1

Polynomial Matrices

1.1 Polynomials

Letting $\mathbb{F}$ be a field, e.g., of the real numbers $\mathbb{R}$, the complex numbers $\mathbb{C}$, the rational numbers $\mathbb{W}$, the rational functions $W(s)$ of a complex variable s, etc.,

$$w(s) = \sum_{i=0}^{n} a_i s^i = a_0 + a_1 s + ... + a_n s^n \tag{1.1.1}$$

is called a polynomial $w(s)$ in the variable s over the field $\mathbb{F}$, where $a_i \in \mathbb{F}$ for $i = 0,1,...,n$ are called the coefficients of this polynomial.

The set of polynomials (1.1.1) over the field $\mathbb{F}$ will be denoted by $\mathbb{F}[s]$.

If $a_n \neq 0$, then the nonnegative integral n is called the degree of a polynomial and is denoted deg $w(s)$, i.e., $n = \deg w(s)$. The polynomial (1.1.1) is called monic, if $a_n = 1$ and zero polynomial, if $a_i = 0$ for $i = 0,1,...,n$. The sum of two polynomials

$$w_1(s) = a_0 + a_1 s + ... + a_n s^n, \tag{1.1.2a}$$

$$w_2(s) = b_0 + b_1 s + ... + b_m s^m, \tag{1.1.2b}$$

is defined in the following way

$$w_1(s) + w_2(s) = \begin{cases} \sum_{i=0}^{m} (a_i + b_i) s^i + \sum_{i=m+1}^{n} a_i s^i, \ n > m \\ \sum_{i=0}^{n} (a_i + b_i) s^i, \ n = m \\ \sum_{i=0}^{n} (a_i + b_i) s^i + \sum_{i=n+1}^{m} b_i s^i, \ m > n \end{cases}. \tag{1.1.3}$$

If $n > m$, then the sum is a polynomial of degree n, if $m > n$ then the sum is a polynomial of degree m. If $n = m$ and $a_n+b_n \neq 0$, then this sum is a polynomial of degree n and a polynomial of degree less than n, if $a_n+b_n = 0$. Thus we have

$$\deg\left[w_1(s)+w_2(s)\right] \leq \max\left[\deg\left[w_1(s)\right], \deg\left[w_2(s)\right]\right]. \tag{1.1.4}$$

In the same vein we define the difference of two polynomials.

A polynomial whose coefficients are the products of the coefficients a_i and the scalar λ, i.e.,

$$\lambda\, w(s) = \sum_{i=0}^{n} \lambda\, a_i s^i \;, \tag{1.1.5}$$

is called the product of the polynomial (1.1.1) and the scalar λ (a scalar can be regarded as a polynomial of zero degree).

A polynomial of the form

$$w_1(s)w_2(s) = \sum_{i=0}^{n+m} c_i s^i \tag{1.1.6a}$$

is called the product of the polynomials (1.1.2), where

$$\begin{gathered} c_i = \sum_{k=0}^{i} a_k b_{i-k}, \quad i = 0, 1, \dots, n+m \\ (a_k = 0 \ \text{ for } k > n, \quad b_k = 0 \ \text{ for } k > m). \end{gathered} \tag{1.1.6b}$$

From (1.1.6a) it follows that

$$\deg\left[w_1(s)w_2(s)\right] = n+m\,, \tag{1.1.7}$$

since $a_n b_m \neq 0$ for $a_n \neq 0$, $b_m \neq 0$.

Let $w_2(s)$ in (1.1.2) be a nonzero polynomial and $n > m$, then there exist exactly two polynomials $q(s)$ and $r(s)$ such that

$$w_1(s) = w_2(s)q(s) + r(s)\,, \tag{1.1.8}$$

where

$$\deg\left[r(s)\right] < \deg\left[w_2(s)\right] = m\,. \tag{1.1.9}$$

The polynomial $q(s)$ is called the integer part when $r(s) \neq 0$ and the quotient when $r(s) = 0$, and $r(s)$ is called the remainder.

If $r(s) = 0$, then $w_1(s) = w_2(s)q(s)$; we say then that polynomial $w_1(s)$ is divisible without remainder by the polynomial $w_2(s)$, or equivalently, that polynomial $w_2(s)$ divides without remainder a polynomial $w_1(s)$, which is denoted by $w_1(s) \mid w_2(s)$. We also say that the polynomial $w_2(s)$ is a divisor of the polynomial $w_1(s)$.

Let us consider the polynomials in (1.1.2). We say that a polynomial $d(s)$ is a common divisor of the polynomials $w_1(s)$ and $w_2(s)$ if there exist polynomials $\overline{w}_1(s)$ and $\overline{w}_2(s)$ such that

$$w_1(s) = d(s)\overline{w}_1(s), \quad w_2(s) = d(s)\overline{w}_2(s) . \tag{1.1.10}$$

Polynomial $d_m(s)$ is called a greatest common divisor (GCD) of the polynomials $w_1(s)$ and $w_2(s)$, if every common divisor of these polynomials is a divisor of the polynomial $d_m(s)$. A GCD $d_m(s)$ of polynomials $w_1(s)$ and $w_2(s)$ is determined uniquely up to multiplication by a constant factor and satisfies the equality

$$d_m(s) = w_1(s)m_1(s) + w_2(s)m_2(s) , \tag{1.1.11}$$

where $m_1(s)$ and $m_2(s)$ are polynomials, which we can determine using Euclid's algorithm or the elementary operations method.

The essence of Euclid's algorithm is as follows. Using division of polynomials we determine the sequences of polynomials $q_1, q_2, ..., q_k$ and $r_1, r_2, ..., r_k$ satisfying the following properties

$$\left\{\begin{array}{l} w_1 = w_2 q_1 + r_1 \\ w_2 = r_1 q_2 + r_2 \\ r_1 = r_2 q_3 + r_3 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ r_{k-2} = r_{k-1} q_k + r_k \\ r_{k-1} = r_k q_{k+1} \end{array}\right\} . \tag{1.1.12}$$

We stop computations when the last nonzero remainder r_k is computed and r_{k-1} is found to be divisible without remainder by r_k. With $r_1, r_2, ..., r_{k-1}$ eliminated from (1.1.12) we obtain (1.1.11) for $d_m(s) = r_k$. Thus the last nonzero remainder r_k is a GCD of the polynomials $w_1(s)$ and $w_2(s)$.

Example 1.1.1.
Let

$$w_1 = w_1(s) = s^3 - 3s^2 + 3s - 1, \; w_2 = w_2(s) = s^2 + s + 1 . \tag{1.1.13}$$

Using Euclid's algorithm we compute

$$\begin{aligned} w_1 &= w_2 q_1 + r_1,\ \ q_1 = s-4,\ \ r_1 = 6s+3,\\ w_2 &= r_1 q_2 + r_2,\ \ q_2 = \frac{1}{6}s + \frac{1}{12},\ \ r_2 = \frac{3}{4}. \end{aligned} \tag{1.1.14}$$

Here we stop because r_1 is divisible without remainder by r_2.

Thus r_2 is a GCD of the polynomials in (1.1.13). Elimination of r_1 from (1.1.14) yields

$$w_1(-q_2) + w_2(1+q_1q_2) = r_2\ ,$$

that is,

$$\left(-s^3+3s^2-3s+1\right)\left(\frac{1}{6}s+\frac{1}{12}\right)+\left(s^2+s+1\right)\left(\frac{1}{6}s^2-\frac{7}{12}s+\frac{2}{3}\right)=\frac{3}{4}.$$

The polynomials in (1.1.2) are called relatively prime (or coprime) if and only if their monic GCD is equal to 1. From (1.1.11) for $d_m(s) = 1$ it follows that polynomials $w_1(s)$ and $w_2(s)$ are coprime if and only if there exist polynomials $m_1(s)$ and $m_2(s)$ such that

$$w_1(s)m_1(s) + w_2(s)m_2(s) = 1\ . \tag{1.1.15}$$

Dividing both sides of (1.1.11) by $d_m(s)$, we obtain

$$1 = \widehat{w}_1(s)m_1(s) + \widehat{w}_2(s)m_2(s)\ , \tag{1.1.16}$$

where

$$\widehat{w}_k(s) = \frac{w_k(s)}{d_m(s)} \quad \text{for} \quad k = 1,2,\dots .$$

Thus if $d_m(s)$ is a GCD of the polynomials $w_1(s)$ and $w_2(s)$, then polynomials $\widehat{w}_1(s)$ and $\widehat{w}_2(s)$ are coprime.

Let $s_1,s_2,\dots,s_p$ be different roots of multiplicities $m_1,m_2,\dots,m_p$ ($m_1+m_2+\dots+m_p = n$), respectively, of the equation $w(s) = 0$. The numbers $s_1,s_2,\dots,s_p$ are called the zeros of polynomial (1.1.1). This polynomial can be uniquely written in the form

$$w(s) = a_n(s-s_1)^{m_1}(s-s_2)^{m_2}\dots(s-s_p)^{m_p}\ . \tag{1.1.17}$$

1.2 Basic Notions and Basic Operations on Polynomial Matrices

A matrix whose elements are polynomials over a field $\mathbb{F}$ is called a polynomial matrix over the field $\mathbb{F}$ (briefly polynomial matrix)

$$\mathbf{A}(s) = \left[a_{ij}(s)\right]_{\substack{i=1,\dots,m \\ j=1,\dots,n}} = \begin{bmatrix} a_{11}(s) & \dots & a_{1n}(s) \\ \vdots & \ddots & \vdots \\ a_{m1}(s) & \dots & a_{mn}(s) \end{bmatrix}, \; a_{ij}(s) \in \mathbb{F}(s). \tag{1.2.1}$$

An ordered pair of the number of rows m and columns n, respectively, is called the dimension of matrix (1.2.1) and is denoted by $m\times n$. A set of polynomial matrices of dimension $m\times n$ over a field $\mathbb{F}$ will be denoted by $\mathbb{F}^{m\times n}[s]$.

The following matrix is an example of a 2×2 polynomial matrix over the field of real numbers

$$\mathbf{A}_0(s) = \begin{bmatrix} s^2+2s+1 & s+2 \\ 2s^2+s+3 & 3s^2+s-3 \end{bmatrix} \in \mathbb{R}^{2\times 2}[s]. \tag{1.2.2}$$

Every polynomial matrix can be written in the form of a matrix polynomial. For example, the matrix (1.2.2) can be written in the form of the matrix polynomial

$$\mathbf{A}_0(s) = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} s^2 + \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} s + \begin{bmatrix} 1 & 2 \\ 3 & -3 \end{bmatrix} = \mathbf{A}_2 s^2 + \mathbf{A}_1 s + \mathbf{A}_0. \tag{1.2.3}$$

Let a matrix of the form (1.2.1) be expressed as the matrix polynomial

$$\mathbf{A}(s) = \mathbf{A}_q s^q + \dots + \mathbf{A}_1 s + \mathbf{A}_0, \; \mathbf{A}_k \in \mathbb{R}^{m\times n}, \quad k = 0, 1, \dots, q. \tag{1.2.4}$$

If $\mathbf{A}_q$ is not a zero matrix, then number q is called its degree and is denoted by $q = \deg \mathbf{A}(s)$. For example, the matrix (1.2.2) (and also (1.2.3)) has the degree two $q = 2$.

If $n = m$ and $\det \mathbf{A}_q \neq 0$, then matrix (1.2.4) is called regular.

The sum of two polynomial matrices

$$\begin{aligned} \mathbf{A}(s) &= \left[a_{ij}(s)\right]_{\substack{i=1,\dots,m \\ j=1,\dots,n}} = \sum_{k=0}^{q} \mathbf{A}_k s^k \quad \text{and} \\ \mathbf{B}(s) &= \left[b_{ij}(s)\right]_{\substack{i=1,\dots,m \\ j=1,\dots,n}} = \sum_{k=0}^{t} \mathbf{B}_k s^k \end{aligned} \tag{1.2.5}$$

of the same dimension $m\times n$ is defined in the following way

$$\mathbf{A}(s)+\mathbf{B}(s)=$$
$$=\left[a_{ij}(s)+b_{ij}(s)\right]_{\substack{i=1,\dots,m\\ j=1,\dots,n}}=\begin{cases}\sum_{k=0}^{t}(\mathbf{A}_k+\mathbf{B}_k)s^k+\sum_{k=t+1}^{q}\mathbf{A}_k s^k & q>t\\ \sum_{k=0}^{q}(\mathbf{A}_k+\mathbf{B}_k)s^k & q=t\\ \sum_{k=0}^{q}(\mathbf{A}_k+\mathbf{B}_k)s^k+\sum_{k=q+1}^{t}\mathbf{B}_k s^k & q<t\end{cases}. \tag{1.2.6}$$

If $q = t$ and $\mathbf{A}_q + \mathbf{B}_q \neq 0$, then the sum in (1.2.6) is a polynomial matrix of degree q, and if $\mathbf{A}_q + \mathbf{B}_q = 0$, then this sum is a polynomial matrix of a degree not greater than q. Thus we have

$$\deg\left[\mathbf{A}(s)+\mathbf{B}(s)\right]\le \max\left[\deg\left[\mathbf{A}(s)\right],\deg\left[\mathbf{B}(s)\right]\right]. \tag{1.2.7}$$

In the same vein, we define the difference of two polynomial matrices.

A polynomial matrix where every entry is the product of an entry of the matrix (1.2.1) and the scalar λ is called the product of the polynomial matrix (1.2.1) and the scalar λ

$$\lambda\,\mathbf{A}(s)=\left[\lambda\, a_{ij}(s)\right]_{\substack{i=1,\dots,m\\ j=1,\dots,n}}.$$

From this definition for $\lambda \neq 0$, we have $\deg[\lambda\mathbf{A}(s)] = \deg[\mathbf{A}(s)]$.

Multiplication of two polynomial matrices can be carried out if and only if the number of columns of the first matrix (1.2.1) is equal to the number of rows of the second matrix

$$\mathbf{B}(s)=\left[b_{ij}(s)\right]_{\substack{i=1,\dots,n\\ j=1,\dots,p}}=\sum_{k=0}^{t}\mathbf{B}_k s^k. \tag{1.2.8}$$

A polynomial matrix of the form

$$\mathbf{C}(s)=\left[c_{ij}(s)\right]_{\substack{i=1,\dots,m\\ j=1,\dots,p}}=\mathbf{A}(s)\mathbf{B}(s)=\sum_{k=0}^{q+t}\mathbf{C}_k s^k \tag{1.2.9}$$

is called the product of these polynomial matrices, where

$$\begin{gathered}\mathbf{C}_k=\sum_{l=0}^{k}\mathbf{A}_l\mathbf{B}_{k-l}\quad k=0,1,\dots,q+t\\ (\mathbf{A}_l=0,\ l>q,\ \mathbf{B}_l=0,\ l>t).\end{gathered} \tag{1.2.10}$$

From (1.2.10) it follows that $\mathbf{C}_{q+t} = \mathbf{A}_q\mathbf{B}_t$ and this matrix is a nonzero one if at least one of the matrices $\mathbf{A}_q$ and $\mathbf{B}_t$ is nonsingular, in other words one of the matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$ is a regular one. Thus we have the relationship

$$\begin{aligned}
&\deg\left[\mathrm{A(s)B(s)}\right] = \deg\left[\mathrm{A(s)}\right] + \deg\left[\mathrm{B(s)}\right] \text{ if at least one of these} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{matrices is regular,} \qquad (1.2.11) \\
&\deg\left[\mathrm{A(s)B(s)}\right] \le \deg\left[\mathrm{A(s)}\right] + \deg\left[\mathrm{B(s)}\right] \text{ otherwise.}
\end{aligned}$$

For example, the product of the polynomial matrices

$$\mathbf{A}(s) = \begin{bmatrix} s^2+s & -2s^2+s+1 \\ -s^2+2s-1 & 2s^2+2 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix}s^2 + \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}s + \begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix},$$

$$\mathbf{B}(s) = \begin{bmatrix} 2s+2 & s+3 \\ s-1 & \frac{1}{2}s+1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & \frac{1}{2} \end{bmatrix}s + \begin{bmatrix} 2 & 3 \\ -1 & 1 \end{bmatrix}$$

is the following polynomial matrix

$$\mathbf{A}(s)\mathbf{B}(s) = \begin{bmatrix} 7s^2+2s-1 & \frac{5}{2}s^2+\frac{9}{2}s+1 \\ 4s-4 & s^2+6s-1 \end{bmatrix} = \begin{bmatrix} 7 & \frac{5}{2} \\ 0 & 1 \end{bmatrix}s^2 + \begin{bmatrix} 2 & \frac{9}{2} \\ 4 & 6 \end{bmatrix}s + \begin{bmatrix} -1 & 1 \\ -4 & -1 \end{bmatrix},$$

whose degree is smaller than the sum deg $[\mathbf{A}(s)]$ + deg $[\mathbf{B}(s)]$, since

$$\mathbf{A}_2\mathbf{B}_1 = \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 1 & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

The matrix (1.2.4) can be written in the form

$$\mathbf{A}(s) = s^q\mathbf{A}_q + \ldots + s\mathbf{A}_1 + \mathbf{A}_0, \tag{1.2.12}$$

since multiplication of the matrix $\mathbf{A}_i$ ($i = 1,2,\ldots,q$) by the scalar s is commutative. Substituting the matrix $\mathbf{S}$ in place of the scalar s into (1.2.4) and (1.2.12), we obtain the following, usually different, matrices

$$\begin{aligned}
\mathbf{A}_p(\mathbf{S}) &= \mathbf{A}_q\mathbf{S}^q + \ldots + \mathbf{A}_1\mathbf{S} + \mathbf{A}_0, \\
\mathbf{A}_l(\mathbf{S}) &= \mathbf{S}^q\mathbf{A}_q + \ldots + \mathbf{S}\mathbf{A}_1 + \mathbf{A}_0.
\end{aligned}$$

The matrix $\mathbf{A}_p(\mathbf{S})$ ($\mathbf{A}_l(\mathbf{S})$) is called the right-sided (left-sided) value of the matrix $\mathbf{A}(s)$ for $s = \mathbf{S}$.

Let

$$\mathbf{C}(s) = \mathbf{A}(s) + \mathbf{B}(s) .$$

It is easy to verify that

$$\mathbf{C}_p(\mathbf{S}) = \mathbf{A}_p(\mathbf{S}) + \mathbf{B}_p(\mathbf{S})$$

and

$$\mathbf{C}_l(\mathbf{S}) = \mathbf{A}_l(\mathbf{S}) + \mathbf{B}_l(\mathbf{S}) .$$

Consider the polynomial matrices in (1.2.5).

Theorem 1.2.1. If the matrix $\mathbf{S}$ commutes with the matrices $\mathbf{A}_i$ for $i = 1,2,...,q$ and $\mathbf{B}_j$ for $j = 1,2,...,t$, then the right-sided and the left-sided value of the product of the matrices in (1.2.5) for $s = \mathbf{S}$ is equal to the product of the right-sided and left-sided values respectively, of these matrices for $s = \mathbf{S}$.

Proof. Taking into account the polynomial matrices in (1.2.5) we can write

$$\mathbf{D}(s) = \mathbf{A}(s)\mathbf{B}(s) = \left(\sum_{i=0}^{q} \mathbf{A}_i s^i\right)\left(\sum_{j=0}^{t} \mathbf{B}_j s^j\right) = \sum_{i=0}^{q}\sum_{j=0}^{t} \mathbf{A}_i \mathbf{B}_j s^{i+j}$$

and

$$\mathbf{D}(s) = \mathbf{A}(s)\mathbf{B}(s) = \left(\sum_{i=0}^{q} s^i \mathbf{A}_i\right)\left(\sum_{j=0}^{t} s^j \mathbf{B}_j\right) = \sum_{i=0}^{q}\sum_{j=0}^{t} s^{i+j} \mathbf{A}_i \mathbf{B}_j .$$

Substituting the matrix $\mathbf{S}$ in place of the scalar s, we obtain

$$\mathbf{D}_p(\mathbf{S}) = \sum_{i=0}^{q}\sum_{j=0}^{t} \mathbf{A}_i \mathbf{B}_j \mathbf{S}^{i+j} = \left(\sum_{i=0}^{q} \mathbf{A}_i \mathbf{S}^i\right)\left(\sum_{j=0}^{t} \mathbf{B}_j \mathbf{S}^j\right) = \mathbf{A}_p(\mathbf{S})\mathbf{B}_p(\mathbf{S}) ,$$

since $\mathbf{B}_j\mathbf{S} = \mathbf{S}\mathbf{B}_j$ for $j = 1,2,...,t$ and

$$\mathbf{D}_l(\mathbf{S}) = \sum_{i=0}^{p}\sum_{j=0}^{q} \mathbf{S}^{i+j} \mathbf{A}_i \mathbf{B}_j = \left(\sum_{i=0}^{p} \mathbf{S}^i \mathbf{A}_i\right)\left(\sum_{j=0}^{q} \mathbf{S}^j \mathbf{B}_j\right) = \mathbf{A}_l(\mathbf{S})\mathbf{B}_l(\mathbf{S}) ,$$

since $\mathbf{S}\mathbf{A}_i = \mathbf{A}_i\mathbf{S}$ for $i = 1,2,...,q$. ■

1.3 Division of Polynomial Matrices

Consider the polynomial matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$ where det $\mathbf{A}(s) \neq 0$ and $\deg \mathbf{A}(s) < \deg \mathbf{B}(s)$. The matrix $\mathbf{A}(s)$ may be not regular, i.e., the matrix of coefficients of the highest power of variable s may be singular.

Theorem 1.3.1. If det $\mathbf{A}(s) \neq 0$, then for the pair of polynomial matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$, $\deg \mathbf{B}(s) > \deg \mathbf{A}(s)$ there exists a pair of matrices $\mathbf{Q}_p(s)$, $\mathbf{R}_p(s)$ such that the following equality is satisfied

$$\mathbf{B}(s) = \mathbf{Q}_p(s)\mathbf{A}(s) + \mathbf{R}_p(s), \quad \deg \mathbf{A}(s) > \deg \mathbf{R}_p(s)\,, \tag{1.3.1a}$$

and there exists a pair of matrices $\mathbf{Q}_l(s)$, $\mathbf{R}_l(s)$ such that the following equality holds

$$\mathbf{B}(s) = \mathbf{A}(s)\mathbf{Q}_l(s) + \mathbf{R}_l(s), \quad \deg \mathbf{A}(s) > \deg \mathbf{R}_l(s)\,. \tag{1.3.1b}$$

Proof. Dividing the elements of matrix $\mathbf{B}(s)$ Adj $\mathbf{A}(s)$ by a polynomial det $\mathbf{A}(s)$, we obtain a pair of matrices $\mathbf{Q}_p(s)$, $\mathbf{R}_l(s)$ such that

$$\begin{aligned} &\mathbf{B}(s)\mathrm{Adj}\,\mathbf{A}(s) = \\ &= \mathbf{Q}_p(s)\det \mathbf{A}(s) + \mathbf{R}_1(s),\ \deg\left[\det \mathbf{A}(s)\right] > \deg \mathbf{R}_1(s). \end{aligned} \tag{1.3.2}$$

Post-multiplication of (1.3.2) by $\mathbf{A}(s)/\det \mathbf{A}(s)$ yields

$$\mathbf{B}(s) = \mathbf{Q}_p(s)\mathbf{A}(s) + \mathbf{R}_p(s)\,, \tag{1.3.3}$$

since Adj $\mathbf{A}(s)$ $\mathbf{A}(s) = \mathbf{I}_n$ det $\mathbf{A}(s)$, where

$$\mathbf{R}_p(s) = \frac{\mathbf{R}_1(s)\mathbf{A}(s)}{\det \mathbf{A}(s)}\,. \tag{1.3.4}$$

From (1.3.4) we have

$$\deg \mathbf{R}_p(s) = \deg \mathbf{R}_1(s) + \deg \mathbf{A}(s) - \deg\left[\det \mathbf{A}(s)\right] < \deg \mathbf{A}(s)\,,$$

since deg [det $\mathbf{A}(s)$] > deg $\mathbf{R}_1(s)$.

The proof of equality (1.3.1b) is similar. ■

Remark 1.3.1. The pairs of matrices $\mathbf{Q}_p(s)$, $\mathbf{R}_p(s)$ and $\mathbf{Q}_l(s)$, $\mathbf{R}_l(s)$ satisfying the equality (1.3.1) are not uniquely determined (are not unique), since

$$\mathbf{B}(s) = [\mathbf{Q}_p(s) + \mathbf{C}(s)]\mathbf{A}(s) + \mathbf{R}_p(s) - \mathbf{A}(s)\mathbf{C}(s) \tag{1.3.5a}$$

and

$$\mathbf{B}(s) = \mathbf{A}(s)[\mathbf{Q}_l(s) + \mathbf{C}(s)] + \mathbf{R}_l(s) - \mathbf{A}(s)\mathbf{C}(s) \tag{1.3.5b}$$

are satisfied for an arbitrary matrix $\mathbf{C}(s)$ satisfying

$$\deg\left[\mathbf{C}(s)\mathbf{A}(s)\right] < \deg \mathbf{A}(s), \quad \deg\left[\mathbf{A}(s)\mathbf{C}(s)\right] < \deg \mathbf{A}(s).$$

Example 1.3.1.
For the matrices

$$\mathbf{A}(s) = \begin{bmatrix} s & 1 \\ -1 & 1 \end{bmatrix}, \quad \mathbf{B}(s) = \begin{bmatrix} s & -s \\ -1 & s^2+1 \end{bmatrix}$$

determine the matrices $\mathbf{Q}_p(s)$, $\mathbf{R}_p(s)$ satisfying the equality (1.3.1a).

In this case, det $\mathbf{A}_1 = 0$ and det $\mathbf{A}(s) = s+1$.
We compute

$$\text{Adj } \mathbf{A}(s) = \begin{bmatrix} 1 & -1 \\ 1 & s \end{bmatrix}, \quad \mathbf{B}(s)\text{Adj } \mathbf{A}(s) = \begin{bmatrix} 0 & -s^2-s \\ s^2 & s^3+s+1 \end{bmatrix},$$

and with (1.3.2) taken into account we have

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

i.e.,

$$\mathbf{Q}_p(s) = \begin{bmatrix} 0 & -s \\ s-1 & s^2-s+2 \end{bmatrix}, \quad \mathbf{R}_1(s) = \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}.$$

According to (1.3.4) we obtain

$$\mathbf{R}_p(s) = \frac{\mathbf{R}_1(s)\mathbf{A}(s)}{\det \mathbf{A}(s)} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Consider two polynomial matrices

$$\mathbf{A}(s) = \mathbf{A}_n s^n + \mathbf{A}_{n-1}s^{n-1} + \ldots + \mathbf{A}_1 s + \mathbf{A}_0, \tag{1.3.6a}$$

$$\mathbf{B}(s) = \mathbf{B}_m s^m + \mathbf{B}_{m-1}s^{m-1} + \ldots + \mathbf{B}_1 s + \mathbf{B}_0. \tag{1.3.6b}$$

Theorem 1.3.2. If $\mathbf{A}(s)$ and $\mathbf{B}(s)$ are square polynomial matrices of the same dimensions, and $\mathbf{A}(s)$ is regular ($\det \mathbf{A}_n \neq 0$), then there exist exactly one pair of polynomial matrices $\mathbf{Q}_p(s)$, $\mathbf{R}_p(s)$ satisfying the equality

$$\mathbf{B}(s) = \mathbf{Q}_p(s)\mathbf{A}(s) + \mathbf{R}_p(s)\,, \tag{1.3.7a}$$

and exactly one pair of polynomial matrices $\mathbf{Q}_l(s)$, $\mathbf{R}_l(s)$ satisfying the equality

$$\mathbf{B}(s) = \mathbf{A}(s)\mathbf{Q}_l(s) + \mathbf{R}_l(s) \tag{1.3.7b}$$

where

$$\deg \mathbf{A}(s) > \deg \mathbf{R}_p(s), \quad \deg \mathbf{A}(s) > \deg \mathbf{R}_l(s)\,.$$

Proof. If $n > m$, then $\mathbf{Q}_p(s) = 0$ and $\mathbf{R}_p(s) = \mathbf{B}(s)$. Assume that $m \geq n$. By the assumption $\det \mathbf{A}_n \neq 0$ there exists the inverse matrix $\mathbf{A}_{n\text{-}1}$. Note that the matrix $\mathbf{B}_m\mathbf{A}_n^{-1}s^{m-n}\mathbf{A}(s)$ has a term in the highest power of s, equal to $\mathbf{B}_m s^m$. Hence

$$\mathbf{B}(s) = \mathbf{B}_m\mathbf{A}_n^{-1}s^{m-n}\mathbf{A}(s) + \mathbf{B}^{(1)}(s)\,,$$

where $\mathbf{B}^{(1)}(s)$ is a polynomial matrix of degree $m_1 \leq m\text{-}1$ of the form

$$\mathbf{B}^{(1)}(s) = \mathbf{B}_{m_1}^{(1)}s^{m_1} + \mathbf{B}_{m_1-1}^{(1)}s^{m_1-1} + \ldots + \mathbf{B}_1^{(1)}s + \mathbf{B}_0^{(1)}\,.$$

If $m_1 \geq n$, then we repeat this procedure, taking the matrix $\mathbf{B}_{m_1}^{(1)}$ instead of the matrix $\mathbf{B}_m$, and obtain

$$\mathbf{B}^{(1)}(s) = \mathbf{B}_{m_1}^{(1)}\mathbf{A}_n^{-1}s^{m_1-n}\mathbf{A}(s) + \mathbf{B}^{(2)}(s)\,,$$

where

$$\mathbf{B}^{(2)}(s) = \mathbf{B}_{m_2}^{(2)}s^{m_2} + \mathbf{B}_{m_2-1}^{(2)}s^{m_2-1} + \ldots + \mathbf{B}_1^{(2)}s + \mathbf{B}_0^{(2)} \quad (m_2 < m_1)\,.$$

Continuing this procedure, we obtain the sequence of polynomial matrices $\mathbf{B}(s)$, $\mathbf{B}^{(1)}(s)$, $\mathbf{B}^{(2)}(s), \ldots,$ of decreasing degrees m, m_1, $m_2, \ldots,$ respectively. In step r, we obtain the matrix $\mathbf{B}^{(r)}(s)$ of degree $m_r < n$ and

$$\mathbf{B}(s) = \left(\mathbf{B}_m\mathbf{A}_n^{-1}s^{m-n} + \mathbf{B}_{m_1}^{(1)}\mathbf{A}_n^{-1}s^{m_1-n} + \ldots + \mathbf{B}_{m_{r-1}}^{(r-1)}\mathbf{A}_n^{-1}s^{m_{r-1}-n}\right)\mathbf{A}(s) + \mathbf{B}^{(r)}(s)\,,$$

that is the equality (1.3.7a) for

$$\mathbf{Q}_p(s) = \mathbf{B}_m \mathbf{A}_n^{-1} s^{m-n} + \mathbf{B}_{m_1}^{(1)} \mathbf{A}_n^{-1} s^{m_1-n} + \ldots + \mathbf{B}_{m_{r-1}}^{(r-1)} \mathbf{A}_n^{-1} s^{m_{r-1}-n},$$
$$\mathbf{R}_p(s) = \mathbf{B}^{(r)}(s). \tag{1.3.8}$$

Now we will show that there exists only one pair $\mathbf{Q}_p(s)$, $\mathbf{R}_p(s)$ satisfying (1.3.7a). Assume that there exist two different pairs $\mathbf{Q}_p^{(1)}(s)$, $\mathbf{R}_p^{(1)}(s)$ and $\mathbf{Q}_p^{(2)}(s)$, $\mathbf{R}_p^{(2)}(s)$ such that

$$\mathbf{B}(s) = \mathbf{Q}_p^{(1)}(s)\mathbf{A}(s) + \mathbf{R}_p^{(1)}(s) \tag{1.3.9a}$$

and

$$\mathbf{B}(s) = \mathbf{Q}_p^{(2)}(s)\mathbf{A}(s) + \mathbf{R}_p^{(2)}(s), \tag{1.3.9b}$$

where deg $\mathbf{A}(s) >$ deg $\mathbf{R}_p^{(1)}(s)$ and deg $\mathbf{A}(s) >$ deg $\mathbf{R}_p^{(2)}(s)$. From (1.3.9) we have

$$\left[\mathbf{Q}_p^{(1)}(s) - \mathbf{Q}_p^{(2)}(s)\right]\mathbf{A}(s) = \mathbf{R}_p^{(2)}(s) - \mathbf{R}_p^{(1)}(s). \tag{1.3.10}$$

For $\mathbf{Q}_p^{(1)}(s) \neq \mathbf{Q}_p^{(2)}(s)$ the matrix $[\mathbf{Q}_p^{(1)}(s) - \mathbf{Q}_p^{(2)}(s)]\mathbf{A}(s)$ is a polynomial matrix of a degree greater than n, and $[\mathbf{R}_p^{(2)}(s) - \mathbf{R}_p^{(1)}(s)]$ is a polynomial matrix of a degree less than n. Hence from (1.3.10) it follows that $\mathbf{Q}_p^{(1)}(s) = \mathbf{Q}_p^{(2)}(s)$ and $\mathbf{R}_p^{(1)}(s) = \mathbf{R}_p^{(2)}(s)$. Similarly one can prove that

$$\mathbf{Q}_l(s) = \mathbf{A}_n^{-1} \mathbf{B}_m s^{m-n} + \mathbf{A}_n^{-1} \mathbf{B}_{m_1}^{(1)} s^{m_1-n} + \ldots + \mathbf{A}_n^{-1} \mathbf{B}_{m_{r-1}}^{(r-1)} s^{m_{r-1}-n},$$
$$\mathbf{R}_l(s) = \mathbf{B}^{(r)}(s). \tag{1.3.11}$$

■

The matrices $\mathbf{Q}_p(s)$, $\mathbf{R}_p(s)$ ($\mathbf{Q}_l(s)$, $\mathbf{R}_l(s)$) are called, respectively: the right (left) quotient and the remainder from division of the matrix $\mathbf{B}(s)$ by the matrix $\mathbf{A}(s)$.

From the proof of Theorem 1.3.2 the following algorithm for determining matrices $\mathbf{Q}_p(s)$ and $\mathbf{R}_p(s)$ ($\mathbf{Q}_l(s)$ and $\mathbf{R}_l(s)$) ensues.

