This is to certify that the

dissertation entitled

Determining the Jordan
normal form of a matrix

presented by

Tianjun Wang

has been accepted towards fulfillment
of the requirements for

Ph.D degree in mathematics

Major professor

Date Aug. 16, 99

DETERMINING THE JORDAN NORMAL FORM OF A MATRIX

By

Tianjun Wang

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Mathematics

1999

ABSTRACT

DETERMINING THE JORDAN NORMAL FORM OF A MATRIX

By

Tianjun Wang

In this dissertation, we present a new algorithm for calculating the Jordan normal form of an $n \times n$ matrix. The algorithm first determines the structure of the Jordan blocks of a matrix without solving for its eigenvalues, and then calculates eigenvalues of the matrix, as zeros of its minimal polynomials, to fill in the diagonal entries of the Jordan blocks. The approximation of eigenvalues of the matrix has no influence on the structure. There are two crucial steps for determining the structure of the Jordan blocks of a matrix in our algorithm. One is to determine the minimal polynomials of cyclic subspaces of the matrix by using a Las Vegas type algorithm with probability one for obtaining the correct answer. The other is to ascertain the structures of the minimal polynomials. The most important feature of our algorithm is that the implementations of these two steps can all be carried out by the symbolic computation, which consists of exact rational operations. Therefore, if the entries of a given matrix are all rational, then the structure of the Jordan blocks in its Jordan normal form can be determined exactly without approximating its eigenvalues.

To my family

ACKNOWLEDGMENTS

I am most indebted to my dissertation advisor, University Distinguished Professor Tien–Yien Li, for his support during my graduate study at Michigan State University. I would like to thank him for suggesting the problem and giving invaluable guidance which makes this dissertation a reality.

Special thanks go to Professor Zhinan Zhang at Xinjiang University for his discussions, suggestions and collaborations when he visited Michigan State University.

I would like to thank my dissertation committee members Professor Chichia Chiu, Professor Richard Hill, Professor Jay Kurtz and Professor Zhengfang Zhou for their valuable suggestions.

My greatest thanks are reserved for my wife, Diane D. Dong, and my family for their enduring patience, constant support and encouragement.

TABLE OF CONTENTS

INTRODUCTION

The numerical computation for the Jordan normal form of a given matrix can be traced back to the 1970's, e.g., see [3], [4], [9] and [11]. Usually, the singular value decomposition (SVD) method is used to ascertain the structure of the Jordan blocks and compute the Jordan normal form of a matrix. The difficulty of this method lies in a critical distinction between equal and unequal eigenvalues for computing the rank and determining finally the block structures. In numerical computation, it is, in general, difficult to determine whether a matrix has exactly equal eigenvalues. Since the sensitivity of computing the rank of a matrix subject to perturbations degrades the reliability in determining the rank, the method can sometimes alter the structure of the Jordan normal form radically[8]. Recently, the computation of the Jordan normal form of a matrix has been studied in several articles. The algebraic theory of matrices is used in [13] to present a parallel algorithm for ascertaining the structures of the Jordan blocks of a matrix.

The Jordan normal form of a matrix consists of two parts: the structure of the Jordan blocks and the eigenvalues of the matrix. In this dissertation, we present a new constructive proof of a fundamental decomposition theorem in [5] which provides the geometric theory of elementary divisors of a matrix. Based on the new proof of the theorem, we develop an algorithm for calculating the Jordan normal form of an $n \times n$ matrix. The algorithm first determines the structure of the Jordan blocks of a matrix without solving for its eigenvalues, and then calculates eigenvalues of the matrix, as zeros of its minimal polynomials, to fill in the diagonal entries of the Jordan blocks. The approximation of eigenvalues of the matrix has no influence on the structure. There are two crucial steps for determining the structure of the Jordan blocks of a matrix in our algorithm. The first one is to determine the minimal polynomials of cyclic subspaces of the matrix by using a Las Vegas type algorithm with probability one for obtaining the correct answer. The other is to ascertain the structures of those minimal polynomials. The most important feature of our algorithm is that the implementations of these two steps can all be carried out by symbolic computation, which consists of exact rational operations. Therefore, if the entries of a given matrix are all rational, then the structure of the Jordan blocks in its Jordan normal form can be determined exactly without approximating its eigenvalues.

The dissertation is organized as follows. In Chapter 1, we introduce some of the basic concepts, results and notations that will be utilized repeatedly in subsequent chapters. In Chapter 2, we study the theory of determining the Jordan normal form of a matrix, and give a new constructive proof of Theorem 3 on page 187 in [5]. In Chapter 3, we develop an algorithm of determining the Jordan normal form of a matrix based on the theoretical results obtained in Chpater 2. In Chapter 4, the implementation of the algorithm presented in Chapter 3 is discussed in detail, meanwhile, computational result is provided to illustrate the effectiveness of the algorithm.

CHAPTER 1

PRELIMINARIES

We use $\mathcal{R}^n$ to denote the n-dimensional vector space over the real number field $\mathcal{R}$, and $\mathcal{R}[\lambda]$ is the polynomial ring on $\mathcal{R}$. Let L be a linear operator mapping $\mathcal{R}^n$ to $\mathcal{R}^n$, then L can be represented by an $n \times n$ matrix A with elements in $\mathcal{R}$ under a basis of $\mathcal{R}^n$. Conversely, an $n \times n$ matrix A with elements in $\mathcal{R}$ can also be regarded as a linear operator from $\mathcal{R}^n$ to $\mathcal{R}^n$ under the standard basis $e_1, e_2, \ldots, e_n$ of $\mathcal{R}^n$, where $e_1 = (1, 0, \ldots, 0)^T$, $e_2 = (0, 1, 0, \ldots, 0)^T$, …, $e_n = (0, \ldots, 0, 1)^T$.

Definition 1.0.1 A subspace S of $\mathcal{R}^n$ is *invariant* under A (or is an *invariant subspace* of A) if $A\mathbf{x} \in S$ whenever $\mathbf{x} \in S$.

For $\mathbf{x} \neq 0$ in $\mathcal{R}^n$, consider the sequence

$$\mathbf{x},\ A\mathbf{x},\ A^2\mathbf{x},\ \ldots,\ A^l\mathbf{x},\ \ldots. \tag{1.1}$$

Since $\mathcal{R}^n$ is an n-dimensional vector space, these vectors defined in (1.1) cannot all be linearly independent, so there exists a positive integer $1 \leq k \leq n$ such that $A^k\mathbf{x}$ is the first vector (i.e., the one with the lowest exponent) which is linearly dependent on its predecessors $\mathbf{x}, A\mathbf{x}, A^2\mathbf{x}, \ldots, A^{k-1}\mathbf{x}$.

Definition 1.0.2 The subspace spanned by $\mathbf{x}, A\mathbf{x}, \ldots, A^{k-1}\mathbf{x}$ is called a *cyclic subspace* of $\mathcal{R}^n$ generated by $\mathbf{x}$.

Definition 1.0.3 If the monic polynomial $\varphi(\lambda) = \lambda^m - (a_m\lambda^{m-1} + \cdots + a_2\lambda + a_1) \in \mathcal{R}[\lambda]$ satisfies

$$\varphi(A)\mathbf{x} = 0 \tag{1.2}$$

for certain vector $\mathbf{x} \in \mathcal{R}^n$, then the polynomial $\varphi(\lambda)$ will be called an *annihilating polynomial* for $\mathbf{x}$ (with respect to the operator A). If for every vector $\mathbf{x} \in \mathcal{R}^n$, $\varphi(\lambda)$ satisfies (1.2), then $\varphi(\lambda)$ is regarded as an *annihilating polynomial for the space* $\mathcal{R}^n$ (with respect to the operator A). The monic polynomial with least degree among all of the annihilating polynomials for a given vector (or $\mathcal{R}^n$) will be called the *minimal polynomial* for the vector (or $\mathcal{R}^n$).

Note that every annihilating polynomial for a given vector (or $\mathcal{R}^n$) is divisible by the minimal polynomial for the vector (or $\mathcal{R}^n$). Obviously, the minimal polynomial for $\mathcal{R}^n$ is divisible by any minimal polynomial for any vector in $\mathcal{R}^n$.

A vector $\mathbf{x} \in \mathcal{R}^n$ is said to be a *regular vector* if its minimal polynomial coincides with the minimal polynomial of $\mathcal{R}^n$.

Theorem 1.0.1 [10] *Almost every vector in $\mathcal{R}^n$ is regular.*

Similar concepts and statements introduced above can be repeated in any subspace $U \subset \mathcal{R}^n$ as well as any quotient space $\mathcal{R}^n/V$ where V is an invariant subspace of $\mathcal{R}^n$ with respect to A. All these concepts repeated in a quotient space will be called *relative*, in contrast to the *absolute* concepts that were introduced above. For instance, a polynomial $\varphi(\lambda)$ is the *relative* minimal polynomial of a vector $\mathbf{y} \in \mathcal{R}^n(\text{mod } V)$ with respect to A if

$$\varphi(A)\mathbf{y} \equiv 0(\text{mod } V);$$

a polynomial $\psi(\lambda)$ is the *relative* minimal polynomial of $\mathcal{R}^n(\text{mod } V)$ with respect to A if

$$\psi(A)\mathbf{y} \equiv 0(\text{mod } V), \qquad \forall\, \mathbf{y} \in \mathcal{R}^n.$$

Notice that the relative minimal polynomial of a vector (or $\mathcal{R}^n/V$) is a divisor of the absolute one.

Definition 1.0.4 When a polynomial $\xi \in \mathcal{R}[\lambda]$ has the decomposition

$$\xi = \xi_1^{m_1}\xi_2^{m_2}\cdots\xi_r^{m_r}, \qquad \xi_i \in \mathcal{R}[\lambda],\ i = 1,\ldots,r$$

where $m_i(i = 1,2,\ldots,r)$ is a positive integer, then the decomposition is called a *squarefree decomposition* if $\gcd(\xi_i,\ \xi_i') = 1$ and $\gcd(\xi_i, \xi_j) = 1$ for $i \neq j$; $i, j = 1,\ldots,r$. Here, gcd stands for the greatest common divisor. Moreover, $\{\xi_1, \xi_2, \cdots, \xi_r\}$ will be called a *squarefree relatively prime basis* of ξ.

Generally, the squarefree decomposition of a polynomial is not unique over $\mathcal{R}[\lambda]$, and the irreducible decomposition of a polynomial is the extreme form of its squarefree decompositions.

Definition 1.0.5 A *common squarefree relatively prime basis* $\{h_1,\ldots,h_l\} \subset \mathcal{R}[\lambda]$ for a set of polynomials $\{g_1,\ldots,g_m\} \subset \mathcal{R}[\lambda]$ is a squarefree relatively prime basis of each g_i for $i = 1,\ldots,m$.

An algorithm involving probability is said to be of *Las Vegas type* if obtaining the correct result eventually is guaranteed, and probability only enters in speed considerations [13]. From Theorem 1.0.1, a regular vector in $\mathcal{R}^n$ can be found by a Las Vegas type algorithm. To begin with, we choose a vector $\mathbf{v}$ in $\mathcal{R}^n$ at random and find its minimal polynomial. This polynomial may fail to coincide with the minimal polynomial of $\mathcal{R}^n$, although, by Theorem 1.0.1, the probability of such failure is zero. When that happens, a different vector $\mathbf{v}'$ in $\mathcal{R}^n$ will be randomly chosen, and its minimal polynomial will be calculated to see if it annihilates the whole space. The probability one ("almost every") in Theorem 1.0.1 leads to extremely early success of this process. In practice, we always find a regular vector in our first attempt.

CHAPTER 2

THEORETICAL BACKGROUND

2.1 Main theorems

Theorem 2.1.1 [5] *Relative to a given* $n \times n$ *matrix* A*, the vector space* $\mathcal{R}^n$ *can always be split into a direct sum of cyclic subspaces* W_1*,* W_2*, ...,* W_t *with minimal polynomials* $\Psi_1(\lambda), \Psi_2(\lambda), \ldots, \Psi_t(\lambda)$ *in* $\mathcal{R}[\lambda]$ *respectively,*

$$\mathcal{R}^n = W_1 \oplus W_2 \oplus \cdots \oplus W_t \tag{2.1}$$

such that $\Psi_1(\lambda)$ *is the minimal polynomial of* $\mathcal{R}^n$ *and*

$$\Psi_t(\lambda)|\Psi_{t-1}(\lambda)|\cdots|\Psi_1(\lambda), \tag{2.2}$$

that is, each $\Psi_j(\lambda)$ *is divisible by* $\Psi_{j+1}(\lambda)$ $(j = 1, 2, \ldots, t-1)$*. Moreover,* $\Psi_{j+1}(\lambda)$ *is the relative minimal polynomial of* $\mathcal{R}^n(mod\ (W_1 \oplus W_2 \oplus \cdots \oplus W_j))$ *for* $j = 1, \ldots, t-1$.

A proof of the above theorem can be found in [5]. Our main goal in this dissertation is to provide a new constructive proof of the theorem which forms the fundamental base for our algorithm in determining the Jordan normal form of A.

The following theorem plays an important role in determining the structure of the Jordan blocks in the Jordan normal form of the matrix A (see also [6], [12] and [13]).

Theorem 2.1.2 [14] *There is a common squarefree relatively prime basis* $\{\varphi_1, \varphi_2, \cdots, \varphi_r\}$ *in* $\mathcal{R}[\lambda]$ *for the minimal polynomials* $\Psi_1, \Psi_2, \ldots, \Psi_t$ *in Theorem 2.1.1. These minimal polynomials can be represented as follows:*

$$\begin{cases} \Psi_1 = \varphi_1^{m_{1,1}} \varphi_2^{m_{1,2}} \cdots \varphi_r^{m_{1,r}}, \\ \Psi_2 = \varphi_1^{m_{2,1}} \varphi_2^{m_{2,2}} \cdots \varphi_r^{m_{2,r}}, \\ \qquad \cdots\cdots \\ \Psi_t = \varphi_1^{m_{t,1}} \varphi_2^{m_{t,2}} \cdots \varphi_r^{m_{t,r}}, \end{cases} \tag{2.3}$$

where $m_{1,k} \geq m_{2,k} \geq \cdots \geq m_{t,k} \geq 0$ *for* $k = 1, 2, \ldots, r$.

Theorem 2.1.3 [5] *For* $\{\varphi_1, \varphi_2, \cdots, \varphi_r\}$ *in Theorem 2.1.2, let* $p_k = \deg \varphi_k$, $k = 1, \ldots, r$. *Then, corresponding to each* $\varphi_k^{m_{ik}}$ *in* (2.3) *for* $k = 1, \ldots, r$ *and* $i = 1, \ldots, t$, *there are* p_k *Jordan blocks of order* m_{ik} *of the form*

$$\begin{pmatrix} \lambda_{k,l} & 1 & & 0 \\ & \lambda_{k,l} & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_{k,l} \end{pmatrix}_{m_{i,k} \times m_{i,k}} , \qquad l = 1, \ldots, p_k \tag{2.4}$$

in the Jordan normal form of A. *Those* $\lambda_{k,l}$*'s are zeros of* φ_k. *In fact, they are eigenvalues of* A, *which may or may not be in* $\mathcal{R}$.

2.2 A constructive proof of Theorem 2.1.1

The theoretical results obtained below form the foundation of our algorithm for determining the minimal polynomials of cyclic subspaces of the $n \times n$ matrix A. Let V be a k-dimensional invariant subspace of A. We define a linear operator G in the quotient space $\mathcal{R}^n/V$ as follows:

$$G \ : \ \mathcal{R}^n/V \longrightarrow \mathcal{R}^n/V,$$

$$\{\mathbf{x}\} \longmapsto \{A\mathbf{x}\},$$

where $\{\mathbf{x}\} = \{\mathbf{y} \in \mathcal{R}^n : \mathbf{y} \equiv \mathbf{x}(\text{mod } V)\}$ denotes an element of $\mathcal{R}^n/V$. The map is well defined since V is an invariant subspace of A.

