{m-1} + \cdots + N_mk - (AN_1k^{m-1} + AN_2k^{m-2} + \cdots + AN_m) = \mathbb{I}\delta_A(k)$$

Therefore we arrive at the following algorithm

$$\begin{array}{rclrcl} N_1 & = & \mathbb{I}, & a_1 & = & -\frac{1}{1}traceAN_1 \\ N_2 & = & AN_1 + a_1\mathbb{I}, & a_2 & = & -\frac{1}{2}traceAN_2 \\ & \vdots & & & \vdots & \\ N_m & = & AN_{m-1} + a_{m-1}\mathbb{I}, & a_m & = & -\frac{1}{m}traceAN_m \\ 0 & = & AN_m + a_m\mathbb{I} & & & \end{array}$$

The last equation can be used for the computation of the inverse of A. Indeed, $AN_m = -a_m\mathbb{I}$ and therefore $N_m = -a_m A^{-1}$. Substituting for N_m at the last equation we get

$$A^{-1} = -\frac{1}{a_m}AN_{m-1} - \frac{a_{m-1}}{a_m}\mathbb{I}$$

□

Remark 6.1 (Computing the minimum and characteristic polynomial). If we want to compute the minimum and the characteristic polynomial we can use the Leverrier - Faddeev Algorithm. We construct the matrix B as we have described (in order to compute the minimum polynomial) and then the reduced row echelon form $\hat{B}$. At this stage we can compute the minimum polynomial of degree r but using the Leverrier - Faddeev Algorithm we can compute also the characteristic. Using this algorithm we can find the coefficients $a_{r+1}, \cdots, a_{m-1}$ of the characteristic $\delta_A(k) = k^m + a_{m-1}k^{m-1} + \cdots + a_0$. Then, by using the reduced row echelon form of B, i.e. the matrix $\hat{B}$, we compute the rest of the coefficients. □

7 Conclusion

In this review article we discuss the computation of the minimum polynomial of a matrix giving a simpler proof than the [1]. Next we show how to compute the nth power of a matrix and the exponential matrix using the Cayley-Hamilton theorem. Using this method it is not necessary the matrix to be diagonizable. By this method we can characterize the elements of A^{-n} when A is invertible and we see that these elements are the $a_{ij}(-n)$ if the $a_{ij}(n)$ are the elements of A^n. But even if the matrix is not invertible the matrix B whose elements are the $a_{ij}(-n)$ is well defined and appears to be the Drazin inverse of the matrix. Finally, we give a simpler presentation of the Leverrier - Faddeev algorithm.

Disclaimer

This paper is an extended version of a preprint document of the same author.
The preprint document is available in this link: file:///C:/Users/Editor-28/Desktop/Downloads/preprints202112.0329.v1.pdf
[As per journal policy, pre-print article can be published as a journal article, provided it is not published in any other journal].

Competing Interests

Author has declared that no competing interests exist.

References

[1] Bialas S, Bialas M. An algorithm for the calculation of the minimal polynomial. Bulletin Polish Academy Sciences. 2008;56.

[2] Higham N. Functions of matrices, theory and computation. SIAM; 2008.

[3] Israel A, Grevill T. Genelalized inverses, theory and applications. Springer; 2003.

[4] Shui-Hung Hou. A simple proof of the Leverrier - Faddeev characteristic polynomial algorithm. SIAM REV. 1998;40(3):706-709.

[5] Brand L. The companion matrix and its Properties. The American Mathematical Monthly. 1964;71(6):629-634.

[6] Berberian S. Linear Algebra, Oxford Science Publications; 1992.

[7] Davis P. Interpolation and approximation. Dover; 1975.

[8] Larson R, Falvo D. Elementary linear algebra. Houghton Mifflin Harcourt Publishing Company; 2009.

[9] Lipschutz S, Lipson M. Linear algebra. McGraw Hill; 2009.

[10] Moler C, Van Loan C. Nineteen dubious ways to compute the exponential of a matrix, twenty - five years later. SIAM Review; 2003.

[11] Penrose R. A Generalized Inverse for matrices., Mathematical Proceedings of the Cambridge Philosophical Society. 1955;406-413:51.

[12] Ricardo H. A Modern Introduction to Linear Algebra. CRC Press; 2010.

[13] Silvester JR. Determinants of block Matrices. Math. Gaz. 2000;84(501):460-467.

[14] Shoup V. Efficient Computation of Minimal Polynomials in Algebraic Extensions of Finite Fields; 1999.

[15] Flowe R, Harris G. A note on genelalized Vandermonde determinants. SIAM J. Matrix Anal. Appl. 1983;14(4):1146-1151.

[16] Cambell S, Meyer C. Generelized Inverses of Linear Trasformations. SIAM; 2009.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
https://www.sdiarticle5.com/review-history/92844