Deanship of Graduate Studies

Al-Quds University

MATRIX POLYNOMIALS

By

Ayed Mohamed Ahmed Abed Al-Ghani

M. Sc. Thesis

2002

Deanship of Graduate Studies

Al-Quds University

MATRIX POLYNOMIALS

By

Ayed Mohamed Ahmed Abed Al-Ghani

M. Sc. Thesis

2002

MATRIX POLYNOMIALS

By

Ayed Mohamed Ahmed Abed Al-Ghani

This thesis submitted in partial fulfillment of the requirements for the degree of Master of Science.

Department of Mathematics/ Graduate Studies

Al-Quds University

August, 2002

Graduate Studies/ Mathematics

Deanship of Graduate Studies

MATRIX POLYNOMIALS

By

Student Name: Ayed Mohamed Ahmed Abed Al-Ghani

Registration No: 9910323

Supervisor: Dr. Amina Afaneh

Co advisor: Dr. Mohamed Khalil

Master thesis submitted and accepted, Date: 21/8/2002

The names and signatures of the examining committee members are as follows:

1- Dr. Mohamed Khalil Head of Committee Signature ....

2- Dr. Tahseen Mughrabi Internal Examiner Signature ....

3- Dr. Nureddin Rabie External Examiner Signature ....

Al-Quds University

Declaration:

I Certify that this thesis submitted for the degree of Master is the result of my own research except where otherwise acknowledged, and that this thesis (or any part of the same) has not been submitted for a higher degree to any other university or institution.

Signed ....

Ayed Mohamed Ahmed Abed Al-Ghani

Date: 21/8/2002

Dedication

To the memory of the late my grandfather.

To my parents.

Acknowledgement

The words are unable to express how I am beholden to those memorable people who helped me to prepare and complete this study. Their names are graven in my memory.

I owe the success of this work to Dr. Amina Afaneh, to her constructive and helpful suggestions regarding the appropriate resources and references of this study. Her on going efforts and final checking of my work helped me a lot to overcome all the difficulties I faced while I was working on this dissertation.

I am also very grateful, despite the inability of these words to express my thanks to Dr. Mohamed Khalil, Dr. Tahseen Mugrabi and Dr. Nureddin Rabie for their contributions in studying, analyzing, checking and providing me with valuable remarks and comments. I have learned so much from them and working with them was a pleasure.

My thanks also to the members of the department of mathematics at Al-Quds University.

Contents

CHAPTER FOUR

JORDAN CHAIN AND DIFFERENTIAL EQUATIONS

Introduction

The matrix polynomial or a λ-matrix is a matrix-valued function of a complex variable of the form $A(\lambda) = A_l\lambda^l + A_{l-1}\lambda^{l-1} + \cdots + A_0$ where $A_l, A_{l-1}, \cdots, A_0$ are $m \times n$ matrices of complex numbers. Matrix polynomial is a generalization of the matrix polynomial $\lambda I - A$ of degree 1, $\lambda I - A$ is very important in finding the eigenvalues and eigenvectors of the constant $n \times n$ matrix A.

We can also write the matrix polynomial $A(\lambda)$ in the form $A(\lambda) = [a_{ij}(\lambda)]_{i,j=1}^n$, so that when the entries of $A(\lambda)$ are evaluated for a particular value of λ, say $\lambda = \lambda_0$, then $A(\lambda_0) \in \mathbf{C}^{n\times n}$ and if we take $n = 1$, we get a scalar polynomial $a(\lambda)$.

In chapter one we study the notion and the kinds of a matrix polynomial, the condition which is necessary for the matrix polynomial to be invertible and the operations on the matrix polynomial.

In chapter two we turn our attention to the invariant polynomials. The importance of the invariant polynomials developed in this chapter allows us to obtain a canonical form of a matrix polynomial without using the elementary row (column) operations to obtain a canonical form. We also study the generally invertible matrix polynomial, generalized inverse, right inverse and left inverse of a matrix polynomial.

In chapter three, we study a factorization of selfadjoint matrix polynomial of the form $A(\lambda) = (M(\overline{\lambda}))^* D(\lambda) M(\lambda)$, where $D(\lambda)$ is a constant matrix or a matrix polynomial and $M(\lambda)$ is a matrix polynomial.

The study of factorization of a selfadjoint matrix polynomial is very important in several applied problems, such as filtering, see [1], chapter 9.

Factorization of matrix polynomials was developed by many researchers as: V. A. Jakubovic, see [9], I. Gohberg, P. Lancaster, and L. Rodman, see [6] and A. C. M. Ran and L. Rodman, see [15].

In chapter four, standard triple and Jordan chain for a matrix polynomial are used to solve the differential equation of the form $\sum_{i=0}^{l} L_i \vec{x}(t)^{(i)} = \vec{f}(t)$, where $L_i \in \mathbf{C}^{n \times n}$, $i = 0,1,\cdots,l$, $\vec{f}(t)$ is a vector-valued function.

In appendix A, we study the Jordan canonical form for a constant matrix and the exponential of a square matrix.

Remarks:

(i) We will use the following system of notion:

Equation j of chapter i in section k is denoted by $(i.k.j)$, similarly definition j of chapter i in section k is denoted by definition $(i.k.j)$ and similar conventions apply to theorems, corollaries and lemmas.

(ii) Here, we mean by a scalar polynomial a polynomial with scalar coefficients not a polynomial as $p(\lambda) = a$, $a \in \mathbf{C}$.

CHAPTER ONE

MATRIX POLYNOMIALS

This chapter contains definitions, theorems and ideas that we shall need in the following chapters. It consists of three sections, section one is about the notion of a matrix polynomial, in section two addition and multiplication of matrix polynomials are introduced and section three about division of matrix polynomials.

1.1 The Notion of a Matrix Polynomial

Definition 1.1.1: A matrix polynomial is a matrix-valued function of a complex variable of the form

$$A(\lambda) = A_l\lambda^l + A_{l-1}\lambda^{l-1} + \cdots + A_1\lambda + A_0$$

where $A_0, A_1, \cdots, A_l$ are $m \times n$ matrices of complex numbers.

Example 1.1.1: The followings are examples of matrix polynomials

(*i*) $A(\lambda) = \begin{bmatrix} \lambda^2 + \lambda + 1 & \lambda^2 - \lambda + 2 \\ 2\lambda & \lambda^2 - 3\lambda - 1 \end{bmatrix}$

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

(*ii*) $B(\lambda) = I_n\lambda - B$, where I_n is the identity matrix and $B \in \mathbf{C}^{n \times n}$.

(*iii*) $C(\lambda) = C$, where $C \in \mathbf{C}^{n \times n}$, *i.e.* matrix with constant entries.

Definition 1.1.2: The **degree** of a matrix polynomial $A(\lambda)$ (denoted by $\deg A(\lambda)$) is the greatest degree of the scalar polynomials appearing as entries of $A(\lambda)$.

In example 1.1.1 $\deg A(\lambda) = 2$, $\deg B(\lambda) = 1$ and $\deg C(\lambda) = 0$.

We call a matrix polynomial $A(\lambda)$

(i) **monic** if the leading coefficient $A_l = I_n$,

(ii) **comonic** if $A_0 = I_n$,

(iii) **regular** if $\det A(\lambda) \not\equiv 0$,

(iv) **unimodular** if $\det A(\lambda)$ is a nonzero constant independent of λ.

Example 1.1.2:

(i) If $\det A_l \neq 0$ then A_l^{-1} exists

$$A_l^{-1}A(\lambda) = A_l^{-1}[A_l\lambda^l + A_{l-1}\lambda^{l-1} + \cdots + A_1\lambda + A_0]$$

$$= I_n\lambda^l + A_l^{-1}A_{l-1}\lambda^{l-1} + \cdots + A_l^{-1}A_1\lambda + A_l^{-1}A_0.$$

Therefore $A_l^{-1}A(\lambda)$ is a monic matrix polynomial.

(ii) If $\det A_0 \neq 0$ then A_0^{-1} exists

$$A_0^{-1}A(\lambda) = A_0^{-1}[A_l\lambda^l + A_{l-1}\lambda^{l-1} + \cdots + A_1\lambda + A_0]$$

$$= A_0^{-1}A_l\lambda^l + A_0^{-1}A_{l-1}\lambda^{l-1} + \cdots + A_0^{-1}A_1\lambda + I_n.$$

Therefore $A_0^{-1}A(\lambda)$ is a comonic matrix polynomial.

(iii) Let $A(\lambda) = \begin{bmatrix} \lambda & -1 \\ 1 & \lambda \end{bmatrix}$.

Since $\det A(\lambda) = \lambda^2 + 1 \neq 0$, then $A(\lambda)$ is a regular matrix polynomial.

$$(iv)\quad A(\lambda) = \begin{bmatrix} 1 & \lambda & -2\lambda^2 \\ 0 & 1 & \lambda^4 \\ 0 & 0 & 1 \end{bmatrix}.$$

Since $\det A(\lambda) = 1$, then $A(\lambda)$ is unimodular matrix polynomial.

We can transform a comonic matrix polynomial into a monic form, as follows:

Let $A(\lambda)$ be a comonic matrix polynomial, that is $A_0 = I_n$

$$\begin{aligned} B(\lambda) &= \lambda^l A(\lambda^{-1}) \\ &= \lambda^l [A_l(\lambda^{-1})^l + A_{l-1}(\lambda^{-1})^{l-1} + \cdots + A_1\lambda^{-1} + I_n] \\ &= \lambda^l [A_l\lambda^{-l} + A_{l-1}\lambda^{-l+1} + \cdots + A_1\lambda^{-1} + I_n] \\ &= A_l + A_{l-1}\lambda + \cdots + A_1\lambda^{l-1} + I_n\lambda^l \\ &= I_n\lambda^l + A_1\lambda^{l-1} + \cdots + A_{l-1}\lambda + A_l. \end{aligned}$$

Therefore $B(\lambda)$ is a monic matrix polynomial.

Definition 1.1.3: An $n \times n$ matrix polynomial $A(\lambda)$ is **invertible** if there exists an $n \times n$ matrix polynomial $B(\lambda)$ such that $A(\lambda)B(\lambda) = I_n$.

We denote $B(\lambda)$ by $(A(\lambda))^{-1}$.

Theorem 1.1.1: A matrix polynomial $A(\lambda)$ is invertible if and only if it is unimodular

Proof:

If $A(\lambda)$ is invertible then there exists $B(\lambda)$ such that, $A(\lambda)B(\lambda) = I_n$.

By taking the determinant of both sides, we obtain $\det A(\lambda)\det B(\lambda) = 1$.

The product of the scalar polynomials $\det A(\lambda)$ and $\det B(\lambda)$ is a nonzero constant.

This is possible only if they are both nonzero constants.

Therefore $A(\lambda)$ is unimodular.

Conversely, if $A(\lambda)$ is unimodular then $\det A(\lambda) = const. \neq 0$.

The entries of the inverse matrix are equal to the cofactors of $A(\lambda)$ (*i.e.* the minors of $A(\lambda)$ of order $n-1$ multiplied by 1 or -1) divided by $const. \neq 0$.

Therefore the inverse matrix is a matrix polynomial in λ. So $B(\lambda) = [A(\lambda)]^{-1}$ is a matrix polynomial. *i.e.* $A(\lambda)$ is invertible. □

1.2 Addition and Multiplication of Matrix Polynomials

Definition 1.2.1: Let $A(\lambda) = A_l\lambda^l + A_{l-1}\lambda^{l-1} + \cdots + A_1\lambda + A_0$ and $B(\lambda) = B_m\lambda^m + B_{m-1}\lambda^{m-1} + \cdots + B_1\lambda + B_0$ be two matrix polynomials of the same order n, then

$$A(\lambda) + B(\lambda) := \begin{cases} \sum_{i=0}^{l}(A_i + B_i)\lambda^i + \sum_{i=l+1}^{m} B_i\lambda^i & \text{if } l < m \\ \sum_{i=0}^{l=m}(A_i + B_i)\lambda^i & \text{if } l = m \\ \sum_{i=0}^{m}(A_i + B_i)\lambda^i + \sum_{i=m+1}^{l} A_i\lambda^i & \text{if } l > m \end{cases}$$

The sum of two matrix polynomials of the same order n can be represented in the form of a matrix polynomial whose degree does not exceed the larger of the degrees of the given matrix polynomials. This fact coincides with the case of scalar polynomials. We mean by a scalar polynomial a polynomial with scalar coefficients not a polynomial as $p(\lambda) = a$.

Definition 1.2.2: Let $A(\lambda) = A_l\lambda^l + A_{l-1}\lambda^{l-1} + \cdots + A_1\lambda + A_0$ and $B(\lambda) = B_m\lambda^m + B_{m-1}\lambda^{m-1} + \cdots + B_1\lambda + B_0$ be two matrix polynomials of the same order n, then

$$A(\lambda)B(\lambda) := A_lB_m\lambda^{l+m} + (A_lB_{m-1} + A_{l-1}B_m)\lambda^{l+m-1} + \cdots + (A_1B_0 + A_0B_1)\lambda + A_0B_0.$$

The product of two matrix polynomials is a matrix polynomial whose degree is less than or equal to the sum of the degrees of the factors.

This fact differs from the case of scalar polynomials, since if $A_l \neq 0$ and $B_m \neq 0$ then A_lB_m, may be a zero matrix.

Example 1.2.1: If $A(\lambda) = \begin{bmatrix} \lambda^2 + \lambda & 2\lambda^2 \\ 2\lambda^2 & 4\lambda^2 + 1 \end{bmatrix}$ and $B(\lambda) = \begin{bmatrix} 4\lambda & -6\lambda \\ -2\lambda & 3\lambda \end{bmatrix}$, then

$$A(\lambda)B(\lambda) = \begin{bmatrix} 4\lambda^2 & -6\lambda^2 \\ -2\lambda & 3\lambda \end{bmatrix}.$$

We note that $\deg(A(\lambda)B(\lambda)) < \deg(A(\lambda) + B(\lambda))$.

Note that if at least one of the two leading coefficients A_l, B_m is invertible, then the degree of the product is always equal to the sum of the degrees of the factors.

Let $\deg A(\lambda) = l$ and $\deg B(\lambda) = m$, this implies that $A_l \neq 0$ and $B_m \neq 0$

Let B_m be invertible. Then if $A_l B_m = 0$ then $(A_l B_m)B_m^{-1} = 0B_m^{-1}$. That is $A_l = 0$, which contradicts the assumption. Therefore $A_l B_m \neq 0$ and $\deg(A(\lambda)B(\lambda)) = l + m$.

1.3 Division of Matrix Polynomials

Definitions 1.3.1: Let $A(\lambda) = \sum_{i=0}^{l} A_i \lambda^i$ and $B(\lambda) = \sum_{i=0}^{m} B_i \lambda^i$ be two $n \times n$ matrix polynomials of degree l and m respectively, and let B_m be invertible.

(i) We shall say that the matrix polynomials $Q(\lambda)$ and $R(\lambda)$ are the **right quotient** and the **right remainder** of $A(\lambda)$, respectively, on division by $B(\lambda)$ if

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

and if the degree of $R(\lambda)$ is less than that of $B(\lambda)$ or $R(\lambda) \equiv 0$.

(ii) Similarly, we shall call the matrix polynomials $\widetilde{Q}(\lambda)$ and $\widetilde{R}(\lambda)$ the **left quotient** and the **left remainder** of $A(\lambda)$ on division by $B(\lambda)$ if

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

and if the degree of $\widetilde{R}(\lambda)$ is less than that of $B(\lambda)$ or $\widetilde{R}(\lambda) \equiv 0$.

Theorem 1.3.1: Let $A(\lambda)=\sum_{i=0}^{l} A_i\lambda^i$ and $B(\lambda)=\sum_{i=0}^{m} B_i\lambda^i$ be $n\times n$ matrix polynomials of degree l and m, respectively, with $\det B_m \neq 0$. Then

(i) there exists a right quotient and right remainder of $A(\lambda)$ on division by $B(\lambda)$, and

(ii) there exists a left quotient and left remainder of $A(\lambda)$ on division by $B(\lambda)$.

Proof:

(i) If $l<m$, then $A(\lambda)=0B(\lambda)+A(\lambda)$ *i.e.* $Q(\lambda)=0$ and $R(\lambda)=A(\lambda)$.

If $l\geq m$. Then,

$$A_l B_m^{-1}\lambda^{l-m}B(\lambda)=A_l B_m^{-1}\lambda^{l-m}[B_m\lambda^m+B_{m-1}\lambda^{m-1}+\cdots+B_0]$$

$$A_l B_m^{-1}\lambda^{l-m}B(\lambda)=A_l\lambda^l+A_l B_m^{-1}B_{m-1}\lambda^{l-1}+\cdots+A_l B_m^{-1}B_0\lambda^{l-m}.$$

The term of highest degree of the matrix polynomial $A_1 B_m^{-1}\lambda^{l-m}B(\lambda)$ is $A_l\lambda^l$, from the previous equation, we get

$$A_l\lambda^l=A_l B_m^{-1}\lambda^{l-m}B(\lambda)-[A_l B_m^{-1}B_{m-1}\lambda^{l-1}+\cdots+A_l B_m^{-1}B_0\lambda^{l-m}]$$

$$A(\lambda)=A_l\lambda^l+A_{l-1}\lambda^{l-1}+\cdots+A_0$$

$$=A_l B_m^{-1}\lambda^{l-m}B(\lambda)-[A_l B_m^{-1}B_{m-1}\lambda^{l-1}+\cdots+A_l B_m^{-1}B_0\lambda^{l-m}]+[A_{l-1}\lambda^{l-1}+\cdots+A_0]$$

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

where

$$A^{(1)}(\lambda)=[A_{l-1}\lambda^{l-1}+\cdots+A_0]-[A_l B_m^{-1}B_{m-1}\lambda^{l-1}+\cdots+A_l B_m^{-1}B_0\lambda^{l-m}]$$

$A^{(1)}(\lambda)$ is a matrix polynomial of degree $=l_1\leq l-1$.

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

If $l_1\geq m$, repeat the process on $A^{(1)}(\lambda)$, to get

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

$$A^{(2)}(\lambda) = A^{(2)}_{l_2} \lambda^{l_2} + A^{(2)}_{l_2 - 1} \lambda^{l_2 - 1} + \cdots + A^{(2)}_0, \ A^{(2)}_{l_2} \neq 0, \ l_2 < l_1.$$

In this manner we can construct a sequence of matrix polynomials

$A(\lambda), \quad A^{(1)}(\lambda), \quad A^{(2)}(\lambda), \quad \cdots$, whose degrees are strictly decreasing, and after a finite number of times we arrive at a matrix polynomial $A^{(r)}(\lambda)$ of degree $l_r < m$ with $l_{r-1} \geq m$.

Write $A(\lambda) =: A^{(0)}(\lambda)$

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

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

The matrix in the parenthesis is a right quotient of $A(\lambda)$ on division by $B(\lambda)$ and $A^{(r)}(\lambda)$ is right remainder.

(ii) $A^T(\lambda) = A_l^T \lambda^l + A_{l-1}^T \lambda^{l-1} + \cdots + A_1^T \lambda + A_0^T$,

$B^T(\lambda) = B_m^T \lambda^m + B_{m-1}^T \lambda^{m-1} + \cdots + B_1^T \lambda + B_0^T$.

Since $\det B_m^T = \det B_m \neq 0$ then by part (i): $A^T(\lambda) = Q_1(\lambda) B^T(\lambda) + R_1(\lambda)$.

Take the transpose of both sides, we obtain

$A(\lambda) = B(\lambda) Q_1^T(\lambda) + R_1^T(\lambda)$

Let $\widetilde{Q}(\lambda) = Q_1^T(\lambda)$ and $\widetilde{R}(\lambda) = R_1^T(\lambda)$

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

Theorem 1.3.2: Let $A(\lambda)=\sum_{i=0}^{l} A_i\lambda^i$, $B(\lambda)=\sum_{i=0}^{m} B_i\lambda^i$ be $n\times n$ matrix polynomials of degree l and m, respectively, with $\det B_m \neq 0$. Then the right quotient, right remainder, left quotient, and left remainder are unique.

Proof:

(i) Right quotient and right remainder are unique.

Suppose that there exist matrix polynomials $Q(\lambda), R(\lambda)$ and $Q_1(\lambda), R_1(\lambda)$ such that

$A(\lambda)=Q(\lambda)B(\lambda)+R(\lambda)$, $A(\lambda)=Q_1(\lambda)B(\lambda)+R_1(\lambda)$,

where $\deg R(\lambda)<m$ and $\deg R_1(\lambda)<m$. Then

$Q(\lambda)B(\lambda)+R(\lambda)=Q_1(\lambda)B(\lambda)+R_1(\lambda)$,

$(Q(\lambda)-Q_1(\lambda))B(\lambda)=R_1(\lambda)-R(\lambda)$.

If $Q(\lambda)\neq Q_1(\lambda)$ then $\deg(Q(\lambda)-Q_1(\lambda))B(\lambda)\geq m$,

but $\deg(R_1(\lambda)-R(\lambda))<m$, this is a contradiction,

then $Q(\lambda)=Q_1(\lambda)$ and $R(\lambda)=R_1(\lambda)$.

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

Example 1.3.1: Find the right quotient and remainder of

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

on division by

$$B(\lambda)=\begin{bmatrix} 2\lambda^2+3 & -\lambda^2+1 \\ -\lambda^2-1 & \lambda^2+2 \end{bmatrix}$$

Solution:

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

$$B(\lambda)=\begin{bmatrix}2 & -1\\ -1 & 1\end{bmatrix}\lambda^2+\begin{bmatrix}3 & 1\\ -1 & 2\end{bmatrix}$$

$$\det B_m=\det\begin{bmatrix}2 & -1\\ -1 & 1\end{bmatrix}=1\neq 0$$

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

$$=\begin{bmatrix}\lambda^3+\lambda & 2\lambda^3+\lambda^2\\ -\lambda^3-2\lambda^2+1 & 3\lambda^3+\lambda\end{bmatrix}-\begin{bmatrix}1 & 2\\ -1 & 3\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 2\end{bmatrix}\lambda\begin{bmatrix}2\lambda^2+3 & -\lambda^2+1\\ -\lambda^2-1 & \lambda^2+2\end{bmatrix}$$

$$=\begin{bmatrix}-3\lambda & \lambda^2-13\lambda\\ -2\lambda^2-\lambda+1 & -11\lambda\end{bmatrix}$$

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

$$=\begin{bmatrix}-3\lambda & \lambda^2-13\lambda\\ -2\lambda^2-\lambda+1 & -11\lambda\end{bmatrix}-\begin{bmatrix}0 & 1\\ -2 & 0\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 2\end{bmatrix}\begin{bmatrix}2\lambda^2+3 & -\lambda^2+1\\ -\lambda^2-1 & \lambda^2+2\end{bmatrix}$$

$$=\begin{bmatrix}-3\lambda-1 & -13\lambda-5\\ -\lambda+5 & -11\lambda+6\end{bmatrix}.$$

Right quotient $Q(\lambda)=A_l B_m^{-1}\lambda^{l-m}+A_{l_1}^{(1)}B_m^{-1}\lambda^{l_1-m}$

$$Q(\lambda)=\begin{bmatrix}1 & 2\\ -1 & 3\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 2\end{bmatrix}\lambda+\begin{bmatrix}0 & 1\\ -2 & 0\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 2\end{bmatrix}=\begin{bmatrix}3\lambda+1 & 5\lambda+2\\ 2\lambda-2 & 5\lambda-2\end{bmatrix}$$

Right remainder $R(\lambda)=\begin{bmatrix}-3\lambda-1 & -13\lambda-5\\ -\lambda+5 & -11\lambda+6\end{bmatrix}.$

Example 1.3.2: Find the left quotient and remainder of

$$A(\lambda)=\begin{bmatrix}\lambda^4+\lambda^2+\lambda-1 & \lambda^3+\lambda^2+\lambda+2\\ 2\lambda^3-\lambda & 2\lambda^2+2\lambda\end{bmatrix}$$

on division by

$$B(\lambda)=\begin{bmatrix}\lambda^2+1 & 1\\ \lambda & \lambda^2+\lambda\end{bmatrix}$$

Solution:

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

$$=\begin{bmatrix}\lambda^4+\lambda^2+\lambda-1 & \lambda^3+\lambda^2+\lambda+2\\ 2\lambda^3-\lambda & 2\lambda^2+2\lambda\end{bmatrix}-\begin{bmatrix}\lambda^2+1 & 1\\ \lambda & \lambda^2+\lambda\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\lambda^2$$

$$=\begin{bmatrix}\lambda-1 & \lambda^3+\lambda^2+\lambda+2\\ \lambda^3-\lambda & 2\lambda^2+2\lambda\end{bmatrix}$$

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

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

$$=\begin{bmatrix}-1 & \lambda^2+2\\ -\lambda^2-\lambda & \lambda^2+2\lambda\end{bmatrix}$$

$$A^{(3)}(\lambda)=A^{(2)}(\lambda)-B(\lambda)B_m^{-1}A_{l_2}^{(2)}\lambda^{l_2-m}$$

$$=\begin{bmatrix}-1 & \lambda^2+2\\ -\lambda^2-\lambda & \lambda^2+2\lambda\end{bmatrix}-\begin{bmatrix}\lambda^2+1 & 1\\ \lambda & \lambda^2+\lambda\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}0 & 1\\ -1 & 1\end{bmatrix}=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$$

Left quotient $\widetilde{Q}(\lambda)=B_m^{-1}A_l\lambda^{l-m}+B_m^{-1}A_{l_1}^{(1)}\lambda^{l_1-m}+B_m^{-1}A_{l_2}^{(2)}\lambda^{l_2-m}$

$$\widetilde{Q}(\lambda)=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\lambda^2+\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\lambda+\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}0 & 1\\ -1 & 1\end{bmatrix}=\begin{bmatrix}\lambda^2 & \lambda+1\\ \lambda-1 & 1\end{bmatrix}$$

Left remainder $\widetilde{R}(\lambda)=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$,

Therefore $A(\lambda)=B(\lambda)\widetilde{Q}(\lambda)$.

Example 1.3.3: Find the right quotient and remainder of

$$A(\lambda)=\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}$$

on division by

$$B(\lambda)=\begin{bmatrix}3 & 1\\ -1 & 2\end{bmatrix}.$$

Solution:

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

$$=\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}-\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}\begin{bmatrix}\frac{2}{7} & \frac{-1}{7}\\ \frac{1}{7} & \frac{3}{7}\end{bmatrix}\begin{bmatrix}3 & 1\\ -1 & 2\end{bmatrix}=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$$

Right quotient $Q(\lambda)=\begin{bmatrix}0 & 0\\ \frac{2}{7} & \frac{-1}{7}\end{bmatrix}$

Right remainder $R(\lambda)=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$.

Definition 1.3.2: If $A(\lambda)$ is a matrix polynomial and $B\in\mathbf{C}^{n\times n}$, we defining the right value $A(B)$ of $A(\lambda)$ at B by

$$A(B)=A_l B^l + A_{l-1}B^{l-1}+\cdots+A_0$$

and the left value $\widetilde{A}(B)$ of $A(\lambda)$ at B by

$$\widetilde{A}(B) = B^l A_l + B^{l-1} A_{l-1} + \cdots + A_0$$

Theorem 1.3.3: The right and left remainders of a matrix polynomial $A(\lambda)$ on division by $\lambda I - B$ are $A(B)$ and $\widetilde{A}(B)$, respectively.

Proof:

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

Premultiply both sides of this equation by A_j and sum the resulting equations for $j = 1, 2, \cdots, l$. Then

$\sum_{j=1}^{l} A_j \lambda^j - \sum_{j=1}^{l} A_j B^j = C(\lambda)(\lambda I - B)$

where $C(\lambda) = A_1(\lambda^0) + A_2(\lambda I + \lambda^0 B) + \cdots + A_l(\lambda^{l-1} I + \lambda^{l-2} B + \cdots + \lambda B^{l-2} + B^{l-1})$ is a matrix polynomial.

