SOME NONLINEAR APPROXIMATIONS FOR MATRIX-VALUED FUNCTIONS [*1)]

Guo-liang Xu
(*Academy of Mathematics and System Sciences, Chinese Academy of Sciences,
Beijing 100080, China*)

Abstract

Some nonlinear approximants, i.e., exponential-sum interpolation with equal distance or at origin, (0,1)-type, (0,2)-type and (1,2)-type fraction-sum approximations, for matrix-valued functions are introduced. All these approximation problems lead to a same form system of nonlinear equations. Solving methods for the nonlinear system are discussed. Conclusions on uniqueness and convergence of the approximants for certain class of functions are given.

Key words: Matrix-valued function, Nonlinear approximation, Interpolation.

1. Introduction

For scalar functions, Baker-Gammel approximation was developed in a series of papers [1, 2, 3, 7] (see [7] for more references) as a method for producing nonlinear approximation with good convergence properties. The main features of these methods can be described as follows: **1**. Use a generating function for the function to be approximated to derive a nonlinear approximant by agreement condition between the function and the approximant. **2**. Draw on the relations between the nonlinear approximants and Padé approximants. **3**. Establish convergence results from this relation and the convergence results of Padé approximation.

In the present paper, we make efforts to generalize these ideas to matrix-valued functions. A few concrete nonlinear approximations, such as exponential sum interpolation at origin and Padé-like approximation, are discussed. Since the multiplications of matrices are not commuting, this generalization is not straightforward. The first problem is how to do partial fraction for a matrix rational function. In the aspect of existence and uniqueness of the nonlinear approximants, there also exist some problems that unlike their scalar counterpart. We shall give solutions to these problems in this paper under certain conditions.

The remaining of the paper is organized as follows: We first introduce the definitions of the nonlinear approximation problems in §2. All these problems lead to a same form system of nonlinear equations. Then we solve this system in §3 using the matrix Padé approximation method and the theory of matrix polynomials. In §4, we establish the relationship between these approximants defined in §2, and then we discuss the uniqueness problem in §5. Finally, utilizing the convergence results of the Padé approximation for the matrix-valued Stieltjes function, we give the convergence conclusion of the nonlinear approximants. Some terminologies and basic facts in the matrix polynomial theory used in this paper are provided in Appendix.

2. Definition of the Approximants

Now we define our nonlinear approximants for matrix-valued functions.

a. Exponential-sum interpolation with equal distance.

Let $f(z) \in \mathbb{C}^{n \times n}[z]$ be a given matrix-valued function. Construct a function in the form

$$F(z) = \sum_{j=1}^{l} A_j e^{s_j z},$$

where A_j, $s_j \in \mathbb{C}^{n \times n}$ are parameters to be determined for $j = 1, \cdots, l$, such that

$$F(iT) = f(iT), \quad i = 0, 1, \cdots, 2l - 1, \tag{2.1}$$

where T is a given constant. Put $c_i = f(iT), S_j = e^{s_j T}$ for $j = 1, \cdots, l$, then (2.1) can be written as

$$\sum_{j=1}^{l} A_j S_j^i = c_i, \quad i = 0, 1, \cdots, 2l - 1. \tag{2.2}$$

Therefore, if the solutions of equation(2.2) are found for the unknowns A_j and S_j, $j = 1, \cdots, l$, then s_j can be determined by

$$s_j = \frac{1}{T} \log S_j, \quad j = 1, 2, \cdots, l.$$

b. Exponential-sum interpolation at the origin.

For a given power series $f(z) = \sum_{i=0}^{\infty} \frac{c_i}{i!} z^i \in \mathbb{C}^{n \times n}[z]$, find

$$E(z) = \sum_{i=0}^{J} \frac{B_i}{i!} z^i + \sum_{j=1}^{l} A_j e^{S_j z} \tag{2.3}$$

such that

$$\left. \frac{d^i E(z)}{dz^i} \right|_{z=0} = c_i, \quad i = 0, 1, \cdots, 2l + J, \quad J \geq -1. \tag{2.4}$$

It follows from (2.3) and (2.4) that

$$\begin{cases} B_i + \sum_{j=1}^{l} A_j S_j^i = c_i, \quad i = 0, 1, \cdots, J, \\ \sum_{j=1}^{l} A_j S_j^i = c_i, \qquad i = J + 1, \cdots, 2l + J. \end{cases} \tag{2.5}$$

If $J = -1$, the system of equations (2.5) is the same as (2.2).

c. (0,1)-type fraction-sum approximation.

Let $f(z) = \sum_{i=0}^{\infty} c_i z^i$ be a given power series with $c_i \in \mathbb{C}^{n \times n}$. The problem of (0,1)-type fraction-sum approximation is to find a function in the form

$$P(z) = \sum_{i=0}^{J} B_i z^i + \sum_{j=1}^{l} A_j (I - z S_j)^{-1}, \quad J \geq -1 \tag{2.6}$$

such that

$$f(z) - P(z) = O(z^{2l + J + 1}), \quad z \to 0. \tag{2.7}$$

Since

$$(I - zS_j)^{-1} = \sum_{i=0}^{\infty} S_j^i z^i \quad \text{as} \quad z \to 0, \tag{2.8}$$

from (2.6) and (2.7) we arrive at equations (2.5) again.

Since the multiplications of $S_j (j = 1, 2, \cdots, l)$ are not commuting in general, $P(z)$ can not be written in a rational function form

$$R(z) = \sum_{i=0}^{m} a_i z^i \left(\sum_{i=0}^{l} b_i z^i \right)^{-1}, \quad a_i, b_i \in \mathbb{C}^{n \times n} \tag{2.9}$$

directly by reducing the function (2.6) to a common denominator. But we shall show in the later that under some conditions $P(z)$ is a right Padé approximant in the following sense:

Determine the parameters a_i, $b_i \in \mathbb{C}^{n \times n}$ in the rational function (2.9) such that

$$\begin{cases} f(z) - R(z) = O(z^{m+l+1}), & z \to 0, \\ b_0 = I. \end{cases} \tag{2.10}$$

An alternative approach to (0,1)-type fraction-sum approximation is to find a function in the form

$$\tilde{P}(z) = \sum_{i=0}^{J} \tilde{B}_i z^i + z^{J+1} \sum_{j=1}^{l} \tilde{A}_j (I - z\tilde{S}_j)^{-1}$$

such that

$$f(z) - \tilde{P}(z) = O(z^{2l+J+1}), \quad z \to 0. \tag{2.11}$$

Hence

$$\begin{cases} \tilde{B}_i = c_i, & i = 0, 1, \cdots, J, \\ \sum_{j=1}^{l} \tilde{A}_j \tilde{S}_j^{i-J-1} = c_i, & i = J+1, \cdots, 2l+J. \end{cases} \tag{2.12}$$

d. (0,2)-type fraction-sum approximation.