Procedure 1.3.1.
Step 1: Given matrix $\mathbf{A}_n$ compute $\mathbf{A}_n^{-1}$.
Step 2: Compute

$$\mathbf{B}_m \mathbf{A}_n^{-1} s^{m-n} \mathbf{A}(s) \quad \left(\mathbf{A}(s)\mathbf{A}_n^{-1}\mathbf{B}_m s^{m-n}\right)$$

and

$$\mathbf{B}^{(1)}(s) = \mathbf{B}(s) - \mathbf{B}_m \mathbf{A}_n^{-1} s^{m-n} \mathbf{A}(s) = \mathbf{B}_{m_1}^{(1)} s^{m_1} + \ldots + \mathbf{B}_1^{(1)} s + \mathbf{B}_0^{(1)}$$

$$\left(\mathbf{B}^{(1)}(s) = \mathbf{B}(s) - \mathbf{A}(s)\mathbf{A}_n^{-1}\mathbf{B}_m s^{m-n} = \mathbf{B}_{m_1}^{(1)} s^{m_1} + \ldots + \mathbf{B}_1^{(1)} s + \mathbf{B}_0^{(1)}\right).$$

Step 3: If $m_1 \geq n$, then compute

$$\mathbf{B}_{m_1}^{(1)}\mathbf{A}_n^{-1}s^{m_1-n}\mathbf{A}(s)\ \left(\mathbf{A}(s)\mathbf{A}_n^{-1}\mathbf{B}_{m_1}^{(1)}s^{m_1-n}\right)$$

and

$$\mathbf{B}^{(2)}(s) = \mathbf{B}^{(1)}(s) - \mathbf{B}_{m_1}^{(1)}\mathbf{A}_n^{-1}s^{m_1-n}\mathbf{A}(s) = \mathbf{B}_{m_2}^{(2)}s^{m_2} + \ldots + \mathbf{B}_1^{(2)}s + \mathbf{B}_0^{(2)}$$
$$\left(\mathbf{B}^{(2)}(s) = \mathbf{B}^{(1)}(s) - \mathbf{A}(s)\mathbf{A}_n^{-1}\mathbf{B}_{m_1}^{(1)}s^{m_1-n} = \mathbf{B}_{m_2}^{(2)}s^{m_2} + \ldots + \mathbf{B}_1^{(2)}s + \mathbf{B}_0^{(2)}\right).$$

Step 4: If $m_2 \geq n$, then substituting in the above equalities m_1 and $\mathbf{B}^{(1)}(s)$ by m_2 and $\mathbf{B}^{(2)}(s)$, respectively, compute $\mathbf{B}^{(3)}(s)$. Repeat this procedure r times until $m_r < n$.

Step 5: Compute the matrices $\mathbf{Q}_p(s)$, $\mathbf{R}_p(s)$ ($\mathbf{Q}_l(s)$, $\mathbf{R}_l(s)$).

Example 1.3.2.
Given the matrices

$$\mathbf{A}(s) = \begin{bmatrix} s^2+1 & -s \\ s & s^2+s \end{bmatrix} \text{ and } \mathbf{B}(s) = \begin{bmatrix} s^4+s^2+1 & s^3-s^2+2s \\ 2s^2+s & s^3+s+2 \end{bmatrix},$$

determine matrices $\mathbf{Q}_p(s)$, $\mathbf{R}_p(s)$ and $\mathbf{Q}_l(s)$, $\mathbf{R}_l(s)$ satisfying (1.3.7).

Matrix $\mathbf{A}(s)$ is regular, since

$$\mathbf{A}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } \mathbf{B}_4 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Using Procedure 1.3.1 we compute the following.

Steps 1–3: In this case,

$$\mathbf{B}_4\mathbf{A}_2^{-1}s^2\mathbf{A}(s) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}s^2\begin{bmatrix} s^2+1 & -s \\ s & s^2+s \end{bmatrix} = \begin{bmatrix} s^4+s & -s^3 \\ 0 & 0 \end{bmatrix}$$

and

$$\mathbf{B}^{(1)}(s) = \mathbf{B}(s) - \mathbf{B}_4\mathbf{A}_2^{-1}s^2\mathbf{A}(s)$$
$$= \begin{bmatrix} s^4+s^2+1 & s^3-s^2+2s \\ 2s^2+s & s^3+s+2 \end{bmatrix} - \begin{bmatrix} s^4+s^2 & -s^3 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2s^3-s^2+2s \\ 2s^2+s & s^3+s+2 \end{bmatrix}.$$

Since $m_1 = 3$, $n = 2$, and

$$\mathbf{B}_3^{(1)} = \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix},$$

we have

$$\mathbf{B}_3^{(1)}\mathbf{A}_2^{-1}s\,\mathbf{A}(s) = \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}s\begin{bmatrix} s^2+1 & -s \\ s & s^2+s \end{bmatrix} = \begin{bmatrix} 2s^2 & 2s^3+2s^2 \\ s^2 & s^3+s^2 \end{bmatrix}$$

and

$$\mathbf{B}^{(2)}(s) = \mathbf{B}^{(1)}(s) - \mathbf{B}_3^{(1)}\mathbf{A}_2^{-1}s\,\mathbf{A}(s) =$$
$$= \begin{bmatrix} 1 & 2s^3-s^2+2s \\ 2s^2+s & s^3+s+2 \end{bmatrix} - \begin{bmatrix} 2s^2 & 2s^3+2s^2 \\ s^2 & s^3+s^2 \end{bmatrix} = \begin{bmatrix} -2s^2+1 & -3s^2+2s \\ s^2+s & -s^2+s+2 \end{bmatrix}.$$

Step 4: We repeat the procedure, since $m_2 = 2 = n$. Taking into account that

$$\mathbf{B}_2^{(2)} = \begin{bmatrix} -2 & -3 \\ 1 & -1 \end{bmatrix},$$

we compute

$$\mathbf{B}_2^{(2)}\mathbf{A}_2^{-1}\mathbf{A}(s) =$$
$$= \begin{bmatrix} -2 & -3 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} s^2+1 & -s \\ s & s^2+s \end{bmatrix} = \begin{bmatrix} -2s^2-3s-2 & -3s^2-s \\ s^2-s+1 & -s^2-2s \end{bmatrix}$$

and

$$\mathbf{B}^{(3)}(s) = \mathbf{B}^{(2)}(s) - \mathbf{B}_2^{(2)}\mathbf{A}_2^{-1}\mathbf{A}(s) = \begin{bmatrix} -2s^2+1 & -3s^2+2s \\ s^2+s & -s^2+s+2 \end{bmatrix}$$
$$- \begin{bmatrix} -2s^2-3s-2 & -3s^2-s \\ s^2-s+1 & -s^2-2s \end{bmatrix} = \begin{bmatrix} 3s+3 & 3s \\ 2s-1 & 3s+2 \end{bmatrix}.$$

Step 5: The degree of this matrix is less than the degree of the matrix $\mathbf{A}(s)$. Hence, according to (1.3.8), we obtain

$$\mathbf{Q}_p(s) = \mathbf{B}_4\mathbf{A}_2^{-1}s^2 + \mathbf{B}_3^{(1)}\mathbf{A}_2^{-1}s + \mathbf{B}_2^{(2)}\mathbf{A}_2^{-1} =$$
$$= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} s^2 + \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix} s + \begin{bmatrix} -2 & -3 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} s^2-2 & 2s-3 \\ 1 & s-1 \end{bmatrix}$$

and

$$\mathbf{R}_p(s) = \mathbf{B}^{(3)}(s) = \begin{bmatrix} 3s+3 & 3s \\ 2s-1 & 3s+2 \end{bmatrix}.$$

We compute $\mathbf{Q}_l(s)$ and $\mathbf{R}_l(s)$ using Procedure 1.3.1.
Steps 1–3: We compute

$$\mathbf{A}(s)\mathbf{A}_2^{-1}\mathbf{B}_4 s^2 = \begin{bmatrix} s^2+1 & -s \\ s & s^2+s \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} s^2 = \begin{bmatrix} s^4+s^2 & 0 \\ s^3 & 0 \end{bmatrix}$$

and

$$\mathbf{B}^{(1)}(s) = \mathbf{B}(s) - \mathbf{A}(s)\mathbf{A}_2^{-1}\mathbf{B}_4 s^2 = \begin{bmatrix} s^4+s^2+1 & s^3-s^2+2s \\ 2s^2+s & s^3+s+2 \end{bmatrix}$$
$$- \begin{bmatrix} s^4+s^2 & 0 \\ s^3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & s^3-s^2+2s \\ -s^3+2s^2+s & s^3+s+2 \end{bmatrix}.$$

Taking into account that $m_1 = 3 > n = 2$ and

$$\mathbf{B}_3^{(1)} = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix},$$

we compute

$$\mathbf{A}(s)\mathbf{A}_2^{-1}\mathbf{B}_3^{(1)}s = \begin{bmatrix} s^2+1 & -s \\ s & s^2+s \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix} s = \begin{bmatrix} s^2 & s^3-s^2+s \\ -s^3-s^2 & s^3+2s^2 \end{bmatrix}$$

and

$$\mathbf{B}^{(2)}(s) = \mathbf{B}^{(1)}(s) - \mathbf{A}(s)A_2^{-1}\mathbf{B}_3^{(1)}s = \begin{bmatrix} 1 & s^3-s^2+2s \\ -s^3+2s^2+s & s^3+s+2 \end{bmatrix} +$$
$$- \begin{bmatrix} s^2 & s^3-s^2+s \\ -s^3-s^2 & s^3+2s^2 \end{bmatrix} = \begin{bmatrix} -s^2+1 & s \\ 3s^2+s & -2s^2+s+2 \end{bmatrix}.$$

Step 4: We repeat the procedure, since $m_2 = 2 = n$. Taking into account that

$$\mathbf{B}_2^{(2)} = \begin{bmatrix} -1 & 0 \\ 3 & -2 \end{bmatrix},$$

we have

$$\mathbf{A}(s)\mathbf{A}_2^{-1}\mathbf{B}_2^{(2)} = \begin{bmatrix} s^2+1 & -s \\ s & s^2+s \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} -1 & 0 \\ 3 & -2 \end{bmatrix} = \begin{bmatrix} -s^2-3s-1 & 2s \\ 3s^2+2 & -2s^2-2s \end{bmatrix}$$

and

$$\mathbf{B}^{(3)}(s) = \mathbf{B}^{(2)}(s) - \mathbf{A}(s)\mathbf{A}_2^{-1}\mathbf{B}_2^{(2)} = \begin{bmatrix} -s^2+1 & s \\ 3s^2+s & -2s^2+s+2 \end{bmatrix}$$
$$-\begin{bmatrix} -s^2-3s-1 & 2s \\ 3s^2+2s & -2s^2-2s \end{bmatrix} = \begin{bmatrix} 3s+2 & -s \\ -s & 3s+2 \end{bmatrix}.$$

Step 5: The degree of this matrix is less than the degree of matrix $\mathbf{A}(s)$. Hence according to (1.3.11), we have

$$\mathbf{Q}_l(s) = \mathbf{A}_2^{-1}\mathbf{B}_4 s^2 + \mathbf{A}_2^{-1}\mathbf{B}_3^{(1)} s + \mathbf{A}_2^{-1}\mathbf{B}_2^{(2)}$$
$$= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} s^2 + \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix} s + \begin{bmatrix} -1 & 0 \\ 3 & -2 \end{bmatrix} = \begin{bmatrix} s^2-1 & s \\ -s+3 & s-2 \end{bmatrix},$$
$$\mathbf{R}_l(s) = \mathbf{B}^{(3)}(s) = \begin{bmatrix} 3s+2 & -s \\ -s & 3s+2 \end{bmatrix}.$$

1.4 Generalized Bezoute Theorem and the Cayley–Hamilton Theorem

Let us consider the division of a square polynomial matrix

$$\mathbf{F}(s) = \mathbf{F}_n s^n + \mathbf{F}_{n-1} s^{n-1} + \cdots + \mathbf{F}_1 s + \mathbf{F}_0 \in \mathbb{C}^{m\times m}[s] \tag{1.4.1}$$

by a polynomial matrix of the first degree $[\mathbf{I}_m s - \mathbf{A}]$, where $\mathbf{F}_k \in \mathbb{C}^{m\times m}$, $k = 0,1,\ldots,n$ and $\mathbf{A} \in \mathbb{C}^{m\times m}$. The right (left) $\mathbf{R}_p$ ($\mathbf{R}_l$) remainder from division of $\mathbf{F}(s)$ by $[\mathbf{I}_m s - \mathbf{A}]$ is a polynomial matrix of zero degree, i.e., it does not depend on s.

Theorem 1.4.1. (Generalised Bezoute theorem). The right (left) remainder $\mathbf{R}_p$ ($\mathbf{R}_l$) from division of the matrix $\mathbf{F}(s)$ by $[\mathbf{I}_m s - \mathbf{A}]$ is equal to $\mathbf{F}_p(\mathbf{A})$ ($\mathbf{F}_l(\mathbf{A})$), i.e.,

$$\mathbf{R}_p = \mathbf{F}_p(\mathbf{A}) = \mathbf{F}_n\mathbf{A}^n + \mathbf{F}_{n-1}\mathbf{A}^{n-1} + \cdots + \mathbf{F}_1\mathbf{A} + \mathbf{F}_0 \in \mathbb{C}^{m\times m} \quad (1.4.2a)$$

$$\left(\mathbf{R}_l = \mathbf{F}_l(\mathbf{A}) = \mathbf{A}^n\mathbf{F}_n + \mathbf{A}^{n-1}\mathbf{F}_{n-1} + \cdots + \mathbf{A}\mathbf{F}_1 + \mathbf{F}_0 \in \mathbb{C}^{m\times m}\right). \quad (1.4.2b)$$

Proof. Post-dividing the matrix $\mathbf{F}(s)$ by $[\mathbf{I}_m s - \mathbf{A}]$, we obtain

$$\mathbf{F}(s) = \mathbf{Q}_p(s)\left(\mathbf{I}_m s - \mathbf{A}\right) + \mathbf{R}_p ,$$

and pre-dividing by the same matrix, we obtain

$$\mathbf{F}(s) = \left[\mathbf{I}_m s - \mathbf{A}\right]\mathbf{Q}_l(s) + \mathbf{R}_l .$$

Substituting the matrix $\mathbf{A}$ in place of the scalar s in the above relationships, we obtain

$$\mathbf{F}_p(\mathbf{A}) = \mathbf{Q}_p(\mathbf{A})(\mathbf{A} - \mathbf{A}) + \mathbf{R}_p = \mathbf{R}_p$$

and

$$\mathbf{F}_l(\mathbf{A}) = (\mathbf{A} - \mathbf{A})\mathbf{Q}_l(\mathbf{A}) + \mathbf{R}_l = \mathbf{R}_l .$$

■

The following important corollary ensues from Theorem 1.4.1.

Corollary 1.4.1. A polynomial matrix $\mathbf{F}(s)$ is post-divisible (pre-divisible) without remainder by $[\mathbf{I}_m s - \mathbf{A}]$ if and only if $\mathbf{F}_p(\mathbf{A}) = 0$ ($\mathbf{F}_l(\mathbf{A}) = 0$).

Let $\varphi(s)$ be the characteristic polynomial of a square matrix $\mathbf{A}$ of degree n, i.e.,

$$\varphi(s) = \det\left[\mathbf{I}_n s - \mathbf{A}\right] = s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0 .$$

From the definition of the inverse matrix we have

$$\left[\mathbf{I}_n s - \mathbf{A}\right]\operatorname{Adj}\left[\mathbf{I}_n s - \mathbf{A}\right] = \mathbf{I}_n\varphi(s) \quad (1.4.3a)$$

and

$$\operatorname{Adj}\left[\mathbf{I}_n s - \mathbf{A}\right]\left[\mathbf{I}_n s - \mathbf{A}\right] = \mathbf{I}_n\varphi(s) . \quad (1.4.3b)$$

It follows from (1.4.3) that a polynomial matrix $\mathbf{I}_n\varphi(s)$ is post-divisible and pre-divisible by $[\mathbf{I}_n s - \mathbf{A}]$. According to Corollary 1.4.1 this is possible if and only if $\mathbf{I}_n\varphi(\mathbf{A}) = \varphi(\mathbf{A}) = 0$. Thus the following theorem has been proved.

Theorem 1.4.2. (Cayley–Hamilton). Every square matrix $\mathbf{A}$ satisfies its own characteristic equation

$$\varphi(\mathbf{A}) = \mathbf{A}^n + a_{n-1}\mathbf{A}^{n-1} + \cdots + a_1\mathbf{A} + a_0\mathbf{I}_n = 0 . \tag{1.4.4}$$

Example 1.4.1.
The characteristic polynomial of the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \tag{1.4.5}$$

is

$$\varphi(s) = \det\left[\mathbf{I}_n s - \mathbf{A}\right] = \begin{bmatrix} s-1 & -2 \\ -3 & s-4 \end{bmatrix} = s^2 - 5s - 2 .$$

It is easy to verify that

$$\varphi(\mathbf{A}) = \mathbf{A}^2 - 5\mathbf{A} - 2\mathbf{I}_2 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^2 - 5\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} - 2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Theorem 1.4.3. Let a polynomial $w(s)\in\mathbb{C}[s]$ be of degree N, and $\mathbf{A}\in\mathbb{C}^{n\times n}$, where $N \geq n$. There exists a polynomial $r(s)$ of a degree less than n, such that

$$w(\mathbf{A}) = r(\mathbf{A}) . \tag{1.4.6}$$

Proof. Dividing the polynomial $w(s)$ by the characteristic polynomial $\varphi(s)$ of the matrix $\mathbf{A}$, we obtain

$$w(s) = q(s)\varphi(s) + r(s) ,$$

where $q(s)$ and $r(s)$ are the quotient and remainder on division of the polynomial $w(s)$ by $\varphi(s)$, respectively, and $\deg \varphi(s) = n > \deg r(s)$. With the matrix $\mathbf{A}$ substituted in place of the scalar s and with (1.4.4) taken into account, we obtain

$$w(\mathbf{A}) = q(\mathbf{A})\varphi(\mathbf{A}) + r(\mathbf{A}) = r(\mathbf{A}) .$$

■

Example 1.4.2.
The following polynomial is given

$$w(s) = s^6 - 5s^5 - 3s^4 + 5s^3 + 2s^2 + 3s + 2 .$$

Using (1.4.6) one has to compute $w(\mathbf{A})$ for the matrix (1.4.5). The characteristic polynomial of the matrix is $\varphi(s) = s^2 - 5s - 2$. Dividing the polynomial $w(s)$ by $\varphi(s)$, we obtain

$$w(s) = \left(s^4 - s^2\right)\left(s^2 - 5s - 2\right) + 3s + 2,$$

that is

$$r(s) = 3s + 2 .$$

Hence

$$w(\mathbf{A}) = r(\mathbf{A}) = 3\mathbf{A} + 2\mathbf{I}_2 = 3\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + 2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 6 \\ 9 & 14 \end{bmatrix}.$$

The above considerations can be generalized to the case of square polynomial matrices.

Theorem 1.4.4. Let $\mathbf{W}(s) \in \mathbb{C}^{n \times n}[s]$ be a polynomial square matrix of degree N, and $\mathbf{A} \in \mathbb{C}^{n \times n}$, where $N \geq n$. There exists, a polynomial matrix $\mathbf{R}(s)$ of a degree less than n such that

$$\mathbf{W}_p(\mathbf{A}) = \mathbf{R}_p(\mathbf{A}) \quad \text{and} \quad \mathbf{W}_l(\mathbf{A}) = \mathbf{R}_l(\mathbf{A}), \tag{1.4.7}$$

where $\mathbf{W}_p(\mathbf{A})$ and $\mathbf{W}_l(\mathbf{A})$ are the right-side and left-side values, respectively, of the matrix $\mathbf{W}(s)$ with $\mathbf{A}$ substituted in place of s.

Proof. Dividing the entries of the matrix $\mathbf{W}(s)$ by the characteristic polynomial $\varphi(s)$ of $\mathbf{A}$, we obtain

$$\mathbf{W}(s) = \mathbf{Q}(s)\varphi(s) + \mathbf{R}(s),$$

where $\mathbf{Q}(s)$ and $\mathbf{R}(s)$ are the quotient and remainder, respectively, of the division of $\mathbf{W}(s)$ by $\varphi(s)$, and $\deg \varphi(s) = n > \deg \mathbf{R}(s)$. With $\mathbf{A}$ substituted in place of the scalar s and with (1.4.4) taken into account, we obtain

$$\mathbf{W}_p(\mathbf{A}) = \mathbf{Q}_p(\mathbf{A})\varphi(\mathbf{A}) + \mathbf{R}_p(\mathbf{A}) = \mathbf{R}_p(\mathbf{A})$$

and

$$\mathbf{W}_l(\mathbf{A}) = \mathbf{Q}_l(\mathbf{A})\varphi(\mathbf{A}) + \mathbf{R}_l(\mathbf{A}) = \mathbf{R}_l(\mathbf{A}) .$$

■

Example 1.4.3.
Given the polynomial matrix

$$\mathbf{W}(s) = \begin{bmatrix} s^6 - 5s^5 - 2s^4 + s^2 - 3s + 1 & s^5 - 5s^4 - 2s^3 - s - 1 \\ s^4 - 5s^3 - 3s^2 + 5s + 3 & -2s^6 + 10s^5 + 4s^4 - s + 2 \end{bmatrix},$$

one has to compute $\mathbf{W}_p(\mathbf{A})$ and $\mathbf{W}_l(\mathbf{A})$ for the matrix (1.4.5) using (1.4.7).

Dividing every entry of $\mathbf{W}(\mathbf{A})$ by the characteristic polynomial $\varphi(s)$ of matrix $\mathbf{A}$, we obtain

$$\mathbf{W}(s) = \begin{bmatrix} s^4 + 1 & s^3 \\ s^2 - 1 & -2s^4 \end{bmatrix} \left(s^2 - 5s - 2\right) + \begin{bmatrix} 2s + 3 & -s - 1 \\ 1 & -s + 2 \end{bmatrix},$$

i.e.,

$$\mathbf{R}(s) = \begin{bmatrix} 2s + 3 & -s - 1 \\ 1 & -s + 2 \end{bmatrix}.$$

Hence

$$\mathbf{W}_p(\mathbf{A}) = \mathbf{R}_p(\mathbf{A}) =$$
$$= \begin{bmatrix} 2 & -1 \\ 0 & -1 \end{bmatrix} \mathbf{A} + \begin{bmatrix} 3 & -1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 3 & -1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -2 & -2 \end{bmatrix}$$

and

$$\mathbf{W}_l(\mathbf{A}) = \mathbf{R}_l(\mathbf{A}) =$$
$$= \mathbf{A} \begin{bmatrix} 2 & -1 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 3 & -1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 3 & -1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 5 & -4 \\ 7 & 5 \end{bmatrix}.$$

1.5 Elementary Operations on Polynomial Matrices

Definition 1.5.1. The following operations are called elementary operations on a polynomial matrix $\mathbf{A}(s) \in \mathbb{C}^{m \times n}[s]$:

1. Multiplication of any i-th row (column) by the number $c \neq 0$.
2. Addition to any i-th row (column) of the j-th row (column) multiplied by any polynomial $w(s)$.
3. The interchange of any two rows (columns), e.g., of the i-th and the j-th rows (columns).

From now on we will use the following notation:

$L[i\times c]$	multiplication of the i-th row by the number $c \neq 0$,
$P[i\times c]$	multiplication of the i-th column by the number $c \neq 0$,
$L[i+j\times w(s)]$	addition to the i-th row of the j-th row multiplied by the polynomial $w(s)$,
$P[i+j\times w(s)]$	addition to the i-th column of the j-th column multiplied by the polynomial $w(s)$,
$L[i, j]$	the interchange of the i-th and the j-th row,
$P[i, j]$	the interchange of the i-th and the j-th column.

It is easy to verify that the above elementary operations when carried out on rows are equivalent to pre-multiplication of the matrix $\mathbf{A}(s)$ by the following matrices:

$$\mathbf{L}_m(i,c)=\begin{array}{c} \qquad\quad i\text{-th column} \\ \begin{bmatrix} 1 & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & c & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 1 \end{bmatrix} \end{array} \begin{array}{l} \\ \\ \\ \\ i\text{-th row} \\ \\ \\ \end{array} \in \mathbb{C}^{m\times m},$$

$$\mathbf{L}_d(i,j,w(s))=\begin{array}{c} \qquad\qquad i \qquad\qquad j \qquad\qquad \\ \begin{bmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & \dots & w(s) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix} \end{array} \in \mathbb{C}^{m\times m}[s],$$

$$\mathbf{L}_z(i,j)=\begin{array}{c} \qquad\quad i \qquad\quad j \qquad\qquad \\ \begin{bmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix} \end{array}. \tag{1.5.1}$$

The same operations carried out on columns are equivalent to post-multiplication of the matrix $\mathbf{A}(s)$ by the following matrices:

$$\mathbf{P}_m(i,c)=\begin{array}{c}\begin{matrix} & & & i\text{-th column} & & & \end{matrix}\\ \left[\begin{matrix} 1 & 0 & \dots & 0 & 0 & \dots & 0\\ 0 & 1 & \dots & 0 & 0 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & c & 0 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & 0 & 0 & \dots & 1 \end{matrix}\right]\end{array}\begin{matrix} \\ \\ \\ \\ i\text{-th row}\\ \\ \\ \end{matrix}\in\mathbb{C}^{n\times n},$$

$$\mathbf{P}_d(i,j,w(s))=\begin{array}{c}\begin{matrix} & & & i & & j & & \end{matrix}\\ \left[\begin{matrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0\\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & 1 & \dots & 0 & \dots & 0\\ 0 & 0 & \dots & w(s) & \dots & 1 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \end{matrix}\right]\end{array}\in\mathbb{C}^{n\times n}, \tag{1.5.2}$$

$$\mathbf{P}_z(i,j)=\begin{array}{c}\begin{matrix} & & & i & & j & & \end{matrix}\\ \left[\begin{matrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0\\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & 1 & \dots & 0 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & 0 & \dots & 0 & \cdots & 1 \end{matrix}\right]\end{array}\in\mathbb{C}^{n\times n}\ .$$

It is easy to verify that the determinants of the polynomial matrices (1.5.1) and (1.5.2) are nonzero and do not depend on the variable s. Such matrices are called unimodular matrices.

1.6 Linear Independence, Space Basis and Rank of Polynomial Matrices

Let $a_i = a_i(s)$, $i = 1,...,n$ be the i-th column of a polynomial matrix $\mathbf{A}(s)\in\mathbb{R}^{m\times n}[s]$. We will consider these columns as m-dimensional polynomial vectors, $a_i\in\mathbb{R}^m[s]$, $i = 1,...,n$.

Definition 1.6.1. Vectors $a_i\in\mathbb{R}^m[s]$ are called linearly independent over the field of rational functions $\mathbb{F}(s)$ if and only if there exist rational functions $w_i=w_i(s)\in\mathbb{F}(s)$ not all equal to zero such that

$$w_1a_1 + w_2a_2 + ... + w_na_n = 0 \quad \text{(zero order)}. \tag{1.6.1}$$

In other words, these vectors are called linearly independent over the field of rational functions, if the equality (1.6.1) implies $w_i = 0$ for $i = 1,...,n$.

For example, the polynomial vectors

$$a_1 = \begin{bmatrix} 1 \\ s \end{bmatrix}, \quad a_2 = \begin{bmatrix} s \\ 1+s^2 \end{bmatrix} \tag{1.6.2}$$

are linearly independent over the field of rational functions, since the equation

$$w_1a_1 + w_2a_2 = \begin{bmatrix} 1 \\ s \end{bmatrix} w_1 + \begin{bmatrix} s \\ 1+s^2 \end{bmatrix} w_2 = \begin{bmatrix} 1 & s \\ s & s^2+1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

has only the zero solution

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 1 & s \\ s & s^2+1 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

We will show that the rational functions w_i, $i = 1,...,n$ in (1.6.1) can be replaced by polynomials $p_i = p_i(s)$, $i = 1,...,n$. To accomplish this, we multiply both sides of (1.6.1) by the smallest common denominator of rational functions w_i, $i = 1,...,n$. We then obtain

$$p_1a_1 + p_2a_2 + ... + p_na_n = 0, \tag{1.6.3}$$

where $p_i = p_i(s)$ are polynomials.

For example, the polynomial vectors

$$a_1 = \begin{bmatrix} 1 \\ s \end{bmatrix}, \quad a_2 = \begin{bmatrix} s+1 \\ s^2+s \end{bmatrix} \tag{1.6.4}$$

are linearly dependent over the field of rational functions, since for

$$w_1 = -1 \quad \text{and} \quad w_2 = \frac{1}{s+1},$$

we obtain

$$w_1 a_1 + w_2 a_2 = -\begin{bmatrix} 1 \\ s \end{bmatrix} + \frac{1}{s+1}\begin{bmatrix} s+1 \\ s^2+s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{1.6.5}$$

Multiplying both sides of (1.6.5) by the smallest common denominator of rational functions w_1 and w_2, which is equal to $s + 1$, we obtain

$$-(s+1)\begin{bmatrix} 1 \\ s \end{bmatrix} + \begin{bmatrix} s+1 \\ s^2+s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

If the number of polynomial vectors of the space $\mathbb{R}^n[s]$ is larger than n, then these vectors are linearly dependent. For example, adding to two linearly independent vectors (1.6.2) an arbitrary vector

$$a = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} \in \mathbb{R}^2[s],$$

we obtain linearly dependent vectors, i.e.,

$$p_1 a_1 + p_2 a_2 + p_3 a = 0, \tag{1.6.6}$$

for $p_1, p_2, p_3 \in \mathbb{R}[s]$ not simultaneously equal to zero.

Assuming, for example, $p_3 = -1$, from (1.6.6) and (1.6.2), we obtain

$$\begin{bmatrix} 1 & s \\ s & s^2+1 \end{bmatrix}\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}$$

and

$$\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} 1 & s \\ s & s^2+1 \end{bmatrix}^{-1}\begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} = \begin{bmatrix} s^2+1 & -s \\ -s & 1 \end{bmatrix}\begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} = \begin{bmatrix} \left(s^2+1\right)a_{11} - s a_{21} \\ -s a_{11} + a_{21} \end{bmatrix}.$$

Thus vectors a_1, a_2, a are linearly dependent for any vector a.

Definition 1.6.2. Polynomial vectors $b_i = b_i(s) \in \mathbb{R}^n[s]$, $i = 1,\ldots,n$ are called a basis of space $\mathbb{R}^n[s]$ if they are linearly independent over the field of rational function and an arbitrary vector $a \in \mathbb{R}^n[s]$ from this space can be represented as a linear combination of these vectors, i.e.,

$$a = p_1 b_1 + p_2 b_2 + \ldots + p_n b_n , \tag{1.6.7}$$

where $p_i \in \mathbb{R}[s]$, $i = 1,\ldots,n$.

There exist many different bases for the same space. For example, for the space $\mathbb{R}^2[s]$ we can adopt the vectors (1.6.2) as a basis. Solving system of equations for an arbitrary vector

$$\begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} \in \mathbb{R}^2[s],$$

$$\begin{bmatrix} a_1 & a_2 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} 1 & s \\ s & s^2+1 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix},$$

we obtain

$$\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} 1 & s \\ s & s^2+1 \end{bmatrix}^{-1} \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} = \begin{bmatrix} \left(s^2+1\right) a_{11} - s a_{21} \\ -s a_{11} + a_{21} \end{bmatrix}.$$

As a basis for this space we can also adopt

$$e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

In this case, $p_1 = a_{11}$ and $p_2 = a_{21}$.

Definition 1.6.3. The number of linearly independent rows (columns) of a polynomial matrix $\mathbf{A}(s) \in \mathbb{R}^{n \times m}[s]$ is called its normal rank (briefly rank).

The rank of a polynomial matrix $\mathbf{A}(s)$ can be also equivalently defined as the highest order of a minor, which is a nonzero polynomial, of this matrix.

The rank of matrix $\mathbf{A}(s) \in \mathbb{R}^{n \times m}[s]$ is not greater than the number of its rows n or columns m, i.e.,

$$\operatorname{rank} \mathbf{A}(s) \leq \min(n, m). \tag{1.6.8}$$

If a square matrix $\mathbf{A}(s) \in \mathbb{R}^{n \times n}[s]$ is of full rank, i.e., $\operatorname{rank} \mathbf{A}(s) = n$, then its determinant is a nonzero polynomial $w(s)$, i.e.,

$$\det \mathbf{A}(s) = w(s) \neq 0 . \tag{1.6.9}$$

Such a matrix is called nonsingular or invertible. It is called singular when $\det \mathbf{A}(s) = 0$ (the zero polynomial). For example, the square matrix built from linearly independent vectors (1.6.2) is nonsingular, since

$$\det \begin{bmatrix} 1 & s \\ s & 1+s^2 \end{bmatrix} = 1$$

and the matrix built from linearly dependent vectors (1.6.4) is singular, since

$$\det \begin{bmatrix} 1 & s+1 \\ s & s^2+s \end{bmatrix} = 0 .$$

Theorem 1.6.1. Elementary operations carried out on a polynomial matrix do not change its rank.

Proof. Let

$$\overline{\mathbf{A}}(s) = \mathbf{L}(s)\mathbf{A}(s)\mathbf{P}(s) \in \mathbb{R}^{n\times m}[s], \tag{1.6.10}$$

where $\mathbf{L}(s)\in\mathbb{R}^{n\times n}[s]$ and $\mathbf{P}(s)\in\mathbb{R}^{m\times m}[s]$ are unimodular matrices of elementary operations on rows and columns, respectively.

From (1.6.10) we immediately have

$$\operatorname{rank} \overline{\mathbf{A}}(s) = \operatorname{rank} \left[\mathbf{L}(s)\mathbf{A}(s)\mathbf{P}(s)\right] = \operatorname{rank} \mathbf{A}(s),$$

since $\mathbf{L}(s)$ and $\mathbf{P}(s)$ are unimodular matrices. ■

For example, carrying out the operation $L_d(2+1\times(-s))$ on rows of the matrix built from the columns (1.6.2), we obtain

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

Both polynomial matrices

$$\begin{bmatrix} 1 & s \\ s & s^2+1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix}$$

are full rank matrices.

1.7. Equivalents of Polynomial Matrices

1.7.1 Left and Right Equivalent Matrices

Definition 1.7.1. Two polynomial matrices $\mathbf{A}(\mathrm{s})$, $\mathbf{B}(s)\in\mathbb{C}^{m\times n}[s]$ are called left (right) or row (column) equivalent if and only if one of them can be obtained from the other as a result of a finite number of elementary operations carried out on its rows (columns)

$$\mathbf{B}(s)=\mathbf{L}(s)\mathbf{A}(s) \qquad \left(\text{or } \mathbf{B}(s)=\mathbf{A}(s)\mathbf{P}(s)\right), \tag{1.7.1}$$

where $\mathbf{L}(s)$ ($\mathbf{P}(s)$) is the product of unimodular matrices of elementary operations on rows (columns).