$\sum_{j=1}^{l} A_j \lambda^j - \sum_{j=1}^{l} A_j B^j = \sum_{j=0}^{l} A_j \lambda^j - \sum_{j=0}^{l} A_j B^j$

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

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

Since the right remainder is unique, this implies that $A(B)$ is the right remainder on division of $A(\lambda)$ by $(\lambda I - B)$.

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

Postmultiply both sides of this equation by A_j and sum the resulting equations for $j = 1, 2, \cdots, l$.

Then $\sum_{j=1}^{l} A_j\lambda^j - \sum_{j=1}^{l} B^j A_j = (\lambda I - B)C(\lambda)$ where $C(\lambda)$ is a matrix polynomial.

$$\sum_{j=0}^{l} A_j\lambda^j - \sum_{j=0}^{l} B^j A_j = (\lambda I - B)C(\lambda)$$

$$A(\lambda) - \widetilde{A}(B) = (\lambda I - B)C(\lambda)$$

then $A(\lambda) = (\lambda I - B)C(\lambda) + \widetilde{A}(B)$.

Since the left remainder is unique,

this implies that $\widetilde{A}(B)$ is the left remainder on division of $A(\lambda)$ by $(\lambda I - B)$. □

We note that:

(*i*) Because of the noncommutativity, we have to distinguish between representation

$$A(\lambda) = Q(\lambda)B(\lambda) + R(\lambda) \quad (1.3.1)$$

(which will be referred to as right division) and the representation

$$A(\lambda) = B(\lambda)\widetilde{Q}(\lambda) + \widetilde{R}(\lambda) \quad (1.3.2)$$

for some matrix polynomials $\widetilde{Q}(\lambda)$ and $\widetilde{R}(\lambda)$, where the degree of $R(\lambda)$ is less than the degree of $B(\lambda)$, or is the zero matrix polynomial. Representation (1.3.2) will be referred to as left division. In general, $Q(\lambda) \neq \widetilde{Q}(\lambda)$ and $R(\lambda) \neq \widetilde{R}(\lambda)$, so we distinguish between right $Q(\lambda)$ and left $\widetilde{Q}(\lambda)$ quotients and between right $R(\lambda)$ and left $\widetilde{R}(\lambda)$ remainders.

(*ii*) The division is not always possible. The simplest example of this situation appears if we take $A(\lambda) = A_0$ as a constant nonsingular matrix and $B(\lambda) = B_0$ as a constant nonzero singular matrix. If the division is possible, then (since the degree of

$B(\lambda)$ is 0) the remainder $R(\lambda)$ and $\widetilde{R}(\lambda)$ must be zeros, and the (1.3.1) and (1.3.2) take the form

$$A_0 = Q(\lambda)B_0,\ A_0 = B_0\widetilde{Q}(\lambda) \tag{1.3.3}$$

respectively. But in view of the invertibility of A_0, neither of (1.3.3) can be satisfied for any $Q(\lambda)$ or $\widetilde{Q}(\lambda)$, so the division is impossible.

CHAPTER TWO

INVARIANT POLYNOMIALS

In this chapter, we shall introduce the concept of invariant polynomials and elementary divisors of matrix polynomials.

In order to do this it is necessary to introduce and examine the elementary row (column) operations on matrix polynomials.

Also, we shall define the generally invertible matrix polynomial, the generalized inverse of the matrix polynomial and the one-sided invertible matrix polynomial.

2.1 Elementary Operations and Canonical Form

This section is devoted to the elementary row (column) operations on matrix polynomials, and using these operations to transform the matrix polynomial to its canonical form.

Definition 2.1.1: An **elementary row (column) operation** on an $n \times n$ matrix polynomial $A(\lambda) = \left[a_{ij}(\lambda)\right]_{i,j=1}^{n}$ is any one of the following operations:

(i) Interchange rows (columns) i and j of $A(\lambda)$.

(ii) Multiply row (column) i of $A(\lambda)$ by $c \neq 0$, where $c \in \mathbf{C}$.

(iii) Add $b(\lambda)$ times row (column) i of $A(\lambda)$ to row (column) j of $A(\lambda)$, $i < j$.

The elementary row operations (i), (ii), and (iii) are equivalent to a multiplication of the matrix polynomial $A(\lambda)$ on the left by the following square matrices of order n:

$$E_1 = \begin{bmatrix} 1 &&&&&&& \\ & \ddots &&&&&& \\ && 0 & \cdots & & \cdots & 1 & \\ && \vdots & 1 & & & \vdots & \\ &&&& \ddots &&& \\ && \vdots & & & 1 & \vdots & \\ && 1 & \cdots & & \cdots & 0 & \\ &&&&&&& \ddots & \\ &&&&&&&& 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 &&&&& \\ & \ddots &&&& \\ && 1 &&& \\ &&& c && \\ &&&& 1 & \\ &&&&& \ddots & \\ &&&&&& 1 \end{bmatrix},$$

and

$$E_r(\lambda) = \begin{bmatrix} 1 &&&&& \\ & \ddots &&&& \\ && 1 &&& \\ && \vdots & \ddots && \\ && b(\lambda) & \cdots & 1 & \\ &&&&& \ddots & \\ &&&&&& 1 \end{bmatrix}.$$

While the elementary column operations (i), (ii), and (iii) are equivalent to a multiplication of the matrix polynomial $A(\lambda)$ on the right by the following square matrices of order n, respectively; E_1, E_2, and

$$E_c(\lambda) = \begin{bmatrix} 1 &&&&& \\ & \ddots &&&& \\ && 1 & \cdots & b(\lambda) & \\ &&& \ddots & \vdots & \\ &&&& 1 & \\ &&&&& \ddots & \\ &&&&&& 1 \end{bmatrix}.$$

Definition 2.1.2: A **canonical matrix polynomial** is a diagonal matrix polynomial of the form

$$A_c(\lambda) = diag[1,1,\cdots,1,a_1(\lambda),a_2(\lambda),\cdots,a_k(\lambda),0,0,\cdots,0],$$

where each $a_j(\lambda)$ is a monic scalar polynomial of degree greater than or equal to one and is divisible by $a_{j-1}(\lambda), \quad j = 2,3,\cdots,k$.

Theorem 2.1.1: Any matrix polynomial $A(\lambda)$ over **C** of order n is equivalent to a canonical matrix polynomial $A_c(\lambda)$.

Proof: See [11] page 257. □

The following example explains how to transform the matrix polynomial $A(\lambda)$ to its equivalent canonical matrix polynomial $A_c(\lambda)$.

Example 2.1.1: Find the canonical matrix polynomial $A_c(\lambda)$ of the matrix polynomial $A(\lambda) = \begin{bmatrix} \lambda & \lambda^2 & 0 \\ \lambda^3 & \lambda^5 & 0 \\ 0 & 0 & 2\lambda \end{bmatrix}$.

Solution:

$$\begin{bmatrix} \lambda & \lambda^2 & 0 \\ \lambda^3 & \lambda^5 & 0 \\ 0 & 0 & 2\lambda \end{bmatrix} \xrightarrow{-\lambda^2 r_1 + r_2 \to r_2} \begin{bmatrix} \lambda & \lambda^2 & 0 \\ 0 & \lambda^5 - \lambda^4 & 0 \\ 0 & 0 & 2\lambda \end{bmatrix} \xrightarrow{c_2 \leftrightarrow c_3} \begin{bmatrix} \lambda & 0 & \lambda^2 \\ 0 & 0 & \lambda^5 - \lambda^4 \\ 0 & 2\lambda & 0 \end{bmatrix}$$

$$\xrightarrow{-\lambda c_1 + c_3 \to c_3} \begin{bmatrix} \lambda & 0 & 0 \\ 0 & 0 & \lambda^5 - \lambda^4 \\ 0 & 2\lambda & 0 \end{bmatrix} \xrightarrow{r_2 \leftrightarrow r_3} \begin{bmatrix} \lambda & 0 & 0 \\ 0 & 2\lambda & 0 \\ 0 & 0 & \lambda^5 - \lambda^4 \end{bmatrix}$$

$$\xrightarrow{\frac{1}{2}r_2 \to r_2} \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda^5 - \lambda^4 \end{bmatrix}$$

Therefore $A_c(\lambda) = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda^5 - \lambda^4 \end{bmatrix}$.

2.2 Invariant Polynomials

In the previous section, the matrix polynomial was transformed to its canonical form by elementary row (column) operations, but now we want to develop another method to transform the matrix polynomial to its canonical form, which is dependent entirely on the minors.

Definition 2.2.1: The **rank** of an $n \times n$ matrix polynomial $A(\lambda)$ is equal to the order of its largest nonzero minor.

The rank of $diag[a_1(\lambda), a_2(\lambda), \cdots, a_r(\lambda), 0, 0, \cdots, 0]$ is equal to r.

Example 2.2.1: The rank of $\begin{bmatrix} \lambda & \lambda+1 \\ \lambda^2 - \lambda & \lambda^2 - 1 \end{bmatrix}$ is equal to 1.

Definition 2.2.2: Two matrix polynomials $A(\lambda)$ and $B(\lambda)$, are said to be **equivalent**, if $B(\lambda)$ can be obtained from $A(\lambda)$ by a finite sequence of elementary operations, or equivalently,

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

where $P(\lambda)$ and $Q(\lambda)$ are matrix polynomials with nonzero constant determinants.

Theorem 2.2.1: Equivalent matrix polynomials have the same rank.

Proof:

Let $A(\lambda)$ be equivalent to $B(\lambda)$, then there are unimodular matrices $P(\lambda)$ and $Q(\lambda)$ such that

$$B(\lambda) = P(\lambda)A(\lambda)Q(\lambda). \tag{2.2.1}$$

By applying the Binet-Cauchy formula[1] twice we obtain:

$$b(\lambda) = \sum_s p_s(\lambda)a_s(\lambda)q_s(\lambda) \tag{2.2.2}$$

where $b(\lambda)$ is a minor of order j of $B(\lambda)$, $a_s(\lambda)$ are minors of order j of $A(\lambda)$, $p_s(\lambda)$ are minors of order j of $P(\lambda)$ and $q_s(\lambda)$ are minors of order j of $Q(\lambda)$.

If $b(\lambda)$ is a nonzero minor of $B(\lambda)$ of the greatest order r, that is, the rank of $B(\lambda)$ is r, then it follows from equation (2.2.2) that at least one minor $a_s(\lambda)$ (of order r) is a nonzero scalar polynomial. Then

$$\text{rank } B(\lambda) \leq \text{rank } A(\lambda) \tag{2.2.3}$$

By multiplying equation (2.2.1) from the left by $[P(\lambda)]^{-1}$ and from the right by $[Q(\lambda)]^{-1}$ then $A(\lambda) = [P(\lambda)]^{-1}B(\lambda)[Q(\lambda)]^{-1}$. Thus

$$\text{rank } A(\lambda) \leq \text{rank } B(\lambda) \tag{2.2.4}$$

[1] **Binet-Cauchy formula:** Let A and B be $m \times n$ and $n \times m$ matrices, respectively. If $m \leq n$ and $C = AB$, then

$$\det C = \sum_j A\begin{pmatrix} 1 & 2 & \cdots & m \\ j_1 & j_2 & \cdots & j_m \end{pmatrix} B\begin{pmatrix} j_1 & j_2 & \cdots & j_m \\ 1 & 2 & \cdots & m \end{pmatrix}$$

Equation (2.2.3) and equation (2.2.4) imply that rank $A(\lambda) =$ rank $B(\lambda)$.

That is, the equivalent matrix polynomials have the same rank. □

Let $A(\lambda)$ be a matrix polynomial of rank r. We denote by $d_j(\lambda)$ the monic greatest common divisor of all the minors of order j in $A(\lambda)$, where, $j = 1,2,\cdots,r$.

Since any minor of order j can be written as a linear combination of minors of order $j-1$, $j = 2,3,\cdots,r$ by cofactor expansion, then each of

$$d_r(\lambda), d_{r-1}(\lambda), \cdots, d_1(\lambda), d_0(\lambda) = 1$$

is divisible by the succeeding one.

Definition 2.2.3: The following scalar polynomials

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

are called the **invariant polynomials** of the matrix polynomial $A(\lambda)$.

Example 2.2.2: Find the invariant polynomials of the matrix polynomial

$$A(\lambda) = \begin{bmatrix} \lambda & \lambda^2 & 0 \\ \lambda^3 & \lambda^5 & 0 \\ 0 & 0 & 2\lambda \end{bmatrix}.$$

Solution:

We can easily find that

$d_0(\lambda) = 1, d_1(\lambda) = \lambda, d_2(\lambda) = \lambda^2$ and $d_3(\lambda) = \lambda^7 - \lambda^6$.

Then

$$i_1(\lambda) = \frac{d_1(\lambda)}{d_0(\lambda)} = \frac{\lambda}{1} = \lambda$$

$$i_2(\lambda) = \frac{d_2(\lambda)}{d_1(\lambda)} = \frac{\lambda^2}{\lambda} = \lambda$$

$$i_3(\lambda) = \frac{d_3(\lambda)}{d_2(\lambda)} = \frac{\lambda^7 - \lambda^6}{\lambda^2} = \lambda^5 - \lambda^4 .$$

Example 2.2.3: Find the invariant polynomials of the canonical matrix polynomial $A_c(\lambda) = diag[a_1(\lambda), a_2(\lambda), \cdots, a_r(\lambda), 0, 0, \cdots, 0]$.

Solution:

Let $d_0(\lambda) = 1, d_1(\lambda) = a_1(\lambda), d_2(\lambda) = a_1(\lambda)a_2(\lambda), \cdots, d_r(\lambda) = a_1(\lambda)a_2(\lambda)\cdots a_r(\lambda)$.

Then

$$i_1(\lambda) = \frac{d_1(\lambda)}{d_0(\lambda)} = \frac{a_1(\lambda)}{1} = a_1(\lambda)$$

$$i_2(\lambda) = \frac{d_2(\lambda)}{d_1(\lambda)} = \frac{a_1(\lambda)a_2(\lambda)}{a_1(\lambda)} = a_2(\lambda)$$

$$\vdots$$

$$i_r(\lambda) = \frac{d_r(\lambda)}{d_{r-1}(\lambda)} = \frac{a_1(\lambda)a_2(\lambda)\cdots a_r(\lambda)}{a_1(\lambda)a_2(\lambda)\cdots a_{r-1}(\lambda)} = a_r(\lambda) .$$

Theorem 2.2.2: Let the matrix polynomials $A(\lambda)$ and $B(\lambda)$ be equivalent. Then the monic greatest common divisor $d_j(\lambda)$ of all minors of order j of $A(\lambda)$ and the monic greatest common divisor $\Delta_j(\lambda)$ of all minors of order j of $B(\lambda)$ are equal, $(j = 1, 2, \cdots, r)$.

Proof:

From equation (2.2.2) in the proof of theorem 2.2.1, we obtain that any common divisor of minors $a_s(\lambda)$ of $A(\lambda)$ of order j, $(1 \le j \le r)$ is a divisor of $b(\lambda)$. Then $\Delta_j(\lambda)$ is divisible by $d_j(\lambda)$.

Also from equation (2.2.4) in the proof of theorem 2.2.1, any common divisor of minors $b_s(\lambda)$ of $B(\lambda)$ of order j $(1 \le j \le r)$ is a divisor of $a(\lambda)$. Then $d_j(\lambda)$ is divisible by $\Delta_j(\lambda)$.

Since $d_j(\lambda)$ and $\Delta_j(\lambda)$ are monic, then $d_j(\lambda) = \Delta_j(\lambda)$, $j = 1,2,\cdots,r$. □

Corollary 2.2.1: Let $d_j(\lambda)$ be the monic greatest common divisor of all minors of order j of the matrix polynomial $A(\lambda)$ of rank r, $j = 1,2,\cdots,r$ then

$$d_j(\lambda) = i_1(\lambda) i_2(\lambda) \cdots i_j(\lambda),$$

where $i_1(\lambda)$, $i_2(\lambda)$,… , $i_j(\lambda)$ are the invariant polynomials of $A(\lambda)$.

Proof:

$i_j(\lambda) = \dfrac{d_j(\lambda)}{d_{j-1}(\lambda)}$ then $d_j(\lambda) = d_{j-1}(\lambda) i_j(\lambda)$

$i_{j-1}(\lambda) = \dfrac{d_{j-1}(\lambda)}{d_{j-2}(\lambda)}$ then $d_{j-1}(\lambda) = d_{j-2}(\lambda) i_{j-1}(\lambda)$.

From the previous equations we obtain:

$d_j(\lambda) = d_{j-2}(\lambda) i_{j-1}(\lambda) i_j(\lambda)$.

$i_{j-2}(\lambda) = \dfrac{d_{j-2}(\lambda)}{d_{j-3}(\lambda)}$ then $d_{j-2}(\lambda) = d_{j-3}(\lambda) i_{j-2}(\lambda)$.

Hence

$d_j(\lambda) = d_{j-3}(\lambda) i_{j-2}(\lambda) i_{j-1}(\lambda) i_j(\lambda)$.

We proceed in the same manner to obtain:

$d_j(\lambda) = d_0(\lambda) i_1(\lambda) i_2(\lambda) \cdots i_j(\lambda) = i_1(\lambda) i_2(\lambda) \cdots i_j(\lambda)$. □

Corollary 2.2.2: Let $i_j(\lambda)$ $(j = 2,3,\cdots,r)$ be the invariant polynomials of the matrix polynomial $A(\lambda)$ of rank r then $i_j(\lambda)$ is divisible by $i_{j-1}(\lambda)$.

Proof:

By theorem 2.1.1, $A(\lambda)$ and $A_c(\lambda)$ are equivalent, where $A_c(\lambda)$ is the canonical matrix polynomial of $A(\lambda)$.

Let $d_j(\lambda)$ be the monic greatest common divisor of all minors of order j of $A(\lambda)$ and $\Delta_j(\lambda)$ be the monic greatest common divisor of all minors of order j of $A_c(\lambda)$.

Then $d_j(\lambda) = \Delta_j(\lambda)$ $(j = 1,2,\cdots,r)$ by theorem 2.2.2. Also

rank $A(\lambda) =$ rank $A_c(\lambda) = r$ by theorem 2.2.1.

$$i_j(\lambda) = \frac{d_j(\lambda)}{d_{j-1}(\lambda)} = \frac{\Delta_j(\lambda)}{\Delta_{j-1}(\lambda)} = \frac{a_1(\lambda)a_2(\lambda)\cdots a_j(\lambda)}{a_1(\lambda)a_2(\lambda)\cdots a_{j-1}(\lambda)} = a_j(\lambda)$$

$$i_{j-1}(\lambda) = \frac{d_{j-1}(\lambda)}{d_{j-2}(\lambda)} = \frac{\Delta_{j-1}(\lambda)}{\Delta_{j-2}(\lambda)} = \frac{a_1(\lambda)a_2(\lambda)\cdots a_{j-1}(\lambda)}{a_1(\lambda)a_2(\lambda)\cdots a_{j-2}(\lambda)} = a_{j-1}(\lambda)$$

since $a_j(\lambda)$ is divisible by $a_{j-1}(\lambda)$ then $i_j(\lambda)$ is divisible by $i_{j-1}(\lambda)$. □

Theorem 2.2.3 (Smith Canonical Form): The matrix polynomial $A(\lambda)$ of rank r is always equivalent to the canonical matrix polynomial $diag[i_1(\lambda), i_2(\lambda), \cdots, i_r(\lambda), 0, 0, \cdots, 0]$, where $i_1(\lambda)$, $i_2(\lambda)$, $\cdots$, $i_r(\lambda)$ are the invariant polynomials of $A(\lambda)$.

Proof:

Let $A_c(\lambda) = diag[a_1(\lambda), a_2(\lambda), \cdots, a_r(\lambda), 0, 0, \cdots, 0]$ be the canonical form of the matrix polynomial $A(\lambda)$.

Then by theorem 2.2.2 and example 2.2.3

$d_j(\lambda) = a_1(\lambda)a_2(\lambda)\cdots a_j(\lambda)$.

$$i_j(\lambda) = \frac{d_j(\lambda)}{d_{j-1}(\lambda)} = \frac{a_1(\lambda)a_2(\lambda)\cdots a_j(\lambda)}{a_1(\lambda)a_2(\lambda)\cdots a_{j-1}(\lambda)} = a_j(\lambda), \quad j = 1, 2, \cdots, r$$

Then $A(\lambda)$ is equivalent to $diag[i_1(\lambda), i_2(\lambda), \cdots, i_r(\lambda), 0, 0, \cdots, 0]$. □

Example 2.2.4: Find the Smith canonical form for $L(\lambda)$, where

$$L(\lambda) = \begin{bmatrix} \lambda(\lambda-1) & 1 \\ 0 & \lambda(\lambda-1) \end{bmatrix}.$$

Solution:

$d_0(\lambda) = 1, d_1(\lambda) = 1$ and $d_2(\lambda) = \lambda^2(\lambda-1)^2$

Then

$$i_1(\lambda) = \frac{d_1(\lambda)}{d_0(\lambda)} = \frac{1}{1} = 1$$

$$i_2(\lambda) = \frac{d_2(\lambda)}{d_1(\lambda)} = \frac{\lambda^2(\lambda-1)^2}{1} = \lambda^2(\lambda-1)^2$$

Therefore the Smith canonical form is $diag[1, \lambda^2(\lambda-1)^2]$.

Corollary 2.2.3: The matrix polynomials $A(\lambda)$ and $B(\lambda)$ are equivalent if and only if they have the same invariant polynomials.

Proof:

Let $A(\lambda)$ and $B(\lambda)$ are equivalent, then $d_j(\lambda) = \Delta_j(\lambda)$, $j = 1,2,\cdots,r$, where $d_j(\lambda)$ is the greatest common divisor of all minors of order j of $A(\lambda)$ and $\Delta_j(\lambda)$ is the greatest common divisor of all minors of order j of $B(\lambda)$.

$$i_j(\lambda) = \frac{d_j(\lambda)}{d_{j-1}(\lambda)} = \frac{\Delta_j(\lambda)}{\Delta_{j-1}(\lambda)} = k_j(\lambda)$$

where $i_j(\lambda)$ is the invariant polynomial of $A(\lambda)$ and $k_j(\lambda)$ is the invariant polynomial of $B(\lambda)$.

Conversely, suppose that $A(\lambda)$ and $B(\lambda)$ have the same invariant polynomials $i_1(\lambda)$, $i_2(\lambda)$, ..., $i_r(\lambda)$, then $A(\lambda)$ and $diag[i_1(\lambda), i_2(\lambda), \cdots, i_r(\lambda), 0, 0, \cdots, 0]$ are equivalent, and also $B(\lambda)$ and $diag[i_1(\lambda), i_2(\lambda), \cdots, i_r(\lambda), 0, 0, \cdots, 0]$ are equivalent. This implies that $A(\lambda)$ and $B(\lambda)$ are equivalent. □

Corollary 2.2.4: If $A \in \mathbf{C}^{n \times n}$. Then the sum of the degrees of the invariant polynomials of $\lambda I - A$ is n.

Proof:

The determinant of the matrix polynomial $\lambda I - A$ is a scalar polynomial with degree n, and the coefficient of λ^n in this scalar polynomial is $\det I = 1 \neq 0$.

Then $\lambda I - A$ is regular, and this implies that the rank of $\lambda I - A$ is equal to n.

By Smith canonical form $\lambda I - A$ is equivalent to $diag[i_1(\lambda), i_2(\lambda), \cdots, i_n(\lambda)]$.

Thus,

there are unimodular matrix polynomials $P(\lambda)$ and $Q(\lambda)$ such that

$P(\lambda)(\lambda I - A)Q(\lambda) = diag[i_1(\lambda), i_2(\lambda), \cdots, i_n(\lambda)]$

$\det P(\lambda)(\lambda I - A)Q(\lambda) = \det diag[i_1(\lambda), i_2(\lambda), \cdots, i_n(\lambda)]$

$c_1 \det(\lambda I - A)c_2 = i_1(\lambda)i_2(\lambda)\cdots i_n(\lambda)$

since $\deg(c_1 \det(\lambda I - A)c_2) = n$, then $\deg(i_1(\lambda)i_2(\lambda)\cdots i_n(\lambda)) = n$.

That is, the sum of the degrees of the invariant polynomials of $\lambda I - A$ is n. □

2.3 Elementary Divisors

In this section we want to define the elementary divisors of a matrix polynomial. Also, we show how to find the elementary divisors from the invariant polynomials and conversely how to find the invariant polynomials from the elementary divisors if the rank of the matrix polynomial is known.

Theorem 2.3.1: If $A, B \in \mathbf{C}^{n\times n}$. Then A and B are similar if and only if $\lambda I - A$ and $\lambda I - B$ are equivalent.

Proof:

Suppose that A and B are similar, then there is a nonsingular matrix S such that $A = SBS^{-1}$, and this implies that $\lambda I - A = S(\lambda I - B)S^{-1}$, this equality proves the equivalence of $\lambda I - A$ and $\lambda I - B$.

Conversely, suppose that $\lambda I - A$ and $\lambda I - B$ are equivalent. Then there are unimodular matrix polynomials $E(\lambda)$ and $F(\lambda)$ such that

$$E(\lambda)(I\lambda - A)F(\lambda) = I\lambda - B. \qquad (2.3.1)$$

Consider the division of $(E(\lambda))^{-1}$ on the left by $I\lambda - A$ and the division of $F(\lambda)$ on the right by $I\lambda - B$, then we obtain

$$(E(\lambda))^{-1} = (I\lambda - A)S(\lambda) + E_0, \tag{2.3.2}$$

$$F(\lambda) = T(\lambda)(I\lambda - B) + F_0. \tag{2.3.3}$$

Substituting equation (2.3.2) and equation (2.3.3) into equation (2.3.1) we obtain

$$(I\lambda - A)(T(\lambda)(I\lambda - B) + F_0) = ((I\lambda - A)S(\lambda) + E_0)(I\lambda - B)$$

$$(I\lambda - A)(S(\lambda) - T(\lambda))(I\lambda - B) = (I\lambda - A)F_0 - E_0(I\lambda - B).$$

If $S(\lambda) \neq T(\lambda)$ then the degree of the left hand side is at least equal to 2, while the degree of the right hand side is equal to 1, we obtain a contradiction.

Hence $S(\lambda) = T(\lambda)$ and $(I\lambda - A)F_0 = E_0(I\lambda - B)$, that is

$F_0\lambda - AF_0 = E_0\lambda - E_0B$.

So $F_0 = E_0$ and $AF_0 = E_0B$, then $AE_0 = E_0B$.

If we prove that E_0 is a nonsingular matrix, then $E_0^{-1}AE_0 = B$,

which means that A and B are similar.

To do this divide $E(\lambda)$ on $I\lambda - B$ from the left, we obtain

$E(\lambda) = (I\lambda - B)U(\lambda) + R_0$

$I = (E(\lambda))^{-1}E(\lambda) = ((I\lambda - A)S(\lambda) + E_0)((I\lambda - B)U(\lambda) + R_0)$

$= (I\lambda - A)S(\lambda)(I\lambda - B)U(\lambda) + (I\lambda - A)S(\lambda)R_0 + E_0(I\lambda - B)U(\lambda) + E_0R_0$

$= (I\lambda - A)S(\lambda)(I\lambda - B)U(\lambda) + (I\lambda - A)S(\lambda)R_0 + (I\lambda - A)F_0U(\lambda) + E_0R_0$

$= (I\lambda - A)\{S(\lambda)(I\lambda - B)U(\lambda) + F_0U(\lambda) + S(\lambda)R_0\} + E_0R_0.$

Thus $S(\lambda)(I\lambda - B)U(\lambda) + F_0U(\lambda) + S(\lambda)R_0 = 0$ and $E_0R_0 = I$,

That is E_0 is a nonsingular. □

Theorem 2.3.2: If $A \in \mathbf{C}^{n\times n}$ and the invariant polynomials of $\lambda I - A$ of nonzero degree are $i_s(\lambda)$, $i_{s+1}(\lambda)$, ..., $i_n(\lambda)$, then the matrix A is similar to the block-diagonal matrix

$C_1 = diag[C_{i_s}, C_{i_{s+1}}, \cdots, C_{i_n}]$,

where C_{i_k} denotes the companion matrix associated with the invariant polynomial $i_k(\lambda)$, $(s \le k \le n)$.

