Universitat de Barcelona

Facultat de Matemàtiques i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

DIAGONALIZATION OF POLYNOMIAL MATRICES

Autora: Clara Molins Rosés

Directora: Dra. Maria Eulàlia Montoro Lopez
Realitzat a: Departament de Matemàtiques i Informàtica

Barcelona, June 13, 2022

Abstract

This work studies the diagonalization of second degree polynomial matrices. First, all the concepts needed to understand the theory on this type of matrix are defined. Then, the most important working tools for solving the problem are introduced: the Smith canonical form and the linearization of polynomial matrices. Finally, it is deduced for which 2nd degree matrices there is a diagonalization.

Resum

En aquest treball s'estudia la diagonalització de matrius polinomials de segon grau. En primer lloc, es defineixen tots els conceptes necessaris per entendre la teoria d'aquest tipus de matrius. A continuació s'introdueixen les eines de treball més importants que permeten resoldre el problema: la forma canònica de Smith i la linearització de matrius polinomials. Finalment, es dedueix per a quines matrius de 2n grau existeix una diagonalització.

2010 Mathematics Subject Classification. 15A18, 15A20, 15A22

Acknowledgements

First of all, I would like to give special thanks to Dr. Eulàlia Montoro for encouraging me to delve deeper into this field of mathematics. Without her support and her easy explanations, this project could not have been possible.

I want to extend my gratitude to my family, especially my parents, for their advice and their unconditional support.

Finally, I would like to thank all the people who have accompanied me for the last four years.

Contents

Chapter 1

Introduction

Polynomial matrices are a very useful tool in many areas. Several systems in a variety of disciplines are described by matrix polynomials [12]. This work is focused on the diagonalization of this type of matrices, particularly in diagonalization of second degree polynomial matrices.

Before digging into the mathematical study, let us name some applications of polynomial matrices to highlight their importance and to understand the motivation of this study. To start, many differential equations can be written as polynomial matrices. Two important areas where second order differential equations arise are the fields of mechanical and electrical oscillation. They are also very useful in signal processing and control theory [12]. Therefore, having a deep understanding of their properties, including their eigenvalue structure, helps solve problems, as can be seen in [12], more straightforwardly.

The term matrix was, according to Encyclopedia Britannica, first introduced by the 19th-century English mathematician James Sylvester. Though it was Arthur Cayley, a friend of his, who developed the algebraic aspect of matrices in two papers in the 1850s. The first theory on polynomial matrices appeared in the following decades. As a reference, we know the Smith canonical form for this type of matrices was obtained by F.G. Frobenius in 1878.

These matrices share a lot of properties with matrices whose elements are coefficients on a field $\mathbb{F}$ and with polynomials with coefficients on a field $\mathbb{F}$ as well. Their common ground and similarities have helped construct most of what we know about them nowadays.

However, the development of the theory is far from perfect and many problems remain unresolved. One of these is their diagonalization. Although this problem has been tackled by various mathematicians in recent years and it has been solved for matrices of first, second, third and fourth degree, it remains open for higher degrees.

This work is focused on second degree diagonalization [10], [13] and is divided mainly into two blocks. We will begin by illustrating the theory on which the problem is based and then, we will focus on its solution.

The general structure of the report is the following: The notation is introduced in Chapter 2. Chapter 3 includes the definition and the main aspects of polynomial matrices and describes the basic arithmetics.

In Chapter 4, the canonical form into which a polynomial matrix can be transformed is presented, precisely the Smith canonical form. It also includes relevant concepts such as equivalence and similarity and introduces invariant polynomials and elementary divisors. In Chapter 5, before digging into the main matter of the work, the linearization of polynomial matrices is explained.

Once the theory is constructed, the diagonalization of second degree polynomial matrices will be discussed in Chapter 6.

Chapter 2

Notation

Before delving into the subject, let us introduce the notation and terminology that will be used.

- Unless otherwise stated we assume that all the matrices are defined in the field $\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$.
- We will denote by $\mathbb{F}[\lambda]$ the set of polynomials with coefficients in $\mathbb{F}$.
- $\mathbb{F}^{n \times m}$ stands for n x m matrices with elements in $\mathbb{F}$ and $\mathbb{F}^{n \times m}[\lambda]$ stands for n x m polynomial matrices whose components are polynomials in $\mathbb{F}$.
- The elements of $A(\lambda) \in \mathbb{F}^{n \times m}[\lambda]$ are denoted by $a_{ij}(\lambda)$, $1 \leq i \leq n$, $1 \leq j \leq m$. If n = m, its *adjoint* matrix is expressed by $\mathrm{Adj}(A(\lambda))$, its *inverse* is denoted by $A^{-1}(\lambda)$ and its *determinant* by $\det(A(\lambda))$. In addition, we will denote by $\deg(A(\lambda))$, $\mathrm{rank}(A(\lambda))$ and $\dim(A(\lambda))$, its degree, rank and dimension respectively.
- Given $A \in \mathbb{F}^{n \times n}[\lambda]$ we represent its *transpose* by A^T. If, in particular, $\mathbb{F} = \mathbb{C}$ then we denote by A^* its *conjugate transpose.*
- We denote by I the n x n identity matrix. When denoting the identity matrix of any other order, say k, we will denote it by I_k.

Chapter 3

Polynomial matrices

3.1 Definition and main aspects

Given $\mathbb{F} = \mathbb{R}, \mathbb{C}$, a *polynomial matrix or λ-matrix*, is a rectangular matrix A(λ) = $(a_{ij}(\lambda)) \in \mathbb{F}^{n\times m}[\lambda]$ whose elements $a_{ij}(\lambda)$ are polynomials in λ. In this work we will only consider square matrices (i.e. n = m).

Now, let us consider

$$A(\lambda) = \begin{pmatrix} a_{11}(\lambda) & a_{12}(\lambda) & \cdots & a_{1n}(\lambda) \\ a_{21}(\lambda) & a_{22}(\lambda) & \cdots & a_{2n}(\lambda) \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}(\lambda) & a_{n2}(\lambda) & \cdots & a_{nn}(\lambda) \end{pmatrix} \in \mathbb{F}^{n\times n}[\lambda]$$

.

Observe that when the elements of A(λ) are evaluated for a particular value of λ, say $\lambda = \lambda_0$, then A(λ_0) $\in \mathbb{F}^{n\times n}$.

This matrices are not only called polynomial matrices because their elements are polynomials but also because they can also be written in the following form

$$A(\lambda) = A_0 + A_1\lambda + ... + A_r\lambda^r \in \mathbb{F}^{n\times n}[\lambda]$$

where $A_i \in \mathbb{F}^{n\times n}$ and r = max $(deg(a_{ij}(\lambda)))$.
The degree of a λ-matrix A(λ) is defined to be the greatest degree of the polynomials it contains and is denoted deg(A(λ)), therefore r = deg(A(λ)).
If $A_r = I$, then A(λ) is said to be monic.

Example 3.1.

$$A(\lambda) = \begin{pmatrix} 3\lambda^3 & \lambda & 1 \\ 0 & \lambda^2 & 0 \\ \lambda^3 & \lambda & \lambda \end{pmatrix} =$$

$$= \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \lambda^3 + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \lambda^2 + \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} \lambda + \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

In this case, $\deg(A(\lambda)) = 3$.

Remark 3.2. Matrices whose elements are scalars, $A \in \mathbb{F}^{n\times n}$, can be viewed as λ-matrices with zero degree.

Definition 3.3. *$A(\lambda) \in \mathbb{F}^{n\times n}[\lambda]$ is said to be* nonsingular *if $det(A(\lambda)) \not\equiv 0$. Else if $det(A(\lambda)) \equiv 0$, then $A(\lambda)$ is said to be* singular.

Example 3.4. Let

$$A(\lambda) = \begin{pmatrix} \lambda+1 & \lambda+3 \\ \lambda^2+3\lambda+2 & \lambda^2+5\lambda+4 \end{pmatrix} \in \mathbb{F}^{2\times 2}[\lambda]$$

then

$$det(A(\lambda)) = (\lambda+1)(\lambda^2+5\lambda+4) - (\lambda^2+3\lambda+2)(\lambda+3) = -2\lambda-2 \not\equiv 0.$$

Therefore, $A(\lambda)$ is nonsingular.

However, the matrix

$$B(\lambda) = \begin{pmatrix} \lambda+1 & \lambda+3 \\ \lambda^2+3\lambda+2 & \lambda^2+5\lambda+6 \end{pmatrix} \in \mathbb{F}^{2\times 2}[\lambda]$$

is singular because

$$det(B(\lambda)) = (\lambda+1)(\lambda^2+5\lambda+6) - (\lambda^2+3\lambda+2)(\lambda+3) \equiv 0.$$

Definition 3.5. *A polynomial matrix $A(\lambda) \in \mathbb{F}^{n\times n}[\lambda]$ is said to be* invertible *if there is a λ-matrix $B(\lambda) \in \mathbb{F}^{n\times n}[\lambda]$ such that $A(\lambda)B(\lambda) = B(\lambda)A(\lambda) = I$, i.e. $B(\lambda) = A(\lambda)^{-1} \in \mathbb{F}^{n\times n}[\lambda]$.*

Proposition 3.6. *([9]) A λ-matrix $A(\lambda) \in \mathbb{F}^{n\times n}[\lambda]$ is invertible if and only if $det(A(\lambda)) \in \mathbb{F} \setminus \{0\}$.*

Proof. If $det(A(\lambda)) = c \neq 0$, then the entries of $A^{-1}(\lambda) = \frac{1}{det(A(\lambda))}\mathrm{Adj}(A(\lambda))^T$ are equal to the minors of $A(\lambda)$ of order n - 1 divided by c $\neq$ 0 and hence are polynomials in λ. Thus $A(\lambda)^{-1} \in \mathbb{F}^{n\times n}[\lambda]$. Conversely, if $A(\lambda)$ is invertible, then

$$A(\lambda)A^{-1}(\lambda) = I \Longrightarrow det(A(\lambda))\frac{1}{det(A(\lambda))} = 1$$

Thus, $\det(A(\lambda)) = c$, with $c \neq 0$. □

A λ-matrix $A(\lambda) \in \mathbb{F}^{n\times n}[\lambda]$ such that $\det(A(\lambda)) \in \mathbb{F} \setminus \{0\}$ is also referred to as a *unimodular* λ-matrix.

Example 3.7. Let

$$A(\lambda) = \begin{pmatrix} 1 & \lambda & -2\lambda^2 \\ 0 & 1 & \lambda^4 \\ 0 & 0 & 1 \end{pmatrix} \in \mathbb{F}^{3\times 3}[\lambda]$$

then, $\det A(\lambda) = 1$ and therefore $A(\lambda)$ is unimodular. In this case,

$$A^{-1}(\lambda) = \begin{pmatrix} 1 & -\lambda & \lambda^5+2\lambda^2 \\ 0 & 1 & -\lambda^4 \\ 0 & 0 & 1 \end{pmatrix} \in \mathbb{F}^{3\times 3}[\lambda].$$

3.2 Arithmetic with matrix polynomials

Let us now define the basic arithmetic operations for polynomial matrices.

3.2.1 Sum

Let A(λ) = $\sum_{i=0}^{l} \lambda^i A_i \in \mathbb{F}^{n\times n}[\lambda]$ with deg(A(λ)) = l, and B(λ) = $\sum_{i=0}^{m} \lambda^i B_i \in \mathbb{F}^{n\times n}[\lambda]$ with deg(B(λ)) = m, then obviously

$$\mathrm{A}(\lambda) + \mathrm{B}(\lambda) = \sum_{i=0}^{max(l,m)} \lambda^i (A_i + B_i) \in \mathbb{F}^{n\times n}[\lambda],$$

Thus, deg(A(λ) + B(λ)) $\leq$ max(l, m).

3.2.2 Product

Considering A(λ) $\in \mathbb{F}^{n\times n}[\lambda]$ and B(λ) $\in \mathbb{F}^{n\times n}[\lambda]$ as above, we define the product of polynomial matrices as follows.

$$A(\lambda)B(\lambda) = \sum_{i=0}^{l} \sum_{j=0}^{m} A_i B_j \lambda^{i+j} \in \mathbb{F}^{n\times n}[\lambda]$$

and deg(A(λ)B(λ)) $\leq$ l + m. Clearly,

$$\det(\mathrm{A}(\lambda)) \neq 0 \text{ or } \det(\mathrm{B}(\lambda)) \neq 0 \implies \deg(\mathrm{A}(\lambda)\mathrm{B}(\lambda)) = \mathrm{l} + \mathrm{m}.$$

As previously stated, we are only considering square matrices on this paper. However, the reader should know the product is also applicable to non-square matrices. In addition, it shall be noted that the product is not commutative.

3.2.3 Division

Let A(λ) $\in \mathbb{F}^{n\times n}[\lambda]$ with deg(A(λ)) = l, and let B(λ) $\in \mathbb{F}^{n\times n}[\lambda]$ with deg(B(λ)) = m and det(B_l) $\neq$ 0. Suppose that there exist Q(λ), R(λ) $\in \mathbb{F}^{n\times n}[\lambda]$, with R($\lambda$) $\equiv$ 0 or deg(R(λ)) $\leq$ m, such that

$$\mathrm{A}(\lambda) = Q(\lambda)B(\lambda) + R(\lambda).$$

We call Q(λ) a *right quotient* of A(λ) on division by B(λ) and R(λ) is a *right remainder* of A(λ) on division by B(λ).

Similarly, $\bar{Q}$, $\bar{R}(\lambda) \in \mathbb{F}^{n\times n}[\lambda]$ are respectively a *left quotient* and *left remainder* of A(λ) on division by B(λ) if

$$A(\lambda) = B(\lambda)\bar{Q}(\lambda) + \bar{R}(\lambda)$$

with $\bar{R}(\lambda) \equiv 0$ or deg($\bar{R}(\lambda)$) $\leq$ m.

If R(λ) $\equiv$ 0, then Q(λ) is said to be a *right divisor* of A(λ) on division by B(λ). Similarly if $\bar{R}(\lambda) \equiv 0$ then $\bar{Q}(\lambda)$ is said to be a *left divisor* of A(λ) on division by B(λ).

Example 3.8. Let us examine the right and left quotients and remainders of A(λ) on division by B(λ), where

$$A(\lambda) = \begin{pmatrix} \lambda^4 + \lambda^2 + \lambda - 1 & \lambda^3 + \lambda^2 + \lambda + 2 \\ 2\lambda^3 - \lambda & 2\lambda^2 + 2\lambda \end{pmatrix} \in \mathbb{F}^{2\times 2}[\lambda]$$

and

$$B(\lambda) = \begin{pmatrix} \lambda^2 + 1 & 1 \\ \lambda & \lambda^2 + \lambda \end{pmatrix} \in \mathbb{F}^{2\times 2}[\lambda]$$

Note first that B(λ) has an invertible leading coefficient.

$$B_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad det(B_2) = 1$$

Now observe let us calculate the right and left quotient and remainders.