For a given series expansion in Chebyshev polynomials of the second kind $f(z) = \sum_{i=0}^{\infty} c_i U_i(z) \in \mathbb{C}^{n \times n}[z]$, find

$$U(z) = \sum_{i=0}^{J} B_i U_i(z) + \sum_{j=1}^{l} A_j (I - 2zS_j + S_j^2)^{-1}. \tag{2.13}$$

such that the series expansion of $U(z)$ in Chebyshev polynomials of the second kind agrees with that of $f(z)$ for the first $2l + J + 1$ terms. Since

$$\frac{1}{1 - 2xt + t^2} = \sum_{n=0}^{\infty} U_n(x) t^n, \quad x \in (-1, 1), \quad |t| < 1.$$

$$(I - 2zS + S^2)^{-1} = \sum_{i=0}^{\infty} U_i(z) S^i, \quad z \in (-1, 1), \quad |\lambda| < 1, \forall \lambda \in \sigma(S) \tag{2.14}$$

substitute (2.14) into (2.13), we arrive at (2.5) again.

e. (1,2)-type fraction-sum approximation.

For a given series expansion in Chebyshev polynomials of the first kind $f(z) = \sum\limits_{i=0}^{\infty} c_i T_i(z) \in$
$\mathbb{C}^{n \times n}[z]$, find

$$T(z) = \sum_{i=0}^{J} B_i T_i(z) + \sum_{j=1}^{l} A_j (I - zS_j)(I - 2zS_j + S_j^2)^{-1}. \tag{2.15}$$

such that the expansion of $T(z)$ in Chebyshev polynomials of the first kind agrees with that of
$f(z)$ for the first $2l + J + 1$ terms. Since

$$\frac{1 - zu}{1 - 2zu + u^2} = \sum_{i=0}^{\infty} T_i(z) u^i, \quad z \in [-1, 1], \quad |u| \leq 1,$$

$$\frac{I - zS}{I - 2zS + S^2} = \sum_{i=0}^{\infty} T_i(z) S^i, \quad z \in [-1, 1], \quad |\lambda| \leq 1, \forall \lambda \in \sigma(S). \tag{2.16}$$

Substituting (2.16) into (2.15), we arrive at the (2.5) once more.

Another source of system(2.2) is the establishment of Gauss-type integral formula

$$\int_a^b f(x) d\sigma(x) \approx \sum_{i=1}^{l} A_i f(S_i), \tag{2.17}$$

where $\sigma(x)$ is a matrix Stieltjes measure (see [9]) on (a, b). We want to determine matrices A_i
and S_i such that (2.17) are equalities for $f(x) = 1, x, \cdots, x^{2l-1}$. These equalities, by letting
$c_i = \int_a^b x^i d\,\sigma(x)$ for $i = 0, 1, \cdots, 2l - 1$, are the same as equation (2.2).

3. The Solutions of the Problems Proposed

We note firstly that equation (2.2) is a special case of equations (2.5) with $J = -1$. Thus
we need only to consider the solution of the nonlinear equations (2.5) for the unknowns A_i, B_i
and S_i. The difficult part in solving the system is to find S_i in (2.5). Once S_i are determined,
finding A_i, B_i becomes a linear problem. These S_i are defined as left solvents of a matrix
polynomial $L(z)$. Hence system (2.5) is solved in the following steps:

1. *Determine the matrix polynomial* $L(z) = \sum_{i=0}^{l} L_i z^i$, $L_i \in \mathbb{C}^{n \times n}$.
2. *Compute the left solvents* S_i *of* $L(z)$.
3. *Compute* A_i, B_i.

Now we detail each of these steps. Some theoretical results are introduced therein.

Step 1. Let $L_i \in \mathbb{C}^{n \times n}$, $i = 0, 1, \cdots, l$ be defined so that $\sum_{i=0}^{l} c_{k+J+i+1} L_i = 0$, $k = 0, 1, \cdots, l - 1$, $L_l = I$. That is, $L_0, \cdots, L_{l-1}$ satisfy the following equation

$$H(J + 1, l, l) \begin{bmatrix} L_0 \\ L_1 \\ \vdots \\ L_{l-1} \end{bmatrix} = - \begin{bmatrix} c_{l+J+1} \\ c_{l+J+2} \\ \vdots \\ c_{2l+J} \end{bmatrix} \tag{3.1}$$

with

$$H(i, j, k) = \begin{bmatrix} c_i & c_{i+1} & \cdots & c_{i+j-1} \\ c_{i+1} & c_{i+2} & \cdots & c_{i+j} \\ \cdots & \cdots & \cdots & \cdots \\ c_{i+k-1} & c_{i+k} & \cdots & c_{i+j+k-2} \end{bmatrix}.$$

Obviously, equation of system (3.1) is solvable if and only if

$$\text{rank } H(J+1, l, l) = \text{ rank } H(J+1, l+1, l). \tag{3.2}$$

Suppose equation (3.1) has solutions and assume that the matrix polynomial $L(z) = \sum\limits_{i=0}^{l} L_i z^i$ has a complete set of left solvents $S_1, \cdots, S_l$. If $J \geq 0$, we further assume that $S_1, \cdots, S_l$ are nonsingular. This assumption is equivalent to $0 \notin \sigma(L)$ (see [6], p.524). Here $\sigma(L)$ denotes the set of latent roots of $L(z)$. The proposition in the following gives necessary and sufficient condition for $0 \notin \sigma(L)$.

Proposition 3.1. *Suppose $H(J+1, l, l)$ is nonsingular, $L_0, \cdots, L_l$ are the solution of (3.1), then $0 \notin \sigma(L)$ if and only if $H(J+2, l, l)$ is nonsingular also.*

Proof. Since $0 \notin \sigma(L)$ if and only if L_0 is nonsingular, it is necessary to show that L_0 is nonsingular if and only if $H(J+2, l, l)$ is nonsingular.

Suppose $H(J+2, l, l)$ is nonsingular, we show that L_0 is nonsingular also. If rank $L_0 < n$, then there exists an invertible matrix P such that $L_0 P$ has at least one zero column, say, the i-th column of $L_0 P$ is zero. Since $L_l P$ is nonsingular, the i-th column C_i of the matrix $\left[(L_0 P)^T \ (L_1 P)^T \ \cdots \ (L_l P)^T \right]^T$ is nonzero. However,

$$H(J+1, l+1, l)C_i = H(J+2, l, l)C_i = 0.$$

This is impossible for $H(J+2, l, l)$ is nonsingular. Hence rank $L_0 = n$. Conversely, if rank $L_0 = n$, then by

$$H(J+1, l+1, l)\begin{bmatrix} I_n \\ L_1 L_0^{-1} \\ \vdots \\ L_l L_0^{-1} \end{bmatrix} = 0,$$

we have rank $H(J+1, l+1, l) = $ rank $H(J+2, l, l)$. Hence by the non-singularity of $H(J+1, l, l)$ we know that $H(J+2, l, l)$ is nonsingular.