Proof:

By corollary 2.2.4 the matrix C_1 is $n \times n$.

The matrix polynomial $\lambda I - C_{i_k}$ $(s \le k \le n)$ is equivalent to the diagonal matrix $D_k(\lambda) = diag[1,1,\cdots,1,i_k(\lambda)]$. This implies that $\lambda I - C_1$ is equivalent to the matrix polynomial $D(\lambda) = diag[D_s(\lambda), D_{s+1}(\lambda), \cdots, D_n(\lambda)]$.

Thus $\lambda I - C_1 = P(\lambda)D(\lambda)Q(\lambda)$, where $P(\lambda)$ and $Q(\lambda)$ are unimodular matrices.

Since by the elementary operations, the matrix polynomial $D(\lambda)$ can be transformed into $diag[1,\cdots,1,i_s(\lambda),\cdots,i_n(\lambda)]$.

Then $\lambda I - C_1$ and $diag[1,\cdots,1,i_s(\lambda),\cdots,i_n(\lambda)]$ are equivalent,

hence $1,\cdots,1,i_s(\lambda),\cdots,i_n(\lambda)$ are the invariant polynomials of both $\lambda I - C_1$ and $\lambda I - A$, which implies that A and C_1 are similar by the previous theorem. □

Example 2.3.1: Find C_1, which is stated in theorem 2.3.2 for the matrix

$$A = \begin{bmatrix} 3 & -1 & 0 \\ -1 & 3 & 0 \\ 1 & -1 & 2 \end{bmatrix}$$

Solution:

$$\lambda I - A = \begin{bmatrix} \lambda-3 & 1 & 0 \\ 1 & \lambda-3 & 0 \\ -1 & 1 & \lambda-2 \end{bmatrix}$$

$d_1(\lambda)=1$, $i_1(\lambda)=1$

$d_2(\lambda)=\lambda-2$, $i_2(\lambda)=\lambda-2$

$d_3(\lambda)=(\lambda-2)^2(\lambda-4)$, $i_3(\lambda)=\lambda^2-6\lambda+8$

$$C_1 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -8 & 6 \end{bmatrix}.$$

Let $A(\lambda)$ be an $n \times n$ matrix polynomial with invariant polynomials $i_1(\lambda),\ i_2(\lambda),\ \cdots,\ i_r(\lambda)$ where $i_s(\lambda),\ ,i_{s+1}(\lambda),\ \cdots,i_r(\lambda)$ are nonconstant. Then,

$$i_j(\lambda) = (f_{j1}(\lambda))^{\alpha_{j1}}(f_{j2}(\lambda))^{\alpha_{j2}}\cdots(f_{jk_j}(\lambda))^{\alpha_{jk_j}},\ j=s,s+1,\cdots,r \tag{2.3.4}$$

where $f_{j1}(\lambda), f_{j2}(\lambda)\cdots, f_{jk_j}(\lambda)$ are irreducible and pairwise relatively prime nonconstant monic scalar polynomials.

Definition 2.3.1: The **elementary divisors** of $A(\lambda)$ are the collection of factors $(f_{jl}(\lambda))^{\alpha_{jl}}\ (l=1,2,\cdots,k_j, j=s,s+1,\cdots,r)$ of $i_j(\lambda)$ where each factor is repeated as many times as it occurs in (2.3.4).

The positive integer α_{jl} is called the order of the elementary divisor $(f_{jl}(\lambda))^{\alpha_{jl}}$.

Example 2.3.2: Find the elementary divisors of the matrix polynomial

$$A(\lambda)=\begin{bmatrix}1 & 0 & 0 & 0 & 0\\ 0 & \lambda & 0 & 0 & 0\\ 0 & 0 & \lambda(\lambda-2) & 0 & 0\\ 0 & 0 & 0 & \lambda^3(\lambda-2)^2(\lambda+5) & 0\\ 0 & 0 & 0 & 0 & 0\end{bmatrix}$$

Solution:

$d_1(\lambda)=1$, $i_1(\lambda)=1$

$d_2(\lambda)=\lambda$, $i_2(\lambda)=\lambda$

$d_3(\lambda)=\lambda^2(\lambda-2)$, $i_3(\lambda)=\lambda(\lambda-2)$

$d_4(\lambda)=\lambda^5(\lambda-2)^3(\lambda+5)$, $i_4(\lambda)=\lambda^3(\lambda-2)^2(\lambda+5)$

The elementary divisors are: $\lambda,\lambda,\lambda^3,\lambda-2,(\lambda-2)^2$ and $\lambda+5$.

In example 2.3.2, we show how to find the elementary divisors from the invariant polynomials, but now we want to show how to construct the invariant polynomials $i_1(\lambda),i_2(\lambda),\cdots,i_r(\lambda)$ from the knowledge of the elementary divisors of $A(\lambda)$ and of the number r of its invariant polynomials or the rank of $A(\lambda)$.

Let $\lambda_1,\lambda_2,\cdots,\lambda_p$ be all the different complex numbers that appear in the elementary divisors, and let $(\lambda-\lambda_i)^{\alpha_{i1}},(\lambda-\lambda_i)^{\alpha_{i2}},\cdots,(\lambda-\lambda_i)^{\alpha_{i,k_i}}$ $(i=1,2,\cdots,p)$ be the elementary divisors containing the number λ_i, ordered in the descending order of the degrees $\alpha_{i1}\geq\alpha_{i2}\geq\cdots\geq\alpha_{ik_i}>0$.

The number r of invariant polynomials must be greater than or equal to $Max\{k_1,k_2,\cdots,k_p\}$.

Under this condition, the invariant polynomials $i_1(\lambda), i_2(\lambda), \cdots, i_r(\lambda)$ are given by the formula

$$i_j(\lambda) = \prod_{i=1}^{p} (\lambda - \lambda_i)^{\alpha_{i,r+1-j}}, j = 1,2,\cdots r$$

where we put $(\lambda - \lambda_i)^{\alpha_{ij}} = 1$ for $j > k_i$.

Example 2.3.3: Determine the invariant polynomials of a matrix polynomial of order 5 having rank 4 and elementary divisors $\lambda,, \lambda, \lambda^3, \lambda - 2, (\lambda - 2)^2, \lambda + 5$

Solution:

By ordering the elementary divisors in the descending order of the degrees, we get

$\lambda^3, \lambda, \lambda$

$(\lambda - 2)^2, \lambda - 2$

$\lambda - 5$.

By using the formula $i_j(\lambda) = \prod_{i=1}^{p} (\lambda - \lambda_i)^{\alpha_{i,r+1-j}}$, $j = 1,2,\cdots,r$,

where p is the number of the different complex numbers that appear in the elementary divisors and r is the rank of the matrix polynomial, we obtain

$$i_1(\lambda) = (\lambda - \lambda_1)^{\alpha_{14}} (\lambda - \lambda_2)^{\alpha_{24}} (\lambda - \lambda_3)^{\alpha_{34}} = (1)(1)(1) = 1$$

$$i_2(\lambda) = (\lambda - \lambda_1)^{\alpha_{13}} (\lambda - \lambda_2)^{\alpha_{23}} (\lambda - \lambda_3)^{\alpha_{33}} = (\lambda)(1)(1) = \lambda$$

$$i_3(\lambda) = (\lambda - \lambda_1)^{\alpha_{12}} (\lambda - \lambda_2)^{\alpha_{22}} (\lambda - \lambda_3)^{\alpha_{32}} = (\lambda)(\lambda - 2)(1) = \lambda(\lambda - 2)$$

$$i_4(\lambda) = (\lambda - \lambda_1)^{\alpha_{11}} (\lambda - \lambda_2)^{\alpha_{21}} (\lambda - \lambda_3)^{\alpha_{31}} = \lambda^3 (\lambda - 2)^2 (\lambda + 5).$$

Note that in the matrix polynomial $A(\lambda)$, the elementary divisors can be founded from the invariant polynomials as in example 2.3.2 and vice versa the invariant polynomials can be founded from the elementary divisors if the rank of $A(\lambda)$ is known as in example 2.3.3.

Now, from corollary 2.2.3 we obtain

Theorem 2.3.3: The matrix polynomials $A(\lambda)$ and $B(\lambda)$ are equivalent if and only if they have the same elementary divisors.

Corollary 2.3.1: A and B are similar if and only if $\lambda I - A$ and $\lambda I - B$ have the same elementary divisors.

Theorem 2.3.4: Let $A(\lambda)$ and $B(\lambda)$ be matrix polynomials, and let $C(\lambda) = diag[A(\lambda), B(\lambda)]$, a block diagonal matrix polynomial. Then the set of elementary divisors of $C(\lambda)$ is the union of the elementary divisors of $A(\lambda)$ and $B(\lambda)$.

Proof:

Let the Smith canonical form of $A(\lambda)$ and $B(\lambda)$ be given by

$A(\lambda) = E_1(\lambda)D_1(\lambda)F_1(\lambda)$ and $B(\lambda) = E_2(\lambda)D_2(\lambda)F_2(\lambda)$,

where $E_1(\lambda), F_1(\lambda), E_2(\lambda)$ and $F_2(\lambda)$ are with constant nonzero determinant. Then

$$C(\lambda) = \begin{bmatrix} A(\lambda) & 0 \\ 0 & B(\lambda) \end{bmatrix} = \begin{bmatrix} E_1(\lambda)D_1(\lambda)F_1(\lambda) & 0 \\ 0 & E_2(\lambda)D_2(\lambda)F_2(\lambda) \end{bmatrix}$$

$$= \begin{bmatrix} E_1(\lambda) & 0 \\ 0 & E_2(\lambda) \end{bmatrix} \begin{bmatrix} D_1(\lambda) & 0 \\ 0 & D_2(\lambda) \end{bmatrix} \begin{bmatrix} F_1(\lambda) & 0 \\ 0 & F_2(\lambda) \end{bmatrix}$$

$= E(\lambda) diag[D_1(\lambda), D_2(\lambda)] F(\lambda)$

where $E(\lambda) = diag[E_1(\lambda), E_2(\lambda)]$ and $F(\lambda) = diag[F_1(\lambda), F_2(\lambda)]$

and $\det E(\lambda) = const. \neq 0$ and $\det F(\lambda) = cons. \neq 0$. Let

$(\lambda-\lambda_0)^{\alpha_1}, (\lambda-\lambda_0)^{\alpha_2}, \cdots, (\lambda-\lambda_0)^{\alpha_p}$, $(\lambda-\lambda_0)^{\beta_1}, (\lambda-\lambda_0)^{\beta_2}, \cdots, (\lambda-\lambda_0)^{\beta_q}$

be the elementary divisors of $D_1(\lambda)$ and $D_2(\lambda)$ respectively, corresponding to the same complex number λ_0.

Arrange the set of exponents $\alpha_1, \alpha_2, \cdots, \alpha_p, \beta_1, \beta_2, \cdots, \beta_q$, in a nondecreasing order:

$\{\alpha_1, \alpha_2, \cdots, \alpha_p, \beta_1, \beta_2, \cdots, \beta_q\} = \{\gamma_1, \gamma_2, \cdots, \gamma_{p+q}\}$,

where $0 < \gamma_1 \leq \gamma_2 \leq \cdots \leq \gamma_{p+q}$.

In the Smith canonical form $D = diag[i_1(\lambda), i_2(\lambda), \cdots, i_r(\lambda), 0, 0, \cdots, 0]$ of $diag[D_1(\lambda), D_2(\lambda)]$, the invariant polynomial $i_r(\lambda)$ is divisible by $(\lambda-\lambda_0)^{\gamma_{p+q}}$ but not by $(\lambda-\lambda_0)^{\gamma_{p+q}+1}$, $i_{r-1}(\lambda)$ 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 $C(\lambda)$ and $\begin{bmatrix} D_1(\lambda) & 0 \\ 0 & D_2(\lambda) \end{bmatrix}$ corresponding to λ_0, are $(\lambda-\lambda_0)^{\gamma_1}, (\lambda-\lambda_0)^{\gamma_2}, \cdots, (\lambda-\lambda_0)^{\gamma_{p+q}}$. □

Example 2.3.4: Determine the elementary divisors of the matrix polynomial

$$C(\lambda) = \begin{bmatrix} \lambda-1 & 0 & 0 \\ 0 & \lambda^2+1 & \lambda^2+1 \\ 0 & -1 & \lambda^2-1 \end{bmatrix}$$

Solution:

Let $C(\lambda)=\begin{bmatrix} A(\lambda) & 0 \\ 0 & B(\lambda) \end{bmatrix}$

where $A(\lambda)=[\lambda-1]$, $B(\lambda)=\begin{bmatrix} \lambda^2+1 & \lambda^2+1 \\ -1 & \lambda^2-1 \end{bmatrix}$

By theorem 2.3.4 the set of elementary divisors of $C(\lambda)$ is the union of the elementary divisors of $A(\lambda)$ and $B(\lambda)$.

The elementary divisor of $A(\lambda)$ are:

$d_1(\lambda)=\lambda-1$, $i_1(\lambda)=\lambda-1$.

The elementary divisors of $B(\lambda)$ are:

$d_1(\lambda)=1$, $i_1(\lambda)=1$

$d_2(\lambda)=\lambda^2(\lambda^2+1)$, $i_2(\lambda)=\lambda^2(\lambda+i)(\lambda-i)$.

So, the elementary divisors of $C(\lambda)$ are: $\lambda-1, \lambda^2, \lambda+i, \lambda-i$.

2.4 One-Sided and Generalized Inverses

In this section we introduce and examine the notion of one-sided and generalized inverse of the matrix polynomial $A(\lambda)$.

Definition 2.4.1: An $m\times n$ matrix polynomial $A(\lambda)$ is called **generally invertible** if all its nonzero invariant polynomials are constant one.

Definition 2.4.2: An $n \times m$ matrix polynomial $N(\lambda)$ is called a **generalized inverse** of the $m \times n$ matrix polynomial $A(\lambda)$ if the followings hold:

$$A(\lambda)N(\lambda)A(\lambda) = A(\lambda),$$

$$N(\lambda)A(\lambda)N(\lambda) = N(\lambda).$$

Example 2.4.1: Show that the matrix polynomial $A(\lambda) = \begin{bmatrix} -2 & -\lambda^2 + \lambda \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$ is generally invertible and find its generalized inverse.

Solution:

$d_1(\lambda) = 1$, $i_1(\lambda) = 1$

$d_2(\lambda) = 1$, $i_2(\lambda) = 1$.

Therefore $A(\lambda)$ is generally invertible.

By the Smith canonical form we can write $A(\lambda)$ as

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

The generalized inverse of $A(\lambda)$ is

$$N(\lambda) = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2}\lambda^2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} -1 & \lambda & -2\lambda^2 \\ 0 & 1 & -4\lambda \\ 0 & 0 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} -\frac{1}{2} & \frac{1}{2}\lambda - \frac{1}{2}\lambda^2 & -\lambda^2 + 2\lambda^3 \\ 0 & 1 & -4\lambda \end{bmatrix}.$$

Theorem 2.4.1: Let $A(\lambda)$ be $m \times n$ matrix polynomial. Then $A(\lambda)$ is generally invertible if and only if $A(\lambda)$ has a generalized inverse.

Proof:

Suppose that the matrix polynomial $A(\lambda)$ is generally invertible and has a rank r.

Then by Smith canonical form $A(\lambda) = E(\lambda)\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}F(\lambda)$, where $E(\lambda)$ and $F(\lambda)$ are $m \times m$ and $n \times n$ matrix polynomials, respectively, with nonzero constant determinants. Let

$$N(\lambda) = F(\lambda)^{-1}\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}E(\lambda)^{-1}.$$

Then

$$N(\lambda)A(\lambda)N(\lambda) = F(\lambda)^{-1}\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}E(\lambda)^{-1}E(\lambda)\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}F(\lambda)F(\lambda)^{-1}\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}E(\lambda)^{-1}$$

$$= F(\lambda)^{-1}\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}E(\lambda)^{-1} = N(\lambda),$$

and

$$A(\lambda)N(\lambda)A(\lambda) = E(\lambda)\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}F(\lambda)F(\lambda)^{-1}\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}E(\lambda)^{-1}E(\lambda)\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}F(\lambda)$$

$$= E(\lambda)\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}F(\lambda) = A(\lambda).$$

Therefore $A(\lambda)$ has $N(\lambda)$ as a generalized inverse.

Conversely, suppose that $A(\lambda)$ has a generalized inverse, *i.e.* there is a matrix polynomial $N(\lambda)$ such that

$$N(\lambda)A(\lambda)N(\lambda) = N(\lambda), \; A(\lambda)N(\lambda)A(\lambda) = A(\lambda).$$

By using the Binet-Cauchy formula twice, any minor $a_k(\lambda)$ and $n_k(\lambda)$ of order k can be expressed as

$$a_k(\lambda)=\sum_k a_k(\lambda)n_k(\lambda)a_k(\lambda) \text{ and } n_k(\lambda)=\sum_k n_k(\lambda)a_k(\lambda)n_k(\lambda),\ k=1,2,\cdots,r.$$

Let $d_k(\lambda)$ be the greatest common divisor of all minors of order k formed from $A(\lambda)$ and $p_k(\lambda)$ be the greatest common divisor of all minors of order k formed from $N(\lambda)$, then $d_k(\lambda)$ is a divisor of $p_k(\lambda)$ and $p_k(\lambda)$ is a divisor of $d_k(\lambda)$.

Since $d_k(\lambda)$ and $p_k(\lambda)$ are monic scalar polynomials, then $d_k(\lambda)=p_k(\lambda)$.

$d_k(\lambda)=d_k^2(\lambda)a(\lambda)$, where $a(\lambda)$ is a scalar polynomial, then there are two cases:

Case 1: $d_k(\lambda)=0$, and this is a trivial case, *i.e.* $A(\lambda)$ is a zero matrix.

Case 2: $d_k(\lambda)=1$

$$d_k(\lambda)=d_{k-1}(\lambda)=\cdots=d_1(\lambda)=1$$

$$i_k(\lambda)=\frac{d_k(\lambda)}{d_{k-1}(\lambda)}=1$$

$$i_k(\lambda)=i_{k-1}(\lambda)=\cdots=i_1(\lambda)=1$$

i.e. all the invariant polynomials of $A(\lambda)$ and $N(\lambda)$ are constant ones. Therefore $A(\lambda)$ is generally invertible. □

Definition 2.4.3: An $m\times n$ matrix polynomial $A(\lambda)$ is said to be **left (respectively, right) invertible** if there exists an $n\times m$ matrix polynomial $A_L^{-1}(\lambda)$ (respectively, $A_R^{-1}(\lambda)$) such that $A_L^{-1}(\lambda)A(\lambda)=I_n$ (respectively $A(\lambda)A_R^{-1}(\lambda)=I_m$).

A matrix polynomial $A_L^{-1}(\lambda)$ (respectively, $A_R^{-1}(\lambda)$) is called **a left (respectively, right inverse)** of $A(\lambda)$.

Theorem 2.4.2: An $m \times n$ matrix polynomial $A(\lambda)$ is a right invertible if and only if all the invariant polynomials are constant one (in particular $m \leq n$)

Proof:

Suppose that the $m \times n$ matrix polynomial $A(\lambda)$ is a right invertible, then there exists an $n \times m$ matrix polynomial $N(\lambda)$ such that $A(\lambda)N(\lambda) = I_m$.

By using the Binet-Cauchy formula:

$$\det I_m = \sum_{1 \le j_1 < j_2 < \cdots < j_m \le n} A\begin{pmatrix} 1 & 2 & \cdots & m \\ j_1 & j_2 & \cdots & j_m \end{pmatrix}; \lambda \Bigg) N\begin{pmatrix} j_1 & j_2 & \cdots & j_m \\ 1 & 2 & \cdots & m \end{pmatrix}; \lambda \Bigg) \tag{2.4.1}$$

Let $d_m(\lambda)$ be the greatest common divisor of all minors of order m formed from $A(\lambda)$.

From equation (2.4.1) we obtain $1 = d_m(\lambda)a(\lambda)$ where $a(\lambda)$ is a scalar polynomial, this happen only if $d_m(\lambda)$ is a constant, since $d_m(\lambda)$ is monic, so, $d_m(\lambda)$ must be 1.

$d_m(\lambda)$ is divisible by $d_{m-1}(\lambda)$, $d_{m-1}(\lambda)$ is divisible by $d_{m-2}(\lambda)$, and so on, so, $d_m(\lambda) = d_{m-1}(\lambda) = \cdots = d_1(\lambda) = 1$.

$i_m(\lambda) = i_{m-1}(\lambda) = \cdots = i_1(\lambda) = 1$ *i.e.* all the invariant polynomials of $A(\lambda)$ are constant 1.

Conversely, suppose that all invariant polynomials of $A(\lambda)$ are constant 1.

By the Smith canonical form $A(\lambda) = E(\lambda)[I_m \quad 0]F(\lambda)$, where $E(\lambda)$ and $F(\lambda)$ are $m \times m$ and $n \times n$ matrix polynomials with constant nonzero determinants. Let

$N(\lambda) = F(\lambda)^{-1}\begin{bmatrix} I_m \\ 0 \end{bmatrix}E(\lambda)^{-1}$.

Then

$$A(\lambda)N(\lambda) = E(\lambda)[I_m \quad 0]F(\lambda)F(\lambda)^{-1}\begin{bmatrix} I_m \\ 0 \end{bmatrix}E(\lambda)^{-1} = I_m.$$

Therefore $A(\lambda)$ is a right invertible. □

Theorem 2.4.3: An $m \times n$ matrix polynomial $A(\lambda)$ is left invertible if and only if all the invariant polynomials are constant one (in particular $n \leq m$).

Proof:

We can prove theorem 2.4.3 similar to the proof of theorem 2.4.2.

CHAPTER THREE

FACTORIZATION OF SELFADJOINT AND SYMMETRIC MATRIX POLYNOMIALS

In this chapter we shall discuss a factorization of selfadjoint and symmetric matrix polynomials. This factorization is developed by Ran and Rodman. See paper [15].

This chapter consists of three sections; in the first section we shall discuss some definitions which are necessary to the subsequent two sections. Factorization of symmetric matrix polynomials over a general field **F** is introduced in section two. In section three we introduce the factorization of selfadjoint matrix polynomials on the real axis.

3.1 Definitions

In this section we want to state some definitions which are necessary to the subsequent sections.

Definition 3.1.1:

(*i*) An $n \times n$ matrix polynomial $A(\lambda)$ is called **selfadjoint** if $A(\lambda) = (A(\overline{\lambda}))^*$, *i.e.* $A(\lambda)$ is a selfadjoint matrix polynomial if the coefficients are Hermitian matrices.

(*ii*) An $n \times n$ matrix polynomial $A(\lambda)$ is called **symmetric** if $A(\lambda) = (A(\lambda))^T$ or $A(\lambda) = (A(-\lambda))^T$.

We note that: If $A(\lambda) = (A(\lambda))^T$ then the coefficients are symmetric matrices, while if $A(\lambda) = (A(-\lambda))^T$ then the coefficients are symmetric matrices if the powers of λ are even and skew symmetric if the powers of λ are odd.

Example 3.1.1: For any $n \times n$ matrix polynomial $M(\lambda)$:

(*i*) If $D \in \mathbf{C}^{n \times n}$ is a Hermitian matrix then $A(\lambda) := (M(\overline{\lambda}))^* DM(\lambda)$ is a selfadjoint matrix polynomial.

(*ii*) If $D \in \mathbf{C}^{n \times n}$ is a symmetric matrix then $A(\lambda) := (M(\lambda))^T DM(\lambda)$ is a symmetric matrix polynomial.

Let $A \in \mathbf{C}^{n \times n}$ is a Hermitian matrix[1]; we denote the number of positive eigenvalues of A by $\nu_+(A)$, the number of negative eigenvalues of A by $\nu_-(A)$, and the number of zero eigenvalues of A by $\nu_0(A)$. Hence $\nu_+(A) + \nu_-(A) + \nu_0(A) = n$.

Definition 3.1.2: For a Hermitian matrix A the **signature** of A, written $sigA$, is the difference between the number of positive eigenvalues and the number of negative eigenvalues (counting multiplicities, of course), that is $sigA = \nu_+(A) - \nu_-(A)$.

Definition 3.1.3: A selfadjoint matrix polynomial $A(\lambda)$ has a **constant signature** if the signature of the selfadjoint matrix polynomial $A(\lambda)$ does not depend on λ.

---

[1] The spectrum of a Hermitian matrix is real.

In the previous chapter we defined the rank of a matrix polynomial $A(\lambda)$ as the order of the largest nonzero minor, which is equivalent to the following:

Definition 3.1.4: The general rank of the matrix polynomial $A(\lambda)$ (denoted by $r(A)$) is defined by $r(A) = \max_{\lambda_0 \in C}\{rankA(\lambda_0)\}$.

The points $\lambda_0 \in \mathbf{C}$ for which $rankA(\lambda_0) = r(A)$ are called **regular points** of $A(\lambda)$; all other points $\lambda_0 \in \mathbf{C}$ are called **singular points** of $A(\lambda)$.

We note that the set of singular points is finite. To prove this, suppose that $r =: r(A)$. By Smith canonical form $A(\lambda)$ is equivalent to $diag[i_1(\lambda), i_2(\lambda), \cdots, i_r(\lambda), 0, 0, \cdots, 0]$ where $i_j(\lambda), j = 1,2,\cdots,r$ are the invariant polynomials of $A(\lambda)$, $\lambda_0 \in \mathbf{C}$ is a singular point of $A(\lambda)$ if $rankA(\lambda_0)$ is less than r, and this occurs if $i_j(\lambda_0) = 0$ for some $j = 1,2,\cdots,r$ *i.e.* λ_0 is a zero of the polynomial $i_j(\lambda)$, for some $j = 1,2,\cdots,r$.

The number of zeros of the invariant polynomials is less than or equal to rl where l is the maximum degree between the degrees of the invariant polynomials. Therefore the number of singular points of $A(\lambda) \le rl$, that means, the set of the singular points of $A(\lambda)$ is finite.

3.2 Factorization of Symmetric Matrix Polynomials Over a General Field

There are some conditions which are necessary to factorize the matrix polynomial $A(\lambda)$ into the form $A(\lambda) = M_*(\lambda) D M(\lambda)$, where D is either a constant matrix or a matrix polynomial, and $M(\lambda)$ is a matrix polynomial, these conditions and others such as the notion of $M_*(\lambda)$ are examined here.

We know from abstract algebra that an automorphism $\sigma: \mathbf{F} \to \mathbf{F}$ is a mapping from a field $\mathbf{F}$ onto itself such that σ is one to one, $\sigma(x+y) = \sigma(x) + \sigma(y)$, and $\sigma(xy) = \sigma(x)\sigma(y)$ for all $x, y \in \mathbf{F}$.

Let $a(\lambda) = \sum_{j=0}^{l} a_j \lambda^j$, be a scalar polynomial over **F**, let $a_*(\lambda) = \sum_{j=0}^{l} \sigma(a_j)\varepsilon^j \lambda^j$, where $\varepsilon = \pm 1$ and $\sigma^2(x) = x$ for all $x \in \mathbf{F}$. Then $a_*(\lambda)$ is also a scalar polynomial over **F**.

For an $m \times n$ matrix polynomial $X(\lambda) = [x_{ij}(\lambda)]_{i,j-1}^{m,n}$ over **F**, define $X_*(\lambda) = [\widetilde{x}_{ij}(\lambda)]_{i,j-1}^{n,m}$ where $\widetilde{x}_{ij}(\lambda) = [x_{ji}(\lambda)]_*$.