Since

$$\dim(\mathcal{R}^n/V) = \dim(\mathcal{R}^{n-k}) = n - k,$$

there is an isomorphism σ from $\mathcal{R}^n/V$ to $\mathcal{R}^{n-k}$. We then have a linear operator $H = \sigma \circ G \circ \sigma^{-1}$ from $\mathcal{R}^{n-k}$ to $\mathcal{R}^{n-k}$ and the following diagram commutes:

$$\begin{tikzcd}
\mathcal{R}^n/V \arrow[r, "G"] \arrow[d, "\sigma"'] & \mathcal{R}^n/V \arrow[d, "\sigma"] \\
\mathcal{R}^{n-k} \arrow[r, "H"] & \mathcal{R}^{n-k}
\end{tikzcd}$$

Lemma 2.2.1 *$\psi(\lambda)$ is the relative minimal polynomial of $\mathcal{R}^n(\text{mod } V)$ with respect to A if and only if $\psi(\lambda)$ is the minimal polynomial of $\mathcal{R}^{n-k}$ with respect to H.*

Proof. For a polynomial $\phi(\lambda) \in \mathcal{R}[\lambda]$, since $H = \sigma \circ G \circ \sigma^{-1}$, we have

$$\phi(H) = \sigma \circ \phi(G) \circ \sigma^{-1}.$$

Therefore,

$$\phi(H)\sigma(\{\mathbf{x}\}) = 0 \Longleftrightarrow \phi(G)(\{\mathbf{x}\}) = 0, \qquad \forall\, \mathbf{x} \in \mathcal{R}^n.$$

By $G(\{\mathbf{x}\}) = \{A\mathbf{x}\}$ for any vector $\mathbf{x} \in \mathcal{R}^n$, we obtain

$$\phi(G)(\{\mathbf{x}\}) = \{\phi(A)\mathbf{x}\}, \qquad \forall\, \mathbf{x} \in \mathcal{R}^n. \tag{2.5}$$

This implies

$$\phi(G)(\{\mathbf{x}\}) = 0 \Longleftrightarrow \phi(A)\mathbf{x} \equiv 0(\text{mod } V), \qquad \forall\, \mathbf{x} \in \mathcal{R}^n.$$

Hence,

$$\phi(H)\sigma(\{\mathbf{x}\}) = 0 \Longleftrightarrow \phi(A)\mathbf{x} \equiv 0(\text{mod } V), \qquad \forall\, \mathbf{x} \in \mathcal{R}^n. \tag{2.6}$$

Let $\psi_1(\lambda)$ be the relative minimal polynomial of $\mathcal{R}^n(\text{mod } V)$ with respect to A and $\psi_2(\lambda)$ be the minimal polynomial of $\mathcal{R}^{n-k}$ with respect to H. Then we have

$$\psi_1(A)\mathbf{x} \equiv 0(\text{mod } V), \qquad \forall\, \mathbf{x} \in \mathcal{R}^n,$$

and

$$\psi_2(H)\sigma(\{\mathbf{x}\}) = 0, \qquad \forall\, \mathbf{x} \in \mathcal{R}^n.$$

It follows from (2.6) that

$$\psi_2(A)\mathbf{x} \equiv 0(\text{mod } V),\ \forall\, \mathbf{x} \in \mathcal{R}^n,$$

and

$$\psi_1(H)\sigma(\{\mathbf{x}\}) = 0, \qquad \forall\, \mathbf{x} \in \mathcal{R}^n.$$

Therefore $\psi_1(\lambda)|\psi_2(\lambda)$ and $\psi_2(\lambda)|\psi_1(\lambda)$. This implies $\psi_1(\lambda) = \psi_2(\lambda)$ and the proof is completed. ■

Lemma 2.2.2 *$\psi(\lambda)$ is the minimal polynomial of a vector $\mathbf{y} \in \mathcal{R}^{n-k}$ with respect to H if and only if $\psi(\lambda)$ is the relative minimal polynomial of every $\mathbf{x} \in \sigma^{-1}(\mathbf{y})(\text{mod } V)$ with respect to A.*

Proof. For a polynomial $\phi \in \mathcal{R}[\lambda]$, $\phi(H) = \sigma \circ \phi(G) \circ \sigma^{-1}$ implies

$$\phi(H)\mathbf{y} = 0 \Longleftrightarrow \phi(G)(\sigma^{-1}(\mathbf{y})) = 0.$$

Let $\mathbf{x} \in \sigma^{-1}(\mathbf{y})$, then $\sigma^{-1}(\mathbf{y}) = \{\mathbf{x}\}$. Thus

$$\phi(H)\mathbf{y} = 0 \Longleftrightarrow \phi(G)(\{\mathbf{x}\}) = 0.$$

By (2.5),

$$\phi(G)(\{\mathbf{x}\}) = 0 \Longleftrightarrow \{\phi(A)\mathbf{x}\} = 0 \Longleftrightarrow \phi(A)\mathbf{x} \equiv 0(\text{mod } V).$$

Hence

$$\phi(H)\mathbf{y} = 0 \Longleftrightarrow \phi(A)\mathbf{x} \equiv 0(\text{mod } V). \tag{2.7}$$

Let $\psi_1(\lambda)$ be the relative minimal polynomial of $\mathbf{x} \in \sigma^{-1}(\mathbf{y})(\text{mod } V)$ with respect to A, and $\psi_2(\lambda)$ be the minimal polynomial of $\mathbf{y}$ with respect to H. Then we have

$$\psi_1(A)\mathbf{x} \equiv 0(\text{mod } V),$$

and

$$\psi_2(H)\mathbf{y} = 0.$$

By (2.7),

$$\psi_2(A)\mathbf{x} \equiv 0(\text{mod } V),$$

and

$$\psi_1(H)\mathbf{y} = 0.$$

Therefore $\psi_1(\lambda)|\psi_2(\lambda)$ and $\psi_2(\lambda)|\psi_1(\lambda)$. This implies $\psi_1(\lambda) = \psi_2(\lambda)$. ■

Lemma 2.2.3 *If V is a k-dimensional cyclic subspace generated by a regular vector $\mathbf{x}_0$ of A, then $\psi(\lambda)$ is the minimal polynomial of a vector $\mathbf{y} \in \mathcal{R}^{n-k}$ with respect to H if and only if $\psi(\lambda)$ is the minimal polynomial of some vector $\mathbf{z} \in \sigma^{-1}(\mathbf{y})$ with respect to A.*

Proof. We first prove the sufficiency. For a given $\mathbf{x} \in \sigma^{-1}(\mathbf{y})$, there exists a vector $\mathbf{v} \in V$ such that $\mathbf{x} = \mathbf{z} + \mathbf{v}$. It follows that

$$\psi(A)\mathbf{x} = \psi(A)\mathbf{z} + \psi(A)\mathbf{v}.$$

Since $\psi(A)\mathbf{z} = 0$, we have $\psi(A)\mathbf{x} = \psi(A)\mathbf{v}$. Note that $\psi(A)\mathbf{v} \in V$. Hence,

$$\psi(A)\mathbf{x} \equiv 0(\text{mod } V), \qquad \forall\, \mathbf{x} \in \sigma^{-1}(\mathbf{y}).$$

This implies $\psi(\lambda)$ is the relative minimal polynomial of every $\mathbf{x} \in \sigma^{-1}(\mathbf{y})(\text{mod } V)$ with respect to A. By Lemma 2.2.2, $\psi(\lambda)$ is the minimal polynomial of $\mathbf{y}$.

Next we prove the necessity. By Lemma 2.2.2 again, $\psi(\lambda)$ is the relative minimal polynomial of every $\mathbf{x} \in \sigma^{-1}(\mathbf{y}) (\text{mod } V)$ with respect to A. In particular, for a fixed $\mathbf{z}_0 \in \sigma^{-1}(\mathbf{y})$, $\psi(A)\mathbf{z}_0 \in V$. This means that there is a polynomial $\xi(\lambda)$ of degree $k-1$ such that

$$\psi(A)\mathbf{z}_0 = \xi(A)\mathbf{x}_0. \tag{2.8}$$

Let $\phi(\lambda)$ be the minimal polynomial of A, then $\phi(\lambda)$ is divisible by $\psi(\lambda)$. This implies that there exists a polynomial $\eta(\lambda)$ such that

$$\phi(\lambda) = \psi(\lambda)\eta(\lambda). \tag{2.9}$$

By (2.8) and (2.9), we have

$$\eta(A)\xi(A)\mathbf{x}_0 = \phi(A)\mathbf{z}_0 = 0.$$

Since $\mathbf{x}_0$ is a regular vector of A, $\phi(\lambda)$ is the minimal polynomial of $\mathbf{x}_0$. Consequently, $\eta(\lambda)\xi(\lambda)$ is divisible by $\phi(\lambda)$. Therefore, there exists a polynomial $\theta(\lambda)$ such that

$$\eta(\lambda)\xi(\lambda) = \phi(\lambda)\theta(\lambda) = \psi(\lambda)\eta(\lambda)\theta(\lambda),$$

and so, $\xi(\lambda) = \psi(\lambda)\theta(\lambda)$. Let $\mathbf{z} = \mathbf{z}_0 - \theta(A)\mathbf{x}_0$. Then,

$$\psi(A)\mathbf{z} \;=\; \psi(A)(\mathbf{z}_0 - \theta(A)\mathbf{x}_0) \;=\; \psi(A)\mathbf{z}_0 - \xi(A)\mathbf{x}_0 \;=\; 0. \tag{2.10}$$

Let $\widetilde{\psi}(\lambda)$ be the minimal polynomial of $\mathbf{z}$. It follows from (2.10), $\widetilde{\psi}(\lambda)|\psi(\lambda)$. Since $\psi(\lambda)$ is the relative minimal polynomial of $\mathbf{z}$, we also have $\psi(\lambda)|\widetilde{\psi}(\lambda)$. Therefore $\psi(\lambda) = \widetilde{\psi}(\lambda)$, and $\psi(\lambda)$ is the minimal polynomial of $\mathbf{z}$. ■

Lemma 2.2.4 *In Lemma 2.2.3, let m be the degree of $\psi(\lambda)$. Then $\mathbf{x}_0$, $A\mathbf{x}_0$, …, $A^{k-1}\mathbf{x}_0$, $\mathbf{z}$, $A\mathbf{z}$, …, $A^{m-1}\mathbf{z}$ are linearly independent.*

Proof. Suppose $\mathbf{x}_0$, $A\mathbf{x}_0$, …, $A^{k-1}\mathbf{x}_0$, $\mathbf{z}$, $A\mathbf{z}$, …, $A^{m-1}\mathbf{z}$ are linearly dependent. Then there are constants $a_0, a_1, \ldots, a_{k-1}$ and b_0, b_1, …, b_{m-1} such that

$$a_0\mathbf{x}_0 + a_1A\mathbf{x}_0 + \cdots + a_{k-1}A^{k-1}\mathbf{x}_0 + b_0\mathbf{z} + b_1A\mathbf{z} + \cdots + b_{m-1}A^{m-1}\mathbf{z} = 0. \tag{2.11}$$

Let

$$\eta(\lambda) = a_0 + a_1\lambda + \cdots + a_{k-1}\lambda^{k-1},$$

$$\xi(\lambda) = -b_0 - b_1\lambda - \cdots - b_{m-1}\lambda^{m-1},$$

then (2.11) becomes

$$\xi(A)\mathbf{z} = \eta(A)\mathbf{x}_0 \in V.$$

By Lemma 2.2.2, $\psi(\lambda)$ is also the relative minimal polynomial of $\mathbf{z}(\text{mod } V)$. Thus $\psi(\lambda)|\xi(\lambda)$. This leads to a contradiction since $\psi(\lambda)$ is of degree m and $\xi(\lambda)$ is of degree $m-1$. ■

Before we give a new constructive proof of Theorem 2.1.1, we first define the following notations recursively.

Let $n_1 = n$ and $L_1 = A$. Suppose for integer $j \geq 1$, we have defined the linear operator L_j in $\mathcal{R}^{n_j}$. Let $\Psi_j(\lambda) \in \mathcal{R}[\lambda]$ be the minimal polynomial of $\mathcal{R}^{n_j}$ with respect to L_j and $\widetilde{\mathbf{w}}_j \in \mathcal{R}^{n_j}$ be a regular vector, let $\widetilde{W}_j$ be the cyclic subspace generated by $\widetilde{\mathbf{w}}_j$ and $k_j = \dim(\widetilde{W}_j)$. If $n_{j+1} = n_j - k_j > 0$, we define

$$\begin{aligned} G_j \ &: \ \mathcal{R}^{n_j}/\widetilde{W}_j \longrightarrow \mathcal{R}^{n_j}/\widetilde{W}_j, \\ &\{\mathbf{x}\} \longmapsto \{L_j\mathbf{x}\}. \end{aligned} \tag{2.12}$$

Since $\dim(\mathcal{R}^{n_j}/\widetilde{W}_j) = \dim(\mathcal{R}^{n_{j+1}})$, there is an isomorphism

$$\sigma_j : \mathcal{R}^{n_j}/\widetilde{W}_j \longrightarrow \mathcal{R}^{n_{j+1}}. \tag{2.13}$$

Let

$$L_{j+1} = \sigma_j \circ G_j \circ \sigma_j^{-1} : \mathcal{R}^{n_{j+1}} \longrightarrow \mathcal{R}^{n_{j+1}}. \tag{2.14}$$

The above procedure will terminate in finite many steps, say t, that is, when $n_t = k_t$.

Remark 2.2.1 *From the above definition, we have* $n = k_1 + k_2 + \cdots + k_t$.

Lemma 2.2.5 *For the above* $\Psi_j(\lambda)(j = 1, \ldots, t)$, *we have*

$$\Psi_t(\lambda)|\Psi_{t-1}(\lambda)|\cdots|\Psi_1(\lambda). \tag{2.15}$$

Proof. For a given j $(2 \leq j \leq t)$, by Lemma 2.2.1, $\Psi_j(\lambda)$ is the relative minimal polynomial of $\mathcal{R}^{n_{j-1}}(\text{mod } \widetilde{W}_{j-1})$ with respect to L_{j-1} and $\Psi_{j-1}(\lambda)$ is the minimal polynomial of $\mathcal{R}^{n_{j-1}}$ with respect to L_{j-1}. Therefore, $\Psi_j(\lambda)|\Psi_{j-1}(\lambda)$. ■

Lemma 2.2.6 *For* $\widetilde{\mathbf{w}}_j \in \mathcal{R}^{n_j}(j = 1, \dots, t)$ *defined above, there are* $\mathbf{w}_j \in \mathcal{R}^n(j = 1, \dots, t)$ *such that* $\mathbf{w}_1 = \widetilde{\mathbf{w}}_1$ *and*

$$\widetilde{\mathbf{w}}_j = \sigma_{j-1}(\{\cdots \sigma_2(\{\sigma_1(\{\mathbf{w}_j\})\})\cdots\}), \quad j = 2, \dots, t; \tag{2.16}$$

moreover, $\Psi_j(\lambda)$ *is the minimal polynomial of* $\mathbf{w}_j$ *with respect to* $A(= L_1)$ *for* $j = 1$, *…,* t.

Proof. Since $\widetilde{\mathbf{w}}_1 \in \mathcal{R}^n$, we simply take $\mathbf{w}_1 = \widetilde{\mathbf{w}}_1$. By definition, $\Psi_1(\lambda)$ is the minimal polynomial of $\mathbf{w}_1$ with respect to L_1. For $2 \leq j \leq t$, since $\Psi_j(\lambda)$ is the minimal polynomial of $\widetilde{\mathbf{w}}_j$ with respect to L_j, it follows from Lemma 2.2.3 that there is a vector $\mathbf{u}_{j-1}^{(j)} \in \sigma_{j-1}^{-1}(\widetilde{\mathbf{w}}_j)$ such that $\Psi_j(\lambda)$ is the minimal polynomial of $\mathbf{u}_{j-1}^{(j)}$ with respect to L_{j-1}. By Lemma 2.2.3 again, $\Psi_j(\lambda)$ is the minimal polynomial of a vector $\mathbf{u}_{j-2}^{(j)} \in \sigma_{j-2}^{-1}(\mathbf{u}_{j-1}^{(j)})$ with respect to L_{j-2}. Repeating the procedure, we eventually obtain a vector $\mathbf{w}_j = \mathbf{u}_1^{(j)} \in \sigma_1^{-1}(\mathbf{u}_2^{(j)}) \subset \mathcal{R}^n$ such that $\Psi_j(\lambda)$ is the minimal polynomial of $\mathbf{w}_j$ with respect to L_1. It is easy to see from $\widetilde{\mathbf{w}}_j = \sigma_{j-1}(\{\mathbf{u}_{j-1}^{(j)}\})$ $(j = 2, \dots, t)$ that $\mathbf{w}_j$ satisfies (2.16). ■

Lemma 2.2.7 *For* $\mathbf{w}_i(i = 1, \dots, t)$ *obtained in Lemma 2.2.6, let* W_i *be the cyclic subspace generated by* $\mathbf{w}_i$ *for* A, *then*

$$\mathcal{R}^n = W_1 \oplus W_2 \oplus \cdots \oplus W_t.$$

Proof. For $i = 1, \dots, t$, let $n_i = \dim(W_i)$. By Remark 2.2.1, $n = n_1 + \cdots + n_t$. Therefore, we only need to show that

$$\mathbf{w}_1, A\mathbf{w}_1, \dots, A^{k_1-1}\mathbf{w}_1, \mathbf{w}_2, A\mathbf{w}_2, \dots, A^{k_2-1}\mathbf{w}_2, \dots, \mathbf{w}_t, A\mathbf{w}_t, \dots, A^{k_t-1}\mathbf{w}_t$$

are linearly independent. We will show this by means of mathematical induction.

Suppose

$$\mathbf{w}_1, A\mathbf{w}_1, \ldots, A^{k_1-1}\mathbf{w}_1, \mathbf{w}_2, A\mathbf{w}_2, \ldots, A^{k_2-1}\mathbf{w}_2, \ldots, \mathbf{w}_j, A\mathbf{w}_j, \ldots, A^{k_j-1}\mathbf{w}_j$$

are linearly independent, we will show that

$$\mathbf{w}_1, A\mathbf{w}_1, \ldots, A^{k_1-1}\mathbf{w}_1, \mathbf{w}_2, A\mathbf{w}_2, \ldots, A^{k_2-1}\mathbf{w}_2, \ldots, \mathbf{w}_{j+1}, A\mathbf{w}_{j+1}, \ldots, A^{k_{j+1}-1}\mathbf{w}_{j+1}$$

are linearly independent. In fact, if

$$\mathbf{w}_1, A\mathbf{w}_1, \ldots, A^{k_1-1}\mathbf{w}_1, \mathbf{w}_2, A\mathbf{w}_2, \ldots, A^{k_2-1}\mathbf{w}_2, \ldots, \mathbf{w}_{j+1}, A\mathbf{w}_{j+1}, \ldots, A^{k_{j+1}-1}\mathbf{w}_{j+1}$$

are linearly dependent, then there are polynomials $\xi_1(\lambda)$, $\xi_2(\lambda)$, …, $\xi_{j+1}(\lambda)$ of degrees $k_1 - 1$, $k_2 - 1$, …, $k_{j+1} - 1$, respectively, such that

$$\xi_{j+1}(A)\mathbf{w}_{j+1} = \xi_1(A)\mathbf{w}_1 + \xi_2(A)\mathbf{w}_2 + \cdots + \xi_j(A)\mathbf{w}_j. \tag{2.17}$$

By the definitions of L_j, G_j and σ_j, we have for any vector $\mathbf{x} \in \mathcal{R}^{n_j}$

$$\begin{aligned} \sigma_j(\{L_j\mathbf{x}\}) &= \sigma_j(G_j(\{\mathbf{x}\})) = \sigma_j \circ G_j \circ \sigma_j^{-1} \circ \sigma_j(\{\mathbf{x}\}) \\ &= L_{j+1}\sigma_j(\{\mathbf{x}\}), \quad j = 1, \ldots, t-1, \end{aligned} \tag{2.18}$$

and

$$\sigma_j(\{\widetilde{\mathbf{w}}_j\}) = 0, \qquad j = 1, \ldots, t-1. \tag{2.19}$$

By $L_1 = A$ and (2.17), we have

$$\begin{aligned} \{\xi_{j+1}(L_1)\mathbf{w}_{j+1}\} &= \{\xi_1(L_1)\mathbf{w}_1\} + \cdots + \{\xi_j(L_1)\mathbf{w}_j\} \\ &= \{\xi_2(L_1)\mathbf{w}_2\} + \cdots + \{\xi_j(L_1)\mathbf{w}_j\}. \end{aligned}$$

Therefore,

$$\sigma_1(\{\xi_{j+1}(L_1)\mathbf{w}_{j+1}\}) = \sigma_1(\{\xi_2(L_1)\mathbf{w}_2\}) + \cdots + \sigma_1(\{\xi_j(L_1)\mathbf{w}_j\}).$$

It follows from (2.18) and $\sigma_1(\{\mathbf{w}_2\}) = \widetilde{\mathbf{w}}_2$ (refer to (2.16)) that

$$\begin{aligned}\xi_{j+1}(L_2)\sigma_1(\{\mathbf{w}_{j+1}\}) &= \xi_2(L_2)\sigma_1(\{\mathbf{w}_2\}) + \cdots + \xi_j(L_2)\sigma_1(\{\mathbf{w}_j\}) \\ &= \xi_2(L_2)\widetilde{\mathbf{w}}_2 + \cdots + \xi_j(L_2)\sigma_1(\{\mathbf{w}_j\}).\end{aligned}$$

So,

$$\{\xi_{j+1}(L_2)\sigma_1(\{\mathbf{w}_{j+1}\})\} = \{\xi_3(L_2)\sigma_1(\{\mathbf{w}_3\})\} + \cdots + \{\xi_j(L_2)\sigma_1(\{\mathbf{w}_j\})\}.$$

Furthermore,

$$\sigma_2(\{\xi_{j+1}(L_2)\sigma_1(\{\mathbf{w}_{j+1}\})\}) = \sigma_2(\{\xi_3(L_2)\sigma_1(\{\mathbf{w}_3\})\}) + \cdots + \sigma_2(\{\xi_j(L_2)\sigma_1(\{\mathbf{w}_j\})\}).$$

From (2.18) and $\sigma_2(\sigma_1(\{\mathbf{w}_3\}) = \widetilde{\mathbf{w}}_3$ (refer to (2.16)), we obtain

$$\xi_{j+1}(L_3)\sigma_2(\{\sigma_1(\{\mathbf{w}_{j+1}\})\}) = \xi_3(L_3)\widetilde{\mathbf{w}}_3 + \cdots + \xi_j(L_3)\sigma_2(\{\sigma_1(\{\mathbf{w}_j\})\}).$$

Repeating the same procedure, eventually we have

$$\xi_{j+1}(L_{j+1})\widetilde{\mathbf{w}}_{j+1} = 0.$$

Obviously, the above procedure is equivalent to acting on both sides of (2.17) by $\sigma_j(\{\cdots\sigma_2(\{\sigma_1(\{\cdot\})\})\cdots\})$.

