CANADIAN THESES

THÈSES CANADIENNES

The University of Alberta

Matrix Padé Approximants

by

George Labahn

A thesis

submitted to the Faculty of Graduate Studies and Research

in partial fulfillment of the requirements for the degree of

Master of Sciences

Department of Computing Science

Edmonton, Alberta

Spring, 1986

ISBN 0-315-30278-X

The University of Alberta

Release Form

Name of Author: George Labahn

Title of Thesis: *Matrix Padé Aproximants*

Degree for Which Thesis was Presented: Master of Sciences

Year This Degree Granted: 1986

George Labahn

Permanent Address: 15612 - 110 Avenue
Edmonton, Alberta
Canada

Dated 3 February 1986

The University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, ~~and~~ recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled *Matrix Padé Approximants* submitted by **George Labahn** in partial fulfillment of the requirements for the degree of Master of Sciences

....
Supervisor

....

....

....

Date ...Feb 12/86.......

Abstract

We study the notion of rational matrix approximants to a formal matrix power series. Various definitions are discussed and a new definition is provided. Solutions are characterized and an algorithm to calculate matrix rational approximants to certain formal matrix power series is given. The algorithm is an extension of one given by Cabay and Choi for calculating scaled Padé fractions for a scalar formal power series. It is found to be less restrictive than existing algorithms. Costs are calculated and compared with other known computational methods. The algorithm turns out to be of lower complexity than existing algorithms.

The relationship between the matrix Padé fractions and the Euclidean algorithm for calculating greatest common divisors of two matrix polynomials is also studied. Two other algorithms are given, including one that provides a fast algorithm for situations where a pseudo Euclidean algorithm can be applied.

Acknowledgements

The author would like to take this opportunity to thank the people who helped make this work possible. As with any research project, numerous people were consulted at all stages of its development. Their time and often insightful criticism of this work sparked new ideas and made old ones clear. In particular, I would like to thank the following people.

My first acknowledgements must go to my supervisor Stan Cabay. During the year I spent under his supervision I grew both academically and professionally. May everyone have such a helpful supervisor.

Second, I would like to thank the members of my examining committee for their time spent in reviewing this thesis.

Without the resources and technical support provided by The Department of Computing Science at the University of Alberta this thesis would not have been possible. In addition a special thanks goes to the Department of Mathematics at the University of Alberta for their support and kindness over the years. Thanks also goes to John Kmet for his job of proof reading and to Martin Dubetz for help with the formatting of this thesis.

Finally, none of this would have been possible without the unlimited patience and kindness of Laura Leslie and the ever present encouragement of my parents.

Table of Contents

Chapter 1

Introduction

Formal power series are expressions of the form

$$A(z) = \sum_{i=0}^{\infty} a_i z^i ,$$

where the coefficients are from some algebraic structure (e.g., a field). The term formal comes from the fact that we are not concerned about the convergence of the power series. The Padé Table of such a power series is a doubly infinite array of rational expressions

$$\frac{U_{m/n}(z)}{V_{m/n}(z)} = \frac{\sum_{i=0}^{m} u_i z^i}{\sum_{i=0}^{n} v_i z^i} .$$

These rational expressions are to agree with as many terms in A(z) as possible. The power series A(z) is said to be normal if, for every such pair (m,n), the agreement is exact through to the z^{m+n} term.

The foundations for Padé theory of rational approximations were laid by Cauchy (1821) in the work "Cours d'Analyse". The concept of Padé table was due to Frobenius [18] who developed the basic algorithmic aspects of the theory. The reference to Padé [24] , whose work followed that of Frobenius, came from his treatment of some abnormal cases arising in the table.

Formal power series and their Padé tables have many applications in mathematics and engineering-related disciplines. Applications include numerical computations for special power series (e.g., Gamma function) (cf Nemeth and Zimanyi [23]), algorithms in the field of numerical analysis (cf Gragg [19]), triangulation of Block Hankel and Toeplitz matrices (cf Rissanen [27]), solving linear systems of equations with Hankel or Toeplitz coefficient matrices (cf Rissanen [25]), in digital filtering theory (cf

Bultheel [12] or Brophy and Salazar [9]) and linear control theory (cf Elgerd [17]).

As an illustration we can single out three areas where formal power series and related rational expressions come into play. In the first case, we can examine the theory of rational approximations in relation to the field of differential equations. There, one has an equation which constrains the behaviour of a particular function y(z) (the constraints may be due to physical laws). Thus, we are given an equation of the form

$$a(z)\cdot\frac{d^{(n)}y}{d^{(n)}z} + \cdots + b(z)\cdot\frac{dy}{dz} + c(z)\cdot y(z) = r(z) \qquad (1.1)$$

with certain initial conditions and wish to determine the function y(z). Numerous methods exist to determine y(z) for various special types of equations. One method used in the 19th century to determine numerical solutions was the power series method. Here one noticed that, as long as a(z) was not zero for a particular value of z, say $z = 0$, then we could divide equation (1.1) by a(z) and isolate $\frac{d^{(n)}y}{d^{(n)}z}$ in a neighborhood of 0. This then implies that the higher derivatives of y(z) exist near this point (given that the coefficients in the equation are analytic, e.g., if the coefficients are polynomials). This then implies that the Taylor expansion of y(z) exists and converges in a neighborhood about the point. One could then determine that y(z) is an infinite power series and calculate as many terms as necessary from equation (1.1), as long as the initial conditions are known.

Sometimes though, we face the inverse problem. Namely, if we have a function y(z) that we can put into the form of an infinite power series, then we want to know the constraints or controls that y(z) must satisfy. Thus, given that y(z) is a formal power series, we may ask for k+1 polynomials

$$P_k(z)\ ,\ P_{k-1}(z)\ ,\ \cdots\ P_1(z)\ ,\ P_0(z)$$

such that

$$P_k(z)\cdot\frac{d^{(k-1)}y}{d^{(k-1)}z} + \cdots + P_1(z)\cdot y(z) + P_0(z) = z^{p_k + \cdots + p_0 + k}\cdot R(z)$$

where R(z) is some power series. Here, p_i denotes the degree of $P_i(z)$. In the particular case where k = 1 , or in the case that all the polynomials except $P_0(z)$ are constant polynomials, the situation looks like

$$V_n(z)\cdot y(z) - U_m(z) = z^{m+n+1}\cdot R(z)$$

and the pair $U_m(z)$ and $V_n(z)$ is called a Padé form of type (m,n) for y(z). In the more general case where k > 1, the polynomials $P_i(z)$ form a Hermite-Padé approximant of y(z). (See for example Baker [3])

A second area where rational expressions have a connection comes up in the area of algebraic computation, more specifically in the general area of calculation of greatest common divisors of two polynomials. If we have two finite-degree polynomials A(z) and B(z), then we can form the formal power series $\frac{A(z)}{B(z)}$. If we find a rational approximation to this quotient, say of degrees m and n respectively, then we get the expression

$$\frac{A(z)}{B(z)} = \frac{U_m(z)}{V_n(z)} + z^{m+n+1}\cdot R(z).$$

We can rewrite this as

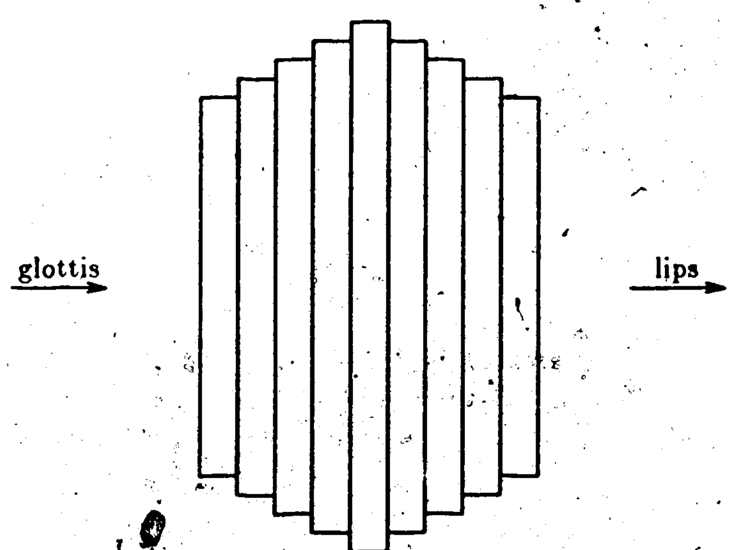
$$A(z)\cdot V_n(z) - B(z)\cdot U_m(z) = z^{m+n+1}\cdot R_{mn}(z)$$

If we replace the variable z by z^{-1} in the above expression and make a wise choice for m and n, then we can multiply both sides of the new equation by a particular power of z to get the new expression

$$A^*(z)\cdot V_n^*(z) - B^*(z)\cdot U_m^*(z) = R_{mn}^*(z)$$

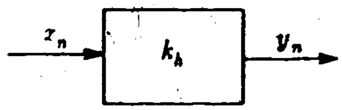
where the superscript * means to reverse the order of the coefficients in the polynomial. But then we have a linear combination of the two reversed polynomials. This is the sort of expression that comes up when applying Euclid's algorithm to the reversed polynomials. By a judicious choice of m and n, we can end up finding the greatest common divisor of two polynomials by calculating certain rational approximants of the power series obtained by reversing the order of the coefficients of the two polynomials and then obtaining the quotient power series. See chapter 5 for more details.

Another example that illustrates the use of formal power series and their rational approximations in a natural setting comes from the field of digital filtering. We consider the problem of modeling human speech production (Bultheel [12]). One such model visualizes the vocal tract as a sequence of coaxial cylinders (cf Wakita [33] or Stevens [31]).



At the left we have the glottis while at the right we have the lips. A sound such as a vowel is produced by a pulse train produced by the input (vocal chords). The vocal tract then produces a periodic soundwave at the output (lips).

In a time-series analyais of this system (cf Makhoul [21]), we sample our output and think of this model of the vocal tract as a discrete system with a discretized input x_n and discretized output y_n. The outputs are determined by convolution of the inputs with an impulse response k_k.



If we think of these three discrete sequences in terms of their z-transforms, i.e., as power series

$$X(z) = \sum x_h z^h \ , \ Y(z) = \sum y_h z^h \ , \ K(z) = \sum k_h z^h ,$$

then convolution of the impulse response function goes over to polynomial multiplication and we get

$$Y(z) = K(z) \cdot X(z) .$$

K(z) is called the transfer function of the system. Note that K(z) has no negative powers since the system we are modeling is causal, that is, there is no output before there is input. The coefficients of the transfer function are usually determined theoretically or experimentally by measuring the output of a signal from an input that isolates on the particular response.

If we now approximate K(z) by a rational expression we obtain a particular model for the vocal tract. For example, if the approximation to K(z) is of the form $\frac{c}{\sum_{i=0}^{n} v_i z^i}$ (with $v_0 = 1$), i.e., a rational approximation where the degrees are 0 and n, then we get an Autoregressive or AR filter (for larger degrees of the numerator, we have ARMA or Autoregressive-moving-average filters). With an AR filter the input and output are related by the form

$$y_t = c \cdot x_t - [v_1 y_{t-1} + \cdots + v_n \cdot y_{t-n}],$$

i.e., a system where the output at time t depends on the input at time t and the n previous outputs. Because of this, these filters are also known as predictive filters. For further information about predictive filters and their uses, see the article by Makhoul.

(To see how filters follow naturally from the model of the vocal tract as a series of coaxial cylinders, we refer the reader to the paper by Wakita [33]).

The previous example is a good illustration of a problem that generalizes to higher dimensions. We have dealt with a single-input-single-output system. Similar problems occur with a multi-input-multi-output system. The formal power series that results now has matrix coefficients. Again the problem of modeling the input-output system comes up and often we look to rational approximations of the form $V_n(z)^{-1} \cdot U_m(z)$ to give us approximations to the system in terms of input-output relationships.

Notice though that there are differences between the scalar case and the multidimensional case that appear even here. Because the rational approximations involve matrix polynomials, an expression of the form $V_n(z)^{-1} \cdot U_m(z)$ differs from an expression of the form $U_m(z) \cdot V_n(z)^{-1}$ (Approximating K(z) by $V_n(z)^{-1} \cdot U_m(z)$ results in the approximation $V_n(z)Y(z) = U_m(z)X(z)$ which relates the output vectors $y_t, \cdots, y_{t-n}$ to the input vectors $x_t, \cdots, x_{t-m}$. The other approximation, the right approximation, does not necessarily imply the same relation since matrices do not necessarily commute.)

The lack of commutativity of matrices is just one of the complications involved in approximating matrix power series by rational approximants. In the calculations to determine the rational approximants themselves, the fact that various coefficients might not be invertible causes problems that must be addressed.

Once there is a desire to approximate a formal power series by a rational expression (called Padé forms), there is the problem of how to go about determining the coefficients for the numerators and denominators. There are a number of algorithms that calculate Padé forms. In the one dimensional case, there are algorithms due to Wynn [34] , Watson and Brezinski [8] , and Rissanen [25]. Algorithms where the coefficients are matrices can be found in the papers of Bultheel [10] , Bose and Basu [6] , and Starkand [30] .

The problem that occurs with most of the above algorithms is that they limit their scope to the normal case . This is a severe restriction. For example, any power series with any zero coefficients is not normal.

One algorithm that does not require the normality restriction was discovered by Cabay and Choi [14]. This algorithm, called the offdiagonal algorithm, is more efficient than previous algorithms and has interesting consequences.

This thesis presents an algorithm to calculate matrix Padé forms, that extends the algorithm of Cabay and Choi for calculating Padé forms in the scalar case. It does not cover all matrix power series, unlike the scalar case. For example, we limit our scope to the situation where the input-ouput system has an equal number of outputs as inputs, i.e. we assume our matrices are square matrices. However our results are more general than those covered by the normal case (which also only deals with square matrices). In addition, it is more efficient in handling the case when the power series is normal than those of found in Bultheel, Bose et al , and Starkand. For example the matrix offdiagonal algorithm computes the Padé forms in complexity $O(n \cdot \log^2 n \cdot p^r)$, while the cost for the algorithms of Bose etal and Starkand are $O(n^3 \cdot p^r)$. Here p is the size of the matrices and p^r represents the cost of multiplying two $p \times p$ matrices. Under normal multiplication r is 3, while using Strassen's method reduces this to r is about 2.81. See, for example, Horowitz and Sahni [20] .

The rest of the thesis is broken into five chapters. Chapter 2 reviews the one dimensional case of Padé forms. A general definition for Padé form of a particular integer type is given and existence of these forms is demonstrated for all power series. Uniqueness is a problem and so we characterize the Padé forms of a power series. This is done by looking at the linear system that results from the order condition. This characterization deals separately with the nonsingular coefficient matrix case (i.e. the normal case) and the singular coefficient matrix case (i.e. the non-normal case). Because of this characterization a special type of Padé form, called the scaled Padé fraction, is identified. These scaled Padé fractions are shown to always exist and are unique up to multiplication by a nonzero coefficient.

Chapter 3 is concerned with the multidimensional case where the formal power series has matrix coefficients. A number of definitions exist to extend the notion of rational approximation to such a power series. We first introduce the broadest possible definition. The equations that result from this definition are examined and existence is proved for all power series and degree conditions. The lack of matrix commutativity and invertiblity is discussed here. Various examples of unexpected (and undesirable) behaviour are included in this chapter. These examples lead us to limit the class of formal matrix power series for which we will calculate Padé forms. This subclass of power series, called nearly normal power series, includes all scalar power series and all normal matrix power series. As in the scalar case, we characterize the Padé form problem for these nearly normal power series both in the normal case and the singular case. Scaled matrix Padé fractions are introduced and some of their properties are discussed. The chapter also discusses definitions of Padé forms used by other authors and comparisons between these definitions and ours are given.

Chapter 4 presents a new algorithm for obtaining the Padé approximant for nearly-normal power series in the multi-dimensional case. When the dimension of our system is one, the algorithm reduces to the offdiagonal algorithm of Cabay and Choi, though in a somewhat different set of steps. Why and how our algorithm works is described in detail. A multidimensional example is included. A detailed cost analysis is provided and results compared with other methods.

Chapter 5 discusses the relationship between the new algorithm and the general

Euclidean algorithm for calculating the greatest common divisor of two matrix polynomials. A second version of the algorithm is presented that displays a duality between the two algorithms. A comparison of costs involved in the calculation of the greatest common divisor via our algorithm and via other algorithms is provided. A third algorithm is also included to provide a fast algorithm of a dual version of a pseudo Euclidean Algorithm where certain invertibility conditions are relaxed.

The final chapter deals with the problems encountered in calculating the matrix Padé forms for arbitrary power series. Possible extensions and new directions of research topics are presented.

Chapter 2

One Dimensional Padé Fractions

2.1. Basic Definitions

Let A(z) be a formal power series with coefficients from some field K (the field is usually the reals or the set of complex numbers). We shall restrict the development to units of the set of formal power series, i.e. those power series having a non-zero constant term. As we shall see later in this chapter, this assumption is made for the sake of convenience since the results obtained can be applied to nonunit power series as well.

For power series we have a notion of degree and order. The degree of a formal matrix power series A(z) is denoted by the symbol $\partial(A)$ and is defined to be the power of z of the highest nonzero coefficient of A(z); if there is no highest power, the degree is set to ∞. The order of a formal matrix power series A(z) is denoted by the usual symbol: ord(A) and stands for the lowest power of z with a nonzero coefficient in A(z). In our situation ord(A) $\geq$ 0 for all power series in consideration.

Definition: 2.1. Let A(z) be a unit formal power series and let m and n be non-negative integers. Then a pair, $U_{m/n}(z)$, $V_{m/n}(z)$, of polynomials is defined to be a **Padé form** of type (m,n) to A(z) if

$$\begin{aligned} &a) \quad \partial(U_{m/n}) \leq m\,, \quad \partial(V_{m/n}) \leq n\,, \\ &b) \quad A(z)\cdot V_{m/n}(z) - U_{m/n}(z) = z^{m+n+1}\cdot R_{m/n}(z) \quad with \quad ord(R_{m/n}) \geq 0\,, \text{ and} \qquad (2.1) \\ &c) \quad V_{m/n}(z) \neq 0\,. \end{aligned}$$

The polynomials $V_{m/n}(z)$, $U_{m/n}(z)$, and $R_{m/n}(z)$ are usually called the denominator, numerator, and residual of the approximation (all of type (m,n)), respectively.

Suppose our power series and polynomials have the following expansions:

$$A(z) = \sum_{i=0}^{\infty} a_i z^i \ , \quad U_{m/n}(z) = \sum_{i=0}^{m} u_i z^i \ , \quad V_{m/n}(z) = \sum_{i=0}^{n} v_i z^i \ , \quad R_{m/n}(z) = \sum_{i=0}^{\infty} r_i z^i \ ,$$

where a_i, u_i, v_i, and r_i are all from the field K. We first mention a notational convenience that we will use throughout this thesis. Namely, for the polynomial

$$U(z) = u_0 + u_1 z + \cdots + u_k z^k$$

we write U , i.e., the same symbol but without the z variable, and the symbol U^- to mean the vectors of coefficients

$$U = \begin{bmatrix} u_k \\ \cdot \\ \cdot \\ u_1 \\ u_0 \end{bmatrix} \quad \text{and} \quad U^- = \begin{bmatrix} u_0 \\ u_1 \\ \cdot \\ \cdot \\ u_k \end{bmatrix}.$$

With the above notation, equation (2.1) can be written as the following system of equations

$$\begin{bmatrix} H^-_{m/n} \\ --- \\ H^+_{m/n} \end{bmatrix} \cdot V_{m/n} = \begin{bmatrix} U^-_{m/n} \\ --- \\ 0 \end{bmatrix}, \tag{2.2}$$

which is shorthand for

$$\begin{bmatrix} a_{-n} & \cdots & \cdots & a_0 \\ \cdots & & & a_1 \\ \cdots & & & \cdots \\ a_{m-n} & a_{m-n+1} & \cdots & a_m \\ \hline a_{m-n+1} & a_{m-n+2} & \cdots & a_{m+1} \\ a_{m-n+2} & a_{m-n+3} & & a_{m+2} \\ \cdots & \cdots & & \cdots \\ a_m & a_{m+1} & \cdots & a_{m+n} \end{bmatrix} \cdot \begin{bmatrix} v_n \\ v_{n-1} \\ \cdots \\ v_0 \end{bmatrix} = \begin{bmatrix} u_0 \\ u_1 \\ \cdots \\ u_m \\ --- \\ 0 \\ \cdots \\ \cdots \\ 0 \end{bmatrix}. \tag{2.3}$$

Here, $a_i = 0$ for $i < 0$.

It is clear that if we can solve the matrix system of equations

$$H^+_{m/n} V_{m/n} = 0 \tag{2.4}$$

for $V_{m/n}$, then we can get $U_{m/n}$ from

$$H^-_{m/n} \cdot V_{m/n} = U^-_{m/n} . \tag{2.5}$$

Thus, we concentrate on solving equation (2.4). A simple count reveals that our system of equations represented by (2.4) consists of n equations in n + 1 unknowns. An elementary result from linear algebra (see for example [1]) ensures that there is at least one non-zero solution to this system (since the nullity of the system must be ≥ 1). Thus, we have

Theorem 2.1 (Frobenius). For any formal unit power series A(z) and any pair of nonzero integers (m,n), there exists a Padé form of type (m,n).

However, as we will show later, Padé forms are not necessarily unique. We expand on this in section 2.3.

2.2. Solutions of Hankel Systems : Normal Case

A system of equations of the form

$$H^+_{m/n} \cdot V_{m/n} = 0 , \tag{2.6}$$

where $H^+_{m/n}$ denotes the $n \times (n+1)$ matrix found in (2.3), is called a Hankel system of equations of type (m,n). Let $H_{m/n}$ denote the square $n \times n$ matrix

$$H_{m/n} = \begin{bmatrix} a_{m-n+1} & a_{m-n+2} & \dots & a_m \\ a_{m-n+2} & a_{m-n+3} & \dots & a_{m+1} \\ \dots & \dots & \dots & \dots \\ a_m & a_{m+1} & \dots & a_{m+n-1} \end{bmatrix}.$$

($H_{m/n}$ is called a Hankel matrix of type (m , n).

Equation (2.6) has a solution if both $H_{m/n}$ and

$$\begin{bmatrix} & a_{m+1}\cdot v_0 \\ H_{m/n} & \cdots \\ & a_{m+n}\cdot v_0 \end{bmatrix},$$

where $v_0 \neq 0$, have the same rank. Clearly this happens if $H_{m/n}$ is invertible.

When the Hankel matrix is nonsingular, then not only do we get a solution but we also can say that the solution is unique up to a choice of v_0. When we place our solution into equation (2.5), we get that $U_{m/n}(z)$ is also unique up to the choice of v_0. If we set $v_0 = 1$, then we get

Theorem 2.2. (cf Gragg [19]) If the Hankel matrix $H_{m/n}$ of the formal power series A(z) is nonsingular, then the Padé form $U_{m/n}(z)$ and $V_{m/n}(z)$ of type (m,n) satisfies

$$GCD(\ U_{m/n}(z)\ ,\ V_{m/n}(z)\ ) = 1\ ,$$

where GCD means the greatest common divisor of two polynomials. In particular, the Padé form of type (m,n) for A(z) is unique up to multiplication by a nonzero constant.