Theorem 3.2.1: Let $X(\lambda)$ and $Y(\lambda)$ be matrix polynomials with appropriate order. Then

(*i*) $[X(\lambda)Y(\lambda)]_* = Y_*(\lambda)X_*(\lambda)$,

(*ii*) $[X(\lambda)]_{**} = X(\lambda)$,

(*iii*) $[x(\lambda)X(\lambda) + y(\lambda)Y(\lambda)]_* = [x(\lambda)]_* X_*(\lambda) + [y(\lambda)]_* Y_*(\lambda)$ where $x(\lambda)$ and $y(\lambda)$ are scalar polynomials,

(*iv*) If $\det X(\lambda) = const. \neq 0$, then $(X_*(\lambda))^{-1} = [(X(\lambda))^{-1}]_*$.

Proof:

(*i*) Let $X(\lambda)$ be an $m \times l$ and $Y(\lambda)$ be an $l \times n$ matrix polynomials, also let

$$Z(\lambda) = X(\lambda)Y(\lambda).$$

The $(j,i)th$ element of $[X(\lambda)Y(\lambda)]_*$ is $[z_{ij}(\lambda)]_* = [\sum_{k=1}^{l} x_{ik}(\lambda) y_{kj}(\lambda)]_*$

where

$$x_{ik}(\lambda) = \overset{(0)}{x}_{ik} + \overset{(1)}{x}_{ik}\lambda + \cdots + \overset{(m_k)}{x}_{ik}\lambda^{m_k}, \quad y_{kj}(\lambda) = \overset{(0)}{y}_{kj} + \overset{(1)}{y}_{kj}\lambda + \cdots + \overset{(n_k)}{y}_{kj}\lambda^{n_k}$$

$$[z_{ij}(\lambda)]_* = [\sum_{k=1}^{l}(\overset{(0)}{x}_{ik} + \overset{(1)}{x}_{ik}\lambda + \cdots + \overset{(m_k)}{x}_{ik}\lambda^{m_k})(\overset{(0)}{y}_{kj} + \overset{(1)}{y}_{kj}\lambda + \cdots + \overset{(n_k)}{y}_{kj}\lambda^{n_k})]_*$$

$$= \sum_{k=1}^{l}[\sigma(\overset{(0)}{x}_{ik})\sigma(\overset{(0)}{y}_{kj}) + \cdots + \sigma(\overset{(m_k)}{x}_{ik})\sigma(\overset{(n_k)}{y}_{kj})\varepsilon^{m_k+n_k}\lambda^{m_k+n_k}] \tag{3.2.1}$$

The $(j,i)th$ element of $Y_*(\lambda)X_*(\lambda)$ is

$$\sum_{k=1}^{l}[y_{kj}(\lambda)]_*[x_{ik}(\lambda)]_*$$

$$= \sum_{k=1}^{l}(\sigma(\overset{(0)}{y}_{kj}) + \cdots + \sigma(\overset{(n_k)}{y}_{kj})\varepsilon^{n_k}\lambda^{n_k})(\sigma(\overset{(0)}{x}_{ik}) + \cdots + \sigma(\overset{(m_k)}{x}_{ik})\varepsilon^{m_k}\lambda^{m_k})$$

$$= \sum_{k=1}^{l}[\sigma(\overset{(0)}{y}_{kj})\sigma(\overset{(0)}{x}_{ik}) + \cdots + \sigma(\overset{(n_k)}{y}_{kj})\sigma(\overset{(m_k)}{x}_{ik})\varepsilon^{n_k+m_k}\lambda^{n_k+m_k}] \tag{3.2.2}$$

From (3.2.1) and (3.2.2) we obtain $[X(\lambda)Y(\lambda)]_* = Y_*(\lambda)X_*(\lambda)$.

(ii) Let $X(\lambda)$ be an $m \times n$ matrix polynomial.

The $(i,j)th$ element of $X(\lambda)$ is

$$x_{ij}(\lambda) = \overset{(0)}{x}_{ij} + \overset{(1)}{x}_{ij}\lambda + \cdots + \overset{(l)}{x}_{ij}\lambda^{l}. \tag{3.2.3}$$

The $(i,j)th$ element of $[X(\lambda)]_{**}$ is

$$[x_{ij}(\lambda)]_{**} = [\sigma(\overset{(0)}{x}_{ij}) + \sigma(\overset{(1)}{x}_{ij})\varepsilon\lambda + \cdots + \sigma(\overset{(l)}{x}_{ij})\varepsilon^{l}\lambda^{l}]_*$$

$$= \sigma(\sigma(\overset{(0)}{x}_{ij})) + \sigma(\sigma(\overset{(1)}{x}_{ij}))\varepsilon^{2}\lambda + \cdots + \sigma(\sigma(\overset{(l)}{x}_{ij}))\varepsilon^{2l}\lambda^{l}$$

$$= \overset{(0)}{x}_{ij} + \overset{(1)}{x}_{ij}\lambda + \cdots + \overset{(l)}{x}_{ij}\lambda^{l}. \tag{3.2.4}$$

From equation (3.2.3) and (3.2.4) we obtain $[X(\lambda)]_{**} = X(\lambda)$.

(iii) Suppose that $X(\lambda)$ and $Y(\lambda)$ are $m \times n$ matrix polynomials. The $(i,j)th$ element of $[x(\lambda)X(\lambda) + y(\lambda)Y(\lambda)]_*$ is

$$[x(\lambda)x_{ji}(\lambda)+y(\lambda)y_{ji}(\lambda)]_* = [x(\lambda)]_*[x_{ji}(\lambda)]_* + [y(\lambda)]_*[y_{ji}(\lambda)]_* \quad (3.2.5)$$

The $(i,j)th$ element of $[x(\lambda)]_* X_*(\lambda)+[y(\lambda)]_* Y_*(\lambda)$ is

$[x(\lambda)]_*[x_{ji}(\lambda)]_* + [y(\lambda)]_*[y_{ji}(\lambda)]_*$

From equation (3.2.5) we obtain that

$$[x(\lambda)X(\lambda)+y(\lambda)Y(\lambda)]_* = [x(\lambda)]_* X_*(\lambda)+[y(\lambda)]_* Y_*(\lambda).$$

(iv) Since $\det X(\lambda) = const. \neq 0$ then $X(\lambda)$ is an $n\times n$ invertible matrix polynomial *i.e.* it has an inverse $(X(\lambda))^{-1}$ and $(X(\lambda))^{-1}$ is a matrix polynomial. Thus $X(\lambda)(X(\lambda))^{-1} = I_n$, and then. $[X(\lambda)(X(\lambda))^{-1}]_* = I_*$.

By (i),

$$[(X(\lambda))^{-1}]_* X_*(\lambda) = I, \quad (3.2.6)$$

also $(X(\lambda))^{-1}X(\lambda) = I_n$ and then $[(X(\lambda))^{-1}X(\lambda)]_* = I_*$, hence.

$$X_*(\lambda)[(X(\lambda))^{-1}]_* = I \quad (3.2.7)$$

From equation (3.2.6) and (3.2.7) we obtain that

$(X_*(\lambda))^{-1} = [(X(\lambda))^{-1}]_*$. □

Theorem 3.2.2: Let $L(\lambda)$ be an $n\times n$ generally invertible matrix polynomial such that

$$L(\lambda) = L_*(\lambda),$$

and let r be the general rank of $L(\lambda)$. Then $L(\lambda)$ can be factorized in the form

$$L(\lambda) = M_*(\lambda)DM(\lambda),$$

where $M(\lambda)$ is an $r\times n$ right invertible matrix polynomial and D is an $r\times r$ constant matrix such that $D = D_*$.

Conversely, if $L(\lambda)=M_*(\lambda)DM(\lambda)$ holds for an $r\times n$ right invertible matrix polynomial $M(\lambda)$ and a constant matrix $D=D_*$, then $L(\lambda)=L_*(\lambda)$, $L(\lambda)$ is generally invertible, and has general rank r.

Proof:

We first prove the converse statement.

Since $M(\lambda)$ is an $r\times n$ right invertible matrix polynomial, then all the invariant polynomials of $M(\lambda)$ are constant one. By the Smith canonical form $M(\lambda)=\widetilde{E}(\lambda)[I_r \quad 0]\widetilde{F}(\lambda)$, where $\widetilde{E}(\lambda)$ and $\widetilde{F}(\lambda)$ are $r\times r$ and $n\times n$ matrix polynomials with nonzero constant determinants.

$$L(\lambda)=M_*(\lambda)DM(\lambda)=(\widetilde{E}(\lambda)[I_r \quad 0]\widetilde{F}(\lambda))_*D\widetilde{E}(\lambda)[I_r \quad 0]\widetilde{F}(\lambda)$$

$$=\widetilde{F}_*(\lambda)\begin{bmatrix} I_r \\ 0 \end{bmatrix}\widetilde{E}_*(\lambda)D\widetilde{E}(\lambda)[I_r \quad 0]\widetilde{F}(\lambda)$$

$$=\widetilde{F}_*(\lambda)\begin{bmatrix} \widetilde{E}_*(\lambda) & 0 \\ 0 & I_{(n-r)} \end{bmatrix}\begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \widetilde{E}(\lambda) & 0 \\ 0 & I_{(n-r)} \end{bmatrix}\widetilde{F}(\lambda)$$

Since

$$\det\left(\widetilde{F}_*(\lambda)\begin{bmatrix} \widetilde{E}_*(\lambda) & 0 \\ 0 & I_{(n-r)} \end{bmatrix}\right)=const.\neq 0$$

and

$$\det\left(\begin{bmatrix} \widetilde{E}(\lambda) & 0 \\ 0 & I_{(n-r)} \end{bmatrix}\widetilde{F}(\lambda)\right)=const.\neq 0.$$

Then by the uniqueness of the Smith canonical form, all the nonzero invariant polynomials of $L(\lambda)$ are constant 1.

Therefore $L(\lambda)$ is generally invertible and has general rank r. Also

For $n=1$ it is trivial. Suppose it is true for $n-1$. Then $A(\lambda)$ can be rewritten as

$$A(\lambda)=\begin{bmatrix} 0 & a_* \\ a & A_1 \end{bmatrix},$$

where $A_1 = A_{1*}$ is $(n-1)\times(n-1)$ matrix polynomial.

Since $\det A(\lambda) = const. \neq 0$, then it has an inverse $(A(\lambda))^{-1}$. Let $(A(\lambda))^{-1}$ be written as

$$(A(\lambda))^{-1} = \begin{bmatrix} \gamma & c_* \\ c & C_1 \end{bmatrix}$$

where $C_1 = C_{1*}$ is $(n-1)\times(n-1)$ matrix polynomial.

Put $y=\frac{1}{2}(1-c_*A_1c)$, $x=-A_1c-ay$, and $Y=\begin{bmatrix} y & x_* \\ c & I_{(n-1)} \end{bmatrix}$.

Since

$$A(\lambda)(A(\lambda))^{-1} = \begin{bmatrix} a_*c & a_*C_1 \\ a\gamma + A_1c & ac_* + A_1C_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & I_{(n-1)} \end{bmatrix} = I_n,$$

then, $a_*c=1$. Now $x=-A_1c-ay=-A_1c-\frac{1}{2}a+\frac{1}{2}ac_*A_1c=-\frac{1}{2}a-\frac{1}{2}A_1c$

$x_* = -\frac{1}{2}c_*A_1 - \frac{1}{2}a_*$ and then

$$x_*c = -\frac{1}{2}c_*A_1c - \frac{1}{2}a_*c = -\frac{1}{2}c_*A_1c - \frac{1}{2}$$

$x_*c = y-1$.

Now,

$$\begin{bmatrix} 1 & -x_* \\ 0 & I_{(n-1)} \end{bmatrix} Y \begin{bmatrix} 1 & 0 \\ -c & I_{(n-1)} \end{bmatrix} = \begin{bmatrix} y-x_*c & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & I \end{bmatrix}.$$

We conclude that $\det Y = 1$. Let

$$Y_*A(\lambda)Y = \begin{bmatrix} 1 & 0 \\ 0 & L_0 \end{bmatrix},$$

where $L_0(\lambda)$ is an $(n-1)\times(n-1)$ matrix polynomial.

Since $L_0(\lambda)$ is $(n-1)\times(n-1)$ matrix polynomial, then by mathematical induction.

$L_0(\lambda) = M_{0*}(\lambda)D_0M_0(\lambda)$.

$$Y_*A(\lambda)Y = \begin{bmatrix} 1 & 0 \\ 0 & M_{0*}(\lambda)D_0M_0(\lambda) \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & M_{0*}(\lambda) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & D_0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & M_0(\lambda) \end{bmatrix}.$$

$Y_*A(\lambda)Y = Y_*X_*(\lambda)L(\lambda)X(\lambda)Y$.

Therefore $L(\lambda) = M_*(\lambda)DM(\lambda)$, where

$$M(\lambda) = \begin{bmatrix} 1 & 0 \\ 0 & M_0(\lambda) \end{bmatrix} Y^{-1}(X(\lambda))^{-1} \text{ and } D = \begin{bmatrix} 1 & 0 \\ 0 & D_0 \end{bmatrix}.$$

Case (ii) if $r < n$.

Then by the Smith canonical form $L(\lambda) = E(\lambda)D_1F(\lambda)$, where $E(\lambda)$ and $F(\lambda)$ are $n\times n$ matrix polynomials with nonzero constant determinants.

Let $\widetilde{L}(\lambda) = (F_*(\lambda))^{-1}L(\lambda)(F(\lambda))^{-1}$. Then

$$\widetilde{L}_*(\lambda) = \left[(F_*(\lambda))^{-1}L(\lambda)(F(\lambda))^{-1}\right]_* = (F_*(\lambda))^{-1}L(\lambda)(F(\lambda))^{-1} = \widetilde{L}(\lambda).$$

Therefore $\widetilde{L}(\lambda) = \widetilde{L}_*(\lambda)$.

$$\widetilde{L}(\lambda) = (F_*(\lambda))^{-1}E(\lambda)D_1F(\lambda)(F(\lambda))^{-1} = (F_*(\lambda))^{-1}E(\lambda)D_1 \qquad (3.2.8)$$

Since $\widetilde{L}(\lambda) = \widetilde{L}_*(\lambda)$, then the last $n-r$ rows and the last $n-r$ columns of $\widetilde{L}(\lambda)$ are zeros.

From equation (3.2.8) $E(\lambda)D_1 = F_*(\lambda)\widetilde{L}(\lambda)$, and hence

$E(\lambda)D_1F(\lambda) = F_*(\lambda)\widetilde{L}(\lambda)F(\lambda)$. That is

$L(\lambda) = F_*(\lambda)\widetilde{L}(\lambda)F(\lambda)$

$$= F_*(\lambda)\begin{bmatrix} N(\lambda) & 0 \\ 0 & 0 \end{bmatrix}F(\lambda) = F_*(\lambda)\begin{bmatrix} I_r \\ 0 \end{bmatrix}N(\lambda)[I_r \quad 0]F(\lambda)$$

where $N(\lambda)$ is an $r \times r$ matrix polynomial formed by the first r rows and the first r columns of $\widetilde{L}(\lambda)$ and from equation (3.2.8) $\det N(\lambda) = const. \neq 0$, Thus, the $r \times r$ matrix polynomial $N(\lambda)$ satisfies the conditions in part (i). Hence $N(\lambda) = M_{1*}(\lambda)DM_1(\lambda)$

Then $L(\lambda) = F_*(\lambda)\begin{bmatrix} I_r \\ 0 \end{bmatrix}M_{1*}(\lambda)DM_1(\lambda)[I_r \quad 0]F(\lambda)$

$$= [M_1(\lambda) \quad [I_r \quad 0] \quad F(\lambda)]_* DM_1(\lambda)[I_r \quad 0]F(\lambda)$$

Let $M(\lambda) = M_1(\lambda)[I_r \quad 0]F(\lambda)$, then $M(\lambda)$ is an $r \times n$ right invertible matrix polynomial. □

If $L(\lambda)$ is not generally invertible, then the following example shows that the representation $L(\lambda) = M_*(\lambda)DM(\lambda)$ with D having the size equal to the general rank of $L(\lambda)$ is not always possible.

Example 3.2.1: Let $L(\lambda) = \begin{bmatrix} \lambda+1 & \lambda+1 \\ \lambda+1 & \lambda+1 \end{bmatrix}$ be a matrix polynomial. Then

$d_1(\lambda) = \lambda + 1$, $i_1(\lambda) = \lambda + 1$ and $i_2(\lambda) = 0$.

Therefore $L(\lambda)$ is not generally invertible. Take $\varepsilon = 1$, then $L_*(\lambda) = L(\lambda)$.

The general rank of $L(\lambda) = 1$. Suppose that $L(\lambda) = M_*(\lambda)DM(\lambda)$.

Then $M(\lambda) = [m_{11}(\lambda) \quad m_{12}(\lambda)]$ and $D = [d_{11}]$, since $r(L) = 1$

$$L(\lambda)=\begin{bmatrix}(m_{11}(\lambda))_* \\ (m_{12}(\lambda))_*\end{bmatrix}[d_{11}][m_{11}(\lambda) \quad m_{12}(\lambda)]$$

$$\begin{bmatrix}\lambda+1 & \lambda+1 \\ \lambda+1 & \lambda+1\end{bmatrix}=\begin{bmatrix}(m_{11}(\lambda))_* d_{11}m_{11}(\lambda) & (m_{11}(\lambda))_* d_{11}m_{12}(\lambda) \\ (m_{12}(\lambda))_* d_{11}m_{11}(\lambda) & (m_{12}(\lambda))_* d_{11}m_{12}(\lambda)\end{bmatrix}$$

If $m_{11}(\lambda)=0$, then $(m_{11}(\lambda))_* d_{11}m_{11}(\lambda)=0\neq\lambda+1$

If $m_{11}(\lambda)=a$ where a is a constant, then $(m_{11}(\lambda))_* d_{11}m_{11}(\lambda)=\sigma(a)d_{11}a$, the last term is of degree 0 and not equal to $\lambda+1$ of degree 1.

If $m_{11}(\lambda)=a+b\lambda$, a,b are constants, then

$(m_{11}(\lambda))_* d_{11}m_{11}(\lambda)=(\sigma(a)+\sigma(b)\lambda)d_{11}(a+b\lambda)$, the last term is of degree 2 and not equal to $\lambda+1$ of degree 1.

Therefore $L(\lambda)$ cannot be factorized as $M_*(\lambda)DM(\lambda)$.

Even if we omit the requirement that $M(\lambda)$ is right invertible. We can, however, obtain a factorization result for not generally invertible $L(\lambda)$ if we allow D to be a matrix polynomial (with special properties).

Before we state this result we shall give the following notes.

Note 3.2.1: If $f(\lambda)$ is monic scalar polynomial then $f_+(\lambda)=:\varepsilon^{\deg f(\lambda)}f_*(\lambda)$ is monic scalar polynomial.

Proof:

Let $f(\lambda)=a_0+a_1\lambda+\cdots+a_l\lambda^l$, $a_l=1$, be a monic scalar polynomial.

$$f_*(\lambda)=\sigma(a_0)+\sigma(a_1)\varepsilon\lambda+\cdots+\sigma(a_{l-1})\varepsilon^{l-1}\lambda^{l-1}+\sigma(a_l)\varepsilon^l\lambda^l,\ \sigma(a_l)=1.$$

$$f_+(\lambda)=\varepsilon^{\deg f(\lambda)}f_*(\lambda)=\varepsilon^l[\varepsilon^l\lambda^l+\sigma(a_{l-1})\varepsilon^{l-1}\lambda^{l-1}+\cdots+\sigma(a_0)]$$

$$=\lambda^l+\sigma(a_{l-1})\varepsilon^{2l-1}\lambda^{l-1}+\cdots+\varepsilon^l\sigma(a_0)$$

Therefore $f_+(\lambda)$ is a monic scalar polynomial. □

Note 3.2.2: $(f_1(\lambda)f_2(\lambda))_+ = f_{1+}(\lambda)f_{2+}(\lambda)$

Proof:

Let $f_1(\lambda) = \overset{(1)}{a_0} + \overset{(1)}{a_1}\lambda + \cdots + \overset{(1)}{a_{l_1}}\lambda^{l_1}$ and $f_2(\lambda) = \overset{(2)}{a_0} + \overset{(2)}{a_1}\lambda + \cdots + \overset{(2)}{a_{l_2}}\lambda^{l_2}$ be scalar polynomials.

$$f_1(\lambda)f_2(\lambda) = \overset{(1)}{a_0}\overset{(2)}{a_0} + (\overset{(1)}{a_0}\overset{(2)}{a_1} + \overset{(1)}{a_1}\overset{(2)}{a_0})\lambda + \cdots + \overset{(1)}{a_{l_1}}\overset{(2)}{a_{l_2}}\lambda^{l_1+l_2}$$

$$(f_1(\lambda)f_2(\lambda))_+ = \varepsilon^{l_1+l_2}\{\sigma(\overset{(1)}{a_0})\sigma(\overset{(2)}{a_0}) + \cdots + \sigma(\overset{(1)}{a_{l_1}})\sigma(\overset{(2)}{a_{l_2}})\varepsilon^{l_1+l_2}\lambda^{l_1+l_2}\}$$

$$= \varepsilon^{l_1}\{\sigma(\overset{(1)}{a_0}) + \cdots + \sigma(\overset{(1)}{a_{l_1}})\varepsilon^{l_1}\lambda^{l_1}\}\varepsilon^{l_2}\{\sigma(\overset{(2)}{a_0}) + \cdots + \sigma(\overset{(2)}{a_{l_2}})\varepsilon^{l_2}\lambda^{l_2}\}$$

$$= \varepsilon^{\deg f_1(\lambda)} f_{1*}(\lambda)\varepsilon^{\deg f_2(\lambda)} f_{2*}(\lambda) = f_{1+}(\lambda)f_{2+}(\lambda).$$ □

Theorem 3.2.3: Let $L(\lambda)$ be an $n \times n$ matrix polynomial such that

$$L(\lambda) = L_*(\lambda)$$

and let r be the general rank of $L(\lambda)$. Furthermore, let

$\{f_1(\lambda)^{\alpha_1}, f_2(\lambda)^{\alpha_2}, \cdots, f_q(\lambda)^{\alpha_q}\}$ be the collection of elementary divisors of $L(\lambda)$. Then $L(\lambda)$ admits a factorization

$$L(\lambda) = M_*(\lambda)D(\lambda)M(\lambda),$$

where $M(\lambda)$ is an $r \times n$ matrix polynomial and $D(\lambda) = D_*(\lambda)$ is an $r \times r$ matrix polynomial. Moreover, the collection of elementary divisors of $D(\lambda)$ is $\{f_j(\lambda), j \in J\}$, where the subset J of $\{1,2,\cdots,q\}$ consists precisely of those indices j for which $f_j(\lambda) = \varepsilon^{\deg f_j(\lambda)} f_{j*}(\lambda)$ and α_j is odd.

Proof:

As in the proof of theorem 3.2.2, we can assume that $r = n$.

Let $L(\lambda) = E(\lambda)D_1(\lambda)F(\lambda)$ be the Smith canonical form for $L(\lambda)$, where $D_1(\lambda) = diag[i_1(\lambda), i_2(\lambda), \cdots, i_n(\lambda)]$, $i_j(\lambda)$ $(j = 1,2,\cdots,n)$ are the invariant polynomials of $L(\lambda)$. Since $L(\lambda) = L_*(\lambda)$, then $L_*(\lambda) = F_*(\lambda)D_{1*}(\lambda)E_*(\lambda)$, where $D_{1*}(\lambda) = diag[i_{1*}(\lambda), i_{2*}(\lambda), \cdots, i_{n*}(\lambda)]$.

By the uniqueness of invariant polynomials, $i_{1*}(\lambda), i_{2*}(\lambda), \cdots, i_{n*}(\lambda)$ are the invariant polynomials of $L_*(\lambda)$.

$$i_j(\lambda) = \overset{(0)}{i}_j + \overset{(1)}{i}_j \lambda + \overset{(2)}{i}_j \lambda^2 + \cdots + \overset{(l_j)}{i}_j \lambda^{l_j}, \text{ where } \overset{(l_j)}{i}_j = 1,$$

while

$$i_{j*}(\lambda) = \sigma(\overset{(0)}{i}_j) + \sigma(\overset{(1)}{i}_j)\varepsilon\lambda + \sigma(\overset{(2)}{i}_j)\varepsilon^2\lambda^2 + \cdots + \sigma(\overset{(l_j)}{i}_j)\varepsilon^{l_j}\lambda^{l_j}, \text{ where } \sigma(\overset{(l_j)}{i}_j) = 1$$

Since $L(\lambda) = L_*(\lambda)$, then we have $i_{j*}(\lambda) = \varepsilon^{\deg i_j(\lambda)} i_j(\lambda)$ $(j = 1,2,\cdots,n)$

$$i_{j+}(\lambda) = \varepsilon^{\deg i_j(\lambda)} i_{j*}(\lambda) = \varepsilon^{2\deg i_j(\lambda)} i_j(\lambda) = i_j(\lambda).$$

Replacing $L(\lambda)$ by $(F_*(\lambda))^{-1} L(\lambda)(F(\lambda))^{1}$ then

$$(F_*(\lambda))^{-1} E(\lambda)D_1(\lambda)F(\lambda)(F(\lambda))^{-1} = (F_*(\lambda))^{-1} E(\lambda)D_1(\lambda) = T(\lambda)D_1(\lambda)$$

where $T(\lambda) = (F_*(\lambda))^{-1} E(\lambda)$.

$$L(\lambda) = \begin{bmatrix} t_{11}(\lambda) & \cdots & t_{1n}(\lambda) \\ \vdots & \ddots & \vdots \\ t_{n1}(\lambda) & \cdots & t_{nn}(\lambda) \end{bmatrix} \begin{bmatrix} i_1(\lambda) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & i_n(\lambda) \end{bmatrix} = \begin{bmatrix} t_{11}(\lambda)i_1(\lambda) & \cdots & t_{1n}(\lambda)i_n(\lambda) \\ \vdots & \ddots & \vdots \\ t_{n1}(\lambda)i_1(\lambda) & \cdots & t_{nn}(\lambda)i_n(\lambda) \end{bmatrix}.$$

The *jth* column of $L(\lambda)$ is divisible by $i_j(\lambda)$ $(j = 1,2,\cdots,n)$. By symmetry the *jth* row of $L(\lambda)$ is divisible by $i_{j*}(\lambda)$.

Let the nonconstant invariant polynomials of $L(\lambda)$ be $i_s(\lambda), i_{s+1}(\lambda), \cdots, i_n(\lambda)$ and factor them. Then

$$i_j(\lambda) = \prod_{l=1}^{k_j} (f_{jl}(\lambda))^{\alpha_{jl}} = \prod_{l=1}^{k_j} (f_{jl+}(\lambda))^{\alpha_{jl}} \quad (j = s, s+1, \cdots, n),$$

and by the uniqueness of the decomposition, the set $\{f_{j1}(\lambda), f_{j2}(\lambda), \cdots, f_{jk_j}(\lambda)\}$ must consist of selfsymmetric scalar polynomials $(f_{jl}(\lambda) = f_{jl+}(\lambda))$ and/or of pairs of mutually symmetric scalar polynomials $f_{jl_1}(\lambda) = f_{jl_2+}(\lambda)$ in this case necessarily $\alpha_{jl_1} = \alpha_{jl_2}$.

Let $f_{j1}(\lambda). f_{j2}(\lambda), \cdots, f_{j,p_j}(\lambda)$ be selfsymmetric and

$$f_{j,p_j+1}(\lambda) = (f_{j,P_j+2}(\lambda))_+, f_{j,P_j+3}(\lambda) = (f_{j,p_j+4}(\lambda))_+, \cdots, f_{j,p_j+2q_j-1}(\lambda) = (f_{j,p_j+2q_j}(\lambda))_+$$

where $p_j + 2q_j = k_j$.

Let j_0 be the smallest index such that $\alpha_{j_0 l} > 1$ for some $l \in \{1,2,\cdots,p_j\}$.