Let us start with the right division. We want to find matrices Q(λ), R(λ) $\in$ $\mathbb{F}^{2\times 2}[\lambda]$ such that

$$A(\lambda) = Q(\lambda)B(\lambda) + R(\lambda)$$

$$\iff A(\lambda) = \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix} B(\lambda) + \begin{pmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{pmatrix}$$

$$\iff \begin{cases} \lambda^4 + \lambda^2 + \lambda - 1 = q_{11}(\lambda^2 + 1) + q_{12}\lambda + r_{11} \\ \lambda^3 + \lambda^2 + \lambda + 2 = q_{11} + q_{12}(\lambda^2 + \lambda) + r_{12} \\ 2\lambda^3 - \lambda = q_{21}(\lambda^2 + 1) + q_{22}\lambda + r_{21} \\ 2\lambda^2 + 2\lambda = q_{21} + q_{22}(\lambda^2 + \lambda) + r_{22} \end{cases}$$

Solving this system it is found that

$$A(\lambda) = \begin{pmatrix} \lambda^2 - 1 & \lambda - 1 \\ 2\lambda & 2 \end{pmatrix} \begin{pmatrix} \lambda^2 + 1 & 1 \\ \lambda & \lambda^2 + \lambda \end{pmatrix} + \begin{pmatrix} 2\lambda & 2\lambda + 3 \\ -5\lambda & -2\lambda \end{pmatrix} = Q(\lambda)B(\lambda) + R(\lambda),$$

therefore $Q(\lambda)$ is not a right divisor of A(λ).

Following a similar process it can be seen that

$$A(\lambda) = \begin{pmatrix} \lambda^2 + 1 & 1 \\ \lambda & \lambda^2 + \lambda \end{pmatrix} \begin{pmatrix} \lambda^2 & \lambda + 1 \\ \lambda - 1 & 1 \end{pmatrix} = B(\lambda)\bar{Q}(\lambda).$$

Thus, $\bar{Q}(\lambda)$ is a left divisor of A(λ).

We must now prove that given two matrices A(λ), B(λ) $\in \mathbb{F}^{n\times n}[\lambda]$, there do exist quotients and remainders as defined. When we have done this we shall also prove their uniqueness.

The proof of the next theorem is a generalization of the division algorithm for scalar polynomials.

Theorem 3.9. *([9]) Let $A(\lambda) = \sum_{i=0}^{l} \lambda^i A_i \in \mathbb{F}^{n\times n}[\lambda]$, $B(\lambda) = \sum_{i=0}^{m} \lambda^i B_i \in \mathbb{F}^{n\times n}[\lambda]$ with deg($A(\lambda)$) = l, deg($B(\lambda)$) = m and det $B_m \neq 0$. Then there exists a right quotient and right remainder of $A(\lambda)$ on division by $B(\lambda)$ and similarly for a left quotient and left remainder.*

Proof. If l < m, we have only to take Q(λ) = 0 and R(λ) = A(λ) to obtain the result.
If l $\geq$ m, we first "divide by" the leading term of B(λ): $B_m\lambda^m$. Observe that the term of highest degree of $A_l B_m^{-1}\lambda^{l-m}B(\lambda)$ is just $A_l\lambda^l$. Hence

$$A(\lambda) = A_l B_m^{-1}\lambda^{l-m}B(\lambda) + A^{(1)}(\lambda),$$

where $A^{(1)}(\lambda) \in \mathbb{F}^{n\times n}[\lambda]$ and $\deg(A^{(1)}(\lambda)) = l_1 \leq l - 1$.
Writing $A^{(1)}(\lambda)$ in decreasing powers, let

$$A^{(1)}(\lambda) = A_{l_1}^{(1)}\lambda^{l_1} + ... + A_0^{(1)}, \quad A_{l_1}^{(1)} \neq 0, \quad l_1 < l.$$

If $l_1 \geq$ m we repeat the process, but on $A^{(1)}(\lambda)$ rather than A(λ) to obtain

$$A^{(1)}(\lambda) = A_{l_1}^{(1)}B_m^{-1}\lambda^{l_1-m}B(\lambda) + A^{(2)}(\lambda)$$

where

$$A^{(2)}(\lambda) = A_{l_2}^{(2)}\lambda^{l2} + ... + A_0^{(2)}, \quad A_{l_2}^{(2)} \neq 0, \quad l_2 < l_1.$$

In this manner we can construct a sequence of matrix polynomials A(λ), $A^{(1)}(\lambda)$, $A^{(2)}(\lambda)$, ... whose degrees are strictly decreasing, and after a finite number of terms we arrive at a matrix polynomial $A^{(r)}(\lambda)$ of degree $l_r <$ m, with $l_{r-1} \geq$ m. Then, if we write A(λ) = $A^{(0)}(\lambda)$, we have that

$$A^{(s-1)}(\lambda) = A_{l_{s-1}}^{(s-1)}B_m^{-1}\lambda^{l_{s-1}-m}B(\lambda) + A^{(s)}(\lambda), \quad s = 1, 2, ..., r.$$

Combining these equations, we obtain

$$A(\lambda) = (A_l B_m^{-1}\lambda^{l-m} + A_{l_1}^{(1)}B_m^{-1}\lambda^{l_1-m} + ... + A_{l_{r-1}}^{(r-1)}B_m^{-1}\lambda^{l_{r-s}-m})B(\lambda) + A^{(r)}(\lambda).$$

The matrix in parentheses can now be identified as the right quotient of A(λ) on division by B(λ), and $A^{(r)}(\lambda)$ as the right remainder.

The existence of a left quotient and left remainder can be similarly proved. □

Theorem 3.10. *([9]) With the hypotheses of Theorem 3.9, the right quotient, right remainder, left quotient and left remainder are each unique.*

Proof. Let us suppose that there exist matrix polynomials Q(λ), R(λ) and $\tilde{Q}(\lambda)$, $\tilde{R}(\lambda)$ such that

$$A(\lambda) = Q(\lambda)B(\lambda) + R(\lambda)$$

and

$$A(\lambda) = \tilde{Q}(\lambda)B(\lambda) + \tilde{R}(\lambda)$$

where deg(R(λ)) < m and deg($\tilde{R}(\lambda)$) < m. Then

$$(Q(\lambda) - \tilde{Q}(\lambda))B(\lambda) = \tilde{R}(\lambda) - R(\lambda).$$

If Q(λ) ≠ $\tilde{Q}(\lambda)$, then deg((Q(λ) - $\tilde{Q}(\lambda)$)$B(\lambda)$) ≥ m. However, deg($\tilde{R}(\lambda) - R(\lambda)$) < m. Therefore, the equation above does not hold and the uniqueness is proved.

A similar argument can be used to establish the uniqueness of the left quotient and left remainder. □

Division by a linear divisor

Now we consider the special case in which a divisor is *linear* (i.e. a matrix polynomial of first degree).
First, note that when discussing a scalar polynomial p(λ) ∈ $\mathbb{F}[\lambda]$, we may write

$$p(\lambda) = a_1\lambda^l + a_{l-1}\lambda^{l-1} + ... + a_0 = \lambda^l a_l + \lambda^{l-1}a_{l-1} + ... + a_0.$$

For a matrix polynomial with a matrix argument, this is not generally possible.
If A(λ) ∈ $\mathbb{F}^{n\times n}[\lambda]$ and B ∈ $\mathbb{F}^{n\times n}$, we define the *right value* A(B) of A(λ) at B by

$$A(B) = A_l B^l + A_{l-1}B^{l-1} + ... + A_0 \in \mathbb{F}^{n\times n}$$

and the *left value* $\bar{A}$(B) of A(λ) at B by

$$\bar{A}(B) = B^l A_l + B^{l-1}A_{l-1} + ... + A_0 \in \mathbb{F}^{n\times n}.$$

The reader should be familiar with the classical *remainder theorem*:

Theorem 3.11. *([9]) On dividing the scalar polynomial p(λ) ∈ $\mathbb{F}[\lambda]$ by λ - b, the remainder is p(b).*

We now prove an extension of this result to matrix polynomials. Note first that λI - B ∈ $\mathbb{F}^{n\times n}[\lambda]$ is monic.

Theorem 3.12. *([9]) The right and left remainders of $A(\lambda) = \sum_{i=0}^{l}\lambda^i A_i \in \mathbb{F}^{n\times n}[\lambda]$ on division by λI - B are A(B) and $\bar{A}$(B), respectively.*

Proof. The factorization

$$\lambda^j I - B^j = (\lambda^{j-1}I + \lambda^{j-2}B + ... + \lambda B^{j-2} + B^{j-1})(\lambda I - B)$$

can be verified by multiplying out the product on the right. Premultiply both sides of this equation by A_j and sum the resulting equation for j = 1, ..., l. The right-hand side of the equation obtained is of the form C(λ)(λI - B), where C(λ) = $\sum_{j=1}^{l} A_j(\lambda^{j-1}I + \lambda^{j-2}B + \cdots + \lambda B^{j-2} + B^{j-1})$. The left-hand side is

$$\sum_{j=1}^{l} A_j\lambda^j - \sum_{j=1}^{l} A_j B^j = \sum_{j=0}^{l} A_j\lambda^j - \sum_{j=0}^{l} A_j B^j = A(\lambda) - A(B).$$

Thus,

$$A(\lambda) = C(\lambda)(\lambda I - B) + A(B)$$

The result now follows from the uniqueness of the right remainder on division of A(λ) by (λI - B). The result for the left remainder is obtained by reversing the factors in the initial factorization, multiplying on the right by A_j, and summing. □

Definition 3.13. *Suppose $A(\lambda) \in \mathbb{F}^{n\times n}[\lambda]$. A matrix $X \in \mathbb{F}^{n\times n}$ such that $A(X)$ = 0 (respectively, $\bar{A}(X)$ = 0) is referred to as a* right (respectively, left) solvent *of $A(\lambda)$.*

Corollary 3.14. *([9]) A polynomial matrix $A(\lambda) \in \mathbb{F}^{n\times n}[\lambda]$ is divisible on the right (respectively, left) by λI - $B \in \mathbb{F}^{n\times n}[\lambda]$ with zero remainder if and only if $B \in \mathbb{F}^{n\times n}$ is a right (respectively, left) solvent of $A(\lambda)$.*

This result provides a proof of the Cayley-Hamilton theorem.

Theorem 3.15. *([9]) Let $A \in \mathbb{F}^{n\times n}$ with characteristic polynomial c(λ), then c(A) = 0.*

Proof. Define B(λ) = Adj(λI - A) $\in \mathbb{F}^{n\times n}[\lambda]$ and observe that deg(B(λ)) = n - 1 and that

$$(\lambda I - A)B(\lambda) = B(\lambda)(\lambda I - A) = c(\lambda)I \in \mathbb{F}^{n\times n}[\lambda]$$

Now, deg(c(λ)I) = n and is divisible on both the left and the right by λI - A, and c(A) = 0. □

Corollary 3.16. *([9]) If $f \in \mathbb{F}[\lambda]$ and $A \in \mathbb{F}^{n\times n}$, then there exists a polynomial $p \in \mathbb{F}[\lambda]$ (depending on A) with deg(p) < n such that f(A) = p(A).*

Proof. Let f(λ) = q(λ)c(λ) + r(λ), where c(λ) is the characteristic polynomial of A and r(λ) ≡ 0 or deg(r(λ)) ≤ n-1. Then f(A) = q(A)c(A) + r(A) and, by Teorema 3.15, f(A) = r(A) □

3.3 Jordan Structure

Definition 3.17. *Let $A(\lambda) \in \mathbb{F}^{n\times n}[\lambda]$ and $\lambda_0 \in \mathbb{F}$. We define rank($A(\lambda_0)$) to be the rank over $\mathbb{F}$ of the matrix $A(\lambda)$ evaluated at λ_0.*

Definition 3.18. *$\lambda_0 \in \mathbb{F}$ is an* eigenvalue *of $A(\lambda)$ if there exist a vector $X \neq 0$ $\in \mathbb{F}^{n\times 1}$ such that $A(\lambda_0)X$ = 0. All vectors $X \in \mathbb{F}^{n\times 1}$ satisfying $A(\lambda_0)X$ = 0 are called* eigenvectors *of $A(\lambda)$ corresponding to the eigenvalue λ_0. In particular, note that*

$$Z = \{\lambda_0 \in \mathbb{F} \mid det(A(\lambda_0)) = 0\}.$$

is the set of all eigenvalues of $A(\lambda)$.

Definition 3.19. *The integer* $\mu = rank(A(\lambda)) - rank(A(\lambda_0))$ *is known as the* geometric multiplicity *of* λ_0*, which is equivalent to the number of linearly independent eigenvectors associated with it. The amount of times* λ_0 *appears as a root of det(*$A(\lambda)$*) is its* algebraic multiplicitity $\bar{\mu}$*.*

Remark 3.20. In general, the algebraic multiplicity and geometric multiplicity of an eigenvalue can differ. However, the geometric multiplicity can never exceed the algebraic multiplicity.

Definition 3.21. *Given an eigenvalue* $\lambda_j \in \mathbb{F}$*, it is called* semisimple *if its algebraic and geometric multiplicities coincide.*

Example 3.22. Let A(λ) $\in \mathbb{F}^{3\times 3}[\lambda]$ be the same matrix as in Example 3.1.

$$det(A(\lambda)) = 3\lambda^6 - \lambda^5 = \lambda^5(3\lambda - 1)$$

therefore, the eigenvalues of A(λ) are $\lambda_1 = 0$ with algebraic multiplicity 5 and $\lambda_2 = \frac{1}{3}$ with algebraic multiplicity 1.

Let us now calculate its corresponding multiplicities and associated eigenvectors.

- We start with $\lambda_1 = 0$. First observe that rank(A(λ)) - rank(A(0)) = 2, the geometric multiplicity of λ_1. This means there exist 2 eigenvectors associated to this eigenvalue. To find them we solve the following equation:

$$A(0) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff z = 0$$

  Therefore, the two eigenvectors are (1, 0, 0) and (0, 1, 0).

- Now consider $\lambda_2 = \frac{1}{3}$. First observe that rank(A(λ)) - rank(A($\frac{1}{3}$)) = 1, the geometric multiplicity of λ_2. This means there exists a unique associated eigenvector. To find it we solve the following equation:

$$A(\frac{1}{3}) = \begin{pmatrix} \frac{1}{9} & \frac{1}{3} & 1 \\ 0 & \frac{1}{9} & 0 \\ \frac{1}{27} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff \begin{cases} \frac{1}{9}x + \frac{1}{3}y + z = 0 \\ \frac{1}{9}y = 0 \end{cases}$$

$$\iff \begin{cases} y = 0 \\ z = -\frac{1}{9}x \end{cases}$$

  Therefore, the unique eigenvector is (1, 0 , $-\frac{1}{9}$).

In Chapter 6 we will explore the diagonalization of second degree polynomial matrices defined as

$$A(\lambda) = A_2\lambda^2 + A_1\lambda + A_0 \in \mathbb{C}^{n\times n}[\lambda]$$

where A_2, A_1, $A_0 \in \mathbb{C}^{n\times n}$.

To accommodate cases where det(A_2) = 0, we may admit the point at infinity as an eigenvalue of A(λ).