Proof:

That a Padé form of type (m,n) exists comes from the Frobenius theorem. Let $U'_{m/n}(z)$ and $V'_{m/n}(z)$ be the numerator and denominator, respectively, of one such Padé form. Suppose that

$$GCD(\ U'_{m/n}(z)\ ,\ V'_{m/n}(z)\ ) = P(z)\ ,$$

where the factor P(z) is nontrivial, that is, $\partial(P) \geq 1$. Factor out the greatest common factor P(z) to get two new polynomials $U^{\#}_{m/n}(z)$ and $V^{\#}_{m/n}(z)$ given by

$$U^{\#}_{m/n}(z) = U'_{m/n}(z)\cdot P(z)^{-1}$$

$$V^{\#}_{m/n}(z) = V'_{m/n}(z)\cdot P(z)^{-1}$$

Let

$$l = \min(\, m - \partial(\, U^{*}_{m/n}(z)\,)\,,\; n - \partial(\, V^{*}_{m/n}(z)\,)\,)\,.$$

Because we factored out a nontrivial divisor from the original Padé form, we know that $l \geq 1$.

Now let

$$U_{m/n}(z) = U^{*}_{m/n}(z) \cdot z^{l}\,,$$

$$V_{m/n}(z) = V^{*}_{m/n}(z) \cdot z^{l}\,.$$

Notice that, since $V'_{m/n}(z)$ is a denominator for the Padé form, then $V'_{m/n}(z) \neq 0$. Hence $V_{m/n}(z) \neq 0$. Also, because of the way l was chosen, it follows that $U_{m/n}(z)$ and $V_{m/n}(z)$ have degrees less than or equal to m and n, respectively. Furthermore,

$$\begin{aligned} A(z)\cdot V_{m/n}(z) - U_{m/n}(z) &= z^{l}\left\{A(z)\cdot V'_{m/n}(z) - U'_{m/n}(z)\right\}\cdot P(z)^{-1} \\ &= \left\{z^{m+n+1}\cdot R'_{m/n}(z)\right\}\cdot z^{l}\cdot P(z)^{-1} \\ &= z^{m+n+1} R_{m/n}(z)\,, \end{aligned}$$

since $\partial(P) \leq l$.

Thus, the pair $U_{m/n}(z)$ and $V_{m/n}(z)$ is also a Padé form of type (m , n) for A(z). In particular, $V_{m/n}$ must satisfy equation (2.6). Since the Hankel matrix $H_{m/n}$ is non-singular, the solutions to (2.6) are unique up to the choice of a constant, namely up to the choice of the constant term v_0. But $l \geq 1$, hence the constant term is zero. Since $V_{m/n}$ is not zero, we have a contradiction. Thus, there can be no nontrivial constant factors and so the greatest common divisor is 1. This completes the proof.

Any time a power series satisfies the condition that the Hankel matrix of type (m,n) is nonsingular for all (m,n), it is said to be normal power series. Most published algorithms which compute Padé forms limit their scope to solving normal power series

(see discussion in chapter 4). However, the normality condition for a unit power series is a very strong condition. For example, even the simple power series

$$A(z) = 1 + z^2 + z^4 + R(z) \quad with\ ord(R) \geq 6$$

does not satisfy the normality condition (Hankel matrix is singular for m=2 and n=3).

2.3. Characterizing Solutions of Hankel Systems : Singular case

Assume now that the Hankel matrix $H_{m/n}$ is a singular matrix. In order to determine the solutions to system (2.6), we row reduce the coefficient matrix $H^{+}_{m/n}$. We do this in the following special way. (We will assume $m \geq n$ for this discussion. This is no problem as we will see later).

Let $H_{(m-l)/(n-l)}$ be the largest non-singular principal minor submatrix of $H_{m/n}$. Since $H_{(m-l)/(n-l)}$ is non-singular, we can solve the Hankel system of type $(m-l, n-l)$ to get the (unique up to constant multiple) Padé form of order $(m-l, n-l)$. Let

$$V^*_{(m-l)/(n-l)}(z) = \sum_{i=0}^{n-l} v_i^* z^i ,$$

$$U^*_{(m-l)/(n-l)}(z) = \sum_{i=0}^{m-l} u_i^* z^i$$

be the unique Padé form for A(z) of type $(m-l, n-l)$ with

$$GCD(\ V^*_{(m-l)/(n-l)}(z)\ ,\ U^*_{(m-l)/(n-l)}(z)\ ) = 1\ .$$

In the case that the largest nonsingular principal minor of $H_{m/n}$ does not exist, then $l = n$, and we set

$$V^*_{(m-n)/0} = 1 \text{ and } U^*_{(m-n)/0} = a_0 + \cdots + a_{m-n} z^{m-n}\ .$$

Since

$$A(z) \cdot V^*_{(m-l)/(n-l)}(z) - U^*_{(m-l)/(n-l)}(z) = z^{m+n-2l+1} \cdot \sum_{i=0}^{\infty} r_i^* z^i\ ,$$

we have the following relationship

$$\left[v_{n-l}^{*}, \cdots, v_{0}^{*}\right] \cdot \begin{bmatrix} a_{m-n+1} & a_{m-l} & a_{m-l+1} & a_{m+1} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m-l+1} & a_{m+n-2l} & & a_{m+n-l+1} \end{bmatrix} = \left[0, \cdots, 0, r_0^{*}, \cdots, r_{l-1}^{*}\right].$$

This relationship can be used to create an important matrix identity. Let

$$L = \left[\begin{array}{ccc|ccc} 1 & & & & & \\ & 1 & & & & \\ & & \ddots \; 0 & & & \\ & & 1 & & & \\ \hline v_{n-l}^{*} & \cdots & & v_0^{*} & & \\ & \cdots & \cdots & & \cdots & \\ & & v_{n-l}^{*} & & & v_0^{*} \end{array}\right],$$

$$H_{m/n}^{+} = \left[\begin{array}{cc|cc} a_{m-n+1} & a_{m-l} & a_{m-l+1} \cdots & a_{m+1} \\ \cdots \quad \cdots & & & \cdots \\ \cdots \quad \cdots & & & \cdots \\ a_{m-l} & a_{m+n-2l-1} & a_{m+n-2l} \cdots & a_{m+n-l} \\ \hline a_{m-l+1} & a_{m+n-2l} & a_{m+n-2l+1} \cdots & a_{m+n-l+1} \\ \cdots \quad \cdots & \cdots & & \cdots \\ a_m \quad \cdots & a_{m+n-l-1} & a_{m+n-l} \cdots & a_{m+n} \end{array}\right],$$

and

$$H_{m/n}^{*} = \left[\begin{array}{cccc|cc} a_{m-n+1} & & & a_{m-l} & a_{m-l+1} & a_{m+1} \\ & \cdots & & & & \\ & & \cdots & & & \\ a_{m-l} & & & & a_{m+n-2l} & a_{m+n-l} \\ \hline 0 & 0 & & 0 & r_0^{*} \cdots & r_l^{*} \\ & & & r_0^{*} & r_1^{*} \cdots & \\ \cdots & & & & & \\ 0 & & r_0^{*} & r_{l-2}^{*} & r_{l-1}^{*} \cdots & r_{2l-1}^{*} \end{array}\right].$$

Then, we have that

$$L \cdot H_{m/n}^{+} = H_{m/n}^{*}. \tag{2.7}$$

We can think of the matrix L as a series of elementary matrices multiplied together representing a series of row operations on the Hankel matrix $H_{m/n}$. This in turn allows us to limit the type of solutions that equation (2.6) will have.

Since $H_{(m-l)/(n-l)}$ is the largest block non-singular principal minor, by examining $H^{*}_{m/n}$ we can deduce that

$$r^{*}_0 = r^{*}_1 = \cdots = r^{*}_{l-1} = 0$$

The above condition implies that

$$A(z) \cdot V^{*}_{(m-l)/(n-l)}(z) - U^{*}_{(m-l)/(n-l)}(z) = z^{m+n-l+k+1} \cdot \sum_{i=0}^{\infty} r_i z^i ,$$

where $k \geq 0$ and $r_i = r^{*}_{l+k+i}$.

There are two cases here. If $k \geq l$, then we already have a solution. If $k < l$, then the right hand side of equation (2.7) becomes

$$\left[\begin{array}{ccc|cccc} a_{m-n+1} & \cdots & a_{m-l} & \cdots & \cdots & & a_{m+1} \\ \cdots & & \cdots & & & & \cdots \\ \cdots & & \cdots & & & & \cdots \\ a_{m-l} & \cdots & & \cdots & \cdots & & a_{m+n-l} \\ \hline & & & 0 & 0 & 0 & 0 \\ & & & & & 0 & \cdots \\ & 0 & & & & \cdots & r_0 \\ & & & & & r_0 & \cdots \\ & & & 0 & r_0 & \cdots & r_{l-k-1} \end{array}\right] \qquad (2.8)$$

Here the top left matrix block is of size $(n-l) \times (n-l)$, the top right hand matrix block is of size $(n-l) \times (l+1)$, the bottom right matrix block is of size $l \times (l+1)$ with the first k rows of this block equal to 0 rows.

Since the $(l-k) \times (l-k)$ matrix

$$\left[\begin{array}{ccccc} 0 & 0 & \cdots & 0 & r_0 \\ 0 & & & r_0 & \\ \cdots & & & & \\ \cdots & r_0 & & & \\ r_0 & & \cdots & & r_{l-k-1} \end{array}\right]$$

is an invertible matrix (since $r_0 \neq 0$), we get that a basis for the solutions of equation (2.6) are given by the set

$$\begin{bmatrix} v^*_{n-l} \\ \cdot \\ \vdots \\ v^*_0 \\ 0 \\ \cdot \\ \vdots \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ v^*_{n-l} \\ \cdot \\ \vdots \\ v^*_0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \cdots \quad \cdots, \quad \begin{bmatrix} \vdots \\ 0 \\ v^*_{n-l} \\ \cdots \\ \vdots \\ v^*_0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{2.8}$$

Here the number of zero elements varies from 0 and l for the first vector, 1 and l-1 for the second, and finally k and l - k for the last vector.

Thus, the general solution to (2.6) is a linear combination of these vectors, namely,

$$V_{m/n} = \begin{bmatrix} v^*_{n-l} \\ \cdots \\ \cdots \\ v^*_0 \\ 0 \\ \cdots \\ 0 \end{bmatrix} \alpha_k + \cdots + \begin{bmatrix} 0 \\ \cdots \\ v^*_{n-l} \\ \cdots \\ v^*_0 \\ 0 \\ \cdots \end{bmatrix} \alpha_0 , \tag{2.9}$$

where α_0 , $\cdots$, α_k are constants from the field K. In polynomial form, the general solution is then given by

$$V_{m/n}(z) = z^{l-k} \left\{ \sum_{i=0}^{n-l} v^*_i z^i \right\} \cdot \left\{ \sum_{i=0}^{k} \alpha_i z^i \right\} . \tag{2.10}$$

From equation (2.5), it then follows that

$$U_{m/n}(z) = z^{l-k} \left\{ \sum_{i=0}^{m-l} u^*_i z^i \right\} \cdot \left\{ \sum_{i=0}^{k} \alpha_i z^i \right\} . \tag{2.11}$$

We summarize our results as follows.

Theorem 2.3: Let A(z) be a unit formal power series. Then there is an entire family of Padé forms of type (m,n). This family of solutions differs only by a greatest common divisor in the numerator and denominator. More specifically, if $U_{m/n}(z)$, $V_{m/n}(z)$ and $U'_{m/n}(z)$, $V'_{m/n}(z)$ are two pairs of Padé forms for A(z) of type (m, n) and if

$$GCD(U_{m/n}(z)\,,\,V_{m/n}(z)\,) = P(z),$$

and

$$GCD(U'_{m/n}(z)\,,\,V'_{m/n}(z)\,) = P'(z),$$

then

$$\frac{U_{m/n}(z)}{P(z)} = \frac{V_{m/n}(z)}{P(z)} \text{ and } \frac{U'_{m/n}(z)}{P'(z)} = \frac{V'_{m/n}(z)}{P'(z)}$$

2.4. One Dimensional Scaled Padé Fractions

The previous section provides us with a family of Padé forms. Furthermore, if we factor out common divisors then we do get uniqueness. This brings us to the following

Definition 2.2. (cf Gragg [19]). Let A(z) be a unit power series. A pair of polynomials $P_{m/n}(z)$ and $Q_{m/n}(z)$ is said to be a **Padé fraction** of A(z) of type (m,n) if

1) the pair is a Padé form for A(z),

2) GCD($P_{m/n}(z)$, $Q_{m/n}(z)$) = 1, and

3) $Q_{m/n}(0) = 1$.

It is clear that Padé fractions are unique. However, existence is another matter. For we have characterized condition 1) by equation (2.10) and (2.11). From these equations we also get that the GCD is given by

$$z^{l-k}\cdot\sum_{i=0}^{k}\alpha_i z^i .$$

In particular, if condition 2) is to hold then this would imply that $k=0$ and $l=0$. Thus, unless we have a normal power series, we do not get existence of a Padé fraction.

Some authors (eg. Baker [2]) have tried to get around this by weaking the order requirement. This approach simply asks for the rational polynomial expression that meets the degree requirements for the numerator and denominator and agrees with the power series A(z) to as many terms as possible.

The approach of scaled Padé fractions is to try to get both uniqueness and existence without lowering any of the original degree or order requirements. The hint for this is theorem 2.3 . This says that Padé forms are unique except for common divisors. The approach of Padé fractions is to eliminate the common divisors and get uniqueness. The approach of scaled Padé fractions is to require the common divisor to be of a specified type, namely z^u for some integer u.

We can get a specific solution for the Padé form problem by returning to equation (2.9) and setting

$$\alpha_0 = \alpha_1 = \cdots = \alpha_{k-1} = 0 \quad \text{and} \quad \alpha_k = 1 .$$

Then we get particular solutions of the form

$$T_{m/n}(z) = z^l \sum_{i=0}^{n-l} v_i^* z^i$$

$$S_{m/n}(z) = z^l \sum_{i=0}^{m-l} u_i^* z^i .$$

This leads us to the following

Definition 2.3. (cf Cabay and Choi [14]). A pair of polynomials $T_{m/n}(z)$ and $S_{m/n}(z)$ with $T_{m/n}(z) \neq 0$ is said to be a scaled Padé fraction for the unit power series A(z) if

a) $\min(m - \partial(T_{m/n}), n - \partial(S_{m/n})) = 0$,

b) $A(z)\cdot T_{m/n}(z) - S_{m/n} = R_{m/n}(z)$ with $ord(R_{m/n}) \geq n+m+1$, and

c) the greatest common factor of $T_{m/n}$ (z) and $S_{m/n}$ (z) is z^l , where l is a non-negative integer.

Notice that condition a) implies that the degrees of $T_{m/n}(z)$ and $S_{m/n}(z)$ are at most n and m, respectively.

This definition thus provides a middle ground between Padé form and Padé fraction. Namely, a scaled Padé fraction is a particular Padé form of type (m,n). In addition, by shifting we can remove the greatest common divisor and determine the unique Padé fraction for A(z) of type (m,n).

We summarize our findings with

Theorem 2.4. Let A(z) be any unit power series and let n and m be two non-negative integers. Then there exists a scaled Padé fraction of type (m, n) for A(z). In addition, this scaled Padé fraction is unique up to multiplication by a non-zero constaht.

Example. Let

$$A(z) = 1 + z^2 + z^4 + R(z) ,$$

where $ord(R) \geq 6$. Suppose we wish to calculate a (2,3) scaled Padé fraction for A(z). If we follow our previous discussions we can proceed as follows.

The (2,3) Hankel matrix is given by

$$H_{2/3} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} .$$

Notice that the matrix is singular. In addition, the largest nonsingular principal minor is given by

$$H_{1/2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} .$$

If we solve for the (1,2) Padé form for A(z) we get

$$V_{1/2}(z) = 1 - z^2 \text{ , and } U_{1/2}(z) = 1 .$$

We are trying to solve

$$H_{2/3}^{+} \cdot \begin{bmatrix} t_3 \\ t_2 \\ t_1 \\ t_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} ,$$

where $H_{2/3}^{+}$ is found to be

$$H_{2/3}^{+} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} .$$

Again, following the [illegible]d of [illegible] previous section, [illegible] form the matrix

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

Multiplying the matrix L to the left of $H_{2/3}^{+}$ gives the matrix

$$H_{2/3}^{*} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} .$$

Solving

$$H_{2/3}^{+} \cdot V_{2/3} = 0$$

is the same as solving

$$H_{2/3}^{*} \cdot V_{2/3} = 0 .$$

But we can read off the solutions to this equations since they are the same as the solution to the (1,2) case, or the shifted versions of this solution. Namely, we have that a basis for the set of solutions is given by

$$\begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix} .$$

Thus, we find our general Padé form is given by

$$V_{2/3}(z) = (1 - z^2)\cdot(\alpha_0 + \alpha_1 z) \ , \text{ and } U_{2/3}(z) = \alpha_0 + \alpha_1 z \ ,$$

where α_0 and α_1 are arbitrary.

From this we get that the Padé fraction for A(z) does not exist since the reduced form

$$\frac{1}{1 - z^2}$$

does not satisfy the order condition. On the other hand the scaled Padé fraction does exist and is given by

$$T_{2/3}(z) = z - z^3 \text{ and } S_{2/3}(z) = z \ .$$

Chapter 3

Matrix Padé Fractions

3.1. Basic Definitions

Let A(z) be a formal power series with coefficients from the ring of p x p matrices over some field K. Then A(z) is a unit matrix power series if A(0), the leading coefficient in A(z), is an invertible p x p matrix. We limit our scope to unit matrix power series. For these unit power series we have the usual notions of degree and order that parallel those of one dimensional power series. As before, the degree and order of a matrix polynomial A(z) is denoted by $\partial(A)$ and ord(A), respectively.

Because of the lack of commutativity for matrices, any definition that specifies a relationship involving matrix multiplication must identify the side on which the multiplication occurs.

Definition 3.1: Let A(z) be a unit formal matrix power series and let m and n be non-negative integers. Then a pair, $U_{m/n}(z)$, $V_{m/n}(z)$, of p x p matrix polynomials is defined to be a Right Matrix Padé Form (RMPF) of type (m,n) to A(z) if

a) $\partial(U_{m/n}) \le m$, $\partial(V_{m/n}) \le n$,

b) $A(z)\cdot V_{m/n}(z) - U_{m/n}(z) = z^{m+n+1}\cdot R_{m/n}(z)$ *with* $ord(R_{m/n}) \ge 0$ and , (3.1)

c) The columns of $V_{m/n}(z)$ are linearly independent over the field K.

The matrix polynomials $V_{m/n}(z)$, $U_{m/n}(z)$, and $R_{m/n}(z)$ are usually called the right denominator, numerator, and residual (all of type (m,n)), respectively.

There is also an equivalent definition for a Left Matrix Padé Form (LMPF). Condition b) is replaced with an equivalent version with matrix multiplication by $V_{m/n}(z)$ being on the left. Condition c) is then replaced with the condition that the rows, instead of the columns, of the denominator must be linearly independent over the base field K.

As in the one dimensional situation, there is a one to one correspondence between matrix polynomials of degree n and (n+1) x 1 vectors of p x p matrix blocks. We again use the notation of the last chapter and denote

$$P_n(z) = p_0 + p_1 z + \dots + p_n z^n \quad by \quad P_n = \begin{bmatrix} p_n \\ \dots \\ p_1 \\ p_0 \end{bmatrix} \quad or \quad P_n^- = \begin{bmatrix} p_0 \\ p_1 \\ \dots \\ p_n \end{bmatrix};$$

3.2. Existence of Matrix Padé Forms

In the equation

$$A(z)\cdot V_{m/n}(z) - U_{m/n}(z) = z^{m+n+1}\cdot R_{m/n}(z),$$

let the matrix polynomials have the following expansions:

$$A(z) = \sum_{i=0}^{\infty} a_i z^i \;,\quad U_{m/n}(z) = \sum_{i=0}^{m} u_i z^i \;,\quad V_{m/n}(z) = \sum_{i=0}^{n} v_i z^i \text{ and } R_{m/n}(z) = \sum_{i=0}^{\infty} r_i z^i \,,$$

where a_i, u_i, v_i, and r_i are all p x p matrices with entries from the field K. Then equation (3.1) can be written as the following block system of equations:

$$\begin{bmatrix} H_{m/n}^- \\ --- \\ H_{m/n}^+ \end{bmatrix} \cdot V_{m/n} = \begin{bmatrix} U_{m/n}^- \\ --- \\ 0 \end{bmatrix}, \tag{3.2}$$

which is shorthand for

$$\begin{bmatrix} a_{-n} & \dots & \dots & a_0 \\ \dots & \dots & \dots & a_1 \\ \dots & \dots & \dots & \dots \\ a_{m-n} & a_{m-n+1} & \dots & a_m \\ \hline a_{m-n+1} & a_{m-n+2} & \dots & a_{m+1} \\ a_{m-n+2} & a_{m-n+3} & \dots & a_{m+2} \\ \dots & \dots & \dots & \dots \\ a_m & a_{m+1} & \dots & a_{m+n} \end{bmatrix} \cdot \begin{bmatrix} v_n \\ v_{n-1} \\ \dots \\ v_0 \end{bmatrix} = \begin{bmatrix} u_0 \\ u_1 \\ \dots \\ u_m \\ \hline 0 \\ \dots \\ \dots \\ 0 \end{bmatrix} \tag{3.3}$$

Here, $a_i = 0$ for $i < 0$.

It is clear that if we can solve the matrix block system of equations

$$H^{+}_{m/n} \cdot V_{m/n} = 0 \tag{3.4}$$

for $V_{m/n}$, then we can get $U_{m/n}$ via

$$H^{-}_{m/n} \cdot V_{m/n} = U^{-}_{m/n} \tag{3.5}$$

Thus, as in the one dimensional case, we concentrate on solving equation (3.4).

For any coefficient matrix v_j, let the ith column be denoted by $v_j^{(i)}$. Replace each occurrence of v_j by $v_j^{(i)}$ in equation (3.3) and set

$$X^{(i)} = \begin{bmatrix} v_n^{(i)} \\ v_{n-1}^{(i)} \\ \cdots \\ v_0^{(i)} \end{bmatrix}.$$

Then the problem becomes one of finding p linearly independent solutions $X^{(1)}, \cdots, X^{(p)}$ of the system of equations

$$H^{+}_{m/n} \cdot X = 0, \tag{3.6}$$

where X represents a $p(n+1) \times 1$ vector. A simple count reveals that this is a homogeneous system of pn equations in $p(n+1)$ unknowns. An elementary result from linear algebra (see for example [1]) ensures that this system has at least p solutions that are linearly independent over the field K . Thus, we have

Theorem 3.2. (Existence of Matrix Padé Forms). For any formal matrix power series A(z) and any pair of nonzero integers (m,n), there exists an (m,n) Matrix Padé Form (either right or left).