Step 2. Now we consider the problem of finding a complete set of left solvents $S_1, \cdots, S_l$ of $L(z)$. Let (X, J) be a Jordan pair of $L^T(z)$. Then (see [6], p.500) the $nl \times nl$ matrix

$$Q = \begin{bmatrix} X^T & (XJ)^T & \cdots & (XJ^{l-1})^T \end{bmatrix}^T \tag{3.3}$$

is nonsingular and

$$L_l^T X J^l + \cdots + L_1^T X J + L_0^T X = 0. \tag{3.4}$$

Denote the degrees of the elementary divisors of $L^T(z)$ (which are also the elementary divisors of $\lambda I - C_{L^T}$) by $k_1, \cdots, k_m$. Then $\sum_{i=1}^{m} k_i = nl$. Suppose $\{k_i\}_{i=1}^{m}$ can be divided into l subsets such that the sum of each subset is n. Then we can arrange J by rows and columns exchange such that $J = \text{diag}[J_1, \cdots, J_l]$, where $J_i \in \mathbb{C}^{n \times n}$, $i = 1, \cdots, l$, are also block diagonal matrices with Jordan blocks as their diagonal elements. Form the partition $X = [X_1 \ \cdots \ X_l]$, where $X_1, \cdots, X_l$ are $n \times n$ matrices, then for $r = 0, 1, \cdots$, we have $XJ^r = [X_1 J_1^r \ X_2 J_2^r \ \cdots \ X_l J_l^r]$. Thus equation (3.4) implies that

$$\sum_{i=0}^{l} L_i^T X_j J_j^i = 0, \quad j = 1, \cdots, l. \tag{3.5}$$

Now we assume that $X_1, \cdots, X_l$ are nonsingular. Then (3.5) implies

$$\sum_{i=0}^{l} L_i^T X_j J_j^i X_j^{-1} = 0, \quad j = 1, \cdots, l.$$

Let $S_j^i = (X_j J_j^i X_j^{-1})^T$ for $j = 1, \cdots, l$. Then $\sum_{i=0}^l S_j^i L_i = 0$ and $Q \operatorname{diag}[X_1^{-1}, \cdots, X_l^{-1}] = V^T$ is nonsingular, since both Q and $\operatorname{diag}[X_1^{-1}, \cdots, X_l^{-1}]$ are nonsingular. Therefore, if the required partition of X and J exists and $X_1, \cdots, X_l$ are invertible, then the complete set of left solvents exists.

In order to get a succeed partition, we should make a good choice among the following two degrees of freedom:

a. J is not unique by considering the order of the diagonal Jordan block.

b. For a fixed J, X is not unique. Note that for any nonsingular matrix $Y \in \mathbb{C}^{nl \times nl}$ that commute with J, (XY, J) is also a Jordan pair.

Here we shall mention a theorem for the conditions of Y that commute with J.

Theorem 3.1. ([6], p.418) *Let $J = \operatorname{diag}[J_1, \cdots, J_m]$ be a Jordan canonical form and k_i be the size of the Jordan block J_i corresponding to an eigenvalue λ_i. Let $Y = [Y_{st}]_{s,t=1}^m$ be the partition of Y consistent with the partition of J into Jordan blocks. Then Y commutes with J if and only if*

a. $Y_{st} = 0$, *for $\lambda_s \neq \lambda_t$,*

b. Y_{st} *is an upper-triangular Toeplitz matrix for $k_s = k_t$ and $\lambda_s = \lambda_t$,*

c. $Y_{st} = [0, Y_{k_s}]$ *for $k_s < k_t$ and $\lambda_s = \lambda_t$,*

d. $Y_{st} = \begin{bmatrix} Y_{k_t} \\ 0 \end{bmatrix}$, *$k_s > k_t$ and $\lambda_s = \lambda_t$,*

where Y_{k_s} and Y_{k_t} are upper-triangular Toeplitz matrices of order k_s and k_t, respectively.

The discussion above gives us the necessary part of the following theorem:

Theorem 3.2. *$L(z)$ has a complete set of left solvents if and only if there exists a Jordan pair (X, J) of $L^T(z)$ and a partition $X = [X_1 \quad \cdots \quad X_l]$, $J = \operatorname{diag}[J_1, \cdots, J_l]$ such that $X_j \in \mathbb{C}^{n \times n}$ are invertible for $j = 1, \cdots, l$.*

Proof. We need only to prove the sufficiency part of the theorem. Let $S_1, \cdots, S_l$ be a complete set of left solvents of $L(z)$. Let

$$S_j^T = X_j J_j X_j^{-1}, \quad j = 1, 2, \cdots, l,$$

where J_j is the Jordan canonical form of S_j^T for $j = 1, 2, \cdots, l$. Then the equalities

$$0 = \sum_{i=0}^l L_i^T (S_j^T)^i = \sum_{i=0}^l L_i^T X_j J_j^i X_j^{-1}, \quad j = 1, 2, \cdots, l$$

implies that

$$0 = \sum_{i=0}^l L_i^T [X_1 J_1^i \quad \cdots \quad X_l J_l^i] = \sum_{i=0}^l L_i^T X J^i,$$

where $X = [X_1 \quad \cdots \quad X_l]$, $J = \operatorname{diag}[J_1, \cdots, J_l]$. On the other hand,

$$\begin{bmatrix} X^T & (XJ)^T & \cdots & (XJ^{l-1})^T \end{bmatrix}^T \operatorname{diag}[X_1^{-1}, \cdots, X_l^{-1}] = V^T.$$

Then $\begin{bmatrix} X^T & (XJ)^T & \cdots & (XJ^{l-1})^T \end{bmatrix}^T$ is an invertible matrix. Therefore, (X, J) is a standard pair of $L^T(z)$ (see [6], p. 495). Since $C_{L^T} V^T = V^T \operatorname{diag}[S_1, \cdots, S_l]$ (i.e., $\operatorname{diag}[S_1, \cdots, S_l]$ is similar to C_{L^T}), J is similar to C_{L^T}. Hence, (X, J) is a Jordan pair of $L^T(z)$.

Step 3. Now the parameters $A_1, \cdots, A_l$ in (2.5) can be determined by the following equations

$$\sum_{j=1}^l A_j S_j^i = c_i, \quad i = J + 1, \cdots, J + l, \quad J \geq -1. \tag{3.6}$$

Since the coefficient matrix of system (3.6) is nonsingular, the system has solution uniquely.