Definition 3.23. *Let $A(\lambda) = \sum_{i=0}^{l} \lambda^i A_i \in \mathbb{F}^{n\times n}[\lambda]$ and consider the matrix*

$$A_*(\lambda) = \sum_{i=0}^{l} \lambda^i A_{l-i} \in \mathbb{F}^{n\times n}[\lambda]$$

This is called the reverse *or* dual *polynomial of $A(\lambda)$, and notice that $A_* = \lambda^l A(1/\lambda)$.*

Remark 3.24. If $\lambda \neq 0$ is an eigenvalue of $A_*(\lambda)$ with geometric and algebraic multiplicities μ and $\bar{\mu}$, then $\frac{1}{\lambda}$ is an eigenvalue of A(λ) with the same multiplicities.

Definition 3.25. *We say that $A(\lambda) \in \mathbb{F}^{n\times n}[\lambda]$ has an* eigenvalue at infinity *if $A_*(\lambda) \in \mathbb{F}^{n\times n}[\lambda]$ has an eigenvalue $\lambda = 0$.*

Remark 3.26. Since $A_*(0) = A_l$, A(λ) has an eigenvalue at infinity if and only if A_l is singular.

Lemma 3.27. *([13]) Let $A(\lambda)\sum_{i=0}^{l} \lambda^i A_{l-i} \in \mathbb{F}^{n\times n}[\lambda]$, and let n_f and n_∞ be the sum of the algebraic multiplicities of its finite and infinite eigenvalues. Then,*

$$ln = n_f + n_\infty \tag{3.3.1}$$

where $l = deg(A(\lambda))$ and $n = dim(A(\lambda))$

Chapter 4

Canonical form of a polynomial matrix

In this chapter our immediate objective is the reduction of a matrix polynomial to a simpler form by means of equivalence transformations, which we will now describe.

4.1 Elementary operations of a polynomial matrix

Now we shall introduce the elementary operations on $A(\lambda) \in \mathbb{F}^{n\times n}[\lambda]$:

1. Multiplication of any row (column) by a number $c \in \mathbb{F} \setminus \{0\}$.
2. Interchange of any two rows (columns).
3. Addition to any row (column) of any other row (column) multiplied by an arbitrary polynomial $b(\lambda) \in \mathbb{F}[\lambda]$.

These three operations are equivalent to a multiplication of the polynomial matrix $A(\lambda)$ on the left by the following square matrices of order n.

We will first define the left elementary operations which are particularly performed on rows.

1. Multiplication of any row by a number $c \in \mathbb{F} \setminus \{0\}$.

$$E' = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & c & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \in \mathbb{F}^{n\times n}[\lambda]$$

2. Interchange of any two rows.

$$E'' = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & \cdots & 1 & & \\ & & \vdots & \ddots & \vdots & & \\ & & 1 & \cdots & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \in \mathbb{F}^{n\times n}[\lambda]$$

3. Addition to any row of any other row multiplied by an arbitrary polynomial b(λ) $\in \mathbb{F}[\lambda]$.

$$E''' = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \cdots & b(\lambda) & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \in \mathbb{F}^{n\times n}[\lambda]$$

Similarly we can define the right elementary operations, which are performed on the columns. The matrices corresponding to them are the same as for the left operations but transposed. We will call them T', T'' and T'''.

The matrices corresponding to either left or right elementary operations are called *elementary matrices*. Note that these matrices are unimodular and therefore their inverses are also elementary matrices.

4.2 Equivalence of polynomial matrices

Definition 4.1. *([5])* Two matrices A(λ), B(λ) $\in \mathbb{F}^{n\times n}[\lambda]$ are called

1. *left-equivalent* if B(λ) can be obtained from A(λ) by means of left elementary operation, i.e. B(λ) = P(λ)A(λ).
2. *right-equivalent* if B(λ) can be obtained from A(λ) by means of right elementary operations, i.e. B(λ) = A(λ)Q(λ).
3. *equivalent* if B(λ) can be obtained from A(λ) by means of left and right elementary operations, i.e. B(λ) = P(λ)A(λ)Q(λ).

where P(λ) and Q(λ) $\in \mathbb{F}^{n\times n}[\lambda]$ are unimodular matrices.

Note that we are using the fact that every P(λ)$\in \mathbb{F}^{n\times n}[\lambda]$ can be represented as a product of elementary matrices. We will see this in Corollary 4.10.

Definition 4.2. *An* equivalence transformation *is the process through which a matrix $A(\lambda) \in \mathbb{F}^{n\times n}[\lambda]$ is transformed to an equivalent matrix $\hat{A}(\lambda) = P(\lambda)A(\lambda)Q(\lambda)$, where $P(\lambda) \in \mathbb{F}^{n\times n}[\lambda]$ and $Q(\lambda) \in \mathbb{F}^{n\times n}[\lambda]$ are unimodular.*

Remark 4.3. Equivalence between matrix polynomials is an equivalence relation.

From this moment on, we will denote equivalence between two matrices $\mathrm{A}(\lambda)$, $\mathrm{B}(\lambda) \in \mathbb{F}^{n\times n}[\lambda]$ by $A(\lambda) \sim B(\lambda)$.

Example 4.4. Let us show that the matrices

$$\mathrm{A}(\lambda) = \begin{pmatrix} \lambda & \lambda+1 \\ \lambda^2-\lambda & \lambda^2-1 \end{pmatrix} \in \mathbb{F}^{2\times 2}[\lambda] \quad \text{and} \quad \mathrm{B}(\lambda) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{F}^{2\times 2}[\lambda]$$

are equivalent and find its transforming unimodular matrices.

Let us start by simplifying $\mathrm{A}(\lambda)$. First, we apply $row_2 = row_2 - (\lambda-1)row_1$:

$$\begin{pmatrix} \lambda & \lambda+1 \\ 0 & 0 \end{pmatrix}.$$

Now, applying $column_2 = column_2 + (-1)column_1$ we obtain

$$\begin{pmatrix} \lambda & 1 \\ 0 & 0 \end{pmatrix}.$$

Finally, we make the following subtraction: $column_1 = column_1 - \lambda column_2$ obtaining the matrix

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \mathrm{B}(\lambda) \sim \mathrm{A}(\lambda).$$

Translating the performed operations into elementary matrices, we can calculate $\mathrm{Q}(\lambda)$ and $\mathrm{P}(\lambda)$.

$$\mathrm{Q}(\lambda) = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -\lambda & 1 \end{pmatrix} = \begin{pmatrix} 1+\lambda & -1 \\ -\lambda & 1 \end{pmatrix} \in \mathbb{F}^{2\times 2}[\lambda]$$

and

$$\mathrm{P}(\lambda) = \begin{pmatrix} 1 & 0 \\ -(\lambda-1) & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1-\lambda & 1 \end{pmatrix} \in \mathbb{F}^{2\times 2}[\lambda].$$

Therefore the equation

$$B(\lambda) = P(\lambda)A(\lambda)Q(\lambda)$$

holds.

Definition 4.5. *We define the* rank *of a matrix polynomial $A(\lambda) \in \mathbb{F}^{n\times n}[\lambda]$ to be the order of its largest minor that is not equal to the zero polynomial. Note that if $A(\lambda) \in \mathbb{F}^{n\times n}[\lambda]$, then it is nonsingular if and only if rank$(A(\lambda)) = n$.*

Proposition 4.6. *([9]) The rank of a matrix polynomial is invariant under equivalence transformations.*

Proof. Let A(λ), B(λ)$\in \mathbb{F}^{n\times n}[\lambda]$ and suppose A(λ)~B(λ). Then there exist unimodular matrices P(λ) $\in \mathbb{F}^{n\times n}[\lambda]$ and Q(λ) $\in \mathbb{F}^{n\times n}[\lambda]$ such that

$$B(\lambda) = P(\lambda)A(\lambda)Q(\lambda).$$

Apply the Binet-Cauchy formula twice to this equation to express a minor $b(\lambda)$ of order j of B(λ) in terms of minors $a_s(\lambda)$ of A(λ) of the same order as follows (after a reordering):

$$b(\lambda) = \sum_s p_s(\lambda)a_s(\lambda)q_s(\lambda) \tag{4.2.1}$$

where $p_s(\lambda)$ and $q_s(\lambda)$ denote the appropriate minors of order j of the matrix polynomials P(λ) and Q(λ), respectively.

If $b(\lambda) \neq 0$ is a minor of B(λ) of the greatest order r (that is, rank(B(λ)) = r), then it follows from Eq. (4.2.1) that at least one minor $a_s(\lambda)$ (of order r) is a nonzero polynomial and hence rank(B(λ)) ≤ rank(A(λ)).

However, applying the same argument to the equation

$$A(\lambda) = P(\lambda)^{-1}B(\lambda)Q(\lambda)^{-1},$$

we see that rank(A(λ)) ≤ rank(B(λ)). Thus, the ranks of equivalent polynomial matrices coincide. □

4.3 The Smith Canonical Form

The main goal of this section is to show how to obtain the simplest form of any A(λ) $\in \mathbb{F}^{n\times n}[\lambda]$ by means of left and right elementary operations. In more detail, it will be shown that any A(λ) $\in \mathbb{F}^{n\times n}[\lambda]$ with rank(A(λ)) = r is equivalent to a diagonal matrix polynomial

$$S(\lambda) = diag[i_1(\lambda), i_2(\lambda), ..., i_r(\lambda), 0, ..., 0] \in \mathbb{F}^{n\times n}[\lambda], \tag{4.3.1}$$

in which i$_j(\lambda)$ is a nonzero monic polynomial for j = 1, 2, ..., r, and i$_{j-1}(\lambda)|$i$_j(\lambda)$, j = 2, 3, ..., r. S(λ) is known as the *Smith canonical form* of A(λ). Note that if some of the polynomials $i_j(\lambda)$ are (nonzero) scalars, then they must be equal to 1 and be placed in the first positions of the canonical matrix. Thus, the Smith canonical form of A(λ) $\in \mathbb{F}^{n\times n}[\lambda]$ with rank(A(λ)) = r is generally

$$S(\lambda) = diag[1, ..., 1, i_k(\lambda), ..., i_r(\lambda), 0, ..., 0] \in \mathbb{F}^{n\times n}[\lambda].$$

.

Example 4.7. An example of a matrix with Smith canonical form is the following:

$$A(\lambda) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \lambda-1 & 0 & 0 \\ 0 & 0 & \lambda-1 & 0 \\ 0 & 0 & 0 & (\lambda-1)^2 \end{pmatrix} \in \mathbb{F}^{4\times 4}[\lambda].$$

Theorem 4.8. *([9]) Any $A(\lambda) \in \mathbb{F}^{n\times n}$ is equivalent to a polynomial matrix of Smith canonical form.*

Proof. We may assume that $A(\lambda) \neq 0$, for otherwise there is nothing to prove. The proof is merely a description of a sequence of elementary transformations needed to reduce successively the rows and columns of A(λ) to the required form.

Step 1: Let $a_{ij}(\lambda) \neq 0$ be an element of A(λ) of least degree; by interchanging rows and columns (elementary operations of type 2) we make it into the element $a_{11}(\lambda)$.

For each element of the first row and column of the resulting matrix, we find the quotient and remainder on division by $a_{11}(\lambda)$:

$$a_{1j}(\lambda) = a_{11}(\lambda)q_{1j}(\lambda) + r_{1j}(\lambda) \qquad j = 2, 3, ..., n$$
$$a_{i1}(\lambda) = a_{11}(\lambda)q_{i1}(\lambda) + r_{i1}(\lambda) \qquad i = 2, 3, ..., n$$

and apply the following transformations (elementary operations of type 3):

$$column_j = column_j - q_{1j}(\lambda)column_1, \qquad (j = 2, ..., n)$$
$$row_i = row_i - q_{i1}(\lambda)row_{1,} \qquad (i = 2, ..., n)$$

Then the elements $a_{1j}(\lambda), a_{i1}(\lambda)$ are replaced by $r_{1j}(\lambda)$ and $r_{i1}(\lambda)$, respectively (i, j = 2, 3, ..., n), all of which are either the zero polynomial or have degree less than that of $a_{11}(\lambda)$. If the polynomials are not all zero, we use an elementary operation of type 2 to interchange $a_{11}(\lambda)$ with an element $r_{1j}(\lambda)$ or $r_{i1}(\lambda)$ of least degree.

Now we repeat the process of reducing the degree of the off-diagonal elements of the first row and column to be less than that of the new $a_{11}(\lambda)$. Clearly, since the deg($a_{11}(\lambda)$) is strictly decreasing at each step, we eventually reduce the λ-matrix to the form:

$$\begin{pmatrix} a_{11}(\lambda) & 0 & \cdots & 0 \\ 0 & a_{22}(\lambda) & \cdots & a_{2n}(\lambda) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}(\lambda) & \cdots & a_{nn}(\lambda) \end{pmatrix} \tag{4.3.2}$$

Step 2: In the form of (4.3.2) there may now be nonzero elements $a_{ij}(\lambda)$, $2 \leq$ i, j $\leq$ n, whose degree is less than that of $a_{11}(\lambda)$. If so, we repeat *Step 1* again and arrive at another matrix of the form (3.1) but with the degree of $a_{11}(\lambda)$ further reduced. Thus, by repeating *Step 1* a sufficient number of times, we can find a matrix of the form (4.3.2) that is equivalent to A(λ) and for which $a_{11}(\lambda)$ is a nonzero element of least degree.

Step 3: Having completed *Step 2*, we now ask whether there are nonzero elements that are not divisible by $a_{11}(\lambda)$. If there is one such, say $a_{ij}(\lambda)$ we do

$$column_1 = column_1 + column_j,$$

find remainders and quotients of the new *column*$_1$ on division by $a_{11}(\lambda)$, and go on to repeat *Steps 1 and 2*, winding up with a form (4.3.2), again with $a_{11}(\lambda)$ replaced by a polynomial of smaller degree.

Again, this process can continue only for a finite number of steps before we arrive at a matrix of the form

$$\begin{pmatrix} i_1(\lambda) & 0 & \cdots & 0 \\ 0 & b_{22}(\lambda) & \cdots & b_{2n}(\lambda) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & b_{n2}(\lambda) & \cdots & b_{nn}(\lambda) \end{pmatrix},$$

where, after an elementary operation of type 1 (if necessary), $a_1(\lambda)$ is monic and all the nonzero elements $b_{ij}(\lambda)$ are divisible by $a_1(\lambda)$ without remainder.