At this point it is worth pointing out that ours is not the only way to define a matrix version of Padé form and indeed others (cf Bose [6] , Starkand [30] , and Bultheel [10]) offer alternate definitions. The problem lies in the matrix version of condition c) of definition 2.1. We could have simply asked that the denominator be nonzero in the matrix definition also. But then, once we had one solution to our system (3.6), we could repeat the solution p times as columns of a p x p matrix and get two matrices that satisfy conditions a) and b) of our definition. This is not what we are looking for. For example, we could never get a unique solution no matter how well behaved the matrix power series is.

On the other hand, a nonzero denominator in the scalar case has additional meaning. It also means, for example, that the denominator can be divided into the numerator to create a rational function. The natural matrix version of this would be to require that the denominator be invertible, i.e.,

$$det(V_{m/n}(z)) \neq 0 .$$

However, this definition creates problems with existence, as we shall see in a later section.

3.3. Solutions of Block Hankel Systems : Nonsingular Case

In the block system of equations

$$H^{+}_{m/n} \cdot V_{m/n} = 0$$

(called a block Hankel system), where $H^{+}_{m/n}$ denotes the $pn \times p(n+1)$ block matrix from the previous section, let $H_{m/n}$ stand for the square $pn \times pn$ block Hankel matrix of type (m,n) given by

$$\begin{bmatrix} a_{m-n+1} & a_{m-n+2} & \dots & a_m \\ a_{m-n+2} & a_{m-n+3} & \dots & a_{m+1} \\ \dots & \dots & \dots & \dots \\ a_m & a_{m+1} & \dots & a_{m+n-1} \end{bmatrix} .$$

Equation (3.4) has a solution if both

$$H_{m/n} \quad \text{and} \quad \begin{bmatrix} & a_{m+1}\cdot v_0^{(i)} \\ H_{m/n} & \cdots \\ & a_{m+n}\cdot v_0^{(i)} \end{bmatrix} \tag{3.7}$$

have the same rank for every column $v_0^{(i)}$ (this is not a necessary condition, just a sufficient one). Clearly this will happen if the block Hankel matrix $H_{m/n}$ is invertible.

Definition 3.3: A matrix power series A(z) is said to be normal if, for every pair of non-negative integers m and n, the (m,n) block Hankel matrix $H_{m/n}$ of A(z) is invertible. A matrix power series is said to be (m,n) - normal if the (i,j) block Hankel matrices $H_{i/j}$ are invertible for every $i \le m$ and $j \le n$.

We note that, since the block Hankel matrices resulted from the introduction of a systems of equations to solve for right matrix Padé forms, we really should be specifying that a matrix power series is right normal or right (m, n) - normal. However, when we are setting up the original system of equations for the right side we come across the system of equations

$$\begin{bmatrix} a_{m-n+1} & & a_m \\ \cdots & \cdots & \cdots \\ a_m & & a_{m+n-1} \end{bmatrix} \cdot \begin{bmatrix} v_n \\ \cdots \\ v_1 \end{bmatrix} = \begin{bmatrix} -a_{m+1}\cdot v_0 \\ \cdots \\ -a_{m+n}\cdot v_0 \end{bmatrix}$$

so invertibility of the Hankel matrix allows us to solve the above system of equations in terms of v_0.

If we now set up the system of equations that result when the multiplication is done on the right, we get the system of equations

$$\begin{bmatrix} v_n, \cdots, v_1 \end{bmatrix} \cdot \begin{bmatrix} a_{m-n+1} & & a_m \\ & \cdots & \\ a_m & & a_{m+n-1} \end{bmatrix} = \begin{bmatrix} -v_0\cdot a_{m+1}, & \cdots, & -v_0\cdot a_{m+n} \end{bmatrix}$$

and the coefficient matrix of this system is just the block transpose of the original Hankel matrix. But, if the Hankel matrix is invertible then so is its block transpose. The block transpose will then have as its inverse the block transpose of the inverse of the Hankel matrix. Thus, if a power series is right normal then it is also left normal and so we need not specify a side when talking about normality. As a matter of record we also mention the fact that if a power series A(z) is normal then so is its transpose power series $A^t(z)$.

If the matrix power series A(z) is normal (actually for the following discussion we need only ask that A(z) be (m, n) - normal), then we can get right (or left) Padé forms by row reducing equation (3.4) to the system

$$\begin{bmatrix} I_{np} & \begin{matrix} a^{\#}_{m+1} \\ \cdots \\ \cdots \\ a^{\#}_{m+n} \end{matrix} \end{bmatrix} \cdot V_{m/n} = 0 \tag{3.8}$$

where I_{np} is the $np \times np$ identity matrix. This implies that all the coefficient matrices $v_n, v_{n-1}, \dots, v_1$, can be written as a multiple of v_0. Since we have the linear independence requirement, condition c), in the definition of Padé form, we limit our choices for v_0 to being an invertible matrix. In particular, if we set v_0 equal to the identity matrix, then we can uniquely determine the denominator up to multiplication by an invertible matrix. Thus, we find a right matrix Padé form that has the property that the constant term of the denominator is the identity matrix. Since a_0 is also assumed to be nonsingular, then u_0 is also a nonsingular matrix.

Finally, we mention that for a normal power series the numerator, denominator and the leading term in the residual power series are all nonsingular polynomial matrices for the (m,n) in discussion.

We summarize our discussion by

Theorem 3.4. Let A(z) be an (m,n) - normal matrix power series with m and n any nonnegative integers. Then there exists a right matrix Padé form $U_{m/n}(z)$ and $V_{m/n}(z)$ of type (m,n) with the properties that

a) $V_{m/n}(0) = I_p$, the p x p identity matrix,

b) $\partial(V_{m/n}) \leq n$, and $\partial(U_{m/n}) \leq m$ and

c) $A(z)\cdot V_{m/n}(z) - U_{m/n}(z) = z^{m+n+1}\cdot R_{m/n}(z)$ and

d) $RGCD(U_{m/n}(z), V_{m/n}(z)) = I_p$,

where $RGCD$ of two matrix polynomials means the greatest right matrix polynomial divisor.

Proof: Parts a), b), and c) follow from the previous discussion. Part d) is proved using arguments that parallel those of theorem 2.2 and will not be included here.

3.4. Solutions of Block Hankel System : Singular Case

The results of the previous section follow those of the one dimensional case. For normal power series, the solutions are unique up to multiplication by invertible constant matrices, the denominator is invertible and there are similar results for both the right side and the left side.

However, as mentioned in the one dimensional situation, the normality condition for a unit power series is a very strong condition. For example, even the simple matrix power series

$$A(z) = \begin{bmatrix} 1+z^2+z^4 & 0 \\ 0 & 1+z^2+z^4 \end{bmatrix} + R(z) \quad with \;\; ord(R) \geq 6$$

does not satisfy the normality condition (block Hankel system is singular for m = 2 and n = 3) Unfortunately, though, there is a big difference in the singular case as opposed to the nonsingular case when it comes to matrix Padé forms.

3.4.1. Singular Matrix Denominators

Equation (3.1) is not necessarily equivalent to solving an equation of the form

$$A(z) = U_{m/n}(z)\cdot V_{m/n}(z)^{-1} + R^{*}_{n+m+1}(z), \tag{3.9}$$

where the order of the residual is at least m + n + 1, even though this is how we visualize the solutions. For, even though we can find $U_{m/n}(z)$ and $V_{m/n}(z)$ of prescribed degrees with the columns of $V_{m/n}(z)$ being linearly independent over the base field, this does not force $V_{m/n}(z)$ to be an invertible matrix. Linear independence over the base field K does not imply linear independence over the ring of polynomials $K[z]$. We say that vectors are algebraically independent if they are linearly independent over the ring $K[z]$.

Example 3.5: Let A(z) be given by

$$A(z) = \begin{bmatrix} 1+z^2+2z^4-z^5 & 0 \\ -z^5 & 1+z^2+z^4 \end{bmatrix} + R(z)$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} z^2 + \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} z^4 + \begin{bmatrix} -1 & 0 \\ -1 & 0 \end{bmatrix} z^5 + R(z),$$

where R(z) has order ≥ 6. Then, if we are looking for the (2,3) right matrix Padé form of A(z), we can solve the block Hankel system of linear equations

$$H^{+}_{2/3}\cdot X = 0$$

to get a basis for the solution space consisting of the vectors

$$X_1 = \begin{bmatrix} 0 \\ -z^2+1 \end{bmatrix}, \quad X_2 = \begin{bmatrix} 0 \\ -z^3+z \end{bmatrix};$$

i.e., the general solution is of the form

$$X = \begin{bmatrix} 0 \\ -az^3-bz^2+az+b \end{bmatrix} = (az+b)\cdot\begin{bmatrix} 0 \\ -z^2+1 \end{bmatrix}$$

with a and b being free choices from K.

A natural choice for $V_{2/3}(z)$ and $U_{2/3}(z)$ is then

$$V_{2/3}(z) = \begin{bmatrix} 0 & 0 \\ -z^2+1 & -z^3+z \end{bmatrix} \quad \text{and} \quad U_{2/3}(z) = \begin{bmatrix} 0 & 0 \\ 1+2z^2 & z+2z^3 \end{bmatrix}.$$

Notice that $V_{2/3}(z)$ has determinant $= 0$ for all z. Also notice that this solution is unique up to a choice of a basis for the system of equations.

In particular, this is a situation where we cannot get equation (3.9) from (3.1). The problem occurs because, although the solution space has dimension 2 when considered as a vector space over the field K, it has only dimension 1 when considered as a module over the ring $K[z]$. For later reference, we note that the rank of the matrix $H^{+}_{2/3}$ is 5. In particular, the rank is not a multiple of the matrix size 2, which prevents the existence of a nonsingular $V_{2/3}(z)$, as will be shown in section 3.6.

Example 3.5 is not surprising for triangular matrices. Indeed, if we desire the (m,n) Padé form for any power series of the form

$$A(z) = \begin{bmatrix} a(z) & b(z) \\ 0 & c(z) \end{bmatrix},$$

where a(z), b(z), and c(z) are all scalar power series, then, when we are solving

$$A(z)\cdot V_{m/n}(z) - U_{m/n}(z) = z^{m+n+1}\cdot R_{m/n}(z)$$

with

$$V_{m/n}(z) = \begin{bmatrix} v_{11}(z) & v_{12}(z) \\ v_{21}(z) & v_{22}(z) \end{bmatrix}, \quad \text{and} \quad U_{m/n}(z) = \begin{bmatrix} u_{11}(z) & u_{12}(z) \\ u_{21}(z) & u_{22}(z) \end{bmatrix},$$

we run into the following situation. If we consider only the last row, we get the equations

$$c(z)\cdot v_{21}(z) - u_{21}(z) = z^{m+n+1}\cdot r_{21}(z)$$

$$c(z)\cdot v_{22}(z) - u_{22}(z) = z^{m+n+1}\cdot r_{22}(z) .$$

To determine the last rows of $V_{m/n}(z)$ and $U_{m/n}(z)$, we need to solve the one dimensional Padé problem for the power series c(z) twice. But our work in chapter 2 tells us that a basis for the solutions to this problem consists of scaled Padé fractions of type (m,n) for c(z) and z^{-l} times these particular scaled Padé fractions, where l is some non-negative integer.

In particular, to determine the rest of the polynomial entries of $V_{m/n}(z)$ and $U_{m/n}(z)$, we need to solve

$$a(z)\cdot v_{11}(z) - u_{11}(z) = -b(z)\cdot v_{21}(z) + z^{m+n+1}\cdot r_{11}(z)$$

$$a(z)\cdot v_{12}(z) - u_{12}(z) = -b(z)\cdot v_{22}(z) + z^{m+n+1}\cdot r_{12}(z) .$$

But then our order condition does not need to hold as it did before. If we examine the system of equations that occurs here, we find that we no longer have a homogeneous system of equations but a specific system. Hence, we need not necessarily find the correct number of linearly independent solutions, and as a matter of fact, we may not have any solution at all. This fact coupled with the result that the entries of the last rows of both $V_{m/n}(z)$ and $U_{m/n}(z)$ differ by a power of z makes algebraically dependent solutions a natural occurrence in this situation.

Algebraically dependent solutions are something that we wish to avoid, since otherwise we can run into the following situation. Suppose that we have two scalar power series a(z) and b(z). Suppose further that for a certain integer type (m,n) we have a Padé form u(z) and v(z) that satisfies

$$a(z)\cdot v(z) - u(z) = z^{m+n+1+k}\cdot r(z) ,$$

where $k \geq 1$ and, in addition, the denominator and numerator are both of a smaller degree than necessary. Then the polynomials $z\cdot u(z)$ and $z\cdot v(z)$ are also Padé forms of type (m,n) for a(z).

If we were now to form the matrix power series

$$A(z) = \begin{bmatrix} a(z) & 0 \\ 0 & b(z) \end{bmatrix}$$

then we would want a right matrix Padé form to give information about both a(z) and b(z). However, as it stands now, we can solve the right matrix Padé problem of type (m,n) for A(z) by simply forming the matrices

$$U_{m/n}(z) = \begin{bmatrix} u(z) & z\cdot u(z) \\ 0 & 0 \end{bmatrix}, \quad V_{m/n}(z) = \begin{bmatrix} v(z) & z\cdot v(z) \\ 0 & 0 \end{bmatrix}$$

That is, we can find a right Padé form by simply ignoring the scalar power series b(z) altogether.

3.4.2. Non-Uniqueness of Matrix Padé Forms

In the singular case, the lack of invertibility for the denominator is not the only problem that stands apart from its scalar counterpart. One problem that comes up in the matrix situation is that there is not necessarily a unique solution of the system (even up to multiplication by an nonsingular constant matrix). This is illustrated by

Example 3.6. Suppose

$$A(z) = \begin{bmatrix} 1+z^2+2z^4 & 0 \\ 0 & 1+z^2+z^4 \end{bmatrix} + R(z)$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} z^2 + \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} z^4 + R(z),$$

where R(z) has order ≥ 6. Solving the block Hankel system of equations

$$H^{+}_{2/3}\cdot X = 0$$

for the (2,3) right matrix Padé form gives the three linearly independent vectors

$$X_1 = \begin{bmatrix} 1-2z^2 \\ 0 \end{bmatrix}, \quad X_2 = \begin{bmatrix} 0 \\ 1-z^2 \end{bmatrix}, \quad \text{and} \quad X_3 = \begin{bmatrix} 0 \\ z-z^3 \end{bmatrix}.$$

There are then three natural choices for $V_{2/3}(z)$ and $U_{2/3}(z)$, namely,

$$V_{2/3}^{(1)}(z) = \begin{bmatrix} 1-2z^2 & 0 \\ 0 & 1-z^2 \end{bmatrix} \quad \text{and} \quad U_{2/3}^{(1)}(z) = \begin{bmatrix} 1-z^2 & 0 \\ 0 & 1 \end{bmatrix}$$

$$V_{2/3}^{(2)}(z) = \begin{bmatrix} 1-2z^2 & 0 \\ 0 & z-z^3 \end{bmatrix} \quad \text{and} \quad U_{2/3}^{(2)}(z) = \begin{bmatrix} 1-z^2 & 0 \\ 0 & z \end{bmatrix}$$

$$V_{2/3}^{(3)}(z) = \begin{bmatrix} 0 & 0 \\ 1-z^2 & z-z^3 \end{bmatrix} \quad \text{and} \quad U_{2/3}^{(3)}(z) = \begin{bmatrix} 0 & 0 \\ 1 & z \end{bmatrix} .$$

Not only do we have too many solutions for our problem, we also find that two of the solutions are preferable in that they result in a nonsingular denominator $V_{2/3}(z)$, while the third solution gives the denominator as a singular matrix. We try to avoid having a singular denominator as this would prevent us from forming the matrix power series $U_{m/n}(z)\cdot V_{m/n}(z)^{-1}$. For example, a singular denominator gives no information about the poles since every point is a pole in this case.

We note that in a situation of multiple solutions, we do not always have a non-singular denominator as one of our choices. For example, if we combine examples 3.5 and 3.6 to form a matrix power series

$$A(z) = \begin{bmatrix} 1+z^2+2z^4 & 0 & 0 & 0 \\ 0 & 1+z^2+z^4 & 0 & 0 \\ 0 & 0 & 1+z^2+2z^4-z^5 & 0 \\ 0 & 0 & -z^5 & 1+z^2+z^4 \end{bmatrix},$$

then the (2,3) block Hankel system will result in more than 4 linearly independent solutions to equation (3.6) but all these solutions, when combined into matrices result in singular denominators.

3.4.3. Left vs Right Matrix Padé Forms

An example that illustrates some more problems with matrix Padé forms in the singular block Hankel case concerns the relationship between right and left matrix Padé forms . There is a one to one correspondence between right and left matrix Padé forms given by the transpose operation. If we take transposes on both sides of equation (3.1)

$$A(z)\cdot V_{m/n}(z) - U_{m/n}(z) = z^{m+n+1}R_{m/n}(z),$$

where $\mathrm{ord}(R_{m/n}) \geq 0$, we get an equation of the form

$$(V_{m/n}(z))^t\cdot A(z)^t - (U_{m/n}(z))^t = z^{m+n+1}(R_{m/n}(z))^t.$$

The degree and order conditions are identical. Also, the rows of $(V_{m/n}(z))^t$ are the same as the columns of $V_{m/n}(z)$, and hence are linearly independent when considered as vectors over the field K . $(A(z))^t$ is the same power series as A(z), except with each coefficient matrix in $(A(z))^t$ being the transpose of the corresponding coefficient matrix of A(z). Thus, there is a one to one correspondence between right matrix Padé forms of a matrix A(z) and left matrix Padé forms of its transpose $(A(z))^t$. In particular, if one wishes to calculate a left matrix Padé form for a power series A(z) and one has an algorithm to calculate a right matrix Padé form for a given power series, then one need only work out the right matrix Padé form for the power series $(A(z))^t$ for the denominator V(z) and numerator U(z) and then the left Padé form for A(z) would be $(V(z))^t$ and $(U(z))^t$.

However, if the Hankel system is singular, then the right and left solutions for a specified degree pair (m,n) may have different properties when the block Hankel matrix is singular as illustrated by

Example 3.7. Let A(z) be the same as the power series from example 3.5. Then the right matrix Padé form is given by

$$V_{2/3}(z) = \begin{bmatrix} 0 & 0 \\ -z^2+1 & -z^3+z \end{bmatrix} \quad \text{and} \quad U_{2/3}(z) = \begin{bmatrix} 0 & 0 \\ 1+2z^2 & z+2z^3 \end{bmatrix}.$$

In particular, the denominator is singular for all values of z. To get the left Padé form for A(z), we find the right Padé form for the transpose power series

$$A^t(z) = \begin{bmatrix} 1+z^2+2z^4-z^5 & -z^5 \\ 0 & 1+z^2+z^4 \end{bmatrix} + R(z)$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} z^2 + \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} z^4 + \begin{bmatrix} -1 & -1 \\ 0 & 0 \end{bmatrix} z^5 + R(z),$$

where R(z) is of order ≥ 6.

A solution to the block Hankel system for the (2,3) Padé form that results from this power series gives the basis

$$X_1 = \begin{bmatrix} z-z^3 \\ 1-z^2 \end{bmatrix}, \; X_2 = \begin{bmatrix} 1+z-2z^2-z^3 \\ 0 \end{bmatrix}, \; \text{and} \; X_3 = \begin{bmatrix} 0 \\ z-z^3 \end{bmatrix}.$$

One solution is then given by

$$V_{2/3}^{(R)}(z) = \begin{bmatrix} z-z^3 & 0 \\ 1-z^2 & z-z^3 \end{bmatrix} = 1-z^2 \cdot \begin{bmatrix} z & 0 \\ 1 & z \end{bmatrix} \quad \text{and} \quad U_{2/3}^{(R)}(z) = \begin{bmatrix} z & 0 \\ 1+z^2 & z \end{bmatrix}.$$

The subsequent left matrix Padé form for the denominator and numerator is then given by

$$V_{2/3}^{(L)}(z) = \begin{bmatrix} z-z^3 & 1-z^2 \\ 0 & z-z^3 \end{bmatrix} \quad \text{and} \quad U_{2/3}^{(L)}(z) = \begin{bmatrix} z & 1+z^2 \\ 0 & z \end{bmatrix}.$$

Note that the denominator is nonsingular. In particular, a solution to the left matrix Padé form problem will result in an approximation to A(z) that that satisfies equation (3.9), whereas the right Padé form will not lead to such a solution.

As a final note, a singular right denominator does not force the left denominator to be nonsingular and conversely. In addition, an example of a power series that has no right or left nonsingular denominator of a specified degree as part of a Padé form can easily be constructed.

3.5. Restricting Matrix Power Series

The previous sections gave us examples of unusual behaviour for the problem of finding matrix Padé forms of various types. The main thing that comes out of all this is not that the solutions are of varied type but rather that our definition is too broad and the type of matrix power series under discussion too general. For example, any matrix Padé form that results in the denominator being a singular matrix does not mesh with our idea that a matrix Padé form should serve as an approximate version of the original power series .

We wish therefore to restrict the definition of Padé forms to matrix power series that at least result in nonsingular denominators. Unfortunately, we cannot do anything about the fact that once we have asked for this type of restriction then we are greatly limiting the class of power series for which Padé forms exist. To get some idea of the limits imposed on us, suppose we have two unit matrix polynomials A(z) and B(z) of degrees n and m, respectively, where $n \geq m$. Then we can create a formal matrix power series by taking a quotient $B(z)^{-1} \cdot A(z)$. Then, if we want the right matrix Padé form of type $(0, n-m)$, we find ourselves with an equation of the form

$$\left\{B(z)^{-1} \cdot A(z)\right\} \cdot V_{n-m/0} - U_{n-m/0}(z) = z^{n-m+1} \cdot R(z) \ .$$

If we multiply by B(z) on the left of both sides of the equation, we get

$$A(z) \cdot V_{n-m/0} = B(z) \cdot U_{n-m/0}(z) + z^{n-m+1} \cdot R^{*}(z) ,$$

where we set $R^{*}(z) = B(z) \cdot R(z) \quad mod\ z^{n+1}$. If we replace every occurence of z by

$1/z$ and then multiply each side by z^n we will get the equation

$$A^*(z)\cdot V^*_{n-m/0} = B^*(z)\cdot U^*_{n-m/0}(z) + R^*(z)\,,$$

where the degree of $R^*(z)$ is less than the degree of $B^*(z)$. Here the superscript * means to reverse the order of the coefficients of the original polynomial (For a further discussion on the * operator see chapter 5.) In particular, if we always ask that we can get an invertible denominator, then we would need to have $V_{n-m/0}$ be an invertible matrix. However this would mean that we could divide $B^*(z)$ into $A^*(z)$. Since the set of matrix polynomials does not form a Euclidean domain when the norm is the degree of a matrix polynomial (unlike the scalar polynomials case - cf [1]), we can see that we have excluded certain matrix power series from consideration.