Once A_j, $S_j (j = 1, \cdots, l)$ are determined, parameters B_i in (2.5) are obtained by

$$B_i = c_i - \sum_{j=1}^{l} A_j S_j^i, \quad i = 0, 1, \cdots, J.$$

The parameters B_i, A_i and S_i given above form a solution of (2.5). In fact, the first $J + l + 1$ equations for $i = 0, 1, \cdots, J, J+1, \cdots, J+l$ hold obviously. For $i = J + l + 1, \cdots, 2l + J$, from (3.1) and induction, we have

$$
\begin{aligned}
c_i &= -\sum_{j=0}^{l-1} c_{i-l+j} L_j = -\sum_{j=0}^{l-1} \sum_{k=1}^{l} A_k S_k^{i-l+j} L_j \\
&= -\sum_{k=1}^{l} A_k S_k^{i-l} \sum_{j=0}^{l-1} S_k^j L_j = \sum_{k=1}^{l} A_k S_k^i.
\end{aligned}
$$

Therefore, the system of equations (2.5) is satisfied.

Example 3.1. Let $n = 2, l = 2$, $J = -1$, $c_k = \begin{bmatrix} 2 & k + 2^k \\ 2 & 1 \end{bmatrix}$, $k = 0, 1, \cdots$. We want to find the solution of system (2.5). Firstly, we find that the solution of (3.1) is

$$[L_0 \ L_1 \ L_2] = \begin{bmatrix} a & b \\ 0 & \frac{5}{4} \end{bmatrix} \ \begin{bmatrix} -1-a & \frac{1}{8} - b \\ 0 & -\frac{5}{2} \end{bmatrix} \ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

where a and b are free parameters. Compute the eigenvalues of C_{L^T}, we have

$$\lambda_1 = 1, \quad \lambda_2 = a, \quad \lambda_3 = \frac{5 + \sqrt{5}}{4}, \quad \lambda_4 = \frac{5 - \sqrt{5}}{4}.$$

Then by computing the eigenvectors of C_{L^T} we get $C_{L^T} = SJS^{-1}$ with

$$S = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 1 & y & 1 & 1 \\ 2 & a & 0 & 0 \\ 1 & ay & \lambda_3 & \lambda_4 \end{bmatrix}, \quad J = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{bmatrix},$$

where $y = -\dfrac{a(\frac{1}{8} - b) + b}{(a - \lambda_3)(a - \lambda_4)}$, and $a \neq 1$. Since the matrix $\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$ is singular, we exchange the column in S and J and get two succeed partitions as follows:

$$
\begin{aligned}
X = [X_1, X_2] &= \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} \ \begin{bmatrix} 1 & 0 \\ y & 1 \end{bmatrix}, \\
J = \text{diag}[J_1, J_2] &= \text{diag}\left[\begin{bmatrix} 1 & 0 \\ 0 & \lambda_3 \end{bmatrix}, \begin{bmatrix} a & 0 \\ 0 & \lambda_4 \end{bmatrix} \right],
\end{aligned}
\tag{3.7}
$$

and

$$
\begin{aligned}
X = [X_1, X_2] &= \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} \ \begin{bmatrix} 1 & 0 \\ y & 1 \end{bmatrix}, \\
J = \text{diag}[J_1, J_2] &= \text{diag}\left[\begin{bmatrix} 1 & 0 \\ 0 & \lambda_4 \end{bmatrix}, \begin{bmatrix} a & 0 \\ 0 & \lambda_3 \end{bmatrix} \right].
\end{aligned}
$$

For partition (3.7) we get

$$S_1 = (X_1 J_1 X_1^{-1})^T = \begin{bmatrix} 1 & \frac{1}{2}(1 - \lambda_3) \\ 0 & \lambda_3 \end{bmatrix}, \quad S_2 = (X_2 J_2 X_2^{-1})^T = \begin{bmatrix} a & (a - \lambda_4)y \\ 0 & \lambda_4 \end{bmatrix}. \tag{3.8}$$

From (3.6) we find

$$A_1 = \begin{bmatrix} 2 & 1 - \frac{2}{\lambda_4 - \lambda_3} \\ 2 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & \frac{2}{\lambda_4 - \lambda_3} \\ 0 & 0 \end{bmatrix}. \tag{3.9}$$

The following result is a corollary of Theorem 3.2.

Theorem 3.3. *If $L(z)$ has only linear elementary divisors, then $L(z)$ has a complete set of left solvents.*

Proof. Let (X, J) be a Jordan pair of $L(z)$. Then Q defined by (3.3) is nonsingular and $J = \text{diag}[\lambda_1, \cdots, \lambda_{nl}]$ is a diagonal matrix. If we expand $\det Q$ by its last n rows, then by Laplace's theorem, there exists at least one $n \times n$ minor of $\det Q$:

$$Q \begin{pmatrix} n(l-1)+1, & n(l-1)+2, & \cdots, & nl \\ j_1, & j_2, & \cdots, & j_n \end{pmatrix}$$

such that both this minor and its complementary minor are non-zero. By exchanging the columns of Q, we may assume $j_i = i$, $i = 1, 2, \cdots, n$. Then

$$X = [X_1 \ \ \tilde{X}], \quad J = \text{diag}[J_1, \ \tilde{J}], \quad X_1, J_1 \in \mathbb{C}^{n \times n}$$

and

$$Q = \begin{bmatrix} X_1 & \tilde{X} \\ X_1 J_1 & \tilde{X}\tilde{J} \\ \cdots & \cdots \\ X_1 J_1^{l-1} & \tilde{X}\tilde{J}^{l-1} \end{bmatrix}$$

with $X_1 J_1^{l-1}$ and $\tilde{Q} = \left[\tilde{X}^T \ \ (\tilde{X}\tilde{J})^T \ \ \cdots \ \ (\tilde{X}\tilde{J}^{l-2})^T \right]^T$ being nonsingular. By induction, we can get that $X_i J_i^{l-i} \in \mathbb{C}^{n \times n}$ for $i = 1, \cdots, l$ and $X_i J_i^{l-i}$ are invertible. Hence $X_1, \cdots, X_l$ are invertible. Therefore, we have achieve a partition (by exchanging the column of Q) of $X = [X_1 \ \cdots \ X_l]$ and $J = \text{diag}[J_1, \cdots, J_l]$ such that $X_1, \cdots, X_l$ are nonsingular. Then the theorem follows from Theorem 3.2.

4. Relation between $P(z)$ and $R(z)$

For the scalar case, the (0,1)-type fraction-sum can be written as a rational function. Hence the problem of (0,1)-type fraction-sum approximation and the problem of Padé approximation are equivalent. However, for the matrix case, the relationship between the two approximations are complicated. This section illustrates this complexity.