Definition 1.7.2. Two polynomial matrices $\mathbf{A}(\mathrm{s})$, $\mathbf{B}(s)\in\mathbb{C}^{m\times n}[s]$ are called equivalent if and only if one of them can be obtained from the other as a result of a finite number of elementary operations carried out on its rows and columns, i.e.,

$$\mathbf{B}(s)=\mathbf{L}(s)\mathbf{A}(s)\mathbf{P}(s)\,, \tag{1.7.2}$$

where $\mathbf{L}(s)$ and $\mathbf{P}(s)$ are the products of unimodular matrices of elementary operations on rows and columns, respectively.

Theorem 1.7.1. A full rank polynomial matrix $\mathbf{A}(s)\in\mathbb{C}^{m\times l}[s]$ is left equivalent to an upper triangular matrix of the form

$$\overline{\mathbf{A}}(s)=\mathbf{L}(s)\mathbf{A}(s)=\begin{cases}\begin{bmatrix}\overline{a}_{11}(s) & \overline{a}_{12}(s) & \dots & \overline{a}_{1l}(s)\\ 0 & \overline{a}_{22}(s) & \dots & \overline{a}_{2l}(s)\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & \overline{a}_{1l}(s)\\ 0 & 0 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & 0\end{bmatrix} & \text{for} \quad m>l\\[2ex] \begin{bmatrix}\overline{a}_{11}(s) & \overline{a}_{12}(s) & \dots & \overline{a}_{1m}(s)\\ 0 & \overline{a}_{22}(s) & \dots & \overline{a}_{2m}(s)\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & \overline{a}_{mm}(s)\end{bmatrix} & \text{for} \quad m=l\end{cases} \tag{1.7.3}$$

$$\bar{\mathbf{A}}(s)=\mathbf{L}(s)\mathbf{A}(s)=\left\{\begin{bmatrix}\bar{a}_{11}(s) & \bar{a}_{12}(s) & \dots & \bar{a}_{1m}(s) & \dots & \bar{a}_{1l}(s)\\ 0 & \bar{a}_{22}(s) & \dots & \bar{a}_{2m}(s) & \dots & a_{2l}(s)\\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & \bar{a}_{mm}(s) & \dots & a_{ml}(s)\end{bmatrix}\right. \quad \text{for } m<l$$

where the elements $\bar{a}_{1i}(s)$, $\bar{a}_{2i}(s),\dots,$ $\bar{a}_{i-1,i}(s)$ are polynomials of a degree less than $\bar{a}_{ii}(s)$ for $i = 1,2,\dots,m$, and $\mathbf{L}(s)$ is the product of the matrices of elementary operations carried out on rows.

Proof. Among nonzero entries of the first columns of the matrix $\mathbf{A}(s)$ we choose the entry that is a polynomial of the lowest degree and carrying out $L[i, j]$, we move this entry to the position (1,1). Denote this entry by $\tilde{a}_{11}(s)$. Then we divide all remaining entries of the first column by $\tilde{a}_{11}(s)$. We then obtain

$$\tilde{a}_{i1}(s)=\tilde{a}_{11}(s)q_{i1}(s)+r_{i1}(s) \quad \text{for} \quad i=2,3,\dots,m\,,$$

where $q_{i1}(s)$ is the quotient and $r_{i1}(s)$ the remainder of division of the polynomial $\tilde{a}_{i1}(s)$ by $\tilde{a}_{11}(s)$. Carrying out $L[i+1\times(-q_{i1}(s))]$, we replace the entry $\tilde{a}_{i1}(s)$ with the remainder $r_{i1}(s)$. If not all remainders are equal to zero, then we choose this one, that is the polynomial of the lowest degree, and carrying out operations $L[i, j]$, we move it to position (1,1). Denoting this remainder by $\tilde{r}_{i1}(s)$, we repeat the above procedure taking the remainder $\tilde{r}_{11}(s)$ instead of $\tilde{a}_{11}(s)$. The degree $\tilde{r}_{11}(s)$ is lower than the degree of $\tilde{a}_{11}(s)$. After a finite number of steps, we obtain the matrix $\tilde{\mathbf{A}}(s)$ of the form

$$\tilde{\mathbf{A}}(s)=\begin{bmatrix}\bar{a}_{11}(s) & \tilde{a}_{12}(s) & \dots & \tilde{a}_{1l}(s)\\ 0 & \tilde{a}_{22}(s) & \dots & a_{2l}(s)\\ \vdots & \vdots & \ddots & \vdots\\ 0 & \tilde{a}_{m2}(s) & \dots & \tilde{a}_{ml}(s)\end{bmatrix}.$$

We repeat the above procedure for the first column of the submatrix obtained from the matrix $\mathbf{A}(s)$ by deleting the first row and the first column. We then obtain a matrix of the form

$$\hat{\mathbf{A}}(s)=\begin{bmatrix}\bar{a}_{11}(s) & \tilde{a}_{12}(s) & \tilde{a}_{13}(s) & \dots & \tilde{a}_{1l}(s)\\ 0 & \bar{a}_{22}(s) & \hat{a}_{23}(s) & \dots & \hat{a}_{2l}(s)\\ 0 & 0 & \hat{a}_{33}(s) & \dots & \hat{a}_{3l}(s)\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \hat{a}_{m3}(s) & \dots & \hat{a}_{ml}(s)\end{bmatrix}.$$

If $\tilde{a}_{12}(s)$ is not a polynomial of lower degree than the one of $\bar{a}_{22}(s)$, then we divide $\tilde{a}_{12}(s)$ by $\bar{a}_{22}(s)$ and carrying out $L[1+2\times(-q_{12}(s))]$, we replace the entry

$\tilde{a}_{12}(s)$ with the entry $\overline{a}_{12}(s) = r_{12}(s)$, where $q_{12}(s)$ and $r_{12}(s)$ are the quotient and the remainder on the division of $\tilde{a}_{12}(s)$ by $\overline{a}_{22}(s)$ respectively.

Next, we consider the submatrix obtained from the matrix $\tilde{\mathbf{A}}(s)$ by removing the first two rows and the first two columns. Continuing this procedure, we obtain the matrix (1.7.3). ■

An algorithm of determining the left equivalent matrix of the form (1.7.3) follows immediately from the above proof.

Example 1.7.1.
The given matrix

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

is to be transformed to the left equivalent form (1.7.3).

To accomplish this, we carry out the following elementary operations:

$$\xrightarrow[L\left[3+1\times(-s^2)\right]]{L\left[2+1\times(-(s+1))\right]} \begin{bmatrix} 1 & -s & 2 \\ 0 & s^2+2 & -2s-1 \\ 0 & 1 & 0 \end{bmatrix} \xrightarrow{L[2,3]} \begin{bmatrix} 1 & -s & 2 \\ 0 & 1 & 0 \\ 0 & s^2+2 & -2s-1 \end{bmatrix} \longrightarrow$$

$$\xrightarrow[L\left[3+2\times(-(s^2+2))\right]]{L[1+2\times s]} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & -2s-1 \end{bmatrix}.$$

Theorem 1.7.2. A full rank polynomial matrix $\mathbf{A}(s)\in\mathbb{C}^{m\times l}[s]$ is right equivalent to a lower triangular matrix of the form

$$\overline{\mathbf{A}}(s) = \mathbf{A}(s)\mathbf{P}(s) = \\ = \begin{cases} \begin{bmatrix} \overline{a}_{11}(s) & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \overline{a}_{21}(s) & \overline{a}_{22}(s) & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \overline{a}_{m1}(s) & \overline{a}_{m2}(s) & \overline{a}_{m3}(s) & \dots & \overline{a}_{mm}(s) & 0 & \dots & 0 \end{bmatrix} & \text{for } n > m, \\ \begin{bmatrix} \overline{a}_{11}(s) & 0 & \dots & 0 \\ \overline{a}_{21}(s) & \overline{a}_{22}(s) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \overline{a}_{m1}(s) & \overline{a}_{m2}(s) & \dots & \overline{a}_{mm}(s) \end{bmatrix} & \text{for } n = m, \end{cases} \tag{1.7.4}$$

$$= \begin{cases} \begin{bmatrix} \bar{a}_{11}(s) & 0 & \dots & 0 \\ \bar{a}_{21}(s) & \bar{a}_{22}(s) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \bar{a}_{l1}(s) & \bar{a}_{l2}(s) & \dots & \bar{a}_{l2}(s) \\ \vdots & \vdots & \ddots & \vdots \\ \bar{a}_{m1}(s) & \bar{a}_{m2}(s) & \dots & \bar{a}_{ml}(s) \end{bmatrix} & \text{for } n < m, \end{cases} \tag{1.7.4}$$

where the elements $\bar{a}_{i1}(s)$, $\bar{a}_{i2}(s), \dots, \bar{a}_{i-1,i}(s)$ are polynomials of lower degree than that of $\bar{a}_{ii}(s)$ for $i = 1,2,\dots,n$, and $\mathbf{P}(s)$ is the product of unimodular matrices of elementary operations carried out on columns.

1.7.2 Row and Column Reduced Matrices

The degree of the i-th column (row) of a polynomial matrix is the highest degree of a polynomial that is an entry of this column (row).

The degree of the i-th column (row) of the matrix $\mathbf{A}(s)$ will be denotedn by $\deg c_i[\mathbf{A}(s)]$ ($\deg r_i[\mathbf{A}(s)]$) or shortly $\deg c_i$ ($\deg r_i$).

Let $\mathbf{L}_c$ ($\mathbf{L}_r$) be the matrix built from the coefficients at the highest powers of variable s in the columns (rows) of the matrix $\mathbf{A}(s)$. For example, for the polynomial matrix

$$\mathbf{A}(s) = \begin{bmatrix} s^2-1 & s & -3s \\ s+2 & -s & 2 \\ s^2 & s-1 & 2s-1 \end{bmatrix}, \tag{1.7.5}$$

we have $\deg \mathbf{A}(s) = 2$

$$\deg c_1 = 2, \quad \deg c_2 = \deg c_3 = 1, \quad \deg r_1 = \deg r_3 = 2, \quad \deg\ r_2 = 1$$

and

$$\mathbf{L}_k = \begin{bmatrix} 1 & 1 & -3 \\ 0 & -1 & 0 \\ 1 & 1 & 2 \end{bmatrix}, \quad \mathbf{L}_w = \begin{bmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

The matrix (1.7.5) can be written, using the above matrices, as follows

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

or

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

In the general case for a matrix $\mathbf{A}(s)\in\mathbb{R}^{m\times n}[s]$, we have

$$\mathbf{A}(s) = \mathbf{L}_c \mathrm{diag}\left[s^{\deg \mathrm{c}_1}, s^{\deg \mathrm{c}_2}, \ldots, s^{\deg \mathrm{c}_l}\right] + \overline{\mathbf{A}}(s) \tag{1.7.6}$$

and

$$\mathbf{A}(s) = \mathrm{diag}\left[s^{\deg \mathrm{r}_1}, s^{\deg \mathrm{r}_2}, \ldots, s^{\deg \mathrm{r}_m}\right] \mathbf{L}_r + \tilde{\mathbf{A}}(s) , \tag{1.7.7}$$

where $\overline{\mathbf{A}}(s)$, $\tilde{\mathbf{A}}(s)$ are polynomial matrices satisfying the conditions

$$\deg \overline{\mathbf{A}}(s) < \deg \mathbf{A}(s), \ \deg \tilde{\mathbf{A}}(s) < \deg \mathbf{A}(s) .$$

If $m = n$ and $\det \mathbf{L}_c \neq 0$, then the determinant of the matrix (1.7.6) is a polynomial of the degree

$$n_k = \sum_{i=1}^{l} \deg c_i ,$$

since

$$\det \mathbf{A}(s) = \det \mathbf{L}_k \det \mathrm{diag}\left[s^{\deg c_1}, s^{\deg c_2}, \ldots, s^{\deg c_l}\right] + \ldots = s^{n_k} \det \mathbf{L}_c + \ldots$$

Similarly, if $\det \mathbf{L}_r \neq 0$, then the determinant of the matrix (1.7.7) is a polynomial of the degree

$$n_r = \sum_{j=1}^{m} \deg r_j .$$

Definition 1.7.3. A polynomial matrix $\mathbf{A}(s)$ is said to be column (row) reduced if and only if $\mathbf{L}_c$ ($\mathbf{L}_r$) of this matrix is a full rank matrix.

Thus, a square matrix $\mathbf{A}(s)$ is column (row) reduced if and only if $\det \mathbf{L}_c \neq 0$ ($\det \mathbf{L}_r \neq 0$).

For example, the matrix (1.7.5) is column reduced but not row reduced, since

$$\det \mathbf{L}_c = \begin{vmatrix} 1 & 1 & -3 \\ 0 & -1 & 0 \\ 1 & 1 & 2 \end{vmatrix} = -5, \quad \det \mathbf{L}_r = \begin{vmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ 1 & 0 & 0 \end{vmatrix} = 0 .$$

From the above considerations and Theorems 1.7.1 and 1.7.1′ the following important corollary immediately follows.

Corollary 1.7.1. Carrying out only elementary operations on rows or columns it is possible to transform a nonsingular polynomial matrix to one of column reduced form and row reduced form, respectively.

1.8 Reduction of Polynomial Matrices to the Smith Canonical Form

Consider a polynomial matrix $\mathbf{A}(s) \in \mathbb{C}^{m \times n}[s]$ of rank r.

Definition 1.8.1. A polynomial matrix of the form

$$\mathbf{A}_S(s) = \begin{bmatrix} i_1(s) & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & i_2(s) & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & i_r(s) & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{bmatrix} \in \mathbb{C}^{m \times n}[s] \cdot \tag{1.8.1}$$

$r \le \min(n,m)$ is called the Smith canonical form of the matrix $\mathbf{A}(s) \in \mathbb{C}^{m \times n}[s]$, where $i_1(s), i_2(s), \dots, i_r(s)$ are nonzero polynomials that are called invariant, with coefficients by the highest powers of the variable s equal to one, such that the polynomial $i_{k+1}(s)$ is divisible without remainder by the polynomial $i_k(s)$, i.e., $i_{k+1} \mid i_k$ for $k = 1,\dots,r$-1.

Theorem 1.8.1. For an arbitrary polynomial matrix $\mathbf{A}(s) \in \mathbb{C}^{m \times n}[s]$ of rank r ($r \le \min(n,m)$) there exists its equivalent Smith canonical form (1.8.1).

Proof. Among the entries of the matrix $\mathbf{A}(s)$ we find a nonzero one, which is a polynomial of the lowest degree in respect to s, and interchanging rows and columns we move it to position (1,1). Denote this entry by $\overline{a}_{11}(s)$. Assume at the beginning that all entries of the matrix $\mathbf{A}(s)$ are divisible without remainder by the entry $\overline{a}_{11}(s)$. Dividing the entries $\overline{a}_{i1}(s)$ of the first column and the first row $\overline{a}_{1j}(s)$ by $\overline{a}_{11}(s)$, we obtain

$$\bar{a}_{i1}(s) = \bar{a}_{11}(s)q_{i1}(s) \qquad (i = 2,3,...,m),$$
$$\bar{a}_{1j}(s) = \bar{a}_{11}(s)q_{1j}(s) \qquad (j = 2,3,...,n),$$

where $q_{i1}(s)$ and $q_{1j}(s)$ are the quotients from division of $\bar{a}_{i1}(s)$ and $\bar{a}_{1j}(s)$ by $\bar{a}_{11}(s)$, respectively.

Subtracting from the i-th row ($i = 2,3,...,m$) the first row multiplied by $q_{i1}(s)$ and, respectively from the j-th column ($j = 2,3,...,m$) the first column multiplied by $q_{1j}(s)$, we obtain a matrix of the form

$$\begin{bmatrix} \bar{a}_{11}(s) & 0 & \dots & 0 \\ 0 & \bar{a}_{22}(s) & \dots & \bar{a}_{2n}(s) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \bar{a}_{m2}(s) & \dots & \bar{a}_{mn}(s) \end{bmatrix}. \tag{1.8.2}$$

If the coefficient by the highest power of s of polynomial $\bar{a}_{11}(s)$ is not equal to 1, then to accomplish this we multiply the first row (or column) by the reciprocal of this coefficient.

Assume next that not all entries of the matrix $\mathbf{A}(s)$ are divisible without remainder by $\bar{a}_{11}(s)$ and that such entries are placed in the first row and the first column. Dividing the entries of the first row and the first column by $\bar{a}_{11}(s)$, we obtain

$$\bar{a}_{1i}(s) = \bar{a}_{11}(s)q_{1i}(s) + r_{1i}(s) \qquad (i = 2,3,...,n),$$
$$\bar{a}_{j1}(s) = \bar{a}_{11}(s)q_{j1}(s) + r_{j1}(s) \qquad (j = 2,3,...,m),$$

where $q_{1i}(s)$, $q_{j1}(s)$ are the quotients and $r_{1i}(s)$, $r_{j1}(s)$ are the remainders of division of $\bar{a}_{1i}(s)$ and $\bar{a}_{j1}(s)$ by $\bar{a}_{11}(s)$, respectively. Subtracting from the j-th row (i-th column) the first row (column) multiplied by $q_{j1}(s)$ (by $\bar{q}_{1i}(s)$), we replace the entry $\bar{a}_{j1}(s)$ ($\bar{a}_{1i}(s)$) by the remainder $r_{j1}(s)$ ($r_{1i}(s)$). Next, among these remainders we find a polynomial of the lowest degree with respect to s and interchanging rows and columns, we move it to the position (1,1). We denote this polynomial by $\bar{r}_{11}(s)$. If not all entries of the first row and the first column are divisible without remainder by $\bar{r}_{11}(s)$, then we repeat this procedure taking the polynomial $\bar{r}_{11}(s)$ instead of the polynomial $\bar{a}_{11}(s)$. The degree of the polynomial $\bar{r}_{11}(s)$ is lower than the degree of $\bar{a}_{11}(s)$. After a finite number of steps, we obtain in the position (1,1) a polynomial that divides without remainder all the entries of the first row and the first column. If the entry $\bar{a}_{ik}(s)$ is not divisible by $\bar{a}_{11}(s)$, then by adding the i-th row (or k-th column) to the first row (the first column), we reduce this case to the previous one.

Repeating this procedure, we finally obtain in the position (1,1) a polynomial that divides without remainder all the entries of the matrix. Further we proceed in the same way as in the first case, when all the entries of the matrix are divisible without remainder by $\bar{a}_{11}(s)$.

If not all entries $\bar{a}_{ij}(s)$ ($i = 2,3,...,m;\ j = 2,3,...,n$) of the matrix (1.8.2) are equal to zero, then we find a nonzero entry among them, which is a polynomial of the lowest degree with respect to s, and interchanging rows and columns, we move it to the position (2,2). Proceeding further as above, we obtain a matrix of the form

$$\begin{bmatrix} \bar{a}_{11}(s) & 0 & 0 & \cdots & 0 \\ 0 & \bar{a}_{22}(s) & 0 & \cdots & 0 \\ 0 & 0 & \bar{a}_{33}(s) & \cdots & \bar{a}_{3n}(s) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \bar{a}_{m3}(s) & \cdots & \bar{a}_{mn}(s) \end{bmatrix},$$

where $\bar{a}_{22}(s)$ is divisible without remainder by $\bar{a}_{11}(s)$, and all elements $\bar{a}_{ij}(s)$ ($i = 3,4,...,m;\ j = 3,4,...,n$) are divisible without remainder by $\bar{a}_{22}(s)$. Continuing this procedure, we obtain a matrix of the Smith canonical form (1.8.1).

■

From this proof the following algorithm for determining of the Smith canonical form follows immediately as, illustrated by the following example.

Example 1.8.1.
To transform the polynomial matrix

$$\mathbf{A}(s) = \begin{bmatrix} (s+2)^2 & (s+2)(s+3) & s+2 \\ (s+2)(s+3) & (s+2)^2 & s+3 \end{bmatrix} \tag{1.8.3}$$

to the Smith canonical form, we carry out the following elementary operations.
Step 1: We carry out the operation $P[1, 3]$

$$\mathbf{A}_1(s) = \begin{bmatrix} s+2 & (s+2)(s+3) & (s+2)^2 \\ s+3 & (s+2)^2 & (s+2)(s+3) \end{bmatrix}.$$

All entries of this matrix are divisible without remainder by $s + 2$ with exception of the entry $s + 3$.

Step 2: Taking into account the equality

$$\frac{s+3}{s+2} = 1 + \frac{1}{s+2},$$

we carry out the operation $L[2+1\times(-1)]$

$$\mathbf{A}_2(s) = \begin{bmatrix} s+2 & (s+2)(s+3) & (s+2)^2 \\ 1 & -(s+2) & s+2 \end{bmatrix}.$$

Step 3: We carry out the operation $L[1, 2]$

$$\mathbf{A}_3(s) = \begin{bmatrix} 1 & -(s+2) & s+2 \\ s+2 & (s+2)(s+3) & (s+2)^2 \end{bmatrix}.$$

Step 4: We carry out the operations $P[2+1\times(s+2)]$ and $P[3+1\times(-s-2)]$

$$\mathbf{A}_4(s) = \begin{bmatrix} 1 & 0 & 0 \\ s+2 & (s+2)(2s+5) & 0 \end{bmatrix}.$$

Step 5: We carry out the operation $L[2+1\times(-s-2)]$ and $P[2\times 1/2]$

$$\mathbf{A}_s(s) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & (s+2)\left(s+\frac{5}{2}\right) & 0 \end{bmatrix}.$$

This matrix is of the desired Smith canonical form of (1.8.3).

From divisibility of the invariant polynomials $i_{k+1} \mid i_k$, $k = 1, ..., r-1$, it follows that there exist polynomials $d_1, d_2, ..., d_r$, such that

$$i_1 = d_1,\ i_2 = d_1 d_2,\ ...,\ i_r = d_1 d_2 ... d_r .$$

Hence the matrix (1.8.1) can be written in the form

$$\mathbf{A}_S(s) = \begin{bmatrix} d_1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & d_1 d_2 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_1 d_2 ... d_r & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{bmatrix}. \tag{1.8.1a}$$

Theorem 1.8.2. The invariant polynomials $i_1(s), i_2(s), ..., i_r(s)$ of the matrix (1.8.1) are uniquely determined by the relationship

$$i_k(s) = \frac{D_k(s)}{D_{k-1}(s)} \quad \text{for} \quad k = 1, 2, ..., r, \tag{1.8.4}$$

where $D_k(s)$ is the greatest common divisor of all minors of degree k of matrix $\mathbf{A}(s)$ ($D_0(s) = 1$).

Proof. We will show that elementary operations do not change $D_k(s)$. Note that elementary operations 1) consisting of multiplying of an i-th row (column) by a number $c \neq 0$ causes multiplication of minors containing this row (column) by this number c. Thus this operation does not change $D_k(s)$. An elementary operation 2) consisting of adding to an i-th row (column) j-th row (column) multiplied by the polynomial $w(s)$ does not change $D_k(s)$, if a minor of the degree k contains either the i-th row and the j-th row or does not contain of them. If the minor of the degree k contains the i-th row, and does not contain the j-th row, then we can represent it as a linear combination of two minors of the degree k of the matrix $\mathbf{A}(s)$. Hence the greatest common divisor of the minors of the degree k does not change. Finally, an operation 3), consisting on the interchange of i-th and j-th rows (columns), does not change $D_k(s)$ either, since as a result of this operation a minor of the degree k either does not change (the both rows (columns) do not belong to this minor), or changes only the sign (both rows belong to the same minor), or it will be replaced by another minor of the degree k of the matrix $\mathbf{A}(s)$ (only one of these rows belongs to this minor).

Thus equivalent matrices $\mathbf{A}(s)$ and $\mathbf{A}_S(s)$ have the same divisors $D_1(s)$, $D_2(s)$, ..., $D_r(s)$. From the Smith canonical form (1.8.1) it follows that

$$\begin{array}{l} D_1(s) = i_1(s), \\ D_2(s) = i_1(s)\, i_2(s), \\ \hline D_r(s) = i_1(s)\, i_2(s) ... i_r(s). \end{array} \tag{1.8.5}$$

From (5) we immediately obtain the formula (4). ■

Using the polynomials $d_1, d_2, ..., d_r$ we can write the relationship (1.8.5) in the form

$$\begin{array}{l} D_1(s) = d_1, \\ D_2(s) = d_1^2 d_2, \\ \hline D_r(s) = d_1^r d_2^{r-1} ... d_r \end{array} . \tag{1.8.6}$$

From definition (1.8.1) and Theorems 1.8.1 and 1.8.2, the following important corollary can be derived.

Corollary 1.8.1. Two matrices $\mathbf{A}(s)$, $\mathbf{B}(s) \in \mathbb{C}^{m \times n}[s]$ are equivalent if and only if they have the same invariant polynomials.

1.9 Elementary Divisors and Zeros of Polynomial Matrices

1.9.1 Elementary Divisors

Consider a polynomial matrix $\mathbf{A}(s)\in\mathbb{C}^{m\times n}[s]$ of the rank r, whose Smith canonical form $\mathbf{A}_S(s)$ is given by the formula (1.8.1).

Let the k-th invariant polynomial of this matrix be of the form

$$i_k(s) = (s-s_1)^{m_{k1}}(s-s_2)^{m_{k2}}\dots(s-s_q)^{m_{kq}} \ . \tag{1.9.1}$$

From divisibility of the polynomial $i_{k+1}(s)$ by the polynomial $i_k(s)$ it follows that

$$\begin{array}{l} m_{r,1} \geq m_{r-1,1} \geq \dots \geq m_{1,1} \geq 0 \\ \hdashline m_{r,q} \geq m_{r-1,q} \geq \dots \geq m_{1,q} \geq 0 \end{array}. \tag{1.9.2}$$

If, for example, $i_1(s) = 1$, then $m_{11} = m_{12} = \dots = m_{1q} = 0$.

Definition 1.9.1. Everyone of the expressions (different from 1)

$$(s-s_1)^{m_{11}},\ (s-s_2)^{m_{12}}, \dots, (s-s_q)^{m_{rq}}$$

appearing in the invariant polynomials (1.9.1) is called elementary divisor of the matrix $\mathbf{A}(s)$.

For example, the elementary divisors of the polynomial matrix (1.8.3) are $(s+2)$ and $(s+2, 5)$.

The elementary divisors of a polynomial matrix are uniquely determined. This follows immediately from the uniqueness of the invariant polynomial of polynomial matrices. Equivalent polynomial matrices possess the same elementary divisors. For a polynomial matrix of known dimensions its rank together with its elementary divisors uniquely determine its Smith canonical form.

For example, knowing the elementary divisors $s-1$, $(s-1)(s-2)$, $(s-2)^2$, $(s-3)$, of a polynomial matrix, its rank $r = 4$ and dimension 4×4, we can write its Smith canonical form of this polynomial matrix

$$\mathbf{A}_s(s) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & s-1 & 0 & 0 \\ 0 & 0 & (s-1)(s-2) & 0 \\ 0 & 0 & 0 & (s-1)(s-2)^2(s-3) \end{bmatrix}. \tag{1.9.3}$$

Consider a polynomial, block-diagonal matrix of the form

$$\mathbf{A}(s) = \mathrm{diag}\,[\mathbf{A}_1(s),\ \mathbf{A}_2(s)] = \begin{bmatrix} \mathbf{A}_1(s) & 0 \\ 0 & \mathbf{A}_2(s) \end{bmatrix}. \qquad (1.9.4)$$

Let $\mathbf{A}_{kS}(s)$ be the Smith canonical form of the matrix $\mathbf{A}_k(s)$, $k = 1,2$, and

$$(s - s_{k_1})^{m_{11}^k}, \ldots, (s - s_{kq})^{m_{rk,qk}^k}$$

its elementary divisors.

Taking into account that equivalent polynomial matrices have the same elementary divisors, we establish that a set of elementary divisors of the matrix (1.9.4) is the sum of the sets of elementary divisors of $\mathbf{A}_k(s)$, $k = 1,2$.

Example 1.9.1.
Determine elementary divisors of the block-diagonal matrix (1.9.4) for

$$\mathbf{A}_1(s) = \begin{bmatrix} s-1 & 1 & 0 \\ 0 & s-1 & 1 \\ 0 & 0 & s-1 \end{bmatrix}, \ \mathbf{A}_2(s) = \begin{bmatrix} s-1 & 1 & 0 \\ 0 & s-1 & 0 \\ 0 & 0 & s-2 \end{bmatrix}. \qquad (1.9.5)$$

It is easy to check that the Smith canonical forms of the matrices (1.9.5) are

$$\mathbf{A}_{1S}(s) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (s-1)^3 \end{bmatrix}, \ \mathbf{A}_{2S}(s) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (s-1)^2(s-2) \end{bmatrix}. \qquad (1.9.6)$$

The elementary divisors of the matrices (1.9.5) are thus equal $(s-1)^3$, $(s-1)^2$, and $(s-2)$, respectively. It is easy to show that the Smith canonical form of the matrix (1.9.4) with the blocks (1.9.5) is equal to

$$\mathbf{A}_S(s) = \mathrm{diag}\begin{bmatrix} 1 & 1 & 1 & 1 & (s-1)^2 & (s-1)^3 & (s-2) \end{bmatrix} \qquad (1.9.7)$$

and its elementary divisors are $(s - 1)^2$, $(s - 1)^3$, $(s - 2)$.

Consider a matrix $\mathbf{A}\in\mathbb{R}^{n\times n}$ and its corresponding polynomial matrix $[\mathbf{I}_n s - \mathbf{A}]$. Let

$$[\mathbf{I}_n s - \mathbf{A}]_S = \mathrm{diag}\left[i_1(s),\ i_2(s),\ \ldots,\ i_n(s)\right], \qquad (1.9.8)$$

where

$$i_k(s) = (s - s_1)^{m_{k1}}(s - s_2)^{m_{k2}} \ldots (s - s_q)^{m_{kq}},\ k = 1,\ \ldots,\ n, \qquad (1.9.9)$$

and $s_1, s_2, \ldots, s_q$, $q \le n$ are the eigenvalues of the matrix **A**.

Definition 1.9.2. Everyone of the expressions (different from 1)

$$(s-s_1)^{m_{11}}, (s-s_2)^{m_{12}}, \ldots, (s-s_q)^{m_{nq}}$$

appearing in the invariant polynomials (1.9.9) is called the elementary divisor of the matrix **A**.

The elementary divisors of the matrix **A** are uniquely determined and they determine its essential structural properties.

1.9.2 Zeros of Polynomial Matrices

Consider a polynomial matrix $\mathbf{A}(s) \in \mathbb{C}^{m\times n}[s]$ of rank r, whose Smith canonical form is equal to (1.8.1). From (1.8.5) it follows that

$$D_r(s) = i_1(s) i_2(s) \ldots i_r(s) . \tag{1.9.10}$$

Definition 1.9.3. Zeros of the polynomial (1.9.10) are called zeros of the polynomial matrix $\mathbf{A}(s)$.

The zeros of the polynomial matrix $\mathbf{A}(s)$ can be equivalently defined as those values of the variable s, for which this matrix loses its full (normal) rank. For example, for the polynomial matrix (1.8.3) we have

$$D_r(s) = (s+2)(s+2.5) .$$

Thus the zeros of the matrix are $s_1^0 = -2$, $s_2^0 = -2.5$.

It is easy to verify that for these values of the variable s, the matrix (1.8.3) (whose normal rank is equal to 2) has a rank equal to 1.

If the polynomial matrix $\mathbf{A}(s)$ is square and of the full rank $r = n$, then

$$\det \mathbf{A}(s) = cD_r(s) \quad \left(c \text{ is a constant coefficient independent of } s\right) \tag{1.9.11}$$

and the zeros of this matrix coincide with the roots of its characteristic equation $\det \mathbf{A}(s) = 0$.

For example, for the first among the matrices (1.9.5) we have

$$\det \mathbf{A}_r(s) = \begin{vmatrix} s-1 & 1 & 0 \\ 0 & s-1 & 1 \\ 0 & 0 & s-1 \end{vmatrix} = (s-1)^3 .$$

Thus this matrix has the zero $s = 1$ of multiplicity 3. The same result will be obtained from (1.9.10), since $D_r(s) = (s - 1)^3$ for $\mathbf{A}_{1S}(s)$.

Theorem 1.9.1. Let a polynomial matrix $\mathbf{A}(s)\in\mathbb{C}^{m\times n}[s]$ have a rank (normal) equal to $r \leq \min(m,n)$. Then

$$\operatorname{rank} \mathbf{A}(s) = \begin{Bmatrix} r & s \notin \sigma_A \\ r - d_i & s = s_i \in \sigma_A \end{Bmatrix}, \tag{1.9.12}$$

where σ_A is a set of the zeros of the matrix $\mathbf{A}(s)$ and d_i is a number of distinct elementary divisors containing s_i.

Proof. By definition of zero, it follows that the matrix $\mathbf{A}(s)$ does not lose its full rank if we substitute in place of the variable s a number that does not belong to the set σ_A, i.e., rank $\mathbf{A}(s) = r$ for $s\notin\sigma_A$. Elementary operations do not change the rank of a polynomial matrix. In view of this rank $\mathbf{A}(s) =$ rank $\mathbf{A}_S(s) = r$, where r is the number of the invariant polynomials (including those equal to 1). If an invariant polynomial contains s_i, then this polynomial is equal to zero for $s = s_i$. Thus we have rank $\mathbf{A}(s_i) = r - d_i$, $s_i\in\sigma_A$, since the number of polynomials containing s_i is equal to the number of distinct elementary divisors containing s_i.

■

For instance, the polynomial matrix (1.9.3) of the full column rank has one elementary divisor containing $s_1^0 = 3$, two elementary divisors containing $s_2^0 = 2$ and three elementary divisors containing $s_3^0 = 1$. In view of this, according to (1.9.12) we have

$$\operatorname{rank} \mathbf{A}_S(3) = 3, \quad \operatorname{rank} \mathbf{A}_S(2) = 2, \quad \operatorname{rank} \mathbf{A}_S(1) = 1 .$$

Remark 1.9.1.
A unimodular matrix $\mathbf{U}(s)\in\mathbb{R}^{n\times n}[s]$ does not have any zeros since det $\mathbf{U}(s) = c$, where c is certain constant independent of the variable s.