3.6. Nearly-Normal Matrix Power Series

In order to obtain solutions to the Padé form problem that approximates rational matrix approximation we ask that the denominator be invertible. As pointed out in section 3.3, we can achieve invertibility of the denominator if we assume that we are dealing with a normal matrix power series. We expand this class by

Definition 3.8: Let m and n be a pair of nonnegative integers. A unit matrix power series A(z) is said to be (m,n) nearly-normal if, for any integer pair (m,n), the sequence of block Hankel matrices

$$H_{(m-n)/0}\,,\; H_{(m-n+1)/1}\,,\;\cdots\;,\; H_{m/n}$$

all have rank a multiple of the matrix size p. A matrix power series is said to be nearly-normal if the power series is (m,n) nearly-normal of any pair of integers m and n.

As was the situation in the normal case there is no need to specify a right or a left side because of the correspondence between right and left via transpose and because rank of a matrix is equal to the rank of its transpose.

It is clear that every formal scalar power series (p = 1) is nearly-normal. Furthermore, any normal formal power series is also nearly-normal. Actually we can be a bit more specific, namely that every (m,n) normal matrix power series is also (m,n) nearly-normal.

The set of all nearly-normal powers series for a given integer pair is more general than the set of all normal power series. For example the power series given at the start of section 3.4 is (2,3) - nearly-normal but it is not (2,3) - normal. Examples of power series that are not nearly-normal can be found by recalling examples 3.5, 3.6 and 3.7, where the integer pair for each of the examples is always (2,3).

The nearly normal condition is helpful when used in conjunction with the following observation. For any given integer pair (m,n)

$$H_{i/j} = \left[\begin{array}{c|c} H_{(i-1)/(j-1)} & \cdots \\ & \cdots \\ \hline \cdots & a_{i+j} \end{array}\right] \tag{3.11}$$

If $H_{(i-1)/(j-1)}$ is of full rank, then if we row reduce the matrix $H_{i/j}$, we get

$$H'_{i/j} = \left[\begin{array}{c|c} I & \cdots \\ & \cdots \\ \hline 0 & r_0 \end{array}\right] \tag{3.12}$$

In particular, if A(z) is (m,n)-nearly-normal and i - j = m - n with i < m then either r_0 is the 0-matrix, or it is of full rank, i.e., r_0 is either the p x p 0-matrix, or it is invertible.

3.7. Padé Forms of Nearly-Normal Power Series

In order to characterize right matrix Padé forms of nearly-normal matrix power series, we return to the vector equation

$$H_{m/n} \cdot X = 0 , \tag{3.6}$$

where X represents an $p(n+1)\times 1$ column vector. When we find p linearly independent solutions we get a right Padé form denominator $V_{m/n}(z)$, and from this also get the numerator $U_{m/n}(z)$.

In order to determine the solutions to (3.6), we row reduce the coefficient block matrix $H_{m/n}$. We do this in a manner similar to the one dimensional case. However, we need to be a little more careful because of the lack of commutativity of matrices.

Let $H_{(m-l)/(n-l)}$ be the largest non-singular principal minor block submatrix of $H_{m/n}$. Since $H_{(m-l)/(n-l)}$ is non-singular, by theorem 3.4, we can solve for both the right and the left matrix Padé forms of type $(m-l,n-l)$. Let

$$V^{(L)}_{(m-l)/(n-l)}(z) = \sum_{i=0}^{n-l} v_i^{(L)} z^i ,$$

$$U^{(L)}_{(m-l)/(n-l)}(z) = \sum_{i=0}^{m-l} u_i^{(L)} z^i$$

be the left matrix Padé form for A(z) of type (m-l,n-l) with $v_0^{(L)} = I_p$, and let

$$V^{(R)}_{(m-l)/(n-l)}(z) = \sum_{i=0}^{n-l} v_i^{(R)} z^i ,$$

$$U^{(R)}_{(m-l)/(n-l)}(z) = \sum_{i=0}^{m-l} u_i^{(R)} z^i$$

be the right matrix Padé form for A(z) of type $(m-l,n-l)$ with $v_0^{(R)} = I_p$. If there is no largest nonsingular principal minor block submatrix, then we set both the right and left denominator to be I_p and the right and left numerator to be

$a_0 + \cdots + a_{m-n} z^{m-n}$.

Since

$$V^{(L)}_{(m-l)/(n-l)}(z) \cdot A(z) - U^{(L)}_{(m-l)/(n-l)}(z) = z^{m+n-2l+1} \cdot \sum_{i=0}^{\infty} r_i^{(L)} z^i ,$$

we have the following block matrix identity

$$\left[v^{(L)}_{n-l}, \cdots, v^{(L)}_0 \right] \cdot \begin{bmatrix} a_{m-n+1} & \cdots & a_{m-l} & \cdots & a_{m+1} \\ \cdots & & \cdots & \cdots & \cdots \\ a_{m-l+1} & \cdots & a_{m+n-2l} & \cdots & a_{m+n-l+1} \end{bmatrix} = \left[0, \ldots, 0 , r_0^{(L)} , \cdots , r_{l-1}^{(L)} \right] .$$

As in the one dimensional case, the special structure of the block Hankel matrix creates the following block matrix identity

$$L \cdot H^{+}_{m/n} = H^{*}_{m/n} , \qquad (3.13)$$

where

$$L = \left[\begin{array}{cccc|ccc} I & & & & & & \\ & I & & & & & \\ & & \cdots & & & & \\ & & & I & & & \\ \hline v^{(L)}_{n-l} & \cdots & \cdots & & v^{(L)}_0 & & \\ & \cdots & \cdots & & & \cdots & \\ & & v^{(L)}_{n-l} & & & & v^{(L)}_0 \end{array} \right] ,$$

$$H^{+}_{m/n} = \left[\begin{array}{ccc|ccc} a_{m-n+1} & & a_{m-l} & a_{m-l+1} & \cdots & a_{m+1} \\ \cdots & \cdots & & & \cdots & \\ \cdots & & \cdots & & \cdots & \\ a_{m-l} & & & a_{m+n-2l} & \cdots & a_{m+n-l} \\ \hline a_{m-l+1} & & a_{m+n-2l} & & \cdots & \\ \cdots & \cdots & \cdots & & \cdots & \\ a_m & \cdots & a_{m+n-l-1} & & \cdots & a_{m+n} \end{array} \right]$$

and

$$H^*_{m/n} = \left[\begin{array}{ccc|ccc} a_{m-n+1} & & a_{m-l} & a_{m-l+1} & & a_{m+1} \\ & \cdots & & & & \\ & & \cdots & & & \\ a_{m-l} & & & a_{m+n-2l} & & a_{m+n-l} \\ \hline 0 & 0 & & r_0^{(L)} & \cdots & r_l^{(L)} \\ \cdots & & \cdots & & \cdots & \\ 0 & & r_0^{(L)} & r_l^{(L)} & \cdots & r_{2l}^{(L)} \end{array}\right]$$

Since $H_{(m-l)/(n-l)}$ is the largest block non-singular principal minor we can deduce that $r_0^{(L)}$ is a singular matrix. But, if the power series was (m,n) nearly-normal, then equation (3.11),and (3.12) shows we can say more, namely, that

$$r_0^{(L)} = r_1^{(L)} = \cdots = r_{l-1}^{(L)} = 0 .$$

The above condition implies that

$$V^{(L)}_{(m-l)/(n-l)}(z)\cdot A(z) - U^{(L)}_{(m-l)/(n-l)}(z) = z^{m+n-l+k+1}\cdot\sum_{i=0}^{\infty} r_i z^i ,$$

where $k \geq 0$, and $r_i = r^{(L)}_{l+k+i}$, with r_0 an invertible p x p matrix (since A(z) is a nearly-normal power series) .

If $k \leq l$ then the right hand side of equation (3.11) becomes

$$\left[\begin{array}{ccc|ccc} a_{m-n+1} & \cdots & a_{m-l} & \cdots & \cdots & a_{m+1} \\ \cdots & & \cdots & & & \cdots \\ \cdots & & & & & \cdots \\ a_{m-l} & \cdots & & \cdots & \cdots & a_{m+n-l} \\ \hline & & & 0 & & \\ & & & & & r_0 \\ & 0 & & & \cdots & \cdots \\ & & & 0 & r_0 \quad \cdots & r_{l-k-1} \end{array}\right] . \qquad (3.14)$$

Since the matrix r_0 is invertible, then so is the matrix

$$\begin{bmatrix} 0 & 0 & \cdots & 0 & r_0 \\ 0 & & \cdots & r_0 & r_1 \\ & & \cdot & & \\ & r_0 & & & \\ r_0 & r_1 & \cdots & & \end{bmatrix} .$$

A basis for the solutions of the system of equations given by the matrix equation (3.14) can now be determined to be the vectors from the columns of the following matrices:

$$\left[v_{n-l}^{(R)} , \cdots , v_0^{(R)} , 0, 0, \cdots , 0, \ldots , 0 \right]^t \tag{3.15}$$

$$\left[0, v_{n-l}^{(R)} , \cdots , v_0^{(R)} , \ldots, 0, \cdots , 0 \right]^t \tag{3.16}$$

$$\cdots \; \cdots \; \cdots$$

$$\left[0 , \cdots , 0 , v_{n-l}^{(R)} , \cdots , v_0^{(R)} , \cdots , 0 \right]^t . \tag{3.17}$$

Here the number of zero block rows varies from 0 and l for (3.14), 1 and l-1 for (3.16), and finally k and l - k for (3.17). Thus, a general solution to equation (3.6) in the case of an (m,n) nearly-normal matrix power series is a linear combination of the $p \cdot (k+1)$ vectors coming from (3.15) through to (3.17).

Since the v's will now all come from the right matrix Padé form we will drop the superscript (R). In its place we let the superscript (i) mean that we are talking about the i th column of a block vector. Then the general solution to (3.6) looks like

$$X = \begin{bmatrix} v_{n-l}^{(1)} \\ \cdots \\ \cdots \\ v_0^{(1)} \\ 0 \\ \cdots \\ 0 \end{bmatrix} \alpha_{1k} + \cdots + \begin{bmatrix} v_{n-l}^{(p)} \\ \cdots \\ \cdots \\ v_0^{(p)} \\ 0 \\ \cdots \\ 0 \end{bmatrix} \alpha_{pk} + \cdots + \begin{bmatrix} 0 \\ \cdots \\ v_{n-l}^{(1)} \\ \cdots \\ v_0^{(1)} \\ 0 \\ \cdots \end{bmatrix} \alpha_{10} + \cdots + \begin{bmatrix} 0 \\ \cdots \\ v_{n-l}^{(p)} \\ \cdots \\ v_0^{(p)} \\ 0 \\ \cdots \end{bmatrix} \alpha_{p0} . \tag{3.18}$$

By combining the columns back into their matrix forms and taking p linearly independent solutions, we see that a general block matrix solution has the form

$$X^{*} = \begin{bmatrix} v_{n-l} \\ \cdots \\ \cdots \\ v_0 \\ 0 \\ \cdots \\ 0 \end{bmatrix} \alpha_k + \cdots + \begin{bmatrix} 0 \\ \cdots \\ v_{n-l} \\ \cdots \\ v_0 \\ 0 \\ \cdots \end{bmatrix} \alpha_0 , \tag{3.19}$$

where the α_i are p x p matrices of constants for $i = 0, \cdots, k$.

Thus, since we are interested in the general solution , we get the following matrix equation

$$V_{m/n}(z) = z^{l-k}\left\{\sum_{i=0}^{n-l} v_i z^i\right\}\cdot\left\{\sum_{i=0}^{k} \alpha_i z^i\right\} . \tag{3.20}$$

From equation (3.5) we also show that

$$U_{m/n}(z) = z^{l-k}\left\{\sum_{i=0}^{m-l} u_i z^i\right\}\cdot\left\{\sum_{i=0}^{k} \alpha_i z^i\right\} . \tag{3.21}$$

Notice that the linear independence requirement condition, condition 3, of definition 3.1 implies that we choose $\alpha(z)$ to be nonsingular. In addition, if we are asking for a specific solution to the right matrix Padé problem for a nearly-normal power series then we can ensure that we ask for one in which the first nonzero term is invertible, and hence we can always find a Padé form where the denominator has nonzero determinant.

We summarize our findings by

Theorem 3.9. Let A(z) be a nearly-normal matrix power series and m and n two non-negative integers. Then there exists a right matrix Padé form $U_{m/n}(z)$ and $V_{m/n}(z)$ for A(z) of type (m,n) with the property that the first nonzero term of the denominator is the identity. Furthermore $U_{m/n}(z)$ and $V_{m/n}(z)$ are unique up to right greatest common divisor by an invertible matrix polynomial.

As with the previous section the proof of this result is a parallel of the result from the previous chapter and will not be included here. In addition, the theorem is true if all matrix multiplication is considered from the left instead of the right.

3.8. Scaled Matrix Padé Fractions

Since we know the general form of a Padé form for a nearly-normal matrix power series, we can restrict our attention to finding a particular Padé form . We call this particular solution a right scaled matrix Padé fraction. To get these fractions we simply take the general form and ask that

$$\alpha_0 = \alpha_1 = \cdots = \alpha_{k-1} = 0 \quad \text{and} \quad \alpha_k = I_p ,$$

where I_p stands for the p x p identity matrix. Then we get particular solutions of the form

$$T_{m/n}(z) = z^l \sum_{i=0}^{n-l} v_i z^i \tag{3.22}$$

$$S_{m/n}(z) = z^l \sum_{i=0}^{m-l} u_i z^i . \tag{3.23}$$

Note that both the numerator and the denominator have invertible matrices for their leading coefficients. This leads us to

Definition 3.10. A pair of matrix polynomials $T_{m/n}(z)$ and $S_{m/n}(z)$ is said to be a right scaled matrix Padé fraction of type (m,n) for the matrix power series A(z) if

a) $\min(\ m - \partial(T_{m/n}),\ n - \partial(S_{m/n})\ ) = 0$,

b) $A(z)\cdot T_{m/n}(z) - S_{m/n}(z) = z^{m+n+1}\cdot R_{m/n}(z) \qquad with\ ord(R_{m/n}) \geq 0$,

c) $RGCD(\ S_{m/n}(z)\ ,\ T_{m/n}(z)\ ) = z^{l}\cdot I_p$, and

d) The first nonzero term of $T_{m/n}(z)$ is invertible.

In the scalar case, where p = 1, we have shown that scaled Padé fractions exist and are unique up to a multiplicative constant. If A(z) is a nearly normal power series then we can get existence and uniqueness up to constant matrix multiple.

As in the scalar case we have

Theorem 3.11. Let A(z) be an (m,n) nearly-normal square matrix power series. Then there are a pair of scaled matrix right (or left) Padé fractions of type (m,n). These scaled matrix Padé fractions are unique up to right (or left) multiplication by an invertible constant matrix.

3.9. Alternate Definitions of Padé Forms

The definition for Padé form given at the start of the chapter here is not the only definition found in the literature. The idea behind any type of Padé form is one of approximating an infinite power series by a rational matrix expression, with the approximation being good to a specified order condition. That is, we are looking for two matrix polynomials N(z) and D(z) of certain degrees that satisfy

$$A(z) = N(z)\cdot D(z)^{-1} + z^{s}\cdot R(z) \tag{3.24}$$

where s is some integer (usually specified in advance).

If s = m+n+1, and we multiply (3.14) by D(z) on the right on both sides we get our condition 3.1. In addition we ask for certain degree limitations on both numerator and denominator along with a predetermined order condition.

These conditions are fairly standard and agree with most definitions of Padé forms of type (m, n). Where our definition differs from the classical definition (cf Bose [6] , Bultheel [10] , Starkand [30]) comes in condition a) . The classical version replaces condition a) with the requirement that $T_{m/n}(z)$ have an invertible constant term. After normalization this condition becomes $T_{m/n}(0) = I_p$, the p x p identity matrix.

The problem with the classical definition is that a Padé form does not always exist for given (m,n), even in the p = 1 scalar case. For purposes of creating a recursive algorithm this is a major stumbling block. In order to get existence the normality condition is then imposed on the power series.

Another definition that appears in the literature (cf Rissanen [26]) takes as its starting point the special nature of matrix inversion that differs from standard polynomial reciprocal taking. For example, if the denominator D(z) is a unimodular matrix , that is, it has a constant determinant, then $D(z)^{-1}$ is also a matrix polynomial and so equation (3.24) is really a polynomial approximation.

If we write $D(z)^{-1}$ in its adjoint form we get

$$D(z)^{-1} = \frac{adj(D(z))}{det(D(z))} . \tag{3.25}$$

so (3.24) becomes

$$A(z) = N(z)\cdot\frac{adj(D(z))}{det(D(z))} + z^{s}\cdot R(z) .$$

Multiplying both sides by the scalar polynomial det(D(z)) gives the equation

$$A(z)\cdot d(z) - N'(z) = z^{s}\cdot R'(z)$$

where d(z) = det(D(z)) and $N'(z) = N(z)\cdot adj(D(z)$.

An alternate definition can now be given that asks to find a scalar polynomial d(z) and a matrix polynomial $N'(z)$ that satisfy certain degree requirements and asks for certain order conditions. For more about this definition and an algorithm to calculate these quantities we refer the reader to the article by Rissanen. Included there is an application to calculating characteristic and minimal polynomials of a square matrix.

Chapter 4

A Matrix Offdiagonal Algorithm

As mentioned in the first chapter, the Padé table of a scalar power series A(z) is a doubly infinite array of rational expressions of the form

$$r_{m/n} = \frac{U_{m/n}(z)}{V_{m/n}(z)} = \frac{\sum_{i=0}^{m} u_i z^i}{\sum_{i=0}^{n} v_i z^i},$$

where the MacLaurin series of the rational expression $r_{m/n}$ is to agree with as many terms of A(z) as possible.

It was shown in chapter 2 that these rational expressions of A(z) are not independent of each other. Various relationships exist between neighboring elements in a Padé table. These relationships have been used to develop algorithms for the calculation of individual elements of the Padé table based on previously calculated entries of the table. In the scalar case, existing order n^2 algorithms include the ϵ-algorithm of Wynn [34] , the η-algorithm of Bauer [4] , and the Q-D algorithm of Rutishauser [28] . For a more complete survey of existing methods, we refer the reader to the articles of Brezinski [8] , Claessens [16] , and Wynn [34] .

These algorithms all suffer from the same disadvantage, namely, the requirement of normality of the power series. Algorithms that calculate Padé fractions in the degenerate case are given by Brent et al [7] , Bultheel [11] , and Rissanen [25] . Another algorithm that could handle the degenerate situation was given by Cabay and Kao [13] . This was an order n^2 algorithm. It was later extended by Cabay and Choi [14] into the offdiagonal algorithm.

The offdiagonal algorithm is superior to the previous algorithms. Besides being able to handle the degenerate case where the power series is nonnormal, offdiagonal is a faster algorithm. The offdiagonal algorithm is an $n \cdot \log^2 n$ algorithm (when m = n), if fast multiplication and division of polynomials via fast Fourier transforms and the Newton's division method are possible in the base field K. If we compare costs, the Anti-Diagonal algorithm of Brent et al uses fast multiplication and division and is also of order $n \log^2 n$. However, it was shown experimentally by Verheijen [32] , that the offdiagonal algorithm has a smaller constant of proportionality than that of the Anti-Diagonal algorithm . Rissanen's algorithm does not use fast multiplication or division and gives an order n^2 algorithm. If fast arithmetic techniques are not used then the offdiagonal algorithm is also of order n^2. However, Cabay and Kao [13] showed that the constant of proportionality is again reduced by a factor of 1.5 over that of Rissanen.

In addition to a smaller number of arithmetic operations, there are other advantages found in the offdiagonal algorithm. It is an iterative algorithm rather than a recursive one, thus allowing significant cost savings during implementation. It also produces intermediate polynomial sequences as a by-product, which is often desirable in practice (see, for example, Chapter 5).

The story for matrix power series and their Padé forms parallels the scalar situation in that most algorithms require invertibility (and hence normality) at every step. Algorithms that required the normality condition include those of Bultheel [10] , Bose and Basu [6] , Starkand [30] , and Rissanen [27] .

In this chapter we develop a matrix algorithm that extends the offdiagonal algorithm from scalar power series to matrix power series. The algorithm is of complexity $O(n \cdot \log^2 n \cdot p^r)$. This compares favorably with the algorithms of Bultheel, Bose and Basu, and Starkand which are all of complexity $O(n^3 \cdot p^r)$ and with the algorithm of Rissanen which is $O(n^2 \cdot \{p^r + p^2\})$. (As mentioned in chapter 1, p^r represents the cost of multiplying two p by p matrices.) The four previously derived algorithms all require normality. The only requirement that we need is that the power series be nearly-normal. Finally, our algorithm is presented in the more general case where we have a quotient of two power series. When the denominator of the quotient power series is set to be the identity, we get Padé fractions for a single power series.

4.1. The Matrix Scaled Padé Table

To explain the workings of this algorithm it is helpful to use a variation of a classical device called the scaled Padé table. If A(z) is a nearly-normal matrix power series then we have both the existence and uniqueness of scaled right matrix Padé fractions. Because of this, we can create a doubly infinite array

$$
\begin{array}{|llll}
\hline
P_{-1/-1} & P_{-1/0} & \cdots & \cdots \\
P_{0/-1} & P_{0/0} & \cdot & \cdot \\
P_{1/-1} & P_{1/0} & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot \\
P_{i/-1} & P_{i/0} & \cdots & P_{i/j} \\
\cdots & \cdots & \cdots & \cdots
\end{array}
$$

In the array, $P_{i/j}$ is the (unique up to right matrix multiplication) right scaled matrix Padé fraction of type (i,j) for A(z). The array is called the extended right scaled matrix Padé table. The term extended comes from the addition of an extra row and column, (row -1 and column -1). The entries in these positions are determined by the right scaled matrix Padé fractions

$$S_{m/-1}(z) = -z^m \cdot I_p \ , \quad T_{m/-1}(z) = 0 \ ,$$

$$S_{-1/n}(z) = 0 \ , \qquad T_{-1/n}(z) = -z^n \cdot I_p \ ,$$

for $m \geq -1$ and $n \geq 0$.

There is of course a similar table for left scaled matrix Padé fractions. Since we will be limiting our discussion to right scaled matrix Padé fractions for the rest of the chapter, we will drop the terms right and matrix along with the modifier extended and just refer to this table as the scaled Padé table for A(z).

Notice that determination of the scaled Padé table will also provide us with a table of Padé fractions for A(z) . Thus, the table of scaled Padé fractions gives us the classical version of Padé table (see Gragg [19]).