Say, $\alpha_{j_0 1} > 1$ (if no such j_0 exists, we put)

$$h_j = \prod_{l=1}^{q_j} (f_{j,p_j+2l}(\lambda))^{\alpha_{j,p_j+2l}} \quad (j = s, s+1, \cdots, n).$$

Define $h_j = \prod_{l=1}^{q_j} (f_{j,p_j+2l}(\lambda))^{\alpha_{j,p_j+2l}}$, $j = s, s+1, \cdots, j_0 - 1$

$$h_j = f_{j1}(\lambda) \prod_{l=1}^{q_j} (f_{j,p_j+2l}(\lambda))^{\alpha_{j,p_j+2l}}, \quad j = j_0, j_0+1, \cdots, n.$$

Where we assume that the elementary divisors are numbered so that $f_{j1}(\lambda) = f_{j_0 1}(\lambda)$ for $j = j_0 + 1, j_0 + 2, \cdots, n$

We define also $h_j = 1$ for $j = 1, 2, \cdots, s-1$

The divisibility relation among the $i_j(\lambda)$'s implies that whenever $f_{j_1 l_1}(\lambda) = f_{j_2 l_2}(\lambda)$ where $j_1 < j_2$, then $\alpha_{j_1 l_1} \le \alpha_{j_2 l_2}$. It follows that h_j divides h_{j+1} $(j = 1,2,\cdots,n-1)$

$i_j(\lambda) = h_{j*} g_j h_j$ $(j = 1,2,\cdots,s,s+1,\cdots,n)$

where $g_j = 1$ for $j = 1,2,\cdots,s-1$

$$g_j = \pm\prod_{l=1}^{p_j}(f_{jl}(\lambda))^{\alpha_{jl}} \text{ for } j = s,s+1,\cdots,j_0 - 1,$$

$$g_j = \pm(f_{j1}(\lambda))^{\alpha_{j1}-2}\prod_{l=2}^{p_j}(f_{jl}(\lambda))^{\alpha_{jl}} \text{ for } j = j_0, j_0+1,\cdots,n.$$

The sign + or – in g_j is chosen so that g_j is monic.

g_j divides g_{j+1} for $j = 1,2,\cdots,n-1$.

$$L(\lambda) = \begin{bmatrix} t_{11}(\lambda) & \cdots & t_{1s}(\lambda)i_s(\lambda) & \cdots & t_{1n}(\lambda)i_n(\lambda) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ t_{s1}(\lambda) & \cdots & t_{ss}(\lambda)i_s(\lambda) & \cdots & t_{sn}(\lambda)i_n(\lambda) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ t_{n1}(\lambda) & \cdots & t_{ns}(\lambda)i_s(\lambda) & \cdots & t_{nn}(\lambda)i_n(\lambda) \end{bmatrix}$$

$$L(\lambda) = H_* \begin{bmatrix} t_{11}(\lambda) & \cdots & t_{1s}(\lambda)h_{s*}g_s & \cdots & t_{1n}(\lambda)h_{n*}g_n \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \frac{t_{s1}(\lambda)}{h_{s*}} & \cdots & t_{ss}(\lambda)g_s & \cdots & \frac{t_{sn}(\lambda)h_{n*}g_n}{h_{s*}} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \frac{t_{n1}(\lambda)}{h_{n*}} & \cdots & \frac{t_{ns}(\lambda)h_{s*}g_s}{h_{n*}} & \cdots & t_{nn}(\lambda)g_n \end{bmatrix} H$$

where $H = diag[1,\cdots,1,h_s,\cdots,h_n]$. Let

$$\widetilde{L}(\lambda) = \begin{bmatrix} t_{11}(\lambda) & \cdots & t_{1s}(\lambda)h_{s*}g_s & \cdots & t_{1n}(\lambda)h_{n*}g_n \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \frac{t_{s1}(\lambda)}{h_{s*}} & \cdots & t_{ss}(\lambda)g_s & \cdots & \frac{t_{sn}(\lambda)h_{n*}g_n}{h_{s*}} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \frac{t_{n1}(\lambda)}{h_{n*}} & \cdots & \frac{t_{ns}(\lambda)h_{s*}g_s}{h_{n*}} & \cdots & t_{nn}(\lambda)g_n \end{bmatrix}$$

Then $L(\lambda) = H_* \widetilde{L}(\lambda) H$ for some matrix polynomial $\widetilde{L}(\lambda) = \widetilde{L}_*(\lambda)$.

Let $\widetilde{i}_1(\lambda), \widetilde{i}_2(\lambda), \cdots, \widetilde{i}_n(\lambda)$ be the invariant polynomials of $\widetilde{L}(\lambda)$. By using the Binet-Cauchy formula: we find that, the determinant of every $l \times l$ submatrix in $L(\lambda)$ is a linear combination of determinants of $l \times l$ submatrices in $\widetilde{L}(\lambda)$ when the determinants are multiplied by $\prod_{j=1}^{l} (h_{j*} h_j)$.

$i_1(\lambda) i_2(\lambda) \cdots i_l(\lambda)$ divides $\widetilde{i}_1(\lambda) \widetilde{i}_2(\lambda) \cdots \widetilde{i}_l(\lambda) \prod_{j=1}^{l} (h_{j*} h_j)$ for $l = 1, 2, \cdots, n$.

Since $i_j(\lambda) = h_{j*} g_j h_j$ $(j = 1, 2, \cdots, n)$ then $g_1 g_2 \cdots g_l$ divides $\widetilde{i}_1(\lambda) \widetilde{i}_2(\lambda) \cdots \widetilde{i}_l(\lambda)$.

On the other hand, for the $(j,l)th$ entries $\widetilde{p}_{jl}(\lambda)$ of $\widetilde{L}(\lambda)$ and $p_{jl}(\lambda)$ of $L(\lambda)$, we obtain $\widetilde{p}_{jl}(\lambda) = h_{j*}^{-1} p_{jl}(\lambda) h_l^{-1} = h_{j*}^{-1} p_{jl}(\lambda) i_l^{-1}(\lambda) h_{l*} g_l$ and since (assuming $j \le l$) both $p_{jl}(\lambda) i_l^{-1}(\lambda)$ and $h_{j*}^{-1} h_{l*}$ are scalar polynomials, $\widetilde{p}_{jl}(\lambda)$ is divisible by g_l. Also, $\widetilde{p}_{jl}(\lambda)$ is divisible by g_j (because g_j divides g_l if $j \le l$).

We see that $\widetilde{p}_{jl}(\lambda)$ is divisible by $g_{\max(j,l)}$. Therefore, the determinant of every $l \times l$ submatrix of $\widetilde{L}(\lambda)$ is divisible by $g_1 g_2 \cdots g_l$. Consequently, $\widetilde{i}_1(\lambda) \widetilde{i}_2(\lambda) \cdots \widetilde{i}_l(\lambda)$ divides $g_1 g_2 \cdots g_l$.

Comparing this result with the previously obtained opposite divisibility relation, we conclude that $g_1, g_2, \cdots, g_n$ are the invariant polynomials of $\widetilde{L}(\lambda)$.

We repeat the procedure given above with $L(\lambda)$ replaced by $\widetilde{L}(\lambda)$, and so on, until (after a finite number of steps) we obtain a matrix polynomial $D(\lambda)$ with the properties required in theorem 3.2.3. □

3.3 Factorization of Selfadjoint Matrix Polynomials

On the Real Axis

In this section we consider the selfadjoint matrix polynomials $L(\lambda)$ over **C**, that is, matrix polynomials with the following property:

$$L(\lambda) = (L(\bar{\lambda}))^*, \lambda \in \mathbf{C}.$$

Theorem 3.3.1: Let $L(\lambda)$ be a selfadjoint $n \times n$ matrix polynomial. Then $L(\lambda)$ admits a factorization

$$L(\lambda) = (M(\bar{\lambda}))^* DM(\lambda),$$

where D is an $m \times m$ constant Hermitian matrix and $M(\lambda)$ an $m \times n$ matrix polynomial, if and only if

$$m \geq m_0$$

where $m_0 = \max_{\lambda \in R} \nu_+(L(\lambda)) + \max_{\lambda \in R} \nu_-(L(\lambda))$.

Moreover, in all factorization $L(\lambda) = (M(\bar{\lambda}))^* DM(\lambda)$ having the minimal size $m_0 \times m_0$ of D, the matrix D is uniquely determined up to congruence[3], that is, D has $\max \nu_+(L(\lambda))$ positive eigenvalues and $\max \nu_-(L(\lambda))$ negative eigenvalues (multiplicities counted).

Proof:

The direct statement can be easily proved as follows:

Suppose that the selfadjoint matrix polynomial $L(\lambda)$ admits a factorization $L(\lambda) = (M(\bar{\lambda}))^* DM(\lambda)$.

[3] Two square matrices A and B are said to be congruent if there exists a nonsingular matrix P such that $A = PBP^*$.

Let λ_0 and λ_1 be real numbers for which $\nu_+(L(\lambda_0)) = \max_{\lambda\in R} \nu_+(L(\lambda))$ and $\nu_-(L(\lambda_1)) = \max_{\lambda\in R} \nu_-(L(\lambda))$. Then

$L(\lambda_0) = (M(\lambda_0))^* DM(\lambda_0) = Y_0^* DY_0$, and

$L(\lambda_1) = (M(\lambda_1))^* DM(\lambda_1) = Y_1^* DY_1$

where $Y_0 = M(\lambda_0)$ and $Y_1 = M(\lambda_1)$.

The Hermitian matrix D must have at least $\nu_+(L(\lambda_0))$ positive eigenvalues and D must have at least $\nu_-(L(\lambda_1))$ negative eigenvalues[4].

$$\begin{aligned} m &\geq \nu_+(D) + \nu_-(D) \\ &\geq \nu_+(L(\lambda_0)) + \nu_-(L(\lambda_1)) \\ &= \max_{\lambda\in R} \nu_+(L(\lambda)) + \max_{\lambda\in R} \nu_-(L(\lambda)) = m_0, \end{aligned}$$

That is, $m \geq m_0$.

If $D^{'}$ is an $m_0 \times m_0$ then, the least number of positive eigenvalues of D is $\max_{\lambda\in R} \nu_+(L(\lambda))$ and the least number of negative eigenvalues of D is $\max_{\lambda\in R} \nu_-(L(\lambda))$.

Next we shall prove the converse statement after introducing the following lemma and proposition:

Lemma 3.3.1: Let $L(\lambda)$ be a selfadjoint $n \times n$ matrix polynomial with $\det L(\lambda) \not\equiv 0$ and let $m_0 \geq n$ be defined by $m_0 = \max_{\lambda\in R} \nu_+(L(\lambda)) + \max_{\lambda\in R} \nu_-(L(\lambda))$.

Then there exists an $m_0 \times m_0$ selfadjoint matrix polynomial $\widetilde{L}(\lambda)$ such that

---

[4] Let A and B be $n \times n$ Hermitian matrices. If $A = MBM^*$ for some matrix M, then $\nu_+(A) = \nu_+(B)$, $\nu_-(A) = \nu_-(B)$ and $\nu_0(A) = \nu_0(B)$. See [11] page 187.

$$\widetilde{L}(\lambda)=\begin{bmatrix} L(\lambda) & 0 \\ 0 & * \end{bmatrix}$$

and such that

$$\max_{\lambda\in R} \nu_+(\widetilde{L}(\lambda))+\max_{\lambda\in R} \nu_-(\widetilde{L}(\lambda))=m_0 .$$

Proof: See [15] page 853. □

Theorem 3.3.2: [4]

Let $L(\lambda)$ be a selfadjoint $n\times n$ matrix polynomial with $\det L(\lambda)\not\equiv 0$. If $L(\lambda)$ has constant signature, then

$$L(\lambda)=(M(\overline{\lambda}))^* DM(\lambda)$$

where $M(\lambda)$ is an $n\times n$ matrix polynomial and D is an $n\times n$ nonsingular Hermitian complex matrix.

Proof:

We use induction on $\delta=\deg\det L(\lambda)$. We shall show that the assertion is true in the nonzero constant case, that is, for $\delta=0$, see [13]. We consider the case for $\delta>0$.

Let $L(\lambda)=E(\lambda)D(\lambda)F(\lambda)$ be the Smith canonical form of $L(\lambda)$, where $E(\lambda)$ and $F(\lambda)$ are of nonzero constant determinants and $D(\lambda)=diag[i_1(\lambda),i_2(\lambda),\cdots,i_n(\lambda)]$ is the diagonal matrix whose diagonal entries $i_j(\lambda)$ are the invariant polynomials of $L(\lambda)$. Thus $i_k(\lambda)$ are monic polynomials and $i_k(\lambda)$ divides $i_{k+1}(\lambda)$ for $k<n$.

As $L^*(\overline{\lambda})=L(\lambda)$, the polynomials $i_k(\lambda)$ have real coefficients. By replacing $L(\lambda)$ with $(F^*(\overline{\lambda}))^{-1}L(\lambda)(F(\lambda))^{-1}$, we may assume that $L(\lambda)=T(\lambda)D(\lambda)$, where $T(\lambda)=(F^*(\overline{\lambda}))^{-1}E(\lambda)$.

As $\delta > 0$, $i_n(\lambda)$ is not a constant. Let λ_0 be a root of $i_n(\lambda)$. The last column of $L(\lambda)$ is divisible by $\lambda - \lambda_0$. Since $L^*(\overline{\lambda}) = L(\lambda)$, the last row of $L(\lambda)$ is divisible by $\lambda - \overline{\lambda}_0$. Hence if either $\lambda_0 \neq \overline{\lambda}_0$, or $\lambda_0 = \overline{\lambda}_0$ and $(\lambda - \lambda_0)^2$ divides $i_n(\lambda)$, we have a factorization

$$L(\lambda) = M_1^*(\overline{\lambda}) D_1(\lambda) M_1(\lambda)$$

where $D_1(\lambda)$ is an $n \times n$ selfadjoint matrix polynomial and $M_1(\lambda) = diag[1, \cdots, 1, \lambda - \lambda_0]$.

Since $D_1(\lambda)$ has constant signature and $\det D_1(\lambda)$ has degree $\delta - 2$, we can use the induction hypothesis to obtain the assertion in this case.

It remains to consider the case in which $\lambda_0 \in \mathbf{R}$ and λ_0 is a simple root of $i_n(\lambda)$. For $\lambda \in \mathbf{R}$ let $f : \mathbf{C}^n \times \mathbf{C}^n \to \mathbf{C}$ be the Hermitian form[5] defined by

$$f(\vec{x}, \vec{y}) = \vec{y}^* L(\lambda_0) \vec{x}.$$

The radical K of f[6] is the nullspace of the matrix $L(\lambda_0)$. If r is the rank of $L(\lambda_0)$, then the last $n - r$ columns of $L(\lambda_0)$ are 0, and so K consists of all vectors in $\mathbf{C}^n$ whose first r coordinates are 0. Let $\vec{v}_1$ be one of these vectors in $\mathbf{C}^n$, then $f(\vec{v}_1, \vec{v}_1) = \vec{v}_1^* L(\lambda_0) \vec{v}_1 = 0$.

Now, let $g : K \times K \to \mathbf{C}$ be another Hermitian form defined by $g(\vec{x}, \vec{y}) = \vec{y}^* L'(\lambda_0) \vec{x}$, where $L'(\lambda)$ is the derivative of $L(\lambda)$ with respect to λ.

Then $g(\vec{v}_1, \vec{v}_1) = \vec{v}_1^* L'(\lambda_0) \vec{v}_1 = 0$.

---

[5] $f : \mathbf{C}^n \times \mathbf{C}^n \to \mathbf{C}$ is in Hermitian form if $f(\vec{x}, \vec{y}) = (f(\vec{y}, \vec{x}))^*$

[6] The radical K of f is defined as $\{\vec{x} : f(\vec{x}, \vec{y}) = 0 \quad for all\ \vec{y} \in \mathbf{C}^n\}$.

Choose a complex unitary matrix V having $\vec{v}_1$ as the first column, and set $B(\lambda) = V^* L(\lambda) V$. As $\vec{v}_1 \in K$, the first column of $B(\lambda_0)$ is 0, and consequently also the first row of this matrix is 0. Since $\vec{v}_1^{\,*} L'(\lambda_0) \vec{v}_1 = g(\vec{v}_1, \vec{v}_1) = 0$, the first entry of $B'(\lambda_0) = V^* L'(\lambda_0) V$ is 0. Hence all elements of $B(\lambda)$ in the first row or column are divisible by $\lambda - \lambda_0$ and the first element is divisible by $(\lambda - \lambda_0)^2$. It follows that we have a factorization

$$B(\lambda) = M_2^*(\bar{\lambda}) D_2(\lambda) M_2(\lambda),$$

where $D_2(\lambda)$ is an $n \times n$ selfadjoint matrix polynomial and $M_2(\lambda) = diag[\lambda - \lambda_0, 1, \cdots, 1]$.

Since $D_2(\lambda)$ has constant signature and $\det D_2(\lambda)$ has degree $\delta - 2$, we can use the inductive hypothesis to complete the proof. □

Proposition 3.3.1: Let $L(\lambda)$ be a selfadjoint $n \times n$ matrix polynomial with $\det L(\lambda) \not\equiv 0$ such that

$$\max_{\lambda \in R} \nu_+(L(\lambda)) + \max_{\lambda \in R} \nu_-(L(\lambda)) = n$$

then $L(\lambda)$ admits a factorization

$$L(\lambda) = (M(\bar{\lambda}))^* D M(\lambda)$$

in which D is $n \times n$.

Now, we want to show that this proposition is equivalent to the previous theorem.

Suppose that $L(\lambda)$ has a constant signature.

Let $\lambda_1, \lambda_2 \in \mathbf{R}$ such that $\nu_+(L(\lambda_1)) = \max_{\lambda \in R} \nu_+(L(\lambda))$ and $\nu_-(L(\lambda_2)) = \max_{\lambda \in R} \nu_-(L(\lambda))$.

$\nu_+(L(\lambda_1)) - \nu_-(L(\lambda_1)) = \nu_+(L(\lambda_2)) - \nu_-(L(\lambda_2))$

then $\nu_+(L(\lambda_1)) + \nu_-(L(\lambda_2)) = \nu_+(L(\lambda_2)) + \nu_-(L(\lambda_1))$

Since $\det L(\lambda) \neq 0$ then there is no zero eigenvalues.

$\nu_+(L(\lambda_1)) + \nu_-(L(\lambda_1)) = n$

$\nu_+(L(\lambda_2)) + \nu_-(L(\lambda_2)) = n$

By summing the two previous equations we obtain:

$\nu_+(L(\lambda_1)) + \nu_-(L(\lambda_2)) + \nu_-(L(\lambda_1)) + \nu_+(L(\lambda_2)) = 2n$

$\nu_+(L(\lambda_1)) + \nu_-(L(\lambda_2)) + \nu_+(L(\lambda_1)) + \nu_-(L(\lambda_2)) = 2n$

Then $\nu_+(L(\lambda_1)) + \nu_-(L(\lambda_2)) = n$ *i.e.* $\max_{\lambda \in R} \nu_+(L(\lambda)) + \max_{\lambda \in R} \nu_-(L(\lambda)) = n$.

Conversely, suppose that $\max_{\lambda \in R} \nu_+(L(\lambda)) + \max_{\lambda \in R} \nu_-(L(\lambda)) = n$

Let $\nu_+(L(\lambda_1)) = \max_{\lambda \in R} \nu_+(L(\lambda))$ and $\nu_-(L(\lambda_2)) = \max_{\lambda \in R} \nu_-(L(\lambda))$

$\nu_+(L(\lambda_1)) + \nu_-(L(\lambda_1)) = n$, since $\det L(\lambda) \neq 0$

$\max_{\lambda \in R} \nu_+(L(\lambda)) + \nu_-(L(\lambda_1)) = n$.

Now $\nu_-(L(\lambda_1)) = \max_{\lambda \in R} \nu_-(L(\lambda))$

Then $\nu_+(L(\lambda_2)) + \nu_-(L(\lambda_2)) = n$, since $\det L(\lambda) \neq 0$

$\nu_+(L(\lambda_2)) + \max_{\lambda \in R} \nu_-(L(\lambda)) = n$.

Thus $\nu_+(L(\lambda_2)) = \max_{\lambda \in R} \nu_+(L(\lambda))$

i.e. the maximum number of positive eigenvalues occurs at λ_1 and λ_2 and also the maximum number of negative eigenvalues occurs at λ_1 and λ_2.

If we take any point λ different from λ_1 and λ_2,

then $\nu_+(L(\lambda)) + \nu_-(L(\lambda)) = n$, since $\det L(\lambda) \neq 0$.

If $\max_{\lambda \in R} \nu_+(L(\lambda))$ does not occur at any $\lambda \in \mathbf{R}$, then $\nu_-(L(\lambda))$ at this point is greater than $\max_{\lambda \in R} \nu_-(L(\lambda))$, and this is a contradiction.

Similarly, we can arrive to a contradiction in the case $\max_{\lambda \in R} \nu_-(L(\lambda))$. □

Now we can easily finish the proof of theorem 3.3.1. Given a selfadjoint matrix polynomial $L(\lambda)$ with $\det L(\lambda) \not\equiv 0$, construct $\widetilde{L}(\lambda)$ as in lemma 3.3.1 and apply proposition 3.3.1 to $\widetilde{L}(\lambda)$:

$$\widetilde{L}(\lambda) = (N(\overline{\lambda}))^* D N(\lambda),$$

where D is a constant $m_0 \times m_0$ Hermitian matrix. Then $L(\lambda) = (M(\overline{\lambda}))^* D M(\lambda)$ holds for $M(\lambda)$ formed by first n columns of $N(\lambda)$. □

Definition 3.3.1: An $n \times n$ selfadjoint matrix polynomial $M(\lambda)$ will be called **elementary** if $r(M) = 1$ and $M(\lambda)$ is positive semidefinite for all real λ.

Lemma 3.3.2: $M(\lambda)$ is elementary if and only if $M(\lambda)$ is of the form $M(\lambda) = \vec{x}(\lambda)(\vec{x}(\overline{\lambda}))^*$ where $\vec{x}(\lambda) \neq \vec{0}$ is an $n \times 1$ column matrix.

Proof:

Suppose that $M(\lambda)$ is elementary, then $r(M) = 1$ and $M(\lambda)$ is positive semidefinite for all real λ *i.e.* $M(\lambda)$ is a Hermitian matrix and all its eigenvalues are nonnegative for all real λ. By corollary 5.5.2 in [11][7]

[7] A Hermitian matrix $H \in C^{n \times n}$ of rank r is positive semidefinite if and only if it is congruent to a matrix of the form $\begin{bmatrix} I_r & 0 \\ 0 & 0_{(n-r)} \end{bmatrix}$.

$$M(\lambda)=P(\lambda)\begin{bmatrix} I_1 & 0 \\ 0 & 0_{(n-1)} \end{bmatrix}(P(\overline{\lambda}))^*$$

$$=\begin{bmatrix} P_{11}(\lambda) & \cdots & P_{1n}(\lambda) \\ \vdots & \ddots & \vdots \\ P_{n1}(\lambda) & \cdots & P_{nn}(\lambda) \end{bmatrix}\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}\begin{bmatrix} \overline{P}_{11}(\overline{\lambda}) & \cdots & \overline{P}_{n1}(\overline{\lambda}) \\ \vdots & \ddots & \vdots \\ \overline{P}_{1n}(\overline{\lambda}) & \cdots & \overline{P}_{nn}(\overline{\lambda}) \end{bmatrix}$$

$$=\begin{bmatrix} P_{11}(\lambda)\overline{P}_{11}(\overline{\lambda}) & \cdots & P_{11}(\lambda)\overline{P}_{n1}(\overline{\lambda}) \\ \vdots & \ddots & \vdots \\ P_{n1}(\lambda)\overline{P}_{11}(\overline{\lambda}) & \cdots & P_{n1}(\lambda)\overline{P}_{n1}(\overline{\lambda}) \end{bmatrix}$$

$$=\begin{bmatrix} P_{11}(\lambda) \\ \vdots \\ P_{n1}(\lambda) \end{bmatrix}\begin{bmatrix} \overline{P}_{11}(\overline{\lambda}) & \cdots & \overline{P}_{n1}(\overline{\lambda}) \end{bmatrix}$$

$$=\vec{x}(\lambda)\,\vec{x}(\overline{\lambda})^*,$$

where $\vec{x}(\lambda)=\begin{bmatrix} P_{11}(\lambda) \\ \vdots \\ P_{n1}(\lambda) \end{bmatrix}\neq\vec{0}$ is an $n\times 1$ column matrix polynomial, Since $\vec{x}(\lambda)=\vec{0}$ implies $M(\lambda)$ is a zero matrix.

Conversely, suppose that $M(\lambda)=\vec{x}(\lambda)(\vec{x}(\overline{\lambda}))^*$ where $\vec{x}(\lambda)\neq\vec{0}$ is an $n\times 1$ column matrix, since $\vec{x}(\lambda)\neq\vec{0}$ then $M(\lambda)\neq 0$.

$$M(\lambda)=\vec{x}(\lambda)I_1(\vec{x}(\overline{\lambda}))^*=P(\lambda)D(P(\overline{\lambda}))^*$$

We choose $P(\lambda)$ as follows:

If $x_1\neq 0$, $P(\lambda)=\begin{bmatrix} x_1 & 0 & \cdots & 0 & 0 \\ x_2 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{n-1} & 0 & \cdots & 1 & 0 \\ x_n & 0 & \cdots & 0 & 1 \end{bmatrix}$, $D=diag[1,0,\cdots,0]$, If $x_1 = 0$ and $x_2 \neq 0$, $P(\lambda) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & x_2 & 0 & \cdots & 0 & 0 \\ 0 & x_3 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & x_{n-1} & 0 & \cdots & 1 & 0 \\ 0 & x_n & 0 & \cdots & 0 & 1 \end{bmatrix}$, $D = diag[0,1,0,\cdots,0]$.

$\vdots$

where $P(\lambda)$ is an $n \times n$ matrix polynomial with columns to be chosen such that $\det P(\lambda) = const. \neq 0$ and D has 1 positive eigenvalue, 0 negative eigenvalue and 0 zero eigenvalue. Therefore $M(\lambda)$ is elementary by Smith canonical form and theorem 3.3.1. □

Theorem 3.3.3: Any selfadjoint $n \times n$ matrix polynomial $L(\lambda)$ admits a representation

$$L(\lambda) = \sum_{j=1}^{m} \varepsilon_j M_j(\lambda)$$

where $\varepsilon_j = \pm 1$ and $M_j(\lambda)$ are elementary matrix polynomial. The number m of terms in $L(\lambda) = \sum_{j=1}^{m} \varepsilon_j M_j(\lambda)$ is greater than or equal to m_0, where m_0 is given by $m_0 = \max_{\lambda \in R} \nu_+(L(\lambda)) + \max_{\lambda \in R} \nu_-(L(\lambda))$. If $m = m_0$, then exactly $\max \nu_+(L(\lambda))$ of ε_j's are equal to $+1$ and exactly $\nu_-(L(\lambda))$ of ε_j's are equal to -1.

Proof:

Theorem 5.5.1 in [11] implies that any Hermitian $n \times n$ matrix is congruent to the matrix $\begin{bmatrix} I_s & 0 & 0 \\ 0 & -I_{r-s} & 0 \\ 0 & 0 & 0_{n-r} \end{bmatrix}$ where r is its rank and s is number of its positive eigenvalues.