Step 4: If all $b_{ij}(\lambda) = 0$, the theorem is proved. If not, the above matrix may be reduced to the form

$$\begin{pmatrix} i_1(\lambda) & 0 & 0 & \cdots & 0 \\ 0 & i_2(\lambda) & 0 & \cdots & b_{2n}(\lambda) \\ 0 & 0 & c_{33}(\lambda) & \cdots & c_{3n}(\lambda) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & c_{n3} & \cdots & c_{nn}(\lambda) \end{pmatrix},$$

where $a_2(\lambda)$ is divisible by $a_1(\lambda)$ and the elements $c_{ij}(\lambda)$, $3 \leq$ i, j $\leq$ n, are divisible by $a_2(\lambda)$. Continuing the process we arrive at the statement of the theorem; a matrix like the following:

$$\begin{pmatrix} i_1(\lambda) & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & i_2(\lambda) & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & i_r(\lambda) & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{F}^{n\times n}[\lambda]. \tag{4.3.3}$$

□

Example 4.9. Let us find the canonical form of the following matrix following the steps shown above.

$$A(\lambda) = \begin{pmatrix} 0 & 1 & \lambda \\ \lambda & \lambda & 1 \\ \lambda^2-\lambda & \lambda^2-1 & \lambda^2-1 \end{pmatrix} \in \mathbb{F}^{3\times 3}[\lambda].$$

First of all we apply *Step 1*: We want to make the element $a_{11}(\lambda)$ become the element of least degree of the matrix. Therefore, we exchange *column*$_1$ and *column*$_2$ of the matrix and obtain:

$$\begin{pmatrix} 1 & 0 & \lambda \\ \lambda & \lambda & 1 \\ \lambda^2-1 & \lambda^2-\lambda & \lambda^2-1 \end{pmatrix} \in \mathbb{F}^{3\times 3}[\lambda]$$

Now let us divide $a_{1j}(\lambda)$ and $a_{i1}(\lambda)$ by $a_{11}(\lambda)$ (j, i = 2, 3):

$$\begin{cases} a_{12}(\lambda) = 0 = 1 \cdot 0 \\ a_{13}(\lambda) = \lambda = 1 \cdot \lambda \\ a_{21}(\lambda) = \lambda = 1 \cdot \lambda \\ a_{31}(\lambda) = \lambda^2 - 1 = 1 \cdot (\lambda^2 - 1). \end{cases}$$

Considering the results of these divisions, we apply the following transformations:

$$\begin{cases} column_2 = column_2 - 0 \cdot (\lambda) column_1 \\ column_3 = column_3 - \lambda \cdot column_1 \\ row_2 = row_2 - \lambda \cdot row_1 \\ row_3 = row_3 - (\lambda^2 - 1) row_1 \end{cases}$$

and obtain the matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 1 - \lambda^2 \\ 0 & \lambda^2 + \lambda & -\lambda^3 + \lambda^2 + \lambda + 1 \end{pmatrix}.$$

Observe that $a_{11}(\lambda)$ is the nonzero polynomial of lest degree, therefore we can skip *Step 2* and, in addition, $a_{22}(\lambda)$, $a_{23}(\lambda)$, $a_{32}(\lambda)$ and $a_{33}(\lambda)$ are divisible by $a_{11}(\lambda)$, so we can also skip *Step 3*.

Now, as not all $a_{ij}(\lambda) = 0$, i = 2, 3, j = 2, 3, we continue with the reduction (*Step 4*). Observe that $a_{22}(\lambda)$ is the polynomial of least degree between $a_{22}(\lambda)$, $a_{23}(\lambda)$, $a_{32}(\lambda)$ and $a_{33}(\lambda)$, as we want. Now we divide $a_{23}(\lambda)$ and $a_{32}(\lambda)$ by $a_{22}(\lambda)$:

$$\begin{cases} a_{23}(\lambda) = -\lambda^2 + 1 = \lambda(-\lambda) + 1 \\ a_{32}(\lambda) = \lambda^2 + \lambda = \lambda(\lambda + 1). \end{cases}$$

Then we apply the following transformations

$$\begin{cases} column_3 = column_3 - (-\lambda) column_2 \\ row_3 = row_3 - (\lambda + 1) row_2 \end{cases}$$

and obtain the matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & 2\lambda^2 \end{pmatrix}.$$

Note that deg$(a_{23}(\lambda)) < deg(a_{11}(\lambda))$. Therefore, we exchange $column_2$ and $column_3$ by applying an elementary operation of second type:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \lambda \\ 0 & 2\lambda^2 & 0 \end{pmatrix}.$$

Now,

$$\begin{cases} a_{23}(\lambda) = \lambda = 1 \cdot \lambda \\ a_{32}(\lambda) = 2\lambda^2 = 1 \cdot 2\lambda^2 \end{cases}$$

and by applying the following transformations

$$\begin{cases} column_3 = column_3 - \lambda \cdot column_2 \\ row_3 = row_3 - 2\lambda^2 \cdot row_2 \end{cases}$$

we finally obtain the canonical form of A(λ):

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

Note that by using left(right) elementary operations only, a matrix polynomial can be reduced to an upper-(lower-)triangular matrix polynomial B(λ) $\in \mathbb{F}^{n\times n}[\lambda]$ with the property that if $\deg(b_{jj}(\lambda)) = l_j$ (i = 1, 2, ..., n) then

1. $l_j = 0$ implies that $b_{ij}(\lambda) = 0$ ($b_{ji}(\lambda) = 0$), i = 1, 2, ..., j - 1, and
2. $l_j > 0$ implies that $\deg(b_{ij}(\lambda)) < l_j$ ($\deg(b_{ji}(\lambda)) < l_i$), i = 1, 2, ..., j - 1.

The reduction of matrix polynomials described above takes on a simple form in the important special case in which A(λ) does not depend explicitly on λ at all, that is, when A(λ)$\equiv$ A $\in \mathbb{F}^{n\times n}$ (i.e. when $\deg(A(\lambda)) = 0$).

Corollary 4.10. *([5]) Consider $A(\lambda) \in \mathbb{F}^{n\times n}[\lambda]$. If $det(A(\lambda)) \in \mathbb{F} \setminus \{0\}$, then the matrix can be represented in the form of a product of a finite number of elementary matrices.*

As we just noted A(λ) can be brought into the form

$$B(\lambda) = \begin{pmatrix} b_{11}(\lambda) & b_{12}(\lambda) & \cdots & b_{1n}(\lambda) \\ 0 & b_{22}(\lambda) & \cdots & b_{2n}(\lambda) \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & b_{nn}(\lambda) \end{pmatrix} \tag{4.3.4}$$

by left elementary operations. Since in the application of elementary operations to a square polynomial matrix the determinant of the matrix is only multiplied by constant nonzero factors, then

$$det(A(\lambda)) = c \cdot det(B(\lambda)) = c \cdot b_{11}(\lambda) b_{22}(\lambda) \cdots b_{nn}(\lambda) \in \mathbb{F} \setminus \{0\},$$

where c $\neq$ 0 $\in \mathbb{F}$. Hence,

$$b_{ii}(\lambda) \in \mathbb{F} \setminus \{0\} \qquad (i = 1, 2, ..., n).$$

We also know that the matrix (4.3.4) is equivalent to the diagonal form (4.3.3) and can therefore be reduced to the identity matrix I by means of left elementary operations of type 1. But then, conversely, the identity matrix I can be transformed into A(λ) by means of the left elementary operations whose matrices are $E_1(\lambda), E_2(\lambda), ..., E_p(\lambda)$. Therefore

$$A(\lambda) = E_p(\lambda) E_{p-1}(\lambda) \cdots E_1(\lambda) I = E_p(\lambda) E_{p-1}(\lambda) \cdots E_1(\lambda).$$

The uniqueness of the polynomials $i_1(\lambda), ..., i_r(\lambda) \in \mathbb{F}[\lambda]$ appearing in the form (4.3.3) is shown in the following section.

4.4 Invariant Polynomials

We start this section by constructing a system of polynomials that is uniquely defined by a given A(λ) $\in \mathbb{F}^{n\times n}[\lambda]$ and that is invariant under equivalence transformations.

Suppose A(λ) $\in \mathbb{F}^{n\times n}[\lambda]$ with rank(A(λ)) = r and let

$$d_j(\lambda) = \text{gcd(all minors of A}(\lambda) \text{ of order j)},$$

where j = 1, 2, ..., r.

Clearly, any minor of order j $\geq$ 2 may be expressed as a linear combination of minors of order j - 1, so that $d_{j-1}(\lambda)|d_j(\lambda)$.

Hence, if we define $d_0(\lambda) \equiv 1$, then in the sequence $d_0(\lambda)$, $d_1(\lambda)$, ..., $d_r(\lambda)$, $d_{j-1}(\lambda)|d_j(\lambda)$, j = 1, 2, ..., r. Note that for the Smith canonical form

$$S(\lambda) = diag[i_1(\lambda), i_2(\lambda), ..., i_r(\lambda), 0, ..., 0]$$

the polynomials described above are respectively

$$d_1(\lambda) = i_1(\lambda),\ d_2(\lambda) = i_1(\lambda)i_2(\lambda),\ ...,\ d_r(\lambda) = \prod_{j=1}^{r} i_j(\lambda)$$

The polynomials $d_0(\lambda), d_1(\lambda), ..., d_r(\lambda)$ are invariant under equivalence transformations. To see this, let $d_j(\lambda)$ and $\delta_j(\lambda)$ denote the (monic) greatest common divisor of all minors of order j of A(λ), B(λ) $\in \mathbb{F}^{n\times n}[\lambda]$, respectively. Note that provided that A(λ)~B(λ), the number of polynomials $d_j(\lambda)$ and $\delta_j(\lambda)$ is the same.

Proposition 4.11. *([9]) Let A(λ), B(λ) $\in \mathbb{F}^{n\times n}[\lambda]$ of rank r be equivalent. Then, with the notation of the previous paragraph, $d_j(\lambda) = \delta_j(\lambda)$ for j = 1, 2, ..., r.*

Proof. Preserving the notation used in the proof of Proposition 4.6, it is easily seen from Eq. (4.2.1) that any common divisor of minors $a_j(\lambda)$ of A(λ) of order j (1 $\leq j \leq$ r) is a divisor of any minor $b_j(\lambda)$ of B(λ) of order j (1 $\leq j \leq$ r). Hence $d_j(\lambda)|\delta_j(\lambda)$. But again, the equation A(λ) = $P(\lambda)^{-1}B(\lambda)Q(\lambda)^{-1}$ implies that $\delta_j(\lambda)|d_j(\lambda)$ and, since both polynomials are assumed to be monic, we obtain

$$\delta_j(\lambda) = d_j(\lambda), \qquad j = 1, 2, ..., r.$$

$\square$

Now consider the quotients

$$i_1(\lambda) = \frac{d_1(\lambda)}{d_0(\lambda)},\ i_2(\lambda) = \frac{d_2(\lambda)}{d_1(\lambda)},\ ...,\ i_r(\lambda) = \frac{d_r(\lambda)}{d_{r-1}(\lambda)}.$$

In view of the divisibility of $d_j(\lambda)$ by $d_{j-1}(\lambda)$, the quotients $i_j(\lambda)$ (j = 1, 2, ..., r) are polynomials. They are called the *invariant polynomials* of A(λ). Note that for j = 2, 3, ..., r, $i_{j-1}(\lambda)|i_j(\lambda)$.

Corollary 4.12. *([9]) Two matrix polynomials $A(\lambda)$ and $B(\lambda)$ are equivalent if and only if they have the same invariant polynomials.*

Proof. The "only if" statement is just Proposition 4.11. If two matrix polynomials have the same invariant polynomials, then Theorem 4.8 implies that they have the same Smith canonical form. The transitive property of equivalence relations then implies that they are equivalent. $\square$

Theorem 4.13. *([5]) If in a block-diagonal matrix*

$$C(\lambda) = \begin{pmatrix} A(\lambda) & 0 \\ 0 & B(\lambda) \end{pmatrix} \in \mathbb{F}^{(n+m)\times(n+m)}[\lambda]$$

every invariant polynomial of $A(\lambda) \in \mathbb{F}^{n\times n}[\lambda]$ divides every invariant polynomial of $B(\lambda) \in \mathbb{F}^{m\times m}[\lambda]$, then the set of invariant polynomials of $C(\lambda)$ is the union of the invariant polynomials of $A(\lambda)$ and $B(\lambda)$.

Proof. We denote by $\mathrm{i}_1^{'}(\lambda)$, $\mathrm{i}_2^{'}(\lambda)$, ..., $\mathrm{i}_r^{'}(\lambda)$ and $\mathrm{i}_1^{''}(\lambda)$, $\mathrm{i}_2^{''}(\lambda)$, ..., $\mathrm{i}_q^{''}(\lambda)$, respectively, the invartiant polynomials of $\mathrm{A}(\lambda) \in \mathbb{F}^{n\times n}[\lambda]$ and $\mathrm{B}(\lambda) \in \mathbb{F}^{m\times m}[\lambda]$.
Then

$$A(\lambda) \sim diag[i_1^{'}(\lambda), ..., i_r^{'}(\lambda), 0, ..., 0], \quad B(\lambda) \sim diag[i_1^{''}(\lambda), ..., i_q^{''}(\lambda), 0, ..., 0]$$

and therefore

$$C(\lambda) \sim diag[i_1^{'}(\lambda), ..., i_r^{'}(\lambda), i_1^{''}(\lambda), ..., i_q^{''}(\lambda), 0, ..., 0].$$

The λ-matrix on the right-hand side of this relation is of canonical diagonal form. The diagonal elements of this matrix that are not identically zero then form a complete system of invariants of the matrix $\mathrm{C}(\lambda) \in \mathbb{F}^{n\times n}[\lambda]$. $\square$

Example 4.14. Let us now compute (from their definition) the invariant polynomials of the following matrix used in Example 4.9:

$$\mathrm{A}(\lambda) = \begin{pmatrix} 0 & 1 & \lambda \\ \lambda & \lambda & 1 \\ \lambda^2 - \lambda & \lambda^2 - 1 & \lambda^2 - 1 \end{pmatrix} \in \mathbb{F}^{3\times 3}[\lambda].$$

First of all we shall calculate the rank of $\mathrm{A}(\lambda)$. In order to do so, we will examine its minors, as we will probably need to calculate them anyways. Let us start with its unique minor of order 3:

$$\det(\mathrm{A}(\lambda)) = \begin{vmatrix} 0 & 1 & \lambda \\ \lambda & \lambda & 1 \\ \lambda^2 - \lambda & \lambda^2 - 1 & \lambda^2 - 1 \end{vmatrix} = 0.$$

This means rank(A(λ)) $\neq$ 3. Now let us compute all minors of order 2.

$$M_1 = \begin{vmatrix} \lambda & 1 \\ \lambda^2-1 & \lambda^2-1 \end{vmatrix} = (\lambda^2-1)(\lambda-1), \quad M_2 = \begin{vmatrix} \lambda & 1 \\ \lambda^2-\lambda & \lambda^2-1 \end{vmatrix} = \lambda^2(\lambda-1),$$

$$M_3 = \begin{vmatrix} \lambda & \lambda \\ \lambda^2-\lambda & \lambda^2-1 \end{vmatrix} = \lambda(\lambda-1), \quad M_4 = \begin{vmatrix} 1 & \lambda \\ \lambda^2-1 & \lambda^2-1 \end{vmatrix} = (\lambda^2-1)(1-\lambda),$$

$$M_5 = \begin{vmatrix} 0 & \lambda \\ \lambda^2-\lambda & \lambda^2-1 \end{vmatrix} = -\lambda^2(\lambda-1), \quad M_6 = \begin{vmatrix} 0 & 1 \\ \lambda^2-\lambda & \lambda^2-1 \end{vmatrix} = -\lambda(\lambda-1),$$

$$M_7 = \begin{vmatrix} 1 & \lambda \\ \lambda & 1 \end{vmatrix} = 1-\lambda^2, \quad M_8 = \begin{vmatrix} 0 & \lambda \\ \lambda & 1 \end{vmatrix} = -\lambda^2, \quad M_9 = \begin{vmatrix} 0 & 1 \\ \lambda & \lambda \end{vmatrix} = \lambda$$

Not all the minors above are null, therefore rank(A(λ)) = 2 and

$$d_2(\lambda) = gcd(M_1, M_2, ..., M_9) = 1.$$

Finally, by looking at the elements of the matrix, we deduce

$$d_1(\lambda) = gcd(0, 1, \lambda, 1, \lambda^2-\lambda, \lambda^2-1) = 1.$$

And we always consider $d_0(\lambda) = 1$.