Theorem 1.9.2. An arbitrary rectangular, polynomial matrix $\mathbf{A}(s)\in\mathbb{R}^{m\times n}[s]$ of full rank that does not have any zeros can be written in the form

$$\mathbf{A}(s) = \begin{Bmatrix} \begin{bmatrix} \mathbf{I}_m & 0 \end{bmatrix} \mathbf{P}(s), & m < n \\ \mathbf{L}(s) \begin{bmatrix} \mathbf{I}_n \\ 0 \end{bmatrix}, & m > n \end{Bmatrix}, \tag{1.9.13}$$

where $\mathbf{P}(s)\in\mathbb{R}^{n\times n}[s]$ and $\mathbf{L}(s)\in\mathbb{R}^{m\times m}[s]$ are unimodular matrices.

Proof. If $m < n$ and the matrix does not have any zeros, then applying elementary operations on columns we can bring this matrix to the form $[\mathbf{I}_m \ 0]$. Similarly, if

$m > n$ and the matrix does not have any zeros, then applying elementary operations on rows we can bring this matrix to the form $\begin{bmatrix} \mathbf{I}_n \\ 0 \end{bmatrix}$.

■

Remark 1.9.2.
From the relationship (1.9.13) it follows that a polynomial matrix built from an arbitrary number of rows or columns of a matrix that does not have any zeros, never has any zeros.

Theorem 1.9.3. An arbitrary polynomial matrix $\mathbf{A}(s) \in \mathbb{R}^{m \times n}[s]$ of rank $r \le \min(m, n)$ having zeros can be presented in the form of the product of matrices

$$\mathbf{A}(s) = \mathbf{B}(s)\mathbf{C}(s), \tag{1.9.14}$$

where the matrix $\mathbf{B}(s) = \mathbf{L}^{-1}(s)\,\mathrm{diag}\,[i_1(s), \ldots, i_r(s), 0, \ldots, 0] \in \mathbb{R}^{m \times m}$ is a matrix containing all the zeros of the matrix $\mathbf{A}(s)$, and

$$\mathbf{C}(s) = \begin{Bmatrix} [\mathbf{I}_m \quad \mathbf{0}]\mathbf{P}^{-1}(s), & n > m \\ \mathbf{P}^{-1}(s), & n = m \\ \begin{bmatrix} \mathbf{I}_n \\ \mathbf{0} \end{bmatrix}\mathbf{P}^{-1}(s), & n < m \end{Bmatrix}. \tag{1.9.15}$$

Proof. Let $\mathbf{L}(s) \in \mathbb{R}^{m \times m}[s]$ and $\mathbf{P}(s) \in \mathbb{R}^{n \times n}[s]$ be unimodular matrices of elementary operations on rows and on columns, respectively, reducing the matrix $\mathbf{A}(s)$ to the Smith canonical form $\mathbf{A}_S(s)$, i.e.,

$$\mathbf{A}_S(s) = \mathbf{L}(s)\mathbf{A}(s)\mathbf{P}(s). \tag{1.9.16}$$

Pre-multiplying (1.9.16) by $\mathbf{L}^{-1}(s)$ and post-multiplying by $\mathbf{P}^{-1}$(s), we obtain

$$\mathbf{A}(s) = \mathbf{L}^{-1}(s)\mathbf{A}_S(s)\mathbf{P}^{-1}(s) = \mathbf{B}(s)\mathbf{C}(s),$$

since

$$\mathbf{A}_S(s) = \begin{Bmatrix} \mathrm{diag}\,[i_1(s), \ldots, i_r(s), 0, \ldots, 0][\mathbf{I}_m \quad 0], & n > m \\ \mathrm{diag}\,[i_1(s), \ldots, i_r(s), 0, \ldots, 0], & n = m \\ \mathrm{diag}\,[i_1(s), \ldots, i_r(s), 0, \ldots, 0]\begin{bmatrix} \mathbf{I}_n \\ 0 \end{bmatrix}, & n < m \end{Bmatrix}.$$

From (1.9.15) it follows that the matrix $\mathbf{C}(s)\in\mathbb{R}^{m\times n}[s]$ does not have any zeros, since the matrix $\mathbf{P}^{-1}(s)$ is a unimodular matrix.

■

1.10 Similarity and Equivalence of First Degree Polynomial Matrices

Definition 1.10.1. Two square matrices **A** and **B** of the same dimension are said to be similar matrices if and only if there exists a nonsingular matrix **P** such that

$$\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} \tag{1.10.1}$$

and the matrix **P** is called a similarity transformation matrix.

Theorem 1.10.1. Similar matrices have the same characteristic polynomials, i.e.,

$$\det\,[s\mathbf{I}-\mathbf{B}] = \det\,[s\mathbf{I}-\mathbf{A}]. \tag{1.10.2}$$

Proof. Taking into account (1.10.1), we can write

$$\begin{aligned}\det\,[s\mathbf{I}-\mathbf{B}] &= \det\left[s\mathbf{P}^{-1}\mathbf{P}-\mathbf{P}^{-1}\mathbf{A}\mathbf{P}\right] = \det\left[\mathbf{P}^{-1}(s\mathbf{I}-\mathbf{A})\mathbf{P}\right]\\ &= \det\mathbf{P}^{-1}\det\,[s\mathbf{I}-\mathbf{A}]\det\mathbf{P} = \det\,[s\mathbf{I}-\mathbf{A}],\end{aligned}$$

since $\det\mathbf{P}^{-1} = (\det\mathbf{P})^{-1}$. ■

Theorem 1.10.2. Polynomial matrices $[s\mathbf{I} - \mathbf{A}]$ and $[s\mathbf{I} - \mathbf{B}]$ are equivalent if and only if the matrices **A** and **B** are similar.

Proof. Firstly, we show that if the matrices **A** and **B** are similar, then the polynomial matrices $[s\mathbf{I} - \mathbf{A}]$ and $[s\mathbf{I} - \mathbf{B}]$ are equivalent. If the matrices **A** and **B** are similar, i.e., they satisfy the relationship (1.10.1), then

$$[s\mathbf{I}-\mathbf{B}] = \left[s\mathbf{I}-\mathbf{P}^{-1}\mathbf{A}\mathbf{P}\right] = \mathbf{P}^{-1}[s\mathbf{I}-\mathbf{A}]\mathbf{P}.$$

This relationship is a special case (for $\mathbf{L}(s) = \mathbf{P}^{-1}$ and $\mathbf{P}(s) = \mathbf{P}$) of the relationship (1.7.2). Thus the polynomial matrices $[s\mathbf{I} - \mathbf{A}]$ and $[s\mathbf{I} - \mathbf{B}]$ are equivalent. We will show now, that if the matrices $[s\mathbf{I} - \mathbf{A}]$ and $[s\mathbf{I} - \mathbf{B}]$ are equivalent, then the matrices **A** and **B** are similar. Assuming that the matrices $[s\mathbf{I} - \mathbf{A}]$ and $[s\mathbf{I} - \mathbf{B}]$ are equivalent, we have

$$[s\mathbf{I}-\mathbf{B}] = \mathbf{L}(s)[s\mathbf{I}-\mathbf{A}]\mathbf{P}(s), \tag{1.10.3}$$

where $\mathbf{L}(s)$ and $\mathbf{P}(s)$ are unimodular matrices. The determinant of the matrix $\mathbf{L}(s)$ is different from zero and does not depend on the variable s. In view of this, the inverse matrix

$$\mathbf{Q}(s) = \mathbf{L}^{-1}(s)$$

is a polynomial, unimodular matrix as well. Pre-dividing the matrix $\mathbf{Q}(s)$ by $[s\mathbf{I} - \mathbf{A}]$ and post-dividing $\mathbf{P}(s)$ by $[s\mathbf{I} - \mathbf{B}]$, we obtain

$$\mathbf{Q}(s) = [s\mathbf{I} - \mathbf{A}]\mathbf{Q}_1(s) + \mathbf{Q}_0\,, \tag{1.10.4}$$
$$\mathbf{P}(s) = \mathbf{P}_1(s)[s\mathbf{I} - \mathbf{B}] + \mathbf{P}_0\,, \tag{1.10.5}$$

where $\mathbf{Q}_1(s)$ and $\mathbf{P}_1(s)$ are polynomial matrices and the matrices $\mathbf{Q}_0$ and $\mathbf{P}_0$ do not depend on the variable s. With (1.10.3) pre-multiplied by $\mathbf{Q}(s) = \mathbf{L}^{-1}(s)$ we obtain

$$\mathbf{Q}(s)[s\mathbf{I} - \mathbf{B}] = [s\mathbf{I} - \mathbf{A}]\mathbf{P}(s) \tag{1.10.6}$$

and after substitution of (1.10.4) and (1.10.5) into (1.10.6)

$$[s\mathbf{I} - \mathbf{A}]\big[\mathbf{Q}_1(s) - \mathbf{P}_1(s)\big][s\mathbf{I} - \mathbf{B}] = [s\mathbf{I} - \mathbf{A}]\mathbf{P}_0 - \mathbf{Q}_0[s\mathbf{I} - \mathbf{B}]\,. \tag{1.10.7}$$

Note that the following equality must hold

$$\mathbf{Q}_1(s) = \mathbf{P}_1(s)\,, \tag{1.10.8}$$

since otherwise the left-hand side of (1.10.7) would be a matrix polynomial of a degree of at least 2, and the right side a matrix polynomial of degree of at most 1. After taking into account the equality (1.10.8) from (1.10.7) we obtain

$$\mathbf{Q}_0[s\mathbf{I} - \mathbf{B}] = [s\mathbf{I} - \mathbf{A}]\mathbf{P}_0\,. \tag{1.10.9}$$

Pre-division of the matrix $\mathbf{L}(s)$ by $[s\mathbf{I} - \mathbf{B}]$ yields

$$\mathbf{L}(s) = [s\mathbf{I} - \mathbf{B}]\mathbf{L}_1(s) + \mathbf{L}_0\,, \tag{1.10.10}$$

where $\mathbf{L}_1(s)$ is a polynomial matrix and $\mathbf{L}_0$ is a matrix independent of the variable s.

We will show that the matrices $\mathbf{Q}_0$ and $\mathbf{L}_0$ are nonsingular matrices satisfying the condition

$$\mathbf{Q}_0\mathbf{L}_0 = \mathbf{I}\,. \tag{1.10.11}$$

Substitution of (1.10.4) and (1.10.10) into the equality

$$\mathbf{Q}(s)\mathbf{L}(s) = \mathbf{I}$$

yields

$$\begin{aligned}\mathbf{I} = \mathbf{Q}(s)\mathbf{L}(s) &= \left[(s\mathbf{I}-\mathbf{A})\mathbf{Q}_1(s)+\mathbf{Q}_0\right]\left[(s\mathbf{I}-\mathbf{B})\mathbf{L}_1(s)+\mathbf{L}_0\right] = \\ &= [s\mathbf{I}-\mathbf{A}]\mathbf{Q}_1(s)[s\mathbf{I}-\mathbf{B}]\mathbf{L}_1(s)+\mathbf{Q}_0[s\mathbf{I}-\mathbf{B}]\mathbf{L}_1(s)+ \\ &+[s\mathbf{I}-\mathbf{A}]\mathbf{Q}_1(s)\mathbf{L}_0+\mathbf{Q}_0\mathbf{L}_0.\end{aligned} \tag{1.10.12}$$

Note that this equality can be satisfied if and only if

$$[s\mathbf{I}-\mathbf{A}]\mathbf{Q}_1(s)[s\mathbf{I}-\mathbf{B}]\mathbf{L}_1(s)+\mathbf{Q}_0[s\mathbf{I}-\mathbf{B}]\mathbf{L}_1(s)+[s\mathbf{I}-\mathbf{A}]\mathbf{Q}_1(s)\mathbf{L}_0 = 0\,. \tag{1.10.13}$$

Otherwise the left-hand side of (1.10.12) would be a matrix polynomial of zero degree and the right-hand side would be a matrix polynomial of at least the first degree. With (1.10.13) taken into account, from (1.10.12) we obtain the equality (1.10.11).

From this equality the nonsingularity of the matrices $\mathbf{Q}_0$ and $\mathbf{L}_0$ as well the equality $\mathbf{L}_0 = \mathbf{Q}_0^{-1}$ follow immediately.

Pre-multiplication of (1.10.9) by $\mathbf{Q}_0^{-1}$ yields

$$[s\mathbf{I}-\mathbf{B}] = \mathbf{L}_0[s\mathbf{I}-\mathbf{A}]\mathbf{P}_0$$

and

$$\mathbf{B} = \mathbf{L}_0\mathbf{A}\mathbf{P}_0\,, \quad \mathbf{L}_0\mathbf{P}_0 = \mathbf{I}\,.$$

From these relationships it follows that the matrices **A** and **B** are similar.

■

Theorem 1.10.3. Matrices **A** and **B** are similar if and only if the matrices $[s\mathbf{I}-\mathbf{A}]$ and $[s\mathbf{I}-\mathbf{B}]$ have the same invariant polynomials.

Proof. According to Corollary 1.8.1 two matrices are equivalent if and only if they have the same invariant polynomials. From Theorem 1.10.2 it follows immediately that the polynomial matrices $[s\mathbf{I}-\mathbf{A}]$ and $[s\mathbf{I}-\mathbf{B}]$ have the same invariant polynomials if and only if the matrices **A** and **B** are similar. Thus the matrices **A** and **B** are similar if and only if the matrices $[s\mathbf{I}-\mathbf{A}]$ and $[s\mathbf{I}-\mathbf{B}]$ have the same invariant polynomials.

■

1.11 Computation of the Frobenius and Jordan Canonical Forms of Matrices

1.11.1 Computation of the Frobenius Canonical Form of a Square Matrix

Consider $n \times n$ matrices of the form

$$\mathbf{F} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-2} & -a_{n-1} \end{bmatrix}, \overline{\mathbf{F}} = \begin{bmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & \dots & 0 & -a_2 \\ 0 & 0 & \dots & 1 & -a_{n-1} \end{bmatrix},$$
$$\hat{\mathbf{F}} = \begin{bmatrix} -a_{n-1} & -a_{n-2} & \dots & -a_1 & -a_0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}, \quad \tilde{\mathbf{F}} = \begin{bmatrix} -a_{n-1} & 1 & 0 & \dots & 0 \\ -a_{n-2} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_1 & 0 & 0 & \dots & 1 \\ -a_0 & 0 & 0 & \dots & 0 \end{bmatrix}. \tag{1.11.1}$$

We say that the matrices in (1.11.1) have Frobenius canonical forms (or normal canonical forms).

Expanding along the row (or the column) containing $a_0, a_1, \dots, a_{n-1}$, it is easy to show that

$$\begin{aligned} &\det\left[\mathbf{I}_n s - \mathbf{F}\right] = \det\left[\mathbf{I}_n s - \overline{\mathbf{F}}\right] = \det\left[\mathbf{I}_n s - \hat{\mathbf{F}}\right] = \det\left[\mathbf{I}_n s - \tilde{\mathbf{F}}\right] = \\ &= s^n + a_{n-1}s^{n-1} + \dots + a_1 s + a_0. \end{aligned} \tag{1.11.2}$$

We will show that the polynomial (1.11.2) is the only invariant polynomial of the matrix (1.11.1) different from 1. Detailed considerations will be given only for the matrix **F**. The proof in the other three cases is similar. Deleting the first column and the n-th row in the matrix

$$\left[\mathbf{I}_n s - \mathbf{F}\right] = \begin{bmatrix} s & -1 & 0 & \dots & 0 & 0 \\ 0 & s & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & s & -1 \\ a_0 & a_1 & a_2 & \dots & a_{n-2} & s + a_{n-1} \end{bmatrix}, \tag{1.11.3}$$

we obtain the minor M_{n1} equal to $(-1)^{n-1}$. With the above in mind, a greatest common devisor of all minors of degree $n - 1$ of this matrix is equal to 1, i.e.,

$D_{n-1}(s) = 1$. From the relationship (1.8.4) it follows that the polynomial (1.11.2) is the only polynomial of the matrix **F** different from 1.

Let $\mathbf{A} \in \mathbb{C}^{n \times n}$ and the monic polynomials

$$i_1(s) = 1, \ ..., \ i_p(s) = 1, \quad i_{p+1}(s), \ ..., \ i_n(s)$$

be the invariant polynomials of the polynomial matrix $[\mathbf{I}_n s - \mathbf{A}]$, where $i_{p+1}(s),\ldots,i_n(s)$ are the polynomials of at least the first degree such that $i_k(s)$ divides (without remainder) $i_{k+1}(s)$ $(k = p+1,...,n-1)$. The matrix $[s\mathbf{I} - \mathbf{A}]$ reduced to the Smith canonical form is of the form

$$[s\mathbf{I} - \mathbf{A}]_S = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \cdots \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & i_{p+1}(s) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & i_n(s) \end{bmatrix}. \tag{1.11.4}$$

Let $\mathbf{F}_{p+1},\ldots,\mathbf{F}_n$ be the matrices of the form (1.11.1) that correspond to the invariant polynomials $i_{p+1}(s),\ldots,i_n(s)$. From considerations of Sect. 1.10 it follows that the quasi-diagonal matrix

$$\mathbf{F}_A = \begin{bmatrix} \mathbf{F}_{p+1} & 0 & \cdots & 0 \\ 0 & \mathbf{F}_{p+2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{F}_n \end{bmatrix} \tag{1.11.5}$$

and **A** have the same invariant polynomials. Thus according to Theorem 1.10.2, the matrices **A** and $\mathbf{F}_A$ are similar. Hence there exists a nonsingular matrix **P** such that

$$\mathbf{A} = \mathbf{PFAP}^{-1}. \tag{1.11.6}$$

The matrix $\mathbf{F}_A$ given by (1.11.5) is called a Frobenius canonical form or a normal canonical form of the square matrix **A**.

Thus the following important theorem has been proved.

Theorem 1.11.1. For every matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ there exists a nonsingular matrix $\mathbf{P} \in \mathbb{C}^{n \times n}$ such that the equality (1.11.6) holds.

Example 1.11.1.
The following matrix is given

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 2 \end{bmatrix}. \tag{1.11.7}$$

Carrying out the elementary operations: $P[1+2\times(s-1)]$, $L[2+1\times(s-1)]$, $P[3+1\times(-s+2)]$, $L[2\times(-1)]$, $L[2+3\times(s-1)^2]$, $L[1\times(-1)]$, $L[2, 3]$, $L[1, 2]$ on the matrix

$$[s\mathbf{I}_3 - \mathbf{A}] = \begin{bmatrix} s-1 & -1 & 0 \\ 0 & s-1 & 0 \\ 1 & 0 & s-2 \end{bmatrix},$$

we transform this matrix to its Smith canonical form

$$[s\mathbf{I}_3 - \mathbf{A}]_S = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (s-1)^2(s-2) \end{bmatrix}.$$

Thus the matrix $\mathbf{A}$ has the only invariant polynomial different from one

$$i_3(s) = (s-1)^2(s-2) = s^3 - 4s^2 + 5s - 2.$$

In view of this, the Frobenius canonical form of the matrix (1.11.7) is the following

$$\mathbf{F}_A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -5 & 4 \end{bmatrix}. \tag{1.11.8}$$

1.11.2 Computation of the Jordan Canonical Form of a Square Matrix

Consider an elementary divisor of the form

$$\left(s - s_0\right)^m. \tag{1.11.9}$$

We will show that the polynomial (1.11.9) is the only elementary divisor of a square matrix of the form

$$\mathbf{J} = \mathbf{J}(s_1^0, m) = \begin{bmatrix} s_0 & 1 & 0 & \cdots & 0 \\ 0 & s_0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & s_0 \end{bmatrix} \in \mathbb{C}^{m\times m}, \tag{1.11.10a}$$

or

$$\mathbf{J}' = \mathbf{J}'(s_1^0, m) = \begin{bmatrix} s_0 & 0 & \cdots & 0 & 0 \\ 1 & s_0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & s_0 & 0 \\ 0 & 0 & \cdots & 1 & s_0 \end{bmatrix} \in \mathbb{C}^{m\times m}. \tag{1.11.10b}$$

The determinant of the polynomial matrix

$$[s\mathbf{I}_m - \mathbf{J}] = \begin{bmatrix} s-s_0 & -1 & 0 & \dots & 0 & 0 \\ 0 & s-s_0 & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & s-s_0 & -1 \\ 0 & 0 & 0 & \dots & 0 & s-s_0 \end{bmatrix} \tag{1.11.11}$$

is equal to the polynomial (1.11.9).

The minor M_{n1} obtained from the matrix (1.11.11) by removing the first and the m-th columns is equal to $(-1)^{m-1}$. Thus a greatest common divisor of all minors of degree m-1 of the matrix (1.11.11) is equal to 1, $D_{m\text{-}1}(s) = 1$. From (1.8.4) it follows that the polynomial (1.11.9) is the only invariant polynomial of the matrix (1.11.11) different from 1. The proof for the matrix $\mathbf{J}'$ is similar.

The matrices $\mathbf{J}$ and $\mathbf{J}'$ are called Jordan blocks of the first and the second type, respectively.

If q elementary divisors correspond to one eigenvalue, then q Jordan blocks correspond to this eigenvalue.

Let $\mathbf{J}_1$, $\mathbf{J}_2,\ldots,\mathbf{J}_p$ be Jordan blocks of the form (1.11.10a) (or (1.11.10b)), corresponding to the elementary divisors of the matrix $\mathbf{A}$, where p is the number of elementary divisors of this matrix. Note that all these elementary divisors of the matrix $\mathbf{A}$ are also the elementary divisors of a quasi-diagonal matrix of the form

$$\mathbf{J}_A = \begin{bmatrix} \mathbf{J}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{J}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{J}_p \end{bmatrix} \in \mathbb{C}^{n\times n}. \tag{1.11.12}$$

Matrices having the same elementary divisors also have the same invariant polynomials. In view of this, according to Theorem 1.10.2, the matrices $\mathbf{A}$ and $\mathbf{J}_A$, being matrices having the same invariant polynomials, are similar. Thus there exists a nonsingular matrix $\mathbf{T}$ such that

$$\mathbf{A} = \mathbf{T}\mathbf{J}_A\mathbf{T}^{-1}. \tag{1.11.13}$$

The matrix (1.11.12) is called the Jordan canonical form of the matrix $\mathbf{A}$, or shortly the Jordan matrix. Thus the following important theorem has been proved.

Theorem 1.11.2. For every matrix $\mathbf{A}\in\mathbb{C}^{n\times n}$ there exists a nonsingular matrix $\mathbf{T}\in\mathbb{C}^{n\times n}$ such that the equality (1.11.13) holds.

If all elementary divisors of the matrix $\mathbf{A}$ are of the first degree (in the relationship (1.11.9) $m = 1$), then a Jordan matrix is a diagonal one. Thus we have the following important corollary.

Corollary 1.11.1. A matrix $\mathbf{A}$ is similar to the diagonal matrix consisting of its eigenvalues if and only if all its elementary divisors are divisors of the first degree.

Example 1.11.2.
The matrix (1.11.7) has only one invariant polynomial different from one and equal to $i(s) = (s-1)^2(s-2)$. Thus this matrix has two elementary divisors $(s-1)^2$ and $(s-2)$. Hence the Smith canonical form of the matrix (1.11.7) is equal to

$$\mathbf{J}_A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

1.12 Computation of Similarity Transformation Matrices

1.12.1 Matrix Pair Method

A cyclic matrix $\mathbf{A}\in\mathbb{R}^{n\times n}$ and its Frobenius form $\mathbf{F}_A$ are given. Compute a nonsingular matrix $\mathbf{P}\in\mathbb{R}^{n\times n}$ such that

$$\mathbf{PAP}^{-1} = \mathbf{F}_A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix}. \tag{1.12.1}$$

For the given matrix **A** we choose a row matrix $c \in \mathbb{R}^{1\times n}$ such that

$$\det \begin{bmatrix} c \\ c\mathbf{A} \\ \vdots \\ c\mathbf{A}^{n-1} \end{bmatrix} \neq 0 . \tag{1.12.2}$$

Almost every matrix c chosen by a "triall and error" method will satisfy the condition (1.12.2), since in the space of parameters the elements of the matrix c lie on a plane.

We choose the matrix **P** in such a way that the condition (1.12.1) holds and

$$c\mathbf{P}^{-1} = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} \in \mathbb{R}^{1\times n} . \tag{1.12.3}$$

Letting p_i ($i = 1,2,...,n$) be the i-th row of the matrix **P**. Using (1.12.1) and (1.12.3), we can write

$$\begin{gathered} \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix} \mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix} \\ \text{and } c = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix}. \end{gathered} \tag{1.12.4}$$

Carrying out the multiplication and comparing appropriate rows from (1.12.4), we obtain

$$p_1 = c,\ p_2 = p_1\mathbf{A},\ p_3 = p_2\mathbf{A}, \dots,\ p_n = p_{n-1}\mathbf{A} . \tag{1.12.5}$$

Using (1.12.5) we can compute the unknown rows $p_1, p_2, ..., p_n$ of the matrix **P**.

Thus we have the following procedure for computation of the matrix **P**.

Procedure 1.12.1.
Step 1: Compute the coefficients $a_0, a_1, \ldots, a_{n-1}$ of the polynomial

$$\det[\mathbf{I}_n s - \mathbf{A}] = s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0 \,. \tag{1.12.6}$$

Step 2: Knowing $a_0, a_1, \ldots, a_{n-1}$ compute the matrix $\mathbf{F}_A$.
Step 3: Choose $c \in \mathbb{R}^{1\times n}$ such that the condition (1.12.2) holds.
Step 4: Using (1.12.5) compute the rows $p_1, p_2, \ldots, p_n$ of the matrix $\mathbf{P}$.

Example 1.12.1.
The following cyclic matrix is given

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 2 \end{bmatrix}. \tag{1.12.7}$$

One has to compute a matrix $\mathbf{P}$ transforming this matrix by similarity to the Frobenius canonical form $\mathbf{F}_A$.

Using Procedure 1.12.1, we obtain the following:
Step 1: The characteristic polynomial of the matrix (1.12.7) has the form:

$$\det[\mathbf{I}_n s - \mathbf{A}] = \begin{vmatrix} s-1 & -1 & 0 \\ 0 & s-1 & 0 \\ 1 & 0 & s-2 \end{vmatrix} = (s-2)(s-1)^2 = s^3 - 4s^2 + 5s - 2. \tag{1.12.8}$$

Step 2: Thus the matrix $\mathbf{F}_A$ has the form

$$\mathbf{F}_A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -5 & 4 \end{bmatrix}. \tag{1.12.9}$$

Step 3: We choose $c = [1 \;\; 0 \;\; 1]$ satisfying the condition (1.12.2), since

$$\det\begin{bmatrix} c \\ c\mathbf{A} \\ c\mathbf{A}^2 \end{bmatrix} = \begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ -2 & 1 & 4 \end{vmatrix} = 4 \,.$$

Step 4: Using (1.12.5), we obtain

$$p_1 = c = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}, \quad p_2 = p_1\mathbf{A} = \begin{bmatrix} 0 & 1 & 2 \end{bmatrix}, \quad p_3 = p_2\mathbf{A} = \begin{bmatrix} -2 & 1 & 4 \end{bmatrix}.$$

Thus the matrix **P** has the form

$$\mathbf{P} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ -2 & 1 & 4 \end{bmatrix}. \tag{1.12.10}$$

If we search for a matrix $\overline{\mathbf{P}}$ that satisfies the condition

$$\overline{\mathbf{P}}^{-1}\mathbf{A}\overline{\mathbf{P}} = \overline{\mathbf{F}}_A = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{bmatrix}, \tag{1.12.11}$$

then it is convenient to choose a column matrix $\mathbf{b} \in \mathbb{R}^n$ in such a way that

$$\det[b, \mathbf{A}b, \cdots, \mathbf{A}^{n-1}b] \neq 0 . \tag{1.12.12}$$

Let $\overline{p}_i$ ($i = 1,\dots,n$) be the i-th column of the matrix $\overline{\mathbf{P}}$. Using (1.12.11) and $\overline{\mathbf{P}}^{-1}\mathbf{b}=[1 \;\; 0 \;\; \dots \;\; 0]^T \in \mathbb{R}^n$, we can write

$$\mathbf{A}\left[\overline{p}_1 \quad \overline{p}_2 \quad \cdots \quad \overline{p}_n\right] = \left[\overline{p}_1 \quad \overline{p}_2 \quad \cdots \quad \overline{p}_n\right] \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{bmatrix}, \tag{1.12.13}$$

$$\mathbf{b} = \left[\overline{p}_1 \quad \overline{p}_2 \quad \cdots \quad \overline{p}_n\right] \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Multiplying and comparing appropriate columns from (1.12.13), we obtain

$$\overline{p}_1 = \mathbf{b}, \overline{p}_2 = \mathbf{A}\overline{p}_1, \overline{p}_3 = \mathbf{A}\overline{p}_2, \cdots, \overline{p}_n = \mathbf{A}\overline{p}_{n-1} . \tag{1.12.14}$$

Using (1.12.14), we can successively compute the columns $\overline{p}_1$, $\overline{p}_2, \dots,$ $\overline{p}_n$ of the matrix $\overline{\mathbf{P}}$. Thus we have the following procedure for computation of the matrix $\overline{\mathbf{P}}$.

Procedure 1.12.2.

Step 1: Is the same as in Procedure 1.12.1.

Step 2: Knowing the coefficients $a_1, a_2, \ldots, a_n$ of the polynomial (1.12.6) compute the matrix $\overline{\mathbf{F}}_A$.

Step 3: Choose $\mathbf{b} \in \mathbb{R}^n$ such that the condition (1.12.12) is satisfied.

Step 4: Using (1.12.14) compute the columns $\overline{p}_1, \overline{p}_2, \ldots, \overline{p}_n$ of the matrix $\overline{\mathbf{P}}$.

Example 1.12.2.

Find a matrix $\overline{\mathbf{P}}$ transforming the matrix (1.12.7) by similarity into its canonical form $\overline{\mathbf{F}}_A$.

Using Procedure 1.12.2 we obtain the following:

Step 1: The characteristic polynomial of the matrix (1.12.7) has the form (1.12.8).

Step 2: Thus the matrix $\overline{\mathbf{F}}_A$ has the form

$$\overline{\mathbf{F}}_A = \begin{bmatrix} 0 & 0 & 2 \\ 1 & 0 & -5 \\ 0 & 1 & 4 \end{bmatrix}. \tag{1.12.15}$$

Step 3: We choose $\mathbf{b} = [0 \;\; 1 \;\; -1]^T$, which satisfies the condition (1.12.12), since

$$\det\left[\mathbf{b}, \mathbf{Ab}, \mathbf{A}^2\mathbf{b}\right] = \begin{vmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \\ -1 & -2 & -5 \end{vmatrix} = 2 .$$

Step 4: Using (1.12.14), we obtain

$$\overline{p}_1 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \quad \overline{p}_2 = \mathbf{A}\overline{p}_1 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \quad \overline{p}_3 = \mathbf{A}\overline{p}_2 = \begin{bmatrix} 2 \\ 1 \\ -5 \end{bmatrix}.$$

Thus the desired matrix has the form

$$\overline{\mathbf{P}} = \left[\overline{p}_1 \quad \overline{p}_2 \quad \overline{p}_3\right] = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \\ -1 & -2 & -5 \end{bmatrix}.$$

The above considerations can be generalised for the remaining canonical Frobenius forms $\hat{\mathbf{F}}_A$ and $\tilde{\mathbf{F}}_A$ of the matrix $\mathbf{A}$.

1.12.2 Elementary Operations Method

Substituting into (1.10.5) and (1.10.10) the matrix $\mathbf{B}$ instead of the variable s, we obtain

$$\mathbf{P}(\mathbf{B}) = \mathbf{P}_0\,, \quad \mathbf{L}(\mathbf{B}) = \mathbf{L}_0\,. \tag{1.12.16}$$

Thus from the relationship $\mathbf{B} = \mathbf{L}_0\mathbf{A}\mathbf{P}_0$ it follows that if the matrices $\mathbf{A}$ and $\mathbf{B}$ are similar, i.e., $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$, then the transformation matrix $\mathbf{P}$ is given by the following formula

$$\mathbf{P} = \mathbf{P}(\mathbf{B}) = \left[\mathbf{L}(\mathbf{B})\right]^{-1}, \tag{1.12.17}$$

where $\mathbf{P}(s)$ and $\mathbf{L}(s)$ are unimodular matrices in the equality

$$[s\mathbf{I} - \mathbf{B}] = \mathbf{L}(s)[s\mathbf{I} - \mathbf{A}]\mathbf{P}(s)\,. \tag{1.12.18}$$

To compute $\mathbf{P}(s)$, using elementary operations, we reduce the matrices $[s\mathbf{I} - \mathbf{A}]$, $[s\mathbf{I} - \mathbf{B}]$ to the Smith canonical form

$$[s\mathbf{I} - \mathbf{A}]_S = \mathbf{L}_1(s)[s\mathbf{I} - \mathbf{A}]\mathbf{P}_1(s)\,, \tag{1.12.19}$$

$$[s\mathbf{I} - \mathbf{B}]_S = \mathbf{L}_2(s)[s\mathbf{I} - \mathbf{B}]\mathbf{P}_2(s) \tag{1.12.20}$$

where

$$\mathbf{P}_1(s) = \mathbf{P}_{11}(s)\mathbf{P}_{12}(s)...\mathbf{P}_{1k_1}(s)\,, \tag{1.12.21}$$

$$\mathbf{P}_2(s) = \mathbf{P}_{21}(s)\mathbf{P}_{22}(s)...\mathbf{P}_{2k_2}(s) \tag{1.12.22}$$

where $\mathbf{P}_{11}(s),\mathbf{P}_{12}(s),\ldots,\mathbf{P}_{1k_2}(s)$ and $\mathbf{P}_{21}(s),\mathbf{P}_{22}(s),\ldots,\mathbf{P}_{2k_2}(s)$ are matrices of elementary operations carried out on columns of matrices $[s\mathbf{I} - \mathbf{A}]$ and $[s\mathbf{I} - \mathbf{B}]$, respectively. The matrices $\mathbf{L}_1(s)$ and $\mathbf{L}_2(s)$ are defined similarly. Similarity of the matrices $\mathbf{A}$ and $\mathbf{B}$ implies that

$$\left[s\mathbf{I} - \mathbf{A}\right]_S = \left[s\mathbf{I} - \mathbf{B}\right]_S\,.$$

Taking into account (1.12.19) and (1.12.20) we obtain

$$\mathbf{L}_2(s)[s\mathbf{I} - \mathbf{B}]\mathbf{P}_2(s) = \mathbf{L}_1(s)[s\mathbf{I} - \mathbf{A}]\mathbf{P}_1(s)$$

i.e.,

$$[s\mathbf{I} - \mathbf{B}] = \mathbf{L}_2^{-1}(s)\mathbf{L}_1(s)[s\mathbf{I} - \mathbf{A}]\mathbf{P}_1(s)\mathbf{P}_2^{-1}(s)\,. \tag{1.12.23}$$

From a comparison of (1.12.18) and (1.12.21) to (1.12.23) and (1.12.22), respectively, we obtain:

$$\mathbf{P}(s) = \mathbf{P}_1(s)\mathbf{P}_2^{-1}(s) = \mathbf{P}_{11}(s)\mathbf{P}_{12}(s)...\mathbf{P}_{1k_1}(s)\mathbf{P}_{2k_2}^{-1}(s)...\mathbf{P}_{22}^{-1}(s)\mathbf{P}_{21}^{-1}(s). \quad (1.12.24)$$

Thus we compute the matrix $\mathbf{P}(s)$ carrying out elementary operations given by the matrices on the identity matrix

$$\mathbf{P}_{11}(s),\ \mathbf{P}_{12}(s), ..., \mathbf{P}_{1k_1}(s),\ \mathbf{P}_{2k_2}^{-1}(s), ..., \mathbf{P}_{22}^{-1}(s),\ \mathbf{P}_{21}^{-1}(s) .$$

When computing the inverse matrices to the matrices of elementary operations we use the following relationships:

$$\begin{aligned}&P^{-1}[i \times c] = P\left[i \times \frac{1}{c}\right],\ \ P^{-1}\left[i + j \times b(s)\right] = P\left[i - j \times w(s)\right],\\ &P^{-1}[i, j] = P[j, i] = P[i, j].\end{aligned} \quad (1.12.25)$$

From the above considerations, the following algorithm for computation of the matrix **P** can be inferred.