Since $\Psi_{j+1}(\lambda)$ is the minimal polynomial of $\widetilde{\mathbf{w}}_{j+1}$ with respect to L_{j+1}, we have

$$\Psi_{j+1}(\lambda)|\xi_{j+1}(\lambda).$$

But the degree of $\xi_{j+1}(\lambda)$ is less than the degree of $\Psi_{j+1}(\lambda)$. Therefore we have

$$\xi_{j+1}(\lambda) = 0.$$

This together with (2.17) contradicts the assumption of the induction. Therefore,

$$\mathbf{w}_1, A\mathbf{w}_1, \ldots, A^{k_1-1}\mathbf{w}_1, \mathbf{w}_2, A\mathbf{w}_2, \ldots, A^{k_2-1}\mathbf{w}_2, \ldots, \mathbf{w}_{j+1}, A\mathbf{w}_{j+1}, \ldots, A^{k_{j+1}-1}\mathbf{w}_{j+1}$$

are linearly independent. ■

Remark 2.2.2 *Note that the relations between* $\widetilde{\mathbf{w}}_j, \mathbf{u}_j^{(j+1)}, \ldots, \mathbf{u}_j^{(t)}$ *in* $\mathcal{R}^{n_j}$ *for* L_j *are the same as the relations between* $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_t$ *in* $\mathcal{R}^n$ *for* A *in the proof of Lemma 2.2.6. Therefore, for* $j = 2, \ldots, t$,

$$\widetilde{\mathbf{w}}_j, L_j\widetilde{\mathbf{w}}_j, \ldots, L_j^{k_j-1}\widetilde{\mathbf{w}}_j, \mathbf{u}_j^{(j+1)}, L_j\mathbf{u}_j^{(j+1)}, \ldots, L_j^{k_{j+1}-1}\mathbf{u}_j^{(j+1)}, \ldots, \mathbf{u}_j^{(t)}, L_j\mathbf{u}_j^{(t)}, \ldots, L_j^{k_t-1}\mathbf{u}_j^{(t)}$$

are also linearly independent.

Lemma 2.2.8 *For* $W_j (j = 1, \ldots, t)$ *defined in Lemma 2.2.7,* $\Psi_{j+1}(\lambda)$ *is the relative minimal polynomial of* $\mathcal{R}^n(\text{mod } (W_1 \oplus W_2 \oplus \cdots \oplus W_j))$ *for* $j = 1, \ldots, t-1$.

Proof. If $j = 1$, then $\mathbf{w}_1 = \widetilde{\mathbf{w}}_1$ and $W_1 = \widetilde{W_1}$. By Lemma 2.2.1, $\Psi_2(\lambda)$ is the relative minimal polynomial of $\mathcal{R}^n(\text{mod } W_1)$ with respect to L_1. Now for a given j $(2 \le j \le t-1)$ and any vector $\mathbf{x} \in \mathcal{R}^n$, there are, by Lemma 2.2.7, polynomials $\xi_1(\lambda)$, $\xi_2(\lambda)$, …, $\xi_t(\lambda)$ with degree $k_1 - 1$, $k_2 - 1$, …, $k_t - 1$, respectively, such that

$$\Psi_{j+1}(A)\mathbf{x} = \xi_1(A)\mathbf{w}_1 + \xi_2(A)\mathbf{w}_2 + \cdots + \xi_t(A)\mathbf{w}_t. \tag{2.20}$$

Let

$$\mathbf{y} = \sigma_j(\{\cdots\sigma_2(\{\sigma_1(\{\mathbf{x}\})\})\cdots\}),$$

then $\mathbf{y} \in \mathcal{R}^{n_{j+1}}$. Recall that $\Psi_{j+1}(\lambda)$ is the minimal polynomial of $\mathcal{R}^{n_{j+1}}$ for L_{j+1}. Acting on both sides of (2.20) by $\sigma_j(\{\cdots\sigma_2(\{\sigma_1(\{\cdot\})\})\cdots\})$ and utilizing (2.18), we have

$$\begin{aligned} 0 \;=\; \Psi_{j+1}(L_{j+1})\mathbf{y} \;&=\; \sigma_j(\{\cdots\sigma_2(\{\sigma_1(\{\Psi_{j+1}(A)\mathbf{x}\})\})\cdots\}) \\ &=\; \xi_{j+1}(L_{j+1})\widetilde{\mathbf{w}}_{j+1} + \xi_{j+2}(L_{j+1})\mathbf{u}_{j+1}^{(j+2)} + \cdots + \xi_t(L_{j+1})\mathbf{u}_{j+1}^{(t)}. \end{aligned}$$

It follows immediately from Remark 2.2.2,

$$\xi_{j+1}(\lambda) = \cdots = \xi_t(\lambda) = 0.$$

Hence (2.20) implies that for all $\mathbf{x} \in \mathcal{R}^n$

$$\Psi_{j+1}(A)\mathbf{x} \equiv 0(\text{mod } (W_1 \oplus W_2 \oplus \cdots \oplus W_j)). \qquad (2.21)$$

Let $\widetilde{\Psi}_{j+1}(\lambda)$ be the relative minimal polynomial of $\mathcal{R}^n(\text{mod } (W_1 \oplus W_2 \oplus \cdots \oplus W_j))$ with respect to A, then $\widetilde{\Psi}_{j+1}(\lambda)|\Psi_{j+1}(\lambda)$. By Lemma 2.2.6, $\Psi_{j+1}(\lambda)$ is the minimal polynomial of $\widetilde{\mathbf{w}}_{j+1}$ with respect to L_{j+1} and

$$\widetilde{\Psi}(L_{j+1})\widetilde{\mathbf{w}}_{j+1} = \sigma_j(\{\cdots \sigma_2(\{\sigma_1(\{\widetilde{\Psi}(A)\mathbf{w}_{j+1}\})\}) \cdots\}) = 0.$$

Thus $\Psi_{j+1}(\lambda)|\widetilde{\Psi}_{j+1}(\lambda)$. Consequently, $\Psi_{j+1}(\lambda) = \widetilde{\Psi}_{j+1}(\lambda)$. Therefore $\Psi_{j+1}(\lambda)$ is the relative minimal polynomial of $\mathcal{R}^n(\text{mod } (W_1 \oplus W_2 \oplus \cdots \oplus W_j))$ with respect to A. ■

Proof of Theorem 2.1.1. Combining Lemma 2.2.5, Lemma 2.2.6, Lemma 2.2.7 and Lemma 2.2.8, the proof of Theorem 2.1.1 follows immediately. ■

Remark 2.2.3 *From the results obtained in this section, it is not necessary to modify each regular vector* $\widetilde{\mathbf{w}}_j$ *of* L_j *in* $\mathcal{R}^{n_j}$ *to a regular vector* $\mathbf{w}_j$ *in* $\mathcal{R}^n (j = 2, \ldots, t)$ *if we only need to determine the minimal polynomials* $\Psi_1(\lambda)$, $\Psi_2(\lambda)$, $\ldots$, $\Psi_t(\lambda)$ *defined in Theorem 2.1.1.*

2.3 A constructive proof of Theorem 2.1.2

In this section, we give a constructive proof of Theorem 2.1.2.

Lemma 2.3.1 [14] *For a given* $\Psi(\lambda) \in \mathcal{R}[\lambda]$, *via finitely many purely rational operations one can always find a squarefree relatively prime basis* $\{\varphi_1, \varphi_2, \ldots, \varphi_s\} \subset \mathcal{R}[\lambda]$ *of* $\Psi(\lambda)$ *such that*

$$\Psi(\lambda) = [\varphi_1(\lambda)]^{m_1}[\varphi_2(\lambda)]^{m_2} \cdots [\varphi_s(\lambda)]^{m_s}, \qquad (2.22)$$

where $m_i > 0\,(i = 1, 2, \ldots, s)$ *are integers and* $m_1 < m_2 < \cdots < m_s$.

Proof. Group all simple factors of $\Psi(\lambda)$ in the complex number field $\mathcal{C}$ which have the same multiple, then $\Psi(\lambda)$ can be represented by

$$\Psi(\lambda) = [\varphi_1(\lambda)]^{m_1}[\varphi_2(\lambda)]^{m_2}\cdots[\varphi_s(\lambda)]^{m_s}, \tag{2.23}$$

with $m_i > 0$ for $i = 1, 2, \ldots, s$ and $m_1 < m_2 < \cdots < m_s$, where $\gcd(\varphi_i, \frac{d}{d\lambda}\varphi_i) = 1$ and $\gcd(\varphi_i, \varphi_j) = 1$ for $i \neq j$; $i, j = 1, 2, \ldots, s$.

Now we only need to show that φ_1, φ_2, ..., φ_s are all in $\mathcal{R}[\lambda]$ and all of them can be obtained via finitely many rational operations in $\mathcal{R}[\lambda]$.

Since $\varphi_i(\lambda)$ is a factor of order m_i of $\Psi(\lambda)$, $\varphi_i(\lambda)$ is a factor of order $(m_i - k)$ of $\frac{d^k}{d\lambda^k}\Psi(\lambda)$ for $k \leq m_i$. Let

$$d_l(\lambda) = \gcd\left(\Psi(\lambda), \frac{d^l}{d\lambda^l}\Psi(\lambda)\right), \qquad l = 1, \ldots, m_s. \tag{2.24}$$

Apparently, $d_l(\lambda)$ has the following form

$$d_l(\lambda) = \varphi_1^{m_1-l}(\lambda)\varphi_2^{m_2-l}(\lambda)\cdots\varphi_s^{m_s-l}(\lambda), \tag{2.25}$$

with the restriction $[\varphi_i(\lambda)]^{m_i-l} \equiv 1$ when $m_i - l \leq 0$, $i = 1, 2, \ldots, s$.

Denote $d_{i-1}(\lambda)/d_i(\lambda)$ by $\eta_i(\lambda)$ for $i = 1, \ldots, m_s - 1$ with $d_0(\lambda) = \Psi(\lambda)$, then

$$\begin{cases} \eta_1(\lambda) & = & d_0(\lambda)/d_1(\lambda) & = & \varphi_1(\lambda)\varphi_2(\lambda)\cdots\varphi_s(\lambda), \\ & & \vdots & & \\ \eta_{m_1}(\lambda) & = & d_{m_1-1}(\lambda)/d_{m_1}(\lambda) & = & \varphi_1(\lambda)\varphi_2(\lambda)\cdots\varphi_s(\lambda), \\ \eta_{m_1+1}(\lambda) & = & d_{m_1}(\lambda)/d_{m_1+1}(\lambda) & = & \varphi_2(\lambda)\cdots\varphi_s(\lambda), \\ & & \vdots & & \\ \eta_{m_2}(\lambda) & = & d_{m_2-1}(\lambda)/d_{m_2}(\lambda) = \varphi_2(\lambda)\cdots\varphi_s(\lambda), & & \\ \eta_{m_2+1}(\lambda) & = & d_{m_2}(\lambda)/d_{m_2+1}(\lambda) & = & \varphi_3(\lambda)\cdots\varphi_s(\lambda), \\ & & \vdots & & \\ \eta_{m_s}(\lambda) & = & d_{m_s-1}(\lambda)/d_{m_s}(\lambda) & = & \varphi_s(\lambda). \end{cases} \tag{2.26}$$

Hence,

$$\begin{cases} \varphi_1(\lambda) & = \eta_{m_1}(\lambda)/\eta_{m_2}(\lambda), \\ \varphi_2(\lambda) & = \eta_{m_2}(\lambda)/\eta_{m_3}(\lambda), \\ & \vdots \\ \varphi_{s-1}(\lambda) & = \eta_{m_{s-1}}(\lambda)/\eta_{m_s}(\lambda), \\ \varphi_s(\lambda) & = \eta_{m_s}(\lambda). \end{cases} \tag{2.27}$$

Note that (2.24), (2.26) and (2.27) can all be accomplished by finitely many rational operations over $\mathcal{R}[\lambda]$, hence φ_1, φ_2, ..., φ_s must belong to $\mathcal{R}[\lambda]$. ∎

Lemma 2.3.2 [14] *For given $\Psi(\lambda)$ and $\Phi(\lambda)$ in $\mathcal{R}[\lambda]$ with $\Phi(\lambda)|\Psi(\lambda)$, one can always find a common squarefree relatively prime basis $\{\varphi_1, \varphi_2, \ldots, \varphi_s\}$ via finitely many purely rational operations such that*

$$\Psi(\lambda) = [\varphi_1(\lambda)]^{m_1}[\varphi_2(\lambda)]^{m_2}\cdots[\varphi_s(\lambda)]^{m_s}, \tag{2.28}$$

and

$$\Phi(\lambda) = [\varphi_1(\lambda)]^{n_1}[\varphi_2(\lambda)]^{n_2}\cdots[\varphi_s(\lambda)]^{n_s} \tag{2.29}$$

with $n_i \le m_i$, for $i = 1, \ldots, s$.

Proof. By Lemma 2.3.1, we may assume $\Psi(\lambda)$ has a squarefree relatively prime basis $\{\eta_1, \eta_2, \ldots, \eta_k\}$ with decomposition

$$\Psi(\lambda) = [\eta_1(\lambda)]^{p_1}[\eta_2(\lambda)]^{p_2}\cdots[\eta_k(\lambda)]^{p_k},$$

and $\Phi(\lambda)$ can be expressed as

$$\Phi(\lambda) = [\eta_1(\lambda)]^{q_1}[\eta_2(\lambda)]^{q_2}\cdots[\eta_k(\lambda)]^{q_k}\Theta_1(\lambda),$$

where each $q_i \le p_i$ is a nonnegative integer and Θ_1 is not divisible by η_i for $i = 1, 2, \ldots, k$.

If $\tau_i = \gcd(\eta_i, \Theta_1)$ has positive degree, then η_i has a nontrivial decomposition

$$\eta_i = \tau_i \, \xi_i. \tag{2.30}$$

For those i's with $\tau_i = 1$, let $\xi_i = \eta_i$. Then $\Theta_1(\lambda)$ can be rewritten as

$$\Theta_1 = \tau_1^{l_1} \tau_2^{l_2} \cdots \tau_k^{l_k} \Theta_2,$$

where each l_i $(i = 1, 2, \ldots, k)$ is a nonnegative integer. Since $\Phi(\lambda)|\Psi(\lambda)$, not all the τ_i's can have zero degree, and hence, $\deg \Theta_2 < \deg \Theta_1$. Replacing τ_i and ξ_i for η_i, a new squarefree relatively prime basis of Ψ contained in $\{\tau_1, \xi_1, \tau_2, \xi_2, \ldots, \tau_k, \xi_k\}$ is obtained such that

$$\Psi = \tau_1^{p_1} \xi_1^{p_1} \tau_2^{p_2} \xi_2^{p_2} \cdots \tau_k^{p_k} \xi_k^{p_k},$$

and

$$\Phi = \tau_1^{q_1 + l_1} \xi_1^{q_1} \tau_2^{q_2 + l_2} \xi_2^{q_2} \cdots \tau_k^{q_k + l_k} \xi_k^{q_k} \Theta_2.$$

Repeating the preceding procedure, a common squarefree relatively prime basis $\{\varphi_1, \ldots, \varphi_s\}$ of $\Psi(\lambda)$ can be obtained, via finitely many steps, for which

$$\Psi(\lambda) = [\varphi_1(\lambda)]^{m_1} [\varphi_2(\lambda)]^{m_2} \cdots [\varphi_s(\lambda)]^{m_s}, \tag{2.31}$$

and

$$\Phi(\lambda) = [\varphi_1(\lambda)]^{n_1} [\varphi_2(\lambda)]^{n_2} \cdots [\varphi_s(\lambda)]^{n_s} \Theta(\lambda) \tag{2.32}$$

with $\gcd(\varphi_i, \Theta) = 1$, $i = 1, 2, \ldots, s$. But then $\Theta(\lambda)$ must be a constant. Otherwise, since $\Phi(\lambda)|\Psi(\lambda)$ implies $\Theta(\lambda)|\Psi(\lambda)$, there exists at least one $\varphi_i (1 \leq i \leq s)$ such that $\gcd(\varphi_i, \Theta) \neq 1$. ■

Remark 2.3.1 *Even if $\Psi(\lambda)$ is not divisible by $\Phi(\lambda)$ in Lemma 2.3.2, a common squarefree relatively prime basis of both $\Psi(\lambda)$ and $\Phi(\lambda)$ can still be obtained via finitely many purely rational operations. In fact, if $\Theta(\lambda)$ in* (2.32) *is not a constant, then, by Lemma 2.3.1, $\Theta(\lambda)$ has a squarefree relatively prime basis $\{\tau_1, \tau_2, \ldots, \tau_l\}$. Since*

$\gcd(\varphi_i(\lambda), \Theta) = 1$ *for* $i = 1, 2, \ldots, s$, $\{\varphi_1, \varphi_2, \ldots, \varphi_s, \tau_1, \tau_2, \ldots, \tau_l\}$ *is a squarefree relatively prime basis of* Φ. *Obviously, it is also a squarefree relatively prime basis of* Ψ.