Therefore the invariant polynomials are

$$i_1(\lambda) = \frac{d_1(\lambda)}{d_0(\lambda)} = \frac{1}{1} = 1, \text{ and } i_2(\lambda) = \frac{d_2(\lambda)}{d_1(\lambda)} = \frac{1}{1} = 1.$$

and the Smith canonical form of A(λ), as seen in Example 4.9, is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathbb{F}^{3\times 3}[\lambda].$$

4.5 Elementary Divisors

Consider a matrix A(λ) $\in \mathbb{F}^{n\times n}[\lambda]$ with rank(A(λ)) = r, invariant polynomials $i_1(\lambda), i_2(\lambda), ..., i_r(\lambda)$ and eigenvalues $\lambda_1, ..., \lambda_t$.

From the Smith canonical form, we deduce

$$det(A(\lambda)) = \prod_{j=1}^{r} i_j(\lambda) = \prod_{k=1}^{t} (\lambda-\lambda_k)^{\bar{\mu}_k} \in \mathbb{F}[\lambda]$$

where $\bar{\mu}_k$ is the algebraic multiplicity of the eigenvalue λ_k.

Moreover, since $i_j(\lambda)|i_{j+1}(\lambda)$ for j = 1, 2, ..., r-1, it follows that there are integers m_{jk}, 1 ≤ j ≤ r and 1 ≤ k ≤ t, such that

$$i_1(\lambda) = (\lambda-\lambda_1)^{m_{11}}(\lambda-\lambda_2)^{m_{12}} \cdots (\lambda-\lambda_s)^{m_{1t}}$$

$$i_2(\lambda) = (\lambda - \lambda_1)^{m_{21}}(\lambda - \lambda_2)^{m_{22}} \cdots (\lambda - \lambda_t)^{m_{2t}}$$

$$\vdots$$

$$i_r(\lambda) = (\lambda - \lambda_1)^{m_{r1}}(\lambda - \lambda_2)^{m_{r2}} \cdots (\lambda - \lambda_t)^{m_{rt}}$$

and for k = 1, 2, ..., t,

$$0 \le m_{1k} \le m_{2k} \le m_{rk} \le \bar{\mu}_k \quad \text{and} \quad \sum_{j=1}^{r} m_{jk} = \bar{\mu}_k.$$

Definition 4.15. *Each factor* $(\lambda - \lambda_k)^{m_{jk}}$ *appearing in the factorization with* $m_{jk} > 0$ *is called an* elementary divisor *of* $A(\lambda)$. *And each integer* m_{jk}, $j = 1, \ldots, r$ *is called a* partial multiplicity *of the eigenvalue* λ_k.

An elementary divisor for which $m_{jk} = 1$ is said to be *linear*; otherwise it is *nonlinear*. We may also refer to the elementary divisors $(\lambda - \lambda_k)^{m_{jk}}$ as those associated with λ_k, with the obvious meaning.

Remark 4.16. The system of all elementary divisors (along with the rank and order) of a matrix polynomial completely defines the set of its invariant polynomials and vice versa. It follows that the elementary divisors are invariant under equivalence transformations.

Theorem 4.17. *([9]) Suppose* $A(\lambda)$, $B(\lambda) \in \mathbb{F}^{n\times n}[\lambda]$. *Then,* $A(\lambda) \sim B(\lambda)$ *if and only if they have the same elementary divisors.*

Theorem 4.18. *([9]) If* $A(\lambda) \in \mathbb{F}^{n\times n}[\lambda]$ *and* $B(\lambda) \in \mathbb{F}^{m\times m}[\lambda]$, *then the set of elementary divisors of the block-diagonal matrix*

$$C(\lambda) = \begin{pmatrix} A(\lambda) & 0 \\ 0 & B(\lambda) \end{pmatrix} \in \mathbb{F}^{(n+m)\times(n+m)}[\lambda]$$

is the union of the sets of elementary divisors of $A(\lambda)$ *and* $B(\lambda)$.

Proof. Let $S_1(\lambda)$ and $S_2(\lambda)$ be the Smith forms of A(λ) and B(λ), respectively. Then clearly

$$\mathrm{C}(\lambda) = \mathrm{E}(\lambda) \begin{pmatrix} S_1(\lambda) & 0 \\ 0 & S_2(\lambda) \end{pmatrix} \mathrm{F}(\lambda)$$

for some nonsingular polynomial matrices E(λ) and F(λ). Let $(\lambda-\lambda_0)^{\alpha_1}, \ldots, (\lambda-\lambda_0)^{\alpha_p}$ and $(\lambda-\lambda_0)^{\beta_1}, \ldots, (\lambda-\lambda_0)^{\beta_q}$ be the elementary divisors of $S_1(\lambda)$ and $S_2(\lambda)$, respectively, corresponding to the same eigenvalue λ_0. Arrange the set of exponents $\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q$, in a nondecreasing order: $\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q = \gamma_1, \ldots, \gamma_{p+q}$, where $0 < \gamma_1 \le \cdots \le \gamma_{p+q}$.

From the definition of invariant polynomials it is clear that in the Smith form

S(λ) $=$ diag$[i_1(\lambda), ..., i_r(\lambda), 0, ..., 0]$ of diag$[S_1(\lambda), S_2(\lambda)]$, the invariant polynomial $i_r(\lambda)$ is divisible by $(\lambda - \lambda_0)^{\gamma_{p+q}}$ but not by $(\lambda - \lambda_0)^{\gamma_{p+q}+1}$; and $i_{r-1}(\lambda)$ is divisible by $(\lambda - \lambda_0)^{\gamma_{p+q-1}}$ but not by $(\lambda - \lambda_0)^{\gamma_{p+q-1}+1}$; and so on. It follows that the elementary divisors of

$$diag[S_1(\lambda), S_2(\lambda)],$$

and therefore also those of C(λ), corresponding to λ_0 are just $(\lambda - \lambda_0)^{\gamma_1}, ..., (\lambda - \lambda_0)^{\gamma_{p+q}}$, and the theorem is proved. $\square$

Example 4.19. Let

$$A(\lambda) = \begin{pmatrix} \lambda - 3 & -1 & 0 & 0 \\ 4 & \lambda + 1 & 0 & 0 \\ -6 & -1 & \lambda - 2 & -1 \\ 14 & 5 & 1 & \lambda \end{pmatrix} \in \mathbb{F}^{4\times 4}[\lambda],$$

we will now calculate its elementary divisors.

First, we apply $row_4 = row_4 + \lambda row_3$:

$$\begin{pmatrix} \lambda - 3 & -1 & 0 & 0 \\ 4 & \lambda + 1 & 0 & 0 \\ -6 & -1 & \lambda - 2 & -1 \\ 14 - 6\lambda & 5 - \lambda & \lambda^2 - 2\lambda + 1 & 0 \end{pmatrix}.$$

Now applying $column_1 = column_1 + (-6)column_4$, $column_2 = column_2 + (-1)column_4$ and $column_3 = column_3 + (\lambda - 2)column_4$ we obtain

$$\begin{pmatrix} \lambda - 3 & -1 & 0 & 0 \\ 4 & \lambda + 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 14 - 6\lambda & 5 - \lambda & \lambda^2 - 2\lambda + 1 & 0 \end{pmatrix}.$$

Then, we make the following addition: $column_1 = column_1 + (\lambda - 3)column_2$ obtaining

$$\begin{pmatrix} 0 & -1 & 0 & 0 \\ \lambda^2 - 2\lambda + 1 & \lambda + 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ -\lambda^1 + 2\lambda - 1 & 5 - \lambda & \lambda^2 - 2\lambda + 1 & 0 \end{pmatrix}.$$

Now we do: $row_2 = row_2 + (\lambda + 1)row_1$ and $row_4 = row_4 + (5 - \lambda)row_1$, and we obtain

$$\begin{pmatrix} 0 & -1 & 0 & 0 \\ \lambda^2 - 2\lambda + 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ -\lambda^1 + 2\lambda - 1 & 0 & \lambda^2 - 2\lambda + 1 & 0 \end{pmatrix}.$$

Finally, we apply $row_2 = row_2 + row_4$; then $row_1 = (-1)row_1$ and $row_3 = (-1)row_3$. After permuting some rows and columns we obtain:

$$\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & (\lambda-1)^2 & 0\\ 0 & 0 & 0 & (\lambda-1)^2\end{pmatrix}.$$

Therefore, its invariant polynomials are 1, 1, $(\lambda-1)^2$ and $(\lambda-1)^2$ and it has two elementary divisors $(\lambda-1)^2$ and $(\lambda-1)^2$.

Chapter 5

Linearization of a matrix polynomial

In this chapter the definition of the companion matrix of a polynomial matrix will be given and after the concept of linearization will be explained in detail.
For a matrix polynomial A(λ) $\in \mathbb{F}^{n\times n}[\lambda]$,

$$A(\lambda) = \sum_{j=0}^{l} A_j\lambda^j, \qquad detA_l \neq 0, \tag{5.0.1}$$

we formulate the generalization of the companion matrix:

Definition 5.1.

$$C_A = \begin{pmatrix} 0 & I_n & 0 & \cdots & 0 \\ 0 & 0 & I_n & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & I_n \\ -\hat{A}_0 & -\hat{A}_1 & -\hat{A}_2 & \cdots & -\hat{A}_{l-1} \end{pmatrix} \in \mathbb{F}^{ln\times ln}$$

where $\hat{A}_j = A_l^{-1}A_j$ *for* j *= 0, 1, ..., l-1.*
C_A *is called the (first)* companion matrix *of A(λ).*

Remark 5.2. The characteristic polynomial of A(λ) satisfies

$$det(A(\lambda)) = det(\lambda I_{ln} - C_A)det(A_1). \tag{5.0.2}$$

This means that the *eigenvalues* of A(λ) coincide with the *eigenvalues* of C_A. In addition, the relation (5.0.2) says that A(λ) and $\lambda I - C_A$ have the same invariant polynomials of highest degree. However, the connection is deeper than this, as the following theorem shows.

Theorem 5.3. *The* $ln \times ln$ *matrix polynomials*

$$\begin{pmatrix} A(\lambda) & 0 \\ 0 & I_{(l-1)n} \end{pmatrix} \qquad and \qquad \lambda I_{ln} - C_A$$

are equivalent.

Proof. First define ln x ln matrix polynomials F(λ) and E(λ) by

$$\mathrm{F}(\lambda) = \begin{pmatrix} I & 0 & 0 & \cdots & 0 \\ -\lambda I & I & 0 & \cdots & 0 \\ 0 & -\lambda I & I & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & -\lambda I & I \end{pmatrix}, \ \mathrm{E}(\lambda) = \begin{pmatrix} B_{l-1}(\lambda) & B_{l-2}(\lambda) & B_{l-3}(\lambda) & \cdots & B_0(\lambda) \\ -I & 0 & 0 & \cdots & 0 \\ 0 & -I & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & -I & 0 \end{pmatrix}$$

where $B_0(\lambda) = A_l$, $B_{r+1}(\lambda) = \lambda B_r(\lambda) + A_{l-r-1}$ for r = 0, 1, ..., l - 2. Clearly,

$$det(E(\lambda)) = \pm det(A_l), \qquad det(F(\lambda)) = 1. \tag{5.0.3}$$

Hence $F(\lambda)^{-1} \in \mathbb{F}^{n\times n}[\lambda]$. It is easily verified that

$$E(\lambda)(\lambda I - C_A) = \begin{pmatrix} A(\lambda) & 0 \\ 0 & I_{(l-1)n} \end{pmatrix} F(\lambda), \tag{5.0.4}$$

and so

$$\begin{pmatrix} A(\lambda) & 0 \\ 0 & I_{(l-1)n} \end{pmatrix} = E(\lambda)(\lambda I - C_A)F(\lambda)^{-1} \tag{5.0.5}$$

determines the equivalence stated in the theorem. □

Remark 5.4. Theorem 5.3 shows that all the invariant polynomials (and hence all of the elementary divisors) of A(λ) and $\lambda I - C_A$ with degree > 0 coincide.

Definition 5.5. *Let $A(\lambda) \in \mathbb{F}^{n\times n}[\lambda]$ with deg(A(λ)) = l and with nonsingular leading coefficient, then for any matrix $\lambda I - L \in \mathbb{F}^{ln\times ln}[\lambda]$ for which $\lambda I - L \sim diag(A(\lambda), I_{(l-1)n})$, is called a* linearization *of A(λ).*

Example 5.6. Let us find a linearization for the matrix polynomial

$$A(\lambda) = \begin{pmatrix} \lambda^2 & -\lambda \\ 0 & \lambda^2 \end{pmatrix}.$$

First of all observe that A(λ) has a nonsingular leading coefficient

$$A_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \in \mathbb{F}^{2\times 2}.$$

Now we want to find a matrix L $\in \mathbb{F}^{4\times 4}$ such that $\lambda I - L \sim \mathrm{diag}[A(\lambda), I_{(l-1)n}]$. We have seen that $\lambda I - C_A \sim \mathrm{diag}[A(\lambda), I_{(l-1)n}]$. Therefore we will compute $C_A \in \mathbb{F}^{4\times 4}$. First, notice that

$$A_2^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A_0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad A_1 = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}.$$

So,

$$\hat{A}_0 = A_2^{-1}A_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

$$\hat{A}_1 = A_2^{-1} A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}.$$

Then

$$C_A = \begin{pmatrix} 0 & I_2 \\ -\hat{A}_0 & -\hat{A}_1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \in \mathbb{F}^{4 \times 4}$$

and we have obtained a linearization $\lambda I - C_A$ of A(λ).

We can find more linearizations $B_1\lambda + B_0$ of A(λ) by applying equivalence transformations to $\lambda I - C_A$. Thus, for any nonsingular matrices P, Q $\in \mathbb{F}^{ln \times ln}$, P$(\lambda I - C_A)Q = B_1\lambda + B_0 \in \mathbb{F}^{ln \times ln}[\lambda]$ is also a linearization of A(λ).

Chapter 6

Second degree diagonalization

In this chapter we will show the characterization of the diagonalizable quadratic matrix polynomials. In other words, we shall see which are the admissible partial multiplicities of the eigenvalues of diagonalizable matrices A(λ) = $A_2\lambda^2 + A_1\lambda + A_0 \in \mathbb{F}^{n\times n}[\lambda]$. The difficulty of the diagonalization of polynomial matrices is that the degree of the matrix must be maintained, we are not only looking for a diagonal matrix with the same Jordan Structure, but with the same degree as well. The presented theory has been extracted from papers [10] and [13].