Let $P(z)$ be a solution of problem (2.7). Then the parameters A_i, B_i and S_i in (2.6) satisfy equations (2.5). Suppose

$$\text{rank} \begin{bmatrix} I & S_1 & \cdots & S_1^{l-1} \\ \cdots & \cdots & \cdots & \cdots \\ I & S_l & \cdots & S_l^{l-1} \end{bmatrix} = \text{rank} \begin{bmatrix} I & S_1 & \cdots & S_1^{l} \\ \cdots & \cdots & \cdots & \cdots \\ I & S_l & \cdots & S_l^{l} \end{bmatrix}. \tag{4.1}$$

Then there exist $L_0, \ L_1, \ \cdots, L_{l-1}$ such that

$$\begin{bmatrix} I & S_1 & \cdots & S_1^{l-1} \\ \cdots & \cdots & \cdots & \cdots \\ I & S_l & \cdots & S_l^{l-1} \end{bmatrix} \begin{bmatrix} L_0 \\ \cdots \\ L_{l-1} \end{bmatrix} = - \begin{bmatrix} S_1^l \\ \cdots \\ S_l^l \end{bmatrix}.$$

Let $L(z) = \sum_{i=0}^{l} L_i z^i$ with $L_l = I$. Then the left value $L(S_j) = \sum_{i=0}^{l} S_j^i L_i$ of $L(z)$ is a zero matrix. Therefore, there exists a monic matrix polynomial $q_j(z)$ of degree $l-1$ for each $j = 1, 2, \cdots, l$, such that $L(z) = (zI - S_j)q_j(z)$ (see [6], page 252). Let $z = t^{-1}$ and

$$\tilde{L}(t) = t^l L(t^{-1}), \quad \tilde{q}_j(t) = t^{l-1} q_j(t^{-1}).$$

Then

$$\tilde{L}(t) = (I - S_j t)\tilde{q}_j(t), \quad \text{or} \quad (I - S_j t)^{-1} = \tilde{q}_j(t)\tilde{L}(t)^{-1}$$

in the neighborhood of the origin. It follows from (2.6) that $P(z)$ can be expressed as $P(z) = \left(\sum_{i=0}^{J+l} b_i z^i \right) \tilde{L}(z)^{-1}$. Therefore, $P(z)$ is a right matrix Padé approximant under condition (4.1).

Now we should mention that (4.1) is always true in the scalar case. However, in the matrix case, (4.1) may not be true. Hence the problem of (0,1)-type fraction-sum approximation and that of Padé approximation are not the same. Condition (4.1) is sufficient for $P(z) = R(z)$. A further conclusion is the following theorem.

Theorem 4.1. Let $P(z) = \sum_{j=0}^{J} B_i z^i + \sum_{j=1}^{l} A_j (I - zS_j)^{-1}$ be a (0,1)-type fraction-sum approximant of $f(z)$. Then it is a Padé approximant of $f(z)$ if and only if

$$
\begin{aligned}
& \mathrm{rank}\left\{ \begin{bmatrix} A_1 S_1^{J+1} & A_2 S_2^{J+1} & \cdots & A_l S_l^{J+1} \\ A_1 S_1^{J+2} & A_2 S_2^{J+2} & \cdots & A_l S_l^{J+2} \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix} \begin{bmatrix} I & S_1 & \cdots & S_1^{l-1} \\ \cdots & \cdots & \cdots & \cdots \\ I & S_l & \cdots & S_l^{l-1} \end{bmatrix} \right\} \\
& = \mathrm{rank}\left\{ \begin{bmatrix} A_1 S_1^{J+1} & A_2 S_2^{J+1} & \cdots & A_l S_l^{J+1} \\ A_1 S_1^{J+2} & A_2 S_2^{J+2} & \cdots & A_l S_l^{J+2} \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix} \begin{bmatrix} I & S_1 & \cdots & S_1^{l} \\ \cdots & \cdots & \cdots & \cdots \\ I & S_l & \cdots & S_l^{l} \end{bmatrix} \right\}.
\end{aligned}
\tag{4.2}
$$

Proof. Let $P(z) = \sum_{i=0}^{\infty} \tilde{c}_i z^i$. Then by the definition of $P(z)$ and (2.8) we have

$$
\tilde{c}_i = c_i, \qquad i = 0, 1, \cdots, 2l + J,
$$
$$
\tilde{c}_i = \sum_{j=1}^{l} A_j S_j^i, \quad i = J+1, J+2, \cdots.
\tag{4.3}
$$

It follows from (4.2) and the equalities above that (3.2) holds and therefore $R(z)$ exists (see [10], Theorem 2.1). Since (4.2) could be written as $H(J+1, l, \infty) = \mathrm{rank}\, H(J+1, l+1, \infty)$, then there exists $\tilde{L}_0, \tilde{L}_1, \cdots, \tilde{L}_l$ (with $\tilde{L}_0 = I$), such that

$$
H(J+1, l, \infty) \begin{bmatrix} \tilde{L}_l \\ \vdots \\ \tilde{L}_1 \end{bmatrix} = - \begin{bmatrix} \tilde{c}_{J+l+1} \\ \tilde{c}_{J+l+2} \\ \vdots \end{bmatrix}.
\tag{4.4}
$$

Let

$$
\begin{bmatrix} b_0 \\ \vdots \\ b_{J+l} \end{bmatrix} = H(-l, l+1, J+l+1) \begin{bmatrix} \tilde{L}_l \\ \vdots \\ \tilde{L}_0 \end{bmatrix},
$$

where we assume $c_i = 0$ if $i < 0$. Then $R(z) := \left(\sum_{i=0}^{J+l} b_i z^i \right) \left(\sum_{i=0}^{l} \tilde{L}_i z^i \right)^{-1}$ is a right matrix Padé approximant and (4.4) implies that the power series expansion of $R(z)$ is the same as that of $P(z)$. This complete the proof of the sufficient part of the theorem. If $P(z) = \left(\sum_{i=0}^{J+l} b_i z^i \right) \tilde{L}(z)^{-1}$, then by $P(z)\tilde{L}(z) = \sum_{i=0}^{J+l} b_i z^i$ and (4.3) we can get (4.2).

It is obvious that (4.1) implies (4.2). However, the inverse conclusion is not valid. The reason to mention condition (4.1) is not only that we know a method how to transform $P(z)$ to $R(z)$ under condition (4.1), but also that condition (4.1) is easier to use than (4.2), which relates to the ranks of infinite matrices.

Example 4.1. Let $n = 2$, $l = 2$, $c_k = \begin{bmatrix} 2 & k + v2^k \\ u & 1 + w2^k \end{bmatrix}$, $k = 0, 1, \cdots$, where u, v and w are parameters. Let

$$
S_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad S_2 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.
$$

Then

$$
\mathrm{rank} \begin{bmatrix} I & S_1 \\ I & S_2 \end{bmatrix} = 3, \quad \mathrm{rank} \begin{bmatrix} I & S_1 & S_1^2 \\ I & S_2 & S_2^2 \end{bmatrix} = 4.
$$

Hence equality (4.1) does not hold. Let $f(z) = \sum_{i=0}^{\infty} c_i z^i$. Then it is easy to check that for $J = -1$, $P(z) = (I - zS_1)^{-1} + A(I - zS_2)^{-1}$ is a (0,1)-type fraction-sum approximant of $f(z)$,

where $A = \begin{bmatrix} 1 & v \\ u & w \end{bmatrix}$. In fact, $P(z)$ and $f(z)$ have the same power series expansion. On the other hand, by computing the rank of the related matrices, we have

1. If $u = 0$, then (4.2) is true.

2. If $u \neq 0$, $w \neq 0$; or $u \neq 0$, $w = 0$, $2v(u + 1) + u = 0$, the relation

$$\mathrm{rank} \begin{bmatrix} c_0 & c_1 & c_2 \\ c_1 & c_2 & c_3 \end{bmatrix} = \mathrm{rank} \begin{bmatrix} c_0 & c_1 \\ c_1 & c_2 \end{bmatrix} \tag{4.5}$$

   does not hold.