4.2. Algorithm Description

For a given integer pair (m,n) the offdiagonal algorithm calculates scaled Padé fractions for any integer pair (M,N) of the scaled Padé table situated on the (m,n) off-diagonal

$$\left\{ (M, N) \,\middle|\, M-N = m-n \right\}$$

with M ≤ m. Note that we can assume that $m \geq n$, since otherwise we can work with the formal power series for the matrix power series $A(z)^{-1}$ (which exists since we have assumed that A(0) is invertible). At any given stage, the algorithm produces two right scaled matrix Padé fractions on this m-n off-diagonal, a predecessor and a present fraction.

The predecessor is a right scaled matrix Padé fraction of type (m', n') that satisfies

$$A(z) \cdot T_{m'/n'}(z) - S_{m'/n'}(z) = z^{m'+n'+1} \cdot R_{m'/n'}(z) \qquad (4.1)$$

where $R_{m'/n'}(0)$ is non-zero. The integers m' and n' satisfy $m' - n' = m - n$. Thus, the predecessor is the unique scaled Padé fraction of type (m',n'), with the property that the order condition is exact.

The present is a right scaled matrix Padé fraction that satisfies

$$A(z)\cdot T_{(m'+1)/(n'+1)}(z) - S_{(m'+1)/(n'+1)}(z) = z^{(m'+1)+(n'+1)+1+k}\cdot R_{(m'+1)/(n'+1)}(z) \qquad (4.2)$$

where $R_{(m'+1)/(n'+1)}(0)$ is non-zero with $k \geq 0$, or $k = \infty$ in which case $R_{(m'+1)/(n'+1)}(z) = 0$. Thus, the present is the next scaled Padé fraction after the predecessor. Its order condition may or may not be exact.

We note that not every scaled Padé fraction can be called a present node, since our definition requires the present node to follow a node which meets the order requirements exactly. However, if $S_{m/n}(z)$ and $T_{m/n}(z)$ is a scaled Padé fraction of type (m,n), with

$$RGCD(S_{m/n}(z), T_{m/n}(z)) = z^h \cdot I_p,$$

then there must be a predecessor/present pair of Padé fractions at nodes (m-h-1,n-h-1) and (m-h,n-h), respectively. For, by scaling backwards, we get that

$$S_{(m-h)/(n-h)}(z) = z^{-h}\cdot S_{m/n}(z)$$

$$T_{(m-h)/(n-h)}(z) = z^{-h}\cdot T_{m/n}(z)$$

satisfies

$$A(z)\cdot T_{(m-h)/(n-h)}(z) - S_{(m-h)/(n-h)}(z) = z^{(m-h)+(n-h)+(h+k)+1}\cdot R_{m/n}(z).$$

Thus, the order condition is met. In addition, the degree requirements are met by construction, so $S_{(m-h)/(n-h)}(z)$, $T_{(m-h)/(n-h)}$ are the scaled Padé fractions of type (m-h,n-h) by uniqueness. There is no reason that the order condition be exact here. However, for the previous node we must have

$$A(z)\cdot T_{(m-h-1)/(n-h-1)}(z) - S_{(m-h-1)/(n-h-1)}(z) \stackrel{?}{=} z^{m+n-2h-l}\cdot R_{(m-h-1)/(n-h-1)}(z),$$

where $R_{(m-h-1)/(n-h-1)}(0) \neq 0$. This is so because, if the order condition were higher, then we could scale forward to the next node (m-h,n-h) and get another pair of matrix polynomials that are scaled right Padé fractions but with right greatest common divisor at least zI_p. Then the pair $S_{(m-h)/(n-h)}$ and $T_{(m-h)/(n-h)}(z)$ would be a scaled Padé fraction with a different right greatest common divisor contradicting the uniqueness of scaled Padé fractions.

Notice that we can write

$$A(z)\cdot T_{(m-n-1)/-1}(z) - S_{(m-n-1)/-1}(z) = z^{m-n-1}R_{(m-n-1)/-1}(z),$$

where

$$T_{(m-n-1)/-1}(z) = 0, \quad S_{(m-n-1)/-1(z)} = -z^{m-n-1}\cdot I_p \quad \text{and } R_{(m-n-1)/-1}(z) = I_p\ ;$$

and,

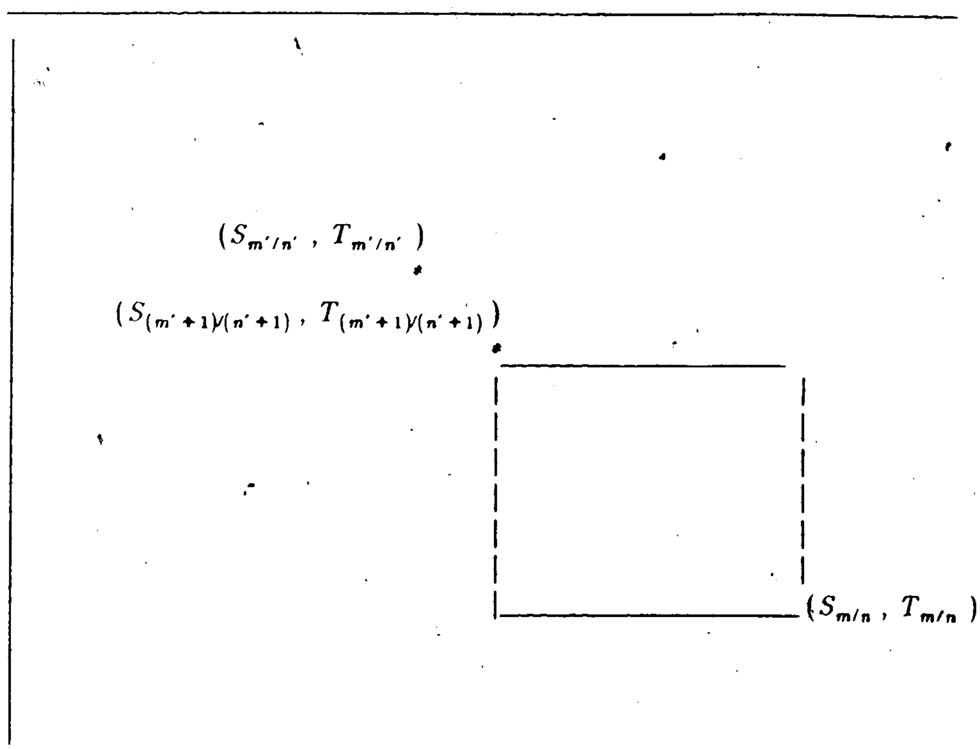
$$A(z)\cdot T_{(m-n)/0}(z) - S_{(m-n)/0}(z) = z^{m-n+1}R_{(m-n)/0}(z),$$

where

$$T_{(m-n)/0}(z) = I_p, \qquad S_{(m-n)/0}(z) = A(z)\ mod\ z^{m-n+1}\ .$$

Thus, we always have an initial value for the predecessor and present for any formal power series.

Graphically we can view the predecessor and present on the scaled Padé table as shown in the following diagram:



We use the predecessor and present fractions to move forward along the $m-n$ offdiagonal until we get the (m, n) node, i.e., until we get the right scaled Padé fraction for A(z) of type (m, n). Once we have a predecessor and present node, there are are two possibilities that face us.

Case 1: $\quad m-(m'+1) \leq k \quad$ (case of simple forward scaling).

If $m-(m'+1) \leq k$, then we can multiply both sides of equation (4.2) by $z^{m-m'-1}$ and get

$$A(z) \cdot T_{m/n}(z) - S_{m/n}(z) = z^{m+n+1} \cdot R_{m/n}(z),$$

where

$$T_{m/n}(z) = z^{m-m'-1} \cdot T_{(m'+1)/(n'+1)}(z),$$
$$S_{m/n}(z) = z^{m-m'-1} \cdot S_{(m'+1)/(n'+1)}(z),$$

and

$$R_{m/n}(z) = R_{(m'+1)/(n'+1)}(z).$$

The order condition holds because

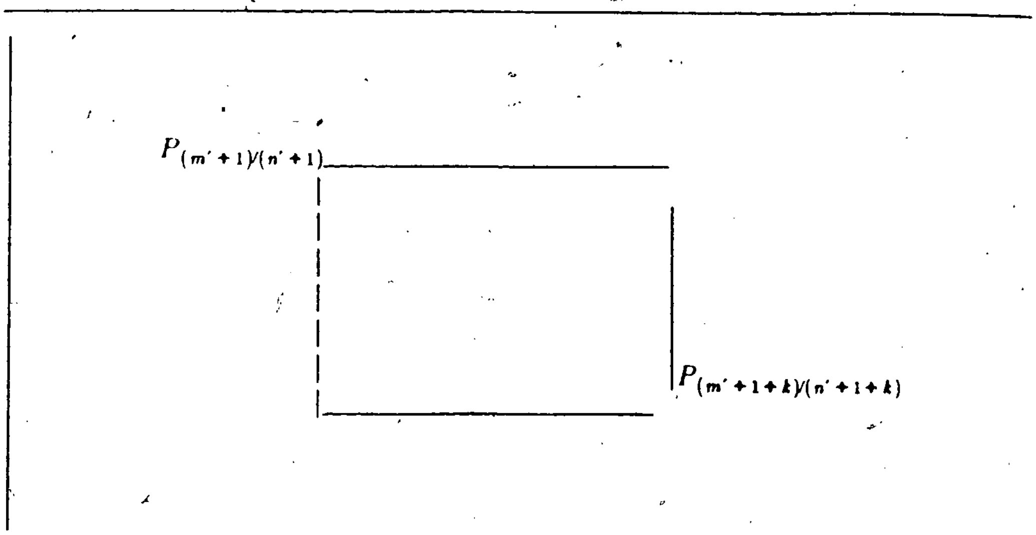
$$\begin{aligned}(m'+1)+(n'+1)+1+k+(m-m'-1) &= m+n'+1+k+1 \\ &= m+n+1+(k+n'-n+1) \\ &\geq m+n+1\,.\end{aligned}$$

We have used the fact that $k \geq m - m' - 1$, so $k \geq n - n' - 1$ (remember $n - n' = m - m'$). The degree condition holds because

$$\begin{aligned}\partial(T_{m/n}) &\leq n' + 1 + m - m' - 1 \\ &\leq n' - m' + m \\ &\leq n - m + m \quad (\textit{since } n' - m' = n - m) \\ &\leq n.\end{aligned}$$

Similarly, we can also show that $\partial(S_{m/n}) \leq m$. In addition, it is easy to show that $\partial(T_{m/n}) = n$ or $\partial(S_{m/n}) = m$. Finally, it is clear that the right greatest common divisor of $T_{m/n}(z)$ and $S_{m/n}(z)$ is z^u, for some non-negative integer u since the polynomials $T_{(m'+1)/(n'+1)}(z)$ and $S_{(m'+1)/(n'+1)}(z)$ have this property.

By the uniqueness of scaled Padé fractions, we have then found the unique (up to scalar multiple) scaled Padé fractions of type (m, n). In particular, we note that this implies that there is a box-like structure in the classical Padé table where the nodes are viewed as Padé fractions. That is, the Padé fraction of nearly-normal power series A(z) is the same for a box starting from the top upper left corner $(m'+1,\ n'+1\ )$ to all nodes k units to the right and k units down.



Case 2: $m - (m'+1) > k$

The second situation arises when $m - (m'+1) > k$. Then, by multiplying equation (4.1) by z^{2+k}, we get the equation

$$A(z)\cdot T_{m'/n'}(z)\cdot z^{2+k} - S_{m'/n'}(z)\cdot z^{2+k} = z^{(m'+1)+(n'+1)+1+k}\cdot R_{m'/n'}(z). \qquad (4.4)$$

Notice that the order condition is now the same as in the present

$$A(z)\cdot T_{(m'+1)/(n'+1)}(z) - S_{(m'+1)/(n'+1)}(z) = z^{(m'+1)+(n'+1)+1+k}\cdot R_{(m'+1)/(n'+1)}(z). \quad (4.5)$$

If we now multiply equation (4.4) on the right by a polynomial $T'(z)$ and multiply equation (4.5) on the right by a polynomial $S'(z)$ and follow this by subtracting the second equation from the first, we get

$$A(z)\cdot T(z) - S(z) = z^{(m'+1)+(n'+1)+1+k}\cdot R'(z), \qquad (4.6)$$

where

$$T(z) = z^{2+k}\cdot T_{m'/n'}(z)\cdot T'(z) - T_{(m'+1)/(n'+1)}(z)\cdot S'(z)$$

and

$$S(z) = z^{2+k} \cdot S_{m'/n'}(z) \cdot T'(z) \; - \; S_{(m'+1)/(n'+1)}(z) \cdot S'(z) \, .$$

Also,

$$\begin{aligned} R^*(z) &= R_{m'/n'}(z) \cdot T'(z) - R_{(m'+1)/(n'+1)}(z) \cdot S'(z) \\ &= R_{(m'+1)/(n'+1)}(z) \left\{ \left\{ R_{(m'+1)/(n'+1)}(z) \right\}^{-1} \cdot R_{m'/n'}(z) \cdot T'(z) - S'(z) \right\} \\ &= R_{(m'+1)/(n'+1)}(z) \left\{ A'(z) \cdot T'(z) - S'(z) \right\}, \end{aligned}$$

where $A'(z)$ is the matrix power series given by

$$A'(z) = \left\{ R_{(m'+1)/(n'+1)}(z) \right\}^{-1} \cdot R_{m'/n'}(z). \tag{4.7}$$

$R_{m'/n'}(z)$ is a unit matrix power series since both $R_{(m'+1)/(n'+1)}(0)$ and $R_{m'/n'}(0)$ are nonsingular. (Recall that A(z) being nearly-normal implies that both predecessor and present residuals will start with invertible matrices. Thus, the inverse of the present residual power series exists.)

The object now is to choose the matrix polynomials $S'(z)$ and $T'(z)$ in equation (4.6) so that S(z) and T(z) become the scaled Padé fractions of type (m,n) for A(z). Let $m'' = m - m' - 1$ and $n'' = n - n' - k - 2$. Then $m'' > n''$ and $n'' = m - m' - k - 2$, since $m - n = m' - n'$. Notice that $n'' \geq 0$ since $m - m' - 1 > k$. We choose $S'(z)$ and $T'(z)$ to be the scaled Padé fraction of type (m'', n'') for $A'(z)$. Then $ord(R^*(z)) \geq m'' + n'' + 1$ and so in equation (4.6)

$$\begin{aligned} ord(A(z) \cdot T(z) - S(z)) &= (m'+1) + (n'+1) + k + 1 + ord(R^*(z)) \\ &\geq m' + m'' + 1 + n' + n'' + k + 2 + 1 \\ &= m + n + 1 \, . \end{aligned}$$

To show that the pair S(z), T(z) in equation (4.6) is a scaled matrix Padé fraction for A(z), we need to prove that the degrees of S(z) and T(z) are at most m and n, respectively, with at least one of the degrees being exact, and that the RGCD of the two matrix polynomials is of the correct type. But,

$$\begin{aligned}
\partial(T) &\le \max\left(\partial(T') + \partial(T_{m'/n'}) + 2 + k\ ,\ \ \partial(S') + \partial(T_{m'+1/n'+1}) \right) \\
&\le \max\left(n'' + n' + k + 2\ ,\ \ m'' + n' + 1 \right) \\
&\le \max\left(n - n' - k - 2 + n' + k + 2\ ,\ \ m - m' - 1 + n' + 1 \right) \\
&\le \max\left(n\ ,\ \ m - (m' - n') \right) \\
&\le \max\left(n\ ,\ \ m - (m - n) \right) \\
&\le \max\left(n\ ,\ n \right) \\
&\le n\ .
\end{aligned}$$

Similarly, $\partial(S) \le m$. In addition, we can show that $\partial(T_{m/n}) = n$ or $\partial(S_{m/n}) = m$ as long as one of $T'(z)$ or $S'(z)$ satisfy this condition.

To see that the resulting pair has a RGCD of the type $z^u \cdot I$ for some nonnegative power u requires an argument that parallels the scalar case and will not be included here. The reader can find a proof for the scalar case in Cabay and Choi [14]. Therefore we have reduced the (m,n) scaled Padé fraction problem for A(z) to solving the (m'',n'') scaled Padé fraction for the power series $A'(z)$ where $m'' < m$ and $n'' < n$.

The computation of the (m,n) scaled Padé fraction required both the present $(m'+1,n'+1)$ and the predecessor (m',n') scaled Padé fractions. If we also required the predecessor of the (m,n) scaled Padé fraction (for example, if we should now decide to calculate scaled Padé fractions beyond (m,n)) then we proceed as follows.

Denote the predecessor for $T'(z)$ and $S'(z)$ by $T'^{*}(z)$ and $S'^{*}(z)$ (every Padé fraction has a predecessor by the same argument as on pages 53-54). The new predecessor for T(z) and S(z) is given by the same formula, i.e.,

$$T^{*}(z) = z^{2+k}\, T_{m'/n'}(z) T'^{*}(z) - T_{(m'+1)/(n'+1)}(z) S'^{*}(z)$$

$$S^{*}(z) = z^{2+k}\, S_{m'/n'}(z) T'^{*}(z) - S_{(m'+1)/(n'+1)}(z) S'^{*}(z)\ .$$

To see that we indeed get the predecessor, notice that if

$$RGCD(T'(z)\,,\, S'(z)) = z^h \cdot I_p \;,$$

then the degree of $S^{\#}(z)$, $T^{\#}(z)$ is $m''-h-1$ and $n''-h-1$, respectively. In addition,

$$A'(z)\cdot T^{\#}(z) - S^{\#}(z) = z^{m''+n''-2h-1}\cdot R^{\#}(z)\,.$$

We can work out that

$$A(z)\cdot T^{\#}(z) - S^{\#}(z)$$

expands to

$$\left\{A(z)\cdot T_{m'/n'}(z) - S_{m'/n'}(z)\right\}\cdot T'^{\#}(z) - \left\{A(z)\cdot T_{(m'+1)/(n'+1)}(z) - S_{(m'+1)/(n'+1)}(z)\right\}\cdot S'^{\#}(z)\,.$$

This in turn reduces to

$$\left\{z^{m'+n'+3+k}\cdot R_{m'/n'}(z)\right\}\cdot T'^{\#}(z) - \left\{z^{m'+n'+3+k}\cdot R_{(m'+1)/(n'+1)}(z)\right\}\cdot S'^{\#}(z)$$

$$= z^{m'+n'+3+k}\cdot R_{(m'+1)/(n'+1)}(z)\left\{R_{(m'+1)/(n'+1)}(z)^{-1}\cdot R_{m'/n'}(z)\cdot T'^{\#}(z) - S'^{\#}(z)\right\}$$

$$= z^{m'+n'+3+k}\cdot R_{(m'+1)/(n'+1)}(z)\left\{A'(z)\cdot T'^{\#}(z) - S'^{\#}(z)\right\}$$

$$= z^{m'+n'+3+k}\cdot R_{(m'+1)/(n'+1)}(z)\left\{z^{m''+n''-2h-1}\cdot R^{\#}(z)\right\}$$

$$= z^{m'+n'+3+k+m''+n''-2h-1}\cdot R_{(m'+1)/(n'+1)}(z)\cdot R^{\#}(z)$$

$$= z^{m'+n'+2+k+(m-m'-1)+(n-n'-k-2)-2h}\cdot R_{(m'+1)/(n'+1)}(z)\cdot R^{\#}(z)$$

$$= z^{m+n-2h-1}\cdot R_{(m-h-1)/(n-h-1)}(z)\,, \quad where\;\; R_{(m-h-1)/(n-h-1)}(0) \neq 0\,.$$

4.2.1. Example of Offdiagonal Calculation

We will construct the scaled Padé fractions for a 2 x 2 matrix power series. To avoid cumbersome fractions we will consider the power series as having entries from the field Z_5, the field of integers mod 5. The formal power series is given by

$$A(z) = \begin{bmatrix} 1+4z+z^4+z^6+3z^8+3z^9 & 4z+z^4+3z^6+4z^8+4z^9 \\ z^4+4z^6+4z^8+3z^9 & 1+4z^6+4z^8+4z^9 \end{bmatrix} + R(z)$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 4 & 4 \\ 0 & 0 \end{bmatrix} z + \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} z^4 + \begin{bmatrix} 1 & 3 \\ 4 & 4 \end{bmatrix} z^6 + \begin{bmatrix} 3 & 4 \\ 4 & 4 \end{bmatrix} z^8 + \begin{bmatrix} 3 & 4 \\ 3 & 4 \end{bmatrix} z^9 + R(z),$$

where R(z) is of order ≥ 10. We will calculate the right scaled matrix Padé fractions for the set of integers (m,n) along the off-diagonal m - n = 1 up to and including entry (5,4).

Initially the (0,-1) and (1,0) entries in the scaled Padé table represent the initial values for the predecessor and present right scaled matrix Padé fractions. These are given by

$$T_{0/-1}(z) = 0 \quad S_{0/-1}(z) = -I$$

$$T_{1/0}(z) = I \text{ and } S_{1/0}(z) = \begin{bmatrix} 1+4z & 4z \\ 0 & 1 \end{bmatrix}.$$

We see that the predecessor and present satisfy

$$A(z)\cdot T_{0/-1}(z) - S_{0/-1}(z) = z^0 \cdot I$$

$$A(z)\cdot T_{1/0}(z) - S_{1/0}(z) = z^4 \cdot R_{1/0}(z),$$

where $R_{1/0}(z)$ is given by

$$R_{1/0}(z) = \begin{bmatrix} 1+z^2+3z^4+3z^5 & 1+3z^2+4z^4+4z^5 \\ 1+4z^2+4z^4+3z^5 & 4z^2+4z^4+4z^5 \end{bmatrix} + \cdots$$

$$= \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 3 \\ 4 & 4 \end{bmatrix} z^2 + \begin{bmatrix} 3 & 4 \\ 4 & 4 \end{bmatrix} z^4 + \begin{bmatrix} 3 & 4 \\ 3 & 4 \end{bmatrix} z^5 + \cdots .$$

Thus, we have the nodes (0,-1) and (1,0) of the scaled Padé table of A(z). Notice that by multiplying the scaled Padé fractions located at (1,0) by z and z^2, we can obtain the right scaled matrix Padé fractions at node locations (2,1) and (3,2), respectively, namely,

$$T_{2/1}(z) = \begin{bmatrix} z & 0 \\ 0 & z \end{bmatrix} \quad \text{and} \quad S_{2/1}(z) = \begin{bmatrix} z+4z^2 & 4z^2 \\ 0 & z \end{bmatrix}$$

$$T_{3/2}(z) = \begin{bmatrix} z^2 & 0 \\ 0 & z^2 \end{bmatrix} \quad \text{and} \quad S_{3/2}(z) = \begin{bmatrix} z^2+4z^3 & 4z^3 \\ 0 & z^2 \end{bmatrix}.$$

If we now align the predecessor and present we get

$$A(z)\cdot 0 - (-z^4\cdot I) = z^4\cdot I \tag{4.8}$$

$$A(z)\cdot I - S_{1/0}(z) = z^4\cdot R_{1/0}(z)\,. \tag{4.9}$$

To get the (5,4) entry in the Padé table we can multiply equation (4.8) by a matrix polynomial $T'(z)$ of degree ≤ 1 and equation (4.9) by a matrix polynomial $S'(z)$ of degree ≤ 4 and stay inside the degree requirements for the denominator and numerator. In addition, to meet our order condition we wish $T'(z)$ and $S'(z)$ to satisfy the condition that the difference

$$I\cdot T'(z) - R_{1/0}(z)\cdot S'(z)$$

be of order 6.