Then by theorem 3.3.1, we obtain:

$$L(\lambda) = (M(\bar{\lambda}))^* DM(\lambda)$$

$$= \begin{bmatrix} (M_{11}(\bar{\lambda}))^* & \cdots & (M_{m1}(\bar{\lambda}))^* \\ \vdots & \ddots & \vdots \\ (M_{1n}(\bar{\lambda}))^* & \cdots & (M_{mn}(\bar{\lambda}))^* \end{bmatrix} \begin{bmatrix} \varepsilon_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \varepsilon_m \end{bmatrix} \begin{bmatrix} M_{11}(\lambda) & \cdots & M_{1n}(\lambda) \\ \vdots & \ddots & \vdots \\ M_{m1}(\lambda) & \cdots & M_{mn}(\lambda) \end{bmatrix}$$

$$= \begin{bmatrix} (\vec{x}_1(\bar{\lambda}))^* & \cdots & (\vec{x}_m(\bar{\lambda}))^* \end{bmatrix} diag[\varepsilon_1 \ \cdots \ \varepsilon_m] \begin{bmatrix} \vec{x}_1(\lambda) \\ \vdots \\ \vec{x}_m(\lambda) \end{bmatrix}$$

where $\vec{x}_i(\lambda) = [M_{i1}(\lambda) \ \cdots \ M_{in}(\lambda)]$, $i = 1,2,\ldots,m$

$$L(\lambda) = \varepsilon_1 (\vec{x}_1(\bar{\lambda}))^* \vec{x}_1(\lambda) + \varepsilon_2 (\vec{x}_2(\bar{\lambda}))^* \vec{x}_2(\lambda) + \cdots + \varepsilon_m (\vec{x}_m(\bar{\lambda}))^* \vec{x}_m(\lambda)$$

$$L(\lambda) = \sum_{j=1}^{m} \varepsilon_j M_j(\lambda)$$

where $M_j(\lambda) = (\vec{x}_j(\bar{\lambda}))^* \vec{x}_j(\lambda)$, $j = 1,2,\cdots,m$, each polynomial matrix $M_j(\lambda)$ is elementary by lemma 3.3.2, also $m \geq m_0$.

If $m = m_0$ then D has $\max_{\lambda \in R} v_+(L(\lambda))$ positive eigenvalues, each of which is equal to $+1$ and $\max_{\lambda \in R} v_-(L(\lambda))$ negative eigenvalues each of which is equal to -1. □

Corollary 3.3.1: Any selfadjoint $n \times n$ matrix polynomial admits a factorization $L(\lambda) = (M(\bar{\lambda}))^* DM(\lambda)$, or a representation $L(\lambda) = \sum_{j=1}^{m} \varepsilon_j M_j(\lambda)$, where $m \leq 2n$.

Proof:

Let $m_0 = \max_{\lambda \in R} v_+(L(\lambda)) + \max_{\lambda \in R} v_-(L(\lambda))$, since $\max_{\lambda \in R} v_+(L(\lambda)) \leq n$ and $\max_{\lambda \in R} v_-(L(\lambda)) \leq n$, then $m_0 \leq 2n$.

By theorem 3.3.1, the selfadjoint $n \times n$ matrix polynomial admit a factorization $L(\lambda) = (M(\bar{\lambda}))^* DM(\lambda)$ at $m = m_0 \leq 2n$, also by theorem 3.3.3, $L(\lambda)$ admits a factorization $L(\lambda) = \sum_{j=1}^{m} \varepsilon_j M_j(\lambda)$ at $m = m_0 \leq 2n$. □

There is a selfadjoint matrix polynomial, which doesn't admit a representation $L(\lambda) = (M(\bar{\lambda}))^* DM(\lambda)$ or $L(\lambda) = \sum_{j=1}^{m} \varepsilon_j M_j(\lambda)$ with $m < 2n$.

The matrix polynomial $L(\lambda) = \lambda I$ does not admit a factorization $L(\lambda) = (M(\bar{\lambda}))^* DM(\lambda)$ or $L(\lambda) = \sum_{j=1}^{m} \varepsilon_j M_j(\lambda)$ with $m < 2n$.

$$L(\lambda) = \begin{bmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{bmatrix}$$

$\max_{\lambda \in R} \nu_+(L(\lambda)) = n$ and $\max_{\lambda \in R} \nu_-(L(\lambda)) = n$.

$m_0 = 2n$. The smallest value of $m = 2n$, so $L(\lambda)$ doesn't admit a factorization at $m < 2n$ but at $m = 2n$ or $m > 2n$.

CHAPTER FOUR

JORDAN CHAIN AND DIFFERENTIAL EQUATIONS

The main goal of this chapter is to solve any constant-coefficient-matrix differential equation of order l given by

$$L_l \vec{x}(t)^{(l)} + L_{l-1} \vec{x}(t)^{(l-1)} + \cdots + L_1 \vec{x}(t)^{(1)} + L_0 \vec{x}(t) = \vec{f}(t),$$

where $L_0, L_1, \cdots, L_l \in \mathbf{C}^{n \times n}$, $\det L_l \neq 0$ and the vector function $\vec{f}(t)$ is piecewise continuous. If $L(\lambda) = L_l \lambda^l + L_{l-1} \lambda^{l-1} + \cdots + L_1 \lambda + L_0 = \sum_{i=0}^{l} L_i \lambda^i$, then the last equation can be written as

$$L(\frac{d}{dt}) \vec{x}(t) = \sum_{i=0}^{l} L_i (\frac{d}{dt})^i \vec{x}(t) = \vec{f}(t).$$

First, we begin with the notion of a standard triple and a Jordan chain of a matrix polynomial.

4.1 Standard Triple

If $a(\lambda) = \sum_{j=0}^{l} a_j \lambda^j$ is a monic scalar polynomial then the related $l \times l$ companion matrix is

$$C_a = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ 0 & 0 & \vdots & \ddots & 0 \\ \vdots & \vdots & \vdots & \vdots & 1 \\ -a_0 & -a_1 & \cdots & \cdots & -a_{l-1} \end{bmatrix}$$

while if $L(\lambda) = \sum_{j=0}^{l} L_j \lambda^j$, is a matrix polynomial with $\det L_l \neq 0$ then the related

$\mathrm{ln} \times \mathrm{ln}$ companion matrix is

$$C_L = \begin{bmatrix} 0 & I_n & 0 & \cdots & 0 \\ 0 & 0 & I_n & \ddots & \vdots \\ 0 & 0 & \vdots & \ddots & 0 \\ \vdots & \vdots & \vdots & \vdots & I_n \\ -\hat{L}_0 & -\hat{L}_1 & \cdots & \cdots & -\hat{L}_{l-1} \end{bmatrix}$$

where $\hat{L}_j = L_l^{-1} L_j$ for $j = 0, 1, \cdots, l-1$

Theorem 4.1.1: The matrix $\begin{bmatrix} L(\lambda) & 0 \\ 0 & I_{(l-1)n} \end{bmatrix}$ is equivalent to the matrix $\lambda I_{\mathrm{ln}} - C_L$

Proof:

Suppose that

$$F(\lambda) = \begin{bmatrix} I_n & 0 & 0 & \cdots & 0 \\ -\lambda I & I_n & 0 & \cdots & 0 \\ 0 & -\lambda I & I_n & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & -\lambda I & I_n \end{bmatrix}$$

and

$$E(\lambda) = \begin{bmatrix} B_{l-1}(\lambda) & B_{l-2}(\lambda) & \cdots & B_1(\lambda) & B_0(\lambda) \\ -I_n & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -I_n & 0 \end{bmatrix}$$

where $B_0(\lambda)=L_l$ and $B_{r+1}(\lambda)=\lambda B_r(\lambda)+L_{l-r-1}$, $r=0,1,\cdots,l-2$.

By using Laplace's theorem[1] we find that $\det E(\lambda)=\det L_l$ and $\det F(\lambda)=1$

$$E(\lambda)(\lambda I-C_L)=\begin{bmatrix} L(\lambda) & 0 & 0 & \cdots & 0 \\ -\lambda I & I_n & 0 & \cdots & 0 \\ 0 & -\lambda I & I_n & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & -\lambda I & I_n \end{bmatrix}$$

and

$$\begin{bmatrix} L(\lambda) & 0 \\ 0 & I_{(l-1)n} \end{bmatrix} F(\lambda)=\begin{bmatrix} L(\lambda) & 0 & 0 & \cdots & 0 \\ -\lambda I & I_n & 0 & \cdots & 0 \\ 0 & -\lambda I & I_n & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & -\lambda I & I_n \end{bmatrix}$$

That is

$$E(\lambda)(\lambda I-C_L)=\begin{bmatrix} L(\lambda) & 0 \\ 0 & I_{(l-1)n} \end{bmatrix} F(\lambda).$$

Since $\det F(\lambda)=1$ then $(F(\lambda))^{-1}$ is a matrix polynomial

---

[1] **Laplace's theorem:** Let A denote an arbitrary $n\times n$ matrix and let any p rows (or columns) of A be chosen. Then $\det A$ is equal to the sum of the products of all $\binom{n}{p}$ minors lying in these rows with their corresponding complementary cofactors:

$$\det A=\sum_j A\begin{pmatrix} i_1 & i_2 & \cdots & i_p \\ j_1 & j_2 & \cdots & j_p \end{pmatrix} A^c\begin{pmatrix} i_1 & i_2 & \cdots & i_p \\ j_1 & j_2 & \cdots & j_p \end{pmatrix},$$

where the summation extends over all $\binom{n}{p}$ distinct sets of (column) indices $j_1, j_2, \cdots, j_p$ $(1\le j_1<j_2<\cdots<j_p\le n$. Or, equivalently, using columns

$$\det A=\sum_i A\begin{pmatrix} i_1 & i_2 & \cdots & i_p \\ j_1 & j_2 & \cdots & j_p \end{pmatrix} A^c\begin{pmatrix} i_1 & i_2 & \cdots & i_p \\ j_1 & j_2 & \cdots & j_p \end{pmatrix},$$

where $1\le i_1<i_2<\cdots<i_p\le n$.

$$\begin{bmatrix} L(\lambda) & 0 \\ 0 & I_{(l-1)n} \end{bmatrix} = E(\lambda)(\lambda I - C_L)(F(\lambda))^{-1}. \square$$

From the definition of C_L, it is clear that $\det L(\lambda) = \det(\lambda I - C_L)(\det L_l)$.

In particular, the latent roots of $L(\lambda)$, that is the zeros of $\det L(\lambda)$ and the eigenvalues of C_L are the same.

Theorem 4.1.2: Let $\hat{L}_j = L_j L_l^{-1}$ for $j = 0,1,\cdots,l-1$, and define the $\ln \times \ln$ matrix C_2 by:

$$C_2 = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & -\hat{L}_0 \\ I & 0 & 0 & \cdots & 0 & -\hat{L}_1 \\ 0 & I & 0 & \cdots & 0 & -\hat{L}_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & I & -\hat{L}_{l-1} \end{bmatrix}$$

also define the symmetrizer for $L(\lambda)$ (denoted by S_L) to be

$$S_L = \begin{bmatrix} L_1 & L_2 & L_3 & \cdots & L_{l-1} & L_l \\ L_2 & L_3 & L_4 & \cdots & L_l & 0 \\ L_3 & L_4 & L_5 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ L_l & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

then S_L is nonsingular and $C_2 = S_L C_L S_L^{-1}$.

Proof:

First, we show that S_L is nonsingular.

We shall prove it by mathematical induction,

for $l = 2$

$$S_L = \begin{bmatrix} L_1 & L_2 \\ L_2 & 0 \end{bmatrix}.$$

By applying Laplace's theorem on the columns of L_2, we obtain:

$$\det S_L = \sum_i S_L \begin{pmatrix} i_{n+1} & i_{n+2} & \cdots & i_{2n} \\ n+1 & n+2 & \cdots & 2n \end{pmatrix} S_L^c \begin{pmatrix} i_{n+1} & i_{n+2} & \cdots & i_{2n} \\ n+1 & n+2 & \cdots & 2n \end{pmatrix}$$

$$= S_L \begin{pmatrix} 1 & 2 & \cdots & n \\ n+1 & n+2 & \cdots & 2n \end{pmatrix} (-1)^s S_L \begin{pmatrix} 1 & 2 & \cdots & n \\ n+1 & n+2 & \cdots & 2n \end{pmatrix}^c$$

$$= (-1)^s (\det L_2)^2 \neq 0.$$

Suppose it is true for $l = k$. That is,

$$\det S_L = (-1)^s (\det L_k)^k \neq 0$$

for $l = k+1$

$$S_L = \begin{bmatrix} L_1 & L_2 & \cdots & L_k & L_{k+1} \\ L_2 & L_3 & \cdots & L_{k+1} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ L_{k+1} & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

By applying Laplace's theorem on the columns of L_{k+1}, we obtain:

$$\det S_L = \sum_i S_L \begin{pmatrix} i_{nk+1} & i_{nk+2} & \cdots & i_{n(k+1)} \\ nk+1 & nk+2 & \cdots & n(k+1) \end{pmatrix} S_L^c \begin{pmatrix} i_{nk+1} & i_{nk+2} & \cdots & i_{n(k+1)} \\ nk+1 & nk+2 & \cdots & n(k+1) \end{pmatrix}$$

$$= S_L \begin{pmatrix} 1 & 2 & \cdots & n \\ nk+1 & nk+2 & \cdots & n(k+1) \end{pmatrix} (-1)^s S_L \begin{pmatrix} 1 & 2 & \cdots & n \\ nk+1 & nk+2 & \cdots & n(k+1) \end{pmatrix}^c$$

$$= (-1)^s (\det L_{k+1}) \det(L_{k+1})^k = (-1)^s (\det L_{k+1})^{k+1} \neq 0.$$

Therefore S_L is nonsingular.

$$C_2S_L = \begin{bmatrix} -L_0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & L_2 & L_3 & \cdots & L_{l-1} & L_l \\ 0 & L_3 & L_4 & \cdots & L_l & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & L_l & 0 & \cdots & 0 & 0 \end{bmatrix}$$

and

$$S_LC_L = \begin{bmatrix} -L_0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & L_2 & L_3 & \cdots & L_{l-1} & L_l \\ 0 & L_3 & L_4 & \cdots & L_l & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & L_l & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

That is

$C_2S_L = S_LC_L$ and hence $C_2 = S_LC_LS_L^{-1}$. □

Definition 4.1.1: Three matrices (U,T,V) are said to be **admissible** for the $n \times n$ matrix polynomial $L(\lambda)$ of degree l if they are of sizes $n \times \ln, \ln \times \ln$ and $\ln \times n$, respectively.

Definition 4.1.2: Two admissible triples (U_1,T_1,V_1) and (U_2,T_2,V_2) for $L(\lambda)$ are called **similar** if there is an invertible matrix S such that

$U_1 = U_2S, \quad T_1 = S^{-1}T_2S, \quad V_1 = S^{-1}V_2$

Note that if $P_1 = [I_n \quad 0 \quad \cdots \quad 0]_{n \times \ln}$ and $R_1 = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ L_l^{-1} \end{bmatrix}_{\ln \times n}$, then the triple (P_1,C_L,R_1) is admissible for $L(\lambda)$, where C_L is the companion matrix of $L(\lambda)$.

Definition 4.1.3: Any triple similar to (P_1, C_L, R_1) is said to be a **standard triple** for $L(\lambda)$.

In the following theorem, we give another characterization of a standard triple.

Theorem 4.1.3: An admissible triple (U,T,V) for $L(\lambda)$ is a standard triple if and only if the following conditions are satisfied:

(i) The matrix $Q = \begin{bmatrix} U \\ UT \\ \vdots \\ UT^{l-1} \end{bmatrix}$ is nonsingular,

(ii) $L_l UT^l + L_{l-1} UT^{l-1} + \cdots + L_1 UT + L_0 U = 0$,

(iii) $V = Q^{-1} R_1$.

Proof:

Suppose that (U,T,V) is a standard triple for $L(\lambda)$

$$\begin{bmatrix} P_1 \\ P_1 C_L \\ \vdots \\ P_1 C_L^{l-1} \end{bmatrix} = I_{\text{ln}}$$

$$\begin{bmatrix} US^{-1} \\ UTS^{-1} \\ \vdots \\ UT^{l-1}S^{-1} \end{bmatrix} = \begin{bmatrix} U \\ UT \\ \vdots \\ UT^{l-1} \end{bmatrix} S^{-1} = QS^{-1} = I_{\text{ln}}.$$

Therefore Q is nonsingular and $Q = S$

$$L_l P_1 C_L^l = L_l (P_1 C_L^{l-1}) C_L = L_l \begin{bmatrix} 0 & \cdots & 0 & I \end{bmatrix} C_L = \begin{bmatrix} -L_0 & -L_1 & \cdots & -L_{l-1} \end{bmatrix} \qquad (4.1.1)$$

$$\sum_{j=0}^{l-1} L_j P_1 C_L^j = [L_0 \quad L_1 \quad \cdots \quad L_{l-1}] \begin{bmatrix} P_1 \\ P_1 C_L \\ \vdots \\ P_1 C_L^{l-1} \end{bmatrix} = [L_0 \quad L_1 \quad \cdots \quad L_{l-1}] \qquad (4.1.2)$$

If we add equation (4.1.1) to equation (4.1.2), then we obtain: $\sum_{j=0}^{l} L_j P_1 C_L^j = 0$

$$U = P_1 S, \quad T = S^{-1} C_L S, \quad V = S^{-1} R_1$$

$$\sum_{j=0}^{l} L_j U T^j = (\sum_{j=0}^{l} L_j P_1 C_L^j) S = 0 .$$

Therefore condition (ii) is satisfied.

$$QV = SS^{-1} R_1 = R_1$$

since Q is nonsingular from condition (i), then $V = Q^{-1} R_1$. Therefore condition (iii) is satisfied.

Conversely, suppose (U,T,V) is admissible for $L(\lambda)$ and satisfies conditions $(i),\ (ii)$ and (iii).

$$P_1 Q = U \text{ , } QT = \begin{bmatrix} UT \\ UT^2 \\ \vdots \\ UT^l \end{bmatrix} \text{ and } C_L Q = \begin{bmatrix} UT \\ UT^2 \\ \vdots \\ UT^l \end{bmatrix},$$

since $QT = C_L Q$ and Q is nonsingular, then $T = Q^{-1} C_L Q$. Therefore (U,T,V) is a standard triple. □

Definition 4.1.4: A standard pair for $L(\lambda)$ is the first two members of any standard triple for $L(\lambda)$.

The following corollary follows from the last theorem.

Corollary 4.1.1: (U,T) is a standard pair for $L(\lambda)$ if and only if the followings hold:

(i) $Q=\begin{bmatrix} U \\ UT \\ \vdots \\ UT^{l-1} \end{bmatrix}$ is nonsingular.

(ii) $L_l UT^l + L_{l-1}UT^{l-1} + \cdots + L_1 UT + L_0 U = 0$. □

Theorem 4.1.4: If (U,T,V) is a standard triple for $L(\lambda)$, S_L is the symmetrizer for $L(\lambda)$, and if Q and R are defined by

$$Q=\begin{bmatrix} U \\ UT \\ \vdots \\ UT^{l-1} \end{bmatrix}, \quad R=\begin{bmatrix} V & TV & \cdots & T^{l-1}V \end{bmatrix},$$

then $RS_L Q = I$.

Proof:

By theorem 4.1.2 we get

$C_2 = S_L C_L S_L^{-1}$ and $C_L = S_L^{-1} C_2 S_L$

$C_L^j R_1 = S_L^{-1} C_2^j (S_L R_1) = S_L^{-1} C_2^j P_1^T$ where C_L^j is the jth power of C_L

$\begin{bmatrix} R_1 & C_L R_1 & \cdots & C_L^{l-1} R_1 \end{bmatrix} = S_L^{-1} \begin{bmatrix} P_1^T & C_2 P_1^T & \cdots & C_2^{l-1} P_1^T \end{bmatrix} = S_L^{-1}$.

Since (U,T,V) is a standard triple, then there is a nonsingular matrix S such that $U = P_1 S$, $T = S^{-1} C_L S$, $V = S^{-1} R_1$, then

$$Q=\begin{bmatrix} U \\ UT \\ \vdots \\ UT^{l-1} \end{bmatrix}=\begin{bmatrix} P_1 \\ P_1C_L \\ \vdots \\ P_1C_L^{l-1} \end{bmatrix}S=S,$$

and

$$R=S^{-1}\begin{bmatrix} R_1 & C_LR_1 & \cdots & C_L^{l-1}R_1 \end{bmatrix}=S^{-1}S_L^{-1}.$$

Hence,

$RS_LQ=(S^{-1}S_L^{-1})S_LS=I$. □

4.2 Jordan Chain

In this section we define the Jordan chain and introduce the notion of Jordan pair for a matrix polynomial $L(\lambda)$. The Jordan chain of a constant matrix can be found in appendix A.

Definition 4.2.1: A standard pair (X,J) is called a Jordan pair if J is in Jordan canonical form.

Since any constant matrix is similar to a matrix J in Jordan canonical form, then we conclude that the $\ln\times\ln$ matrix C_L is similar to a matrix J in Jordan canonical form, that is there exists a nonsingular matrix S such that $J=S^{-1}C_LS$. Let $X=P_1S$ then (X,J) is a Jordan pair of the matrix polynomial $L(\lambda)$.

Now, suppose that (X,J) is a Jordan pair, where J is a diagonal matrix and $\vec{x}_j \in \mathbf{C}^n$ is the jth column of X for $j = 1,2,\cdots,\ln$, then from theorem 4.1.3 we get

$$Q = \begin{bmatrix} X \\ XJ \\ \vdots \\ XJ^{l-1} \end{bmatrix}$$

is nonsingular matrix, and

$$L_l XJ^l + L_{l-1} XJ^{l-1} + \cdots + L_1 XJ + L_0 X = 0 \qquad (4.2.1)$$

If $\vec{x}_j = \vec{0}$ for some $j = 1,2,\cdots,\ln$ then Q has a column of zeros, and this contradicts the fact that Q is nonsingular matrix. Therefore $\vec{x}_j \neq \vec{0}$ for all $j = 1,2,\cdots,\ln$.

Definition: 4.2.2: If $\vec{x}$ is a nonzero vector such that $L(\lambda_j)\vec{x} = \vec{0}$ ($i.e.$ $\vec{x} \in \ker L(\lambda_j)$). Then $\vec{x}$ is called a **latent vector** of $L(\lambda)$ corresponding to the **latent root** λ_j.

Note that λ_j is a latent root of $L(\lambda)$ if and only if $\det L(\lambda_j) = 0$.

We note that if (X,J) is a Jordan pair and J is a diagonal matrix then every column of X is a latent vector of $L(\lambda)$. For: Pick out the jth column of each term of equation (4.2.1), then we obtain

$$L_l \lambda_j^l \vec{x}_j + L_{l-1} \lambda_j^{l-1} \vec{x}_j + \cdots + L_1 \lambda_j \vec{x}_j + L_0 \vec{x}_j = \vec{0}$$

That is, $L(\lambda_j)\vec{x}_j = \vec{0}$ and since $\vec{x}_j \neq \vec{0}$, then $\vec{x}_j$ (jth column of X) is a latent vector of $L(\lambda)$ corresponding to the latent root λ_j.

Definition 4.2.3: The set of vectors $\vec{x}_0, \vec{x}_1, \cdots, \vec{x}_k$, with $\vec{x}_0 \neq \vec{0}$, is a **Jordan chain** of length $k+1$ for $L(\lambda)$ corresponding to the latent root λ_0 if the following $k+1$ relations hold

$$L(\lambda_0)\vec{x}_0 = \vec{0}$$

$$L(\lambda_0)\vec{x}_1 + \frac{1}{1!}L^{(1)}(\lambda_0)\vec{x}_0 = \vec{0}$$

$$L(\lambda_0)\vec{x}_2 + \frac{1}{1!}L^{(1)}(\lambda_0)\vec{x}_1 + \frac{1}{2!}L^{(2)}(\lambda_0)\vec{x}_0 = \vec{0}$$

$$\vdots$$

$$L(\lambda_0)\vec{x}_k + \cdots + \frac{1}{(k-2)!}L^{(k-2)}(\lambda_0)\vec{x}_2 + \frac{1}{(k-1)!}L^{(k-1)}(\lambda_0)\vec{x}_1 + \frac{1}{k!}L^{(k)}(\lambda_0)\vec{x}_0 = \vec{0}$$

where $L^{(i)}(\lambda_0)$ means the ith derivative of the matrix polynomial $L(\lambda)$ at λ_0.

In the following, we show that if (X, J) is a Jordan pair of the $n \times n$ matrix polynomial $L(\lambda)$, then each column of X is made up of Jordan chain for $L(\lambda)$:

Let $J = diag[J_1, J_2, \cdots, J_s]$, where J_j is a Jordan block of size n_j, for $j = 1,2,\cdots,s$.

Form a partition $X = [X_1 \quad X_2 \quad \cdots \quad X_s]$, where X_j is $n \times n_j$ for $j = 1,2,\cdots,s$.

Now for $r = 0,1,2,\cdots$, we obtain

$$XJ^r = [X_1 J_1^r \quad X_2 J_2^r \quad \cdots \quad X_s J_s^r]$$

Thus,

$$0 = L_l XJ^l + L_{l-1} XJ^{l-1} + \cdots + L_1 XJ + L_0 X ,$$

$$0 = L_l \begin{bmatrix} X_1 J_1^l & X_2 J_2^l & \cdots & X_s J_s^l \end{bmatrix} + \cdots + L_0 \begin{bmatrix} X_1 & X_2 & \cdots & X_s \end{bmatrix},$$

$$0 = \begin{bmatrix} L_l X_1 J_1^l & L_l X_2 J_2^l & \cdots & L_l X_s J_s^l \end{bmatrix} + \cdots + \begin{bmatrix} L_0 X_1 & L_0 X_2 & \cdots & L_0 X_s \end{bmatrix}$$

this implies:

$$L_l X_1 J_1^l + L_{l-1} X_1 J_1^{l-1} + \cdots + L_1 X_1 J_1 + L_0 X_1 = 0$$

$$L_l X_2 J_2^l + L_{l-1} X_2 J_2^{l-1} + \cdots + L_1 X_2 J_2 + L_0 X_2 = 0$$

$$\vdots$$

$$L_l X_s J_s^l + L_{l-1} X_s J_s^{l-1} + \cdots + L_1 X_s J_s + L_0 X_s = 0 .$$

That is for $j = 1,2,\cdots,s$

$$L_l X_j J_j^l + L_{l-1} X_j J_j^{l-1} + \cdots + L_1 X_j J_j + L_0 X_j = 0 \tag{4.2.2}$$

Let the columns of X_j be denoted by $\vec{x}_1^{\,j}, \vec{x}_2^{\,j}, \cdots, \vec{x}_{n_j}^{\,j}$, that is

$$X_j = \begin{bmatrix} \vec{x}_1^{\,j} & \vec{x}_2^{\,j} & \cdots & \vec{x}_{n_j}^{\,j} \end{bmatrix}.$$

Let $\binom{r}{p} \lambda_j^{r-p} = 0$ if $r < p$, then

$$J_j^r = \begin{bmatrix} \lambda_j^r & \binom{r}{1}\lambda_j^{r-1} & \cdots & \binom{r}{n_j - 1}\lambda_j^{r-n_j+1} \\ 0 & \lambda_j^r & \cdots & \binom{r}{n_j - 2}\lambda_j^{r-n_j+2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_j^r \end{bmatrix}$$

using (4.2.2) and examining the columns in the order $1,2,\cdots,n_j$, a chain of the form in definition 4.2.3 is obtained with λ_0 replaced by λ_j.

$$L(\lambda_j)\vec{x}_1^{\,j} = \vec{0}$$

$$L(\lambda_j)\vec{x}_2^{\,j} + \frac{1}{1!}L^{(1)}(\lambda_j)\vec{x}_1^{\,j} = \vec{0}$$

$\vdots$

$$L(\lambda_j)\vec{x}_{n_j}^{\,j} + \frac{1}{1!}L^{(1)}(\lambda_j)\vec{x}_{n_j-1}^{\,j} + \cdots + \frac{1}{(n_j-1)!}L^{(n_j-1)}(\lambda_j)\vec{x}_1^{\,j} = \vec{0}.$$

Thus, for each j, the columns of X_j form a Jordan chain of length n_j corresponding to the latent root λ_j.