Proof of Theorem 2.1.2[14]. The assertion of the theorem is a direct consequence of the above lemma by mathematical induction.

Suppose the assertion of the theorem holds for $2 \leq j < t$, that is, $\{\Psi_i\}_{i=1}^{j}$ has a common squarefree relatively prime basis $\{\eta_1, \eta_2, \ldots, \eta_s\}$. Since $\{\eta_1, \eta_2, \ldots, \eta_s\}$ is a squarefree relatively prime basis of Ψ_j and $\Psi_{j+1} | \Psi_j$, a new squarefree relatively prime basis of Ψ_j and Ψ_{j+1} can be obtained by repeating the same procedure described in Lemma 2.3.2. Moreover, each η_i is either in the new squarefree relatively prime basis or can be rewritten as a product of polynomials in the new squarefree relatively prime basis. It follows that the new squarefree relatively prime basis $\{\varphi_1, \varphi_2, \ldots, \varphi_r\}$ of Ψ_j and Ψ_{j+1} is also a common squarefree relatively prime basis of $\{\Psi_i\}_{i=1}^{j-1}$. ■

Remark 2.3.2 *Squarefree decompositions of the minimal polynomials* $\Psi_1, \Psi_2, \ldots, \Psi_t$ *in Theorem 2.1.2 can be implemented by using symbolic computation. From Theorem 2.1.3, as long as those decompositions in* (2.3) *are achieved, the structure of the Jordan normal form of* A *can be determined exactly, without knowing the zeros of* φ_k*'s. The approximation of those zeros of* φ_k, $k = 1, \ldots, r$, *has no influence on this structure.*

CHAPTER 3

THE ALGORITHM

Following what was described in the last chapter, we outline our algorithm for determining the Jordan normal form of an $n \times n$ matrix A in $\mathcal{R}^n$.

- **Step A.** Find the minimal polynomials $\Psi_1(\lambda), \Psi_2(\lambda), \ldots, \Psi_t(\lambda)$ in Theorem 2.1.1.

  A0. $j = 1$, $n_1 = n$, $A_1 = A$.

  A1. Find a regular vector $\widetilde{\mathbf{w}}_j$ in $\mathcal{R}^{n_j}$ and its minimal polynomial $\Psi_j(\lambda)$ with respect to A_j. Denote the degree of $\Psi_j(\lambda)$ by k_j.

  A2. Let $n_{j+1} = n_j - k_j$. If $n_{j+1} = 0$, then Step **A** is finished.

  A3. Extend $\widetilde{\mathbf{w}}_j, A_j\widetilde{\mathbf{w}}_j, \ldots, A_j^{k_j-1}\widetilde{\mathbf{w}}_j$ to a basis $\widetilde{\mathbf{w}}_j, A_j\widetilde{\mathbf{w}}_j, \ldots, A_j^{k_j-1}\widetilde{\mathbf{w}}_j$, $\mathbf{v}_{k_j+1}$, $\mathbf{v}_{k_j+2}$, $\ldots$, $\mathbf{v}_{n_j}$ of $\mathcal{R}^{n_j}$. Then $\{\mathbf{v}_{k_j+1}\}$, $\{\mathbf{v}_{k_j+2}\}$, $\cdots$, $\{\mathbf{v}_{n_j}\}$ form a basis of the quotient space $\mathcal{R}^{n_j}/\widetilde{W}_j$, where $\widetilde{W}_j$ is the cyclic subspace generated by $\widetilde{\mathbf{w}}_j$. With this basis, we may define the isomorphism σ_j as follows:

$$\begin{aligned} \sigma_j \ : \ \mathcal{R}^{n_j}/\widetilde{W}_j &\longrightarrow \mathcal{R}^{n_{j+1}}, \\ \mathbf{y} &\longmapsto (c_1, \ldots, c_{n_{j+1}})^T, \end{aligned}$$

  where $\mathbf{y} = c_1\{\mathbf{v}_{k_j+1}\} + c_2\{\mathbf{v}_{k_j+2}\} + \cdots + c_{n_{j+1}}\{\mathbf{v}_{k_j+n_{j+1}}\}$.

A4. Compute the matrix A_{j+1} of the linear operator $L_{j+1} = \sigma_j \circ G_j \circ \sigma_j^{-1}$: $\mathcal{R}^{n_{j+1}} \longrightarrow \mathcal{R}^{n_{j+1}}$ under the standard basis $\{e_1^{(j+1)}, e_2^{(j+1)}, \ldots, e_{n_{j+1}}^{(j+1)}\}$ of $\mathcal{R}^{n_{j+1}}$.

A5. Let $j = j + 1$ and go to **A1**.

- **Step B.** Determine the structure of the Jordan blocks in the Jordan normal form of A.

B1. Find a common squarefree relatively prime basis $\{\varphi_1, \varphi_2, \ldots, \varphi_r\}$ of Ψ_1, Ψ_2, $\ldots$, Ψ_{t-1}, Ψ_t in $\mathcal{R}[\lambda]$ such that

$$\begin{cases} \Psi_1 & = \varphi_1^{m_{1,1}} \varphi_2^{m_{1,2}} \cdots \varphi_r^{m_{1,r}}, \\ \Psi_2 & = \varphi_1^{m_{2,1}} \varphi_2^{m_{2,2}} \cdots \varphi_r^{m_{2,r}}, \\ & \cdots\cdots \\ \Psi_{t-1} & = \varphi_1^{m_{t-1,1}} \varphi_2^{m_{t-1,2}} \cdots \varphi_r^{m_{t-1,r}}, \\ \Psi_t & = \varphi_1^{m_{t,1}} \varphi_2^{m_{t,2}} \cdots \varphi_r^{m_{t,r}}; \end{cases} \tag{3.1}$$

where $m_{1,k} \geq m_{2,k} \geq \cdots \geq m_{t,k} \geq 0$ for $k = 1, 2, \ldots, r$. To achieve this, first use Lemma 2.3.1 to obtain a squarefree relatively prime basis of Ψ_1; secondly, modify the squarefree relatively prime basis of Ψ_1 so that it becomes a common squarefree relatively prime basis for both Ψ_1 and Ψ_2; then apply Lemma 2.3.2 inductively to modify the known squarefree relatively prime basis of $\Psi_1, \Psi_2, \cdots, \Psi_{k-1}$ to a new squarefree relatively prime basis of $\Psi_1, \Psi_2, \cdots, \Psi_k$ until $k = t$.

B2. From Theorem 2.1.3, corresponding to each $\varphi_k^{m_{i,k}}$ of (3.1) for $k = 1, \ldots, r$ and $i = 1, \ldots, t$; there are $p_k = \deg \varphi_k$ Jordan blocks of order $m_{i,k}$ of the form

$$\begin{pmatrix} \lambda_{k,l} & 1 & & 0 \\ & \lambda_{k,l} & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_{k,l} \end{pmatrix}_{m_{i,k} \times m_{i,k}} , \qquad l = 1, \ldots, p_k, \tag{3.2}$$

in the Jordan normal form of A, where the $\lambda_{k,l}$'s are zeros of φ_k. Notice that the *structure* of the Jordan normal form of A has been determined without knowing those $\lambda_{k,l}$'s.

- **Step C.** Calculate the Jordan normal form of A.

After the structure of the Jordan normal form of A has been determined in Step **B**, the remaining task is to approximate all zeros of the polynomials $\varphi_1, \varphi_2, \cdots, \varphi_r$ obtained in Step **B1** to fill in the diagonal entries of the matrices in (3.2).

CHAPTER 4

IMPLEMENTATION OF THE ALGORITHM

In this chapter, we will discuss in detail the implementation of the algorithm outlined in Chapter 3, and present an example to demonstrate the implementation of the algorithm.

4.1 Implementation of Step A

Let $n_1 = n$ and $A_1 = A$. For $j = 1, 2, \ldots$, choose a vector $\mathbf{w} \in \mathcal{R}^{n_j}$ at random. Consider the $n_j \times (2n_j + 1)$ matrix

$$H_j = (\mathbf{w}, A_j\mathbf{w}, \ldots, A_j^{n_j}\mathbf{w},\ I_{n_j}), \tag{4.1}$$

where $I_{n_j} = (e_1^{(j)}, \cdots, e_{n_j}^{(j)})$ is the $n_j \times n_j$ identity matrix in $\mathcal{R}^{n_j}$. In general, we denote the $k \times k$ identity matrix in $\mathcal{R}^k$ by I_k.

Suppose r_j is the largest integer for which $\mathbf{w}, A_j\mathbf{w}, \ldots, A_j^{r_j-1}\mathbf{w}$ are linearly independent. In general, Gaussian elimination with row pivoting can be used to reduce H_j to the following form

$$H_j^* = \left(\begin{array}{cccccccc} h_{1,1} & \cdots & h_{1,r_j} & h_{1,r_j+1} & * & \cdots & * & \\ & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \\ 0 & & h_{r_j,r_j} & h_{r_j,r_j+1} & * & \cdots & * & I^*_{n_j} \\ & & & 0 & 0 & \cdots & 0 & \\ & 0 & & \vdots & \vdots & \ddots & \vdots & \\ & & & 0 & 0 & \cdots & 0 & \\ \multicolumn{3}{c}{\underbrace{\hphantom{h_{1,1}\;\cdots\;h_{1,r_j}}}_{r_j}} & & \multicolumn{3}{c}{\underbrace{\hphantom{0\;\cdots\;0}}_{n_j-r_j}} & \end{array}\right) \tag{4.2}$$

where $h_{i,i} \neq 0$ for $i = 1, \ldots, r_j$, and $I^*_{n_j}$ is the $n_j \times n_j$ matrix resulting from applying the above Gaussian elimination process on I_{n_j}. Notice that if computation is symbolic, then pivoting can be replaced by finding a nonzero element in Gaussian elimination process. In Appendix, an efficient algorithm of reducing H_j to H_j^* is developed to reduce the memory storage and drastically speed up the computation.

Solving the linear system

$$\begin{pmatrix} h_{1,1} & h_{1,2} & \cdots & h_{1,r_j} \\ & h_{2,2} & \cdots & h_{2,r_j} \\ & & \ddots & \vdots \\ 0 & & & h_{r_j,r_j} \end{pmatrix} \begin{pmatrix} a_0^{(j)} \\ a_1^{(j)} \\ \vdots \\ a_{r_j-1}^{(j)} \end{pmatrix} = \begin{pmatrix} h_{1,r_j+1} \\ h_{2,r_j+1} \\ \vdots \\ h_{r_j,r_j+1} \end{pmatrix} \tag{4.3}$$

yields the minimal polynomial of $\mathbf{w}$

$$\Psi(\lambda) = \lambda^{r_j} - (a_{r_j-1}^{(j)} \lambda^{r_j-1} + \cdots + a_1^{(j)} \lambda + a_0^{(j)}). \tag{4.4}$$

Let $n_{j+1} = n_j - r_j$. If $n_{j+1} = 0$, then $\widetilde{\mathbf{w}}_j = \mathbf{w}$, $\Psi_j(\lambda) = \Psi(\lambda)$ is the minimal polynomial of $\mathcal{R}^{n_j}$ with $k_j = r_j$, and

$$\mathcal{R}^{n_j} = \widetilde{W}_j = \text{Span}\{\widetilde{\mathbf{w}}_j, A_j \widetilde{\mathbf{w}}_j, \ldots, A_j^{r_j-1} \widetilde{\mathbf{w}}_j\}.$$

Step **A** is finished.

Now let us assume $n_{j+1} > 0$. For the $n_j \times n_j$ matrix $I^*_{n_j}$ in (4.2), write $I^*_{n_j} = (u_1^{(j)}, \cdots, u_{n_j}^{(j)})$. It is easy to see that there are n_{j+1} columns of $I^*_{n_j}$, say $u_{i_1}^{(j)}, \ldots, u_{i_{n_{j+1}}}^{(j)}$, for which $u_{i_1}^{(j)} = e_{r_j+1}^{(j)}, u_{i_2}^{(j)} = e_{r_j+2}^{(j)}, \ldots, u_{i_{n_{j+1}}}^{(j)} = e_{n_j}^{(j)}$. For instance, when H_j^* is achieved by applying Gaussian elimination without requiring row pivoting to H_j then

$$I^*_{n_j} = (u_1^{(j)}, \cdots, u_{r_j}^{(j)}, e_{r_j+1}^{(j)}, \cdots, e_{n_j}^{(j)}),$$

namely,

$$u_{i_p}^{(j)} = u_{r_j+p}^{(j)} = e_{r_j+p}^{(j)}, \qquad p = 1, \ldots, n_{j+1}.$$

The same consequence takes place if the pivoting only occur within the first r_j rows. If, during the elimination process, say only the pth row ($p \leq r_j$) and the $(r_j + k)$th row have been interchanged, then

$$I^*_{n_j} = (\, u_1^{(j)}, \cdots, u_{p-1}^{(j)}, e_{r_j+k}^{(j)}, u_{p+1}^{(1)}, \cdots, u_{r_j}^{(j)}, e_{r_j+1}^{(j)}, \cdots, e_{r_j+k-1}^{(j)}, u_{r_j+k}^{(j)}, e_{r_j+k+1}^{(j)}, \cdots, e_{n_j}^{(j)}\,),$$

that is,

$$u_{i_p}^{(j)} = u_{r_j+p}^{(j)} = e_{r_j+p}^{(j)} \quad \text{for } p = 1, \ldots, n_{j+1} \quad \text{except } p = k,$$

while $u_{i_k}^{(j)} = u_p^{(j)} = e_{r_j+k}^{(j)}$.

When Gaussian elimination is applied to transform H_j to H_j^*, in essence, a nonsingular matrix P_j is constructed making $P_j H_j = H_j^*$. Accordingly,

$$P_j(\mathbf{w}, A_j\mathbf{w}, \cdots, A_j^{r_j-1}\mathbf{w}) = \begin{pmatrix} h_{1,1} & \cdots & h_{1,r_j} \\ & \ddots & \vdots \\ 0 & & h_{r_j,r_j} \\ & 0 & \end{pmatrix} = \begin{pmatrix} C \\ 0 \end{pmatrix}$$

where

$$C = \begin{pmatrix} h_{1,1} & \cdots & h_{1,r_j} \\ & \ddots & \vdots \\ 0 & & h_{r_j,r_j} \end{pmatrix},$$

and

$$P_j = P_j I_{n_j} = P_j(e_1^{(j)},\cdots,e_{n_j}^{(j)}) = (u_1^{(j)},\cdots,u_{n_j}^{(j)}) = I_{n_j}^{*} \tag{4.5}$$

with $P_j e_i^{(j)} = u_i^{(j)}$, $i = 1,\ldots,n_j$. In particular,

$$\begin{cases} P_j e_{i_1}^{(j)} = u_{i_1}^{(j)} = e_{r_j+1}^{(j)}, \\ P_j e_{i_2}^{(j)} = u_{i_2}^{(j)} = e_{r_j+2}^{(j)}, \\ \qquad\vdots \\ P_j e_{i_{n_{j+1}}}^{(j)} = u_{i_{n_{j+1}}}^{(j)} = e_{n_j}^{(j)}. \end{cases} \tag{4.6}$$

It follows that $\mathbf{w}, A_j\mathbf{w},\ldots,A_j^{r_j-1}\mathbf{w}$ and $e_{i_1}^{(j)},\ldots,e_{i_{n_{j+1}}}^{(j)}$ are linearly independent, since P_j is nonsingular and

$$P_j(\mathbf{w},\cdots,A_j^{r_j-1}\mathbf{w},\ e_{i_1}^{(j)},\cdots,e_{i_{n_{j+1}}}^{(j)}) = \begin{pmatrix} h_{1,1} & \cdots & h_{1,r_j} & & & \\ & \ddots & \vdots & P_j\,e_{i_1}^{(j)} & & P_j\,e_{i_{n_{j+1}}}^{(j)} \\ 0 & & h_{r_j,r_j} & \Vert & \cdots & \Vert \\ & & & u_{i_1}^{(j)} & & u_{i_{n_{j+1}}}^{(j)} \\ & 0 & & & & \end{pmatrix}$$

$$= \begin{pmatrix} h_{1,1} & \cdots & h_{1,r_j} & \\ & \ddots & \vdots & \\ 0 & & h_{r_j,r_j} & e_{r_j+1}^{(j)},\cdots,e_{n_j}^{(j)} \\ & 0 & & \end{pmatrix} = \begin{pmatrix} C & 0 \\ 0 & I_{n_{j+1}} \end{pmatrix} \tag{4.7}$$

is also nonsingular. Consequently, $\mathbf{w}, A_j\mathbf{w},\cdots,A_j^{r_j-1}\mathbf{w},\ e_{i_1}^{(j)},\cdots,e_{i_{n_{j+1}}}^{(j)}$ form a new basis of $\mathcal{R}^{n_j}$; and if there exists certain $e_{i_s}^{(j)}$, $1 \le s \le n_{j+1}$ for which $\Psi(A_j)e_{i_s}^{(j)} \neq 0$, then $\Psi(\lambda)$ is not the minimal polynomial of $\mathcal{R}^{n_j}$, in other words, $\mathbf{w}$ fails to be a regular vector in $\mathcal{R}^{n_j}$. In this case, a different vector in $\mathcal{R}^{n_j}$ will be chosen at random and the above procedure will be repeated, making our algorithm one of the Las Vegas type, until $\Psi(A_j)e^{(j)}_{i_s} = 0$ for all $s = 1, 2, \ldots, n_{j+1}$. An extremely early success, usually the first try in practice, of this procedure can be warranted by Theorem 1.0.1.