Solving the quadratic eigenvalue problem is critical in several applications in control and systems. As noted in the introduction of this work, several systems in a variety of disciplines are described by quadratic matrix polynomials

$$A(\lambda) = A_2\lambda^2 + A_1\lambda + A_0$$

where A_2, A_1, $A_0 \in \mathbb{C}^{n\times n}$.

Before we begin this chapter let us briefly comment the case for polynomial matrices of first degree.

Suppose A(λ) = $\lambda A_1 + A_0 \in \mathbb{C}^{n\times n}[\lambda]$.

- First, if $det(A_1) \neq 0$ we can transform A(λ) to an equivalent matrix $\hat{A}(\lambda)$:

  $$A_1^{-1}A(\lambda) = A_1^{-1}(A_1\lambda + A_0) = I\lambda + B = \hat{A}(\lambda).$$

  Then, A(λ) is diagonalizable if and only if B $\in \mathbb{C}^{n\times n}$ is diagonalizable.
- If det(A_1) = 0, then we consider the infinite eigenvalues of A(λ). Suppose $\lambda_1, ..., \lambda_t, \lambda_\infty \in \mathbb{F}$ are the eigenvalues of A(λ) with corresponding algebraic multiplicities $\bar{\mu}_i$, i = 1, ..., t, ∞ and partial multiplicities m_{ij}, i = 1, ..., t, ∞. Then

  $$\sum_{i=1}^{t}\bar{\mu}_i + \bar{\mu}_\infty = n \quad \text{and}$$

  $$m_{ij} = m_{\infty j} = 1, \quad i = 1, ..., t, \infty.$$

  We want to show that there exists a diagonal matrix $\hat{A}(\lambda)$ with the same infinite structure as A(λ). We know that all elementary divisors of A(λ) have degree =

1, therefore, if we associate each of them to one of the entries of $\hat{A}(\lambda)$, we obtain a diagonal matrix isospectral to A(λ). Therefore $A(\lambda)$ is diagonalizable.

and therefore A(λ) is diagonalizable.

Let us now focus on the diagonalization of polinomial matrices A(λ) $= A_2\lambda^2 + A_1\lambda + A_0 \in \mathbb{C}^{n\times n}[\lambda]$.

Definition 6.1. A system is said to be

- *real* if $A_2,\ A_1,\ A_0 \in \mathbb{R}^{n\times n}$.
- *hermitian* or *real symmetric* if $A_2,\ A_1,\ A_0$ are all Hermitian, or all real and symmetric.
- *diagonal* or *decoupled* if $A_2,\ A_1,\ A_0$ are diagonal matrices or equivalently if it admits an isospectral diagonal system.

Definition 6.2. Two systems will be called *isospectral* if they share the same Jordan form; i.e. the same eigenvalues and the same partial multiplicities.

One alternative to solve this diagonalization problem is to reduce the matrix to a diagonal form so that its eigenvalue structure can be recognized in the diagonal of the equivalent matrix. There are two major categories of diagonalizable systems.

- The first category, which we will study in Sections 6.1, 6.2 and 6.3, consists of systems that can be directly decoupled to a diagonal system by applying congruence or strict equivalence transformations.

  In the first case, we are talking about systems A(λ) for which there exists a non-singular matrix U such that L(λ)~U$^*L(\lambda)U$ where $U^*L(\lambda)U$ is diagonal.

  The second case refers to systems A(λ) $\in \mathbb{C}^{n\times n}[\lambda]$ for which there exist non-singular matrices U, V $\in \mathbb{C}^{n\times n}$ for which $L(\lambda)$~UL(λ)V and $UL(\lambda)V$ is diagonal.

  Turns out these relatively simple cases require one of the coefficients $A_2,\ A_1,\ A_0$ to be expressed in terms of the other two, and their natural independence is lost.

- The second category, which we will study in Section 6.4, is much wider and concerns systems for which their linearizations (acting on the larger space $\mathbb{C}^{2n\times 2n}$) are strictly equivalent, meaning systems which can be decoupled by applying congruence or strict equivalence transformations to the isospectral "linearization" λB_1 - B_0 of A(λ), where

$$B_1 = \begin{pmatrix} A_1 & A_2 \\ A_2 & 0 \end{pmatrix}, \quad B_0 = \begin{pmatrix} -A_0 & 0 \\ 0 & A_2 \end{pmatrix}, \tag{6.0.1}$$

  while preserving the structure of A(λ).

6.1 Diagonalization without linearization

Let us first explore the polynomial matrices of the first category.

6.1.1 Symmetric and Hermitian systems: reduction by congruence

When reducing a Hermitian or real-symmetric polynomial matrix real eigenvalues can arise, then knowledge of the *sign characteristic* of each real eigenvalue is required.

Each real eigenvalue has one or more partial multiplicities and a +1 or -1 is associated with each of them. The eigenvalue has

1. *positive type* if all the associated numbers are +1,
2. *negative type* if all the associated numbers are -1.
3. If all the associated numbers are +1 or -1 we say it has *definite type*, otherwise it has *mixed type.*

Here, we first admit semisimple real eigenvalues with no restriction on the type.

Lemma 6.3. *([10]) Let A_2, $A_0 \in \mathbb{C}^{n\times n}$ with det$(A_2) \neq 0$, $A_2^* = A_2$, $A_0^* = A_0$. Assume that $\lambda A_2 + A_0$ is diagonalizable with all eigenvalues real. Let*

$$\Delta = diag[\lambda_1 I_1, \lambda_2 I_2, ..., \lambda_s I_s], \qquad S = diag[\pm I_1, \pm I_2, ..., \pm I_s] \tag{6.1.1}$$

where the size of the identity matrix I_i is a partial multiplicity of eigenvalue λ_i for each i, and the sign of each term in S is determined by the corresponding +1 or -1 in the sign characteristic. Then there exists a family of nonsingular matrices $V \in \mathbb{C}^{n\times n}$ such that

$$V^* A_2 V = S, V^* A_0 V = S\Delta. \tag{6.1.2}$$

If V is one such matrix, then so is any matrix VD where $D = diag[D_1, D_2, ..., D_s]$ and each D_i is unitary with the size of I_i.

An analogous result hols in the case A_2, $A_0 \in \mathbb{R}^{n\times n}$. It is only necessary to use congruence over $\mathbb{R}$ and to replace the unitary matrices A_i by real orthogonal matrices. (This is a special case of Theorem 9.2 of [11].)

Remark 6.4. If, in addition, A_2 is positive definite, all eigenvalues of A(λ) would be real and of positive type, then S = I. This is the original case in the paper of Caughey and O'Kelly [1]. We shall now show a generalization of their theorem.

Theorem 6.5. *([10]) Let the hypothesis of Lemma 6.3 hold, assume also that all eigenvalues of $\lambda A_2 + A_0$ have definite type, and that $A_1^T = A_1$. Then there exists a nonsingular $U \in \mathbb{C}^{n\times n}$ such that $U^* A_2 U$, $U^* A_1 U$, and $U^* A_0 U$ are diagonal if and only if $A_1 M^{-1} A_0 = A_0 M^{-1} A_1$.*
If, in particular, A_2, A_1, A_0 are real and symmetric, then there is a corresponding nonsingular matrix $U \in \mathbb{R}^{n\times n}$ such that $U^T A_2 U$, $U^T A_1 U$, and $U^T A_0 U$ are diagonal if and only if $A_1 A_2^{-1} A_0 = A_0 A_2^{-1} A_1$.

Proof. Usin the notation in Lemma 6.3, from Eq. (6.1.2) we obtain

$$A_2^{-1} A_0 = (VSV^*)(V^{-*} S\Delta V^{-1}) = VS^2 \Delta V^{-1} = V^{-1}. \tag{6.1.3}$$

$$A_0 A_2^{-1} = (V^{-*} S\Delta V^{-1})(VSV^*) = V^{-*} S\Delta S V^* = V^{-*} \Delta U^*. \tag{6.1.4}$$

If also $A_1A_2^{-1}A_0 = A_0A_2^{-1}A_1$ then $A_1V\Delta V^{-1} = V^{-*}\Delta V^*$ and so

$$(V^*A_1V)\Delta = \Delta(V^*A_1V).$$

The assumption that all eigenvalues have definite type means $\lambda_1, \lambda_2, ..., \lambda_s$ in Eq. (6.1.1) are *distinct*. And it follows that V^*A_1V is block-diagonal. But, as the blocks of V^*A_1V are Hermitian, the unitary blocks of matrix D of Lemma 6.3 can be chosen to further reduce V^*A_1V to diagonal form.

Conversely, if U^*A_2U, U^*A_1U, U^*A_0U are diagonal, it is easily verified that $A_1A_2^{-1}A_0 = A_0A_2^{-1}A_1$.

The case of real-symmetric matrices A_2, A_1, A_0 is very similar. □

6.1.2 No symmetry systems: reduction by strict equivalence

For systems which are not Symmetric or Hermitian it is natural to replace the congruence transformations of $A(\lambda)$ by strict equivalence transformations.

Lemma 6.6. *([10]) Let A_2, $A_0 \in \mathbb{C}^{n\times n}$ with $det(A_2) \neq 0$, assume that $\lambda A_2 + A_0$ is semisimple, and write a diagonal matrix of the eigenvalues of $\lambda A_2 + A_0$ in the form*

$$\Delta = diag[\lambda_1 I_1, \lambda_2 I_2, ..., \lambda_s I_s],$$

where $\lambda_i \neq \lambda_j$ when $i \neq j$.

Then there is a family of nonsingular matrices U, $V \in \mathbb{C}^{n\times n}$ such that

$$UA_2V = I \quad and \quad UA_0V = \Delta \tag{6.1.5}$$

If $A = diag[A_1, A_2, ..., A_s]$ is nonsingular and A_j has the size of I_j, then U, V can be replaced by $A^{-1}U$, VA, respectively.

Theorem 6.7. *([10]) Let A_2, A_1, $A_0 \in \mathbb{C}^{n\times n}$ with $det(A_2) \neq 0$ and assume that $\lambda A_2 + A_0$ has n distinct eigenvalues. Then there exist nonsingular U, $V \in \mathbb{C}^{n\times n}$ such that $UA_2V = I$, and UA_1V, UA_0V are diagonal if and only if $A_1A_2^{-1}A_0 = A_0A_2^{-1}A_1$.*

Proof. First use Lemma 4.3 to obtain nonsingular U, V $\in \mathbb{C}^{n\times n}$ such that UA_2V = I and UA_0V = Δ, a diagonal matrix. Then

$$A_2^{-1}A_0 = (VU)(U^{-1}\Delta V^{-1}) = V\Delta V^{-1}, \tag{6.1.6}$$
$$A_0A_2^{-1} = (U^{-1}\Delta V^{-1})(VU) = U^{-1}\Delta U. \tag{6.1.7}$$

If also $A_1A_2^{-1}A_0 = A_0A_2^{-1}A_1$ then $A_1(V\Delta V^{-1})$ =$(U^{-1}\Delta U)A_1$ and hence

$$(UA_1V)\Delta = \Delta(UA_1V) \tag{6.1.8}$$

Since Δ is diagonal with distinct diagonal entries, this implies that UA_1V is also diagonal, as required. Conversely, if A_{2_0} = UA_2V, A_{1_0} = UA_1V, A_{0_0} = UA_0V are diagonal, it is easily verified that $A_1A_2^{-1}A_0 = A_0A_2^{-1}A_1$. □

We have seen that many systems cannot be diagonalized by strict equivalence or congruence because a very strong commutativity condition must be satisfied. Nonetheless, most systems admit a diagonal isospectral system through which they can be decoupled, will explore this in more detail in the following section.

6.2 Systems with nonsingular leading coefficient

In this section the spectrum of a diagonal matrix $A(\lambda) \in \mathbb{C}^{n\times n}[\lambda]$ with $\deg(A(\lambda)) = 2$ and $\det(A_2)\neq 0$ is characterized with the aim of determining necessary and sufficient conditions for it to admit an isospectral diagonal system.

Suppose $A(\lambda) \in \mathbb{C}^{n\times n}[\lambda]$ with $\deg(A(\lambda)) = 2$ and $\det(A_2) \neq 0$ and let $\lambda_1, ..., \lambda_t$ be its eigenvalues with corresponding algebraic, geometric and partial multiplicities $\bar{\mu}_i$, μ_i and m_{ij}, i = 1, ..., t and j = 1, ..., μ_i respectively. Then:

$$\sum_{i=1}^{t} \bar{\mu}_i = 2n. \tag{C1}$$

Example 6.8. Let

$$A(\lambda) = \lambda^2 \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} + \lambda \begin{pmatrix} -1 & -1 & -3/2 \\ -1 & 0 & -2 \\ -3/2 & -2 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 1/2 \\ 0 & 0 & 1 \\ 1/2 & 1 & 0 \end{pmatrix}$$

First note that $\det(A_2) = 1 \neq 0$. Now, $\det(A(\lambda)) = \lambda^2(\lambda-1)^4$. Therefore $A(\lambda)$ has eigenvalues $\lambda_1 = 0$ and $\lambda_2 = 1$ with corresponding algebraic multiplicities $\bar{\mu}_1 = 2$ and $\bar{\mu}_2 = 4$. And condition (C1) holds.

If in addition, $A(\lambda)$ is diagonal then:

1. μ_i coincides with the number of entries in the diagonal of $A(\lambda)$ containing at least one linear factor $(\lambda - \lambda_i)$ and
2. the partial multiplicities of λ_i coincide with the multiplicities of λ_i in the entries of $A(\lambda)$. Therefore,
3. $$1 \le m_{ij} \le 2, \quad i = 1, ..., t, \quad j = 1, ..., \mu_i, \tag{C2}$$

   and in consequence, $\mu_i \ge \frac{\bar{\mu}_i}{2}$.

Example 6.9. Consider the matrix

$$A(\lambda) = \begin{pmatrix} \lambda^2+1 & 0 \\ 0 & \lambda^2 \end{pmatrix} \in \mathbb{C}^{2\times 2}[\lambda].$$

$A_2 = I_2$ is nonsingular and $\det(A(\lambda)) = (\lambda^2+1)\lambda^2$. Therefore, its eigenvalues are $\lambda_1 = 0$, $\lambda_2 = i$ and $\lambda_3 = -i$. Note that the Smith canonical form of $A(\lambda)$ is

$$S(\lambda) = \begin{pmatrix} 1 & 0 \\ 0 & \lambda^2(\lambda-i)(\lambda+i) \end{pmatrix} \in \mathbb{C}^{2\times 2}[\lambda] \tag{6.2.1}$$

and its elementary divisors are λ^2, $(\lambda - i)$ and $(\lambda + i)$. Then the algebraic multiplicities of λ_1, λ_2 and λ_3 are $\bar{\mu}_1 = 2$, $\bar{\mu}_2 = 1$ and $\bar{\mu}_3 = 1$, the geometric multiplicities are $\mu_i = 1$, i = 1, 2, 3 and partial multiplicities $m_{11} = 2$, $m_{21} = 1$, $m_{31} = 1$. The above conditions 1, 2, 3 stand and $\mu_i \ge \frac{\bar{\mu}_i}{2}$, i = 1, 2, 3.