3. For $u \neq 0$, $w = 0$, $v = 0$, (4.2) is true.

4. For $u \neq 0$, $w = 0$, $2v(u + 1) + u \neq 0$, $v \neq 0$, (4.2) does not hold, but (4.5) holds. The discussions above are summarized in the Fig 4.1.
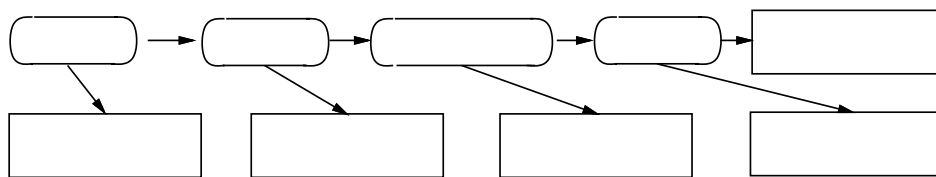

Fig 4.1: The relations between $P(z)$ and $R(z)$ for Example 4.1

This example shows that these two problems considered are very different.

Although $P(z) = R(z)$ in certain cases, we do not know how to get $R(z)$ from $P(z)$ and vice versa. For example, for $u = 0$, $w = 1$,

$$R(z) = N(z)L^{-1}(z) = \left(\begin{bmatrix} 2 & v \\ 0 & 2 \end{bmatrix} + z \begin{bmatrix} -2a & 2 - 2b - v \\ 0 & -3 \end{bmatrix} \right) *$$
$$\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + z \begin{bmatrix} -a - 1 & 0.5 - b \\ 0 & -3 \end{bmatrix} + z^2 \begin{bmatrix} a & b \\ 0 & 2 \end{bmatrix} \right)^{-1},$$

where a and b are free parameters. It seems difficult to get this form from $P(z)$. On the other hand, if we let $\tilde{L}(t) = t^2 L(t^{-1})$. Then the left values of $\tilde{L}(t)$ at S_1 and S_2 are

$$\tilde{L}(S_1) = \begin{bmatrix} 0 & -0.5 \\ 0 & 0 \end{bmatrix}, \quad \tilde{L}(S_2) = \begin{bmatrix} 0 & 0.5 \\ 0 & 0 \end{bmatrix},$$

respectively. That is, S_1 and S_2 are not left solvents of $\tilde{L}(t)$. This implies that S_1 and S_2 can not be found by the methods given in §3.

Let $u = 2$, $v = 1$, $w = 0$. Then by Example 3.1, we get another solution of (0,1)-type fraction-sum interpolation problem as follows

$$P_1(z) = A_1(I - zS_1)^{-1} + A_2(I - zS_2)^{-1},$$

where A_1, A_2, S_1 and S_2 are given in (3.8)–(3.9). Since (4.1) holds in this case, $P_1(z)$ is also the solution of Padé problem. But $\left. \frac{\partial^k P_1(z)}{\partial z} \right|_{z=0} \neq \left. \frac{\partial^k P(z)}{\partial z} \right|_{z=0}$ for $k = 4, 5, \cdots$, since

$$\left. \frac{\partial^k P_1(z)}{\partial z} \right|_{z=0} = k! \begin{bmatrix} 2 & 1 + 2(\lambda_4^k - \lambda_3^k) \\ 2 & 1 \end{bmatrix}, \quad \left. \frac{\partial^k P(z)}{\partial z} \right|_{z=0} = k! \begin{bmatrix} 2 & k + 2^k \\ 2 & 1 \end{bmatrix}, \quad k = 4, 5, \cdots.$$

5. Uniqueness

We consider the uniqueness of the approximants defined in §2 in the following sense:

a. For any solution $F(z)$ of (2.1), $F(iT)$ is uniquely defined for any $i = 0, 1, \cdots$.

b. For any solution of (2.4) or (2.7) or (2.13) or (2.15), its series expansion is unique.

It follows from the definition of uniqueness that the uniqueness of the approximants defined in §2 is equivalent to

$$\sum_{j=1}^{l} A_j S_j^i \quad \text{are uniquely defined for } i > 2l + J \tag{5.1}$$

and for any solution of system (2.5). Since (2.5) is a system of nonlinear equations, it may have many solutions. We have noticed in Example 4.1 that, if $u = 2$, $v = 1$ and $w = 0$, the (0,1)-type fraction-sum approximation problem has at least two different solutions $P(z)$ and $P_1(z)$. But the solution of Padé problem (2.10), which is $P_1(z)$, is unique since condition (5.2) holds. What we have achieved now about the uniqueness is that we can establish uniqueness conclusion about these solutions of (2.5) which make $P(z)$ to be a solution of the Padé problem.

Theorem 5.1. *Suppose that system (2.5) is solvable. Then its solution, which makes (4.2) valid, is unique in the sense of (5.1) if and only if*

$$\text{rank } H(J + 1, l, l) = \text{ rank } H(J + 1, l, l + 1). \tag{5.2}$$

Proof. Let $f(z) = \sum_{i=0}^{\infty} c_i z^i$ and c_i be given in each problem proposed in §2. Then it follows from the validity of (4.2) that the right matrix Padé approximant of $f(z)$ exists. It is unique (see [10]) if and only if (5.2) holds. The uniqueness of Padé approximant and (4.2) is equivalent to (5.1).

Corollary 5.1. *The solution of (2.5) obtained from section 3 is unique if and only if (5.2) holds.*

Proof. Since the solution given in §3 always satisfy relation (4.1), and (4.1) implies (4.2), the corollary follows from the theorem.

6. The Convergence of Baker-Gammel Approximants

We establish now a convergence conclusion about $K_l(z) = \sum_{j=1}^{l} A_j k(z, S_j)$ defined in §2 $(J = -1)$ for a special class of functions. The conclusion is established based on the convergence results about matrix Padé approximants for Stieltjes functions.