Now, $\Psi(A_j)\mathbf{y} = 0$ for any vector $\mathbf{y} \in \mathcal{R}^{n_j}$; therefore $\Psi(\lambda)$ is the minimal polynomial of $\mathcal{R}^{n_j}$ and $\mathbf{w}$ is a regular vector in $\mathcal{R}^{n_j}$. Let

$$\begin{aligned} k_j &= r_j, \\ \widetilde{\mathbf{w}}_j &= \mathbf{w}, \\ \Psi_j(\lambda) &= \Psi(\lambda), \end{aligned}$$

and

$$\widetilde{W}_j = \operatorname{span}\{\widetilde{\mathbf{w}}_j, A_j\widetilde{\mathbf{w}}_j, \cdots, A_j^{k_j-1}\widetilde{\mathbf{w}}_j\}.$$

Evidently, $\{e^{(j)}_{i_1}\}$, …, $\{e^{(j)}_{i_{n_{j+1}}}\}$ form a basis of the quotient space $\mathcal{R}^{n_j}/\widetilde{W}_j$. We thus complete the parts **A1**, **A2** and **A3**.

Next we consider the implementation of **A4**. Rewrite $P_j = I^*_{n_j}$ in (4.5) as

$$P_j = I^*_{n_j} = \begin{array}{cc} \left(\begin{array}{cc} M_1 & M_3 \\ M_2 & M_4 \end{array}\right) & \begin{array}{l} \}\ k_j \\ \}\ n_{j+1} \end{array} \\ \begin{array}{cc} \underbrace{}_{k_j} & \underbrace{}_{n_{j+1}} \end{array} & \end{array} \tag{4.8}$$

where M_2 is an $n_{j+1} \times k_j$ matrix and M_4 is an $n_{j+1} \times n_{j+1}$ matrix. Write the nonsingular matrix

$$\begin{aligned} T_j &= (\widetilde{\mathbf{w}}_j, \cdots, A_j^{k_j-1}\widetilde{\mathbf{w}}_j, e^{(j)}_{i_1}, \cdots, e^{(j)}_{i_{n_{j+1}}}) \\ &= \begin{array}{cc} \left(\begin{array}{cc} Q_1 & Q_3 \\ Q_2 & Q_4 \end{array}\right) & \begin{array}{l} \}\ k_j \\ \}\ n_{j+1} \end{array} \\ \begin{array}{cc} \underbrace{}_{k_j} & \underbrace{}_{n_{j+1}} \end{array} & \end{array} \end{aligned} \tag{4.9}$$

where Q_2 is an $n_{j+1} \times k_j$ matrix and Q_4 is an $n_{j+1} \times n_{j+1}$ matrix. Since, from (4.7),

$$P_jT_j = P_j(\widetilde{\mathbf{w}}_j, \cdots, A_j^{k_j-1}\widetilde{\mathbf{w}}_j, e^{(j)}_{i_1}, \cdots, e^{(j)}_{i_{n_{j+1}}}) = \left(\begin{array}{cc} C & 0 \\ 0 & I_{n_{j+1}} \end{array}\right), \tag{4.10}$$

so,

$$\begin{pmatrix} C^{-1} & 0 \\ 0 & I_{n_{j+1}} \end{pmatrix} P_j T_j = \begin{pmatrix} C^{-1} & 0 \\ 0 & I_{n_{j+1}} \end{pmatrix} \begin{pmatrix} C & 0 \\ 0 & I_{n_{j+1}} \end{pmatrix} = I_{n_j}. \tag{4.11}$$

Hence,

$$\begin{aligned} T_j^{-1} &= \begin{pmatrix} C^{-1} & 0 \\ 0 & I_{n_{j+1}} \end{pmatrix} P_j = \begin{pmatrix} C^{-1} & 0 \\ 0 & I_{n_{j+1}} \end{pmatrix} \begin{pmatrix} M_1 & M_3 \\ M_2 & M_4 \end{pmatrix} \\ &= \begin{pmatrix} C^{-1}M_1 & C^{-1}M_3 \\ M_2 & M_4 \end{pmatrix}. \end{aligned} \tag{4.12}$$

We now generate the matrix A_{j+1} of the linear operator L_{j+1} under the standard basis $e_1^{(j+1)}$, $e_2^{(j+1)}$, $\ldots$, $e_{n_{j+1}}^{(j+1)}$ of $\mathcal{R}^{n_{j+1}}$. Let $A_j = (a_1^{(j)}, a_2^{(j)}, \cdots, a_{n_j}^{(j)})$, where $a_i^{(j)}$ is the ith column of A_j, for $1 \leq i \leq n_j$, then

$$(A_j e_{i_1}^{(j)}, \cdots, A_j e_{i_{n_{j+1}}}^{(j)}) = (a_{i_1}^{(j)}, \cdots, a_{i_{n_{j+1}}}^{(j)}) = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \begin{matrix} \} \; k_j \\ \} \; n_{j+1} \end{matrix} \tag{4.13}$$

where B_1 is a $k_j \times n_{j+1}$ matrix and B_2 is an $n_{j+1} \times n_{j+1}$ matrix. Since $\{\widetilde{\mathbf{w}}_j$, $\cdots$, $A_j^{k_j-1}\widetilde{\mathbf{w}}_j$, $e_{i_1}^{(j)}$, $\cdots$, $e_{i_{n_{j+1}}}^{(j)}\}$ is a basis of $\mathcal{R}^{n_j}$, for $p = 1, \ldots, n_{j+1}$, there are constants $b_{1,p}$, $b_{2,p}$, $\ldots$, $b_{n_j,p}$ such that

$$\begin{aligned} A_j e_{i_p}^{(j)} &= b_{1,p}\widetilde{\mathbf{w}}_j + \cdots + b_{k_j,p}A_j^{k_j-1}\widetilde{\mathbf{w}}_j + b_{k_j+1,p}e_{i_1}^{(j)} + \cdots + b_{n_j,p}e_{i_{n_{j+1}}}^{(j)} \\ &= (\widetilde{\mathbf{w}}_j, \cdots, A_j^{k_j-1}\widetilde{\mathbf{w}}_j, e_{i_1}^{(j)}, \cdots, e_{i_{n_{j+1}}}^{(j)}) \begin{pmatrix} b_{1,p} \\ \vdots \\ b_{k_j,p} \\ b_{k_j+1,p} \\ \vdots \\ b_{n_j,p} \end{pmatrix} \\ &= T_j(b_{1,p}, \cdots, b_{n_j,p})^T. \end{aligned} \tag{4.14}$$

This implies

$$\{A_j e_{i_p}\} = b_{k_j+1,p}\{e^{(j)}_{i_1}\} + \cdots + b_{n_j,p}\{e^{(j)}_{i_{n_{j+1}}}\}, \qquad p = 1, \ldots, n_{j+1}. \tag{4.15}$$

Since σ_j is defined by

$$\sigma_j \ : \ \mathcal{R}^{n_j}/\widetilde{W}_j \longrightarrow \mathcal{R}^{n_{j+1}},$$

$$\mathbf{y} \longmapsto (c_1, \ldots, c_{n_{j+1}})^T,$$

where $\mathbf{y} = c_1\{e^{(j)}_{i_1}\} + c_2\{e^{(j)}_{i_2}\} + \cdots + c_{n_{j+1}}\{e^{(j)}_{i_{n_{j+1}}}\}$, we have

$$\sigma_j(\{e^{(j)}_{i_p}\}) = e^{(j+1)}_p, \qquad p = 1, \ldots, n_{j+1}.$$

For $p = 1, \ldots, n_{j+1}$, it follows from the definition of L_{j+1} that

$$L_{j+1}e^{(j+1)}_p = \sigma_j \circ G_j \circ \sigma_j^{-1}(e^{(j+1)}_p) = \sigma_j \circ G_j(\{e_{i_p}\}) = \sigma_j(\{A_j e_{i_p}\}). \tag{4.16}$$

By (4.15) and (4.16), $(b_{k_j+1,p}, \cdots, b_{n_j,p})^T$ is the p-th column of A_{j+1} for $p = 1, \ldots, n_{j+1}$. That is,

$$A_{j+1} = \begin{pmatrix} b_{k_j+1,1} & \cdots & b_{k_j+1,n_{j+1}} \\ \vdots & \ddots & \vdots \\ b_{n_j,1} & \cdots & b_{n_j,n_{j+1}} \end{pmatrix}. \tag{4.17}$$

In the following, we deduce a simple representation of A_{j+1}. By (4.14), we have

$$T_j^{-1}(A_j e_{i_1}, \cdots, A_j e_{i_{n_{j+1}}}) = \begin{pmatrix} b_{1,1} & \cdots & b_{1,n_{j+1}} \\ \vdots & \ddots & \vdots \\ b_{k_j+1,1} & \cdots & b_{k_j+1,n_{j+1}} \\ \vdots & \ddots & \vdots \\ b_{n_j,1} & \cdots & b_{n_j,n_{j+1}} \end{pmatrix}. \tag{4.18}$$

It follows from (4.12) and (4.13),

$$\begin{aligned} A_{j+1} &= \Big(\underbrace{0}_{k_j}, I_{n_{j+1}}\Big) T_j^{-1}(A_j e_{i_1}, \cdots, A_j e_{i_{n_{j+1}}}) \\ &= \Big(0, I_{n_{j+1}}\Big) \begin{pmatrix} C^{-1}M_1 & C^{-1}M_3 \\ M_2 & M_4 \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \\ &= \Big(M_2, M_4\Big) \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = M_2 B_1 + M_4 B_2. \end{aligned} \tag{4.19}$$

From (4.8), both matrices M_2 and M_4 are available after Gaussian elimination is applied in the first step. Therefore, A_{j+1} can be obtained with minimal computation effort. In particular, if no pivoting is employed in the elimination process, $M_4 = I_{n_{j+1}}$ and $A_{j+1} = M_2 B_1 + B_2$.

Remark 4.1.1 *It follows from* $G_j(\{e_{i_p}\}) = \{A_j e_{i_p}\}$ *and (4.15) that*

$$G_j(\{e_{i_p}\}) = b_{k_j+1,p}\{e^{(j)}_{i_1}\} + \cdots + b_{n_j,p}\{e^{(j)}_{i_{n_{j+1}}}\}, \qquad p = 1, \ldots, n_{j+1}.$$

Thus, A_{j+1} *is also the matrix of* G_j *under the basis* $\{e_{i_1}\}, \ldots, \{e_{i_{n_{j+1}}}\}$ *of* $\mathcal{R}^{n_j}/\widetilde{W}_j$.

4.2 Implementation of Step B

Now we discuss the detail of determining the structure of the Jordan blocks in the Jordan normal form of A.

A common squarefree relatively prime basis $\{\varphi_1, \varphi_2, \cdots, \varphi_r\}$ in $\mathcal{R}[\lambda]$ for the minimal polynomials $\Psi_1, \Psi_2, \ldots, \Psi_t$ constructed in Step A satisfying the decomposition,

$$\begin{aligned} \Psi_1 &= \varphi_1^{m_{1,1}} \varphi_2^{m_{1,2}} \cdots \varphi_r^{m_{1,r}}, \\ \Psi_2 &= \varphi_1^{m_{2,1}} \varphi_2^{m_{2,2}} \cdots \varphi_r^{m_{2,r}}, \\ &\cdots\cdots \\ \Psi_t &= \varphi_1^{m_{t,1}} \varphi_2^{m_{t,2}} \cdots \varphi_r^{m_{t,r}}, \end{aligned} \tag{4.20}$$

where $m_{1,k} \geq m_{2,k} \geq \cdots \geq m_{t,k} \geq 0$ for $k = 1, 2, \ldots, r$, can be generated by means of symbolic computation as follows.

For $k_1 = \deg \Psi_1(\lambda)$, let $d_0(\lambda) = \Psi_1(\lambda)$ and

$$d_l(\lambda) = \gcd\left(\Psi_1(\lambda), \frac{d^l}{d\lambda^l}\Psi_1(\lambda)\right), \qquad l = 1, \ldots, k_1.$$

In the proof of Lemma 2.3.1, it was shown that if $\{\theta_1, \ldots, \theta_s\}$ is a squarefree relatively prime basis of $\Psi_1(\lambda)$ such that

$$\Psi_1(\lambda) = [\theta_1(\lambda)]^{m_1}[\theta_2(\lambda)]^{m_2} \cdots [\theta_s(\lambda)]^{m_s}$$

with $m_i > 0$ $(i = 1, \ldots, s)$ and $m_1 < m_2 < \cdots < m_s$, then for $\eta_i(\lambda) = d_{i-1}(\lambda)/d_i(\lambda)$, $i = 1, \ldots, m_s - 1$, we have, ignoring the constant factors,

$$\left\{ \begin{array}{lcl}
\eta_1(\lambda) & = & d_0(\lambda)/d_1(\lambda) = \theta_1(\lambda)\theta_2(\lambda)\cdots\theta_s(\lambda), \\
 & & \vdots \\
\eta_{m_1}(\lambda) & = & d_{m_1-1}(\lambda)/d_{m_1}(\lambda) = \theta_1(\lambda)\theta_2(\lambda)\cdots\theta_s(\lambda), \\
\eta_{m_1+1}(\lambda) & = & d_{m_1}(\lambda)/d_{m_1+1}(\lambda) = \theta_2(\lambda)\cdots\theta_s(\lambda), \\
 & & \vdots \\
\eta_{m_2}(\lambda) & = & d_{m_2-1}(\lambda)/d_{m_2}(\lambda) = \theta_2(\lambda)\cdots\theta_s(\lambda), \\
\eta_{m_2+1}(\lambda) & = & d_{m_2}(\lambda)/d_{m_2+1}(\lambda) = \theta_3(\lambda)\cdots\theta_s(\lambda), \\
 & & \vdots \\
\eta_{m_s}(\lambda) & = & d_{m_s-1}(\lambda)/d_{m_s}(\lambda) = \theta_s(\lambda) \\
\eta_{m_s+1}(\lambda) & = & d_{m_s}(\lambda)/d_{m_s+1}(\lambda) = 1.
\end{array} \right.$$

This implies

$$\left\{ \frac{\eta_1}{\eta_2}, \frac{\eta_2}{\eta_3}, \cdots, \frac{\eta_{m_s}}{\eta_{m_s+1}} \right\} = \{ \underbrace{1, \ldots, 1}_{m_1-1}, \theta_1(\lambda), \underbrace{1, \ldots, 1}_{m_2-m_1-1}, \theta_2(\lambda), \cdots, \underbrace{1, \ldots, 1}_{m_s-m_{s-1}-1}, \theta_s(\lambda) \}.$$

Therefore, to generate $\{\theta_1, \ldots, \theta_s\}$ and find $m_1, m_2, \ldots, m_s$, we may compute

$$\frac{\eta_1}{\eta_2}, \ \frac{\eta_2}{\eta_3}, \ \frac{\eta_3}{\eta_4}, \ \cdots$$

yielding the sequence

$$\underbrace{1, \ldots, 1}_{n_1}, \ \frac{\eta_{n_1+1}}{\eta_{n_1+2}} \neq 1, \ \underbrace{1, \ldots, 1}_{n_2}, \ \frac{\eta_{n_1+n_2+2}}{\eta_{n_1+n_2+3}} \neq 1, \ 1, \ \cdots,$$

and

$$\begin{array}{ll}
m_1 = n_1 + 1, & \theta_1(\lambda) = \frac{\eta_{m_1}}{\eta_{m_1+1}}, \\
m_2 = m_1 + n_2 + 1, & \theta_2(\lambda) = \frac{\eta_{m_2}}{\eta_{m_2+1}}, \\
\vdots & \vdots \\
m_j = m_{j-1} + n_j + 1, & \theta_j(\lambda) = \frac{\eta_{m_j}}{\eta_{m_j+1}}, \\
\vdots & \vdots
\end{array} .$$

The process stops when $m_1 \deg \theta_1 + \cdots + m_s \deg \theta_s = k_1$ for certain s, and clearly, $\{\theta_1, \ldots, \theta_s\}$ constitutes a squarefree relatively prime basis of $\Psi_1(\lambda)$ for which

$$\Psi_1(\lambda) = [\theta_1(\lambda)]^{m_1} [\theta_2(\lambda)]^{m_2} \cdots [\theta_s(\lambda)]^{m_s}. \tag{4.21}$$

By the polynomial "long division" in $\mathcal{R}[\lambda]$[15], $\Psi_2(\lambda)$ can be decomposed as

$$\Psi_2(\lambda) = [\theta_1(\lambda)]^{b_1} [\theta_2(\lambda)]^{b_2} \cdots [\theta_s(\lambda)]^{b_s} \Theta_1(\lambda), \tag{4.22}$$

where $0 \leq b_i \leq m_i$ and $\Theta_1(\lambda)$ is not divisible by any θ_i, for $i = 1, \ldots, s$. With this decomposition, we calculate $\tau_i(\lambda) = \gcd(\theta_i, \Theta_1)$, $i = 1, \ldots, s$ by symbolic computation [18]. For those i's with $\deg \tau_i(\lambda) > 0$, θ_i has a non-trivial decomposition

$$\theta_i(\lambda) = \tau_i(\lambda) \, \xi_i(\lambda), \qquad i = 1, \ldots, s.$$

For $\tau_i(\lambda) = 1$, we let $\xi_i(\lambda) = \theta_i(\lambda)$. Writing

$$\Theta_1(\lambda) = \tau_1^{l_1}(\lambda) \tau_2^{l_2}(\lambda) \cdots \tau_s^{l_s}(\lambda) \Theta_2(\lambda),$$

and replacing $\theta_i(\lambda)$ in (4.21) and (4.22) by $\tau_i(\lambda)\xi_i(\lambda)$, for $i = 1, \ldots, s$; yields

$$\begin{aligned} \Psi_1 &= \tau_1^{m_1}\xi_1^{m_1}\tau_2^{m_2}\xi_2^{m_2}\cdots\tau_s^{m_s}\xi_s^{m_s}, \\ \Psi_2 &= \tau_1^{b_1+l_1}\xi_1^{b_1}\tau_2^{b_2+l_2}\xi_2^{b_2}\cdots\tau_s^{b_s+l_s}\xi_s^{b_s}\Theta_2(\lambda). \end{aligned} \tag{4.23}$$

Obviously, ignoring those $\tau_i(\lambda)$'s with degree 0 in $\{\tau_1, \xi_1, \ldots, \tau_s, \xi_s\}$, a new squarefree relatively prime basis of $\Psi_1(\lambda)$ is constructed. By repeating the same procedure for (4.23) until the factor $\Theta_i(\lambda)$ of Ψ_2 which is not divisible by polynomials in the basis becomes a constant, a common squarefree relatively prime basis of both $\Psi_1(\lambda)$ and $\Psi_2(\lambda)$ is produced.