Remark 6.10. Note that each linear factor of the elements in the diagonal of A(λ), coincides with an elementary divisor of its Smith canonical form. And each quadratic elementary divisor is associated with just one entry in the diagonal of A(λ).

For each distinct eigenvalue $\lambda_i \in \mathbb{C}$, i = 1, ..., t, we define the positive integers $s_i \geq 0$ by writing

$$m_{ij} = \begin{cases} 2 & \text{for } j = 1, 2, ..., s_i \\ 1 & \text{for } j = s_i + 1, ..., \mu_i. \end{cases}$$

The total number of quadratic elementary divisors is then

$$p = \sum_{i=1}^{t} s_i, \tag{6.2.2}$$

and, the n - p remaining entries of A(λ) contain distinct eigenvalues. So, the number of linear elementary divisors associated with λ_i is smaller than or equal to n - p, finally

$$\frac{\bar{\mu}_i}{2} \leq \mu_i \leq n - p + s_i. \tag{C3}$$

Theorem 6.11. *([13]) Let $A(\lambda) = A_2\lambda^2 + A_1\lambda + A_0 \in \mathbb{C}^{n\times n}[\lambda]$ represent a regular system with no eigenvalue at infinity ($det(A_2) \neq 0$). There exists a diagonal system $\hat{A}(\lambda)$ isospectral to $A(\lambda)$, if and only if (C1), (C2) and (C3) hold.*

Proof. The necessity has been established. To prove the sufficiency, suppose that (C1), (C2) and (C3) hold and consider the Jordan structure of A(λ) as above.

From (C2) we know that the largest degree of the elementary divisors is 2, so, each of these divisors could be associated with one of the n entries in the diagonal of $\bar{A}(\lambda)$. To finish the proof we must verify that the number

$$q = \sum_{i=1}^{t}(\mu_i - s_i) \tag{6.2.3}$$

of remaining linear elementary divisors of A(λ) is even.

Suppose there are s_i quadratic divisors associated with λ_i then,

$$2s_i + \mu_i - s_i = \mu_i + s_i = \bar{\mu}_i \tag{6.2.4}$$

and because of (C1) and (6.2.2)

$$\sum_{i=1}^{t} \mu_i + p = 2n$$

so finally

$$\sum_{i=1}^{t} \mu_i + p = 2n - 2p$$

$$\sum_{i=1}^{t} (\mu_i - s_i) = q = 2(n - p)$$

Now, condition (C3) ensures that the maximum number of linear elementary divisors associated with λ_i for i = 1, ..., t is n - p. Therefore, the linear elementary divisors of A(λ) can be organized in n - p pairs with distinct eigenvalues so that the remaining n - p entries in the diagonal of $\bar{A}(\lambda)$ can be constructed. $\square$

Example 6.12. Let

$$\mathrm{A}(\lambda) = \begin{pmatrix} \lambda^2+\lambda+1 & \lambda+1 \\ \lambda+1 & \lambda^2+2\lambda+1 \end{pmatrix} = \lambda^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \lambda \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

In this case det(A(λ)) $= (\lambda+1)^4$ so there is a unique eigenvalue: $\lambda_1 =$ -1, with $\bar{\mu} = 4$ and $\mu = 1$. Therefore, condition (C3) does not hold ($\frac{4}{2} \nleq 1$) and A(λ) is not diagonalizable.

Example 6.13. Now consider

$$A(\lambda) = \lambda^2 \begin{pmatrix} 0 & 1 \\ 1 & 3 \end{pmatrix} + \lambda \begin{pmatrix} -1 & -3 \\ -3 & -7 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}.$$

In this case, det(A(λ)) = -$\lambda(\lambda-1)^3$. So its eigenvalues are $\lambda_1 = 0$ and $\lambda_2 = 1$ with algebraic multiplicities $\bar{\mu}_1 = 1$ and $\bar{\mu}_2 = 3$ and geometric multiplicities $\mu_1 = 0$ and $\mu_2 = 2$. Now observe that its Smith canonical form is

$$S(\lambda) = \begin{pmatrix} \lambda-1 & 0 \\ 0 & \lambda(\lambda-1)^2 \end{pmatrix}.$$

Therefore its elementary divisors are λ, $\lambda-1$ and $(\lambda-1)^2$. Thus, all the conditions of Theorem 6.11 are satisfied and $A(\lambda)$ is isospectral to the following diagonal matrix:

$$\begin{pmatrix} (\lambda-1)^2 & 0 \\ 0 & (\lambda-1)\lambda \end{pmatrix}$$

.

6.3 Systems with singular leading coefficient

Now we extend the results Section 6.3 to admit systems with a singular leading coefficient. In particular, we will see that the results of Theorem 6.11 can be applied to any regular matrix even if A_2 is singular ($\det(A_2) = 0$).

Let us begin by illustrating some possible cases that could arise through examples.

Example 6.14. Consider the matrix

$$\mathrm{A}(\lambda) = \begin{pmatrix} 1-2\lambda+\lambda^2 & 1-2\lambda+\lambda^2 & -1+2\lambda-\lambda^2 & 0 \\ 0 & \lambda^2 & 0 & 0 \\ 0 & 0 & 1 & \lambda^2 \\ 0 & 0 & 0 & -1+\lambda \end{pmatrix} \in \mathbb{C}^{4\times 4}[\lambda]$$

with eigenvalues $\lambda_1 = 0$ and $\lambda_2 = 1$ and partial multiplicities $m_{11} = 2, m_{12} = 2$ and $m_{22} = 1$. Note that the elementary divisors of A(λ) are $(\lambda-1)^2$, λ^2 and $(\lambda-1)$. Clearly the conditions of Theorem 6.11 hold (regardless of the eigenvalue at infinity), and a diagonal matrix sharing the same finite eigenvalue structure as A(λ) is

$$\hat{A}(\lambda) = \begin{pmatrix} (\lambda-1)^2 & 0 & 0 & 0 \\ 0 & \lambda^2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \lambda-1 \end{pmatrix}$$

Notice however that the structure at infinity is not preserved. The partial multiplicity of the eigenvalue at infinity in A(λ) is $m_{\infty 1} = 3$, whereas in $\hat{A}(\lambda)$ we have $m_{\infty 1} = 2$ and $m_{\infty 2}$= 1. In fact, as we will now show, there does not exist a diagonal matrix of degree 2 sharing the same finite and infinite structure as A(λ).

Example 6.15. Now consider the matrix

$$A(\lambda) = \begin{pmatrix} -1+\lambda & 0 & 0 \\ 2\lambda-\lambda^2 & 0 & -2+\lambda \\ -9+6\lambda-\lambda^2 & 9-6\lambda+\lambda^2 & 0 \end{pmatrix} \in \mathbb{C}^{3\times 3}[\lambda].$$

Det(A(λ)) = -$(\lambda-1)(\lambda-2)(\lambda-3)^2$. Therefore its eigenvalues are $\lambda_1 = 1$, $\lambda_2 = 2$ and $\lambda_3 = 3$ with algebraic multiplicities $\bar{\mu}_1 = 1$, $\bar{\mu}_2 = 1$ and $\bar{\mu}_3 = 2$ and geometric multiplicities $\mu_i = 1$, i = 1, 2, 3.

Similarly, the algebraic and geometric multiplicities of the eigenvalue at infinity are equal to 2 and 1, respectively. Clearly the conditions of Theorem 6.11 hold, and two diagonal matrices sharing the same finite eigenvalue structure as A(λ) are

$$\hat{A}_1(\lambda) = \begin{pmatrix} -1+\lambda & 0 & 0 \\ 0 & -2+\lambda & 0 \\ 0 & 0 & (-3+\lambda)^2 \end{pmatrix} \quad \text{and}$$

$$\hat{A}_2(\lambda) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & (-2+\lambda)(-1+\lambda) & 0 \\ 0 & 0 & (-3+\lambda)^2 \end{pmatrix}.$$

Notice that only $\hat{A}_2(\lambda)$ preserves the infinite eigenvalue structure of A(λ).

Remark 6.16. We will say that the matrix $\hat{A}_2(\lambda)$ in the Example 6.15 above is a diagonal matrix strongly equivalent to A(λ), i.e. that A(λ) admits a strongly isospectral diagonal system $\hat{A}_2(\lambda)$.

Next we derive the conditions for a system to admit a strongly equivalent diagonal matrix. Let us start by analyzing the structure of systems A(λ) which are already diagonal. Note that, in this case, the diagonal could include not only quadratic terms, but also linear and constant terms.

Suppose that, in addition to the distinct eigenvalues λ_i for i = 1, ..., t, the diagonal matrix A(λ) has an eigenvalue at infinity with partial multiplicities $m_{\infty 1} \geq m_{\infty 2} \geq \cdots \geq m_{\infty \mu_\infty}$ where $\mu_\infty \leq n$ is its geometric multiplicity, and let $\bar{\mu}_\infty$ denote is its algebraic multiplicity. Then,

$$\sum_{i=1}^{t} \bar{\mu}_i + \bar{\mu}_\infty = 2n. \qquad \text{(C1')}$$

Remark 6.17. Note that the geometric multiplicity of the eigenvalue at infinity is the number of constant and linear entries in the diagonal of A(λ), and its algebraic multiplicity is the multiplicity of $\lambda = 0$ in the polynomial $\det(A_*(\lambda))$, where $A_*(\lambda)$ is the reverse matrix of A(λ).

Clearly then, the partial multiplicities of λ_i are the multiplicities of λ_i in the entries of A(λ), and the partial multiplicities of the point at infinity are equal to 2 for any constant entry in the diagonal of A(λ) and equal to 1 for any linear entry. Then

$$\begin{aligned} 1 \leq m_{ij} \leq 2, \quad i = 1, .., t, \quad j = 1, ..., \mu_i, \\ 1 \leq m_{\infty j} \leq 2, \quad j = 1, ..., \mu_\infty, \end{aligned} \tag{C2'}$$

and, in consequence,

$$\mu_i \geq \frac{\bar{\mu}_i}{2}, \qquad \mu_\infty \geq \frac{\bar{\mu}_\infty}{2}.$$

Each quadratic finite elementary divisor (i.e. each partial multiplicity $m_{ij} = 2$) is associated with just one quadratic entry in the diagonal of A(λ), and each quadratic infinite elementary divisor (i.e. each partial multiplicity $m_{\infty j} = 2$) is associated with just one constant entry. Define also

$$m_{\infty j} = \begin{cases} 2 & \text{for } j = 1, 2, ..., s_\infty \\ 1 & \text{for } j = s_\infty + 1, ..., \mu_\infty. \end{cases}$$

The number of constant entries in A(λ) is then s_∞ and, in consequence, the $n - p - s_\infty$ other entries of A(λ) contain two distinct finite eigenvalues or one finite eigenvalue and the infinite. So, the number of linear elementary divisors associated with λ_i is smaller than or equal to $n - p - s_\infty$, finally

$$\begin{aligned} \frac{\bar{\mu}_i}{2} \leq \mu_i \leq n - p - s_\infty + s_i, \\ \frac{\bar{\mu}_\infty}{2} \leq \mu_\infty \leq n - p. \end{aligned} \tag{C3'}$$

Theorem 6.18. *([13]) Let $A(\lambda) = A_2\lambda^2 + A_1\lambda + A_0$ be a regular system with an eigenvalue at infinity. There exists a diagonal system $\hat{A}(\lambda)$ with the same finite and infinite structure as $A(\lambda)$, if and only if (C1'), (C2') and (C3') hold.*

Proof. The necessity is already established.
To prove the sufficiency, suppose that (C1'), (C2') and (C3') hold. From (C2') we know that the largest degree of the elementary finite and infinite divisors is 2, so, each of these divisors could be associated to one of the n entries in the diagonal of $\hat{A}(\lambda)$ as a quadratic entry or as a constant entry respectively. So to finish the proof we must verify that the number

$$\bar{q} = \sum_{i=1}^{t} (\mu_i - s_i) + \mu_\infty - s_\infty$$

of linear elementary finite and infinite divisors of A(λ) is even and, moreover, they can be organized in $n - p - s_\infty$ pairs with distinct eigenvalues (possibly infinite)

λ_{k1} and λ_{k2} for k = 1, ..., $n-p-s_\infty$, so that the remaining $n-p-s_\infty$ entries in the diagonal of $\hat{A}(\lambda)$ can be constructed.
If there are s_i quadratic divisors associated with λ_i and s_∞ quadratic infinite divisors then,

$$2s_i + \mu_i - s_i = \mu_i + s_i = \bar{\mu}_i, \quad \text{for } i = i, ..., t,$$

$$2s_\infty + \mu_\infty - s_\infty = \mu_\infty + s_\infty = \bar{\mu}_\infty$$

and because of (C1') and (6.2.2)

$$\sum_{i=1}^{t} \mu_i + p + \mu_\infty + s_\infty = 2n,$$

so finally

$$\sum_{i=1}^{t} \mu_i - p + \mu_\infty - s_\infty = 2n - 2p - 2s_\infty$$

$$\sum_{i=1}^{t} (\mu_i - s_i) + \mu_\infty - s_\infty = \bar{q} = 2(n - p - s_\infty).$$

Now condition (C3') ensures that the maximum number of linear elementary divisors associated with λ_i for i = 1, ..., t, and the maximum number of linear elementary divisors at infinity is $n-p-s_\infty$ and, in consequence, that the $\bar{q}$ linear elementary finite and infinite divisors of A(λ) can be organized in $n-p-s_\infty$ pairs with distinct eigenvalues (possibly including the point at infinity). □

6.4 Diagonalization through linearization

Let us suppose A(λ) $\in \mathbb{C}^{n\times n}[\lambda]$ admits a diagonal isospectral system. We will now dig into the problem of generating an isospectral diagonal system by the application of strict equivalence or congruence transformations to the linearization $\lambda B_1 - B_0$ defined in (6.0.1). The idea is to find nonsingular matrices U, V $\in \mathbb{C}^{2n\times 2n}$ such that

$$\lambda\hat{B}_1 - \hat{B}_0 = U(\lambda B_1 - B_0)V \tag{6.4.1}$$

is the linearization of a diagonal system

$$\hat{A}(\lambda) = \hat{A}_2\lambda^2 + \hat{A}_1\lambda + \hat{A}_0.$$

The following specific classes of transformations will be considered. Notice they all preserve the Jordan structures.

Definition 6.19. *([10])*

1. *A system $A(\lambda) = A_2\lambda^2 + A_1\lambda + A_0 \in \mathbb{C}^{n\times n}[\lambda]$ is DE$\mathbb{C}$ (diagonalizable by strict equivalence over $\mathbb{C}$) if there exist nonsingular U, $V \in \mathbb{C}^{2n\times 2n}$ such that*

$$U(\lambda B_1 - B_0)V = \lambda\hat{B}_1 - \hat{B}_0,$$

where $\lambda\hat{B}_1 - \hat{B}_0$ is the linearization of a (generally complex) diagonal system $\hat{A}(\lambda) = \lambda^2\hat{A}_2 + \lambda\hat{A}_1 + \hat{A}_0$.