Algorithm 1.12.1.

Step 1: Transforming the matrices $[s\mathbf{I} - \mathbf{B}]$, $[s\mathbf{I} - \mathbf{A}]$ to the Smith canonical forms, determine the sequence of elementary operations given by the matrices

$$\mathbf{P}_{11}(s),\ \mathbf{P}_{12}(s), ..., \mathbf{P}_{1k_1}(s),\ \mathbf{P}_{21}(s),\ \mathbf{P}_{22}(s), ..., \mathbf{P}_{2k_2}(s) .$$

Step 2: Carrying out elementary operations given by the matrices $\mathbf{P}_{11}(s),\mathbf{P}_{12}(s),\ldots,\mathbf{P}_{1k_2}(s),\mathbf{P}_{2k_2}^{-1}(s)$ $\mathbf{P}_{22}^{-1}(s)$, $\mathbf{P}_{21}^{-1}(s)$ on the identity matrix, compute the matrix $\mathbf{P}(s)$.

Step 3: Substituting in the matrix $\mathbf{P}(s)$ in place of s the matrix **B**, compute the matrix $\mathbf{P} = \mathbf{P}(\mathbf{B})$.

Example 1.12.3.

Compute a matrix **P** that transforms the matrix (1.11.7) to the Frobenius canonical form (1.11.8).

In this case, the matrix $\mathbf{F}_A$ is the matrix **B**.

Step 1: To reduce the matrix

$$s\mathbf{I} - \mathbf{F}_A = \begin{bmatrix} s & -1 & 0 \\ 0 & s & -1 \\ -2 & 5 & s-4 \end{bmatrix}$$

to its Smith canonical form

$$[s\mathbf{I}-\mathbf{F}_A]_S = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (s-1)^2(s-2) \end{bmatrix},$$

the following elementary operations need to be carried out

$$L\left[3+2\times(s-4)\right],\ P\left[2+3\times(s)\right],\ P\left[1+2\times(s)\right],$$
$$L\left[3+1\times\left(s^2-4s+5\right)\right],\ L\left[1\times(-1)\right],\ L\left[2\times(-1)\right],$$
$$P\left[2,3\right],\ P\left[1,2\right].$$

Step 2: In Example 1.11.1 to reduce the matrix $[\mathbf{I}_n s - \mathbf{A}]$ to the Smith canonical form, the following elementary operations are applied

$$P[1+2\times(s-1)],\ L[2+1\times(s-1)],\ P[3+1\times(2-s)],\ L[2\times(-1)],$$
$$L[2+3\times(s-1)^2],\ L[1\times(-1)],\ L[2,3],\ L[1,2].$$

To compute the matrix $\mathbf{P}(s)$ the following elementary operations have to be carried out on the columns of the identity matrix of the third degree

$$P\left[1+2\times(s-1)\right],\ P\left[3+1\times(2-s)\right],\ P\left[2,3\right],\ P\left[1,2\right],$$
$$P\left[1+2\times(-s)\right],\ P\left[2+3\times(-1)\right].$$

Then we obtain

$$\mathbf{P}(s) = \begin{bmatrix} 2(1-s) & 1 & 0 \\ -2(s-1)^2 & -1 & 1 \\ 1 & 0 & 0 \end{bmatrix} =$$
$$= \begin{bmatrix} 0 & 0 & 0 \\ -2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} s^2 + \begin{bmatrix} -2 & 0 & 0 \\ 4 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} s + \begin{bmatrix} 2 & 1 & 0 \\ -2 & -1 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

Step 3: We substitute into this matrix the matrix

$$\mathbf{F}_A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -5 & 4 \end{bmatrix}$$

in place of the variable s. We obtain

$$\mathbf{P} = \mathbf{P}(\mathbf{F}_A) = \begin{bmatrix} 2 & -1 & 0 \\ -2 & 3 & -1 \\ 1 & 0 & 0 \end{bmatrix}.$$

It is easy to check that this matrix transforms the matrix (1.11.7) to the form $\mathbf{F}_A$.

1.12.3 Eigenvectors Method

Let a matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ and its Jordan canonical form (1.11.12), containing p blocks of the form (1.11.10a), be given. Let the i-th block, corresponding to the eigenvalue s_i, have the dimensions $m_i \times m_i$ ($i = 1,\ldots,p$). The following matrix

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_1 \ \mathbf{T}_2 \ \ldots \ \mathbf{T}_p \end{bmatrix}, \ \ \mathbf{T}_i = \begin{bmatrix} t_{i1} \ t_{i2} \ \ldots \ t_{im_i} \end{bmatrix} \tag{1.12.26}$$

satisfying (1.11.13) is to be computed.

Post-multiplying (1.11.13) by $\mathbf{T}$ we obtain

$$\mathbf{AT} = \mathbf{TJ}_A$$

and after taking into account (1.12.26), (1.11.10) and (1.11.12)

$$\mathbf{AT}_i = \mathbf{T}_i \mathbf{J}_i \quad \text{for} \quad i = 1,\ldots,p,$$

and

$$\begin{gathered} [\mathbf{A} - \mathbf{I}s_i] t_{i1} = 0, \ \ [\mathbf{A} - \mathbf{I}s_i] t_{i2} = t_{i1}, \ \ldots, \ [\mathbf{A} - \mathbf{I}s_i] t_{im_i} = t_{i,m_i-1}, \\ i = 1,\ldots,p. \end{gathered} \tag{1.12.27}$$

For the eigenvalue s_i from the first among the equations in (1.12.27) we compute the column t_{i1}, knowing t_{i1} we compute from the second equation the column t_{i2} and finally from the last equation we compute the column t_{im_i}.

Repeating these computations successively for $i = 1,2,\ldots,p$, we obtain the desired matrix (1.12.26).

Example 1.12.4.
Compute the matrix $\mathbf{T}$ transforming the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 & 0 \\ -1 & 3 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{1.12.28}$$

to its Jordan canonical form

$$\mathbf{J}_A = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{1.12.29}$$

From (1.12.29) it follows that the matrix (1.12.28) has one eigenvalue $s_1 = 2$ of multiplicity 3 and one eigenvalue $s_2 = 1$ of multiplicity 1.

In this case, the matrix (1.12.26) is of the form

$$\mathbf{T} = [\mathbf{T}_1 \;\; \mathbf{T}_2] = [t_{11} \;\; t_{12} \;\; t_{13} \;\; t_{21}].$$

For $i = 1$ the equations (1.12.27) take the form

$$[\mathbf{A} - \mathbf{I}s_1]t_{11} = \begin{bmatrix} -1 & 2 & 1 & 0 \\ -1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} t_{11} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$[\mathbf{A} - \mathbf{I}s_1]t_{12} = \begin{bmatrix} -1 & 2 & 1 & 0 \\ -1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} t_{12} = t_{11},$$

$$[\mathbf{A} - \mathbf{I}s_1]t_{13} = \begin{bmatrix} -1 & 2 & 1 & 0 \\ -1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} t_{13} = t_{12},$$

and for $i = 2$

$$[\mathbf{A} - \mathbf{I}s_2]t_{21} = \begin{bmatrix} 0 & 2 & 1 & 0 \\ -1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} t_{21} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Solving these equations successively, we obtain

$$t_{11} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad t_{12} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad t_{13} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad t_{21} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

and the desired matrix has the form

$$\mathbf{T} = \begin{bmatrix} t_{11} & t_{12} & t_{13} & t_{21} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

If blocks have the form (1.11.10b) considerations are similar.

1.13 Matrices of Simple Structure and Diagonalisation of Matrices

1.13.1 Matrices of Simple Structure

Consider a matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ whose characteristic polynomial has the form

$$\psi(\lambda) = \det\left[\mathbf{I}_n \lambda - \mathbf{A}\right] = \lambda^n + a_{n-1}\lambda^{n-1} + \ldots.. + a_1\lambda + a_0 . \tag{1.13.1}$$

The roots $\lambda_1, \lambda_2, \ldots, \lambda_p$ ($p \le n$) of the equation $\psi(\lambda) = 0$ are called eigenvalues of the matrix **A**, and the set of these eigenvalues is called the spectrum of this matrix.

Definition 1.13.1. We say that an eigenvalue λ_i has algebraic multiplicity n_i, if λ_i is the n_i-fold root of the equation $\psi(\lambda) = 0$, i.e.,

$$\begin{aligned} &\psi(\lambda_i) = \psi'(\lambda_i) = \ldots. = \psi^{(n_i-1)}(\lambda_i) = 0, \\ &\text{but} \quad \psi^{(n_i)}(\lambda_i) \ne 0, \ i = 1, \ldots, p, \\ &\text{where} \quad \psi^{(k)}(\lambda) = \frac{d^k\psi(\lambda)}{d\lambda^k}, \ \text{ i.e.,} \end{aligned} \tag{1.13.2}$$

$$\psi(\lambda) = (\lambda - \lambda_1)^{n_1} (\lambda - \lambda_2)^{n_2} \ldots. (\lambda - \lambda_p)^{n_p} . \tag{1.13.3}$$

We say that an eigenvalue λ_i has geometrical multiplicity m_i if

$$\text{rank}\left[\mathbf{I}_n \lambda_i - \mathbf{A}\right] = n - m_i, \ \ i = 1, \ldots, p . \tag{1.13.4}$$

From the Jordan canonical form of the matrix $\mathbf{A}$, it follows that $n_i > m_i$ for $i = 1,\dots,p$.

Definition 1.13.2. A matrix $\mathbf{A}\in\mathbb{C}^{n\times n}$ for which $n_i = m_i$ for $i = 1,\dots,p$, is called a matrix of simple structure. Otherwise we say that the matrix has a complex structure.

For example, the matrix

$$\mathbf{A} = \begin{bmatrix} 2 & a \\ 0 & 2 \end{bmatrix} \tag{1.13.5}$$

for $a = 0$ is a matrix of simple structure, since $n_1 = m_1 = 2$ and for $a \neq 0$ it is a matrix of complex structure, since $n_1 = 2$, $m_1 = 1$ (rank $[\mathbf{I}_2 2 - \mathbf{A}] = 1$)).

Theorem 1.13.1. The similar matrices $\mathbf{A}\in\mathbb{C}^{n\times n}$ and $\mathbf{B} = \mathbf{PAP}^{-1}$, $\det \mathbf{P} \neq 0$, have eigenvalues of the same algebraic and geometric multiplicities.

Proof. According to Theorem 1.10.1, the similar matrices $\mathbf{A}$ and $\mathbf{B}$ share the same characteristic polynomial, i.e.,

$$\det\left[\mathbf{I}_n\lambda - \mathbf{A}\right] = \det\left[\mathbf{I}_n\lambda - \mathbf{B}\right]. \tag{1.13.6}$$

The equality (1.13.6) implies that the matrices $\mathbf{A}$ and $\mathbf{B}$ have the same eigenvalues of the same algebraic multiplicities.

From the relationship

$$\begin{aligned} &\operatorname{rank}\left[\mathbf{I}_n\lambda_i - \mathbf{B}\right] = \operatorname{rank}\left[\mathbf{P}\left[\mathbf{I}_n\lambda_i - \mathbf{A}\right]\mathbf{P}^{-1}\right] \operatorname{rank}\left[\mathbf{I}_n\lambda_i - \mathbf{A}\right], \\ &\text{for} \quad i = 1, \dots, p, \end{aligned} \tag{1.13.7}$$

it follows that eigenvalues of the matrices $\mathbf{A}$ and $\mathbf{B}$ also have the same geometrical multiplicities.

■

From the Jordan canonical structure and (1.13.4) the following important corollary ensues.

Corollary 1.13.1. Geometrical multiplicity m_i of an eigenvalue λ_i, $i = 1,\dots,p$ of the matrix $\mathbf{A}$ is equal to a number of blocks corresponding to this eigenvalue.

Theorem 1.13.2. A matrix $\mathbf{A}\in\mathbb{C}^{n\times n}$ is of simple structure if and only if all its elementary divisors are of the first degree.

Proof. According to Corollary 1.11.1, the matrix **A** is similar to the diagonal matrix consisting of eigenvalues of this matrix if and only if all its elementary divisors are of the first degree. In this case

$$\operatorname{rank}\left[\mathbf{I}_n\lambda_i - \mathbf{A}\right] = n - n_i \quad \text{for} \quad i = 1,\ldots,p . \tag{1.13.8}$$

In view of this, $m_i = n_i$ for $i = 1,\ldots,p$, and **A** is a matrix of simple structure if and only if all its divisors are of the first degree.

■

Example 1.13.1.
The matrix (1.13.5) is a matrix of simple structure if and only if $a = 0$, since the Smith canonical form of the matrix

$$\left[\mathbf{I}_2 s - \mathbf{A}\right] = \begin{bmatrix} s-2 & -a \\ 0 & s-2 \end{bmatrix}$$

is equal to

$$\left[\mathbf{I}_2 s - \mathbf{A}\right]_s = \begin{bmatrix} s-2 & 0 \\ 0 & s-2 \end{bmatrix}, \quad \text{for} \quad a = 0,$$

$$\left[\mathbf{I}_2 s - \mathbf{A}\right]_s = \begin{bmatrix} 1 & 0 \\ 0 & (s-2)^2 \end{bmatrix}, \quad \text{for} \quad a \neq 0 .$$

For $a = 0$, the matrix (1.13.5) has two elementary divisors of the first degree, and for $a \neq 0$, it has one elementary divisor $(s - 2)^2$. According to Theorem 1.13.2, the matrix (1.13.5) is thus of simple structure if and only if $a = 0$.

1.13.2 Diagonalisation of Matrices of Simple Structure

Theorem 1.13.3. For every matrix $\mathbf{A}\in\mathbb{C}^{n\times n}$ of simple structure there exists a non-singular matrix $\mathbf{P}\in\mathbb{C}^{n\times n}$ such that

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \operatorname{diag}\left[\lambda_1, \lambda_2, \ldots, \lambda_n\right] \tag{1.13.9}$$

where some eigenvalues λ_i, $i = 1,\ldots,p$ can be equal.

Proof. From the fact that **A** is a matrix of simple structure it follows that for every eigenvalue λ_i there are as many corresponding eigenvectors $\mathbf{P}_i$ as the multiplicity of the eigenvalue amounts to

$$\mathbf{A}\mathbf{P}_i = \lambda_i \mathbf{P}_i, \quad \text{for} \quad i = 1, \ldots, n . \tag{1.13.10}$$

The eigenvectors $\mathbf{P}_1,\mathbf{P}_2,\ldots,\mathbf{P}_n$ are linearly independent. Hence the matrix $\mathbf{P} = [\mathbf{P}_1,\mathbf{P}_2,\ldots,\mathbf{P}_n]$ is nonsingular.

From (1.13.10) for $i = 1,\ldots,n$, we have

$$\mathbf{AP} = \mathbf{P}\,\text{diag}\left[\lambda_1, \lambda_2, \ldots, \lambda_n\right]. \tag{1.13.11}$$

Pre-multiplying (1.13.11) by $\mathbf{P}^{-1}$, we obtain (1.13.9). ■

In particular, in the case when $\mathbf{A}$ has distinct eigenvalues $\lambda_1, \lambda_2,\ldots,\lambda_n$, the following important corollary ensues from Theorem 1.13.3.

Corollary 1.13.2. Every matrix $\mathbf{A}\in\mathbb{C}^{n\times n}$ with distinct eigenvalues $\lambda_1, \lambda_2,\ldots,\lambda_n$ can be transformed by similarity to the diagonal form diag $[\lambda_1, \lambda_2,\ldots,\lambda_n]$.

To compute the eigenvectors $\mathbf{P}_1,\mathbf{P}_2,\ldots,\mathbf{P}_n$, we solve the equation

$$\left[\mathbf{I}_n\lambda_i - \mathbf{A}\right]\mathbf{P}_i = 0, \quad \text{for} \quad i = 1, \ldots, n \tag{1.13.12}$$

or taking instead of $\mathbf{P}_i$ any nonzero column of Adj $[\mathbf{I}_n\lambda_i - \mathbf{A}]$.

From definition of the inverse matrix

$$\left[\mathbf{I}_n\lambda - \mathbf{A}\right]^{-1} = \frac{\text{Adj}\left[\mathbf{I}_n\lambda - \mathbf{A}\right]}{\det\left[\mathbf{I}_n\lambda - \mathbf{A}\right]},$$

we have

$$\left[\mathbf{I}_n\lambda - \mathbf{A}\right]\text{Adj}\left[\mathbf{I}_n\lambda - \mathbf{A}\right] = \mathbf{I}_n \det\left[\mathbf{I}_n\lambda - \mathbf{A}\right]. \tag{1.13.13}$$

Substituting $\lambda = \lambda_i$ into (1.13.13) and taking into account that det $[\mathbf{I}_n\lambda_i - \mathbf{A}] = 0$, we obtain

$$\left[\mathbf{I}_n\lambda_i - \mathbf{A}\right]\text{Adj}\left[\mathbf{I}_n\lambda_i - \mathbf{A}\right] = 0, \quad \text{for} \quad i = 1, \ldots, p\,. \tag{1.13.14}$$

From (1.13.14) it follows that every nonzero column of Adj $[\mathbf{I}_n\lambda_i - \mathbf{A}]$ is the eigenvector of the eigenvalue λ_i of the matrix $\mathbf{A}$.

Example 1.13.2.
Compute a matrix $\mathbf{P}$ that transforms the matrix

$$\mathbf{A} = \frac{1}{2}\begin{bmatrix} -3 & -1 & 1 \\ 1 & -5 & -1 \\ 2 & -2 & -4 \end{bmatrix} \tag{1.13.15}$$

to the diagonal form.

The characteristic equation of the matrix (1.13.15)

$$\det\left[\mathbf{I}_n\lambda-\mathbf{A}\right]=\begin{vmatrix}\lambda+\frac{3}{2} & \frac{1}{2} & -\frac{1}{2}\\ -\frac{1}{2} & \lambda+\frac{5}{2} & \frac{1}{2}\\ -1 & 1 & \lambda+2\end{vmatrix}=\lambda_3+6\lambda^2+11\lambda+6=0$$

has three real roots $\lambda_1=-1$, $\lambda_2=-2$, $\lambda_3=-3$. To compute the eigenvectors $\mathbf{P}_1,\mathbf{P}_2,\mathbf{P}_3$, we compute the adjoint (adjugate) matrix

$$\operatorname{Adj}\left[\mathbf{I}_n\lambda-\mathbf{A}\right]=\begin{bmatrix}\lambda^2+\frac{9}{2}\lambda+\frac{9}{2} & -\frac{1}{2}\lambda-\frac{3}{2} & \frac{1}{2}\lambda+\frac{3}{2}\\ \frac{1}{2}\lambda+\frac{1}{2} & \lambda^2+\frac{7}{2}\lambda+\frac{5}{2} & -\frac{1}{2}\lambda-\frac{1}{2}\\ \lambda+2 & -\lambda-2 & \lambda^2+4\lambda+4\end{bmatrix}. \tag{1.13.16}$$

As the eigenvectors $\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$ of the matrix (1.13.15) we take the third column of the adjoint matrix successively for $\lambda_1=-1$, $\lambda_2=-2$, $\lambda_3=-3$. The matrix built from these vectors (after multiplication of the third column for $\lambda_2=-2$ by 2) has the form

$$\mathbf{P}=\left[\mathbf{P}_1,\ \mathbf{P}_2,\ \mathbf{P}_3\right]=\begin{bmatrix}1 & 1 & 0\\ 0 & 1 & 1\\ 1 & 0 & 1\end{bmatrix}$$

and its inverse is

$$\mathbf{P}^{-1}=\frac{1}{2}\begin{bmatrix}1 & -1 & 1\\ 1 & 1 & -1\\ -1 & 1 & 1\end{bmatrix}.$$

Hence

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P}=\frac{1}{2}\begin{bmatrix}1 & -1 & 1\\ 1 & 1 & -1\\ -1 & 1 & 1\end{bmatrix}\frac{1}{2}\begin{bmatrix}-3 & -1 & 1\\ 1 & -5 & -1\\ 2 & -2 & -4\end{bmatrix}\begin{bmatrix}1 & 1 & 0\\ 0 & 1 & 1\\ 1 & 0 & 1\end{bmatrix}=\begin{bmatrix}-1 & 0 & 0\\ 0 & -2 & 0\\ 0 & 0 & -3\end{bmatrix}.$$

Example 1.13.3.
Compute a matrix **P** that reduces the matrix

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 1 & -1 & 1 \end{bmatrix} \tag{1.13.17}$$

to the diagonal form.

The characteristic equation of the matrix (1.13.17)

$$\det\left[\mathbf{I}_3\lambda - \mathbf{A}\right] = \begin{vmatrix} \lambda-2 & 0 & 0 \\ 0 & \lambda-2 & 0 \\ -1 & 1 & \lambda-1 \end{vmatrix} = (\lambda-2)^2(\lambda-1) = 0$$

has one double root $\lambda_1 = 2$ and one root of multiplicity 1, $\lambda_2 = 1$. The matrix (1.13.17) is a matrix of simple structure, since

$$\operatorname{rank}\left[\mathbf{I}_3\lambda_1 - \mathbf{A}\right] = \operatorname{rank}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 1 & 1 \end{bmatrix} = 1 .$$

Thus using similarity transformation the matrix (1.13.17) can be reduced to the diagonal form.

From the equation

$$\left[\mathbf{I}_3\lambda_1 - \mathbf{A}\right]\mathbf{P}_i = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 1 & 1 \end{bmatrix}\mathbf{P}_i = 0 \qquad (i = 1, 2)$$

it follows that as the eigenvectors $\mathbf{P}_1$ and $\mathbf{P}_2$ we can adopt

$$\mathbf{P}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{P}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

Solving the equation

$$\left[\mathbf{I}_3\lambda_2 - \mathbf{A}\right]\mathbf{P}_3 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ -1 & 1 & 0 \end{bmatrix}\mathbf{P}_3 = 0 ,$$

we obtain $\mathbf{P}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. Thus

$$\mathbf{P} = \left[\mathbf{P}_1,\ \mathbf{P}_2,\ \mathbf{P}_3\right] = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

It is easy to verify that

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

1.13.3 Diagonalisation of an Arbitrary Square Matrix by the Use of a Matrix with Variable Elements

Let a square matrix **A** and a diagonal matrix **Λ** of the same dimension be given. We will show that an arbitrary matrix **A** can be transformed to the diagonal form **Λ** by use of a transformation of a matrix with variable elements.

Theorem 1.13.4. For an arbitrary matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ and a given diagonal matrix $\mathbf{\Lambda} \in \mathbb{C}^{n \times n}$ there exists a nonsingular matrix

$$\mathbf{T} = \mathbf{T}(t) = e^{(\mathbf{A}-\mathbf{\Lambda})t} \tag{1.13.18}$$

such that

$$(\mathbf{A}\mathbf{T} - \dot{\mathbf{T}})\mathbf{T}^{-1} = \mathbf{\Lambda}\,. \tag{1.13.19}$$

Proof. From (1.13.18) it follows that this matrix is nonsingular for arbitrary matrices **A** and **Λ**. Taking into account that

$$\dot{\mathbf{T}} = (\mathbf{A} - \mathbf{\Lambda})e^{(\mathbf{A}-\mathbf{\Lambda})t} = (\mathbf{A} - \mathbf{\Lambda})\mathbf{T}\,,$$

we obtain

$$\left(\mathbf{A}\mathbf{T} - \dot{\mathbf{T}}\right)\mathbf{T}^{-1} = \left(\mathbf{A}\mathbf{T} - \left(\mathbf{A} - \mathbf{\Lambda}\right)\mathbf{T}\right)\mathbf{T}^{-1} = \mathbf{\Lambda}\,.$$

■

Example 1.13.4.
Compute a matrix $\mathbf{T}$ that transforms the matrix

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

to the diagonal form

$$\mathbf{\Lambda} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

Note that the given matrix $\mathbf{A}$ is of the Jordan canonical form and one cannot transform it to a diagonal form using similarity transformation (with a matrix $\mathbf{P}$ with constant elements) since it does not have a simple structure.

Using (1.13.18) we compute

$$\mathbf{T} = e^{(\mathbf{A}-\mathbf{\Lambda})t} = \exp\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} t = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}.$$

Taking into account that

$$\mathbf{T}^{-1} = \begin{bmatrix} 1 & -t \\ 0 & 1 \end{bmatrix}, \quad \dot{\mathbf{T}} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

it is easy to check that

$$\mathbf{\Lambda} = \left(\mathbf{AT} - \dot{\mathbf{T}}\right)\mathbf{T}^{-1} = \left\{\begin{bmatrix} 2 & 2t+1 \\ 0 & 2 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right\}\begin{bmatrix} 1 & -t \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

These considerations can be generalised into a matrix $\mathbf{A}(t)$ whose elements depend on time t.

We will show that a square matrix $\mathbf{A}(t)$ of dimension $n \times n$ with elements being continuous functions of time t can be transformed to the diagonal form

$$\mathbf{\Lambda}(t) = \operatorname{diag}\left[\lambda_1(t), \lambda_2(t), ..., \lambda_n(t)\right]. \tag{1.13.20}$$

Let matrix $\phi(t)$ be the solution of the matrix differential equation

$$\dot{\phi}(t) = \mathbf{A}\phi(t), \tag{1.13.21}$$

Satisfying, for example, the initial condition $\phi(0) = \mathbf{I}_n$.

Let

$$\mathbf{T}(t) = \phi(t)e^{\int_0^t \mathbf{\Lambda}(\tau)d\tau} . \tag{1.13.22}$$

It is known that the matrix (1.13.22) is a nonsingular matrix for every $t \geq 0$.

We will show that the matrix (1.13.22) satisfies the equation

$$\dot{\mathbf{T}}(t) = \mathbf{A}(t)\mathbf{T}(t) - \mathbf{T}(t)\mathbf{\Lambda}(t) . \tag{1.13.23}$$

Differentiating the matrix (1.13.22) with respect to t and taking into account (1.13.21), we obtain

$$\begin{aligned}\dot{\mathbf{T}}(t) &= \dot{\phi}(t)e^{-\int_0^t \mathbf{\Lambda}(t)dt} - \phi(t)e^{-\int_0^t \mathbf{\Lambda}(t)dt}\lambda(t) = \\ &= \mathbf{A}(t)\phi(t)e^{-\int_0^t \mathbf{\Lambda}(t)dt} - \phi(t)e^{-\int_0^t \mathbf{\Lambda}(t)dt}\mathbf{\Lambda}(t) = \mathbf{A}(t)\mathbf{T}(t) - \mathbf{T}(t)\mathbf{\Lambda}(t).\end{aligned}$$

From (1.13.23), we obtain

$$\mathbf{\Lambda}(t) = \mathbf{T}^{-1}(t)\left[\mathbf{A}(t)\mathbf{T}(t) - \dot{\mathbf{T}}(t)\right] = \mathrm{diag}\left[\lambda_1(t), \lambda_2(t), ..., \lambda_n(t)\right].$$

Thus the desired matrix is given by the relationship (1.13.22), where the matrix $\phi(t)$ is a solution to the equation (1.13.21).

1.14 Simple Matrices and Cyclic Matrices

1.14.1 Simple Polynomial Matrices

Consider a polynomial matrix $\mathbf{A}(s) \in \mathbb{C}^{m\times n}[s]$ of rank $r \leq \min(m,n)$.

Definition 1.14.1. A polynomial matrix $\mathbf{A}(s) \in \mathbb{C}^{m\times n}$ of rank r is called a simple one if and only if it has only one invariant polynomial distinct from 1.

Taking (1.8.4) into account, we can equivalently define a simple matrix as a polynomial matrix satisfying the conditions

$$D_1(s) = D_2(s) = ... = D_{r-1}(s) = 1 \quad \text{and} \quad D_r(s) = i_r(s) , \tag{1.14.1}$$

where $D_k(s)$, $k = 1,\dots,r$ is a greatest common divisor of all minors of size k of the matrix $\mathbf{A}(s)$.

Thus the Smith canonical form of the simple matrix $\mathbf{A}(s)$ is equal to

$$\mathbf{A}_S(s) = \begin{cases} \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & i_r(s) & 0 & \cdots & 0 \end{bmatrix} & \text{for } n > m = r \\ \operatorname{diag}[1, \cdots, 1, i_r(s)] & \text{for } n = m = r \\ \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & i_r(s) \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix} & \text{for } m > n = r \end{cases} . \tag{1.14.2}$$

Theorem 1.14.1. A polynomial matrix $\mathbf{A}(s)\in\mathbb{C}^{m\times n}[s]$ of rank r is simple if and only if

$$\operatorname{rank} \mathbf{A}\left(s_i^0\right) = r - 1 \quad \text{for} \quad s_i^0 \in \sigma_A , \tag{1.14.3}$$

where σ_A is the set of zeros of the matrix $\mathbf{A}(s)$.

Proof. The normal rank of the matrix $\mathbf{A}(s)$ and of its Smith canonical form $\mathbf{A}_S(s)$ is the same, i.e., rank $\mathbf{A}(s)$ = rank $\mathbf{A}_S(s) = r$. From (1.14.2) it follows that the defect of the rank of the matrix $\mathbf{A}(s)$ is equal to 1 if and only if s is a zero of this matrix.

■

From Definition 1.14.1 one obtains the following important corollary.

Corollary 1.14.1. A polynomial matrix $\mathbf{A}(s)$ is simple if and only if only one elementary divisor corresponds to each zero.

Example 1.14.1.
In Example 1.8.1 it was shown that to the polynomial matrix

$$\mathbf{A}(s) = \begin{bmatrix} (s+2)^2 & (s+2)(s+3) & s+2 \\ (s+2)(s+3) & (s+2)^2 & s+3 \end{bmatrix} \tag{1.14.4}$$

the Smith canonical form

$$\mathbf{A}_S(s) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & (s+2)(s+2.5) & 0 \end{bmatrix} \tag{1.14.5}$$

corresponds.

From (1.14.5) it follows that $i_1(s) = 1$, $i_2(s) = (s+2)(s+2.5)$ and thus the matrix (1.14.4) is simple.

It is easy to check that the matrix (1.14.4) loses its full rank equal to 2 for zeros $s_1 = -2$ and $s_2 = -2.5$, since

$$\mathbf{A}(-2) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{A}(-2.5) = \begin{bmatrix} 0.25 & -0.25 & -0.5 \\ -0.25 & 0.25 & 0.5 \end{bmatrix}.$$

We obtain the same result from the matrix (1.14.5).

1.14.2 Cyclic Matrices

Consider a matrix $\mathbf{A} \in \mathbb{C}^{n \times m}$.

Definition 1.14.2. A matrix $\mathbf{A} \in \mathbb{C}^{n \times m}$ is called cyclic if and only if the polynomial matrix $[\mathbf{I}_n s - \mathbf{A}]$ corresponding to it is simple.

Consider the following matrices

$$\mathbf{F} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-2} & -a_{n-1} \end{bmatrix}, \quad \overline{\mathbf{F}} = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \ddots & 0 & -a_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{bmatrix}, \tag{1.14.16}$$

$$\hat{\mathbf{F}} = \begin{bmatrix} -a_{n-1} & -a_{n-2} & \cdots & -a_1 & -a_0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}, \quad \tilde{\mathbf{F}} = \begin{bmatrix} -a_{n-1} & 1 & 0 & \cdots & 0 \\ -a_{n-2} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_1 & 0 & 0 & \cdots & 1 \\ -a_0 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

We say that the matrices (1.14.6) have Frobenius canonical form.

Expanding the determinant along the row (or column) containing $a_0, a_1, \ldots, a_{n-1}$ it is easy to show that the following equality holds

$$\det\left[\mathbf{I}_n s - \mathbf{F}\right] = \det\left[\mathbf{I}_n s - \overline{\mathbf{F}}\right] = \det\left[\mathbf{I}_n s - \hat{\mathbf{F}}\right] = \det\left[\mathbf{I}_n s - \tilde{\mathbf{F}}\right] \\ = s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0. \tag{1.14.7}$$

Theorem 1.14.2. The matrices (1.14.6) are cyclic for arbitrary values of the coefficients $a_0, a_1, \dots, a_{n-1}$.

Proof. We prove the theorem in detail only for the matrix $\mathbf{F}$, since in other cases the proof is similar.

After deleting the first column and the n-th row from the matrix

$$[\mathbf{I}_n s - \mathbf{F}] = \begin{bmatrix} s & -1 & 0 & \cdots & 0 & 0 \\ 0 & s & -1 & \ddots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & s & -1 \\ a_0 & a_1 & a_2 & \cdots & a_{n-2} & a_{n-1} \end{bmatrix}, \tag{1.14.8}$$

we obtain the minor M_{n1} equal to $(-1)^{n-1}$. Thus the greatest common divisor $D_{n-1}(s)$ of the all minors of degree $n-1$ of the matrix (1.14.8) is equal to 1, i.e., $D_{n-1}(s) = 1$. The condition (1.14.1) is thus satisfied and the matrix $\mathbf{F}$ is cyclic.