Inductively, when a common squarefree relatively prime basis of $\Psi_1(\lambda)$, ..., $\Psi_{j-1}(\lambda)$ has been constructed, a common squarefree relatively prime basis of both $\Psi_{j-1}(\lambda)$ and $\Psi_j(\lambda)$ can be constructed by the above procedure to serve as a common squarefree relatively prime basis of $\Psi_1(\lambda)$, ..., $\Psi_j(\lambda)$, and the implementation of the part **B1** is accomplished.

Now we consider the implementation of the part **B2**. Corresponding to each $\varphi_k^{m_{i,k}}$ in (4.20) for $k = 1, \ldots, r$ and $i = 1, \ldots, t$; there are $p_k = \deg \varphi_k$ Jordan blocks of order $m_{i,k}$ of the form

$$\begin{pmatrix} \lambda_{k,l} & 1 & & 0 \\ & \lambda_{k,l} & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_{k,l} \end{pmatrix}_{m_{i,k}\times m_{i,k}}, \quad l = 1, \ldots, p_k, \tag{4.24}$$

in the Jordan normal form of A, where the $\lambda_{k,l}$'s are zeros of φ_k. Without knowing those zeros, the structure of the Jordan normal form of A is determined at this stage.

Remark 4.2.1 *The computation in Step A involves only Gaussian eliminations and divisions between polynomials, and the major computation in Step B is to generate a* *common squarefree relatively prime basis. Those computations can all be carried out by symbolic computation. Therefore, if all the entries in the given matrix A consist of only rational numbers, then the structures of the Jordan blocks in the Jordan normal form of A can be determined exactly.*

4.3 Implementation of Step C

In order to fill in the diagonal entries of the matrix in (4.24), the zeros of $\varphi_1, \ldots, \varphi_r$ in (4.20), which may not be in $\mathcal{R}$, need to be evaluated. The approximation of those zeros can be computed by well established subroutines, such as the Jenkins-Traub routine [2] in the IMSL library. When those zeros are available, the complete Jordan normal form of A is achieved.

Although those $\lambda_{k,l}$'s are eigenvalues of A, it is inappropriate to approximate those numbers by calculating the eigenvalues of A using standard codes such as EISPACK [16], or more advanced LAPACK [1], because if $\lambda_{k,l}$ is obtained as an eigenvalue of A, it might fail to provide the information of the location to which $\lambda_{k,l}$ belongs.

4.4 Computational result

Golub and Van Loan presented a 10×10 matrix [7] as shown below to illustrate the difficulty of calculating the Jordan normal form of a matrix by numerical computation:

$$A = \begin{pmatrix} 1 & 1 & 1 & -2 & 1 & -1 & 2 & -2 & 4 & -3 \\ -1 & 2 & 3 & -4 & 2 & -2 & 4 & -4 & 8 & -6 \\ -1 & 0 & 5 & -5 & 3 & -3 & 6 & -6 & 12 & -9 \\ -1 & 0 & 3 & -4 & 4 & -4 & 8 & -8 & 16 & -12 \\ -1 & 0 & 3 & -6 & 5 & -4 & 10 & -10 & 20 & -15 \\ -1 & 0 & 3 & -6 & 2 & -2 & 12 & -12 & 24 & -18 \\ -1 & 0 & 3 & -6 & 2 & -5 & 15 & -13 & 28 & -21 \\ -1 & 0 & 3 & -6 & 2 & -5 & 12 & -11 & 32 & -24 \\ -1 & 0 & 3 & -6 & 2 & -5 & 12 & -14 & 37 & -26 \\ -1 & 0 & 3 & -6 & 2 & -5 & 12 & -14 & 36 & -25 \end{pmatrix}.$$

Here, we use the algorithm described in the above sections to obtain the Jordan normal form of A. For symbolic computation used in Step A and Step B, we take advantage of routines available in Mathematica[18]. The main procedure and result are described as follows.

Let $A_1 = A$ and $I_{10} = (e_1, e_2, \cdots, e_{10})$ be the identity matrix of order 10. Randomly choose a non-zero vector $\mathbf{w} \in \mathcal{R}^{10}$, for example, we choose

$$\mathbf{w} = (1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \frac{1}{7}, \frac{1}{8}, \frac{1}{9}, \frac{1}{10}).$$

Then $K[\mathbf{w}] = (\mathbf{w}, A_1\mathbf{w}, \ldots, A_1^{10}\,\mathbf{w})$ is of the form

$$\begin{pmatrix}
1 & \frac{1949}{1260} & \frac{2587}{1260} & \frac{2167}{1260} & -\frac{263}{126} & -\frac{21689}{1260} & -\frac{20792}{315} & -\frac{7597}{36} & -\frac{79013}{126} & -\frac{2275877}{1260} & -\frac{1622597}{315} \\
\frac{1}{2} & \frac{269}{630} & -\frac{601}{1260} & -\frac{1279}{315} & -\frac{977}{63} & -\frac{30929}{630} & -\frac{45364}{315} & -\frac{7357}{18} & -\frac{72293}{63} & -\frac{2033957}{630} & -\frac{2889034}{315} \\
\frac{1}{3} & \frac{2}{35} & -\frac{563}{420} & -\frac{2593}{420} & -\frac{293}{14} & -\frac{26729}{420} & -\frac{6464}{35} & -\frac{6317}{12} & -\frac{62885}{42} & -\frac{1800997}{420} & -\frac{434839}{35} \\
\frac{1}{4} & -\frac{46}{315} & -\frac{2357}{1260} & -\frac{4801}{630} & -\frac{1576}{63} & -\frac{23789}{315} & -\frac{69308}{315} & -\frac{5701}{9} & -\frac{115018}{63} & -\frac{1671077}{315} & -\frac{4911188}{315} \\
\frac{1}{5} & -\frac{67}{252} & -\frac{2773}{1260} & -\frac{10837}{1260} & -\frac{17663}{630} & -\frac{107101}{1260} & -\frac{15752}{63} & -\frac{131173}{180} & -\frac{1340957}{630} & -\frac{7895737}{1260} & -\frac{5871193}{315} \\
\frac{1}{6} & -\frac{37}{105} & -\frac{69}{28} & -\frac{992}{105} & -\frac{3233}{105} & -\frac{19729}{210} & -\frac{9764}{35} & -\frac{8219}{10} & -\frac{84947}{35} & -\frac{1516069}{210} & -\frac{2275274}{105} \\
\frac{1}{7} & -\frac{1019}{2520} & -\frac{821}{315} & -\frac{24889}{2520} & -\frac{10121}{315} & -\frac{35341}{360} & -\frac{73657}{252} & -\frac{2176921}{2520} & -\frac{1612159}{630} & -\frac{3847211}{504} & -\frac{28936991}{1260} \\
\frac{1}{8} & -\frac{221}{504} & -\frac{6791}{2520} & -\frac{5113}{504} & -\frac{83129}{2520} & -\frac{254381}{2520} & -\frac{759197}{2520} & -\frac{321407}{360} & -\frac{6682673}{2520} & -\frac{19984037}{2520} & -\frac{60255653}{2520} \\
\frac{1}{9} & -\frac{577}{1260} & -\frac{1147}{420} & -\frac{4297}{420} & -\frac{6977}{210} & -\frac{8537}{84} & -\frac{31847}{105} & -\frac{53933}{60} & -\frac{560719}{210} & -\frac{3353641}{420} & -\frac{2527876}{105} \\
\frac{1}{10} & -\frac{197}{420} & -\frac{691}{252} & -\frac{2581}{252} & -\frac{10469}{315} & -\frac{128069}{1260} & -\frac{191089}{630} & -\frac{161801}{180} & -\frac{841082}{315} & -\frac{10060937}{1260} & -\frac{15167263}{630}
\end{pmatrix}.$$

Utilizing Gaussian elimination with row pivoting for $(\,K[\mathbf{w}], I_{10}\,)$, we obtain $(\,K^{*}[\mathbf{w}],\ I_{10}^{*}\,)$ where

$$K^{*}[\mathbf{w}] = \begin{pmatrix}
1 & \frac{1949}{1260} & \frac{2587}{1260} & \frac{2167}{1260} & -\frac{263}{126} & -\frac{21689}{1260} & -\frac{20792}{315} & -\frac{7597}{36} & -\frac{79013}{126} & -\frac{2275877}{1260} & -\frac{1622597}{315} \\
0 & -\frac{97}{280} & -\frac{421}{280} & -\frac{4133}{840} & -\frac{405}{28} & -\frac{34009}{840} & -\frac{11656}{105} & -\frac{7277}{24} & -\frac{23351}{28} & -\frac{1953317}{840} & -\frac{461719}{70} \\
0 & 0 & -\frac{367}{10476} & -\frac{7405}{31428} & -\frac{5713}{5238} & -\frac{135833}{31428} & -\frac{123928}{7857} & -\frac{1722455}{31428} & -\frac{322241}{1746} & -\frac{19126645}{31428} & -\frac{5179783}{2619} \\
0 & 0 & 0 & -\frac{47}{26424} & -\frac{77}{4404} & -\frac{2809}{26424} & -\frac{1702}{3303} & -\frac{57799}{26424} & -\frac{37523}{4404} & -\frac{826661}{26424} & -\frac{242519}{2202} \\
0 & 0 & 0 & 0 & -\frac{331}{19740} & -\frac{832}{4935} & -\frac{4085}{3948} & -\frac{16519}{3290} & -\frac{27921}{1316} & -\frac{19253}{235} & -\frac{5868517}{19740} \\
0 & 0 & 0 & 0 & 0 & -\frac{5}{1986} & -\frac{65}{1986} & -\frac{250}{993} & -\frac{495}{331} & -\frac{5025}{662} & -\frac{22945}{662} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}$$

and

$$I_{10}^* = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{860}{2619} & -\frac{3466}{2619} & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{169}{1101} & \frac{2581}{2202} & -\frac{1503}{734} & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{5283}{3290} & \frac{6571}{658} & -\frac{4491}{329} & \frac{6711}{1645} & 0 & 0 & 1 & 0 & 0 & 0 \\ \frac{4743}{6620} & -\frac{20743}{6620} & \frac{4671}{3310} & \frac{3857}{1655} & 0 & 0 & -\frac{1519}{662} & 1 & 0 & 0 \\ -\frac{1}{5} & \frac{4}{5} & 0 & -\frac{8}{5} & 1 & 0 & 0 & 0 & 0 & 0 \\ -\frac{2}{5} & \frac{7}{5} & \frac{2}{5} & -\frac{12}{5} & 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{39}{125} & -\frac{392}{375} & -\frac{152}{375} & \frac{632}{375} & 0 & 0 & \frac{28}{75} & -\frac{48}{25} & 1 & 0 \\ \frac{164}{375} & -\frac{539}{375} & -\frac{78}{125} & \frac{844}{375} & 0 & 0 & \frac{42}{25} & -\frac{248}{75} & 0 & 1 \end{pmatrix}. \qquad (4.25)$$

It follows from (4.25) that $\{\mathbf{w}_1, A_1\mathbf{w}_1, \ldots, A_1^5\mathbf{w}_1,\ e_5, e_6, e_9, e_{10}\}$ is a basis of $\mathcal{R}^{10}$. Solving the linear system

$$\begin{pmatrix} 1 & \frac{1949}{1260} & \frac{2587}{1260} & \frac{2167}{1260} & -\frac{263}{126} & -\frac{21689}{1260} \\ 0 & -\frac{97}{280} & -\frac{421}{280} & -\frac{4133}{840} & -\frac{405}{28} & -\frac{34009}{840} \\ 0 & 0 & -\frac{367}{10476} & -\frac{7405}{31428} & -\frac{5713}{5238} & -\frac{135833}{31428} \\ 0 & 0 & 0 & -\frac{47}{26424} & -\frac{77}{4404} & -\frac{2809}{26424} \\ 0 & 0 & 0 & 0 & -\frac{331}{19740} & -\frac{832}{4935} \\ 0 & 0 & 0 & 0 & 0 & -\frac{5}{1986} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{pmatrix} = \begin{pmatrix} -\frac{20792}{315} \\ -\frac{11656}{105} \\ -\frac{123928}{7857} \\ -\frac{1702}{3303} \\ -\frac{4085}{3948} \\ -\frac{65}{1986} \end{pmatrix}$$

yields the minimal polynomial of $\mathbf{w}$

$$\begin{aligned} \Psi(\lambda) &= \lambda^6 - (a_5\lambda^5 + a_4\lambda^4 + a_3\lambda^3 + a_2\lambda^2 + a_1\lambda + a_0) \\ &= \lambda^6 - 13\lambda^5 + 69\lambda^4 - 191\lambda^3 + 290\lambda^2 - 228\lambda + 72. \end{aligned}$$

Since $\Psi(A_1)e_i = 0$ for $i = 5, 6, 9, 10$, $\Psi(\lambda)$ is the minimal polynomial of $\mathcal{R}^{10}$ and $\mathbf{w}$ is a regular vector. Therefore,

$$\begin{aligned} k_1 &= \deg \Psi \;=\; 6, \\ \widetilde{\mathbf{w}}_1 &= \mathbf{w}, \\ \Psi_1(\lambda) &= \Psi(\lambda), \end{aligned}$$

$$\widetilde{W}_1 = \{\widetilde{\mathbf{w}}_1, A_1\widetilde{\mathbf{w}}_1, \cdots, A_1^5\widetilde{\mathbf{w}}_1\},$$

moreover, $\{e_5\}$, $\{e_6\}$, $\{e_9\}$, $\{e_{10}\}$ form a basis of the quotient space $\mathcal{R}^{10}/\widetilde{W}_1$.

From (4.8) and (4.25), we have

$$(M_2, M_4) = \begin{pmatrix} -\frac{1}{5} & \frac{4}{5} & 0 & -\frac{8}{5} & 1 & 0 & \vdots & 0 & 0 & 0 & 0 \\ -\frac{2}{5} & \frac{7}{5} & \frac{2}{5} & -\frac{12}{5} & 0 & 1 & \vdots & 0 & 0 & 0 & 0 \\ \frac{39}{125} & -\frac{392}{375} & -\frac{152}{375} & \frac{632}{375} & 0 & 0 & \vdots & \frac{28}{75} & -\frac{48}{25} & 1 & 0 \\ \frac{164}{375} & -\frac{539}{375} & -\frac{78}{125} & \frac{844}{375} & 0 & 0 & \vdots & \frac{42}{25} & -\frac{248}{75} & 0 & 1 \end{pmatrix}.$$

According to (4.13), we have

$$(A_1e_5, A_1e_6, A_1e_9, A_1e_{10}) = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 4 & -3 \\ 2 & -2 & 8 & -6 \\ 3 & -3 & 12 & -9 \\ 4 & -4 & 16 & -12 \\ 5 & -4 & 20 & -15 \\ 2 & -2 & 24 & -18 \\ \cdots & \cdots & \cdots & \cdots \\ 2 & -5 & 28 & -21 \\ 2 & -5 & 32 & -24 \\ 2 & -5 & 37 & -26 \\ 2 & -5 & 36 & -25 \end{pmatrix}.$$

Therefore,

$$A_2 = (M_2, M_4) \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = M_2\, B_1 + M_4\, B_2 =$$

$$\begin{pmatrix} -\frac{1}{5} & \frac{4}{5} & 0 & -\frac{8}{5} & 1 & 0 \\ -\frac{2}{5} & \frac{7}{5} & \frac{2}{5} & -\frac{12}{5} & 0 & 1 \\ \frac{39}{125} & -\frac{392}{375} & -\frac{152}{375} & \frac{632}{375} & 0 & 0 \\ \frac{164}{375} & -\frac{539}{375} & -\frac{78}{125} & \frac{844}{375} & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 & 4 & -3 \\ 2 & -2 & 8 & -6 \\ 3 & -3 & 12 & -9 \\ 4 & -4 & 16 & -12 \\ 5 & -4 & 20 & -15 \\ 2 & -2 & 24 & -18 \end{pmatrix} +$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{28}{75} & -\frac{48}{25} & 1 & 0 \\ \frac{42}{25} & -\frac{248}{75} & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & -5 & 28 & -21 \\ 2 & -5 & 32 & -24 \\ 2 & -5 & 37 & -26 \\ 2 & -5 & 36 & -25 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -4 & 4 & 0 & 0 \\ \frac{199}{75} & -\frac{76}{75} & 1 & 1 \\ \frac{86}{25} & -\frac{39}{25} & -4 & 5 \end{pmatrix}.$$

Let I_4 be the identity matrix in $\mathcal{R}^4$, we choose a non-zero vector

$$\mathbf{w} = (\tfrac{1}{2}, \tfrac{2}{3}, \tfrac{3}{4}, \tfrac{4}{5}) \in \mathcal{R}^4.$$

Using Gaussian elimination with row pivoting on $(\, K[\mathbf{w}], I_4 \,)$ where

$$K[\mathbf{w}] = (\mathbf{w}, A_2\mathbf{w}, A_2^2\mathbf{w}, A_2^3\mathbf{w}, A_2^4\mathbf{w}) = \begin{pmatrix} 1 & -\frac{1}{2} & -2 & -10 & -32 \\ \frac{1}{2} & -2 & -10 & -32 & -88 \\ \frac{1}{3} & \frac{273}{10} & \frac{433}{50} & \frac{2029}{100} & \frac{857}{25} \\ \frac{1}{4} & \frac{773}{300} & \frac{2041}{300} & \frac{2429}{300} & -\frac{7547}{300} \end{pmatrix},$$

yields the matrix

$$(\, K^*[\mathbf{w}], I_4^* \,) = \left(\begin{array}{ccccc|cccc} 1 & -\frac{1}{2} & -2 & -10 & -32 & 1 & 0 & 0 & 0 \\ 0 & -\frac{9}{4} & -9 & -27 & -72 & -\frac{1}{2} & 1 & 0 & 0 \\ 0 & 0 & -\frac{139}{150} & -\frac{2141}{300} & -\frac{927}{25} & -\frac{1219}{1350} & \frac{769}{675} & -1 & 0 \\ 0 & 0 & 0 & -\frac{507}{1112} & \frac{2535}{556} & \frac{2743}{1668} & -\frac{829}{417} & -\frac{751}{278} & 1 \end{array}\right).$$

Solving

$$\begin{pmatrix} 1 & -\frac{1}{2} & -2 & -10 \\ 0 & -\frac{9}{4} & -9 & -27 \\ 0 & 0 & -\frac{139}{150} & -\frac{2141}{300} \\ 0 & 0 & 0 & -\frac{507}{1112} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} -32 \\ -72 \\ -\frac{927}{26} \\ \frac{2535}{556} \end{pmatrix},$$

we obtain the minimal polynomial $\Psi(\lambda)$ of $\mathbf{w}$ in $\mathcal{R}^4$

$$\begin{aligned} \Psi(\lambda) &= \lambda^4 - (a_3\lambda^3 + a_2\lambda^2 + a_1\lambda + a_0) \\ &= \lambda^4 - 10\lambda^3 + 37\lambda^2 - 60\lambda + 36. \end{aligned}$$

Since $\deg \Psi(\lambda) = 4 = 10 - \deg \Psi_1(\lambda)$, we have $\Psi_2(\lambda) = \Psi(\lambda)$.