Example 4.2.1: Consider the following matrix polynomial

$$L(\lambda) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\lambda^2 + \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}\lambda + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \lambda^2 & 0 \\ \lambda+1 & \lambda(\lambda-1) \end{bmatrix}.$$

Then,

$$\det L(\lambda) = \lambda^3(\lambda-1) = 0$$

The latent roots are: $\lambda_1 = 0$ and $\lambda_2 = 1$.

$\begin{bmatrix} 0 \\ \alpha \end{bmatrix}$ is a latent vector corresponding to $\lambda_2 = 1$.

$\vec{x}_0 = \begin{bmatrix} 0 \\ \alpha \end{bmatrix}$ is a latent vector corresponding to $\lambda_1 = 0$. Now we find $\vec{x}_1$:

$$L(0)\vec{x}_1 + L^{(1)}(0)\vec{x}_0 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\vec{x}_1 + \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 0 \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Then $\vec{x}_1 = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$, where β is any complex number.

To find $\vec{x}_2$:

$$L(0)\vec{x}_2+L^{(1)}(0)\vec{x}_1+\frac{1}{2}L^{(2)}(0)\vec{x}_0=\begin{bmatrix}0&0\\1&0\end{bmatrix}\vec{x}_2+\begin{bmatrix}0&0\\1&-1\end{bmatrix}\begin{bmatrix}\alpha\\\beta\end{bmatrix}+\begin{bmatrix}1&0\\0&1\end{bmatrix}\begin{bmatrix}0\\\alpha\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$$

Then $\vec{x}_2=\begin{bmatrix}-2\alpha+\beta\\\gamma\end{bmatrix}$, where γ is any complex number.

Take $\alpha=1$, $\beta=\gamma=0$ then $\vec{x}_0=\begin{bmatrix}0\\1\end{bmatrix}$, $\vec{x}_1=\begin{bmatrix}1\\0\end{bmatrix}$, $\vec{x}_2=\begin{bmatrix}-2\\0\end{bmatrix}$.

Therefore $X=\begin{bmatrix}0&0&1&-2\\1&1&0&0\end{bmatrix}$ and $J=\begin{bmatrix}1&0&0&0\\0&0&1&0\\0&0&0&1\\0&0&0&0\end{bmatrix}$.

Example 4.2.2: To find a Jordan pair for the matrix polynomial $L(\lambda)=\begin{bmatrix}\lambda^2-2\lambda-2&\lambda+2\\\lambda+2&\lambda^2-2\lambda-2\end{bmatrix}$, we first find $\det L(\lambda)$.

$\det L(\lambda)=(\lambda^2-2\lambda-2)^2-(\lambda+2)^2=0$.

The latent roots of $L(\lambda)$ are $-1,0,1,4$.

$L(-1)=\begin{bmatrix}1&1\\1&1\end{bmatrix}$, then $\begin{bmatrix}1\\-1\end{bmatrix}$ is a latent vector corresponding to $\lambda_1=-1$,

$L(0)=\begin{bmatrix}-2&2\\2&-2\end{bmatrix}$, then $\begin{bmatrix}1\\1\end{bmatrix}$ is a latent vector corresponding to $\lambda_2=0$,

$L(1)=\begin{bmatrix}-3&3\\3&-3\end{bmatrix}$, then $\begin{bmatrix}1\\1\end{bmatrix}$ is a latent vector corresponding to $\lambda_3=1$,

$L(4)=\begin{bmatrix}6&6\\6&6\end{bmatrix}$, then $\begin{bmatrix}1\\-1\end{bmatrix}$ is a latent vector corresponding to $\lambda_4=4$.

Therefore $X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & -1 \end{bmatrix}$, $J = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$.

4.3 Homogeneous Differential Equations

Consider the constant-coefficient-matrix differential equation of order l given by

$$L_l \vec{x}(t)^{(l)} + L_{l-1} \vec{x}(t)^{(l-1)} + \cdots + L_1 \vec{x}(t)^{(1)} + L_0 \vec{x}(t) = \vec{0} \qquad (4.3.1)$$

where $L_0, L_1, \cdots, L_l \in \mathbf{C}^{n \times n}$ and $\det L_l \neq 0$.

If we consider the matrix polynomial $L(\lambda) = \sum_{j=0}^{l} L_j \lambda^j$ with $\det L_l \neq 0$, then equation (4.3.1) can be abbreviated to the form:

$$L(\frac{d}{dt}) \vec{x}(t) = \sum_{j=0}^{l} L_j \frac{d^j \vec{x}}{dt^j} = \vec{0}.$$

In the next theorem we see that, the properties of the solution of equation (4.3.1) are closely related to the spectral properties of $L(\lambda)$.

Theorem 4.3.1: If $\det L(\lambda) \not\equiv 0$ then the solution space of equation (4.3.1) has dimension equal to the degree of $\det L(\lambda)$.

Proof:

Since $\det L(\lambda) \not\equiv 0$ then $L(\lambda) = E(\lambda)D(\lambda)F(\lambda)$, where $E(\lambda), F(\lambda)$ have nonzero constant determinants and $D(\lambda) = diag[i_1(\lambda), i_2(\lambda), \cdots, i_n(\lambda)]$.

Equation (4.3.1) can be written in the form:

$$L(\frac{d}{dt})\vec{x}(t) = E(\frac{d}{dt})D(\frac{d}{dt})F(\frac{d}{dt})\vec{x}(t) = \vec{0}.$$

If we let $\vec{y}(t) = F(\frac{d}{dt})\vec{x}(t)$ and multiply the above equation from the left by $E^{-1}(\frac{d}{dt})$ then we obtain:

$$diag\left[i_1(\frac{d}{dt}), i_2(\frac{d}{dt}), \cdots, i_n(\frac{d}{dt})\right]\vec{y}(t) = \vec{0}.$$

That is, $i_k(\frac{d}{dt})y_k(t) = 0$, for $k = 1, 2, \cdots, n$.

The scalar differential equation $i_k(\frac{d}{dt})y_k(t) = 0$ has exactly $d_k = \deg i_k(\lambda)$ linearly independent solutions. Thus the dimension of the solution space of equation (4.3.1) is $d_1 + d_2 + \cdots + d_n$, where $d_k = \deg i_k(\lambda), \quad k = 1, 2, \cdots, n$.

From the Smith form: we conclude that

$$\det L(\lambda) = \gamma \det D(\lambda) = \gamma \prod_{k=1}^{n} i_k(\lambda),$$

where $\gamma = \det E(\lambda) \det F(\lambda)$. Hence

$$\deg(\det L(\lambda)) = \deg(\det D(\lambda)) = \deg(\prod_{k=1}^{n} i_k(\lambda)) = \sum_{k=1}^{n} \deg i_k(\lambda) = \sum_{k=1}^{n} d_k.$$

Now, $\vec{x}(t)$ can be easily found from $\vec{x}(t) = F^{-1}(\frac{d}{dt})\vec{y}(t)$.

Therefore, the dimension of the solution space of equation (4.3.1) is equal to the degree of $\det L(\lambda)$. □

Example 4.3.1: Consider the system $\frac{dx_1(t)}{dt} + x_2(t) = 0$,

$$x_1(t) + \frac{d^2 x_2(t)}{dt^2} = 0.$$

The corresponding matrix polynomial is $L(\lambda) = \begin{bmatrix} \lambda & 1 \\ 1 & \lambda^2 \end{bmatrix}$.

By computing the Smith canonical form for a matrix polynomial $L(\lambda)$, we obtain:

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

Then, the system becomes:

$$L(\frac{d}{dt}) \vec{x}(t) = \begin{bmatrix} 1 & 0 \\ \frac{d^2}{dt^2} & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{d^3}{dt^3} - 1 \end{bmatrix} \begin{bmatrix} \frac{d}{dt} & 1 \\ -1 & 0 \end{bmatrix} \vec{x}(t) = \vec{0},$$

where $\vec{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$. Let $\vec{y}(t) = \begin{bmatrix} \frac{d}{dt} & 1 \\ -1 & 0 \end{bmatrix} \vec{x}(t)$ and multiply the above equation from

the left by $\begin{bmatrix} 1 & 0 \\ \frac{d^2}{dt^2} & 1 \end{bmatrix}^{-1}$ then we obtain:

$$\begin{bmatrix} 1 & 0 \\ 0 & \frac{d^3}{dt^3} - 1 \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

where $\vec{y}(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix}$

Then

$y_1(t) = 0,$

$$\frac{d^3 y_2(t)}{dt^3} - y_2(t) = 0.$$

The linearly independent solutions of the last equation are:

$$y_{2,1}(t) = e^t, \; y_{2,2}(t) = e^{\frac{-1+i\sqrt{3}}{2}t}, \; y_{2,3}(t) = e^{\frac{-1-i\sqrt{3}}{2}t}.$$

The general solution is

$$y_2(t) = c_1 e^t + c_2 e^{\frac{-1+i\sqrt{3}}{2}t} + c_3 e^{\frac{-1-i\sqrt{3}}{2}t}$$

$$= c_1 e^t + c_2 e^{-\frac{1}{2}t} e^{i\frac{1}{2}\sqrt{3}t} + c_3 e^{-\frac{1}{2}t} e^{-i\frac{1}{2}\sqrt{3}t}$$

$$= c_1 e^t + c_2 e^{-\frac{1}{2}t}(\cos\frac{1}{2}\sqrt{3}t + i\sin\frac{1}{2}\sqrt{3}t) + c_3 e^{-\frac{1}{2}t}(\cos-\frac{1}{2}\sqrt{3}t + i\sin-\frac{1}{2}\sqrt{3}t$$

$$= c_1 e^t + (c_2 + c_3)e^{-\frac{1}{2}t}\cos\frac{1}{2}\sqrt{3}t + (ic_2 - ic_3)e^{-\frac{1}{2}t}\sin\frac{1}{2}\sqrt{3}t$$

$$= C_1 e^t + C_2 e^{-\frac{1}{2}t}\cos\frac{1}{2}\sqrt{3}t + C_3 e^{-\frac{1}{2}t}\sin\frac{1}{2}\sqrt{3}t,$$

where $C_1 = c_1$, $C_2 = c_2 + c_3$ and $C_3 = ic_2 - ic_3$.

Thus, the linearly independent solutions of the system are:

$$\begin{bmatrix} -e^t \\ e^t \end{bmatrix}, \begin{bmatrix} -e^{\frac{-1+i\sqrt{3}}{2}t} \\ \frac{-1+i\sqrt{3}}{2}e^{\frac{-1+i\sqrt{3}}{2}t} \end{bmatrix}, \begin{bmatrix} -e^{\frac{-1-i\sqrt{3}}{2}t} \\ \frac{-1-i\sqrt{3}}{2}e^{\frac{-1-i\sqrt{3}}{2}t} \end{bmatrix}.$$

Theorem 4.3.2: Let $\vec{x}_0, \vec{x}_1, \vec{x}_2, \cdots, \vec{x}_{k-1}$ be a Jordan chain for $L(\lambda)$ at λ_0.

Then the k functions

$$\vec{u}_0(t) = \vec{x}_0 e^{\lambda_0 t}$$

$$\vec{u}_1(t) = (t\vec{x}_0 + \vec{x}_1)e^{\lambda_0 t}$$

$$\vec{u}_2(t) = (\frac{t^2}{2!}\vec{x}_0 + t\vec{x}_1 + \vec{x}_2)e^{\lambda_0 t}$$

$$\vdots$$

$$\vec{u}_{k-1}(t) = (\sum_{j=0}^{k-1}\frac{t^j}{j!}\vec{x}_{k-1-j})e^{\lambda_0 t}$$

are linearly independent solutions of equation (4.3.1).

Proof:

Since $(\frac{d}{dt})^j(\vec{u}_0(t)) = (\frac{d}{dt})^j\vec{x}_0 e^{\lambda_0 t} = \lambda_0^j\vec{x}_0 e^{\lambda_0 t}$, then

$$L(\frac{d}{dt})(\vec{u}_0(t)) = \sum_{j=0}^{l} L_j(\frac{d}{dt})^j(\vec{x}_0 e^{\lambda_0 t}) = L(\lambda_0)\vec{x}_0 e^{\lambda_0 t} = \vec{0}.$$

Therefore $\vec{u}_0(t) = \vec{x}_0 e^{\lambda_0 t}$ is a solution of equation (4.3.1). Also

$$(\frac{d}{dt} - \lambda_0 I)\vec{u}_0(t) = (\frac{d}{dt} - \lambda_0 I)\vec{x}_0 e^{\lambda_0 t} = \vec{0},$$

$$(\frac{d}{dt} - \lambda_0 I)\vec{u}_1(t) = (\frac{d}{dt} - \lambda_0 I)(t\vec{x}_0 + \vec{x}_1)e^{\lambda_0 t} = \vec{x}_0 e^{\lambda_0 t}.$$

That is.

$$(\frac{d}{dt} - \lambda_0 I)\vec{u}_1(t) = \vec{u}_0(t),$$

and hence

$$(\frac{d}{dt} - \lambda_0 I)^2\vec{u}_1(t) = \vec{0}.$$

By using Taylor expansion of $L(\lambda)$ about λ_0, we obtain:

$$L(\lambda) = L(\lambda_0) + \frac{1}{1!}L^{(1)}(\lambda_0)(\lambda - \lambda_0) + \frac{1}{2!}L^{(2)}(\lambda_0)(\lambda - \lambda_0)^2 + \cdots + \frac{1}{l!}L^{(l)}(\lambda_0)(\lambda - \lambda_0)^l$$

$$L(\frac{d}{dt}) = L(\lambda_0) + \frac{1}{1!}L^{(1)}(\lambda_0)(\frac{d}{dt} - \lambda_0 I) + \frac{1}{2!}L^{(2)}(\lambda_0)(\frac{d}{dt} - \lambda_0 I)^2 + \cdots + \frac{1}{l!}L^{(l)}(\lambda_0)(\frac{d}{dt} - \lambda_0 I)^l$$

Since $(\frac{d}{dt} - \lambda_0 I)^j \vec{u}_1(t) = \vec{0}$ for $j = 2,3,\cdots$, then

$$L(\frac{d}{dt})\vec{u}_1(t) = L(\lambda_0)\vec{u}_1(t) + \frac{1}{1!}L^{(1)}(\lambda_0)(\frac{d}{dt} - \lambda_0 I)\vec{u}_1(t)$$

$$= L(\lambda_0)\vec{u}_1(t) + \frac{1}{1!}L^{(1)}(\lambda_0)\vec{u}_0(t)$$

$$= L(\lambda_0)(t\vec{x}_0 + \vec{x}_1)e^{\lambda_0 t} + \frac{1}{1!}L^{(1)}(\lambda_0)\vec{x}_0 e^{\lambda_0 t}$$

$$= (L(\lambda_0)\vec{x}_0)te^{\lambda_0 t} + (L(\lambda_0)\vec{x}_1 + \frac{1}{1!}L^{(1)}(\lambda_0)\vec{x}_0)e^{\lambda_0 t} = \vec{0}.$$

Therefore $\vec{u}_1(t) = (t\vec{x}_0 + \vec{x}_1)e^{\lambda_0 t}$ is a solution of equation (4.3.1).

$$(\frac{d}{dt} - \lambda_0 I)\vec{u}_1(t) = (\frac{d}{dt} - \lambda_0 I)(t\vec{x}_0 + \vec{x}_1)e^{\lambda_0 t} = \vec{u}_0(t)$$

$$(\frac{d}{dt} - \lambda_0 I)\vec{u}_2(t) = (\frac{d}{dt} - \lambda_0 I)(\frac{1}{2!}t^2\vec{x}_0 + t\vec{x}_1 + \vec{x}_2)e^{\lambda_0 t} = (t\vec{x}_0 + \vec{x}_1)e^{\lambda_0 t}$$

That is

$$(\frac{d}{dt} - \lambda_0 I)\vec{u}_2(t) = \vec{u}_1(t)$$

Hence

$$(\frac{d}{dt} - \lambda_0 I)^3\vec{u}_2(t) = \vec{0}$$

By using Taylor expansion of $L(\lambda)$ about λ_0, we obtain:

$$L(\lambda) = L(\lambda_0) + \frac{1}{1!}L^{(1)}(\lambda_0)(\lambda - \lambda_0) + \frac{1}{2!}L^{(2)}(\lambda_0)(\lambda - \lambda_0)^2 + \frac{1}{3!}L^{(3)}(\lambda_0)(\lambda - \lambda_0)^3$$

$$+ \cdots + \frac{1}{l!}L^{(l)}(\lambda_0)(\lambda - \lambda_0)^l$$

$$L(\frac{d}{dt}) = L(\lambda_0) + \frac{1}{1!}L^{(1)}(\lambda_0)(\frac{d}{dt} - \lambda_0 I) + \frac{1}{2!}L^{(2)}(\lambda_0)(\frac{d}{dt} - \lambda_0 I)^2 + \frac{1}{3!}L^{(3)}(\lambda_0)(\frac{d}{dt} - \lambda_0 I)^3$$

$$+ \cdots + \frac{1}{l!}L^{(l)}(\lambda_0)(\frac{d}{dt} - \lambda_0 I)^l$$

Since

$$(\frac{d}{dt} - \lambda_0 I)^j \vec{u}_2(t) = \vec{0} \text{ for } j = 3,4,\cdots,$$

Then

$$L(\frac{d}{dt})\vec{u}_2(t)$$

$$= L(\lambda_0)\vec{u}_2(t) + \frac{1}{1!}L^{(1)}(\lambda_0)(\frac{d}{dt} - \lambda_0 I)\vec{u}_2(t) + \frac{1}{2!}L^{(2)}(\lambda_0)(\frac{d}{dt} - \lambda_0 I)^2 \vec{u}_2(t)$$
$$= L(\lambda_0)\vec{u}_2(t) + \frac{1}{1!}L^{(1)}(\lambda_0)\vec{u}_1(t) + \frac{1}{2!}L^{(2)}(\lambda_0)\vec{u}_0(t)$$
$$= L(\lambda_0)(\frac{1}{2!}t^2 \vec{x}_0 + t\vec{x}_1 + \vec{x}_2)e^{\lambda_0 t} + \frac{1}{1!}L^{(1)}(\lambda_0)(t\vec{x}_0 + \vec{x}_1)e^{\lambda_0 t} + \frac{1}{2!}L^{(2)}(\lambda_0)\vec{x}_0 e^{\lambda_0 t}$$

$$= \frac{1}{2!}(L(\lambda_0)\vec{x}_0)t^2 e^{\lambda_0 t} + (L(\lambda_0)\vec{x}_1 + \frac{1}{1!}L^{(1)}(\lambda_0)\vec{x}_0)te^{\lambda_0 t} + (L(\lambda_0)\vec{x}_2 +$$

$$\frac{1}{1!}L^{(1)}(\lambda_0)\vec{x}_1 + \frac{1}{2!}L^{(2)}(\lambda_0)\vec{x}_0)e^{\lambda_0 t} = \vec{0}.$$

Therefore $\vec{u}_2(t) = (\frac{1}{2!}t^2 \vec{x}_0 + t\vec{x}_1 + \vec{x}_2)e^{\lambda_0 t}$ is a solution of equation (4.3.1).

In this manner we deduce that:

$\vec{u}_{k-1}(t) = (\sum_{j=0}^{k-1} \frac{t^j}{j!} \vec{x}_{k-1-j})e^{\lambda_0 t}$ is a solution of equation (4.3.1).

Now, we show that

$\vec{u}_0(t),\ \vec{u}_1(t),\ ,\vec{u}_2(t),\cdots,\ \vec{u}_{k-1}(t)$ are linearly independent: If

$$\vec{0} = \alpha_0 \vec{u}_0(t) + \alpha_1 \vec{u}_1(t) + \alpha_2 \vec{u}_2(t) + \cdots + \alpha_{k-1} \vec{u}_{k-1}(t)$$

$$\vec{0} = \alpha_0 \vec{x}_0 e^{\lambda_0 t} + \alpha_1 (t\vec{x}_0 + \vec{x}_1) e^{\lambda_0 t} + \alpha_2 (\frac{1}{2!} t^2 \vec{x}_0 + t\vec{x}_1 + \vec{x}_2) e^{\lambda_0 t} + \cdots +$$

$$\alpha_{k-1} (\frac{1}{(k-1)!} t^{k-1} \vec{x}_0 + \frac{1}{(k-2)!} t^{k-2} \vec{x}_1 + \frac{1}{(k-3)!} t^{k-3} \vec{x}_2 + \cdots + \vec{x}_{k-1}) e^{\lambda_0 t}.$$

By equating the coefficients of powers of t and canceling $e^{\lambda_0 t}$, since $e^{\lambda_0 t} \neq 0$, we obtain

$$\frac{1}{(k-1)!} \alpha_{k-1} \vec{x}_0 = \vec{0}$$

and $\alpha_{k-1} = 0$, since $\vec{x}_0 \neq \vec{0}$.

$$\frac{1}{(k-2)!} \alpha_{k-1} \vec{x}_1 + \frac{1}{(k-2)!} \alpha_{k-2} \vec{x}_0 = \frac{1}{(k-2)!} \alpha_{k-2} \vec{x}_0 = \vec{0}$$

$\alpha_{k-2} = 0$, since $\vec{x}_0 \neq \vec{0}$.

$$\frac{1}{(k-3)!} \alpha_{k-1} \vec{x}_2 + \frac{1}{(k-3)!} \alpha_{k-2} \vec{x}_1 + \frac{1}{(k-3)!} \alpha_{k-3} \vec{x}_0 = \frac{1}{(k-3)!} \alpha_{k-3} \vec{x}_0 = \vec{0}$$

$\alpha_{k-3} = 0$, since $\vec{x}_0 \neq \vec{0}$.

In this manner we deduce that:

$\alpha_0 = \alpha_1 = \alpha_2 = \cdots = \alpha_{k-1} = 0$.

Therefore $\vec{u}_0(t)$, $\vec{u}_1(t)$, $\vec{u}_2(t)$, $\cdots$, $\vec{u}_{k-1}(t)$ are linearly independent. □

Theorem 4.3.3: Let $\vec{x}_0, \vec{x}_1, \cdots, \vec{x}_{k-1} \in \mathbf{C}^n$ with $\vec{x}_0 \neq \vec{0}$. If the vector-valued function $\vec{u}_{k-1}(t) = (\frac{t^{k-1}}{(k-1)!}\vec{x}_0 + \frac{t^{k-2}}{(k-2)!}\vec{x}_1 + \cdots + t\vec{x}_{k-2} + \vec{x}_{k-1})e^{\lambda_0 t}$ is a solution of equation (4.3.1), then λ_0 is a latent root of $L(\lambda)$ and $\vec{x}_0, \vec{x}_1, \cdots, \vec{x}_{k-1}$ is a Jordan chain for $L(\lambda)$ corresponding to λ_0.

Proof:

Let $\vec{u}_0(t) = \vec{x}_0 e^{\lambda_0 t}$

$$\vec{u}_1(t) = (t\vec{x}_0 + \vec{x}_1)e^{\lambda_0 t}$$

$$\vec{u}_2(t) = (\frac{1}{2!}t^2\vec{x}_0 + t\vec{x}_1 + \vec{x}_2)e^{\lambda_0 t}$$

$$\vdots$$

$$\vec{u}_{k-2}(t) = (\frac{t^{k-2}}{(k-2)!}\vec{x}_0 + \frac{t^{k-3}}{(k-3)!}\vec{x}_1 + \cdots + t\vec{x}_{k-3} + \vec{x}_{k-2})e^{\lambda_0 t}$$

$(\frac{d}{dt} - \lambda_0 I)^j \vec{u}_{k-1}(t) = \vec{u}_{k-j-1}(t)$ for $j = 0,1,2,\cdots$

$$L(\frac{d}{dt}) = L(\lambda_0) + \frac{1}{1!}L^{(1)}(\lambda_0)(\frac{d}{dt} - \lambda_0 I) + \frac{1}{2!}L^{(2)}(\lambda_0)(\frac{d}{dt} - \lambda_0 I)^2 + \cdots + \frac{1}{l!}L^{(l)}(\lambda_0)(\frac{d}{dt} - \lambda_0 I)^l$$

Since $\vec{u}_{k-1}(t)$ is a solution of equation (4.3.1), then

$$\vec{0} = L(\frac{d}{dt})\vec{u}_{k-1}(t)$$

$$= L(\lambda_0)\vec{u}_{k-1}(t) + \frac{1}{1!}L^{(1)}(\lambda_0)\vec{u}_{k-2}(t) + \frac{1}{2!}L^{(2)}(\lambda_0)\vec{u}_{k-3}(t) + \cdots + \frac{1}{l!}L^{(l)}(\lambda_0)\vec{u}_{k-1-l}(t).$$

Equating the vector coefficients of powers of t to zero, we obtain the Jordan chain. □

Example 4.3.2: Solve the following system of differential equations:

$$\frac{d^2x_1}{dt^2} = 0$$

$$\frac{d^2x_2}{dt^2} + \frac{dx_1}{dt} - \frac{dx_2}{dt} + x_1 = 0$$

Solution:

Let $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \vec{x}^{(2)}(t) + \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} \vec{x}^{(1)}(t) + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \vec{x}(t) = \vec{0}$$

the corresponding $L(\lambda) = \begin{bmatrix} \lambda^2 & 0 \\ \lambda + 1 & \lambda^2 - \lambda \end{bmatrix}$.

By example 4.2.1, there are a latent roots $\lambda_1 = 1$ with latent vector $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $\lambda_2 = 0$ with latent vectors $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} -2 \\ 0 \end{bmatrix}$. Then

$\vec{v}_0(t) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} e^t$ is the solution corresponding to λ_1 and

$\vec{u}_0(t) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\vec{u}_1(t) = \begin{bmatrix} 1 \\ t \end{bmatrix}$ and $\vec{u}_2(t) = \begin{bmatrix} -2 + t \\ \frac{1}{2}t^2 \end{bmatrix}$ are the solutions corresponding to λ_2.

4.4 Nonhomogeneous Differential Equations

In this section we want to find the solution of the constant-coefficient-matrix differential equation of order l given by

$$L_l \vec{x}(t)^{(l)} + L_{l-1} \vec{x}(t)^{(l-1)} + \cdots + L_1 \vec{x}(t)^{(1)} + L_0 \vec{x}(t) = \vec{f}(t) \qquad (4.4.1)$$

where $L_0, L_1, \cdots, L_l \in \mathbf{C}^{n\times n}$, $\det L_l \neq 0$ and the vector function $\vec{f}(t)$, whose values in $\mathbf{C}^n$, is piecewise continuous.

Let $\vec{x}_0(t) = \vec{x}(t)$, $\vec{x}_1(t) = \vec{x}(t)^{(1)}$, $\cdots$, $\vec{x}_{l-1}(t) = \vec{x}(t)^{(l-1)}$, then equation (4.4.1) can be rewritten as

$$L_l \frac{d}{dt}(\vec{x}_{l-1}(t)) + L_{l-1} \vec{x}_{l-1}(t) + \cdots + L_1 \vec{x}_1(t) + L_0 \vec{x}_0(t) = \vec{f}(t),$$

Define

$$\vec{\tilde{x}}(t) := \begin{bmatrix} \vec{x}_0(t) \\ \vec{x}_1(t) \\ \vdots \\ \vec{x}_{l-1}(t) \end{bmatrix}, \quad \vec{g}(t) := \begin{bmatrix} \vec{0} \\ \vec{0} \\ \vdots \\ L_l^{-1} \vec{f}(t) \end{bmatrix},$$

the above equations can be condensed to the form

$$\frac{d\vec{\tilde{x}}(t)}{dt} = C_L \vec{\tilde{x}}(t) + \vec{g}(t), \qquad (4.4.2)$$

where C_L is the companion matrix for $L(\lambda)$.