Thus, we now look for matrix polynomials, denoted here $P_{4/1}(z)$ and $Q_{4/1}(z)$ to avoid confusion with the original scaled Padé fractions, that satisfy the condition of being the (4,1) right scaled matrix Padé fraction for the power series

$$\left\{R_{1/0}(z)\right\}^{-1} = \begin{bmatrix} z^2+3z^4+z^5 & 1+2z^4+z^5 \\ 1+z^2+4z^4 & 4+2z^2+3z^4 \end{bmatrix} + \cdots$$

$$= \begin{bmatrix} 0 & 1 \\ 1 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix} z^2 + \begin{bmatrix} 3 & 2 \\ 4 & 3 \end{bmatrix} z^4 + \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} z^5 + \cdots .$$

If we accept for the present time that the (4,1) right scaled matrix Padé fraction $P_{4/1}(z)$ and $Q_{4/1}(z)$ for $R_{1/0}(z)^{-1}$ are given by

$$P_{4/1}(z) = \begin{bmatrix} 4+4z+2z^2+2z^3+4z^4 & 2+z+2z^2+3z^3 \\ 3+3z & 2z+z^2+4z^4 \end{bmatrix}$$

$$Q_{4/1}(z) = \begin{bmatrix} 2+2z & 2+3z \\ 4+4z & 2+z \end{bmatrix},$$

then we can get that the (5,4) right scaled matrix Padé fraction for A(z) is given by

$$S_{5/4}(z) = z^4 S_{0/-1}(z)\cdot Q_{4/1}(z) - S_{1/0}(z)\cdot P_{4/1}(z)$$

$$= \begin{bmatrix} 1+3z+z^4+2z^5 & 3+z+z^2+z^4+z^5 \\ 2+2z+z^4+z^5 & 3z+4z^2+4z^4+4z^5 \end{bmatrix}.$$

Similarly, we get

$$T_{5/4}(z) = z^4 T_{0/-1}(z)\cdot Q_{4/1}(z) - T_{1/0}(z)\cdot P_{4/1}(z)$$

$$= \begin{bmatrix} 1+z+3z^2+3z^3+z^4 & 3+4z+3z^2+2z^3 \\ 2+2z & 3z+4z^2+z^4 \end{bmatrix}.$$

To see how we got the (4,1) right scaled matrix Padé fractions for $R_{1/0}(z)^{-1}$, we again use our method of the previous section. Let $A'(z) = R_{1/0}(z)^{-1}$. We get the (2,-1) and (3,0) nodes of the scaled Padé table for $A'(z)$. These results are given by

$$Q_{2/-1}(z) = 0 \quad , \quad P_{2/-1}(z) = -z^2 \cdot I$$

$$Q_{3/0}(z) = I \quad , \quad P_{3/0}(z) = \begin{bmatrix} z^2 & 1 \\ 1+z^2 & 4+2z^2 \end{bmatrix}.$$

These matrix polynomials satisfy the equations

$$A'(z) \cdot 0 - P_{2/-1}(z) = z^2 \cdot I$$

$$A'(z) \cdot I - P_{3/0}(z) = z^4 \cdot R'_{3/0}(z) ,$$

where

$$R'_{3/0}(z) = \begin{bmatrix} 3+z & 2+z \\ 4 & 3 \end{bmatrix}.$$

As before, all we need is the (1,0) right scaled matrix Padé fraction for the inverse of this residual matrix power series . Our calculations give

$$P'_{1/0}(z) = \begin{bmatrix} 3+3z & 3+2z \\ 1+z & 3+4z \end{bmatrix} \text{ and } \quad Q'_{1/0}(z) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Thus, returning from our recursive descent, we calculate

$$P_{4/1}(z) = z^2(-z^2 \cdot I) - \begin{bmatrix} z^2 & 1 \\ 1+z^2 & 4+2z^2 \end{bmatrix} \cdot \begin{bmatrix} 3+3z & 3+2z \\ 1+z & 3+4z \end{bmatrix}$$

$$= \begin{bmatrix} 4+4z+2z^2+2z^3+4z^4 & 2+z+2z^2+3z^3 \\ 3+3z & 2z+z^2+4z^4 \end{bmatrix}.$$

A similar equation gives the denominator as

$$Q_{4/1}(z) = \begin{bmatrix} 2+2z & 2+3z \\ 4+4z & 2+z \end{bmatrix}.$$

If we also wished to calculate the (4,3) Padé table entry for our original A(z) (which we would want to do in case we later wished to proceed further along the offdiagonal), then we would determine the predecessor of the (4,1) entry in the Padé table for $A'(z)$. Since this is just the (3,0) entry, we do not need to descend any further. We then use the (3,0) entry to get the (4,3) entry by means of

$$S_{4/3}(z) = z^4 S_{0/-1}(z)\cdot Q_{3/0}(z) - S_{1/0}(z)\cdot P_{3/0}(z)$$

$$= z^4\cdot I - \begin{bmatrix} 1+4z & 4z \\ 0 & 1 \end{bmatrix}\cdot\begin{bmatrix} z^2 & 1 \\ 1+z^2 & 4+2z^2 \end{bmatrix}$$

$$= \begin{bmatrix} z+4z^2+2z^3+z^4 & 4+2z^3 \\ 4+4z^2 & 1+3z^2+z^4 \end{bmatrix},$$

and

$$T_{4/3}(z) = z^4\cdot T_{0/-1}(z)\cdot Q_{3/0}(z) - T_{1/0}(z)\cdot P_{3/0}(z)$$

$$= \begin{bmatrix} 4z^2 & 4 \\ 4+4z^2 & 4+2z^2 \end{bmatrix}.$$

4.3. The MATRIX-OFFDIAG Algorithm

The results derived in the sections are summarized in the algorithm - MATRIX-OFFDIAG below. The algorithm computes the (m,n) scaled matrix Padé fraction for a quotient matrix power series $-B(z)^{-1}\cdot A(z)$ by solving

$$A(z)\cdot T_{m/n}(z) + B(z)\cdot S_{m/n}(z) = z^{m+n+1}\cdot R_{m/n}(z)\ .$$

The algorithm computation proceeds iteratively along the (m,n) offdiagonal calculating the Padé fractions in steps a power of 2. We require that the quotient power series be (m,n)-nearly normal. We then solve our original Padé problem by setting $B(z) = I_p$.

The algorithm for calculating the right scaled matrix Padé fraction is then given by an extension of the second algorithm of Choi and Cabay . (Choi , Cabay [14]). Essentially we use the results obtained in the last section. We give the version for right Padé form. The left version is easily determined from the right version.

ALGORITHM (MATRIX-OFFDIAG)

INPUT: A, B, m, n, p where

1) m , n , p are nonnegative integers with $m \geq n$ and $p \geq 1$.

2) A and B are $p \times p$ matrix power series (we note that we only require A mod z^{m+n+1} and B mod z^{m+n+1}). B must be a unit power series.

OUTPUT: Matrix polynomials S_1, S_0, T_1, T_0, and an integer IER where:

1) The pair S_1 and T_1 is the (m, n) right scaled matrix Padé fraction for, $-B^{-1} \cdot A$.

2) The pair S_0 and T_0 is the (m-h-1,n-h-1) right scaled matrix Padé fraction for $-B^{-1} \cdot A$, where h is the integer determined by:

$$z^h \cdot I_p = RGCD(S_1 , T_1)$$

(Thus S_0, T_0 and $z^{-h} S_1$, $z^{-h} T_1$ form a predecessor/present pair, the first such pair before (m,n).)

3) IER , an integer variable that is used to indicate if invertibility requirements have been met (0 if success, 1 if failure).

Step 1: # initialization #

IER ← 0

i ← -1

M ← (m - n)

$N \leftarrow 0$

$S_1 \leftarrow -B^{-1} \cdot A \bmod z^{M+1}$, $S_0 \leftarrow z^{M-1} \cdot I_p$,

$T_1 \leftarrow I_p$, $T_0 \leftarrow 0$

Step 2: # calculation of step-size #

$i \leftarrow i+1$

$s \leftarrow \min\{2^i - N, n - N\}$

Step 3: # Termination criterion #

If $s = 0$ exit.

Step 4: # Calculation of closest present node before the pair S_1, T_1 #

Determine h such that $z^h \cdot I_p = \mathrm{RGCD}(S_1, T_1)$

Set

$$S_1 \leftarrow z^{-h} \cdot S_1 \ , \quad T_1 \leftarrow z^{-h} \cdot T_1$$

Step 5: # Calculation of residual for $S_1 \cdot T_1^{-1}$ #

Compute k and the power series R_1 such that

$$(A \cdot T_1 + B \cdot S_1) \bmod z^{M+N+2s+1} = z^{M+N-h+k+1} R_1,$$

where $R_1(0) \neq 0$ if $k < 2s + h$.

Step 6: # Separating into different cases #

If $k \geq s$ then

$S_1 \leftarrow z^{s+h} \cdot S_1$,

$T_1 \leftarrow z^{s+h} \cdot T_1$

go to step 11

$R_1(0)$ is singular, then set IER = 1 and exit.

Else go to step 7

Step 7: # Calculation of degrees for residual scaled matrix Padé fraction #

$m' \leftarrow s + h$

$n' \leftarrow s - k - 1$

Step 8: # Computation of residual for $S_0 \cdot T_0^{-1}$ #

Compute R_0 so that

$$(A \cdot T_0 + B \cdot S_0) \; mod \; z^{M+N+m'+n'-2h} = z^{M+N-2h-1} R_0 ,$$

where $R_0(0) \neq 0$.

Step 9: # Computation of residual right scaled matrix Padé fractions #

$$S'_1 , S'_0 , T'_1 , T'_0 ,$$

determined from MATRIX-OFFDIAG (R_0, R_1, m' , n' , p,IER)

If IER = 1 then exit.

Step 10: # Advance along offdiagonal for scaled fractions #

$$S_1 \leftarrow S_1 \cdot T'_1 + z^{h+k+2} S_0 \cdot S'_1$$

$$T_1 \leftarrow T_1 \cdot T'_1 + z^{h+k+2} T_0 \cdot S'_1$$

$$S_0 \leftarrow S_1 \cdot T'_0 + z^{h+k+2} S_0 \cdot S'_0$$

$$T_0 \leftarrow T_1 \cdot T'_0 + z^{h+k+2} T_0 \cdot S'_0$$

Step 11: # Calculation of degrees of $S_1 \cdot T_1^{-1}$ #

$$M \leftarrow M + s$$
$$N \leftarrow N + s$$

go to step 2

4.4. Cost of Matrix-Offdiagonal Algorithm

Let $C_p(m,n)$ represent the cost of calculating the (m,n) right scaled matrix Padé fraction of the nearly-normal p x p matrix power series $B(z)^{-1} \cdot A(z)$. We will determine an asymptotic formula for this cost in terms of m,n, and p. The determination of the cost essentially follows the arguments from Cabay and Choi [14] into two parts, either $n \leq m \leq 2n$ or $m > 2n$. We first consider the case where $n \leq m \leq 2n$.

The first step that requires a nontrivial calculation is step 1, the calculation of the quotient power series $-B^{-1} \cdot A \ mod \ z^{m-n+1}$. To determine a quotient we can use fast inversion via Newton's method. This requires the calculation of the inverse of the leading term, an order p^r, operation and then on the order of n log n matrix multiplies since the number of terms desired is less than 3n. Thus the total cost of this step is of the order of $p^r \cdot n \log n$ operations. As mentioned previously, the superscript r is the order condition for matrix multiplication of two p by p matrices (this is the same order condition as inverting a p by p matrix).

To determine the cost of the rest of the steps we first notice that if we set q = $\lceil \log n \rceil$, where $\lceil \quad \rceil$ represents the ceiling of a real number, then the algorithm terminates after q iterations. At the beginning of the ith iteration the value of N, M, and the increment s in step 2 is found to be

$$s = \begin{cases} 1 & , i = 0, \\ 2^{i-1} & , 0 < i < q, \\ n - 2^{q-1} & , i = q, \end{cases} \tag{4.10}$$

$$N = \begin{cases} 0 & , i = 0, \\ 2^{i-1} & , i > 0, \end{cases} \tag{4.11}$$

and

$$M = N + (m - n). \tag{4.12}$$

Suppose that we are at the ith iteration. We will determine the cost of completing the (i+1) st iteration. To do this, we assume that the right scaled matrix Padé fraction S_1 and T_1 of type (M,N) and its predecessor S_0 and T_0 are known.

During the iterative step, the first nontrivial step is step 5 which calculates the residual $R_1(z)$ of the present. Let

$$A_1(z) = \sum_{j=0}^{M} a_j z^j, \quad B_1(z) = \sum_{j=0}^{N} b_j z^j \tag{4.13}$$

and

$$A_2(z) = \sum_{j=0}^{N+2s-1} a_{M+j+1} z^j, \quad B_2(z) = \sum_{j=0}^{M+2s-1} b_{N+j+1} z^j .$$

In the equation

$$(A \cdot T_1 + B \cdot S_1) \bmod z^{M+N+2s+1} = z^{M+N-h+k+1} R_1$$

(where h is determined from step 4), the left side is the same as

$$(A_1 T_1 + z^{M+1} A_2 T_1 + B_1 S_1 + z^{N+1} B_2 S_1) \bmod z^{M+N+2s+1}$$

and this is of order M + N − h + k + 1. Note that, since $A_1 T_1$ and $B_1 S_1$ are both matrix polynomials of degree at most M + N − h , they do not enter into the calculation of the residual. $R_1(z)$ and the integer k will be determined by the multiplication of the matrix polynomial T_1 with A_2 and S_1 with B_2. We can determine the residual by twice multiplying two matrix polynomials and getting two matrix polynomials, each of degree at most 2N + 2s − 1 < 4N. We can calculate this quantity by fast matrix polynomial multiplication with a cost of the order of N log N matrix multiplies, i.e of the order of $p^r N \log N$ operations.

If we do not terminate in step 6 then our next nontrivial step is the calculation of the residual for the predecessor. A similar argument to the one above results in a cost of the order of $p^r\ N \log N$ operations.

Step 9, the recursive call to compute the right scaled matrix fractions of type (m',n') for $A'(z)$ requires $C_p(m',n')$ operations. Notice that $n' \leq m' \leq 2N$.

The final nontrivial step is step 10 where we determine the new predecessor and present scaled Padé fractions. This step requires 8 matrix polynomial multiplications. Since $M \leq 2N$ each of these polynomial products are of degree at most m+s and so at most of degree 3N. By fast multiplication methods this requires on the order of $p^r\ N \log N$ operations.

Therefore the total cost of the i-th iteration (without step 1) is bounded by

$$\begin{aligned} &C_p(m',n') + c(2N)log(2N)\cdot p^r \\ &\leq C_p(2N,N) + c(2N)log(2N)p^r \\ &\leq C_p(2^i,2^{i-1}) + ci2^ip^r\ . \end{aligned}$$

But then the total cost (including step 1)in the situation where we have k iterations is bounded by

$$\begin{aligned} C_p(2^{k+1},2^k) &= \sum_{i=0}^{k}\left\{C_p(2^i,2^{i-1}) + ci2^ip^r\right\} + p^r\ n \log n \\ &= \sum_{i=0}^{k-1}\left\{C_p(2^i,2^{i-1}) + ci2^ip^r\right\} + C_p(2^k,2^{k-1}) + ck2^kp^r + p^r\ n \log n \\ &= 2C_p(2^k,2^{k-1}) + ck2^kp^r \\ &= 2^k\left\{C_p(2,1) + cp^r\sum_{i=0}^{k} i\right\} \\ &= 2^k\left\{C_p(2,1) + cp^r\frac{k(k+1)}{2}\right\}. \end{aligned}$$

If n = 2^k, then this is just of order

$$p^r\ n \log^2 n$$

Since for $0 \leq n \leq m \leq 2n$

$$C_p(m,n) \leq \sum_{i=0}^{k} \left\{ C_p(2^i, 2^{i-1}) + ci2^i p^r \right\} + p^r \cdot n \log n ,$$

we can get

Theorem 4.1. If $0 \leq n \leq m \leq 2n$, the MATRIX-OFFDIAG can compute the right scaled matrix Padé fraction of type (m,n) for a nearly-normal p x p matrix power series $B(z)^{-1} \cdot A(z)$ in time of $O(p^r \cdot n \cdot \log^2 n)$ operations where p^r is the cost of multiplying two p by p matrices.

Thus we can determine the cost for finding the (m,n) scaled Padé fraction as long as $m \leq 2n$. Suppose now that $m > 2n$. For ease of discussion we limit ourselves to the case where $B(z) = -I$. We can write A(z) as

$$A(z) = A_1(z) + z^{m-n+\mu} A_2(z) ,$$

where

$$A_1(z) = A(z) \bmod z^{m-n+1}$$

and where $A_2(0)$ is an invertible matrix if $\mu < \infty$ (since A is a nearly normal power series and $A_2(z)$ is just the residual of the (m-n,0) Padé fraction). If $\mu \leq n$, let the pair $S_{n/(n-\mu)}(z)$ and $T_{n/(n-\mu)}(z)$ be the right scaled matrix Padé fraction for $A_2(z)^{-1}$. Then we can determine that the right scaled matrix Padé fraction of type (m,n) for A(z) is given by (see Choi [15])

$$S_{m/n}(z) = \begin{cases} A_1(z) S_{n/(n-\mu)}(z) + z^{m-n+\mu} T_{n/(n-\mu)}(z), & \text{if } \mu \leq n \\ A_1(z) z^n, & \textit{otherwise} \end{cases} \tag{4.14}$$

$$T_{m/n}(z) = \begin{cases} S_{n/(n-\mu)}(z), & \text{if } \mu \leq n \\ z^n \cdot I_p , & \textit{otherwise.} \end{cases}$$

To see this in the case that $\mu \leq n$ we simply notice that

$$A(z) T_{m/n}(z) - S_{m/n}(z) = A_1(z) T_{m/n}(z) - S_{m/n}(z) + z^{m-n+\mu} A_2(z) T_{m/n}(z)$$

$$= z^{m-n+\mu} A_2(z) S_{n/(n-\mu)}(z) - z^{m-n+\mu} T_{n/(n-\mu)}(z)$$

$$= z^{m-n+\mu} A_2(z) \left\{ S_{n/(n-\mu)}(z) - A_2(z)^{-1} T_{n/(n-\mu)}(z) \right\}$$

$$= - z^{m-n+\mu} A_2(z) z^{n+n-\mu+1} R_{n/(n-\mu)}(z)$$

$$= - z^{m+n+1} R_{n/(n-\mu)}(z) .$$

It is an easy matter to check the degree and right greatest common divisor are of the correct type. Similarly it is not hard to check the case where $\mu > n$

The above result is helpful when we are determining the cost of the MATRIX-OFFDIAG algorithm for the scaled Padé fraction of type (m,n) in the case that $m > 2n$. For, then we can return the problem to one of finding a different scaled Padé fraction of type $(n,n-\mu)$ for the power series for $A_2(z)^{-1}$. The number of terms of the matrix polynomial $A_2(z)$ that enter into the calculations is at most 2n. Because of this, the calculation of the inverse power series is accomplished in order $p^r \cdot n \log n$ operations using fast division techniques. The cost of determining the $(n,n-\mu)$ scaled Padé fraction is bounded by $C_p(n,n)$ and hence is of the order $p^r n \log^2 n$. Finally, the cost of determining $S_{m/n}(z)$ from equation 4.14 using fast multiplication involves $m \log m$ matrix multiplies. Thus, we get

Theorem 4.2. Let (m,n) be arbitrary nonnegative integers, and let A(z) be a nearly-normal p x p matrix power series. Then the cost of calculating the (m,n) right scaled matrix Padé fraction of type (m,n) is in time of the order of $\left\{ m \log m + n \log^2 n \right\} p^r$ where p^r is the cost of multiplying two p by p matrices.

Chapter 5

Scaled Padé Fractions and Greatest Common Divisors

One of the more interesting properties of Padé forms is their relationship to the greatest common divisors of two polynomials. Relationships between the Padé forms of a quotient power series and the greatest common divisor of two polynomials have been studied by many other authors. See for example, Berlekamp [5] and McEliece and Shearer [22] . One of the more elegant relationships was discovered in 1984 by Cabay and Choi [14] . They proved that a second version of their offdiagonal algorithm, when applied to a quotient power series a(z)/b(z) for a specific off-diagonal was the same as the extended Euclidean algorithm as applied to the pair of polynomials $a^*(z)$ and $b^*(z)$. Because of the speed of the first version of their offdiagonal algorithm they provide an order $n \log^2 n$ algorithm for finding the greatest common divisor of two polynomials (if fast arithmetic is allowed - otherwise the algorithm finds the greatest common divisor in order n^2 operations).

In the matrix case less is known about any relationships between the greatest common divisor of two matrix polynomials and the matrix Padé forms of the related quotient matrix power series. However, any relationships found would be interesting, given the importance to engineers of finding the greatest common divisor of two matrix polynomials. For example, in linear systems theory, one often faces the task of modeling a system where the transfer matrix T(z) of the problem is known, or theoretically determined. This transfer matrix is a matrix of rational polynomials with p rows and q columns representing a p-input, q-output system. Since the elements are rational polynomials, the lowest common multiples of the denominators of each column can be determined. If these lowest common multiples are placed into a diagonal matrix D(z) (lowest common multiple of column i of the original matrix will be the i-th diagonal entry of D(z)), then, if N(z) is the original matrix with the lowest common divisors removed from T(z), we have a representation for the transfer function as

$$T(z) = N(z)\cdot D(z)^{-1}.$$

Notice that both N(z) and D(z) are matrix polynomials. It is clear that this representation is not always the desired way to represent T(z) as a matrix fraction of two polynomial matrices. In order to find a simpler representation one finds the greatest common divisor of the two matrix polynomials and divides out the common factors to get a new representation for T(z) (factors will be invertible since D(z) is invertible).