For $\Psi_1(\lambda)$ and $\Psi_2(\lambda)$, a common squarefree relatively prime basis $\{ (\lambda - 1), (\lambda - 2), (\lambda - 3) \}$ in $\mathcal{R}[\lambda]$ is found (by means of symbolic computation) for which

$$\begin{aligned} \Psi_1(\lambda) &= (\lambda-2)^3(\lambda-3)^2(\lambda-1), \\ \Psi_2(\lambda) &= (\lambda-2)^2(\lambda-3)^2. \end{aligned}$$

It follows that the Jordan normal form J of A is of the form

$$J = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix}.$$

APPENDIX

APPENDIX: An elimination algorithm

1 An elimination algorithm

For a nonzero vector $\mathbf{w} \in \mathcal{R}^n$ and an $n \times n$ matrix A, we consider the following $n \times (2n+1)$ matrix

$$H = (\mathbf{w}, A\mathbf{w}, \ldots, A^n\mathbf{w},\ I_n).$$

In this appendix, we develop an efficient elimination algorithm to reduce H to the following form

$$H^* = \left(\begin{array}{cccccccc}
h_{1,1} & \cdots & h_{1,r} & h_{1,r+1} & * & \cdots & * & \\
 & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \\
0 & & h_{r,r} & h_{r,r+1} & * & \cdots & * & I^*_n \\
 & & & 0 & 0 & \cdots & 0 & \\
 & 0 & & \vdots & \vdots & \ddots & \vdots & \\
 & & & 0 & 0 & \cdots & 0 & \\
\multicolumn{3}{c}{\underbrace{\hspace{3cm}}_{r}} & & \multicolumn{3}{c}{\underbrace{\hspace{2cm}}_{n-r}} &
\end{array}\right)$$

where $r > 0$ and $h_{i,i} \neq 0$ for $i = 1, \ldots, r$. Recall that we are only interested in obtaining the first $r+1$ columns and the I^*_n part in the above matrix. Our algorithm is therefore designed towards this main goal.

Let $\mathbf{w} = (w_1, \cdots, w_n) \in \mathcal{R}^n$ with $w_i \neq 0$ $(1 \leq i \leq n)$. The vector

$$\mathbf{g}_i = (\underbrace{0, \cdots, 0}_{i},\ \frac{w_{i+1}}{w_i}, \cdots, \frac{w_n}{w_i})^T,$$

is known as the *Gauss vector* and

$$G_i = \left(\frac{w_{i+1}}{w_i}, \cdots, \frac{w_n}{w_i} \right)^T$$

is the *pseudo Gauss vector* of $\mathbf{g}_i$. The matrix of the form

$$E_i = I_n - \mathbf{g}_i \, e_i^T$$

is known as the *elimination* matrix, for which we have

$$E_i \, \mathbf{w} = \begin{pmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & -\frac{w_{i+1}}{w_i} & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & -\frac{w_n}{w_i} & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_i \\ w_{i+1} \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} w_1 \\ \vdots \\ w_i \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Here, the elimination procedure $E_i \, \mathbf{w} = (I_n - \mathbf{g}_i \, e_i^T) \, \mathbf{w}$ is simply called *elimination on* $\mathbf{w}$ *by* $\mathbf{g}_i$.

An $n \times n$ *permutation* matrix $P_{i,j}$ is obtained by interchanging the i-th and j-th rows of the identity matrix I_n. Therefore, $P_{i,j}$ is of the form

$$P_{i,j} = (e_1, \cdots, e_j, \cdots, e_i, \cdots, e_n) = \left(\begin{array}{ccccccccccc} 1 & & & & & & & & & & \\ & \ddots & & & & & & & & & \\ & & 1 & & & & & & & & \\ & & & 0 & 0 & \cdots & 0 & 1 & & & \\ & & & 0 & 1 & \cdots & 0 & 0 & & & \\ & & & \vdots & \vdots & \ddots & \vdots & \vdots & & & \\ & & & 0 & 0 & \cdots & 1 & 0 & & & \\ & & & 1 & 0 & \cdots & 0 & 0 & & & \\ & & & & & & & & 1 & & \\ & & & & & & & & & \ddots & \\ & & & & & & & & & & 1 \end{array} \right) \begin{array}{c} \\ \\ \\ i \\ \\ \\ \\ j \\ \\ \\ \\ \end{array} .$$

Obiously, $P_{i,j} = I_n$ when $i = j$.

We now describe our elimination algorithm. Denote the vector

$$\mathbf{w}_1 = \begin{pmatrix} w_{1,1}^{(0)} \\ \vdots \\ w_{n,1}^{(0)} \end{pmatrix},$$

and

$$\mathbf{w}_{i+1} = \begin{pmatrix} w_{1,i+1}^{(0)} \\ \vdots \\ w_{n,i+1}^{(0)} \end{pmatrix} = A\,\mathbf{w}_i, \qquad 1 \le i \le n.$$

For $\mathbf{w}_1$, let

$$|w_{i_1,1}^{(0)}| = \max_{1 \le i \le n} |w_{i,1}^{(0)}|,$$

and for the permutation matrix $P_{i_1,1}$ write

$$P_{i_1,1}\mathbf{w}_1 = \begin{pmatrix} w_{1,1}^{(1)} \\ \vdots \\ w_{n,1}^{(1)} \end{pmatrix}$$

where $w_{1,1}^{(1)} = w_{i_1,1}^{(0)}$. For $E_1 = I_n - \mathbf{g}_1\, e_1^T$ where

$$\mathbf{g}_1 = \begin{pmatrix} 0 \\ \frac{w_{2,1}^{(1)}}{w_{1,1}^{(1)}} \\ \vdots \\ \frac{w_{n,1}^{(1)}}{w_{1,1}^{(1)}} \end{pmatrix} = \begin{pmatrix} 0 \\ G_1 \end{pmatrix},$$

we have

$$E_1\, P_{i_1,1}\, \mathbf{w}_1 = E_1 \begin{pmatrix} w_{1,1}^{(1)} \\ \vdots \\ w_{n,1}^{(1)} \end{pmatrix} = \begin{pmatrix} V_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

where $V_1 = (w^{(0)}_{i_1,1})$. Denote the first column of E_1 by

$$T_1 = \begin{pmatrix} 1 \\ -G_1 \end{pmatrix}.$$

Now, for $\mathbf{w}_2$, we write

$$E_1 \, P_{i_1,1} \, \mathbf{w}_2 = \begin{pmatrix} w^{(0)}_{i_1,2} \\ B_1 \end{pmatrix}$$

with

$$B_1 = (w^{(1)}_{2,2}, \cdots, w^{(1)}_{n,2})^T.$$

If $B_1 = \mathbf{0}$, then $V_2 = (w^{(0)}_{i_1,2})$ and the elimination procedure terminates. Otherwise, let

$$|w^{(1)}_{i_2,2}| = \max_{2 \le i \le n} |w^{(1)}_{i,2}|,$$

then for the permutation matrix $P_{i_2,2}$ we write

$$P_{i_2,2} \begin{pmatrix} w^{(0)}_{i_1,2} \\ B_1 \end{pmatrix} = \begin{pmatrix} w^{(0)}_{i_1,2} \\ w^{(2)}_{2,2} \\ \vdots \\ w^{(2)}_{n,2} \end{pmatrix}$$

where $w^{(2)}_{2,2} = w^{(1)}_{i_2,2}$. Let $E_2 = I_n - \mathbf{g}_2 \, e_2^T$ where

$$\mathbf{g}_2 = \begin{pmatrix} 0 \\ 0 \\ \frac{w^{(2)}_{3,2}}{w^{(2)}_{2,2}} \\ \vdots \\ \frac{w^{(2)}_{n,2}}{w^{(2)}_{2,2}} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ G_2 \end{pmatrix}.$$

Then,

$$E_2 \, P_{i_2,2} \, E_1 \, P_{i_1,1} \, \mathbf{w}_2 = E_2 \, P_{i_2,2} \begin{pmatrix} w^{(0)}_{i_1,2} \\ B_1 \end{pmatrix} = E_2 \begin{pmatrix} w^{(0)}_{i_1,2} \\ w^{(1)}_{i_2,2} \\ \vdots \\ w^{(2)}_{n,2} \end{pmatrix} = \begin{pmatrix} V_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

where $V_2 = (w_{i_1,2}^{(0)}, w_{i_2,2}^{(1)})^T$. Denote the second column of E_2 by

$$T_2 = \begin{pmatrix} 0 \\ 1 \\ -G_2 \end{pmatrix},$$

and update T_1 by

$$T_1 \longleftarrow E_2 \, P_{i_2,2} \, T_1.$$

In general, for $1 < k < n+1$, suppose $P_{i_1,1}$, E_1, …, $P_{i_k,k}$, E_k, and T_1, …, T_k are computed, let us consider the following elimination procedure on $\mathbf{w}_{k+1}$. Write

$$E_k \, P_{i_k,k} \cdots E_1 \, P_{i_1,1} \, \mathbf{w}_{k+1} = \begin{pmatrix} w_{i_1,k+1}^{(0)} \\ \vdots \\ w_{i_k,k+1}^{(k-1)} \\ B_k \end{pmatrix}$$

with

$$B_k = (w_{k+1,k+1}^{(k)}, \cdots, w_{n,k+1}^{(k)})^T.$$

If $B_k = \mathbf{0}$, then $V_{k+1} = (w_{i_1,k+1}^{(0)}, \cdots, w_{i_k,k+1}^{(k-1)})^T$ and the elimination procedure terminates. Otherwise, let

$$|w_{i_{k+1},k+1}^{(k)}| = \max_{k+1 \le i \le n} |w_{i,k+1}^{(k)}|.$$

For matrices $P_{i_{k+1},k+1}$ and $E_{k+1} = I_n - \mathbf{g}_{k+1} \, e_{k+1}^T$ where

$$\mathbf{g}_{k+1} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \frac{w_{k+2,k+1}^{(k)}}{w_{i_{k+1},k+1}^{(k)}} \\ \vdots \\ \frac{w_{n,k+1}^{(k)}}{w_{i_{k+1},k+1}^{(k)}} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ G_{k+1} \end{pmatrix},$$

we have

$$E_{k+1}\,P_{i_{k+1},k+1}\begin{pmatrix} w^{(0)}_{i_1,k+1} \\ \vdots \\ w^{(k-1)}_{i_k,k+1} \\ B_k \end{pmatrix} = \begin{pmatrix} w^{(0)}_{i_1,k+1} \\ \vdots \\ w^{(k-1)}_{i_k,k+1} \\ w^{(k)}_{i_{k+1},k+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} V_{k+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

with $V_{k+1} = (w^{(0)}_{i_1,k+1}, \cdots, w^{(k-1)}_{i_k,k+1}, w^{(k)}_{i_{k+1},k+1})^T$. Denote the $(k+1)$-th column of E_{k+1} by

$$T_{k+1} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ -G_{k+1} \end{pmatrix}$$

and update $T_1, \ldots, T_k$ as follows,

$$\begin{cases} T_1 & \longleftarrow \quad E_{k+1}\,P_{i_{k+1},k+1}\,T_1, \\ & \vdots \\ T_k & \longleftarrow \quad E_{k+1}\,P_{i_{k+1},k+1}\,T_k. \end{cases}$$

Obviously, if the elimination procedure is terminated at the $(k+1)$-th step, then I^*_n can be expressed as

$$I^*_n = E_{k+1}\,P_{i_{k+1},k+1} \cdots E_1\,P_{i_1,1} = M_{k+1}\,P_{k+1}$$

where

$$P_{k+1} = P_{i_1,1}\,P_{i_2,2} \cdots P_{i_{k+1},k+1}$$

and

$$M_{k+1} = \begin{pmatrix} T_1, & \cdots, & T_{k+1}, & \begin{matrix} 0 \\ I_{n-k-1} \end{matrix} \end{pmatrix}.$$

Remark *In implementing the above elimination algorithm, we only need to store the integer list* $\{i_1, i_2, \ldots\}$*, the vector lists* $\{V_1, V_2, \ldots\}$*,* $\{G_1, G_2, \ldots\}$ *and* $\{T_1, T_2, \ldots\}$*.*

2 The algorithm of determining the minimal polynomials in Theorem 2.1.1 by using the elimination algorithm

Based on the above elimination algorithm, we use the following pseudo code to describe our algorithm of determining the minimal polynomials $\Psi_1(\lambda)$, $\Psi_2(\lambda)$, …, $\Psi_t(\lambda)$ of an $n \times n$ matrix A in Theorem 2.1.1.

Input: An $n \times n$ matrix A.

Output: The minimal polynomials $\Psi_1(\lambda)$, $\Psi_2(\lambda)$, …, $\Psi_t(\lambda)$.

```
    $t = 0$ (the number of minimial polynomials);
    $m = n$;
(♠) $P = \emptyset$ (the list of permutation indices $\{i_1, i_2, \ldots\}$);
    $L = \emptyset$ (the list of the pseudo Gauss vectors $\{G_1, G_2, \ldots\}$);
    $R = \emptyset$ (the list of $\{V_1, V_2, \ldots\}$ generated by elimination);
    $T = \emptyset$ (the list of the columns $\{T_1, T_2, \ldots\}$ of the inverse matrix );

    generate a random vector $\mathbf{x} = (x_1, \cdots, x_m)^T \in \mathcal{R}^m$;
    $\mathbf{v} = \mathbf{w} = \mathbf{x}$;

    $r = 0$ (the degree of a minimal polynomial);
    For $i = 1$ to $m + 1$
        For $j = 1$ to $i - 1$
            If $i_j > j$, then
                interchange $w_j$ and $w_{i_j}$, i.e., $w_j \Longleftrightarrow w_{i_j}$
            Endif
            update $\mathbf{w}$ by using $G_j$;
            $V_i = V_i \cup \{w_j\}$
```

```
      End

      If $i < m + 1$, then

        search for the first nonzero element among $\{w_i, \ldots, w_m\}$;

        If no nonzero element is found, then

          goto $(\clubsuit)$

        Else

          there is a $w_{k_i} \neq 0$ $(i \leq k_i \leq m)$

        Endif

        $P = P \cup \{k_i\}$;

        If $k_i > i$, then

          interchange $w_i$ and $w_{k_i}$, i.e., $w_i \Longleftrightarrow w_{k_i}$,

          and the corresponding elements in $T_1, \ldots, T_{i-1}$

        Endif

        $V_i = V_i \cup \{w_i\}$, and $R = R \cup \{V_i\}$;

        generate $G_i$, and $L = L \cup \{G_i\}$;

        update $T_1, \ldots, T_{i-1}$ by using $G_i$;

        generate $T_i$ by using $G_i$, and $T = T \cup \{T_i\}$;

        generate the next vector $\mathbf{w} = A\mathbf{v}$;

        $\mathbf{v} = \mathbf{w}$;

        $r = r + 1$

      Endif

    End

    $t = t + 1$;

$(\clubsuit)$ generate the minimal polynomial $\Psi_t(\lambda)$ of $\mathbf{x}$;

    If $m - r > 0$, then

      If $\Psi_t(\lambda)$ is not the minimal polynomial of $\mathcal{R}^m$, then
```

```
        goto (♠)
    Endif
    form the matrices $M_2, M_4$ and $B_1, B_2$;
    generate the new matrix $A = M_2 B_1 + M_4 B_2$ in $\mathcal{R}^{m-r}$;
    $m = m - r$;
    goto (♠)
Endif
stop.
```

3 Computational example

We use the following 8×8 Hadamard matrix

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

to demonstrate our elimination algorithm described above in computing its Jordan normal form. Since all the entries of this matrix are integers, we may proceed our computation by the symbolic computation. And, as we mentioned before, in this case the decision of row pivoting will base on whether or not the pivot element is zero rather than how big its magnitute is.

Let $A_1 = A$. Choose a nonzero vector

$$\mathbf{x} = \left(1, \tfrac{1}{2}, \tfrac{1}{3}, \tfrac{1}{4}, \tfrac{1}{5}, \tfrac{1}{6}, \tfrac{1}{7}, \tfrac{1}{8}\right)^T \in \mathcal{R}^8,$$

and let $\mathbf{v} = \mathbf{w} = \mathbf{x}$.