2. *A real system $A(\lambda) = A_2\lambda^2 + A_1\lambda + A_0 \in \mathbb{R}^{n\times n}[\lambda]$ is DE$\mathbb{R}$ (diagonalizable by strict equivalence over $\mathbb{R}$) if there exist nonsingular U, $V \in \mathbb{C}^{2n\times 2n}$ such that*

$$U(\lambda B_1 - B_0)V = \lambda\hat{B}_1 - \hat{B}_0,$$

where $\lambda\hat{B}_1 - \hat{B}_0$ is the linearization of a real diagonal *system $\hat{A}(\lambda) = \lambda^2\hat{A}_2 + \lambda\hat{A}_1 + \hat{A}_0$.*

3. *A system $A(\lambda) = A_2\lambda^2 + A_1\lambda + A_0 \in \mathbb{C}^{n\times n}[\lambda]$ is DC$\mathbb{R}$ (diagonalizable by congruence) if there exists a nonsingular $U \in \mathbb{C}^{2n\times 2n}$ such that*

$$U(\lambda B_1 - B_0)U^T = \lambda\hat{B}_1 - \hat{B}_0,$$

where $\lambda\hat{B}_1 - \hat{B}_0$ is the linearization of a real diagonal *system $\hat{A}(\lambda) = \lambda^2\hat{A}_2 + \lambda\hat{A}_1 + \hat{A}_0$.*

Remark 6.20. In the third case, if the system is Hermitian (or real and symmetric), then so are $\hat{B}_1, \hat{B}_0$ and, in particular, because $\hat{A}_0$ and $\hat{A}_2$ are diagonal, so is $\hat{B}_0$.

Let us denote by $\mathbb{J}_{n,\mathbb{C}}$ and $\mathbb{J}_{n,\mathbb{R}}$ the classes of 2n x 2n canonical Jordan matrices for n x n diagonal complex and real systems.

Theorem 6.21. *([10])*

1. *$A(\lambda) \in \mathbb{C}^{n\times n}[\lambda]$ of deg($A(\lambda)$) $= 2$ with Jordan form $J \in \mathbb{C}^{2n\times 2n}$ is DE$\mathbb{C}$ if and only if $J \in \mathbb{J}_{n,\mathbb{C}}$.*
2. *$A(\lambda) \in \mathbb{R}^{n\times n}[\lambda]$ of deg($A(\lambda)$) $= 2$ with Jordan form $J \in \mathbb{C}^{2n\times 2n}$ is DE$\mathbb{R}$ if and only if $J \in \mathbb{J}_{n,\mathbb{R}}$.*
3. *An Hermitian system $A(\lambda) \in \mathbb{C}^{n\times n}[\lambda]$ with Jordan form $J \in \mathbb{C}^{2n\times 2n}$ is DC$\mathbb{R}$ if and only if $J \in \mathbb{J}_{n,\mathbb{R}}$.*

Proof. 1. Following definition 6.19, if A(λ) is DE$\mathbb{C}$ then $\lambda B_1 - B_0$ and $\lambda\hat{B}_1 - \hat{B}_0$ have the same Jordan form J, and since $\lambda\hat{B}_1 - \hat{B}_0$ is the linearization of a diagonal system, J $\in \mathbb{J}_{n,\mathbb{C}}$.
Conversely, A(λ) has Jordan form J $\in \mathbb{J}_{n,\mathbb{C}}$ implies that there exists an strictly isospectral diagonal system $\hat{A}(\lambda)$. Thus, $\lambda B_1 - B_0$ and $\lambda\hat{B}_1 - \hat{B}_0$ are isospectral. But then it follows from the Kronecker reduction of regular polynomial matrices ([11], Theorem 3.1) that the systems are strictly equivalent to the same canonical form and hence to one another.

2. Suppose A(λ) is DE$\mathbb{R}$. Then, in definition 6.19, $\lambda B_1 - B_0$ and $\lambda\hat{B}_1 - \hat{B}_0$ have the same Jordan form J and, since $\hat{A}(\lambda)$ is real diagonal, J $\in \mathbb{J}_{n,\mathbb{R}}$. The converse argument is as in (1) but over the real field. ([11], Theorem 3.2)

3. Suppose A(λ) is a DC$\mathbb{R}$ hermitian system. Then, in definition 6.19, $\lambda B_1 - B_0$ and $\lambda\hat{B}_1 - \hat{B}_0$ have the same Jordan form J and, because $\hat{A}$ is diagonal, J $\in \mathbb{J}_{n,\mathbb{R}}$ as required.
Conversely, let the Hermitian system A(λ) have Jordan form J $\in \mathbb{J}_{n,\mathbb{R}}$ and let ϵ be its sign characteristic. Then, J $\in \mathbb{J}_{n,\mathbb{R}}$ implies that there exist isospectral real diagonal systems $\hat{A}(\lambda)$. According to (C2) the partial multiplicities of

the eigenvalues of A(λ) (and then of $\hat{A}(\lambda)$) are either 2 or 1. By Proposition 10.12 in [6] it can be concluded that for the real semisimple eigenvalues the number of +1's and -1's in ϵ is equal. Furthermore, the diagonal terms of (λ) with distinct real zeros necessarily combine pairs of eigenvalues with opposite signs.
Given one such $\hat{A}(\lambda)$, multiplication by a diagonal of +1's and -1's and exchanging factors corresponding to semisimple real eigenvalues along the diagonal generates another isospectral diagonal system. Using this freedom, and knowing the signs attached to the real eigenvalues of A(λ), corresponding signs can be associated with the real eigenvalues of $\hat{A}(\lambda)$. In this way an $\hat{A}(\lambda)$ is determined which is strictly isospectral with A(λ).
Now let $\lambda B_1 - B_0$, $\lambda \hat{B}_1 - \hat{B}_0$ be the linearizations of A(λ) and $\hat{A}(\lambda)$, respectively, and note that each one inherits both the spectrum and sign characteristic of the parent polynomial. Then it follows that the systems $\lambda B_1 - B_0$ and $\lambda \hat{B}_1 - \hat{B}_0$ have the same canonical forms and are therefore congruent. Thus, A(λ) is DC$\mathbb{R}$.

□

6.4.1 Algorithms

Now we could ask ourselves how the matrices U and V in Eq. (6.4.1) in the case A(λ) $= A_2\lambda^2 + A_2\lambda + A_0 \in \mathbb{C}^{n\times n}[\lambda]$ has an isospectral diagonal system $\hat{A}(\lambda) \in \mathbb{C}^{n\times n}[\lambda]$.

This problem has been tackled by Chu and Del Buono in [3] and [4], but only for systems with singular leading coefficient. In particular they found two different algorithms to calculate U and V. The first, the *spectral decomposition algorithm*, is based on the Jordan structure of A(λ) whereas the second, the so called *isospectral flow algorithm* does not need the computation of the eigenvalues of A(λ).

It should be noted that Zúñiga explored both these algorithms in [13] and concluded none of them is practically useful, therefore the development of practical methods to calculate U and V is still open.

6.5 Third degree diagonalization

The diagonalization problem has also been recently solved for polynomial matrices

$$A(\lambda) = A_3\lambda^3 + A_2\lambda^2 + A_1\lambda + A_0$$

where $A_3, A_2, A_1, A_0 \in \mathbb{C}^{n\times n}[\lambda]$ in ([14]).

In this case, the complexity is higher and A(λ) must meet even more conditions. Let us show that the conditions found for second degree polynomial matrices are not enough to conclude if third degree polynomial matrix is diagonalizable. We will see it through a counterexample.

Before, let us replicate conditions (C1), (C2) and (C3) for third degree polynomial matrices.

Suppose $A(\lambda) = A_3\lambda^3 + A_2\lambda^2 + A_1\lambda + A_0 \in \mathbb{C}^{n\times n}[\lambda]$ with deg(A_3)$\neq 0$ has eigenvalues $\lambda_1, ..., \lambda_t$ with corresponding algebraic, geometric and partial multiplicities $\bar{\mu}_i$, μ_i and m_{ij}, i = 1, ..., t, j = 1, ..., μ_i.
Then, from Lemma 3.27 we know that

$$\sum_{i=1}^{t} \bar{\mu}_i + \bar{\mu}_\infty = 3n. \tag{C1''}$$

and it can easily be seen that

$$1 \leq m_{ij} \leq 3, \quad i = 1, ..., t, \quad j = 1, ..., \mu_i. \tag{C2''}$$

For each eigenvalue $\lambda_i \in \mathbb{C}$, i = 1, ..., t, we define the positive integers r_i and p_i by writing

$$m_{ij} = \begin{cases} 3 & \text{for } j = 1,\ 2,\ ...,\ r_i \\ 2 & \text{for } j = r_i + 1,\ ...,\ r_i + p_i \\ 1 & \text{for } j = r_i + p_i + 1,\ ...,\ \mu_i. \end{cases}$$

Then the total number of cubic elementary divisors is

$$k = \sum_{i=1}^{t} r_i, \quad i = 1, ..., t$$

and

$$\mu_i \leq n - k + k_i, \quad i = 1, ..., t. \tag{C3''}$$

Now let us see a counterexample. Consider the matrix polynomial

$$A(\lambda) = \begin{pmatrix} (\lambda-1)^2(\lambda-3) & -2(\lambda-1)(\lambda-3) & -(\lambda-1)(\lambda-3) \\ 0 & (\lambda-1)(\lambda-2)(\lambda-3) & -1(\lambda-1)(\lambda-3) \\ 0 & 0 & (\lambda-1)(\lambda-2)(\lambda-3) \end{pmatrix}.$$

Det(A(λ)) = $(\lambda-1)^4(\lambda-2)^2(\lambda-3)^3$. Therefore, its eigenvalues are $\lambda_1 = 1$, $\lambda_2 = 2$ and $\lambda_3 = 3$.
Note that the Smith canonical form of A(λ) is

$$S(\lambda) = \begin{pmatrix} (\lambda-1)(\lambda-3) & 0 & 0 \\ 0 & (\lambda-1)(\lambda-3) & 0 \\ 0 & 0 & (\lambda-1)^2(\lambda-2)^2(\lambda-3) \end{pmatrix},$$

then its elementary divisors are

$$(\lambda-\lambda_1)^2, (\lambda-\lambda_1), (\lambda-\lambda_1), (\lambda-\lambda_2)^2, (\lambda-\lambda_3), (\lambda-\lambda_3), (\lambda-\lambda_3). \tag{6.5.1}$$

It can easily be seen that conditions (C1'') and (C2'') and (C3'') hold.

However, the only possible distributions for the elementary divisors in products of degree 3 are:

- $(\lambda-\lambda_1)^2(\lambda-\lambda_1), (\lambda-\lambda_2)^2(\lambda-\lambda_1), (\lambda-\lambda_3)(\lambda-\lambda_3)(\lambda-\lambda_3)$
- $(\lambda-\lambda_1)^2(\lambda-\lambda_3), (\lambda-\lambda_2)^2(\lambda-\lambda_1), (\lambda-\lambda_1)(\lambda-\lambda_3)(\lambda-\lambda_3)$
- $(\lambda-\lambda_1)^2(\lambda-\lambda_1), (\lambda-\lambda_2)^2(\lambda-\lambda_3), (\lambda-\lambda_1)(\lambda-\lambda_3)(\lambda-\lambda_3)$
- $(\lambda-\lambda_1)^2(\lambda-\lambda_3), (\lambda-\lambda_2)^2(\lambda-\lambda_3), (\lambda-\lambda_1)(\lambda-\lambda_1)(\lambda-\lambda_3)$,

and no diagonal matrix with these diagonal elements has the polynomials of (6.5.1) as elementary divisors.

Chapter 7

Conclusions

The diagonalization of polynomial matrices has plenty of applications in science and engineering problems such as signal processing and control theory mainly.

The criteria for diagonalization of first degree polynomial matrices, as has been shown, is quite obvious and no additional conditions are needed. For this reason, the main subject of the work is the diagonalization of second degree polynomial matrices. First, systems that can be directly decoupled to a diagonal system by applying congruence or strict equivalence transformations have been studied. The study of this first category has been divided between systems with nonsingular and singular leading coefficients. Secondly, the conditions for the diagonalization of systems for which their linearizations are strictly equivalent have been specified.

As shown in the last section of the work, for the diagonalization of third degree systems the established conditions for second degree systems are not enough to decide whether a matrix is diagonalizable or not. However, though it hasn't been shown in this work, this case has been solved, as well as for fourth degree polynomials matrices.

With respect to matrices with fifth degree or above, the problem remains open. The main difficulty lies in the increase of possible permutations of the elementary divisors in the diagonal matrix.

To sum up, the main goal of the work, the study of the second degree case, has been achieved.

Bibliography

[1] CAUGHEY, T.K.; O'KELLY, M.E.J. *Classical normal modes in damped linear dynamic systems*, ASME J. Appl. Mech. 32 (1965) 583-588.

[2] CAUGHEY, T.K.; MA, F. *Analysis of linear nonconservative vibrations*, ASME J. Appl. Mech. 62 (1995) 685-691.

[3] CHU, M.; DEL BUONO, N. *Total decoupling of general quadratic pencils, Part 1: Theory*, J. Sound Vibration 309 (2008) 96-111.

[4] CHU, M.; DEL BUONO, N. *Total decoupling of general quadratic pencils, Part 2: Structure preserving isospectral flows*, J. Sound Vibration 309 (2008) 112-128.

[5] GANTMATCHER, F.R. *The theory of matrices*, vol. 1, K.A. Hirsch, 130-174, 2000.

[6] GOHBERG, I.; LANCASTER, P.; RODMAN, L. *Matrix Polynomials*, Academic Press, New York, 1982.

[7] KAILATH, T. *Linear systems*, Prentice-Hall, Inc., Englewood Cliffs, N.J.; 1980.

[8] LANCASTER, P. *Linearization of regular matrix polynomials*, Electron. J. Linear Algebra 17 (2008) 21-27.

[9] LANCASTER, P.; TISMENETSKY, M. *The Theory of Matrices Second edition with applications*, Academic Press, Inc.; 1985.

[10] LANCASTER, P.; ZABALLA, I. *Diagonalizable quadratic eigenvalue problems*, Mechanical Systems and Signal Processing 23 (2009) 1134-1144.

[11] LANCASTER, P.; RODMAN, L., *Canonical forms for Hermitian matrix pairs under strict equivalence and congruence*, SIAM Rev. 47, 2005.

[12] TISSEUR, F.; MEERBERGEN, K. *The quadratic eigenvalue problem*, SIAM 43(2) (2001) 235–286

[13] ZÚÑIGA, J.C. *Diagonalization of quadratic matrix polynomials*, Systems Control Letters 59 (2010) 105-113.

[14] YU, P; ZHANG, G; ZHANG, Y. *Decoupling of cubic polynomial matrix systems*, Numerical algebra, control and optimitzation, 11(1) (2021) 13-26.