■

Theorem 1.14.3. A matrix $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{n\times n}$ is cyclic if the following conditions are satisfied

$$a_{ij} \begin{cases} = 0 & \text{for } j > i+1 \\ \neq 0 & \text{for } j = i+1 \end{cases}, \quad i, j = 1, \dots, n, \tag{1.14.9a}$$

$$a_{ij} \begin{cases} = 0 & \text{for } i > j+1 \\ \neq 0 & \text{for } i = j+1 \end{cases}, \quad i, j = 1, \dots, n. \tag{1.14.9b}$$

Proof. If the conditions (1.14.9a) are satisfied then after deleting the first column and the n-th row from the matrix

$$[\mathbf{I}_n s - \mathbf{A}] = \begin{bmatrix} s-a_{11} & -a_{12} & 0 & \cdots & 0 & 0 \\ a_{21} & s-a_{22} & -a_{23} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -a_{n-1,1} & -a_{n-1,2} & -a_{n-2,3} & \dots & s-a_{n-1,n-1} & -a_{n-1,n} \\ -a_{n1} & -a_{n2} & -a_{n3} & \cdots & -a_{n,n-1} & s-a_{nn} \end{bmatrix}, \tag{1.14.10}$$

we obtain the minor M_{n1} equal to $M_{n1}=(-1)^{n-1}a_{12}a_{23}\dots a_{n-1,n} \neq 0$. Thus $D_{n-1}(s) = 1$ and the condition (1.14.1) is satisfied. In the case of (1.14.9b) the proof is similar.

■

Example 1.14.2.
Determine the conditions under which the matrix

$$\mathbf{A}_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \tag{1.14.11}$$

is or is not a cyclic matrix.

If $a_{21} \neq 0$, then carrying out the elementary operations: $L[1+2\times 1/a_{21}(s - a_{11})]$, $L[2\times(-a_{21})]$, $L[1,2]$ and $L[2\times a_{21}]$ on the matrix

$$[\mathbf{I}_2 s - \mathbf{A}_2] = \begin{bmatrix} s - a_{11} & -a_{12} \\ -a_{21} & s - a_{22} \end{bmatrix},$$

we obtain its Smith canonical form, which is equal to

$$\begin{gathered}\left[\mathbf{I}_2 s - \mathbf{A}_2\right] = \begin{bmatrix} 1 & 0 \\ 1 & \varphi(s) \end{bmatrix}, \\ \varphi(s) = \det\left[\mathbf{I}_2 s - \mathbf{A}\right] == s^2 - \left(a_{11} + a_{22}\right)s + a_{11}a_{22} - a_{12}a_{21}.\end{gathered} \tag{1.14.12}$$

From (1.14.12) it follows that for $a_{21} \neq 0$, the matrix (1.14.11) is cyclic for any values of other elements.

We obtain a similar result for $a_{12} \neq 0$.

It is easy to check that for $a_{12} = a_{21} = 0$ the diagonal matrix

$$\mathbf{A}_2 = \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix},$$

is cyclic if and only if $a_{11} \neq a_{22}$.

Theorem 1.14.4. A matrix $\mathbf{A}\in\mathbb{C}^{n\times n}$ is cyclic if and only if only one Jordan block corresponds to it every distinct eigenvalue, i.e.,

$$\mathbf{J}_A = \begin{bmatrix} \mathbf{J}_{(s_1,n_1)} & 0 & \dots & 0 \\ 0 & \mathbf{J}_{(s_2,n_2)} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{J}_{(s_p,n_p)} \end{bmatrix} \in \mathbb{C}^{n\times n} \tag{1.14.13a}$$

where

$$\mathbf{J}(s_k, n_k) = \begin{bmatrix} s_k & 1 & 0 & \dots & 0 & 0 \\ 0 & s_k & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & s_k & 1 \\ 0 & 0 & 0 & \dots & 0 & s_k \end{bmatrix} \in \mathbb{C}^{n_k \times n_k} \tag{1.14.13b}$$

or

$$\mathbf{J}(s_k, n_k) = \begin{bmatrix} s_k & 0 & \dots & 0 & 0 & 0 \\ 1 & s_k & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & s_k & 0 \\ 0 & 0 & \dots & 0 & 1 & s_k \end{bmatrix} \in \mathbb{C}^{n_k \times n_k}, \quad k = 1, \dots, p \cdot$$

Proof. The polynomial matrix

$$\mathbf{I}_n s - \mathbf{J}_A = \operatorname{diag}[\mathbf{I}_{n_1} s - \mathbf{J}(s_1, n_1), \dots, \mathbf{I}_{n_p} s - \mathbf{J}(s_p, n_p)] \tag{1.14.14}$$

is simple since

$$\operatorname{rank}\left[\mathbf{I}_n s - \mathbf{J}_A\right]_{|s=s_k} = n - 1 \quad \text{for} \quad k = 1, \dots, p\,. \tag{1.14.15}$$

By virtue of Theorem 1.14.1 and Definition 1.14.2 the matrices (1.14.3) and **A** are cyclic. If at least two blocks correspond to one eigenvalue s_k then defect of the rank of the matrix (1.14.14) is greater than 1 and the matrix **A** is not cyclic.

■

Example 1.14.3.
From Theorem 1.14.4 it follows that the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & a \end{bmatrix} \tag{1.14.16}$$

is a cyclic one for $a \neq 1$. However, it is not cyclic for $a = 1$, since two Jordan blocks correspond to its eigenvalue that is equal to 1

$$\mathbf{J}(1,2) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{J}(1,1) = [1]\,.$$

From Theorem 1.14.4 for $\mathbf{J}(s_k, n_k) = a_k$, $n_k = 1$, $k = 1,\dots,n$ one obtains the following important corollary.

Corollary 1.14.1. The diagonal matrix

$$\mathbf{A} = \operatorname{diag}\left[a_1, a_2, ..., a_n\right] \in \mathbb{C}^{n\times n} \tag{1.14.17}$$

is cyclic if and only if $a_i \neq a_j$ for $i \neq j$.

Theorem 1.14.5. Let $\lambda_1, \lambda_2, ..., \lambda_p$ be the eigenvalues of multiplicities $n_1, n_2, ..., n_p$, respectively, of the matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$. This matrix is cyclic if and only if

$$\operatorname{rank}\left[\mathbf{I}_n\lambda_i - \mathbf{A}\right]^{n_i} = \operatorname{rank}\left[\mathbf{I}_n\lambda_i - \mathbf{A}\right]^{n_i+1} \quad \text{for } i = 1, ..., p. \tag{1.14.18}$$

Proof. It is known that similarity transformation does not change the rank of the matrix

$$\operatorname{rank}\left[\mathbf{I}_n\lambda_i - \mathbf{A}\right]^{n_i} = \operatorname{rank}\left[\mathbf{I}_n\lambda_i - \mathbf{J}_A\right]^{n_i} \quad \text{for } i = 1, ..., p, \tag{1.14.19}$$

where $\mathbf{J}_A$ is a Jordan canonical form of the matrix $\mathbf{A}$.

Taking into account (1.11.10) it is easy to verify that

$$\left[\mathbf{I}_{n1}\lambda_i - \mathbf{J}(\lambda_i, n_i)\right]^{n_i} = 0 \quad \text{for } i = 1, ..., p\ . \tag{1.14.20}$$

From the Jordan canonical form $\mathbf{J}_A$ of the matrix $\mathbf{A}$ and (1.14.20) it follows that only one block corresponds to every eigenvalue λ_i if and only if the condition (1.14.18) is satisfied. Thus by virtue of Theorem 1.14.4, the matrix $\mathbf{A}$ is cyclic if and only if the condition (11.14.8) is satisfied.

■

Example 1.14.4.

The matrix (1.14.16) for $a \neq 1$ has only one eigenvalue $\lambda_1 = 1$ of multiplicity $n_1 = 2$ and one eigenvalue $\lambda_2 = a$ of multiplicity 1.

It is easy to check that

$$\left[\mathbf{I}_3\lambda_1 - \mathbf{A}\right]^2 = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1-a \end{bmatrix}^2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & (1-a)^2 \end{bmatrix},$$

$$\left[\mathbf{I}_3\lambda_1 - \mathbf{A}\right]^3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & (1-a)^3 \end{bmatrix},$$

$$\operatorname{rank}\left[\mathbf{I}_3\lambda_1 - \mathbf{A}\right]^2 = \operatorname{rank}\left[\mathbf{I}_3\lambda_1 - \mathbf{A}\right]^3 = \begin{cases} 1 & \text{for } a \neq 1 \\ 0 & \text{for } a = 1 \end{cases}$$

and

$$[\mathbf{I}_3\lambda_2 - \mathbf{A}] = \begin{bmatrix} a-1 & -1 & 0 \\ 0 & a-1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad [\mathbf{I}_3\lambda_2 - \mathbf{A}]^2 = \begin{bmatrix} (a-1)^2 & -2(a-1) & 0 \\ 0 & (a-1)^2 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\operatorname{rank}[\mathbf{I}_3\lambda_2 - \mathbf{A}] = \operatorname{rank}[\mathbf{I}_3\lambda_2 - \mathbf{A}]^2 = \begin{cases} 2 & \text{for } a \neq 1 \\ 0 & \text{for } a = 1 \end{cases}.$$

Thus the condition (1.14.18) is satisfied and the matrix is cyclic if and only if $a \neq 1$.

Theorem 1.14.6. A matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$ can be transformed by similarity to the Frobenius canonical form (1.14.6) or to the Jordan canonical form (1.14.13) if and only if the matrix $\mathbf{A}$ is a cyclic one.

Proof. It is known that there exist nonsingular matrices $\mathbf{P}_1$ and $\mathbf{P}_2$ of similarity transformation such that

$$\mathbf{A}_F = \mathbf{P}_1\mathbf{A}\mathbf{P}_1^{-1} \text{ such that } \mathbf{J}_A = \mathbf{P}_2\mathbf{A}\mathbf{P}_2^{-1} \tag{1.14.21}$$

if and only if the polynomial matrices $[\mathbf{I}_n s - \mathbf{A}]$, $[\mathbf{I}_n s - \mathbf{A}_F]$ and $[\mathbf{I}_n s - \mathbf{J}_A]$ are equivalent, i.e., they have the same invariant polynomials. This takes place if and only if the matrix $\mathbf{A}$ is cyclic. The sufficiency follows immediately by virtue of Theorems 1.11.1 and 1.11.2.

■

Example 1.14.5.
Consider the matrix (1.14.16). This matrix for $a \neq 1$ is cyclic and can be transformed by similarity into the Frobenius canonical form $\mathbf{A}_F$ that is equal to

$$\mathbf{A}_F = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ a & -2a-1 & 2+a \end{bmatrix}, \tag{1.14.22}$$

since

$$\det[\mathbf{I}_3 s - \mathbf{A}] = \begin{bmatrix} s-1 & -1 & 0 \\ 0 & s-1 & 0 \\ 0 & 0 & s-a \end{bmatrix}$$
$$= (s-1)^2(s-a) = s^3 - (2+a)s^2 + (2a+1)s - a.$$

For $a = 1$, the matrix (1.14.16) has the Jordan canonical form with two blocks corresponding to an eigenvalue equal to 1 and is not cyclic. The matrix (1.14.16) for $a = 1$ cannot be transformed by similarity into the Frobenius canonical form.

1.15 Pairs of Polynomial Matrices

1.15.1 Greatest Common Divisors and Lowest Common Multiplicities of Polynomial Matrices

Let $\mathbb{C}^{m\times n}[s]$ be the set of $m\times n$ polynomial matrices with complex coefficients in the variable s.

Definition 1.15.1. A matrix $\mathbf{B}(s)\in\mathbb{C}^{m\times q}[s]$ is called a left divisor (LD) of the matrix $\mathbf{A}(s)\in\mathbb{C}^{m\times l}[s]$ if and only if there exists a matrix $\mathbf{C}(s)\in\mathbb{C}^{q\times l}[s]$ such that

$$\mathbf{A}(s) = \mathbf{B}(s)\mathbf{C}(s)\,. \tag{1.15.1}$$

A matrix $\mathbf{C}(s)\in\mathbb{C}^{m\times l}[s]$ is called a right divisor (RD) of $\mathbf{A}(s)\in\mathbb{C}^{m\times l}[s]$ if and only if there exists a matrix $\mathbf{B}(s)\in\mathbb{C}^{m\times q}[s]$ such that (1.15.1) holds.

Definition 1.15.2. A matrix $\mathbf{A}(s)\in\mathbb{C}^{q\times l}[s]$ is called a right multiplicity (RM) of a matrix $\mathbf{B}(s)\in\mathbb{C}^{m\times q}[s]$ if and only if there exists a matrix $\mathbf{C}(s)\in\mathbb{C}^{q\times n}[s]$ such that (1.15.1) holds.

A matrix $\mathbf{A}(s)\in\mathbb{C}^{m\times l}[s]$ is called a left multiplicity (LM) of a matrix $\mathbf{C}(s)\in\mathbb{C}^{q\times l}[s]$ if and only if there exists a matrix $\mathbf{B}(s)\in\mathbb{C}^{m\times l}[s]$ such that (1.15.1) holds.

Consider the two polynomial matrices $\mathbf{A}(s)\in\mathbb{C}^{m\times l}[s]$ and $\mathbf{B}(s)\in\mathbb{C}^{m\times p}[s]$.

Definition 1.15.3. A matrix $\mathbf{L}(s)\in\mathbb{C}^{m\times q}[s]$ is called a left common divisor (LCD) of matrices $\mathbf{A}(s)\in\mathbb{C}^{m\times l}[s]$ and $\mathbf{B}(s)\in\mathbb{C}^{m\times p}[s]$ if and only if there exist matrices $\mathbf{A}_1(s)\in\mathbb{C}^{q\times l}[s]$ and $\mathbf{B}_1(s)\in\mathbb{C}^{q\times p}[s]$ such that

$$\mathbf{A}(s) = \mathbf{L}(s)\mathbf{A}_1(s) \quad \text{and} \quad \mathbf{B}(s) = \mathbf{L}(s)\mathbf{B}_1(s)\,. \tag{1.15.2}$$

A matrix $\mathbf{P}(s)\in\mathbb{C}^{q\times l}[s]$ is called a right common divisor (RCD) of matrices $\mathbf{A}(s)\in\mathbb{C}^{m\times l}[s]$ and $\mathbf{B}(s)\in\mathbb{C}^{p\times l}[s]$ if and only if there exist matrices $\mathbf{A}_2(s)\in\mathbb{C}^{m\times q}[s]$ and $\mathbf{B}_2(s)\in\mathbb{C}^{p\times q}[s]$ such that

$$\mathbf{A}(s) = \mathbf{A}_2(s)\mathbf{P}(s) \quad \text{and} \quad \mathbf{B}(s) = \mathbf{B}_2(s)\mathbf{P}(s)\,. \tag{1.15.3}$$

Definition 1.15.4. A matrix $\mathbf{D}(s)\in\mathbb{C}^{p\times l}[s]$ is called a common left multiplicity (CLM) of matrices $\mathbf{A}(s)\in\mathbb{C}^{m\times l}[s]$ and $\mathbf{B}(s)\in\mathbb{C}^{q\times l}[s]$ if and only if there exist matrices $\mathbf{D}_1(s)\in\mathbb{C}^{m\times m}[s]$ and $\mathbf{D}_2(s)\in\mathbb{C}^{p\times q}[s]$ such that

$$\mathbf{D}(s)=\mathbf{D}_1(s)\mathbf{A}(s) \quad \text{and} \quad \mathbf{D}(s)=\mathbf{D}_2(s)\mathbf{B}(s) . \tag{1.15.4}$$

A matrix $\mathbf{F}(s)\in\mathbb{C}^{m\times p}[s]$ is called a common right multiplicity (CRM) of matrices $\mathbf{A}(s)\in\mathbb{C}^{m\times l}[s]$ and $\mathbf{B}(s)\in\mathbb{C}^{m\times q}[s]$ if and only if there exist matrices $\mathbf{F}_1(s)\in\mathbb{C}^{l\times p}[s]$ and $\mathbf{F}_2(s)\in\mathbb{C}^{q\times p}[s]$ such that

$$\mathbf{F}(s)=\mathbf{A}(s)\mathbf{F}_1(s) \quad \text{and} \quad \mathbf{F}(s)=\mathbf{B}(s)\mathbf{F}_2(s) . \tag{1.15.5}$$

Definition 1.15.5. A matrix $\mathbf{L}(s)\in\mathbb{C}^{m\times q}[s]$ is called a greatest common left divisor (GCLD) of matrices $\mathbf{A}(s)\in\mathbb{C}^{m\times l}[s]$ and $\mathbf{B}(s)\in\mathbb{C}^{m\times p}[s]$ if and only if

- the matrix $\mathbf{L}(s)$ is a common left divisor of the matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$;
- the matrix $\mathbf{L}(s)$ is a common right multiplicity of every common left divisor of the matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$, i.e., if $\mathbf{A}(s)=\mathbf{L}_1(s)\mathbf{A}_3(s)$ and $\mathbf{B}(s)=\mathbf{L}_1(s)\mathbf{B}_3(s)$, then $\mathbf{L}(s)=\mathbf{L}_1(s)\mathbf{T}(s)$, where $\mathbf{L}_1(s)$, $\mathbf{A}_3(s)$, $\mathbf{B}_3(s)$ and $\mathbf{T}(s)$ are polynomial matrices of appropriate dimensions.

A matrix $\mathbf{P}(s)\in\mathbb{C}^{q\times l}[s]$ is called a greatest common right divisor (GCRD) of matrices $\mathbf{A}(s)\in\mathbb{C}^{m\times l}[s]$ and $\mathbf{B}(s)\in\mathbb{C}^{p\times l}[s]$ if and only if

1. the matrix $\mathbf{P}(s)$ is a common right divisor of the matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$;
2. the matrix $\mathbf{P}(s)$ is a common left multiplicity of every common right divisor of the matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$, i.e., if $\mathbf{A}(s)=\mathbf{A}_4(s)\mathbf{P}_1(s)$ and $\mathbf{B}(s)=\mathbf{B}_4(s)\mathbf{P}_1(s)$, then $\mathbf{P}(s)=\mathbf{T}(s)\mathbf{P}_1(s)$, where $\mathbf{A}_4(s)$, $\mathbf{P}_1(s)$, $\mathbf{B}_4(s)$ and $\mathbf{T}(s)$ are polynomial matrices of appropriate dimensions.

Definition 1.15.6. A matrix $\mathbf{D}(s)\in\mathbb{C}^{p\times l}[s]$ is called a smallest common left multiplicity (SCLM) of matrices $\mathbf{A}(s)\in\mathbb{C}^{m\times l}[s]$ and $\mathbf{B}(s)\in\mathbb{C}^{q\times l}[s]$ if and only if

1. the matrix $\mathbf{D}(s)$ is a common left multiplicity of the matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$;
2. the matrix $\mathbf{D}(s)$ is a right devisor of every common multiplicity of the matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$, i.e., if $\overline{\mathbf{D}}(s)=\mathbf{D}_3(s)\mathbf{A}(s)$ and $\overline{\mathbf{D}}(s)=\mathbf{D}_4(s)\mathbf{B}(s)$, then $\overline{\mathbf{D}}(s)=\mathbf{T}(s)\mathbf{D}(s)$, where $\overline{\mathbf{D}}(s)$, $\mathbf{D}_3(s)$, $\mathbf{D}_4(s)$ and $\mathbf{T}(s)$ are polynomial matrices of appropriate dimensions.

A matrix $\mathbf{F}(s)\in\mathbb{C}^{m\times p}[s]$ is called a smallest common right multiplicity (SCRM) of matrices $\mathbf{A}(s)\in\mathbb{C}^{m\times l}[s]$ and $\mathbf{B}(s)\in\mathbb{C}^{m\times q}[s]$ if and only if

1. the matrix $\mathbf{F}(s)$ is a common right multiplicity of the matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$;
2. the matrix $\mathbf{F}(s)$ is a left divisor of every common multiplicity of the matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$, i.e., if $\overline{\mathbf{F}}(s)=\mathbf{A}(s)\mathbf{F}_3(s)$ and $\overline{\mathbf{F}}(s)=\mathbf{B}(s)\mathbf{F}_4(s)$, then $\overline{\mathbf{F}}(s)=\mathbf{F}(s)\mathbf{T}(s)$, where $\overline{\mathbf{F}}(s)$, $\mathbf{F}_3(s)$, $\mathbf{F}_4(s)$ and $\mathbf{T}(s)$ are polynomial matrices of appropriate dimensions.

1.15.2 Computation of Greatest Common Divisors of a Polynomial Matrix

Problem 1.15.1. Given $\mathbf{C}\in\mathbb{C}^{l\times m}[s]$, $\mathbf{L}\in\mathbb{C}^{l\times l}[s]$ a matrix $\mathbf{C}_1$ is to be computed such that

$$\mathbf{C} = \mathbf{L}\mathbf{C}_1, \tag{1.15.6}$$

where $\mathbf{L}$ is a lower triangular matrix and rank $\mathbf{L} \geq$ rank $\mathbf{C}$.

Solution. Assume that the matrix $\mathbf{L}$ of rank r has the form

$$\mathbf{L} = \begin{bmatrix} g_{11} & 0 & \dots & 0 & 0 & \dots & 0 \\ g_{21} & g_{22} & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ g_{r1} & g_{r2} & \dots & g_{rr} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ g_{l1} & g_{l2} & \dots & g_{lr} & 0 & \dots & 0 \end{bmatrix} \tag{1.15.7}$$

and the matrix $\mathbf{C}_1$

$$\mathbf{C}_1 = \begin{bmatrix} x_{11} & \dots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{l1} & \dots & x_{lm} \end{bmatrix}. \tag{1.15.8}$$

The equality (1.15.6) can be written in the form

$$\begin{bmatrix} c_{11} & \dots & c_{1m} \\ \vdots & \ddots & \vdots \\ c_{l1} & \dots & c_{lm} \end{bmatrix} = \begin{bmatrix} g_{11} & 0 & \dots & 0 & 0 & \dots & 0 \\ g_{21} & g_{22} & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ g_{r1} & g_{r2} & \dots & g_{rr} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ g_{l1} & g_{l2} & \dots & g_{lr} & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} x_{11} & \dots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{l1} & \dots & x_{lm} \end{bmatrix}. \tag{1.15.9}$$

Carrying out the multiplication and comparing appropriate elements from the equality (1.15.9), we obtain

$$c_{1j} = g_{11}x_{1j} \quad \text{i.e.,} \quad x_{1j} = \frac{c_{1j}}{g_{11}}, \quad j = 1,\dots,m$$

and

$$c_{2j} = g_{21}x_{1j} + g_{22}x_{2j}, \quad x_{2j} = \frac{1}{g_{22}}\left(c_{2j} - g_{21}x_{1j}\right).$$

Thus in the general case for $i \le r$ we obtain

$$x_{ij} = \frac{1}{g_{ii}}\left(c_{ij} - \sum_{k=1}^{i-1} g_{ik}x_{kj}\right). \tag{1.15.10}$$

Entries of rows of the matrix $\mathbf{C}_1$ with indices (i, j), $i = r+1,...,l$, $j = 1,...,m$ can be chosen arbitrarily.

Example 1.15.1.
Given the matrices

$$\mathbf{C} = \begin{bmatrix} 1+s & 1+s & 1-s^2 \\ 1+s & 1-s & 1 \\ 2 & 0 & 2-s \end{bmatrix}, \qquad \mathbf{L} = \begin{bmatrix} 1+s & 0 & 0 \\ 1 & s & 0 \\ 1 & 1 & 0 \end{bmatrix},$$

one has to compute a matrix $\mathbf{C}_1$ that satisfies (1.15.6). In this case, rank $\mathbf{L} = 2$. According to (1.15.10), to compute x_{1j}, we divide the first row of the matrix $\mathbf{C}$ by $g_{11} = 1+s$, and then we subtract the first row of the matrix $\mathbf{C}_1$ from the second row of the matrix $\mathbf{C}$ and we divide the result by s. We thus obtain:

$$\mathbf{C}_1 = \begin{bmatrix} 1 & 1 & 1-s \\ 1 & -1 & 1 \\ 2 & 0 & 2-s \end{bmatrix}.$$

Example 1.15.2.
Given $\mathbf{C}\in\mathbb{C}^{l\times m}[s]$, $\mathbf{P}\in\mathbb{C}^{m\times m}[s]$, one has to compute a matrix $\mathbf{C}_2$ such that

$$\mathbf{C} = \mathbf{C}_2\mathbf{P}, \tag{1.15.11}$$

where $\mathbf{P}$ is an upper triangular matrix and rank $\mathbf{P} \ge$ rank $\mathbf{C}$.

Using the transposition, the solution of the dual problem can be reduced to the solution of 1.15.1

1.15.3 Computation of Greatest Common Divisors and Smallest Common Multiplicities of Polynomial Matrices

Theorem 1.15.1. A matrix $\mathbf{L}(s)\in\mathbb{C}^{m\times m}[s]$ is a GCLD of matrices $\mathbf{A}(s)\in\mathbb{C}^{m\times l}[s]$ and $\mathbf{B}(s)\in\mathbb{C}^{m\times q}[s]$ $(m \le l+q)$ if and only if

$$\begin{bmatrix}\mathbf{A}(s) & \mathbf{B}(s)\end{bmatrix} \quad \text{and} \quad \begin{bmatrix}\mathbf{L}(s) & 0\end{bmatrix} \tag{1.15.12}$$

are right equivalent matrices.

Proof. If the matrices (1.15.12) are right equivalent then there exists a unimodular matrix

$$\mathbf{U}(s) = \begin{bmatrix}\mathbf{U}_{11}(s) & \mathbf{U}_{12}(s) \\ \mathbf{U}_{21}(s) & \mathbf{U}_{22}(s)\end{bmatrix},$$

such that

$$\begin{bmatrix}\mathbf{A}(s) & \mathbf{B}(s)\end{bmatrix}\begin{bmatrix}\mathbf{U}_{11}(s) & \mathbf{U}_{12}(s) \\ \mathbf{U}_{21}(s) & \mathbf{U}_{22}(s)\end{bmatrix} = \begin{bmatrix}\mathbf{L}(s) & 0\end{bmatrix} \tag{1.15.13}$$

and

$$\begin{bmatrix}\mathbf{A}(s) & \mathbf{B}(s)\end{bmatrix} = \begin{bmatrix}\mathbf{L}(s) & 0\end{bmatrix}\begin{bmatrix}\mathbf{V}_{11}(s) & \mathbf{V}_{12}(s) \\ \mathbf{V}_{21}(s) & \mathbf{V}_{22}(s)\end{bmatrix}, \tag{1.15.14}$$

where

$$\begin{bmatrix}\mathbf{V}_{11}(s) & \mathbf{V}_{12}(s) \\ \mathbf{V}_{21}(s) & \mathbf{V}_{22}(s)\end{bmatrix} = \mathbf{U}^{-1}(s).$$

From (1.15.14) we have

$$\mathbf{A}(s) = \mathbf{L}(s)\mathbf{V}_{11}(s) \quad \text{and} \quad \mathbf{B}(s) = \mathbf{L}(s)\mathbf{V}_{12}(s).$$

Thus the matrix $\mathbf{L}(s)$ is a CLD of the matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$. To show that the matrix $\mathbf{L}(s)$ is a GCLD of the matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$, we take into account the relationship

$$\mathbf{A}(s)\mathbf{U}_{11}(s) + \mathbf{B}(s)\mathbf{U}_{21}(s) = \mathbf{L}(s), \tag{1.15.15}$$

which ensues from (1.15.13). Hence it follows that every CLD of the matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$ is also an LD of the matrix $\mathbf{L}(s)$. Thus the matrix $\mathbf{L}(s)$ is a RM of every CLD of the matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$, i.e., a GCLD of these matrices.

Now we will show that if a matrix $\mathbf{L}(s)$ is a GCLD of the matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$, then the matrices in (1.15.12) are right equivalent. By assumption we have

$$\mathbf{A}(s) = \mathbf{L}(s)\mathbf{A}_1(s), \quad \mathbf{B}(s) = \mathbf{L}(s)\mathbf{B}_1(s), \tag{1.15.16}$$

where a GCLD of the matrices $\mathbf{A}_1(s)$ and $\mathbf{B}_1(s)$ is the identity matrix $\mathbf{I}_m$.

From (1.15.16) we have

$$[\mathbf{A}(s) \quad \mathbf{B}(s)] = [\mathbf{L}(s) \quad 0]\begin{bmatrix} \mathbf{A}_1(s) & \mathbf{B}_1(s) \\ \mathbf{N}(s) & \mathbf{M}(s) \end{bmatrix}, \tag{1.15.17}$$

where $\mathbf{N}(s)$ and $\mathbf{M}(s)$ are arbitrary polynomial matrices.

We will show that there exist matrices $\mathbf{N}(s)$ and $\mathbf{M}(s)$ such that the matrix

$$\begin{bmatrix} \mathbf{A}_1(s) & \mathbf{B}_1(s) \\ \mathbf{N}(s) & \mathbf{M}(s) \end{bmatrix} \tag{1.15.18}$$

is a unimodular matrix. A GCLD of the matrices $\mathbf{A}_1(s)$ and $\mathbf{B}_1(s)$ is the identity matrix $\mathbf{I}_m$. In view of this, there exists a unimodular matrix $\mathbf{U}_1(s)$ such that

$$[\mathbf{A}_1(s) \quad \mathbf{B}_1(s)]\mathbf{U}_1(s) = [\mathbf{I}_m \quad \mathbf{0}].$$

The matrix $\mathbf{U}_1^{-1}(s)$ is also a unimodular matrix. Thus from the last relationship we have

$$[\mathbf{A}_1(s) \quad \mathbf{B}_1(s)] = [\mathbf{I}_m \quad \mathbf{0}]\mathbf{U}_1^{-1}(s) = [\mathbf{I}_m \quad \mathbf{0}]\begin{bmatrix} \mathbf{A}_1(s) & \mathbf{B}_1(s) \\ \mathbf{N}(s) & \mathbf{M}(s) \end{bmatrix}.$$

Thus the matrix (1.15.18) is unimodular and from (1.15.17) it follows that the matrices (1.15.12) are right equivalent.

■

Corollary 1.15.1. If a matrix $\mathbf{L}(s)$ is a GCLD of the matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$, then there exist polynomial matrices $\mathbf{U}_{11}$, (s) $\mathbf{U}_{21}(s)$ such that (1.15.15) holds.

The matrix $\mathbf{L}(s)$ can be a lower triangular matrix.

Corollary 1.15.2. If the GCLD of the matrices $\mathbf{A}_1(s)$ and $\mathbf{B}_1(s)$, is equal to $\mathbf{L}(s) = \mathbf{I}$, then there exist polynomial matrices $\mathbf{N}(s)$ and $\mathbf{M}(s)$ such that the square matrix (1.15.18) is a unimodular one.

From (1.15.13) it follows that

$$\mathbf{A}(s)\mathbf{U}_{12}(s) = -\mathbf{B}(s)\mathbf{U}_{22}(s) = \mathbf{F}(s). \tag{1.15.19}$$

Theorem 1.15.2. The matrix $\mathbf{F}(s)$ given by the equality (1.15.19) is a SCRM of the matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$.

Proof. From Definition 1.15.4 and (1.15.19) it follows that the matrix $\mathbf{F}(s)$ is a CRM of the matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$. One has still to show that the matrix $\mathbf{F}(s)$ is a left divisor of every CRM of the matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$. To show this, it suffices to note that the GCRD of the matrices $\mathbf{U}_{12}(s)$, $\mathbf{U}_{22}(s)$ is an identity matrix $\mathbf{I}_{m-1-q}$.

■

To compute a GCLD and SCRM of matrices $\mathbf{A}(s)\in\mathbb{C}^{m\times l}[s]$ and $\mathbf{B}(s)\in\mathbb{C}^{m\times q}[s]$, one can apply the following algorithm.

Algorithm 1.15.1.
Step 1: Write the matrices $\mathbf{A}(s)$, $\mathbf{B}(s)$ and the identity matrices $\mathbf{I}_l$, $\mathbf{I}_q$ as

$$\begin{bmatrix} \mathbf{A}(s) & \mathbf{B}(s) \\ \mathbf{I}_l & 0 \\ 0 & \mathbf{I}_q \end{bmatrix}.$$

Step 2: Carrying out appropriate elementary operations on the columns of the matrix $[\mathbf{A}(s)\ \mathbf{B}(s)]$ reduce it to the form $[\mathbf{L}(s)\ 0]$. Carry out the same elementary operations on the columns of the matrix $\mathbf{I}_{l+q}$. Partition the resulting matrix $\mathbf{U}(s)$ into the submatrices $\mathbf{U}_{11}(s)$, $\mathbf{U}_{12}(s)$, $\mathbf{U}_{21}(s)$, $\mathbf{U}_{22}(s)$ of dimensions corresponding to those of the matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$, i.e.,

$$\begin{bmatrix} \mathbf{A}(s) & \mathbf{B}(s) \\ \mathbf{I}_l & 0 \\ 0 & \mathbf{I}_q \end{bmatrix} \xrightarrow{R} \begin{bmatrix} \mathbf{L}(s) & 0 \\ \mathbf{U}_{11}(s) & \mathbf{U}_{12}(s) \\ \mathbf{U}_{21}(s) & \mathbf{U}_{22}(s) \end{bmatrix}. \tag{1.15.20}$$

Step 3: The GCLD and SCRM we seek are equal to $\mathbf{L}(s)$ in (1.15.20) and $\mathbf{F}(s)$ in (1.15.19), respectively.

Example 1.15.1.
Compute a GCLD and a GCRD of the matrices

$$\mathbf{A}(s) = \begin{bmatrix} s^2-2s & s \\ s-2 & 1 \end{bmatrix}, \quad \mathbf{B}(s) = \begin{bmatrix} s-2 \\ 1 \end{bmatrix}.$$

In this case, $m = l = 2$, $q = 1$. In order to compute $\mathbf{L}(s)$ and $\mathbf{U}(s)$, we write the matrices $\mathbf{A}(s)$, $\mathbf{B}(s)$ and $\mathbf{I}_2$, $\mathbf{I}_1$ as follows

$$\begin{bmatrix} \mathbf{A}(s) & \mathbf{B}(s) \\ \mathbf{I}_2 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} s^2-2s & s & s-2s \\ s-2 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and we perform the following elementary operations

$$\xrightarrow[P[3+2\times(-1)]]{P[1+2\times(2-s)]} \begin{bmatrix} 0 & s & -2 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 2-s & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow[P[2,3]]{P[1,2]} \begin{bmatrix} s & -2 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 2-s \\ 0 & 1 & 0 \end{bmatrix}.$$

Thus we have

$$\mathbf{L}(s) = \begin{bmatrix} s & -2 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{U}(s) = \begin{bmatrix} \mathbf{U}_{11}(s) & \mathbf{U}_{12}(s) \\ \mathbf{U}_{21}(s) & \mathbf{U}_{22}(s) \end{bmatrix} = \left[\begin{array}{cc:c} 0 & 0 & 1 \\ 1 & -1 & 2-s \\ \hdashline 0 & 1 & 0 \end{array}\right].$$

We compute the SCRM of the matrices **A**(s) and **B**(s) using (1.15.19)

$$\mathbf{F}(s) = \mathbf{A}(s)\mathbf{U}_{12}(s) = \begin{bmatrix} s^2-2s & s \\ s-2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 2-s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Theorem 1.15.3. A matrix $\mathbf{P}(s) \in C^{q\times l}[s]$ is the GCRD of matrices $\mathbf{A}(s) \in C^{m\times l}[s]$ and $\mathbf{B}(s) \in C^{p\times l}[s]$ ($m+p \geq l$) if and only if the matrices

$$\begin{bmatrix} \mathbf{A}(s) \\ \mathbf{B}(s) \end{bmatrix} \text{ and } \begin{bmatrix} \mathbf{P}(s) \\ 0 \end{bmatrix} \tag{1.15.21}$$

are left equivalent.