The solution[2] of the equation (4.4.2) is:

$$\vec{\tilde{x}}(t) = e^{C_L t}\vec{z} + \int_0^t e^{C_L(t-\tau)}\vec{g}(\tau)d\tau,\ \vec{z} \in \mathbf{C}^{ln}.$$

Consider the standard triple (P_1, C_L, R_1) for $L(\lambda)$, where

$$P_1 = [I \quad 0 \quad \cdots \quad 0],\quad C_L = \begin{bmatrix} 0 & I_n & 0 & \cdots & 0 \\ 0 & 0 & I_n & \cdots & 0 \\ 0 & 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & I_n \\ -\hat{L}_0 & -\hat{L}_1 & \cdots & -\hat{L}_{l-2} & -\hat{L}_{l-1} \end{bmatrix},\quad R_1 = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ L_l^{-1} \end{bmatrix}$$

and $\hat{L}_j = L_l^{-1}L_j,\quad j = 0,1,\cdots,l-1$, thus

$$P_1\vec{\tilde{x}}(t) = [I \quad 0 \quad \cdots \quad 0]\begin{bmatrix} \vec{x}_0(t) \\ \vec{x}_1(t) \\ \vdots \\ \vec{x}_{l-1}(t) \end{bmatrix} = \vec{x}(t).$$

Then

$$\vec{x}(t) = P_1 e^{C_L t}\vec{z} + P_1\int_0^t e^{C_L(t-\tau)}R_1\vec{f}(\tau)d\tau.$$

From the above discussion, we conclude the following lemma.

---

[2] If $\vec{f}(t)$ is a piecewise-continuous function of t with values in $\mathbf{C}^n$, $t_0 \in \mathbf{R}$, and $A \in \mathbf{C}^{n \times n}$, then every solution of the differential equation

$$\frac{d\vec{x}(t)}{dt} = A\vec{x}(t) + \vec{f}(t)$$

has the form

$$\vec{x}(t) = e^{A(t-t_0)}\vec{x}_0 + \int_{t_0}^t e^{A(t-\tau)}\vec{f}(\tau)d\tau,$$

and $\vec{x}_0 = \vec{x}(t_0)$. See [11] page 339.

Lemma 4.4.1: Every solution of the following equation $L_l \vec{x}(t)^{(l)} + L_{l-1} \vec{x}(t)^{(l-1)} + \cdots + L_1 \vec{x}(t)^{(1)} + L_0 \vec{x}(t) = \vec{f}(t)$

has the form

$$\vec{x}(t) = P_1 e^{C_L t} \vec{z} + P_1 \int_0^t e^{C_L (t-\tau)} R_1 \vec{f}(\tau) d\tau . \square$$

Theorem 4.4.1: There is a unique solution of the differential equation $L_l \vec{x}(t)^{(l)} + L_{l-1} \vec{x}(t)^{(l-1)} + \cdots + L_1 \vec{x}(t)^{(1)} + L_0 \vec{x}(t) = \vec{f}(t)$ satisfying initial conditions

$$\vec{x}^{(j)}(0) = \vec{x}_j, \quad j = 0,1,\cdots,l-1$$

for any given vectors $\vec{x}_0, \vec{x}_1, \cdots, \vec{x}_{l-1} \in \mathbf{C}^n$. This solution is given by $\vec{x}(t) = P_1 e^{C_L t} \vec{z} + P_1 \int_0^t e^{C_L (t-\tau)} R_1 \vec{f}(\tau) d\tau$

with

$$\vec{z} = \begin{bmatrix} R_1 & C_L R_1 & \cdots & C_L^{\,l-1} R_1 \end{bmatrix} S_L \begin{bmatrix} \vec{x}_0 \\ \vec{x}_1 \\ \vdots \\ \vec{x}_{l-1} \end{bmatrix}$$

and S_L is the symmetrizer for $L(\lambda)$.

Proof:

$\vec{x}(0) \quad = P_1 \vec{z}$,

$\vec{x}^{(1)}(0) = P_1 C_L \vec{z}$,

$$\vec{x}^{(2)}(0) = P_1 C_L{}^2 \vec{z},$$

$$\vdots$$

$$\vec{x}^{(l-1)}(0) = P_1 C_L{}^{l-1} \vec{Z}.$$

$$\begin{bmatrix} \vec{x}(0) \\ \vec{x}^{(1)}(0) \\ \vec{x}^{(2)}(0) \\ \vdots \\ \vec{x}^{(l-1)}(0) \end{bmatrix} = \begin{bmatrix} P_1 \\ P_1 C_L \\ P_1 C_L{}^2 \\ \vdots \\ P_1 C_L{}^{l-1} \end{bmatrix} \vec{z} = Q \vec{z}.$$

Recall the proof of theorem 4.1.4, we note that $Q = S$ and $R = S^{-1} S_L^{-1}$.

Therefore $QRS_L = SS^{-1} S_L^{-1} S_L = I$

$$\begin{bmatrix} \vec{x}(0) \\ \vec{x}^{(1)}(0) \\ \vec{x}^{(2)}(0) \\ \vdots \\ \vec{x}^{(l-1)}(0) \end{bmatrix} = Q \vec{z} = Q \begin{bmatrix} R_1 & C_L R_1 & \cdots & C_L{}^{l-1} R_1 \end{bmatrix} S_L \begin{bmatrix} \vec{x}_0 \\ \vec{x}_1 \\ \vec{x}_2 \\ \vdots \\ \vec{x}_{l-1} \end{bmatrix}$$

$$= QRS_L \begin{bmatrix} \vec{x}_0 \\ \vec{x}_1 \\ \vec{x}_2 \\ \vdots \\ \vec{x}_{l-1} \end{bmatrix} = \begin{bmatrix} \vec{x}_0 \\ \vec{x}_1 \\ \vec{x}_2 \\ \vdots \\ \vec{x}_{l-1} \end{bmatrix}.$$

Therefore $\vec{x}(t)$ satisfies the initial conditions. Since Q is an invertible matrix then $\vec{z}$ is unique. □

Appendix A

1. Jordan Canonical Forms

Definition (1): A matrix $J \in \mathbf{C}^{n\times n}$ is in Jordan canonical form if it is a block diagonal matrix. That is,

$$J = \begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_k \end{bmatrix}$$

in which every diagonal block $J_i; i = 1,2,\cdots,k$ is a Jordan block. That is,

$$J_i = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \ddots & \vdots \\ \vdots & \ddots & \lambda_i & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_i \end{bmatrix}.$$

Definition (2): Let A be $n\times n$ matrix. A vector $\vec{x}_m \in \mathbf{C}^n$ is a generalized eigenvector of type m corresponding to the eigenvalue λ for the matrix A if

$$(A-\lambda I)^m \vec{x}_m = \vec{0} \text{ and } (A-\lambda I)^{m-1} \vec{x}_m \neq \vec{0} \tag{1}$$

The chain propagated by $\vec{x}_m$, a generalized eigenvector of type m corresponding to the eigenvalue λ for A, is the set of vectors in $\mathbf{C}^n$ $\{\vec{x}_m, \vec{x}_{m-1}, \cdots, \vec{x}_1\}$ given by

$$\vec{x}_{m-1} = (A - \lambda I)\vec{x}_m$$

$$\vec{x}_{m-2} = (A - \lambda I)\vec{x}_{m-1} = (A - \lambda I)^2 \vec{x}_m$$

$$\vec{x}_{m-3} = (A - \lambda I)\vec{x}_{m-2} = (A - \lambda I)^3 \vec{x}_m \tag{2}$$

$$\vdots$$

$$\vec{x}_1 = (A - \lambda I)\vec{x}_2 = (A - \lambda I)^{m-1} \vec{x}_m$$

In general, for $j = 1,2,\cdots,m-1$

$$\vec{x}_j = (A - \lambda I)\vec{x}_{j+1} = (A - \lambda I)^{m-j} \vec{x}_m \tag{3}$$

Theorem (1): The jth vector in a chain, $\vec{x}_j$ as defined by (3), is a generalized eigenvector of type j corresponding to the same matrix and eigenvalue associated with the generalized eigenvector of type m that propagated the chain.

Proof:

Let $\vec{x}_m$ be a generalized eigenvector of type m for a matrix A with eigenvalue λ. Then, $(A - \lambda I)^m \vec{x}_m = \vec{0}$ and $(A - \lambda I)^{m-1} \vec{x}_m \neq \vec{0}$. Using (3), we conclude that

$$(A - \lambda I)^j \vec{x}_j = (A - \lambda I)^j [(A - \lambda I)^{m-j} \vec{x}_m] = (A - \lambda I)^m \vec{x}_m = \vec{0}$$

$$(A - \lambda I)^{j-1} \vec{x}_j = (A - \lambda I)^{j-1} [(A - \lambda I)^{m-j} \vec{x}_m] = (A - \lambda I)^{m-1} \vec{x}_m \neq \vec{0}$$

Thus, $\vec{x}_j$ is a generalized eigenvector of type j corresponding to the eigenvalue λ for A. □

Theorem (2): A chain is a linearly independent set of vectors.

Proof:

Let $\{\vec{x}_m, \vec{x}_{m-1}, \cdots, \vec{x}_1\}$ be a chain propagated from $\vec{x}_m$, a generalized eigenvector of type m corresponding to the eigenvalue λ for A. We consider the vector equation

$$c_m \vec{x}_m + c_{m-1} \vec{x}_{m-1} + \cdots + c_1 \vec{x}_1 = \vec{0}. \tag{4}$$

To prove that this chain is linearly independent, we must show that the only solution to (4) is the trivial solution $c_m = c_{m-1} = \cdots = c_1 = 0$.

We shall do this iteratively. First, we multiply both sides of (4) by $(A-\lambda I)^{m-1}$. Note that for $j = 1, 2, \cdots, m-1$,

$$(A-\lambda I)^{m-1} c_j \vec{x}_j = c_j (A-\lambda I)^{m-j-1}[(A-\lambda I)^j \vec{x}_j]$$

$$= c_j (A-\lambda I)^{m-j-1}[\vec{0}] = \vec{0}.$$

Thus, (4) becomes $c_m (A-\lambda I)^{m-1} \vec{x}_m = \vec{0}$. But $\vec{x}_m$ is a generalized eigenvector of type m, so the vector $(A-\lambda I)^{m-1} \vec{x}_m \neq \vec{0}$. It then follows that $c_m = 0$.

Substituting $c_m = 0$ into (4) and then multiplying the resulting equation by $(A-\lambda I)^{m-2}$, we find, by similar reasoning, that $c_{m-1} = 0$.

Continuing this process, we find iteratively that $c_m = c_{m-1} = \cdots = c_1 = 0$, which implies that the chain is linearly independent. □

A generalized eigenvector $\vec{x}_m$ of type m corresponding to an eigenvalue λ of an $n \times n$ matrix A has the property that

$$(A-\lambda I)^m \vec{x}_m = \vec{0} \text{ and } (A-\lambda I)^{m-1} \vec{x}_m \neq \vec{0}.$$

Thus $\vec{x}_m$ is in the kernel of $(A-\lambda I)^m$ but not in the kernel of $(A-\lambda I)^{m-1}$. If $\vec{x} \in \ker[(A-\lambda I)^{m-1}]$, then $\vec{x} \in \ker[(A-\lambda I)^m]$. Consequently, the dimension of $\ker[(A-\lambda I)^{m-1}]$ is less than the dimension of $\ker[(A-\lambda I)^m]$ or, in terms of the rank $rank[(A-\lambda I)^{m-1}] > rank[(A-\lambda I)^m]$.

The difference

$$\rho_m = rank[(A-\lambda I)^{m-1}] - rank[(A-\lambda I)^m] \qquad (5)$$

is the number of linearly independent generalized eigenvectors of type m corresponding to A and its eigenvalue λ. The differences $\rho_m, m = 1,2,\cdots$ are called index numbers.

Definition (3): A canonical basis for a matrix $A \in \mathbf{C}^{n\times n}$ is the set of n linearly independent set of generalized eigenvectors of type $1,2,\cdots,m$.

How to create a canonical basis?

For each distinct eigenvalue of a matrix A do the following:

Step 1: Using the index numbers, determine the number of linearly independent generalized eigenvectors of highest type, say type m, corresponding to λ.

Determine one such set, $\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_r\}$.

Step 2: If $m = 1$, stop; otherwise continue.

Step 3: For each vector $\vec{v}_i, i = 1,2,\cdots,r$, calculate $(A-\lambda I)\vec{v}_i$, the next vector in its chain.

Step 4: Using the index numbers, determine the number of linearly independent generalized eigenvectors of the type $m-1$. If this number coincides with the number of vectors obtained in step 3, replace $\vec{v}_i$ by $(A-\lambda I)\vec{v}_i, i=1,2,\cdots,r$ and go to step 6; otherwise continue.

Step 5: Find additional generalized eigenvectors of type $m-1$ so that when these new vectors are adjoint to the set $\{(A-\lambda I)\vec{v}_i, i=1,2,\cdots,r\}$, they form a set of linearly independent vectors.

Step 6: Decrement m by 1 and return to step 2.

Example (1): The matrix $A=\begin{bmatrix} 4 & 0 & -1 & -1 \\ -4 & 2 & 2 & 2 \\ 2 & 1 & 2 & 0 \\ 2 & -1 & -2 & 0 \end{bmatrix}$ has an eigenvalue 2 of multiplicity 4.

$$A-2I=\begin{bmatrix} 2 & 0 & -1 & -1 \\ -4 & 0 & 2 & 2 \\ 2 & 1 & 0 & 0 \\ 2 & -1 & -2 & -2 \end{bmatrix},\ rank[(A-2I)]=2$$

$$(A-2I)^2=\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},\ rank[(A-2I)^2]=0$$

$$(A-2I)^3=\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},\ rank[(A-2I)^3]=0$$

The associated index numbers are:

$\rho_1 = rank[(A-2I)^0] - rank[(A-2I)] = 4-2 = 2$,

$\rho_2 = rank[(A-2I)] - rank[(A-2I)^2] = 2-0 = 2$,

$\rho_3 = rank[(A-2I)^2] - rank[(A-2I)^3] = 0-0 = 0$.

Corresponding to $\lambda = 2$, A has two linearly independent generalized eigenvectors of type 1 and two linearly independent generalized eigenvectors of type 2.

To find the two linearly independent generalized eigenvectors of type 2, we solve the equations:

$(A-2I)^2 \vec{x}_2 = \vec{0}$ and $(A-2I)\vec{x}_2 \neq \vec{0}$.

Therefore the two linearly independent generalized eigenvectors of type 2 are:

$\vec{x}_2 = [1 \quad 0 \quad 0 \quad 0]^T$ and $\vec{y}_2 = [0 \quad 1 \quad 0 \quad 0]^T$.

The two linearly independent generalized eigenvectors of type 1 are:

$$\vec{x}_1 = (A-2I)\vec{x}_2 = \begin{bmatrix} 2 & 0 & -1 & -1 \\ -4 & 0 & 2 & 2 \\ 2 & 1 & 0 & 0 \\ 2 & -1 & -2 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ -4 \\ 2 \\ 2 \end{bmatrix}$$

and

$$\vec{y}_1 = (A-2I)\vec{y}_2 = \begin{bmatrix} 2 & 0 & -1 & -1 \\ -4 & 0 & 2 & 2 \\ 2 & 1 & 0 & 0 \\ 2 & -1 & -2 & -2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}.$$

Therefore

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

Note (1): In a canonical basis, all vectors from the same chain are grouped together, and generalized eigenvectors in each chain are ordered by increasing type.

2. The Exponential Function

The exponential of a square matrix A is defined as:

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} = I + \frac{A}{1!} + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots$$

A matrix J is in Jordan canonical form if it has the block diagonal pattern

$$J = \begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_r \end{bmatrix}$$

where each J_i $(i = 1,2,\cdots,r)$ being a Jordan blocks of the form

$$J_i = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \ddots & \vdots \\ \vdots & \ddots & \lambda_i & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_i \end{bmatrix}.$$

Any matrix $A \in \mathbf{C}^{n\times n}$ is similar to a matrix J in Jordan canonical form. That is, there exists a nonsingular matrix M such that

$$A = MJM^{-1} \text{ and } e^A = Me^J M^{-1}.$$

Now, we find e^J,

since

$$J^k = \begin{bmatrix} J_1^k & 0 & \cdots & 0 \\ 0 & J_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_r^k \end{bmatrix}$$

Then

$$e^J = \sum_{k=0}^{\infty} \frac{J^k}{k!} = \sum_{k=0}^{\infty} \frac{1}{k!} \begin{bmatrix} J_1^k & 0 & \cdots & 0 \\ 0 & J_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_r^k \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{k=0}^{\infty} \frac{J_1^k}{k!} & 0 & \cdots & 0 \\ 0 & \sum_{k=0}^{\infty} \frac{J_2^k}{k!} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_{k=0}^{\infty} \frac{J_r^k}{k!} \end{bmatrix}$$

$$= \begin{bmatrix} e^{J_1} & 0 & \cdots & 0 \\ 0 & e^{J_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{J_r} \end{bmatrix}.$$

Since all Jordan blocks have superdiagonal element, which are all ones, for a $p \times p$ Jordan block, we can show by direct calculations that each successive power has one additional diagonal of nonzero entries, until all elements above the main diagonal become nonzero. On each diagonal, the entries are identical.

If we designate the nth power of a Jordan block as the matrix $[a_{ij}^n]$, then the entries can be expressed compactly in terms of derivatives as

$$a_{i,i+j}^n = \begin{cases} \frac{1}{j!} \frac{d^j}{d\lambda_i^j} (\lambda_i^n) & for \quad j = 0,1,\cdots,\min\{n, p-i\} \\ 0 & otherwise \end{cases}$$

then,

$$e^{J_i} = e^{\lambda_i} \begin{bmatrix} 1 & \frac{1}{1!} & \frac{1}{2!} & \frac{1}{3!} & \cdots & \frac{1}{(p-1)!} \\ 0 & 1 & \frac{1}{1!} & \frac{1}{2!} & \cdots & \frac{1}{(p-2)!} \\ 0 & 0 & 1 & \frac{1}{1!} & \cdots & \frac{1}{(p-3)!} \\ 0 & 0 & 0 & 1 & \cdots & \frac{1}{(p-4)!} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Example (2): Compute e^{At} when $A = \begin{bmatrix} 3 & 2 & 1 \\ -1 & 3 & 2 \\ 1 & -3 & -2 \end{bmatrix}$

Solution:

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

The matrix A has two different eigenvalues:

$\lambda_1 = 0$ with multiplicity 1 and $\lambda_2 = 2$ with multiplicity 2.

$(A - 0I)\vec{x}_1 = \vec{0}$

$$\begin{bmatrix} 3 & 2 & 1 \\ -1 & 3 & 2 \\ 1 & -3 & -2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\vec{x}_1 = [1 \quad -7 \quad 11]^T$,

$A - 2I = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 1 & 2 \\ 1 & -3 & -4 \end{bmatrix}$, $rank[(A-2I)] = 2$,

$$(A-2I)^2 = \begin{bmatrix} 0 & 1 & 1 \\ 0 & -7 & -7 \\ 0 & 11 & 11 \end{bmatrix},\ rank[(A-2I)^2] = 1,$$

$$(A-2I)^3 = \begin{bmatrix} 0 & -2 & -2 \\ 0 & 14 & 14 \\ 0 & -22 & -22 \end{bmatrix},\ rank[(A-2I)^3] = 1,$$

$$\rho_1 = rank[(A-2I)^0] - rank[(A-2I)] = 3 - 2 = 1,$$

$$\rho_2 = rank[(A-2I)] - rank[(A-2I)^2] = 2 - 1 = 1,$$

$$\rho_3 = rank[(A-2I)^2] - rank[(A-2I)^3] = 1 - 1 = 0.$$

There is one generalized eigenvector of type 2 corresponding to $\lambda_2 = 2$ and one generalized eigenvector of type 1 (that is, there is one eigenvector) corresponding to $\lambda_2 = 2$

$$(A-2I)^2 \vec{v}_2 = \vec{0} \text{ and } (A-2I)\vec{v}_2 \neq \vec{0}$$

$$\begin{bmatrix} 0 & 1 & 1 \\ 0 & -7 & -7 \\ 0 & 11 & 11 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 2 & 1 \\ -1 & 1 & 2 \\ 1 & -3 & -4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\vec{v}_2 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T$$

$$\vec{v}_1 = (A-2I)\vec{v}_2 = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 1 & 2 \\ 1 & -3 & -4 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

$$M = \begin{bmatrix} 1 & 1 & 1 \\ -7 & -1 & 0 \\ 11 & 1 & 0 \end{bmatrix},\ M^{-1} = \begin{bmatrix} 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{-11}{4} & \frac{-7}{4} \\ 1 & \frac{5}{2} & \frac{3}{2} \end{bmatrix}$$

$$J = M^{-1}AM = \begin{bmatrix} [0] & 0 \\ 0 & \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} J_1 & 0 \\ 0 & J_2 \end{bmatrix}$$

$$e^{At} = Me^{Jt}M^{-1}$$

$$e^{tJ_1} = [1]$$

$$e^{tJ_2} = e^{2t}\begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} e^{2t} & te^{2t} \\ 0 & e^{2t} \end{bmatrix}$$

$$e^{Jt} = \begin{bmatrix} e^{J_1 t} & 0 \\ 0 & e^{J_2 t} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{2t} & te^{2t} \\ 0 & 0 & e^{2t} \end{bmatrix}$$

$$e^{At} = \begin{bmatrix} 1 & 1 & 1 \\ -7 & -1 & 0 \\ 11 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{2t} & te^{2t} \\ 0 & 0 & e^{2t} \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{-11}{4} & \frac{-7}{4} \\ 1 & \frac{5}{2} & \frac{3}{2} \end{bmatrix}$$

$$= \begin{bmatrix} e^{2t} + te^{2t} & \frac{1}{4} - \frac{1}{4}e^{2t} + \frac{5}{2}te^{2t} & \frac{1}{4} - \frac{1}{4}e^{2t} + \frac{3}{2}te^{2t} \\ -te^{2t} & \frac{-7}{4} + \frac{11}{4}e^{2t} - \frac{5}{2}te^{2t} & \frac{-7}{4} + \frac{7}{4}e^{2t} - \frac{3}{2}te^{2t} \\ te^{2t} & \frac{11}{4} - \frac{11}{4}e^{2t} + \frac{5}{2}te^{2t} & \frac{11}{4} - \frac{7}{4}e^{2t} + \frac{3}{2}te^{2t} \end{bmatrix}.$$

Symbols

R .... the set of real numbers

C .... the set of complex numbers

F .... field

$A(\lambda)$ .... matrix polynomial

$a(\lambda)$ .... scalar polynomial (polynomial with scalar coefficients)

I_n .... identity matrix of order n

$\mathbf{C}^{n\times n}$ .... $n\times n$ matrices with complex entries

$\deg A(\lambda)$ .... degree of $A(\lambda)$

$\det A(\lambda)$ .... determinant of $A(\lambda)$

$(A(\lambda))^{-1}$ .... the inverse of $A(\lambda)$

const. .... constant (*i.e.* element from **C**)

$(A(\lambda))^T$ .... the transpose of $A(\lambda)$

$A_c(\lambda)$ .... canonical matrix polynomial

diag[ ] .... diagonal matrix

$\det A(\lambda) \not\equiv 0$ .... the determinant of $A(\lambda)$ is not identically zero

$A_L^{-1}(\lambda)$ .... a matrix polynomial satisfies $A_L^{-1}(\lambda)A(\lambda) = I_n$

$A_R^{-1}(\lambda)$ .... a matrix polynomial satisfies $A(\lambda)A_R^{-1}(\lambda) = I_m$

$A\begin{pmatrix} 1 & 2 & \cdots & m \\ j_1 & j_2 & \cdots & j_m \end{pmatrix}; \lambda$ .... minor of order m of the matrix polynomial $A(\lambda)$

$(A(\lambda))^*$ .... conjugate transpose of $A(\lambda)$

$\overline{\lambda}$ .... conjugate of λ

$v_+(A)$ .... the number of positive eigenvalues of A

$v_-(A)$ .... the number of negative eigenvalues of A

$v_0(A)$ .... the number of zero eigenvalues of A

$sigA$ .... signature of A

$r(A)$ .... the general rank of $A(\lambda)$

$\sigma^2(x)$ .... $(\sigma o \sigma)(x)$

$f_+(\lambda)$ .... $\varepsilon^{\deg f(\lambda)} f_*(\lambda)$

$\vec{x}(t)$ .... the vector $x(t)$ *i.e.* $\vec{x}(t) = [x_1(t) \quad x_2(t) \quad \cdots \quad x_n(t)]^T$

$\vec{x}(t)^{(i)}$ .... the *ith* derivative of $\vec{x}(t)$

C_a .... the $l \times l$ companion matrix of the scalar polynomial $a(\lambda)$

C_L .... the $\ln \times \ln$ companion matrix of the matrix polynomial $L(\lambda)$

S_L .... symmetrizer for the matrix polynomial $L(\lambda)$

$\ker L(\lambda_j)$ .... the kernel of the $n \times n$ constant matrix $L(\lambda_j)$

$L^{(i)}(\lambda)$ .... the *ith* derivative of the matrix polynomial $L(\lambda)$

$\binom{r}{p}$ .... the combination of r over p *i.e.* $\dfrac{r!}{(r-p)!\,p!}$

$\dfrac{d^j \vec{x}}{dt^j}$ .... the *jth* derivative of $\vec{x}(t)$ with respect to t

e^A .... the exponential of a square matrix A

$\Delta_j(\lambda)$ .... the monic greatest common divisor of all minors of order j of $B(\lambda)$.

References

[1] Anderson, B. D. O. and Moore, J. B., Optimal Filtering. Englewood Cliffs: Prentice-Hall, 1979.

[2] Anton, H., Elementary Linear Algebra. New York: John Wiley and Sons, 1987.

[3] Boyce, E., William and DiPrima C. Richard, Elementary Differential Equations and Boundary Value Problems. New York: John Wiley and Sons, Inc., 1997.

[4] Bronson, Richard, Linear Algebra: an Introduction. New York: Academic Press, Inc., 1995.

[5] Dragomir, Ž. Đokovic, Factorization of Hermitian Matrix Polynomials with Constant Signature. Linear Algebra Appl., 194, pp. 85-90, 1993.

[6] Gantmacher, F. R., The Theory of Matrices. United States of America: American Mathematical Society, 1998.

[7] Gohberg, I., Lancaster, P. and Rodman, L., Factorization of Selfadjoint Matrix Polynomials with Constant Signature. Linear and Multilinear Algebra, Vol. 11, pp. 209-224, 1982.

[8] Gohberg, I., Lancaster, P. and Rodman, L., Matrices and Indefinite Scalar Products. Basel: Birkhäuser Verlag, 1983.

[9] Gohberg, I., Lancaster, P. and Rodman, L., Matrix Polynomials. New York: Academic Press, 1982.

[10] Jakubovic, V. A., Factorization of Symmetric Matrix Polynomials. Soviet Math. Dokl. Vol. 11, No. 5, pp. 1261-1264, 1970.

[11] Kolman, B., Introductory Linear Algebra with Applications. New York: Macmillan Publishing Co., Inc., 1976.

[12] Lancaster, P. and Tismenetsky, M., The Theory of Matrices. New York: Academic Press, 1985.

[13] Lancaster, Peter, Common Eigenvalues, Divisors, and Multiples of Matrix Polynomials. Elsevier Science Publishing Co., Inc., pp. 139-160, 1986.

[14] Lyubachevskii, B. D., Factorization of Symmetric Matrices with Elements from a Ring with Involution. I, II (English transl.), Siberian Math. J. 14: 233-246, 423-433, 1973.

[15] Miller, R. K. and Michel, A. N., Ordinary Differential Equations. New York: Academic Press, 1982.

[16] Ran, A. C. M. and Rodman, L., Factorization of Matrix Polynomials with Symmetries. SIAM J. MATRIX ANAL. APPL., Vol. 15, No. 3, pp. 845-864, 1994.

[17] Zill, G. Dennis, A First Course in Differential Equations with Applications. Boston: Prindle, Weber and Schmidt, 1982.