Lemma 6.1. (see [4]) *Let $f(z) = \sum_{i=0}^{\infty} c_i z^i$, $c_i \in \mathbb{R}^{n \times n}$ be symmetric matrix. Suppose*

$$H(0, l, l) \quad \text{and} \quad H(1, l, l) \text{ are positive defined for all } l. \tag{6.1}$$

Let $N(z)\tilde{L}^{-1}(z)$ be the right $(l - 1, l)$ type matrix Padé approximants of $f(z)$. Then

(i) The zeros of $\det(z^l \tilde{L}(z^{-1}))$ are real and positive.

(ii) All the elementary divisors of $z^l \tilde{L}(z^{-1})$ are linear.

Lemma 6.2. (see [8]) *Let $\sigma(x)$ be a bounded, real symmetric and increasing matrix-valued function. Then if $\sigma(x)$ is constant for $x \geq 1$*

$$\limsup_{l \to \infty} \|R_l(z) - f(z)\|^{\frac{1}{2l}} \leq \frac{|\sqrt{1 - z} - 1|}{|\sqrt{1 - z} + 1|}, \quad z \notin [1, +\infty),$$

where $R_l(z)$ is the right $(l - 1, l)$ type matrix Padé approximant of $f(z)$ and $f(z)$ is defined by

$$f(z) = \int_0^1 \frac{1}{1 - zt} d\sigma(t), \tag{6.2}$$

$\| \cdot \|$ *stands for the Frobenius norm for matrix and* $\sqrt{}$ *denotes the principal branch of the square root.*

From these lemmas, we can get our theorem about the convergence of $K_l(z)$.

Theorem 6.1. *Let* $\sigma(x)$ *be a measure defined as in Lemma 6.2. Assume* $\sigma(x)$ *is constant for* $x \geq 1$ *and (6.1) holds for* $c_i = \int_0^1 x^i d\sigma(x)$, $i = 0, 1, \cdots$. *Let*

$$g(z) = \int_0^\infty k(z, x) d\sigma(x) = \int_0^1 k(z, x) d\sigma(x) = \sum_{i=0}^\infty c_i k_i(z),$$

where $k(z, u)$ *is an analytic function of* u *at least within a distance* $\Delta > 0$ *from* $[0, +\infty)$ *for all* $z \in \mathcal{D}$, *a compact region in the* $z-$*plane. Furthermore let*

$$\int_\Gamma \frac{|k(z, \omega)|}{|\omega|} d\omega \leq M < \infty, \quad z \in \mathcal{D},$$

where Γ *is a contour at distance* Δ *from* $[0, +\infty)$. *Then*

(i). *there exist* A_i, $S_i(i = 1, \cdots, l)$ *such that (4.1) holds,* $K_l(z) = \sum_{i=1}^l A_i k(z, S_j)$ *is a solution of problem (2.16) and the solution which make (4.2) holds is unique.*

(ii). *for* $K_l(z)$, *a solution of (2.16) with* $S_1, \cdots, S_l$ *satisfy (4.1), we have*

$$\limsup_{l \to \infty} \|K_l(z) - g(z)\|^{\frac{1}{2l}} \leq \max_{\omega \in \Gamma} \frac{|\sqrt{1 - \omega^{-1}} - 1|}{|\sqrt{1 - \omega^{-1}} + 1|}, \quad z \in \mathcal{D}. \tag{6.3}$$

Proof. (i). Under the assumption of the theorem, the right $(l - 1, l)$ type matrix Padé approximant $R_l(z) = N(z)\tilde{L}^{-1}(z)$ exists for $f(z)$ defined by (6.2). It follows from Lemma 6.1 and Theorem 3.2 that $L(z) = z^l \tilde{L}^{-1}(z^{-1})$ has a complete set of left solvents $S_1, \cdots, S_l$. Therefore, problem (2.16) is solvable. Furthermore, since (4.1) and (5.2) hold, $K_l(z)$ is unique for those $S_1, \cdots, S_l$, which satisfy (4.1).

(ii). Since $\bigcup_{j=1}^l \sigma(S_j) = \sigma(L)$ (see [6], p.524), we have by Lemma 6.1 that the eigenvalues of S_j are real and positive. Then we have (see [6], p.332)

$$
\begin{aligned}
K_l(z) \quad &= \sum_{j=1}^l A_j k(z, S_j) = \frac{1}{2\pi i} \sum_{j=1}^l A_j \int_\Gamma k(z, \omega)(\omega I - S_j)^{-1} d\omega \\
&= \frac{1}{2\pi i} \int_\Gamma \frac{k(z, \omega)}{\omega} \sum_{j=1}^l A_j \left(I - \omega^{-1} S_j\right)^{-1} d\omega = \frac{1}{2\pi i} \int_\Gamma \frac{k(z, \omega)}{\omega} P_l(\omega^{-1}) d\omega \\
&= \frac{1}{2\pi i} \int_\Gamma \frac{k(z, \omega)}{\omega} R_l(\omega^{-1}) d\omega.
\end{aligned}
$$

The last equality follows from Theorem 4.1.

On the other hand, by Cauchy formulas

$$g(z) = \int_0^\infty \left(\frac{1}{2\pi i} \int_{\Gamma_1} \frac{k(z, \omega)}{\omega - x} d\omega\right) d\sigma(x),$$

where Γ_1 is a circular contour with center x and radius Δ. It follows from the analyticity of $k(z, \omega)$ that

$$g(z) = \frac{1}{2\pi i} \int_\Gamma \frac{k(z, \omega)}{\omega} \left(\int_0^\infty \frac{1}{1 - x\omega^{-1}} d\sigma(x)\right) d\omega = \frac{1}{2\pi i} \int_\Gamma \frac{k(z, \omega)}{\omega} f(\omega^{-1}) d\omega.$$

Hence

$$
\begin{aligned}
\|g(z) - K_l(z)\| \quad &= \left\| \frac{1}{2\pi i} \int_\Gamma \frac{k(z, \omega)}{\omega} \left[f(\omega^{-1}) - R_l(\omega^{-1})\right] d\omega \right\| \\
&\leq \frac{1}{2\pi} \max_{\omega \in \Gamma} \|f(\omega^{-1}) - R_l(\omega^{-1})\| \int_\Gamma \frac{|k(z, \omega)|}{|\omega|} d\omega.
\end{aligned}
$$

Therefore,

$$\limsup_{l\to\infty} \|g(z) - K_l(z)\| \le \frac{M}{2\pi} \limsup_{l\to\infty} \max_{\omega\in\Gamma} \|f(\omega^{-1}) - R_l(\omega^{-1})\|$$

and (6.3) follows from Lemma 6.2.

Since $|\sqrt{1-z} - 1|/|\sqrt{1-z} + 1| = 1$ if and only if z is real and $z \ge 1$, we have

$$\max_{\omega\in\Gamma} \frac{|\sqrt{1-\omega^{-1}} - 1|}{|\sqrt{1-\omega^{-1}} + 1|} = r < 1.$$

Therefore, we have

Corollary 6.1. *Under the conditions of Theorem 6.1. $K_l(z)$ converge uniformly to $g(z)$ on $\mathcal{D}$ as $l \to \infty$.*

7. Appendix: Some Terminologies and Facts

In this appendix, we introduce some used terminologies and basic facts in the matrix polynomial theory, more related details could be found in [6].