The proof of this theorem is similar to that of Theorem 1.15.1. Carrying out elementary operations on the rows, we make the following transformation

$$\begin{bmatrix} \mathbf{A}(s) & \mathbf{I}_m & 0 \\ \mathbf{B}(s) & 0 & \mathbf{I}_p \end{bmatrix} \xrightarrow{L} \begin{bmatrix} \mathbf{P}(s) & \mathbf{U}'_{11}(s) & \mathbf{U}'_{12}(s) \\ 0 & \mathbf{U}'_{21}(s) & \mathbf{U}'_{22}(s) \end{bmatrix}. \tag{1.15.22}$$

Carrying out elementary operations on the rows of the matrix $\mathbf{I}_{m+p}$ transforming the matrix $\begin{bmatrix}\mathbf{A}(s)\\ \mathbf{B}(s)\end{bmatrix}$ to the form $\begin{bmatrix}\mathbf{P}(s)\\ 0\end{bmatrix}$, we compute the unimodular matrix

$$\mathbf{U}'(s) = \begin{bmatrix}\mathbf{U}'_{11}(s) & \mathbf{U}'_{12}(s)\\ \mathbf{U}'_{21}(s) & \mathbf{U}'_{22}(s)\end{bmatrix}.$$

Corollary 1.15.3. If the matrix $\mathbf{P}(s)$ is a GCRD of the matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$, then there exist polynomial matrices $\mathbf{U}_{11}'(s)$ and $\mathbf{U}_{12}'(s)$ such that the equality

$$\mathbf{U}'_{11}(s)\mathbf{A}(s) + \mathbf{U}'_{12}(s)\mathbf{B}(s) = \mathbf{P}(s) \tag{1.15.23}$$

holds.

Corollary 1.15.4. If a GCRD of the matrices $\mathbf{A}_1(s)$ and $\mathbf{B}_1(s)$ is equal to $\mathbf{P}(s) = \mathbf{I}$, then there exist polynomial matrices $\mathbf{N}'(s)$ and $\mathbf{M}'(s)$ such that the square matrix

$$\begin{bmatrix}\mathbf{A}_1(s) & \mathbf{N}'(s)\\ \mathbf{B}_1(s) & \mathbf{M}'(s)\end{bmatrix} \tag{1.15.24}$$

is a unimodular one.

Theorem 1.15.4. The matrix $\mathbf{D}(s)$ given by

$$\mathbf{D}(s) = \mathbf{U}'_{21}(s)\mathbf{A}(s) = -\mathbf{U}'_{22}(s)\mathbf{B}(s) \tag{1.15.25}$$

is an SCLM of the matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$.

Proof of this theorem is similar to that of Theorem 1.15.2.

An algorithm for computing a GCRD and a SCLM of matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$ is different from Algorithm 1.15.1 only in that instead of the transformation (1.15.20), we carry out the transformation (1.15.22) and instead of elementary operations on columns, we carry out elementary operations on rows. The GCRD we seek is equal to the matrix $\mathbf{P}(s)$, and the SCLM that is equal to the matrix $\mathbf{D}(s)$ is computed from (1.15.25).

Remark 1.15.1.
Greatest common divisors and smallest common multiplicities are computed uniquely up to multiplication by a unimodular matrix. In this sense, they are not unique, therefore we usually put the indefinite article *a* before these notions.

1.15.4 Relatively Prime Polynomial Matrices and the Generalised Bezoute Identity

Definition 1.15.7. Matrices $\mathbf{A}(s)\in\mathbb{C}^{m\times l}[s]$ and $\mathbf{B}(s)\in\mathbb{C}^{m\times q}[s]$ are called relatively left prime (RLP) if and only if only unimodular matrices are their left common divisors.

Matrices $\mathbf{A}(s)\in\mathbb{C}^{m\times l}[s]$ and $\mathbf{B}(s)\in\mathbb{C}^{p\times l}[s]$ are called relatively right prime (RRP) if and only if only unimodular matrices are their right common divisors.

Theorem 1.15.5. Matrices $\mathbf{A}(s)\in\mathbb{C}^{m\times l}[s]$, $\mathbf{B}(s)\in\mathbb{C}^{m\times q}[s]$ are RLP if and only if the matrices

$$\begin{bmatrix}\mathbf{A}(s) & \mathbf{B}(s)\end{bmatrix} \text{ and } \begin{bmatrix}\mathbf{I}_m & 0\end{bmatrix} \tag{1.15.26}$$

are right equivalent.

Matrices $\mathbf{A}(s)\in\mathbb{C}^{m\times l}[s]$, $\mathbf{B}(s)\in\mathbb{C}^{p\times l}[s]$ are RRP if and only if the matrices

$$\begin{bmatrix}\mathbf{A}(s) \\ \mathbf{B}(s)\end{bmatrix} \text{ and } \begin{bmatrix}\mathbf{I}_l \\ 0\end{bmatrix} \tag{1.15.27}$$

are left equivalent.

Proof. If the matrices (1.15.26) are right equivalent then according to Theorem 1.15.1, the GCLD of the matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$ is $\mathbf{I}_m$, i.e., these matrices are RLP.

If the matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$ are RLP, then the GLCD is a unimodular matrix, which by use of elementary operations on the columns can by reduced to the form $[\mathbf{I}_m \ \ 0]$, i.e., the matrices (1.15.26) are right equivalent. The proof of the second part of the theorem is similar.

■

From Corollary 1.15.1 for $\mathbf{L}(s) = \mathbf{I}_m$ and from Corollary 1.15.3 for $\mathbf{P}(s) = \mathbf{I}_1$ we obtain the following.

Corollary 1.15.5. If the matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$ are RLP, then there exist unimodular matrices $\mathbf{U}_{11}(s)$ and $\mathbf{U}_{21}(s)$ such that

$$\mathbf{A}(s)\mathbf{U}_{11}(s) + \mathbf{B}(s)\mathbf{U}_{21}(s) = \mathbf{I}_m \,. \tag{1.15.28}$$

If the matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$ are RRP, then there exist polynomial matrices $\mathbf{U}_{11}'(s)$ and $\mathbf{U}_{12}'(s)$ such that

$$\mathbf{U}_{11}'(s)\mathbf{A}(s) + \mathbf{U}_{12}'(s)\mathbf{B}(s) = \mathbf{I}_l \,. \tag{1.15.29}$$

The matrices $\mathbf{U}_{11}(s)$, $\mathbf{U}_{21}(s)$ and $\mathbf{U}_{11}'(s)$, $\mathbf{U}_{12}'(s)$ can be computed using Algorithm 1.15.1

Example 1.15.2.
Show that the matrices

$$\mathbf{A}(s) = \begin{bmatrix} s^2 & s \\ s+1 & 1 \end{bmatrix}, \quad \mathbf{B}(s) = \begin{bmatrix} s^2 + 2 \\ s \end{bmatrix}$$

are RLP and compute polynomial matrices $\mathbf{U}_{11}(s)$, and $\mathbf{U}_{21}(s)$ for them such that (28) holds.

We will show that the given matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$ have a GCLD equal to $\mathbf{I}_2$. To accomplish this, we write down these matrices and matrix $\mathbf{I}_3$ in the form

$$\begin{bmatrix} \mathbf{A}(s) & \mathbf{B}(s) \\ \mathbf{I}_l & 0 \\ 0 & \mathbf{I}_q \end{bmatrix} = \begin{bmatrix} s^2 & s & s^2+2 \\ s+1 & 1 & s \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and we carry out the following elementary operations

$$\xrightarrow{\substack{P[1+2\times(-s)] \\ P[3+2\times(-s)]}} \left[\begin{array}{ccc} 0 & s & 2 \\ 1 & 1 & 0 \\ \hdashline 1 & 0 & 0 \\ -s & 1 & -s \\ 0 & 0 & 1 \end{array}\right] \xrightarrow{\substack{P[2+1\times(-1)] \\ P\left[2+3\times\left(-\frac{1}{2}s\right)\right]}} \left[\begin{array}{ccc} 0 & 0 & 2 \\ 1 & 0 & 0 \\ \hdashline 1 & -1 & 0 \\ -s & 1+s+\frac{1}{2}s^2 & -s \\ 0 & -\frac{1}{2}s & 1 \end{array}\right] \to$$

$$\xrightarrow{\substack{P\left[3\times\frac{1}{2}\right] \\ P\{2,3\} \\ P[1,2]}} \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \hdashline 0 & 1 & -1 \\ -\frac{1}{2}s & -s & 1+s+\frac{1}{2}s^2 \\ \frac{1}{2} & 0 & -\frac{1}{2}s \end{array}\right].$$

Thus the given matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$ have a GCLD equal to $\mathbf{I}_2$. Thus these matrices are RLP.

From the matrix

$$\mathbf{U}(s) = \left[\begin{array}{cc|c} 0 & 1 & -1 \\ -\frac{1}{2}s & -s & 1+s+\frac{1}{2}s^2 \\ \hline \frac{1}{2} & 0 & -\frac{1}{2}s \end{array}\right],$$

we obtain

$$\mathbf{U}_{11}(s) = \begin{bmatrix} 0 & 1 \\ -\frac{1}{2} & -s \end{bmatrix}, \qquad \mathbf{U}_{21}(s) = \begin{bmatrix} \frac{1}{2} & 0 \end{bmatrix}.$$

It is easy to verify that the matrices **A**(s), **B**(s), $\mathbf{U}_{11}(s)$, $\mathbf{U}_{21}(s)$ satisfy (1.15.28).

1.15.5 Generalised Bezoute Identity

Consider the polynomial RLP matrices $\mathbf{A}(s)\in\mathbb{R}^{m\times n}[s]$, $\mathbf{B}(s)\in\mathbb{R}^{m\times p}[s]$, ($n+p\geq m$).

Theorem 1.15.6. If polynomial matrices $\mathbf{A}(s)\in\mathbb{R}^{m\times n}[s]$ and $\mathbf{B}(s)\in\mathbb{R}^{m\times p}[s]$ are RLP, then there exist polynomial matrices **C**(s), **D**(s), $\mathbf{M}_1(s)$, $\mathbf{M}_2(s)$, $\mathbf{M}_3(s)$ and $\mathbf{M}_4(s)$ of appropriate dimensions such that

$$\begin{bmatrix} \mathbf{A}(s) & \mathbf{B}(s) \\ \mathbf{C}(s) & \mathbf{D}(s) \end{bmatrix} \begin{bmatrix} \mathbf{M}_1(s) & \mathbf{M}_2(s) \\ \mathbf{M}_3(s) & \mathbf{M}_4(s) \end{bmatrix} = \begin{bmatrix} \mathbf{I}_m & 0 \\ 0 & \mathbf{I}_{n+p-m} \end{bmatrix} \tag{1.15.30}$$

and

$$\begin{bmatrix} \mathbf{M}_1(s) & \mathbf{M}_2(s) \\ \mathbf{M}_3(s) & \mathbf{M}_4(s) \end{bmatrix} \begin{bmatrix} \mathbf{A}(s) & \mathbf{B}(s) \\ \mathbf{C}(s) & \mathbf{D}(s) \end{bmatrix} = \begin{bmatrix} \mathbf{I}_m & 0 \\ 0 & \mathbf{I}_{n+p-m} \end{bmatrix}. \tag{1.15.31}$$

Proof. By the assumption that the matrices **A**(s) and **B**(s) are RLP it follows that there exists a unimodular matrix of elementary operations on columns

$$\mathbf{U}(s) = \begin{bmatrix} \mathbf{U}_1(s) & \mathbf{U}_2(s) \\ \mathbf{U}_3(s) & \mathbf{U}_4(s) \end{bmatrix} \in \mathbb{R}^{(n+p)\times(n+p)}[s]$$

such that $\begin{bmatrix} \mathbf{A}(s) & \mathbf{B}(s) \end{bmatrix}\mathbf{U}(s) = \begin{bmatrix} \mathbf{I}_m & 0 \end{bmatrix}$.

Post-multiplying the latter equality by the matrix

$$\mathbf{U}^{-1}(s) = \begin{bmatrix} \mathbf{V}_1(s) & \mathbf{V}_2(s) \\ \mathbf{V}_3(s) & \mathbf{V}_4(s) \end{bmatrix}, \tag{1.15.32}$$

we obtain

$$\begin{bmatrix} \mathbf{A}(s) & \mathbf{B}(s) \end{bmatrix} = \begin{bmatrix} \mathbf{I}_m & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V}_1(s) & \mathbf{V}_2(s) \\ \mathbf{V}_3(s) & \mathbf{V}_4(s) \end{bmatrix}$$

and

$$\begin{bmatrix} \mathbf{A}(s) & \mathbf{B}(s) \end{bmatrix} = \begin{bmatrix} \mathbf{V}_1(s) & \mathbf{V}_2(s) \end{bmatrix}.$$

The matrix (1.15.32) is unimodular and the following equality holds

$$\mathbf{B}^{-1}(s)\mathbf{U}(s) = \begin{bmatrix} \mathbf{A}(s) & \mathbf{B}(s) \\ \mathbf{V}_3(s) & \mathbf{V}_4(s) \end{bmatrix} \begin{bmatrix} \mathbf{U}_1(s) & \mathbf{U}_2(s) \\ \mathbf{U}_3(s) & \mathbf{U}_4(s) \end{bmatrix} = \begin{bmatrix} \mathbf{I}_m & 0 \\ 0 & \mathbf{I}_{n+p-m} \end{bmatrix}.$$

Thus $[\mathbf{C}(s)\ \ \mathbf{D}(s)] = [\mathbf{V}_3(s)\ \ \mathbf{V}_4(s)]$ and $\mathbf{M}_k(s) = \mathbf{U}_k(s)$, for $k = 1,2,3,4$. The identity (1.15.31) follows from the equality $\mathbf{U}(s)\mathbf{U}(s)^{-1} = \mathbf{U}(s)^{-1}\mathbf{U}(s) = \mathbf{I}_{n+p}$.

■

The following dual theorem can be proved in a similar way.

Theorem 1.15.7. If polynomial matrices $\mathbf{A}'(s)\in\mathbb{R}^{m\times n}[s]$ and $\mathbf{B}'(s)\in\mathbb{R}^{p\times n}[s]$ are RRP, then there exist polynomial matrices $\mathbf{C}'(s)$, $\mathbf{D}'(s)$, $\mathbf{N}_1(s)$, $\mathbf{N}_2(s)$, $\mathbf{N}_3(s)$ and $\mathbf{N}_4(s)$ of appropriate dimensions, such that

$$\begin{bmatrix} \mathbf{A}'(s) & \mathbf{C}'(s) \\ \mathbf{B}'(s) & \mathbf{D}'(s) \end{bmatrix} \begin{bmatrix} \mathbf{N}_1(s) & \mathbf{N}_2(s) \\ \mathbf{N}_3(s) & \mathbf{N}_4(s) \end{bmatrix} = \begin{bmatrix} \mathbf{I}_n & 0 \\ 0 & \mathbf{I}_{m+p-n} \end{bmatrix}, \tag{1.15.33}$$

$$\begin{bmatrix} \mathbf{N}_1(s) & \mathbf{N}_2(s) \\ \mathbf{N}_3(s) & \mathbf{N}_4(s) \end{bmatrix} \begin{bmatrix} \mathbf{A}'(s) & \mathbf{C}'(s) \\ \mathbf{B}'(s) & \mathbf{D}'(s) \end{bmatrix} = \begin{bmatrix} \mathbf{I}_n & 0 \\ 0 & \mathbf{I}_{m+p-n} \end{bmatrix}. \tag{1.15.34}$$

1.16 Decomposition of Regular Pencils of Matrices

1.16.1 Strictly Equivalent Pencils

Definition 1.16.1. A pencil $[\mathbf{E}s - \mathbf{A}]$ (or a pair of matrices $(\mathbf{E}, \mathbf{A})$) is called regular if the matrices $\mathbf{E}$ and $\mathbf{A}$ are square and

$$\det[\mathbf{E}s - \mathbf{A}] \neq 0 \text{ for some } s \in \mathbb{C} \tag{1.16.1}$$

Definition 1.16.2. Let $\mathbf{E}_k, \mathbf{A}_k \in \mathbb{C}^{m\times n}$ for $k = 1,2$. The pencils $[\mathbf{E}_1 s - \mathbf{A}_1]$ and $[\mathbf{E}_2 s - \mathbf{A}_2]$ (or the pairs of the matrices $(\mathbf{E}_1, \mathbf{A}_1)$ and $(\mathbf{E}_2, \mathbf{A}_2)$) are called strictly equivalent if there exist nonsingular matrices $\mathbf{P} \in \mathbb{C}^{m\times m}$, $\mathbf{Q} \in \mathbb{C}^{n\times n}$ (with elements independent of the variable s) such that

$$\mathbf{P}\left[\mathbf{E}_1 s - \mathbf{A}_1\right]\mathbf{Q} = \mathbf{E}_2 s - \mathbf{A}_2 . \tag{1.16.2}$$

Let $D_k(s, t)$ $(k = 1, ..., n)$ be the greatest common divisor of the all minors of degree k of the matrix $[\mathbf{E}s - \mathbf{A}t]$. According to (1.8.4) the invariant polynomials of the matrix $[\mathbf{E}s - \mathbf{A}t]$ are uniquely determined by

$$i_k(s,t) = \frac{D_{n-k+1}(s,t)}{D_{n-k}(s,t)} \text{ for } k = 1, 2, ..., r . \tag{1.16.3}$$

Factoring the polynomials (1.16.3) into appropriate polynomials that cannot be factored in a given field, we obtain elementary divisors $e_i(s, t)$ $(i = 1, ..., p)$ of the matrix $[\mathbf{E}s - \mathbf{A}t]$. Substituting $t = 1$ into $e_i(s, t)$, we obtain appropriate elementary divisors $e_i(s) = e_i(s,1)$ of the $[\mathbf{E}s - \mathbf{A}]$. Knowing $e_i(s)$ of the matrix $[\mathbf{E}s - \mathbf{A}]$, we can also compute elementary divisors $e_i(s,t)$ of the $[\mathbf{E}s - \mathbf{A}t]$ using the relationship $e_i(s,t) = t^q e_i(s/t)$, where q is the degree of the polynomial $e_i(s)$.

In this way, we can find all finite elementary divisors of the matrix $[\mathbf{E}s - \mathbf{A}t]$ with exception of elementary divisors of the form t^q. Elementary divisors of the form t^q are called infinite elementary divisors of the matrix $[\mathbf{E}s - \mathbf{A}]$. Infinite elementary divisors appear if and only if $\det \mathbf{E} = 0$.

For instance, bringing the pencil

$$[\mathbf{E}s - \mathbf{A}] = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} s + \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

into the Smith canonical form

$$[\mathbf{E}s - \mathbf{A}]_S = \begin{bmatrix} 1 & 0 \\ 0 & s+1 \end{bmatrix},$$

we assess that this pencil possesses the finite elementary divisor $s + 1$ and the infinitive elementary divisor t, since $e(s) = s+1$, $q = 1$ and $te(s/t) = s+t$.

Consider two square pencils of the same size

$$\left[\mathbf{E}_1 s - \mathbf{A}_1\right] \text{ and } \left[\mathbf{E}_2 s - \mathbf{A}_2\right] \text{ such that } \det \mathbf{E}_1 \neq 0 \text{ and } \det \mathbf{E}_2 \neq 0. \tag{1.16.4}$$

Theorem 1.16.1. If the condition (1.16.4) is satisfied, then the pencils $[\mathbf{E}_1 s - \mathbf{A}_1]$ and $[\mathbf{E}_2 s - \mathbf{A}_2]$ are equivalent if and only if they are strictly equivalent, i.e., unimodular matrices $\mathbf{L}(s)$ and $\mathbf{P}(s)$ in the equation

$$\mathbf{E}_1 s - \mathbf{A}_1 = \mathbf{L}(s)\left[\mathbf{E}_2 s - \mathbf{A}_2\right]\mathbf{P}(s) \tag{1.16.5}$$

can be replaced with matrices $\mathbf{L}$ and $\mathbf{P}$, which are both independent of the variable s,

$$\mathbf{E}_1 s - \mathbf{A}_1 = \mathbf{L}\left[\mathbf{E}_2 s - \mathbf{A}_2\right]\mathbf{P}\,. \tag{1.16.6}$$

Proof. The inverse matrix $\mathbf{M}(s) = \mathbf{L}^{-1}(s)$ of a unimodular matrix $\mathbf{L}(s)$ is also a unimodular matrix. Pre-multiplying (1.16.5) by $\mathbf{M}(s)$, we obtain

$$\mathbf{M}(s)\left[\mathbf{E}_1 s - \mathbf{A}_1\right] = \left[\mathbf{E}_2 s - \mathbf{A}_2\right]\mathbf{P}(s)\,. \tag{1.16.7}$$

Pre-dividing the matrix $\mathbf{M}(s)$ by $[\mathbf{E}_2 s - \mathbf{A}_2]$ and post-dividing the matrix $\mathbf{P}(s)$ by $[\mathbf{E}_2 s - \mathbf{A}_2]$, we obtain

$$\mathbf{M}(s) = \left[\mathbf{E}_2 s - \mathbf{A}_2\right]\mathbf{Q}(s) + \mathbf{M},\quad \mathbf{P}(s) = \mathbf{T}(s)\left[\mathbf{E}_1 s - \mathbf{A}_1\right] + \mathbf{P}, \tag{1.16.8}$$

where $\mathbf{M}$ and $\mathbf{P}$ are matrices independent of the variable s.

Substituting (1.16.8) into (1.16.7), we obtain

$$\left[\mathbf{E}_2 s - \mathbf{A}_2\right]\left[\mathbf{T}(s) - \mathbf{Q}(s)\right]\left[\mathbf{E}_1 s - \mathbf{A}_1\right] = \mathbf{M}\left[\mathbf{E}_1 s - \mathbf{A}_1\right] - \left[\mathbf{E}_2 s - \mathbf{A}_2\right]\mathbf{P}\,. \tag{1.16.9}$$

This equality holds only for $\mathbf{T}(s) = \mathbf{Q}(s)$; otherwise the left-hand side of this equation would be a polynomial matrix of at least second degree, and the right-hand side would be a polynomial matrix of at most first degree.

Taking into account $\mathbf{T}(s) = \mathbf{Q}(s)$ in (1.16.9), we obtain

$$\mathbf{M}\left[\mathbf{E}_1 s - \mathbf{A}_1\right] = \left[\mathbf{E}_2 s - \mathbf{A}_2\right]\mathbf{P}\,. \tag{1.16.10}$$

We will show that det $\mathbf{M} \neq 0$. Pre-dividing the matrix $\mathbf{L}(s)$ by $\mathbf{E}_1 s$ - $\mathbf{A}_1$, we obtain

$$\mathbf{L}(s) = \left[\mathbf{E}_1 s - \mathbf{A}_1\right]\mathbf{R}(s) + \mathbf{L}\,, \tag{1.16.11}$$

where $\mathbf{L}$ is independent of the variable s.

Using (1.16.11), (1.16.7) and (1.16.8) successively, we obtain

$$\begin{aligned}\mathbf{I} &= \mathbf{M}(s)\mathbf{L}(s) = \mathbf{M}(s)\left(\left[\mathbf{E}_1 s - \mathbf{A}_1\right]\mathbf{R}(s) + \mathbf{L}\right) = \\ &= \mathbf{M}(s)\left[\mathbf{E}_1 s - \mathbf{A}_1\right]\mathbf{R}(s) + \mathbf{M}(s)\mathbf{L} = \\ &= \left[\mathbf{E}_2 s - \mathbf{A}_2\right]\mathbf{P}(s)\mathbf{R}(s) + \left[\mathbf{E}_2 s - \mathbf{A}_2\right]\mathbf{Q}(s)\mathbf{L} + \mathbf{ML} = \\ &= \left[\mathbf{E}_2 s - \mathbf{A}_2\right]\left[\mathbf{P}(s)\mathbf{R}(s) + \mathbf{Q}(s)\mathbf{L}\right] + \mathbf{ML}.\end{aligned} \tag{1.16.12}$$

The right-hand side of (1.16.12) is a matrix of zero degree (equal to an identity matrix) if and only if

$$\mathbf{P}(s)\mathbf{R}(s) + \mathbf{Q}(s)\mathbf{L} = 0\ . \tag{1.16.13}$$

With the above taken into account, from (1.16.12) we have

$$\mathbf{ML} = \mathbf{I}\ .$$

Thus the matrix $\mathbf{M}$ is nonsingular and $\mathbf{L} = \mathbf{M}^{-1}$. Pre-multiplying (1.16.10) by $\mathbf{L} = \mathbf{M}^{-1}$, we obtain (1.16.6).

■

From Theorem 1.16.1 we have the following important corollary.

Corollary 1.16.1. If the condition (1.16.4) is satisfied, then notions of equivalence and strict equivalence of pencils $[\mathbf{E}_1 s - \mathbf{A}_1]$ and $[\mathbf{E}_2 s - \mathbf{A}_2]$ are the same.

From the fact that two polynomial matrices are equivalent if and only if they have the same elementary divisors and from Corollary 1.16.1, the following theorem ensues immediately.

Theorem 1.16.2. If the condition (1.16.4) is satisfied, then pencils $[\mathbf{E}_1 s - \mathbf{A}_1]$ and $[\mathbf{E}_2 s - \mathbf{A}_2]$ are strictly equivalent if and only if they have the same finite elementary divisors.

If the condition (1.16.4) is not satisfied, then the pencils $[\mathbf{E}_1 s - \mathbf{A}_1]$ and $[\mathbf{E}_2 s - \mathbf{A}_2]$ might not be equivalent in spite of the fact that they have the same elementary devisors.

$$\begin{aligned}\left[\mathbf{E}_1 s - \mathbf{A}_1\right] &= \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix} s + \begin{bmatrix} 2 & 1 & 3 \\ 3 & 2 & 5 \\ 3 & 2 & 6 \end{bmatrix}, \\ \left[\mathbf{E}_2 s - \mathbf{A}_2\right] &= \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} s + \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix},\end{aligned} \tag{1.16.14}$$

are not strictly equivalent (since rank $\mathbf{E}_1 = 2$, rank $\mathbf{E}_2 = 1$), although they have the same elementary divisor $s+1$, because they have different infinite elementary divisors. Performing elementary operations on the pencil $[\mathbf{E}_1 s - \mathbf{A}_1 t]$, we obtain assertion of this.

Theorem 1.16.3. Two regular pencils $[\mathbf{E}_1 s - \mathbf{A}_1]$ and $[\mathbf{E}_2 s - \mathbf{A}_2]$ are strictly equivalent if and only if they have the same finite and infinite elementary divisors.

Proof. The strict equivalence of the pencils $[\mathbf{E}_1 s - \mathbf{A}_1]$ and $[\mathbf{E}_2 s - \mathbf{A}_2]$ implies strict equivalence of the pencils $[\mathbf{E}_1 s - \mathbf{A}_1 t]$ and $[\mathbf{E}_2 s - \mathbf{A}_2 t]$. In view of this, the pencils $[\mathbf{E}_1 s - \mathbf{A}_1]$ and $[\mathbf{E}_2 s - \mathbf{A}_2]$ should have the same finite and infinite elementary divisors. Conversely, let two regular pencils $[\mathbf{E}_1 s - \mathbf{A}_1]$ and $[\mathbf{E}_2 s - \mathbf{A}_2]$, which have the same finite and infinite elementary divisors, be given. Let

$$s = a\lambda + b\mu, \quad t = c\lambda + d\mu \quad (ad - bc \neq 0). \tag{1.16.15}$$

Substituting (1.16.15) into $[\mathbf{E}_1 s - \mathbf{A}_1 t]$ and $[\mathbf{E}_2 s - \mathbf{A}_2 t]$ yields

$$\begin{aligned} \left[\mathbf{E}_1 s - \mathbf{A}_1 t\right] &= \left[\mathbf{E}_1 (a\lambda + b\mu) - \mathbf{A}_1 (c\lambda + d\mu)\right] = \left[\overline{\mathbf{E}}_1 \lambda - \overline{\mathbf{A}}_1 \mu\right], \\ \left[\mathbf{E}_2 s - \mathbf{A}_2 t\right] &= \left[\mathbf{E}_2 (a\lambda + b\mu) - \mathbf{A}_2 (c\lambda + d\mu)\right] = \left[\overline{\mathbf{E}}_2 \lambda - \overline{\mathbf{A}}_2 \mu\right], \end{aligned} \tag{1.16.16}$$

where

$$\overline{\mathbf{E}}_1 = a\mathbf{E}_1 - c\mathbf{A}_1, \ \overline{\mathbf{A}}_1 = d\mathbf{A}_1 - b\mathbf{E}_1, \ \overline{\mathbf{E}}_2 = a\mathbf{E}_2 - c\mathbf{A}_2, \ \overline{\mathbf{A}}_2 = d\mathbf{A}_2 - b\mathbf{E}_2. \tag{1.16.17}$$

By assumption of regularity of the pencils $[\mathbf{E}_1 s - \mathbf{A}_1 t]$ and $[\mathbf{E}_2 s - \mathbf{A}_2 t]$, one can choose numbers a and c such that

$$\det \overline{\mathbf{E}}_1 \neq 0 \quad \text{and} \quad \det \overline{\mathbf{E}}_2 \neq 0. \tag{1.16.18}$$

If the condition (1.16.18) is satisfied, then the pencils

$$\left[\overline{\mathbf{E}}_1 \lambda - \overline{\mathbf{A}}_1 \mu\right] \quad \text{and} \quad \left[\overline{\mathbf{E}}_2 \lambda - \overline{\mathbf{A}}_2 \mu\right]$$

are strictly equivalent and this fact implies that the pencils $[\mathbf{E}_1 s - \mathbf{A}_1 t]$ and $[\mathbf{E}_2 s - \mathbf{A}_2 t]$, as well as the starting-point pencils $[\mathbf{E}_1 s - \mathbf{A}_1]$, $[\mathbf{E}_2 s - \mathbf{A}_2]$, are strictly equivalent.

■

1.16.2 Weierstrass Decomposition of Regular Pencils

Assume at the beginning that rectangular matrices **E**, $\mathbf{A}\in\mathbb{C}^{q\times n}$ are such that

$$\operatorname{rank}[\mathbf{E}s-\mathbf{A}]=q \quad \text{for some} \quad s\in\mathbb{C}. \tag{1.16.19}$$

Theorem 1.16.4. If the condition (1.16.19) is satisfied, then there exist full-rank matrices $\mathbf{P}\in\mathbb{C}^{q\times n}$ and $\mathbf{Q}\in\mathbb{C}^{n\times n}$ such that

$$[\mathbf{E}s-\mathbf{A}]=\mathbf{P}\begin{bmatrix}\mathbf{I}_{n_1}s-\mathbf{A}_1 & 0\\ 0 & \mathbf{N}s-\mathbf{I}_{n_2}\end{bmatrix}\mathbf{Q}, \tag{1.16.20}$$

where n_1 is the greatest degree of the polynomial of the variable s, which is a minor of degree q of the matrix $[\mathbf{E}s-\mathbf{A}]$, $n_1+n_2=n$, and **N** is a nilpotent matrix of index v ($\mathbf{N}^v=0$).

Proof. If the condition (1.16.19) is satisfieds then there exists a number $c\in\mathbb{C}$ such that the matrix $\mathbf{F}=[\mathbf{E}c-\mathbf{A}]$ has full row rank. In this case, there exists the inverse of this matrix

$$\mathbf{F}_p=\mathbf{F}^T\left[\mathbf{F}\mathbf{F}^T\right]^{-1}\in\mathbb{C}^{n\times q}, \tag{1.16.21}$$

which satisfies the condition $\mathbf{F}\mathbf{F}_p=\mathbf{I}_q$.

Note that

$$\left[\mathbf{E}s-\mathbf{A}\right]=\left[\mathbf{E}(s-c)+\mathbf{E}c-\mathbf{A}\right]=\left[\mathbf{E}(s-c)+\mathbf{F}\right]=\mathbf{F}\left[\mathbf{F}_p\mathbf{E}(s-c)+\mathbf{I}_n\right]. \tag{1.16.22}$$

According to the considerations in Sect. 4.2.2, there exists a nonsingular matrix $\mathbf{T}\in\mathbb{C}^{n\times n}$ such that

$$\mathbf{F}_p\mathbf{E}=\mathbf{T}\left[\operatorname{diag}\left(\mathbf{J}_1,\mathbf{J}_0\right)\right]\mathbf{T}^{-1}, \tag{1.16.23}$$

where $\mathbf{J}_1=\mathbb{C}^{n_1\times n_1}$ is a nonsingular matrix and $\mathbf{J}_0=\mathbb{C}^{n_2\times n_2}$ is a nilpotent matrix with index v. The matrix **T** can be chosen in such a way that diag ($\mathbf{J}_1$, $\mathbf{J}_0$) has the Jordan canonical form.