However, greatest common divisors for matrix polynomials are not as simple as they are in the one dimensional scalar case. In the scalar case, a nontrivial common factor is any factor that divides both polynomials and is nontrivial in the sense that it has degree at least one . For the matrix case, a common divisor can be of degree larger than one and yet be trivial in the sense that it does not just divide the two matrix polynomials in question, but divides every matrix polynomial. For example, in the 2×2 case the matrix polynomial

$$P(z) = \begin{bmatrix} 1 & z \\ 0 & 1 \end{bmatrix}$$

divides every matrix polynomial of compatible size since its inverse is also a matrix polynomial,

$$P(z)^{-1} = \begin{bmatrix} 1 & -z \\ 0 & 1 \end{bmatrix}.$$

Any polynomial matrix P(z) that has as its inverse another polynomial matrix is then called a trivial divisor. For if Q(z) is any other matrix polynomial of compatible size, then we would have

$$\begin{aligned} Q(z) &= Q(z)\cdot P(z)^{-1}\cdot P(z) \\ &= Q'(z)\cdot P(z), \end{aligned}$$

where $Q'(z)$ is a matrix polynomial. Thus, such matrix polynomials divide into every other matrix polynomial. Matrix polynomials that have the property that their inverses are also matrix polynomials are called unimodular matrices. They are characterized by the property that their determinant is a constant. Nontrivial divisors are then matrix polynomials with the property that the degree of the determinant is at least one.

As in the scalar case, there is a correspondence between greatest common divisors and scaled Padé fractions of quotient polynomials. We include the algorithm that results from this correspondence as an application of MATRIX-OFFDIAG. The problem with the algorithm is its requirement that the quotient power series be nearly-normal. Thus the algorithm gives a greatest common divisor algorithm but only for restricted pairs of matrix polynomials.

5.1. The MATRIX-OFFDIAG#2 Algorithm

Before investigating the relationship between the calculating of GCDs and the offdiagonal algorithm of chapter 4 we would like to redo the offdiagonal algorithm.

In our initial description of the offdiagonal algorithm in chapter 4, we had a node called the predecessor and a node called the present. The predecessor was a node on the (m,n) offdiagonal of the scaled Padé table, where the degree was exact, i.e.

$$A(z)\cdot T_{m'/n'}(z) + B(z)\cdot S_{m'/n'}(z) = z^{m'+n'+1}\cdot R_{m'/n'}(z) . \tag{5.1}$$

The present was then the next node on the offdiagonal; i.e. the scaled Padé fraction of type $(m'+1,\ n'+1)$ and so satisfied

$$A(z)\cdot T_{(m'+1)/(n'+1)}(z) + B(z)\cdot S_{(m'+1)/(n'+1)}(z) = z^{m'+n'+k+3}\cdot R_{(m'+1)/(n'+1)}(z) , \tag{5.2}$$

where $k \geq 0$. Notice that the present cannot have a nontrivial common divisor since then we would be able to factor out a power of z (since these are the only common factors allowed) and get two distinct scaled Padé fractions of type (m',n'), one pair having exact order condition the other not. Since scaled Padé fractions are unique up to nonsingular matrix multiplication, we would have a contradiction.

We then noticed that if we were interested in getting nodes $(m'+h, n'+h)$ for any $1 < h \le k+1$ then we could simply multiply the present node by z^{h-1}. If we wanted a node (m,n) beyond these nodes we did a recursive call.

However,

$$A(z)\cdot T_{(m'+1)/(n'+1)}\cdot z^{k} + B(z)\cdot S_{(m'+1)/(n'+1)}(z) = z^{(m'+1+k)+(n'+1+k)+1}\cdot R_{(m'+1)/(n'+1)}(z)$$

so that node $(m'+1+k, n'+1+k)$ is exact and hence would make a next predecessor node along the (m,n)-offdiagonal. Thus, to continue with this sort of iteration we would need to align equations (5.1) with (5.2) to get node $(m'+k+2, n'+k+2)$.

$$A(z)\cdot T_{m'/n'}(z)\cdot z^{k+2} + B(z)\cdot S_{m'/n'}(z)\cdot z^{k+2} = z^{m'+n'+k+3}\cdot R_{m'/n'}(z)$$

$$A(z)\cdot T_{(m'+1)/(n'+1)}(z) \quad + \quad B(z)\cdot S_{(m'+1)/(n'+1)}(z) \quad = z^{m'+n'+k+3}\cdot R_{(m'+1)/(n'+1)}(z)\,.$$

To get the $(m'+k+2, n'+k+2)$ scaled Padé fraction, we then simply find the (k+1,0) scaled Padé fraction for the quotient made from the two residual power series

$$R_{m'/n'}(z) \quad \text{and} \quad R_{(m'+1)/(n'+1)}(z)\,.$$

We get a new algorithm for determining the scaled Padé fractions along the (m,n) offdiagonal by iterating from predecessor/present until we reach the node that we are interested in. If we return to our algorithm, we see that the above description for the algorithm can be implemented by changing our step size. In light of the fact that the new version of the algorithm is totally iterative, we keep track of our iterations by a variable i and let $S_i(z)$ and $T_i(z)$ denote the predecessor at step i and $S_{i+1}(z)$ and $T_{i+1}(z)$ the subsequent predecessor (instead of a present node).

The details of the new algorithm are then given by (Cabay-Choi [14]).

MATRIX-OFFDIAG #2

INPUT: A, B, m, n, p where

1) m, n, and p are nonnegative integers with $p \geq 1$.

2) A and B are unit p x p matrix power series. Actually we simply require A mod z^{m+n+1} and B mod z^{m+n+1}.

OUTPUT: Matrix polynomials S_{i+1} , T_{i+1} , S_i , T_i and an integer IER where

1) The pair S_{i+1} and T_{i+1} is the right scaled matrix Padé fraction of type (m,n) of the matrix power series $-B^{-1} \cdot A$.

2) The pair S_i and T_i is the right scaled Padé fraction of type (m-h-1,n-h-1) of the matrix power series $-B^{-1} \cdot A$ where h is determined from

$$RGCD(S_{i+1}, T_{i+1}) = z^h \cdot I_p$$

3) IER is an error indicator. If IER = 1 then no such scaled Padé fractions were found using this algorithm. If IER = 0 then algorithm was successful.

Step 1: #Initialization#

$$
\begin{aligned}
i &\leftarrow -1 \\
M &\leftarrow (m-n) \\
N &\leftarrow 0 \\
IER &\leftarrow 0 \\
S_{i+1} &= -B^{-1} \cdot A \ mod \ z^{M+1} \\
S_i &= z^{M-1} \cdot I_p \\
T_{i+1} &= I_p \\
T_i &= 0
\end{aligned}
$$

Step 2: #Termination Criterion #

If N = n then exit, else $i \leftarrow i+1$

Step 3: # Computation of residuals of the pair S_i and T_i.

Compute R_i and an integer k from

$$A \cdot T_i + B \cdot S_i \ mod \ z^{M+N+2k+3} = z^{M+N+k+1} \cdot R_i$$

where $R_i(0) \neq 0$ is invertible if k < 2(n - N). If $R_i(0)$ is singular set IER = 1 and exit.

Step 4: #Calculation of step size #

$s \leftarrow \min\{ k+1 , n - N \}$

Step 5: # Identification of cases #

If $k \geq s$ then let

$$S_{i+1} = z^s \cdot S_i$$
$$T_{i+1} = z^s \cdot T_i$$
$$S_i = S_{i-1}$$
$$T_i = T_{i-1}$$

and go to step 8. Else go to step 6.

Step 6: # Computation of residual for S_{i-1} and T_{i-1} #

Compute R_{i-1} such that

$$A \cdot T_{i-1} + B \cdot S_{i-1} \ mod \ z^{M+N+k+1} = z^{M+N-1} \cdot R_{i-1} ,$$

where $R_{i-1}(0) \neq 0$.

Step 7: # Advancement of scaled Padé fraction computation#

$$S_{i+1} = -S_i \cdot (R_i^{-1} \cdot R_{i-1} \ mod \ z^{k+2}) + z^{k+2} \cdot S_{i-1} ,$$
$$T_{i+1} = -T_i \cdot (R_i^{-1} \cdot R_{i-1} \ mod \ z^{k+2}) + z^{k+2} \cdot T_{i-1} ,$$

$$S_{i+1} = S_i$$
$$T_{i+1} = T_i$$

Step 8: # Calculation of degrees of S_{i+1} and T_{i+1} #

$$N \leftarrow N + s$$
$$M \leftarrow M + s$$

go to step 2.

5.2. Duality of Right Scaled Matrix Padé Fractions and LGCDs

When dealing with matrix polynomials, whether for greatest common divisors or for Padé fractions, the lack of commutativity requires the specification of right or left matrix multiplication to be included in all definitions.

As mentioned previously there is a relationship between the scalar offdiagonal algorithm and the extended Euclidean algorithm for calculating greatest common factors of two polynomials. This duality can be extended naturally to the matrix polynomial case if certain invertibility assumptions are made. These correspond to nearly-normal power series.

The main ingredient of this duality is the correspondence between a polynomial and its reciprocal polynomial.

Definition 5.1: Let p(z) be any polynomial with coefficients from a ring R. Then the reciprocal of p(z), denoted by $p^*(z)$, is defined by:

$$p^*(z) = z^n \cdot p(z^{-1}),$$

where n is the degree of p(z).

If

$$p(z) = p_0 + p_1 z + \cdots + p_n z^n$$

then

$$p^*(z) = p_n + p_{n-1}z + \cdots + p_0 z^n ,$$

i.e., $p^*(z)$ is the same as p(z) with the coefficients in reverse order. The name reciprocal is due to the fact that the roots of p(z) are the reciprocals of the roots of $p^*(z)$. Clearly $p(z) = (p^*)^*(z)$ when $p_0 \cdot p_n \neq 0$.

Recall that when we have a polynomial ring with coefficients from a field, then we have the extended Euclidean algorithm which is an iterative procedure for finding the greatest common divisor of two polynomials a(z) and b(z). It produces a sequence of four polynomials

$$t_i(z) \;,\; s_i(z) \;,\; r_i(z) \;, \text{ and } \; q_i(z)$$

that satisfy the initial conditions

$$s_{-1}(z) = 1, \quad t_{-1}(z) = 0, \quad r_{-1}(z) = a(z)$$

$$s_0(z) = 0, \quad t_0(z) = 1, \quad r_0(z) = b(z) .$$

Here we assume $\partial(a) \geq \partial(b)$.

The method to determine higher values of $r_i(z)$, and $q_i(z)$ involves a simple division algorithm

$$r_{i-2}(z) = r_{i-1}(z) \cdot q_{i-1}(z) + r_i(z)$$

with $\partial(r_i) < \partial(r_{i-1})$. The higher values of $s_i(z)$, and $t_i(z)$ are then determined by working backwards via

$$s_i(z) = s_{i-2}(z) - s_{i-1}(z) \cdot q_i(z)$$

and

$$t_i(z) = t_{i-2}(z) - t_{i-1}(z) \cdot q_i(z) ,$$

which results in the identity

$$a(z)\cdot t_i(z) + b(z)\cdot s_i(z) = r_i(z)$$

for $0 \le i \le n+1$.

Eventually this sequence of remainders must stop since the degrees are decreasing at every step. The last nonzero remainder, $r_n(z)$, is then the greatest common divisor and we get the equation

$$r_n(z) = a(z)\cdot t_n(z) + b(z)\cdot s_n(z)\,.$$

In addition to the above relations between t_i, s_i, r_i and q_i there are the following intermediate results that fall out of this sequence (McEliece and Shearer [22]):

$$r_{i-1}(z)s_i(z) - r_i(z)s_{i-1}(z) = (-1)^i b(z)\,, \qquad \text{for } 0 \le i \le n+1\,, \tag{P1}$$

$$r_{i-1}(z)t_i(z) - r_i(z)t_{i-1}(z) = (-1)^{i+1} a(z)\,, \qquad \text{for } 0 \le i \le n+1\,, \tag{P2}$$

$$t_i(z)s_{i-1}(z) - t_{i-1}(z)s_i(z) = (-1)^{i+1}\,, \qquad \text{for } 0 \le i \le n+1\,, \tag{P3}$$

$$\partial(t_i) + \partial(r_{i-1}) = \partial(a)\,, \qquad \text{for } 1 \le i \le n+1\,, \tag{P4}$$

$$\partial(s_i) + \partial(r_{i-1}) = \partial(b)\,, \quad 0 \le i \le n+1\,. \tag{P5}$$

In the situation where we are dealing with matrix polynomials rather than scalar polynomials, we have problems with using a Euclidean algorithm. The first involves the lack of commutativity of matrices and hence a need to spell out if we are finding the right or the left greatest common divisor. In the above description care has been taken to divide the remainder on the left throughout and the algorithm results in the finding of the left greatest common divisor.

A second problem is that a division algorithm does not exist for the ring of matrix polynomials when the norm is taken to be the degree of the matrix polynomial because of the lack of matrix invertibility for the coefficients. Thus, a division algorithm exists for only a subset of the entire ring, and so a Euclid's algorithm will only be available for a similar subclass.

For those matrix polynomials for which we can apply Euclid's algorithm, there is a duality between finding a left greatest divisor via Euclid's algorithm and a right scaled Padé fraction in the matrix case via MATRIX-OFFDIAG#2. We illustrate this by

Example 5.1. Let

$$A(z) = \begin{bmatrix} z^7+z^5+z^4+1 & z^7+z^5+z+1 \\ z^5+z^4+z^2 & z^7+z^6+z^3+z^2 \end{bmatrix}$$

and

$$B(z) = \begin{bmatrix} z^4+z^3+1 & z^6+z^4+z^3+z \\ z^6+z^2 & z^5+z^3 \end{bmatrix}.$$

For ease of calculation we will work over the field Z_2, the field of integers modulo 2. If we go through the Euclidean algorithm on the left then we get that the greatest common divisor on the left is

$$r_3(z) = \begin{bmatrix} z^2 & 1 \\ 0 & z^2 \end{bmatrix}.$$

To get an idea of how the left Euclidean algorithm is a dual to the right MATRIX-OFFDIAG algorithm we do not write out all the individual divisions with their respective quotients but rather concern ourselves with the matrices s_i and t_i in the context of property P4. For i = 0, 1, 2 and 3 this property translates into

$$A(z)\cdot 0 + B(z)\cdot I = B(z) \qquad (5.1)$$

$$A(z)\cdot I + B(z)\cdot\begin{bmatrix} 1 & z \\ z & z \end{bmatrix} = \begin{bmatrix} z^4+z^3+z^2 & z^5+z^2+1 \\ z^5 & z^4+z^2 \end{bmatrix} \qquad (5.2)$$

$$A(z)\cdot\begin{bmatrix} z & 0 \\ 1 & z \end{bmatrix} + B(z)\cdot\begin{bmatrix} 1 & z^2 \\ z^2+z & z^2+1 \end{bmatrix} = \begin{bmatrix} z^2 & z^4 \\ z^4 & 0 \end{bmatrix} \qquad (5.3)$$

$$A(z)\cdot\begin{bmatrix} z^2+1 & z \\ 0 & z^2 \end{bmatrix} + B(z)\cdot\begin{bmatrix} z^2+z+1 & z^3+z+1 \\ z^3+z+1 & z^3+z^2+z \end{bmatrix} = \begin{bmatrix} z^2 & 1 \\ 0 & z^2 \end{bmatrix} \qquad (5.4)$$

Notice, that if at any time the matrix on the right did not have an invertible leading coefficient then the algorithm would have to stop.

If we replace all occurrences of z with z^{-1} in the above equations and then multiply each equation by z^l, where l is the smallest power that makes the expressions into polynomials in z, we will get the following series of expressions.

$$A^*(z)\cdot 0 \quad + \quad B^*(z)\cdot z\cdot I \quad = z\cdot B^*(z) \qquad (5.1^*)$$

$$A^*(z)\cdot I \quad + \quad B^*(z)\cdot\begin{bmatrix} z & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} z^3+z^4+z^5 & z^2+z^5+z^7 \\ z^2 & z^3+z^5 \end{bmatrix} \qquad (5.2^*)$$

$$= z^2\cdot\begin{bmatrix} z+z^2+z^3 & 1+z^3+z^5 \\ 1 & z+z^3 \end{bmatrix}$$

$$A^*(z)\cdot\begin{bmatrix} 1 & 0 \\ z & 1 \end{bmatrix} + B^*(z)\cdot\begin{bmatrix} z^2 & 1 \\ 1+z & 1+z^2 \end{bmatrix} = \begin{bmatrix} z^4+z^6 & z^4 \\ z^4 & 0 \end{bmatrix} \qquad (5.3^*)$$

$$= z^4\cdot\begin{bmatrix} 1+z^2 & 1 \\ 1 & 0 \end{bmatrix}$$

$$A^*(z)\cdot\begin{bmatrix} 1+z^2 & z \\ 0 & 1 \end{bmatrix} + B^*(z)\cdot\begin{bmatrix} z+z^2+z^3 & 1+z^2+z^3 \\ 1+z^2+z^3 & 1+z+z^2 \end{bmatrix} = \begin{bmatrix} z^7 & z^9 \\ 0 & z^7 \end{bmatrix} \qquad (5.4^*)$$

$$= z^7\begin{bmatrix} 1 & z^2 \\ 0 & 1 \end{bmatrix}$$

Notice that the above four equations give a sequence of 3 polynomials $t^*_i(z)$, $s^*_i(z)$ and $r^*_i(z)$ that satisfy the condition of being right scaled matrix Padé fractions for $a^*(z)$ and $b^*(z)$ in the sense of our version of MATRIX-OFFDIAG. These right scaled matrix Padé fractions are of type (0,-1), (1,0), (2,1), and (3,2), respectively. Also, we can also calculate the (4,3) right scaled Padé fraction by multiplying equation (5.4*) by z.

The formal statement of the duality result can now be stated as

Theorem (Duality) Let A(z) and B(z) be finite matrix polynomials with invertible leading term. Assume that the formal matrix power series $B(z)^{-1} \cdot A(z)$ is a nearly-normal power series. Let $T_i(z)$, $S_i(z)$ and $R_i(z)$ be the right denominators, right numerators, and right residuals that result from calculating the right scaled matrix Padé fractions via the MATRIX-OFFDIAG algorithm of Chapter 4. Then the matrix polynomials $T_i^*(z)$, $S_i^*(z)$, and $R_i^*(z)$, that result from the algorithm are the same as the multipliers that result from the extended Euclidean algorithm for left greatest common divisor (LGCD) as applied to the matrix polynomials $A^*(z)$ and $B^*(z)$.

Proof:

The formal proof of this result parallels the scalar result found in the paper of Cabay and Choi [14] and is not given here.

5.3. A Fast Pseudo-Euclidean Algorithm

The last section gave a duality between Euclid's algorithm for determining the left greatest common divisor of two matrix polynomials and the MATRIX-OFFDIAG#2 algorithm for determining the right scaled Padé fraction of a quotient power series. The advantage of this of course is that, if we were interested in the greatest common divisor, then we need not use MATRIX-OFFDIAG#2 (i.e. Euclid's Algorithm) but rather the much faster algorithm MATRIX-OFFDIAG.

This speed up is impressive and useful in the scalar case since all scalar power series are nearly-normal. Thus, we can always apply the off-diagonal algorithm just like we can always apply Euclid's algorithm. However, this is not the case in the matrix polynomial situation. Here we are always faced with the problem of perhaps lacking invertibility, making for a situation where Euclid's algorithm cannot be applied. On the other hand, when Euclid's algorithm can be applied then, by duality, MATRIX-OFFDIAG#2 can be applied, and consequently the faster MATRIX-OFFDIAG algorithm can also be used.

In the case where we are determining the greatest common divisor of two matrix polynomials A(z) and B(z), Euclid's algorithm breaks down when there is a remainder that does not have an invertible matrix as the coefficient of the highest power. To counter this, at least in the case that the remainder is invertible even though the highest term is not, we can use a type of pseudo division to get a pseudo Euclidean algorithm. To understand this pseudo division we first prove the following lemma.

Lemma 5.2: Let R(z) be a p x p matrix polynomial of degree $\leq$ n. Suppose that R(z) is nonsingular, i.e., det(R(z)) $\neq$ 0. Then there exists a matrix factor V(z) with the property that

$$R(z) = Q(z) \cdot V(z),$$

where Q(z) is a unit polynomial, i.e., det(Q(0)) $\neq$ 0. The factor V(z) is of the form

$$S(z) = D_1 \cdot C_1 \cdot \ \ldots \ \cdot D_k \cdot C_k, \ \ with \ det(S(z)) = K \cdot z^h ,$$

where the matrices C_i are constant matrices and the D_i are diagonal matrices with only powers of z on the diagonal. The integer h is the leading nonzero power in det(R(z)). In addition, V(z) has degree less than n.

Proof: Let

$$R(z) = R_0 + R_1 z + \cdots + R_n z^n .$$

Suppose det(R(z)) = r(z), where $r(z) = r_h z^h + r_{h+1} z^{h+1} + \cdots$. If h = 0, then R(z) is a unit matrix polynomial and so we set Q(z) = R(z) and V(z) = I. Otherwise, since det(R_0) = det(R(0)) = 0, R_0 is a singular matrix. Thus we can column reduce R_0, via multiplication on the right by a nonsingular matrix C_1, to the matrix with the last k columns, $k \geq 1$, being zero columns. Let $C'(z^{-1})$ be the diagonal matrix with z^{-1} on the diagonal for the last k columns and with 1 along the remaining entries. If

$$R'(z) = R(z) \cdot C_1 \cdot C'(z^{-1});$$

then $R'(z)$ is still a matrix polynomial by the construction of C_1 and $C'(z)$. Furthermore, we can write

$$R(z) = R'(z) \cdot V_1(z), \tag{5.6}$$

where

$$V_1(z) = C'(z) \cdot C_1^{-1}.$$

If we take determinants in equation (5.6), then we see that $R'(z)$ is a factor of R(z) and

$$det(R'(z)) = r'_u z^u + \cdots,$$

where $0 \leq u \leq h - p \cdot k < h$. If u is 0, then we are done. Otherwise, we repeat the above construction on $R'(z)$. Eventually we must reach a stage where the leading term of the determinant is nonzero. Working backwards, we can then factor R(z) into the desired form!

We note that there is a parallel result where the unit factor is on the right instead of the left.

To see how the lemma is used, suppose that while calculating our division sequence we come across the equation

$$R_{i-1}(z) = R_i(z) \cdot Q_i(z) + R_{i+1}(z).$$

Any matrix polynomial that divides A(z) and B(z) and the remainders $R_0(z)$, ..., $R_i(z)$ will then divide $R_{i+1}(z)$. If the highest coefficient is not invertible then we can reverse the order of the coefficients, use lemma 5.2 to factor the reversed polynomial into a unit times a "scale" matrix, and reverse the orders again to get

$$R_{i+1}(z)\cdot V_{i+1}(z) = R'_{i+1}(z),$$

where $R'_{i+1}(z)$ has an invertible matrix as its highest coefficient.

If the remainder becomes singular, then we must stop without determining the largest divisor. Otherwise, we can determine the largest unit divisor of A(z) and B(z).

In the standard case, when we write out this greatest unit common divisor as a linear combination of both A(z) and B(z), we will get the equation by working the division algorithm in reverse order. Thus, we get

$$R_{k+1}(z) = R_k(z)\cdot Q_k(z) - R_{k-1}(z) \quad (5.9)$$

$$= \left\{R_{k-1}(z)\cdot Q_{k-1}(z) - R_{k-2}(z)\right\}\cdot Q_k(z) - R_{k-1}(z) \quad (5.10)$$

$$= R_{k-1}(z)\, Q'_{k-1}(z) + R_{k-2}(z)\cdot Q'_{k-2}(z), \quad (5.11)$$

and so on.