Matrix Polynomial. A *matrix polynomial* $L(z)$ is matrix whose elements are polynomials in z. An $n \times n$ matrix polynomial $L(z)$ of degree l over the complex field $\mathbb{C}$ could be written as

$$L(z) = L_0 + zL_1 + \cdots + z^l L_l, \quad L_i \in \mathbb{C}^{n\times n}$$

where l is the degree of $L(z)$. If $L_l = I$, $L(z)$ is called monic.

Solvent. Let $S \in \mathbb{C}^{n\times n}$. The *left value* of $L(z)$ at S is defined as

$$L(S) = L_0 + SL_1 + \cdots + S^l L_l.$$

If $L(S) = 0$, then S is referred to as a *left solvent* of $L(z)$. $L(z)$ is divisible on the left by $zI - S$ with zero remainder if and only if S is a left solvent of $L(z)$. A set of left solvents $S_1, \cdots, S_l$ of $L(z)$ is *complete* if the corresponding Vandermonde matrix

$$V = \begin{bmatrix} I & S_1 & \cdots & S_1^{l-1} \\ \cdots & \cdots & \cdots & \cdots \\ I & S_l & \cdots & S_l^{l-1} \end{bmatrix}$$

is nonsingular (see [6], p.524).

Latent Roots. Let $L(\lambda)$ be an $n \times n$ matrix polynomial, the zeros of $\det L(\lambda)$ are said to be *latent roots* of $L(\lambda)$.

Elemntary Divisors. Suppose that the $n \times n$ matrix polynomial $L(\lambda)$ over $\mathbb{C}$ has rank r and let $d_j(\lambda)$ be the greatest common divisor of all minors of $L(\lambda)$ of order j, $j = 1, \cdots, r$. If we define $d_0(\lambda) \equiv 1$, then $d_j(\lambda)$ is divisible by $d_{j-1}(\lambda)$, $j = 1, \cdots, r$. The quotients $i_j(\lambda) = d_j(\lambda)/d_{j-1}(\lambda)$ are called *invariant polynomials* of $L(\lambda)$. Let $\lambda_1, \lambda_2, \cdots, \lambda_s$ be the distinct latent roots of $L(\lambda)$ with multiplicities $m_1, m_2, \cdots, m_s$. Then there are integers α_{jk}, $1 \le j \le r$ and $1 \le k \le s$, such that

$$i_j(\lambda) = (\lambda - \lambda_1)^{\alpha_{j1}}(\lambda - \lambda_2)^{\alpha_{j2}} \cdots (\lambda - \lambda_s)^{\alpha_{js}}, \quad j = 1, \cdots, r \tag{7.1}$$

and $0 \le \alpha_{1k} \le \alpha_{2k} \le \cdots \le \alpha_{rk} \le m_k$, $\sum_{j=1}^{r} \alpha_{jk} = m_k$ for $k = 1, \cdots, s$. Each factor $(\lambda - \lambda_k)^{\alpha_{jk}}$ appearing in the factorizations (7.1) with $\alpha_{jk} > 0$ is called an *elementary divisor* of $L(\lambda)$. An elementary divisor for which $\alpha_{jk} = 1$ is said to be *linear*, otherwise it is *nonlinear*.

Companion Matrix. Let $L(z)$ be an $n \times n$ monic matrix polynomial of degree l. Then

$$C_L = \begin{bmatrix} 0 & I & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & I \\ -L_0 & -L_1 & -L_2 & \cdots & -L_{l-1} \end{bmatrix}.$$

is said to be the *first companion matrix* for $L(z)$. It is known that all of the elementary divisors of $L(z)$ and $zI - C_L$ coincide.

Jordan Pair. Let $L(z)$ be an $n \times n$ monic matrix polynomial of degree l. (X, J) is a *Jordan pair* of $L(z)$ if J is the Jordan canonical form of C_L (i.e., $C_L = SJS^{-1}$), and $X \in \mathbb{C}^{n \times nl}$ consists of the first n rows of S, here C_L is the first companion matrix for $L(z)$. Let (X, J) be a Jordan pair for $L(z)$. Then (see [6], p.500) the $nl \times nl$ matrix $Q = \begin{bmatrix} X^T & (XJ)^T & \cdots & (XJ^{l-1})^T \end{bmatrix}^T$ is nonsingular.

Matrix Value of Scalar Function. For a scalar function $g(z) \in \mathbb{C}[z]$, the matrix value $g(A)$ in terms of matrix $A \in \mathbb{C}^{n \times n}$ and function g is defined as $g(A) := p(A)$, where $p(z)$ is any polynomial that assume the same values as $g(z)$ on the spectrum $\sigma(A) = \{\lambda_1, \cdots, \lambda_s\}$. One computation method for $g(A)$, that uses the Jordan canonical form of A, can be found in [6], p.311. Another approach for computing $g(A)$ is to use the spectral resolution of $g(A)$ (see [6], p.314). An explicit formula for $g(A)$ that do not use the spectrum of A can be found in [5].

References

[1] G. A. Baker, The Padé approximant method and some related generalizations, In G. A. Baker and J. L. Gammel, editors, The Padé approximant in Theoretical Physics, Academic Press, New York, 1970.

[2] G. A. Baker and J. L. Gammel, The Padé approximant, *J. Math., Anal. Appl.*, **2** (1961), 21-30.

[3] G. A. Baker and P. Graves-Morris, Padé Approximants, Part II: Extension and Applications, Addison-Weslely Publishing Company, London, Amsterdam, 1981.

[4] S. Basu and N. K. Bose, Matrix Stieltjes series and network models, *SIAM Math. Anal.*, **14**:2 (1983), 209–222.

[5] M. Bruschi and P. E. Ricci, An explicit formula for $f(\mathcal{A})$ and the generating functions of the generalized Lucas polynomials, *SIAM Math. Anal.*, **13**:1 (1982), 162–165.

[6] P. Lancaster and M. Tismenetsky, The Theory of Matrices second Edition with Application, Academic Press, New York, London, second edition, 1985.

[7] R. D. Small, The generating function method of nonlinear approximation, *SIAM Numer. Anal.*, **25**:1 (1988), 235–244.

[8] B. V. Sydow, Padé approximation of matrix-values series of Stieltjes, *Ark. Mat.*, **15** (1977), 200–210.

[9] N. Wiener and P. Masani, The prediction theory of multivariate stochastic processes, Part I, *Acta. Math.*, **98** (1957), 111–157.

[10] G. Xu, The Existence and Uniqueness of Matrix Padé Approximants, *J. Comp. Math.*, **8**:1 (1990), 65–74.