Substitution of (1.16.23) into (1.16.22) yields

$$\begin{aligned}[\mathbf{E}s-\mathbf{A}]&=\mathbf{F}\mathbf{T}\left[\operatorname{diag}\,(\mathbf{J}_1(s-c)+\mathbf{I}_{n_1},\ \mathbf{J}_0(s-c)+\mathbf{I}_{n_2})\right]\mathbf{T}^{-1}\\ &=\mathbf{F}\mathbf{T}\operatorname{diag}\left(\mathbf{J}_1,\mathbf{J}_0c-\mathbf{I}_{n_2}\right)\left[\operatorname{diag}\ \left(\mathbf{I}_{n_1}s+\mathbf{J}_1^{-1}(\mathbf{I}_{n_1}-\mathbf{J}_1c),\right.\right.\\ &\left.\left.\left(\mathbf{J}_0c-\mathbf{I}_{n_2}\right)^{-1}\mathbf{J}_0s-\mathbf{I}_{n_2}\right)\mathbf{T}^{-1}=\mathbf{P}\left[\operatorname{diag}\left(\mathbf{I}_{n_1}s-\mathbf{A}_1,\ \mathbf{N}s-\mathbf{I}_{n_2}\right)\right]\mathbf{Q},\end{aligned} \tag{1.16.24}$$

where

$$\mathbf{P} = \mathbf{FT} \text{ diag } \left(\mathbf{J}_1, \mathbf{J}_0 c - \mathbf{I}_{n_2}\right), \quad \mathbf{A}_1 = \mathbf{J}_1^{-1}\left(\mathbf{J}_1 c - \mathbf{I}_{n_1}\right),$$
$$\mathbf{N} = \left(\mathbf{J}_0 c - \mathbf{I}_{n_2}\right)^{-1} \mathbf{J}_0, \quad \mathbf{Q} = \mathbf{T}^{-1}. \tag{1.16.25}$$

Note that $\mathbf{N}^v = 0$, since $\mathbf{J}_0{}^v = 0$ and $\mathbf{N}v = (\mathrm{J}_0 c - \mathbf{I}_{n_2})^{-v}\mathbf{J}_0{}^v = 0$.

Remark 1.16.3.
Transforming $\mathbf{A}_1$ and $\mathbf{N}$ to the Jordan canonical form, we obtain

$$\text{diag}\left[\mathbf{H}_{m_1}s - \mathbf{I}_{m_1}, \ldots, \mathbf{H}_{m_t}s - \mathbf{I}_{m_t}, \; \mathbf{I}_{n_1}s - \mathbf{J}\right],$$

where

$$\mathbf{H}_{m_i} = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{bmatrix} \in \mathbb{R}^{m_i \times m_i} \quad (i = 1, \ldots, t)$$

and $\mathbf{J}$ is the Jordan canonical form of the matrix $\mathbf{A}_1$ and $m_1+m_2+\ldots+m_t+n_1=n$.

Theorem 1.16.4 generalises the classical Weierstrass theorem for the case of a rectangular pencil, which satisfies the condition (1.16.19).

If $q = n$, then the matrix $\mathbf{P}$ is square and nonsingular

$$\mathbf{P}^{-1}\left[\mathbf{E}s - \mathbf{A}\right]\mathbf{Q}^{-1} = \begin{bmatrix} \mathbf{I}_{n_1}s - \mathbf{A}_1 & 0 \\ 0 & \mathbf{N}s - \mathbf{I}_{n_2} \end{bmatrix}, \tag{1.16.26}$$

and n_1 is equal to the degree of the polynomial det [$\mathbf{E}s$ - $\mathbf{A}$].

Theorem 1.16.5. If [$\mathbf{E}s - \mathbf{A}$] is a regular pencil, then there exist two nonsingular matrices $\mathbf{P}$, $\mathbf{Q} \in \mathbb{C}^{n \times n}$ such that (1.16.26) holds.

The transformation matrices $\mathbf{P}$ and $\mathbf{Q}$ appearing in (1.16.26) can be computed by use of (1.16.25). Another method of computing these matrices is provided below.

Let s_i be the i-th root of the equation

$$\det\left[\mathbf{E}s - \mathbf{A}\right] = 0 \tag{1.16.27}$$

and

$$m_i = \dim \operatorname{Ker}\left[\mathbf{E}s_i - \mathbf{A}\right]. \tag{1.16.28}$$

Compute finite eigenvectors $v_{ij}{}^1$ using the equation

$$\left[\mathbf{E}s_i - \mathbf{A}\right]v_{ij}^1 = 0, \quad \text{for} \quad j = 1, \ ..., \ m_i, \tag{1.16.29}$$

and then (finite) eigenvectors $v_{ij}{}^{k+1}$ from the equation

$$\left[\mathbf{E}s_i - \mathbf{A}\right]v_{ij}^{k+1} = -\mathbf{E}v_{ij}^k, \quad \text{for} \quad k \geq 1. \tag{1.16.30}$$

Let

$$m_\infty = \dim \operatorname{Ker} \mathbf{E} = n - \operatorname{rank} \mathbf{E}. \tag{1.16.31}$$

We compute infinite eigenvectors $v_{\infty j}{}^1$ from the equations

$$\mathbf{E}v_{\infty j}^1 = 0, \quad \text{for} \quad j = 1, \ ..., \ m_\infty, \tag{1.16.32}$$

and then eigenvectors $v_{\infty j}{}^{k+1}$ from the equation

$$\mathbf{E}v_{\infty j}^{k+1} = \mathbf{A}v_{\infty j}^k, \quad \text{for} \quad k \geq 1. \tag{1.16.33}$$

The computed vectors are columns of the desired matrices

$$\mathbf{P} = \left[\mathbf{E}v_{ij}^k \ \vdots \ \mathbf{A}v_{\infty j}^k\right], \quad \mathbf{Q}^{-1} = \left[v_{ij}^k \ \vdots \ v_{\infty j}^k\right]. \tag{1.16.34}$$

Using (1.16.29)–(1.16.33) one can easily verify that

$$\begin{aligned} &\left[\mathbf{E}s - \mathbf{A}\right]\left[v_{ij}^k \quad \vdots \quad v_{\infty j}^k\right] = \\ &= \left[\mathbf{E}v_{ij}^k \quad \vdots \quad \mathbf{A}v_{\infty j}^k\right]\begin{bmatrix} \mathbf{I}_{n_1}s - \mathbf{A}_1 & 0 \\ 0 & \mathbf{N}s - \mathbf{I}_{n_2} \end{bmatrix}. \end{aligned} \tag{1.16.35}$$

Pre-multiplying (1.16.35) by $[\mathbf{E}v_{ij}{}^k \ \vdots \ \mathbf{A}v_{\infty j}{}^k]^{-1}$, we obtain (26) for **P** and **Q** given by (1.16.34).

Example 1.16.1.
Compute the matrices **P** and **Q** for a regular pencil whose matrices **E** and **A** have the form

$$\mathbf{E} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & -1 \end{bmatrix}.$$

In this case,

$$\det[\mathbf{E}s - \mathbf{A}] = \begin{vmatrix} s-1 & 0 & -1 \\ 0 & s-1 & 0 \\ 1 & 0 & 1 \end{vmatrix} = s(s-1)$$

and $n_1 = 2$, $n_2 = 1$, $s_1 = 1$, $s_2 = 0$, $m_1 = \dim \operatorname{Ker} [\mathbf{E}s_1 - \mathbf{A}] = 1$.

Using (1.16.29), (1.16.30), (1.16.32) and (1.16.33), we compute successively

$$[\mathbf{E}s_1 - \mathbf{A}]v_1 = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} v_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad v_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{E}v_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},$$

$$[\mathbf{E}s_2 - \mathbf{A}_2]v_2 = \begin{bmatrix} -1 & 0 & -1 \\ 0 & -1 & 0 \\ 1 & 0 & 1 \end{bmatrix} v_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad \mathbf{E}v_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},$$

$$\mathbf{E}v_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} v_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{A}v_3 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}.$$

Thus from (1.16.33) we have

$$\mathbf{P} = [\mathbf{E}v_1, \mathbf{E}v_2, \mathbf{A}v_3] = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \ \mathbf{Q}^{-1} = [v_1, v_2, v_3] = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix}.$$

1.17 Decomposition of Singular Pencils of Matrices

1.17.1 Weierstrass–Kronecker Theorem

Definition 1.17.1. A pencil **[E***s* – **A] (E, A**$\in\mathbb{C}^{m\times n}$) is said to be singular if $m \neq n$ or det **[E***s* – **A]** for all $s\in\mathbb{C}$ when $m = n$.

Let rank $[\mathbf{E}s - \mathbf{A}] = r \le \min(m, n)$ for almost every $s \in \mathbb{C}$.

Assume that $r < n$. In this case, the columns of the matrices $[\mathbf{E}s - \mathbf{A}]$ are linearly dependent and the equation

$$\left[\mathbf{E}s - \mathbf{A}\right] x = 0 \tag{1.17.1}$$

has a nonzero solution $x = x(s)$.

Among the polynomial solutions to (1.17.1) we seek solutions of the minimal degree p with respect to s having the form

$$x(s) = x_0 + x_1 s + x_2 s^2 + \cdots + x_p s^p . \tag{1.17.2}$$

Substituting (1.17.2) into (1.17.1) and comparing coefficients by the same powers of the variable s, we obtain the equations

$$-\mathbf{A}x_0 = 0,\ \ \mathbf{E}x_{i-1} - \mathbf{A}x_i = 0, \quad \text{for} \ \ i = 1, ..., p \ \ \text{and} \ \ \mathbf{E}x_p = 0,$$

which can be written in the form

$$\begin{bmatrix} -\mathbf{A} & 0 & 0 & \dots & 0 & 0 \\ \mathbf{E} & -\mathbf{A} & 0 & \dots & 0 & 0 \\ 0 & \mathbf{E} & -\mathbf{A} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \mathbf{E} & -\mathbf{A} \\ 0 & 0 & 0 & \dots & 0 & \mathbf{E} \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_{p-1} \\ x_p \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}. \tag{1.17.3}$$

Note that (1.17.3) has a solution if and only if the matrix

$$\mathbf{G}_p = \begin{bmatrix} -\mathbf{A} & 0 & 0 & \dots & 0 & 0 \\ \mathbf{E} & -\mathbf{A} & 0 & \dots & 0 & 0 \\ 0 & \mathbf{E} & -\mathbf{A} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \mathbf{E} & -\mathbf{A} \\ 0 & 0 & 0 & \dots & 0 & \mathbf{E} \end{bmatrix} \in \mathbb{C}^{(p+2)m \times (p+1)n}$$

does not have full column rank. By assumption p is minimal, thus we have rank $\mathbf{G}_i = (i+1)n$, for $i = 0,1,...,p-1$ and rank $\mathbf{G}_p < (p+1)n$.

Lemma 1.17.1. If (1.17.1) has the solution (1.17.2) of the minimal degree $p > 0$, then the pencil $[\mathbf{E}s - \mathbf{A}]$ is strictly equivalent to the pencil

$$\begin{bmatrix} \mathbf{L}_p & 0 \\ 0 & \overline{\mathbf{E}}s - \overline{\mathbf{A}} \end{bmatrix}, \tag{1.17.4}$$

where

$$\mathbf{L}_p = \begin{bmatrix} s & 1 & 0 & \cdots & 0 & 0 \\ 0 & s & 1 & \cdots & 0 & 0 \\ 0 & 0 & s & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & s & 1 \end{bmatrix} \in \mathbb{C}^{p\times(p+1)}$$

and the equation

$$\left[\overline{\mathbf{E}}s - \overline{\mathbf{A}}\right]x = 0 \tag{1.17.5}$$

does not have polynomial solutions of degree smaller than p.

Proof. Consider a linear operator [$\mathbf{E}s$ – $\mathbf{A}$] mapping $\mathbb{C}^n$ into $\mathbb{C}^m$. We will show that one can choose bases in $\mathbb{C}^n$ and $\mathbb{C}^m$ in such a way that the corresponding pencil [$\mathbf{E}s$ – $\mathbf{A}$] has the form

$$\begin{bmatrix} \mathbf{L}_p & \mathbf{B}s + c \\ 0 & \overline{\mathbf{E}}s - \overline{\mathbf{A}} \end{bmatrix}. \tag{1.17.6}$$

The linear operator equation corresponding to (1.17.1) is

$$\left[\mathbf{E}s - \mathbf{A}\right]x = 0\,, \tag{1.17.7}$$

where

$$x = x(s) = x_0 + x_1 s + x_2 s^2 + \cdots + x_p s^p\,.$$

Similarly as for (1.17.1) we obtain

$$\mathbf{A}x_0 = 0, \quad \mathbf{E}x_{i-1} = \mathbf{A}x_i \quad \text{and} \quad i = 1, ..., p, \quad \mathbf{E}x_p = 0\,. \tag{1.17.8}$$

We will show that the vectors

$$\mathbf{A}x_1,\ \mathbf{A}x_2, ..., \mathbf{A}x_p \tag{1.17.9}$$

are linearly independent.

Suppose that vector $\mathbf{A}x_k$ linearly depends on vectors $\mathbf{A}x_1,\ldots,\mathbf{A}x_{k-1}$ ($k \le p$), that is

$$\mathbf{A}x_k = a_1\mathbf{A}x_1 + \ldots + a_{k-1}\mathbf{A}x_{k-1} \quad \text{for certain} \quad a_i \in \mathbb{C}\,.$$

Using (1.17.8), we obtain

$$\mathbf{A}x_k = \mathbf{E}x_{k-1} = a_1\mathbf{E}x_0 + a_2\mathbf{E}x_1 + \ldots + a_{k-1}\mathbf{E}x_{k-2}$$

and

$$\mathbf{E}\hat{x}_{k-1} = 0\,,$$

where

$$\hat{x}_{k-1} = x_{k-1} - a_1x_0 - a_2x_1 - \ldots - a_{k-1}x_{k-2}\,.$$

Note that

$$\begin{aligned}\mathbf{A}x_{k-1} &= \mathbf{A}x_{k-1} - a_1\mathbf{A}x_0 - a_2\mathbf{A}x_1 - \ldots - a_{k-1}\mathbf{A}x_{k-2}\\ &= \mathbf{E}\left(x_{k-2} - a_2x_0 - a_3x_1 - \ldots - a_{k-1}x_{k-3}\right) = \mathbf{E}\hat{x}_{k-2}\end{aligned}$$

where

$$\hat{x}_{k-2} = x_{k-2} - a_2x_0 - a_3x_1 - \ldots - a_{k-1}x_{k-3}\,.$$

Similarly,

$$\begin{aligned}\mathbf{A}\hat{x}_{k-2} &= \mathbf{A}x_{k-2} - a_2\mathbf{A}x_0 - a_3\mathbf{A}x_1 - \ldots - a_{k-1}\mathbf{A}x_{k-3}\\ &= \mathbf{E}\left(x_{k-3} - a_3x_1 - \ldots - a_{k-1}x_{k-4}\right) = \mathbf{E}\hat{x}_{k-3},\end{aligned}$$

where

$$\hat{x}_{k-3} = x_{k-3} - a_3x_1 - \ldots - a_{k-1}x_{k-4}\,.$$

Continuing this procedure, we obtain

$$\mathbf{A}\hat{x}_{k-3} = \mathbf{E}\hat{x}_{k-4}\ ,\ldots,\ \mathbf{A}\hat{x}_1 = \mathbf{E}\hat{x}_0,\ \mathbf{A}\hat{x}_0 = 0\,,$$

where

$$\hat{x}_{k-4} = x_{k-4} - a_4x_1 - \ldots - a_{k-1}x_{k-5}\ ,\ldots,\ \hat{x}_1 = x_1 - a_{k-1}x_0,\ \hat{x}_0 = x_0\,.$$

Taking into account the above relationships one can easily verify that the vector

$$x = \hat{x}(s) = \hat{x}_0 + \hat{x}_1 s + \hat{x}_2 s^2 + \ldots + \hat{x}_{k-1} s^{k-1} \quad \text{and} \quad k \le p$$

is a solution to (1.17.7) of degree smaller than p. This contradiction proves that the vectors (1.17.9) are linearly independent.

We will show by contradiction that the vectors $x_0, x_1, \ldots, x_p$ are also linearly independent.

Suppose that these vectors are linearly dependent, that is

$$b_0 x_0 + b_1 x_1 + \ldots + b_p x_p = 0 \quad \text{for some} \quad b_i \in \mathbb{C}. \tag{1.17.10}$$

In this case, we obtain

$$b_1 \mathbf{A} x_1 + b_2 \mathbf{A} x_2 + \ldots + b_p \mathbf{A} x_p = 0,$$

since $\mathbf{A}x_0 = 0$.

The vectors (1.17.9) are linearly independent. In view of this, $b_1 = b_2 = \ldots = b_p = 0$ and from (1.17.10) we obtain $b_0 x_0 = 0$. Note that $x_0 \neq 0$, since otherwise $s^{-1}x(s)$ would also be a solution of the equation. Hence $b_0 x_0 = 0$ implies $b_0 = 0$ and the vectors $x_0, x_1, \ldots, x_p$ are linearly independent. We choose the vectors (1.17.9) to be the first basis vectors of the space $\mathbb{C}^n$ and the vectors $x_0, x_1, \ldots, x_p$ to be the first basis vectors of $\mathbb{C}^m$, respectively. Using (1.17.8), it is easy to verify that in this case, the pencil $[\mathbf{E}s - \mathbf{A}]$ has the form (1.17.6). Note that (1.17.4) can be obtained from (1.17.6) by adding to $[\mathbf{B}s + \mathbf{C}]$ an appropriate linear combination with coefficients independent of s from columns $\mathbf{L}_p$ and rows $[\overline{\mathbf{E}}s - \overline{\mathbf{A}}]$.

We will show that (1.17.5) has no solutions of degree smaller than p.

Taking into account (1.17.4), we can write down

$$\begin{bmatrix} \mathbf{L}_p & 0 \\ 0 & \overline{\mathbf{E}}s - \overline{\mathbf{A}} \end{bmatrix} \begin{bmatrix} z \\ y \end{bmatrix} = 0, \tag{1.17.11}$$

which is equivalent to

$$\mathbf{L}_p z = 0, \quad \left[\overline{\mathbf{E}}s - \overline{\mathbf{A}}\right] y = 0. \tag{1.17.12}$$

From the special form of $\mathbf{L}_p$ it follows that the equation ($\mathbf{L}_p z = 0$) has a solution of degree p of the form

$$z_i = (-1)^{i-1} s^{i-1} z_1 \quad (i = 1, \ldots, p+1)$$

for arbitrary z_1, where z_1 is the i-th component of vector z. Thus the matrix $\mathbf{G}_{p-1}$ in (1.17.3) has full column rank equal to pn.

The equation $[\bar{\mathbf{E}}s - \bar{\mathbf{A}}]y = 0$ has solution of the minimal degree p if and only if the matrix $\mathbf{G}_{p-1}$ in the equation

$$\bar{\mathbf{G}}_{p-1} = \begin{bmatrix} -\bar{\mathbf{A}} & 0 & 0 & \dots & 0 & 0 \\ \bar{\mathbf{E}} & -\bar{\mathbf{A}} & 0 & \dots & 0 & 0 \\ 0 & \bar{E} & -\bar{\mathbf{A}} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \bar{\mathbf{E}} & -\bar{\mathbf{A}} \\ 0 & 0 & 0 & \dots & 0 & \bar{\mathbf{E}} \end{bmatrix} \in \mathbb{C}^{(p+1)(n-p)\times p(n-p-1)}$$

has full column rank, equal to $p(n-p-1)$. From (1.17.4) it follows that the matrix $\mathbf{G}_{p-1}$ in (1.17.3), after the appropriate interchange of rows and columns, can be written in the form

$$\mathbf{G}_{p-1} = \begin{bmatrix} \hat{\mathbf{G}}_{p-1} & 0 \\ 0 & \bar{\mathbf{G}}_{p-1} \end{bmatrix},$$

where $\hat{\mathbf{G}}_{p-1}$ has dimensions $p(p+1)\times p(p+1)$ and corresponds to the equation $\mathbf{L}_p z = 0$. Note that the condition rank $\mathbf{G}_{p-1} = np$ implies that rank $\hat{\mathbf{G}}_{p-1} = p(p+1)$ and rank $\bar{\mathbf{G}}_{p-1} = p(n-p-1)$. Hence the equation $\mathbf{L}_p z = 0$ has no solution of degree smaller than p. ■

In the general case we assume that

1. rank $[\mathbf{E}s - \mathbf{A}] = r < \min(m,n)$;
2. columns and rows of $[\mathbf{E}s - \mathbf{A}]$ are linearly dependent over $\mathbb{C}$, i.e., there exist $x\in\mathbb{C}^n$ and $v\in\mathbb{C}^m$ (independent of s) such that

$$\left[\mathbf{E}s - \mathbf{A}\right]x = 0 \tag{1.17.13}$$

and

$$\left[\mathbf{E}s - \mathbf{A}\right]^T v = 0. \tag{1.17.14}$$

Let (1.17.13) have p_0 linearly independent solutions $x_1, x_2, \dots, x_{p_0}$. Choosing these solutions as the first p_0 basis vectors of the space $\mathbb{C}^n$, we obtain a strictly equivalent pencil that has the form

$$\left[\mathbf{0}_{np_0} \;\vdots\; \bar{\mathbf{E}}s - \bar{\mathbf{A}}\right], \tag{1.17.15}$$

where 0_{np_0} is a zero-matrix of dimension $n\times p_0$.

Similarly, let (1.17.14) have q_0 linearly independent solutions $v_1, v_2, \dots, v_{q_0}$.Choosing these solutions as the first q_0 basis vectors of the space $\mathbb{C}^m$, we obtain a strictly equivalent pencil that has the form

$$\left[\begin{array}{c|c} & \mathbf{0}_{q_0, n-p_0} \\ \mathbf{0}_{n,p_0} & \tilde{\mathbf{E}}s - \tilde{\mathbf{A}} \end{array}\right], \tag{1.17.16}$$

where $[\tilde{\mathbf{E}}s - \tilde{\mathbf{A}}]$ has rows and columns linearly independent over $\mathbb{C}$.

Let the columns of $[\tilde{\mathbf{E}}s - \tilde{\mathbf{A}}]$ be linearly dependent over the field of rational functions $C(s)$ and let the equation

$$\left[\tilde{\mathbf{E}}s - \tilde{\mathbf{A}}\right]x = 0$$

Have a polynomial solution of the minimal degree p_1. Applying Lemma 1.17.1 to the pencil $[\tilde{\mathbf{E}}s - \tilde{\mathbf{A}}]$, we obtain a strictly equivalent pencil that has the form

$$\left[\begin{array}{c|cc} & \mathbf{0}_{q_0, n-p_0} & \\ \mathbf{0}_{n,p_0} & \mathbf{L}_{p1} & 0 \\ & 0 & \mathbf{E}_1 s - \mathbf{A}_1 \end{array}\right], \tag{1.17.17}$$

and the equation

$$\left[\mathbf{E}_1 s - \mathbf{A}_1\right]x = 0 \tag{17.18}$$

that has no polynomial solutions of degree smaller than p_1.

If (1.17.18) has a polynomial solution of the minimal degree p_2, then continuing this procedure, we obtain a strictly equivalent pencil that has the form

$$\left[\begin{array}{c|c} & \mathbf{0}_{q_0, n-p_0} \\ \mathbf{0}_{np_0} & \operatorname{diag}\left[\mathbf{L}_{p_1}, \mathbf{L}_{p_2}, \dots, \mathbf{L}_{p_w}, \ \mathbf{E}_w s - \mathbf{A}_w\right] \end{array}\right], \tag{1.17.19}$$

where $p_1 \le p_2 \le \ldots \le p_w$, and the equation $[\mathbf{E}_w s - \mathbf{A}_w]x = 0$ has no nonzero polynomial solutions.

If the pencil $[\mathbf{E}_w s - \mathbf{A}_w]$ has linearly dependent rows over the field $\mathbb{C}(s)$ and the equation

$$\left[\mathbf{E}_w s - \mathbf{A}_w\right]^T v = 0$$

has polynomial solution of the minimal degree q_1, then applying Lemma 1.17.1 to $[\mathbf{E}_w s - \mathbf{A}_w]^T$, we obtain a strictly equivalent pencil that has the form

$$\left[\begin{array}{c|c} \mathbf{0}_{np_0} & \begin{array}{c} \mathbf{0}_{q_0,n-p_0} \\ \operatorname{diag}\left[\mathbf{L}_{p_1}, \mathbf{L}_{p_{2_q}}, \ldots, \mathbf{L}_{p_w}, \ \mathbf{L}_{q1}^T, \ \mathbf{E}_1' s - \mathbf{A}_1'\right] \end{array} \end{array}\right], \tag{1.17.20}$$

where the equation

$$\left[\mathbf{E}_1' s - \mathbf{A}_1'\right]^T v = 0 \tag{1.17.21}$$

has no polynomial solutions of degree smaller than q_1.

If (1.17.21) has polynomial solution of the minimal degree q_2, then continuing this procedure, we obtain a strictly equivalent pencil that has the form

$$\left[\begin{array}{c|c} \mathbf{0}_{np_0} & \begin{array}{c} \mathbf{0}_{q_0,n-p_0} \\ \operatorname{diag}\left[\mathbf{L}_{p_1}, \ldots, \mathbf{L}_{p_w}, \ \mathbf{L}_{q1}^T, \ldots, \mathbf{L}_{q_s}^T, \ \mathbf{E}_0 s - \mathbf{A}_0\right] \end{array} \end{array}\right], \tag{1.17.22}$$

where $[\mathbf{E}_0 s - \mathbf{A}_0]$ is a regular pencil.

Applying Theorem 1.16.4 to the pencil $[\mathbf{E}_0 s - \mathbf{A}_0]$, we obtain the Weierstrass–Kronecker canonical form of a singular pencil, that is

$$\left[\begin{array}{c|c} \mathbf{0}_{np_0} & \begin{array}{c} \mathbf{0}_{q_0,n-p_0} \\ \operatorname{diag}\left[\mathbf{L}_{p_1}, \ldots, \mathbf{L}_{p_w}, \mathbf{L}_{q1}^T, \ldots, \mathbf{L}_{q_s}^T, \mathbf{H}_{n_1} s - \mathbf{I}_{n_1}, \ldots, \mathbf{H}_{n_t} s - \mathbf{I}_{n_t}, \mathbf{I}_r s - \mathbf{J}\right] \end{array} \end{array}\right], \tag{1.17.23}$$

where the pencil

$$\operatorname{diag}\left[\mathbf{H}_{n_1} s - \mathbf{I}_{n_1}, \ldots, \mathbf{H}_{n_t} s - \mathbf{I}_{n_t}, \mathbf{I}_r s - \mathbf{J}\right]$$

corresponds to the regular pencil $\mathbf{E}_0 s - \mathbf{A}_0$.

Thus we have proven the following Weierstrass–Kronecker theorem about decomposition of a singular pencil.

Theorem 1.17.1. An arbitrary singular pencil $[\mathbf{E}s - \mathbf{A}]$ is strictly equivalent to the pencil (1.17.23).

1.17.2 Kronecker Indices of Singular Pencils and Strict Equivalence of Singular Pencils

Let us consider a pencil $[\mathbf{E}s - \mathbf{A}]$ for $\mathbf{E}, \mathbf{A} \in \mathbb{C}^{m \times n}$. Let $x_1(s)$ be a nonzero polynomial solution of minimal degree p_1 of the equation

$$\left[\mathbf{E}s - \mathbf{A}\right] x = 0 . \tag{1.17.24}$$

Among polynomial solutions of the equation, linearly which are independent of $x_1(s)$ over $\mathbb{C}(s)$, we choose a solution $x_2(s)$ of minimal degree p_2 ($p_2 \geq p_1$). Then among polynomial solutions of (1.17.24) which are linearly independent of $x_1(s)$ and $x_2(s)$ over $\mathbb{C}(s)$, we choose solutions $x_3(s)$ of minimal degree p_3 ($p_3 \geq p_2$). Continuing this procedure we obtain a sequence of linearly independent polynomial solutions of (1.17.24) of the form

$$x_1(s),\ x_2(s), \ldots,\ x_w(s) \quad (w \leq n) \tag{1.17.25}$$

with degrees

$$p_1 \leq p_2 \leq \cdots \leq p_w\,. \tag{1.17.26}$$

In the general case, for a given pencil $[\mathbf{E}s - \mathbf{A}]$ there exist many sequences of the polynomial solutions (1.17.25) to (1.17.24). We will show that all these sequences of polynomial solutions have the same sequence of degrees (1.17.26).

Suppose that $\overline{x}_1(s), \overline{x}_2(s), \ldots, \overline{x}_w(s)$ with degrees $\overline{p}_1 \leq \overline{p}_2 \leq \ldots \leq \overline{p}_w$ is another sequence of polynomial solutions to (1.17.24). Let

$$p_1 = \ldots = p_{n_1} < p_{n_1+1} = \ldots = p_{n_2} < p_{n_2+1} = \ldots$$

and

$$\overline{p}_1 = \ldots = \overline{p}_{\overline{n}_1} < \overline{p}_{\overline{n}_1+1} = \ldots = \overline{p}_{\overline{n}_2} < \overline{p}_{\overline{n}_2+1} = \ldots$$

From this choice of $x_1(s)$ and $\overline{x}_1(s)$ it follows that $p_1 = \overline{p}_1$. Note that $\overline{x}_1(s)$ for $i = 1, \ldots, \overline{n}_1$ is a linear combination $x_1(s), \ldots, x_{n_1}(s)$, since otherwise $x_{n_1+1}(s)$ in (1.17.25) could be replaced with a polynomial vector of degree smaller than p_{n_1+1}. Similarly, $x_i(s)$ for $i = 1, \ldots, n_1$ is a linear combination $\overline{x}_1(s), \ldots, \overline{x}_{n_1}(s)$. In view of this, $n_1 = \overline{n}_1$ and $p_{n_1+1} = \overline{p}_{\overline{n}_1+1}$. Similarly it is easy to show that $p_{n_2+1} = \overline{p}_{\overline{n}_2+1}$ $p_{n_2+1} = \overline{p}_{\overline{n}_2+1}$.

Definition 1.17.2. Nonnegative integers $p_1, p_2, \ldots, p_w$ are called minimal column (Kronecker) indices of the pencil $[\mathbf{E}s - \mathbf{A}]$.

Let $v_1(s)$ be the nonzero polynomial solution of the minimal degree q_1 of the equation

$$[\mathbf{E}s - \mathbf{A}]^T v = 0\,. \tag{1.17.27}$$

Among the polynomial solutions of this equation, which are linearly independent over $\mathbb{C}(s)$ of $v_1(s)$, we choose a solution $v_2(s)$ of minimal degree q_2 ($q_2 \geq q_1$).

Continuing this procedure, we obtain a sequence of polynomial solutions to (1.17.27) of the form

$$v_1(s),\ v_2(s), ...,\ v_s(s)\ \ (s \le n) \tag{1.17.28}$$

with degrees

$$q_1 \le q_2 \le \cdots \le q_s\ . \tag{1.17.29}$$

Similarly to (1.17.25) and (1.17.26) one can show that all sequences of polynomial solutions (1.17.28) to (1.17.27) have the same sequences of minimal degrees (1.17.29).

Definition 1.17.3. Nonnegative integers $q_1,q_2,...,q_s$ are called minimal row (Kronecker) indices of the pencil $[\mathbf{E}s - \mathbf{A}]$.

Lemma 1.17.2. Strictly equivalent pencils have the same minimal column and row Kronecker indices.

Proof. Take strictly equivalent pencils $[\mathbf{E}_1 s - \mathbf{A}_1]$ and $[\mathbf{E}_2 s - \mathbf{A}_2]$, i.e., related by the relationship $[\mathbf{E}_2 s - \mathbf{A}_2] = \mathbf{P}[\mathbf{E}_1 s - \mathbf{A}_1]\mathbf{Q}$.

Pre-multiplying the equation

$$\left[\mathbf{E}_1 s - \mathbf{A}_1\right]x = 0 \tag{1.17.30}$$

by a nonsingular matrix $\mathbf{P}$ and defining a new vector $z = \mathbf{Q}^{-1}x$ ($\mathbf{Q}$ is a nonsingular matrix), we obtain

$$\mathbf{P}\left[\mathbf{E}_1 s - \mathbf{A}_1\right]\mathbf{Q}\mathbf{Q}^{-1}x = \left[\mathbf{E}_2 s - \mathbf{A}_2\right]z = 0\,. \tag{1.17.31}$$

Thus these pencils have the same minimal column indices, since the degree of x in (1.17.30) is equal to the degree of z in (1.17.31). Similarly we can prove that these pencils have the same minimal row indices.

■

Lemma 1.17.3. The Weierstrass–Kronecker canonical form (1.17.23) of the pencil $[\mathbf{E}s - \mathbf{A}]$ is completely determined by p_0 minimal column indices, which are equal to zero, nonzero minimal column indices $p_1,p_2,...,p_w,q_0$, minimal row indices equal to zero, nonzero minimal row indices $q_1,q_2,...,q_s$ and by its finite and infinite elementary divisors.

Proof. The matrix $\mathbf{L}_{p_i}$ ($i = 1,...,w$) has only one minimal column index p_i, since the equation $\mathbf{L}_{p_i} z = 0$ has only one polynomial solution of degree p_i and the rows of the matrix $\mathbf{L}_{p_i}$ are linearly independent. Similarly, the matrix $\mathbf{L}^T_{q_j}$ ($j = 1,...,s$)

has only one minimal zero index q_j, since the equation $\mathbf{L}^T_{q_j} v = 0$ has only one polynomial solution of degree q_j and columns of the matrix $\mathbf{L}^T_{q_j}$ are linearly independent. It is easy to check that the matrix $\mathbf{L}_{p_i}$ (or $\mathbf{L}^T_{q_j}$) does not have any elementary divisors, since one of its minors, of the greatest degree p_i (respectively q_j), is equal to 1 and the other one is equal to s^{p_i} (s^{q_j}).

The first p_0 columns of the matrix (1.17.23) correspond to polynomial solutions of (1.17.13). In view of this, the first p_0 minimal column indices of $[\mathbf{E}s - \mathbf{A}]$ are equal to 0. Dually, the first q_0 minimal row indices of $[\mathbf{E}s - \mathbf{A}]$ are equal to zero.

Note that the pencil $[\mathbf{E}_0 s - \mathbf{A}_0]$ in (1.17.22) is regular, hence it is completely determined by its finite and infinite elementary divisors. From the block-diagonal form (1.17.23) it follows that the canonical form of the pencil $[\mathbf{E}s - \mathbf{A}]$ is completely determined by minimal column and row indices, and finite and infinite elementary divisors of every diagonal block.

■

From Lemmas 1.17.2 and 1.17.3 and from the fact that two singular pencils having the same canonical forms are strictly equivalent, the following Kronecker theorem can be inferred.

Theorem 1.17.2. (Kronecker) Two singular pencils $[\mathbf{E}_1 s - \mathbf{A}_1]$, $[\mathbf{E}_2 s - \mathbf{A}_2]$ for $\mathbf{E}_k, \mathbf{A}_k \in \mathbb{C}^{m \times n}$ $(k = 1,2)$ are strictly equivalent if and only if they have the same minimal column and row indices, as well the same finite and infinite elementary divisors.