If we come to a stage where pseudo division is required, then we get a small difference in our process. For example, suppose that the remainder $R_{k-1}(z)$ is the result of a pseudo division by the remainder $R'_{k-2}(z)$ so that

$$R_{k-1}(z) = R'_{k-2}(z)\cdot Q_{k-2}(z) + R_{k-3}(z).$$

Since

$$R_{k-2}(z)\cdot V_{k-2}(z) = R'_{k-2}(z),$$

we still get

$$R_{k-1}(z) = R_{k-2}(z)\cdot Q'_{k-2}(z) + R_{k-3}(z),$$

where

$$Q'_{k-2}(z) = V_{k-2}(z) \cdot Q_{k-2}(z) .$$

Therefore, we can continue the process begun in equation (5.9), obtaining eventually

$$R_{k+1}(z) = A(z) \cdot T_{k+1}(z) + B(z) \cdot S(z) .$$

Thus, as long as we never have a remainder that is singular, we can determine the greatest common divisor for A(z) and B(z). Note that, should we attempt to reverse the order of the coefficients at every step, we would lose the duality with scaled Padé fractions because we lose our control of the degrees of the comultipliers T(z) and S(z).

However, the MATRIX-OFFDIAG algorithm can still be used to accomplish a speed up to this pseudo-Euclidean algorithm. For, up until we run into a remainder without an invertible leading term, we are really just doing the standard Euclidean algorithm. Since this is dual to MATRIX-OFFDIAG#2 we simply use the faster MATRIX-OFFDIAG algorithm for this stage. Once a pseudo division has taken place, we are just applying Euclid's algorithm to the original remainder and the new pseudo remainder so again we just speed things up by using MATRIX-OFFDIAG until pseudo division is required again. This will give us a fast pseudo Euclidean algorithm for finding the left greatest common divisor of two matrix polynomials A(z) and B(z).

To see that we can use MATRIX-OFFDIAG when we are required to pseudo divide, we return to our original description of the algorithm that appeared in chapter 4. The MATRIX-OFFDIAG algorithm breaks down when the predecessor satisfies

$$A(z) \cdot T_{m/n}(z) + B(z) \cdot S_{m/n}(z) = z^{m+n+1} \cdot R_{m/n}(z) ,$$

while the present satisfies

$$A(z) \cdot T_{(m+1)/(n+1)}(z) + B(z) \cdot S_{(m+1)/(n+1)}(z) = z^{m+n+k+3} \cdot R_{(m+1)/(n+1)}(z) , \qquad (5.12)$$

but where $R_{(m+1)/(n+1)}(z)$ is not a unit power series. Thus, we have a problem with the fact that we can no longer form the power series $\left\{R_{(m+1)/(n+1)}(z)\right\}^{-1}$, since the leading term is no longer invertible. However, if the residual is at least invertible, even though the leading term is not, then we can factor the residual into

$$R_{(m+1)/(n+1)}(z) = R'_{(m+1)/(n+1)}(z) \cdot V(z),$$

where $R'_{(m+1)/(n+1)}(z)$ is a unit matrix power series and $V(z)$ has the special form described in lemma 5.2.

If we now invert $V(z)$ and multiply equation (5.12) by this matrix we get the two equations

$$A(z) \cdot T_{m/n}(z) \cdot z^{2+k} + B(z) \cdot S_{m/n}(z) \cdot z^{2+k} = z^{m+n+k+3} \cdot R_{m/n}(z) \quad (5.13)$$

$$A(z) \cdot T_{(m+1)/(n+1)}(z) \cdot \left\{V(z)\right\}^{-1} + B(z) \cdot S_{(m+1)/(n+1)}(z) \cdot \left\{V(z)\right\}^{-1} \quad (5.14)$$

$$= z^{m+n+k+3} \cdot R'_{(m+1)/(n+1)}(z).$$

We can then work with the quotient $\left\{R'_{(m+1)/(n+1)}(z)\right\}^{-1} \cdot R_{m/n}(z)$ to further our order condition but at the cost of having denominators and numerators with negative powers.

On the other hand, when we work in the dual situation, these negative powers do not matter. We still get linear combinations of the matrix polynomials and so we can say something about the greatest common divisors of the two reversed matrix polynomials. To see this, assume A(z) and B(z) are of degree M and N, respectively, with $M \geq N$ and where M - N = m - n. Suppose that we wish to calculate the right scaled Padé fraction of type (m',n'). Furthermore, assume that we can solve the right scaled Padé problem for the residual quotient series without any further matrix scaling. Then, we have

$$A(z)\cdot T'_{m'/n'}(z) + B(z)\cdot S'_{m'/n'}(z) = z^{m'+n'+k'+1}\cdot R_{m'/n'}(z) \tag{5.15}$$

where

$$T'_{m'/n'}(z) = T_{m/n}(z)\cdot P(z)\cdot z^{2+k} + T_{(m+1)/(n+1)}(z)\cdot\left\{V(z)\right\}^{-1}\cdot Q(z) \tag{5.16}$$

$$S'_{m'/n'}(z) = T_{m/n}(z)\cdot P(z)\cdot z^{2+k} + S_{(m+1)/(n+1)}\cdot\left\{V(z)\right\}^{-1}\cdot Q(z), \tag{5.17}$$

and where the pair Q(z) and P(z) forms the right scaled matrix Padé fraction of type $(m'-m-k-2, m'-m-1)$ for the quotient of the residues. Substitute z^{-1} in place of z in the the above two equations. Since our construction of $V(z)$ implied that we have a polynomial in z^{-1}, we get that

$$\left\{V(z^{-1})\right\}^{-1}$$

is just a polynomial in z. If we now multiply equation (5.16) by $z^{n'}$ and equation (5.17) by $z^{m'}$, then we get that $T'^{*}_{m'/n'}(z)$ and $S'^{*}_{m'/n'}(z)$ where

$$T'^{*}_{m'/n'}(z) = z^{n'}\cdot T'_{m'/n'}(z^{-1})$$

and

$$S'^{*}_{m'/n'}(z) = z^{m'}\cdot S'_{m'/n'}(z^{-1})$$

are matrix polynomials (even though the original T' and S' were not matrix polynomials). If we now replace all occurrences of z by z^{-1} in equation (5.15) and then multiply the resulting equation by $z^{N+m'}$ we get

$$A^{*}(z)\cdot T'^{*}_{m'/n'}(z) + B^{*}(z)\cdot S'^{*}_{m'/n'}(z) = R^{*}_{m'/n'}(z) \tag{5.18}$$

where the remainder $R^{*}_{m'/n'}(z)$ is a matrix polynomial of smaller degree than the previous remainder. This process can be continued at every recursive call with the resulting pseudo numerator and denominator being a polynomial after the reversal process.

We summarize our discussion by

ALGORITHM (PMATRIX-OFFDIAG).

INPUT: A,B,m,n,p where

1) m, n, p are nonnegative integers with m ≥ n and p ≥ 1.

2) A and B are pxp matrix power series (as before we only need the polynomials A mod z^{m+n+1} and B mod z^{m+n+1}). B must be a unit power series.

OUTPUT:

1) Matrix pseudo-polynomials (i.e. negative powers are allowed) S_1, S_0, T_1, T_0 and an integer IER where:

$$A \cdot T_1 + B \cdot S_1 = z^{m+n+1} \cdot R_{m/n}$$

$$A \cdot T_0 + B \cdot S_0 = z^{m+n-2h-1} \cdot R_{(m-h-1)/(n-h-1)}$$

where $R_{m/n}$ and $R_{(m-h-1)/(n-h-1)}$ are matrix polynomials and h is determined via

$$z^h = RGCD(S_1 ; T_1)$$

In addition the degrees of S_1, T_1, S_0, T_0 are at most m, n, m-h-1, n-h-1, respectively (By the degree of a pseudo-polynomial we mean the highest power of z in the expression).

2) IER, an integer variable that is used to indicate invertibility has been met (0 if success, 1 if failure).

Step 1: #initialization#

$IER \leftarrow 0$
$i \leftarrow -1$

$M \leftarrow (m-n)$

$N \leftarrow 0$

$S_1 \leftarrow -B^{-1} \cdot A \bmod z^{M+1}$, $S_0 \leftarrow z^{M-1} \cdot I_p$

$T_1 \leftarrow I_p$, $T_0 \leftarrow 0$

Step 2: #calculation of step-size#

$i \leftarrow i+1$

$s \leftarrow \min\{2^i - N, n - N\}$

Step 3: #Termination criterion#

If $s = 0$ exit

Step 4: #calculation of closest present node before the pair S_1, T_1#

Determine h such that

$$z^h \cdot I_p = RGCD(S_1, T_1)$$

Set

$$S_1 \leftarrow z^{-h} \cdot S_1 , \quad T_1 \leftarrow z^{-h} \cdot T_1$$

Step 5: #Calculation of residual for $S_1 \cdot T_1^{-1}$ #

Compute k and the power series R_1 such that

$$(A \cdot + B \cdot S_1) \bmod z^{M+N-h+k+1} \cdot R_1 ,$$

where $R_1 \neq 0$ if $k < 2s + 1$.

Step 6: #Calculation of matrix scaling polynomial#

Calculate matrix polynomials V_1 and R'_1 such that

$$R_1 = R'_1 \cdot V_1$$

and where $R'_1(0)$ is nonsingular. If such a matrix cannot be found set IER = 1 and exit.

Step 7: #seperating into different cases#

If $k \geq s$ then

$$S_1 \leftarrow z^{s+h} \cdot S_1 \quad , \quad T_1 \leftarrow z^{s+h} \cdot T_1$$

then go to step 12. Else go to step 8.

Step 8: #calculation of degrees for residual scaled matrix Padé fraction #

$m' \leftarrow s+h$

$n' \leftarrow s-k-1$

Step 9: #computation of residual for $S_0 \cdot T_0^{-1}$#

Compute R_0 so that

$$(A \cdot T_0 + B \cdot S_0) \; mod \; z^{M+N+m'+n'-2h} = z^{M+N-2h-1} \cdot R_0$$

where $R_0(0) \neq 0$.

Step 10: #Computation of residual right pseudo-polynomials#

$$S'_1 , S'_0 , T'_1 , T'_0$$

determined from PMATRIX-OFFDIAG(R_0, R'_1, m', n', p, IER). If IER = 1 then exit.

Step 11: #Advance along offdiagonal for scaled fractions #

$$S_1 \leftarrow S_1 \cdot V_1^{-1} \cdot T'_1 + z^{h+k+2} \cdot S_0 \cdot S'_1$$

$$T_1 \leftarrow T_1 \cdot V_1^{-1} \cdot T'_1 + z^{h+k+2} \cdot T_0 \cdot S'_1$$

$$S_0 \leftarrow S_1 \cdot V_1^{-1} \cdot T'_0 + z^{h+k+2} \cdot S_0 \cdot S'_0$$

$$T_0 \leftarrow T_1 \cdot V_1^{-1} \cdot T'_0 + z^{h+k+2} \cdot T_0 \cdot S'_0$$

Step 12: #Calculation of degrees of $S_1 \cdot T_1^{-1}$ #

$$M \leftarrow M + s$$
$$N \leftarrow N + s$$

go to step 2.

The algorithm, when interpreted in the dual situation, gives us a fast pseudo Euclidean algorithm.

Chapter 6

Conclusions

Research into the problem of determining rational approximants to formal power series has been going on for well over a hundred years. Emphasis has centered about the calculation of these rational approximants for scalar power series. A similar problem occurs when higher-dimensional power series are considered. This thesis has considered the problem of determining an adequate definition for a rational approximant of a formal matrix power series and also, given a suitable definition, the problem of computation of these matrix fractions. In attempting to extend the notion of Padé approximant we have limited our study to that of square matrix power series.

The classical theory of Padé approximants for the scalar situation centers about the concept of a Padé form, which always exists but is not unique, and Padé fraction, which is unique but does not always exist. A characterization of the system of equations that comes up when solving for Padé forms leads one to the concept of scaled Padé fraction. These always exist and, in addition, are unique.

When the problem of Padé approximants is extended to the multidimensional case there is again the problem of existence and/or uniqueness of these rational fractions. Even extending the definitions of Padé form, Padé fraction and scaled Padé fraction leads to difficulties. For example, we must specify the side that the matrix multiplication must be on since matrices are not commutative. This is easily accomplished. A more complex problem involves the extension of a fundamental requirement for a scalar power series, namely that the denominator of any rational expression be nonzero. This is a straightforward condition in the scalar case, but can be extended in numerous ways in the matrix case. Asking that the denominator be nonzero is too broad while requiring that the denominator be invertible (equivalent in the scalar case) is very restrictive. A middle ground is chosen and a definition for Padé form is given in the matrix case. This definition is meant to be as broad as possible and to extend the definition of Padé form in the scalar case. As in the scalar case, existence is demonstrated for matrix Padé forms. However, the broadness of our definition results in unexpected, and unwanted, behaviour. We find that, for some particular power series, we can end up with multiple answers, some with invertible denominator others with singular ones. For a particular matrix power series a Padé form on the right may have different fundamental properties than those found on the left. There are even situations where a Padé form can be found but where the denominator will not be invertible.

These undesirable fundamental properties force us to restrict the type of power series that we study. A subclass of matrix power series, the nearly-normal matrix power series, is then introduced. These nearly-normal power series provide a more natural extension of scalar power series in terms of the type of Padé approximants that one can calculate for the series. These power series are more general than the concept of normal power series which are the type of matrix power series that are most often found in the literature. In addition, all scalar power series are nearly-normal. Nearly-normal power series also lead to a type of matrix fraction that extends the notion of scaled Padé fraction found in the scalar case. As in the scalar case, these exist and are unique up to multiplication by a nonzero constant matrix.

The existence and uniqueness of these scaled matrix Padé fractions for nearly-normal matrix power series allows us to determine an algorithm that will calculate these quantities. An Off-diagonal algorithm is presented in the matrix case. This algorithm calculates the scaled matrix Padé fractions and extends an algorithm that was first presented by Cabay and Choi [14] in the scalar case. In particular, it calculates the fractions for any scalar power series and any normal matrix power series in addition to others.

A slow algorithm is also presented that calculates the matrix scaled Padé fractions along the same offdiagonal. This second algorithm calculates the same quantities as the first. It is useful in that, when the algorithm is applied to the formal power series that results from the quotient of two finite degree matrix polynomials, one gets a dual algorithm to Euclid's algorithm for calculating the greatest common divisor to two matrix polynomials. This duality is accomplished by reversing the order of the coefficients in each polynomial. Thus, the first algorithm can be used as a fast algorithm for calculating greatest common factors.

In the matrix polynomial case, Euclid's algorithm is limited because of strong invertibility restrictions. It can be extended, however, to produce a pseudo-Euclidean algorithm (instead of calculating a remainder sequence one calculates a sequence of pseudo remainders). This can also determine a greatest common divisor but with fewer invertibility restrictions. The dual version of this does not produce an algorithm to calculate Padé fractions because the degrees of the comultipliers of the remainder sequence are not kept under as tight a control as they are in Euclid's algorithm. Nonetheless, the ideas involved in the algorithm to calculate a fast GCD algorithm also extend to produce a fast pseudo algorithm for computing GCD's.

The biggest drawback to these algorithms is that they can halt when certain invertibility conditions are not met. Thus, they can be used to reduce a GCD calculation but not always solve it. In the case that the algorithm is halted, other algorithms must be applied where possible.

Given the examples of unusual behaviour found in chapter 3, it is probably unlikely that the present definitions for matrix Padé approximant will suffice for all matrix power series. This is not to say that we might not find a situation where we simply end up not calculating all the Padé approximants along a specific offdiagonal. By stepping off the diagonal paths, we might succeed in calculating a subset of those along the offdiagonal in terms of previous successful calculations.

To extend the notion of Padé approximant to all unit matrix power series there must likely be a different notion of the degree of a matrix polynomial, and/or the notion of the order of a residual sequence. One likely candidate for the degree of a matrix polynomial would be the determinant degree, i.e., the degree of the polynomial that results from taking the determinant of the matrix polynomial. There are plausible arguments for changing from degree to determinant degree both from the fields of algebraic computation and from the study of systems theory.

The evidence from the field of algebraic computation comes from consideration of the dual situation, namely the problem of determining greatest common divisors of two matrix polynomials. The notion of a trivial divisor uses the concept of determinant degree instead of polynomial degree. Also, for certain sets of matrix polynomials there is a quotient algorithm as long as the norm used is that of a determinant degree. See for example Sanov [29].

There is also evidence from systems theory that the determinant degree is at least as important as the ordinary degree of a matrix polynomial. For a matrix polynomial, the roots are uninformative. However, the roots of the determinant provide the values where the numerator or denominator is singular. This is important in systems theory.

Along with changes in the notion of a degree of a Padé form one can also alter the notion of order of a matrix power series. The order condition could be altered for each row or column. The row or column orders could then add to a value greater than some lower bound. The invertibility of a residual could then be similar to the notion of row or column proper (or at least its reversed polynomial form could be proper). We leave these ideas for further research efforts.

References

[1] H. Anton, *Elementary Linear Algebra*, Wiley, New York, 1984.

[2] G.A. Baker, *Essentials of Padé Approximants*, Academic Press, New York, 1975.

[3] G.A. Baker, "Invariance Properties in Hermite-Padé Approximation Theory", *J. of Comp. and Appl. Math*, Vol. 11, No. 1, 1984, pp. 49-56.

[4] F.L. Bauer, The Quotient-Difference and Epsilon Algorithms, in *On Numerical Approximation*, R.E. Langer (ed.), University of Wisconsin Press, Madison, 1959, pp. 361-370.

[5] E.R. Berlekamp, *Algebraic Coding Theory*, McGraw-Hill, New York, 1968.

[6] N.K. Bose and S. Basu, "Theory and Recursive Computation of 1-D Matrix Padé Approximants", *IEEE Trans. on Circuits and Systems*, Vol. 4, 1980, pp. 323-325 .

[7] R. Brent, F.G. Gustavson and D.Y.Y. Yun, "Fast Solution of Toeplitz Systems of Equations and Computation of Padé Approximants ", *J. of Algorithms* , Vol. 1, 1980, pp. 259-295.

[8] C. Brezinski, "Computation of Padé Approximants and Continued Fractions", *J. Comput. Appl. Math*, Vol. 2, 1976, pp. 113-123.

[9] F. Brophy and A.C. Salazar, "Considerations of the Padé Approximant Technique in the Synthesis of Recursive Digital Filters", *IEEE Trans. Audio and Electroacoustics*, Vol. AU-21, No. 6, Dec 1973, pp. 500-505.

[10] A. Bultheel, "Recursive Algorithms for the Matrix Padé Problem", *Math. of Computation*, Vol. 35, 1980 , pp. 875-892.

[11] A. Bultheel, "Recursive Algorithms for Nonnormal Padé Tables", *SIAM J. Appl. Math*, Vol. 39, No. 1, 1980, pp. 106-118.

[12] A. Bultheel, "Algorithms To Compute The Reflection Coefficients of Digital Filters", *Num. Methods of Approx Theory*, Vol. 7, 1983.

[13] S. Cabay and T.T.C. Kao, The Diagonal Padé Table and the Triangular Decomposition of Hankel Matrices, Tech. Report TR83-10, Dep't of Computing Science, University of Alberta, 1983.

[14] S. Cabay and D.K. Choi, Algebraic Computations of Scaled Padé Fractions, *SIAM J. of Computation*, 1986.

[15] D.K. Choi, *Algebraic Compuations of Scaled Padé Fractions*, Ph.D. Thesis, University of Alberta, 1984.

[16] G. Claessens, "A new look at the Padé table and the different methods for computing its elements", *J. of Comp and Appl Math*, Vol. 1, 1975, pp. 141-152.

[17] O.I. Elgerd, *Control Systems Theory*, McGraw-Hill, New York, 1967.

[18] Frobenius, "Uber Relationel zwischen den Naeherungsbruchen von Potenzreihen", *J. fur Math*, Vol. 90, 1881, pp. 1-17.

[19] W.B. Gragg, "The Padé Table and its Relation to Certain Algorithms of Numerical Analysis", *SIAM Rev.*, Vol. 14, 1972, pp. 1-61.

[20] E. Horowitz and S. Sahni, *Fundamentals of Computer Algorithms*, Computer Science Press, Maryland, 1978.

[21] J. Makhoul, "Linear Prediction: A Tutorial Review", *Proc. IEEE*, Vol. 63, April 1974, pp. 561-580.

[22] R.J. McEliece and J.B. Shearer, "A Property of Euclid's Algorithm and an Application to Padé Approximation", *SIAM J. Appl. Math*, Vol. 34, 1978, pp. 611-617.

[23] G. Nemeth and M. Zimanyi, "Polynomial Type Padé Approximants", *Math. of Comp.*, Vol. 38, 1982, pp. 553-565.

[24] H. Padé, Sur la Representation Approchee d'une Fonction par des Fractions Rationelles, *Ann Ecole Nor.*, 1892, pp. 1-93, supplement.

[25] J. Rissanen, "Solution of Linear Equations with Hankel and Toeplitz Matrices", *Numer. Math*, Vol. 22, 1974, pp. 361-366.

[26] J. Rissanen, Recursive Evaluation of Padé Approximants for Matrix Sequences, *IBM J. Res. Develop.*, 1972, pp. 401-406.

[27] J. Rissanen, "Algorithms for Triangular Decomposition of Block Hankel and Toeplitz Matrices with Application to Factoring Positive Matrix Polynomials", *Math. Comp.*, Vol. 27, 1973, pp. 147-154.

[28] H. Rutishauser, "Der Quotient-Differenzen-Algorithm", *Z. Angew. Math. Phys.*, Vol. 5, 1954, pp. 233-251.

[29] I.N. Sanov, "Euclid's Algorithm and One-sided Decomposition into Prime Factors for Matrix Rings", *Siberian Math. J.*, Vol. 8, No. 4, 1967, pp. 640-645

[30] Y. Starkand, "Explicit Formulas for Matrix-valued Padé Approximants", *J. of Comp. and Appl. Math*, Vol. 5, No. 1, 1979, pp. 63-65.

[31] K. Stevens, "Synthesis of Speech by Electical Analog Devices", *J. Audio Engineering Soc.*, Vol. 4, No. 1, Jan 1956, pp. 2-8.

[32] A. Verheijen, *Evaluation of Diagonal Padé Algorithms*," M.Sc. Thesis, University of Alberta, 1983.

[33] H. Wakita, "Direct Estimation of the Vocal Tract Shape by Inverse Filtering of Acoustic Speech Waveforms", *IEEE Trans. Audio and Electroacoustics*, Vol. AU-21, No. 5, Oct 1973, pp. 417-427.

[34] P. Wynn, "The Rational Approximation of Functions which are formally defined by a power series expansion", *Math. Comp.*, Vol. 14, 1960, pp. 147-186.