Since the first entry of $\mathbf{w}$ is 1, pivoting is not necessary. Using Gauss vector

$$\mathbf{g}_1 = \begin{pmatrix} 0 \\ G_1 \end{pmatrix}$$

with the pesudo Gauss vector

$$G_1 = \left(\tfrac{1}{2}, \tfrac{1}{3}, \tfrac{1}{4}, \tfrac{1}{5}, \tfrac{1}{6}, \tfrac{1}{7}, \tfrac{1}{8}\right)^T$$

on the elimination of $\mathbf{w}$, *i.e.*, $\mathbf{w} \leftarrow (I_8 - \mathbf{g}_1 e_1^T)\mathbf{w}$, we have $\mathbf{w} = (1, 0, 0, 0, 0, 0, 0, 0)^T$ and

$$V_1 = (1), \qquad T_1 = \begin{pmatrix} 1 \\ -G_1 \end{pmatrix}.$$

Compute

$$\mathbf{w} = A_1 \mathbf{v} = \left(\tfrac{761}{280}, \tfrac{533}{840}, \tfrac{853}{840}, \tfrac{121}{280}, \tfrac{1217}{840}, \tfrac{149}{280}, \tfrac{229}{280}, \tfrac{337}{840}\right)^T$$

and let $\mathbf{v} = \mathbf{w}$. followed by updating $\mathbf{w}$ by using $\mathbf{g}_1$

$$\mathbf{w} \longleftarrow (I_8 - \mathbf{g}_1 e_1^T)\,\mathbf{w} = \left(\tfrac{761}{280}, -\tfrac{1217}{1680}, \tfrac{23}{210}, -\tfrac{277}{1120}, \tfrac{1901}{2100}, \tfrac{19}{240}, \tfrac{421}{980}, \tfrac{59}{960}\right)^T.$$

Since the second entry of $\mathbf{w}$ is $-\frac{1217}{1680}$, pivoting is not needed. Elimination on $\mathbf{w}$ by the Gauss vector

$$\mathbf{g}_2 = \begin{pmatrix} 0 \\ 0 \\ G_2 \end{pmatrix}$$

where

$$G_2 = \left(-\tfrac{184}{1217}, \tfrac{831}{2434}, -\tfrac{7604}{6085} - \tfrac{133}{1217}, -\tfrac{5052}{8519}, -\tfrac{413}{4868}\right)$$

and updating T_1 by $I_8 - \mathbf{g}_2 e_2^T$, yields

$$\begin{aligned} \mathbf{w} &\longleftarrow (I_8 - \mathbf{g}_2 e_2^T)\,\mathbf{w} = \left(\tfrac{761}{280}, -\tfrac{1217}{1680}, 0, 0, 0, 0, 0, 0\right)^T, \\ T_1 &\longleftarrow (I_8 - \mathbf{g}_2 e_2^T)\,T_1 = \left(1, -\tfrac{1}{2}, -\tfrac{1493}{3651}, -\tfrac{193}{2434}, -\tfrac{5019}{6085}, -\tfrac{808}{3651}, -\tfrac{3743}{8519}, -\tfrac{815}{4868}\right)^T, \end{aligned}$$

and hence

$$V_2 = \begin{pmatrix} \frac{761}{280} \\ -\frac{1217}{1680} \end{pmatrix}, \quad \text{and} \quad T_2 = \begin{pmatrix} 0 \\ 1 \\ -G_2 \end{pmatrix}.$$

Next, we calculate

$$\mathbf{w} = A_1 \mathbf{v} = \left(8, 4, \tfrac{8}{3}, 2, \tfrac{8}{5}, \tfrac{4}{3}, \tfrac{8}{7}, 1\right)^T$$

and let $\mathbf{v} = \mathbf{w}$.

Updating $\mathbf{w}$ by $\mathbf{g}_1$ and $\mathbf{g}_2$, we have

$$\mathbf{w} \longleftarrow (I_8 - \mathbf{g}_2 e_2^T)(I_8 - \mathbf{g}_1 e_1^T)\, \mathbf{w} = (8,0,0,0,0,0,0,0)^T,$$

and, therefore the elimination procedure for A_1 should be terminated here. Based on V_1 and V_2, we solve the linear system

$$\begin{pmatrix} 1 & \frac{761}{280} \\ 0 & -\frac{1217}{1680} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} 8 \\ 0 \end{pmatrix}.$$

to obtain the minimal polynomial $\Psi(\lambda) = \lambda^2 - 8$ of $\mathbf{x}$.

Since pivoting is not used in the elimination process of the above two vectors, we may choose $\{e_3\}, \ldots, \{e_8\}$ as the basis of the quotient space $\mathcal{R}^8/W_1$ where $W_1 = \text{span}\{\mathbf{x}, A_1\mathbf{x}\}$. Since $\Psi(A_1)e_i = 0$ for $i = 3,4,5,6,7,8$, $\Psi_1(\lambda) = \Psi(\lambda)$ is the minimal polynomial of $\mathcal{R}^8$. From

$$I_8^* = \begin{pmatrix} T_1, \ T_2, & \begin{matrix} 0 \\ I_6 \end{matrix} \end{pmatrix} = \left(\begin{array}{cc|c} 1 & 0 & \\ -\frac{1}{2} & 1 & 0 \\ -\frac{1493}{3651} & \frac{184}{1217} & \\ -\frac{193}{2434} & -\frac{831}{2434} & \\ -\frac{5019}{6085} & \frac{7604}{6085} & \\ -\frac{808}{3651} & \frac{133}{1217} & I_6 \\ -\frac{3743}{8519} & \frac{5052}{8519} & \\ -\frac{815}{4868} & \frac{413}{4868} & \end{array} \right),$$

and the notations in (4.8), we have

$$(M_2, M_4) = \begin{pmatrix} -\frac{1493}{3651} & \frac{184}{1217} & \\ -\frac{193}{2434} & -\frac{831}{2434} & \\ -\frac{5019}{6085} & \frac{7604}{6085} & \\ -\frac{808}{3651} & \frac{133}{1217} & I_6 \\ -\frac{3743}{8519} & \frac{5052}{8519} & \\ -\frac{815}{4868} & \frac{413}{4868} & \end{pmatrix}.$$

It follows from (4.19)

$$(A_1e_3, A_1e_4, A_1e_5, A_1e_6, A_1e_7, A_1e_8) = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 \\ -1 & -1 & 1 & 1 & -1 & -1 \\ -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & -1 & 1 & -1 & 1 \\ -1 & -1 & -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 & 1 & -1 \end{pmatrix}.$$

Therefore,

$$\begin{aligned} A_2 &= (M_2, M_4) \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = M_2\, B_1 + B_2 \\ &= \begin{pmatrix} -\frac{4592}{3651} & -\frac{5696}{3651} & \frac{2710}{3651} & \frac{1606}{3651} & -\frac{4592}{3651} & -\frac{5696}{3651} \\ -\frac{1729}{1217} & \frac{1536}{1217} & \frac{705}{1217} & -\frac{898}{1217} & -\frac{1729}{1217} & \frac{1536}{1217} \\ \frac{1734}{1217} & -\frac{6538}{6085} & -\frac{700}{1217} & -\frac{18708}{6085} & -\frac{700}{1217} & -\frac{18708}{6085} \\ \frac{3242}{3651} & -\frac{4858}{3651} & -\frac{4060}{3651} & \frac{2444}{3651} & -\frac{4060}{3651} & \frac{2444}{3651} \\ -\frac{1030}{1217} & -\frac{17314}{8519} & -\frac{1030}{1217} & -\frac{17314}{8519} & \frac{1404}{1217} & -\frac{276}{8519} \\ -\frac{2635}{2434} & \frac{910}{1217} & -\frac{2635}{2434} & \frac{910}{1217} & \frac{2233}{2434} & -\frac{1524}{1217} \end{pmatrix}. \end{aligned}$$

Next, choose a nonzero vector

$$\mathbf{x} = \left(1, \tfrac{1}{2}, \tfrac{1}{3}, \tfrac{1}{4}, \tfrac{1}{5}, \tfrac{1}{6}\right)^T \in \mathcal{R}^6,$$

and let $\mathbf{v} = \mathbf{w} = \mathbf{x}$.

Since the first entry of $\mathbf{w}$ is 1, again no pivoting is needed. Elimination on $\mathbf{w}$ by the Gauss vector

$$\mathbf{g}_1 = \begin{pmatrix} 0 \\ G_1 \end{pmatrix}$$

where

$$G_1 = \left(\tfrac{1}{2}, \tfrac{1}{3}, \tfrac{1}{4}, \tfrac{1}{5}, \tfrac{1}{6}\right)^T,$$

i.e., $\mathbf{w} \leftarrow (I_6 - \mathbf{g}_1 e_1^T)\mathbf{w}$, we obtain $\mathbf{w} = (1,0,0,0,0,0)^T$, and hence

$$V_1 = (1), \quad \text{and} \quad T_1 = \begin{pmatrix} 1 \\ -G_1 \end{pmatrix}.$$

Compute

$$\mathbf{w} = A_2\mathbf{v} = \left(-\tfrac{80029}{36510}, -\tfrac{10403}{12170}, -\tfrac{12782}{18255}, -\tfrac{334}{3651}, -\tfrac{620377}{2555570}, -\tfrac{16573}{18255}\right)^T$$

and let $\mathbf{v} = \mathbf{w}$. Updating $\mathbf{w}$ by $\mathbf{g}_1$ yields

$$\mathbf{w} \longleftarrow (I_6 - \mathbf{g}_1 e_1^T)\mathbf{w} = \left(-\tfrac{80029}{36510}, \tfrac{17611}{73020}, \tfrac{3337}{109530}, \tfrac{22223}{48680}, -\tfrac{1270841}{638925}, -\tfrac{118847}{219060}\right)^T.$$

Since the second entry of $\mathbf{w}$ is $\tfrac{17611}{73020}$, again no pivoting is needed. Elimination on $\mathbf{w}$ by

$$\mathbf{g}_2 = \begin{pmatrix} 0 \\ 0 \\ G_2 \end{pmatrix}$$

with

$$G_2 = \left(\tfrac{6674}{52833}, \tfrac{66669}{35222}, -\tfrac{462124}{56035}, -\tfrac{118847}{52833}\right)^T$$

and updating T_1 by $I_6 - \mathbf{g}_2 e_2^T$ yields

$$\begin{aligned} \mathbf{w} &\longleftarrow (I_6 - \mathbf{g}_2 e_2^T)\,\mathbf{w} = \left(-\tfrac{80029}{36510}, \tfrac{17611}{73020}, 0, 0, 0, 0\right), \\ T_1 &\longleftarrow (I_6 - \mathbf{g}_2 e_2^T)\,T_1 = \left(1, -\tfrac{1}{2}, -\tfrac{4758}{17611}, \tfrac{24592}{35222}, -\tfrac{242269}{56035}, -\tfrac{22743}{17611}\right)^T. \end{aligned}$$

It follows that

$$V_2 = \begin{pmatrix} -\frac{80029}{36510} \\ \frac{17611}{73020} \end{pmatrix}, \quad \text{and} \quad T_2 = \begin{pmatrix} 0 \\ 1 \\ -G_2 \end{pmatrix}.$$

Compute

$$\mathbf{w} = A_2 \mathbf{v} = \left(8, 4, \tfrac{8}{3}, 2, \tfrac{8}{5}, \tfrac{4}{3}\right)$$

and let $\mathbf{v} = \mathbf{w}$. Updating $\mathbf{w}$ by $\mathbf{g}_1$ and $\mathbf{g}_2$ yields

$$\mathbf{w} \longleftarrow (I_6 - \mathbf{g}_2\, e_2^T)(I_6 - \mathbf{g}_1\, e_1^T)\,\mathbf{w} = (8, 0, 0, 0, 0, 0)^T.$$

Solving the linear system

$$\begin{pmatrix} 1 & -\frac{80029}{36510} \\ 0 & \frac{17611}{73020} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} 8 \\ 0 \end{pmatrix}$$

results in the minimal polynomial $\Psi(\lambda) = \lambda^2 - 8$ of $\mathbf{x}$.

It is easy to see that the basis of the quotient space $\mathcal{R}^6/W_1$ is $\{e_3\}, \{e_4\}, \{e_5\}, \{e_6\}$ where $W_1 = \operatorname{span}\{\mathbf{x}, A_2\mathbf{x}\}$. Since $\Psi(A_2)e_i = 0$ for $i = 3, 4, 5, 6$, $\Psi_2(\lambda) = \Psi(\lambda)$ is the minimal polynomial of $\mathcal{R}^6$. Again, from

$$I_6^* = \begin{pmatrix} T_1, & T_2, & \begin{matrix} 0 \\ I_4 \end{matrix} \end{pmatrix} = \left(\begin{array}{ccc} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & \\ -\frac{4758}{17611} & -\frac{6674}{52833} & \\ \frac{24529}{35222} & -\frac{66669}{35222} & \\ -\frac{242269}{56035} & \frac{462124}{56035} & I_4 \\ -\frac{22743}{17611} & \frac{118847}{52833} & \end{array}\right),$$

we have

$$(M_2, M_4) = \left(\begin{array}{ccc} -\frac{4758}{17611} & -\frac{6674}{52833} & \\ \frac{24529}{35222} & -\frac{66669}{35222} & I_4 \\ -\frac{242269}{56035} & \frac{462124}{56035} & \\ -\frac{22743}{17611} & \frac{118847}{52833} & \end{array}\right)$$

in (4.8), and from (4.19),

$$(A_2e_3, A_2e_4, A_2e_5, A_2e_6) = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} \frac{2710}{3651} & \frac{1606}{3651} & -\frac{4592}{3651} & -\frac{5696}{3651} \\ \frac{705}{1217} & -\frac{898}{1217} & -\frac{1729}{1217} & \frac{1536}{1217} \\ -\frac{700}{1217} & -\frac{18708}{6085} & -\frac{700}{1217} & -\frac{18708}{6085} \\ -\frac{4060}{3651} & \frac{2444}{3651} & -\frac{4060}{3651} & \frac{2444}{3651} \\ -\frac{1030}{1217} & -\frac{17314}{8519} & \frac{1404}{1217} & -\frac{276}{8519} \\ -\frac{2635}{2434} & \frac{910}{1217} & \frac{2233}{2434} & -\frac{1524}{1217} \end{pmatrix}.$$

Therefore,

$$\begin{aligned} A_3 &= (M_2, M_4)\begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = M_2\,B_1 + B_2 \\ &= \begin{pmatrix} -\frac{14950}{17611} & -\frac{818932}{264165} & -\frac{2954}{52833} & -\frac{247644}{88055} \\ -\frac{178745}{105666} & \frac{125342}{52833} & \frac{74095}{105666} & -\frac{148252}{52833} \\ \frac{24272}{33621} & -\frac{48124}{4803} & -\frac{24616}{4803} & \frac{2878228}{168105} \\ -\frac{25995}{35222} & -\frac{78202}{52833} & -\frac{69125}{105666} & \frac{63428}{17611} \end{pmatrix}. \end{aligned}$$

Following exactly the same procedures as described above, the minimal polynomial $\Psi_3(\lambda) = \lambda^2 - 8$ of $\mathcal{R}^4$ with respect to A_3 and the matrix

$$A_4 = \begin{pmatrix} -\frac{1702102}{398853} & \frac{688850416}{41879565} \\ -\frac{47165}{75972} & -\frac{1702102}{398853} \end{pmatrix}$$

can be achieved. And the minimal polynomial $\Psi_4(\lambda) = \lambda^2 - 8$ of $\mathcal{R}^2$ with respect to A_4 can be determined similarly.

For $\Psi_1(\lambda)$, $\Psi_2(\lambda)$, $\Psi_3(\lambda)$ and $\Psi_4(\lambda)$, a common squarefree relatively prime basis $\{\,(\lambda - 2\sqrt{2}),\ (\lambda + 2\sqrt{2})\,\}$ in $\mathcal{R}[\lambda]$ can be found by means of symbolic computation such that

$$\Psi_1(\lambda) = \Psi_2(\lambda) = \Psi_3(\lambda) = \Psi_4(\lambda) = (\lambda - 2\sqrt{2})(\lambda + 2\sqrt{2}).$$

Therefore, it follows from Theorem 2.1.3 that the Jordan normal form J of A is of the form

$$J = \begin{pmatrix} 2\sqrt{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -2\sqrt{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2\sqrt{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2\sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2\sqrt{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -2\sqrt{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2\sqrt{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2\sqrt{2} \end{pmatrix}.$$

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] E. Anderson, et al., *LAPACK user's guide*, SIAM Philadelphia, 1992.

[2] A. Ralston and P. Rabinowitz, *A first course in numerical analysis* (Second edition), McGraw-Hill, 1978.

[3] T. Beelen and P. Van Dooren, *Computational aspects of the Jordan canonical form*, in Reliable Numerical Computation, M.G. Cox and S. Hammerling, eds., Clarendon Press, Oxford, 1990, pp. 57-72.

[4] J. Demmel, *A numerical analyst's Jordan canonical form*, Ph.D. Dissertation, Computer Science Division, University of California, Berkeley, 1983.

[5] F.R. Gantmacher, *The theory of matrices*, Vol. 1, Chelsea, New York, 1959.

[6] J. von zur Gathen, *Parallel algorithms for algebraic problems*, SIAM J. Comput., **13** (1984), pp. 802-824.

[7] G.H. Golub and C.F. Van Loan, *Matrix Computation* (First edition), The Johns Hopkins University Press, 1983.

[8] G.H. Golub and C.F. Van Loan, *Matrix Computation* (Third edition), The Johns Hopkins University Press, 1996.

[9] G.H. Golub and J.H. Wilkinson, *Ill-conditioned eigensystems and the computation of the Jordan canonical form*, SIAM Rev., **18** (1976), pp. 578-619.

[10] A.S. Householder, *The theory of matrices in numerical analysis*, Ginn (Blaisdell), Boston, 1964.

[11] B. Kågström and A. Ruhe, *An algorithm for numerical computation of the Jordan normal form of a complex matrix*, ACM Trans. Math. Soft., **6** (1980), pp. 398-419.

[12] E. Kaltofen, *Sparse Hensel lifting*, in Proceedings of EUROCAL'85, Vol. 2, Lecture Notes in Comput. Sci. 204, Springer-Verlag, pp. 4-17, 1985.

[13] E. Kaltofen and M.F. Krishnamoorthy and B.D. Saunders, *Parallel algorithms for matrix normal forms*, Linear algebra appl., **136** (1990), July 15, pp. 189-208.

[14] T.Y. Li, Zhinan Zhang and Tianjun Wang, *Determine the structure of the Jordan normal form of a matrix by symbolic computation*, Linear Algebra and Its Applications, **252** (1997), pp. 221-259.

[15] S. MacLane and C. Birkhoff, *Algebra*, The Macmillan Company, New York, 1967.

[16] B.T. Smith, et al., *Matrix eigensystem routines - EISPACK guide* (Second edition), Springer, New York, 1976.

[17] J.H. Wilkinson, *The algebraic eigenvalue problem*, Clarendon Press, Oxford, 1965.

[18] S. Wolfram, *Mathematica: A system for doing mathematics by computer*, Addison-Wesley, Redwood City, California, 1991.